# Nuclear mRNA quality control factors Gbp2 and Hrb1 take part in cytoplasmic mRNA surveillance through nonsense-mediated decay

## Dissertation

for the award of the degree
"Doctor rerum naturalium"
of the Georg-August-Universität Göttingen

within the doctoral program "IMPRS Molecular Biology"
of the Georg-August University School of Science (GAUSS)

submitted by
**Yen-Yun Lu**
from Taipei, Taiwan

Göttingen 2021

**Thesis Committee**

Prof. Dr. Heike Krebber
    Department of Molecular Genetics
    Institute for Microbiology and Genetics
    Georg August University of Göttingen

Prof. Dr. Jörg Großhans
    Campus Institute for Dynamics of Biological Networks (CIDBN)
    Georg August University of Göttingen

Prof. Dr. Reinhard Lührmann
    Department of Cellular Biochemistry
    Max Planck Institute for Biophysical Chemistry

**Members of the Examination Board**

Referee:
Prof. Dr. Heike Krebber
    Department of Molecular Genetics
    Institute for Microbiology and Genetics
    Georg August University of Göttingen

2nd Referee:
Prof. Dr. Jörg Großhans
    Campus Institute for Dynamics of Biological Networks (CIDBN)
    Georg August University of Göttingen

**Further members of the Examination Board**

Prof. Dr. Reinhard Lührmann
    Department of Cellular Biochemistry
    Max Planck Institute for Biophysical Chemistry

Prof. Dr. Stefanie Pöggeler
    Department of Genetics of Eukaryotic Microorganisms
    Institute for Microbiology and Genetics
    Georg August University of Göttingen

Prof. Dr. Kai Heimel
    Department of Molecular Microbiology and Genetics
    Institute for Microbiology and Genetics
    Georg August University of Göttingen

Prof. Dr. Ralf Ficner
    Department of Molecular Structural Biology
    Institute for Microbiology and Genetics
    Georg August University of Göttingen

**Date of oral examination: 23.09.2021**

# Affidavit

I hereby declare that I prepared this doctoral thesis titled "Nuclear mRNA quality control factors Gbp2 and Hrb1 take part in cytoplasmic mRNA surveillance through nonsense-mediated decay" independently and with no other sources and aids than quoted.

Göttingen, July 2021

_____

Yen-Yun Lu

# Table of Contents

# 1. Abstract

Yeast SR-like proteins Gbp2 and Hrb1 shuttle between the nucleus and the cytoplasm, and serve as quality control factors for pre-mRNA splicing. They are loaded onto the RNA co-transcriptionally, and upon correct splicing, recruit mRNA export receptors to facilitate transport through the nuclear pore complex. When errors in splicing occur, they recruit instead the nuclear exosome cofactor TRAMP complex, thereby retaining the faulty transcripts in the nucleus and supporting their degradation. In the cytoplasm, Gbp2 and Hrb1 were found to co-purify with polysome fractions and therefore conceivably serve cytoplasmic functions related to translation. Their exact role therein was however unknown. Nonsense-mediated mRNA decay (NMD) is a cytoplasmic quality control pathway primarily targeting transcripts that contain premature termination codons. Considering the surveillance activity of Gbp2 and Hrb1 in splicing, and the fact that splicing defects are a major source of substrates for NMD, it was investigated whether the two proteins are involved in NMD. Despite a great number of studies, the complete list of participating factors and the exact mechanisms of NMD have not been fully defined. The intimate link between NMD and numerous human diseases calls for our greater understanding of this conserved quality control pathway. The results of this work demonstrate that Gbp2 and Hrb1 are part of the NMD mRNP, where they ensure the efficient removal of the target transcript. They seem not to be essential for substrate recognition and NMD activation, but actively participate in subsequent events. Presumably, they mediate mRNP remodeling that may support a direct connection between Upf1, the main factor of NMD, and the 5' end of the transcript, the site where downstream processes majorly occur. They bind to the 5'-associating translation initiation factor eIF4G to repress translation of the target RNA. Further, they help to recruit decay machineries to promote proper elimination of the defective transcript. Identification of these novel roles of Gbp2 and Hrb1 in the cytoplasm demonstrates a molecular link between the nuclear and cytoplasmic mRNA surveillance systems, which likely contributes to a collective maintenance of transcriptome integrity. Importantly, the functions of Gbp2 and Hrb1 suggest conserved characteristics with the mammalian exon junction complex and its associating SR proteins, presenting a framework for future exploration of so far unregarded functions of human SR proteins in both cytoplasmic and nuclear mRNA quality control. Continued research on yeast SR-like proteins will bring further insights and provide exciting possibilities to unravel the more complex mechanisms in human cells.

# 2. Introduction

## 2.1. Quality control of messenger RNAs

Gene expression is a process carefully regulated to support proper development and growth of organisms as well as their rapid response to diverse environmental challenges. Regulation occurs at various levels, and among those the integrity of key genetic molecules is monitored in the cell. As templates for protein synthesis, the major output of gene expression, messenger RNAs (mRNAs) are one of the main targets subject to such quality control. Indeed, several mechanisms have been discovered that act at different stages of the mRNA life cycle to ensure that only correct mRNAs continue down the pathway and that aberrant ones are eliminated (van Hoof and Wagner, 2011). Many of the mechanisms and responsible protein factors are conserved across eukaryotic species (Fasken and Corbett, 2009; Kilchert and Vasiljeva, 2013; Simms et al., 2017), and similar strategies have been found in prokaryotes (Buskirk and Green, 2017; Inada, 2020). Moreover, defects in mRNA surveillance mechanisms manifest as diseases (Moraes, 2010; Wolin and Maquat, 2019), underscoring the importance of such quality control pathways.

### 2.1.1. Nuclear mRNA quality control

In eukaryotic cells, quality control of mRNA takes place both in the nucleus and cytoplasm, and at the nuclear pore complex (NPC). During transcription, newly synthesized transcripts undergo a series of modifications, which include 5' capping, 3' polyadenylation, and the removal of intron sequences in a process termed splicing (Proudfoot et al., 2002; Bentley, 2014). Quality control occurs in parallel to these events, targeting defective transcripts that arise from errors during processing to rapid degradation by nuclear exoribonucleases (Eberle and Visa, 2014; Bresson and Tollervey, 2018; Schmid and Jensen, 2018) (**Figure 1**). Many proteins are involved, either associating for processing per se, or have additional roles in regulating RNA stability, quality control, and/or connecting upstream and downstream events. mRNA biogenesis is thus accompanied by elaborate remodeling of the messenger ribonucleoprotein (mRNP) structure (Singh et al., 2015).

Pre-mRNA processing begins at the 5' end, where a 7-methylguanosine ($m^7G$) cap is added to the first nucleotide and is subsequently bound by the cap-binding complex (CBC,

Cbc1–Cbc2 in yeast and CBP80–CBP20 in human), protecting the 5' end from degradation. Incorrectly capped transcripts become substrates of the 5'–3' exonuclease Rat1 (XRN2 in human), which degrades the transcript with the help of its cofactor, the pyrophosphatase Rai1 (Jiao et al., 2010) (**Figure 1b**). A protein homologous to Rai1, Dxo1 (DXO in human), was also found to function in the quality control and degradation of transcripts that are incorrectly capped (Chang et al., 2012; Jiao et al., 2013).

Splicing is carried out by a macromolecular machinery called spliceosome, comprised of multiple small nuclear RNAs (snRNAs) and proteins that form highly dynamic interactions, directing the excision of intron sequences and the ligation of flanking exons (Will and Lührmann, 2011). In higher eukaryotes, splicing is often marked by the binding of a large multi-protein exon junction complex (EJC) at about 20–24 nucleotides (nt) upstream of the exon–exon junction (Le Hir et al., 2000). The EJCs accompany the mRNA to the cytoplasm and have roles in mRNA export, translation, and cytoplasmic quality control (Woodward et al., 2017; Schlautmann and Gehring, 2020) (see *2.3.3.1*). Splicing is linked to subsequent 3' end formation, for which most transcripts are endonucleolytically cleaved followed by the addition of a polyadenosine (poly(A)) tail by specialized enzymes and factors (Proudfoot, 2011). The poly(A) tail is rapidly bound by poly(A)-binding proteins, which provide control over the poly(A) tail length and protection against exonucleolytic attack (Eckmann et al., 2011; Schmid et al., 2015). Failure in proper splicing or 3' end processing predominantly leads to degradation of the transcript by the exosome, the major 3'–5' exonucleolytic machinery in the cell (**Figure 1c** and **1d**), although the mammalian XRN2 was also indicated to play a role (Davidson et al., 2012). Other errors in transcription elongation, premature transcription termination, and aberrations in mRNP packaging can also result in exosome-mediated degradation (Eberle and Visa, 2014; Bresson and Tollervey, 2018) (**Figure 1e**).

The exosome is a multi-subunit complex that is composed of a nine-protein, barrel-shaped catalytically inactive core and additional exonucleases that provide catalytic activities (Kilchert et al., 2016; Schmid and Jensen, 2019). In both the nucleus and the cytoplasm, the processive 3'–5' exo- and endonuclease Dis3 (also known as Rrp44) situates at the bottom of the exosome and degrades RNAs that have been threaded through the core. On the other hand, the distributive 3'–5' exonuclease Rrp6 binds at the top of the exosome exclusively in the nucleus. The activity of the exosome further relies heavily on different cofactors, depending on the pathway in which it is engaged. The main cofactor of the exosome in the nucleus is the Trf4/5–Air1/2–Mtr4 polyadenylation (TRAMP) complex.
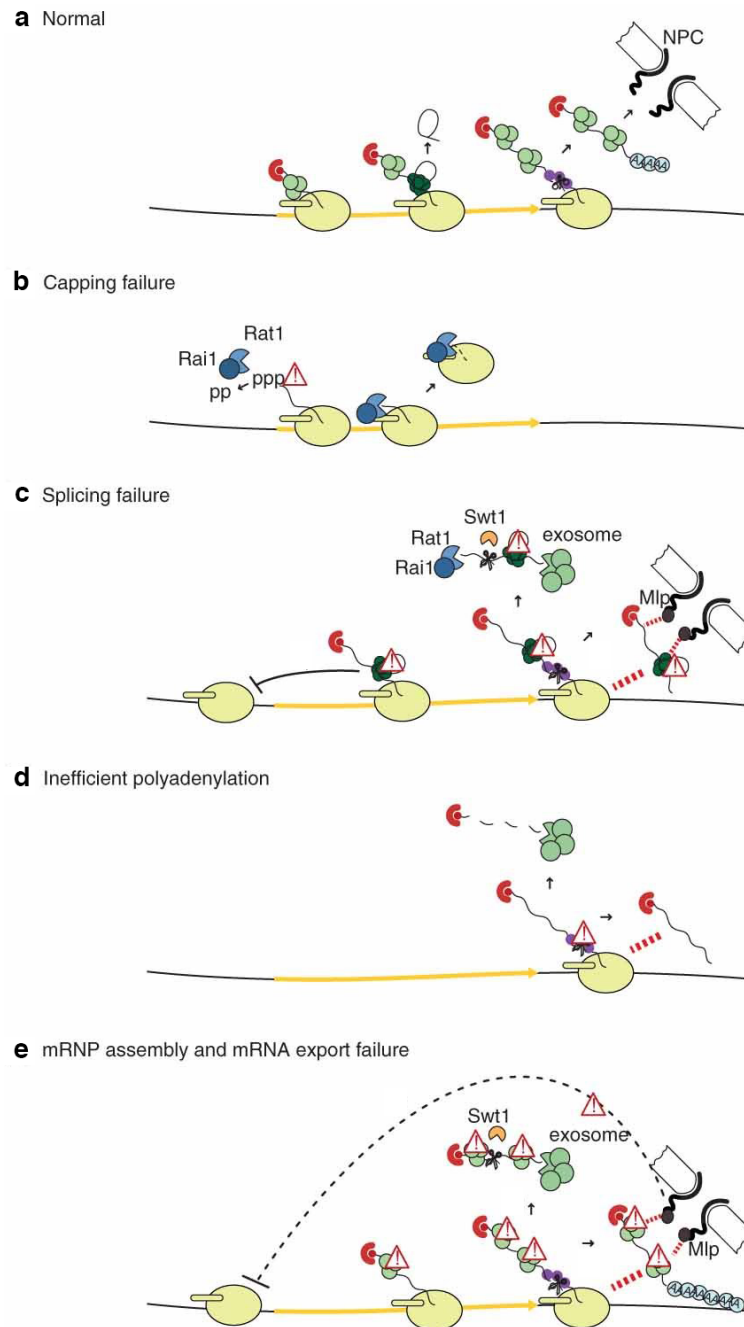
**Figure 1. mRNA quality control in the nucleus**

**a.** pre-mRNAs undergo co-transcriptional processing, including 5' capping, splicing, and 3' polyadenylation. These processes are coupled with dynamic recruitment and dissociation of a plethora of proteins that finally leads to the formation of a mature messenger ribonucleoprotein (mRNP) structure. Correctly processed and packaged mRNPs can be exported through the nuclear pore complex (NPC). **b.** Transcripts that are incorrectly capped are degraded by the 5'–3' exoribonuclease Rat1 with its cofactor Rai1. **c.** Defects in splicing lead to retention of the transcript at the NPC and degradation by the 3'–5' decay machinery exosome or the endonuclease Swt1. Degradation from the 5' end may also be involved. **d.** Problematic 3' end formation is targeted by the nuclear exosome. **e.** Proper mRNP assembly is quality controlled at the NPC by Mlp proteins. Incorrect mRNP formation leads to degradation by the nuclear exosome or Swt1. Red circle and half-circle: 5' cap and cap-binding protein; light green circles: export factors; dark green circles: spliceosome; purple circles: 3' processing factors; light blue circles: poly(A)-binding proteins; warning signs: defects in processing. (Adapted from Schmid and Jensen, 2010).

It associates on the top of the exosome and helps to deliver RNA substrates to the active exonucleases. Trf4 (or its paralog Trf5) is thought to add a short oligo(A) tail to the RNA substrate, which promotes threading of the RNA 3' end into the exosome core by the RNA helicase Mtr4. This helicase activity is key to proper function of the exosome. Air2 (or its paralog Air1) is thought to stabilize the complex by binding to both Mtr4 and Trf4, but was also shown to contribute to the direct binding of some RNA substrates. In yeast, another cofactor of the nuclear exosome is the Nrd1–Nab3–Sen1 (NNS) complex, which is closely linked to RNA decay during alternative pathways of transcription termination (Rondón et al., 2009; Ghazal et al., 2009; Schulz et al., 2013).

How the 5' and 3' exonucleases selectively target aberrant transcripts is not completely understood. An emerging model suggests that decay is the default pathway, which is kinetically competed against by proper processing, mRNP assembly, and timely export of the mRNA (Jensen et al., 2003; Saguez et al., 2005; Eberle and Visa, 2014; Bresson and Tollervey, 2018). Recent research in the yeast *Saccharomyces cerevisiae* proposed novel functions for a group of RNA-binding proteins that support this model (Zander and Krebber, 2017). These shuttling proteins, including Nab2, Npl3, Gbp2, and Hrb1, are recruited to the RNA co-transcriptionally and earlier identified as adaptors for the export receptor heterodimer, Mex67–Mtr2, linking transcription to export (Stewart, 2010; Soheilypour and Mofrad, 2018). It was then observed that deletion of these proteins resulted in the leakage of potentially faulty transcripts into the cytoplasm, which are usually retained and degraded in the nucleus in wild-type cells (Hackmann et al., 2014; Zander et al., 2016). This suggests that the proteins may serve a quality control function prior to supporting export. Although the exact mechanisms of each protein require further investigation, there is accumulating evidence showing that they promote recruitment of degradation machineries to their bound transcripts when processing is defective (Hackmann et al, 2014; unpublished data, laboratory of Heike Krebber). When processing is carried out properly, they would bind instead to export receptors to facilitate export. Thus, the quality control function of these proteins may provide a molecular basis to the competition between mRNA decay and continuation in downstream biogenesis events (see also **2.2.1**).

Before transport into the cytoplasm, the mRNP undergoes a final quality control step at the NPC (**Figure 1c** and **1e**). Several NPC-associated proteins were found to play a role, and one of the best characterized factors among them is Mlp1, a yeast protein localized at the nuclear face of the NPC (Soheilypour and Mofrad, 2018). Mlp1 was shown to retain

unspliced intron-containing transcripts in the nucleus through a mechanism that depends on the 5' splice site (Galy et al., 2004). The mammalian homolog, TPR, was also reported to regulate export of unspliced transcripts (Coyle et al., 2011; Rajanala and Nandicoori, 2012). In addition, Mlp1 physically interacts with the export receptor Mex67 and its adaptor proteins, which is thought to monitor completion of mRNA maturation, retaining those transcripts that are not properly packaged, hence likely not correctly processed (Green, D.M. et al., 2003; Galy et al., 2004; Vinciguerra et al., 2005; Hackmann et al., 2014). Binding of export receptors is necessary for export as they coat the highly-charged mRNA and interact with the hydrophobic meshwork of phenylalanine–glycine (FG) nucleoporins in the interior of the NPC, facilitating passage of the mRNP macromolecule (Paci et al., 2021). Aberrant transcripts that are retained are degraded by the TRAMP– exosome machinery. In yeast, an endonuclease, Swt1, associates with the NPC and likely also plays a role in degrading defective transcripts (Skruzný et al., 2009; Fasken and Corbett, 2009).

The mRNA biogenesis process in the nucleus fascinatingly shows how the many different events, from transcription to export, are closely connected and that each step is coupled with quality control. An upstream defect, which signals surveillance, often results in downstream defects that activate additional levels of quality control. This intricate network allows the process to be constantly monitored and prevents faulty transcripts from entering the cytoplasm, sparing the translation machinery from wasted production of potentially harmful proteins. Notably, mRNA-binding proteins that participate at multiple steps throughout the pathway appear to have a central role in providing communication between the events.

### 2.1.2. Cytoplasmic mRNA quality control

Despite nuclear quality control, mRNAs may encounter damage in the cytoplasm or have errors within their coding sequences. These defects can give rise to aberration during the translation process, resulting in altered dynamics or even stalling and collision of the decoding ribosome(s) (Powers et al., 2020; D'Orazio and Green, 2021). To cope with these problematic situations, several cytoplasmic mRNA quality control mechanisms have evolved.

On transcripts with strong secondary structures, a stretch of rare codons, or when the cell is deficient in certain aminoacyl-tRNAs, the ribosome may stall in the open reading frame (ORF) and has to be resolved by the no-go mRNA decay (NGD) pathway. This involves

endonucleolytic cleavage of the transcript, likely by a protein called Cue2 (Doma and Parker, 2006; D'Orazio et al., 2019), and recruitment of the proteins Dom34–Hbs1 (PELOTA–HBS1L in human) to dissociate the ribosome (Shoemaker et al., 2010; Pisareva et al., 2011) (**Figure 2**). Dom34–Hbs1 are structurally similar to canonical termination factors eRF1–eRF3 (Graille et al., 2008; Kobayashi et al., 2010) and both act together with Rli1 (ABCE1 in human) to separate ribosomal subunits (Pisarev et al., 2010; Barthelme et al., 2011; Shoemaker and Green, 2011) (see also **2.3.1**). The transcript is thereafter degraded by 5' and 3' exonucleases.
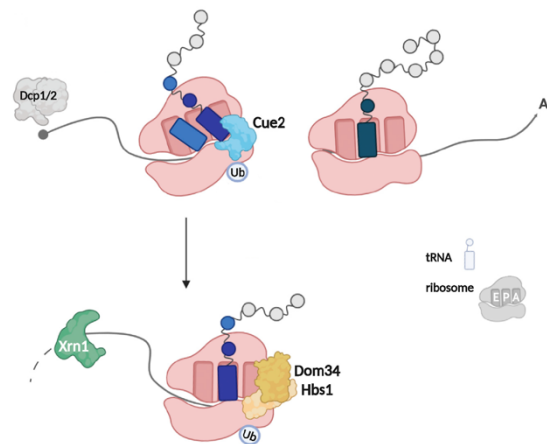


**Figure 2. No-go mRNA decay**
Ribosomes that stall in the open reading frame (ORF) are rescued by the no-go mRNA decay pathway. The transcript is endonucleolytically cleaved by Cue2 and subsequently degraded from the 5' and 3' ends. Ribosome recycling requires Dom34 and Hbs1, factors that are structurally similar to canonical translation termination factors eRF1 and eRF3. (Adapted from D'Orazio and Green, 2021).

Alternatively, transcripts may lack an in-frame stop codon, which can arise from aberrant polyadenylation or a mutation that disrupts the original stop codon. The ribosome translates into the poly(A) tail and eventually stalls at the 3' end of the mRNA. This activates nonstop mRNA decay (NSD), which relies strongly on the cytoplasmic exosome cofactor, the SKI complex, for ribosome rescue and subsequent RNA degradation by the exosome (van Hoof et al., 2002; Zinoviev et al., 2020) (**Figure 3**). Dom34–Hbs1 and Rli1 may then be used to separate the ribosomal subunits. In both NGD and NSD, the nascent polypeptide, which remains associated with the large ribosomal subunit after subunit splitting, is targeted by the ribosome-associated quality control (RQC) system and removed through the ubiquitylation–proteasome pathway (Joazeiro, 2019).
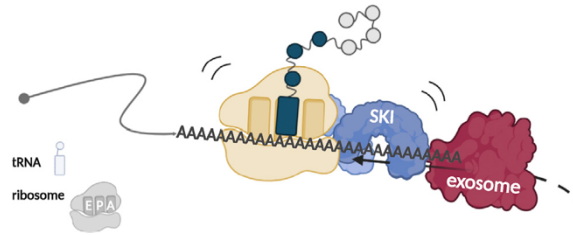
**Figure 3. Nonstop mRNA decay**
Ribosomes that translate into the poly(A) tail due to lack of an in-frame stop codon are rescued by the cytoplasmic exosome cofactor SKI complex in nonstop mRNA decay. The transcript is degraded by the exosome and ribosome splitting likely depends on Dom34–Hbs1 (not depicted). (Adapted from D'Orazio and Green, 2021).

Another type of error is a premature termination codon (PTC), which may arise from nonsense mutations or defects in splicing that escaped nuclear quality control. Translation of PTC-containing transcripts usually results in truncated polypeptides that are potentially toxic to the cell. Termination at a PTC thus activates the nonsense-mediated mRNA decay (NMD) pathway to repress further translation and eliminate the transcript (Kurosaki et al., 2019) (see **2.3**). Together, the cytoplasmic quality control pathways delete error-prone transcripts and release stalled ribosomes so that they can re-engage in translation of functional proteins.

## 2.2. The yeast shuttling proteins Gbp2 and Hrb1

As mentioned previously, Gbp2 and Hrb1 are two of the yeast RNA-binding proteins that help to couple transcription to mRNA export. Gbp2 (G-strand binding protein) was first identified, as its name suggests, as a telomeric G-strand binding protein and implicated in telomere maintenance (Lin and Zakian, 1994; Konkel et al., 1995; Pang et al., 2003). Hrb1 (Hypothetical RNA-binding protein) is a paralog of Gbp2 (Byrne and Wolfe, 2005), and both proteins consist of three RNA-recognition motifs (RRMs) as well as an SR/RGG domain at the N-terminus, which is rich in serine–arginine or arginine–glycine–glycine motifs (Windgassen and Krebber, 2003; Häcker and Krebber, 2004) (**Figure 4**). A third yeast protein, Npl3, is also highly homologous and contains both RRM and SR domains (Bossie et al., 1992; Häcker and Krebber, 2004). Interestingly, these domain features resemble that of human SR proteins, a family of proteins typified by the presence of one or two N-terminal RRM domains and a C-terminal SR/RS domain (Birney et al., 1993; Wegener and Müller-McNicoll, 2019). Human SR proteins were initially discovered as key

factors for constitutive and alternative pre-mRNA splicing, but in following research found to have more diverse roles in RNA metabolism. Gbp2, Hrb1, and Npl3, the three SR-like proteins known in yeast, were thus investigated and found to also participate in several RNA metabolic pathways (Lee et al., 1996; Windgassen et al., 2004; Bucheli and Buratowski, 2005; Dermody et al., 2008; Kress et al., 2008; Estrella et al., 2009). In general, SR proteins bind to RNA with their RRMs and can mediate protein–protein interactions at the same time through their SR/RS domains. Further, the SR/RS motifs are (putative) sites for post-translational modifications, providing a potential mechanism for regulation (Wegener and Müller-McNicoll, 2019). These characteristics make them ideal factors to facilitate the complex and dynamic RNA–protein or protein–protein interactions that occur in the cell, which has indeed been demonstrated in many reports.
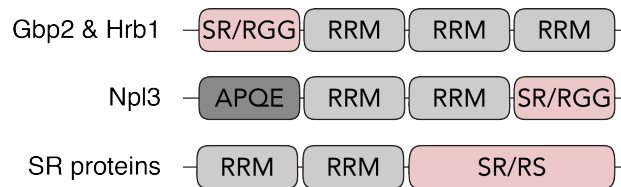


**Figure 4. Domain structures of yeast SR-like proteins and classical SR proteins**
The yeast SR-like proteins Gbp2, Hrb1, and Npl3 have domain structures similar to classical SR proteins in higher eukaryotes. They each consist of at least one RNA-recognition motif (RRM) and an SR domain that is rich in serine–arginine repeats. The SR domains of the yeast SR-like proteins also contain several arginine–glycine–glycine (RGG) motifs. Note that the domains are not depicted in proportion to their actual sizes. (Adapted from Häcker and Krebber, 2004).

### 2.2.1. The nuclear phase of Gbp2 and Hrb1

Gbp2 and Hrb1 join the journey of mRNAs in the nucleus, where they are co-transcriptionally recruited. This was discovered by the observation that Gbp2 and Hrb1 genetically and physically interact with the RNA polymerase II C-terminal domain (CTD)-kinase Ctk1 and specifically associate with components of the transcription/export (TREX) complex (Hurt et al., 2004; Meinel et al., 2013; Xie et al., 2021b). Ctk1 phosphorylates serine 2 of the CTD and is required for transcription elongation (Cho et al., 2001). The TREX complex, comprising the THO complex (Tho2, Hpr1, Mft1, Thp2, Tex1), the DEAD-box helicase Sub2, and export factor Yra1, is recruited to phosphorylated CTD and functions to connect transcription with mRNA export by recruiting appropriate factors for the formation of an export competent mRNP (Meinel and Sträßer, 2015). Recently, biochemical and structural studies on the THO/TREX complex confirmed previous

observations and revealed that the SR/RS as well as the RRM domains of Gbp2 are required for its interaction with multiple domains of the THO complex (Martínez-Lumbreras et al., 2016; Xie et al., 2021a).

The co-transcriptional recruitment of Gbp2 and Hrb1 is thought to contribute to transcription-coupled export, supported by later findings that Gbp2 and Hrb1 bind to the mRNA export receptor Mex67, require it for their export (Windgassen and Krebber, 2003; Häcker and Krebber, 2004; Hackmann et al., 2014), and that other export receptor adaptor proteins are also co-transcriptionally recruited, although through different pathways. In particular, Npl3 is recruited earlier by RNA polymerase II and associate both with the 5' and 3' ends of the transcript, while Nab2 is recruited later and binds to the poly(A) tail (Lei et al., 2001; Green et al., 2002; Wende et al., 2019).

Subsequent research indicated a role for Gbp2 and Hrb1 in quality control of splicing in addition to supporting export. First, it was observed that the co-transcriptional recruitment of Gbp2 and Hrb1 is also closely linked to splicing. Several reports have demonstrated that recruitment and stable binding of the THO/TREX complex is connected to splicing, both in yeast and in human (Abruzzi et al., 2004; Masuda et al., 2005; Lardelli et al., 2010; Chanarat et al., 2011; Gromadzka et al., 2016). It was then shown that Gbp2 and Hrb1 co-purify with the spliceosome, specifically with the late splicing factors Prp17 and Prp43, and their binding to mRNAs depends on functional splicing (Warkocki et al., 2009; Hackmann et al., 2014). This is further supported by transcriptome-wide analyses: the two proteins, in particular Gbp2, associate preferentially with transcripts that derived from intron-containing genes (Hackmann et al., 2014), and a transcriptome-wide binding profile of Gbp2 showed that it binds mRNAs mostly at the 5' proximal region (Tuck and Tollervey, 2013), which corresponds to the position of most yeast introns (Mourier and Jeffares, 2003; Neuvéglise et al., 2011). In another analysis using the PAR-CLIP method, Gbp2 showed distributed binding on mRNAs throughout the ORF, while Hrb1 showed a higher tendency to bind toward the 5' end (Baejen et al., 2014).

Functional splicing is further linked to export of Gbp2 and Hrb1 with the mRNA, as binding of these proteins to Mex67 and their export to the cytoplasm were disrupted when splicing factor genes were mutated (Hackmann et al., 2014). Interestingly, following research revealed that Gbp2 and Hrb1 genetically and/or physically interact with several factors involved in the nuclear quality control pathway: the helicase Mtr4 of the TRAMP complex, the 3'–5' exonuclease Rrp6 of the nuclear exosome, and Mlp1 and Mlp2, two quality control factors at the NPC (Hackmann et al., 2014; Bretes et al., 2014). These findings led

to the idea that Gbp2 and Hrb1 may function in nuclear quality control, preventing transcripts that have not gone through proper splicing from being exported, and potentially facilitating their degradation by the TRAMP–exosome machinery. This was tested by monitoring export of aberrant transcripts through *in situ* hybridization experiments as well as qPCR analyses following cell fractionation. The results showed that faulty, intron-containing transcripts, which accumulated in the nucleus when *MTR4* or *RRP6* was deleted, were increasingly transported into the cytoplasm when Gbp2 or Hrb1 was also absent (Hackmann et al., 2014). Furthermore, Gbp2 and Hrb1 showed increased binding to faulty transcripts when *RRP6* is mutated or when *MLP1* is deleted. These results demonstrate that Gbp2 and Hrb1 retain improperly spliced transcripts in the nucleus for surveillance. In line with this, deletion of *GBP2* or *HRB1* together with late splicing factor genes is toxic, while a combination of three deletions is lethal to the cell.

A model was proposed for the nuclear quality control function of Gbp2 and Hrb1, suggesting that these proteins direct RNAs either towards decay or export. The model posits that Gbp2 and Hrb1 are recruited to the nascent RNA by the TREX complex and they monitor pre-mRNA splicing. When splicing is somehow defective, Gbp2 and Hrb1 would interact with Mtr4, which leads to the degradation of the aberrant transcript by the nuclear exosome. If splicing is carried out correctly, Gbp2 and Hrb1 would recruit instead Mex67–Mtr2 through interactions with Mex67, preventing degradation and promoting export. This is supported by studies showing that the presence of Gbp2 and Hrb1 is required for the association of Mtr4, and prevent the association of Mex67, with unspliced transcripts (Hackmann et al., 2014). Surveillance occurs prior to export, as interactions between Gbp2, Hrb1 and Mex67 or Mlp1 is impaired when *MTR4* is mutated. Finally, the binding of Mtr4 and Mex67 to Gbp2 and Hrb1 was found to be mutually exclusive, supporting the "decay or export" notion. Therefore, Gbp2 and Hrb1 act as quality control factors in the nucleus to help distinguish aberrant from correct splicing and direct transcripts to the proper downstream pathway – either retention and degradation or further processing and nuclear export. In this way, they link mRNA processing with surveillance and export, thereby contributing to a regulated system for the delivery of proper transcripts to the cytoplasm. Quality control of splicing by Gbp2 and Hrb1 is summarized in **Figure 5**.
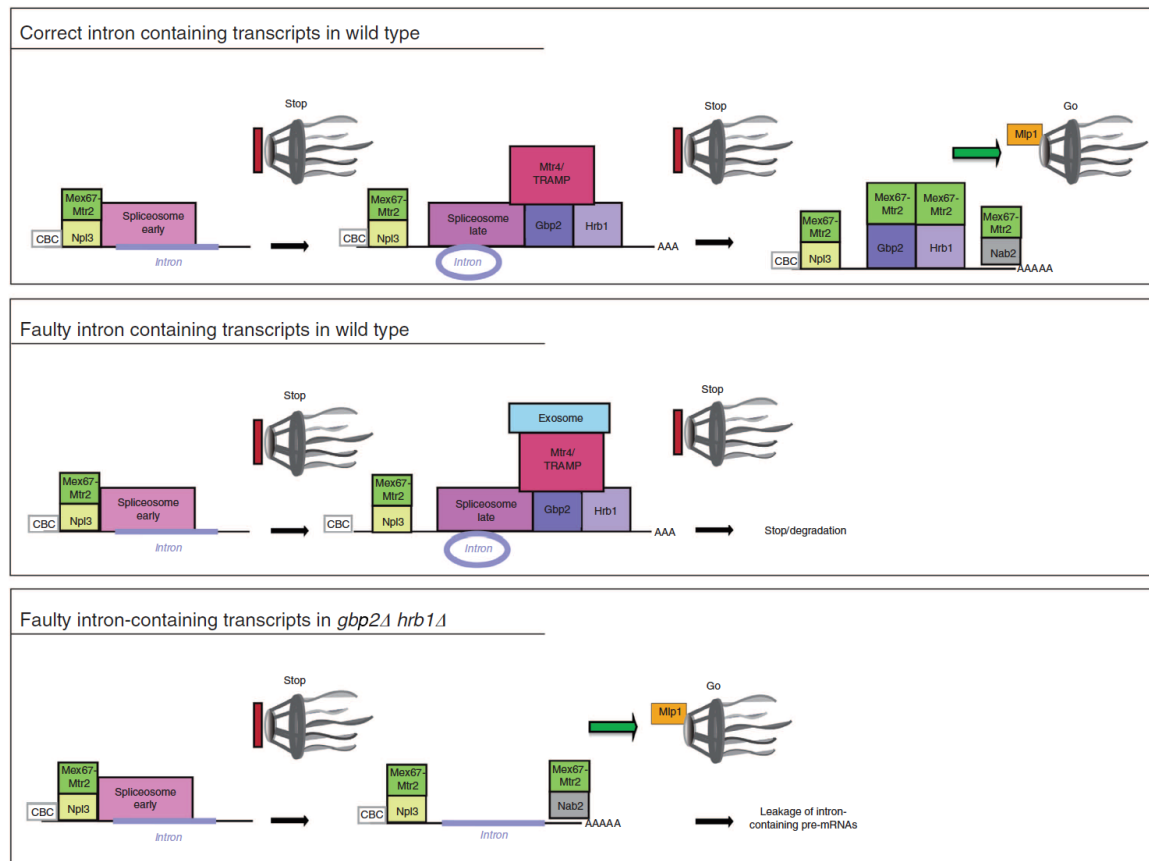
**Figure 5. Nuclear quality control of splicing by Gbp2 and Hrb1**
Gbp2 and Hrb1 are recruited to the mRNA during transcription and associate with the late spliceosome. Upon proper completion of splicing, they recruit the mRNA export receptor heterodimer Mex67–Mtr2 to support nuclear export. Mlp1 and other proteins at the NPC check for proper mRNP packaging. Transcripts bound by several molecules of Mex67–Mtr2 are efficiently transported through the NPC (top). In the case of aberrant splicing, Gbp2 and Hrb1 recruit the Trf4/5–Air1/2–Mtr4 polyadenylation (TRAMP) complex by interacting with Mtr4, which facilitates transcript degradation by the nuclear exosome (middle). When Gbp2 and Hrb1 are absent, incorrect splicing is not properly detected and faulty transcripts may escape quality control at the NPC and leak into the cytoplasm (bottom). (Adapted from Hackmann et al., 2014).

Notably, despite their involvement in quality control of splicing, the absence of Gbp2 and Hrb1 does not lead to splicing defects and therefore they are not bona fide splicing factors (Kress et al., 2008; Hackmann et al., 2014). In addition, deletion of these genes resulted in no significant defects in bulk mRNA localization, indicating that they are not essential for mRNA export (Hackmann et al., 2014). In fact, the overexpression of these proteins is toxic to the cell, likely due to the retention of bulk poly(A)$^+$ RNAs in the nucleus (Windgassen and Krebber, 2003; Häcker and Krebber, 2004; Zander et al., 2016; unpublished data, laboratory of Heike Krebber). The fact that overexpression of Gbp2 and Hrb1 leads to nuclear mRNA accumulation and that faulty intron-containing transcripts

that cannot be degraded leak into the cytoplasm when Gbp2 and Hrb1 are absent rather highlight the role of these proteins as nuclear mRNA quality control and retention factors.

It is perhaps also noteworthy to point out that under stress conditions, nuclear quality control by Gbp2 and Hrb1 is neglected on stress-responsive transcripts to allow their rapid export (Zander et al., 2016). In the case of heat stress, Gbp2, Hrb1, and other Mex67 adaptor proteins were shown to dissociate from bulk mRNAs, while Mex67 binds directly to heat-shock mRNAs to facilitate their export, demonstrating how cells may adapt cellular mechanisms to accommodate to change.

Similar functions in quality control have not yet been fully established for other proteins that recruit Mex67–Mtr2 in yeast. These proteins include Hpr1, Yra1, two components of the TREX complex, and Npl3, Nab2, all shown to directly bind Mex67–Mtr2 (Wende et al., 2019). They are recruited to the nascent transcript during different stages of transcription and associate primarily with different structures of the mRNA. Npl3, for example, is recruited at an early stage and has been shown to associate with the 5' cap (Shen et al., 2000; Lei et al., 2001). It has been implicated in many processes, including transcription elongation, splicing, 3' end processing, mRNA export, translation, and there are evidences supporting its quality control role for 5' capping (unpublished data, laboratory of Heike Krebber). Nab2 is the nuclear poly(A)-binding protein and functions in poly(A) tail length control in addition to mRNA export (Hector et al., 2002; Soucek et al., 2012). Interestingly, Nab2 associates with Rrp6 for 3' end control (Roth et al., 2009; Schmid et al., 2012; Soucek et al., 2016) and the exosome is involved in the autoregulation of the *NAB2* mRNA level through Nab2 itself (Roth et al., 2005). Whether these Mex67-recruiting proteins or other RNA-binding proteins that are part of the export mRNP have further quality control functions remains to be determined. In human cells, a competition between decay by the exosome and mRNA export was also indicated, and current knowledge suggests likewise a mutual exclusive association of the mRNA with exosome cofactors and export factors (Schmid and Jensen, 2019). The detail mechanisms and proteins that mediate this are however unclear.

Proper completion of splicing together with other mRNA maturation steps leads to the association of different export adaptor proteins and in turn the recruitment of several molecules of Mex67–Mtr2 along the transcript, allowing efficient export. Some proteins are released from the mRNP before or immediately after passage through the NPC, while others accompany the transcript into the cytoplasm. The two proteins that are associated with the TREX complex, Hpr1 and Yra1, are released from the mRNP prior to export

through mechanisms mediated by ubiquitylation (Gwizdek et al., 2006; Iglesias et al., 2010; Tutucci and Stutz, 2011). On the cytoplasmic side, the DEAD-box helicase Dbp5, tethered to the NPC through the nucleoporin Nup159 (Schmitt et al., 1999; Hodge et al., 1999; Weirich et al., 2004), triggers remodeling of the export mRNP by dissociating Mex67–Mtr2 and Nab2 (and potentially more proteins) from the mRNA (Lund and Guthrie, 2005; Tran et al., 2007; Xie and Ren, 2019). For this, energy from ATP hydrolysis is required, and Dbp5 relies on the additional factors Gle1 and inositol hexakisphosphate (IP$_6$) to stimulate its ATPase activity (York et al., 1999; Weirich et al., 2006; Alcázar-Román et al., 2006). Removal of export receptors and adaptor proteins at the cytoplasmic side of the NPC is thought to prevent the mRNP from diffusing back through the NPC, thus providing directionality for mRNA export (Stewart, 2010; Tieg and Krebber, 2013).

## 2.2.2. The cytoplasmic phase of Gbp2 and Hrb1

In contrast, Gbp2 and Hrb1, as well as Npl3, enter the cytoplasm along with the mRNAs and likely remain associated. They were found in polysome-containing fractions in sucrose density gradient fractionation experiments (Windgassen et al., 2004), suggesting that they are bound to the mRNA during translation. In comparison, Nab2 and Mex67, which are displaced by Dbp5 following export, were mainly present in ribosome-free fractions. While Npl3 has been found to be involved in translation (Windgassen et al., 2004; Estrella et al., 2009; Rajyaguru et al., 2012; Baierlein et al., 2013), whether Gbp2 and Hrb1 have additional functions there or in the cytoplasm in general is not known. Current understanding of their dissociation from the mRNA and subsequent nuclear import is also limited. It was discovered that nuclear re-import of these proteins depends on the import receptor Mtr10 and at least partially on their SR/RGG domains and the yeast SR protein kinase Sky1, a homolog of the human serine/arginine-rich protein-specific kinase SRPK1 (Siebel et al., 1999). Both Gbp2 and Hrb1 localize mainly to the nucleus at steady state, and deletion or mutation of *MTR10* or their SR/RGG domains rendered the mislocalization of these proteins to the cytoplasm (Windgassen and Krebber, 2003; Häcker and Krebber, 2004; unpublished data, laboratory of Heike Krebber). This phenotype could partially be rescued by overexpression of *MTR10*. Deletion of *SKY1* also caused Gbp2 to mislocalize to the cytoplasm, therefore it seems likely that phosphorylation of the Gbp2 SR domain by Sky1 affects its import, either directly or indirectly. Subsequent analyses on mRNA binding revealed that deletion or mutation of *MTR10* led to the increased binding of Gbp2 and Hrb1 to poly(A)$^+$ RNAs, whereas deletion of *SKY1* had a lesser effect, indicating that Mtr10 plays a more direct role in the

release of these proteins from mRNAs. Further details regarding how these events are regulated are however lacking. Interestingly, import of Npl3 involves a similar mechanism, which depends on both Mtr10 and Sky1. It was suggested that phosphorylation of Npl3 by Sky1 promotes its dissociation from mRNAs and subsequently its import (Yun and Fu, 2000; Gilbert et al., 2001), but some data indicate that, like for Gbp2 and Hrb1, Mtr10 is important for its release from mRNA and Sky1 acts afterwards (Windgassen et al., 2004).

The shuttling activity of Gbp2, Hrb1, and Npl3 is conserved in higher eukaryotes. The human SR protein family consists of 12 members (SRSF1–SRSF12), most of which were shown to remain associated with the mRNA after splicing and may shuttle between the nucleus and the cytoplasm (Singh et al., 2012; Müller-McNicoll et al., 2016; Wegener and Müller-McNicoll, 2019). Their shuttling activities vary and the differences are closely related to the phosphorylation state of their SR/RS domains (Cáceres et al., 1998; Cazalla et al., 2002; Lin et al., 2005). Three of the proteins, SRSF1, SRSF3, and SRSF7, were identified as prominent shuttling proteins and export adaptor proteins that directly bind to the export receptor NXF1 (human homolog of Mex67) (Cáceres et al., 1998; Huang et al., 2003; Lai and Tarn, 2004; Müller-McNicoll et al., 2016). Their shuttling depends on both their SR/RS and RRM domains. On the other hand, SRSF2 usually does not shuttle, but was observed to exhibit enhanced shuttling activity in murine pluripotent cells as a result of a lower phosphorylation level of the protein (Botti et al., 2017). Phosphorylation level of SR proteins therefore appears to serve as a mechanism to mediate selective export of transcripts (Huang and Steitz, 2005; Wegener and Müller-McNicoll, 2019). Interestingly, this was also implicated in the retention of improperly spliced transcripts in the nucleus by SR proteins, which prevent recruitment of NXF1 in their hyperphosphorylated states (Wegener and Müller-McNicoll, 2018). In the cytoplasm, human SR proteins associate with the translating ribosome and were suggested to play a role in translational regulation (Sanford et al., 2004; Howard and Sanford, 2015; Wegener and Müller-McNicoll, 2019). Finally, similar to the yeast SR-like proteins, SR proteins are phosphorylated by SRPK1/2 and re-imported into the nucleus by the transportin-SR (TRN-SR) proteins (Lai et al., 2001), the mammalian homolog of Mtr10 (Kataoka et al., 1999), which was suggested to occur after the SR proteins dissociate from the mRNA (Wegener and Müller-McNicoll, 2019).

## 2.3. Nonsense-mediated mRNA decay

Nonsense-mediated decay (NMD) is a well-conserved cytoplasmic mRNA quality control mechanism that has been found in species across the eukaryotic kingdom (Karousis and Mühlemann, 2019; Kurosaki et al., 2019). It was originally identified to function in detection and removal of RNAs that harbor premature termination codons (PTCs). Therefore, it was long thought to be a specific quality control pathway that targets only defective RNAs. However, later research revealed that NMD also targets a large portion of RNAs that have normal, complete open reading frames from which functional proteins can be made. Recently, studies both in yeast and human cells further demonstrated that the examined non-coding RNA species are also regulated by NMD (Tani et al., 2013; Wery et al., 2016; de Andres-Pablo et al., 2017). Thus, NMD has functions beyond quality control, and is now generally considered as a mechanism for post-transcriptional gene regulation (Nickless et al., 2017; Nasif et al., 2018).

The best-known NMD targets are transcripts that contain PTCs. This could arise from mutations during transcription (nonsense mutations), errors during splicing, or, in higher eukaryotes, regulated alternative splicing. Reports have shown that NMD is responsible for removing transcripts that are suboptimally spliced both in yeast and higher eukaryotes (Jaillon et al., 2008; Sayani et al., 2008). Further, alternative splicing was estimated to commonly result in PTCs (Lewis et al., 2003; Green, R.E. et al., 2003), and its coupling to NMD has been observed in several different species (Hansen et al., 2009; Barberan-Soler et al., 2009; McIlwain et al., 2010; García-Moreno and Romão, 2020). The link between alternative splicing and NMD is generally considered as a mechanism to regulate the abundance of affected mRNAs in addition to surveillance, including many that encode factors involved in the splicing process itself. Besides PTCs, a codon coding for selenocysteine (UGA) might be interpreted as a stop codon when the cellular concentration of selenocysteine is low, causing translation to terminate earlier and may thus activate NMD (Moriarty et al., 1998; Karousis and Mühlemann, 2019; Kurosaki et al., 2019). Transcripts that contain an upstream open reading frame (uORF) can also result in translation termination that can be recognized as premature (Hurt et al., 2013). In addition, it was found that NMD is preferentially initiated on transcripts with a long 3' untranslated region (UTR), which can occur naturally or through regulated selection of alternative polyadenylation sites. A recent transcriptome-wide study revealed that human transcripts with long 3' UTRs resulting from alternative cleavage and polyadenylation are globally targeted by NMD (Kishor et al., 2020). Apart from these, error-free transcripts without any

NMD-activating features and non-coding RNAs can also be NMD targets. Recently, a study in yeast demonstrated that NMD targets lacking PTCs show a higher rate of out-of-frame translation, which could result in PTCs and might explain why they are recognized as targets (Celik et al., 2017a; Celik et al., 2017b). In short, NMD affects a broad range of targets, demonstrating its widespread impact on the transcriptome.

The function of NMD in control of mRNA quality and gene expression becomes critical as early as in cell development and differentiation, and is important for the immediate response of cells to a variety of environmental stimuli and/or stress. Although NMD is not essential for the yeast *S. cerevisiae*, non-functional NMD has been shown to result in developmental abnormalities for all organisms examined and embryonic lethality for fruit flies and mice (Hwang and Maquat, 2011; Kurosaki et al., 2019). In fact, NMD is linked to many human diseases, including cancer, neurodegenerative, and inherited disorders. Moreover, viral infection can be defended through NMD, while viruses have also developed various strategies to evade this surveillance system (Pawlicka et al., 2020; Leon and Ott, 2021; May and Simon, 2021). NMD has thus been the topic of much research over the years. Nonetheless, the discovery of contradictory results and the difficulty to study this pathway in a step-by-step manner have made our thorough understanding of its detailed mechanism challenging.

### 2.3.1. Eukaryotic translation

NMD is a translation-dependent pathway, and is intimately linked to translation termination, which in turn affects translation initiation. In this section, some general aspects of eukaryotic translation are described, focusing on factors associated with initiation and termination, especially those that may be involved in NMD and thus will be relevant to this work.

The majority of eukaryotic translation initiation relies on a $m^7G$ cap-dependent mechanism (Jackson et al., 2010; Hinnebusch and Lorsch, 2012; Merrick and Pavitt, 2018). As mentioned previously, the 5' end of mature mRNAs are protected by the $m^7G$ cap and cap-binding proteins. In the cytoplasm, the cap is recognized and bound by the translation initiation factor eIF4E. eIF4E is under regulation of eIF4E-binding proteins (4E-BPs), which prevents it from binding to eIF4G. When 4E-BPs dissociate, eIF4E can bind eIF4G, resulting in an increase in its affinity to the $m^7G$ cap. Exactly when and how the cap-binding complex (CBC) recruited in the nucleus is replaced by eIF4E is so far unclear. eIF4E, eIF4G, and a third factor, eIF4A, are components of the eIF4F complex, which

associates with the cap to facilitate translation initiation. eIF4G is a large scaffold protein that mediates interactions with eIF4E, eIF4A, the mRNA, and poly(A)-binding proteins, forming a circularized structure of the mRNA that is thought to facilitate efficient translation (see below). eIF4A is an RNA helicase that removes secondary structures or binding proteins near the 5' end of the mRNA to prepare a stretch of single stranded mRNA that can fit into the cleft of the small ribosomal subunit (40S).

Prior to binding mRNA, the 40S subunit is joined by the initiator methionyl-transfer RNA (Met-tRNAi) in the P site. Several other initiation factors are also associated and together they make up the 43S pre-initiation complex (PIC) (**Figure 6**). The 43S PIC binds to the eIF4F-bound mRNA at, or in the vicinity of, the $m^7G$ cap. Through mediation by different factors and at the expense of energy from ATP hydrolysis, the 43S PIC subsequently scans the mRNA in the 3' direction until an AUG start codon is recognized and base-pairs with the anticodon of Met-tRNAi. This triggers a reorganization of the initiation complex, with the release and joining of different factors as well as recruitment of the large ribosomal subunit (60S). The 80S ribosome is formed and ready for translation elongation, which begins when a corresponding aminoacyl-tRNA binds in the ribosome A site, base-pairing with the next codon (Jackson et al., 2010; Hinnebusch and Lorsch, 2012; Merrick and Pavitt, 2018).
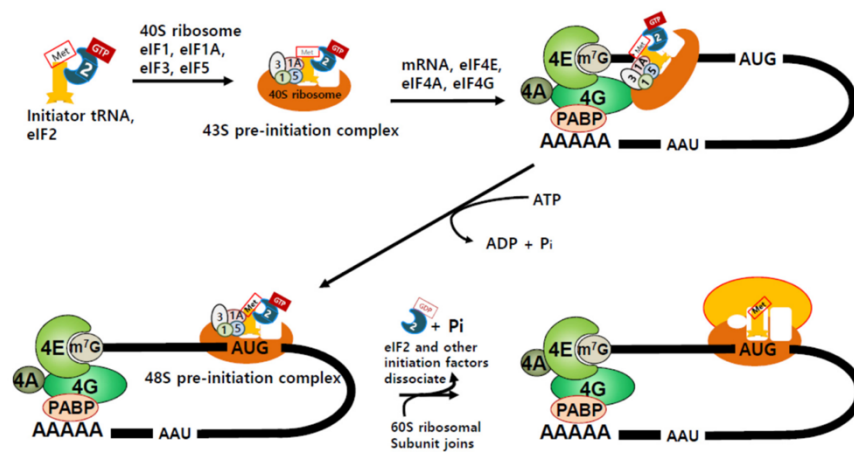


**Figure 6. Eukaryotic cap-dependent translation initiation**
The initiation factor eIF2-GTP associates with the initiator methionyl-tRNA (Met-tRNAi) and together they join the small ribosomal subunit (40S) with several additional initiation factors, forming the 43S pre-initiation complex (PIC). The 43S PIC is recruited to the 5' cap of the mRNA, bound by the eIF4F complex consisting of cap-binding protein eIF4E, scaffold protein eIF4G, and RNA helicase eIF4A. eIF4G also interacts with the poly(A)-binding protein (PABP), circularizing the mRNA to support efficient translation. With energy from ATP hydrolysis, the 43S PIC scans downstream for the first AUG start codon. Upon recognition of the start codon, several initiation factors dissociate and the large ribosomal subunit (60S) joins to form the complete, translating 80S ribosome. (Adapted from Kim, 2019).

Translation termination is initiated when the codon that enters the A site is a termination codon. This is recognized by the release factor eRF1, which is mainly facilitated by the GTP-binding protein eRF3 (Alkalaeva et al., 2006; Hellen, 2018). It is generally thought that eRF1 and eRF3 in its GTP-bound state interact to form an inactive ternary complex already before contacting the ribosome (**Figure 7a**). They enter the ribosome, where recognition of the stop codon by eRF1 as well as the proteins' multiple interactions with the ribosomal subunits stimulate GTP hydrolysis of eRF3. This in turn results in conformational changes that allow eRF1 to catalyze peptidyl-tRNA hydrolysis, leading to peptide release (Frolova et al., 1999; Song et al., 2000). eRF1 and eRF3 are thought to promote each other's activities. eRF1 promotes binding of GTP to eRF3 and later its hydrolysis by eRF3 (Hauryliuk et al., 2006; Cheng et al., 2009; Wada and Ito, 2014); eRF3 may support stop codon recognition and its GTP hydrolysis enhances the peptide-release activity of eRF1 (Alkalaeva et al., 2006; Eyler et al., 2013).
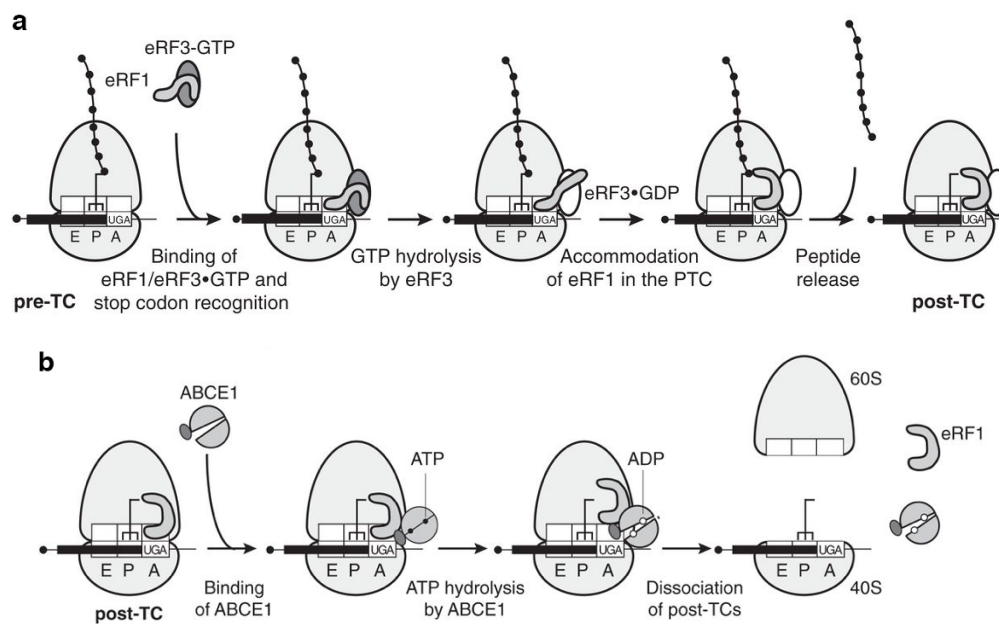


**Figure 7. Eukaryotic translation termination**
**a.** The inactive release factor eRF1–eRF3-GTP complex is recruited to the stop codon at the ribosome A site. GTP hydrolysis by eRF3 is triggered, which leads to a conformational change in eRF1 that promotes the hydrolysis of the peptidyl-tRNA bond. **b.** After peptide release, ABCE1 binds to eRF1. Energy from ATP hydrolysis by ABCE1 is used to split the ribosomal subunits. Pre-TC: pre-termination complex; post-TC: post-termination complex. (Adapted from Hellen, 2018).

A study performed in yeast *S. cerevisiae* suggested a different model regarding the ribosome entry of these release factors (Beißel et al., 2019). Their results indicated that eRF3 is already in association with the ribosome when eRF1, accompanied by the ATP-bound DEAD-box helicase Dbp5, joins (**Figure 8**). ATP hydrolysis and subsequent dissociation of Dbp5-ADP would allow eRF1 to interact with eRF3, which then leads to GTP hydrolysis of eRF3 and following peptide release. The involvement of Dbp5 was proposed to provide control over the eRF1–eRF3 interaction and prevent premature binding before proper positioning of the proteins.

After the peptide is released, the ribosome needs to be disassembled before the subunits can engage in a new round of translation. This requires the protein Rli1 (ABCE1 in human), which interacts with eRF1 and binds in the inter-subunit space of the ribosome, splitting the subunits through conformational changes that is driven by the energy of ATP hydrolysis (**Figure 7b** and **8**). Binding of Rli1/ABCE1 requires that eRF3 is previously dissociated from the ribosome. The ribosomal subunits are separated and, after release of the mRNA and tRNA from the 40S subunit, are ready for further rounds of translation. Rli1/ABCE1 is also involved in ribosome splitting in cytoplasmic quality control pathways NGD and NSD (see **2.1.2**).
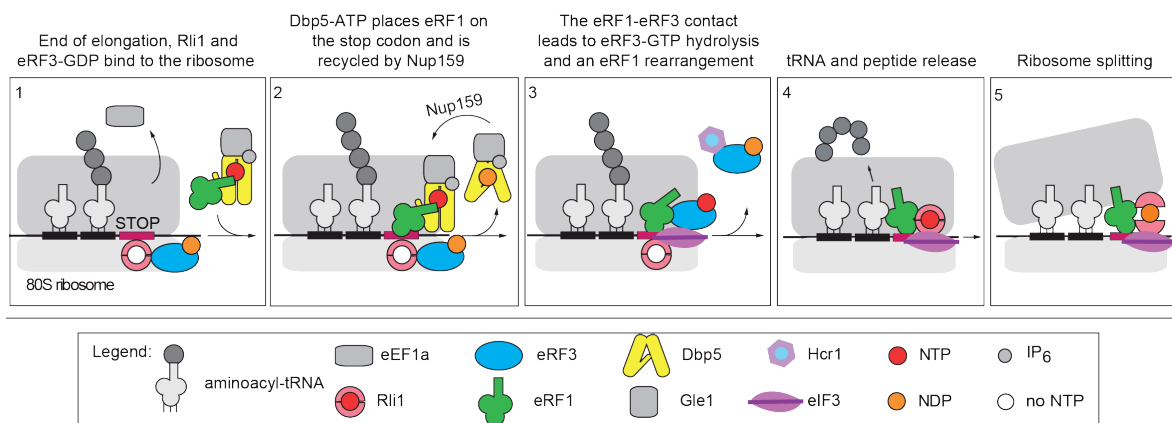


**Figure 8. Dbp5-mediated translation termination**
Rli1 and GDP-bound eRF3 are associated with the terminating ribosome. ATP-bound Dbp5 subsequently delivers eRF1 to the stop codon. ATP hydrolysis and dissociation of Dbp5-ADP allow proper interaction between eRF1 and eRF3. GTP is then recruited to eRF3 and its hydrolysis as well as dissociation of eRF3-GDP promotes peptide release. Finally, ATP-associated Rli1 interacts with eRF1 and separates ribosomal subunits to recycle the ribosome. (Adapted from Beißel et al., 2019).

Regulation of translation can occur through different mechanisms. The cytoplasmic poly(A)-binding protein Pab1 (PABPC1 in human), for example, was shown to interact with eRF3 and positively regulate termination (Uchida et al., 2002; Cosson et al., 2002; Ivanov et al., 2008; Ivanov et al., 2016). In addition, contact between the 5' and 3' ends of a transcript is thought to play a role in regulating translation (Jackson et al., 2010; Hinnebusch and Lorsch, 2012; Merrick and Pavitt, 2018). Several reports have suggested that the mRNA is circularized during translation, forming a so-called "closed-loop" structure that may promote the recruitment of the 43S PIC. This was supported by physical interactions demonstrated between the cap-binding eIF4E, eIF4G and PABPC1. The closed-loop formation was also suggested to accelerate translation re-initiation by (1) tethering eIF4F to the 5' end, thereby neglecting the need for it to re-associate for each round of translation and (2) enhancing re-binding of freed ribosomal subunits to the 5' end after termination, due to the close proximity. The idea that transcript circularization is a general mechanism in translation was however challenged by later findings. In recent research it was observed that not all mRNAs are translated in a closed-loop form and this structure was suggested to either be unstable or occurs according to transcript type and/or cellular condition (Adivarahan et al., 2018; Khong and Parker, 2018).

It is generally believed that newly exported mRNAs undergo a pioneer round of translation before they are engaged in the more efficient, steady-state translation described above (Maquat et al., 2010b). The major difference between the two translation pathways lies in the mRNP composition, as in the pioneer round the transcript is bound by the CBC and other RNA-binding proteins, including EJCs, which is replaced by the eIF4F complex and are removed, respectively, before steady-state translation begins. Despite these differences, the two pathways rely on mostly the same translation factors, including eIF4G and eRF1–eRF3. The main purpose of the less-efficient, pioneer round translation is thought to be its coupling to quality control, supported by the findings that EJCs as well as the CBC promote activation of NMD when the transcript harbors a PTC (see *2.3.3.1*).

## 2.3.2. Core factors of NMD – the UPF proteins

The complete NMD pathway is carried out by multiple protein factors specific to NMD as well as factors of the general degradation pathways. In particular, three proteins were found to be conserved in all species tested for NMD (Gupta and Li, 2018; Kurosaki et al., 2019). These are the Upf (Up-frameshift) proteins – Upf1, Upf2, and Upf3 – which form a complex at the PTC-associated ribosome upon NMD target recognition. Among them,

Upf1 is the main factor that, with the support of Upf2 and Upf3, modulates target recognition and degradation. These proteins play specific and important roles in NMD and are thus considered as the core factors of NMD.

Upf1 is an ATPase-dependent RNA helicase that belongs to the Superfamily (SF)-1 helicases (Fairman-Williams et al., 2010). It comprises an amino-terminal cysteine/histidine (CH)-rich domain and a carboxy-terminal ATPase/RNA helicase domain. Upf1 was shown in human cells to bind to all transcripts promiscuously at random positions, but dissociates with a higher probability from non-targets, depending on its ATPase activity (Kurosaki et al., 2014; Lee et al., 2015). This correlates with observations across species that Upf1 is significantly enriched on PTC-containing mRNAs in comparison to PTC-less transcripts at steady state (Johns et al., 2007; Johansson et al., 2007; Hwang et al., 2010). Upf1 was also shown to be actively removed from the transcript by the translating ribosome, resulting in its relative enrichment in the 3' UTR of RNAs, which was suggested to have an impact on NMD activation (Hogg and Goff, 2010; Hurt et al., 2013; Kurosaki and Maquat, 2013; Zünd et al., 2013) (see *2.3.3.2*). In NMD, Upf1 is known to bind to the translation termination complex at the stop codon through interactions with the release factors eRF1 and eRF3 (Czaplinski et al., 1998; Ivanov et al., 2008). The CH domain of yeast Upf1 was also found to interact with the small ribosomal subunit protein Rps26, with implications in NMD initiation as well as ribosome release (Ghosh et al., 2010; Min et al., 2013).

The association of Upf1 with the terminating ribosome is the first step and the prerequisite for activation of NMD. Upon NMD activation, Upf1 uses energy from ATP hydrolysis to drive its helicase activity and translocate in the 5' to 3' direction on the single-stranded target RNA (Czaplinski et al., 1995; Weng et al., 1996a; Bhattacharya et al., 2000; Shigeoka et al., 2012). This helicase activity is critical for proper completion of NMD and has been implicated in multiple steps of the process. It was shown to be important for NMD target discrimination (Kurosaki et al., 2014; Lee et al., 2015), to enhance translation termination at a PTC (Weng et al., 1998; Serdar et al., 2016), and to unwind secondary RNA structures and dissociate bound proteins, including the ribosome, to allow complete degradation of the transcript (Franks et al., 2010; Fiorini et al., 2015; Serdar et al., 2016; Serdar et al., 2020). Moreover, Upf1 forms interactions with several other NMD and degradation factors, serving as a platform for the recruitment of these proteins (Karousis and Mühlemann, 2019; Kurosaki et al., 2019). Notably, the cellular abundance of Upf1 is several folds higher than that of Upf2 and Upf3. This on the one hand highlights it as the

main factor of NMD, while on the other hand indicates that it may have functions independent of the other two Upf proteins. Indeed, it has been observed in higher eukaryotes that some NMD events are Upf2- or Upf3-independent, and Upf1 has been discovered to participate in additional pathways that may not require the other Upf proteins (Kim and Maquat, 2019; Lavysh and Neu-Yilik, 2020).

Upf2 acts as a bridge in the Upf1–Upf2–Upf3 complex by interacting with both Upf1 and Upf3, and stimulates the enzyme activity of Upf1 (Chamieh et al., 2008; Gupta and Li, 2018). It interacts with the CH domain of Upf1, which, together with ATP binding, induces a structural change of Upf1 from the tight RNA-clamping to the relatively open RNA-unwinding conformation (Chakrabarti et al., 2011). In the latter conformation, the affinity of Upf1 to RNA is reduced, while its helicase activity is enhanced, promoting the execution of downstream events. Upf2 was also shown to interact with components of both ribosomal subunits, eRF3, the EJC, and SMG1 (see **2.3.5**), reflecting its support throughout the NMD pathway (Wang et al., 2001; López-Perrote et al., 2016).

Upf3 is the least conserved of all three Upf proteins (Karousis and Mühlemann, 2019). It is also different from the other two proteins because it was suggested to shuttle between the nucleus and the cytoplasm. Human UPF3 localizes to the nucleus at steady state, where it joins the EJC (Lykke-Andersen et al., 2000; Serin et al., 2001; Kim et al., 2001). It remains associated with the EJC in the cytoplasm until translation, where it can then be involved in target recognition and activation of NMD (see *2.3.3.1*). In yeast, although Upf3 is localized primarily in the cytoplasm, it contains sequence elements that resemble nuclear localization and nuclear export signals, therefore may also shuttle (Shirley et al., 1998; Shirley et al., 2002). However, whether its presence in the nucleus could be linked to its function in NMD requires further investigation. Vertebrates have two paralogs of UPF3: UPF3 and UPF3X (or UPF3A and UPF3B), which is thought to provide an additional layer of regulation to the NMD pathway. They function in NMD in a rather antagonistic way: UPF3X is the stronger activator of NMD, while UPF3 poorly activates or even inhibits NMD (Kunz et al., 2006; Shum et al., 2016). UPF3 might compete with UPF3X for binding of UPF2 and in this way regulate NMD as the cell responds to different cellular conditions.

### 2.3.3. NMD activation

Considering the variety of substrates that are targeted to NMD and the fact that translation termination in NMD shares the same termination factors (eRF1 and eRF3) as normal termination, it is a long-standing question how a specific transcript is recognized as an

NMD target. To date, research has pointed to different models for NMD activation, but among supporting evidences were also studies showing disagreeing results to each model. The fact that one model cannot explain all known NMD cases suggests that indeed multiple pathways are used to activate NMD and/or a unified underlying mechanism still needs to be identified. Since NMD covers diverse substrates, it is conceivable that different mechanisms have developed, potentially for different subgroups of substrates, which converge at the activation of Upf1 and later rely on the same set of degradation machineries.

### 2.3.3.1. EJC-dependent NMD activation

The most prominent and efficient feature for NMD activation in higher eukaryotes is the presence of an EJC in the 3' UTR of a transcript, ≥ 50–55 nt downstream of a stop codon (Nagy and Maquat, 1998; Colombo et al., 2017). The EJC consists of core factors eIF4A3, RBM8A, MAGOH, and is joined by UPF3X, as well as many other auxiliary factors (Buchwald et al., 2010; Woodward et al., 2017; Schlautmann and Gehring, 2020). It remains bound to the RNA throughout export and is normally dissociated by the translocating ribosome during the pioneer round of translation (Lejeune et al., 2002; Sato and Maquat, 2009) (**Figure 9c**). However, since the ribosome disassembles at the stop codon and does not translate further into the 3' UTR, any EJC that is in the 3' UTR would generally not be removed and remain associated with the mRNA. Given that stop codons are typically present in the last exon, the presence of an EJC downstream of the stop codon implies that the stop codon is likely premature. In these cases, UPF2 is thought to be recruited to the EJC-bound UPF3X and subsequently interacts with UPF1 that associates with the terminating ribosome (Gehring et al., 2003; Chamieh et al., 2008; Buchwald et al., 2010). The formation of such an UPF1–UPF2–UPF3 complex was shown to stimulate the helicase activity of UPF1, thereby promoting downstream events of NMD (**Figure 9a**).

It became apparent, however, that an EJC downstream of the stop codon is not the only feature that can activate NMD because EJCs are not present in some lower eukaryotes, like in yeast *S. cerevisiae* (Boisramé et al., 2019), where NMD still functions properly. Besides, this model limits NMD targets to intron-containing transcripts, which has been proven to be untrue. Furthermore, since EJCs are removed during translation, this model implies that NMD is only activated in the pioneer round of translation. Indeed, studies showed that NMD largely occurs immediately after mRNA export, during the first round of
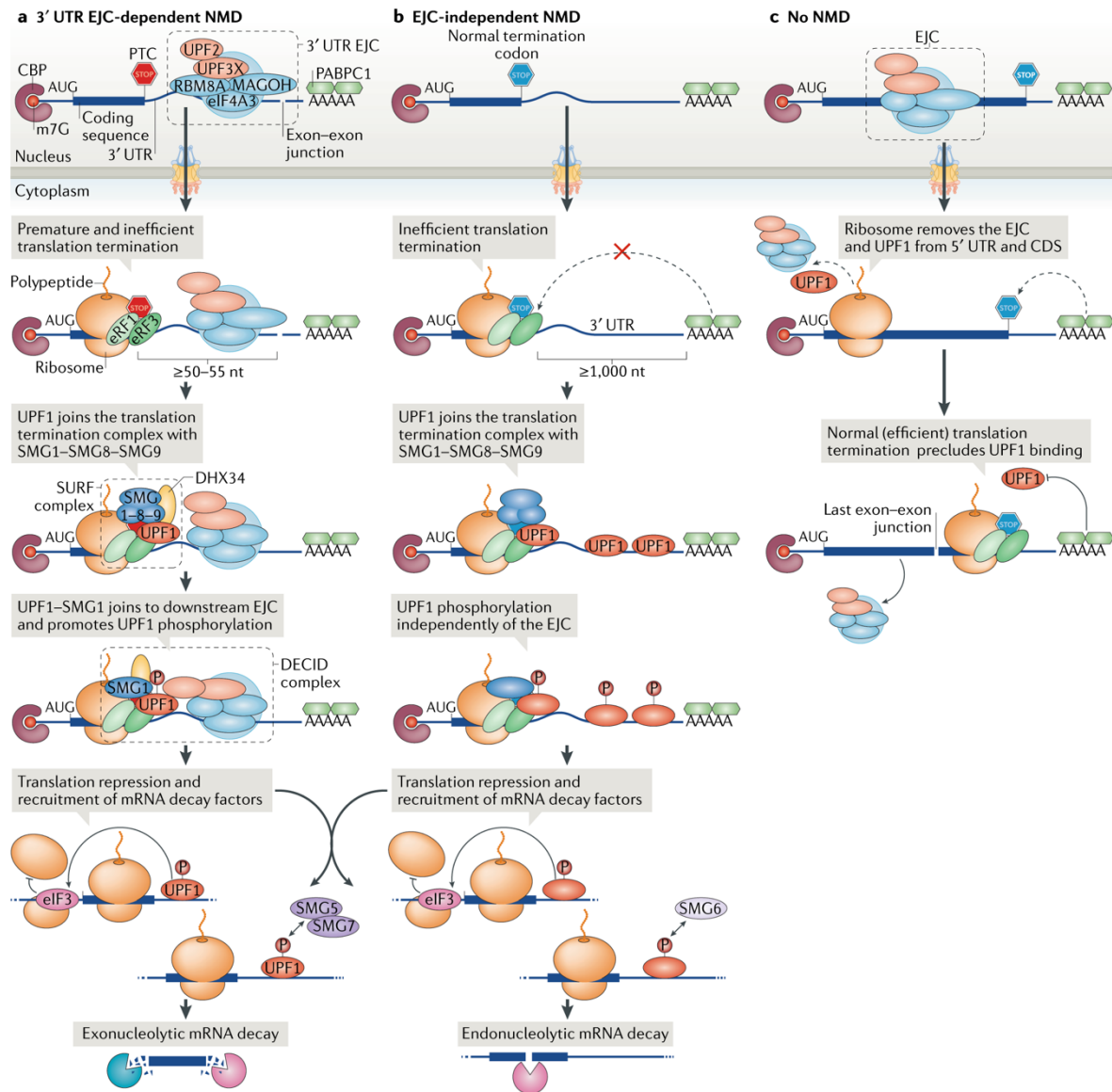
**Figure 9. Nonsense-mediated mRNA decay in higher eukaryotes**

**a.** Exon junction complexes (EJCs), containing eIF4A3, RBM8A, MAGOH, and NMD factors, are deposited onto the transcript upon splicing and later disassembled by the translating ribosome. In the presence of a premature termination codon (PTC) more than 50–55 nt upstream of an EJC, the ribosome stops at the PTC and does not remove the downstream EJC. The interaction between poly(A)-binding protein PABPC1 and termination factor eRF3, which normally promotes termination, is hindered and may cause inefficient termination. UPF1, in a complex with the serine/threonine kinase and its regulators SMG1–8–9, is recruited to the terminating ribosome, forming the SMG1–UPF1–eRFs (SURF) complex. UPF1 and SMG1 contacts UPF2 and UPF3 that are associated with the downstream EJC, forming the decay-inducing (DECID) complex. This promotes phosphorylation of UPF1 by SMG1, leading to translation repression and mRNA degradation. Translation repression may be mediated through the interaction between phosphorylated UPF1 and translation initiation factor eIF3. Phosphorylated UPF1 on the other hand recruits SMG6 for endonucleolytic cleavage of the transcript and SMG5–SMG7, which promotes deadenylation and decapping followed by exonucleolytic decay from both the 5' and 3' ends. **b.** On transcripts with a long 3' untranslated region (3' UTR), PABPC1 is too distant from the terminating ribosome and cannot interact with eRF3, leading to less efficient translation termination. UPF1, which is enriched in the 3' UTR, may bind to eRF1–eRF3 and become phosphorylated, activating NMD. Subsequent translation repression and mRNA decay events occur as described in **a**. **c.** EJCs upstream of the

stop codon as well as UPF1 molecules that bind promiscuously to the transcript are removed by the ribosome during translation. The correct mRNP structure at a normal stop codon allows efficient termination and counteracts NMD activation. (Adapted from Kurosaki et al., 2019).

translation while the 5' cap is still associated with the CBC (Barberan-Soler et al., 2009; Maquat et al., 2010a; Trcek et al., 2013; Heyer and Moore, 2016; Kim et al., 2017). Reports in mammalian cells also demonstrated that CBP80, a component of the CBC, associates with UPF1 and promotes NMD activation (Hosoda et al., 2005; Hwang et al., 2010). Once the CBC is replaced by the eIF4F complex, EJC-mediated NMD is largely over and the mRNA becomes relatively immune to NMD. Further supporting this is the fact that steady-state translation favors translation re-initiation on the same RNA through the closed-loop structure, and enhanced translation efficiency is thought to counteract effective NMD activation (Zhang and Maquat, 1997; Neu-Yilik et al., 2011). In agreement with that, several reports suggested that the interaction between PABPC1 and eIF4G, which mediates the close-loop formation, is responsible for stabilizing NMD reporter transcripts (Peixeiro et al., 2012; Joncourt et al., 2014; Fatscher et al., 2014). Later studies, however, showed that NMD is initiated also in subsequent rounds of translation and that EJC-dependent NMD occurs also on eIF4E-bound transcripts (Durand and Lykke-Andersen, 2013; Rufener and Mühlemann, 2013). In yeast, it was likewise found that NMD targets both Cbc1–Cbc2- and eIF4E-bound transcripts (Maderazo et al., 2003; Gao et al., 2005), although the fraction of NMD targets belonging to each group is unknown and the possible differences in mechanisms are unclear. Given these findings, it is generally thought that EJC-dependent activation is the basis of a large proportion, but not all cases, of NMD.

Notably, although yeast lacks EJCs, early research had identified the RNA-binding protein Hrp1 as a factor that promotes NMD in a fashion somehow analogous to the EJC (González et al., 2000). Hrp1 was shown to bind specifically to a sequence motif termed downstream sequence element (DSE) originally identified in the *PGK1* mRNA (Zhang et al., 1995). Further results suggested that binding of Hrp1 to DSE within ~200 nt downstream of a PTC promotes the activation of NMD by recruiting Upf1 to the vicinity of the terminating ribosome (Ruiz-Echevarría et al., 1998; González et al., 2000). However, the sequence of a DSE is rather degenerate (Zhang et al., 1995) and in-depth research on DSE- and Hrp1-dependent NMD is so far lacking. Confoundingly, Hrp1 was also indicated to bind to the 5' UTR of some transcripts and promote Upf1-dependent decay (Kebaara et al., 2003; Miller et al., 2018). Further research might better characterize Hrp1-mediated

NMD on a global scale and reveal whether the binding of Hrp1 to a DSE alone can induce NMD or if it is only part of a more complex mechanism that signals NMD activation.

Interestingly, a recent study in yeast suggests that despite lacking EJCs, an intron in proximity to a PTC enhances NMD (Wen et al., 2020 [preprint]), raising the possibility that the connection between splicing and NMD might be conserved from simple eukaryotic species. That said, only less than 5% of yeast transcripts contain introns, and approximately 27% of the yeast transcriptome undergoes splicing (Spingola et al., 1999; Lopez and Séraphin, 1999; Ares et al., 1999), but NMD substrates were estimated to result from almost 50% of all expressed genes (Malabat et al., 2015). This indicates that NMD must also be activated through alternative, splicing- and EJC-independent mechanisms.

### 2.3.3.2. EJC-independent NMD activation – the long 3' UTR model

The molecular mechanism of NMD activation by the EJC-dependent model is relatively well-characterized, but it fails to account for NMD events where an EJC is apparently not present. Besides in yeast, which seems to lack EJCs, NMD that is independent of EJC was also observed in other species, including human (Bühler et al., 2006; Singh et al., 2008; Wen and Brogna, 2010; Zahdeh and Carmel, 2016; Carvalho et al., 2017; Tian et al., 2017).

Several reports indicate that the physical distance between the termination codon and the poly(A) tail, namely the length of the 3' UTR, might be the key factor that determines whether or not NMD is activated on a specific transcript. This is known as the EJC-independent or the long 3' UTR model (Kurosaki et al., 2019). In early studies using the *PGK1* reporter in yeast, it was shown that stop codons positioned within the first two-thirds of the ORF strongly affected the half-life and steady-state level of the reporter transcript through NMD (Peltz et al., 1993; Hagan et al., 1995). Several following reports both in yeast and human corroborated the link between long 3' UTRs and NMD activation (Muhlrad and Parker, 1999a; Singh et al., 2008; Kebaara and Atkin, 2009; Tani et al., 2012; Wu et al., 2020).

The long 3' UTR model posits that the canonical 3' UTR of a transcript provides the correct mRNP formation for proper translation termination, thereby preventing NMD activation (Amrani et al., 2004). In particular, the interaction between the poly(A)-binding protein Pab1/PABPC1 and eRF3 was shown to promote termination (Uchida et al., 2002; Cosson et al., 2002; Ivanov et al., 2008; Ivanov et al., 2016). Therefore, it is thought that

when the physical distance between the poly(A) tail and the termination codon is too long, Pab1–eRF3 interaction is hindered, while Upf1 is enriched in the 3' UTR and thus has a higher chance to bind to the ribosome and activate NMD (Amrani et al., 2004) (**Figure 9b**). It was further shown that human UPF1 competes with PABPC1 for the interaction with eRF3 (Singh et al., 2008), supporting the idea that the availability of specific termination-promoting factors, which depends largely on the 3' UTR structure of the transcript, affects the probability of NMD activation. In line with this, tethering Pab1/PABPC1 or eRF3 between the PTC and the poly(A) tail could suppress NMD (Amrani et al., 2004; Ivanov et al., 2008; Silva et al., 2008; Eberle et al., 2008; Singh et al., 2008; Fatscher et al., 2014). Reciprocally, artificially extending the 3' UTR of a non-target transcript could render it subjective to the NMD pathway (Eberle et al., 2008; Singh et al., 2008).

Inconsistently, transcriptome-wide studies in mammalian cells showed that the 3' UTR length of an mRNA does not correlate well with its susceptibility to NMD (Singh et al., 2008; Tani et al., 2012; Hurt et al., 2013). This discrepancy could at least be partially explained, however, by the discovery of additional *cis*-acting regulatory elements in the 3' UTR of mammalian mRNAs that make NMD regulation more complicated. For example, an AU-rich element within the first 200 nt downstream of the termination codon in a long 3' UTR was shown to be sufficient to inhibit NMD (Toma et al., 2015). In addition, binding of hnRNP L or the polypyrimidine tract binding protein PTBP1 close to an NMD-inducing stop codon could protect transcripts with long 3' UTRs from NMD (Ge et al., 2016; Kishor et al., 2019a; Fritz et al., 2020)**.**

The importance of the Pab1–eRF3 interaction in NMD activation was also challenged, as later reports in yeast revealed that this interaction had no effect on NMD (Kervestin et al., 2012; Roque et al., 2015). Furthermore, the absence of the poly(A) tail or Pab1 does not hinder NMD (Meaux et al., 2008), revealing the fact that competitive binding of Pab1 and Upf1 to eRF3 is probably not as crucial as proposed. In line with that, it was discovered that the carboxy-terminal domain of the human PABPC1, with which it binds to eRF3, is dispensable for NMD suppression (Joncourt et al., 2014), and that NMD could be inhibited by a tethered PABPC1 mutant that cannot interact with eRF3 (Fatscher et al., 2014). These findings suggest that the poly(A)-binding proteins may contribute to NMD activation through additional mechanisms or that other 3' UTR-associated elements are also involved.

Neither the EJC-dependent nor the long 3' UTR model can perfectly explain the mechanism of NMD activation, although they do indicate similar key features, such as

aberrant termination and the increased presence or recruitment of critical NMD factors. Studies are devoted to reveal further insights and to integrate the current knowledge for a unifying mechanism.

### 2.3.4. The major mRNA degradation machineries in the cytoplasm

Degradation of NMD targets relies on the same machineries as canonical mRNA turnover, with some differences in mechanism as well as involvement of additional specific factors (see **2.3.5**). Regular decay of mRNAs in the cytoplasm typically begins with deadenylation and in turn the displacement of poly(A)-binding proteins (**Figure 10**). The two major deadenylase complexes are the conserved Ccr4–Not and Pan2–Pan3 complexes, whereby their main catalytic activities come from Ccr4 and Pan2, respectively (Łabno et al., 2016). It is thought that deadenylation generally occurs in a bi-phasic process, where Pan2–Pan3 starts by trimming the poly(A) tail in a distributive manner, which is later taken over by the processive Ccr4–Not complex. How the deadenylases are recruited to the substrate is not well-understood, but likely involves interactions mediated by poly(A)-binding proteins as well as additional factors.
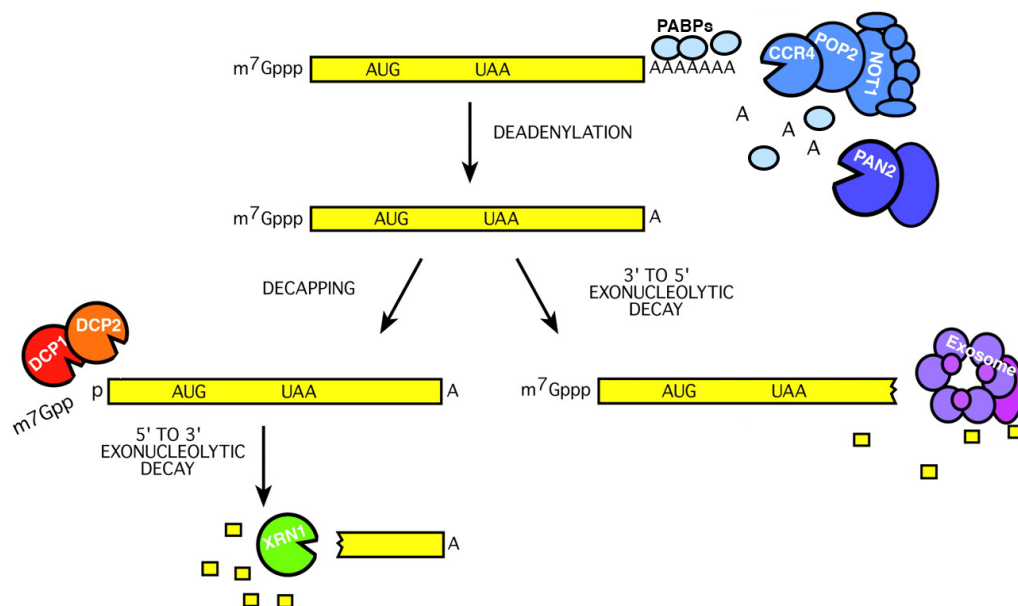


**Figure 10. Canonical cytoplasmic mRNA decay**
Normal mRNA turnover in the cytoplasm starts with shortening of the 3' end poly(A) tail, carried out by the Pan2–Pan3 and Ccr4–Not deadenylase complexes. Deadenylation promotes removal of the 5' cap by Dcp2 and its cofactor Dcp1, and the decapped transcript is subsequently degraded by the 5'–3' exoribonuclease Xrn1. The deadenylated 3' end is targeted by the cytoplasmic exosome, which requires the cofactor Ski complex (not depicted) and degrades the transcript in the 3'–5' direction. (Adapted from Decker and Parker, 2002).

Deadenylated transcripts are unprotected at the 3' end and thus vulnerable to 3'–5' exonucleolytic attack, exerted by the cytoplasmic exosome. In yeast, the inactive exosome core relies on Dis3(Rrp44) for exonuclease activity (Chlebowski et al., 2013), whereas human has three proteins homologous to Dis3: DIS3L, DIS3L2 in the cytoplasm, and DIS3 in the nucleus (Tomecki et al., 2010). The exosome is activated by its cofactor, the Ski complex, consisting of Ski3, Ski8, and the DExD/H RNA helicase Ski2 (Halbach et al., 2013). In yeast, an additional protein, Ski7, functions to connect the exosome with the Ski complex and thereby promotes target recognition and degradation (Araki et al., 2001).

In addition to allowing the exosome to access the transcript from the 3' end, deadenylation also activates degradation from the 5' end by promoting removal of the $m^7G$ cap (Łabno et al., 2016). The main decapping enzyme is Dcp2, which in yeast interacts with its cofactor Dcp1 to form a holoenzyme that stimulates its catalytic activity. Further, decapping is mediated by numerous enhancers of decapping (EDCs) both in yeast and higher eukaryotes. Dcp2 interacts with the EDCs and they form a decapping mRNP structure, which interacts with the major cytoplasmic 5'–3' exonuclease Xrn1. Dcp2 hydrolyzes the cap structure, leaving a 5'-monophosphate that is favorable for Xrn1-mediated nucleolytic attack. In yeast, decay from the 5' end by Xrn1 is the predominant pathway for cytoplasmic mRNA turnover, while in other eukaryotes the exosome may contribute more equally as Xrn1 to the overall mRNA decay activity.

### 2.3.5. NMD-mediated transcript degradation

Upon recognition of an NMD target by the Upf proteins, RNA degradation is rapidly initiated to remove the defective transcript. In yeast, it was found that NMD targets are mostly degraded from the 5' end by Xrn1 following decapping by Dcp1–Dcp2 and EDCs (Muhlrad and Parker, 1994; Hagan et al., 1995; Cao and Parker, 2003; He et al., 2003; Nissan et al., 2010). Different from canonical 5'–3' decay, where decapping of the RNA is often dependent on 3' deadenylation (Tucker and Parker, 2000; Ghosh and Jacobson, 2010), it was shown that decapping in the NMD process occurs independently of deadenylation (Muhlrad and Parker, 1994; He and Jacobson, 2015), likely for the purpose of accelerating degradation during quality control. The molecular mechanism for such a speedy response of the decapping enzymes is however unclear. Besides 5'–3' degradation, it was demonstrated that NMD targets could also be degraded in the 3' to 5' direction, albeit to a lesser extent, by the cytoplasmic exosome (Muhlrad and Parker, 1994; Mitchell and Tollervey, 2003). This may be promoted by the physical interaction between

Upf1 and Pab1, which was shown to stimulate accelerated deadenylation (Richardson et al., 2012). Together, the 5' and 3' degradation machineries carry out rapid decay to ensure the timely removal of NMD substrate RNAs. In a recent study that utilized affinity purification coupled with mass spectrometry, Upf1 was found to be in a complex with Dcp1 and Dcp2 along with additional factors that are involved in degradation (Dehecq et al., 2018), compatible with earlier observations that Upf1 physically interacts with several decapping factors (Tarassov et al., 2008; Swisher and Parker, 2011). This may be the core of a decay complex that acts on NMD targets in yeast, which may be joined by other factors that mediate the degradation process.

In higher eukaryotes, several additional factors that facilitate RNA degradation were identified. Moreover, the phosphorylation and dephosphorylation cycle of UPF1 was found to play an essential role in NMD-mediated degradation in metazoan (Karousis and Mühlemann, 2019; Kurosaki et al., 2019). Coordination of UPF1 phosphorylation with effective degradation is mediated by a group of SMG proteins. In human cells, UPF1 is recruited to the terminating ribosome along with SMG1 (**Figure 9a**), which is a serine/threonine kinase that phosphorylates UPF1 at multiple residues. The SMG1–UPF1–eRF1–eRF3 complex formed at the terminating ribosome is known as the SURF complex (Kashima et al., 2006). SMG1 is initially bound by SMG8 and SMG9, which inhibit its kinase activity until the SURF complex comes into contact with the downstream EJC, forming the decay-inducing (DECID) complex (Kashima et al., 2006; Yamashita et al., 2009). Formation of the DECID complex promotes the phosphorylation of UPF1 by SMG1, which represents the first and committing step of RNA degradation in human NMD.

Phosphorylated UPF1 directly recruits SMG6, an endonuclease that cleaves the RNA at a site close to the termination codon (Gatfield and Izaurralde, 2004; Huntzinger et al., 2008; Eberle et al., 2009) (**Figure 9** bottom right). The resulting 5' and 3' fragment ends are unprotected and subjected to degradation by the cytoplasmic exosome and XRN1, respectively. The 3' cleavage fragment likely contains UPF factors and EJC components, which are removed by UPF1 through its helicase activity prior to 5' to 3' degradation by XRN1 (Franks et al., 2010). Phosphorylated UPF1 also directly recruits the SMG5–SMG7 heterodimer (Ohnishi et al., 2003; Okada-Katsuhata et al., 2012) (**Figure 9** bottom left), which subsequently recruits the CCR4–NOT complex for deadenylation (Loh et al., 2013). Phosphorylated UPF1, SMG5, as well as the CCR4–NOT complex recruit also the decapping complex composed of decapping enzyme DCP2 and its cofactor DCP1a. The deadenylated and decapped transcript is then subjected to 5'- and 3'-mediated RNA

decay by XRN1 and the cytoplasmic exosome. After initiation of degradation, SMG5, SMG6, and SMG7 recruit the protein phosphatase 2A (PP2A) holoenzyme to dephosphorylate UPF1, allowing it to participate in a new round of NMD (Chiu et al., 2003; Anders et al., 2003; Ohnishi et al., 2003; Kurosaki et al., 2014).

Although a similar Smg-mediated regulation of Upf1 has not been defined in yeast, two additional factors that are involved in degradation of NMD substrates were identified. Nmd4 and Ebs1 were shown to associate with Upf1 and, importantly, were co-purified with Upf1 in samples that contained also decapping factors but not Upf2 and Upf3. They were thus suggested to be part of the mRNP undergoing decay in yeast NMD (He and Jacobson, 1995; Dehecq et al., 2018). Interestingly, these proteins share domain structure similarities with the SMG5–7 proteins (Clissold and Ponting, 2000; Luke et al., 2017; Dehecq et al., 2018). Moreover, Ebs1 was implicated in the degradation of an NMD reporter and further analyses showed that *EBS1* and *NMD4* deletion hindered NMD-mediated decay of both endogenous and reporter NMD substrates (Luke et al., 2007; Dehecq et al., 2018). These proteins were therefore suggested to be potential homologs of the human SMG5–7 proteins. In light of that, factors involved in NMD might be more conserved among species than previously thought. It remains to be further investigated how Ebs1 and Nmd4 facilitate NMD exactly, and whether more proteins that could be homologs of higher eukaryote NMD factors exist. Also, yeast Upf1 was shown to undergo phosphorylation that is likely functionally related to NMD (Wang et al., 2006; Lasalde et al., 2014), but details of those modifications and how they mediate yeast NMD remains to be clarified.

### 2.3.6. NMD-mediated translational repression

The activation of NMD leads not only to decay of the affected transcript, but also to a decrease in the translational efficiency of the targeted mRNA. It was shown in yeast using a reporter transcript that the Upf1-mediated identification of an mRNA as nonsense-containing induces translational repression and subsequently decapping (Muhlrad and Parker, 1999b). How the nonsense signal is transmitted to the 5' end for translation inhibition and decapping is however unknown. Similarly, UPF1 was shown to decrease translation initiation of a reporter mRNA in human cells (Isken et al., 2008). There, the authors observed that phosphorylated UPF1 interacts with eIF3, a translation initiation factor that promotes the formation of active 80S ribosomes, and presumably in this way directly prevents further translation initiation (**Figure 9** bottom). Since translation initiation

usually depends on the 5' cap, which is also the target of decapping factors, it is conceivable that NMD substrates undergo repression of translation initiation prior to or concurrent with decapping and decay. In fact, several studies have demonstrated that translation repressors that target initiation often also promote decapping (Coller and Parker, 2005; Roy and Rajyaguru, 2018). It remains to be further understood how these events are coordinated in NMD and which factors are involved. Together, inhibition of translation and RNA degradation both serve to achieve the goal of NMD: minimizing aberrant gene products.

## 2.4. Aim of the study

The yeast SR-like proteins Gbp2 and Hrb1 control the quality of mRNAs in the nucleus and support the export of properly spliced transcripts (Hackmann et al., 2014). By doing so, they contribute to the coupling of multiple processes during mRNA biogenesis. Moreover, they could potentially couple nuclear processes with cytoplasmic mRNA metabolism, as their association with mRNAs continue until translation (Windgassen et al., 2004). However, their function in the cytoplasm is so far unknown. The third SR-like protein, Npl3, similarly plays a role in co-transcriptional mRNA processing and export, and has been shown to regulate translation (Windgassen et al., 2004; Estrella et al., 2009; Rajyaguru et al., 2012; Baierlein et al., 2013). Their human counterparts, the SR proteins, are also known to accompany transcripts from the nucleus to the cytoplasm and have roles in translational regulation (Wegener and Müller-McNicoll, 2019). We postulate that Gbp2 and Hrb1 also function in the cytoplasm. Given that splicing defects often result in PTCs, we investigated if the two proteins are also involved in the cytoplasmic quality control pathway NMD. Although NMD is to date the best-studied mRNA surveillance pathway, several parts of the process remain ambiguous. Interestingly, human SR proteins have been implicated in NMD. Many of them are auxiliary factors of the exon junction complex and can potentially promote NMD through EJC-dependent activation. Furthermore, it was demonstrated that overexpression of SRSF1 and SRSF2 enhances NMD, although this is not dependent on their shuttling activities (Zhang and Krainer, 2004). Recently, it was shown that SRSF1 stimulates NMD by promoting Upf1 binding and that SRSF2 affects the association of key NMD factors (Aznarez et al., 2018; Rahman et al., 2020), demonstrating direct links between SR proteins and cytoplasmic surveillance. Nevertheless, the exact role of human SR proteins in NMD is difficult to unravel, since these factors affect the splicing process per se, which in turn influences EJC assembly (Singh et al., 2012) and may result in secondary effects when studying NMD downstream events. Moreover, the transcript level of several genes encoding splicing factors and SR proteins are regulated by NMD, making it complicated to examine the pathway independently (Ni et al., 2007; Lareau and Brenner, 2015). Gbp2 and Hrb1 are not splicing factors (Kress et al., 2008; Hackmann et al., 2014), but otherwise show functional similarities with SR proteins. Using the simple yeast system to characterize the possible function of these proteins in NMD, we wish to obtain new insights for some of the unsolved questions in this surveillance pathway. Supported by the recent finding that Gbp2 is enriched in purified Upf1 samples (Dehecq et al., 2018), this study explores the continued quality control function of Gbp2 and Hrb1 from the nucleus to the cytoplasm.

# 3. Results

*Partial results presented in this work have been published in Grosse et al., 2021 (see Appendix).*

## 3.1. Gbp2 and Hrb1 physically associate with core factors of NMD

### 3.1.1. Gbp2 and Hrb1 co-purify with the Upf proteins

To understand the possible role of Gbp2 and Hrb1 in NMD, it was first investigated whether Gbp2 and Hrb1 physically interact with the core NMD factors. For this, co-immunoprecipitation (co-IP) experiments using strains that genomically expressed Upf-GFP fusion proteins were carried out. GFP tagging did not cause deleterious effects, as growth of these strains were comparable to that of wild-type cells (**Figure 11**). Whole cell lysates were prepared from these strains and the Upf proteins were immunoprecipitated. The eluate samples were then analyzed on western blots with antibodies specific to Gbp2 and Hrb1. As shown in **Figure 12**, both Gbp2 and Hrb1 co-immunoprecipitated with all three Upf proteins. However, the interactions were sensitive to RNase treatment. This could mean that the proteins associate with the same set of RNAs but do not directly interact. On the other hand, the interactions may be transient and RNA-binding may be required for the proteins to be in certain conformations that enhance their binding affinities to each other. Interestingly, our other data suggest that Gbp2 and Hrb1 may be involved in NMD only on a subset of, and not all, NMD targets (see **4.1.2**). This would also result in the low abundance of possible interactions between these proteins and the Upf factors.

In an attempt to capture interactions that are short-lived or require RNA-binding, the co-IP experiment was repeated with additional application of formaldehyde, a protein–protein and protein–nucleic acids cross-linker (Ramanathan et al., 2019). Nevertheless, several trials with adjustments in formaldehyde concentration and incubation time so far only improved the pull-down signals and did not demonstrate co-precipitation between the examined proteins (**Figure 13**).
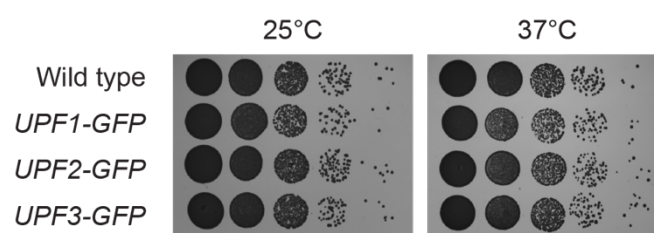


*(see next page for figure legend)*

**Figure 11. Fusion of the GFP tag to Upf proteins does not affect cell growth**
Growth analysis of the *GFP*-tagged *UPF* strains was carried out. Ten-fold serial dilutions of the wild-type, *UPF1-*, *UPF2-*, and *UPF3-GFP* strains, from $10^7$ to $10^3$ cells/ml, were spotted on full YPD agar plates and grown at the indicated temperatures for 2 days.
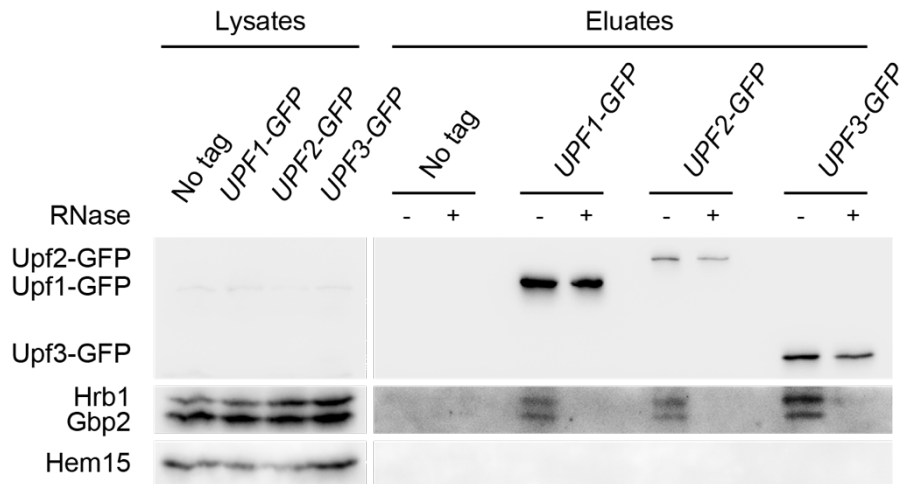


**Figure 12. Gbp2 and Hrb1 associate with Upf proteins**
Co-precipitation of Gbp2 and Hrb1 with Upf1-, Upf2-, or Upf3-GFP in the presence or absence of RNase was analyzed on western blots. Hem15 served as a control for unspecific binding.
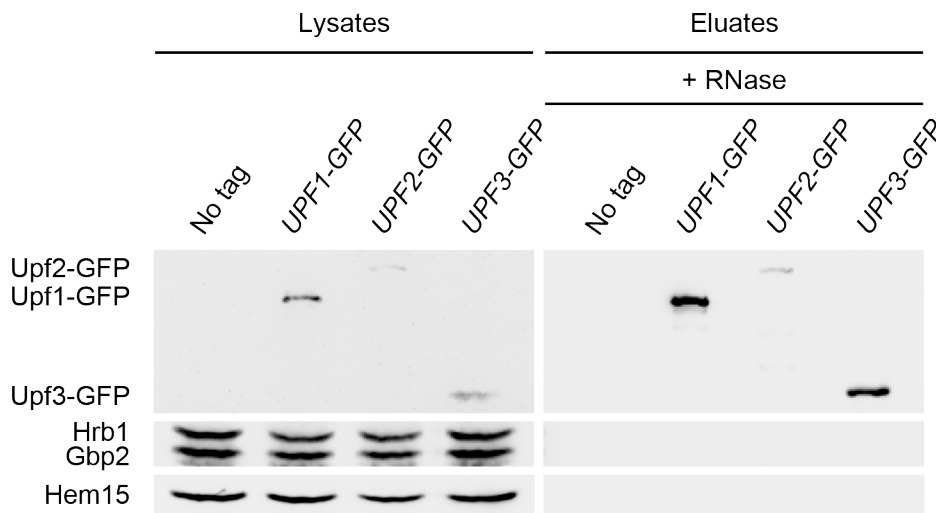


**Figure 13. Gbp2 and Hrb1 did not co-elute with Upf proteins after formaldehyde cross-link**
Co-precipitation of Gbp2 and Hrb1 with Upf proteins after formaldehyde and RNase treatment was analyzed on western blots. Formaldehyde was added to the cell cultures to a final concentration of 0.5% and incubated for 10 minutes before harvesting. Excess formaldehyde was trenched with 0.5 M glycine. Hem15 served as a control for unspecific binding.

**3.1.2. Gbp2 shows enriched co-purification with an ATP hydrolysis mutant of Upf1**

To further investigate whether Gbp2 and Hrb1 associate with the Upf proteins in an NMD-dependent manner, we made use of a mutant form of Upf1 that is unable to facilitate completion of the NMD pathway. The ATPase-dependent helicase activity of Upf1 is critical for functional NMD (see **2.3.2**). In early studies, conserved amino acids in the ATPase and helicase domain of Upf1 were mutated and the resulting mutant proteins were characterized (Weng et al., 1996a; Weng et al., 1996b). Mutations that alter an aspartate and a glutamate to two alanine residues (*upf1-DE572AA*) were found to abolish the ATPase and helicase activities of Upf1. This mutant was however still able to bind ATP, and its RNA-binding was also not affected (Weng et al., 1996a; Cheng et al., 2007). Later reports revealed that in cells expressing this ATP hydrolysis mutant upf1, NMD is still initiated, but degradation of the NMD substrate becomes defective (Franks et al., 2010; Serdar et al., 2016). It was demonstrated that upf1-DE572AA fails to displace the ribosome from the transcript after initiation of NMD. Xrn1 presumably begins to degrade the mRNA from the 5' end, but stalls at the persisting ribosome, resulting in incomplete degradation and eventually the accumulation of 3' decay fragments. Furthermore, several NMD factors, including SMG5–7, UPF2, and UPF3, showed increased co-purification with upf1-DE572AA in a complex with the 3' fragments, suggesting that release of these factors from the target RNA is dependent on the ATPase activity of UPF1 (Franks et al., 2010).

To understand how Gbp2 and Hrb1 would associate with the ATP hydrolysis mutant upf1, co-IP assays were carried out using the *UPF1* deletion strain transformed with either an *UPF1-HA* or an *upf1-DE572AA-HA* plasmid. Upf1 was immunoprecipitated and the co-purification of Gbp2 and Hrb1 was analyzed on western blots. However, signals of co-precipitated Gbp2 or Hrb1 could not be detected, even in the absence of RNase treatment (**Figure 14**). Since we observed clear interactions between the proteins and wild-type Upf1 in the previous experiment (**Figure 12**), we had reason to assume that the pull-down efficiency of the HA tag might have been too low for the detection of these intrinsically low-abundant interactions.

|  | upf1Δ | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | Lysates | | | Eluates | | |
| p*UPF1-HA* | - | + | - | - | + | - |
| p*upf1-DE572AA-HA* | - | - | + | - | - | + |
| HA | | | | | ▬ | ▬ |
| Hrb1<br>Gbp2 | ▬ | ▬ | ▬ | | | |
| Tdh1 | ▬ | ▬ | ▬ | | | |
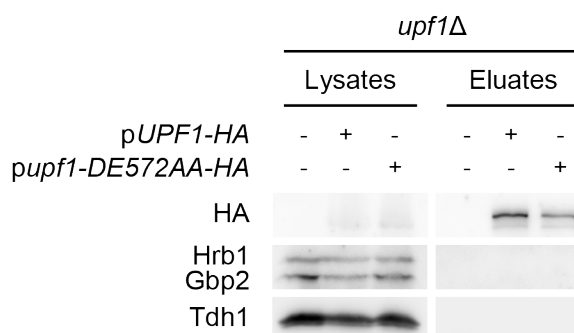
**Figure 14. Interaction of Gbp2 and Hrb1 with Upf1 is likely low-abundant**
Co-precipitation of Gbp2 and Hrb1 with Upf1- or upf1-DE572AA-HA was analyzed on western blots. Tdh1 served as a control for unspecific binding.

To rule out this possibility, the tagged *UPF1* genes were subcloned to replace the HA tag with GFP. The GFP-tagged proteins were first tested for their functionality via growth analysis. In previous reports, it was shown that deletion of the *UPF1* gene led to hypersensitivity of the cell to translation inhibitors, likely reflecting the translation-dependent nature of the NMD pathway (Leeds et al., 1992; Cui et al., 1995; Estrella et al., 2009). Consistent with these reports, the *upf1Δ* strain showed decreased growth on plates supplemented with cycloheximide, an antibiotic that inhibits translation elongation (**Figure 15**). Cells expressing the plasmid-encoded Upf1-GFP rescued the growth defect, whereas the upf1-DE572AA-GFP protein did not (**Figure 15**), indicating that the GFP-tagged Upf1 is functional.
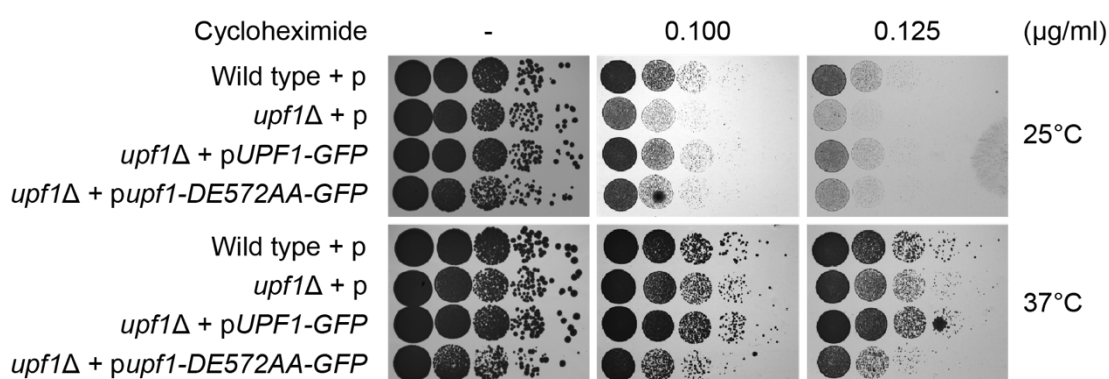


**Figure 15. The ATP hydrolysis mutant upf1-DE572AA is functionally defective**
Growth tests were carried out to evaluate the functionality of the Upf1- and upf1-DE572AA-GFP proteins. Ten-fold serial dilutions, from $10^7$ to $10^3$ cells/ml, of wild-type or *upf1Δ* cells transformed with an empty vector (p) or one of the *UPF1* plasmids were spotted on selective agar plates and grown at the indicated temperatures for 3 days. The plates contained different concentrations of cycloheximide as indicated.

Interestingly, subsequent co-IP experiments with these proteins showed that co-purification of Gbp2 with the mutant upf1 increased about 1.5-fold over wild-type Upf1 (**Figure 16a** and **16b**). In contrast, Hrb1 showed no difference in its binding to wild-type and mutant upf1, suggesting that Gbp2 enrichment was not an unspecific effect for general RNA-binding proteins. This indicates that Gbp2 probably remains present on NMD targets after initiation of degradation and may require proper function of Upf1 to be dissociated from the mRNP. This is analogous to the aforementioned NMD factors that are enriched in upf1-DE572AA mRNPs and supports the notion that Gbp2 has cytoplasmic functions specific to NMD. Interestingly, the result also points out a possible divergence in the functions of the two highly homologous proteins Gbp2 and Hrb1, which requires to be further explored.
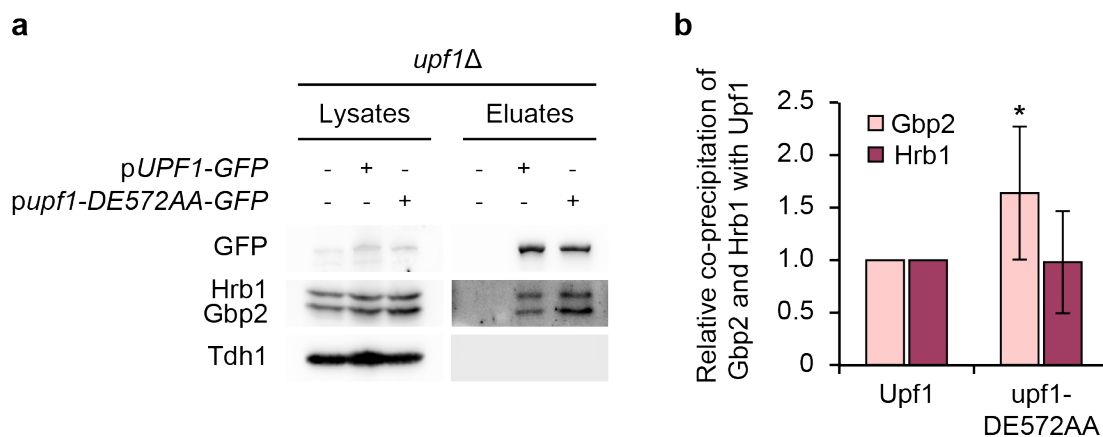


**Figure 16. Gbp2 associates to a greater extent with the ATP hydrolysis mutant of Upf1**
**a.** Co-precipitation of Gbp2 and Hrb1 with wild-type Upf1 or upf1-DE572AA was analyzed on western blots. Tdh1 served as a control for unspecific binding. **b.** Relative binding of Gbp2 and Hrb1 to wild-type or mutant upf1 was quantified. The intensities of the Gbp2 and Hrb1 bands in **a.** were measured and related to that of the corresponding Upf1 or upf1-DE572AA pull-down signals. Results were obtained from 7 independent experiments and are shown as means with error bars indicating the standard deviations.

### 3.1.3. Gbp2 is likely part of the Upf1 mRNP complex on targeted NMD substrates

Since previous co-IP experiments provided mostly indirect evidence for the physical interactions between Gbp2, Hrb1, and the Upf proteins, we employed the bimolecular fluorescence complementation (BiFC) system, which has the capacity to detect protein–protein interactions that are weak or transient in living cells (Miller et al., 2015; Zhang et al., 2016). The *GFP* sequence was split into an N-terminal and a C-terminal fragment (Baierlein et al., 2013; Miller et al., 2015) and fused to the C-terminus of the *UPF1* (*UPF1-*

*N-GFPsplit*) and *GBP2* gene (*GBP2-C-GFPsplit*), respectively. Physical interaction between the proteins of interest was studied in cells that express both split-GFP constructs, while expression of individual constructs served as negative controls. If the two proteins interact, the separated *GFP* fragments are brought into close proximity and may reassemble into a functional GFP molecule that emits fluorescence. Therefore, detection of GFP signals would indicate that these proteins closely, if not directly, interact in the cell (Miller et al., 2015).

The *gbp2Δ upf1Δ* strain was transformed with *GBP2-C-GFPsplit, UPF1-N-GFPsplit*, or both constructs, and the cells were examined under a fluorescence microscope. However, initially no evident GFP signal was detected in cells expressing both proteins as well as in the controls (**Figure 17a–c**). The position at which the tag is fused to the proteins may affect protein–protein interaction or the assembly of the chromophore. A plasmid encoding *UPF1* N-terminally tagged with *N-GFPsplit* was therefore constructed and used in the experiment. Nevertheless, this construct also did not result in strong green fluorescence when expressed together with the *GBP2* split-GFP construct (**Figure 17d–f**). It is conceivable that even in the presence of Upf1–Gbp2 interaction, the signal may be below detection limit, given that the interaction is presumably rare under normal situations.
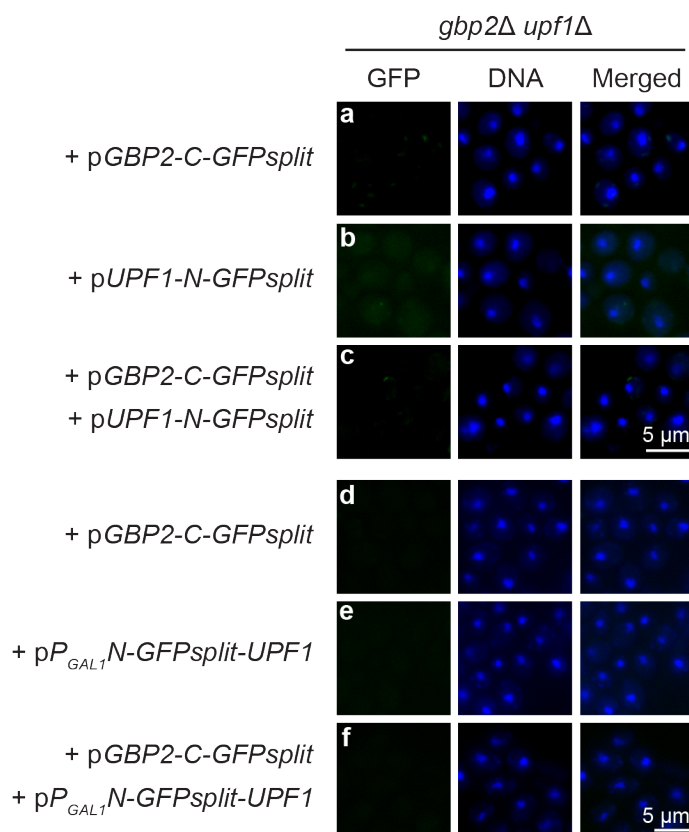


**Figure 17. Initial split-GFP experiments did not indicate binding between Gbp2 and Upf1** Physical interaction between Gbp2 and Upf1 was studied using the split-GFP system. The *gbp2Δ upf1Δ* strain was transformed with the indicated plasmids and analyzed under the fluorescent microscope. DNA was stained with DAPI (4',6-diamidino-2-phenylindole) (blue). **a., b., d., e.** Cells transformed with only one of the split-GFP constructs served as negative controls. **f.** Expression of the *P_{GAL1}N-GFPsplit-UPF1* plasmid was induced with galactose for 2 hours.

Conditions in which NMD substrates were enriched in the cell were thus tested. The experiment was first repeated in the *xrn1Δ gbp2Δ upf1Δ* strain to suppress degradation of NMD targets, and slightly increased GFP fluorescence was observed when both split-GFP constructs were expressed compared to the controls (**Figure 18a–c**). The cells were further transformed with a reporter plasmid that overexpresses a PTC-containing mRNA with a galactose-inducible promoter to increase NMD. The reporter construct ($P_{GAL1}CBP80^{PTC}$) (**Figure 19**) was created by introducing a point mutation in the intron-containing *CBP80* gene, resulting in a PTC downstream of the intron. Previous experiments had shown that the reporter is efficiently targeted to the NMD pathway in wild-type cells (Grosse et al., 2021). Additionally, to take advantage of the ATP hydrolysis mutant upf1-DE572AA, which showed an increased interaction with Gbp2 compared to wild type (**Figure 16**), a mutant *upf1* split-GFP construct was also created (*upf1-DE572AA-N-GFPsplit*). With these conditions, co-expressing split-GFP constructs of *upf1-DE572AA* and *GBP2* resulted in fluorescent signals that were significantly higher than in the negative controls (**Figure 18d**). This indicates that Gbp2 indeed physically contacts Upf1 *in vivo*. Since the interaction could be enriched by increasing the amount of NMD targets in the cell, Gbp2 is presumably present in Upf1-containing mRNP complexes on transcripts that undergo NMD, and is likely associated with the 3' decay fragments that accumulate with the mutant upf1. Despite this interaction being naturally rare, the difficulty to detect it in wild-typical situations may also reflect the rapid degradation of NMD targets after their recognition.
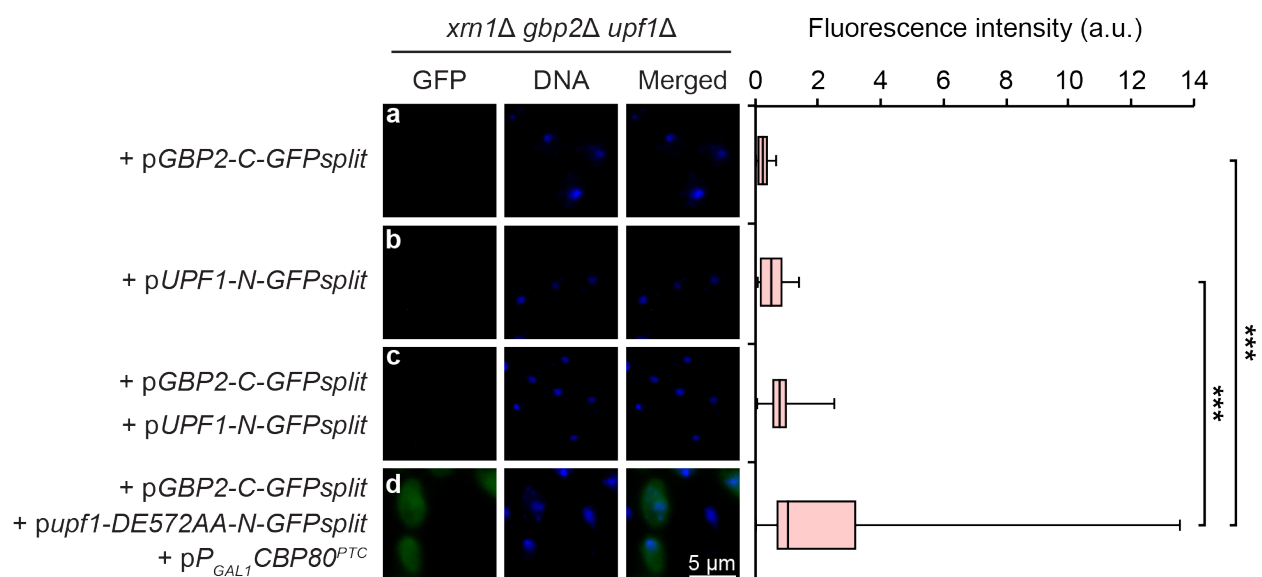


**Figure 18. Gbp2 and Upf1 physically interact *in vivo***
Physical interaction between Gbp2 and Upf1 or upf1-DE572AA in *xrn1Δ gbp2Δ upf1Δ* cells was studied using the split-GFP system as in **Figure 17**. GFP fluorescent signals of 100 cells per strain

per experiment was quantified from 3 independent experiments and the results are summarized in the box plot. **a-b.** Cells transformed with only one of the split-GFP constructs served as negative controls. **d.** Expression of the *CBP80^PTC* NMD reporter (*P_GAL1_CBP80^PTC*) was induced with galactose for 2 hours to increase NMD.



**Figure 19. The *P_GAL1_CBP80^PTC* NMD reporter contains a PTC downstream of its intron**
Scheme of the NMD reporter construct (Grosse et al., 2021) used throughout this study.

The established split-GFP system was subsequently applied to investigate the physical interaction between Upf1 and Hrb1. A similar set-up was used, except the *GBP2* split-GFP construct was replaced with the *HRB1-C-GFPsplit* plasmid. However, we did not observe evident fluorescent signal when both *UPF1* and *HRB1* constructs were expressed (**Figure 20a–c**). Using the mutant upf1 and additionally overexpressing the *CBP80^PTC* reporter led to similar results (**Figure 20d**). As Hrb1 was not enriched in the precipitated upf1-DE572AA samples (**Figure 16**), this finding was not too surprising. Association of Upf1 and Hrb1 may be mediated through other proteins and thus is indirect, or it may occur only early in the NMD pathway, prior to Xrn1-mediated degradation, and therefore was not enriched by the deletion of *XRN1*. These possibilities require further investigation for confirmation.
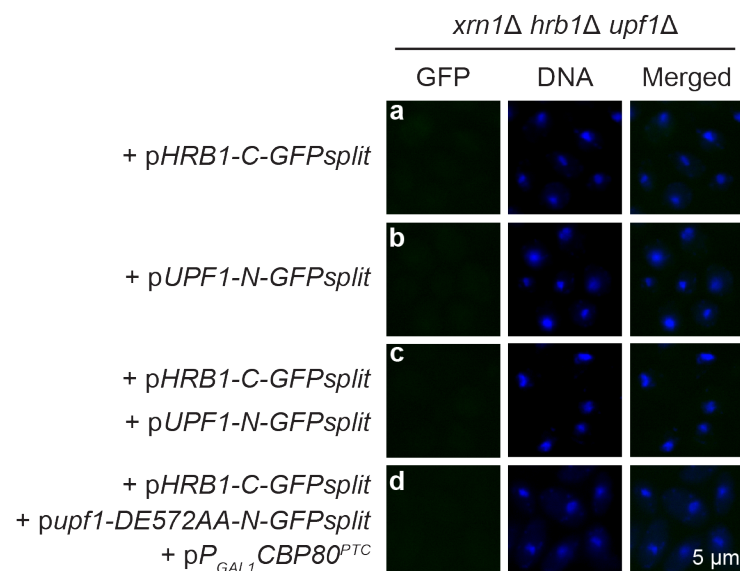


**Figure 20. Association of Hrb1 with Upf1 is likely transient or indirect**
Physical interaction between Hrb1 and Upf1 or upf1-DE572AA was studied using the split-GFP system. The *xrn1Δ hrb1Δ upf1Δ* strain was transformed with the indicated plasmids and analyzed with the fluorescent microscope.

## 3.2. Gbp2 and Hrb1 appear to function in NMD after target recognition

### 3.2.1. Binding of Upf1 to the *CBP80^{PTC}* NMD reporter mRNA is independent of Gbp2 and Hrb1

To gain insight into how Gbp2 and Hrb1 may function in NMD, we first asked if they are involved in an early step during substrate recognition. In yeast, the long 3' UTR model and the binding of Hrp1 downstream of a PTC have been proposed for NMD activation, but do not account for all cases of NMD (see **2.3.3**). The prominent EJC-dependent mechanism discovered in higher eukaryotes was also not described in yeast. Additional factors that may facilitate target recognition and activation of NMD are otherwise poorly defined. Considering the similarities between Gbp2, Hrb1, and human shuttling SR proteins, which are associated with the EJC, we wondered if Gbp2 and Hrb1 help recruit and stabilize Upf1 on spliced NMD substrates, similar to an EJC downstream of a PTC.

To study this, RNA co-immunoprecipitation (RIP) experiments were carried out to examine the association of Upf1 with the *CBP80^{PTC}* reporter in the presence or absence of Gbp2 and Hrb1. Cells were transformed with a plasmid that encodes *UPF1-GFP* and the reporter plasmid, which was highly expressed upon galactose induction. Following Upf1 immunoprecipitation (**Figure 21a**), co-precipitated RNA was purified, reverse transcribed, and the level of *CBP80^{PTC}* reporter mRNA was detected via qPCR. Background signals were deducted by relating to the no tag control and results were normalized to the levels in whole cell lysates as well as an endogenous, non-target RNA to eliminate NMD unrelated effects. It was previously shown that factors involved in NMD substrate recognition, such as UPF2, promote the stable binding of UPF1 to reporter substrate RNAs (Kurosaki et al., 2014). Therefore, the *upf2Δ* strain was used as a control for the decreased association between Upf1 and the reporter RNA. Indeed, qPCR results showed that in cells depleted of Upf2, association of Upf1 with the reporter RNA significantly decreased compared to wild type (**Figure 21b**). On the contrary, simultaneous deletion of *GBP2* and *HRB1* did not show an effect on the association of Upf1 with the NMD substrate. This indicates that stable binding of Upf1 to the reporter RNA is most likely independent of Gbp2 and Hrb1 and argues that these two proteins are not essential for target recognition in the NMD pathway.
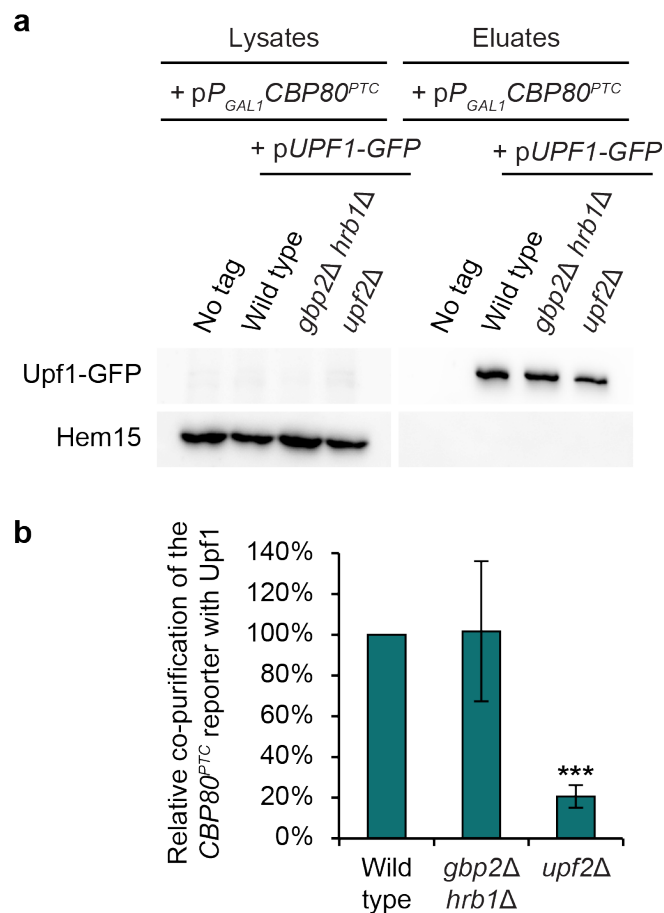
**Figure 21. Stable binding of Upf1 to the *CBP80^PTC* reporter does not require Gbp2 and Hrb1**
RNA co-immunoprecipitation (RIP) experiments were carried out with the indicated strains to study the association of Upf1 with the NMD reporter in the presence or absence of Gbp2 and Hrb1. Expression of the *CBP80^PTC* reporter was induced with galactose for 2 hours. **a.** Immunoprecipitation of Upf1-GFP was visualized on a western blot. Hem15 served as a negative control. **b.** The relative levels of co-precipitated *CBP80^PTC* mRNA in each strain were analyzed through qPCR and normalized to endogenous *RPS6A* mRNA levels. Results are shown as means with standard deviations. *gbp2Δ hrb1Δ* n = 8; *upf2Δ* n = 4.

## 3.2.2. Binding of Gbp2 and Hrb1 to NMD substrates is reduced in the absence of Upf1

According to the previous result, possible functions of Gbp2 and Hrb1 in the NMD process would rather occur downstream of target recognition by Upf1. It would thus be conceivable that the cytoplasmic association of Gbp2 and Hrb1 with NMD substrates may be influenced by Upf1. To test this, RIP experiments similar to the experiment described above were carried out and the binding of the two proteins to the NMD reporter was examined. The *CBP80^PTC* reporter was overexpressed in cells that contain either genomic *GBP2-GFP* or plasmid-encoded *HRB1-GFP*. Gbp2- or Hrb1-GFP was precipitated from whole cell lysates of wild-type or *upf1Δ* cells (**Figure 22a** and **22b**) and the co-purified

RNA was analyzed via qPCR. A no tag control was used and the levels of the NMD reporter were related to that of the *21S* rRNA as an internal control. Interestingly, we found that binding of both Gbp2 and Hrb1 to the *CBP80^{PTC}* reporter decreased on average about 50% in cells lacking Upf1 relative to wild type (**Figure 22c**1). Since the loading of both proteins onto mRNAs occurs co-transcriptionally in the nucleus (Hurt et al., 2004) and is therefore independent of Upf1, these results more likely suggest an earlier dissociation of the two proteins from the RNA in the cytoplasm when Upf1 is missing. Presumably, because NMD could not be activated in the absence of Upf1, the NMD reporter was regarded as a normal mRNA and Gbp2 and Hrb1 were displaced earlier during regular rounds of translation.
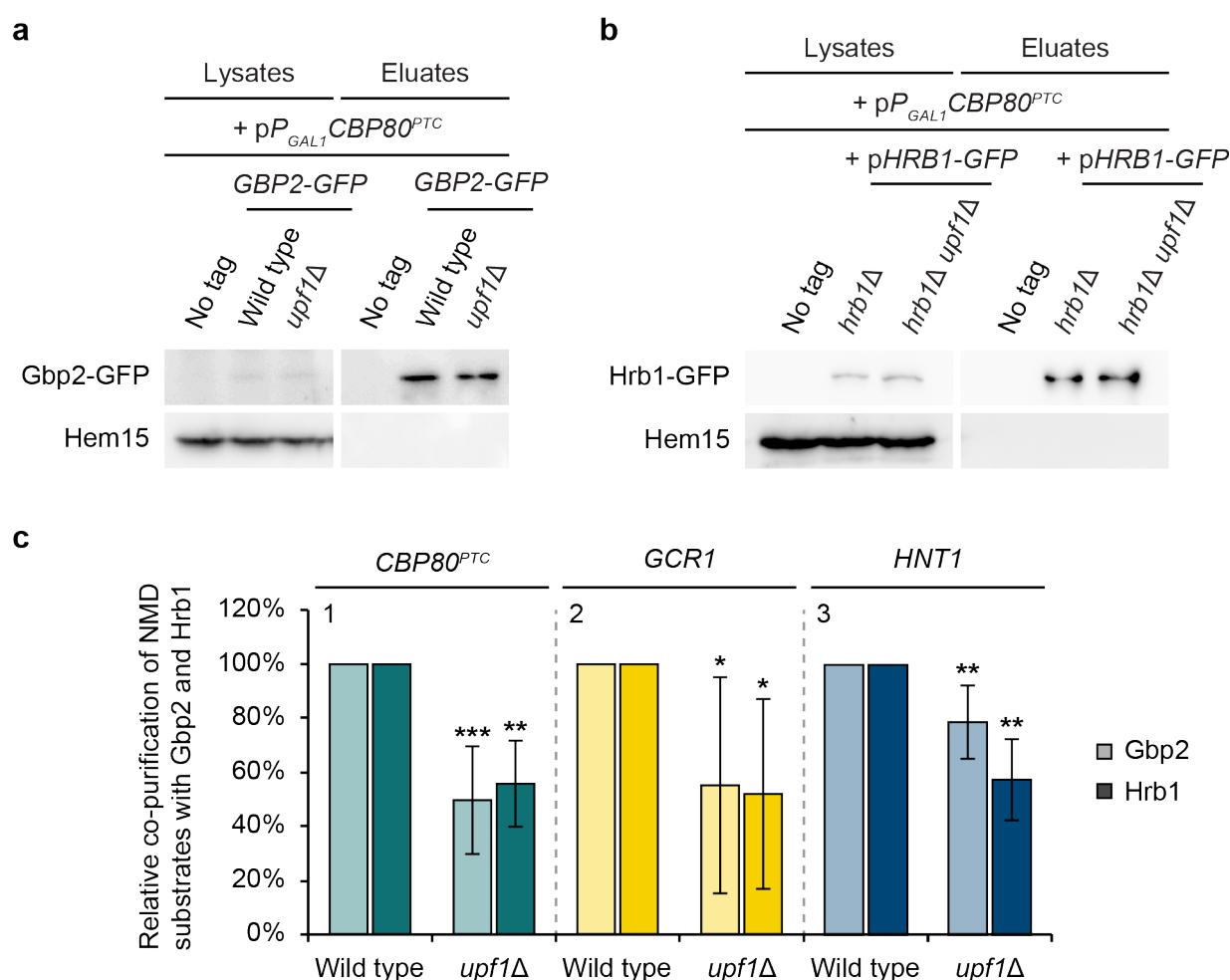


**Figure 22. Cytoplasmic association of Gbp2 and Hrb1 with mRNAs is influenced by NMD**
RIP experiments were carried out to study the association of Gbp2 and Hrb1 with NMD substrates in the presence and absence of Upf1. **a-b.** Immunoprecipitation of Gbp2-GFP (**a**) or Hrb1-GFP (**b**) was visualized on western blots. Hem15 served as a negative control. **c.** The relative levels of co-precipitated *CBP80^{PTC}* or the indicated endogenous mRNAs were analyzed through qPCR and normalized to *21S* rRNA. Results are shown as means with standard deviations. *CBP80^{PTC}*: Gbp2 n = 7, Hrb1 n = 6; *GCR1*: n = 6; *HNT1*: Gbp2 n = 7, Hrb1 n = 5.

To check whether this effect can also be demonstrated on endogenous NMD substrates, two intron-containing mRNAs, *GCR1* and *HNT1,* which were identified as putative NMD targets (Kawashima et al., 2014; Dehecq et al., 2018), were additionally analyzed. Consistent with the *CBP80^{PTC}* reporter, binding of Gbp2 and Hrb1 to *GCR1* decreased on average almost 50% in *upf1Δ* cells (**Figure 22c**2). The association of Hrb1 with the *HNT1* mRNA likewise decreased nearly half in the absence of Upf1, while Gbp2 was found to bind >20% less (**Figure 22c**3). These findings support the idea that Gbp2 and Hrb1 play specific roles in the NMD pathway that could lead to their extended association with NMD substrates. Considering that only Gbp2 was found to be enriched with the upf1-DE572AA mutant, it could be speculated that its involvement in NMD continues longer than that of Hrb1. In combination with previous results, we propose that Gbp2 and Hrb1 are involved in NMD for a subset of transcripts (presumably those that are spliced). They are dispensable for NMD target recognition, but likely act in downstream processes after initial activation of the NMD pathway.

## 3.3. Gbp2 and Hrb1 promote Ski2 recruitment to the *CBP80^{PTC}* reporter transcript

The major activity following NMD substrate recognition is the degradation of the target RNA. In yeast, this is mainly carried out by the exonuclease Xrn1 from the 5' end and to a lesser extent by the cytoplasmic exosome from the 3' end (He and Jacobson, 2015) (see **2.3.5**). For their role as mRNA quality control factors in the nucleus, Gbp2 and Hrb1 interact with Mtr4, a component of the TRAMP complex, the cofactor of the nuclear exosome. Through recruitment of the TRAMP complex, they are thought to promote binding of the exosome to an RNA that should be eliminated (Hackmann et al., 2014) (see **2.2.1** and **Figure 5**). We explored the idea that Gbp2 and Hrb1 also help to recruit the cytoplasmic exosome to the NMD substrate through an analogous mechanism. To test this, we chose to study Ski2, a member of the cytoplasmic exosome cofactor Ski complex, as it is the cytosolic structural and functional homolog of Mtr4 (Johnson and Jackson, 2013). First, it was investigated whether Gbp2 and Hrb1 physically interact with Ski2. With co-IP experiments, it was shown that Gbp2 and Hrb1 both co-purified with Ski2-GFP and the interactions were resistant to RNase treatment (Grosse et al., 2021).

Following that, the physical interaction between Upf1 and Ski2 was studied. Genomic Ski2-GFP was immunoprecipitated and the co-elution of Upf1-HA was analyzed via

western blot. In agreement with previous findings that exosome-mediated degradation of yeast NMD substrates is dependent on Ski complex components (Mitchell and Tollervey, 2003), Upf1 co-purified with Ski2 despite RNase treatment (**Figure 23a** bottom). This suggests that the proteins physically interact, probably in a complex that signals 3' degradation. To understand if Gbp2 and Hrb1 are relevant for the Upf1–Ski2 contact, the interaction was examined in the *gbp2Δ hrb1Δ* strain, which showed that deletion of *GBP2* and *HRB1* led to a significantly decreased interaction between the proteins when compared to wild type (**Figure 23a** and **23b**). This indicates that association of Ski2 with Upf1 depends, at least in part, on Gbp2 and Hrb1, supporting the hypothesis that these two proteins help to recruit the Ski complex to the faulty RNA.
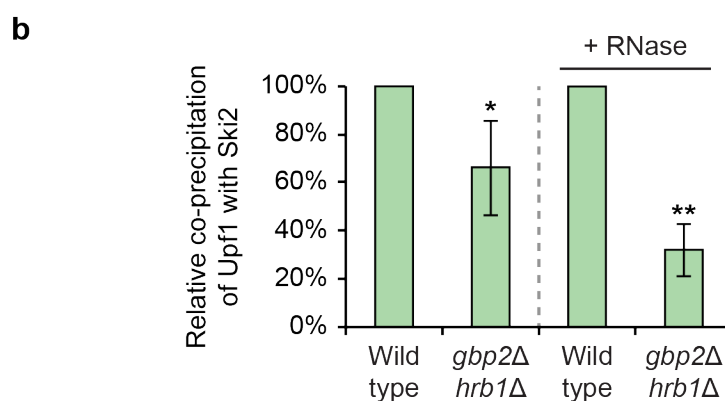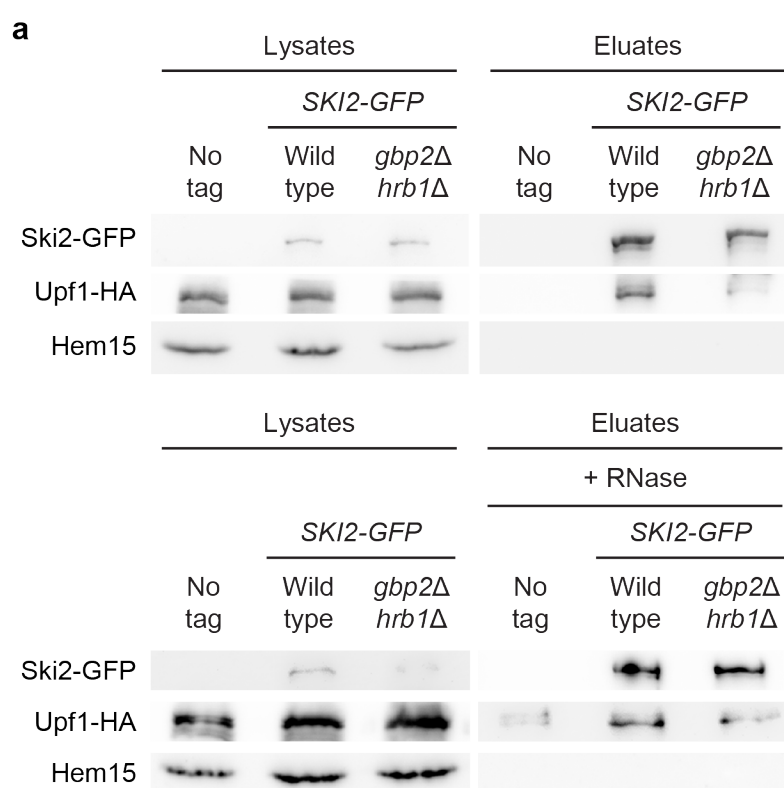


*(see next page for figure legend)*

**Figure 23. The Upf1–Ski2 interaction is partially dependent on Gbp2 and Hrb1**
**a.** Co-precipitation of Upf1-HA with Ski2-GFP in the presence or absence of Gbp2 and Hrb1 was analyzed on western blots. All strains were deleted for *UPF1* and transformed with the *UPF1-HA* plasmid. Hem15 served as a control for unspecific binding. RNase was added during immunoprecipitation for the bottom blot. **b.** Relative binding of Upf1 to Ski2 in the different strains was quantified from the western blots shown in **a**. The intensities of the Upf1 bands were related to that of the corresponding Ski2 pull-down signals. Results were obtained from 3 independent experiments each and are shown as means with standard deviations.

To further verify this hypothesis, the association of Ski2 with the *CBP80^{PTC}* reporter RNA was analyzed in RIP experiments using different mutant strains. Ski2-GFP was immunoprecipitated (**Figure 24a**) and the associating *CBP80^{PTC}* levels were determined by qPCR. The levels were normalized to that of the endogenous *CBP80* mRNA to eliminate non-NMD effects. We saw that in cells lacking Upf1, where NMD is non-functional, binding of Ski2 to the NMD reporter reduced approximately 35% compared to wild type (**Figure 24b**). This is a relatively mild effect and probably reflects the fact that 3'-mediated decay is the minor pathway used in yeast NMD. Interestingly, in cells lacking Gbp2 and Hrb1, the binding of Ski2 to the *CBP80^{PTC}* reporter decreased on average more than 20%, while the single deletion of *GBP2* or *HRB1* did not cause significant differences. Notably, we observed in other experiments that the double deletion of *GBP2* and *HRB1* constantly led to approximately half of the effect of *upf1Δ* in the context of NMD (Grosse et al., 2021). The Ski2 RIP result correlates with those findings and shows that Gbp2 and Hrb1 together affect the association of Ski2 with the NMD substrate. This and previous results are in support of a function of Gbp2 and Hrb1 in the recruitment of the Ski complex and conceivably, as a consequence, the cytoplasmic exosome to an RNA undergoing NMD. They may thereby promote efficient RNA decay, similar to their function in nuclear mRNA quality control.

## 3.4. Gbp2 and Hrb1, like Nmd4 and Ebs1, are auxiliary factors of NMD

We have obtained evidence supporting a role of Gbp2 and Hrb1 in NMD-mediated degradation. Recently, two other yeast proteins were shown to be involved in NMD and specifically implicated in target degradation. Nmd4 and Ebs1, owing to similar domain structures and related functions in NMD, were proposed to be yeast homologs of the human SMG5–7 proteins (Luke et al., 2007; Dehecq et al., 2018) (see **2.3.5**). As deletion of *GBP2* and *HRB1*, as well as deletion of *NMD4* and *EBS1*, resulted only in partial defects in the degradation of NMD substrates (Dehecq et al., 2018; Grosse et al., 2021),

we wondered if their combined depletion would more severely disrupt the degradation of NMD targets.

**a**



**b**
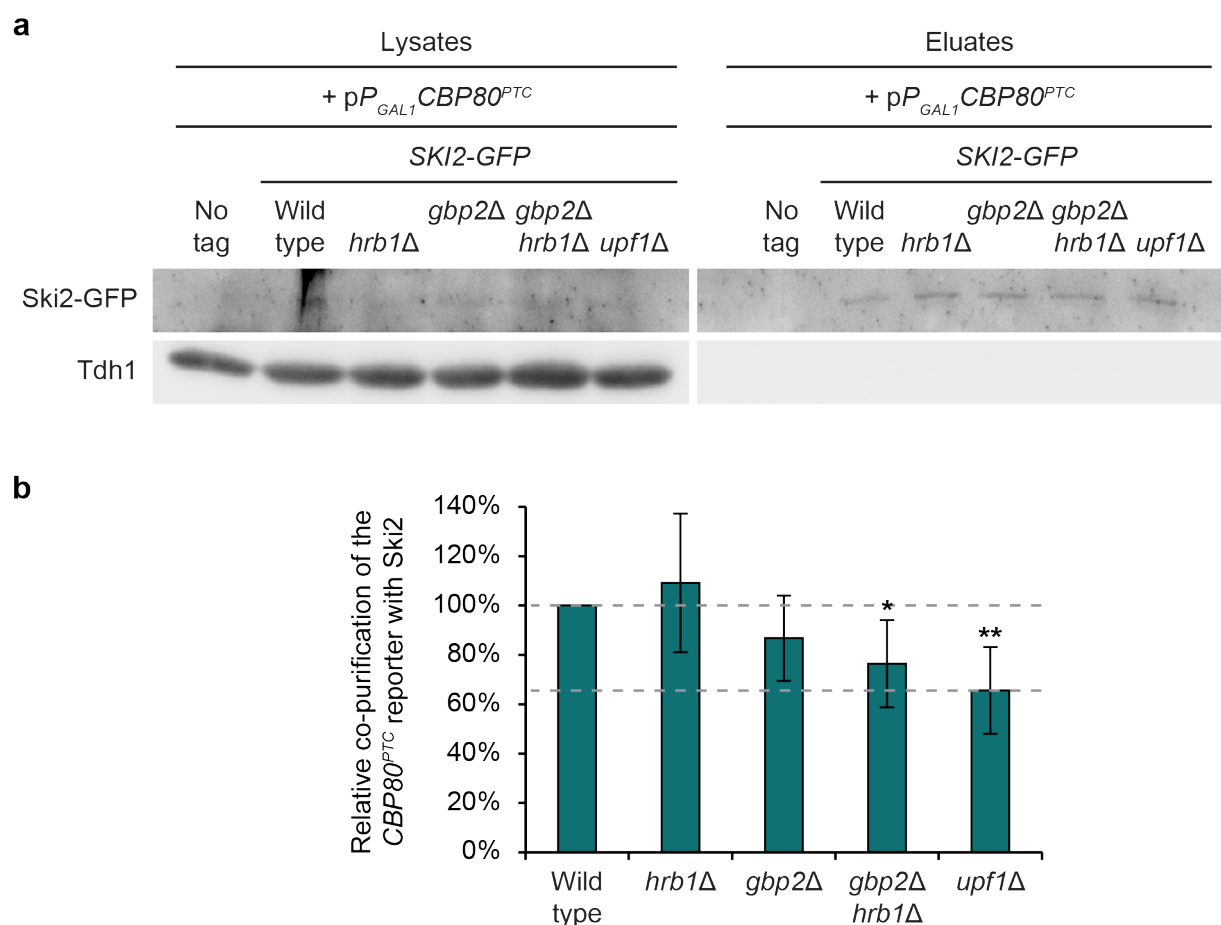


**Figure 24. Association of Ski2 with the *CBP80^PTC* reporter is promoted by Gbp2 and Hrb1**
RIP experiments were carried out with the indicated strains to study the association of Ski2 with the NMD reporter and its dependence on Gbp2 and Hrb1. Except for the *upf1Δ* strain, all strains were deleted for *UPF1* and transformed with the *UPF1-HA* plasmid. *CBP80^PTC* expression was induced with galactose for 2 hours. **a.** Immunoprecipitation of Ski2-GFP was visualized on a western blot. Tdh1 served as a negative control. **b.** The relative levels of co-precipitated *CBP80^PTC* mRNA in each strain were analyzed through qPCR and normalized to endogenous *CBP80* mRNA levels. Results are shown as means with standard deviations. Dashed lines indicate the results of the wild-type (top line) and the *upf1Δ* (bottom line) strain. *hrb1Δ* n = 6; *gbp2Δ* n = 8; *gbp2Δ hrb1Δ* n = 7; *upf1Δ* n = 6.

To investigate this, a *CBP80^PTC* NMD reporter or its PTC-less counterpart *CBP80* was expressed in wild type, *gbp2Δ hrb1Δ*, the quadrupole knockout *nmd4Δ ebs1Δ gbp2Δ hrb1Δ*, or *upf1Δ* cells. Different from the reporter (**Figure 19**) used in other experiments in this work, both the PTC-less and PTC-containing *CBP80* reporters used here were expressed through the endogenous *CBP80* promoter to better reflect the natural

phenomena. RNA was purified from whole cell lysates and the relative abundance of the NMD reporter in each strain was analyzed via qPCR. The levels of the $CBP80^{PTC}$ reporter were related to that of the PTC-less $CBP80$ reporter to exclude non-NMD effects. We observed, as expected, that the NMD reporter was significantly enriched in cells lacking Upf1 (**Figure 25**). In cells lacking Gbp2 and Hrb1, the reporter level increased about 1.5-fold compared to wild type. The additional deletion of $NMD4$ and $EBS1$, however, appear not to cause further stabilization of the $CBP80^{PTC}$ reporter in comparison to the double knockout strain. This indicates that the four proteins are presumably involved in the same process in NMD and may have redundant roles in facilitating target degradation. In addition, it also suggests that the proteins serve as auxiliary factors in NMD and likely work in concert with several, if not many, other factors to contribute to robust degradation of NMD targets.
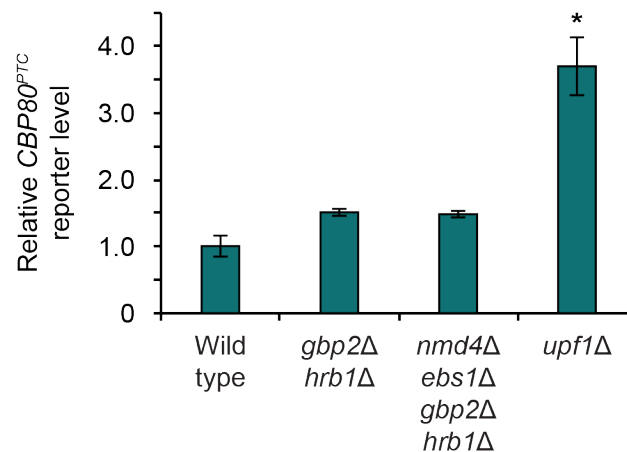


**Figure 25. Deletion of *GBP2*, *HRB1*, *NMD4*, and *EBS1* causes partial defects in NMD**
The $CBP80^{PTC}$ reporter was expressed from the endogenous $CBP80$ promoter in the indicated strains and its relative levels were analyzed via whole cell RNA purification and subsequent qPCR. The results were normalized to the PTC-less $CBP80$ reporter and the endogenous $TDH1$ mRNA to eliminate non-NMD and internal effects. Results were obtained from 3 independent experiments and shown as means with standard deviations.

## 3.5. Gbp2 and Hrb1 may play a role in translational repression on target RNAs subjected to NMD

While an NMD substrate is rapidly subjected to degradation, translation initiation on the RNA is concomitantly suppressed (Muhlrad and Parker, 1999b; Isken et al., 2008; Kurosaki et al., 2019) (see **2.3.6**). This prevents further production of aberrant proteins and is also believed to promote decay. Earlier, we discovered that deletion of *GBP2* and

*HRB1* led to increased translation of intron-containing NMD reporters in an Upf1-dependent manner (Grosse et al., 2021), suggesting that in addition to facilitating degradation, these proteins may directly or indirectly inhibit translation of NMD substrates. Interestingly, a group of RGG domain-containing proteins are known to repress translation by directly binding to the translation initiation factor eIF4G through their RGG motifs (Rajyaguru et al., 2012). Gbp2 and Hrb1 also contain RGG domains (**Figure 4**) and could potentially regulate translation initiation in a similar fashion. It was also suggested that such a repression activity could involve more than one RGG motif translation repressor (Bhatter et al., 2019). Thus, Gbp2 and Hrb1 may act independently or together with known RGG domain-containing translation repressors to inhibit translation of NMD substrates.

To test this hypothesis, co-IP assays were first carried out to study the physical interactions between Gbp2, Hrb1, and two RGG domain-containing translation repressors, Scd6 and Sbp1 (Rajyaguru et al., 2012). Yeast strains that express genomic Scd6-GFP and Sbp1-GFP proteins were used, and growth test showed that the GFP tags did not cause deleterious effects to the cells (**Figure 26**). Scd6- and Sbp1-GFP were immunoprecipitated and co-elution of Gbp2 and Hrb1 was analyzed on western blots. As shown in **Figure 27**, the proteins do not seem to associate with Scd6. On the other hand, while they co-purified with Sbp1 in the presence of RNA, the interaction was largely lost upon RNase treatment. Furthermore, it was later found that the absence of neither Scd6 nor Sbp1 had an effect on the translation of the $CBP80^{PTC}$ reporter transcript (Grosse et al., 2021). These results negate the idea that Gbp2 and Hrb1 regulate translation of NMD substrates together with these two RGG domain-containing translation repressors.

Alternatively, they may require other unknown factors or are themselves specific translation repressors of targeted NMD substrates. Like Scd6 and Sbp1, they may inhibit translation directly by binding translation initiation factors. In fact, physical interactions between Gbp2, Hrb1 and eIF4G as well as eIF4E was observed in subsequent experiments. Cells expressing genomic eIF4G- or eIF4E-GFP proteins, which showed similar growth as wild-type cells (**Figure 28**), were used in co-IP experiments. Co-precipitation of Gbp2 and Hrb1 with these initiation factors was examined via western blot analysis. Both proteins appear to physically interact with eIF4G and eIF4E, as they were co-precipitated both in the presence and absence of RNA (**Figure 29**). This provides a first hint for the possible activity of Gbp2 and Hrb1 as direct translation repressors. However, further studies are required to confirm this role and reveal whether this function is specific for their targeted NMD substrates.
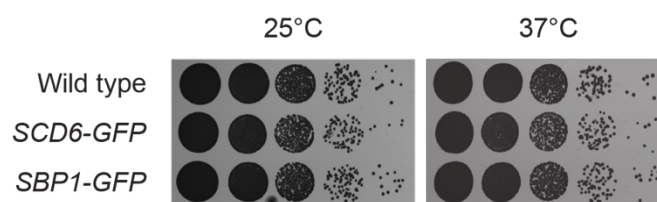
**Figure 26. Fusion of the GFP tag to Scd6 and Sbp1 does not lead to growth defects**
Growth analysis of the *SCD6-* and *SBP1-GFP* strains was carried out. Ten-fold serial dilutions of cell suspensions, from $10^7$ to $10^3$ cells/ml, were spotted on full YPD agar plates and grown at the indicated temperatures for 2 days.
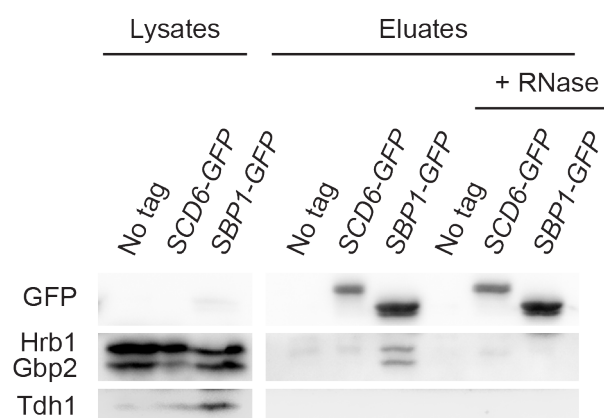


**Figure 27. Gbp2 and Hrb1 weakly associate with Scd6 and Sbp1, two RGG motif-containing translation repressors**
Co-precipitation of Gbp2 and Hrb1 with Scd6- or Sbp1-GFP in the presence and absence of RNase was analyzed on western blots. Tdh1 served as a control for unspecific binding.
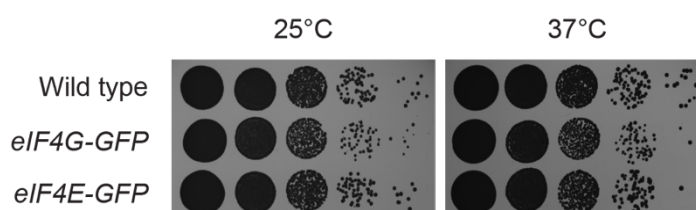


**Figure 28. The *eIF4G-* and *eIF4E-GFP* strains show similar growth as wild type**
Growth analysis of the *eIF4G-* and *eIF4E-GFP* strains was carried out. Ten-fold serial dilutions of the cells, from $10^7$ to $10^3$ cells/ml, were spotted on full YPD agar plates and grown at the indicated temperatures for 2 days.
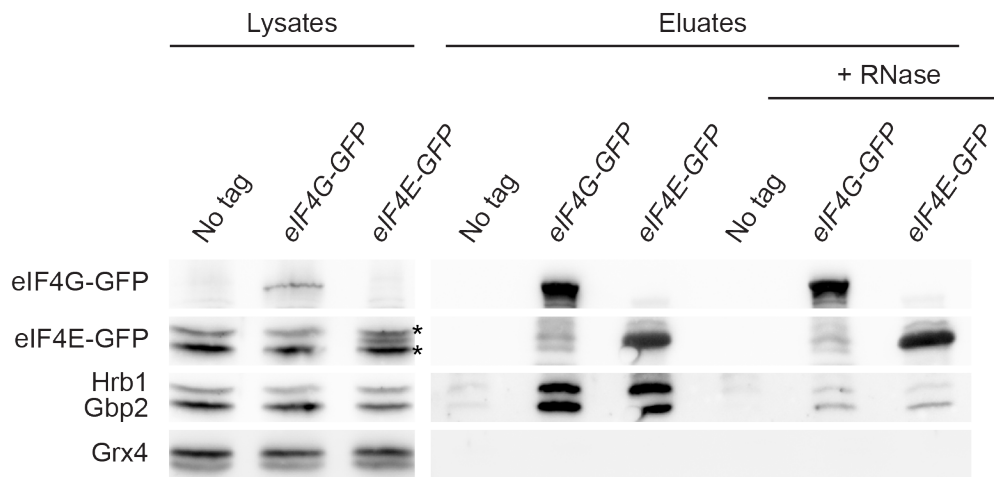
**Figure 29. Gbp2 and Hrb1 physically interact with the cap-binding translation initiation factors eIF4G and eIF4E**

Co-precipitation of Gbp2 and Hrb1 with eIF4G- or eIF4E-GFP in the presence and absence of RNase was analyzed on western blots. Grx4 served as a control for unspecific binding. The asterisks indicate Hrb1 and Gbp2 bands that were detected prior to detection of GFP signals.

## 3.6. Gbp2 and Hrb1 may facilitate NMD-induced mRNP structuring for rapid RNA degradation

### 3.6.1. Gbp2 and Hrb1 help to mediate interactions between the PTC and the 5' end of the transcript

We discovered physical interactions between Gbp2, Hrb1 and diverse NMD-related factors, including Upf proteins (**Figure 12, 16,** and **18**), translation initiation factors (**Figure 29**), and decay factors (Grosse et al., 2021). The main functions of these proteins are carried out at different sites of the target RNA and it is intriguing how their interactions with Gbp2, Hrb1, and with each other could all take place in the course of NMD. It seems plausible that the RNA folds into certain structures to allow these interactions to occur as part of one bigger NMD-mRNP complex, where main and auxiliary factors of early and late steps of NMD come into contact. Such a model would be favorable, as it also provides a basis for understanding how recognition of a PTC in the middle of an RNA could be communicated to the ends of the transcript for control of translation initiation and activation of mRNA decay.

In metazoan, phosphorylated Upf1 at the PTC serves as a binding platform for SMG proteins, which in turn recruit the decay factors (Kurosaki et al., 2019) (see **2.3.5**). Although endonucleolytic cleavage by SMG6 creates unprotected fragments that are immediately accessible to degradation machineries, accelerated 5' decapping and 3'

deadenylation, which take place at the ends of a transcript, are also stimulated (Cao and Parker, 2003). Curiously, the responsible factors, which appear to associate with UPF1 and SMG proteins at the PTC, are apparently able to reach at least the 5' end of the transcript and initiate degradation, while Upf1 is still bound close to the PTC (Serdar et al., 2016). How this is mechanistically achieved is nonetheless unclear. In yeast, endonucleolytic cleavage has not been discovered and degradation seems to occur exclusively from the 5' or 3' end (He and Jacobson, 2015; Karousis and Mühlemann, 2019). Yeast Upf1 was found to be in a complex with Dcp1, Dcp2, and other decapping factors (Dehecq et al., 2018). However, it is also unclear how this complex efficiently contacts the 5' end to mediate decapping and subsequent degradation. Considering the physical interactions between Gbp2, Hrb1 and multiple factors along the NMD pathway, it is tempting to speculate that the two proteins may help in mRNP rearrangements that would connect the ends of the RNA to the PTC-associated Upf1. An example of mRNA structuring mediated by dimerization of RNA-binding proteins has previously been established (Aibara et al., 2017). Moreover, the ability of Gbp2 and Hrb1 to bind each other and to another molecule of themselves was validated through co-IP assays (Grosse et al., 2021). Presumably, multiple copies of Gbp2 and Hrb1 that are located at different positions on the RNA could interact with each other or various other mRNA-binding proteins, thereby promoting formation of secondary mRNP structures that bring distant proteins into close proximity.

To test if Gbp2 and Hrb1 play a role in bridging the distance between Upf1 and the 5' end of the transcript, the physical interaction between Upf1 and the 5' binding eIF4G was studied, and whether the interaction would be impaired in the absence of Gbp2 and Hrb1 was examined. Following precipitation of genomically expressed eIF4G-GFP, co-elution of Upf1-HA was visualized on western blots. We observed an RNase-resistant interaction between Upf1 and eIF4G, which appear to be unaffected by deletion of *GBP2* and *HRB1* (**Figure 30**). To increase NMD events in the cell, the *CBP80^PTC* reporter was additionally overexpressed. Interestingly, co-precipitation of Upf1 with eIF4G was still similar between wild type and *gbp2Δ hrb1Δ* in the presence of RNA, but was severely reduced upon RNase treatment (**Figure 31a** and **31b**). This implies that loading of eIF4G and Upf1 onto the same target RNA is independent of Gbp2 and Hrb1, whereas their direct association seems to be promoted by the two proteins, supporting our hypothesis. In agreement, preliminary result from an eIF4E–Upf1 co-precipitation assay showed a decreased interaction of the two proteins in the *gbp2Δ hrb1Δ* strain compared to wild type (**Figure 32**). We therefore postulate that Gbp2 and Hrb1 contribute to mRNP remodeling upon

activation of NMD. By promoting contact between the PTC and the 5' end of the transcript, they may help to convey the signal of nonsense recognition efficiently to the site of downstream processes.

| | Lysates | | | Eluates | | | + RNase | | |
|---|---|---|---|---|---|---|---|---|---|
| | | *eIF4G-GFP* | | | *eIF4G-GFP* | | | *eIF4G-GFP* | |
| | No tag | Wild type | *gbp2Δ hrb1Δ* | No tag | Wild type | *gbp2Δ hrb1Δ* | No tag | Wild type | *gbp2Δ hrb1Δ* |
| eIF4G-GFP | | | | | | | | | |
| Upf1-HA | | | | | | | | | |
| Hem15 | | | | | | | | | |
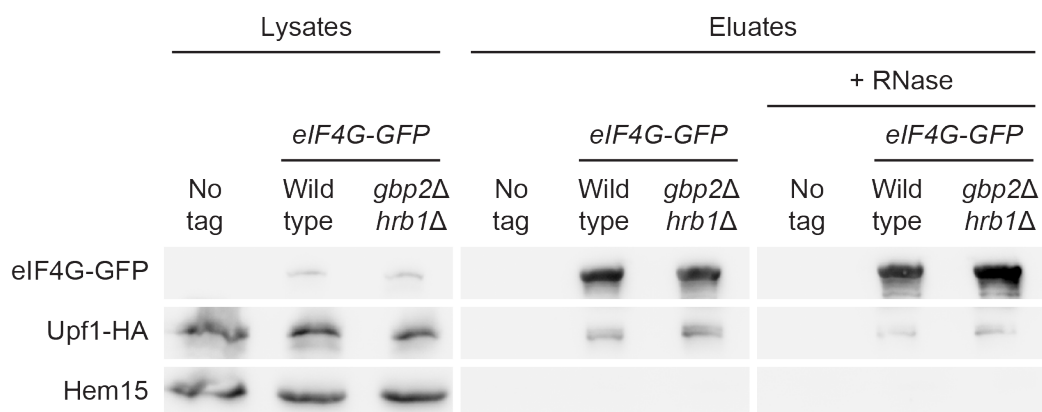
**Figure 30. Upf1 physically interacts with eIF4G in an RNase-insensitive manner**
Co-precipitation of Upf1-HA with eIF4G-GFP in the presence or absence of Gbp2 and Hrb1 was analyzed on western blots. All strains were transformed with a *UPF1-HA* plasmid. Hem15 served as a control for unspecific binding.

**a**

| | Lysates | | | Eluates | | | + RNase | | |
|---|---|---|---|---|---|---|---|---|---|
| | + p$P_{GAL1}$CBP80$^{PTC}$ | | | + p$P_{GAL1}$CBP80$^{PTC}$ | | | + p$P_{GAL1}$CBP80$^{PTC}$ | | |
| | | *eIF4G-GFP* | | | *eIF4G-GFP* | | | *eIF4G-GFP* | |
| | No tag | Wild type | *gbp2Δ hrb1Δ* | No tag | Wild type | *gbp2Δ hrb1Δ* | No tag | Wild type | *gbp2Δ hrb1Δ* |
| eIF4G-GFP | | | | | | | | | |
| Upf1-HA | | | | | | | | | |
| Hem15 | | | | | | | | | |

**b**
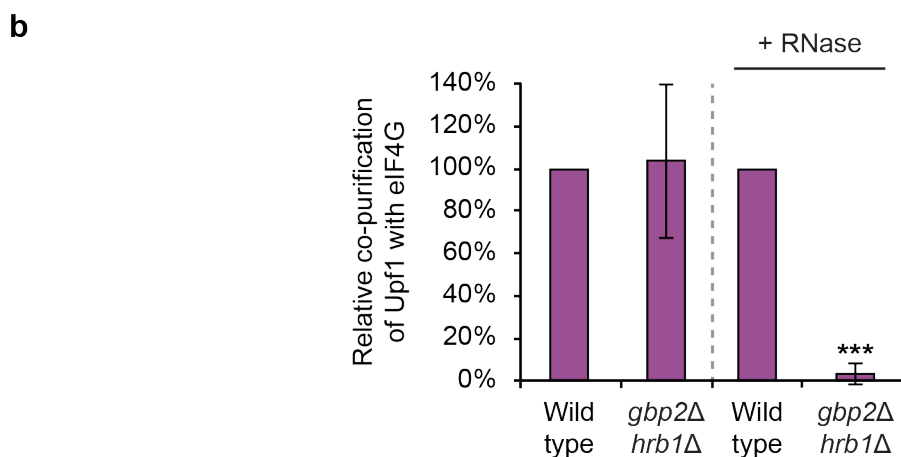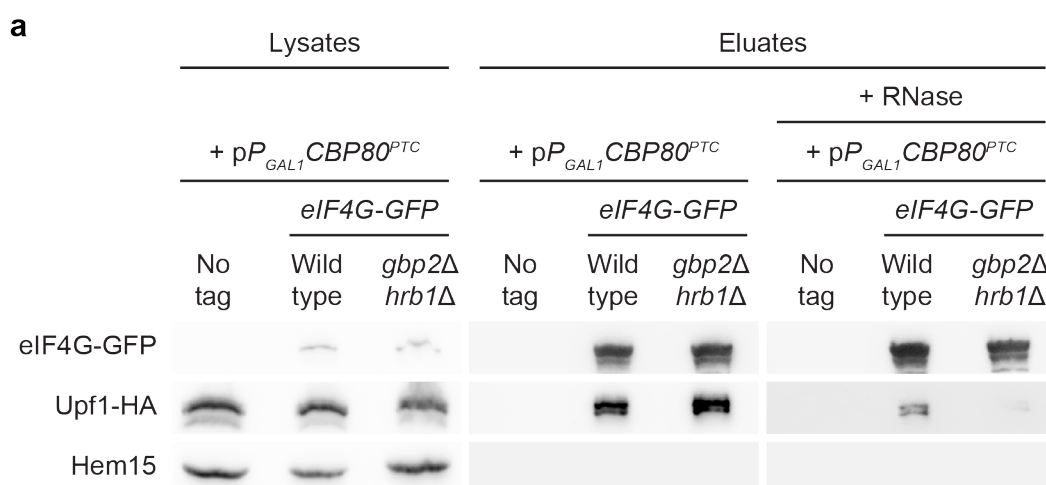
*(see next page for figure legend)*

**Figure 31. Gbp2 and Hrb1 may mediate the direct interaction between eIF4G and Upf1 on NMD targets**

**a.** Co-precipitation of Upf1-HA with eIF4G-GFP in the presence or absence of Gbp2 and Hrb1 was analyzed on western blots. All strains were transformed with a *UPF1-HA* plasmid. The *CBP80^PTC* NMD reporter was expressed upon 2 hours of galactose induction. Hem15 served as a control for unspecific binding. **b.** Relative binding of Upf1 to eIF4G was quantified from **a**. The intensities of the Upf1 bands were related to that of the corresponding eIF4G pull-down signals. Upf1 signals in the non-RNase treated samples were quantified using less-exposed figures. Results are shown as means with standard deviations. No RNase n = 5; + RNase n = 3.

| | Lysates | | | Eluates | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | + RNase | | |
| | | *eIF4E-GFP* | | | *eIF4E-GFP* | | | *eIF4E-GFP* | |
| | No tag | Wild type | *gbp2Δ hrb1Δ* | No tag | Wild type | *gbp2Δ hrb1Δ* | No tag | Wild type | *gbp2Δ hrb1Δ* |
| eIF4E-GFP | | — | — | | ▬ | ▬ | | ▬ | ▬ |
| Upf1-HA | ▬ | ▬ | ▬ | | — | ▬ | | — | — |
| Nop1 | — | — | — | | — | — | | — | — |
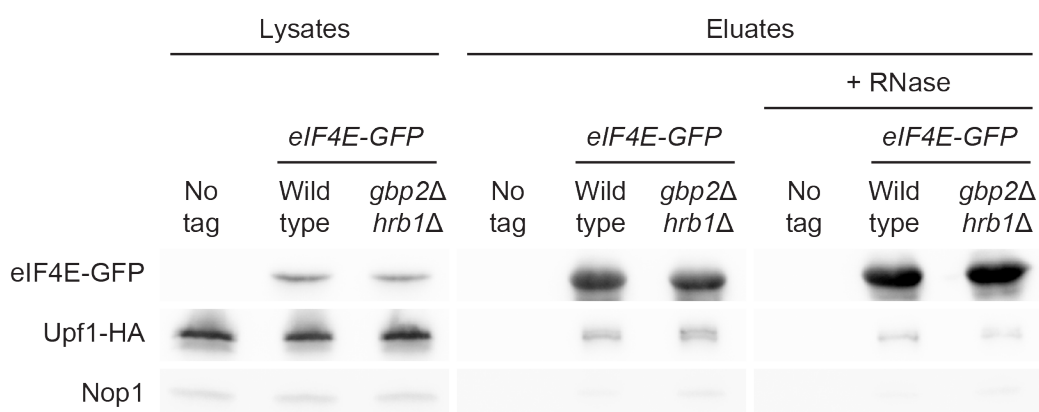
**Figure 32. Direct Upf1–eIF4E interaction may depend partially on Gbp2 and Hrb1**

Co-precipitation of Upf1-HA with eIF4E-GFP in the presence or absence of Gbp2 and Hrb1 was analyzed on western blots. All strains were transformed with a *UPF1-HA* plasmid. Nop1 served as a control for unspecific binding. The experiment was only done once.

## 3.6.2. Communication of PTC recognition to 3' degradation may involve a different mechanism

It was further investigated if a similar mechanism is used for the communication with degradation from the 3' end. Degradation in the 3'–5' direction by the cytoplasmic exosome is preceded by deadenylation, which, in both general mRNA degradation and NMD, is primarily regulated by the poly(A)-binding protein Pab1 (Richardson et al., 2012; Brambilla et al., 2019) (see **2.3.4** and **2.3.5**). Physical interaction between Upf1 and Pab1 has been demonstrated through mass spectrometry and immunoprecipitation assays in both human and yeast (Schell et al., 2003; Richardson et al., 2012). Further data indicated that association of Upf1 with the RRM1 domain of Pab1 is critical for accelerated 3' deadenylation in NMD (Richardson et al., 2012). This interaction may therefore also serve to transfer the PTC signal to the 3' end for the onset of 3' degradation.

To study whether Gbp2 and Hrb1 are involved, it was first checked if there are physical interactions between these proteins and Pab1. A plasmid encoding *PAB1-GFP* was used,

and the functionality of the Pab1-GFP protein was confirmed in growth tests showing that expression of the plasmid rescued the defective growth phenotype of *pab1* temperature sensitive strains (**Figure 33**). GFP-tagged Pab1 was expressed in wild-type cells and immunoprecipitated, followed by the detection of Gbp2 and Hrb1 signals on western blots. Gbp2 and Hrb1 both co-eluted with Pab1, although RNase treatment severely interfered with the interactions (**Figure 34**). This contradicts previous mass spectrometry analyses that showed association of Gbp2 with Pab1 independent of RNA (Collins et al., 2007; Richardson et al., 2012). Possibly, RNA-binding of these proteins promote their direct interaction that favors detection with the less sensitive chemiluminescence-based detection method used here. On the other hand, evidence for a direct Hrb1–Pab1 interaction is lacking. Thus, it remains possible that Hrb1 and Pab1 simply locate on the same mRNA and do not specifically associate.
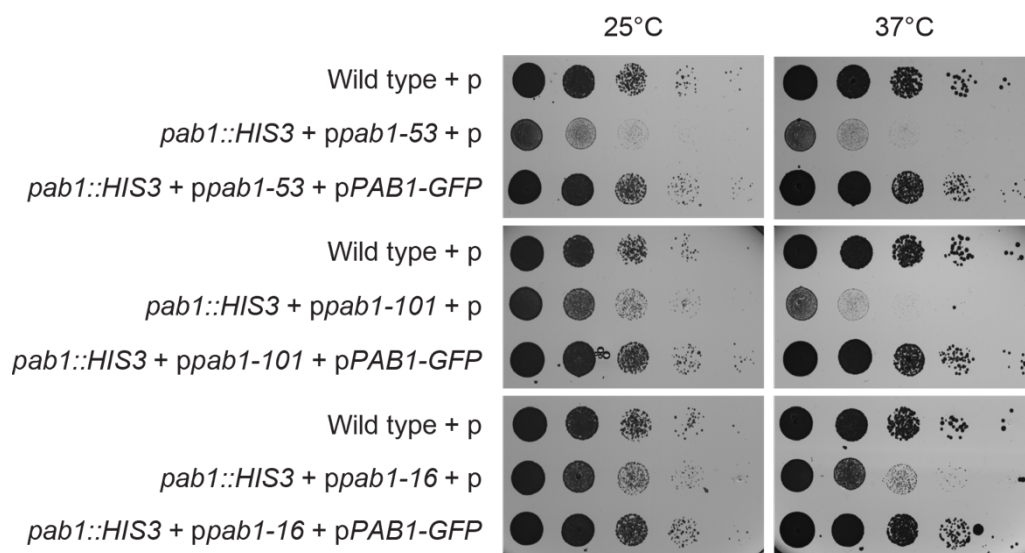


**Figure 33. GFP-tagged Pab1 protein is functional**
Functionality of the plasmid-encoded Pab1-GFP protein was tested in growth analyses. Ten-fold serial dilutions, from $10^7$ to $10^3$ cells/ml, of wild type or *pab1* temperature-sensitive mutant strains transformed with an empty vector (p) or the *PAB1-GFP* plasmid were spotted on selective agar plates and grown at the indicated temperatures for 2 days.

| | Lysates | | Eluates | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | | | | | + RNase | |
| p*PAB1-GFP* | - | + | - | + | - | + |
| GFP | | | | | | |
| Hrb1 | | | | | | |
| Gbp2 | | | | | | |
| Tdh1 | | | | | | |
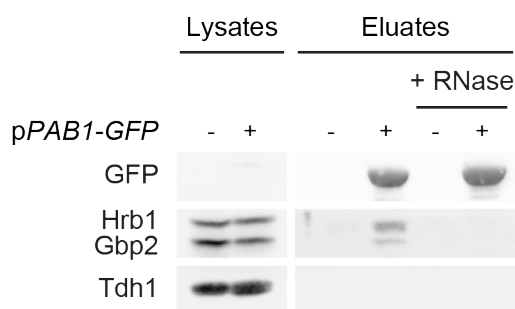
**Figure 34. Gbp2 and Hrb1 physically associate with Pab1 in the presence of RNAs**
Co-precipitation of Gbp2 and Hrb1 with Pab1-GFP in the presence or absence of RNase was analyzed on western blots. Tdh1 served as a control for unspecific binding.

To understand if the possible association between Gbp2, Hrb1, and Pab1 is relevant for their function in NMD, the interaction between Upf1 and Pab1 was studied and whether it would show a dependence on Gbp2 and Hrb1 was investigated. Upf1-HA and Pab1-GFP were both expressed from plasmids in *upf1Δ* cells. Following Pab1 pull-down, co-precipitation of Upf1 was examined. The two proteins physically interact both in the presence and absence of RNA (**Figure 35**). Accordingly, an RNase-insensitive co-purification of yeast Pab1 and Upf1 has been demonstrated (Richardson et al., 2012). However, the Upf1–Pab1 interaction did not seem to be altered when the cells were additionally deleted for *GBP2* and *HRB1* (**Figure 35**). The result was therefore not sufficient to show an NMD-related interaction between Upf1 and Pab1 that may be mediated by Gbp2 and Hrb1. Nevertheless, since 3' degradation is the minor pathway used in yeast NMD and Upf1 may also interact with Pab1 on bulk, normal mRNAs, it remains possible that the effect was too mild to be detected.



| | Lysates | | | Eluates | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | | + p*PAB1-GFP* | | | + p*PAB1-GFP* | | | + RNase | |
| | | | | | | | | + p*PAB1-GFP* | |
| | No tag | Wild type | *gbp2Δ hrb1Δ* | No tag | Wild type | *gbp2Δ hrb1Δ* | No tag | Wild type | *gbp2Δ hrb1Δ* |
| Pab1-GFP | | | | | | | | | |
| Upf1-HA | | | | | | | | | |
| Hem15 | | | | | | | | | |
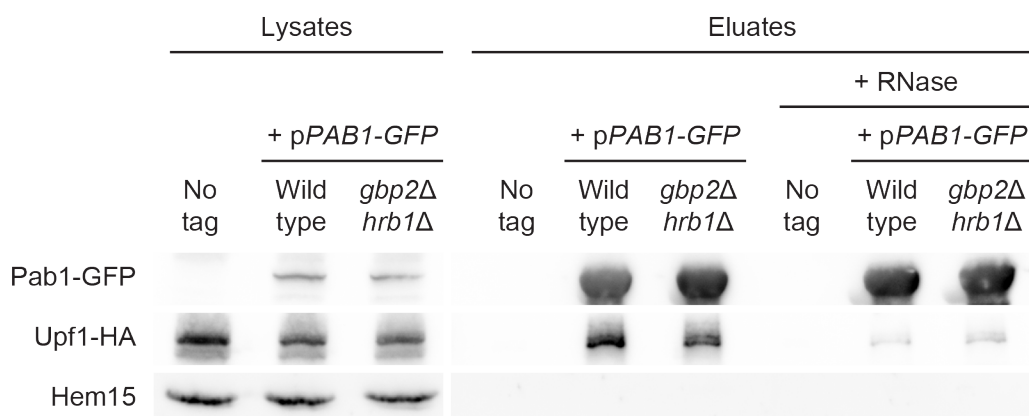
**Figure 35. Pab1 physically contacts Upf1 independently of Gbp2 and Hrb1**
Co-precipitation of Upf1-HA with Pab1-GFP in the presence or absence of Gbp2 and Hrb1 was analyzed on western blots. All strains were deleted for *UPF1* and expressed *UPF1-HA* from a plasmid. Hem15 served as a control for unspecific binding.

The upf1-DE572AA mutant was therefore used and the interaction between mutant upf1 and Pab1 was investigated. We reasoned that 3' degradation, hence association of the PTC with the 3' end, might be enhanced when 5' degradation cannot be completed. A possible effect of Gbp2 and Hrb1 on the Upf1–Pab1 interaction might thus be more evident using the mutant upf1. Co-IP experiments were carried out as above using cells that expressed *UPF1-* or *upf1-DE572AA-HA*. In addition, the *CBP80^{PTC}* reporter was overexpressed to increase NMD. Nevertheless, co-purification of upf1-DE572AA with Pab1 was likewise unaffected by deletion of *GBP2* and *HRB1* (**Figure 36a** and **36b**). This indicates either that the effects were still too transient for detection or that communication between the core NMD factors and the 3' end relies on a different mechanism.

**a**



**b**
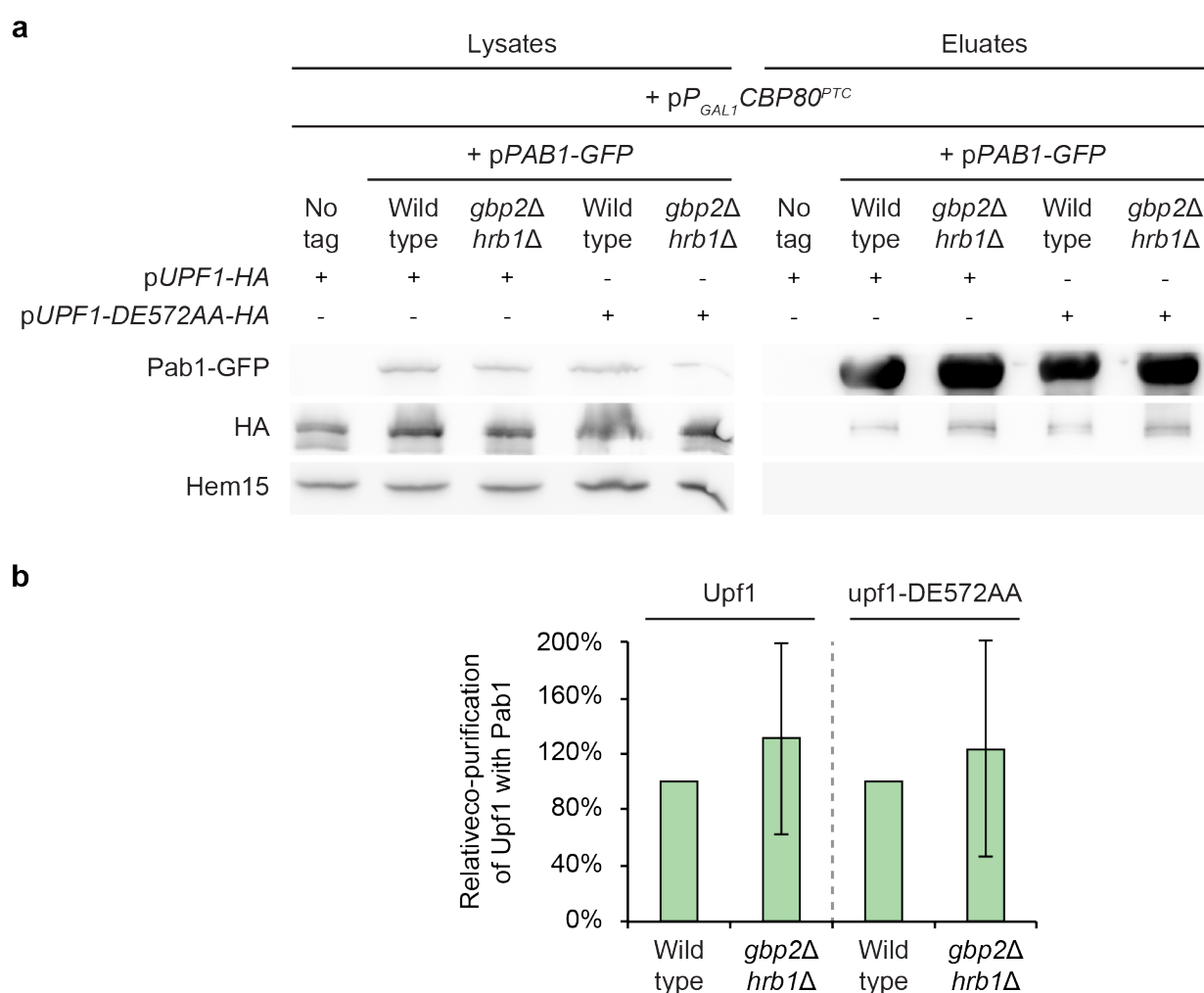


**Figure 36. Gbp2 and Hrb1 appear dispensable for the Pab1-Upf1 interaction**
**a.** Co-precipitation of Upf1- or upf1-DE572AA-HA with Pab1-GFP in the presence or absence of Gbp2 and Hrb1 was analyzed on western blots. All strains were deleted for *UPF1* and expressed wild-type or mutant Upf1 from plasmids. The *CBP80^{PTC}* reporter was expressed upon 2 hours of galactose induction. Hem15 served as a control for unspecific binding. **b.** Relative binding of wild-type or mutant Upf1 to Pab1 shown in **a** was quantified. The intensities of the Upf1 bands were related to that of the corresponding Pab1 pull-down signals. Results were obtained from four independent experiments and are shown as means with standard deviations.

Taken together, the results shown in this study provide evidences for the involvement of Gbp2 and Hrb1 in the cytoplasmic mRNA quality control pathway NMD and begin to unravel mechanistic details regarding their possible functions therein. They appear to act downstream of substrate recognition by Upf1 and are relevant for target degradation as well as translational repression. Remarkably, they seem to continue their quality control roles, initially identified in nuclear events, in the cytoplasm and may thus demonstrate a conserved connection between cellular compartments in eukaryotic species.

# 4. Discussion

## 4.1. Gbp2 and Hrb1 are involved in NMD

### 4.1.1. Cytoplasmic in addition to nuclear quality control

Over the past decades, identification and functional characterization of the yeast SR-like proteins have demonstrated how these proteins, like human SR proteins, can play diverse roles in mRNA metabolism and therefore contribute to the maintenance of transcriptome integrity. Nevertheless, our understanding of the functions of these shuttling proteins is as yet incomplete, and especially little is known regarding the cytoplasmic function of Gbp2 and Hrb1. The results from this study, in combination with other findings, reveal the involvement of these proteins in the cytoplasmic mRNA quality control pathway NMD. Deletion of *GBP2* and *HRB1* led to accumulation of PTC-containing NMD reporters both on the RNA and protein level in an Upf1-dependent manner (Grosse et al., 2021). This was likely a result of less efficient degradation and reduced translation inhibition after NMD activation, as the absence of Gbp2 and Hrb1 did not affect the stable binding of Upf1 to the *CBP80$^{PTC}$* reporter (**Figure 21**), but restricted its association with decapping factor Dcp1 (Grosse et al., 2021), Ski complex factor Ski2 (**Figure 23**), and the cap-binding translation initiation factor eIF4G (**Figure 31**), potentially hindering a speedy response to silence the faulty transcript.

Since Gbp2 and Hrb1 were shown to be involved in the quality control of pre-mRNA splicing and are adaptor proteins for the mRNA export receptor Mex67–Mtr2 (Hackmann et al., 2014), we needed to take into account that the nuclear functions of these proteins may indirectly affect downstream events that occur in the cytoplasm. In regard to splicing, microarray analyses had shown that deletion of *GBP2*, *HRB1*, or both genes did not significantly alter the ratio of pre-/mature mRNA levels of intron-containing genes (Kress et al., 2008), whereas the ratio evidently increased when a known bona fide splicing factor was defective. In the same study, the other SR-like protein, Npl3, showed contrastingly an impact on splicing of numerous genes. Unlike Npl3, which interacts with and presumably promotes the recruitment of spliceosome assembly factors to support splicing (Kress et al., 2008), Gbp2 and Hrb1 associate mainly with late splicing factors and probably have a more specific role in detecting errors during the process rather than affecting the process itself (Hackmann et al., 2014). These findings suggest that in cells deleted for *GBP2* and *HRB1*, splicing should be carried out efficiently, at least to a similar extent as in wild-type cells, and

therefore should not cause additional splicing defects that may potentially result in an increase of natural NMD targets. In addition, although these proteins promote mRNA export by recruiting Mex67–Mtr2 to the transcript, their absence did not result in general defects in bulk mRNA export (Hackmann et al., 2014). Thus, in cells lacking Gbp2 and Hrb1, processed mature mRNAs, including the reporter transcripts that were used in our experiments, should be normally transported into the cytoplasm.

Gbp2 and Hrb1, however, are known to promote degradation of transcripts in case of splicing defects. Since the *CBP80^PTC* reporter contains an intron and therefore also undergoes splicing, Gbp2 and Hrb1 may affect degradation of *CBP80^PTC* in the nucleus. Thus, to study decay of the reporter transcript through the NMD pathway, as in **Figure 25**, we could not rely on mRNA half-life measurements, as we could not easily distinguish nuclear and cytoplasmic effects. We analyzed instead the steady-state level of the transcript, and to exclude nuclear effects, used a PTC-less but otherwise identical version of the reporter as control. Effects related to the intron, splicing, and nuclear decay should be eliminated by normalizing to the PTC-less reporter. Since a PTC is only read in the cytoplasm during translation, any difference between the PTC-containing and PTC-less reporters should reflect consequences from NMD in the cytoplasm. Yet, depletion of Gbp2 and Hrb1 was shown to cause leakage of incorrectly spliced transcripts into the cytoplasm (Hackmann et al., 2014). This raises the possibility that some intrinsic splicing errors, which could potentially lead to PTCs, may escape nuclear surveillance in *gbp2Δ hrb1Δ* cells and increase the amount of endogenous NMD targets in the cytoplasm. These additional targets may sequester NMD factors or decay factors, thereby leading to indirect effects when we examine the association of these factors with the *CPB80^PTC* reporter (as in **Figure 21** and **24**). These results thus should be interpreted with care.

### 4.1.2. Gbp2 and Hrb1 affect a subset of NMD targets

Gbp2 and Hrb1 appear to play a role in NMD on a subgroup of targets and not all of them. In the beginning of our study, the *PGK1*-based NMD reporter (*PGK1^PTC*) was initially analyzed, but its steady-state level showed no dependence on Gbp2 and Hrb1, leading us to conclude that the proteins are not involved in the NMD of this reporter (Grosse et al., 2021). Although it is a well-established reporter frequently used in yeast, the *PGK1* gene does not contain an intron. Considering the functional relationship between Gbp2, Hrb1 and splicing, we tested two NMD reporters derived from the intron-containing *DBP2* and *CBP80* genes. Gbp2 and Hrb1 indeed contributed to the destabilization of these two reporters via

the NMD pathway (Grosse et al., 2021). Similar results were discovered on the protein level, showing that translation of the two intron-containing reporters were in part regulated by Gbp2 and Hrb1 in an Upf1-dependent manner, while that of the *PGK1^PTC^* reporter was not (Grosse et al., 2021). Based on these findings, it appears plausible that Gbp2 and Hrb1 have a role in NMD specifically on targets that are derived from intron-containing genes. This would agree with previous findings that Gbp2 and Hrb1 are enriched on transcripts of intron-containing genes compared with those derived from intron-less genes (Hackmann et al., 2014). In comparison, Npl3 did not show preferential binding regarding the presence of introns and seems to bind all mRNAs equally. This also correlates with the cellular abundance of these SR-like proteins: there are approximately 4,000 molecules of Gbp2 or Hrb1 in the cell, but more than 40,000 molecules of Npl3 (Ho et al., 2018), at least 10-fold that of the two other proteins. These observations further support a specific link between Gbp2, Hrb1, and intron-containing transcripts. As Gbp2 and Hrb1 participate in the quality control of splicing, and splicing defects can often lead to PTCs, it is tempting to speculate that Gbp2 and Hrb1 continue their quality control function on spliced mRNAs to ensure removal of problematic transcripts that may escape or are not recognized by nuclear quality control (see **4.8**).

In yeast, a relatively low proportion of genes contain introns, and most of them have only one intron (Spingola et al., 1999), while metazoan genes typically contain several introns and therefore undergo more extensive and complex splicing events (Mourier and Jeffares, 2003; Kim et al., 2007; Park and Graveley, 2007). Together with the lack of major EJC components and thus an independence from EJC-mediated NMD activation, a link between splicing and NMD was thought to be rather unimportant in yeast (Boisramé et al., 2019; Kurosaki et al., 2019). Nevertheless, an estimated 27% of the mRNAs produced by the yeast cell are derived from intron-containing genes, indicating that spliced transcripts make up a larger part of the yeast transcriptome than may be expected (Lopez and Séraphin, 1999; Ares et al., 1999). Also, several reports have demonstrated that incorrect splicing of pre-mRNAs is a main source of PTC-containing transcripts that are targeted to NMD in yeast (He et al., 1993; Sayani et al., 2008; Celik et al., 2017a). According to one study, about 16% of intron-containing pre-mRNAs potentially become NMD substrates (Celik et al., 2017a). These data suggest that transcripts derived from intron-containing genes are an essential part of yeast NMD, and shed light on the potential importance of Gbp2 and Hrb1 in this pathway.

In light of these ideas, we conducted our study further using the intron-containing reporters. In this work, we mainly focused on the *CBP80^PTC* reporter, as its 5' proximal intron position is more representative of yeast intron-containing transcripts (**Figure 19**). Contrarily, the intron in the *DBP2* gene is situated in the second half of the ORF, relatively atypical for yeast. That said, it remains possible that Gbp2 and Hrb1 take part in NMD on targets based on different criteria. To assess the link between intron-containing substrates and the two proteins, we could construct intron-less versions of the *DBP2^PTC* and *CBP80^PTC* reporters and examine whether Gbp2 and Hrb1 would still affect the stability of these reporters through NMD. Reciprocally, insertion of an intron into the *PGK1^PTC* reporter may also provide some insight.

The observed effects of Gbp2 and Hrb1 in NMD were not only found on the constructed reporters, but also relevant for endogenous NMD targets. This was demonstrated in the Gbp2 and Hrb1 RIP experiments, in which co-precipitation of two intron-containing, putative endogenous target mRNAs with Gbp2 and Hrb1 decreased in the absence of Upf1 (**Figure 22c**2–3). The same results were found for the *CBP80^PTC* reporter, showing less binding of Gbp2 and Hrb1 when Upf1 was depleted (**Figure 22c**1). These findings suggest an earlier dissociation of the proteins from target RNAs when NMD was not activated, agreeing with them having NMD-related functions on these RNAs. The milder effects for the endogenous targets in comparison to the reporter transcript probably point out that these endogenous mRNAs are not always PTC-containing and thus not always NMD targets. In support of this, it was suggested that intron identity and splicing signals, which eventually influence splicing efficiency, can affect NMD targeting of transcripts that contain introns (Sayani et al., 2008). The selected endogenous targets may be spliced at different efficiencies and thus be affected by NMD to different extents. To verify Gbp2 and Hrb1's function in physiological NMD, it may be helpful to analyze additional intron-containing as well as intron-less endogenous targets and check if the effects correlate with splicing efficiencies. Detection of the protein–NMD substrate interactions in *upf2Δ* or *upf3Δ* strains may also serve to strengthen a connection between NMD and the altered RNA-binding of these proteins.

Apart from these, Gbp2 and Hrb1's engagement in endogenous NMD was also demonstrated through co-IP experiments. In the Gbp2–Upf1 co-IP experiment (**Figure 16**), an increased interaction between the proteins was detected when the ATPase of Upf1 was defective. In another experiment, a decrease of the Ski2–Upf1 interaction in the absence of Gbp2 and Hrb1 was observed (**Figure 23**). These experiments were done in cells that did not express the NMD reporter, therefore the results should represent effects that originate

from endogenous NMD events. The relatively mild effect observed in the former experiment matches the idea that the proteins act on a subpopulation of endogenous targets. Together, these data agree with a role of Gbp2 and Hrb1 in the NMD pathway on a subset of targets under natural conditions. To gain further insight into how Gbp2 and Hrb1 participate in NMD, we may benefit from transcriptome-wide or other high-throughput methods that would allow us to globally identify the substrates that they affect and analyze the extent of their impact.

Previous research had also shown that several factors involved in NMD are not always required for the pathway. For example, the EJC, known to be important for NMD activation, is limited to transcripts that are spliced, which do not represent all NMD targets. The core factors Upf2 and Upf3 appear to be important, or necessary, at least for the early stage of NMD, but several reports in higher eukaryotes have demonstrated that they are dispensable in some cases (Lavysh and Neu-Yilik, 2020). The SMG5–SMG7 and SMG6 proteins were likewise suggested to have redundant roles in facilitating target decay (Colombo et al., 2017; Boehm et al., 2021). These data exemplify how the NMD pathway may be supported by different sets of proteins in addition to the core factors. Which proteins are involved presumably depends on the substrate type and the cellular condition in each case.

## 4.2. Gbp2 and Hrb1 are part of the NMD mRNP

To find out if Gbp2 and Hrb1 directly participate in the NMD process, we investigated the physical interactions between these proteins and the Upf factors. Co-IP of the two SR-like proteins with Upf1, Upf2, and Upf3 suggested that these proteins can be found on a common set of RNAs (**Figure 12**). Co-precipitation was however not detected when samples were additionally treated with RNase, suggesting RNA-dependent or indirect interactions. Similarly, it was shown that assembly of the yeast Upf1–Upf2–Upf3 complex requires RNA (Dehecq et al., 2018), demonstrating that formation of direct protein–protein contacts in the pathway may occur only in the presence of RNA. That Gbp2 and Hrb1 may affect just a subset of NMD targets (see **4.1.2**) is possibly also reflected in the difficulty to detect strong physical interactions between these proteins and factors involved in NMD. Taking into consideration the relatively low abundance of Upf proteins – about 7,000 molecules/cell of Upf1 and about 2,000 molecules/cell of Upf2 and Upf3 each (Ho et al., 2018) –, detection of Upf2 and Upf3 in NMD complexes is especially challenging. This likely also accounted for failures in previous attempts (Gavin et al., 2006; Krogan et al., 2006; Collins et al., 2007).

Gbp2, Hrb1, and the Upf proteins may also be part of the same mRNP complex and do not necessarily interact directly. Split-GFP experiments demonstrated that at least Gbp2 closely associates with Upf1 on NMD substrates *in vivo*. GFP fluorescence, which would indicate the close proximity of Upf1 and Gbp2, increased as the amount of NMD substrates was enriched in the cell (**Figure 17** and **18**). Besides overexpressing the *CBP80*<sup>PTC</sup> reporter to increase substrate levels, we also disrupted the NMD-triggered decay pathway by depletion of the main 5'–3' exonuclease Xrn1, thereby accumulating substrates that could not be efficiently degraded. Under wild-typical conditions, NMD substrates are presumably degraded rapidly upon recognition, likely leading to the low-abundant and transient nature of physical interactions that are formed during the process. Indeed, it had been shown in yeast that many NMD-sensitive RNAs are eliminated within a few minutes (Guan et al., 2006). As there is also an inherent delay in the appearance of split-fluorescence signals due to the time required for complementation and maturation of the chromophore (Romei and Boxer, 2019), too transient interactions may not allow the assembly of a functional fluorescence protein in time. By preventing or slowing down degradation, we were likely able to capture more of those short-lived interactions and visualize them through the methods that were applied.

This was also demonstrated by using the ATP hydrolysis mutant upf1-DE572AA in both co-IP and split-GFP analyses. The ATPase activity of Upf1 is required for removal of RNA secondary structures as well as dissociation of the ribosome and other bound proteins for full target degradation (Franks et al., 2010; Fiorini et al., 2015; Serdar et al., 2016). This mutant upf1 therefore causes incomplete degradation and results in accumulation of 3' decay intermediates on which the ribosome, the upf1-DE572AA protein, and several other NMD factors appear to concomitantly remain (Franks et al., 2010; Serdar et al., 2016). Further analysis of the resulting 3' RNA decay intermediates surprisingly revealed that the 5' end of these fragments align to sequences several nucleotides downstream of the PTC (Serdar et al., 2020). Based on additional supporting data, a model was proposed in which the ATP hydrolysis mutant upf1 interferes with Rli1's activity to separate ribosomal subunits. The ribosome migrates into the 3' UTR of the NMD substrate, where it blocks 5'–3' degradation by Xrn1 (Serdar et al., 2020). The increased association of Gbp2 with upf1-DE572AA shown in both co-IP (**Figure 16**) and split-GFP experiments (**Figure 18**) suggests that Gbp2 is likely part of the Upf1-containing decay fragment mRNP and one of the factors that dissociates from the target depending on proper function of Upf1. This supports NMD-related association of Gbp2 with Upf1 and argues for NMD-specific functions of this protein.

Consistent with these results and providing evidence for the physical involvement of Hrb1 in NMD, it was discovered that enrichment of NMD targets in Xrn1-depleted cells shifted the localization of both Gbp2 and Hrb1 towards the cytoplasm in an Upf1- and PTC-dependent manner (Grosse et al., 2021), whereas the two proteins normally localize to the nucleus at steady-state (Windgassen and Krebber, 2003; Häcker and Krebber, 2004). This illustrates a further link between the NMD pathway and the two proteins, in particular in how disruption of the degradation process probably affects the association of the proteins with the target mRNP. In line with this, Gbp2 and Hrb1 appear to dissociate later from putative NMD substrates when NMD is activated (**Figure 22c**). Collectively, these findings support the physical involvement of Gbp2 and Hrb1 in the NMD pathway on certain substrates, at least during the degradation process.

Interestingly, in both experiments that showed an increased association of Gbp2 with the mutant upf1, Hrb1 did not show the same phenotype (**Figure 16** and **20**). This indicates that, unlike Gbp2, Hrb1 is probably a more peripheral component and no longer present in the NMD mRNP retained by upf1-DE572AA. Since elevated substrates that could not be degraded from the 5' end delayed the re-import of Hrb1 (see previous paragraph), we speculate that Hrb1 completes its function(s) early in the pathway and dissociates from the substrate shortly prior to or upon degradation by Xrn1. Supporting this, we observed that Hrb1, but not Gbp2, promotes the binding of the decapping factor Dcp1 to the $CBP80^{PTC}$ reporter (Grosse et al., 2021; see **4.4**), which acts before Xrn1. Therefore, Hrb1 seems to participate until the early stage of target degradation, while Gbp2 remains bound for a longer time. These data show first evidences of a functional divergence between the two highly homologous proteins, at the same time revealing how they may contribute to a common pathway through different means.

Potentially transient or RNA-dependent interactions such as those studied in this work call for methods that can better capture biological complexes within short time frames and in their real-time physiological conformation. One approach that could be implemented is cross-linking. As an example, formaldehyde cross-linking locks interactions between proteins and nucleic acids and has been used widely in molecular biological research (Hoffman et al., 2015; Srinivasa et al., 2015; Ramanathan et al., 2019). By applying this in a co-IP experiment, protein–protein interactions of interest may be preserved while the RNA-binding proteins retain their RNA-bound state. RNase treatment would subsequently remove protein-free segments of RNA and thereby prevent co-precipitation of non-interacting proteins from the same transcript. Our attempt with this method in the Upf IP

experiment did not show co-elution of Gbp2 and Hrb1 with the Upf proteins (**Figure 13**), although formaldehyde cross-linking assisted the detection of RNase-insensitive interactions between Dcp1 and Gbp2 as well as Hrb1 (Grosse et al., 2021). Either the SR-like proteins weakly interact with the Upf proteins, or there are technical issues that need to be solved in the experimental set-up requires further investigation. It should be noted that formaldehyde-mediated reactions are strongly dependent on reaction time, formaldehyde concentration, and several other variables (Hoffman et al., 2015), and therefore requires optimization for each experiment. Besides formaldehyde cross-linking, UV irradiation is regularly used to cross-link protein and RNA molecules (Ramanathan et al., 2019). Although the cross-linking efficiency is lower, the resulting irreversible covalent bonds allow stringent washing in co-precipitation assays, which would favor detection of genuine and direct interactions. This approach also enabled us to analyze the binding of Ski2, Dcp1, and Xrn1 to the $CBP80^{PTC}$ reporter transcript in different mutants (**Figure 24**; Grosse et al., 2021). For future investigation, the more sensitive mass spectrometry analysis could be done for Gbp2 and Hrb1 that are purified from $xrn1\Delta$, $dcp1\Delta$, or cells expressing the ATP hydrolysis mutant upf1. Accumulation of NMD mRNPs in these cells would support the identification and verification of protein interactions involved in those complexes.

Despite the challenge to identify physical interaction partners of Gbp2 and Hrb1 within the NMD pathway, the accumulated results indicate that the proteins are physically associated with NMD target mRNPs. These co-transcriptionally recruited proteins accompany mRNAs into the cytoplasm and are presumably released upon normal translation or, in the case of NMD activation, proper completion of target degradation.

## 4.3. NMD activation – new perspectives

Although the association of human SR proteins with EJCs and the indication that Gbp2 and Hrb1 may promote NMD on transcripts that are derived from intron-containing genes made it tempting to postulate that the two proteins also help anchor Upf proteins to NMD substrates like the EJCs, our study suggested that this is probably not the case. In the absence of both Gbp2 and Hrb1, Upf1 was able to associate with the $CBP80^{PTC}$ reporter transcript as well as it did in wild-type cells (**Figure 21**). On the contrary, Upf1 failed to stably associate with the reporter RNA in the absence of Upf2, consistent with previous findings in human cells and the indication that UPF2 is important for NMD target recognition (Kurosaki et al., 2014). This suggests that substrate recognition by Upf1 and consequent

NMD activation on the *CBP80^PTC* reporter RNA is largely independent of Gbp2 and Hrb1. Additionally, we could analyze endogenous NMD substrates in the same experiment to find out if similar effects could be observed. Further, it may be tested if the absence of Gbp2 and Hrb1 would alter the physical interactions between the three Upf proteins, i.e., the formation of the Upf1–Upf2–Upf3 complex, one of the key steps to NMD activation. Although the two SR-like proteins did not appear to affect Upf1's binding to a substrate transcript, it remains possible that they promote recruitment of Upf2 and/or Upf3 and thus facilitate NMD activation. Nevertheless, as we were not able to show direct physical interactions between these two proteins and Upf2 or Upf3 (**Figure 12** and **13**), a major role of these proteins in the Upf1–Upf2–Upf3 complex formation seems relatively unlikely.

In yeast, the most prominent NMD activation mechanism is described by the long 3' UTR model (see *2.3.3.2*). Considering that most yeast introns are situated close to the 5' end of the transcript, PTCs that arise from incorrect splicing would also be relatively 5' proximal and thus far away from the poly(A) tail. This naturally creates a PTC in a long 3' UTR context that might sufficiently favor NMD activation and therefore neglects the need for facilitation from additional *cis*-acting factors. On the other hand, it is so far not completely understood how a long 3' UTR actually leads to activation of NMD. The original hypothesis, which emphasizes the pro-termination interaction between Pab1 and eRF3 as a key determinant, has been challenged and has led to the quest for alternative or additional mechanisms. Gbp2 and Hrb1, despite their apparent activities in supporting NMD, probably do not play significant roles there.

Recently, a new explanation for NMD activation has emerged that aims to combine current findings and provide a more general perspective on the subject matter. This model suggests that NMD activation, like many other cellular events, is the outcome of competition between different molecular signals, reminiscent to the model proposed for nuclear RNA surveillance that was described in **2.1.1**. According to these models, all transcripts, both in the nucleus and the cytoplasm, are targeted to decay by default, but evade degradation if they could engage timely in other functional processes (Bresson and Tollervey, 2018; Kishor et al., 2019b). In the case of NMD, the competition between various pro-termination and pro-NMD features would ultimately determine the fate of a transcript. An EJC downstream of the stop codon would represent a strong NMD-inducing feature, while the loss of eRF3–Pab1 interaction and the presence of Hrp1 bound to DSE may represent milder, but effective NMD-inducing features. Any other factor that could influence the molecular environment or

the kinetics of translation termination and Upf1 activation may all contribute to a shift towards or against commitment to the NMD pathway (Kishor et al., 2019b).

In light of this, NMD activation may be categorized into roughly two situations: for transcripts with a strong NMD-activating feature, e.g. an EJC downstream of the stop codon, termination is evidently aberrant and NMD likely occurs upon the first round of translation; for transcripts with mild NMD-inducing features, as may be the case for many normal-looking endogenous targets, translation terminates in a less optimal environment but may still succeed, and NMD might be activated only after several rounds of translation (He and Jacobson, 2015; Karousis and Mühlemann, 2019). This led to a new notion that NMD may occur stochastically and may monitor every mRNA at each round of translation. Supporting this is a recent study that examined individual NMD events in human cells using a single-molecule imaging approach and estimated that each ribosome has a probability of 0.11 to induce NMD (Hoek et al., 2019).

The new model provides a basis for understanding why almost all of the prominent NMD-inducing features identified so far has been found to be dispensable in certain cases of NMD. It would also explain why many endogenous NMD targets seem only moderately affected by NMD (Colombo et al., 2017; **Figure 22c**3). In a broader sense, Gbp2 and Hrb1 may also serve as weaker pro-NMD factors that work in concert with other signals to eventually lead to degradation of the target. For instance, they may have an impact on the mRNP structure and contribute to a termination environment that is in better favor of NMD. Nevertheless, this would arguably not be their main function and not be significant for yeast NMD activation. In addition, it was shown that dissociation of Gbp2 and Hrb1 from mRNAs is dependent on the import receptor Mtr10, and presumably followed directly by nuclear import of these proteins (Windgassen and Krebber, 2013; Häcker and Krebber, 2004). Therefore, although it is possible that displacement of Gbp2 and Hrb1 from NMD target transcripts uses a somehow different mechanism, we assume that Gbp2 and Hrb1-mediated NMD occurs principally in the first round of translation.

## 4.4. Gbp2 and Hrb1 facilitate target RNA degradation

If Gbp2 and Hrb1 are not critical for the activation of NMD, their impact on reporter RNA levels may reflect an involvement in transcript degradation. Several experiments have implicated these proteins as part of the mRNP that is undergoing NMD-activated degradation (see **4.2**). Moreover, Gbp2 and Hrb1 physically interact with factors of the

cytoplasmic RNA turnover machineries, including the decapping factor Dcp1, the exonuclease Xrn1, and the exosome cofactor Ski2 (Grosse et al., 2021), supporting their role in cytoplasmic transcript decay. Interaction of Gbp2 and Hrb1 with the helicase Ski2 of the Ski complex is reminiscent of their binding to Mtr4, which is the helicase component of the nuclear exosome cofactor TRAMP complex and in this sense a nuclear counterpart of Ski2. Gbp2 and Hrb1 were suggested to help recruit the TRAMP complex through interactions with Mtr4 and thereby guide the nuclear exosome to faulty transcripts during quality control of splicing (Hackmann et al., 2014). Similarly, efficient recruitment of Ski2 to the NMD substrate seems to depend on Gbp2 and Hrb1. Following detection of interactions between Ski2 and the two proteins, we observed that Upf1 also co-purified with Ski2, and this physical interaction significantly decreased upon depletion of Gbp2 and Hrb1, both in the presence and absence of RNA (**Figure 23**). Assuming that this interaction essentially occurs on transcripts that are targeted for NMD, the result implies that recruitment of the Ski complex to NMD substrates is likely hindered in the absence of Gbp2 and Hrb1. Given that Ski2 is required for the 3' degradation of NMD substrates in yeast (Mitchell and Tollervey, 2003), reduced Ski2 association implies less effective degradation by the exosome. Examination of this phenomenon through protein–RNA interaction, in our case the association of Ski2 with the $CBP80^{PTC}$ reporter, revealed a reduced binding of the protein to the substrate RNA in cells lacking Gbp2 and Hrb1 (**Figure 24**). This corresponds to the decreased interaction between Ski2 and Upf1. However, as noted earlier, the result of this experiment needs to be assessed carefully because we cannot exclude the possibility that increased natural NMD substrates in the cytoplasm may sequester Ski2 molecules and cause an indirect decrease in the binding of Ski2 to the $CBP80^{PTC}$ reporter. Nevertheless, the previous co-IP experiments and the nuclear TRAMP complex-recruiting activities of Gbp2 and Hrb1 support their involvement in the recruitment of Ski2 to the NMD substrate. The RIP result may indeed reflect the same effect.

Single deletion of *GBP2* or *HRB1* did not affect the binding of Ski2 to $CBP80^{PTC}$ (**Figure 24**), inferring that both Gbp2 and Hrb1 facilitate recruitment of the Ski complex. To obtain additional evidence, we could test the single knockout strains in the Ski2–Upf1 co-IP experiment. Furthermore, to comprehensively analyze the impact of Gbp2 and Hrb1 on 3' degradation, we could examine other factors of the Ski complex, the Ski7 protein, and components of the exosome. As the cytoplasmic exosome requires the Ski complex and Ski7 to be recruited to the target RNA (Araki et al., 2001; Halbach et al., 2013), it could be expected that interactions between Upf1 and exosome factors, if there are any, likewise depend partially on Gbp2 and Hrb1.

The two proteins may also affect 3' decay at an earlier stage, namely during deadenylation. Although the major 5' degradation of yeast NMD targets was shown to be independent of deadenylation (Muhlrad and Parker, 1994), degradation from the 3' end still requires that the poly(A) tail is previously removed (Mitchell and Tollervey, 2003). This likely involves deadenylase complexes common to canonical mRNA turnover, and Pab1, a major regulator of transcript 3' end stability (see **2.3.4**). In higher eukaryotes, the SMG5–SMG7 proteins also recruit deadenylases to support 3' decay (Loh et al., 2013). Association of Gbp2 and Hrb1 with components of the deadenylase complexes is not known. However, the proteins were shown to interact with Pab1 (Collins et al., 2007; Richardson et al., 2012; **Figure 34**), although the functional significance is so far unclear. Interestingly, both Gbp2 and Upf1 appear to bind to the RRM1 domain of Pab1 *in vivo*, and the Upf1–Pab1 interaction was implicated in accelerated Pab1 removal, which promotes deadenylation and 3' decay (Richardson et al., 2012). Whether there is competitive or sequential binding of Gbp2 and Upf1 with Pab1 and whether there are relevant roles of Gbp2 and Hrb1 in deadenylation of NMD targets remain to be investigated. Methods that allow detection of the RNA 3' end, such as northern blot and qPCR analyses, could be carried out in 5' decay-deficient cells to directly examine 3' NMD in the presence and absence of the SR-like proteins.

In addition to 3' decay, Gbp2 and Hrb1 were also found to facilitate the 5' degradation pathway (Grosse et al., 2021). Both Gbp2 and Hrb1 promote proper targeting of the *CBP80^{PTC}* reporter to Xrn1. However, they likely do not directly support the exonuclease, but act upstream of it: Hrb1 seems to help recruit the decapping machinery to the transcript and thereby ensures target accessibility to Xrn1, while Gbp2 probably supports effective decapping through its translation repression activity (see **4.5**), which acts beforehand (Grosse et al., 2021). Collectively, our data allowed us to conclude that Gbp2 and Hrb1 facilitate NMD substrate elimination through recruitment of the 5' and 3' decay machineries. While both Gbp2 and Hrb1 may assist the Ski–exosome complex, Dcp1 recruitment appear to be specifically promoted by Hrb1. These activities parallel their nuclear function in supporting exosome-mediated degradation of incorrectly spliced transcripts and also suggest some functional similarity with the human SMG5–SMG7 proteins.

Two other yeast proteins, Nmd4 and Ebs1, were suggested to be functional homologs of SMG6 and SMG5–SMG7, respectively (Luke et al., 2007; Dehecq et al., 2018). In addition to their physical interactions with Upf1, with decapping factors, and their functional implications in NMD target degradation (He and Jacobson, 1995; Luke et al., 2007; Dehecq et al., 2018), these proteins contain 14-3-3 and PIN domains that resemble those of the

SMG5–7 proteins (Luke et al., 2007; Dehecq et al., 2018). As none of Gbp2, Hrb1, Nmd4, or Ebs1 alone was necessary or sufficient for fully active NMD, we tested if these proteins may complement each other and altogether be required for substrate degradation. Nonetheless, neither the double knockout of *GBP2* and *HRB1* nor the quadruple knockout of all four proteins stabilized the *CBP80^PTC* reporter to the extent of the *UPF1* knockout strain (**Figure 25**). This indicates that despite these proteins contribute to target decay, their absence does not completely abolish the decay activities in the NMD pathway. The degradation process likely relies on additional auxiliary factors, or Upf1 alone has the capacity to eventually trigger decay. Further, it suggests different possibilities regarding how the four examined proteins may participate in target degradation. As Nmd4 and Ebs1 appear to take part in NMD also for a subgroup of targets (Dehecq et al., 2018), they may not be relevant for NMD of the *CBP80^PTC* reporter. If, however, they are indeed involved in NMD on this transcript, they may have overlapping functions with Gbp2 and Hrb1 in regard to promoting degradation. Alternatively, Gbp2 and Hrb1 may act upstream of Nmd4 and Ebs1, therefore deletion of *NMD4* and *EBS1* had no further effect in addition to *gbp2Δ hrb1Δ*. The identification of other factors and the interplay between these auxiliary proteins would be interesting subjects for future research. Also, studies that address the ability of Nmd4 and Ebs1 to recruit decay machineries would allow us to further compare their mechanisms with the SMG5–SMG7 proteins.

Through a series of affinity purification-coupled mass spectrometry analyses, it was found that yeast Upf1 appear to modulate NMD through two distinct and sequential complexes (Dehecq et al., 2018). The first complex contains Upf1, Upf2, and Upf3 (named "Upf1-23") and is responsible for recognition of NMD targets. Subsequently, the second complex (named "Upf1-decapping"), which includes Upf1, decapping factors, Nmd4, Ebs1, but not Upf2 and Upf3, mediates decapping and decay. Gbp2 was identified in the Upf1 purifications, but also co-purified with Dcp1, Nmd4, and Upf2 in their analyses (Dehecq et al., 2018). It is thus ambiguous whether it is part of the Upf1-23 or the Upf1-decapping complex. Based on our findings, it is tempting to consider that Gbp2 is part of the latter complex that carries out target degradation. Hrb1 also weakly co-purified with Upf1 (Dehecq et al., 2018) and may be part of the Upf1-decapping complex, although conceivably as a more peripheral component. Since the proteins are thought to be present on target transcripts already before activation of NMD, it is also not unlikely that they are part of both Upf1-23 and Upf1-decapping complexes. Together, the results provide interesting insights into the cytoplasmic roles of Gbp2 and Hrb1. Their activities in promoting NMD target decay

establish their overall quality control character and their recruitment of decay machineries, reminiscent of SMG5–SMG7, suggests a further level of conservation in the NMD pathway.

## 4.5. Gbp2 and Hrb1 may also exhibit translation repression activities

Besides removal of error-prone transcripts, the important goal of mRNA quality control mechanisms is to prevent the production of aberrant proteins. This is also reflected in the NMD pathway, as deletion of *UPF1* leads to the accumulation of protein products translated from NMD target transcripts (Muhlrad and Parker, 1999b; Kim et al., 2017). Although enrichment in protein levels can also result from RNA stabilization and reduced proteasomal degradation of the translated polypeptides, which were both shown to be consequences of inhibiting NMD, there is evidence that translation is directly affected (Kervestin and Jacobson, 2012; see **2.3.6**).

Deletion of *GBP2* and *HRB1* resulted in increased protein products of the intron-containing NMD reporters, although to a lesser extent when compared with *UPF1* deletion (Grosse et al., 2021). The effects were Upf1- and PTC-dependent, therefore specific to the NMD pathway. Protein level per reporter mRNA level was calculated to estimate translation rates, which increased in *gbp2Δ hrb1Δ* cells, suggesting the possibility that Gbp2 and Hrb1 suppress translation of NMD substrates in addition to influencing substrate RNA degradation. One efficient way to inhibit translation is to target translation initiation, as it is the first step of the process. Further, most eukaryotic translation initiation depends on the 5' cap and it's binding proteins, leading to an intrinsic link between inhibition of translation initiation and 5' decapping and decay. In view of this and considering Hrb1's role in promoting recruitment of the decapping complex (see **4.4**), it seems likely that Gbp2 and Hrb1 together affect both translation initiation and subsequent 5' degradation of NMD substrates.

Several members of the group of RGG motif-containing proteins were found to exhibit translation repression activities in yeast (Segal et al., 2006; Rajyaguru et al., 2012). These proteins, which include Npl3, Scd6, and Sbp1, appear to interfere with translation initiation through their direct binding to eIF4G (Rajyaguru et al., 2012), a component of the cap-binding eIF4F complex. Moreover, both their binding to eIF4G as well as their ability to repress translation of reporter mRNAs were shown to be dependent on their RGG motifs (Rajyaguru et al., 2012). The fact that Gbp2 and Hrb1 also contain RGG motifs and the finding that they physically interact with eIF4G and another factor of the eIF4F complex,

eIF4E (**Figure 29**), allowed us to consider similar roles of these proteins in translation repression during NMD. We also considered the possibility that these proteins do not suppress translation directly, but facilitate the process by supporting one or more of the identified RGG domain-containing translation repressors mentioned above. However, Gbp2 and Hrb1 did not co-immunoprecipitate with Scd6 and did not appear to directly interact with Sbp1 (**Figure 27**), despite the implication of a Gbp2–Sbp1 interaction from a genome-wide mass spectrometry analysis (Gavin et al., 2006). Gbp2 and Hrb1 have been found to physically associate with Npl3, but this likely occurs already in the nucleus during their co-transcriptional recruitment to the nascent mRNA (Inoue et al., 2000; Hurt et al., 2004; Tardiff et al., 2007; Erce et al., 2013). Furthermore, deletion of neither *NPL3*, *SCD6*, nor *SBP1* resulted in an increase of Cbp80$^{PTC}$ protein level as in *gbp2Δ* and *upf1Δ* cells (Grosse et al., 2021). These findings suggest that Gbp2 and Hrb1 unlikely work together with the three proteins for translational inhibition in NMD and may specifically and directly repress translation of their targeted NMD substrates. That said, it remains possible that Gbp2 and Hrb1 cooperate with other yet unknown factors to modulate translation in NMD or are somehow engaged in the proteasomal degradation pathway. It also does not exclude the probability that Npl3, Scd6, or Sbp1 may be involved in translation repression of other NMD substrates.

Very recently, a study showed that Gbp2, similar to the three translation repressors described above, can directly bind to eIF4G via its RGG domain and represses the translation of reporter mRNAs in an RGG motif-dependent manner (Poornima et al., 2021). This corroborates our findings and supports the function of Gbp2 as a translation repressor. Their study focused on general mRNAs and the examined reporters do not contain introns and are normally not NMD targets. Gbp2 affected translation of those reporters without influencing mRNA levels (Poornima et al., 2021), different from what we observed for the NMD reporters. Presumably, Gbp2 can act generally as a translation repressor in the cytoplasm, but have additional functions on NMD substrates to facilitate transcript decay. It remains to be demonstrated whether translation repression by Gbp2 on NMD-sensitive and -insensitive transcripts operates through the same mechanism. Moreover, it would be interesting to find out if Hrb1 also have similar activities. In any event, the current studies have expanded the cellular functions of Gbp2 and further illustrates the versatility of the SR protein family in regulating gene expression.

## 4.6. Gbp2 and Hrb1 may contribute to mRNP architecture in NMD

Despite successful identification of most main NMD factors, it is still open to question how premature termination, taking place in the middle of the transcript, is physically linked to degradation, occurring from the two transcript ends in yeast. The core factor, Upf1, was found to interact with numerous factors involved at either the target recognition or degradation stage of NMD, but it seems to stay associated with the terminating ribosome and close to the PTC throughout the pathway (Franks et al., 2010; Serdar et al., 2016). One reasonable possibility that could explain this discrepancy is that RNA folding enables spatial proximity between Upf1 and the ends of the transcript during NMD. Such a conformation would favor a fluent transition from the initial PTC recognition phase into subsequent actions carried out by different factors, while all still mediated by Upf1. Folding and compaction is an inherent nature of RNA molecules, and the dynamic structural organization of RNPs is known to serve functional purposes (Khong and Parker, 2020). Previous studies have suggested contact between Upf1 and both ends of the transcript. It was shown in human cells that CBP80 interacts with UPF1 and promotes the formation of the SURF and DECID complexes to stimulate NMD (Hwang et al., 2010; Maquat et al., 2010a). In the yeast system, Upf1 was found to interact with Pab1 and the interaction was shown to support 3' decay (Richardson et al., 2012).

We propose that Gbp2 and Hrb1 contribute to the proper formation of the NMD-mRNP structure in light of several observations. First, the two proteins are likely part of the NMD-mRNP from target recognition until degradation, as discussed previously. Second, these proteins associate with the Upf proteins, as well as factors of the 5' and 3' decay machineries. Third, Gbp2 and Hrb1 were also found to physically interact with eIF4G, eIF4E and Pab1, demonstrating contacts with the 5' and 3' ends of the transcript. Fourth, these proteins can form hetero- or homomers with each other and themselves (Grosse et al., 2021), while dimerization of another yeast RNA-binding protein, Nab2, was previously demonstrated to contribute to transcript compaction and mRNA structure (Aibara et al., 2017). Npl3 had also been shown to dimerize, which is required for ribosome subunit joining to form monosomes for translation initiation (Baierlein et al., 2013). Fifth, depletion of Gbp2 and Hrb1 resulted in an evident reduction of the RNase-insensitive interaction between Upf1 and eIF4G (**Figure 31**). This effect was more evidently observed after we overexpressed the *CBP80^PTC* reporter to increase NMD substrates in the cells (**Figure 30** and **31**), suggesting that the effect is NMD-specific. Since eIF4G binds most transcripts for translation initiation and Upf1 also associates with transcripts regardless of NMD activation,

it is not surprising that eIF4G and Upf1 co-elute in the presence of RNAs. During NMD-induced translation repression and degradation, however, they may come into direct or close contact, which appeared to depend on Gbp2 and Hrb1. Together, these findings support a possible role of Gbp2 and Hrb1 in the remodeling and organization of the NMD-mRNP structure, which may contribute to an efficient coordination of upstream and downstream events, also compatible with their subsequent activities in translation repression and target degradation. Further, dimerization, or potentially oligomerization, of RNA-binding proteins may underlie one of the molecular mechanisms for which RNA folding and compaction is achieved during NMD, although this requires further confirmation.

The interaction between Upf1 and Pab1, on the contrary, did not appear to be affected by depletion of Gbp2 and Hrb1 (**Figure 35**). Given that the 3' decay pathway has a minor role, we may only be able to detect an effect after enriching NMD, for example by overexpressing the reporter construct. The mutant upf1-DE572AA that restricts 5' degradation also associated similarly with Pab1 regardless of Gbp2 and Hrb1 (**Figure 36**). In this experiment, however, the samples were not treated with RNase, and therefore may not reflect the direct interactions that we would like to address. Alternatively, the mRNP structure of the 3' UTR might not necessarily be arranged through the same mechanism as the 5' end. As Gbp2 and Hrb1 were found mostly 5' proximal on the transcript (Tuck and Tollervey, 2013; Baejen et al., 2014), the possibility to form dimers and support RNA folding in the 3' UTR may be limited. Upf1 may directly contact Pab1, independent of additional factors, or the 3' pathway involves other yet unknown factors. Further experiments are required to clearly define the organization of the 3' as well as the 5' mRNP conformation during NMD. To visualize mRNP architecture *in vivo*, traditional electron microscopy (EM) or the more modern cryo-EM analyses may be useful. Single-molecule fluorescence *in situ* hybridization (smFISH) assays that use differently labeled probes for specific regions of the transcript might also allow a more straightforward observation of the mRNP structure and its dynamics (Khong and Parker, 2018).

## 4.7. Nuclear and cytoplasmic mRNA quality control by Gbp2 and Hrb1: a model

Combining the presented data, a model for the function of Gbp2 and Hrb1 in NMD is proposed and depicted together with their role in nuclear mRNA quality control in **Figure 37**.
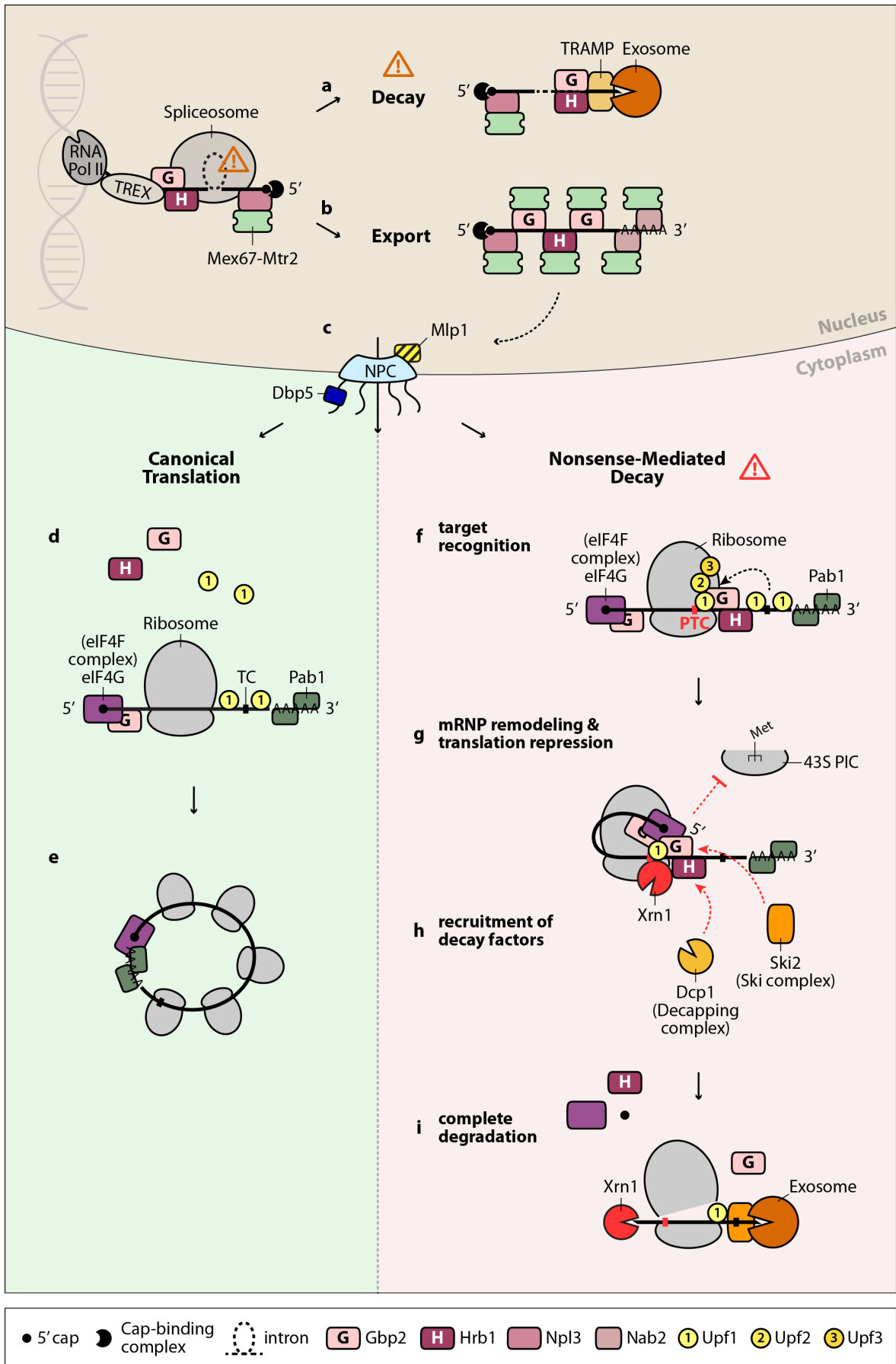
**Figure 37. Model for the mRNA quality control functions of Gbp2 and Hrb1**

Gbp2 and Hrb1 take part in mRNA quality control both in the nucleus and the cytoplasm. For details refer to the main text in section **4.7**. RNA Pol II: RNA polymerase II; TREX: transcription/export complex; NPC: nuclear pore complex; TC: termination codon; PTC: premature termination codon; 43S PIC: 43S pre-initiation complex.

In the nucleus, Gbp2 and Hrb1 are loaded onto the nascent transcript through the THO/TREX complex and associate with late splicing factors of the spliceosome. They are enriched on intron-containing mRNAs and are mostly found towards the 5' end of the transcript. In nuclear mRNA quality control, they recruit the TRAMP complex through interactions with the helicase component Mtr4, which, upon errors in splicing, guides the faulty transcript to the nuclear exosome for degradation (**Figure 37a**). In the event of correct splicing, Gbp2 and Hrb1 recruit instead the export receptor heterodimer Mex67–Mtr2, supporting proper packaging of the mature mRNP (**Figure 37b**), which is quality controlled by Mlp1 and other NPC-associated proteins before the mRNP is transported through the NPC (**Figure 37c**). Gbp2 and Hrb1 enter the cytoplasm together with the transcript and remain bound until translation, while Mex67–Mtr2 and Nab2 are removed by Dbp5.

In the cytoplasm, mRNAs undergo translation during which proteins that bind to the ORF, including Upf1, Gbp2, and Hrb1, are displaced by the ribosome (**Figure 37d**). Correct ORFs offer favorable environments at the stop codon that allow translation to terminate in an efficient manner. Formation of a closed-loop structure mediated by the Pab1–eIF4G interaction protects the transcript ends from exonucleolytic attack and supports efficient rounds of subsequent translation, which further counteracts decay (**Figure 37e**).

On PTC-containing transcripts, translation terminates less efficiently. A combination of NMD-inducing features results in the stable association of Upf1 with the ribosome and the formation of the Upf1–Upf2–Upf3 complex, marking the activation of the NMD pathway (**Figure 37f**). Once the substrate is committed to the NMD pathway, downstream events likely occur concurrently or at least very rapidly. Gbp2 and Hrb1, potentially by forming dimers (or oligomers) with each other or with other RNA-binding proteins, facilitate mRNP remodeling that brings the 5' end of the transcript into proximity with Upf1, allowing direct contact of Upf1 with eIF4G (**Figure 37g**). Gbp2 and Hrb1 also binds to eIF4G, and at least Gbp2 represses translation initiation of the target transcript. Hrb1 promotes recruitment of Dcp1 (decapping complex), while both proteins promote recruitment of Ski2 (Ski complex) (**Figure 37h**). Nmd4 and Ebs1 may also be part of the mRNP and facilitate target degradation (not depicted).

Once the transcript is decapped, Xrn1, which is presumably already associated with Upf1, degrades the RNA from the 5' end (**Figure 37i**). Hrb1 dissociates from the mRNP prior to or soon after the start of Xrn1-mediated degradation. The ATPase/helicase activity of Upf1 is required for ribosome disassembly and displacement of bound NMD factors, which allows complete elimination of the transcript by Xrn1. Gbp2 is also removed from the mRNP at this step. Deadenylation and exosome-mediated decay from the 3' end can also occur, but is the minor pathway in yeast NMD.

Dissociation of Gbp2 and Hrb1 from mRNAs and their re-import to the nucleus is dependent on the import receptor Mtr10 and likely supported by Sky1-mediated phosphorylation of their SR/RGG domains (not depicted). Further details of the mechanisms thereof await to be defined.

## 4.8. Gbp2 and Hrb1 provide dual layers of quality control

In the nucleus, Gbp2 and Hrb1 help to couple co-transcriptional processing, quality control, and nuclear export of mRNA. Their cytoplasmic functions in NMD and translation repression extends the connection of mRNA metabolic events beyond the nuclear envelope. Such nucleus–cytoplasm connections had in fact been established earlier. For instance, nuclear capping, polyadenylation, and the proteins that bind to the resulting 5' cap and poly(A) tail structures are required for nuclear export and later critical for translation. Npl3, the other yeast SR-like protein, is known to be involved in both transcription and translation processes. More analogous to Gbp2 and Hrb1, the metazoan EJC and the associated SR proteins link pre-mRNA splicing to export and further to translation and cytoplasmic quality control. Correlation of nuclear and cytoplasmic events appears to be a common theme in the mRNA life cycle. It likely enables communication between different processes and, importantly, provides the basis for various levels of regulation and surveillance. Gbp2 and Hrb1, like Npl3 and their human relatives, appear to be part of this system, and the dynamic mRNP composition that they take part in has a strong influence on the fate of the transcript.

The continued quality control function of Gbp2 and Hrb1 in the cytoplasm offers an extra level of protection and could in fact play a crucial role as a fail-safe mechanism for the cell (Sayani and Chanfreau, 2012). It was found that quite a surprising proportion of intron-containing transcripts escape the nucleus and are targeted to NMD in the cytoplasm (Sayani et al., 2008; Celik et al., 2017a). Nuclear quality control in higher eukaryotes was likewise found to be leaky and insufficient for the elimination of all error-prone transcripts (Eberle

and Visa, 2014). This rationalizes the need for multiple surveillance systems in both cellular compartments, especially because cytoplasmic quality control detects errors in the open reading frame, which can be neglected in the nucleus. Perhaps based on a similar rationale, transcript decay involves functional redundancies by often triggering both the 5' and 3' exonucleolytic machineries, and in certain cases an addition of endonucleolytic activities, to degrade a single target. In fact, earlier work that looked into the interacting network of yeast and human degradation systems corroborated the joint action of multiple decay machineries in cellular processes (Eberle and Visa, 2014). Either in canonical turnover or quality control, the cell appears to apply mechanisms with certain overlapping activities to ensure effective removal of cellular waste and strive for a careful maintenance of genome integrity.

## 4.9. Gbp2 and Hrb1 as prototypes of human proteins

The accumulated effort in the functional characterization of Gbp2 and Hrb1 revealed multiple similarities between these proteins and the human SR proteins. In the nucleus, they are recruited to the transcript during transcription and are involved in splicing (Wegener and Müller-McNicoll, 2019; see **2.2.1**). Human SR proteins were found to regulate the splicing process per se, while Gbp2 and Hrb1 were found to act as surveillance factors. Through splicing, the proteins become preferentially associated with transcripts derived from intron-containing genes (Hackmann et al., 2014). Further, some SR proteins, like Gbp2 and Hrb1, act as adaptor proteins for the export receptors and facilitate mRNA export (Müller-McNicoll et al., 2016). In the cytoplasm, both Gbp2, Hrb1 and SR proteins participate in NMD and are also implicated in regulating translation. Gbp2 can act as a translation repressor (Poornima et al., 2021), while some SR proteins were shown to either suppress or enhance translation of specific mRNAs (Swartz et al., 2007; Kim et al., 2014). An additional correlation is provided by the discovery of homologous proteins that phosphorylates these proteins (the kinases Sky1 and SRPK1, respectively) (Siebel et al., 1999) and mediates their import (the import receptors Mtr10 and TRN-SR, respectively) (Kataoka et al., 1999). Despite these similarities, currently identified functions of these proteins in NMD are not completely identical: SR proteins are mainly involved by generating substrates through regulation of splicing and by promoting activation of the pathway through the EJC-dependent mechanism (Wegener and Müller-McNicoll, 2019). Gbp2 and Hrb1, on the other hand, contribute mostly in processes downstream.

In view of their activities in promoting target degradation, Gbp2 and Hrb1 seem to resemble the SMG5–SMG7 proteins in human NMD. However, the identification of Nmd4 and Ebs1 as potential homologs of SMG5, SMG6, and SMG7 based on domain structural as well as functional similarities (Dehecq et al., 2018) suggests that Gbp2 and Hrb1 are probably less related to the SMG proteins. That said, the finding that these auxiliary proteins assist target decay in NMD like SMG proteins in higher eukaryotes implies that the pathway may rely on more conserved mechanisms between species than previously thought.

In light of these observations and the potential that human SR proteins still have so far unidentified activities in NMD, it is perhaps reasonable to consider Gbp2 and Hrb1 as the prototype of the EJC. Although present results do not support strict functional equivalences between these factors, Gbp2 and Hrb1 appear to serve as molecular "marks" for nuclear splicing and create a "memory" that can be linked to NMD in the cytoplasm, analogous to EJCs. It is conceivable that a less distinguished form of the EJC suffices for the relatively simple genome and modest intronome in yeast. As the complexity of gene expression increased through evolution, the functions of Gbp2 and Hrb1, which appear to be more general and primitive, may have elaborated into the larger family of SR proteins and their more specialized and specific activities in splicing and in NMD, accommodating also more possibilities for intricate regulation. In addition, the complex nature of alternative splicing results in high rates of errors (Jaillon et al., 2008; Saudemont et al., 2017), which may have prompted a stronger link between splicing and NMD to improve the stringency of surveillance, implemented through EJC's direct role in NMD activation. Future studies are expected to reveal more similarities as well as differences between the yeast SR-like proteins and their human counterparts, which will offer a more complete picture on how the significance and versatility of these RNA-binding proteins have developed early but continued to advance through evolution.

## 4.10. Concluding remarks and future aspects

Through this work, the cytoplasmic functions of yeast SR-like proteins Gbp2 and Hrb1 were identified for the first time. These functions are related to translation, in agreement with their association with polysomes (Windgassen et al., 2004). They are involved in NMD, presumably for transcripts derived from intron-containing genes, indicating that they safeguard transcripts beyond the nucleus in a more comprehensive manner. It also demonstrates that the yeast NMD pathway requires the assistance of previously

unidentified auxiliary factors and that the participation of these proteins may reflect an underlying mechanism of this well-conserved quality control system.

Along with the new insights, several questions opened up and other unsolved topics also remain. For example, the fact that Gbp2 and Hrb1, as well as Nmd4 and Ebs1, appear to affect a subset of NMD targets suggests that there may be more accessory proteins involved if not Upf1 alone is sufficient to eventually mediate completion of the pathway. The full list of yeast NMD factors and their functions, the coordination of substrate type with specific auxiliary factors, the interplay between Gbp2, Hrb1, and different factors, and finally the incorporation of findings across species into a universal mechanism of NMD are subjects that remain to be further investigated. Another interesting aspect of concern is the possible relevance of post-translational modifications (PTMs) of Gbp2 and Hrb1 to their functions in NMD and beyond. The activities of human SR proteins are known to be substantially regulated by PTMs (Wegener and Müller-McNicoll, 2019). These include the most common and significant phosphorylation of their SR/RS domains, but also methylation of arginine residues and some cases of acetylation among other less frequent modifications. Unsurprisingly, phosphorylation and methylation of Npl3 have been linked to its functions both in the nucleus and cytoplasm and to its cellular localization (Yun and Fu, 2000; Gilbert et al., 2001; Gilbert and Guthrie, 2004; Dermody et al., 2008; Yeh et al., 2021). Deletion of *SKY1*, the gene that encodes the yeast SR protein kinase, and the SR/RGG domains of Gbp2 and Hrb1 resulted in mislocalization of the proteins to the cytoplasm (Windgassen and Krebber, 2003; Häcker and Krebber, 2004; unpublished data, laboratory of Heike Krebber). Although it seems quite plausible that PTMs may play crucial roles in the nuclear and cytoplasmic functions of Gbp2 and Hrb1, evidence for these mechanisms is currently lacking. When and how PTMs are linked to the two proteins' shuttling life cycles are so far poorly explored areas of research.

In addition to mRNA quality control, Gbp2 and Hrb1 may potentially be involved in other processes. Transcriptome-wide analysis of yeast RNPs strikingly revealed that Gbp2 is also strongly associated with small nucleolar RNAs (snoRNAs) and long non-coding RNAs (lncRNAs) (Tuck and Tollervey, 2003). Later, several reports demonstrated that lncRNAs that engage in translation may be targeted to NMD, and this is functionally significant for regulation of gene expression (Smith and Baker, 2015; Wery et al., 2016; Colombo et al., 2017; Cai et al., 2020). These findings, in combination with our results, raise the idea that Gbp2 might participate in the NMD-mediated control of lncRNA stability. Further, Gbp2 may have a more general role in translational regulation. It was found that Gbp2 accumulates in

stress granules or distinct cytoplasmic foci under different types of cellular stress (Buchan et al., 2008; Poornima et al., 2021), which is a common feature for factors that are involved in translation inhibition and RNA decay. Recent experimental data illustrated its activity as a translation repressor (Poornima et al., 2021), supporting it likely having a more widespread effect on cytoplasmic gene expression. The function of Gbp2 on non-coding RNA species, the multiple tasks of Gbp2 as well as their associated cellular locations, and the role of Hrb1 in these or other unique activities are some of the additional questions that could be addressed next.

Our understanding of the transcriptome and its previously underestimated significance in gene expression is currently expanding rapidly. The continued research on Gbp2, Hrb1, as well as other yeast RNA-binding proteins will deepen our knowledge about this class of proteins and the processes that they are involved in. Hopefully, this will provide valuable insights into related systems in human cells and eventually culminate in the development of improved therapies for disease curing or other aspects of medicine and biotechnology.

# 5. Materials and Methods

## 5.1. Materials

### 5.1.1. Equipment and Software

**Table 1. Equipment used in this study**

| Equipment | Manufacturer/Source |
| --- | --- |
| Milli-Q® Water purification system | Millipore |
| Improved Neubauer counting chamber | Carl Roth |
| Biophotometer | Eppendorf |
| Quatographic IntelliScan 1600 scanner | Quato Technology |
| Heraeus™ Pico™ 21 Microcentrifuge | Thermo Fisher Scientific |
| Heraeus™ Fresco™ 21 Microcentrifuge | Thermo Fisher Scientific |
| Heraeus™ Multifuge™ X3R with TX-750 or F15-8x50cy rotor | Thermo Fisher Scientific |
| MIKRO 20 Centrifuge | Hettich |
| Primo Star light microscope | ZEISS |
| Eclipse E400 tetrad microscope | Nikon |
| Fluorescence microscope DMI6000B with Leica DFC360 FX camera | Leica Microsystems |
| LABOPORT® N 86 KT.18 mini diaphragm vacuum pump | KNF Neuberger |
| T100™ Thermal Cylcer | Bio-Rad Laboratories |
| MyCycler™ Thermal Cycler | Bio-Rad Laboratories |
| CFX Connect Real-Time PCR Detection System | Bio-Rad Laboratories |
| NanoDrop™ Spectrophotometer | Thermo Fisher Scientific |
| Thermomixer comfort | Eppendorf |
| FastPrep-24™ Homogenizer | MP Biomedicals |
| PerfectBlue™ Semi-Dry Electro Blotter, Sedec™ M | Peqlab |
| *GEL* iX20 Imager | Intas Science Imaging |
| Fusion SL 3500.WL | Peqlab |
| Bio-Link 254 UV-crosslinking chamber | Vilber Lourmat |

Glassware and pipette tips used were autoclaved at 121 °C for 20 minutes or sterilized at 180 °C for 5 hours.

**Table 2. Software used in this study**

| Software | Source |
|---|---|
| Silver Fast Quato XFU | Quato Technology |
| Leica AF 2.7.3.9723 | Leica Microsystems |
| ImageJ | NIH |
| Photoshop | Adobe |
| Illustrator | Adobe |
| SnapGene Viewer | SnapGene |
| ApE | M. Wayne Davis |
| Intas-Capture-Software | Intas Science Imaging |
| NanoDrop2000 Operating Software | Thermo Fisher Scientific |
| Fusion Software | Peqlab |
| Bio1D | Vilber Lourmat |
| CFX manager 3.1 | Bio-Rad Laboratories |
| Microsoft® Excel | Microsoft Corporation |
| Microsoft® Word | Microsoft Corporation |

### 5.1.2. Chemicals and consumable materials

**Table 3. Particular chemicals and materials used in this study**

| Chemicals/Materials | Manufacturer/Source |
|---|---|
| 5-Fluoroorotic acid (FOA) | Apollo Scientific |
| Filtropur BT25 Bottle top filter, 250 ml, PES, 0.2 µm | Sarstedt |
| Cycloheximide | Carl Roth |
| Salmon Sperm DNA | Sigma-Aldrich |
| Formaldehyde 37% | Carl Roth |
| Diagnostic microscope slides, 12 well, 5.2 mm, PTFE-coating | Thermo Fisher Scientific |
| Poly-L-lysine hydrobromide | Sigma-Aldrich |
| DAPI | Carl Roth |
| dNTPs | Thermo Fisher Scientific |
| Agarose NEEO Ultra-Quality | Carl Roth |
| HDGreen™ Plus Safe DNA Dye | Intas Science Imaging |
| ROTI® Phenol/Chloroform/Isoamyl alcohol | Carl Roth |
| cOmplete™, EDTA-free Protease Inhibitor Cocktail | Roche |
| Glass beads type S 0.4 – 0.6 mm | Carl Roth |
| GFP-Trap®_A beads | Chromotek |
| GFP Selector beads | NanoTag Biotechnologies |

| Chemicals/Materials (continued) | Manufacturer/Source |
|---|---|
| MYC-Trap®_A beads | Chromotek |
| IgG Sepharose 6 FastFlow | GE Healthcare |
| Albumin Fraction V | Carl Roth |
| ROTIPHORESE® Gel 30 (37.5:1) acrylamide | Carl Roth |
| Ponceau S | Applichem |
| Difco™ Skim Milk | Becton, Dickinson and Company |
| Amersham™ Protran® 0.45 µm NC Nitrocellulose Blotting Membrane | GE Healthcare |
| Whatman® Blotting Paper | Hahnemühle |
| WesternBright™ Quantum™ Western Blotting HRP Substrate | Advansta |
| DEPC (diethyl pyrocarbonate) | Carl Roth |
| RiboLock RNAse inhibitor | Thermo Fisher Scientific |
| TRIzol™ Reagent | Thermo Fisher Scientific |
| GlycoBlue™ Coprecipitant | Thermo Fisher Scientific |
| 2x qPCRBIO SyGreen Mix Lo-ROX | PCR Biosystems |
| Oligo (dT)$_{18}$ Primer | Sigma-Aldrich |
| Random Hexamer Primer | Thermo Fisher Scientific |

**Table 4. Size standards used in this study**

| Size Standards | Manufacturer/Source |
|---|---|
| Lambda DNA/EcoRI plus HindIII Marker | Thermo Fisher Scientific |
| GeneRuler 50bp DNA Ladder | Thermo Fisher Scientific |
| Cozy™ Prestained Protein Ladder | highQu |
| PageRuler™ Prestained Protein Ladder | Thermo Fisher Scientific |

**Table 5. Kits used in this study**

| Kits | Manufacturer/Source |
|---|---|
| Nucleospin® Gel and PCR Clean-up | MACHEREY-NAGEL |
| Nucleospin® Plasmid | MACHEREY-NAGEL |
| NucleoBond® Xtra Midi/Maxi | MACHEREY-NAGEL |
| Nucleospin® RNA | MACHEREY-NAGEL |
| TURBO DNA-*free™* Kit | Thermo Fisher Scientific |
| Maxima First Strand cDNA Synthesis Kit | Thermo Fisher Scientific |
| FastGene® Scriptase II cDNA Synthesis Kit | NIPPON Genetics |

### 5.1.3. Enzymes and antibodies

**Table 6. Enzymes used in this study**

| Enzymes | Manufacturer/Source |
| --- | --- |
| Zymolyase 20T | Zymo Research |
| DreamTaq DNA Polymerase | Thermo Fisher Scientific |
| Q5® High-Fidelity DNA polymerase | New England BioLabs |
| KAPAHiFi™ DNA Polymerase | Kapa Biosystems |
| FastAP™ Thermosensitive Alkaline Phosphatase | Thermo Fisher Scientific |
| T4 DNA Ligase | Thermo Fisher Scientific |
| T5 Exonuclease | New England BioLabs |
| Taq DNA Ligase | New England BioLabs |
| RNase A | Qiagen |
| RNase-free DNase I | Qiagen |
| Proteinase K, lyophilized | Carl Roth |

| Restriction enzymes | Buffer for double digestion | Manufacturer/Source |
| --- | --- | --- |
| PacI<br>AscI | CutSmart® Buffer | New England BioLabs |
| BamHI<br>XhoI | Buffer G | Thermo Fisher Scientific |
| PstI<br>SacI | Tango Buffer | Thermo Fisher Scientific |
| PstI<br>BcuI | Tango Buffer | Thermo Fisher Scientific |
| AjiI<br>EheI | BamHI Buffer | Thermo Fisher Scientific |

**Table 7. Antibodies used in this study**

| Primary antibodies | Organism | Dilution (western blot) | Source |
| --- | --- | --- | --- |
| GFP antibody (FL) | rabbit | 1 : 1 000 | Santa Cruz |
| GFP antibody (polyclonal) | rabbit | 1 : 2 000 | Laboratory of Heike Krebber |
| GFP antibody (monoclonal, GF28R) | mouse | 1 : 5 000 | Thermo Fisher Scientific |
| GFP antibody (polyclonal, PABG1) | rabbit | 1 : 4 000 | Chromotek |
| HA antibody (mooclonal, F-7) | mouse | (used in IP) | Santa Cruz |
| HA antibody (Y-11) | rabbit | 1 : 200 | Santa Cruz |
| HA antibody (12CA5) | mouse | 1 : 750 | Santa Cruz |

| Primary antibodies (continued) | Organism | Dilution (western blot) | Source |
|---|---|---|---|
| Gbp2 antibody | rabbit | 1 : 50 000 | Laboratory of Heike Krebber |
| Hrb1 antibody | rabbit | 1 : 20 000 | Laboratory of Heike Krebber |
| Hem15 antibody | rabbit | 1 : 10 000 | Gift from Roland Lill |
| Tdh1 antibody (monoclonal, GA1R) | mouse | 1 : 5 000 | Thermo Fisher Scientific |
| Grx4 antibody | rabbit | 1 : 1 000 | Gift from Ulrich Mühlenhoff |
| Nop1 antibody | mouse | 1 : 4 000 | Gift from Ulrich Mühlenhoff |

| Secondary antibodies | Organism | Dilution (western blot) | Source |
|---|---|---|---|
| Anti-Rabbit IgG (H+L) HRPO | goat | 1 : 25 000 | Dianova |
| Anti-Mouse IgG (H+L) HRPO | goat | 1 : 25 000 | Dianova |

### 5.1.4. *Saccharomyces cerevisiae* strains

**Table 8. Yeast strains used in this study**

| Strain number | Genotype | Full genotype | Source |
|---|---|---|---|
| HKY171 | *pab1-53* | *MATα ade2-1 his3-11 leu2-3,112 trp1-1 ura3-1 can1-100 **pab1::HIS3** + p**pab1-53** CEN TRP1* | (Morrissey et al., 1999) |
| HKY298 | *hrb1Δ* | *MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 **hrb1::KanMX4*** | Euroscarf |
| HKY314 | wild type | *MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0* | Euroscarf |
| HKY449 | *pab1-101* | *MATα ade2-1 his3-11 leu2-3,112 trp1-1 ura3-1 can1-100 **pab1::HIS3** + p**pab1-101** CEN TRP1* | Laboratory of Heike Krebber |
| HKY450 | *pab1-16* | *MATα ade2-1 his3-11 leu2-3,112 trp1-1 ura3-1 can1-100 **pab1::HIS3** + p**pab1-16** CEN TRP1* | Laboratory of Heike Krebber |
| HKY492 | *upf1Δ* | *MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 **upf1::KanMX4*** | Euroscarf |
| HKY493 | *upf2Δ* | *MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 **upf2::KanMX4*** | Euroscarf |
| HKY502 | *GBP2-GFP* | *MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 **GBP2-GFP:HIS3MX6*** | Invitrogen |

| Strain number | Genotype | Full genotype | Source |
|---|---|---|---|
| HKY706 | *gbp2Δ hrb1Δ* | *MATa his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0* **hrb1::KanMX4 gbp2::KanMX4 +** p*NPL3 CEN URA3* | (Hackmann et al., 2014) |
| HKY1164 | *UPF1-GFP* | *MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0* **UPF1-GFP:HIS3MX6** | Invitrogen |
| HKY1165 | *UPF2-GFP* | *MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0* **UPF2-GFP:HIS3MX6** | Invitrogen |
| HKY1166 | *UPF3-GFP* | *MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0* **UPF3-GFP:HIS3MX6** | Invitrogen |
| HKY1699 | *upf1Δ gbp2Δ hrb1Δ* | *MATa his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0* **upf1::KanMX4 gbp2::KanMX4 hrb1::KanMX4** | (Grosse et al., 2021) |
| HKY1734 | *upf1Δ gbp2Δ* | *MATa his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0* **upf1::KanMX4 gbp2::KanMX4** | Laboratory of Heike Krebber |
| HKY1735 | *upf1Δ hrb1Δ* | *MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0* **upf1::KanMX4 hrb1::KanMX4** | (Grosse et al., 2021) |
| HKY1787 | *eIF4G-GFP* | *MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0* **TIF4631-GFP:HIS3MX6** | Invitrogen |
| HKY1794 | *GBP2-GFP upf1Δ* | *MATa his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0* **GBP2-GFP:HIS3MX6 upf1::KanMX4** | (Grosse et al., 2021) |
| HKY1834 | *xrn1Δ upf1Δ* | *MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0* **xrn1::KanMX4 upf1::KanMX4** | (Grosse et al., 2021) |
| HKY1881 | *SCD6-GFP* | *MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0* **SCD6-GFP:HIS3MX6** | Invitrogen |
| HKY1886 | *SBP1-GFP* | *MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0* **SBP1-GFP:HIS3MX6** | Invitrogen |
| HKY1887 | *xrn1Δ gbp2Δ upf1Δ* | *MATa his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0* **xrn1::KanMX4 gbp2::KanMX4 upf1::KanMX4** | (Grosse et al., 2021) |
| HKY1957 | *eIF4E-GFP* | *MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0* **CDC33-GFP:HIS3MX6** | Invitrogen |
| HKY1980 | *SKI2-GFP upf1Δ* | *MATα his3Δ1 leu2Δ0 ura3Δ0* **SKI2-GFP:HIS3MX6 upf1::KanMX4** | (Grosse et al., 2021) |
| HKY1981 | *SKI2-GFP upf1Δ gbp2Δ hrb1Δ* | *MATa his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0* **SKI2-GFP:HIS3MX6 upf1::KanMX4 gbp2::KanMX4 hrb1::KanMX4** | (Grosse et al., 2021) |

| Strain number | Genotype | Full genotype | Source |
|---|---|---|---|
| HKY2046 | eIF4G-GFP gbp2Δ hrb1Δ | MATα his3Δ1 leu2Δ0 ura3Δ0 **TIF4631-GFP:HIS3MX6 gbp2::KanMX4 hrb1::KanMX4** | (Grosse et al., 2021) |
| HKY2071 | SKI2-GFP upf1Δ gbp2Δ | MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 **SKI2-GFP:HIS3MX6 upf1::KanMX4 gbp2::KanMX4** | (Grosse et al., 2021) |
| HKY2072 | SKI2-GFP upf1Δ hrb1Δ | MATa his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 **SKI2-GFP:HIS3MX6 upf1::KanMX4 hrb1::KanMX4** | (Grosse et al., 2021) |
| HKY2140 | xrn1Δ hrb1Δ upf1Δ | MAT? his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 **xrn1::KanMX4 hrb1::KanMX4 upf1::KanMX4** | This study |
| HKY2147 | nmd4Δ ebs1Δ gbp2Δ hrb1Δ | MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 **nmd4::KanMX4 ebs1::KanMX4 gbp2::KanMX4 hrb1::KanMX4** | Laboratory of Heike Krebber |

### 5.1.5. Plasmids

**Table 9. Plasmids used in this study**

| Plasmid number | Features | Source |
|---|---|---|
| pHK87 | CEN LEU2 AMP[R] | (Sikorski and Hieter, 1989) |
| pHK88 | CEN URA3 AMP[R] | (Sikorski and Hieter, 1989) |
| pHK101 | 2μ HIS3 AMP[R] | (Sikorski and Hieter, 1989) |
| pHK235 | $P_{GAL1}$**HRB1(Δ1-417)-GFP** 2μ URA3 AMP[R] | (Shen et al., 1998) |
| pHK367 | **GBP2-GFP** CEN URA3 AMP[R] | (Windgassen and Krebber, 2003) |
| pHK537 | **HRB1-GFP** CEN URA3 AMP[R] | (Häcker and Krebber, 2004) |
| pHK811 | **PAB1-GFP** CEN URA3 AMP[R] | Laboratory of Heike Krebber |
| pHK1321 | **N-GFPsplit-NPL3** CEN URA3 AMP[R] | (Baierlein et al., 2013) |
| pHK1322 | **C-GFPsplit-NPL3** CEN LEU2 AMP[R] | (Baierlein et al., 2013) |
| pHK1574 | **CBP80-MYC** CEN URA3 AMP[R] | (Becker et al., 2019) |
| pHK1578 | **CBP80[PTC]-MYC** CEN URA3 AMP[R] | (Grosse et al., 2021) |
| pHK1592 | **UPF1-HA** CEN LEU2 HIS3 AMP[R] | (Serdar et al., 2016) |

| Plasmid number | Features | Source |
|---|---|---|
| pHK1593 | *upf1-DE572AA-HA* CEN LEU2 HIS3 AMP<sup>R</sup> | (Serdar et al., 2016) |
| pHK1600 | *P<sub>GAL1</sub>MYC-CBP80<sup>PTC</sup>* CEN URA3 AMP<sup>R</sup> | (Grosse et al., 2021) |
| pHK1625 | *UPF1-GFP* CEN LEU2 HIS3 AMP<sup>R</sup> | This study |
| pHK1626 | *upf1-DE572AA-GFP* CEN LEU2 HIS3 AMP<sup>R</sup> | This study |
| pHK1628 | *UPF1-N-GFPsplit* CEN LEU2 HIS3 AMP<sup>R</sup> | This study |
| pHK1629 | *GBP2-C-GFPsplit* CEN URA3 AMP<sup>R</sup> | This study |
| pHK1630 | *HRB1-C-GFPsplit* CEN URA3 AMP<sup>R</sup> | This study |
| pHK1633 | *upf1-DE572AA-N-GFPsplit* CEN LEU2 HIS3 AMP<sup>R</sup> | This study |
| pHK1636 | *P<sub>GAL1</sub>(His)<sub>6</sub>-NPL3* CEN LEU2 AMP<sup>R</sup> | Laboratory of Heike Krebber |
| pHK1642 | *P<sub>GAL1</sub>MYC-CBP80<sup>PTC</sup>* CEN HIS3 AMP<sup>R</sup> | This study |
| pHK1644 | *P<sub>GAL1</sub>N-GFPsplit-UPF1* CEN LEU2 AMP<sup>R</sup> | This study |
| pHK1649 | *UPF1-HA* CEN LEU2 AMP<sup>R</sup> | Laboratory of Heike Krebber |
| pHK1650 | *upf1-DE572AA-HA* CEN LEU2 AMP<sup>R</sup> | Laboratory of Heike Krebber |
| pHK1667 | *CDC33-GFP* CEN URA3 AMP<sup>R</sup> | Laboratory of Heike Krebber |
| pHK1686 | *upf1-DE572AA-N-GFPsplit* CEN LEU2 AMP<sup>R</sup> | This study |

| Plasmid number | Features | Source |
|---|---|---|
| pHK1593 | *upf1-DE572AA-HA* CEN LEU2 HIS3 AMP$^R$ | (Serdar et al., 2016) |
| pHK1600 | *P$_{GAL1}$MYC-CBP80$^{PTC}$* CEN URA3 AMP$^R$ | (Grosse et al., 2021) |
| pHK1625 | *UPF1-GFP* CEN LEU2 HIS3 AMP$^R$ | This study |
| pHK1626 | *upf1-DE572AA-GFP* CEN LEU2 HIS3 AMP$^R$ | This study |
| pHK1628 | *UPF1-N-GFPsplit* CEN LEU2 HIS3 AMP$^R$ | This study |
| pHK1629 | *GBP2-C-GFPsplit* CEN URA3 AMP$^R$ | This study |
| pHK1630 | *HRB1-C-GFPsplit* CEN URA3 AMP$^R$ | This study |
| pHK1633 | *upf1-DE572AA-N-GFPsplit* CEN LEU2 HIS3 AMP$^R$ | This study |
| pHK1636 | *P$_{GAL1}$(His)$_6$-NPL3* CEN LEU2 AMP$^R$ | Laboratory of Heike Krebber |
| pHK1642 | *P$_{GAL1}$MYC-CBP80$^{PTC}$* CEN HIS3 AMP$^R$ | This study |
| pHK1644 | *P$_{GAL1}$N-GFPsplit-UPF1* CEN LEU2 AMP$^R$ | This study |
| pHK1649 | *UPF1-HA* CEN LEU2 AMP$^R$ | Laboratory of Heike Krebber |
| pHK1650 | *upf1-DE572AA-HA* CEN LEU2 AMP$^R$ | Laboratory of Heike Krebber |
| pHK1667 | *CDC33-GFP* CEN URA3 AMP$^R$ | Laboratory of Heike Krebber |
| pHK1686 | *upf1-DE572AA-N-GFPsplit* CEN LEU2 AMP$^R$ | This study |

## 5.1.6. Oligonucleotides

**Table 10. Oligonucleotides used in this study for cloning**

| Oligo number | Sequence | Target | Direction |
|---|---|---|---|
| HK95 | 5'- AGGCGTTTAGCGTAC -3' | *HRB1* | reverse |
| HK164 | 5'- ATCGAAGTAGATCGAGG -3' | *HRB1 promoter* | forward |
| HK2828 | 5'- TGGCCATGAAGTTCCAGGTG -3' | *XRN1* | forward |
| HK2829 | 5'- TAGGAGGTGGAGGTGGGAAG -3' | *XRN1* | reverse |
| HK2935 | 5'- AGCAATTTGGGAATACGGATCCCCGGGTTAATTAAAGCTATGGCTAGCAAAGG -3' | *GFP* | forward |
| HK2936 | 5'- TAAGAAATTCGCTTATTTAGAAGTGGCGCGCCCTTTGTTAGCAGCCGGATC -3' | *GFP* | reverse |
| HK2987 | 5'- CGGGTTAATTAAgcctATGGTGAGCAAGGGC -3' | *N-GFPsplit* | forward |
| HK2988 | 5'- AAGTGGCGCGCCTTAGCCTCCTGGGGCCATG -3' | *N-GFPsplit* | reverse |

| Oligo number | Sequence | Target | Direction |
|---|---|---|---|
| HK2989 | 5'-CGCCCTCGAGGGCCTATGGACAAGCAGAAG -3' | *C-GFPsplit* | forward |
| HK2990 | 5'- GGTGGATCCTGGCTTGTACAGCTCGTC -3' | *C-GFPsplit* | reverse |
| HK3193 | 5'-CCAGTATTCTTAACCCAACTGCACAGAACAA AAACCTGCAGCCTCCTCTAGTACACTC -3' | *HIS3* | forward |
| HK3194 | 5'-TGAAAGTTCCATCTAGAGCGGCCGCCACCG CGGTGGAGCTCGGGTGATGGTTCACGTAG -3' | *HIS3* | reverse |
| HK4427 | 5'-CCTCTATACTTTAACGTCAAGGAGAAAAAAC TATAT CTAGAAATGGTGAGCAAGGGCGAGG -3' | *N-GFPsplit* | forward |
| HK4428 | 5'-GAGAACCGGAACCGACCATAGGCCTACCCC CGCCTCC -3' | *N-GFPsplit* | reverse |
| HK4429 | 5'-CCAGGAGGCGGGGGTAGGCCTATGGTCGG TTCCGGTTCTC -3' | *UPF1* | forward |
| HK4430 | 5'-GAGGTCGACGGTATCGATAAGCTTGATATC GAATTCCTGCATTATATTCCCAAATTGCTGAA G -3' | *UPF1* | reverse |

**Table 11. Oligonucleotides used in this study for qPCR**

| Oligo number | Sequence | Target | Direction |
|---|---|---|---|
| HK634 | 5'- CGAAAGCTGAATGAAGG -3' | *GCR1* | forward |
| HK635 | 5'- CGATTCCTTTAAACTCTTCG -3' | *GCR1* | reverse |
| HK843 | 5'- CGTCATCTTCCTTGGACAAACC -3' | *RPS6A* | reverse |
| HK877 | 5'- AAGGGTGAGCAAGAATTGGAAGG -3' | *RPS6A* | forward |
| HK1002 | 5'- TGCTAAGGCTGTCGGTAAGG -3' | *TDH1* | forward |
| HK1003 | 5'- TCAGAGGAGACAACGGCATC -3' | *TDH1* | reverse |
| HK2696 | 5'- CGCTATTCCACACGAATCCA -3' | *CBP80* | forward |
| HK2697 | 5'- ACTCTGGATGATCCGTTCAAGTC -3' | *MYC* | reverse |
| HK2782 | 5'- ACTTGAACGGATCATCCAGAGTCG -3' | *MYC-CBP80* | forward |
| HK2783 | 5'- AGGTGGGATTCTCTGTCATTTAGG -3' | *CBP80$^{PTC}$* | reverse |
| HK3089 | 5'- AGTTACGCTAGGGATAACAGGG -3' | *21S* rRNA | forward |
| HK3090 | 5'- TGACGAACAGTCAAACCCTTC -3' | *21S* rRNA | reverse |
| HK3152 | 5'- AAGCTTCCTTTCGGGCTTTG -3' | *NUF2* terminator | reverse |
| HK3153 | 5'- CGCTATTCCACACGAATCCAC -3' | *CBP80* | forward |

| Oligo number | Sequence | Target | Direction |
|---|---|---|---|
| HK3154 | 5'- GCGGAGTGATAACGAATGTAGTC -3' | *CBP80* 3' UTR | reverse |
| HK4124 | 5'- GTGGTTTGATTGTAGGGTGG -3' | *HNT1* | forward |
| HK4125 | 5'- TCGGAGCCTTCTAGTTTGG -3' | *HNT1* | reverse |

## 5.2. Cell cultivation

Cell culture media and double-distilled water (Milli-Q® Water purification system) were autoclaved (121 °C, 20 min) before usage. Heat-sensitive components, such as antibiotics, sucrose, and galactose, were sterile-filtrated (Filtropur BT25) and added to the respective autoclaved media after they had cooled to at least approximately 60 °C.

### 5.2.1. Cultivation of *Escherichia coli* cells

*E. coli* cells were cultivated according to standard protocols (Green and Sambrook, 2012; Elbing and Brent, 2019) using the LB medium. For selection of plasmids carrying a gene resistant to the antibiotic ampicillin, the media was supplemented with ampicillin to a final concentration of 100 µg/ml before usage.

| **LB medium (low salt) (pH 7.5)** | |
|---|---|
| Tryptone | 1% (w/v) |
| Yeast extract | 0.5 % (w/v) |
| NaCl | 0.5% (w/v) |
| Agar-Agar (for plates only) | 1.5% (w/v) |

(Bertani, 1951; Lennox, 1955; Green and Sambrook, 2012)

For growth in solid media, bacterial cell suspension was spread on a LB agar plate and incubated at 37 °C. For liquid cultures, liquid LB medium was inoculated with single bacterial colonies from LB plates and incubated at 37 °C overnight with agitation.

### 5.2.2. Cultivation of *Saccharomyces cerevisiae* cells

Yeast cells were cultivated according to standard protocols (Sherman, 2002) in YPD (yeast extract peptone dextrose) or selective media when not otherwise stated. Selective media were chosen according to the selection marker gene(s) of the plasmid(s) that the cells carried.

**YPD medium**

| Yeast extract | 1% (w/v) |
| --- | --- |
| Peptone | 2% (w/v) |
| Glucose | 2% (w/v) |
| Agar-Agar (for plates only) | 1.8% (w/v) |

**Selective medium**

| Nitrogen base | 1.7 g/l |
| --- | --- |
| Ammonium sulphate | 5.1 g/l |
| L-Alanine | 80 mg/l |
| L-Arginine | 80 mg/l |
| L-Asparagine | 80 mg/l |
| L-Aspartic acid | 80 mg/l |
| L-Cysteine | 80 mg/l |
| L-Glutamine | 80 mg/l |
| L-Glutamic acid | 80 mg/l |
| L-Glycine | 80 mg/l |
| Inositol | 80 mg/l |
| L-Isoleucine | 80 mg/l |
| L-Methionine | 80 mg/l |
| Para-aminobenzoic acid | 8 mg/l |
| L-Phenylalanine | 80 mg/l |
| L-Proline | 80 mg/l |
| L-Serine | 80 mg/l |
| L-Threonine | 80 mg/l |
| L-Tyrosine | 80 mg/l |
| L-Valine | 80 mg/l |
| Glucose* | 2% (w/v) |
| Agar-Agar* (for plates only) | 1.8% (w/v) |
| **optional components** | |
| L-Adenine | 20 mg/l |
| L-Histidine | 80 mg/l |
| L-Leucine | 400 mg/l |
| L-Lysine | 80 mg/l |
| L-Tryptophan | 80 mg/l |
| Uracil | 80 mg/l |

(Sherman, 2002)

*Components were autoclaved separately.

All yeast strains were suspended in 50% glycerol and kept at −80 °C for long-term storage. The required strains were streaked onto agar plates from the storage stocks and incubated at 25 °C for 2 – 3 days before they were used in experiments.

Yeast cells were generally grown at 25 °C unless otherwise indicated. For growth in solid media, yeast cells were streaked on the appropriate agar plates and incubated at 25 °C for 2 – 3 days before storage at 4 °C. Cells were re-streaked onto new agar plates regularly. For liquid cultures, cell material was taken from agar plates and resuspended in corresponding liquid media, followed by incubation at 25 °C with agitation.

For cells transformed (see **5.4**) with a plasmid that encodes a gene under the control of the *GAL1* promoter, gene expression was induced with galactose. Small volumes of glucose-containing media were inoculated with cells and these pre-cultures were incubated overnight until cells reached the stationary phase. New cultures with sucrose-containing media were inoculated with the pre-cultures to cell densities of $OD_{600} \sim 0.05$ (see *5.2.3.2*) and incubated overnight. When the cells achieved logarithmic phase (log phase, $1 – 3 \times 10^7$ cells/ml or $OD_{600}$ 0.6 – 1.2), galactose was added to a final concentration of 2% (w/v) and the cultures were incubated further at 25 °C for 2 hours with agitation to induce gene expression.

FOA (5-Fluoroorotic acid) plates were used to select for the absence of the *URA3* gene (Boeke et al., 1987). The FOA plates were prepared as the selective plates (see above), except containing only 50 mg uracil and in addition 1 g of FOA. All components except agar was sterile filtrated and combined with the autoclaved agar solution after it has cooled down to about 60 °C. In this work, the strain HKY706 was incubated on FOA plates before usage in experiments to select for cells that did not contain the *NPL3* plasmid.


## 5.2.3. Determination of cell density in liquid cultures

Cell densities were determined for control of cell growth phase, preparation for growth analyses (see **5.3**), or estimation of the resulting cell pellet size after centrifugation.


*5.2.3.1. Cell counting*

Cell suspensions were diluted in distilled water and 10 µl was pipetted into an improved Neubauer counting chamber. Cells were directly counted under a light microscope and the approximate cell densities were calculated according to the volume of the counting chamber.

*5.2.3.2. Measurement of optical density*

Cell suspensions were pipetted into plastic cuvettes and their optical densities at 600 nm wavelength ($OD_{600}$) were measured with a biophotometer. $OD_{600}$ values between 0.1 and 1.0 correlate linearly with cell densities. Cultures with higher $OD_{600}$ values were diluted in respective media and re-measured. Pure media served as blanks and were used to calibrate the $OD_{600}$ value to zero for each measurement.

## 5.3 Yeast cell growth analysis

Yeast cells were taken from agar plates and suspended in sterile distilled water. Cell densities were determined by cell counting (see *5.2.3.1*) and the cell suspensions were serially diluted to densities of $10^7$, $10^6$, $10^5$, $10^4$, and $10^3$ cells/ml. Subsequently, 10 μl cell suspension of each density was pipetted onto the appropriate agar plates in linear order, with suitable distances, and incubated at the indicated temperatures. Plates were monitored and scanned on the following days with a Quatographic IntelliScan 1600 scanner and the Silver Fast Quato XFU software.

## 5.4. Yeast transformation

Transformation of yeast cells with plasmids was carried out according to a lithium acetate-based protocol (Gietz et al., 1992). Yeast cells were grown to log phase and collected from 5 ml cultures for each transformation via centrifugation at 2 000x g for 5 minutes. The cell pellets were resuspended in 1 ml sterile distilled water, transferred to 1.5 ml tubes, and collected via centrifugation at 12 000x g for 1 minute. The cell pellets were washed (resuspension in buffer and centrifugation at 12 000x g for 1 minute) once in 1 ml sterile TE/Lithium acetate (LiOAc) buffer and resuspended in 50 μl sterile TE/LiOAc. The cell suspensions were each combined with 1 μg plasmid DNA, 50 μg single-stranded salmon sperm carrier DNA, which was heated at 95 °C for 5 minutes and cooled down on ice for 2 minutes before use, as well as 300 μl sterile PEG/TE/LiOAc. Two plasmids (1 μg each) were added for transformation with two plasmids at once. The components were mixed thoroughly and the samples were incubated with rotation at 25 °C for 30 minutes, followed by incubation at 42 °C for 15 minutes. Cells were collected by centrifugation at 12 000x g for 1 minute, washed once with 1 ml sterile distilled water, and resuspended in

100 µl sterile distilled water to be plated on the corresponding selective agar plates. The cells were incubated at 25 °C for 2 – 3 days and single colonies were selected and streaked onto new selective plates for further cultivation and later use.

**TE/LiOAc (pH7.5)**

| | |
|---|---|
| Tris | 10 mM |
| EDTA (Ethylenediaminetetraacetic acid) | 1 mM |
| Lithium acetate | 100 mM |

**PEG/TE/LiOAc (pH7.5)**

| | |
|---|---|
| PEG (Polyethylene glycol) 4 000 | 40% (v/v) |
| Tris | 10 mM |
| EDTA | 1 mM |
| Lithium acetate | 100 mM |

## 5.5. Generation of yeast strains

Yeast strains with desired genotypes from two different strains were generated by mating of yeast (Sherman and Hicks, 1991; Sherman, 2002).

### 5.5.1. Mating and sporulation

Two haploid strains with opposite mating types were mixed on a YPD agar plate and incubated at 25 °C for 2 days. If the two strains carried different selection marker genes, the mixed cell material was re-streaked onto the corresponding selective plate and incubated at 25 °C for 2 – 3 days to select for diploids. Some mixed cell material was subsequently added to 2 ml of Super-SPO medium and incubated at 25 °C with agitation for 3 – 7 days to induce sporulation.

**Super-SPO medium**

**Solution 1**

| | |
|---|---|
| Yeast extract | 0.5% (w/v) |
| Potassium acetate | 306 mM |

(continued on next page)

**Super-SPO medium (continued)**

| Solution 2 | |
| --- | --- |
| Glucose | 0.1% (w/v) |
| Arginine | 20 mg/l |
| Histidine | 20 mg/l |
| Leucine | 20 mg/l |
| Lysine | 20 mg/l |
| Methionine | 20 mg/l |
| Tryptophan | 20 mg/l |
| Adenine | 40 mg/l |
| Tyrosine | 40 mg/l |
| Uracil | 40 mg/l |
| Phenylalanine | 100 mg/l |
| Threonine | 350 mg/l |

(Rose et al., 1990)

Solution 1 was autoclaved, Solution 2 was sterile filtrated, and the two solutions were mixed in a 1:1 ratio.

### 5.5.2. Tetrad dissection

Samples of the Super-SPO cultures were taken and observed under the light microscope to check for the appearance of tetrads. When sufficient amounts of tetrads were seen, cells were collected from 100 µl of the Super-SPO culture by centrifugation at 12 000x g for 1 minute. Pelleted cells were washed (resuspension and centrifugation at 12 000x g for 1 minute) once in 100 µl sterile distilled water and resuspended in 50 µl P-solution. The asci walls were digested by adding 50 µg of zymolyase to the cell suspension and incubation at room temperature for 6 – 7 minutes. The cell suspension was centrifuged at 12 000x g for 1 minute and the supernatant was removed. The cell pellet, kept on ice, was washed once with 100 µl P-solution and resuspended in 200 µl P-solution. From this, 2.5 µl was mixed with 100 µl sterile distilled water and altogether transferred onto a YPD agar plate, evenly spread on one side of the plate over about one-third of the area, and let dried. Using a tetrad microscope, tetrads were selected and the four spores were separately placed on the remaining two-thirds of the YPD plate area. The spores were incubated at 25 °C for 2 – 3 days and re-streaked onto a new YPD plate to increase cell material. Cell material of each spore was transferred into 200 µl 50% glycerol in 96-well plates and used for following tetrad analysis or stored for long-term at −80 °C.

**P-solution (pH 6.5)**

| | |
|---|---|
| Potassium phosphate buffer (pH 6.5) | 0.1 M |
| Sorbitol | 1.2 M |

**0.1 M Potassium phosphate buffer (pH 6.5)**

| | |
|---|---|
| $K_2HPO_4$ | 33.7 mM |
| $KH_2PO_4$ | 66.3 mM |

### 5.5.3. Tetrad analysis

The tetrad spores were analyzed to identify the desired genotype. Cells were stamped with a microplate replicator from the 96-well plate onto YPD plates and incubated at different temperatures to identify temperature sensitive genotypes or onto different selective plates to identify the selection marker gene(s) that they carried. For selection of the *KanMX4* cassette, cells were stamped onto YPD plates that were previously plated with 100 µl of geneticin (80 µg/µl).

For determination of mating type, cells were stamped onto YPD plates that were covered with *MATa* or *MATα* reference strains. The stamped reference plates were replica-plated onto B plates on the following day and incubated at 25 °C for up to 3 days. Spores were only able to mate with the reference strain that had an opposite mating type. Since the reference strains carry different auxotrophic markers from the spores that were tested, only diploid cells were able to grow on the selective B plates.

**B plates**

| | |
|---|---|
| Nitrogen base | 1.7 g/l |
| Ammonium sulphate | 5.1 g/l |
| Agar-Agar* | 3% (w/v) |
| Glucose* | 2% (w/v) |

*Components were autoclaved separately

To distinguish genotypes that used the same selection marker, for example multiple gene deletions that used the *KanMX4* marker, yeast colony PCR was carried out to directly analyze the specific genes of interest. Cells from the spores were suspended in 50 µl sterile PBS and the cell walls were digested by incubation with 200 µg zymolyase at 37 °C for 1 hour then at 95 °C for 10 minutes. Cell debris was collected via centrifugation at 12 000x g for 1 minute and 2 µl of the supernatant was used as the DNA template in PCR reactions carried out with the DreamTaq DNA Polymerase (see **5.7.1**).

**PBS (pH 7.4)**

| | |
|---|---|
| NaCl | 137 mM |
| KCl | 2.7 mM |
| $KH_2PO_4$ | 1.8 mM |
| $Na_2HPO_4$ | 10 mM |

### 5.5.4. Generation of strain HKY2140

For generation of strain HKY2140 (genotype *xrn1Δ hrb1Δ upf1Δ*), a spore from the mating of HKY492 and HKY706 that has the genotype *MATa hrb1Δ upf1Δ* was mated with strain HKY1834 (genotype *MATα xrn1Δ upf1Δ*). Deletion of *XRN1* and *HRB1* in the obtained spores was verified through yeast colony PCR. For *XRN1*, the primers HK2828 and HK2829 (**Table 10**) were used at an annealing temperature of 55 °C, both of which anneal to the coding sequence of *XRN1* and did not result in a PCR fragment from the *xrn1::KanMX4* allele. For *HRB1*, the primers HK164 and HK95 (**Table 10**) were used at an annealing temperature of 43 °C, both of which anneal to the coding sequence of *HRB1* and did not result in a PCR fragment from the *hrb1::KanMX4* allele.

## 5.6. Fluorescence microscopy

### 5.6.1. Split-GFP analysis

The *eGFP* sequence was divided into N-terminal (amino acids 1 – 155) and C-terminal (amino acids 156 – 239) fractions (Baierlein et al., 2013) and fused to the *UPF1* or *upf1-DE572AA* and the *GBP2 or HRB1* genes on yeast plasmids, respectively (see **5.7.11**). Yeast strains were transformed (see **5.4**) with the split-GFP plasmid(s) and the *CBP80^{PTC}* reporter (pHK1642) as indicated. Expression of *CBP80^{PTC}* was induced with galactose for 2 hours as described in **5.2.2**.

GFP fluorescence microscopy was done principally as previously described (Hackmann et al., 2014) with minor modifications. For each sample, a 5 ml yeast culture was grown to log phase and formaldehyde was added to a final concentration of 1.5% (v/v) for fixation. Cells were immediately collected by centrifugation at 2 000x g for 5 minutes at 4 °C and for following steps kept on ice. Pellets were resuspended in 1 ml 0.1 M potassium phosphate buffer (see **5.5.2**), transferred to 1.5 ml tubes, and collected again by centrifugation at 12 000x g for 1 minute at 4 °C. This washing step was repeated once

with P-solution (see **5.5.2**). According to the size of the pellet, the cells were resuspended in 100 µl – 1 ml P-solution and kept on ice.

Twelve-well diagnostic microscope slides were used. The wells were first coated with polylysine – 20 µl of 0.3% poly-L-lysine solution was applied to each well and incubated at room temperature for 5 minutes. Polylysine was rinsed off with distilled water and the slide was let air-dried. For all following steps, 20 µl of the respective solution or cell suspension was carefully pipetted onto each well, and after the indicated incubation time, removed carefully with the vacuum pump without completely drying the wells. The cell suspensions in P-solution were pipetted onto the wells and incubated for 15 minutes at room temperature. To decrease bleaching of the fluorophores, the slide was covered to avoid light when possible. Excess cells were removed carefully and remaining cells were permeabilized with 20 µl fresh 0.5% Triton X-100 (in P-solution) for 30 – 60 seconds. Excess solution was removed carefully. The cells were washed once with P-solution and once with Aby wash 2, with 30 – 60 seconds incubation time for each. DAPI solution (4',6-diamidino-2-phenylindole; 1 µg/ml in Aby wash 2) was applied and incubated for 5 minutes at room temperature to stain DNA, followed by three times washing with Aby wash 2, each time incubated for 5 minutes. Slides were air-dried, covered with the mounting medium, and sealed with a cover slide using nail polish.

**Aby wash 2**

| | |
|---|---|
| Tris (pH 9.5) | 0.1 M |
| NaCl | 0.1 M |

**Mounting medium**

| | |
|---|---|
| n-Propyl gallate | 1% (w/v) |
| Glycerol | 40% (v/v) |
| PBS (see **5.5.3**) | 20% (v/v) |

The slide was examined with the Leica DMI6000B microscope with 63x magnification in glycerol and 1.6x ocular magnification, using the L5 filter cube for GFP and the 405 filter cube for DAPI. Images were taken using a Leica DFC360FX camera with the Leica AF 2.7.3.9723 software, and processed with ImageJ and Adobe Photoshop.

### 5.6.2. Quantification of fluorescent signals

GFP signal intensities were quantified using ImageJ. For each strain in each experiment, 100 cells were chosen randomly and the mean signal intensity of the whole area of each

cell was measured by the software. An area of the background without any cells was measured for each image and served as the blank value for cells quantified from that image. Significant differences between strains were analyzed by comparing all 300 values from the 3 biological repeats for each strain.

## 5.7. Recombinant DNA construction

Plasmids encoding specific genes with the desired tags and promoter sequences were constructed with the following strategy that was based on standard methods (Green and Sambrook, 2012). The required DNA insert fragments were amplified with polymerase chain reaction (PCR) using high-fidelity proofreading DNA polymerases and joined to a restriction enzyme-digested vector either through ligation or Gibson Assembly. After transformation of *E. coli* cells with the recombinant DNA, the plasmids were isolated from *E. coli* colonies and their sequences were verified via PCR, restriction digestion, and DNA sequencing. The following sections describe each step in detail. A brief description of the cloning strategy for each constructed plasmid is summarized in **5.7.11**. The SnapGene Viewer and the ApE plasmid editor were used for strategy design and construct analysis.

### 5.7.1. Polymerase chain reaction (PCR)

DNA fragments of the required gene sequences plus 5' and 3' overhangs were produced by PCR with the Q5® High-Fidelity DNA Polymerase or the KAPAHiFi™ DNA Polymerase following the manufacturers' protocols (**Table 12**). For DNA fragments that were later joined to a vector via ligation, the 5' and 3' overhangs contained restriction enzyme cleavage sites corresponding to the respective vectors (see **5.7.4** and **5.7.5**); for DNA fragments that were later joined to a vector by Gibson Assembly, the overhangs contain sequences complementary to the respective vectors (see **5.7.6**). Forward and reverse primers (see **Table 10**) were designed manually or with the help of an online primer design tool, such as Primer-BLAST (NCBI, NIH), and ordered from Sigma-Aldrich in dry form in tubes. The oligonucleotides were dissolved in sterile, nuclease-free water to 100 µM and stored at −80 °C. From these stock solutions, the oligonucleotides were diluted into 10 µM aliquots and stored at −20 °C for regular usage.

PCRs were carried out in T100™ or MyCycler™ Thermo Cyclers with thermocycling parameters suggested by the manufacturers of the respective DNA polymerases (**Table 13**). Each primer pair was usually tested at three different annealing temperatures to find

the ideal reaction condition. For PCRs used to verify genotypes (analytic PCRs), as in yeast or *E. coli* colony PCRs (see **5.5.3** and **5.7.8**, respectively), only the sequences complimentary to the primers and the sizes of the PCR products were important. Since complete error-free sequences of the DNA fragments were not required, the non-proofreading DreamTaq DNA Polymerase was used, according to the manufacturer's suggestions (**Table 12 and 13**).

**Table 12. Components of PCR samples**

| DNA Polymerase: | Q5® | KAPAHiFi™ | DreamTaq |
|---|---|---|---|
| **Primers** | 0.5 µM each | 0.3 µM each | 0.2 µM each |
| **dNTPs** | 200 µM each | 300 µM each | 200 µM each |
| **Polymerase** | 0.02 U/µl | 0.02 U/µl | 0.04 U/µl |
| **Template DNA** | < 1 000 ng | 250 ng | < 1 000 ng |

**Table 13. PCR parameters**

| Q5® | | | |
|---|---|---|---|
| | Temperature | Time | Repeat |
| **Initial denaturation** | 98 °C | 30 sec | 1x |
| **Denaturation** | 98 °C | 10 sec | |
| **Annealing** | 50 – 60 °C | 30 sec | 35x |
| **Extension** | 72 °C | 30 sec/kb | |
| **Final Extension** | 72 °C | 10 min | 1x |

| KAPAHiFi™ | | | |
|---|---|---|---|
| | Temperature | Time | Repeat |
| **Initial denaturation** | 95 °C | 5 min | 1x |
| **Denaturation** | 98 °C | 20 sec | |
| **Annealing** | 50 – 60 °C | 15 sec | 35x |
| **Extension** | 72 °C | 30 sec/kb | |
| **Final Extension** | 72 °C | 5 min | 1x |

| DreamTaq | | | |
|---|---|---|---|
| | Temperature | Time | Repeat |
| **Initial denaturation** | 95 °C | 3 min | 1x |
| **Denaturation** | 95 °C | 30 sec | |
| **Annealing** | 50 – 60 °C | 30 sec | 25x |
| **Extension** | 72 °C | 1 min/kb | |
| **Final Extension** | 72 °C | 10 min | 1x |

### 5.7.2. Agarose gel electrophoresis

For the separation and visualization of DNA, agarose gel electrophoreses were carried out. Agarose gels were prepared by dissolving 1% (w/v) (or 2% for DNA fragments smaller than 500 bp) agarose in TAE buffer through heating. For staining of the DNA, 33 µg/100 ml ethidium bromide or 5 µl/100 ml HDGreen™ Plus Safe DNA Dye was added to the agarose solution after it has cooled down to approximately 60 °C. The agarose solution was poured into self-made casting devices and combs were inserted to create sample wells before gel polymerization. Agarose gels were stored at 4 °C or placed in self-made electrophoresis chambers in TAE buffer for electrophoresis.

**TAE buffer**

| | |
|---|---|
| Tris | 40 mM |
| Acetic acid | 0.114% (v/v) |
| EDTA | 1 mM |

DNA samples were mixed with 1/5x volume of 6x DNA sample buffer and loaded onto the gels. The Lambda DNA/EcoRI plus HindIII Marker was used as a size standard for 1% gels and the GeneRuler 50 bp DNA Ladder was used for 2% gels. Electrophoresis was carried out under constant voltage of 100 – 120 V for 60 minutes or until the samples migrated to the desired distance. DNA was visualized and gel images were captured with the *GEL* iX20 Imager system, which includes a UV transilluminator, an EXview HAD CCD™ camera, and the Intas-Capture-Software, following the manufacturer's instructions.

**6x DNA sample buffer**

| | |
|---|---|
| Tris (pH 7.6) | 10 mM |
| EDTA | 60 mM |
| Glycerol | 60% (v/v) |
| Bromophenol blue | 0.03% (w/v) |
| Xylene cyanol blue | 0.03% (w/v) |

### 5.7.3. DNA purification and extraction from gel

For DNA fragments that were separated on an agarose gel and further required for cloning, a piece of gel containing the DNA fragment was cut out carefully with a scalpel and DNA was extracted and purified from the gel with the Nucleospin® Gel and PCR Clean-up kit following the protocol provided by the manufacturer. DNA was finally eluted in 15 – 30 µl Buffer NE (5 mM Tris/HCl, pH 8.5) and later stored at −20 °C. DNA

concentration was measured with a NanoDrop™ Spectrophotometer and purity of the DNA was assessed by the A280/A260 value.

### 5.7.4. Restriction enzyme digestion and dephosphorylation of DNA ends

Plasmids and PCR products were cleaved using different restriction endonucleases to linearize vectors, create complementary sequences for ligation (see **5.7.11**), or analyze constructed plasmids. Suitable restriction enzymes were chosen and corresponding reaction buffers were used according to the manufacturers' suggestions. For double digestion, online tools provided by the manufacturers – DoubleDigest Calculator from Thermo Fisher Scientific and NEBcloner from New England BioLabs – were used to check the compatibility of two enzymes and to find the ideal reaction buffer (**Table 6**). In principle, 10 µl reaction samples were prepared following the manufacturer's protocols and restriction digestion was performed at 37 °C overnight except when a restriction enzyme known to have star activity was used. In those cases, digestion was done for 2 hours.

Dephosphorylation of DNA ends was performed to prevent re-annealing of the digested vectors during ligation by directly adding 1U of FastAP™ Thermosensitive Alkaline Phosphatase to the reaction after restriction digestion and incubating at 37 °C for 10 minutes. The restriction enzymes were then heat-inactivated when possible using the temperature and time suggested by the manufacturers. Digested DNA was separated on an agarose gel (see **5.7.2**) and, when necessary, purified from the gel as described in **5.7.3**.

### 5.7.5. DNA ligation

DNA insert fragments and linearized vectors that were processed by the same restriction enzymes and therefore had compatible ends were ligated with the T4 DNA Ligase using the provided buffer. For each sample, 100 ng vector was mixed with the insert fragment at a molar ratio of 1:2 – 1:4 (vector to insert) in a total volume of 10 µl. Ligation was performed at 16 °C overnight, followed by inactivation of the ligase by incubation at 65 °C for 10 minutes. The samples were stored at −20 °C or directly used to transform *E. coli* cells (see **5.7.7**).

### 5.7.6. Gibson Assembly

Gibson Assembly was performed according to the published principles (Gibson et al., 2009; Gibson, 2011). Primers were designed so that amplified insert fragments had at

least 40 bp of overlapping sequences with the vector at both the 5' and 3' ends. For each reaction, 100 ng of linearized vector was combined with insert fragments in 1:2 – 1:3 molar ratio, or 1:5 when the insert was smaller than 200 bp, in 10 µl of nuclease-free water. This was then mixed with 10 µl self-made 2x Gibson Assembly Master Mix and incubated at 50 °C for 1 hour. The samples were stored at −20 °C or directly used to transform *E. coli* cells (see **5.7.7**).

**2x Gbison Assembly Master Mix**

| | |
|---|---|
| 5x ISO buffer | 100 µl |
| T5 Exonuclease (1 U/µl) | 2 µl |
| Q5® DNA Polymerase ( 2 U/µl) | 6.3 µl |
| Taq DNA Ligase (40 U/µl) | 25 µl |
| add DEPC water to 375 µl | |

**5x Isothermal (ISO) reaction buffer**

| | |
|---|---|
| Tris (pH 7.5) | 500 mM |
| $MgCl_2$ | 50 mM |
| DTT (dithiothreitol) | 50 mM |
| NAD (Nicotinamide adenine dinucleotide) | 5 mM |
| dNTPs | 1 mM each |
| PEG 8 000 | 25% (v/v) |

### 5.7.7. *E. coli* cell transformation

For transformation of *E. coli* cells with plasmid DNA, the DH5α *Escherichia coli* strain was used. Chemically competent DH5α *E. coli* cells were prepared according to Inoue et al. (1990) and stored at −80 °C. For each transformation, 100 µl of chemically competent DH5α cells was thawed on ice and mixed with the complete ligation sample (see **5.7.5**), the complete Gibson Assembly sample (see **5.7.6**), or 100 ng of plasmid DNA. The mixture was incubated on ice for 30 minutes. Cells were then treated with heat shock at 42 °C for 2 minutes, immediately followed by addition of 1 ml LB medium (see **5.2.1**) and incubation at 37 °C for 45 – 90 minutes. Cells were collected by centrifugation at 600x g for 5 minutes and most of the supernatant was removed. The pellet was resuspended in the residual, approximately 100 µl, LB medium and plated on an LB agar plate that contained the necessary antibiotic, in this study always ampicillin. The plates were kept overnight at 37 °C and growth of transformed cells were checked on the next day.

### 5.7.8. Colony PCR of *E. coli*

Cells of multiple *E. coli* colonies transformed with the cloning product were used directly in PCRs to initially screen if the constructs were correct. Primers were chosen that would (or would not) yield a PCR product only if the plasmids contained the correct recombinant sequences. The non-proofreading DreamTaq DNA Polymerase was used (see **5.7.1**). Cell material was picked from the LB agar plate with a sterile toothpick and smeared on the inner wall of the PCR reaction tube. The toothpick with remaining cell material was placed in a tube that contained 5 ml LB medium (with 100 µg/ml ampicillin; see **5.2.1**) and incubated at 37 °C with rotation overnight to further cultivate the *E. coli* cells for later use (see *5.7.9.1*). Components of the PCR sample, except for the template DNA, were added to the reaction tube according to the manufacturer's protocol (**Table 12**). For the thermocycling program (**Table 13**), the denaturation time was increased to 15 minutes. PCR products were analyzed via agarose gel electrophoresis (see **5.7.2**).

### 5.7.9. Extraction and purification of plasmids from *E. coli* cells

#### 5.7.9.1. Small-scale (mini) preparation

Small amounts of plasmids were purified from *E. coli* cultures with two different methods.

For organic extraction, cells were collected from 5 ml *E. coli* cultures by centrifugation at 2 000x g for 15 minutes and the supernatant was removed. Cells were resuspended in 100 µl Solution 1, mixed thoroughly with 200 µl Solution 2 by inverting the tube several times, and incubated at room temperature for 5 minutes. Subsequently, 100 µl ice-cold Solution 3 was added, mixed thoroughly by inverting the tube several times, and the mixture was incubated on ice for 5 minutes. The mixture was centrifuged at 12 000x g for 5 minutes and the supernatant was transferred into a fresh tube. To this, 400 µl Phenol/Chloroform/Isoamyl alcohol was added and the sample was mixed thoroughly before centrifugation at 12 000x g for 5 minutes to separate the organic and aqueous phases. The upper, aqueous phase was transferred into a new tube, mixed with 800 µl 100% ethanol, and incubated on ice for 10 minutes to precipitate DNA. The mixture was centrifuged at 12 000x g for 5 minutes and the supernatant was removed. The DNA pellet was washed once with 1 ml 70% ethanol and air-dried at room temperature. The pellet was resuspended in nuclease-free water and 1 µl RNase A (10mg/ml) was added. DNA concentration was measured with NanoDrop™ and the purified plasmid was used for PCR, restriction enzyme digestion, and transformation of yeast cells, or stored at −20 °C.

**Solution 1**

| Tris (pH 8) | 25 mM |
|---|---|
| EDTA | 10 mM |
| Glucose | 1% (w/v) |

**Solution 2**

| NaOH | 0.2 N |
|---|---|
| SDS (sodium dodecyl sulfate ) | 1% (w/v) |

**Solution 3 (pH 5.2)**

| Potassium acetate | 2.5 M |
|---|---|
| Acetic acid | 9.6% (v/v) |

For higher DNA purity, small-scale purifications of plasmids were done with the Nucleospin® Plasmid kit following the manufacturer's protocol. DNA was finally eluted in Buffer AE (5 mM Tris/HCl, pH 8.5) and DNA concentration was measured with NanoDrop™. The purified plasmids were used for DNA sequencing (see **5.7.10**) in addition to PCR, restriction enzyme digestion, and transformation of yeast cells, or stored at −20 °C.

*5.7.9.2. Larger-scale (Midi) preparation*
To obtain larger amounts of plasmids, the NucleoBond® Xtra Midi/Maxi kit was used following the manufacturer's instructions. DNA was finally eluted in sterile nuclease-free water and DNA concentration was measured, adjusted to about 1 µg/µl, and stored at −20 °C or used in further applications.

**5.7.10. DNA sequencing**
After analysis through colony PCR of *E. coli* (see **5.7.8**) and restriction digestion (see **5.7.4**), the constructed plasmids were finally verified through DNA sequencing. Samples that contained 0.5 – 1 µg of a purified plasmid and 20 pmol of the selected sequencing primer were provided to LGC Genomics for sequencing. After validation of the sequences, the constructs were used in yeast transformation (see **5.4**) for other experiments or used further in cloning.

### 5.7.11. Constructed plasmids

The cloning strategy of each constructed plasmid is shortly summarized. Detail features of the used plasmids can be found in **Table 9**, and the full sequences of the used oligonucleotides can be found in **Table 10**. A list of the used restriction enzymes and restriction digestion buffers can be found in **Table 6**. PCR was described in **5.7.1**, restriction enzyme digestion was described in **5.7.4**, ligation was described in **5.7.5**, and Gibson Assembly was described in **5.7.6**.

**pHK1625** *UPF1-GFP* was constructed with Gibson Assembly. The plasmid pHK1592 was used as the vector and was digested with the restriction enzymes PacI and AscI to remove the HA tag sequence. The insert fragment (*–GFP–*) was amplified with the Q5® polymerase using pHK235 as template with the primers HK2935 and HK2936.

**pHK1626** *upf1-DE572AA-GFP* was constructed with Gibson Assembly. The plasmid pHK1593 was used as the vector and was digested with the restriction enzymes PacI and AscI to remove the HA tag sequence. The insert fragment (*–GFP–*) was amplified with the Q5® polymerase using pHK235 as template with the primers HK2935 and HK2936.

**pHK1628** *UPF1-N-GFPsplit* was constructed by ligation of the vector and insert fragments. The plasmid pHK1592 was used as the vector and the insert fragment (*–N-GFPsplit–*) was amplified with the Q5® polymerase using pHK1321 as template with the primers HK2987 and HK2988. The vector and the insert fragment were both digested with the restriction enzymes PacI and AscI and ligated together, removing the HA tag sequence.

**pHK1629** *GBP2-C-GFPsplit* was constructed by ligation of the vector and insert fragments. The plasmid pHK367 was used as the vector and the insert fragment (*–C-GFPsplit–*) was amplified with the Q5® polymerase using pHK1322 as template with the primers HK2989 and HK2990. The vector and the insert fragment were both digested with the restriction enzymes BamHI and XhoI and ligated together, removing the *GFP* sequence.

**pHK1630** *HRB1-C-GFPsplit* was constructed by ligation of the vector and insert fragments. The plasmid pHK537 was used as the vector and the insert fragment (*–C-GFPsplit–*) was amplified with the Q5® polymerase using pHK1322 as template with the primers HK2989 and HK2990. The vector and the insert fragment were both digested with the restriction enzymes BamHI and XhoI and ligated together, removing the *GFP* sequence.

**pHK1633** *upf1-DE572AA-N-GFPsplit (HIS)* was constructed by ligation of the vector and insert fragments. The plasmid pHK1593 was used as the vector and the insert fragment (–*N-GFPsplit*–) was amplified with the Q5® polymerase using pHK1321 as template with the primers HK2987 and HK2988. The vector and the insert fragment were both digested with the restriction enzymes PacI and AscI and ligated together, removing the HA tag sequence. The pHK1633 plasmid contains a *HisMX4* cassette that was already present in the pHK1593 vector backbone, and was later removed to create pHK1686 (see below).

**pHK1642** *P~GAL1~MYC-CBP80^PTC^ (HIS)* was constructed with Gibson Assembly. The plasmid pHK1600 was used as the vector and was digested with the restriction enzymes PstI and SacI to remove the *URA3* sequence. The insert fragment (–*HIS3*–) was amplified with the KAPAHiFi™ polymerase using pHK101 as template with the primers HK3193 and HK3194.

**pHK1644** *P~GAL1~N-GFPsplit-UPF1* was constructed by Gibson Assembly with two insert fragments. The plasmid pHK1636 was used as the vector and was digested with the restriction enzymes PstI and BcuI to remove the *NPL3* sequence. The first insert fragment (–*N-GFPsplit*–) was amplified with the KAPAHiFi™ polymerase using pHK1321 as template with the primers HK4427 and HK4428. The second insert fragment (–*UPF1*–) was amplified with the KAPAHiFi™ polymerase using pHK1628 as template with the primers HK4429 and HK4430.

**pHK1686** *upf1-DE572AA-N-GFPsplit* was constructed by removal of the *HisMX4* cassette of pHK1633. The plasmid pHK1633 was digested with the restriction enzymes AjiI and EheI to remove the *HisMX4* sequence. The resulting blunt ends of the digested plasmid were directly ligated together.

## 5.8. Biochemical methods for protein analysis

### 5.8.1. Preparation of yeast whole cell lysates

In general, 400 ml, or larger volumes when required, yeast cultures were grown to log phase and cells were collected by centrifugation at 2 000x g for 5 minutes at 4 °C. The pellets were resuspended in distilled water and transferred into 2 ml screw cap tubes. Cells were collected by centrifugation at 12 000x g for 1 minute, the supernatant was removed, and the pellets were used directly or frozen in liquid nitrogen and stored at −20 °C. In the case of formaldehyde cross-linking (**Figure 13**), prior to centrifugation to collect cells, formaldehyde was added to the cultures to a final concentration of 0.5%, mixed well, and incubated at room temperature with shaking for 10 minutes (based on Klockenbusch and Kast, 2010; Nilsen, 2014, with modification). Glycine was immediately added to a final concentration of 0.5 M and incubated at room temperature for 5 minutes with shaking to quench excess formaldehyde. Cells were then collected as described above.

For the following lysis procedure, cells and lysates were always kept on ice. Pellets were resuspended in one pellet-volume of ice-cold PBSKMT buffer with 5 µl cOmplete™, EDTA-free Protease Inhibitor Cocktail for every 100 µl pellet. One pellet-volume of glass beads (0.4 – 0.6 mm) were added and cells were lysed with the FastPrep-24™ homogenizer at 5 m/s for 30 s, three times, with 5 minutes on ice in between. Cell debris and glass beads were precipitated by centrifugation at 12 000x g for 5 minutes at 4 °C. The cell lysates (supernatant) were transferred into new tubes and cleared by centrifugation at 12 000x g for 10 minutes at 4 °C, repeatedly when required. For formaldehyde cross-linked samples (**Figure 13**), the lysates were sonicated after homogenization to release the proteins of interest from insoluble cellular material (Spencer and Davie, 2002). For this, the tubes were placed in an ultrasonic bath and sonicated for 1 minute, 3 times, with 1 minute on ice in between. Cell lysates were then separated from cell debris and glass beads and cleared by several centrifugation steps as described above.

Small samples of cleared lysates were taken for lysate controls used in SDS-PAGE and western blot (see **5.8.3** and **5.8.4**). The rest was used for protein immunoprecipitation (see **5.8.2**).

**PBSKMT buffer**

| | |
|---|---|
| PBS (see **5.5.3**) | |
| KCl | 3 mM |
| MgCl₂ | 2.5 mM |
| Triton X-100* | 0.5% (v/v) or 0.1% (v/v) for study of cytoplasmic proteins |

*Added freshly before use

### 5.8.2. Protein co-immunoprecipitation

For all following steps, beads and cell samples were kept on ice. To immunoprecipitate GFP-tagged proteins, the GFP-Trap®_A or GFP Selector beads were used. For each IP reaction, 10 µl slurry of beads were washed by adding 1 ml ice-cold PBSKMT to the beads, mixing, and removing most of the buffer after centrifugation at 400x g for 2 minutes, repeated 5 times. The beads were incubated with 1 ml 3% (w/v) bovine serum albumin (BSA, Albumin Fraction V) in PBSKMT with rotation at 25 °C for 1 hour to block the beads to reduce later unspecific binding. Beads were re-collected by centrifugation at 400x g for 2 minutes and washed 2 times with 1 ml PBSKMT to remove excess BSA. BSA-blocked beads were combined with cleared cell lysates and incubated at 4 °C with rotation for 1.5 – 2 hours. When indicated, RNase A was added to a final concentration of 200 µg/ml for the last 30 minutes of incubation.

For **Figure 29**, to reduce unspecific binding of Gbp2 and Hrb1 to the beads, the cleared cell lysates were incubated with MYC-Trap®_A beads prior to incubation with GFP Selector beads. For this, 10 µl slurry of MYC-Trap®_A beads for each sample were washed 5 times with PBSKMT as described above, combined with cleared lysates, and incubated at 4 °C for 1 hour with rotation. MYC-Trap®_A beads were precipitated with centrifugation at 400x g for 2 minutes and lysates were transferred to a new tube and combined with previously washed and BSA-blocked GFP Selector beads. Immunoprecipitation was carried out further as described above.

For **Figure 14**, immunoprecipitation was carried out with an anti-HA antibody and IgG-Sepharose beads. The IgG-Sepharose beads were prepared as described for the GFP beads, without blocking with BSA. For immunoprecipitation, the cleared lysates were incubated with 1 µl of Anti-HA-Tag Antibody (F-7) with rotation at 4 °C for 1.5 hours. Subsequently, 10 µl of previously washed IgG-Sepharose beads were added and the samples were incubated with rotation at 4 °C for 4 hours.

After the final incubation steps, beads were collected by centrifugation at 400x g for 2 minutes at 4 °C and washed 4 – 7 times with 1 ml PBSKMT. For this, 1 ml buffer was added to the beads and the tubes were gently inverted 5 times. Beads were re-collected by centrifugation at 400x g for 2 minutes at 4 °C and the supernatant was carefully discarded. After the last wash, 20 µl of 2x SDS sample buffer was added to the beads and the samples were used for SDS-PAGE analysis (see **5.8.3**) or stored at −20 °C.

### 5.8.3. SDS-PAGE

Proteins were separated according to size on denaturing SDS polyacrylamide gels in SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) as previously described (Garfin, 2009). The polyacrylamide gels contained a 5% gel for protein stacking on top of a 10% gel for protein resolving (**Table 14**), and were casted using self-made casting devices. Self-made combs were inserted into the stacking gels before polymerization to create sample wells. The gels were placed in self-made electrophoresis chambers, and the chambers were filled with SDS electrophoresis buffer.

**Table 14. Components of the SDS polyacrylamide gels**

|  | Stacking gel | Resolving gel |
|---|---|---|
| ROTIPHORESE® Gel 30 acrylamide | 5% (v/v) | 10% (v/v) |
| Tris (pH 8.8) | - | 375 mM |
| Tris (pH 6.8) | 125 mM | - |
| SDS | 0.1% (w/v) | 0.1% (w/v) |
| APS (Ammonium persulfate) | 0.1% (w/v) | 0.1% (w/v) |
| TEMED (Tetramethylethylenediamine)* | 0.1% (v/v) | 0.04% (v/v) |

*Added last for polymerization

| SDS electrophoresis buffer | |
|---|---|
| Tris | 25 mM |
| Glycine | 192 mM |
| SDS | 0.1% (w/v) |

For SDS-PAGE samples, cleared whole cell lysates and immunoprecipitation eluates, prepared as described in **5.8.1** and **5.8.2**, were mixed with 2x SDS sample buffer and heated at 95 °C for 5 minutes to denature the proteins. The formaldehyde cross-linked samples were heated at 95 °C for 20 minutes at this step to reverse cross-linking. The

samples were briefly mixed, centrifuged, and loaded into the wells of the polyacrylamide gels. The complete eluate samples were loaded and 15 – 30 µl of the lysate controls were loaded. Cozy™ Prestained Protein Ladder or PageRuler™ Prestained Protein Ladder was used as size standards. Electrophoresis was performed under constant current of 6 – 35 mA until the proteins had migrated for the desired distance.

**2x SDS sample buffer**

| | |
|---|---|
| Tris (pH 6.8) | 125 mM |
| SDS | 4% (w/v) |
| Glycerol | 20% (v/v) |
| Bromophenol blue | 0.02% (w/v) |
| 2-Mercaptoethanol* | 10% (v/v) |

*Added freshly before use

## 5.8.4. Western blot

The proteins that were separated on the polyacrylamide gels were transferred onto nitrocellulose membranes by electroblotting (Towbin et al., 1979; Burnette, 1981). One piece of nitrocellulose membrane and two pieces of Whatman paper were prepared for each polyacrylamide gel in corresponding sizes slightly bigger than the gel. In a semi-dry electro blotter, blots were assembled as follows: anode – Whatman paper – nitrocellulose membrane – polyacrylamide gel – Whatman paper – cathode. All materials were completely soaked in blotting buffer prior to blot assembly and maintained wet. Air bubbles were carefully removed and blotting was performed at a constant current of 1.5 mA/cm$^2$ membrane area for 1.5 – 2 hours.

**Blotting buffer**

| | |
|---|---|
| Tris | 25 mM |
| Glycine | 192 mM |
| Methanol* | 20% (v/v) |

(Towbin et al., 1979)

*Added freshly before use

After blotting, membranes were immersed in Ponceau S solution for approximately 5 minutes and de-stained with distilled water for visualization of protein bands and control of transfer efficiency. If required, the membranes were cut and Ponceau staining was further removed by washing with water or TBST buffer. Membranes were blocked with 5% (w/v) non-fat milk powder in TBST at room temperature with gentle agitation for 1 hour.

Subsequently, membranes were immersed in the respective primary antibodies (**Table 7**) diluted in 2% (w/v) non-fat milk powder in TBST and incubated at 4 °C overnight with gentle agitation. Primary antibody solutions were removed and membranes were washed with TBST under gentle agitation for 10 minutes, 3 times, at room temperature. Corresponding secondary antibodies (**Table 7**) were diluted in 2% (w/v) non-fat milk powder in TBST and applied to the membranes for 2 hours of incubation at room temperature with gentle agitation. Membranes were washed 3 times with TBST as described above and finally washed once in TBS (TBST without Tween-20) or distilled water to remove Tween-20. The WesternBright™ HRP substrate solution was applied to the membranes and protein signals were detected through chemiluminescence using the Fusion machine and its software.

**Ponceau S solution**

| | |
|---|---|
| Ponceau S | 0.2% (w/v) |
| Acetic acid | 5% (v/v) |

**TBST buffer (pH 7.4)**

| | |
|---|---|
| Tris | 50 mM |
| NaCl | 150 mM |
| Tween-20* | 0.1% (v/v) |

*Added freshly before use

### 5.8.5. Quantification of western blot signals

Western blot signals of the proteins of interest were quantified with the Bio1D software. The optical densities of the protein bands were measured from selected western blot images in which the signal intensities were not saturated. The rolling ball method was used and the detection threshold was adjusted for subtraction of unspecific background signals.

# 5.9. Biochemical methods for RNA analysis

## 5.9.1. RNA co-immunoprecipitation

### 5.9.1.1. Preparation of whole cell lysates

Whole cell lysates for RNA co-immunoprecipitation (RIP) experiments were prepared as described in **5.8.1**. For UV cross-linking (**Figure 24**), cells were treated with UV radiation before being collected (Sei and Conrad, 2014). Cells were pelleted by centrifugation at 2 000x g for 5 minutes at 4 °C. Pellets were resuspended in 50 ml of the respective media, poured into 15 cm petri dishes, and exposed to 254 nm UV radiation for 3 minutes 30 seconds, twice, in a UV-crosslinking chamber while placed on ice-cold cooling blocks. The petri dishes were shortly shaken to resuspend cells in between the two treatments. Cell suspension was collected into falcon tubes and centrifuged at 2 000x g for 5 minutes at 4 °C to pellet the cross-linked cells. Pellets were washed with distilled water and collected in 2 mL screw cap tubes as described in **5.8.1**. Pellets were directly used in further experiments or rapidly frozen in liquid nitrogen and stored at −20 °C.

Cell lysis was done principally as described in **5.8.1**. Instead of PBSKMT, pellets were resuspended in RNA co-immunoprecipitation (RIP) buffer with 5 µl cOmplete™, EDTA-free Protease Inhibitor Cocktail for every 100 µl pellet. In addition, 0.5 µl RNase inhibitor RiboLock was added for every 400 µl of pellet. Cells were lysed with FastPrep-24™ as described in **5.8.1**. Small samples of cleared lysates were kept for lysate controls in SDS-PAGE (see **5.8.3**) and 50 µl of cleared lysates were transferred into new tubes for RNA isolation (see below).

| RNA co-immunoprecipitaiton (RIP) buffer | |
|---|---|
| Tris (pH 7.5) | 25 mM |
| NaCl | 150 mM |
| MgCl$_2$ | 2 mM |
| PMSF (phenylmethylsulfonyl fluoride)* | 0.2 mM |
| DTT* | 0.5 mM |
| Triton X-100* | 0.2% (v/v) |

*Components were added freshly before use. 100mM stock PMSF solution in isopropanol was prepared and heated at 65 °C for 5 minutes to dissolve PMSF before adding to the RIP buffer to the final concentration.

*5.9.1.2. Protein immunoprecipitation*

Immunoprecipitation was done in principle similarly to protein co-IP experiments (see **5.8.2**). Cleared cell lysates were combined with GFP beads (previously washed 5 times with RIP buffer) and incubated with rotation at 4 °C for 2 hours. The 50 µl lysate samples prepared in the previous section (see *5.9.1.1*) were combined with 5 µl of DNase I and incubated together with the IP samples at 4 °C for 2 hours with rotation. Afterwards, RIP buffer was used for washing of the beads, while the DNase-treated lysate samples were kept on ice. During the last wash, the beads were split into two portions: after the beads were resuspended in buffer by inverting the tubes, 200 – 300 µl of the suspension was transferred with a cut tip into a new tube. All samples were centrifuged at 400x g at 4 °C for 2 minutes to precipitate the beads and excess buffer was removed. The smaller portions of the split beads were combined with 20 µl of 2x SDS sample buffer (see **5.8.3**), and together with the lysate controls (see *5.9.1.1*), used for SDS-PAGE and western blot analysis (see **5.8.3** and **5.8.4**) to verify protein pull-down. The larger portions were resuspended in 100 µl RIP buffer and 5 µl of DNase I was added. Both the eluate and the lysate samples that contained DNase I were incubated at 25 °C for 30 minutes with rotation for removal of DNA. Subsequently, 1 ml TRIzol™ was directly added to the samples for RNA isolation (see *5.9.3.1*).

For UV cross-linked samples, additional proteinase treatment was carried out to remove proteins before RNA isolation. After immunoprecipitation and 4 times washing of the beads with RIP buffer, the beads were further washed two times with 1 ml Proteinase K buffer without SDS. In the second wash the beads were split into two portions as described in the previous paragraph. The smaller portions were prepared as SDS-PAGE samples. The bigger portions were resuspended in 100 µl Proteinase K buffer and incubated with 40 µg Proteinase K at 50 °C with agitation for 1.5 hours. For the 50 µl lysate samples, 5 mM EDTA and 0.5% SDS were added and the samples were incubated with 80 µg Proteinase K as described above. Afterwards, 1 ml TRIzol™ was directly added to the samples for RNA isolation (see *5.9.3.1*).

**Proteinase K buffer** (always prepared freshly)

| | |
|---|---|
| Tris (pH 7.5) | 50 mM |
| EDTA | 5 mM |
| NaCl | 50 mM |
| DTT | 0.5 mM |
| Triton X-100 | 0.2% (v/v) |
| SDS | 0.5% (w/v) |

## 5.9.2. DEPC water

For all samples and experiments concerning purified RNA, RNase-free diethyl pyrocarbonate-treated water (DEPC water) was used. For preparation of DEPC water, 0.1% (v/v) DEPC was added to double-distilled water and mixed thoroughly by rigorous shaking and incubation with constant stirring at room temperature overnight. The water was then autoclaved to inactivate residual DEPC.

## 5.9.3. RNA isolation

### 5.9.3.1. RNA purification from RIP lysate and eluate samples

RNA was isolated from cleared yeast whole cell lysates or RNA co-immunoprecipitation eluates using the TRIzol™ Reagent (Chomczyński and Sacchi, 1987) according to the manufacturer's instructions. 1 ml TRIzol™ was added to the lysates or IP beads and the samples were incubated with agitation at 65 °C for 10 minutes. Afterwards, 200 µl chloroform was added, mixed thoroughly, and the samples were centrifuged at 12 000x g for 15 minutes at room temperature. The aqueous (upper) phase was transferred into a new tube to which an equal volume of 100% isopropanol was added. For the lysate and eluate samples, respectively, 10 µg glycogen and 1 µl GlycoBlue™ Coprecipitant were also added. The samples were mixed thoroughly and incubated at −20 °C overnight for RNA precipitation, followed by centrifugation at 12 000x g for 30 minutes at 4 °C. The supernatant was removed and the RNA pellets were washed carefully with 1 ml ice-cold 75% ethanol (in DEPC water) at least twice. After each wash the supernatant was removed as completely as possible. The RNA pellets were dried at 65 °C and dissolved in DEPC water, and RNA concentration was measured with the NanoDrop™ spectrophotometer. The RNA was used further (see **5.9.4** and **5.9.5**) or stored at −20 °C for short-term and at −80 °C for long-term.

### 5.9.3.2. Total RNA extraction from yeast cells

For isolation of total RNA from yeast cells (**Figure 25**), the Nucleospin® RNA kit was used according to the manufacturer's instructions except for the lysate preparation steps. For this, 15 − 20 ml log phase yeast cultures were centrifuged at 2 000x g for 5 minutes at 4 °C. The cells were lysed in 350 µl Buffer RA1 (provided in the kit) with 3.5 µl 2-mercaptoethanol and 200 µl glass beads (0.4 − 0.6 mm) using the FastPrep-24™ homogenizer at 5 m/s for 30 seconds, three times, with 5 minutes incubation on ice in

between. The samples were centrifuged at 12 000x g at 4 °C for 1 minute and the lysates (supernatant) were collected for RNA isolation with the kit. RNA was finally eluted in 60 μl RNase-free water and the concentration was measured with NanoDrop™ before further usage or storage at −20 or −80 °C.

### 5.9.4. TURBO DNase treatment

In order to remove residual DNA in the purified RNA samples, DNase treatment was performed using the TURBO DNA-*free*™ kit following the manufacturer's guidelines. Briefly, the RNA samples were diluted to 200 ng/μl or lower concentrations, mixed with TURBO DNase buffer and TURBO DNase, and incubated at 37 °C for 30 minutes. Samples were treated with DNase Inactivation Reagent, centrifuged, and purified RNA in the supernatant was transferred into new tubes.

### 5.9.5. cDNA synthesis (reverse transcription)

Isolated RNA was reverse transcribed into cDNA with Maxima Reverse Transcriptase or the FastGene® Scriptase II cDNA synthesis kit following the manufacturers' protocols. Briefly, equal amounts of RNA from each sample were mixed with oligo (dT)$_{18}$ primers or random hexamer primers, dNTPs, reverse transcriptase buffer, DTT, RiboLock RNase inhibitor, reverse transcriptase and sterile nuclease-free water. For reverse transcription with Maxima Reverse Transcriptase, samples were incubated at 50 °C for 30 minutes followed by reaction termination at 85 °C for 5 minutes. For FastGene® Scriptase II, samples were incubated at 42 °C for 50 minutes for reverse transcription followed by incubation at 70 °C for 15 minutes to deactivate the enzyme.

The cDNA samples were diluted appropriately with DEPC water for optimal $C_T$ values in qPCRs (see **5.9.6**). For each cDNA sample, a corresponding NRT control sample was prepared that contained identical components without the reverse transcriptase.

### 5.9.6. Quantitative polymerase chain reaction (qPCR) and data analysis

Real-Time PCRs were carried out with the cDNA samples to detect relative amounts of certain RNA species using CFX Connect Real-Time PCR Detection Systems. qPCR samples were prepared on ice with the 2x qPCRBIO SyGreen Mix Lo-ROX master mix according to the manufacturer's protocol with modifications: 10 μl instead of 20 μl reactions were used and typically 80 nM of forward and reverse primers were added. For

template DNA, 2 µl of cDNA was used. PCR primers were designed according to standard guidelines and listed in **Table 11**. The thermocycling program is shown in **Table 15**.

All cDNA samples were pipetted in triplicates and the mean $C_T$ from three values was used for data analysis. When one value was an extreme outlier, it was neglected. For each cDNA sample, the NRT control was pipetted once and its $C_T$ value was used to control reverse transcription efficiency and identify DNA contamination. The melting curves and amplification efficiencies were used to control the validity of the primer pairs. When they were not optimal, the concentration of the primers and the temperature of the annealing and extension step were adjusted. The obtained $C_T$ data were analyzed according to the $2^{-\Delta\Delta C_T}$ method (Livak and Schmittgen, 2001).

**Table 15. qPCR parameters**

|  | Temperature | Time | Repeat |
|---|---|---|---|
| **Initial denaturation** | 95 °C | 3 min | 1x |
| **Denaturation** | 95 °C | 10 sec | |
| **Annealing and Extension** | 50 – 65 °C | 20 sec | 45x |
| **SYBR Green measurement** | | | |
| **Melting curve** | 65 – 95 °C / 0.5 °C, 5 sec | | 1x |

## 5.10. Statistical analysis and figures

Bar graphs show mean values of the biological repeats and error bars represent standard deviations from the mean. Unpaired, two-tailed Student's *t*-tests were performed with Microsoft® Excel for statistical analyses. The *p*-values were used to determine statistical significance and indicated in the figures by asterisks. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. The asterisks in the bar graphs represent significant differences compared with the respective wild type or wild-typical sample, whose value was set to 1 or 100%.

Figures shown in the Results and Discussion sections were made and arranged using Microsoft® Excel and Adobe Illustrator.

# 6. References

Abruzzi, K.C., Lacadie, S., and Rosbash, M. (2004). Biochemical analysis of TREX complex recruitment to intronless and intron-containing yeast genes. EMBO J. *23*, 2620–2631.

Adivarahan, S., Livingston, N., Nicholson, B., Rahman, S., Wu, B., Rissland, O.S., and Zenklusen, D. (2018). Spatial Organization of Single mRNPs at Different Stages of the Gene Expression Pathway. Mol. Cell *72*, 727–738.e5.

Aibara, S., Gordon, J.M., Riesterer, A.S., McLaughlin, S.H., and Stewart, M. (2017). Structural basis for the dimerization of Nab2 generated by RNA binding provides insight into its contribution to both poly(A) tail length determination and transcript compaction in *Saccharomyces cerevisiae*. Nucleic Acids Res. *45*, 1529–1538.

Alcázar-Román, A.R., Tran, E.J., Guo, S., and Wente, S.R. (2006). Inositol hexakisphosphate and Gle1 activate the DEAD-box protein Dbp5 for nuclear mRNA export. Nat. Cell Biol. *8*, 711–716.

Alkalaeva, E.Z., Pisarev, A.V., Frolova, L.Y., Kisselev, L.L., and Pestova, T.V. (2006). In Vitro Reconstitution of Eukaryotic Translation Reveals Cooperativity between Release Factors eRF1 and eRF3. Cell *125*, 1125–1136.

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*, 112–118.

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,* 641–650.

Araki, Y., Takahashi, S., Kobayashi, T., Kajiho, H., Hoshino, S., and Katada, T. (2001). Ski7p G protein interacts with the exosome and the Ski complex for 3'-to-5' mRNA decay in yeast. EMBO J. *20*, 4684–4693.

Ares, M., Jr., Grate, L., and Pauling, M.H. (1999). A handful of intron-containing genes produces the lion's share of yeast mRNA. RNA *5*, 1138–1139.

Aznarez, I., Nomakuchi, T.T., Tetenbaum-Novatt, J., Rahman, M.A., Fregoso, O., Rees, H., and Krainer, A.R. (2018). Mechanism of Nonsense-Mediated mRNA Decay Stimulation by Splicing Factor SRSF1. Cell Rep. *23*, 2186–2198.

Baejen, C., Torkler, P., Gressel, S., Essig, K., Söding, J., and Cramer, P. (2014). Transcriptome Maps of mRNP Biogenesis Factors Define Pre-mRNA Recognition. Mol. Cell *55*, 745–757.

Baierlein, C., Hackmann, A., Gross, T., Henker, L., Hinz, F., and Krebber, H. (2013). Monosome Formation during Translation Initiation Requires the Serine/Arginine-Rich Protein Npl3. Mol. Cell. Biol. *33*, 4811–4823.

Barberan-Soler, S., Lambert, N.J., and Zahler, A.M. (2009). Global analysis of alternative splicing uncovers developmental regulation of nonsense-mediated decay in *C. elegans*. RNA *15*, 1652–1660.

Barthelme, D., Dinkelaker, S., Albers, S.-V., Londei, P., Ermler, U., and Tampé, R. (2011). Ribosome recycling depends on a mechanistic link between the FeS cluster domain and a conformational switch of the twin-ATPase ABCE1. Proc. Natl. Acad. Sci. USA *108*, 3228–3233.

Becker, D., Hirsch, A.G., Bender, L., Lingner, T., Salinas, G., and Krebber, H. (2019). Nuclear Pre-snRNA Export Is an Essential Quality Assurance Mechanism for Functional Spliceosomes. Cell Rep. *27*, 3199–3214.e3.

Beißel, C., Neumann, B., Uhse, S., Hampe, I., Karki, P., and Krebber, H. (2019). Translation termination depends on the sequential ribosomal entry of eRF1 and eRF3. Nucleic Acids Res. *47*, 4798–4813.

Bentley, D.L. (2014). Coupling mRNA processing with transcription in time and space. Nat. Rev. Genet. *15*, 163–175.

Bertani, G. (1951). STUDIES ON LYSOGENESIS. I. The Mode of Phage Liberation by Lysogenic Escherichia coli. J. Bacteriol. *62*, 293–300.

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,* 1226–1235.

Bhatter, N., Roy, R., Shah, S., Sastry, S.P., Parbin, S., Iyappan, R., Kankaria, S., and Rajyaguru, P.I. (2019). Arginine methylation augments Sbp1 function in translation repression and decapping. FEBS J. *286*, 4693–4708.

Birney, E., Kumar, S., and Krainer, A.R. (1993). Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. Nucleic Acids Res. *21*, 5803–5816.

Boehm, V., Kueckelmann, S., Gerbracht, J.V., Kallabis, S., Britto-Borges, T., Altmüller, J., Krüger, M., Dieterich, C., and Gehring, N.H. (2021). SMG5-SMG7 authorize nonsense-mediated mRNA decay by enabling SMG6 endonucleolytic activity. Nat. Commun. *12*, 3965.

Boeke, J.D., Trueheart, J., Natsoulis, G., and Fink, G.R. (1987). 5-Fluoroorotic Acid as a Selective Agent in Yeast Molecular Genetics. Methods Enzymol. *154*, 164–175.

Boisramé, A., Devillers, H., Onésime, D., Brunel, F., Pouch, J., Piot, M., and Neuvéglise, C. (2019). Exon junction complex components Y14 and Mago still play a role in budding yeast. Sci. Rep. *9*, 849.

Bossie, M.A., DeHoratius, C., Barcelo, G., and Silver, P. (1992). A Mutant Nuclear Protein with Similarity to RNA Binding Proteins Interferes with Nuclear Import in Yeast. Mol. Biol. Cell *3*, 875–893.

Botti, V., McNicoll, F., Steiner, M.C., Richter, F.M., Solovyeva, A., Wegener, M., Schwich, O.D., Poser, I., Zarnack, K., Wittig, I., Neugebauer, K.M., and Müller-McNicoll, M. (2017). Cellular differentiation state modulates the mRNA export activity of SR proteins. J. Cell Biol. *216*, 1993–2009.

Brambilla, M., Martani, F., Bertacchi, S., Vitangeli, I., and Branduardi, P. (2019). The *Saccharomyces cerevisiae* poly (A) binding protein (Pab1): Master regulator of mRNA metabolism and cell physiology. Yeast *36*, 23–34.

Bresson, S., and Tollervey, D. (2018). Surveillance-ready transcription: nuclear RNA decay as a default fate. Open Biol. *8*, 170270.

Bretes, H., Rouviere, J.O., Leger, T., Oeffinger, M., Devaux, F., Doye, V., and Palancade, B. (2014). Sumoylation of the THO complex regulates the biogenesis of a subset of mRNPs. Nucleic Acids Res. *42*, 5043–5058.

Buchan, J.R., Muhlrad, D., and Parker, R. (2008). P bodies promote stress granule assembly in *Saccharomyces cerevisiae*. J. Cell Biol. *183*, 441–455.

Bucheli, M.E., and Buratowski, S. (2005). Npl3 is an antagonist of mRNA 3' end formation by RNA polymerase II. EMBO J. *24*, 2150–2160.

Buchwald, G., Ebert, J., Basquin, C., Sauliere, J., Jayachandran, U., Bono, F., Le Hir, H., and Conti, E. (2010). Insights into the recruitment of the NMD machinery from the crystal structure of a core EJC-UPF3b complex. Proc. Natl. Acad. Sci. USA *107*, 10050–10055.

Bühler, M., Steiner, S., Mohn, F., Paillusson, A., and Mühlemann, O. (2006). EJC-independent degradation of nonsense immunoglobulin-μ mRNA depends on 3' UTR length. Nat. Struct. Mol. Biol. *13*, 462–464.

Burnette, W.N. (1981). "Western Blotting": Electrophoretic Transfer of Proteins from Sodium Dodecyl Sulfate–Polyacrylamide Gels to Unmodified Nitrocellulose and Radiographic Detection with Antibody and Radioiodinated Protein A. Anal. Biochem. *112,* 195–203.

Buskirk, A.R., and Green, R. (2017). Ribosome pausing, arrest and rescue in bacteria and eukaryotes. Phil. Trans. R. Soc. B *372*, 20160183.

Byrne, K.P., and Wolfe, K.H. (2005). The Yeast Gene Order Browser: Combining curated homology and syntenic context reveals gene fate in polyploid species. Genome Res. *15*, 1456–1461.

Cáceres, J.F., Screaton, G.R., and Krainer, A.R. (1998). A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm. Genes Dev. *12,* 55–66.

Cai, B., Li, Z., Ma, M., Zhang, J., Kong, S., Abdalla, B.A., Xu, H., Jebessa, E., Zhang, X., Lawal, R.A., and Nie, Q. (2020). Long noncoding RNA *SMUL* suppresses *SMURF2* production-mediated muscle atrophy via nonsense-mediated mRNA decay. Mol. Ther. Nucleic Acids. *23*, 512–526.

Cao, D., and Parker, R. (2003). Computational Modeling and Experimental Analysis of Nonsense-Mediated Decay in Yeast. Cell *113*, 533–545.

Carvalho, T., Martins, S., Rino, J., Marinho, S., and Carmo-Fonseca, M. (2017). Pharmacological inhibition of the spliceosome subunit SF3b triggers exon junction complex-independent nonsense-mediated decay. J. Cell Sci. *130*, 1519–1531.

Cazalla, D., Zhu, J., Manche, L., Huber, E., Krainer, A.R., and Cáceres, J.F. (2002). Nuclear Export and Retention Signals in the RS Domain of SR Proteins. Mol. Cell. Biol. *22*, 6871–6882.

Celik, A., Baker, R., He, F., and Jacobson, A. (2017a). High-resolution profiling of NMD targets in yeast reveals translational fidelity as a basis for substrate selection. RNA *23*, 735–748.

Celik, A., He, F., and Jacobson, A. (2017b). NMD monitors translational fidelity 24/7. Curr. Genet. *63*, 1007–1010.

Chakrabarti, S., Jayachandran, U., Bonneau, F., Fiorini, F., Basquin, C., Domcke, S., Le Hir, H., and Conti, E. (2011). Molecular Mechanisms for the RNA-Dependent ATPase Activity of Upf1 and its Regulation by Upf2. Mol. Cell *41*, 693–703.

Chamieh, H., Ballut, L., Bonneau, F., and Le Hir, H. (2008). NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. Nat. Struct. Mol. Biol. *15*, 85–93.

Chanarat, S., Seizl, M., and Sträßer, K. (2011). The Prp19 complex is a novel transcription elongation factor required for TREX occupancy at transcribed genes. Genes Dev. *25*, 1147–1158.

Chang, J.H., Jiao, X., Chiba, K., Oh, C., Martin, C.E., Kiledjian, M., and Tong, L. (2012). Dxo1 is a new type of eukaryotic enzyme with both decapping and 5'-3' exoribonuclease activity. Nat. Struct. Mol. Biol. *19*, 1011–1017.

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*, 253–264.

Cheng, Z., Saito, K., Pisarev, A.V., Wada, M., Pisareva, V.P., Pestova, T.V., Gajda, M., Round, A., Kong, C., Lim, M., Nakamura, Y., Svergun, D.I., Ito, K., and Song, H. (2009).

Structural insights into eRF3 and stop codon recognition by eRF1. Genes Dev. *23*, 1106–1118.

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.

Chlebowski, A., Lubas, M., Jensen, T.H., and Dziembowski, A. (2013). RNA decay machines: The exosome. Biochim. Biophys. Acta *1829*, 552–560.

Cho, E.-J., Kobor, M.S., Kim, M., Greenblatt, J., and Buratowski, S. (2001). Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. Genes Dev. *15*, 3319–3329.

Chomczyński, P., and Sacchi, N. (1987). Single-Step Method of RNA Isolation by Acid Guanidinium Thiocyanate–Phenol–Chloroform Extraction. Anal. Biochem. *162*, 156–159.

Clissold, P.M., and Ponting, C.P. (2000). PIN domains in nonsense-mediated mRNA decay and RNAi. Curr. Biol. *10*, R888–890.

Coller, J., and Parker, R. (2005). General Translational Repression by Activators of mRNA Decapping. Cell *122*, 875–886.

Collins, S.R., Kemmeren, P., Zhao, X.-C., Greenblatt, J.F., Spencer, F., Holstege, F.C.P., Weissman, J.S., and Krogan, N.J. (2007). Toward a Comprehensive Atlas of the Physical Interactome of *Saccharomyces cerevisiae*. Mol. Cell. Proteomics *6*, 439–450.

Colombo, M., Karousis, E.D., Bourquin, J., Bruggmann, R., and Mühlemann, O. (2017). Transcriptome-wide identification of NMD-targeted human mRNAs reveals extensive redundancy between SMG6- and SMG7-mediated degradation pathways. RNA *23*, 189–201.

Cosson, B., Couturier, A., Chabelskaya, S., Kiktev, D., Inge-Vechtomov, S., Philippe, M., and Zhouravleva, G. (2002). Poly(A)-Binding Protein Acts in Translation Termination via Eukaryotic Release Factor 3 Interaction and Does Not Influence [*PSI*+] Propagation. Mol. Cell. Biol. *22*, 3301–3315.

Coyle, J.H., Bor, Y.-C., Rekosh, D., and Hammarskjold, M.-L. (2011). The Tpr protein regulates export of mRNAs with retained introns that traffic through the Nxf1 pathway. RNA *17*, 1344–1356.

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*, 423–436.

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*, 610–623.

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*, 1665–1677.

Davidson, L., Kerr, A., and West, S. (2012). Co-transcriptional degradation of aberrant pre-mRNA by Xrn2. EMBO J. *31*, 2566–2578.

de Andres-Pablo, A., Morillon, A., and Wery, M. (2017). LncRNAs, lost in translation or licence to regulate? Curr. Genet. *63*, 29–33.

Decker, C.J., and Parker, R. (2002). mRNA decay enzymes: Decappers conserved between yeast and mammals. Proc. Natl. Acad. Sci. USA *99*, 12512–12514.

Dehecq, M., Decourty, L., Namane, A., Proux, C., Kanaan, J., Le Hir, H., Jacquier, A., and Saveanu, C. (2018). Nonsense-mediated mRNA decay involves two distinct Upf1-bound complexes. EMBO J. *37*, e99278.

Dermody, J.L., Dreyfuss, J.M., Villén, J., Ogundipe, B., Gygi, S.P., Park, P.J., Ponticelli, A.S., Moore, C.L., Buratowski, S., and Bucheli, M.E. (2008). Unphosphorylated SR-Like Protein Npl3 Stimulates RNA Polymerase II Elongation. PLoS ONE *3*, e3273.

Doma, M.K., and Parker, R. (2006). Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. Nature *440*, 561–564.

D'Orazio, K.N., Wu, C.C.-C., Sinha, N., Loll-Krippleber, R., Brown, G.W., and Green, R. (2019). The endonuclease Cue2 cleaves mRNAs at stalled ribosomes during No Go Decay. eLife *8*, e49117.

D'Orazio, K.N., and Green, R. (2021). Ribosome states signal RNA quality control. Mol. Cell *81*, 1372–1383.

Durand, S., and Lykke-Andersen, J. (2013). Nonsense-mediated mRNA decay occurs during eIF4F-dependent translation in human cells. Nat. Struct. Mol. Biol. *20*, 702–709.

Eberle, A.B., Stalder, L., Mathys, H., Orozco, R.Z., and Mühlemann, O. (2008). Posttranscriptional Gene Regulation by Spatial Rearrangement of the 3' Untranslated Region. PLoS Biol. *6*, e92.

Eberle, A.B., Lykke-Andersen, S., Mühlemann, 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., and Visa, N. (2014). Quality control of mRNP biogenesis: Networking at the transcription site. Semin. Cell Dev. Biol. *32*, 37–46.

Eckmann, C.R., Rammelt, C., and Wahle, E. (2011). Control of poly(A) tail length. Wiley Interdiscip. Rev. RNA *2*, 348–361.

Elbing, K.L., and Brent, R. (2019). Growth of *E. coli* in Liquid Medium. Curr. Protoc. Mol. Biol. *125*, e81.

Erce, M.A., Abeygunawardena, D., Low, J.K.K., Hart-Smith, G., and Wilkins, M.R. (2013). Interactions Affected by Arginine Methylation in the Yeast Protein–Protein Interaction Network. Mol. Cell. Proteomics *12*, 3184–3198.

Estrella, L.A., Wilkinson, M.F., and González, C.I. (2009). The Shuttling Protein Npl3 Promotes Translation Termination Accuracy in *Saccharomyces cerevisiae*. J. Mol. Biol. *394*, 410–422.

Eyler, D.E., Wehner, K.A., and Green, R. (2013). Eukaryotic Release Factor 3 Is Required for Multiple Turnovers of Peptide Release Catalysis by Eukaryotic Release Factor 1. J. Biol. Chem. *288*, 29530–29538.

Fairman-Williams, M.E., Guenther, U.-P., and Jankowsky, E. (2010). SF1 and SF2 helicases: family matters. Curr. Opin. Struct. Biol. *20*, 313–324.

Fasken, M.B., and Corbett, A.H. (2009). Mechanisms of nuclear mRNA quality control. RNA Biol. *6*, 237–241.

Fatscher, T., Boehm, V., Weiche, B., and Gehring, N.H. (2014). The interaction of cytoplasmic poly(A)-binding protein with eukaryotic initiation factor 4G suppresses nonsense-mediated mRNA decay. RNA *20*, 1579–1592.

Fiorini, F., Bagchi, D., Le Hir, H., and Croquette, V. (2015). Human Upf1 is a highly processive RNA helicase and translocase with RNP remodelling activities. Nat. Commun. *6*, 7581.

Franks, T.M., Singh, G., and Lykke-Andersen, J. (2010). Upf1 ATPase-Dependent mRNP Disassembly Is Required for Completion of Nonsense-Mediated mRNA Decay. Cell *143*, 938–950.

Fritz, S.E., Ranganathan, S., Wang, C.D., and Hogg, J.R. (2020). The RNA-binding protein PTBP1 promotes ATPase-dependent dissociation of the RNA helicase UPF1 to protect transcripts from nonsense-mediated mRNA decay. J. Biol. Chem. *295*, 11613–11625.

Frolova, L.Y., Tsivkovskii, R.Y., Sivolobova, G.F., Oparina, N.Y., Serpinsky, O.I., Blinov, V.M., Tatkov, S.I., and Kisselev, L.L. (1999). Mutations in the highly conserved GGQ motif of class 1 polypeptide release factors abolish ability of human eRF1 to trigger peptidyl-tRNA hydrolysis. RNA *5*, 1014–1020.

Galy, V., Gadal, O., Fromont-Racine, M., Romano, A., Jacquier, A., and Nehrbass, U. (2004). Nuclear Retention of Unspliced mRNAs in Yeast is Mediated by Perinuclear Mlp1. Cell *116*, 63–73.

Gao, Q., Das, B., Sherman, F., and Maquat, L.E. (2005). Cap-binding protein 1-mediated and eukaryotic translation initiation factor 4E-mediated pioneer rounds of translation in yeast. Proc. Natl. Acad. Sci. USA *102*, 4258–4263.

García-Moreno, J.F., and Romão, L. (2020). Perspective in Alternative Splicing Coupled to Nonsense-Mediated mRNA Decay. Int. J. Mol. Sci. *21*, 9424.

Garfin D.E. (2009). One-Dimensional Gel Electrophoresis. Methods Enzymol. *463*, 497–513.

Gatfield, D., and Izaurralde, E. (2004). Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in *Drosophila*. Nature *429*, 575–578.

Gavin, A.C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L.J., Bastuck, S., Dümpelfeld, B., et al. (2006). Proteome survey reveals modularity of the yeast cell machinery. Nature *440*, 631–636.

Ge, Z., Quek, B.L., Beemon, K.L., and Hogg, J.R. (2016). Polypyrimidine tract binding protein 1 protects mRNAs from recognition by the nonsense-mediated mRNA decay pathway. eLife *5*, e11155.

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*, 939–949.

Ghazal, G., Gagnon, J., Jacques, P.-E., Landry, J.-R., Robert, F., and Elela, S.A. (2009). Yeast RNase III Triggers Polyadenylation-Independent Transcription Termination. Mol. Cell *36*, 99–109.

Ghosh, S., Ganesan, R., Amrani, N., and Jacobson, A. (2010). Translational competence of ribosomes released from a premature termination codon is modulated by NMD factors. RNA *16*, 1832–1847.

Ghosh, S., and Jacobson, A. (2010). RNA decay modulates gene expression and controls its fidelity. Wiley Interdiscip. Rev. RNA *1*, 351–361.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., III, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345.

Gibson, D.G. (2011). Enzymatic Assembly of Overlapping DNA Fragments. Methods Enzymol. *498*, 349–361.

Gietz, D., St. Jean, A., Woods, R.A., and Schiestl, R.H. (1992). Improved method for high efficiency transformation of intact yeast cells. Nucleic Acids Res. *20*, 1425.

Gilbert, W., Siebel, C.W., and Guthrie, C. (2001). Phosphorylation by Sky1p promotes Npl3p shuttling and mRNA dissociation. RNA *7*, 302–313.

Gilbert, W., and Guthrie, C. (2004). The Glc7p Nuclear Phosphatase Promotes mRNA Export by Facilitating Association of Mex67p with mRNA. Mol. Cell *13*, 201–212.

González, C.I., Ruiz-Echevarría, M.J., Vasudevan, S., Henry, M.F., and Peltz, S.W. (2000). The yeast hnRNP-like Protein Hrp1/Nab4 Marks a Transcript for Nonsense-Mediated mRNA Decay. Mol. Cell *5*, 489–499.

Graille, M., Chaillet, M., and van Tilbeurgh, H. (2008). Structure of Yeast Dom34: A Protein Related to Translation Termination Factor Erf1 and Involved in No-Go Decay. J. Biol. Chem. *283*, 7145–7154.

Green, D.M., Marfatia, K.A., Crafton, E.B., Zhang, X., Cheng, X., and Corbett, A.H. (2002). Nab2p Is Required for Poly(A) RNA Export in *Saccharomyces cerevisiae* and Is Regulated by Arginine Methylation via Hmt1p. J. Biol. Chem. *277*, 7752–7760.

Green, D.M., Johnson, C.P., Hagan, H., and Corbett, A.H. (2003). The C-terminal domain of myosin-like protein 1 (Mlp1p) is a docking site for heterogeneous nuclear ribonucleoproteins that are required for mRNA export. Proc. Natl. Acad. Sci. USA *100*, 1010–1015.

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–i121.

Green, M.R., and Sambrook, J. (2012). Molecular Cloning: A Laboratory Manual (Fourth Edition). (New York: Cold Spring Habor Laboratory Press).

Gromadzka, A.M., Steckelberg, A.-L., Singh, K.K., Hofmann, K., and Gehring, N.H. (2016). A short conserved motif in ALYREF directs cap- and EJC-dependent assembly of export complexes on spliced mRNAs. Nucleic Acids Res. *44*, 2348–2361.

Grosse, S., Lu, Y.-Y., Coban, I., Neumann, B., and Krebber, H. (2021). Nuclear SR-protein mediated mRNA quality control is continued in cytoplasmic nonsense-mediated decay. RNA Biol. 1–18. Advance online publication.

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.

Gupta, P., and Li, Y.-R. (2018). Upf proteins: highly conserved factors involved in nonsense mRNA mediated decay. Mol. Biol. Rep. *45*, 39–55.

Gwizdek, C., Iglesias, N., Rodriguez, M.S., Ossareh-Nazari, B., Hobeika, M., Divita, G., Stutz, F., and Dargemont, C. (2006). Ubiquitin-associated domain of Mex67 synchronizes recruitment of the mRNA export machinery with transcription. Proc. Natl. Acad. Sci. USA *103*, 16376–16381.

Häcker, S., and Krebber, H. (2004). Differential Export Requirements for Shuttling Serine/Arginine-type mRNA-binding Proteins. J. Biol. Chem. *279*, 5049–5052.

Hackmann, A., Wu, H., Schneider, U.-M., Meyer, K., Jung, K., and Krebber, H. (2014). Quality control of spliced mRNAs requires the shuttling SR proteins Gbp2 and Hrb1. Nat. Commun. *5*, 3123.

Hagan, K.W., Ruiz-Echevarria, M.J., Quan, Y., and Peltz, S.W. (1995). Characterization of *cis*-Acting Sequences and Decay Intermediates Involved in Nonsense-Mediated mRNA Turnover. Mol. Cell. Biol. *15*, 809–823.

Halbach, F., Reichelt, P., Rode, M., and Conti, E. (2013). The Yeast Ski Complex: Crystal Structure and RNA Channeling to the Exosome Complex. Cell *154*, 814–826.

Hansen, K.D., Lareau, L.F., Blanchette, M., Green, R.E., Meng, Q., Rehwinkel, J., Gallusser, F.L., Izaurralde, E., Rio, D.C., Dudoit, S., and Brenner, S.E. (2009). Genome-Wide Identification of Alternative Splice Forms Down-Regulated by Nonsense-Mediated mRNA Decay in *Drosophila*. PLoS Genet. *5*, e1000525.

Hauryliuk, V., Zavialov, A., Kisselev, L., and Ehrenberg, M. (2006). Class-1 release factor eRF1 promotes GTP binding by class-2 release factor eRF3. Biochimie *88*, 747–757.

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. USA *90*, 7034–7038.

He, F., and Jacobson, A. (1995). Identification of a novel component of the nonsense-mediated mRNA decay pathway by use of an interacting protein screen. Genes Dev. *9*, 437–454.

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*, 1439–1452.

He, F., and Jacobson, A. (2015). Nonsense-Mediated mRNA Decay: Degradation of Defective Transcripts Is Only Part of the Story. Annu. Rev. Genet. *49*, 339–366.

Hector, R.E., Nykamp, K.R., Dheur, S., Anderson, J.T., Non, P.J., Urbinati, C.R., Wilson, S.M., Minvielle-Sebastia, L., and Swanson, M.S. (2002). Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export. EMBO J. *21*, 1800–1810.

Hellen, C.U.T. (2018). Translation Termination and Ribosome Recycling in Eukaryotes. Cold Spring Harb. Perspect. Biol. *10*, a032656.

Heyer, E.E., and Moore, M.J. (2016). Redefining the Translational Status of 80S Monosomes. Cell *164*, 757–769.

Hinnebusch, A.G., and Lorsch, J.R. (2012). The Mechanism of Eukaryotic Translation Initiation: New Insights and Challenges. Cold Spring Harb. Perspect. Biol. *4*, a011544.

Ho, B., Baryshnikova, A., and Brown, G.W. (2018). Unification of Protein Abundance Datasets Yields a Quantitative *Saccharomyces cerevisiae* Proteome. Cell Syst. *6*, 192–205.e3.

Hodge, C.A., Colot, H.V., Stafford, P., and Cole, C.N. (1999). Rat8p/Dbp5p is a shuttling transport factor that interacts with Rat7p/Nup159p and Gle1p and suppresses the mRNA export defect of *xpo1-1* cells. EMBO J. *18*, 5778–5788.

Hoek, T.A., Khuperkar, D., Lindeboom, R.G.H., Sonneveld, S., Verhagen, B.M.P., Boersma, S., Vermeulen, M., and Tanenbaum, M.E. (2019). Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay. Mol. Cell *75*, 324–339.e11.

Hoffman, E.A., Frey, B.L., Smith, L.M., and Auble, D.T. (2015). Formaldehyde Crosslinking: A Tool for the Study of Chromatin Complexes. J. Biol. Chem. *290*, 26404–26411.

Hogg, J.R., and Goff, S.P. (2010). Upf1 Senses 3'UTR Length to Potentiate mRNA Decay. Cell *143*, 379–389.

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.

Howard, J.M., and Sanford, J.R. (2015). The RNAissance family: SR proteins as multifaceted regulators of gene expression. Wiley Interdiscip. Rev. RNA *6*, 93–110.

Huang, Y., Gattoni, R., Stévenin, J., and Steitz, J.A. (2003). SR Splicing Factors Serve as Adapter Proteins for TAP-Dependent mRNA Export. Mol. Cell *11*, 837–843.

Huang, Y., and Steitz, J.A. (2005). SRprises along a Messenger's Journey. Mol. Cell *17*, 613–615.

Huntzinger, E., Kashima, I., Fauser, M., Saulière, J., and Izaurralde, E. (2008). SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. RNA *14*, 2609–2617.

Hurt, E., Luo, M.-J., Röther, S., Reed, R., and Sträßer, K. (2004). Cotranscriptional recruitment of the serine-arginine-rich (SR)-like proteins Gbp2 and Hrb1 to nascent mRNA via the TREX complex. Proc. Natl. Acad. Sci. USA *101*, 1858–1862.

Hurt, J.A., Robertson, A.D., and Burge, C.B. (2013). Global analyses of UPF1 binding and function reveal expanded scope of nonsense-mediated mRNA decay. Genome Res. *23*, 1636–1650.

Hwang, J., Sato, H., Tang, Y., Matsuda, D., and Maquat, L.E. (2010). UPF1 Association with the Cap-Binding Protein, CBP80, Promotes Nonsense-Mediated mRNA Decay at Two Distinct Steps. Mol. Cell *39*, 396–409.

Hwang, J., and Maquat, L.E. (2011). Nonsense-mediated mRNA decay (NMD) in animal embryogenesis: to die or not to die, that is the question. Curr. Opin. Genet. Dev. *21*, 422–430.

Iglesias, N., Tutucci, E., Gwizdek, C., Vinciguerra, P., Von Dach, E., Corbett, A.H., Dargemont, C., and Stutz, F. (2010). Ubiquitin-mediated mRNP dynamics and surveillance prior to budding yeast mRNA export. Genes Dev. *24*, 1927–1938.

Inada, T. (2020). Quality controls induced by aberrant translation. Nucleic Acids Res. *48*, 1084–1096.

Inoue, H., Nojima, H., and Okayama, H. (1990). High efficiency transformation of *Escherichia coli* with plasmids. Gene *96*, 23–28.

Inoue, K., Mizuno, T., Wada, K., and Hagiwara, M. (2000). Novel RING Finger Proteins, Air1p and Air2p, Interact with Hmt1p and Inhibit the Arginine Methylation of Npl3p. J. Biol. Chem. *275*, 32793–32799.

Isken, O., Kim, Y.K., Hosoda, N., Mayeur, G.L., Hershey, J.W.B., and Maquat, L.E. (2008). Upf1 Phosphorylation Triggers Translational Repression during Nonsense-Mediated mRNA Decay. Cell *133*, 314–327.

Ivanov, P.V., Gehring, N.H., Kunz, J.B., Hentze, M.W., and Kulozik, A.E. (2008). Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. EMBO J. *27*, 736–747.

Ivanov, A., Mikhailova, T., Eliseev, B., Yeramala, L., Sokolova, E., Susorov, D., Shuvalov, A., Schaffitzel, C., and Alkalaeva, E. (2016). PABP enhances release factor recruitment and stop codon recognition during translation termination. Nucleic Acids Res. *44*, 7766–7776.

Jackson, R.J., Hellen, C.U.T., and Pestova, T.V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nat. Rev. Mol. Cell Biol. *11*, 113–127.

Jaillon, O., Bouhouche, K., Gout, J.-F., Aury, J.-M., Noel, B., Saudemont, B., Nowacki, M., Serrano, V., Porcel, B.M., Ségurens, B., et al. (2008). Translational control of intron splicing in eukaryotes. Nature *451*, 359–362.

Jensen, T.H., Dower, K., Libri, D., and Rosbash, M. (2003). Early Formation of mRNP: License for Export or Quality Control? Mol. Cell *11*, 1129–1138.

Jiao, X., Xiang, S., Oh, C., Martin, C.E., Tong, L., and Kiledjian, M. (2010). Identification of a quality-control mechanism for mRNA 5'-end capping. Nature *467*, 608–611.

Jiao, X., Chang, J.H., Kilic, T., Tong, L., and Kiledjian, M. (2013). A Mammalian Pre-mRNA 5' End Capping Quality Control Mechanism and an Unexpected Link of Capping to Pre-mRNA Processing. Mol. Cell *50*, 104–115.

Joazeiro, C.A.P. (2019). Mechanisms and functions of ribosome-associated protein quality control. Nat. Rev. Mol. Cell Biol. *20*, 368–383.

Johansson, M.J.O., He, F., Spatrick, P., Li, C., and Jacobson, A. (2007). Association of yeast Upf1p with direct substrates of the NMD pathway. Proc. Natl. Acad. Sci. USA *104*, 20872–20877.

Johns, L., Grimson, A., Kuchma, S.L., Newman, C.L., and Anderson, P. (2007). *Caenorhabditis elegans* SMG-2 Selectively Marks mRNAs Containing Premature Translation Termination Codons. Mol. Cell. Biol. *27*, 5630–5638.

Johnson, S.J., and Jackson, R.N. (2013). Ski2-like RNA helicase structures. RNA Biol. *10*, 33–43.

Joncourt, R., Eberle, A.B., Rufener, S.C., and Mühlemann, O. (2014). Eukaryotic Initiation Factor 4G Suppresses Nonsense-Mediated mRNA Decay by Two Genetically Separable Mechanisms. PLoS ONE *9*, e104391.

Karousis, E.D., and Mühlemann, O. (2019). Nonsense-Mediated mRNA Decay Begins Where Translation Ends. Cold Spring Harb. Perspect. Biol. *11*, a032862.

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*, 355–367.

Kataoka, N., Bachorik, J.L., and Dreyfuss, G. (1999). Transportin-SR, a Nuclear Import Receptor for SR Proteins. J. Cell Biol. *145*, 1145–1152.

Kawashima, T., Douglass, S., Gabunilas, J., Pellegrini, M., and Chanfreau, G.F. (2014). Widespread Use of Non-productive Alternative Splice Sites in *Saccharomyces cerevisiae*. PLoS Genet. *10*, e1004249.

Kebaara, B., Nazarenus, T., Taylor, R., Forch, A., and Atkin, A.L. (2003). The Upf-dependent decay of wild-type PPR1 mRNA depends on its 5'-UTR and first 92 ORF nucleotides. Nucleic Acids Res. *31*, 3157–3165.

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. *37*, 2771–2778.

Kervestin, S., and Jacobson, A. (2012). NMD: a multifaceted response to premature translational termination. Nat. Rev. Mol. Cell Biol. *13*, 700–712.

Kervestin, S., Li, C., Buckingham, R., and Jacobson, A. (2012). Testing the *faux*-UTR model for NMD: Analysis of Upf1p and Pab1p competition for binding to eRF3/Sup35p. Biochimie *94*, 1560–1571.

Khong, A., and Parker, R. (2018). mRNP architecture in translating and stress conditions reveals an ordered pathway of mRNP compaction. J. Cell Biol. *217*, 4124–4140.

Khong, A., and Parker, R. (2020). The landscape of eukaryotic mRNPs. RNA *26*, 229–239.

Kilchert, C., and Vasiljeva, L. (2013). mRNA quality control goes transcriptional. Biochem. Soc. Trans. *41*, 1666–1672.

Kilchert, C., Wittmann, S., and Vasiljeva, L. (2016). The regulation and functions of the nuclear RNA exosome complex. Nat. Rev. Mol. Cell Biol. *17*, 227–239.

Kim, V.N., Kataoka, N., and Dreyfuss, G. (2001). Role of the Nonsense-Mediated Decay Factor hUpf3 in the Splicing-Dependent Exon-Exon Junction Complex. Science *293*, 1832–1836.

Kim, E., Magen, A., and Ast, G. (2007). Different levels of alternative splicing among eukaryotes. Nucleic Acids Res. *35*, 125–131.

Kim, J., Park, R.Y., Chen, J.-K., Kim, J., Jeong, S., and Ohn, T. (2014). Splicing factor SRSF3 represses the translation of programmed cell death 4 mRNA by associating with the 5'-UTR region. Cell Death Differ. *21*, 481–490.

Kim, W.K., Yun, S., Kwon, Y., You, K.T., Shin, N., Kim, J., and Kim, H. (2017). mRNAs containing NMD-competent premature termination codons are stabilized and translated under UPF1 depletion. Sci. Rep. *7*, 15833.

Kim, H.-J. (2019). Cell Fate Control by Translation: mRNA Translation Initiation as a Therapeutic Target for Cancer Development and Stem Cell Fate Control. Biomolecules *9*, 665.

Kim, Y.K., and Maquat, L.E. (2019). UPFront and center in RNA decay: UPF1 in nonsense-mediated mRNA decay and beyond. RNA *25*, 407–422.

Kishor, A., Ge, Z., and Hogg, J.R. (2019a). hnRNP L-dependent protection of normal mRNAs from NMD subverts quality control in B cell lymphoma. EMBO J. *38*, e99128.

Kishor, A., Fritz, S.E., and Hogg, J.R. (2019b). Nonsense-mediated mRNA decay: The challenge of telling right from wrong in a complex transcriptome. Wiley Interdiscip. Rev. RNA *10*, e1548.

Kishor, A., Fritz, S.E., Haque, N., Ge, Z., Tunc, I., Yang, W., Zhu, J., and Hogg, J.R. (2020). Activation and inhibition of nonsense-mediated mRNA decay control the abundance of alternative polyadenylation products. Nucleic Acids Res. *48*, 7468–7482.

Klockenbusch, C., and Kast, J. (2010). Optimization of Formaldehyde Cross-Linking for Protein Interaction Analysis of Non-Tagged Integrin β1. J. Biomed. Biotechnol. *2010*, 927585.

Kobayashi, K., Kikuno, I., Kuroha, K., Saito, K., Ito, K., Ishitani, R., Inada, T., and Nureki, O. (2010). Structural basis for mRNA surveillance by archaeal Pelota and GTP-bound EF1α complex. Proc. Natl. Acad. Sci. USA *107*, 17575–17579.

Konkel, L.M.C., Enomoto, S., Chamberlain, E.M., McCune-Zierath, P., Iyadurai, S.J.P., and Berman, J. (1995). A class of single-stranded telomeric DNA-binding proteins required for Rap1p localization in yeast nuclei. Proc. Natl. Acad. Sci. USA *92*, 5558–5562.

Kress, T.L., Krogan, N.J., and Guthrie, C. (2008). A Single SR-like Protein, Npl3, Promotes Pre-mRNA Splicing in Budding Yeast. Mol. Cell *32*, 727–734.

Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P., et al. (2006). Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. Nature *440*, 637–643.

Kunz, J.B., Neu-Yilik, G., Hentze, M.W., Kulozik, A.E., and Gehring, N.H. (2006). Functions of hUpf3a and hUpf3b in nonsense-mediated mRNA decay and translation. RNA *12*, 1015–1022.

Kurosaki, T., and Maquat, L.E. (2013). Rules that govern UPF1 binding to mRNA 3' UTRs. Proc. Natl. Acad. Sci. USA *110*, 3357–3362.

Kurosaki, T., Li, W., Hoque, M., Popp, M.W.-L., Ermolenko, D.N., Tian, B., and Maquat, L.E. (2014). A post-translational regulatory switch on UPF1 controls targeted mRNA degradation. Genes Dev. *28*, 1900–1916.

Kurosaki, T., Popp, M.W., and Maquat, L.E. (2019). Quality and quantity control of gene expression by nonsense-mediated mRNA decay. Nat. Rev. Mol. Cell Biol. *20*, 406–420.

Łabno, A., Tomecki, R., and Dziembowski, A. (2016). Cytoplasmic RNA decay pathways - Enzymes and mechanisms. Biochim. Biophys. Acta *1863*, 3125–3147.

Lai, M.-C., Lin, R.-I., and Tarn, W.-Y. (2001). Transportin-SR2 mediates nuclear import of phosphorylated SR proteins. Proc. Natl. Acad. Sci. USA *98*, 10154–10159.

Lai, M.-C., and Tarn, W.-Y. (2004). Hypophosphorylated ASF/SF2 Binds TAP and Is Present in Messenger Ribonucleoproteins. J. Biol. Chem. *279*, 31745–31749.

Lardelli, R.M., Thompson, J.X., Yates, J.R., III, and Stevens, S.W. (2010). Release of SF3 from the intron branchpoint activates the first step of pre-mRNA splicing. RNA *16*, 516–528.

Lareau, L.F., and Brenner, S.E. (2015). Regulation of Splicing Factors by Alternative Splicing and NMD is Conserved between Kingdoms Yet Evolutionarily Flexible. Mol. Biol. Evol. *32*, 1072–1079.

Lasalde, C., Rivera, A.V., León, A.J., González-Feliciano, J.A., Estrella, L.A., Rodríguez-Cruz, E.N., Correa, M.E., Cajigas, I.J., Bracho, D.P., Vega, I.E., Wilkinson, M.F., and González, C.I. (2014). Identification and functional analysis of novel phosphorylation sites in the RNA surveillance protein Upf1. Nucleic Acids Res. *42*, 1916–1929.

Lavysh, D., and Neu-Yilik, G. (2020). UPF1-Mediated RNA Decay—Danse Macabre in a Cloud. Biomolecules *10*, 999.

Le Hir, H., Izaurralde, E., Maquat, L.E., and Moore, M.J. (2000). The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon–exon junctions. EMBO J. *19*, 6860–6969.

Lee, M.S., Henry, M., and Silver, P.A. (1996). A protein that shuttles between the nucleus and the cytoplasm is an important mediator of RNA export. Genes Dev. *10*, 1233–1246.

Lee, S.R., Pratt, G.A., Martinez, F.J., Yeo, G.W., and Lykke-Andersen, J. (2015). Target Discrimination in Nonsense-Mediated mRNA Decay Requires Upf1 ATPase Activity. Mol. Cell *59*, 413–425.

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*, 2165–2177.

Lei, E.P., Krebber, H., and Silver, P.A. (2001). Messenger RNAs are recruited for nuclear export during transcription. Genes Dev.*15*, 1771–1782.

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*, 3536–3545.

Lennox, E.S. (1955). Transduction of Linked Genetic Characters of the Host by Bacteriophage P1. Virology *1*, 190–206.

Leon, K., and Ott, M. (2021). An 'Arms Race' between the Nonsense-mediated mRNA Decay Pathway and Viral Infections. Semin. Cell Dev. Biol. *111*, 101–107.

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. USA *100*, 189–192.

Lin, J.-J., and Zakian, V.A. (1994). Isolation and characterization of two *Saccharomyces cerevisiae* genes that encode proteins that bind to $(TG_{1-3})_n$ single strand telomeric DNA *in vitro*. Nucleic Acids Res. *22*, 4906–4913.

Lin, S., Xiao, R., Sun, P., Xu, X., and Fu, X.-D. (2005). Dephosphorylation-Dependent Sorting of SR Splicing Factors during mRNP Maturation. Mol. Cell *20*, 413–425.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta C_T}$ Method. Methods *25*, 402–408.

Loh, B., Jonas, S., and Izaurralde, E. (2013). The SMG5–SMG7 heterodimer directly recruits the CCR4–NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. Genes Dev. *27*, 2125–2138.

Lopez, P.J., and Séraphin, B. (1999). Genomic-scale quantitative analysis of yeast pre-mRNA splicing: Implications for splice-site recognition. RNA *5*, 1135–1137.

López-Perrote, A., Castaño, R., Melero, R., Zamarro, T., Kurosawa, H., Ohnishi, T., Uchiyama, A., Aoyagi, K., Buchwald, G., Kataoka, N., Yamashita, A., and Llorca, O. (2016). Human nonsense-mediated mRNA decay factor UPF2 interacts directly with eRF3 and the SURF complex. Nucleic Acids Res. *44*, 1909–1923.

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*, 7688–7697.

Lund, M.K., and Guthrie, C. (2005). The DEAD-Box Protein Dbp5p Is Required to Dissociate Mex67p from Exported mRNPs at the Nuclear Rim. Mol. Cell *20*, 645–651.

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*, 1121–1131.

Maderazo, A.B., Belk, J.P., He, F., and Jacobson, A. (2003). Nonsense-Containing mRNAs That Accumulate in the Absence of a Functional Nonsense-Mediated mRNA Decay Pathway Are Destabilized Rapidly upon Its Restitution. Mol. Cell. Biol. *23*, 842–851.

Malabat, C., Feuerbach, F., Ma, L., Saveanu, C., and Jacquier, A. (2015). Quality control of transcription start site selection by nonsense-mediated-mRNA decay. eLife *4*, e06722.

Maquat, L.E., Hwang, J., Sato, H., and Tang, Y. (2010a). CBP80-Promoted mRNP Rearrangements during the Pioneer Round of Translation, Nonsense-Mediated mRNA Decay, and Thereafter. Cold Spring Harb. Symp. Quant. Biol. *75*, 127–134.

Maquat, L.E., Tarn, W.-Y., and Isken, O. (2010b). The Pioneer Round of Translation: Features and Functions. Cell *142*, 368–374.

Martínez-Lumbreras, S., Taverniti, V., Zorrilla, S., Séraphin, B., and Pérez-Cañadillas, J.M. (2016). Gbp2 interacts with THO/TREX through a novel type of RRM domain. Nucleic Acids Res. *44*, 437–448.

Masuda, S., Das, R., Cheng, H., Hurt, E., Dorman, N., and Reed, R. (2005). Recruitment of the human TREX complex to mRNA during splicing. Genes Dev. *19*, 1512–1517.

May, J.P., and Simon, A.E. (2021). Targeting of viral RNAs by Upf1-mediated RNA decay pathways. Curr. Opin. Virol. *47*, 1–8.

McIlwain, D.R., Pan, Q., Reilly, P.T., Elia, A.J., McCracken, S., Wakeham, A.C., Itie-Youten, A., Blencowe, B.J., and Mak, T.W. (2010). Smg1 is required for embryogenesis and regulates diverse genes via alternative splicing coupled to nonsense-mediated mRNA decay. Proc. Natl. Acad. Sci. USA. *107*, 12186–12191.

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*, 134–140.

Meinel, D.M., Burkert-Kautzsch, C., Kieser, A., O'Duibhir, E., Siebert, M., Mayer, A., Cramer, P., Söding, J., Holstege, F.C.P., and Sträßer, K. (2013). Recruitment of TREX to the Transcription Machinery by Its Direct Binding to the Phospho-CTD of RNA Polymerase II. PLoS Genet. *9*, e1003914.

Meinel, D.M., and Sträßer, K. (2015). Co-transcriptional mRNP formation is coordinated within a molecular mRNP packaging station in *S. cerevisiae*. BioEssays *37*, 666–677.

Merrick, W.C., and Pavitt, G.D. (2018). Protein Synthesis Initiation in Eukaryotic Cells. Cold Spring Harb. Perspect. Biol. *10*, a033092.

Miller, K.E., Kim, Y., Huh, W.-K., and Park, H.-O. (2015). Bimolecular Fluorescence Complementation (BiFC) Analysis: Advances and Recent Applications for Genome-Wide Interaction Studies. J. Mol. Biol. *427*, 2039–2055.

Miller, D., Brandt, N., and Gresham, D. (2018). Systematic identification of factors mediating accelerated mRNA degradation in response to changes in environmental nitrogen. PLoS Genet. *14*, e1007406.

Min, E.E., Roy, B., Amrani, N., He, F., and Jacobson, A. (2013). Yeast Upf1 CH domain interacts with Rps26 of the 40S ribosomal subunit. RNA *19*, 1105–1115.

Mitchell, P., and Tollervey, D. (2003). An NMD Pathway in Yeast Involving Accelerated Deadenylation and Exosome-Mediated 3'→5' Degradation. Mol. Cell *11*, 1405–1413.

Moraes, K.C.M. (2010). RNA surveillance: Molecular Approaches in Transcript Quality Control and their Implications in Clinical Diseases. Mol. Med. *16*, 53–68.

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*, 2932–2939.

Morrissey, J.P., Deardorff, J.A., Hebron, C., and Sachs, A.B. (1999). Decapping of Stabilized, Polyadenylated mRNA in Yeast *pab1* Mutants. Yeast *15*, 687–702.

Mourier, T., and Jeffares, D.C. (2003). Eukaryotic Intron Loss. Science *300*, 1393.

Muhlrad, D., and Parker, R. (1994). Premature translational termination triggers mRNA decapping. Nature *370*, 578–581.

Muhlrad, D., and Parker, R. (1999a). Aberrant mRNAs with extended 3' UTRs are substrates for rapid degradation by mRNA surveillance. RNA *5*, 1299–1307.

Muhlrad, D., and Parker, R. (1999b). Recognition of Yeast mRNAs as "Nonsense Containing" Leads to Both Inhibition of mRNA Translation and mRNA Degradation: Implications for the Control of mRNA Decapping. Mol. Biol. Cell *10*, 3971–3978.

Müller-McNicoll, M., Botti, V., de Jesus Domingues, A.M., Brandl, H., Schwich, O.D., Steiner, M.C., Curk, T., Poser, I., Zarnack, K., and Neugebauer, K.M. (2016). SR proteins are NXF1 adaptors that link alternative RNA processing to mRNA export. Genes Dev. *30*, 553–566.

Nagy, E., and Maquat, L.E. (1998). A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. Trends Biochem. Sci. *23*, 198–199.

Nasif, S., Contu, L., and Mühlemann, O. (2018). Beyond quality control: The role of nonsense-mediated mRNA decay (NMD) in regulating gene expression. Semin. Cell Dev. Biol. *75*, 78–87.

Neu-Yilik, G., Amthor, B., Gehring, N.H., Bahri, S., Paidassi, H., Hentze, M.W., and Kulozik, A.E. (2011). Mechanism of escape from nonsense-mediated mRNA decay of human β-globin transcripts with nonsense mutations in the first exon. RNA *17*, 843–854.

Neuvéglise, C., Marck, C., and Gaillardin, C. (2011). The intronome of budding yeasts. C. R. Biol. *334*, 662–670.

Ni, J.Z., Grate, L., Donohue, J.P., Preston, C., Nobida, N., O'Brien, G., Shiue, L., Clark, T.A., Blume, J.E., and Ares, M., Jr. (2007). Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. Genes Dev. *21*, 708–718.

Nickless, A., Bailis, J.M., and You, Z. (2017). Control of gene expression through the nonsense-mediated RNA decay pathway. Cell Biosci. *7*, 26.

Nilsen T.W. (2014). Preparation of Cross-Linked Cellular Extracts with Formaldehyde. Cold Spring Harb. Protoc. *2014*, 1001–1003.

Nissan, T., Rajyaguru, P., She, M., Song, H., and Parker, R. (2010). Decapping Activators in *Saccharomyces cerevisiae* Act by Multiple Mechanisms. Mol. Cell *39*, 773–783.

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*, 1187–1200.

Okada-Katsuhata, Y., Yamashita, A., Kutsuzawa, K., Izumi, N., Hirahara, F., and Ohno, S. (2012). N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. Nucleic Acids Res. *40*, 1251–1266.

Paci, G., Caria, J., and Lemke, E.A. (2021). Cargo transport through the nuclear pore complex at a glance. J. Cell Sci. *134*, jcs247874.

Pang, T.-L., Wang, C.-Y., Hsu, C.-L., Chen, M.-Y., and Lin, J.-J. (2003). Exposure of Single-stranded Telomeric DNA Causes $G_2$/M Cell Cycle Arrest in *Saccharomyces cerevisiae*. J. Biol. Chem. *278*, 9318–9321.

Park, J.W., and Graveley, B.R. (2007). Complex Alternative Splicing. Adv. Exp. Med. Biol. *623*, 50–63.

Pawlicka, K., Kalathiya, U., and Alfaro, J. (2020). Nonsense-Mediated mRNA Decay: Pathologies and the Potential for Novel Therapeutics. Cancers *12*, 765.

Peixeiro, I., Inácio, Â., Barbosa, C., Silva, A.L., Liebhaber, S.A., and Romão, L. (2012). Interaction of PABPC1 with the translation initiation complex is critical to the NMD resistance of AUG-proximal nonsense mutations. Nucleic Acids Res. *40*, 1160–1173.

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*, 1737–1754.

Pisarev, A.V., Skabkin, M.A., Pisareva, V.P., Skabkina, O.V., Rakotondrafara, A.M., Hentze, M.W., Hellen, C.U.T, Pestova, T.V. (2010). The Role of ABCE1 in Eukaryotic Posttermination Ribosomal Recycling. Mol. Cell *37*, 196–210.

Pisareva, V.P., Skabkin, M.A., Hellen, C.U.T., Pestova, T.V., and Pisarev, A.V. (2011). Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. EMBO J. *30*, 1804–1817.

Poornima, G., Srivastava, G., Roy, B., Kuttanda, I.A., Kurbah, I., and Rajyaguru, P.I. (2021). RGG-motif containing mRNA export factor Gbp2 acts as a translation repressor. RNA Biol. 1–12. Advance online publication.

Powers, K.T., Szeto, J.-Y.A., and Schaffitzel, C. (2020). New insights into no-go, non-stop and nonsense-mediated mRNA decay complexes. Curr. Opin. Struct. Biol. *65*, 110–118.

Proudfoot, N.J., Furger, A., and Dye, M.J. (2002). Integrating mRNA Processing with Transcription. Cell *108*, 501–512.

Proudfoot, N.J. (2011). Ending the message: poly(A) signals then and now. Genes Dev. *25*, 1770–1782.

Rahman, M.A., Lin, K.-T., Bradley, R.K., Abdel-Wahab, O., and Krainer, A.R. (2020). Recurrent SRSF2 mutations in MDS affect both splicing and NMD. Genes Dev. *34*, 413–427.

Rajanala, K., and Nandicoori, V.K. (2012). Localization of Nucleoporin Tpr to the Nuclear Pore Complex Is Essential for Tpr Mediated Regulation of the Export of Unspliced RNA. PLoS ONE *7*, e29921.

Rajyaguru, P., She, M., and Parker, R. (2012). Scd6 Targets eIF4G to Repress Translation: RGG Motif Proteins as a Class of eIF4G-Binding Proteins. Mol. Cell *45*, 244–254.

Ramanathan, M., Porter, D.F., and Khavari, P.A. (2019). Methods to study RNA–protein interactions. Nat. Methods *16*, 225–234.

Richardson, R., Denis, C.L., Zhang, C., Nielsen, M.E.O., Chiang, Y.-C., Kierkegaard, M., Wang, X., Lee, D.J., Andersen, J.S., and Yao, G. (2012). Mass spectrometric identification of proteins that interact through specific domains of the poly(A) binding protein. Mol. Genet. Genomics *287*, 711–730.

Romei, M.G., and Boxer, S.G. (2019). Split Green Fluorescent Proteins: Scope, Limitations, and Outlook. Annu. Rev. Biophys. *48*, 19–44.

Rondón, A.G., Mischo, H.E., Kawauchi, J., and Proudfoot, N.J. (2009). Fail-Safe Transcriptional Termination for Protein-Coding Genes in *S. cerevisiae*. Mol. Cell *36*, 88–98.

Roque, S., Cerciat, M., Gaugué, I., Mora, L., Floch, A.G., de Zamaroczy, M., Heurgué-Hamard, V., and Kervestin, S. (2015). Interaction between the poly(A)-binding protein Pab1 and the eukaryotic release factor eRF3 regulates translation termination but not mRNA decay in *Saccharomyces cerevisiae*. RNA *21*, 124–134.

Rose, M.D., Winston, F., and Hieter, P. (1990). Methods in Yeast Genetics, A Laboratory Course Manual. (New York: Cold Spring Harbor Laboratory Press).

Roth, K.M., Wolf, M.K., Rossi, M., and Butler, J.S. (2005). The Nuclear Exosome Contributes to Autogenous Control of *NAB2* mRNA Levels. Mol. Cell. Biol. *25*, 1577–1585.

Roth, K.M., Byam, J., Fang, F., and Butler, J.S. (2009). Regulation of *NAB2* mRNA 3'-end formation requires the core exosome and the Trf4p component of the TRAMP complex. RNA *15*, 1045–1058.

Roy, D., and Rajyaguru, P.I. (2018). Suppressor of clathrin deficiency (Scd6)—An emerging RGG-motif translation repressor. Wiley Interdiscip. Rev. RNA *9*, e1479.

Rufener, S.C., and Mühlemann, O. (2013). eIF4E-bound mRNPs are substrates for nonsense-mediated mRNA decay in mammalian cells. Nat. Struct. Mol. Biol. *20*, 710–717.

Ruiz-Echevarría, M.J., González, C.I., and Peltz, S.W. (1998). Identifying the right stop: determining how the surveillance complex recognizes and degrades an aberrant mRNA. EMBO J. *17*, 575–589.

Saguez, C., Olesen, J.R., and Jensen, T.H. (2005). Formation of export-competent mRNP: escaping nuclear destruction. Curr. Opin. Cell Biol. *17*, 287–293.

Sanford, J.R., Gray, N.K., Beckmann, K., and Cáceres, J.F. (2004). A novel role for shuttling SR proteins in mRNA translation. Genes Dev. *18*, 755–768.

Sato, H., and Maquat, L.E. (2009). Remodeling of the pioneer translation initiation complex involves translation and the karyopherin importin β. Genes Dev. *23*, 2537–2550.

Saudemont, B., Popa, A., Parmley, J.L., Rocher, V., Blugeon, C., Necsulea, A., Meyer, E., and Duret, L. (2017). The fitness cost of mis-splicing is the main determinant of alternative splicing patterns. Genome Biol. *18*, 208.

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*, 360–370.

Sayani, S., and Chanfreau, G.F. (2012). Sequential RNA degradation pathways provide a fail-safe mechanism to limit the accumulation of unspliced transcripts in *Saccharomyces cerevisiae*. RNA *18*, 1563–1572.

Schell, T., Köcher, T., Wilm, M., Séraphin, 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*, 775–783.

Schlautmann, L.P., and Gehring, N.H. (2020). A Day in the Life of the Exon Junction Complex. Biomolecules *10*, 866.

Schmid, M., and Jensen, T.H. (2010). Nuclear quality control of RNA polymerase II transcripts. Wiley Interdiscip. Rev. RNA *1*, 474–485.

Schmid, M., Poulsen, M.B., Olszewski, P., Pelechano, V., Saguez, C., Gupta, I., Steinmetz, L.M., Moore, C., and Jensen, T.H. (2012). Rrp6p Controls mRNA Poly(A) Tail Length and Its Decoration with Poly(A) Binding Proteins. Mol. Cell *47*, 267–280.

Schmid, M., Olszewski, P., Pelechano, V., Gupta, I., Steinmetz, L.M., and Jensen, T.H. (2015). The Nuclear PolyA-Binding Protein Nab2p Is Essential for mRNA Production. Cell Rep. *12*, 128–139.

Schmid, M., and Jensen, T.H. (2018). Controlling nuclear RNA levels. Nat. Rev. Genet. *19*, 518–529.

Schmid, M., and Jensen, T.H. (2019). The Nuclear RNA Exosome and Its Cofactors. Adv. Exp. Med. Biol. *1203*, 113–132.

Schmitt, C., von Kobbe, C., Bachi, A., Panté, N., Rodrigues, J.P., Boscheron, C., Rigaut, G., Wilm, M., Séraphin, B., Carmo-Fonseca, M., and Izaurralde, E. (1999). Dbp5, a DEAD-box protein required for mRNA export, is recruited to the cytoplasmic fibrils of nuclear pore complex via a conserved interaction with CAN/Nup159p. EMBO J. *18*, 4332–4347.

Schulz, D., Schwalb, B., Kiesel, A., Baejen, C., Torkler, P., Gagneur, J., Soeding, J., and Cramer, P. (2013). Transcriptome Surveillance by Selective Termination of Noncoding RNA Synthesis. Cell *155*, 1075–1087.

Segal, S.P., Dunckley, T., and Parker, R. (2006). Sbp1p Affects Translational Repression and Decapping in *Saccharomyces cerevisiae*. Mol. Cell. Biol. *26*, 5120–5130.

Sei, E., and Conrad, N.K. (2014). UV Cross-Linking of Interacting RNA and Protein in Cultured Cells. Methods Enzymol. *539*, 53–66.

Serdar, L.D., Whiteside, D.L., and Baker, K.E. (2016). ATP hydrolysis by UPF1 is required for efficient translation termination at premature stop codons. Nat. Commun. *7,* 14021.

Serdar, L.D., Whiteside, D.L., Nock, S.L., McGrath, D., and Baker, K.E. (2020). Inhibition of post-termination ribosome recycling at premature termination codons in UPF1 ATPase mutants. eLife *9*, e57834.

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*, 209–223.

Shen, E.C., Henry, M.F., Weiss, V.H., Valentini, S.R., Silver, P.A., and Lee, M.S. (1998). Arginine methylation facilitates the nuclear export of hnRNP proteins. Genes Dev. *12*, 679–691.

Shen, E.C., Stage-Zimmermann, T., Chui, P., and Silver, P.A. (2000). 7The Yeast mRNA-binding Protein Npl3p Interacts with the Cap-binding Complex. J. Biol. Chem. *275*, 23718–23724.

Sherman, F., and Hicks, J. (1991). Micromanipulation and Dissection of Asci. Methods Enzymol. *194*, 21–37.

Sherman, F. (2002). Getting Started with Yeast. Methods Enzymol. *350,* 3–41.

Shigeoka, T., Kato, S., Kawaichi, M., and Ishida, Y. (2012). Evidence that the Upf1-related molecular motor scans the 3'-UTR to ensure mRNA integrity. Nucleic Acids Res. *40*, 6887–6897.

Shirley, R.L., Lelivelt, M.J., Schenkman, L.R., Dahlseid, J.N., and Culbertson, M.R. (1998). A factor required for nonsense-mediated mRNA decay in yeast is exported from the nucleus to the cytoplasm by a nuclear export signal sequence. J. Cell Sci. *111*, 3129–3143.

Shirley, R.L., Ford, A.S., Richards, M.R., Albertini, M., and Culbertson, M.R. (2002). Nuclear Import of Upf3p Is Mediated by Importin-α/-β and Export to the Cytoplasm Is Required for a Functional Nonsense-Mediated mRNA Decay Pathway in Yeast. Genetics *161*, 1465–1482.

Shoemaker, C.J., Eyler, D.E., and Green, R. (2010). Dom34:Hbs1 Promotes Subunit Dissociation and Peptidyl-tRNA Drop-Off to Initiate No-Go Decay. Science *330*, 369–372.

Shoemaker, C.J., and Green, R. (2011). Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. Proc. Natl. Acad. Sci. USA *108*, E1392–1398.

Shum, E.Y., Jones, S.H., Shao, A., Dumdie, J., Krause, M.D., Chan, W.-K., Lou, C.-H., Espinoza, J.L., Song, H.-W., Phan, M.H., Ramaiah, M., Huang, L., McCarrey, J.R., Peterson, K.J., De Rooij, D.G., Cook-Andersen, H., and Wilkinson, M.F. (2016). The Antagonistic Gene Paralogs *Upf3a* and *Upf3b* Govern Nonsense-Mediated RNA Decay. Cell *165*, 382–395.

Siebel, C.W., Feng, L., Guthrie, C., and Fu, X.-D. (1999). Conservation in budding yeast of a kinase specific for SR splicing factors. Proc. Natl. Acad. Sci. USA *96*, 5440–5445.

Sikorski, R.S., and Hieter, P. (1989). A System of Shuttle Vectors and Yeast Host Strains Designed for Efficient Manipulation of DNA in *Saccharomyces cerevisiae*. Genetics *122*, 19–27.

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

Simms, C.L., Thomas, E.N., and Zaher, H.S. (2017). Ribosome-based quality control of mRNA and nascent peptides. Wiley Interdiscip. Rev. RNA *8*, e1366.

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.

Singh, G., Kucukural, A., Cenik, C., Leszyk, J.D., Shaffer, S.A., Weng, Z., and Moore, M.J. (2012). The Cellular EJC Interactome Reveals Higher-Order mRNP Structure and an EJC-SR Protein Nexus. Cell *151*, 750–764.

Singh, G., Pratt, G., Yeo, G.W., and Moore, M.J. (2015). The Clothes Make the mRNA: Past and Present Trends in mRNP Fashion. Annu. Rev. Biochem. *84*, 325–354.

Skružný, M., Schneider, C., Rácz, A., Weng, J., Tollervey, D., and Hurt, E. (2009). An Endoribonuclease Functionally Linked to Perinuclear mRNP Quality Control Associates with the Nuclear Pore Complexes. PLoS Biol. *7*, e1000008.

Smith, J.E., and Baker, K.E. (2015). Nonsense-mediated RNA decay – a switch and dial for regulating gene expression. BioEssays *37*, 612–623.

Soheilypour, M., and Mofrad, M.R.K. (2018). Quality control of mRNAs at the entry of the nuclear pore: Cooperation in a complex molecular system. Nucleus *9*, 202–211.

Song, H., Mugnier, P., Das, A.K., Webb, H.M., Evans, D.R., Tuite, M.F., Hemmings, B.A., and Barford, D. (2000). The Crystal Structure of Human Eukaryotic Release Factor eRF1 – Mechanism of Stop Codon Recognition and Peptidyl-tRNA Hydrolysis. Cell *100*, 311–321.

Soucek, S., Corbett, A.H., and Fasken, M.B. (2012). The long and the short of it: The role of the zinc finger polyadenosine RNA binding protein, Nab2, in control of poly(A) tail length. Biochim. Biophys. Acta *1819*, 546–554.

Soucek, S., Zeng, Y., Bellur, D.L., Bergkessel, M., Morris, K.J., Deng, Q., Duong, D., Seyfried, N.T., Guthrie, C., Staley, J.P., Fasken, M.B., and Corbett, A.H. (2016). Evolutionarily Conserved Polyadenosine RNA Binding Protein Nab2 Cooperates with Splicing Machinery to Regulate the Fate of Pre-mRNA. Mol. Cell. Biol. *36*, 2697–2714.

Spencer V.A., and Davie J.R. (2002). Isolation of Proteins Cross-linked to DNA by Formaldehyde. In The Protein Protocols Handbook, J.M. Walker, ed. (Totowa, New Jersey: Humana Press), pp. 753–757.

Spingola, M., Grate, L., Haussler, D., and Ares, M., Jr. (1999). Genome-wide bioinformatic and molecular analysis of introns in *Saccharomyces cerevisiae*. RNA *5*, 221–234.

Srinivasa, S., Ding, X., and Kast, J. (2015). Formaldehyde cross-linking and structural proteomics: Bridging the gap. Methods *89*, 91–98.

Stewart, M. (2010). Nuclear export of mRNA. Trends Biochem. Sci. *35*, 609–617.

Swartz, J.E., Bor, Y.-C., Misawa, Y., Rekosh, D., and Hammarskjold, M.L. (2007). The Shuttling SR Protein 9G8 Plays a Role in Translation of Unspliced mRNA Containing a Constitutive Transport Element. J. Biol. Chem. *282*, 19844–19853.

Swisher, K.D., and Parker, R. (2011). Interactions between Upf1 and the Decapping Factors Edc3 and Pat1 in *Saccharomyces cerevisiae*. PLoS ONE *6*, e26547.

Tani, H., Imamachi, N., Salam, K.A., Mizutani, R., Ijiri, K., Irie, T., Yada, T., Suzuki, Y., and Akimitsu, N. (2012). Identification of hundreds of novel UPF1 target transcripts by direct determination of whole transcriptome stability. RNA Biol. *9*, 1370–1379.

Tani, H., Torimura, M., and Akimitsu, N. (2013). The RNA Degradation Pathway Regulates the Function of GAS5 a Non-Coding RNA in Mammalian Cells. PLoS ONE *8*, e55684.

Tarassov, K., Messier, V., Landry, C.R., Radinovic, S., Serna Molina, M.M., Shames, I., Malitskaya, Y., Vogel, J., Bussey, H., and Michnick, S.W. (2008). An in Vivo Map of the Yeast Protein Interactome. Science *320*, 1465–1470.

Tardiff, D.F., Abruzzi, K.C., and Rosbash, M. (2007). Protein characterization of *Saccharomyces cerevisiae* RNA polymerase II after *in vivo* cross-linking. Proc. Natl. Acad. Sci. USA *104*, 19948–19953.

Tian, M., Yang, W., Zhang, J., Dang, H., Lu, X., Fu, C., and Miao, W. (2017). Nonsense-mediated mRNA decay in *Tetrahymena* is EJC independent and requires a protozoa-specific nuclease. Nucleic Acids Res. *45*, 6848–6863.

Tieg, B., and Krebber, H. (2013). Dbp5 – From nuclear export to translation. Biochim. Biophys. Acta *1829*, 791–798.

Toma, K.G., Rebbapragada, I., Durand, S., and Lykke-Andersen, J. (2015). Identification of elements in human long 3' UTRs that inhibit nonsense-mediated decay. RNA *21*, 887–897.

Tomecki, R., Kristiansen, M.S., Lykke-Andersen, S., Chlebowski, A., Larsen, K.M., Szczesny, R.J., Drazkowska, K., Pastula, A., Andersen, J.S., Stepien, P.P., Dziembowski, A., and Jensen, T.H. (2010). The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. EMBO J. *29*, 2342–2357.

Towbin, H., Staehelin, T., and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. Proc. Natl. Acad. Sci. USA *76*, 4350–4354.

Tran, E.J., Zhou, Y., Corbett, A.H., and Wente, S.R. (2007). The DEAD-Box Protein Dbp5 Controls mRNA Export by Triggering Specific RNA:Protein Remodeling Events. Mol. Cell *28*, 850–859.

Trcek, T., Sato, H., Singer, R.H., and Maquat, L.E. (2013). Temporal and spatial characterization of nonsense-mediated mRNA decay. Genes Dev. *27*, 541–551.

Tuck, A.C., and Tollervey, D. (2013). A Transcriptome-wide Atlas of RNP Composition Reveals Diverse Classes of mRNAs and lncRNAs. Cell *154*, 996–1009.

Tucker, M., and Parker, R. (2000). Mechanisms and Control of mRNA Decapping in *Saccharomyces cerevisiae*. Annu. Rev. Biochem. *69*, 571–595.

Tutucci, E., and Stutz, F. (2011). Keeping mRNPs in check during assembly and nuclear export. Nat. Rev. Mol. Cell Biol. *12*, 377–384.

Uchida, N., Hoshino, S., Imataka, H., Sonenberg, N., and Katada, T. (2002). A Novel Role of the Mammalian GSPT/eRF3 Associating with Poly(A)-binding Protein in Cap/Poly(A)-dependent Translation. J. Biol. Chem. *277*, 50286–50292.

van Hoof, A., Frischmeyer, P.A., Dietz, H.C., and Parker, R. (2002). Exosome-Mediated Recognition and Degradation of mRNAs Lacking a Termination Codon. Science *295*, 2262–2264.

van Hoof, A., and Wagner, E.J. (2011). A brief survey of mRNA surveillance. Trends Biochem. Sci. *36*, 585–592.

Vinciguerra, P., Iglesias, N., Camblong, J., Zenklusen, D., and Stutz, F. (2005). Perinuclear Mlp proteins downregulate gene expression in response to a defect in mRNA export. EMBO J. *24*, 813–823.

Wada, M., and Ito, K. (2014). A genetic approach for analyzing the co-operative function of the tRNA mimicry complex, eRF1/eRF3, in translation termination on the ribosome. Nucleic Acids Res. *42*, 7851–7866.

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*, 880–890.

Wang, W., Cajigas, I.J., Peltz, S.W., Wilkinson, M.F., and González, C.I. (2006). Role for Upf2p Phosphorylation in *Saccharomyces cerevisiae* Nonsense-Mediated mRNA Decay. Mol. Cell. Biol. *26*, 3390–3400.

Warkocki, Z., Odenwälder, P., Schmitzová, J., Platzmann, F., Stark, H., Urlaub, H., Ficner, R., Fabrizio, P., and Lührmann, R. (2009). Reconstitution of both steps of *Saccharomyces*

*cerevisiae* splicing with purified spliceosomal components. Nat. Struct. Mol. Biol. *16*, 1237–1243.

Wegener, M., and Müller-McNicoll, M. (2018). Nuclear retention of mRNAs – quality control, gene regulation and human disease. Semin. Cell Dev. Biol. *79*, 131–142.

Wegener, M., and Müller-McNicoll, M. (2019). View from an mRNP: The Roles of SR Proteins in Assembly, Maturation and Turnover. Adv. Exp. Med. Biol. *1203*, 83–112.

Weirich, C.S., Erzberger, J.P., Berger, J.M., and Weis, K. (2004). The N-terminal Domain of Nup159 Forms a β-Propeller that Functions in mRNA Export by Tethering the Helicase Dbp5 to the Nuclear Pore. Mol. Cell *16*, 749–760.

Weirich, C.S., Erzberger, J.P., Flick, J.S., Berger, J.M., Thorner, J., and Weis, K. (2006). Activation of the DExD/H-box protein Dbp5 by the nuclear-pore protein Gle1 and its coactivator InsP$_6$ is required for mRNA export. Nat. Cell Biol. *8*, 668–676.

Wen, J., and Brogna, S. (2010). Splicing-dependent NMD does not require the EJC in *Schizosaccharomyces pombe*. EMBO J. *29*, 1537–1551.

Wen, J., He, M., Petric, M., Marzi, L., Wang, J., Piechocki, K., McLeod, T., Singh, A.K., Dwivedi, V., and Brogna, S. (2020). An intron proximal to a PTC enhances NMD in *Saccharomyces cerevisiae*. bioRxiv. https://doi.org/10.1101/149245.

Wende, W., Friedhoff, P., and Sträßer, K. (2019). Mechanism and Regulation of Co-transcriptional mRNP Assembly and Nuclear mRNA Export. Adv. Exp. Med. Biol. *1203*, 1–31.

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*, 5477–5490.

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*, 5491–5506.

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*, 205–214.

Wery, M., Descrimes, M., Vogt, N., Dallongeville, A.-S., Gautheret, D., and Morillon, A. (2016). Nonsense-Mediated Decay Restricts LncRNA Levels in Yeast Unless Blocked by Double-Stranded RNA Structure. Mol. Cell *61*, 379–392.

Will, C.L., and Lührmann, R. (2011). Spliceosome Structure and Function. Cold Spring Harb. Perspect. Biol. *3*, a003707.

Windgassen, M., and Krebber, H. (2003). Identification of Gbp2 as a novel poly(A)[+] RNA-binding protein involved in the cytoplasmic delivery of messenger RNAs in yeast. EMBO Rep. *4*, 278–283.

Windgassen, M., Sturm, D., Cajigas, I.J., González, C.I., Seedorf, M., Bastians, H., and Krebber, H. (2004). Yeast Shuttling SR Proteins Npl3p, Gbp2p, and Hrb1p Are Part of the Translating mRNPs, and Npl3p Can Function as a Translational Repressor. Mol. Cell. Biol. *24*, 10479–10491.

Wolin, S.L., and Maquat, L.E. (2019). Cellular RNA surveillance in health and disease. Science *366*, 822–827.

Woodward, L.A., Mabin, J.W., Gangras, P., and Singh, G. (2017). The exon junction complex: a lifelong guardian of mRNA fate. Wiley Interdiscip. Rev. RNA *8*, e1411.

Wu, C., Roy, B., He, F., Yan, K., and Jacobson, A. (2020). Poly(A)-Binding Protein Regulates the Efficiency of Translation Termination. Cell Rep. *33*, 108399.

Xie, Y., and Ren, Y. (2019). Mechanisms of nuclear mRNA export: A structural perspective. Traffic *20*, 829–840.

Xie, Y., Clarke, B.P., Kim, Y.J., Ivey, A.L., Hill, P.S., Shi, Y., and Ren, Y. (2021a). Cryo-EM structure of the yeast TREX complex and coordination with the SR-like protein Gbp2. eLife *10*, e65699.

Xie, Y., Lord, C.L., Clarke, B.P., Ivey, A.L., Hill, P.S., McDonald, W.H., Wente, S.R., and Ren, Y. (2021b). Structure and activation mechanism of the yeast RNA Pol II CTD kinase CTDK-1 complex. Proc. Natl. Acad. Sci. USA *118*, e2019163118.

Yamashita, A., Izumi, N., Kashima, I., Ohnishi, T., Saari, B., Katsuhata, Y., Muramatsu, R., Morita, T., Iwamatsu, A., Hachiya, T., Kurata, R., Hirano, H., Anderson, P., and Ohno, S. (2009). SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay. Genes Dev. *23*, 1091–1105.

Yeh, F.-L., Chang, S.-L., Ahmed, G.R., Liu, H.-I., Tung, L., Yeh, C.-S., Lanier, L.S., Maeder, C., Lin, C.-M., Tsai, S.-C., Hsiao, W.-Y., Chang, W.-H., and Chang, T.-H. (2021). Activation of Prp28 ATPase by phosphorylated Npl3 at a critical step of spliceosome remodeling. Nat. Commun. *12*, 3082.

York, J.D., Odom, A.R., Murphy, R., Ives, E.B., and Wente, S.R. (1999). A Phospholipase C-Dependent Inositol Polyphosphate Kinase Pathway Required for Efficient Messenger RNA Export. Science *285*, 96–100.

Yun, C.Y., and Fu, X.-D. (2000). Conserved SR Protein Kinase Functions in Nuclear Import and Its Action Is Counteracted by Arginine Methylation in *Saccharomyces cerevisiae*. J. Cell Biol. *150*, 707–717.

Zahdeh, F., and Carmel, L. (2016). The role of nucleotide composition in premature termination codon recognition. BMC Bioinformatics *17*, 519.

Zander, G., Hackmann, A., Bender, L., Becker, D., Lingner, T., Salinas, G., and Krebber, H. (2016). mRNA quality control is bypassed for immediate export of stress-responsive transcripts. Nature *540*, 593–596.

Zander, G., and Krebber, H. (2017). Quick or quality? How mRNA escapes nuclear quality control during stress. RNA Biol. *14*, 1642–1648.

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*, 2231–2244.

Zhang, J., and Maquat, L.E. (1997). Evidence that translation reinitiation abrogates nonsense-mediated mRNA decay in mammalian cells. EMBO J. *16,* 826–833.

Zhang, Z., and Krainer, A.R. (2004). Involvement of SR Proteins in mRNA Surveillance. Mol. Cell *16*, 597–607.

Zhang, X.-E., Cui, Z., and Wang, D. (2016). Sensing of biomolecular interactions using fluorescence complementing systems in living cells. Biosens. Bioelectron. *76*, 243–250.

Zinoviev, A., Ayupov, R.K., Abaeva, I.S., Hellen, C.U.T., and Pestova, T.V. (2020). Extraction of mRNA from Stalled Ribosomes by the Ski Complex. Mol. Cell *77*, 1340–1349.e6.

Zünd, D., Gruber, A.R., Zavolan, M., and Mühlemann, O. (2013). Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3' UTRs. Nat. Struct. Mol. Biol. *20,* 936–943.

# Acknowledgement

# Appendix

RESEARCH PAPER

# Nuclear SR-protein mediated mRNA quality control is continued in cytoplasmic nonsense-mediated decay

Sebastian Grosse*, Yen-Yun Lu ⓘ*, Ivo Coban, Bettina Neumann, and Heike Krebber

Abteilung Für Molekulare Genetik, Institut Für Mikrobiologie Und Genetik, Göttinger Zentrum Für Molekulare Biowissenschaften (GZMB), Georg-August Universität Göttingen, Göttingen, Germany

## ABSTRACT

One important task of eukaryotic cells is to translate only mRNAs that were correctly processed to prevent the production of truncated proteins, found in neurodegenerative diseases and cancer. Nuclear quality control of splicing requires the SR-like proteins Gbp2 and Hrb1 in *S. cerevisiae*, where they promote the degradation of faulty pre-mRNAs. Here we show that Gbp2 and Hrb1 also function in nonsense mediated decay (NMD) of spliced premature termination codon (PTC)-containing mRNAs. Our data support a model in which they are in a complex with the Upf-proteins and help to transmit the Upf1-mediated PTC recognition to the transcripts ends. Most importantly they appear to promote translation repression of spliced transcripts that contain a PTC and to finally facilitate degradation of the RNA, presumably by supporting the recruitment of the degradation factors. Therefore, they seem to control mRNA quality beyond the nuclear border and may thus be global surveillance factors. Identification of SR-proteins as general cellular surveillance factors in yeast will help to understand the complex human system in which many diseases with defects in SR-proteins or NMD are known, but the proteins were not yet recognized as general RNA surveillance factors.

## Introduction

All cellular processes depend on the correct and effective translation of mRNAs into proteins. Eukaryotic cells control the correctness of the mRNAs both in the nucleus and in the cytoplasm [1–5]. While the nuclear quality control recognizes rather structural defects of the transcripts, the cytoplasmic quality control detects incorrect open reading frames through the decoding capability of the ribosome. Structural defects include transcripts that have retained their introns or mRNAs that are uncapped or non-polyadenylated [4,5]. Such defects can create problems in translation, as both transcript ends are usually connected in the cytoplasm to allow a repeated cycling of the ribosomes on transcripts to increase translation efficiency [6].

Several proteins contribute to generating correctly matured mRNAs [4,7]. However, one group of proteins is particularly important for the nuclear mRNA surveillance in yeast, because their absence results in the leakage of faulty mRNAs into the cytoplasm [8]. These guard proteins include Npl3, Gbp2, Hrb1 and Nab2, the first three of which are highly homologous with human serine arginine (SR)-proteins [5]. Among them, Gbp2 and Hrb1 preferentially bind to transcripts that undergo splicing, as they interact with late splicing factors [9]. In case splicing does not occur correctly, Gbp2 and Hrb1 recruit the TRAMP-complex, which fetches the nuclear exosome to degrade the faulty transcript. On correctly

spliced mRNAs Gbp2 and Hrb1 interact with Mex67-Mtr2 (TAP-p15 in human) instead, promoting nuclear export [9]. Other guard proteins control different maturation steps, but they operate similarly in principle. Proper packaging of the RNA into a ribonucleoparticle (RNP) supports transit through the nuclear pore complex (NPC) [4,5]. At the NPC, the nuclear basket protein Mlp1 controls proper Mex67-coverage of the guard proteins on the mRNA [4,5]. Thus, also in the absence of Mlp1, faulty transcripts are not retained in the nucleus and leak into the cytoplasm [9,10].

Remarkably, while Mex67 is removed upon transport, the guard proteins remain bound on the transcript until translation [11]. This suggests that they may have additional functions in the cytoplasm and might continue their roles as quality control factors. In particular, because after dissociation of Mex67 they are free for new interactions. However, such a cytoplasmic quality control function for the guard proteins has not been explored to date. In contrast to the nuclear quality control system, the cytoplasmic quality control checks the encoded sequence. During translation, intact or faulty open reading frames can be distinguished. In this way, broken mRNAs that lack a stop codon or mRNAs with strong secondary structures that stall the ribosome are eliminated by the no-stop- (NSD) or no-go-decay (NGD), respectively [7]. Another severe defect is premature termination, often resulting from improper splicing and transcripts that escaped nuclear quality control [1,12]. PTC-containing transcripts

are recognized and eliminated by the nonsense-mediated decay (NMD). In its centre are the Upf-proteins. Upon interacting with eRF1 and eRF3 at the terminating ribosome, they signal the cell to inhibit translation and degrade affected mRNAs [1,3,13]. Upf1 is the central ATP-dependent RNA helicase and required for the recognition of PTC-containing mRNAs. Upf2 and Upf3 support NMD by formation of a Upf1-Upf2-Upf3 complex signalling degradation [1,3,13].

How NMD is initiated is still debated. In metazoans, Upf2 and Upf3 bind to the exon-junction complex (EJC) downstream of the PTC, and are thought to communicate somehow with the ribosome to trigger RNA degradation [1,3,13]. Interestingly, some human SR proteins, which are known to be important mediators of splicing, are also associated with the EJC [14]. However, their function in NMD is rather nebulous, not the least because of their upstream role in splicing, which is difficult to study independently from their potential subsequent function in NMD. Although EJCs are characteristic for multicellular organisms and have not been discovered in *S. cerevisiae*, a similar system was described in yeast, involving Hrp1. Hrp1 is a yeast RNA-binding protein that binds to a downstream sequence element (DSE), originally identified in the *PGK1* mRNA. Hrp1 was shown to interact with the Upf-proteins when the *PGK1* contained a PTC [15]. Another proposed trigger of NMD is a long 3′-untranslated region (UTR). A long distance between the terminating ribosome and the poly(A) tail impedes the interaction between the poly(A) binding protein Pab1 (PABP in human) and the terminating ribosome, which under regular conditions promotes efficient termination. In this model, multiple copies of Upf1 are distributed on the RNA in a loosely bound fashion and removed by the passing ribosome. However, when termination occurs prematurely, Upf1-binding is stabilized. As Pab1 is far away, formation of a stable Upf1-2-3 complex is promoted and NMD is elicited [1,16].

While the exact mechanism of NMD activation is relatively vague, our understanding of the downstream events is even less clear. It seems evident that PTC-containing mRNAs are translationally repressed and mainly degraded from the 5′ ends in yeast [17]. Yet, how this is mediated is currently unclear. It is known that the ATPase activity of Upf1 is required to disassemble the ribosome and allow complete degradation of the transcript [18,19]. It is also known that Upf1 can be found in a complex with decay enzymes [20]. Dcp1 and Dcp2 are required for decapping of NMD targets and Xrn1 for the subsequent 5′ to 3′ exonucleolytic RNA decay. However, how the degradation factors are recruited from the PTC to the distant ends of the transcript is unclear. Although the 5′-end mediated degradation pathway is mainly used, decay can also occur from the 3′-end via the exosome and its cytoplasmic co-factor complex, containing Ski2 [1,21]. In metazoans degradation of the recognized NMD targets is supported by additional factors, such as SMG6, which cleaves NMD-transcripts endonucleolytically, and SMG5-SMG7, which bridge interactions between Upf1 and degrading enzymes at the PTC [1,22]. In yeast, similar auxiliary factors are poorly understood, and likely more factors participate in NMD than currently known.

Our study presented here suggests that the nuclear and the cytoplasmic mRNA quality control systems may be coupled.

We found the nuclear guard proteins Gbp2 and Hrb1 to be players in NMD. They continue their guarding function, originally discovered in the nucleus, further in the cytoplasm. For a subset of targets, those that contained intron sequences, these SR-like proteins appear to help repress translation upon detection of a PTC and to recruit the cytoplasmic degradation machineries to the faulty transcript. Importantly, by bridging the Upf1-bound PTC to the 5′ end of the mRNA Gbp2 and Hrb1 may help to transmit the signal of the mRNA defect directly to the starting point of translational repression and degradation.

## Materials and methods

*Saccharomyces cerevisiae* strains used in this study are listed in Table S1, plasmids in Table S2 and oligonucleotides in Table S3. Yeast strains and plasmids were generated by conventional methods. Yeast strains were cultivated in standard media at 25°C and harvested in log phase at $1–3 \times 10^7$ cells/ml or $OD_{600}$ 0.5–1.3.

## Method details

### Induction of NMD reporters with galactose responsive promoters

For induction of NMD reporters under the control of the *GAL1* promoter, yeast cells were grown in media containing sucrose instead of glucose. The promoter was induced by addition of 2% galactose for 2 hours before harvesting, with the following exceptions: In Fig. 1E reporters with the endogenous *CBP80* and *DBP2* promoters were used. In Fig. 4G the NMD reporter was induced for 20 min after which transcription was stopped with 2% glucose. Cells were harvested after another 30 min of growth. In Fig. A, C, G $DBP2^{PTC}$ was not induced, cells were grown in sucrose to maintain a low transcription rate. In Fig. 2B, E, G, I $CBP80^{PTC}$ was induced for 4 h.

### Co-Immunoprecipitation (IP)

GFP fusion proteins were purified using GFP-Trap_A beads (Chromotek, gta-400) or GFP-selector beads (Nanotag Biotechnologies, N0310), following the manufacturer's instructions. Cell pellets were lysed in 1x volume cold PBSKMT buffer (137 mM NaCl, 5.7 mM KCl, 10 mM $KH_2PO_4$, 2 mM $Na_2HPO_4$ 2.5 mM $MgCl_2$, 0.5% Triton X-100) with protease inhibitor (5 µl per 100 µl cell pellet, Merck, 11,697,498,001). One pellet-volume of glass beads (0.4–0.6 mm) was added and cells were lysed in a FastPrep-24 (MP Biomedicals) at 4 m/s for 30 s twice. Glass beads and cell debris were removed by centrifugation at 16000x g for 1 min at 4°C and the supernatant was further cleared by centrifugation at 16000x g for 10 min at 4°C. Approx. 2% of the cleared lysate was kept as lysate sample for western blot analysis. The remaining lysate was incubated with equilibrated GFP-Trap_A beads (Chromotek) (Figs. 3A, B and 4E) or GFP-selector beads (Nanotag Biotechnologies) (Figs. 4A, B, D, F, 5A, B, D and 6A–D) for 2 h at 4°C. Where indicated,

200 μg/ml RNase A was added. The conditions for RNA removal were verified by qPCR. The beads were washed 4–8 times with PBSKMT and resuspended in SDS sample buffer (125 mM Tris – pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol, 0.05% (w/v) Bromophenol blue and 5% (v/v) 2-mercaptoethanol)). The complete eluate and the lysate samples were used for SDS-PAGE and western blot analysis with the indicated antibodies (GFP (GF28R, Thermo Fischer Scientific, MA5-15256) 1/50000, Zwf1 (Merck, A9521) 1/50000, Tdh1 (GA1R, Thermo Fischer, MA5-15738) 1/50000, Hem15 (U. Mühlenhoff) 1/5000, Gbp2 (self-made) 1/50000, Hrb1 (self-made) 1/20000, c-MYC (9E10, Santa Cruz, Sc-40) 1/750, HA (F-7, Santa Cruz, sc-7392) 1/750, Grx4 (U. Mühlenhoff) 1/1000.

For formaldehyde crosslinking (Fig. 4A), yeast cells were treated with 1% formaldehyde for 10 min at 25°C prior to harvesting. The formaldehyde was quenched by adding 0.5 M glycine. Immunoprecipitation was performed as described above with 20 min decrosslinking at 95°C in SDS sample buffer before gel loading.

### Yeast cell lysis for western blot analysis

For experiments shown in Fig. 2, log phase yeast cells were lysed in SDS sample buffer (125 mM Tris – pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol, 0.05% (w/v) Bromophenol blue and 5% (v/v) 2-mercaptoethanol) with one pellet volume (or 200 μl for smaller cell pellets) of glass beads (0.4–0.6 mm) and heated at 95°C for 5 min. Glass beads and cell debris were removed by centrifugation at 16000x g for 1 min. The supernatant was used for western blot experiments. For experiments shown in Fig. 2C–H, yeast cell cultures were split before harvesting. One half was used for western blot analysis, the other half for RNA isolation (see below).

### Quantification of western blot signals

Western blot signals were quantified with the Bio-1D software (Vilber Lourmat) by measuring the optical densities of western blot bands. With the Bio-1D software, images were selected in which the analysed signal intensities were not saturated. Background subtraction was performed using the rolling ball method and setting a detection threshold.

### Fluorescence microscopy

Logarithmic yeast cells were fixated with 2.6% formaldehyde and immediately harvested by centrifugation at 3500x g for 5 min at 4°C. The cells were washed once with 0.1 M potassium phosphate buffer pH 6.5, once with P solution (0.1 M potassium phosphate buffer pH 6.5, 1.2 M Sorbitol) and resuspended in P solution. The cells were incubated 15 min on polylysine coated microscope slides and excess cells were removed. The cells were permeabilized with 0.5% Triton X-100 in P-solution for approx. 1 min and washed once with P solution and once with Aby wash 2 (0.1 M Tris – pH 9.5, 0.1 M NaCl). DNA was stained with DAPI (1 μg/ml in Aby wash 2) for 5 min and washed three times for 5 min with Aby wash 2. Microscope slides were dried and the cells

mounted in 40% (v/v) glycerol, 20% (v/v) PBS (137 mM NaCl, 2.7 mM KCl, 10 mM $KH_2PO_4$, 2 mM $Na_2HPO_4$ and 1% (w/v) n-propyl gallate). Microscopy images were taken with a Leica AF6000 microscope and a LEICA DFC360FX camera with the LEICA AF 2.7.3.9 software. In Fig. 1A, Z-stacks (10 images, 0.2 μm) were deconvoluted (blind, 3 iterations, with the LEICA AF 2.7.3.9 software).

### Split GFP analysis

Proteins of interest were fused with either the N-terminal (amino acids 1–155) or the C-terminal (amino acids 156–239) part of eGFP. Fluorescence microscopy experiments were performed as described above, except cells were treated with 1.5% formaldehyde for fixation. For quantification of the fluorescence signal, 100 cells from each strain from each experiment were randomly chosen and the mean signal intensity of each cell was measured using Image J. Significant differences between strains were calculated by comparing all 300 signal intensity values.

### RNA co-immunoprecipitation (RIP)

Immunoprecipitation (IP) was performed essentially as described above followed by RNA isolation. A no-tag control was always used to verify RNA-binding to the precipitated proteins. RIP experiments shown in Fig 1F, G. were performed without UV crosslinking. In the other RIP experiments (Figs. 4D, F and 5D) protein-RNA complexes were crosslinked by UV irradiation. For this the cells were treated two times for 3.5 min (0.6 J/cm in a 50 ml suspension in a 15 cm petri dish) with 254 nM UV light on a cold metal block, with light shaking in between. For RIP experiments the cells were lysed in two pellet volumes RIP buffer (150 mM NaCl, 2 mM $MgCl_2$, 0.2 mM PMSF, 0.5 mM DTT, 0.2% (v/v) Triton X-100, 25 mM Tris/HCl – pH 7.5) with protease inhibitor (5 μl per 100 μl cell pellet, Merck, 11697498001) and RNase inhibitor (0.12 μl/100 μl pellet-volume RiboLock, Thermo Scientific, EO0381). Approx. 2% of the cleared lysate was kept as lysate sample for western blot analysis and 5–10% lysate was used as lysate sample for RNA isolation. DNaseI was added to the lysate sample (14 Kunitz units per 100 μl, Qiagen, 79256) and to the remaining RIP sample (6.5 Kunitz units per 100 μl). GFP tagged proteins were precipitated with GFP-Trap_A beads (Chromotek) (Fig. 1F, 1G (Gbp2)) or GFP-Selector beads (Nanotag Biotechnologies) (Fig. 1G (Hrb1), 4D, 4F, 5D) and the beads were washed 5–7 times with RIP buffer. Approx. 20% of the beads were resuspended in SDS sample buffer and used for western blot analysis. The remaining beads were used for RNA isolation with TRIzol (Thermo Fisher Scientific, 15596018), following the manufacturer's instructions. In UV-crosslinked RIP experiments, the beads were washed two more times with proteinase K buffer (50 mM NaCl, 0.5 mM DTT, 0.2% Triton X-100, 50 mM Tris/HCl pH 7.5) and resuspended in 100 μl proteinase K buffer. Afterwards, 0.5% SDS, 5 mM EDTA and 80 μg (lysate sample) or 40 μg (eluate sample) Proteinase K was added and incubated 90 min at 55°C with shaking prior to RNA isolation.

DNA was further removed with the TURBO DNA-free DNase kit (Thermo Fischer Scientific, AM1907).

### Total RNA isolation

Total RNA was isolated from log phase yeast cultures using the NucleoSpin® RNA isolation kit (Macherey-Nagel, 740,955). DNA was further removed with the TURBO DNA-free DNase kit (Thermo Fischer Scientific, AM1907).

### Reverse transcription and quantitative PCR

One microgram total RNA or 50–100 ng eluted RNA was reverse transcribed with the Maxima First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, EP0741) or the FastGene Scriptase II kit (NIPPON Genetics, LS63). Reverse transcription was performed with either Oligo $(dT)_{18}$ (Fig. 1D–G (Gbp2), Figs 2D, F, H, and 4G) or random primers (Fig. 1G (Hrb1), 4D, 4F, 5D). qPCR was performed in triplicates, using the qPCRBIO SyGreen Mix Lo-ROX (NIPPON Genetics, PB20.11–50) in a CFX Connect 96FX2 qPCR cycler (BIO RAD).

### Quantification and statistical analysis

All experiments were performed in independent biological replicates as indicated in the figure legends. All bar graphs show arithmetic mean values and error bars illustrate the standard deviation of biological replicates. P-values were calculated by unpaired, two tailed, homo- or heteroscedastic Student's t-test and are indicated by * ($p < 0.05$), ** ($p < 0.01$) and *** ($p < 0.001$). The number of biological replicates is indicated as 'n' in the figure legends.

## Results

### The nuclear guard proteins Gbp2 and Hrb1 show features of NMD factors

Gbp2 and Hrb1 are established nuclear quality control proteins that shuttle with the mRNA into the cytoplasm [9,23]. Thus, it seems conceivable that they might also participate in the cytoplasmic mRNA quality control. It was shown earlier that defects in the NMD, NGD and NSD pathways result in increased protein aggregation due to progressive abnormal association of misfolded proteins in insoluble protein structures, central to the pathology of neurodegenerative diseases such as Alzheimer's and Parkinson's [24]. In their studies, Jamar et al. analysed protein aggregation of RFP tagged Hsp104 by visualizing and quantifying fluorescent Hsp104-RFP foci, which increased when mRNA surveillance is defective, such as in upf1Δ cells. We used their established Hsp104-RFP microscopy assay and compared foci formation in wild type and upf1Δ cells to the situation in the gbp2Δ hrb1Δ double knock out strain. While in wild type these protein aggregates were only visible in ~5% of the cells, both upf1Δ and gbp2Δ hrb1Δ strains showed dot-like protein aggregates in more than 15% of the cells, comparable with the values obtained in the original publication (Fig. 1A, B). This

accumulation didn't significantly increase in the triple knock out strain gbp2Δ hrb1Δ upf1Δ, suggesting that these proteins might act in one pathway. As Gbp2 and Hrb1 are involved in the nuclear surveillance, we investigated whether defects in nuclear quality control factors per se lead to increased protein aggregation due to an increased leakage of defective mRNAs, which might overwhelm the cytoplasmic surveillance systems. For this we analysed cells that were deleted for the NPC gatekeeper MLP1. As shown in Fig. 1A, B, mlp1Δ showed no increased protein aggregations of Hsp104, indicating that the nuclear escape of faulty mRNAs is not sufficient to create protein aggregates. This suggests that Gbp2 and Hrb1 might have a yet undiscovered function in the cytoplasmic quality control.

Because Gbp2 and Hrb1 are nuclear quality control factors for splicing, a processing step that is a source for PTCs when not carried out correctly, we investigated whether Gbp2 and Hrb1 might function in NMD. We used the well-established reporter assay in which PTC-containing PGK1 transcript is highly expressed using a galactose-inducible promoter [25] (Fig. 1C, S1A). However, while an ~8-fold increase of the $PGK1^{PTC}$ mRNA was detectable in upf1Δ by qPCRs from the isolated total RNA, no increased level was measured for gbp2Δ or hrb1Δ or the double mutant (Fig. 1D). This suggests that Gbp2 and Hrb1 are not involved in the NMD-induced degradation of this intron-less reporter transcript. However, these guard proteins are preferentially loaded onto spliced pre-mRNAs and have roles in nuclear mRNA quality control specific for these transcripts [9], which reminds of the EJC in metazoans. Thus, a function of these proteins in cytoplasmic quality control may also be specific to the subset of mRNAs that are spliced. Therefore, we constructed two intron-containing reporter mRNAs (Fig. 1C, S1A). We chose the DBP2 gene, which possesses a long open reading frame before the intron sequence and placed the PTC upstream of the intron, which is rather atypical for yeast, but represents the frequently found EJC-model in human cells. Additionally, we selected the CBP80 gene in which the intron sequence is located very close to the AUG start codon, which is common in yeast, and placed the PTC shortly downstream of the intron. We detected the steady-state RNA levels of these intron-containing reporters, expressed upon galactose induction, and observed a ~ 3.6-fold ($DBP2^{PTC}$) and ~5.6-fold ($CBP80^{PTC}$) increase in upf1Δ compared to wild type (Fig. 1D), showing that they are targeted for the Upf1-dependent NMD pathway under wild-typical conditions.

We next asked if deletion of GBP2 and HRB1 would have an effect on the intron-containing $DBP2^{PTC}$ and $CBP80^{PTC}$ reporters. It should be noted that Gbp2 and Hrb1 are nuclear retention factors and are dispensable for mRNA export and splicing [9]. Further, to exclude effects from their nuclear quality control function, we used wild-typical reporters as controls that are identical but lack the PTCs (Fig S1A). As PTCs can only be recognized during translation, the difference between PTC-containing and PTC-less reporters have to be a consequence of mRNA stability through the NMD pathway. We expressed these reporters using the transcripts' endogenous promoters in order to reflect natural conditions as much as possible. Results of qPCR analyses show that the

NMD reporters are enriched ~2.8-fold ($DBP2^{PTC}$) and ~2.3-fold ($CBP80^{PTC}$) in $upf1\Delta$ (Fig. 1E). Interestingly, with a ~ 1.8-fold enrichment for both reporters, $gbp2\Delta$ $hrb1\Delta$ cells showed approximately half of the $upf1\Delta$ effect. In the absence of Upf1, the additional loss of Gbp2 and Hrb1 had no further effect on either reporter (Fig. 1E, S1B). Together, these observations suggest that Gbp2 and Hrb1 may act in the Upf1-mediated pathway on transcripts derived from intron-containing genes.

Upf1 is stabilized on NMD targets [26]. Studies on human UPF1 showed that NMD factors that are relevant for the initial detection of NMD targets, such as UPF2, are important for the interaction of UPF1 with PTC-mRNAs [27]. To investigate whether Gbp2 and Hrb1 might affect the initial detection of NMD or rather act after Upf1 has triggered the pathway, we carried out similar experiments and tested via RNA-co-immunoprecipitation (RIP) whether the interaction of Upf1 with the $CBP80^{PTC}$ reporter is affected in the absence of Gbp2 and Hrb1. As an internal control to rule out NMD unrelated effects, we normalized reporter RNA levels to an endogenous wild-typical mRNA. All precipitated RNA levels were normalized to their relative levels from whole-cell lysates. We found that, unlike $upf2\Delta$, the double knock out of $GBP2$ and $HRB1$ did not affect the binding of Upf1 to the $CBP80^{PTC}$ NMD reporter (Fig. 1F, S1C). In contrast, about 50% less PTC-containing $CBP80$ was bound to either Gbp2 or Hrb1 when Upf1 was missing (Fig. 1G, S1D). This suggests that Gbp2 and Hrb1 are likely not involved in NMD substrate recognition and Upf1 recruitment, but might rather help in downstream events of NMD. When Upf1 is absent and PTCs are not identified as false, Gbp2 and Hrb1 may dissociate earlier from the NMD reporter during normal rounds of translation, resulting in the observed decreased association.

To analyse the relevance of Gbp2 and Hrb1 to NMD in physiological conditions, we studied the binding of these proteins to natural NMD substrates. We chose the intron-containing $GCR1$ and $HNT1$ transcripts that were identified as putative NMD targets in a genome wide analysis [20]. We found that interactions of Gbp2 and Hrb1 with these natural NMD substrates were reduced when $UPF1$ was deleted, similar to the findings with the reporter construct (Fig. 1G). Together, our first results uncover the involvement of Gbp2 and Hrb1 in NMD and suggest that they likely act downstream of Upf1 in the effective elimination of a subset of NMD targets (presumably those that are spliced).

### Gbp2 and Hrb1 repress translation of NMD reporter transcripts

NMD prevents the expression of prematurely terminated and thus potentially harmful proteins after PTC detection by two distinct, yet intertwined, mechanisms: a) the degradation of the transcripts, and b) the repression of further translation initiation on such faulty mRNAs. For NMD it is known that this repression requires Upf1, as NMD substrates show an increased translation in $upf1\Delta$ cells [17]. To investigate whether our intron-containing PTC-reporter constructs also undergo translational repression, we analysed their expression in wild type and $upf1\Delta$ cells. For this purpose, we created variants of the reporter mRNAs that encode N-terminally MYC-tagged proteins (Fig S1A), which allowed detection, even if they were only translated up to the PTC. With western blot analyses, we show that translation of the $MYC\text{-}DBP2^{PTC}$ reporter was terminated at the PTC, resulting in a 45 kDa truncated protein (Fig. 2A), which was substantially enriched when Upf1 was missing (Fig. 2C). Likewise, a Upf1-dependent translational repression of $MYC\text{-}CBP80^{PTC}$ was also observed (Fig. 2E). Interestingly, the translated product of $MYC\text{-}CBP80^{PTC}$ was not terminated at the PTC, but rather at the original stop codon, producing a full-length protein (Fig. 2B). However, translation of $CBP80^{PTC}$ in wild type was 10-fold lower than in $upf1\Delta$ (Fig. 2F), comparable to other described NMD reporters [17,20,28], entailing that the PTC is indeed recognized by the NMD machinery efficiently. Moreover, the read-through product in $upf1\Delta$ is still several magnitudes lower than the normal PTC-less $CBP80$ translation (Fig S2A), suggesting that read through of the PTC is extremely rare in wild type cells and presumably occurs only when NMD fails and not vice versa. Consistently, PTCs are described to be susceptible to readthrough, especially when NMD is impaired [29]. In fact, the widely used $PGK1^{PTC}$ reporter showed an identical behaviour when fused to an N-terminal MYC-tag (Fig S2D) and although the PTC is apparently read through in $upf1\Delta$ cells, the mRNA remains susceptible to NMD [30]. In line with this, it was recently demonstrated that each round of translation has an equal probability to initiate NMD [31]. Thus, read through of a PTC by one ribosome does not render the mRNA immune to NMD in subsequent rounds of translation.

To study whether the guard proteins Gbp2 and Hrb1 also function in translational repression of the NMD targets, we analysed the expression of PTC-containing reporter constructs in the single and double knock out strains. To obtain an estimate of the translation rate of the reporters, we measured the relative RNA level in each strain by qPCR (Fig S2B, S2C) and related the protein signals to the respective RNA levels. We found that in the case of the $DBP2^{PTC}$ reporter in which the PTC is in the middle of the transcript, both Gbp2 and Hrb1 were necessary for functional translational repression, as their absence increased translation more than two-fold (Fig. 2C, D). Interestingly, in case of the rather yeast-typical $CBP80^{PTC}$ reporter in which the PTC is shortly after the start codon, only Gbp2 seemed to be relevant with its absence leading to a ~ 5.5-fold increase in protein level (Fig S2C) and a ~ 3.5-fold increase of protein per mRNA (Fig. 2E, F). The effects are PTC-dependent, as protein levels of PTC-less reporters remain similar in all strains (Fig. 2C, E). Consistently, such Gbp2- and/or Hrb1-mediated translational repression was not observed with the $PGK1^{PTC}$ reporter that was derived from an unspliced gene (Fig S2D). To manifest that this translational repression is Upf1-dependent, we compared the NMD reporter translation obtained in the $upf1\Delta$ strain with that detected in the $gbp2\Delta$ $hrb1\Delta$ $upf1\Delta$ triple knock out strain. Similar to the effect on PTC-containing mRNA degradation, the loss of Gbp2 and Hrb1 had no further effect if Upf1 was absent (Fig. 2G, H). The measured protein levels in our analyses may also be affected by differences in protein stability, since Upf1 also causes
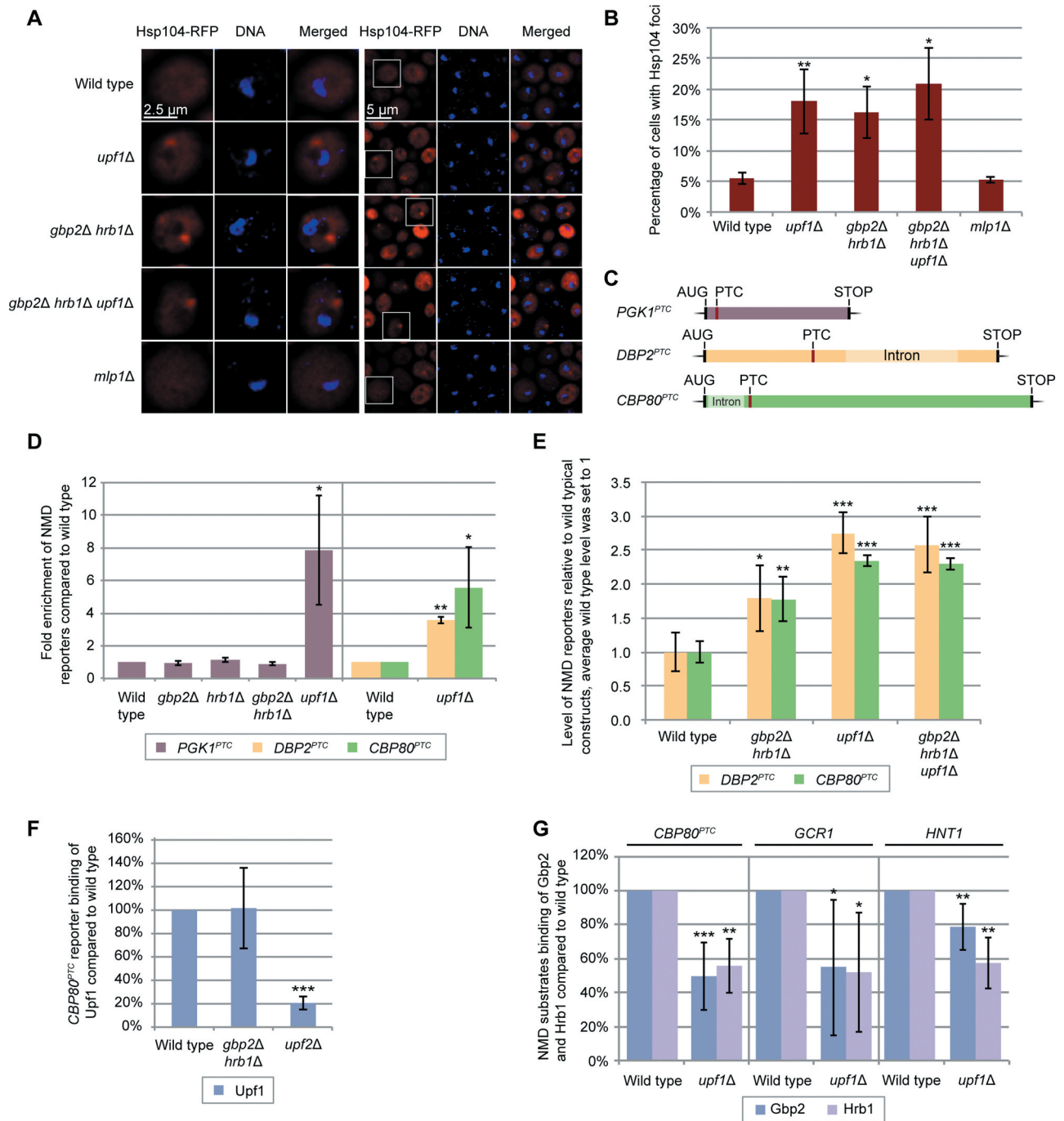
**Figure 1.** The nuclear guard proteins Gbp2 and Hrb1 show features of NMD factors. (A) Protein aggregation is increased in cells lacking *GBP2* and *HRB1*. Localization of RFP-tagged Hsp104 is shown in the indicated strains that were grown to the logarithmic growth phase at 25°C and shifted to 37°C for 1 h. (B) Cells that contain Hsp104-RFP foci were counted and the percentage of cells with aggregates is shown. 300 cells were counted per experiment and error bars represent the standard deviation between different experiments. $n = 3$ (wild type and *upf1Δ* $n = 6$). (C) Scheme of the used reporter constructs. See also Fig S1A. (D) Gbp2 and Hrb1 do not function in the Upf1-mediated decay of the intron-less *PGK1* transcript. PTC-containing transcripts were expressed by 2 h galactose induction and monitored by qPCR. Newly generated *DBP2^PTC* and *CBP80^PTC* reporters were expressed in wild type and *upf1Δ* to compare them to the established *PGK1^PTC* reporter. $n = 3$ (*PGK1^PTC*), $n = 5$ (*DBP2^PTC*, *CBP80^PTC*). (E) Gbp2 and Hrb1 are required for the effective degradation of the PTC-containing, spliced *DBP2^PTC* and *CBP80^PTC* transcripts. Transcripts were expressed using their endogenous promoters. qPCRs from RNA of the indicated strains were carried out in the presence or the absence of the PTC and are shown in relation. The average wild type level of PTC-containing NMD reporter per PTC-less reporter was set to 1 and other data are shown in relation. $n = 4$ and $n = 4$ (*gbp2Δ hrb1Δ* $n = 7$), respectively. See also Fig S1B. (F) The binding of Upf1 to the PTC-containing reporter RNA is independent of Gbp2 and Hrb1. RNA-co-immunoprecipitation (RIP) experiments of Upf1-GFP were carried out in the indicated strains and the amount of bound PTC-reporter transcript was normalized to *RPS6A* mRNA. $n = 8$ (*gbp2Δ hrb1Δ*), $n = 4$ (*upf2Δ*). See also Fig S1C. (G) The binding of Gbp2 and Hrb1 to *CBP80^PTC* and endogenous NMD substrates is reduced in the absence of Upf1. RIP experiments with Gbp2 and Hrb1 were done in wild type and *upf1Δ* cells and qPCR results are shown. RNA levels were normalized to *21S* rRNA. *CBP80^PTC*: Gbp2 $n = 7$, Hrb1 $n = 6$; *GCR1*: $n = 6$; *HNT1*: Gbp2 $n = 7$, Hrb1 $n = 5$. See also Fig S1D.
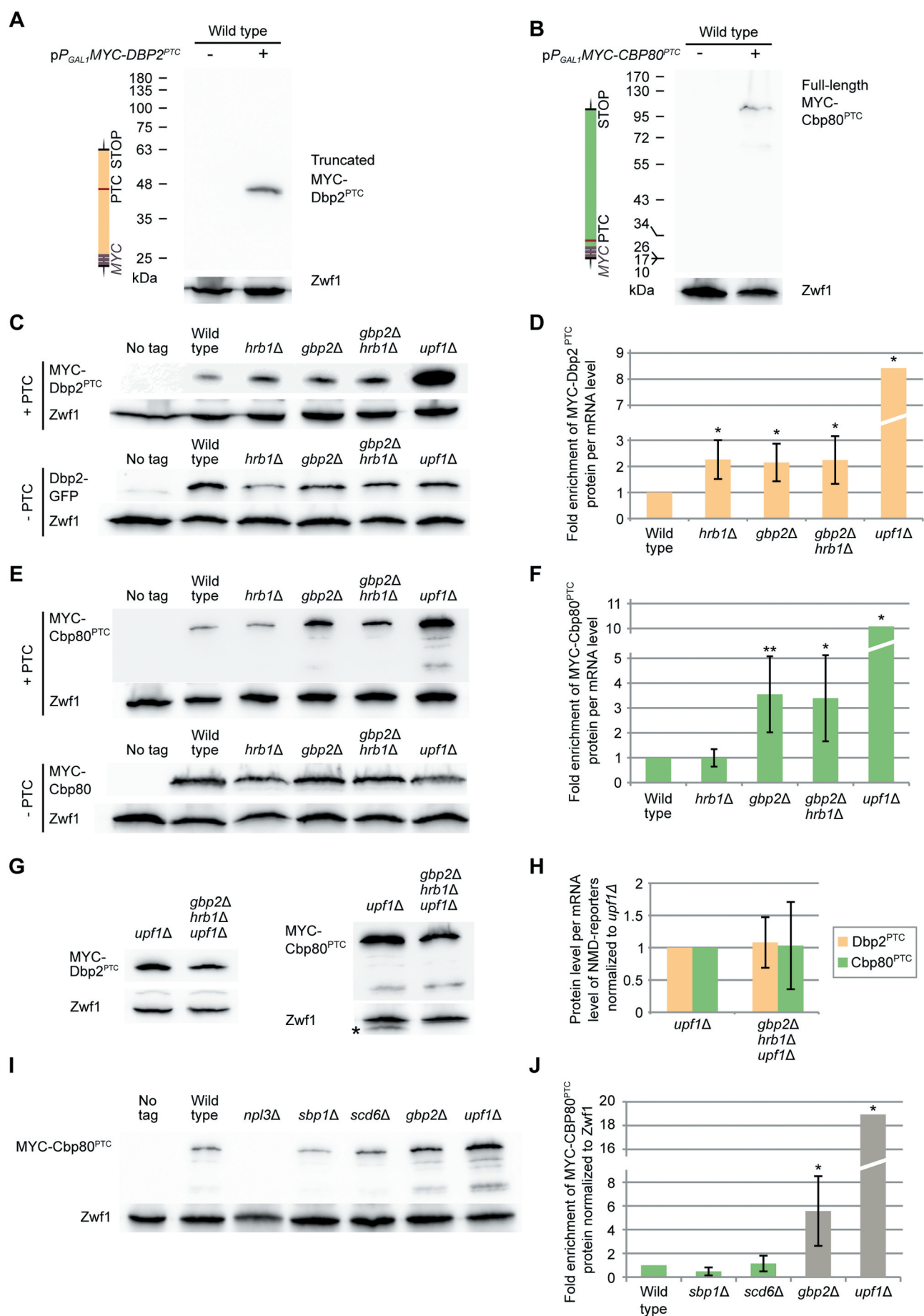
Figure 2. Gbp2 and Hrb1 are involved in translation repression of NMD targets. (A) Translation of the *DBP2^PTC* reporter results in a truncated protein, shown on

destabilization of the nascent polypeptide [28,32]. Nevertheless, this is also a Upf1-mediated effect as part of NMD, hence the observed effects of Gbp2 and Hrb1 on protein level per PTC-mRNA are most probably effects within the NMD pathway. Further, the full-length MYC-Cbp80 protein appears to be a consequence of failed NMD at the PTC followed by normal translation termination at the regular stop codon. It seems unlikely that Upf1 causes destabilization of such normal translation products, indicating that the difference in MYC-Cbp80 levels resulted essentially from differences in translation. Therefore, while the effects of Gbp2 and Hrb1 on $DBP2^{PTC}$ may partially be a consequence of protein stability, we have to assume that Gbp2 is indeed involved in the translation repression of the $CBP80^{PTC}$ reporter.

Both guard proteins contain a serine/arginine (SR)-rich domain, which is also comprised of several arginine/glycine/glycine (RGG)-motifs. The RGG domain was described to be important for a group of proteins, Scd6, Sbp1 and Npl3, involved in inhibition of translation initiation by directly binding eIF4G via their RGG-motifs [33–35]. This makes Gbp2 and Hrb1 potential candidates for NMD-dependent translation repressors. However, to investigate if the known RGG-motif translation repressors can also inhibit the translation of NMD targets, we investigated the expression of our reporters in the respective knock out strains. As shown in Fig. 2I and J, the absence of none of the three proteins increased the translation of the NMD reporter, suggesting that translational repression of NMD substrates could be a specific function of Gbp2 and Hrb1. Interestingly, protein expression in $npl3\Delta$ was completely abolished (undetectable even with long exposure times) while the RNA level was ~10% of that in wild type cells (Fig S2E), suggesting a more general function for Npl3 in translation, which would fit to its proposed role in ribosomal subunit joining [33]. The fact that RGG domain-containing Gbp2 and Hrb1 specifically affected translational repression of $DBP2^{PTC}$ and $CBP80^{PTC}$ but not of the $PGK1^{PTC}$ reporter raises the possibility that these proteins may directly repress translation on specific NMD substrates downstream of Upf1.

Although in previous analyses we observed relatively mild effects on the RNA levels of our intron-containing reporters (Fig. 1D, E), more significant effects were seen on the protein level. Functional Upf1 reduced the amount of translated protein from the reporters on average ~19- and ~25-fold (Fig S2B, S2C), comparable to other established NMD reporters [17,20,28]. As a quality control pathway, one of the main

functions of NMD is the repression of aberrant protein production and in this regard, NMD seems to function normally on the $CBP80^{PTC}$ and $DBP2^{PTC}$ reporters.

## Gbp2 and Hrb1 presumably take part in NMD in the cytoplasm

While we could see that Gbp2 and Hrb1 are relevant for NMD on our reporter constructs on both the RNA and protein-level, it is unclear if the two proteins are physically involved in NMD in the cytoplasm. To investigate whether Gbp2 and Hrb1 physically interact with the Upf-proteins, we carried out co-immunoprecipitation (co-IP) analyses. GFP-tagged Upf1, Upf2 and Upf3 were pulled down from yeast cell lysates and the co-precipitation of Gbp2 and Hrb1 was investigated using specific antibodies (Fig S3A, S3B). Both guard proteins co-purified with all three Upf-proteins (Fig. 3A), although the interactions were sensitive to RNase. This could mean that the proteins are present on the same RNA but not in the same complex, or that the interactions occur only when Gbp2 and Hrb1 are bound to RNA. To further understand the interactions between these proteins, we performed co-IP experiments with strains expressing the wild-type or an ATP-hydrolysis defective mutant of Upf1, upf1-DE572AA (Fig. 3B). In the $upf1\text{-}DE572AA$ mutant, RNA-binding of upf1 is not affected [36], but the ribosome cannot disassemble after NMD has been initiated and the Xrn1-mediated 5′ decay stops at the stalled ribosome, resulting in accumulation of a 3′ decay fragment [18,19]. Moreover, several NMD factors showed increased co-purification with mutant upf1 on the decay fragments in human cells [18]. To test the functionality of the UPF1- and upf1-DE572AA-GFP plasmids, we transformed $upf1\Delta$ cells and analysed cell growth on cycloheximide-containing plates (Fig S3C). UPF1 deletion was shown to result in increased sensitivity of the cell to the translation inhibitor [37,38], an effect that was attributed to the fact that NMD is translation-dependent. This growth defect could be rescued by the wild-typical UPF1- but not the upf1-DE572AA-GFP plasmid (Fig S3C). Subsequently, we found that co-precipitation of Gbp2 with upf1-DE572AA selectively increased more than 1.5-fold compared to wild-type Upf1 (Fig. 3B, C). Since Hrb1 did not show an increased association, it cannot be an unspecific enrichment of general RNA-binding proteins. This suggests that Gbp2 is likely still bound to the RNA decay fragments, while Hrb1 might dissociate at an earlier point in time.
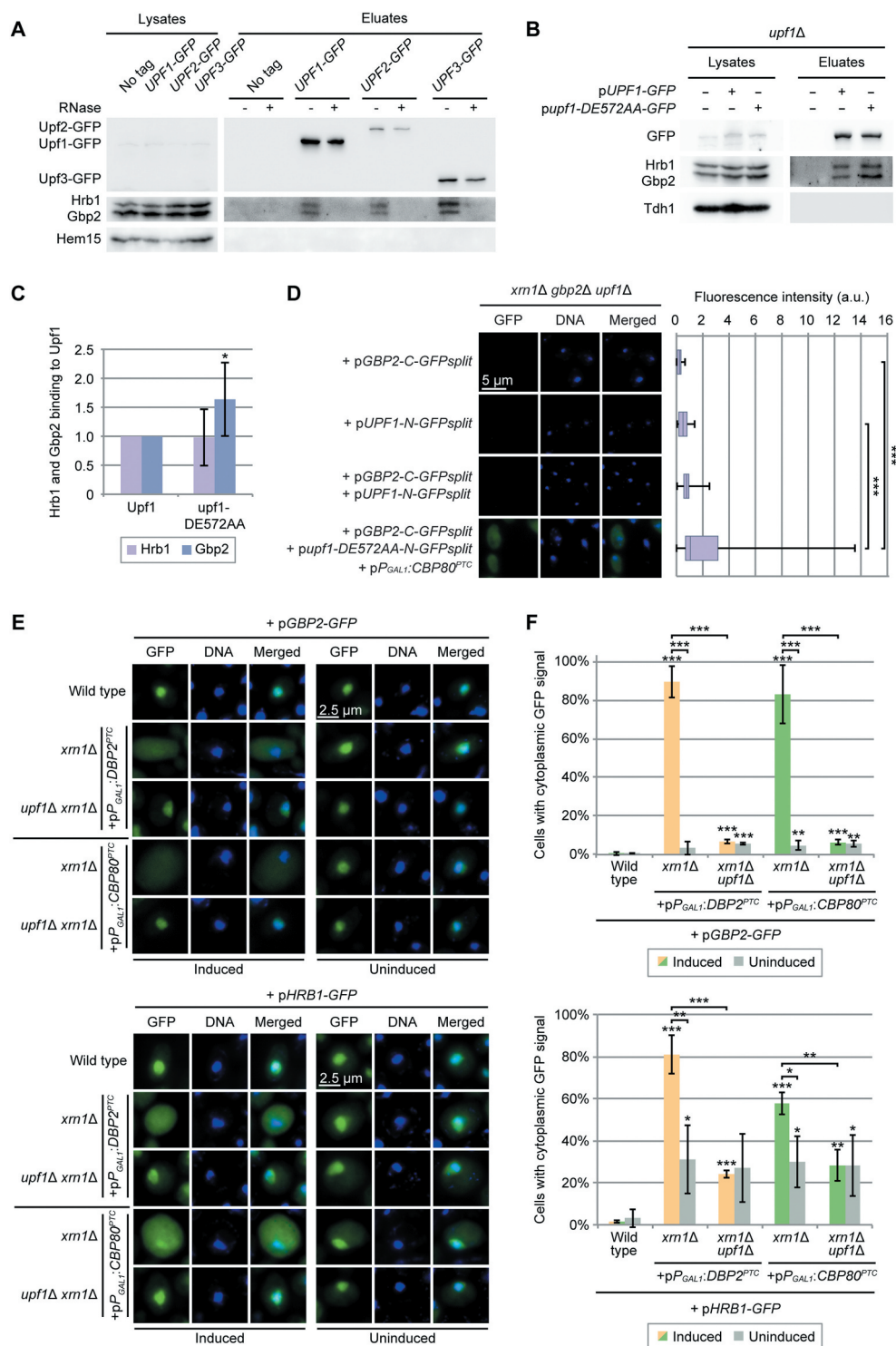
**Figure 3.** Gbp2 and Hrb1 seem to take part in NMD in the cytoplasm. (A) Gbp2 and Hrb1 co-precipitate with all three Upf proteins. Western blot analysis of co-IPs of Gbp2 and Hrb1 with GFP-tagged Upf1, Upf2 and Upf3 are shown. GFP-tagged Upf-proteins were not detectable in the lysates. Hem15 served as a negative control. (B) The interaction of Gbp2 and Upf1 increases when the ATPase activity of the helicase is defective. A western blot of a Upf1-GFP and upf1-DE572AA-GFP IP and Gbp2 and Hrb1 co-precipitation is shown. (C) The binding of Gbp2 and Hrb1 with upf1-DE572AA shown in (B) was quantified from independent experiments. $n = 7$. The signal intensities of the Gbp2 and Hrb1 bands were related to the corresponding Upf1- or upf1-DE572AA-GFP pull-down signals. (D) Upf1 and Gbp2 physically interact. Split-GFP experiments with the indicated plasmids are shown. Cells expressing either N-GFPsplit or C-GFPsplit alone were used as negative controls. The experiments were performed in $xrn1\Delta$ cells to reduce the degradation of PTC-containing transcripts after NMD initiation. If indicated, p$P_{GAL1}$:$CBP80^{PTC}$ was induced for 2 h to increase the presence of PTC-mRNAs. The signal of 100 cells was quantified per experiment. $n = 3$. (E) Both Gbp2 and Hrb1 mislocalize to the cytoplasm when PTC-containing transcripts cannot be degraded efficiently. GFP-tagged Gbp2 and Hrb1 were localized by fluorescence microscopy in wild type, $xrn1\Delta$ and $upf1\Delta$ $xrn1\Delta$ cells in the presence or absence of the indicated PTC-reporter plasmids. Cell cultures were split in two and expression of the reporter constructs was induced for 2 h in one sample. (F) Quantification of the experiments shown in (E). Error bars represent the standard deviation between independent experiments with 100–200 analysed cells per experiment. $n = 3$. See also Fig S3D.

To gain further insight into the interaction between Upf1 and Gbp2, we used the split-GFP system, which allows detection of transient protein-protein interactions [39]. The proteins of interest were expressed with N-terminal or C-terminal parts of GFP. In case of close proximity, the GFP fragments assemble and emit fluorescent light in living cells [39]. Although no significant amount of GFP-signal was detectable under wild-typical conditions, clear GFP-signals were measured in the presence of elevated levels of NMD substrates when the C-terminal GFP (*C-GFPsplit*) was tagged to Gbp2 and N-terminal GFP (*N-GFPsplit*) fused with the upf1-DE572AA mutant (Fig. 3D). This shows that Gbp2 comes into close proximity with Upf1 in the cell, presumably in the same complex at the site of the PTC, as the upf1-DE572AA protein is stalled there. However, this analysis suggests also that such complexes are low abundant and rather labile in wild-typical situations, possibly due to the immediate degradation of the PTC-containing mRNA and the simultaneous disassembly of the associated protein complexes.

To get further evidence for a cytoplasmic involvement of the guard proteins in NMD, we impaired NMD at an earlier point in time, by deletion of *XRN1*, to prevent the initial 5′-degradation and analysed, whether this delay would visibly affect re-import of Gbp2 and Hrb1 into the nucleus. Clearly, in the presence of increased levels of NMD-substrates, we detected both guard proteins in the cytoplasm of *xrn1Δ* (Fig. 3E, F). To ascertain that this is indeed caused by NMD we additionally deleted *UPF1*. In fact, the cytoplasmic localization of both guard proteins disappeared in *xrn1Δ* when Upf1 was absent (Fig. 3E, F), despite the reporter levels being even higher in these cells (Fig S3D), suggesting an NMD-specific effect. In agreement, overexpression of PTC-less reporters did not result in the cytoplasmic localization of either Gbp2 or Hrb1 (Fig S3E). This shows that ongoing NMD delays the nuclear reimport of Gbp2 and Hrb1, possibly because the proteins remain associated with the RNAs that hold out for NMD degradation. Together, these results imply that both Gbp2 and Hrb1 are present in NMD-complexes. The stronger mislocalization of Gbp2 and its persistent binding to stalled Upf1-complexes furthermore supports the idea that Hrb1 might leave the NMD-identified mRNA earlier than Gbp2.

## Hrb1 promotes the recruitment of the 5'-degradation machinery to NMD targets

A nuclear function of the guard proteins is to load the degradation machinery to faulty transcripts and in this way initiate their elimination [5,9]. It is conceivable that Gbp2 and Hrb1 might have a similar function in the cytoplasm. Degradation of NMD-targets mainly depends on the Dcp1/Dcp2-mediated de-capping and the subsequent Xrn1-mediated exonucleolytic RNA decay [18,19,25,40]. Dcp1, as well as Dcp2, co-purified with Upf1-bound complexes that also contain other decay factors [20]. Therefore, we first investigated whether Gbp2 and Hrb1 interact with Dcp1. Co-IPs with GFP-tagged Dcp1 showed an interaction of Dcp1 with both Gbp2 and Hrb1 (Fig. 4A, S4A). While we initially observed that the co-precipitation of Gbp2 with Dcp1 was lost upon RNase treatment (Fig S4B), this co-precipitation was visible under

crosslinking conditions with formaldehyde (Fig. 4A) in which effective RNA removal was verified via qPCR (Fig S4C).

To ensure that the interaction of the guard proteins with Dcp1 is relevant for NMD, we investigated whether the Dcp1-Upf1 interaction was affected by the absence of the two guard proteins. Indeed, their interaction was reduced to ~67% in the *gbp2Δ hrb1Δ* strain (Fig. 4B, C). This could indicate that Dcp1 is not properly targeted to NMD-substrates when Gbp2 and Hrb1 are missing. While our studies on the $CBP80^{PTC}$ and $DBP2^{PTC}$ reporters suggest that the Upf1-mediated degradation is diminished approximately by half in *gbp2Δ hrb1Δ* (Fig. 1E, S1B), we can only see an average reduction of one third in the overall Upf1-Dcp1 interaction. However, this analysis was performed without expressing an NMD reporter and relies on the interaction of Upf1 with Dcp1 on endogenous NMD targets. As Gbp2 and Hrb1 appear to be relevant for a subset rather than all NMD targets, a milder effect would be expected in this analysis. Consequently, we would expect stronger effects by directly analysing the Dcp1 binding to an NMD target that is affected by Gbp2 and Hrb1. Indeed, RIP-experiments revealed a significantly reduced binding of Dcp1 to $CBP80^{PTC}$ when the guard proteins were missing (Fig. 4D, S4D). Loss of Gbp2 and Hrb1 reduces Dcp1 binding approximately half as much as Upf1, agreeing with our analysis shown in Fig. 1E. Interestingly, while the single knock out of *HRB1* showed the same decrease in Dcp1 recruitment to the NMD-target as the double knock out, we detected no effect for *gbp2Δ*. This supports a model in which mostly Hrb1 is involved in proper Dcp1 recruitment to a subset of NMD targets.

RNAs with removed caps are substrates for Xrn1, which also physically interacts with both Gbp2 and Hrb1 (Fig. 4E, S4A) and Upf1 [20]. The interaction between Hrb1 and Xrn1 remained intact upon addition of RNase A, while the interaction of Gbp2 strongly decreased, suggesting that Gbp2 may require RNA binding for interaction with the 5′-degradation machinery. Nevertheless, the Hrb1-Xrn1 interaction was RNase insensitive, which indicates a physical complex of Hrb1 and the 5′ degradation machinery. Interestingly, subsequent RIP-experiments of Xrn1 to the NMD-targets revealed that the interaction of Xrn1 was unaffected in *gbp2Δ hrb1Δ* (Fig. 4F, S4E), suggesting that Xrn1 is not recruited by the guard proteins, but might rather wait in the NMD-complexes for uncapped substrates. As it cannot degrade capped RNAs, there is no necessity for a regulated recruitment of Xrn1. To further test if decapping, and thereby Xrn1 degradation, is defective without Gbp2 and Hrb1, we performed an in vitro Xrn1 digestion experiment. We observed that Xrn1 readily degrades $CBP80^{PTC}$ RNA purified from cells deleted for *XRN1* (Fig. 4G), indicating that the purified reporter RNAs are mostly decapped. The additional deletion of *GBP2, HRB1* or *UPF1* strongly impairs the in vitro degradation, suggesting that decapping is defective in these strains. This effect is PTC dependent, as wild-typical *CBP80* showed no differences between the mutants (Fig. 4G). The remaining fractions of the $CBP80^{PTC}$ RNA vary strongly in the mutant strains in this analysis. This doesn't allow quantitative comparison between the different mutants; however, all mutants do appear to have
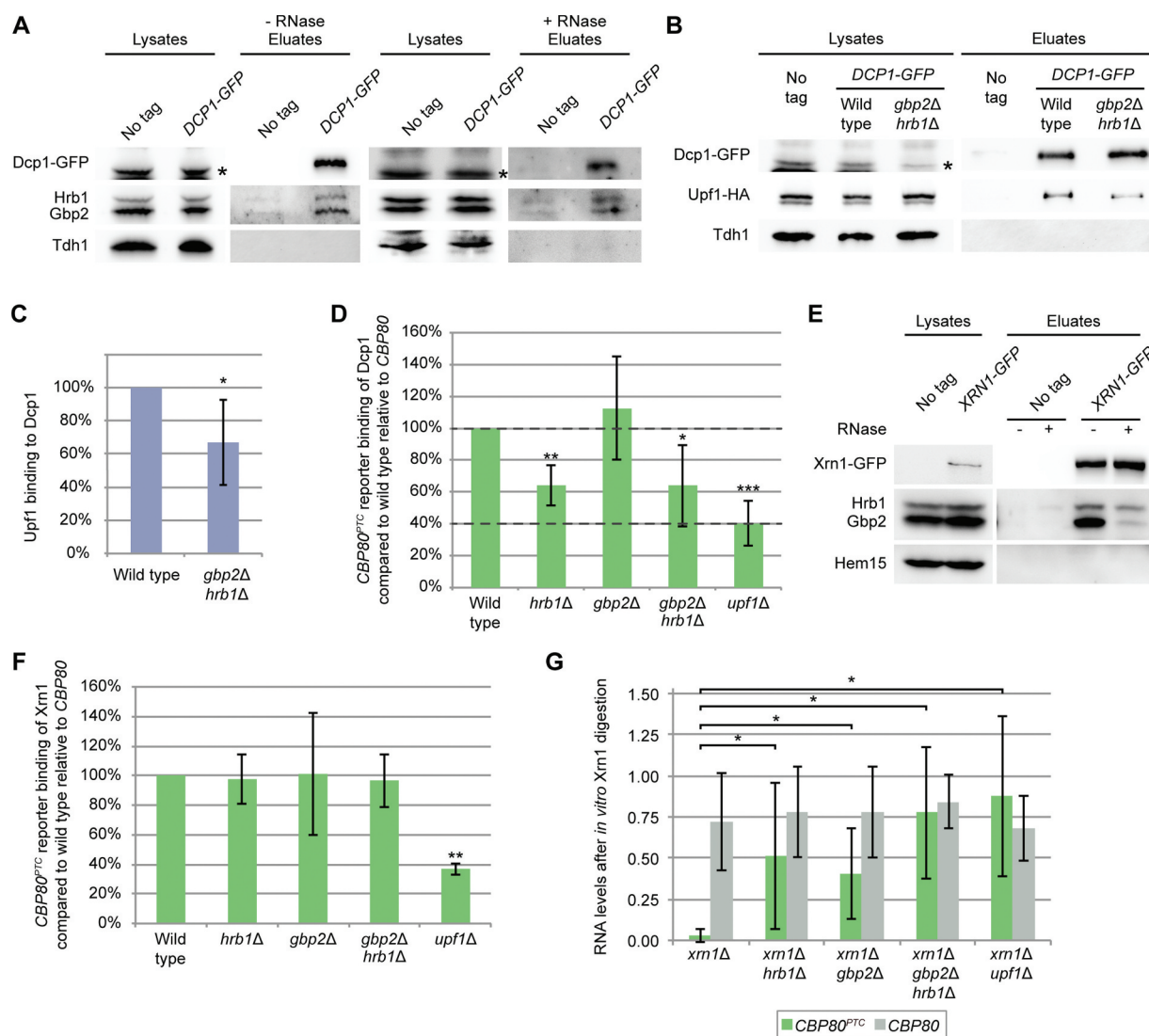
**Figure 4.** Hrb1 is involved in the recruitment of the 5′-end degradation machinery. (A) Gbp2 and Hrb1 co-precipitate with Dcp1. Western blots of co-IPs of Gbp2 and Hrb1 with Dcp1-GFP are shown. The asterisks indicate bands of Hrb1 from previous detection with the Hrb1 antibody. Tdh1 served as a negative control. For the RNase treated IP the cells were treated with 1% formaldehyde for 10 min at 25°C. After the precipitation, proteins were de-crosslinked for 20 min at 95°C. Dcp1-GFP was not detectable in the lysates. See also Fig S4B and S4C. (B) The interaction of Upf1 and Dcp1 is promoted by Gbp2 and Hrb1. Co-IPs of Upf1-HA with GFP-tagged Dcp1 in *GBP2 HRB1* and *gbp2Δ hrb1Δ* cells are shown on western blots. All cells are deleted for *UPF1*, and express p*UPF1-HA*. The asterisk indicates an unspecific cross-reaction with the GFP antibody. Dcp1-GFP was not detectable in the lysates. Tdh1 served as a negative control. (C) Quantification of seven independent co-IPs shown in (B). Signal intensities of the Upf1-HA bands were related to the corresponding Dcp1-GFP pull-down signals. (D) The binding of Dcp1 to a PTC-containing transcript is disturbed in the *HRB1* knock out. Dcp1 RIP experiments and subsequent qPCRs were carried out in the indicated strains. All strains express genomic *DCP1-GFP*. n = 5 (*hrb1Δ* n = 6). Co-purified RNA levels were normalized to the endogenous wild-typical *CBP80* mRNA and the total levels from whole-cell lysates. Dashed lines indicate the level of wild type and average level of *upf1Δ*. See also Fig S4D. (E) Xrn1 interacts with Gbp2 and Hrb1. Gbp2 and Hrb1 co-IPs with Xrn1-GFP are shown on western blots. Hem15 served as a negative control. (F) Xrn1 recruitment to PTC-containing substrates is Upf1- but not Gbp2- or Hrb1-dependent. Xrn1 RIP experiments and subsequent qPCRs with the PTC-containing reporter are shown in the indicated strains. All strains express genomic *XRN1-GFP*. n = 3. See also Fig S4E. (G) Decapping of *CBP80^{PTC}* RNA is defective without Gbp2 or Hrb1. RNA was isolated in the indicated strains containing the *CBP80^{PTC}* reporter. A sample of this RNA was used for *in vitro* Xrn1 digestion, which can only degrade decapped RNAs. *CBP80^{PTC}* and endogenous *CBP80* were detected after Xrn1 digestion via qPCR and normalized to control samples without Xrn1 digestion. n = 6 (*upf1Δ* n = 4).

an obvious decapping defect compared to the *xrn1Δ* single mutant in which the PTC-reporter RNA was consistently removed almost completely throughout all repetitions. The *in vitro* Xrn1 digestion of *CBP80^{PTC}* indicates that decapping of this reporter is also defective in *gbp2Δ* cells (Fig. 4G), although Dcp1 recruitment was unaffected (Fig. 4D).

Presumably, this is the consequence of Gbp2's involvement in translation inhibition (Fig. 2E, F), as active translation initiation counteracts decapping [41].

Together, our findings suggest that Hrb1 functions in the NMD-induced 5′ degradation of PTC-containing mRNAs by promoting recruitment of Dcp1. Once de-capping is initiated,

Hrb1 probably leaves the PTC-containing transcript, while Gbp2 is still part of the Upf1-complex.

### Gbp2 and Hrb1 help to recruit the 3'-end degradation machinery

In addition to the major 5′ degradation pathway, the Ski-complex and the cytoplasmic exosome degrade NMD targets from the 3′-end [21,25,40]. Co-IPs with Ski2-GFP revealed physical interactions with both guard proteins, which persisted when RNase A was added (Fig. 5A, S5A). However, the interaction with Gbp2 was again decreased, suggesting that RNA binding enables protein interaction (Fig. 5A, S5B). That these interactions could be relevant for NMD is shown in the co-IP experiment between Ski2 and Upf1, where a ~ 70% decreased interaction between these proteins was observed when Gbp2 and Hrb1 were missing (Fig. 5B, C). To analyse whether Ski2 recruitment is promoted by the two guard proteins, we compared its binding to the $CBP80^{PTC}$ transcript in RIP-experiments. The absence of Upf1 resulted only in a ~ 30% decrease in the interaction of Ski2 with the NMD-target (Fig. 5D, S5C), which likely reflects the subordinate role of the 3′-mediated degradation of NMD targets [1,21]. Interestingly, in the absence of Gbp2 and Hrb1, the interaction of Ski2 with the NMD-target was more than 20% decreased, more than half of the effect in $upf1\Delta$, suggesting that the two guard proteins likely promote Ski complex recruitment. As Gbp2 is in close contact with Upf1 and shows an increased binding in stalled Upf-complexes, it might play a more important role in NMD-induced 3′-mediated mRNA degradation.

### Gbp2 and Hrb1 may help connect the 5'-end with the PTC

The discovered functions of Gbp2 and Hrb1 in translational repression and NMD-mediated degradation of the target RNAs occur at the ends of the transcripts, while detection of the PTC happens within the open reading frame. To communicate premature termination to the transcript ends, the pathway must be able to bridge this distance. In human cells Upf1 was already suggested to contact the 5′ end somehow [42]. However, so far it was not possible to get a clear picture. In order to investigate whether Gbp2 and Hrb1 could contribute to forming a higher ordered structure of the mRNA, we first checked if the two proteins can interact with each other. By using differently tagged guard proteins in co-IPs we were able to show that Gbp2 interacts with Hrb1 independently of RNA and both proteins interact with themselves (Fig. 6A, B). Secondly, we analysed their ability to contact the 5′ cap through interaction with the cap-binding proteins eIF4E and eIF4G. Co-IPs showed physical interactions of both cap-binders with Gbp2 and Hrb1 (Fig. 6C, S6A). As the guard proteins associate with both Upf1 and eIF4G, we tested if they would promote an interaction between these two proteins. With co-IP experiments, we could detect a physical interaction between eIF4G and Upf1, but it seemed not to be affected in the absence of Gbp2 and Hrb1 (Fig. S6B). We then over-expressed the $CBP80^{PTC}$ reporter to enhance NMD in the cells, and observed that the eIF4G-Upf1 interaction was

evidently reduced in $gbp2\Delta\ hrb1\Delta$ when RNase was added (Fig. 6D, E). This indicates that eIF4G and Upf1 probably bind to the same transcript independently of the two guard proteins, as shown by the unchanged co-purification without RNase treatment. However, their direct physical interaction is likely promoted by Gbp2 and Hrb1, as in the absence of RNA these proteins were less co-purified in $gbp2\Delta\ hrb1\Delta$ than in wild type.

The signals with RNase treatment were close to the detection limit. Therefore, the actual reduction might be smaller than suggested by the quantified values (Fig. 6E, rightmost bar). Nonetheless, a significant decrease of the interaction was evident, hinting at a possible role of Gbp2 and Hrb1 in transferring the information that a PTC was detected to the ends of the mRNA. Through interactions with each other and themselves, multiple copies of Gbp2 and Hrb1 at different positions on the RNA may promote formation of mRNP structures that bring proteins along the mRNA into spatial proximity.

Taken together, our findings indicate that Gbp2 and Hrb1 are involved in NMD. Similar to their guarding function in the nucleus, where they recruit the export receptor Mex67 upon successful splicing or, instead, the degradation machinery when splicing fails, they monitor gene expression also in the cytoplasm: From correct mRNAs, they dissociate during early translation [11], but in case of Upf1-mediated detection of a PTC, the guard proteins remain mRNA bound, promote repression of new rounds of translation and presumably the recruitment of degradation machineries (Fig. 7). We propose a model in which the guard proteins bridge the PTC-bound Upf-complex to the 5′-end of the transcript, thereby facilitating the information flow of the need for rapid translational repression and exonucleolytic degradation to the place of action. Thus, their guarding function continues in the cytoplasm after nuclear quality control.

## Discussion

### The splicing guard proteins are also cytoplasmic mRNA surveillance factors

Gbp2 and Hrb1 were identified as nuclear quality control factors [9]. Both guard proteins accompany the mRNAs into the cytoplasm and remain bound during translation [11,23,43], which might be relevant for the cytoplasmic surveillance system, similar to the EJC in humans, where the nuclear information from splicing is preserved in the cytoplasm. Indeed, after constructing intron-containing reporter genes, we could identify a role of Gbp2 and Hrb1 as auxiliary factors in NMD (Fig. 1E), which also appears to be relevant for endogenous NMD targets under natural physiological conditions (Figs. 1G, 3B, C, 4B, C and 5B, C). Since Gbp2 and Hrb1 are involved in the regulation of nuclear mRNA export, we could consider the possibility that the nuclear export of reporter RNAs is impaired. However, it was shown that Gbp2 and Hrb1 can retain RNAs in the nucleus but are no mRNA export factors, as their loss shows no mRNA export defects [8,23,43]. Moreover, we can see increased protein levels translated from the reporter constructs in $gbp2\Delta$

*hrb1Δ* cells (Fig. 2C-F), thus the reporters appear to be efficiently exported from the nucleus. Because Gbp2 and Hrb1 also affect degradation of mRNAs in the nucleus [9], we would not be able to identify cytoplasmic-specific effects from RNA half-life measurements. Therefore, we had to rely on steady-state RNA levels and relation to PTC-less control reporters initially to demonstrate effects that are specific to the cytoplasm and to NMD. Nevertheless, the *in vitro* Xrn1 digestion experiment showed clearly that indeed degradation of the reporter construct is defective in cells depleted of *GBP2* and *HRB1* (Fig. 4G).

The observation that Gbp2 and Hrb1 only affected the intron-containing reporters but not the (intron-less) *PGK1* reporter might be explained by the fact that Gbp2 and Hrb1 only stably bind to spliced transcripts. In yeast only 5% of all genes contain introns, but since many of them are highly expressed, such as genes encoding ribosomal proteins, 25% of all mRNAs are spliced [44]. Thus, intron-containing transcripts could contribute to a considerable portion of NMD targets. If Gbp2 and Hrb1 are indeed involved in NMD, specifically for spliced targets, this would also include correctly spliced transcripts when premature termination is caused by other means. It would, however, also be a failsafe mechanism to remove incorrectly spliced transcripts that escaped nuclear quality control. Previously shown severe sickness or lethality of *gbp2Δ hrb1Δ* cells when splicing is affected

[9] may be a consequence of the two proteins removing aberrant transcripts in the nucleus and the cytoplasm. That said, it is possible that Gbp2 and Hrb1 affect a subset of transcripts that is defined by other RNA features than splicing. Similarly, it was shown that Ebs1 and Nmd4, potential yeast homologs of human SMG5-7, have partial effects on NMD compared to Upf1 [20,45] and are presumably involved on a subset of targets. Further, cases of NMD have been reported that are independent of Upf2 and Upf3 [46,47], supporting the idea that as auxiliary factors, Gbp2 and Hrb1 may affect only a subgroup of NMD substrates. This also suggests that more players act in NMD and likely multiple factors together contribute to efficient NMD.

## Gbp2 and Hrb1 could be precursors of the EJC

We have shown earlier that the stable transcript association of Gbp2 and Hrb1 is a consequence of splicing [9]. Further, their rather 5′ proximal binding pattern of mRNAs correlates with the typical intron position in yeast [48,49]. Thus, these SR-like proteins might represent precursors of the human EJC. In fact, several human shuttling SR-proteins were reported to be part of the EJC [14]. Moreover, reports have demonstrated effects of this group of proteins on NMD, but the mechanisms are not understood. For example, overexpression of either SRSF1 or SRSF2 induces NMD [50]. Furthermore, SRSF1
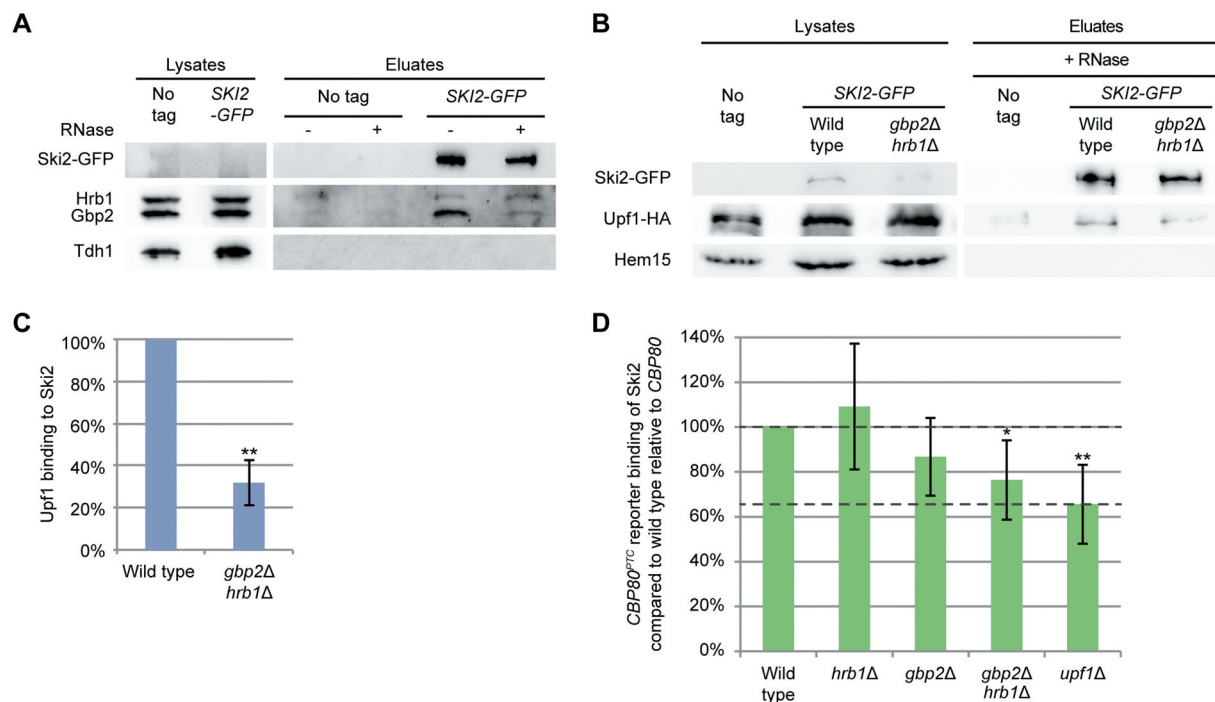


**Figure 5.** Gbp2 and Hrb1 are involved in the recruitment of the 3′-end degradation machinery. (A) Ski2 co-precipitates Gbp2 and Hrb1. Western blot of Gbp2 and Hrb1 co-IPs with Ski2-GFP is shown. Ski2-GFP was not detectable in the lysate. See also Fig S5B. (B) Proper interaction of Upf1 and Ski2 requires Gbp2 and Hrb1. Upf1-HA co-IPs with Ski2-GFP are shown on a western blot in the indicated strains. All cells are deleted for *UPF1*, and express p*UPF1-HA*. (C) The Ski2 and Upf1 interaction shown in (B) was quantified. Signal intensities of Upf1-HA bands were related to the corresponding Ski2-GFP pull-down signals from three independent co-IPs. (D) Gbp2 and Hrb1 promote the Ski2 interaction with the *CBP80^{PTC}* transcript. Ski2 RIP experiments and subsequent qPCRs with the PTC-containing reporter are shown in the indicated strains. All strains express genomic *SKI2-GFP*. n = 6 (*gbp2Δ* n = 8, *gbp2Δ hrb1Δ* n = 7). Dashed lines indicate the level of wild type and average level of *upf1Δ*. See also Fig S5C.
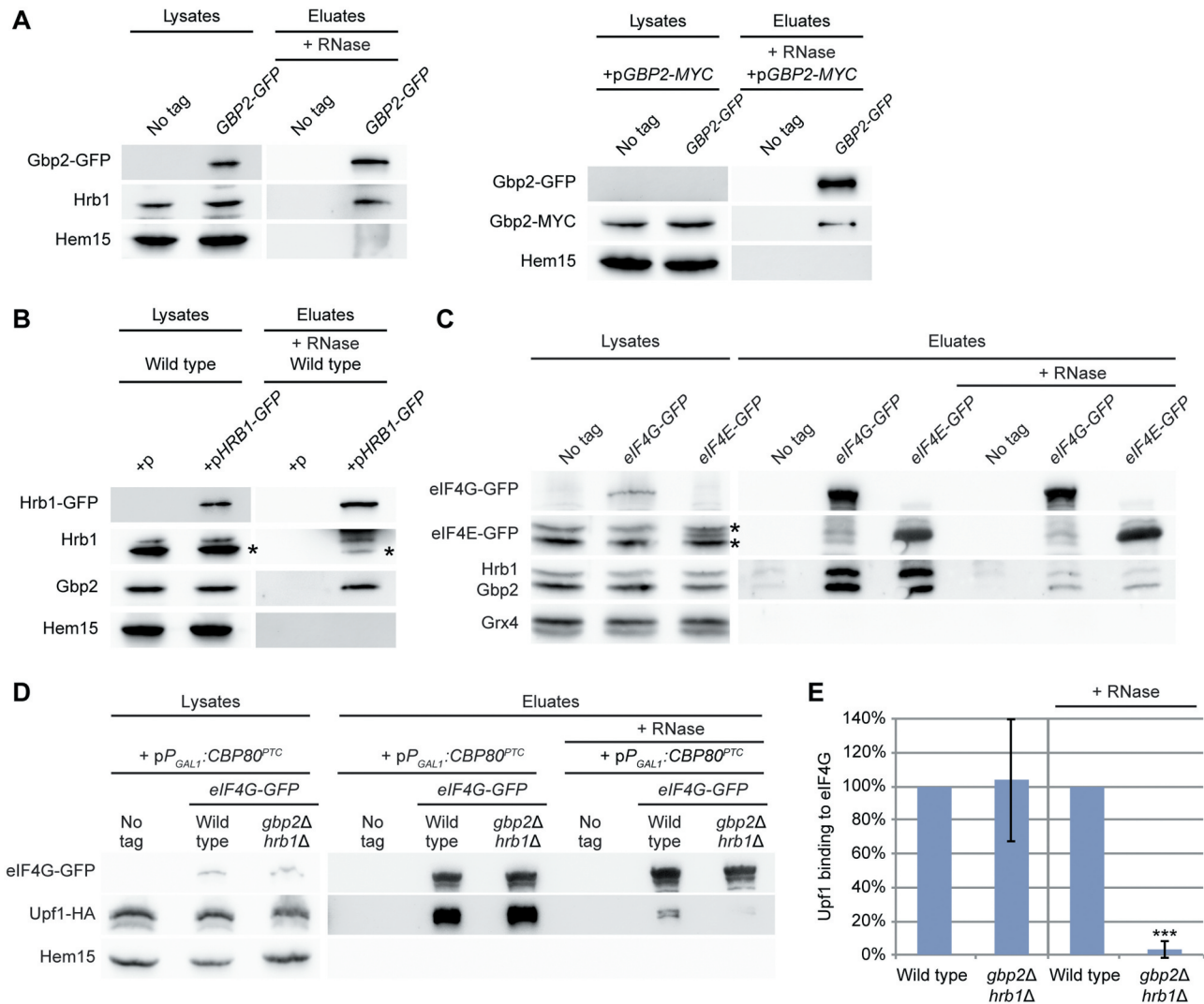
**Figure 6.** Gbp2 and Hrb1 might help to transmit the Upf1-mediated PTC alert to the 5′-end of the mRNA. (A) Gbp2 and (B) Hrb1 interact with each other and themselves. Co-IPs of differently tagged and untagged Gbp2 and Hrb1 versions upon RNase treatment are shown. Hem15 served as a negative control. Gbp2-GFP was not always detectable in the lysates. The asterisks indicate Gbp2 bands. (C) Gbp2 and Hrb1 interact with eIF4E and eIF4G. Co-IP of Gbp2 and Hrb1 with GFP tagged versions of the 5′ mRNA-binding proteins is shown. The asterisks indicate Hrb1 (top) and Gbp2 (bottom) bands. (D) The Upf1 interaction with eIF4G is significantly reduced in *gbp2Δ hrb1Δ* upon RNase treatment. Co-IP of Upf1 with eIF4G is shown in the indicated strains. p$P_{GAL1}$:$CBP80^{PTC}$ was induced for 2 h. All cells express p*UPF1-HA*. (E) Quantification of IP experiments shown in (D). Signal intensities of the Upf1-HA bands were related to the corresponding eIF4G-GFP pull-down signals. Upf1-HA signals without RNase treatment were quantified using less-exposed figures than shown in Fig. 6D. No RNase $n = 5$, + RNase $n = 3$. See also Fig S6.

was suggested to induce NMD indirectly by promoting translation [51,52], but also directly by contacting Upf1 [53]. Together, these findings from metazoans suggest that SR-proteins are involved in NMD, but their exact cytoplasmic functions remain rather nebulous. Also, up to date only the SR-proteins from yeast were described as nuclear guard proteins that prevent the leakage of faulty transcripts into the cytoplasm. But metazoan shuttling SR-proteins were also noticed as nuclear export factors, because they promote splicing and the subsequent recruitment of the Mex67 homolog TAP for nuclear export, similar to Gbp2 and Hrb1 [54,55]. Thus, yeast and human shuttling SR-proteins show many similarities in their behaviour and future studies are required to define roles of the human SR-proteins as potential nuclear guard proteins and specify their role as cytosolic NMD-

factors, either as part of the EJC or as additional and independent regulators of NMD.

## NMD-target degradation

We discovered a function of the guard proteins in degradation of NMD-targets, which is dependent on Upf1 (Fig. 1). In fact, both SR-proteins co-precipitated with all three Upf-proteins (Fig. 3). Endogenous NMD events, which are normally rare, measurably increased the association of Gbp2 with mutant upf1 (Fig. 3B, C), and this association could further be enhanced by the overexpression of an NMD substrate (Fig. 3D). As the interaction of Gbp2, but not Hrb1 increases in the presence of the stalled upf1-DE572AA complex and is detectable with the split GFP system, we suggest a direct physical
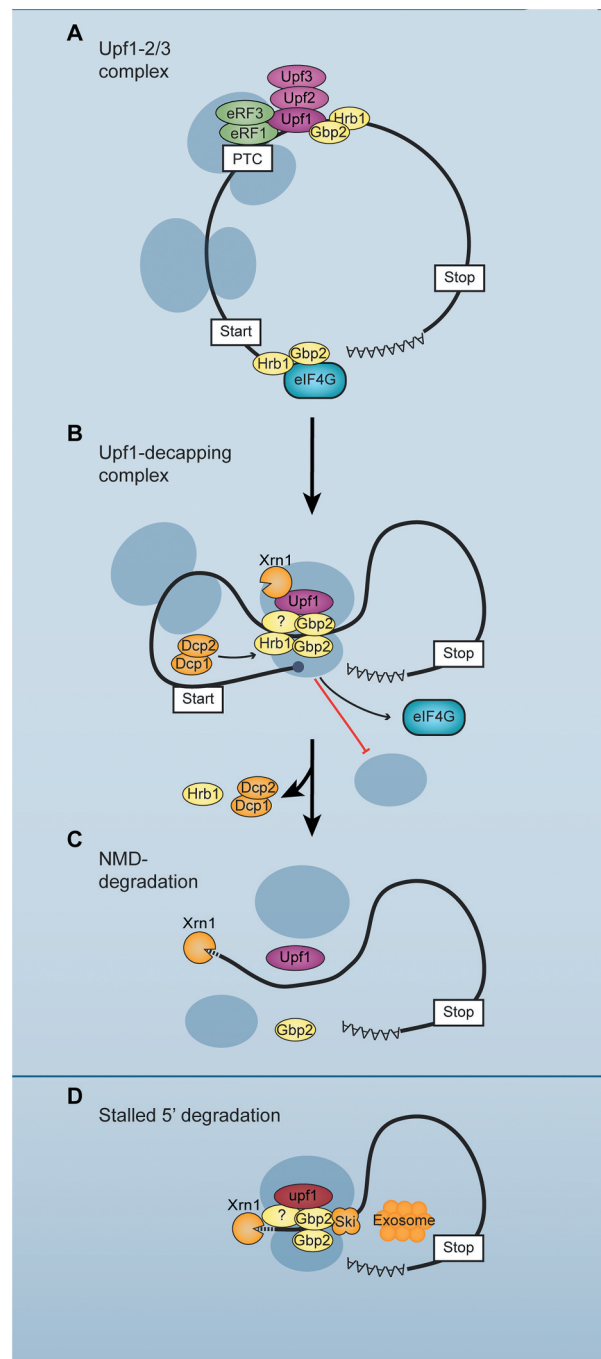
**Figure 7.** Model for the functions of the guard proteins Gbp2 and Hrb1 in NMD. (A) Gbp2 and Hrb1 are bound to the translated mRNA with a preference towards the 5' UTR, where the introns are located in yeast. Upf1 binds to a PTC and is joined by Upf2 and Upf3 forming the *Upf1-2/3 complex*. (B) Through interactions with themselves and possibly additional factors, the guard proteins help to restructure the RNP and transmit the PTC recognition from Upf1 to the 5' end of the mRNA, where they inhibit translation initiation. Hrb1 promotes Dcp1 recruitment to the *Upf1-decapping* complex. Decapping can occur after translation inhibition and dissociation of translation initiation factors. Xrn1 binds to the *Upf1-decapping* complex independently of Gbp2 and Hrb1. (C) After decapping, Xrn1 can degrade the mRNA. Hrb1 dissociates upon decapping or the onset of Xrn1-mediated degradation. Gbp2 dissociates when the helicase activity of Upf1 detaches the ribosome from the PTC. (D) In the minor 3'-5' degradation pathway, Gbp2 recruits the Ski complex to Upf1. The Ski complex facilitates degradation by the exosome.

contact between Upf1 and Gbp2 that is very transient under normal conditions. In this late NMD-complex, Hrb1 might already have fulfilled its function at the 5'-end and left the NMD-substrate. Similarly, some but not all human NMD factors are enriched in the mutant upf1-complex [18].

Such Gbp2- and Hrb1-containing NMD complexes, whose formation depends on Upf1, are further supported by the cytoplasmic localization of the usually nuclear guard proteins at steady state. In cells where NMD-substrates accumulate, such as in *xrn1Δ* cells, both guard proteins were enriched in the cytoplasm (Fig. 3E, F). Importantly, for this accumulation

Upf1 is required, which clearly indicates that cytoplasmic function of these guard proteins is linked to NMD. Further, we found physical interactions of these proteins with cytoplasmic degradation factors (Figs. 4A, E and 5A, S4B, S5B). Notably, the interactions of Hrb1 appear rather RNase resistant, while Gbp2 shows reduced interactions upon RNase treatment. This could suggest that Gbp2 can only fold properly to interact with the 5′ degradation machinery when bound to RNA. In line with that, a crosslinking reagent was required for a visible interaction between Dcp1 and Gbp2 with RNase treatment (Fig. 4A, S4B, S4C). Since co-precipitation of Gbp2 and Hrb1 was visible with the tested degradation factors upon RNase treatment, this suggests that Gbp2 and Hrb1 associate with the degrading complexes and are not simply present on the same RNA.

One reason why the NMD-targets are stabilized in the absence of the two guard proteins might be that the proteins help to recruit RNA degrading factors in the cytoplasm (Figs. 4 and 5), similar to their nuclear quality control function [9]. There, they recruit the nuclear Mtr4 protein, a part of the TRAMP-complex, which is a co-factor for the nuclear exosome. In the cytoplasm, Gbp2 and Hrb1 are required for the effective recruitment of Ski2, which is the cytoplasmic counterpart of Mtr4, a highly homologous RNA-helicase that is necessary for the exosomal RNA degradation [56]. Given that Gbp2 accumulated on PTC-containing transcripts on which 5′ degradation stalled due to the upf1-DE572AA mutant, Ski2 might only act after the ribosome is dissociated. Ribosome dissociation upon utilizing the ATPase activity of Upf1 might lead to rearrangements of the NMD complex and allow Gbp2 to promote Ski2-mediated degradation. Hrb1 also seems to be relevant, although it doesn't seem to accumulate on the 3′ degradation fragments in upf1-DE572AA (Fig. 3B, C).

For the main degradation pathway from the 5′-end [17,25,40], our results show that Hrb1 is required for efficient Dcp1 recruitment (Fig. 4D). Nevertheless, the *in vitro* Xrn1 digestion of $CBP80^{PTC}$ indicates that decapping of this reporter is also defective in $gbp2\Delta$ (Fig. 4G), although Dcp1 recruitment was unaffected (Fig. 4D). In addition to a potential, combined action of Gbp2 and Hrb1 in structuring the RNP, Gbp2 appears to promote access of the decapping enzyme to the cap concomitantly through its function in translation initiation inhibition. Such roles have also been suggested for the RGG proteins Sbp1 and Scd6 [35,57]. This is interesting, because it shows for the first time that the highly homologous guard proteins Gbp2 and Hrb1 affect the same pathway but do so via different mechanisms.

### Translational repression of NMD-substrates

For NMD it is not only important to degrade a faulty transcript, but also to repress new rounds of translation in order to prevent the expression of potentially toxic truncated proteins. Upf1 was shown to repress translation of NMD-targets [20,25,32] and we found that the guard proteins are also involved in the translational repression of NMD substrates that are intron-containing (Fig. 2). The proteins had no influence on translation when Upf1 was missing or no PTC was present, suggesting that this effect is NMD-specific (Fig. 2C–

H). The fact that NMD seems to have a much greater effect on the protein level than on the RNA level of the new NMD reporters might reflect the fact that the main function of this quality control pathway is to prevent the production of potentially harmful polypeptides. This makes the removal of the PTC-containing mRNA rather subordinate as long as the cell effectively prevents the protein production. That said, NMD was also described to function in regulation of RNA levels for certain targets apart from quality control [1,3]. There, regulation of the RNA stability per se is presumably the main function.

Both guard proteins were previously detected to be associated with polysomes [11]. Also, Gbp2 was found to accumulate in P-bodies, in which RNAs accumulate and are translationally repressed after starvation [58]. However, they have not been analysed for their potential to repress translation. Interestingly, both Gbp2 and Hrb1 contain arginine, glycine, glycine (RGG)-repeat motifs that have the potential to inhibit translation initiation. Other RGG-containing proteins, Scd6, Sbp1 and Npl3, were shown to interact with eIF4G via the RGG-motif and inhibit translation *in vivo* and *in vitro* [11,34,35]. Gbp2 and Hrb1 also interact with the cap-binder eIF4E and its interacting scaffolding protein eIF4G (Fig. 6C), but in contrast to Npl3, Sbp1 and Scd6, they appear to specifically be involved in the translation of NMD-substrates (Fig. 2C, E, I, J), suggesting that Gbp2 and Hrb1 are potentially specific translational repressors of their bound NMD-targets.

### Gbp2 and Hrb1 transmit the PTC-recognition alert to the transcript ends

How the Upf-proteins, bound to the PTC, communicate to the ends of the transcripts that translation on this mRNA should be suppressed and degradation initiated was unclear. At least Gbp2 gets into close proximity with PTC-bound Upf1 (Fig. 3D) and both Gbp2 and Hrb1 associate with the 5′- and 3′-degradation machineries (Figs. 4A, E and 5A) as well as the cap-binding eIF4E and eIF4G (Fig. 6C). Furthermore, the proteins interact with each other and themselves (Fig. 6A, B). These characteristics make them excellent candidates for establishing contact between the PTC-bound Upf-proteins and the 5′ end of the transcript. RNA commonly folds into variable secondary structures and restructuring of mRNA promoted by protein-protein interactions has also been demonstrated previously [59]. By such RNP complex rearrangements the alert for PTC-recognition could be transmitted to the 5′-end, where the consequential repression of translation initiation and mRNA degradation are executed. We found indeed a significant reduction of the RNA-independent interaction between Upf1 and eIF4G in the absence of the guard proteins, which supports our model that Gbp2 and Hrb1 mediate the connection of the PTC with the 5′-end of the transcript, thereby bringing the PTC-alert to the site where further action is required (Fig. 6D, E).

Taken together, we have identified the nuclear splicing guard proteins Gbp2 and Hrb1 as auxiliary NMD-factors for intron-containing transcripts. Upon detection of a PTC by

Upf1, they seem to be involved in directing this information to the ends of the transcript, translational repression and degradation of the faulty RNA (Fig. 7). Their splicing-mediated binding to transcripts appears analogous to the loading of EJCs in higher eukaryotes and it is tempting to speculate that they might be the yeast counterpart or precursor of the EJC. Most importantly, to date human SR-proteins have not been in the focus of nuclear and cytoplasmic mRNA quality control. However, due to the fact that these proteins are mutated in many neurodegenerative diseases and cancer (http://www.cbioportal.org/), further understanding of their functions in human would provide valuable knowledge for the future. In particular, human SR-proteins are bona fide splicing factors, which can indirectly affect NMD, and the expression of some SR-proteins is auto-regulated via the NMD pathway [60–64], making it complicated to sort out the function of these proteins in mRNA quality control. The identification of the yeast SR-proteins Gbp2 and Hrb1 not only as nuclear but also cytoplasmic quality control factors, required for the degradation and translational repression of PTC-containing transcripts and connecting both surveillance mechanisms in the cell, offers new perspectives for the understanding of human SR-proteins and related diseases.

## Disclosure statement

The Authors declare no competing interests.

## ORCID

Yen-Yun Lu 🔟 http://orcid.org/0000-0002-6818-9028

## References

[1] He F, Jacobson A. Nonsense-Mediated mRNA Decay: degradation of Defective Transcripts Is Only Part of the Story. Annu Rev Genet. 2015;49(1):339–366.

[2] Inada T. The Ribosome as a Platform for mRNA and Nascent Polypeptide Quality Control. Trends Biochem Sci. 2017;42(1):5–15.

[3] Kurosaki T, Popp MW, Maquat LE. Quality and quantity control of gene expression by nonsense-mediated mRNA decay. Nat Rev. 2019;20:406–420.

[4] Soheilypour M, Mofrad MRK. Quality control of mRNAs at the entry of the nuclear pore: cooperation in a complex molecular system. Nucleus. 2018;9(1):202–211.

[5] Zander G, Krebber H. Quick or quality? How mRNA escapes nuclear quality control during stress. RNA Biol. 2017;14 (12):1642–1648.

[6] Amrani N, Ghosh S, Mangus DA, et al. Translation factors promote the formation of two states of the closed-loop mRNP. Nature. 2008;453(7199):1276–1280.

[7] Doma MK, Parker R. RNA quality control in eukaryotes. Cell. 2007;131(4):660–668.

[8] Zander G, Hackmann A, Bender L, et al. mRNA quality control is bypassed for immediate export of stress-responsive transcripts. Nature. 2016;540(7634):593–596.

[9] Hackmann A, Wu H, Schneider UM, et al. Quality control of spliced mRNAs requires the shuttling SR proteins Gbp2 and Hrb1. Nat Commun. 2014;5(1):3123.

[10] Galy V, Gadal O, Fromont-Racine M, et al. Nuclear retention of unspliced mRNAs in yeast is mediated by perinuclear Mlp1. Cell. 2004;116(1):63–73.

[11] Windgassen M, Sturm D, Cajigas IJ, et al. Yeast shuttling SR proteins Npl3p, Gbp2p, and Hrb1p are part of the translating mRNPs, and Npl3p can function as a translational repressor. Mol Cell Biol. 2004;24(23):10479–10491.

[12] Shoemaker CJ, Green R. Translation drives mRNA quality control. Nat Struct Mol Biol. 2012;19(6):594–601.

[13] Karousis ED, Muhlemann O. Nonsense-Mediated mRNA Decay Begins Where Translation Ends. Cold Spring Harb Perspect Biol. 2019;11(2):a032862.

[14] Singh G, Kucukural A, Cenik C, et al. The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. Cell. 2012;151(4):750–764.

[15] Gonzalez CI, Ruiz-Echevarria MJ, Vasudevan S, et al. The yeast hnRNP-like protein Hrp1/Nab4 marks a transcript for nonsense-mediated mRNA decay. Mol Cell. 2000;5 (3):489–499.

[16] Amrani N, Ganesan R, Kervestin S, et al. A faux 3′-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. Nature. 2004;432(7013):112–118.

[17] Muhlrad D, Parker R. Recognition of yeast mRNAs as "nonsense containing" leads to both inhibition of mRNA translation and mRNA degradation: implications for the control of mRNA decapping. Mol Biol Cell. 1999;10:3971–3978.

[18] Franks TM, Singh G, Lykke-Andersen J. Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense-mediated mRNA decay. Cell. 2010;143(6):938–950.

[19] Serdar LD, Whiteside DL, Baker KE. ATP hydrolysis by UPF1 is required for efficient translation termination at premature stop codons. Nat Commun. 2016;7(1):14021.

[20] Dehecq M, Decourty L, Namane A, et al. Nonsense-mediated mRNA decay involves two distinct Upf1-bound complexes. EMBO J. 2018;37(21):e99278.

[21] Mitchell P, Tollervey D. An NMD pathway in yeast involving accelerated deadenylation and exosome-mediated 3′–>5′ degradation. Mol Cell. 2003;11(5):1405–1413.

[22] Unterholzner L, Izaurralde E. SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. Mol Cell. 2004;16(4):587–596.

[23] Windgassen M, Krebber H. Identification of Gbp2 as a novel poly(A)+ RNA-binding protein involved in the cytoplasmic delivery of messenger RNAs in yeast. EMBO Rep. 2003;4(3):278–283.

[24] Jamar NH, Kritsiligkou P, Grant CM. Loss of mRNA surveillance pathways results in widespread protein aggregation. Sci Rep. 2018;8(1):3894.

[25] Muhlrad D, Parker R. Premature translational termination triggers mRNA decapping. Nature. 1994;370(6490):578–581.

[26] Johansson MJ, He F, Spatrick P, et al. Association of yeast Upf1p with direct substrates of the NMD pathway. Proc Natl Acad Sci U S A. 2007;104(52):20872–20877.

[27] Kurosaki T, Li W, Hoque M, et al. A post-translational regulatory switch on UPF1 controls targeted mRNA degradation. Genes Dev. 2014;28(17):1900–1916.

[28] Kuroha K, Tatematsu T, Inada T. Upf1 stimulates degradation of the product derived from aberrant messenger RNA containing a specific nonsense mutation by the proteasome. EMBO Rep. 2009;10(11):1265–1271.

[29] Keeling KM, Bedwell DM. Suppression of nonsense mutations as a therapeutic approach to treat genetic diseases. Wiley Interdisciplinary Reviews: RNA. 2011;2(6):837–852.

[30] Maderazo AB, Belk JP, He F, et al. Nonsense-containing mRNAs that accumulate in the absence of a functional nonsense-mediated mRNA decay pathway are destabilized rapidly upon its restitution. Mol Cell Biol. 2003;23(3):842–851.

[31] Hoek TA, Khuperkar D, Lindeboom RGH, et al. Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay. Mol Cell. 2019;75(2):324–39 e11.

[32] Kim WK, Yun S, Kwon Y, et al. mRNAs containing NMD-competent premature termination codons are stabilized and translated under UPF1 depletion. Sci Rep. 2017;7(1):15833.

[33] Baierlein C, Hackmann A, Gross T, et al. Monosome formation during translation initiation requires the serine/arginine-rich protein Npl3. Mol Cell Biol. 2013;33(24):4811–4823.

[34] Rajyaguru P, She M, Parker R. Scd6 targets eIF4G to repress translation: RGG motif proteins as a class of eIF4G-binding proteins. Mol Cell. 2012;45(2):244–254.

[35] Segal SP, Dunckley T, Parker R. Sbp1p affects translational repression and decapping in Saccharomyces cerevisiae. Mol Cell Biol. 2006;26(13):5120–5130.

[36] Weng Y, Czaplinski K, Peltz SW. Genetic and biochemical characterization of mutations in the ATPase and helicase regions of the Upf1 protein. Mol Cell Biol. 1996;16(10):5477–5490.

[37] Estrella LA, Wilkinson MF, Gonzalez CI. The Shuttling Protein Npl3 Promotes Translation Termination Accuracy in Saccharomyces cerevisiae. J Mol Biol. 2009;394(3):410–422.

[38] Leeds P, Wood JM, Lee BS, et al. Gene products that promote mRNA turnover in Saccharomyces cerevisiae. Mol Cell Biol. 1992;12(5):2165–2177.

[39] Magliery TJ, Wilson CG, Pan W, et al. Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: scope and mechanism. J Am Chem Soc. 2005;127 (1):146–157.

[40] He F, Jacobson A. Upf1p, Nmd2p, and Upf3p regulate the decapping and exonucleolytic degradation of both nonsense-containing mRNAs and wild-type mRNAs. Mol Cell Biol. 2001;21 (5):1515–1530.

[41] Coller J, Parker R. Eukaryotic mRNA decapping. Annu Rev Biochem. 2004;73(1):861–890.

[42] Maquat LE, Hwang J, Sato H, et al. CBP80-promoted mRNP rearrangements during the pioneer round of translation, nonsense-mediated mRNA decay, and thereafter. Cold Spring Harb Symp Quant Biol. 2010;75:127–134.

[43] Hacker S, Krebber H. Differential Export Requirements for Shuttling Serine/Arginine-type mRNA-binding Proteins. J Biol Chem. 2004;279(7):5049–5052.

[44] Ares M Jr., Grate L, Pauling MH. A handful of intron-containing genes produces the lion's share of yeast mRNA. RNA. 1999;5 (9):1138–1139. (New York, NY).

[45] Luke B, Azzalin CM, Hug N, et al. Saccharomyces cerevisiae Ebs1p is a putative ortholog of human Smg7 and promotes nonsense-mediated mRNA decay. Nucleic Acids Res. 2007;35 (22):7688–7697.

[46] Chan WK, Huang L, Gudikote JP, et al. An alternative branch of the nonsense-mediated decay pathway. EMBO J. 2007;26 (7):1820–1830.

[47] Gehring NH, Kunz JB, Neu-Yilik G, et al. Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements. Mol Cell. 2005;20 (1):65–75.

[48] Baejen C, Torkler P, Gressel S, et al. Transcriptome Maps of mRNP Biogenesis Factors Define Pre-mRNA Recognition. Mol Cell. 2014;55(5):745–757.

[49] Tuck AC, Tollervey D. A Transcriptome-wide Atlas of RNP Composition Reveals Diverse Classes of mRNAs and lncRNAs. Cell. 2013;154(5):996–1009.

[50] Zhang Z, Krainer AR. Involvement of SR proteins in mRNA surveillance. Mol Cell. 2004;16(4):597–607.

[51] Sanford JR, Gray NK, Beckmann K, et al. A novel role for shuttling SR proteins in mRNA translation. Genes Dev. 2004;18 (7):755–768.

[52] Sato H, Hosoda N, Maquat LE. Efficiency of the pioneer round of translation affects the cellular site of nonsense-mediated mRNA decay. Mol Cell. 2008;29(2):255–262.

[53] Aznarez I, Nomakuchi TT, Tetenbaum-Novatt J, et al. Mechanism of Nonsense-Mediated mRNA Decay Stimulation by Splicing Factor SRSF1. Cell Rep. 2018;23(7):2186–2198.

[54] Huang Y, Steitz JA. SRprises along a messenger's journey. Mol Cell. 2005;17(5):613–615.

[55] Müller-McNicoll M, Botti V, de Jesus Domingues AM, et al. SR proteins are NXF1 adaptors that link alternative RNA processing to mRNA export. Genes Dev. 2016;30(5):553–566.

[56] Johnson SJ, Jackson RN. Ski2-like RNA helicase structures: common themes and complex assemblies. RNA Biol. 2013;10 (1):33–43.

[57] Nissan T, Rajyaguru P, She M, et al. Decapping activators in Saccharomyces cerevisiae act by multiple mechanisms. Mol Cell. 2010;39(5):773–783.

[58] Buchan JR, Muhlrad D, Parker R. P bodies promote stress granule assembly in Saccharomyces cerevisiae. J Cell Biol. 2008;183 (3):441–455.

[59] Aibara S, Gordon JM, Riesterer AS, et al. Structural basis for the dimerization of Nab2 generated by RNA binding provides insight into its contribution to both poly(A) tail length determination and transcript compaction in Saccharomyces cerevisiae. Nucleic Acids Res. 2017;45(3):1529–1538.

[60] Lejeune F, Maquat LE. Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. Curr Opin Cell Biol. 2005;17(3):309–315.

[61] Sikorski RS, Hieter P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics. 1989;122:19–27.

[62] Becker D, Hirsch AG, Bender L, et al. Nuclear Pre-snRNA Export Is an Essential Quality Assurance Mechanism for Functional Spliceosomes. Cell Rep. 2019;27(11):3199–214 e3.

[63] Lee RE, Brunette S, Puente LG, et al. Metacaspase Yca1 is required for clearance of insoluble protein aggregates. Proc Natl Acad Sci U S A. 2010;107(30):13348–13353.

[64] Fabrizio P, Dannenberg J, Dube P, et al. The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome. Mol Cell. 2009;36(4):593–608.

# Evolution of intron splicing towards optimized gene expression is based on various *Cis*- and *Trans*-molecular mechanisms

Idan Frumkin[1]*, Ido Yofe[1], Raz Bar-Ziv[1], Yonat Gurvich[1], Yen-Yun Lu[2], Yoav Voichek[1], Ruth Towers[1], Dvir Schirman[1], Heike Krebber[2], Yitzhak Pilpel[1]*

**1** Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel, **2** Abteilung für Molekulare Genetik, Institut für Mikrobiologie und Genetik, Göttinger Zentrum für Molekulare Biowissenschaften (GZMB), Georg-August Universität Göttingen, Göttingen, Germany

ⓧ These authors contributed equally to this work.

* frumkin.idan@gmail.com (IF); pilpel@weizmann.ac.il (YP)

## Abstract

Splicing expands, reshapes, and regulates the transcriptome of eukaryotic organisms. Despite its importance, key questions remain unanswered, including the following: Can splicing evolve when organisms adapt to new challenges? How does evolution optimize inefficiency of introns' splicing and of the splicing machinery? To explore these questions, we evolved yeast cells that were engineered to contain an inefficiently spliced intron inside a gene whose protein product was under selection for an increased expression level. We identified a combination of mutations in *Cis* (within the gene of interest) and in *Trans* (in mRNA-maturation machinery). Surprisingly, the mutations in *Cis* resided outside of known intronic functional sites and improved the intron's splicing efficiency potentially by easing tight mRNA structures. One of these mutations hampered a protein's domain that was not under selection, demonstrating the evolutionary flexibility of multi-domain proteins as one domain functionality was improved at the expense of the other domain. The *Trans* adaptations resided in two proteins, Npl3 and Gbp2, that bind pre-mRNAs and are central to their maturation. Interestingly, these mutations either increased or decreased the affinity of these proteins to mRNA, presumably allowing faster spliceosome recruitment or increased time before degradation of the pre-mRNAs, respectively. Altogether, our work reveals various mechanistic pathways toward optimizations of intron splicing to ultimately adapt gene expression patterns to novel demands.

## Introduction

Throughout evolution, cells acquired regulatory mechanisms to tune gene expression, which have been the subject of intensive investigations—focusing mainly on transcription and translation. Among other known mechanisms, when cells are challenged to increase protein expression levels, the DNA sequence of genes can change so as to increase transcription [1,2], support more efficient mRNA translation [3,4], or result in greater mRNA transcript stability

[5,6]. Additionally, the transcription and translation machineries themselves have been shown to adapt to environmental challenges by altering the cellular pools of transcription factors [7] or tRNAs [8,9].

In evolving expression programs, adaptation often occurs either directly on the genes under pressure ("evolution in *Cis*") [10] or indirectly, e.g., on the expression machineries, typically transcription and translation ("evolution in *Trans*")[11,12]. These two routes of evolution are profoundly different [13], as the first (*Cis*) provides a localized solution that in principle can affect only a certain gene, while the later (*Trans*) could be the method of choice if a coordinated change in many genes is needed.

Surprisingly, although the process of splicing is central to the maturation and regulation of mRNAs in eukaryotes [14–18], its role in adapting to novel demands on gene expression has not been thoroughly investigated. During mRNA splicing, precursor mRNAs are processed to remove introns while fusing exons together to create the mature transcript. This process can provide evolutionary means to diversify the proteome towards phenotypic novelty, as the choice of intron to be excluded, as well as the exons which are found in the mature transcript, can both be regulated based on the cell's needs [16,19,20]. An aspect of splicing evolution that has been extensively studied is gain and loss of introns, for which several molecular models have been proposed, mainly reverse transcription and recombination-mediated intron loss, intron transposition, and also exonization and intronization via mutations [21–25].

While intron loss and gain have been demonstrated experimentally [26,27], other forms of evolution through changes in splicing, such as alterations in splicing efficiency under changing conditions, have not. Adaptation of splicing efficiency is presumably essential to cellular evolution given a recent finding that splicing efficiency increases with transcription rate [28], therefore making it likely that splicing efficiency of introns is under constant selection during evolution. Yet, the mechanisms that allow this adaptation are unknown.

Here, we set to reveal whether introns or the splicing apparatus can evolve so as to alter the expression levels of genes in an adaptive manner. To this end, we engineered yeast cells to express a reporter gene, to which we inserted an inefficiently spliced intron that was fused to an antibiotic resistance gene. We then carried out a lab-evolution experiment in which cells were exposed to the drug and followed their adaptation.

Our results demonstrate that adaptations were related to splicing, and they appear to have not addressed directly the transcription or translation of the gene under selection. Two alternative adaptive routes for evolution of splicing were observed. First, we found *Cis*-acting solutions in the form of adaptive mutations that occurred in the intron itself but also, surprisingly, in an upstream exon. These mutations resulted in increased splicing efficiency and higher expression levels of the antibiotic resistance gene. We then show how one such *Cis* mutation alters the predicted RNA structure of the intron to better support splicing.

Yet, in some other evolved cells there were no mutations in *Cis*, i.e., in the gene or in its surrounding regions, but rather *Trans*-acting adaptations that have increased cellular availability of the splicing machinery. Sequencing the genomes of *Trans*-evolved colonies revealed nonsynonymous mutations in the RNA recognition motifs of two SR-like proteins that are known to have a diverse set of cellular functions related to RNA splicing and maturation. In particular, SR-like proteins were shown to support splicing by co-transcriptional recruitment of splicing factors [29–32] and were also shown to be involved in quality control of nascent mRNAs by selectively exporting from the nucleus spliced mRNAs upon completion of splicing [33,34]. Here, we show that adaptations in *Trans* that occurred through this experiment have modified the affinity of these proteins to the transcript under selection in a way that could allow its more efficient splicing.

## Results

### Low splicing efficiency of a drug resistance gene leads to stressed cells in presence of antibiotics

We hypothesized that splicing efficiency of genes could serve as a means to optimize their expression levels. To test this hypothesis, we used the yeast *Saccharomyces cerevisiae*, in which approximately 30% of the transcriptome is spliced to form mature mRNAs [35] at a range of splicing efficiencies [18,36]. We built a construct that consists of two fused domains. The first is a fluorescent reporter (YFP) that was engineered to include one of two alternative natural introns—with either high or low splicing efficiency. The intron was localized near the YFP's fluorescence site [36]. Downstream to this protein was fused an antibiotics resistance gene (kanamycin resistance gene [kan]). This general design consisted of three alternative strains: (i) "Control" with a YFP-Kan construct without an intron; (ii) "Splicing[High]" with a YFP-Kan gene that harbors the natural intron of *OSH7* that was previously reported to have high splicing efficiency within this YFP context [36]; and (iii) "Splicing[Low]" with a YFP-Kan gene that harbors the natural intron of *RPS26B*, with a low splicing efficiency [36] (see Fig 1A and S1 Table for a list of strains used in this study).

We first hypothesized that cellular growth of each strain in the presence of the antibiotic G418 will depend on YFP-Kan expression levels. We followed the growth of the three strains in the presence of the antibiotics and found that Control cells had the highest fitness, Splicing[High] grew slower, and Splicing[Low] demonstrated a severe growth defect compared with the two other strains (Fig 1B and 1C). We also measured fluorescence intensity of the YFP-Kan reporter in the presence of the drug and observed that Control cells demonstrated the highest fluorescence levels, followed by Splicing[High], and with Splicing[Low] cells showing the lowest YFP-Kan levels (Fig 1D). These results demonstrate that the inefficiently spliced intron in Splicing[Low] reduces cellular levels of YFP-Kan and hence, presumably, leads to a reduced fitness.

Because YFP-Kan expression levels in Splicing[Low] were significantly lower compared with the other strains, we hypothesized that Splicing[Low] cells did not reach the needed concentration of the resistance protein to sufficiently neutralize the antibiotics, and hence resulted in stressed cells. To test this hypothesis, we performed mRNA sequencing of exponentially growing Control and Splicing[Low] cells in an antibiotics-containing medium and analyzed the transcriptome profiles of these cells. Indeed, we observed that ribosomal genes were down-regulated in Splicing[Low] compared with Control cells—a clear signature of stressed cells [37] (Fig 1E). Notably, we observed an averaged 8% reduction for mRNA levels of ribosomal proteins between Control and Splicing[Low] cells and, correspondingly, an 18% reduction in growth rate. Interestingly, this is consistent with the correlation observed in a recent study between growth rate and ribosomal expression levels in yeast cells [38]. In parallel, stress-related genes [39] were up-regulated in Splicing[Low] cells compared with Control cells (Fig 1E). We thus concluded that the general stress response was activated in Splicing[Low] cells.

### Rapid evolutionary adaptation increases expression level of the resistance gene

Our experimental system mimics an evolutionary scenario in which there is an immediate and continuous selection pressure to up-regulate the expression level of a specific gene in a particular environment. How would the system evolve to better resist the antibiotics? Possible means to adapt include mutations in the gene's promoter to increase transcription, mutations that increase translation initiation or efficiency, or mutations inside the gene itself that could increase the specific activity of the protein (Fig 1A). Additionally, the splicing machinery may
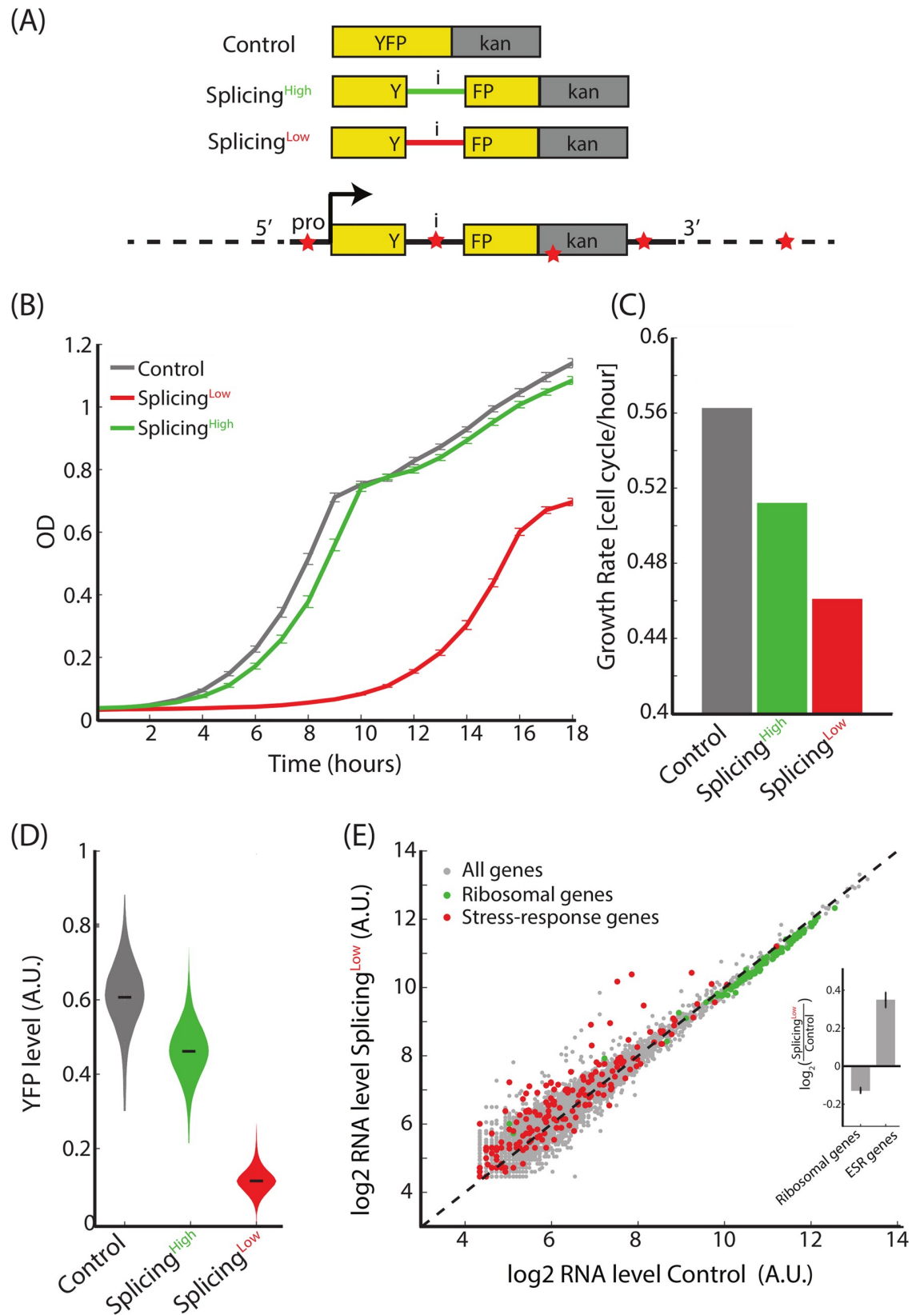
**Fig 1. Inefficient intron splicing leads to lower gene expression levels and compromised antibiotics resistance. (A)** We introduced two alternative introns into a YFP domain that was fused to a kanamycin resistance domain, to generate three strains: (i) Control without an intron; (ii) Splicing^High with an efficiently spliced intron; and (iii) Splicing^Low with an inefficiently spliced intron. Evolving cells at the presence of the antibiotics could adapt by mutating different parts of the YFP-Kan construct (evolution in *Cis*) or other loci, evolution in *Trans* (red stars represent potential locations of such putative mutation sites). **(B,C)** Splicing^Low suffers from a severe growth defect compared with Control or Splicing^High cells when the antibiotic is supplemented to the medium. The growth defect is manifested as both a longer lag phase and a lower maximal growth rate. **(D)** Florescence intensity of the YFP-Kan reporter for all three strains shows that Splicing^Low cells have lower expression levels of YFP-Kan. This observation links between YFP-Kan expression levels and cellular fitness. **(E)** Transcriptome profiling shows that ribosomal genes were down-regulated (green dots, $p = 4.62 \times 10^{-26}$, paired *t* test) and stress-response genes were up-regulated (red dots, $p = 3.40 \times 10^{-5}$, paired *t* test) in Splicing^Low compared with Control cells. This observation suggests that Splicing^Low cells experience stress because of compromised resistance to the antibiotics and that the general stress response was activated in them. **(Inset)** Mean $\log_2$ ratio of ribosomal and ESR gene groups. See numerical data for this figure in S1 Data.

https://doi.org/10.1371/journal.pbio.3000423.g001

also take part in adaptation of gene expression levels. To find which evolutionary tracks are used by cells as they adapt, we evolved the three strains by daily serial dilution on a medium supplemented with G418 for approximately 560 generations, in four independent cultures for each strain. Interestingly, only the cultures of Splicing^Low cells demonstrated a significant improvement in fitness when grown under the drug at the end of the experiment (Fig 2A and 2B). This observation suggests that only Splicing^Low experienced a sufficiently strong selective pressure to adapt to the presence of the antibiotics in the medium, in contrast to the Control and Splicing^High strains, which originally had much higher levels of the resistance proteins.

Consistent with the fitness measurements, YFP measurements of the evolved cultures showed that expression levels of the YFP-Kan fusion gene increased in all four evolved cultures of Splicing^Low compared with the ancestral strain (Fig 2C). Conversely, the increase in YFP-Kan expression levels in the evolved Control and Splicing^High populations was significantly smaller (Fig 2C). These results further indicate that Splicing^Low cells experienced the strongest selective pressure to adapt rapidly to the presence of the antibiotics, and that they achieved this goal by increasing the levels of the YFP-Kan reporter. We next moved to reveal the molecular mechanisms underlying this evolutionary adaptation.

## Adaptations in both *Cis* and *Trans* lead to increased splicing efficiency

We hypothesized that improving the low splicing efficiency of the intron in Splicing^Low could be natural selection's means to adapt towards increasing the resistance gene expression levels. We therefore sequenced the YFP-Kan locus, covering the entire gene from promoter to terminator, in 16 colonies from two evolved populations (termed here population A and population B) of Splicing^Low. Interestingly, we found that the colonies were split into two types—either with or without a mutation in the YFP-Kan locus (Fig 3A). In population A, we found that the same mutation occurred in four out of eight colonies, changing adenine to cytosine inside the intron, 97 nucleotides upstream to its 3′ end (Fig 3B). In population B, we identified an exonic nonsynonymous mutation that changed a thymine to cytosine 14 nucleotides upstream of the intron (V61A change in the YFP protein) in three out of eight colonies. In five other colonies from this population there were no mutations in the YFP-Kan locus.

Notably, none of the colonies demonstrated a mutation in the construct's promoter, terminator, or in the sequence of the Kan resistance gene itself. These results propose that different mutations in the intron, or its vicinity, were adaptive and might affect splicing efficiency of the intron. Surprisingly, the observed mutations did not occur in the 5′ donor, 3′ acceptor, nor in the intron branch point—suggesting that other positions of the intron can also be functional by affecting splicing efficiency.

While the intron- and exon-mutated colonies represent an evolutionary adaptation in *Cis*, the colonies that showed no mutation in the entire gene construct, that coexist with the *Cis*-
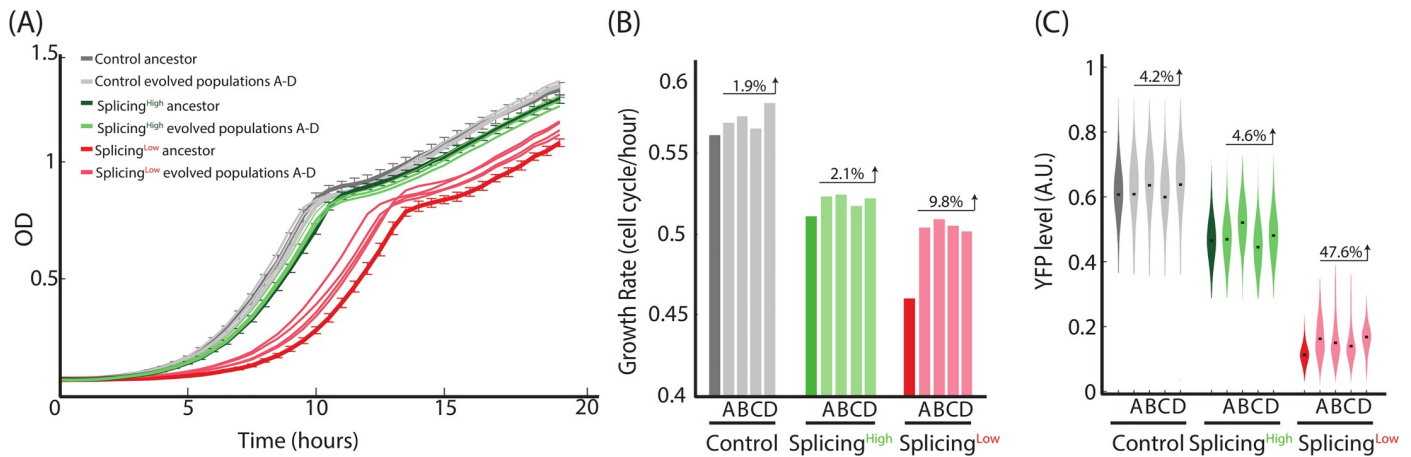
**Fig 2. Rapid adaptation to the presence of the antibiotics is observed only for Splicing^Low cells. (A,B)** We evolved Control, Splicing^High, and Splicing^Low cells for approximately 560 generations with the presence of the antibiotics in four independent cultures for each strain. Growth measurements of evolved populations compared with the three ancestors shows that only evolved Splicing^Low cells demonstrate significant improvement in growth for all four independent evolution lines. The number above each group of evolved populations represents the average improvement in growth rate compared with these populations' ancestor strains. These observations suggest that the inefficiently spliced intron led to a rapid adaptation of Splicing^Low cells. **(C)** Fluorescence intensities of the YFP-Kan reporter for all evolved cultures show that expression levels were much increased in all four evolved cultures of Splicing^Low compared with the ancestral strain (effect sizes = 78.67, 79.54, 75.17, 83.19). Conversely, the increase in expression levels in the evolved Control and Splicing^High populations were smaller (Control effect sizes = 64.66, 68.44, 63.51, 67.74; Splicing^High effect sizes = 54.33, 70.66, 52.43, and 58.27). The number above each group of evolved populations represents the average increase in YFP-Kan levels compared with these populations' ancestor strains. These observations suggest that adaptation of Splicing^Low cells was based on their ability to increase expression levels of the resistance proteins. See numerical data for this figure in S1 Data.

https://doi.org/10.1371/journal.pbio.3000423.g002

evolved colonies in the same populations, potentially found adaptive solutions in *Trans* that may have occurred elsewhere in the genome. We thus randomly chose six colonies: four colonies with a *Cis* mutation and two colonies that showed no mutations in *Cis*, as we reasoned that such colonies may have adapted in *Trans*. We termed these colonies according to the evolution lines from which they were derived: A-cis1, A-cis2, B-cis1, B-cis2, A-trans, and B-trans. We followed the growth of these evolved colonies in the presence of G418 and found, as expected, that all grew faster than the Splicing^Low ancestor (Fig 3C).

We then performed mRNA sequencing and transcriptome analyses of all colonies and found that they indeed demonstrate relaxation of the stress state that was featured in the ancestor. In particular, the general stress response genes were reduced relative to their high levels in the Splicing^Low ancestor, and ribosomal proteins were up-regulated relative to their low levels in this ancestral strain (Fig 3D for colony A-cis1). These dynamics were shared by all the *Cis*-evolved colonies (Fig 3E), demonstrating the robustness of this observation. Specifically, in all four *Cis*-evolved colonies, the expression levels of most ribosomal genes were increased compared with the Splicing^Low ancestor (percentage of up-regulated ribosomal genes in the four colonies, 86%, 63%, 83%, and 84%) and levels of the ESR genes were mostly reduced compared with the ancestor (percentage of down-regulated ESR genes in four colonies, 90%, 76%, 87%, and 84%). These findings suggest that indeed these colonies adapted to the presence of the antibiotics in the environment and that the stress experienced by them was partially alleviated.

We next hypothesized that cellular fitness might correlate with mRNA levels of the YFP-Kan construct because increased transcript levels should result in higher concentrations of the YFP-Kan protein. Indeed, maximal growth rates of the Control and Splicing^Low ancestors and of the six evolved colonies correlate with mRNA levels of their YFP-Kan construct, as deduced from the RNA-seq (Fig 3F).

Because the observed *Cis* mutations occurred at the vicinity of the intron, we hypothesized that they increased splicing efficiency of the YFP-Kan transcript. To test this possibility, we
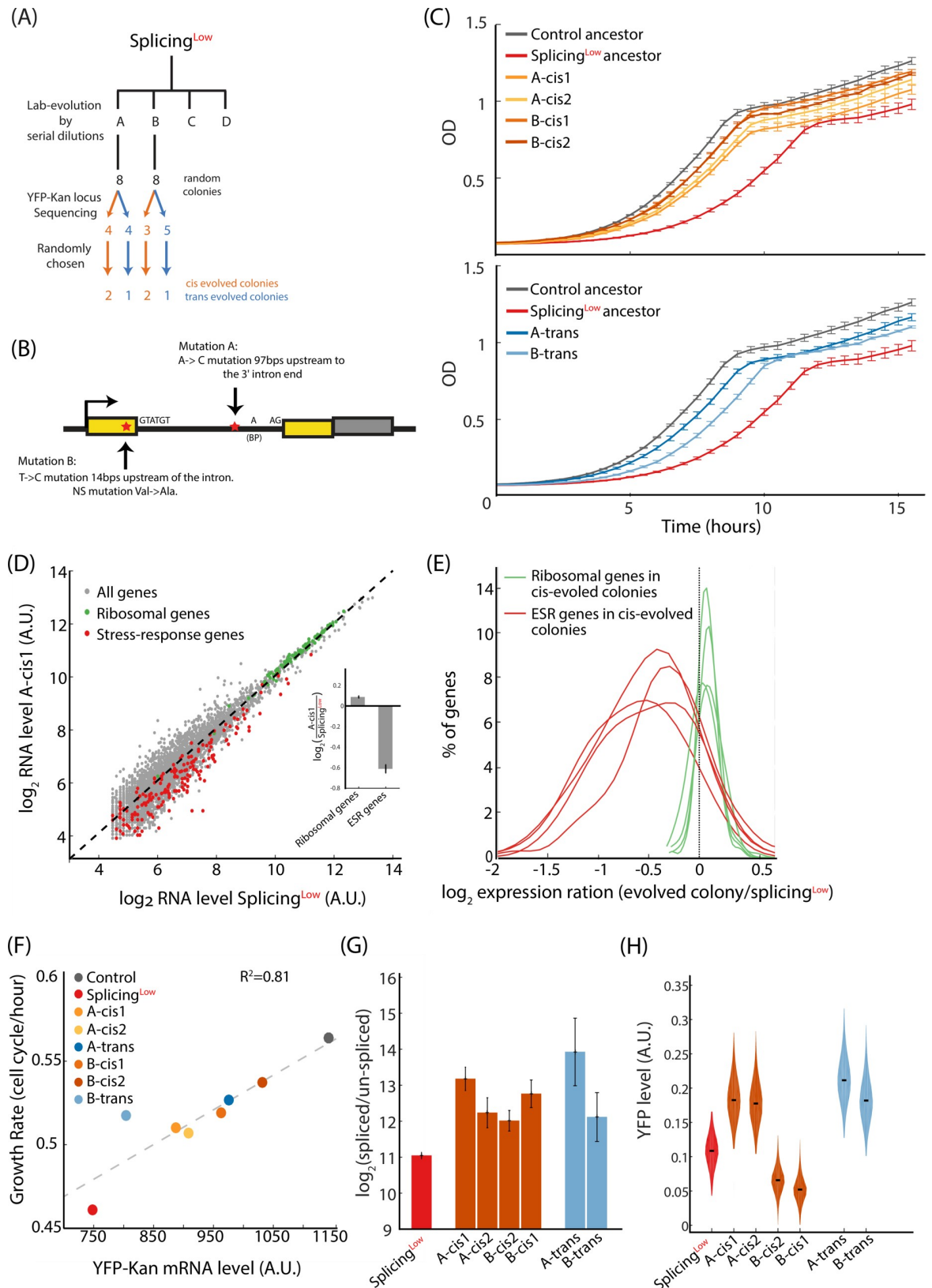
**Fig 3. Evolved colonies demonstrate increased splicing efficiency that results in higher transcript levels and relieved stress. (A)** We randomly chose 16 colonies in total from two evolved lines of Splicing^Low and sequenced the YFP-Kan locus of those colonies. We found

that approximately half had mutations in the YFP-Kan construct (indication of evolution in *Cis*) and the other half did not (indication of evolution in *Trans*). Of those colonies, we randomly chose two *Cis*-evolved and one *Trans*-evolved colonies from each evolved population for further examination. **(B)** Sequencing of the YFP-Kan construct in the evolved colonies revealed two mutation types: (i) in the intron itself and (ii) in the upstream exon. These mutations did not occur in the intron 5′ donor, 3′ acceptor, or the branching point—suggesting that other positions of the intron and its vicinity are functional and may affect splicing. **(C)** All *Cis*-evolved colonies (upper graph) and *Trans*-evolved colonies (lower graph) show increased fitness compared with the Splicing$^{Low}$ ancestor, yet still lower than the Control ancestor. **(D)** Transcriptome profiling reveals that ribosomal genes were up-regulated (green dots, $p = 4.94 \times 10^{-18}$, paired $t$ test) and stress-related genes were down-regulated (red dots, $p = 3.64 \times 10^{-15}$, paired $t$ test) in the evolved colony A-cis1 compared with the Splicing$^{Low}$ ancestor. **(Inset)** Mean log$_2$ ratio of ribosomal and ESR gene groups. **(E)** The four *Cis*-evolved colonies show similar trends, i.e., increased expression levels of ribosomal genes and decreased expression levels of stress-response genes ($p$-values for all cases < 0.005, paired $t$ test). These observations suggest that the stress experienced by the evolved colonies was alleviated during their adaptation to the antibiotics in the medium. **(F)** mRNA levels of YFP-Kan transcripts correlate with growth rates (R$^2$ = 0.82, $p = 0.0023$), suggesting that cellular fitness in our setup is indeed determined by the availability of kanamycin resistance proteins to overcome the antibiotics. **(G)** All *Cis*- and *Trans*-evolved colonies demonstrate increased splicing efficiency of the YFP-Kan mRNA compared with the Splicing$^{Low}$ ancestor ($p < 0.05$ for all colonies compared with Splicing$^{Low}$ ancestor). This result suggests that all adaptation trajectories led to the adaptation of the splicing process to better mature the un-spliced YFP-Kan transcript. **(H)** Fluorescence intensity of the YFP-Kan reporter shows increased levels for the two *Cis*-evolved colonies with the mutation in the intron and for the two *Trans*-evolved colonies. In contrast, the two *Cis*-evolved colonies with the nonsynonymous mutation in the exon demonstrate decreased YFP-Kan levels. This observation suggests that the nonsynonymous mutation hampered the ability of the YFP domain to fluorescent and reduced the fluorescence intensity per protein molecule (see text for full explanation). See numerical data for this figure in S1 Data.

https://doi.org/10.1371/journal.pbio.3000423.g003

assayed splicing efficiency for both *Cis*- and *Trans*-evolved colonies with qPCR, targeting the un-spliced or spliced transcript versions. Interestingly, the ratio of spliced to un-spliced transcripts, a measure of splicing efficiency, was higher in all evolved colonies compared with the Splicing$^{Low}$ ancestor, suggesting that at least some of the increase in mRNA level we observed in the evolved colonies results from higher splicing efficiency (Fig 3G).

To prove that adaptation of the colonies actually led to higher protein levels of the fluorescence-resistance fused protein, we measured fluorescence intensity using flow cytometry. We found that the two Cis-evolved colonies from population A (A-cis1 and A-cis2) and the two *Trans*-evolved colonies (A-trans and B-trans) showed higher YFP-Kan levels compared with the ancestor. In contrast, the two *Cis*-evolved colonies from population B (B-cis1 and B-cis2) demonstrated decreased fluorescence intensity values relative to the ancestor (Fig 3H). These observations indicate that the nonsynonymous, exon mutation reduced the per-protein fluorescence value of the YFP component of the YFP-Kan construct in these colonies. Indeed, this exonic mutation occurred in a position that was recently reported to reduce fluorescence when mutated in the highly similar GFP [40]. Because YFP's functionality, i.e., fluorescence, was not selected for or against in our setup, it appears to have been free to obtain mutations that help achieve a higher expression level of the entire fusion construct. It thus seems that a modular domain architecture of a protein may increase its evolvability under relevant conditions, as it allows the trade-off and optimization of one domain at the expense of another.

We next wanted to confirm that the fitness gained in the *Cis*-evolved colonies was indeed due to these mutations. It is still possible that additional beneficial mutations exist in the genome of the *Cis*-evolved colonies, which account for at least part of the improved fitness we observed in these colonies. We thus generated two rescue strains, termed *cis*-Rescue-A and *cis*-Rescue-B, in which these *Cis*-acting mutations were introduced individually to the ancestral Splicing$^{Low}$ background. Notably, the two rescue strains grew better than Splicing$^{Low}$ cells in the presence of the antibiotics, although not as well as Control cells (Fig 4A). Additionally, the stress experienced by Splicing$^{Low}$ cells, as observed by changes in expression levels of ribosomal and stress genes, was alleviated upon insertion of each individual *Cis* mutation (Fig 4B). Then, we measured splicing efficiencies and fluorescence intensity levels for both rescue strains and found that they resembled the levels shown in the evolved single colonies (Fig 4C and 4D, in comparison with Fig 3G and 3H).
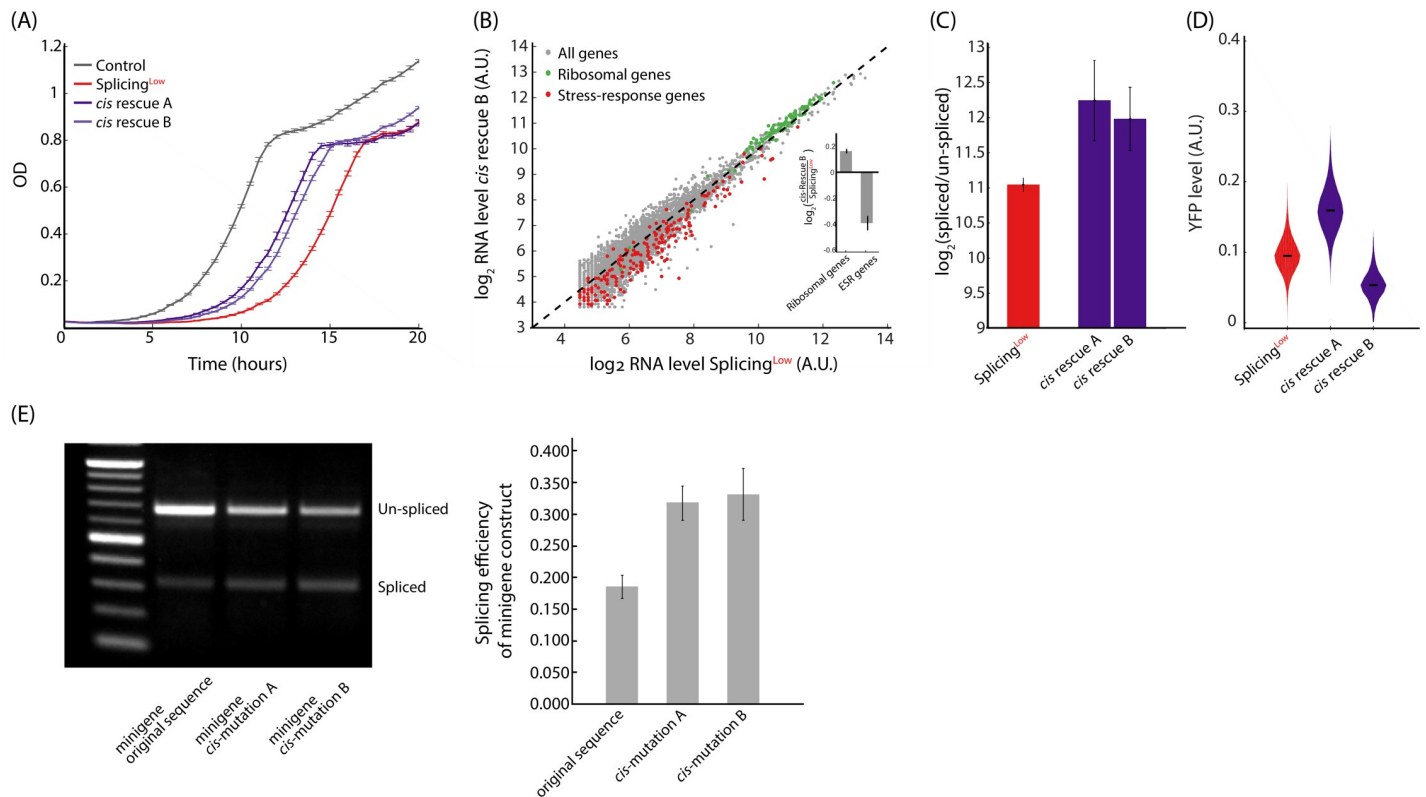
**Fig 4. *Cis*-acting mutations are sufficient to increase fitness by elevating splicing efficiency. (A)** We created two *Cis*-rescue strains, each harboring one of the mutations that appeared spontaneously in the evolved populations. Growth of the two *Cis*-rescue strains show that a single mutation in the YFP-Kan construct is sufficient to increase fitness compared with Splicing^Low. **(B)** The exonic mutation is also sufficient to alleviate stress, as ribosomal genes were up-regulated (green dots, $p = 1.02 \times 10^{-18}$, paired *t* test) and stress-related genes were down-regulated (red dots, $p = 9.02 \times 10^{-12}$, paired *t* test) in *cis*-Rescue-B compared with Splicing^Low. The same trend was also observed for the intronic mutation for *cis*-Rescue-A cells. (**Inset**) Mean log₂ ratio of ribosomal and ESR gene groups. **(C)** The two *Cis*-rescue strains demonstrate higher splicing efficiency of the YFP-Kan mRNA compared with the Splicing^Low ancestor ($p < 0.05$). This result suggests that a single mutation is sufficient to improve splicing efficiency. **(D)** Fluorescence intensity of the YFP-Kan reporter for the *cis*-Rescue-A and *cis*-Rescue-B strains show similar trends as the colonies in Fig 3D—supporting earlier conclusions. **(E)** The effects of *Cis* mutations on splicing tested with a mini-gene approach. We cloned the intron's original sequence and its two mutated versions together with 200 bps surrounding it into a high–copy number plasmid. RT-PCR assays for WT (BY4741) cells transformed with these plasmids show that, even in this context, the *Cis* mutations are sufficient to increase splicing efficiency of the intron for both mutated intron versions compared with the original sequence ($p < 0.05$, *t* test). See numerical data for this figure in S1 Data. WT, wild-type.

https://doi.org/10.1371/journal.pbio.3000423.g004

Our data so far demonstrate that the *Cis* mutations we identified in our lab-evolution experiment are beneficial because they improve splicing efficiency in the context of the entire YFP-Kan transcript. To further demonstrate that indeed these mutations increase splicing efficiency and to better exclude other mechanisms that might affect the YFP-Kan transcript, we used a mini-gene approach, a widely used methodology in splicing studies [41]. In this approach, the intron together with its adjacent exons are cloned to a new context, and splicing efficiency is measured. To this end, we cloned the original sequence of the intron plus 200 bases up- and downstream of it into a 2μ plasmid to a locus driven by a strong promoter. We additionally created two mutated versions of this construct, each with one of the two *Cis* mutations we discovered. Because there are only two short exons surrounding the intron in this system, it is ideal to explore the effects of the mutations on splicing. We thus harvested mRNA from exponentially growing cells and measured the splicing efficiency of each variant using PCR, with primers flanking the intron that simultaneously amplify both spliced and un-spliced versions of the YFP-Kan transcripts. In agreement with the splicing assay done for the full gene (Fig 4C), in the mini-gene assay too, both variants with either mutation *cis* A or B showed increased splicing efficiency compared with the ancestral versions (Fig 4E).

Our results thus far provide direct evidence that intron splicing takes part in the adaptation and optimization of gene expression patterns to environmental needs. Although intron sequences are much less conserved, compared with exons, and are believed to be less functional, we demonstrate that their sequence can be used by natural selection as a molecular mechanism to regulate splicing efficiency and adjust gene expression patterns.

## *Cis* mutation is adaptive through effects on mRNA structure that make 5′ donor and branching point sites more accessible to splicing

How can the mutations we identified in the intron facilitate splicing? One possibility is that these changes favorably alter the RNA structure of the un-spliced YFP-Kan transcript by making the splicing sites more accessible to the spliceosome. Indeed, RNA structures can inhibit or facilitate binding of spliceosome components to the pre-mRNA and affect splicing efficiency [42,43]. We thus computationally modeled the RNA structure of the intron and 50 bases on both its sides using the ViennaRNA algorithm [44]. We performed this analysis for the original and the two mutated sequences of the intron. Interestingly, mutation *cis* B, located near the 5′ donor site, leads to massive changes in the predicted RNA structure, notably causing the structure near the 5′ donor and the branching point sites to loosen (Fig 5A). Specifically, the pairing probability, a prediction for how likely it is for a position along the RNA molecule to associate with other positions, is decreased at the 5′ donor and branching point positions between the original and the mutated sequence (Fig 5B). How likely is a single point mutation to change so drastically the predicted structure of an RNA? To ask that, we constructed a simple null model in which we calculated the predicted pair-probability difference between the original sequence and each of the other possible single nucleotide mutations in the intron. Notably, mutation *cis* B falls among the 5.5% of all mutations with the highest predicted potential to affect the secondary structure; i.e., it is among the 5.5% of point mutations that loosen the RNA secondary structure near the 5′ donor and branching point the most (Fig 5C). These observations suggest a model in which mutation *cis* B may facilitate splicing due to increased accessibility of the splicing machinery to the functional splicing sites. In contrast, mutation *cis* A did not show similar patterns to mutation *cis* B (Fig 5A and 5B), raising the possibility that a different, still obscure, mechanism is causing the beneficial effects of this mutation.

## Evolution in *Trans*: Increasing cellular availability of the splicing machinery can be adaptive

We next aimed to decipher the mechanism behind the increased YFP-Kan levels in the *Trans*-evolved colonies that showed no mutations in *Cis*, i.e., within the reporter gene or in its vicinity. We reasoned that elevating availability of the splicing machinery could be a means to increase splicing efficiency of the YFP-Kan transcript and thus could be used as an adaptive mechanism to the antibiotics challenge. As with other cellular machineries whose functioning depends on supply-to-demand economy [4,8,45–47], increased splicing availability could be achieved by either increasing the expression of the splicing machinery genes or decreasing expression levels of other intron-containing genes, which collectively constitute the "demand" for the splicing machinery.

To test if any of these evolutionary routes were indeed taken by the evolved cells, we calculated the expression level ratio of genes between the evolved colonies and their Splicing^Low ancestor. In colony A-*trans*, we observed increased expression ratio of splicing machinery genes (the "supply") and decreased expression ratio of non-ribosomal intron-containing genes (the "demand," Fig 6B, and S1 Fig for other colonies). In contrast, both supply and demand genes show similar expression levels between Splicing^Low and Control ancestors, suggesting
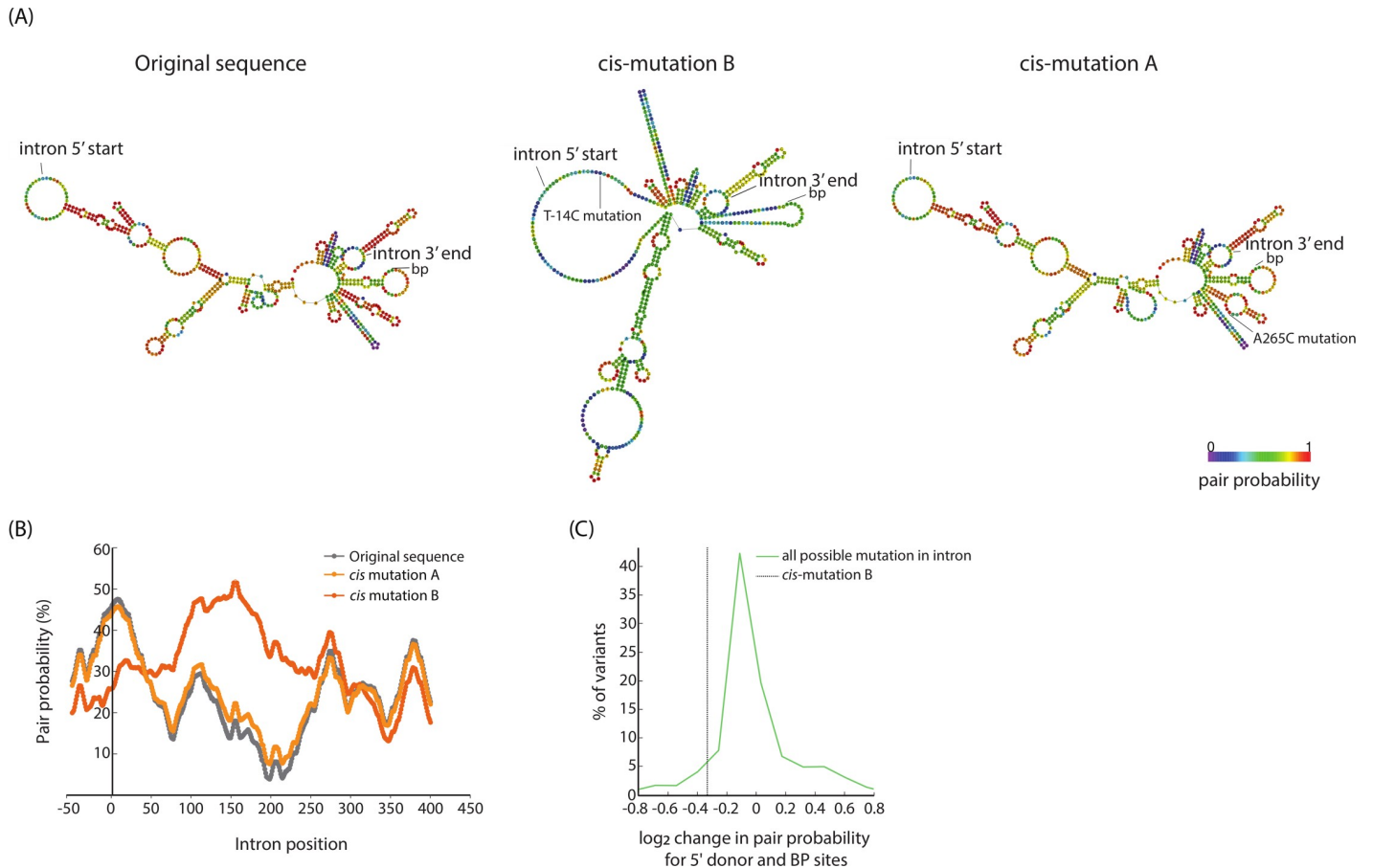
(A)



(B)

(C)

**Fig 5. A secondary structure model of the effects of *Cis* mutation on splicing efficiency. (A)** RNA secondary structure predictions using Vienna algorithm of the intron plus 50 bps of its surrounding exons for the original sequence and its two mutated versions. *Cis*-mutation B reshapes the predicted secondary structure and lowers the base pairing probability around the 5′ donor and branching point sites. **(B)** Base pairing probability at all positions as calculated by the Vienna algorithm for the original intron sequence and its two mutated versions. Position of 5′ donor site is 0, position of branching point site is 328, position of 3′ acceptor site is 361, position of mutation A is 265, position of mutation B is −14. **(C)** Probability distribution of the $\log_2$ of change in base pairing probability along a 20-base window at the 5′ donor and branching point sites for all possible single nucleotide variations on the intron's sequence compared with the original sequence. Vertical line denotes the value of the observed *cis*-mutation B, which is at the bottom 5.5 percentage of the histogram; i.e., it is among mutations that mostly reduce the base pair probability at these sites.

that a mere physiological adaptation did not confer upon the Splicing^Low ancestor the ability to adjust supply and demand to the intron burden (Fig 6A and see Discussion).

We then computed changes in "supply-to-demand splicing availability" by summing up for each strain the expression levels of the splicing machinery genes and intron-containing genes, separately, then dividing these values and normalizing the ratio with that of the control strain. Interestingly, the supply-to-demand difference has increased appreciably in the *Trans*-evolved colonies compared with the Splicing^Low ancestor. The *Cis*-evolving colonies have also improved their supply-to-demand relative to the ancestor, indicating that they too may have evolved in *Trans* in addition to their *Cis* adaptations. Thus, we concluded that both *Cis* and *Trans* adaptation routes can co-occur in the same genome towards optimization of its gene expression patterns. Notably, the group of non-evolved strains (Splicing^Low and the two *Cis*-rescue strains) showed lower supply-to-demand ratios compared with the group of evolved strains (both *Cis*- and *Trans*-evolved colonies) (Fig 6C).
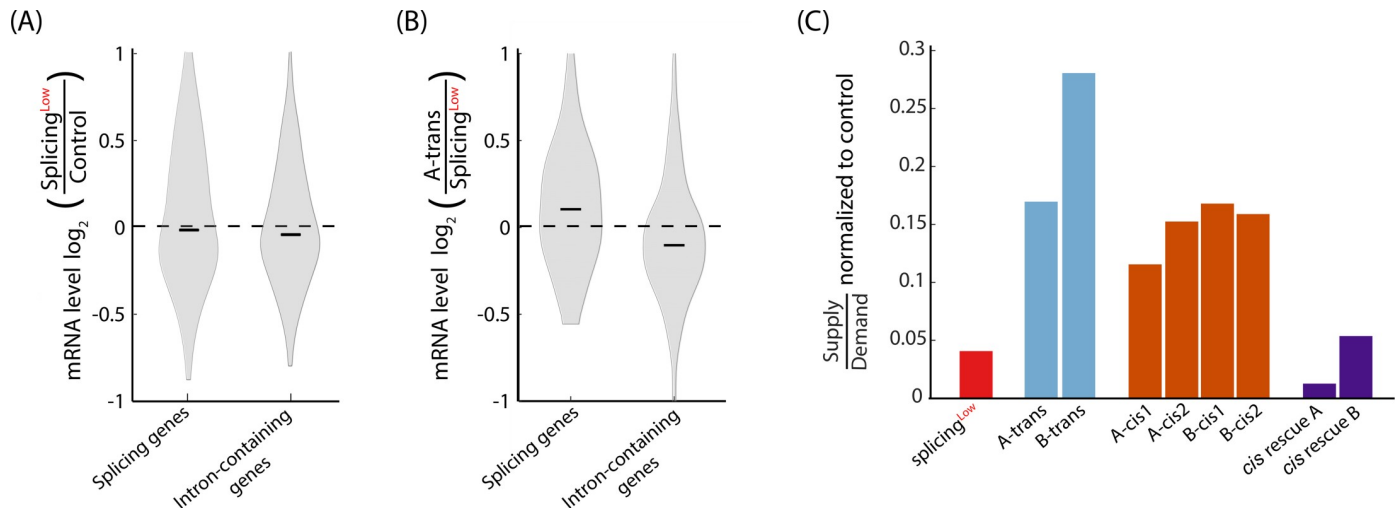
**Fig 6. Increasing cellular availability of the splicing machinery is an adaptive mechanism of splicing. (A)** The groups of splicing genes (splicing supply) and intron-containing genes (splicing demand) show similar levels between Splicing$^{Low}$ and Control ancestors ($p > 0.05$, paired $t$ test for both gene groups). **(B)** The groups of splicing genes and intron-containing genes were increased ($p = 1.36 \times 10^{-3}$, paired $t$ test) and decreased ($p = 1.67 \times 10^{-2}$, paired $t$ test), respectively, in colony A-*trans* compared with the Splicing$^{Low}$ ancestor. This observation suggests that the supply-to-demand ratio of the splicing machinery was increased in the A-*trans* colony, which allowed its increased splicing efficiency of the YFP-Kan transcript. **(C)** Supply-to-demand ratios for the splicing machinery were calculated to Control and Splicing$^{Low}$ ancestors, to all evolved colonies, and to the *Cis*-rescue strains as the sum expression level of all splicing-related genes over the sum expression level intron-containing genes. Values of all strains were then normalized to the value of the Control ancestor. Importantly, supply-to-demand ratios are similar for all strains that did not evolve (Control and Splicing$^{Low}$ ancestors and the two rescue strains) and were increased for all evolved colonies ($p = 0.005$ for difference in supply-to-demand ratios between evolved and non-evolved strains, $t$ test). These results suggest that indeed the cellular availability of the splicing machinery was elevated due to an evolutionary adaptation process and not because of other physiological mechanisms (see Discussion). See numerical data for this figure in S1 Data.

Taken together, we argue that increased cellular availability of the splicing machinery is the result of an evolutionarily adaptive process, which might have allowed for the improved splicing efficiency of the YFP-Kan gene in the evolved colonies. While genomic sequencing revealed interesting mutations in *Trans* (see below), we did not find mutations that can explain these supply-to-demand changes in a straightforward way.

### Adaptation in *Trans* is also based on mutations in SR-like proteins

We next aimed to reveal which genetic changes may have occurred in *Trans* to the fusion gene. To that end, we fully sequenced the genomes of the two *Trans*-evolved colonies (A-*trans* and B-*trans*) and identified, respectively, three and two mutations in each of their genomes (see S2 Table for list of mutations). Interestingly, in both colonies we found a mutation in a member of the SR-protein group in *S. cerevisiae*, which are splicing-related genes. Specifically, in colony A-*trans* we found a nonsynonymous mutation in *GBP2*, changing histidine into tyrosine, H160Y, and in colony B-*trans* we found a nonsynonymous mutation in *NPL3*, changing phenylalanine into valine, F160V (Fig 7A).

Npl3 was previously shown to be loaded onto pre-mRNAs during early stages of transcription [48,49] and to support the recruitment of the spliceosome [29] after it has completed its quality control function to monitor correct 5′ capping (Schneider and Krebber, in preparation). Unlike for Npl3, deletion of Gbp2 does not result in accumulation of pre-mRNAs [29]. In contrast, Gbp2 was shown to associate with the late-stage spliceosome and appears to function as a quality control factor that oversees successful splicing of pre-mRNAs, which facilitates the recruitment of the nuclear exosome to incorrectly spliced pre-mRNAs in order to eliminate these faulty transcripts [33].
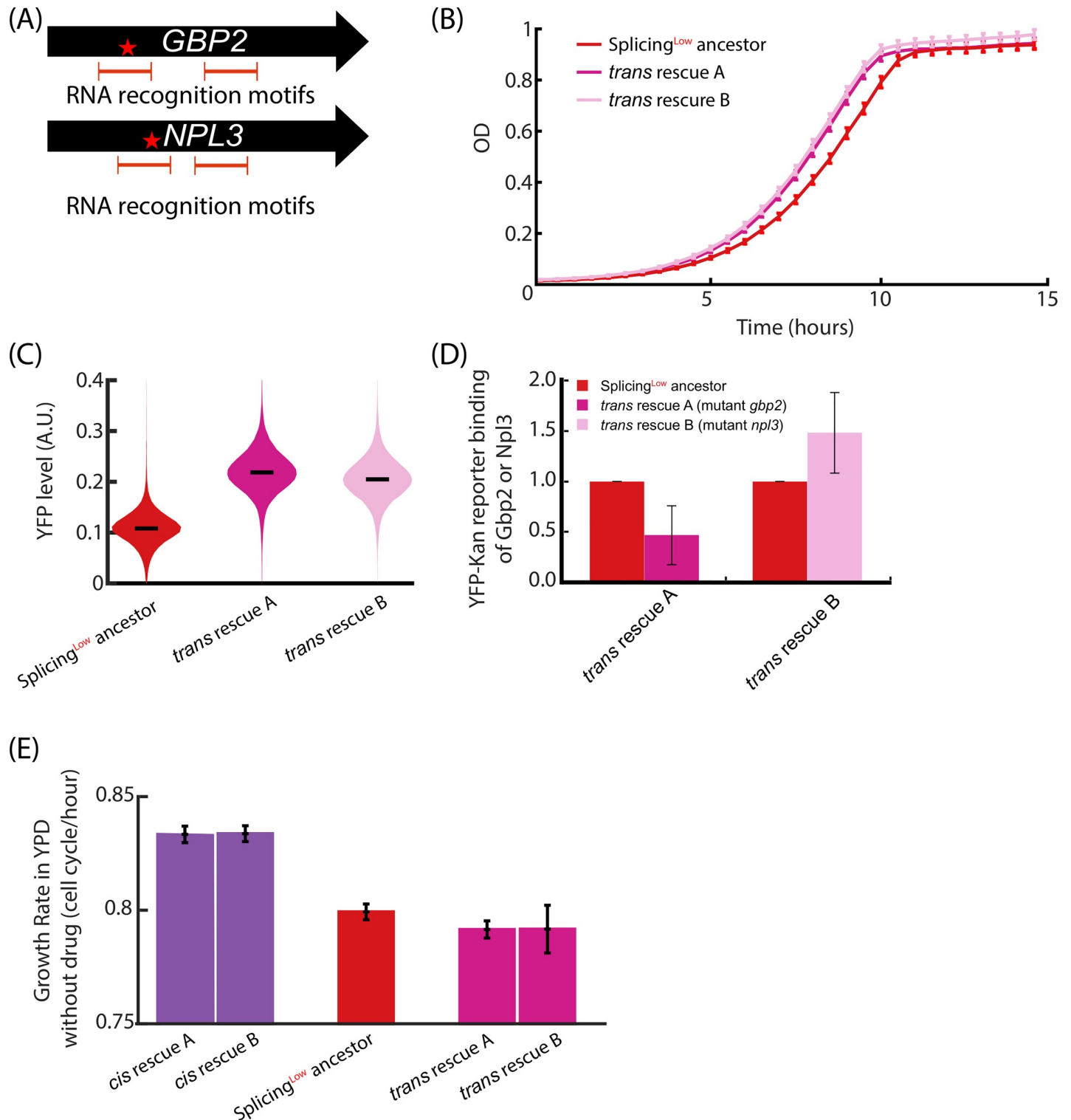
**Fig 7. *Trans* adaptation is based on mutations in SR-like proteins that change their affinity to the YFP-Kan transcript. (A)** Genome sequencing of the *Trans*-evolved colonies revealed two mutations in the SR-like proteins *GBP2* and *NPL3*. Both nonsynonymous mutations, H160Y in *GBP2* and F160V in *NPL3*, occurred in an RNA recognition motif. **(B)** We created two *Trans*-rescue strains, termed *trans*-Rescue-A and *trans*-Rescue-B, each harboring one of the two mutations as described in A. Growth of these two *Trans*-rescue strains in the presence of the drug show that the *Trans* mutations are sufficient to increase fitness compared with Splicing^Low. **(C)** Fluorescence intensities of the YFP-Kan reporter for *trans*-Rescue-A and *trans*-Rescue-B strains are higher compared with Splicing^Low cells. **(D)** Binding of mutant gbp2 and npl3 to the YFP-Kan mRNA was examined and compared with the respective WT proteins. RT-PCR results showed that mutant gbp2 has a reduced binding affinity

to the reporter ($p = 2.97 \times 10^{-3}$), whereas mutant npl3 binds with a slightly higher affinity ($p = 9.54 \times 10^{-2}$). These opposite effects on binding can be rationalized by the different functions of *GBP2* and *NPL3*. See text for full details. **(E)** Growth rates of all four rescue strains and the Splicing$^{Low}$ ancestor under permissive conditions, without the drug. Notably, *Cis*-rescue strains demonstrate higher growth rates compared with Splicing$^{Low}$—indicating that they alleviate the burden of the inefficiently spliced intron that is independent of the drug. In contrast, both *Trans*-rescue strains show similar growth rates compared with Splicing$^{Low}$—indicating that mutations in SR-like proteins are associated with additional costs to cells that cancel out the decreased burden of the inefficiently spliced intron. See numerical data for this figure in S1 Data. WT, wild-type.

To examine whether these mutations in the two SR-like proteins are sufficient to increase fitness and YFP fluorescence of cells that harbor the low-efficiency spliced intron as part of the YFP-Kan construct, we used CRISPR technology (see Methods) to introduce these point mutations, each individually, to the background of the Splicing$^{Low}$ ancestor. We termed these engineered strains *trans*-Rescue-A and *trans*-Rescue-B. Indeed, these two new rescue strains grew faster on the G418 antibiotics compared with ancestor cells (Fig 7B), and they both showed higher fluorescence intensity levels of YFP (Fig 7C). Hence, it appears that these *Trans* mutations result in increased levels of the YFP-Kan proteins.

Notably, both mutations we identified in *GBP2* and *NPL3* occurred in one of the two RNA recognition motifs of their encoded proteins (Fig 7A), suggesting that their affinity to the YFP-Kan transcript may have changed. We therefore tested if the mutant proteins would show modified YFP-Kan RNA-binding affinities compared with their wild counterparts. Mutant or wild-type (WT) Gbp2 and Npl3 were each immunoprecipitated with specific antibodies and the associated RNAs were purified (see Methods). Following reverse transcription, the levels of YFP-Kan mRNA bound to the different proteins were analyzed with RT-PCR. Interestingly, we obtained distinct results for Gbp2 and Npl3. While mutant gbp2 showed significantly decreased binding to YFP-Kan compared with WT, mutant npl3 seemed to have a slightly higher affinity to the reporter RNA (Fig 7D). These results are in agreement with our current understanding of the functions of the two proteins. The increase in YFP-Kan binding of the mutant npl3 might result in better assembly of the spliceosome and consequently increased splicing efficiency. The decreased binding of the mutant gbp2 to the YFP-Kan transcript might delay the assembly of the nuclear RNA degradation machinery and hence provide more time and higher chances to the YFP-Kan transcript to be successfully spliced and exported to the cytoplasm.

### *Cis* and *Trans* mutations show distinct phenotypes under permissive conditions

If the *Trans* mutations in the SR-like proteins are beneficial, why are they not the WT sequence? We speculated that these mutations are beneficial only under the drug pressure, and that they otherwise come with a cost when the antibiotic is absent. We thus measured the effects of these mutations on fitness in a medium lacking the antibiotics (rich YPD medium). We found that cells with either of the *Cis* mutations grew modestly faster than the Splicing$^{Low}$ ancestor strain (Fig 7E), indicating that they alleviate some drug-independent burden of the inefficiently spliced intron in Splicing$^{Low}$ cells. In contrast, cells with either of the *Trans* mutations grew similarly to Splicing$^{Low}$ cells when the drug was not supplemented to the medium (Fig 7E). It seems then that the *Trans* mutations do not alleviate cellular fitness in the absence of the drug, i.e., they do not reduce the drug-independent burden of the inefficiently spliced intron.

### Discussion

In this work, we study the role of the splicing machinery in optimization of gene expression programs by placing selective pressure on cells to improve the splicing efficiency of a specific

gene. Our results provide molecular evidence for the relevance of splicing as another instrument in the cellular toolbox towards adjusting its gene expression patterns. To the best of our knowledge, we demonstrate the first experimental evidence of splicing efficiency adaptation, confirming that this adaptation can occur in *Cis* and *Trans* similarly to adaptations of other stages of the gene expression regulation process. Adaptation of splicing efficiency might be very common among species evolution, given the observed correlation between splicing efficiency and transcription level of genes [28]. Hence, it is of interest that we elucidate more molecular mechanism that allow optimization of intron splicing.

Two potential solutions to the burden that we imposed on our Splicing^Low ancestor lines were surprisingly not realized during our lab-evolution. First, considering previous studies of splicing evolution, one could have expected the intron to be lost by a genomic deletion or through reverse transcription [50,51]. Such a solution could have been an ideal evolutionary adaptation to alleviate the burden, as we show that the intron-less strain has the highest fitness. The fact that we did not observe an intron-loss event suggests that this is a less accessible solution in this case, in agreement with previous evidence in yeast that nucleotide mutations are 33 times more frequent than deletion events [52]. It further seems that our evolving strains did not evolve a transcription- or translation-based solution that in principle could have elevated the expression of the YFP-Kan fusion gene. Instead, adaptation appears to have affected the splicing of the gene. Given that some of the adaptations were found within the intron, another surprise was that none of the mutations occurred within any of the intron's three functional sites: the 5′ donor, 3′ acceptor, or the branch point of splicing. Indeed, one mutation that was verified here to affect splicing resides in a region of the intron not known to exert a major effect on splicing, and another splicing-improving mutation happened in the upstream exon. These observations indicate that various positions in the intron and its proximity may facilitate the splicing rate and take part in the evolution of introns, a phenomenon that was previously discussed only in regards to alternative splicing [53,54].

Interestingly, the *Cis* mutation that resides inside the intron results in considerable changes in the predicted RNA secondary structure of the intron, presumably lowering the association probability of the 5′ donor and branching point sites. This finding goes in line with previous observations that show how RNA structures can inhibit or facilitate binding of spliceosome components to the pre-mRNA and affect splicing efficiency [42,43], specifically at the edges of introns [36,42,43,55,56]. Thus, it is tempting to speculate that many more positions than previously recognized of introns' sequences, rather than only its edges and the branching point, are under selection and are biologically relevant to the intron function.

Notably, the fluorescence intensity per protein molecule of the YFP domain was decreased due to the nonsynonymous mutation in the YFP first exon, suggesting that under certain evolutionary constraints, selection may hamper superfluous functions of certain protein domains so as to increase availability of the entire protein.

Adaptive changes also occurred in *Trans* to the YFP-Kan locus and increased availability of the splicing machinery to this gene. An alternative explanation for evolutionarily adaptive changes in splicing availability could be based on cellular physiological response. In such a model, the introduction of an inefficiently spliced intron into a gene that is on high demand may occupy a larger portion of splicing machinery, which in turn prevents splicing of other intron-containing genes, which ultimately leads to degradation at the pre-mRNA level, e.g., by the mRNA quality control machinery [33]. Although valid, this physiological adaptation appears less likely to explain our results. First, the physiological model predicts that the expression level changes that we see following the evolutionary adaptations will also be seen when comparing the expression of the splicing genes and intron-containing genes between Splicing^Low and the Control strain (that harbors no intron in its YFP-Kan construct). However, we

did not observe such a difference, suggesting that a physiological regulatory model is less likely (Fig 6A). Second, the two *Cis*-rescue strains, which did not evolve and only harbor our artificially introduced *Cis*-acting mutation, did not demonstrate changes in splicing availability. Generally, evolved strains showed higher supply-to-demand rations for splicing compared with non-evolved strains in our study (Fig 6C). These observations more forcefully support an evolutionarily adaptive process that relies on additional genetic or epigenetic changes that are accumulated during the continuous growth of the evolved populations.

Recently, the competition of pre-mRNAs for the splicing machinery was shown to affect cellular function, as splicing efficiency of multiple introns was influenced by changes in the composition of the transcript pool [57]. While this mechanism was elegantly shown to take part in physiological adaptation by maintaining the separation between meiotic and vegetative gene-expression states, it is also possible that it can be used as an adaptive mechanism in evolution to optimize gene expression levels of cells.

Our findings demonstrate how availability of the splicing apparatus may have been adaptively increased both by elevating the expression level of the machinery's genes and/or by reducing expression of other intron-containing genes that probably compete with the antibiotic resistance un-spliced RNA for the spliceosome. Thus, increase in supply-to-demand ratio, analogous to the case in translation systems [8,58], appears to have evolved in this case.

Additionally, we revealed a role for SR-like proteins, *GBP2* and *NPL3*, when the splicing machinery adapts to a new need of optimizing splicing of a specific intron. Notably, these proteins have various functions, including early recruitment of the spliceosome, quality surveillance of nascent mRNA quality, association with the nuclear RNA degradation machinery of faulty transcripts, and finally, assistance with nuclear export for mature mRNAs.

Interestingly, the nonsynonymous mutations we found in these proteins occurred in one of their RNA recognition motifs. Using an RNA immunoprecipitation assay, we showed that the mutations in Gbp2 and Npl3 decrease and increase, respectively, the binding capacities of the YFP-Kan construct. Yet, these surprising opposite effects can be explained by the different roles of Gbp2 and Npl3—the former is a quality control factor of splicing that elicits degradation of un- or mis-spliced transcripts, and the latter is a spliceosome-recruitment agent. On the one hand, the lower binding capacity of the mutated Gbp2 probably provides more time for the spliceosome to complete the maturation of the YFP-Kan pre-mRNA before it is degraded. The improved binding capacity of the mutated Npl3, on the other hand, facilitates the recruitment of the spliceosome and hence might improve splicing efficiency of the transcript. The *Trans*-evolution mechanism we reveal here is intriguing because it shows that under acute selection, a cellular machinery can evolve for the need of one gene only.

Ultimately, we revealed a fundamental difference for *Cis* and *Trans* evolution of the splicing machinery when this cellular process faces a need to adapt. While *Cis*-based adaptations are "local" and lowered the burdens of splicing for the intron under selection, *Trans*-based adaptations showed wider cellular effects that may be costly to cells when the original evolutionary challenge is lifted. Further investigations will reveal which of these solutions, *Cis* or *Trans*, proves to be more evolutionarily stable—to fully reveal the dynamics of splicing adaptation when cells optimize their gene expression.

## Methods

### Yeast strains and plasmids

All *S. cerevisiae* strains in this study have the following genetic background: his3Δ1::TEF2-m-Cherry::URA3::RPS28Ap-YiFP-KAN::NAT; canΔ1::STE2pr-Sp_his5; lypΔ1::STE3pr-LEU2; leu2Δ0; ura3Δ0.

Strains of Y-intron-FP were taken from Yofe and colleagues [36] and were introduced with a Kan resistance gene fused 3′ terminally to the YFP. To reconstitute the mutations discovered after lab evolution (rescue strains), we amplified cassettes of Y-$i_{mut}$-FP-KAN and transformed these into the ancestor Control strain, selecting with KAN. Notably, all strains also carry an mCherry-fluorescent protein driven by an independent *TEF2* promoter that was used to normalize cell-to-cell variability for the YFP-Kan expression levels.

## Media

Cultures were grown at 30˚C in rich medium (1% bacto-yeast extract, 2% bacto-peptone, and 2% dextrose [YPD]). Throughout all experiments, G418 was supplemented to the medium at a concentration of 3 mg/mL, which is 10-fold higher than the standard.

## Evolution experiments

Lab-evolution experiments were carried out by daily serial dilution for 80 days. Cells were grown on 1.2 mL of YPD+G418 at 30˚C until reaching stationary phase and then diluted by a factor of 1:120 into fresh media (approximately 7 generations per dilution, a total of about 560 generations).

## Liquid growth measurements

Cells were grown in YPD+G418 at 30˚C overnight. The following day, they were diluted to an OD = 0.05 in YPD+G418, and optical density (600 nm) measurements were taken at 30-minute intervals. Growth comparisons were performed using 96-well plates, and the growth curve for each strain was obtained by averaging at least 15 wells.

## Flow cytometry measurements of YFP-Kan levels

Cells were grown in YPD+G418 at 30˚C overnight. The following day, they were diluted to an OD = 0.05 in YPD+G418, placed at 30˚C, and followed until they reached logarithmic growth phase at an optical density of between 0.4 and 0.5. Then, YFP and mCherry levels were measured for between 20,000 and 50,000 cells for each culture with flow cytometry. Gating was performed according to side and forward scatters, and YFP levels were normalized with the mCherry signal for each cell individually.

## Quantitative PCR measurements of splicing efficiency

Cultures were grown in YPD+G418 at 30˚C until cells reached the logarithmic growth phase at an optical density of approximately 0.4. Then, RNA was extracted using MasterPure kit (Epicentre) and were reverse transcribed to cDNA using random primers. A total of 2 μL of cDNA was added to each reaction as template for qPCR using light cycler 480 SYBR I master kit and the LightCycler 480 system (Roche Applied Science), according to the manufacturer's instructions. For each strain, qPCRs were performed with two to three biological repetitions and three technical repeats. A first qPCR was performed targeting the transcript-spliced version, with a forward primer complementing the exon-exon junction and a downstream reverse primer. A second PCR targeted the un-spliced version of the transcript, with a forward primer complementing the intron and the same reverse primer of the first reaction: $F_{exon-exon}$ = 5′-CACTACTTTAGGTTATGGTTT-3′; $F_{intron}$ = 5′-CTTCAATTTACTGAATTTGTATG-3′; $R_{both}$ = 5′-GTCTTGTAGTTACCGTCA-3′.

Splicing efficiency is reported as the average Cp of the spliced transcript minus the average Cp of the un-spliced version.

## mRNA deep sequencing

Cultures were grown in YPD+G418 at 30˚C until cells reached the logarithmic growth phase at an optical density of approximately 0.4. Cells were then harvested by centrifugation and flash-frozen in liquid nitrogen. RNA was extracted using a modified protocol of nucleospin 96 RNA kit (Machery-Nagel). Specifically, cell lysis was done in a 96-deep-well plate by adding to each well 450 μL of lysis buffer containing 1 M sorbitol, 100 mM EDTA, and 0.45 μL lyticase (10 IU/μL). The plate was incubated at 30˚C for 30 minutes to break cell walls and centrifuged for 10 minutes at 3,000 rpm, followed by the removal of the supernatant. Then, extraction continued as in the protocol of nucleospin 96 RNA kit, only using β-mercaptoethanol instead of DTT. Poly(A)-selected RNA extracts of approximate size of 200 bps were reverse transcribed to cDNA using poly(T) primers that were bar coded with a unique molecular identifier (UMI). cDNA was then amplified and sequenced with an Illumina HiSeq 2500.

## Analysis of mRNA deep sequencing

Processing of RNA-seq data was performed as described in Voichek and colleagues [59]. Shortly, reads were aligned using Bowtie [60] (parameters:—best–a–m 2 –strata -5 10) to the genome of *S. Cerevisiae* (R64 from SGD) with an additional chromosome containing the sequence of the YFP-Kan construct. For each sequence, we normalized for PCR bias using UMIs, as described in Kivioja and colleagues [61]. Next, reads for each gene end (400 bp upstream to 200 bp downstream of the ORF's 3′ end) were summed up to estimate the gene's expression level. Genes with coverage lower than 10 reads were excluded. To normalize for differences in coverage among samples, we divided each gene expression by the total read count of each sample and then multiplied by $10^6$. Then, the expression ratio was calculated between an evolved/rescue colony to the ancestor, and a $log_2$ operation was performed on that ratio. These values were used to compare expression levels of gene groups (ribosomal genes, general stress response genes, splicing machinery genes, intron-containing genes) and of the YFP-Kan mRNA levels as described in the manuscript. When calculating the expression levels of splicing machinery and intron-containing gene groups, the ribosomal and general stress response genes were excluded from the analysis in order to avoid bias from cellular regulation due to changes in physiology and growth rate of the cells.

## CRISPR genome engineering

CRISPR protocol for *S. cerevisiae* was performed as described previously [62]. Shortly, gRNA sequences targeting a locus near (<50 bps) the position of the desired mutations were cloned into the vector bRA89 that allows co-expression of Cas9 and gRNA. CRISPR plasmid (1 μg) was co-transformed to yeast cells (20 μL) along with a repair dsDNA cassette with the desired mutations (500 ng) using the standard yeast transformation protocol. Cells were plated on selective plates (YPD+hygromycin, 50 mg/mL), colonies were screened with PCR and Sanger sequencing for the desired mutations, positive colonies were grown on permissive media for 24 hours to ascertain loss of the CRISPR plasmid, and copies were stored at −80˚C.

The repair cassette holds 45-bp homology on each end, flanking the desired mutation. Because the mutation did not modify the PAM sequence, additional synonymous mutations were also introduced alongside the desired mutation to stop the cutting of the locus by the Cas9+gRNA complex. These synonymous mutations were also introduced to the relevant genetic background without the desired mutations and were confirmed to have no effect on cellular growth or YFP expression levels.

## RNA co-immunoprecipitation

Yeast cells were grown in YPD at 25˚C until logarithmic phase and collected via centrifugation. Cells were resuspended in 1× pellet volume RNA-IP buffer (25 mM Tris-HCl [pH 7.5], 150 mM NaCl, 2 mM MgCl₂, 0.5% [v/v] Triton X-100, 0.2 mM PMSF in isopropanol, 0.5 mM DTT, 40 units of RiboLock RNase Inhibitor [Thermo Fisher Scientific], cOmplete EDTA-free Protease Inhibitor Cocktail [Roche Diagnostics and Sigma-Aldrich]) and lysed with vigorous vortexing three times, 20 s, 6 m/s with the FastPrep-24 Instrument (MP Biomedicals) in the presence of 1× pellet volume of glass beads. For immunoprecipitation, cleared lysates were incubated with or without Gbp2 or Npl3 antibody (rabbit polyclonal, self-made) at 4˚C for 1 hour followed by the addition of 10 μL Protein G sepharose beads (Amersham Biosciences) and further incubation at 4˚C for 2 hours. RNase-Free DNase I (QIAGEN) was also added to digest DNA during incubation. A total of 100 μL of the lysate was treated with DNase I and incubated at 4˚C in parallel. The beads were washed five times with RNA-IP buffer. RNA was isolated from both lysate and eluate samples using the TRIzol Reagent (Thermo Fisher Scientific) and eluted in 20 μL DEPC-treated ddH₂O. Prior to reverse transcription, the RNA samples were treated with TURBO DNA-*free* Kit (Thermo Fisher Scientific) to eliminate residual DNA.

## Reverse transcription and qRT-PCR of precipitated RNA

The same concentration of RNA from all samples was reverse transcribed using the FastGene Scriptase II cDNA Kit (NIPPON Genetics EUROPE) with Random Hexamer Primer (Thermo Fisher Scientific). PCR samples were prepared with qPCRBIO Sygreen Mix Lo-ROX (PCR Biosystems), and qRT-PCR was performed using the CFX Connect Real-Time PCR Detection System (Bio-Rad Laboratories) for 45 cycles with an annealing temperature of 60˚C. Forward primer (5′-TTATTCACTGGTGTTGTCCC-3′) and reverse primer (5′-CATGGAACTGGC AATTTACC-3′) that bind specifically to *YFP-Kan* were used. Each PCR reaction was performed in triplicates, and the average Cq value was used in further analysis. From each eluate Cq value, the corresponding lysate Cq value was substracted (ΔCq). The ΔCq of the no-antibody control sample was subtracted from the ΔCq of the pull-down sample (ΔΔCq). Binding of the RNA was calculated by $2^{-\Delta\Delta Cq}$.

## Supporting information

**S1 Fig. Supply-to-demand ratio in each of the evolved colonies.** Folding change ratios in $\log_2$ are shown for splicing genes (left) and intron-containing genes (right). Black line represents the median of the distribution. See numerical data for this figure in S1 Data.
(PDF)

**S1 Data. Supporting data for all main figures.**
(XLSX)

**S1 Table. List of strains used in this work.**
(XLSX)

**S2 Table. List of genomic mutations in *Trans*-evolved colonies.**
(XLSX)

## Acknowledgments

## Author Contributions

**Conceptualization:** Idan Frumkin, Ido Yofe, Yitzhak Pilpel.

**Data curation:** Idan Frumkin, Ido Yofe, Raz Bar-Ziv, Yonat Gurvich, Yen-Yun Lu, Ruth Towers, Dvir Schirman, Heike Krebber.

**Formal analysis:** Idan Frumkin, Yen-Yun Lu, Yoav Voichek, Dvir Schirman, Heike Krebber.

**Funding acquisition:** Yitzhak Pilpel.

**Investigation:** Idan Frumkin, Ido Yofe, Yitzhak Pilpel.

**Methodology:** Idan Frumkin, Ido Yofe, Raz Bar-Ziv, Yonat Gurvich, Yen-Yun Lu, Yoav Voichek, Heike Krebber, Yitzhak Pilpel.

**Project administration:** Idan Frumkin.

**Resources:** Yen-Yun Lu.

**Supervision:** Yitzhak Pilpel.

**Validation:** Idan Frumkin, Heike Krebber, Yitzhak Pilpel.

**Visualization:** Idan Frumkin.

**Writing – original draft:** Idan Frumkin, Yitzhak Pilpel.

**Writing – review & editing:** Idan Frumkin, Yitzhak Pilpel.

## References

1. Weingarten-Gabbay S, Segal E. The grammar of transcriptional regulation. Hum Genet. 2014; https://doi.org/10.1007/s00439-013-1413-1 PMID: 24390306

2. Amorós-Moya D, Bedhomme S, Hermann M, Bravo IG. Evolution in regulatory regions rapidly compensates the cost of nonoptimal codon usage. Mol Biol Evol. 2010; 27: 2141–51. https://doi.org/10.1093/molbev/msq103 PMID: 20403964

3. Quax TEF, Claassens NJ, Söll D, van der Oost J. Codon Bias as a Means to Fine-Tune Gene Expression. Mol Cell. 2015; 59: 149–161. https://doi.org/10.1016/j.molcel.2015.05.035 PMID: 26186290

4. Gingold H, Pilpel Y. Determinants of translation efficiency and accuracy. Mol Syst Biol. 2011; 7: 481. https://doi.org/10.1038/msb.2011.14 PMID: 21487400

5. Presnyak V, Alhusaini N, Chen Y-H, Martin S, Morris N, Kline N, et al. Codon optimality is a major determinant of mRNA stability. Cell. 2015; 160: 1111–24. https://doi.org/10.1016/j.cell.2015.02.029 PMID: 25768907

6. Bazzini AA, Del Viso F, Moreno-Mateos MA, Johnstone TG, Vejnar CE, Qin Y, et al. Codon identity regulates mRNA stability and translation efficiency during the maternal-to-zygotic transition. EMBO J. 2016; 35: 2087–2103. https://doi.org/10.15252/embj.201694699 PMID: 27436874

7. Babu M, Teichmann SA. Evolution of transcription factors and the gene regulatory network in Escherichia coli. Nucleic Acids Res. 2003; 31: 1234–44. https://doi.org/10.1093/nar/gkg210 PMID: 12582243

8. Yona AH, Bloom-Ackermann Z, Frumkin I, Hanson-Smith V, Charpak-Amikam Y, Feng Q, et al. tRNA genes rapidly change in evolution to meet novel translational demands. Elife. 2013; 2: e01339. https://doi.org/10.7554/eLife.01339 PMID: 24363105

**9.** Kirchner S, Ignatova Z. Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. Nat Rev Genet. 2014; https://doi.org/10.1038/nrg3861 PMID: 25534324

**10.** Wray GA. The evolutionary significance of cis-regulatory mutations. Nat Rev Genet. 2007; 8: 206–16. https://doi.org/10.1038/nrg2063 PMID: 17304246

**11.** Lemos B, Araripe LO, Fontanillas P, Hartl DL. Dominance and the evolutionary accumulation of cis- and trans-effects on gene expression. Proc Natl Acad Sci U S A. 2008; 105: 14471–6. https://doi.org/10.1073/pnas.0805160105 PMID: 18791071

**12.** Wittkopp PJ, Haerum BK, Clark AG. Evolutionary changes in cis and trans gene regulation. Nature. 2004; 430: 85–8. https://doi.org/10.1038/nature02698 PMID: 15229602

**13.** Wong ES, Schmitt BM, Kazachenka A, Thybert D, Redmond A, Connor F, et al. Interplay of cis and trans mechanisms driving transcription factor binding and gene expression evolution. Nat Commun. 2017; 8: 1092. https://doi.org/10.1038/s41467-017-01037-x PMID: 29061983

**14.** Matera a. G, Wang Z. A day in the life of the spliceosome. Nat Rev Mol Cell Biol. 2014; 15: 108–121. https://doi.org/10.1038/nrm3742 PMID: 24452469

**15.** Merkin J, Russell C, Chen P, Burge CB. Evolutionary dynamics of gene and isoform regulation in Mammalian tissues. Science. 2012; 338: 1593–9. https://doi.org/10.1126/science.1228186 PMID: 23258891

**16.** Barbosa-Morais NL, Irimia M, Pan Q, Xiong HY, Gueroussov S, Lee LJ, et al. The evolutionary landscape of alternative splicing in vertebrate species. Science. 2012; 338: 1587–93. https://doi.org/10.1126/science.1230612 PMID: 23258890

**17.** Petibon C, Parenteau J, Catala M, Elela SA. Introns regulate the production of ribosomal proteins by modulating splicing of duplicated ribosomal protein genes. Nucleic Acids Res. 2016; 44: gkw140. https://doi.org/10.1093/nar/gkw140 PMID: 26945043

**18.** Parenteau J, Durand M, Morin G, Gagnon J, Lucier J-F, Wellinger RJ, et al. Introns within ribosomal protein genes regulate the production and function of yeast ribosomes. Cell. 2011; 147: 320–31. https://doi.org/10.1016/j.cell.2011.08.044 PMID: 22000012

**19.** Reyes a., Anders S, Weatheritt RJ, Gibson TJ, Steinmetz LM, Huber W. Drift and conservation of differential exon usage across tissues in primate species. Proc Natl Acad Sci. 2013; 2–7. https://doi.org/10.1073/pnas.1307202110 PMID: 24003148

**20.** Bush SJ, Chen L, Tovar-Corona JM, Urrutia AO. Alternative splicing and the evolution of phenotypic novelty. Philos Trans R Soc Lond B Biol Sci. 2017; 372. https://doi.org/10.1098/rstb.2015.0474 PMID: 27994117

**21.** Irimia M, Roy SW. Origin of spliceosomal introns and alternative splicing. Cold Spring Harb Perspect Biol. 2014; 6. https://doi.org/10.1101/cshperspect.a016071 PMID: 24890509

**22.** Roy SW, Irimia M. Mystery of intron gain: new data and new models. Trends Genet. 2009; 25: 67–73. https://doi.org/10.1016/j.tig.2008.11.004 PMID: 19070397

**23.** Hooks KB, Delneri D, Griffiths-Jones S. Intron evolution in Saccharomycetaceae. Genome Biol Evol. 2014; https://doi.org/10.1093/gbe/evu196 PMID: 25364803

**24.** Shabalina S a Ogurtsov AY, Spiridonov AN Novichkov PS, Spiridonov N a, Koonin E V. Distinct patterns of expression and evolution of intronless and intron-containing mammalian genes. Mol Biol Evol. 2010; 27: 1745–9. https://doi.org/10.1093/molbev/msq086 PMID: 20360214

**25.** Mitrovich QM, Tuch BB, De La Vega FM, Guthrie C, Johnson AD. Evolution of yeast noncoding RNAs reveals an alternative mechanism for widespread intron loss. Science. 2010; 330: 838–41. https://doi.org/10.1126/science.1194554 PMID: 21051641

**26.** Lee S, Stevens SW. Spliceosomal intronogenesis. Proc Natl Acad Sci U S A. 2016; 201605113. https://doi.org/10.1073/pnas.1605113113 PMID: 27217561

**27.** Derr LK, Strathern JN, Garfinkel DJ. RNA-mediated recombination in S. cerevisiae. Cell. 1991; 67: 355–64. https://doi.org/10.1016/0092-8674(91)90187-4 PMID: 1655280

**28.** Ding F, Elowitz MB. Constitutive splicing and economies of scale in gene expression. Nat Struct Mol Biol. 2019; 26. https://doi.org/10.1038/s41594-019-0226-x PMID: 31133700

**29.** Kress TL, Krogan NJ, Guthrie C. A Single SR-like Protein, Npl3, Promotes Pre-mRNA Splicing in Budding Yeast. Mol Cell. 2008; 32: 727–734. https://doi.org/10.1016/j.molcel.2008.11.013 PMID: 19061647

**30.** Chen Y-C, Milliman EJ, Goulet I, Côté J, Jackson CA, Vollbracht JA, et al. Protein arginine methylation facilitates cotranscriptional recruitment of pre-mRNA splicing factors. Mol Cell Biol. 2010; 30: 5245–56. https://doi.org/10.1128/MCB.00359-10 PMID: 20823272

**31.** Moehle EA, Ryan CJ, Krogan NJ, Kress TL, Guthrie C. The yeast SR-like protein Npl3 links chromatin modification to mRNA processing. PLoS Genet. 2012; 8: e1003101. https://doi.org/10.1371/journal.pgen.1003101 PMID: 23209445

**32.** Muddukrishna B, Jackson CA, Yu MC. Protein arginine methylation of Npl3 promotes splicing of the SUS1 intron harboring non-consensus 5' splice site and branch site. Biochim Biophys Acta. 2017; 1860: 730–739. https://doi.org/10.1016/j.bbagrm.2017.04.001 PMID: 28392442

**33.** Hackmann A, Wu H, Schneider U-M, Meyer K, Jung K, Krebber H. Quality control of spliced mRNAs requires the shuttling SR proteins Gbp2 and Hrb1. Nat Commun. 2014; 5: 3123. https://doi.org/10.1038/ncomms4123 PMID: 24452287

**34.** Zander G, Hackmann A, Bender L, Becker D, Lingner T, Salinas G, et al. MRNA quality control is bypassed for immediate export of stress-responsive transcripts. Nature. 2016; 540: 593–596. https://doi.org/10.1038/nature20572 PMID: 27951587

**35.** Ares M, Grate L, Pauling MH. A handful of intron-containing genes produces the lion's share of yeast mRNA. RNA. 1999; 5: 1138–9. https://doi.org/10.1017/s1355838299991379 PMID: 10496214

**36.** Yofe I, Zafrir Z, Blau R, Schuldiner M, Tuller T, Shapiro E, et al. Accurate, Model-Based Tuning of Synthetic Gene Expression Using Introns in S. cerevisiae. PLoS Genet. 2014; 10: e1004407. https://doi.org/10.1371/journal.pgen.1004407 PMID: 24968317

**37.** de Nadal E, Ammerer G, Posas F. Controlling gene expression in response to stress. Nat Rev Genet. 2011; 12: 833–45. https://doi.org/10.1038/nrg3055 PMID: 22048664

**38.** Metzl-Raz E, Kafri M, Yaakov G, Soifer I, Gurvich Y, Barkai N. Principles of cellular resource allocation revealed by condition-dependent proteome profiling. Elife. 2017; 6. https://doi.org/10.7554/eLife.28034 PMID: 28857745

**39.** Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, et al. Genomic expression programs in the response of yeast cells to environmental changes. Mol Biol Cell. 2000; 11: 4241–57. https://doi.org/10.1091/mbc.11.12.4241 PMID: 11102521

**40.** Sarkisyan KS, Bolotin DA, Meer M V., Usmanova DR, Mishin AS, Sharonov G V., et al. Local fitness landscape of the green fluorescent protein. Nature. 2016; 533: 397–401. https://doi.org/10.1038/nature17995 PMID: 27193686

**41.** Webb M, editor. Cancer Susceptibility [Internet]. Totowa, NJ: Humana Press; 2010. https://doi.org/10.1007/978-1-60761-759-4

**42.** Warf MB, Berglund JA. Role of RNA structure in regulating pre-mRNA splicing. Trends Biochem Sci. 2010; 35: 169–78. https://doi.org/10.1016/j.tibs.2009.10.004 PMID: 19959365

**43.** Gahura O, Hammann C, Valentová A, Půta F, Folk P. Secondary structure is required for 3' splice site recognition in yeast. Nucleic Acids Res. 2011; 39: 9759–67. https://doi.org/10.1093/nar/gkr662 PMID: 21893588

**44.** Lorenz R, Bernhart SH, Höner Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, et al. ViennaRNA Package 2.0. Algorithms Mol Biol. 2011; 6: 26. https://doi.org/10.1186/1748-7188-6-26 PMID: 22115189

**45.** Qian W, Yang JR, Pearson NM, Maclean C, Zhang J. Balanced codon usage optimizes eukaryotic translational efficiency. PLoS Genet. 2012; 8. https://doi.org/10.1371/journal.pgen.1002603 PMID: 22479199

**46.** Brackley CA, Romano MC, Thiel M. The dynamics of supply and demand in mRNA translation. PLoS Comput Biol. 2011; 7: e1002203. https://doi.org/10.1371/journal.pcbi.1002203 PMID: 22022250

**47.** Frumkin I, Lajoie MJ, Gregg CJ, Hornung G, Church GM, Pilpel Y. Codon usage of highly expressed genes affects proteome-wide translation efficiency. Proc Natl Acad Sci U S A. 2018; 115: E4940–E4949. https://doi.org/10.1073/pnas.1719375115 PMID: 29735666

**48.** Lei EP, Krebber H, Silver PA. Messenger RNAs are recruited for nuclear export during transcription. Genes Dev. 2001; 15: 1771–82. https://doi.org/10.1101/gad.892401 PMID: 11459827

**49.** Lee MS, Henry M, Silver PA. A protein that shuttles between the nucleus and the cytoplasm is an important mediator of RNA export. Genes Dev. 1996; 10: 1233–46. https://doi.org/10.1101/gad.10.10.1233 PMID: 8675010

**50.** Sverdlov A V, Babenko VN, Rogozin IB, Koonin E V. Preferential loss and gain of introns in 3' portions of genes suggests a reverse-transcription mechanism of intron insertion. Gene. 2004; 338: 85–91. https://doi.org/10.1016/j.gene.2004.05.027 PMID: 15302409

**51.** Cohen NE, Shen R, Carmel L. The role of reverse transcriptase in intron gain and loss mechanisms. Mol Biol Evol. 2012; 29: 179–86. https://doi.org/10.1093/molbev/msr192 PMID: 21804076

**52.** Zhu YO, Siegal ML, Hall DW, Petrov D a. Precise estimates of mutation rate and spectrum in yeast. Proc Natl Acad Sci U S A. 2014; 111: E2310–8. https://doi.org/10.1073/pnas.1323011111 PMID: 24847077

**53.** Rosenberg AB, Patwardhan RP, Shendure J, Seelig G. Learning the sequence determinants of alternative splicing from millions of random sequences. Cell. 2015; 163: 698–711. https://doi.org/10.1016/j.cell.2015.09.054 PMID: 26496609

54. Julien P, Miñana B, Baeza-Centurion P, Valcárcel J, Lehner B. The complete local genotype-phenotype landscape for the alternative splicing of a human exon. Nat Commun. 2016; 7: 11558. https://doi.org/10.1038/ncomms11558 PMID: 27161764

55. Soemedi R, Cygan KJ, Rhine CL, Glidden DT, Taggart AJ, Lin C-L, et al. The effects of structure on pre-mRNA processing and stability. Methods. 2017; 125: 36–44. https://doi.org/10.1016/j.ymeth.2017.06.001 PMID: 28595983

56. Hiller M, Zhang Z, Backofen R, Stamm S. Pre-mRNA secondary structures influence exon recognition. PLoS Genet. 2007; 3: e204. https://doi.org/10.1371/journal.pgen.0030204 PMID: 18020710

57. Munding EM, Shiue L, Katzman S, Donohue JP, Ares M. Competition between Pre-mRNAs for the Splicing Machinery Drives Global Regulation of Splicing. Mol Cell. 2013; 51: 338–348. https://doi.org/10.1016/j.molcel.2013.06.012 PMID: 23891561

58. Kudla G, Murray AW, Tollervey D, Plotkin JB. Coding-sequence determinants of gene expression in Escherichia coli. Science. 2009; 324: 255–8. https://doi.org/10.1126/science.1170160 PMID: 19359587

59. Voichek Y, Bar-Ziv R, Barkai N. Expression homeostasis during DNA replication. Science. 2016; 351: 1087–90. https://doi.org/10.1126/science.aad1162 PMID: 26941319

60. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009; 10: R25. https://doi.org/10.1186/gb-2009-10-3-r25 PMID: 19261174

61. Kivioja T, Vähärautio A, Karlsson K, Bonke M, Enge M, Linnarsson S, et al. Counting absolute numbers of molecules using unique molecular identifiers. Nat Methods. 2011; 9: 72–4. https://doi.org/10.1038/nmeth.1778 PMID: 22101854

62. Dekker B, Anand R, Memisoglu G, Haber J. Cas9-mediated gene editing in Saccharomyces cerevisiae. Protoc Exch. 2017; https://doi.org/10.1038/protex.2017.021a