Cytoplasmic control of sense-antisense mRNA pairs in Saccharomyces cerevisiae

Albertas Navickas

Université Pierre et Marie Curie

Doctoral School 515 *Complexité du vivant*

UMR 8226 – Laboratoire de biologie moléculaire et cellulaire des eucaryotes

Institut de biologie physico-chimique

13 rue Pierre et Marie Curie, 75005 Paris

Cytoplasmic control of sense-antisense mRNA pairs in *Saccharomyces cerevisiae*

Thesis submitted to attain the PhD degree in Biology

Presented by Albertas NAVICKAS

Supervised by Lionel BENARD

Defended on 23 September 2016

Thesis defense committee:

Mr NAMY Olivier, *directeur de recherche*, rapporteur

Mr PALANCADE Benoît, *chargé de recherche*, rapporteur

Mr DEVAUX Frédéric, *professeur d'université*, examiner

Mr JACQUIER Alain, *directeur de recherche*, examiner

Mr CONDON Ciarán, *directeur de recherche*, examiner

Mr BENARD Lionel, *chargé de recherche*, thesis supervisor

ACKNOWLEDGEMENTS

I would like to start by acknowledging the members of my thesis defense committee that have accepted to accord their time and attention to the work presented in this manuscript.

This thesis has been done in the research unit UMR8226, directed by Stéphane Lemaire, at *Institut de Biologie Physico-Chimique*, and funded by a doctoral grant from LABEX Dynamo. I would like thus to acknowledge all the members of our unit for their everyday help and support throughout the years I have spent at IPBC. I am particularly grateful to Isabelle Krempholtz, who has solved all the administrational problems before they even arose.

My experience within the RNA metabolism group has been an incredibly interesting and rich journey, and definitely my happiest scientific time so far—thanks to my thesis supervisor Lionel Benard, whose restless implication in the world of discoveries will remain an example for me. I have also had the best senior colleague I could wish for, Claire Torchet, who has accompanied me daily from the very beginning to the end of this work. I shall thank all the past and current members of our group—Flore Sinturel, Rénette Saint-Fort, and many interns—for their diverse contributions to this work.

I would like to thank the members of my thesis committee—Alice Lebreton, Cosmin Saveanu, and Ciarán Condon—for their support and useful discussions during my time at IBPC. I also thank our collaborators Antonin Morillon, Maxime Wery, and Marc Descrimes whose implication in the genome-wide experiments has been crucial for this work.

Finally, I would like to express my gratefulness to many other colleagues from the institute, my family and friends, whose daily support has been priceless throughout these years. I wish to specially thank my mother and my grandfather, who granted me the access to the wonderful world of science and have led me through it ever since.

ABSTRACT

Recent transcriptome analyses have revealed that convergent gene transcription can produce many 3' overlapping mRNAs in diverse organisms. This phenomenon has been studied in the context of nuclear RNA interference (RNAi) pathway, however little is known about the cytoplasmic fate of 3' overlapping messengers or their impact on gene expression. In this work, we address the outcomes of interaction between sense-antisense mRNA pairs in *Saccharomyces cerevisiae*, a model organism naturally devoid of RNAi. We demonstrate that the complementary tails of 3' overlapping mRNAs can interact in the cytoplasm in a sequence-specific manner and promote post-transcriptional remodeling of mRNA stability and translation. Our findings are based on the detailed analysis of a convergent mRNA pair, POR1 and OCA2, subsequently generalized using the reconstituted RNAi approach in *S. cerevisiae*. Genome-wide experiments confirm that in wild-type cells, sense-antisense mRNA pairs form RNA duplexes *in vivo* and thus have potential roles in modulating the respective mRNA or protein levels under different growth conditions. We show that the fate of hundreds of messenger-interacting messengers is controlled by Xrn1, revealing the extent to which this conserved 5'-3' cytoplasmic exoribonuclease plays an unexpected but key role in the post-transcriptional control of convergent gene expression. In sum, our work opens a perspective to consider an additional, cytoplasmic mechanism of interaction between sense-antisense mRNA pairs, in both RNAi-positive and negative organisms.

RÉSUMÉ

Les récentes études transcriptomiques chez divers organismes ont montré que la transcription des gènes convergents peut produire des ARN messagers (ARNm) chevauchants. Ce phénomène a été analysé dans le contexte de l'interférence par ARN (ARNi) nucléaire, et peu d'information existe quant au destin cytoplasmique des messagers 3' chevauchants ou leur impact sur l'expression des gènes. Dans ce travail, nous avons abordé les conséquences potentielles de l'interaction entre des paires d'ARNm sens-antisens chez *Saccharomyces cerevisiae*, un organisme modèle naturellement dépourvu de l'ARNi. Nous avons démontré que les extrémités 3' complémentaires des ARNm peuvent interagir dans le cytoplasme et moduler la stabilité ainsi que la traduction d'ARNm. Nos résultats sont issus d'une étude détaillée d'une paire d'ARNm convergents, POR1 et OCA2, ensuite généralisée par l'approche de l'ARNi reconstituée chez *S. cerevisiae*. L'analyse globale a confirmé que dans les cellules sauvages, les paires d'ARNm sens-antisens forment des duplexes d'ARN *in vivo* et ont un rôle potentiel à moduler l'expression d'ARNm ou de protéines respectifs, dans des différentes conditions de croissance. Nous avons montré que le destin de centaines des messagers convergents est contrôlé par Xrn1, révélant l'importance de cette exoribonucléase 5'-3' cytoplasmique très conservée dans la régulation post-transcriptionnelle des gènes convergents. Notre travail ouvre donc la perspective de considérer un nouveau mécanisme de l'interaction entre les paires d'ARNm sens-antisens dans le cytoplasme, chez les organismes contenant ou non la voie de l'interférence par ARN.

ABBREVIATIONS

aa-tRNA – aminoacyl-tRNA
ABC – ATP-binding cassette
ADAR – adenosine deaminase acting on RNA
ADP – adenosine diphosphate
AGO – Argonaute
ARE – A/U-rich element
ARE-BP – A/U-rich element binding protein
ATP – adenosine triphosphate
cAMP – cyclic adenosine monophosphate
CAT tails – C-terminal alanine threonine tails
CBC – cap binding complex
ChIP – chromatin immunoprecipitation
circRNA – circular RNA
CLIP – cross-link and immunoprecipitation
CPF-CF – cleavage and polyadenylation factor – cleavage factor
CTD – C-terminal domain of RNA Pol II
CUT – cryptic unstable transcript
disiRNA – Dicer-independent siRNA
DNA – deoxyribonucleic acid
dsRNA – double-stranded RNA
eEF – eukaryotic translation elongation factor
eIF – eukaryotic translation initiation factor
EJC – exon-junction complex
EMT – epithelial-mesenchymal transition
ER – endoplasmic reticulum
eRF – eukaryotic translation release factor
GDP – guanosine diphosphate
GO – gene ontology
GTP – guanosine triphosphate
IRES – internal ribosome entry site
lncRNA – long non-coding RNA

MAPK – mitogen-activated protein kinase
mimRNA – messenger-interacting messenger RNA
miRISC – miRNA-induced silencing complex
miRNA – micro RNA
mRNA – messenger RNA
mRNP – messenger ribonucleoprotein
ncRNA – non-coding RNA
NET-seq – native elongating transcript sequencing
NGD – no-go decay
NMD – nonsense-mediated decay
NNS – Nrd1-Nab3-Sen1 termination complex
NPC – nuclear pore complex
NSD – non-stop decay
ORF – open reading frame
P-bodies – processing bodies
pAp – adenosine 5',3' bisphosphate
PIC – pre-initiation complex
piRISC – piRNA-induced silencing complex
piRNA – PIWI-interacting RNA
PKA – protein kinase A
Pol II – RNA polymerase II
PPH – pyrophosphohydrolase
PTC – premature termination codon
PTGS – post-transcriptional gene silencing
RDRC – RNA-dependant RNA polymerase complex
RISC – RNA-induced silencing complex
RITS – RNA-induced transcriptional silencing
RNA – ribonucleic acid
RNA-seq – RNA sequencing
RNAi – RNA interference
RQC – ribosome quality control complex
rRNA – ribosomal RNA
SINE – short interspersed nuclear element

siRISC – siRNA-induced silencing complex

siRNA – small interfering RNA

SMD – spliceosome-mediated decay

smFISH – single-molecule fluorescence *in situ* hybridization

snoRNA – small nucleolar RNA

snRNA – small nuclear RNA

ssRNA – single-stranded RNA

StMD – STAU1-mediated decay

SUT – stable unannotated transcript

TE – transposable element

TF – transcription factor

TOR – target of rapamycin

TORC – target of rapamycin complex

TRAMP – Trf4/5-Air1/2-Mtr4 polyadenylation complex

tRNA – transfer RNA

tRNAi – initiator Methionine tRNA

TSS – transcription start site

TTS – transcription termination site

uORF – upstream open reading frame

UPR – unfolded protein response

UTR – untranslated region

VDAC – voltage-dependent anion channel

WGD – whole genome duplication

XUT – Xrn1-sensitive unstable stranscript

CHAPTER I | INTRODUCTION

Outline

The scope of this thesis is focused on the cytoplasmic fate of 3' complementary messenger RNAs (mRNAs) in a model organism *Saccharomyces cerevisiae*. To accompany a fluent reading of the results presented in this manuscript, the Chapter I is aimed to (1) provide the reader with the current view of the life cycle of a eukaryotic mRNA and to (2) guide through the origin and potential outcomes of mRNA-mRNA interactions currently described in the literature. As the work presented here originated at the intersection of mRNA degradation and translation studies, these two aspects of mRNA metabolism are at the core of the first part of the chapter. The mechanisms described and factors involved are mostly based on budding yeast studies and are extended to other model organisms where relevant. The second part of the chapter aims to systemize the proofs of existence and the mechanistic implications of mRNA-mRNA interactions in a broad spectrum of model organisms and to reposition these observations in the vast field of RNA-RNA interaction studies. Finally, the presentation of objectives of this study follows, encompassing a detailed case analysis of a convergent mRNA pair and a further attempt to generalize the principles of mRNA-mRNA interactions.

PART I | The life cycle of a eukaryotic mRNA

From the very beginning of its existence, an mRNA molecule undergoes a multitude of highly regulated processes that shape its entire life cycle. In this part, the nuclear events (mRNA synthesis, maturation and export to the cytoplasm) are briefly introduced and then followed by a more detailed presentation of the cytoplasmic processes (mRNA translation and turnover).

1 | The basics of mRNA metabolism in the nucleus

A brief outline of mRNA transcription, maturation and export

The transcription—synthesis of an mRNA molecule in the nucleus—is a multistage process that is usually viewed as a sequence of transcription initiation, elongation and termination. It is mediated by a twelve-subunit complex RNA polymerase II (Pol II), highly conserved throughout all eukaryotes (*1*). At the stage of initiation, a multitude of transcription factors (TFs) interact with the promoter sequence of a gene, to form the pre-initiation complex (PIC). This step is the most tightly controlled of the entire process, as in concert with general TFs, a high number of specific TFs act responding to distinct stages of the cellular program, environmental cues, various stresses, etc. The PIC, once bound to the promoter, undergoes specific conformational changes that result in the opening of the double-stranded DNA structure and the formation of the transcription bubble (*1*). This structure is thought to be transcription elongation-competent and engages in the production of a nascent mRNA (Figure 1). It is noteworthy that in contrast to the highly specific recruitment of TFs to a promoter, the directionality of the transcribing Pol II is often leaky and results in the widespread divergent transcription, generating various types of non-coding RNA (ncRNA, detailed further) (*2*). Moreover, it was equally shown that the exact selection of the transcription start site (TSS) in yeast has a low specificity. This generates an extensive 5' mRNA end heterogeneity, which is however actively controlled by the cytoplasmic quality control pathways (discussed further) (*3*).

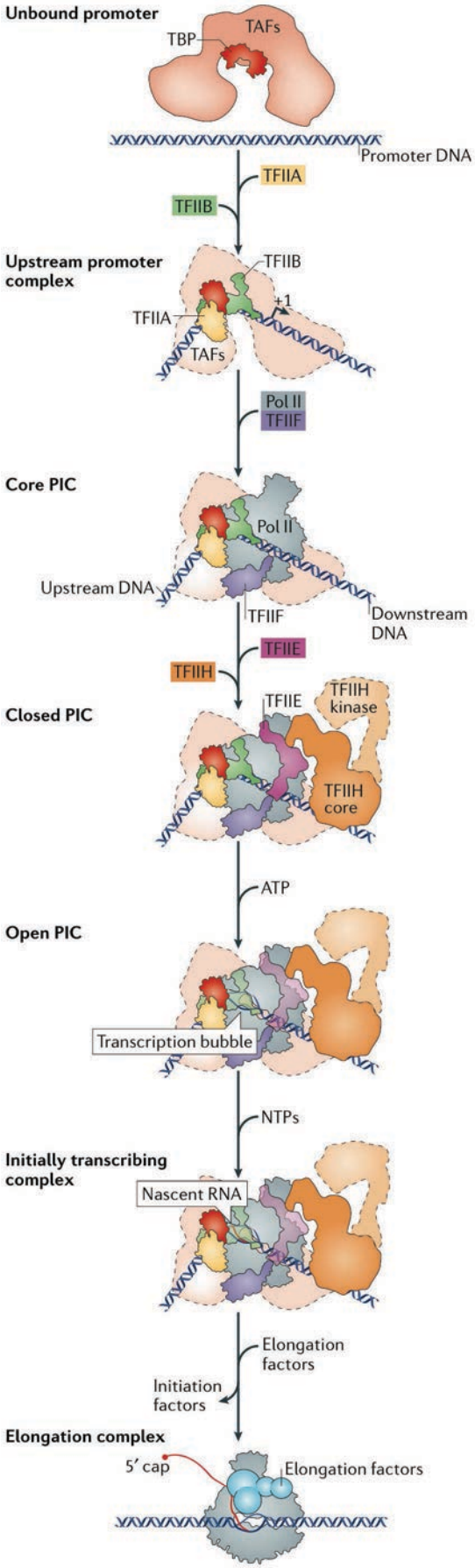
Unbound promoter
TBP
TAFs
Promoter DNA
TFIIA
TFIIB
Upstream promoter complex
TFIIB
+1
TFIIA
TAFs
Pol II
TFIIF
Core PIC
Pol II
Upstream DNA
TFIIF
Downstream DNA
TFIIE
TFIIH
Closed PIC
TFIIE
TFIIH kinase
TFIIH core
ATP
Open PIC
Transcription bubble
NTPs
Initially transcribing complex
Nascent RNA
Elongation factors
Initiation factors
Elongation complex
5′ cap
Elongation factors

Figure 1. The schematic of Pol II transcription initiation.

The canonical model for stepwise PIC assembly from general TFs (various colors) and RNA Pol II (grey) on promoter DNA is depicted. TFIID or its TATA box-binding protein (TBP) subunit binds to promoter DNA, inducing a bend. The TBP–DNA complex is then stabilized by TFIIB and TFIIA, which flank TBP on both sides. The resulting upstream promoter complex is joined by the Pol II–TFIIF complex, leading to the formation of the core PIC. Subsequent binding of TFIIE and TFIIH complete the PIC. In the presence of ATP, the DNA is opened (forming the "transcription bubble") and RNA synthesis begins. Finally, dissociation of initiation factors enables the formation of the Pol II elongation complex, which is associated with transcription elongation factors (blue). NTP, nucleoside triphosphate; TAF, TBP-associated factor. Adapted from (*1*).

The C-terminal domain (CTD) of Pol II plays a major role throughout the elongation step of transcription. Its phosphorylation status differs significantly at the early and late stages of elongation and a collection of elongation factors—CTD kinases, phosphatases and cyclins—modify it throughout the process. This status has an important role in adjusting CTD's affinity to various mRNA maturation factors, including capping, splicing and termination complexes, and thus mediates the crosstalk between the transcription elongation and RNA processing. The interplay between various chromatin remodelers and the differentially phosphorylated CTD adds an additional layer to the control of the elongating Pol II (*4*). Interestingly, the widespread promoter-proximal Pol II pausing, initially described in metazoan organisms and highly implicated in the regulation of transcription (*5*), is thought to be absent in *S. cerevisiae*, probably due to the lack of the negative elongation factor (NELF) (*6*).

When the elongating Pol II passes the poly(A) signal located in the 3' untranslated region (UTR) of a nascent mRNA, the CPF-CF (cleavage and polyadenylation factor – cleavage factor) complex is loaded and mediates the termination of transcription. Several CPF-CF subunits were shown to recognize directly termination and processing signals within the RNA sequence; the Pcf11 subunit mediates the interaction with a particular phosphorylation status of Pol II CTD. These interactions lead to the mRNA cleavage by the CPF nuclease Ysh1 and the rapid addition of a poly(A) tail at the 3' end of the mRNA by the CPF poly(A) polymerase, Pap1 (*7*). It was recently reported that the early association of the poly(A) tail with the nuclear poly(A)-binding protein, Nab2, is essential for the mRNA production (*8*). However, the major, primarily cytoplasmic poly(A) binding protein

Pab1 also binds mRNA poly(A) tails in the nucleus, protects it from exonucleolytic attacks and is important for the export to the cytoplasm (*7*). It is important to note that, contrary to the metazoan counterparts, the budding yeast poly(A) signals do not have a clear consensus sequence which gives rise to a significant mRNA 3' end heterogeneity; some regulatory roles have been proposed for different isoforms of the same mRNA (*9*).

After the mRNA is cleaved at the poly(A) site, Pol II continues the transcription and currently two mutually inclusive models propose the possible mechanisms of Pol II dissociation from the DNA. The allosteric model suggests that after the cleavage, binding of the termination complex induces a conformational change of the elongation complex and its subsequent release. On the other hand, the torpedo model is based on the importance of the nuclear 5'-3' exoribonuclease Rat1 that attacks the nascent RNA from the cleavage site up to the transcribing Pol II and contributes to the termination. The possibility that Rat1, reaching Pol II, might act as the allosteric factor has already been evoked (Figure 2) (*7*).

Rapidly after the exit of a nascent mRNA from the elongating Pol II, the capping enzyme is recruited to the 5' end of the transcript, via the interaction with the phosphorylated CTD of Pol II, as mentioned above (*4*). In *S. cerevisiae*, this enzyme is composed of Cet1, Ceg1 (Cet1 and Ceg1 forming a heterotetramer) and Abd1, conferring RNA 5'-triphosphatase, guanylyltransferase and guanine-N7 methyltransferase activities, respectively. After the conversion of 5'-triphosphate to 5'-diphosphate, a subsequent addition of GMP and its methylation at position N7, the m7GpppN cap structure is formed (*10*). This 5' structure is a specific feature of Pol II transcripts, acts as a protection against 5'-3' exonucleolytic attacks and ensures an efficient transport of an mRNA to the cytoplasm. In the nucleus, the cap-binding complex (CBC), a heterodimer that consists of Cbp20 and Cbp80, binds the cap and accompanies it during the export (*11*).

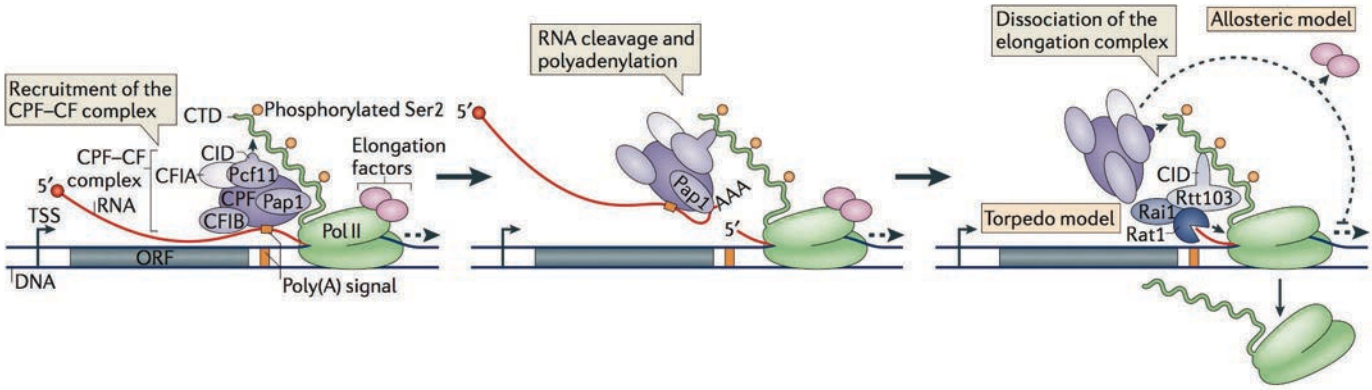

Figure 2. The schematic of Pol II transcription termination.

During termination at mRNA-coding genes, components of the cleavage and polyadenylation factor (CPF) and cleavage factor (CF) complexes—CPF, CFIA and CFIB—recognize specific sequences in the 3' UTR of the transcript. The CFIA component Pcf11 interacts with the Ser2-phosphorylated form of the CTD of RNA Pol II through its CTD-interaction domain (CID). Upon endonucleolytic cleavage of the transcript at the poly(A) site, poly(A) tails are added by the CPF-associated poly(A) polymerase Pap1. The 5' end of the downstream portion of the transcript is then targeted by the Rat1 5'-3' exonuclease. Two alternative models are proposed for the mechanism of termination after transcript cleavage. The allosteric model posits that loss of elongation factors and/or conformational changes in the polymerase after transcription of the poly(A) signal destabilizes the elongation complex. The torpedo model postulates that the Rat1 exonuclease (alone or in complex with its cofactor Rai1) is recruited by the CTD-interacting protein Rtt103 and degrades the nascent RNA after cleavage. The interaction of Rat1 with the polymerase leads to the dissociation of the elongation complex. Adapted from (*7*).

Along the 5' cap structure formation, mRNA splicing and export-competent mRNP (mRNA particle) assembly are essential steps of co-transcriptional mRNA maturation. Similarly to the capping enzyme recruitment, splicing machinery is also linked to the differentially phosphorylated CTD (*4*). The correct splicing is of a particular importance in metazoans as nearly all genes are intronic in these organisms. The exon-junction complex (EJC), deposited on mRNA during splicing events, is a core element of a mature mRNP and is crucial during the quality control of a messenger in the cytoplasm (*12*). In *S. cerevisiae*, introns are far less prevalent; the EJC complex is not conserved, although it was shown that splicing occurs co-transcriptionally, and the SR-like (serine-arginine rich) factor Npl3 links it to the further steps of mRNA assembly. Npl3 binds mRNA directly and, depending on its phosphorylation status, can recruit the general mRNA export receptor, Mex67-Mtr2 complex. However, this complex—the key factor of mRNA export—is usually

recruited by a multifactorial THO/TREX complex, associating co-transcriptionally with a nascent mRNP (*13*).

The subunits of the THO/TREX complex act as a bridge between the mRNA (via Sub2-Yra1), the chromatin (via the components of THO subcomplex) and the export receptor Mex67-Mtr2. Interestingly, the nuclear poly(A) binding protein Nab2 was shown to interact with Mex67 and Yra1, emphasizing the link between the mRNA 3' end formation and export. An additional complex, TREX-2, interacting with the chromatin remodeler complex SAGA and Mex67-Mtr2, is implicated in the coupling of mRNA synthesis and export (*14*). Subsequently, the mature mRNP is addressed to the nuclear pore complex (NPC), a megadalton-grade multiprotein assembly, that accompanies its transfer to the cytoplasm (Figure 3). Importantly, several NPC associated factors ensure the quality control of the mRNP prior the final export step (e.g. Mlp1-Mlp2 complex causes the retention of misspliced mRNA and targets it to the nuclear degradation via Swt1, a NPC associated endoribonuclease) (*15*). The multitude of factors (frequently showing evolutionary conservation) involved in mRNP maturation and export witnesses the high complexity of the process and enables its significant regulatory potential (*16*).

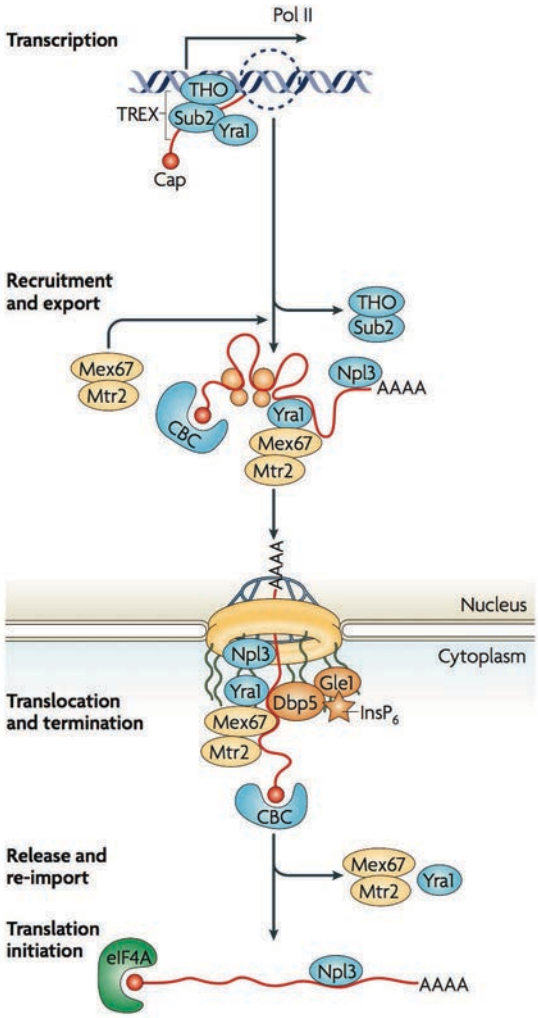

Figure 3. The schematic of nuclear mRNP maturation and export.

The nascent transcript generated by RNA Pol II is co-transcriptionally folded and assembled into a pre-mRNP by the THO/TREX complex (Yra1 and Sub2 are subunits of TREX, and THO is a subcomplex) that accompanies the elongating Pol II. After co-transcriptional association with additional pre-mRNA factors (e.g. the cap-binding complex (CBC) and RNA-binding proteins, shown in orange), the Mex67-Mtr2 mRNA export receptor is recruited to the mRNP via adaptor proteins such as Yra1. When emerging at the cytoplasmic side of the nuclear pore complex (NPC), the mRNP encounters a remodeling machinery on the NPC fibrils that consists of the ATP-dependent RNA helicase Dbp5 and its activators Gle1 and the signaling molecule inositol hexakisphosphate (InsP6). After dissociation of Mex67-Mtr2 and other export factors from the RNA cargo, the mRNA may still contain a few shuttling RNA-binding proteins, such as Npl3, that influence translation. Transport factors are re-imported, which in some cases requires phosphorylation. Adapted from (*13*).

Nuclear RNA decay pathways

Along the multiple steps of mRNA synthesis, maturation and mRNP assembly, a potent nuclear decay machinery acts to assure the timely processing, quality control and turnover of the RNA pool. Nuclear 5'-3' and 3'-5' exoribonuclease complexes together with several endoribonucleases are of the particular importance for ribosomal (rRNA), small nuclear (snRNA), small nucleolar (snoRNA), transfer (tRNA) and other RNA species maturation. On the other hand, messenger RNA is thought to be rapidly protected from nuclear degradation pathways by the cap structure and associated CBC at the 5' end, and by the poly(A) tail and associated Nab2 or Pab1 at the 3' end, as mentioned above. However, nuclear exoribonucleases contribute significantly to the surveillance of incorrectly maturated mRNA (e.g. containing defective 5' cap, misspliced, non-canonically terminated and thus poly(A) tailless, etc.) and also have a role in controlling gene expression (*17*). Finally, nuclear decay machinery is a key mechanism in limiting pervasive transcription, a eukaryotic phenomenon shown to produce a multitude of ncRNA genome-wide (*18*).

5'-3' exoribonucleolytic activity in the nucleus is mediated by Rat1, an evolutionarily conserved and essential protein (known as Xrn2 in other organisms). It is important for transcription termination, rRNA and other ncRNA processing. This factor also has a role in telomere maintenance, as Rat1 is a key nuclease in eliminating telomeric repeat-containing RNA (TERRA), a ncRNA species inhibiting the activity of the telomere elongating enzyme, telomerase. Several co-factors of Rat1 have been described: Rai1 was shown to form a complex with Rat1 and enhance its nucleolytic activity *in vivo* and *in vitro*; Rtt103 was discovered to bind the Rat1-Rai1 complex and to recruit it to the terminating Pol II via its phosphorylated CTD (*19*). However, a very important function of Rai1 was revealed when it appeared to have a pyrophosphohydrolase (PPH) as well as decapping activity. Kiledjian and colleagues provided significant data showing that mRNA capping is not a uniform process, and a detectable portion of coding transcripts has an aberrant 5' end. Specifically, mRNAs bearing unmethylated cap (GpppN) or triphosphate (pppN) structure undergo its removal by Rai1, yielding a 5'-monophosphate extremity that is subsequently degraded by Rat1 (*20, 21*).

Recently, a yeast specific Rai1 homologue protein Dxo1 was described. It has a low sequence similarity to Rai1 although shares highly similar structure. Differently from Rai1, Dxo1 does not show any detectable PPH activity but is active at removing unmethylated 5' caps. Interestingly, it also possesses a distributive 5'-3' exoribonucleolytic activity and was equally detected in the cytoplasm. Function-wise, Dxo1 is partially redundant with Rai1, having a major role in mRNA 5' structure surveillance; it is however unknown to what extent its exoribonuclease activity is significant *in vivo* (*22*).

Differently from individually active nuclear 5'-3' exoribonucleases, 3'-5' decay pathway is mediated by a large multiprotein complex, RNA exosome. It is composed of nine-subunit catalytically inactive barrel-like core (also called Exo9) and two catalytic subunits, Dis3 (Rrp44) and Rrp6 (Figure 4). Exo9-Dis3 complex (also referred to as Exo10) is found both in the nucleus and in the cytoplasm. Rrp6 localizes exclusively to the nucleus and thus Exo10-Rrp6 (or Exo11) complex is usually referred to as the nuclear RNA exosome (*17*). Both Dis3 and Rrp6 are 3'-5' exoribonucleases, and Dis3 possesses additional endoribonucleolytic activity (*23*); these three activities were shown to be partially redundant but also have some specific targets (*24*). Generally, RNA exosome is thought to have a relatively low substrate specificity and relies on multiple accessory factors for its diverse functions. It is also known that Dis3 and Rrp6 association with the Exo9 complex *in vitro* lowers its processivity and increases its sensitivity to 3' proximal secondary structures, emphasizing the importance of its co-factors (*17*).

The discovery of the TRAMP complex, an essential co-factor of the nuclear RNA exosome, lead to a significant advance in the understanding of its activity. TRAMP contains one of two homologous non-canonical poly(A) polymerases, Trf4 and Trf5, an DExD/H family RNA helicase Mtr4, and one of two homologous zinc knuckle domain proteins, Air1 and Air2. It polyadenylates substrate RNA 3' end and makes it accessible to the core exosome (*25, 26*). TRAMP complex is essential for the 3' end maturation of snRNA and snoRNA, as the transcription of these ncRNAs is terminated by the NNS (Nrd1-Nab3-Sen1) pathway, an alternative mechanism of Pol II dissociation (discussed below) (*7*). Recently, two additional co-factors recruiting TRAMP and subsequently the nuclear RNA exosome were identified: Utp18 and

Nop53 recognize a specific domain of Mtr4 and target the entire complex to distinct rRNA maturation intermediates (*27*).

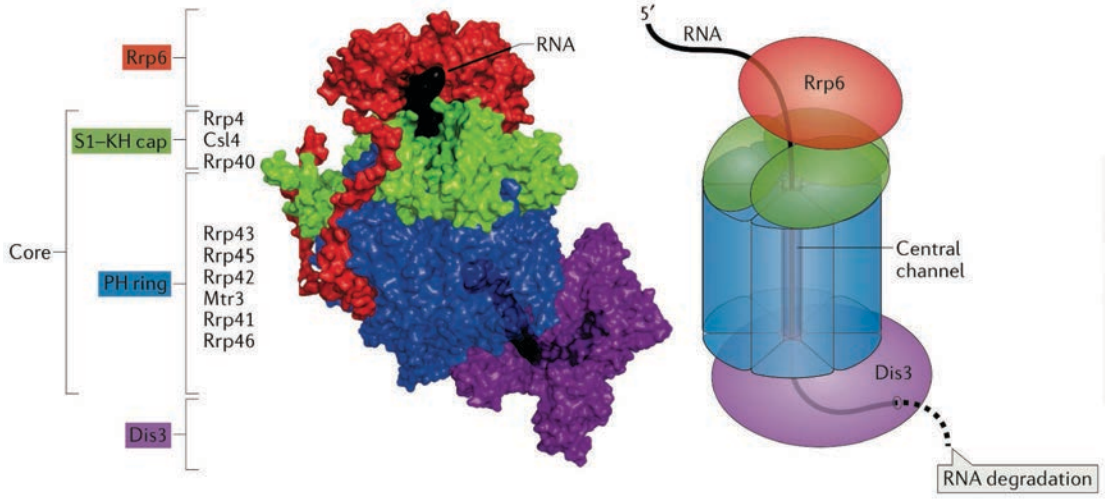

Figure 4. The structure of the nuclear RNA exosome.

On the left, a model of the exosome Exo11 complex, containing the universal Dis3 catalytic subunit and the nuclear Rrp6 subunit. On the right, schematics of an RNA molecule threading 3' to 5' through the central channel of the cap and PH ring to the Dis3 exonucleolytic center, where it is degraded. Adapted from (*17*).

The NNS transcription termination pathway is an interesting example of a direct link between RNA synthesis and processing. Nrd1 and Nab3, the subunits of the NNS complex, are RNA binding proteins that recognize distinct A/U rich consensus sequences on a nascent transcript. RNA helicase Sen1 is then recruited which promotes termination by dissociating Pol II. Importantly, NNS complex binds to the Pol II CTD via the CID domain of Nrd1 that has a high affinity to the phosphorylation status of CTD specific for the early stage of transcription. It is coherent with the fact that the pathway is primarily involved in the 3' end biogenesis of short snRNA and snoRNA. However, a large amount of unstable transcripts, known as CUTs (cryptic unstable transcripts) in *S. cerevisiae*, also undergoes an early termination by the NNS complex. Subsequently, Ndr1 interaction with the CTD is replaced by an interaction with Trf4 and the products of early termination are readily polyadenylated by the TRAMP complex and degraded (the case of CUTs) or processed (the case of snRNA/snoRNA) by the exosome (*7*).

The nuclear 3'-5' decay machinery is equally important for the early quality control of mRNA. In the case of the poly(A) signal read-through, mRNA is not cleaved by the PCF-CF complex and the transcription terminates by a non-canonical mechanism downstream. Most commonly, the NNS complex is recruited at the Nrd1-Nab3 binding sites but alternative pathways, including the roadblock mechanism (when transcription is blocked by a DNA binding factor, such as Reb1) and a Rnt1 (a double-stranded RNA (dsRNA) endonuclease) mediated cleavage, have been described. Independently of the termination mechanism, such RNAs are immediately targeted to degradation by the exosome. As mentioned above, misspliced mRNA or incorrectly assembled mRNP are also targeted to nuclear 3'-5' degradation, and the role of certain SR proteins, THO subcomplex and TRAMP was suggested in this pathway (*17*).

The regulatory potential of gene expression by the nuclear RNA exosome has been also addressed, although only few examples have been described in details. An interesting case is the feedback regulation of *NRD1* expression, as its mRNA contains multiple NNS binding sites. Accordingly, when the Nrd1 level is low, NRD1 mRNA transcription is canonically terminated and leads to the expression of Nrd1 protein. However, when the Nrd1 level is high, it mediates its own mRNA premature termination and rapid nuclear decay. Another example is the regulation of bromodomain factor *BDF2* expression. BDF2 mRNA is intronless but contains a cryptic splicing site, and spliceosomes are recruited to this site. However, the splicing is inefficient and its intermediates are targeted to degradation by the nuclear exosome. The process has been referred to as spliceosome-mediated decay (SMD) and associated to the stress response. Finally, the nuclear 3'-5' decay factor Rrp6 has been implicated in the cellular program related processes. It was shown to have a key role in eliminating histone mRNAs at the exit of the S phase (*17*), and act in the onset of the meiotic program (*28*).

2 | The essentials of eukaryotic translation

A brief outline of translation initiation, elongation, termination and ribosome recycling

As soon as a mature mRNP reaches cytoplasm it is extensively remodeled. NPC cytoplasmic fibrils associated factors mediate this mRNP composition change, and the DEAD-box RNA helicase Dbp5 plays a key role in the process. The dissociated mRNA export receptor Mex67-Mtr2, along with the shuttling TREX complex factors, are then reimported to the nucleus, and the differential phosphorylation status of Npl3 acts as a sensor of the shuttling direction (*13*). The subsequent events discussed hereafter follow the distinct stages of translation, the major cytoplasmic event in the life cycle of an mRNA (Figure 5).

The nuclear CBC is replaced by the cytoplasmic cap binding protein, the translation initiation factor eIF4E. Together with eIF4G and eIF4A, it forms the eIF4F complex, known to activate an mRNA for translation. eIF4G is a multidomain scaffold protein, that recruits not only its eIF4F partners but is also important for the association with other initiation factors. eIF4G interaction with Pab1, bringing mRNA 5' and 3' extremities together, has been the basis of a long standing "mRNA loop" model, suggesting the circular flow of translation processes. However, as the eIF4G-Pab1 interaction is domain specific and thus can be experimentally disrupted, it appeared that the cellular significance of this association is less general and rather cell type specific (*30*).

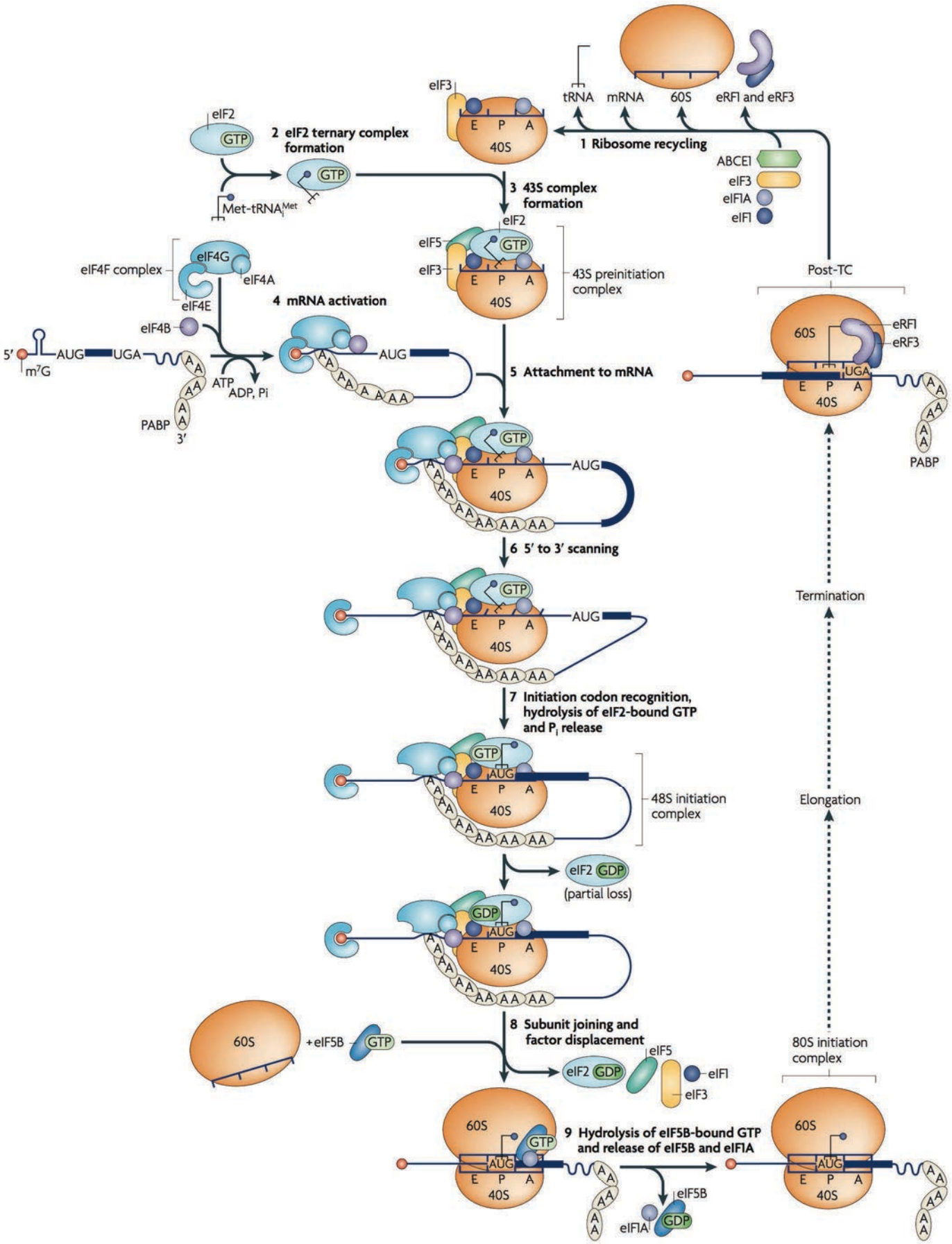
eIF3
tRNA
mRNA
60S
eRF1 and eRF3
eIF2
GTP
2 eIF2 ternary complex formation
1 Ribosome recycling
ABCE1
eIF3
eIF1A
eIF1
Met-tRNA$_i^{Met}$
3 43S complex formation
eIF2
eIF5
eIF3
43S preinitiation complex
Post-TC
eIF4F complex
eIF4G
eIF4A
eIF4E
4 mRNA activation
eIF4B
5′
m^{7}G
AUG
UGA
PABP
3′
ATP
ADP, Pi
5 Attachment to mRNA
60S
eRF1
eRF3
6 5′ to 3′ scanning
7 Initiation codon recognition, hydrolysis of eIF2-bound GTP and P$_i$ release
48S initiation complex
eIF2 GDP
(partial loss)
Termination
Elongation
8 Subunit joining and factor displacement
+eIF5B
eIF5
eIF1
80S initiation complex
9 Hydrolysis of eIF5B-bound GTP and release of eIF5B and eIF1A
E P A
40S

Figure 5. The schematic of the canonical pathway of eukaryotic translation initiation.

The canonical pathway of eukaryotic translation initiation is divided into eight stages. These stages follow the recycling of post-termination complexes (1) to yield separated 40S and 60S ribosomal subunits, and result in the formation of an 80S ribosomal initiation complex, in which initiator tRNA is base-paired with the initiation codon in the ribosomal P-site and which is competent to start the translation elongation stage. These stages are: eIF2-GTP-$tRNA_i$ ternary complex formation (2); formation of a 43S PIC comprising a 40S subunit, eIF1, eIF1A, eIF3, eIF2-GTP-$tRNA_i$ and probably eIF5 (3); mRNA activation, during which the mRNA cap-proximal region is unwound in an ATP-dependent manner by eIF4F with eIF4B (4); attachment of the 43S complex (6); recognition of the initiation codon and 48S initiation complex formation, which switch the scanning complex to a "closed" conformation and leads to displacement of eIF1 to allow eIF5-mediated hydrolysis of eIF2-bound GTP and phosphate release (7); joining of 60S subunits to 48S complexes and concomitant displacement of eIF2-GDP and other factors (eIF1, eIF3, eIF4B, eIF4F and eIF5) mediated by eIF5B (8); and GTP hydrolysis by eIF5B and release of eIF1A and GDP-bound eIF5B from assembled elongation-competent 80S ribosomes (9). Translation is a cyclical process, in whish termination follows elongation and leads to recycling (1), which generates separated ribosomal subunits. Adapted from (*29*).

An integral part of the eIF4F complex is a DEAD-box RNA helicase eIF4A, equally recruited by the scaffold subunit eIF4G. eIF4A has a key role in disrupting potential secondary structures present within the 5' UTR of an mRNA, an obstacle for the subsequent 48S scanning (discussed below). The helicase activity is ATP dependent, and another initiation factor, eIF4B, enhances the activity of eIF4A. Among the eIF4G interactants, an additional DEAD-box RNA helicase Ded1 was shown to be important for the translation of long 5' UTR bearing mRNA. Ded1 is an essential protein in budding yeast and has only partial target redundancy with eIF4A, suggesting its importance in the translation initiation. Moreover, several metazoan specific helicases (e.g. DHX29) and eIF4G interacting factors (e.g. p97), important for the initiation step, have been described (*29*).

In parallel to the eIF4F association to the 5' end of an mRNA, the 43S pre-initiation complex (PIC) assembly takes place. It begins with the formation of the eIF2-GTP-Met-$tRNA^{Met}$ (initiator methionine tRNA) ternary complex that brings the initiator tRNA ($tRNA_i$) to the 40S ribosomal subunit. GTP binding to the eIF2γ subunit confers the specificity to the methionine moiety of $tRNA_i$, and the presence of eIF2α-eIF2β in the complex significantly increases the affinity of this interaction.

The highly specific interaction between the eIF2-GTP and $tRNA_i$ is thought to protect 40S from the elongator tRNA association. However, the ternary complex cannot bind 40S alone and multiple additional initiation factors accompany this binding. eIF1 and eIF1A stabilize 40S in an "open" conformation that can easily bind the ternary complex, and eIF1A was shown to directly recruit eIF2. eIF1A also interacts with a multiprotein eIF3 that is important for the 43S PIC recruitment to the activated mRNA via eIF4G. Additionally, 43S PIC contains eIF5, a GTPase-activating factor that is important for the later steps of mRNA scanning and start codon recognition (discussed further) (*30*).

Once the eIF2-GTP-Met-$tRNA^{Met}$ ternary complex is recruited to the eIF1, eIF1A, eIF3 and eIF5 containing 40S, the 43S PIC is formed. It is still unclear if the subsequent activated (eIF4F containing) mRNA recognition by the 43S PIC is a distinct stage of translation initiation or if eIF4F recruitment to the PIC is happens simultaneously with the 43S assembly. However, this association is a key step in initiating the translation process as it enables 40S scanning of the 5' UTR. Upon the start codon (AUG) recognition, eIF2-bound GTP is hydrolyzed to GDP and this process is activated by eIF5. eIF2-GTP is later recovered by the guanosine nucleotide exchange factor eIF2B. GDP-bound form of the ternary complex in the 43S PIC induces its conformational change into the "closed" state, stabilizing AUG and $tRNA_i$ anticodon interaction. The importance of the sequence context surrounding mRNA start codon has been also described. The actual complex is called 48S initiation complex, containing the 40S and initiation factors assembled on the mRNA, with the $tRNA_i$ positioned in the ribosome P site and paired with the AUG codon (*30*).

As evoked before, eIF2-GDP has a low affinity to the $tRNA_i$ and thus dissociates from the 48S. This dissociation is promoted by the eIF5B, a translational GTPase that also accelerates 60S joining. It was suggested that eIF5 leaves the 48S together with the eIF2-GDP complex and frees the eIF1A domain that serves for the eIF5B-GTP recruitment. Upon the 40S and 60S joining, eIF5B-bound GTP is hydrolyzed and the low affinity of eIF5B-GDP to 80S triggers its release. Subsequently, eIF1A dissociates from the 80S assembly; however, the order and timing of eIF3 or eIF1 release remain unclear (*30*).

The dissociation of translation initiation factors results in the translation elongation competent 80S ribosome assembled at the mRNA start codon. The elongation cycle starts with the aminoacyl-tRNA (aa-tRNA) activation by its interaction with the GTP-bound translation elongation factor eEF1A (Figure 6). This factor directs an activated aa-tRNA to the vacant A site of the ribosome. Upon codon recognition, eEF1A-bound GTP hydrolysis is triggered, leading to eEF1A-GDP release and aa-tRNA accommodation in the A site. eEF1A-GTP is later recovered by the guanosine nucleotide exchange factor eEF1B. Rapidly after the aa-tRNA positioning, peptide bond formation with the P-site $tRNA_i$ (later on—peptidyl-tRNA) occurs. This reaction takes place in the peptidyl transferase center, located within the 60S ribosomal subunit and composed primarily of evolutionarily conserved rRNA elements (*31*).

Following the peptide bond formation, the ratcheting of the ribosome positions the A and P site tRNAs into the hybrid A/P and P/E states, respectively. In this hybrid conformation, the acceptor ends of tRNAs move into P and E sites of the ribosome while the anticodon loops remain in the A and P sites. To translocate the tRNAs into the canonical P and E sites, the eEF2 elongation factor is required. The GTP-bound eEF2 is thought to stabilize the hybrid state of the ribosome, triggering GTP hydrolysis and leading to the subsequent conformational changes. These changes position the ribosome in the post-translocation state (with deacylated tRNA in the E site and peptidyl-tRNA in the P site) and promote sequential P_i and eEF2-GDP release. The post-translocation state ribosome has a vacant A site, ready to accommodate the next eIF1A-activated aa-tRNA (*31*).

Importantly, *S. cerevisiae* ribosomes require a yeast-specific translation elongation factor eEF3, that is not present in metazoan organisms neither bacteria nor archaea. No homologous proteins were detected in these taxa during comparative studies, and eEF3 is thought to be a distinct case of fungal translation mechanism. Differently from eEF1 and eEF2 associating with the A site, eEF3 was found to interact at the proximity to the E site, and it was suggested that this factor contributes to the deacylated tRNA release from the ribosome. However, it is unclear why eEF3 is essential to yeast translation elongation but is dispensable to otherwise highly conserved metazoan or prokaryotic ribosomes (*31*).

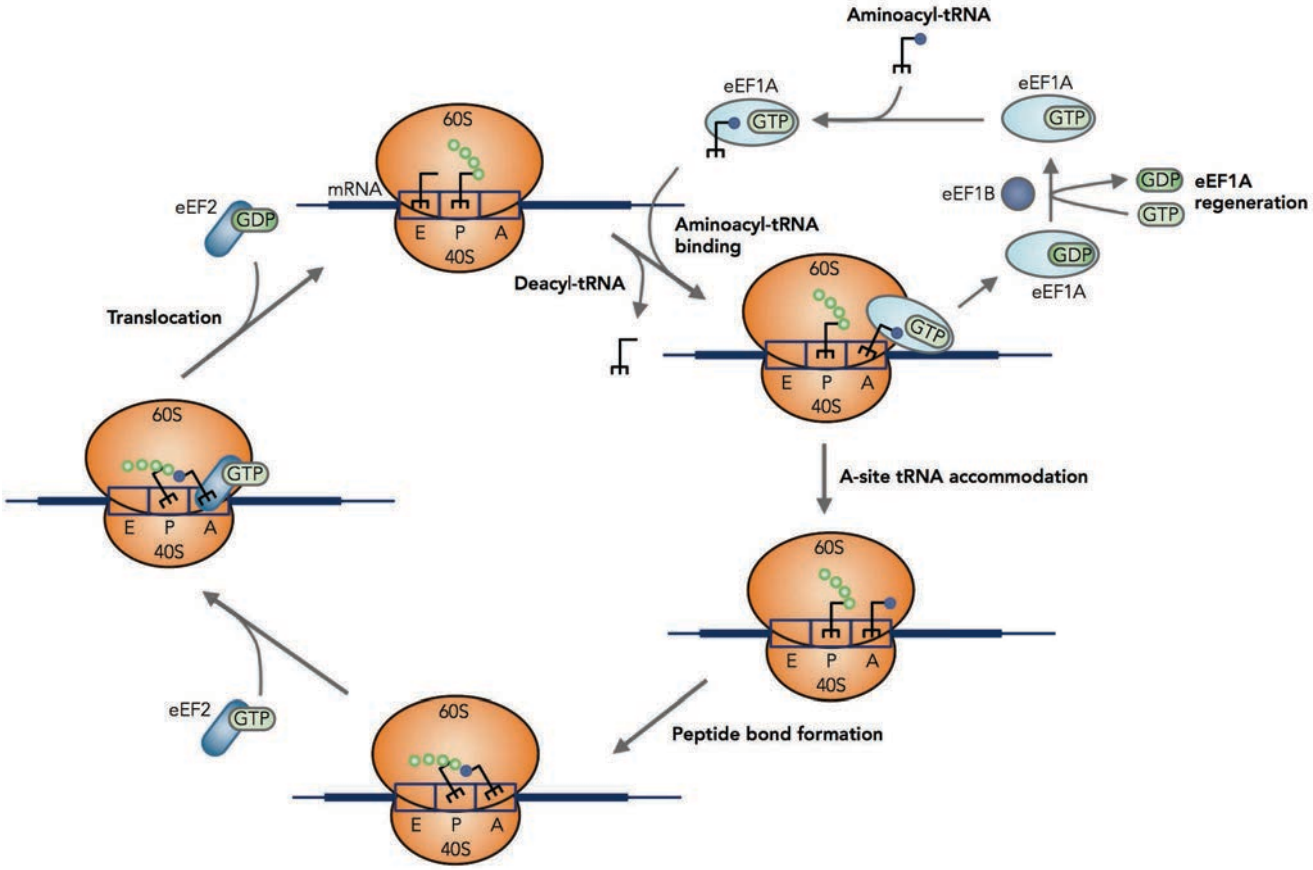

Figure 6. The model of the eukaryotic translation elongation pathway.

Starting at the top, an eEF1A-GTP-aminoacyl-tRNA ternary complex binds the aminoacyl-tRNA to the 80S ribosome with the anticodon loop of the tRNA in contact with the mRNA in the A site of the small subunit. Following release of eEF1A-GDP, the aminoacyl-tRNA is accommodated into the A site, and the eEF1A-GDP is recycled to eEF1A-GTP by the exchange factor eEF1B. Peptide bond formation is accompanied by transition of the A- and P-site tRNAs into hybrid states with the acceptors ends of the tRNAs moving to the P and E sites, respectively. Binding of eEF2-GTP promotes translocation of the tRNAs into the canonical P and E sites, and is followed by release of eEF2-GDP, which unlike eEF1A does not require an exchange factor. The ribosome is now ready for the next cycle of elongation with release of the deacylated tRNA from the E site and binding of the appropriate eEF1A-GTP-aminoacyl-tRNA to the A site. Adapted from (*31*).

At the beginning identified as a translation initiation factor, eIF5A proved to be an essential to eukaryotic translation elongation. It bears a unique post-translational modification on its conserved lysine residue that is transformed into hypusine during a two-stage mechanism. Recent structural studies revealed that eIF5A associates with the E site of the ribosome, and its hypusine moiety contacts the peptidyl transferase center (*32*). It was shown that eIF5A is crucial for the translation of poly-proline tracts as the ribosome translocates inefficiently due to the proline-proline bond

stereochemistry. As multiple essential proteins contain more than three subsequent proline residues (sufficient to induce ribosome stalling), it explains the eukaryotic necessity of eIF5A (*33*). Accordingly, as bacteria code less poly-proline containing essential proteins, eIF5A prokaryotic homologue EF-P was shown to be dispensable in *Escherichia coli* (*34*).

When a translating ribosome encounters a stop codon (UGA, UAG or UAA) in its A site, translation termination step occurs (Figure 7). It is mediated by two termination factors, eRF1 and eRF3, structurally similar to the tRNA-eEF1 complex. eRF1 is responsible for the stop codon recognition and subsequently catalyzes peptidyl-tRNA hydrolysis. Its three-domain structure permits the separation of these functions: at the N-terminus, the highly conserved NIKS motif mimics the codon-anticodon interactions; the middle domain contains a universal GGQ motif, essential for the polypeptide release; finally, the C-terminal domain recruits the eRF3. The latter factor is a translational GTPase, and the eRF1-eRF3-bound GTP hydrolysis positions eRF1 within the A site, triggering eRF3-GDP release (*31*).

Upon eRF3 dissociation, the 80S ribosome is still bound to the mRNA, and contains deacylated tRNA in the P site and eRF1 in the A site. This configuration is described as a post-termination complex, competent for the recycling of ribosomal subunits. An ATP-binding cassette (ABC) protein Rli1, universally conserved throughout eukaryotes (ABCE1 in metazoans), plays a key role in the recycling step of translation. After eRF3 release, Rli1 interacts with eRF1-bound ribosome and stimulates the peptidyl-tRNA hydrolysis. It subsequently uses ATP hydrolysis to dissociate 80S into subunits and restart the translation cycle. Although it is known that eRF1-eRF3 are capable of dissociating ribosomal subunits *in vitro*, this activity is highly stimulated by Rli1 and is ATP dependent (*31*).

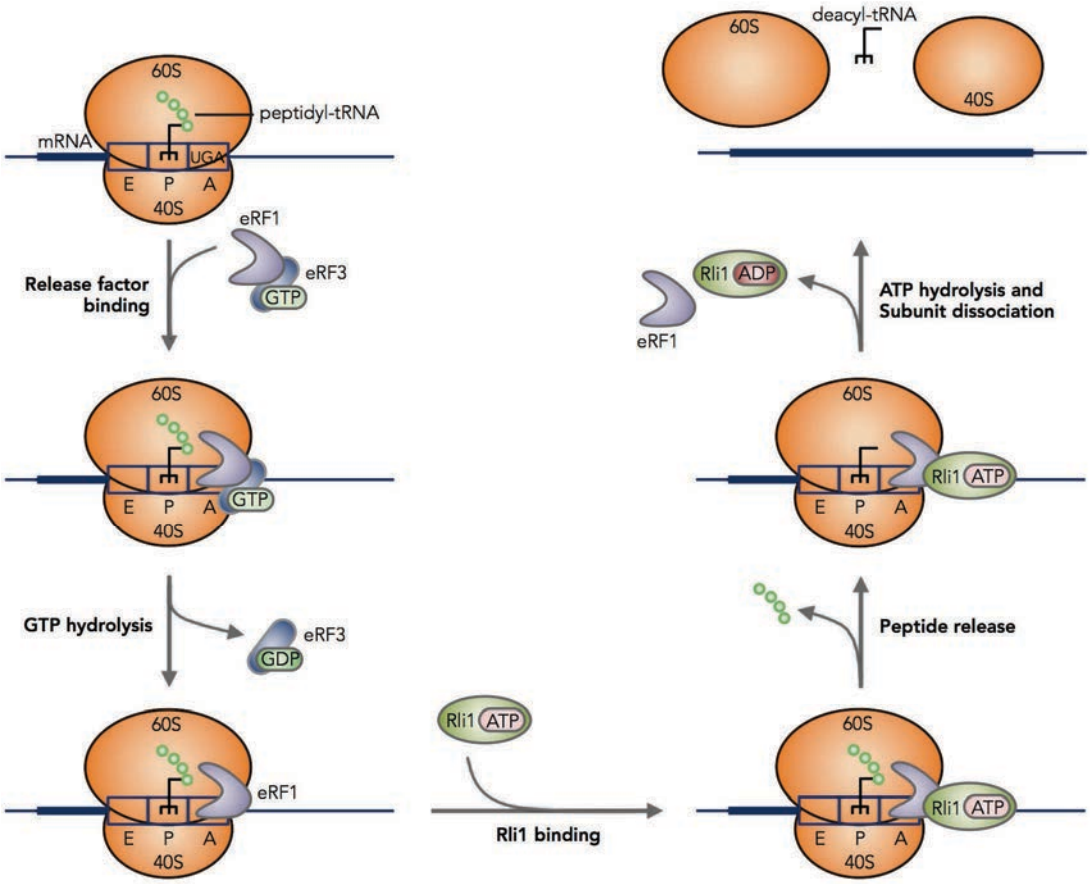

Figure 7. The model of the eukaryotic translation termination and recycling pathways.

Upon recognition of a stop codon, the eRF1-eRF3-GTP ternary complex binds to the A site of the ribosome in a pre-accommodated state, GTP hydrolysis occurs, and eRF3 is released. Rli1 binds and facilitates the accommodation of eRF1 into an optimally active configuration. Subsequently, peptidyl-tRNA is hydrolyzed, and, upon ATP hydrolysis, ribosomal subunits are split, to restart a new translation cycle. Adapted from (*31*).

The mechanisms of translational control

Along the various stresses and environmental cues-responding transcriptional control, translational regulation adds an effective and dynamic layer to controlling gene expression. Various studies have demonstrated that upon glucose deprivation, amino acid starvation, heat shock, etc., the global landscape of cellular translation changes significantly; in most cases, the further detailed analysis of distinct translational outcomes lead to the identification of general translational control mechanisms (*35*).

At the translation initiation stage, the most extensively studied regulatory mechanism has been eIF2α phosphorylation. This highly conserved post-translational modification decreases the ternary complex supply to the 43S PIC formation, as it inhibits the eIF2-GTP recycling by eIF2B. In budding yeast, Gcn2 phosphorylates eIF2α upon general amino acid starvation; in metazoans, additional kinases modify eIF2α in response to endoplasmic reticulum (ER) stress (PERK), exogenous (viral) dsRNA presence (PKR), or low heme concentration (HRI). Gcn2 interacts with Gcn1 and Gcn20, a large scaffold protein and an ABC ATPase, respectively, and a current model proposes that Gcn1-Gcn20 acts as a sensor of a deacylated tRNA presence in the A site of the ribosome. The Gcn1-Gcn20 complex then passes the deacylated tRNA to Gcn2, and this interaction activates its kinase domain (*35*).

The main outcome of eIF2α phosphorylation, as outlined above, is the general decrease of translation initiation. However, the reduced 43S PIC availability acts as a key advantage for GCN4 mRNA translation, coding for a basic leucine-zipper (bZIP) transcription factor that activates general amino acid biosynthesis pathways. The mechanism is based on the presence of four upstream open reading frames (uORFs) within the 5' UTR of GCN4 (Figure 8). The translation is initiated at uORF1 and the sequence context of this three amino acid-long ORF favors the 40S subunit to stay on the mRNA after 60S dissociation. Under normal growth conditions (when eIF2α is not phosphorylated and the ternary complex supply is high), translation is reinitiated at the subsequent uORFs but terminated before the actual *GCN4* ORF, due to the nucleotide context at the 3' of uORF4. Under starvation condition (when eIF2α is phosphorylated), the remaining 40S continues scanning downstream of uORF1 and avoids reinitiation at uORF2-4, due to the limited ternary complex supply. The reinitiation finally takes place at the main ORF, Gnc4 protein is translated and activates the transcription of its target genes (*35*). Interestingly, a recent genome-wide study in *S. cerevisiae* using ribosome profiling revealed that *GCN4* has the highest mRNA to ribosome ratio (within the coding sequence) in the cell, reflecting its predominant translational control (*36*).

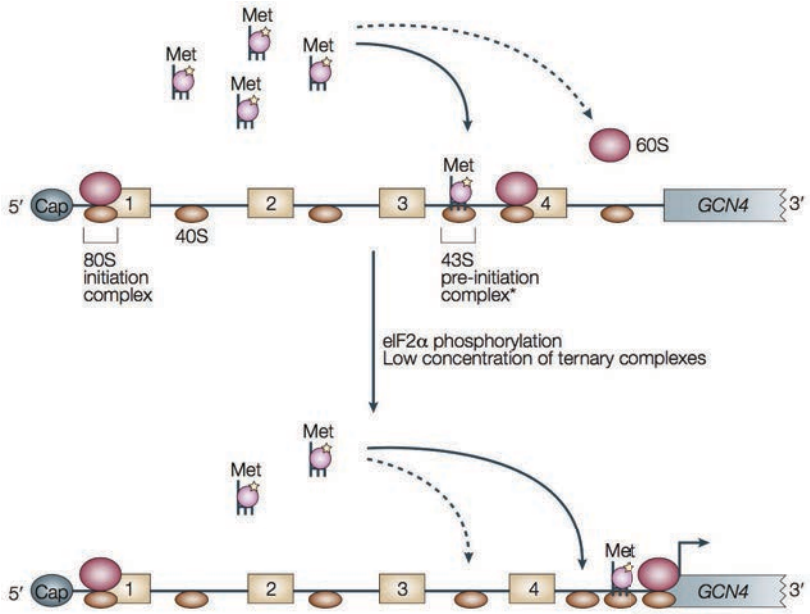

Figure 8. The mechanism of regulation of GCN4 mRNA translation.

GCN4 mRNA contains four upstream open reading frames (uORF1–4). Under conditions of amino-acid sufficiency (upper panel), reinitiation occurs more frequently after each uORF (continuous arrow), because of an increased probability of recharging the scanning 40S subunits that traverse the regions between the uORFs with active ternary complexes. As a result, reinitiation at the *GCN4* ORF becomes infrequent (dashed arrow). Under conditions of amino-acid scarcity, which induces eIF2α phosphorylation, and low levels of ternary complex (lower panel), reinitiation is unlikely to occur at the uORFs. This increases the probability of scanning 40S subunits reaching the region downstream of uORF4 and, subsequently, the *GCN4* AUG initiation codon. The asterisk indicates that the exact composition of the 43S pre-initiation complex, in the context of reinitiation, is still unknown. Adapted from (*37*).

eIF2α phosphorylation by Gcn2 engages in a crosstalk with other cellular signaling pathways. The TOR (target of rapamycin) complex 1 (TORC1), implicated in nutrient sensing and stress response, has been shown to mediate an inhibitory phosphorylation of Gcn2. This modification is thought to connect nitrogen sensing with general amino acid biosynthesis, as inhibiting TORC1 triggers Gcn2 dephosphorylation and leads to phosphorylated eIF2α-dependent control (*35*). In metazoan organisms, signaling-modulated translation is extensively prevalent, and TOR complexes, together with mitogen-activated protein kinases (MAPKs), play key roles in this process (*38*).

Along the eIF2α phosphorylation, translation initiation step can be modulated via eIF4E binding proteins (4E-BPs), a large group of factors containing a consensus

motif required for the interaction with eIF4E. 4E-BPs are of particular importance in metazoan organisms where they participate in a broad spectrum of cellular processes (*39*). In budding yeast, two 4E-BPs exist, Caf20 and Eap1, and were shown to associate to a large part of the transcriptome, probably via additional RNA binding proteins conferring target specificity. These 4E-BPs, similarly to their metazoan counterparts, compete with eIF4G for eIF4E binding and thus repress translation initiation by limiting 43S PIC recruitment to a target mRNA (*35*).

In contrast to a canonical cap-dependent translation initiation, various alternative cap-independent initiation mechanisms have been described. The classical example of this type of eIF4F-independent 43S PIC recruitment is a multitude of viral RNAs, synthesized by viral RNA polymerases and thus devoid of 5' cap. These RNAs usually contain a variant of a characteristic secondary structure within their 5' UTR, known as internal ribosome entry site (IRES) and capable of recruiting 43S PIC directly. However, many cellular mRNAs were found to contain IRES in metazoan organisms, and use cap-independent initiation as additional means of translational control. In *S. cerevisiae*, this mechanism is thought to be of a lesser prevalence, but the case of eIF4E-independent translation initiation from the internal part of URE2 mRNA, creating a N-terminally truncated protein, has been described (*35*). Intriguingly, two recent reports associated the presence of N^6-methyladenosine (m^6A) within the 5' UTR of mammalian mRNAs implicated in the heat-shock response, and its cap-independent translation upon general translational arrest during heat shock (*40, 41*).

Later stages of translation are equally targeted by regulatory mechanisms. In metazoan organisms, diverse factors have been associated with inhibiting or activating 60S recruitment (e.g. in *Drosophila melanogaster*, VASA interacts with eIF5B and activates translation of a subset of germline mRNA), as well as interfering with translation elongation (e.g. mammalian FMRT associates with a particular pool of synapse-located mRNA and stalls its translation). In budding yeast, the 80S-associating factor Stm1 has been described to stall elongating ribosomes and to activate decay of the translated mRNA (*39*). An interesting case is the translational regulation of *CPA1* expression, coding for an enzyme implicated in the arginine biosynthesis pathway. CPA1 mRNA contains a uORF, and at low arginine concentration both uORF and, due to a leaky scanning, *CPA1* ORF are translated.

The uORF codes for a short arginine attenuator peptide (AAP), and at high arginine concentration the ribosome is stalled while translating this peptide, reducing the translation of the downstream ORF (*35*).

The translation elongation factor eEF2 is targeted by multiple post-translational modifications, and some of it were associated with regulatory roles. eEF2 bears a highly conserved modification of a particular histidine residue (H699 in *S. cerevisiae*) by diphthamide, which is a target of diphtheria toxin. The abrogation of diphthamide modification pathway in yeast does not cause observable effects, however similar mutations in mice resulted in embryonic lethality, suggesting the specific role of modified eEF2 in the developmental program (*31*). eEF2 is also a target of multiple methylations, and in *S. cerevisiae* these modifications have been associated with an increased rate of ribosomal frameshifting and differential response to translation inhibitors (*35*). Moreover, eEF2 phosphorylation was reported to inhibit the elongation step in mouse brain cells and, paradoxically, favor the translation of a particular subset of synapse-located mRNAs (*31*).

A distinct subcellular localization of mRNA is an additional and powerful way of translational control. Pioneering studies in *D. melanogaster* revealed that the majority of mRNA in this organism is non-randomly localized (*42*). This concept was subsequently generalized to other model organisms, constituting a significant layer in controlling gene expression. In budding yeast, a classical example of localization-driven translational control is ASH1 mRNA, transported to the bud tip of a dividing haploid cell (Figure 9). *ASH1* codes for a DNA binding protein, repressing the synthesis of HO endonuclease and inactivating in turn the mating type switch. As the ASH1 mRNA is localized to the bud tip, the mating type switch is inhibited in the daughter cell only, and remains active in the mother cell. Moreover, ASH1 mRNA translation is repressed in the mother cell by a dedicated mechanism. A complex, containing She2, She3, Myo4, and termed locasome, recruits ASH1 mRNA co-transcriptionally and directs it to the actin cytoskeleton, further transporting it to the bud tip. Another factor, Puf6, binds ASH1 mRNA in the nucleus and accompanies it to the cytoplasm where, together with Hek2, an RNA binding protein, it represses ASH1 translation by associating with eIF5B and preventing 60S recruitment. Once at the bud tip, membrane localized casein kinase Yck1 phosphorylates Puf6 and Hek2, leading to eIF5B release and ASH1 translation (*43*).

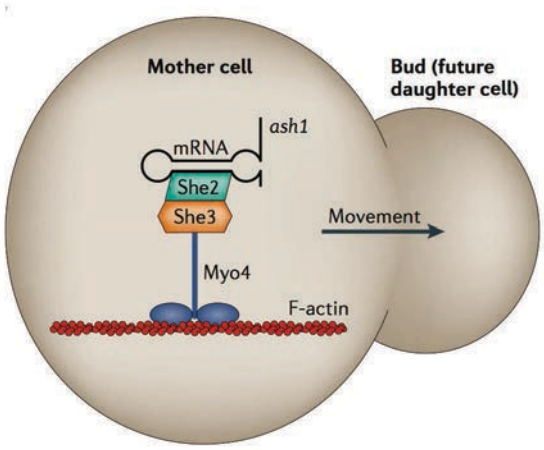

Figure 9. The mechanism of ASH1 mRNA localization.

Localization of ASH1 mRNA to the bud tip in dividing *Saccharomyces cerevisiae* cells involves myosin (Myo4)-driven transport along F-actin filaments. Two proteins, called She2 and She3, mediate the linkage; She2 binds to a stem-loop region within the 3' UTR of ASH1 mRNA. Adapted from (*39*).

Recently, an unexpected link between cellular signaling and mRNA localization has been reported in *S. cerevisiae* (*44*). As mentioned above, upon glucose depletion, the global translation status of a cell changes significantly: translation initiation is strongly inhibited, elongating ribosomes run off of the mRNAs, and within a time lapse of several minutes mostly vacant 80S accumulate. It has been shown that protein kinase A (PKA) signaling pathway is involved in this nutrient sensing response, however the exact mechanism leading to a prompt inhibition of translation initiation is unknown (*45*). Interestingly, glucose starvation triggers the transcriptional upregulation of genes involved in heat-shock response and glucose metabolism. In contrast, only mRNAs coding for heat-shock genes are actively translated upon glucose depletion; those coding for glucose metabolism genes follow the general trend and are devoid of ribosomes. Moreover, mRNAs that remain in the translationally active pool under starvation condition are diffusely distributed in the cytoplasm, while those that are ribosome-free localize in the cytoplasmic foci, known as processing bodies (P-bodies) and stress granules (detailed below). Surprisingly, the mRNA exclusion from the starvation-induced cytoplasmic foci is dependent on the presence of heat-shock factor 1 (Hsf1) binding sites within its promoter sequence, linking the transcriptional history of an mRNA with its cellular localization (*44*).

Various studies have reported the aggregation of mRNA and its accompanying translation and decay machinery into microscopically visible mRNP granules, evoked above. Two classes of cytoplasmic foci have been determined according to their composition: P-bodies, containing various cytoplasmic RNA decay factors, and stress granules, composed of mainly translation initiation-related proteins. However, the two classes form a compositional gradient, and P-bodies have been proposed to serve as a structural platform for stress granule assembly. P-bodies have been detected under standard growth conditions and are currently thought to be the hubs of cytoplasmic RNA decay (discussed further, in section 3 of this chapter). In contrast, stress granules (as the term suggests) appear as a part of stress- or starvation-induced response and have been suggested to participate in the stress-triggered translational repression, sequestering essential initiation factors (*46*). Recent findings indicate that mRNP granule assembly strongly relies on the biophysical properties of its components, as phase-separated liquid droplets have been readily observed in RNA and RNA-binding protein containing solutions (*47*). Moreover, its dynamics is dependent on the ATP-hydrolyzing chaperones for remodeling and Cdc48/autophagy for clearance (*48, 49*).

A separate field of translational control studies addresses the mechanisms of microRNA (miRNA) mediated translational repression. miRNAs are a class of short ncRNA, initially described in *Caenorhabditis elegans*, but now known to be universally prevalent in metazoans and plants. These RNAs are 20-25 nucleotides long and are produced during a multistage cascade of nuclear and cytoplasmic processing events. miRNAs associate with Argonaute (AGO) family proteins to form RNA-induced silencing complex (RISC), and subsequently recognize its target mRNA via complete or partial base pairing. The interaction between an mRNA and a miRNA-containing RISC induces a broad spectrum of consequences, among which mRNA cleavage, enhanced mRNA degradation by 3'-5' or 5'-3' cytoplasmic pathways, translational repression, etc. have been identified (Figure 10). A conserved RISC-interacting protein, known as GW182 in *D. melanogaster*, mediates the recruitment of the cytoplasmic decay machinery; however, the exact details of miRNA-mediated translational repression are still unclear but the involvement of eIF4A inhibition has been reported (*50*). Importantly, miRNA processing and interacting factors are not present in *S. cerevisiae* and, accordingly, no short miRNA-like RNA species were

detected in this model organism. Nevertheless, other closely related budding yeast species as well as distant fungi are capable of producing short ncRNAs, termed small interfering RNAs (siRNAs), that are mainly implicated in transposon and other repeated sequence-containing element silencing. It is noteworthy that siRNA-mediated gene silencing pathway, also known as RNA interference (RNAi), can be artificially reconstituted in *S. cerevisiae*, and this technique has been of particular importance for the results presented in this manuscript (*51*).

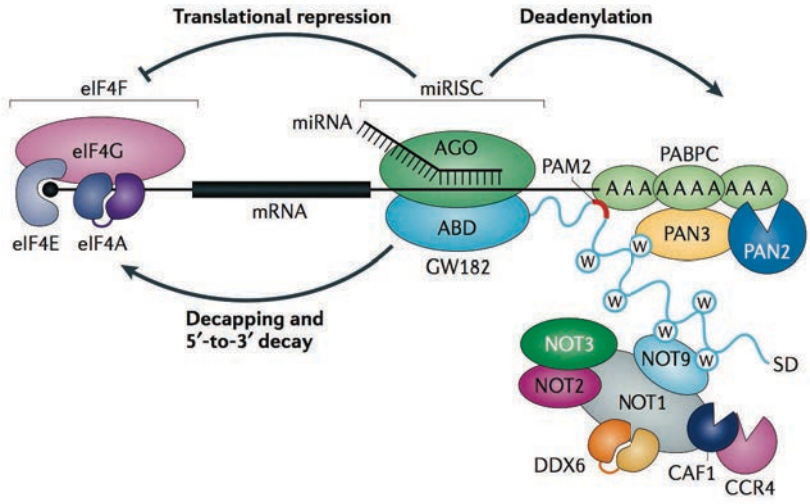

Figure 10. The overview of miRNA-mediated gene silencing in metazoans.

Metazoan miRNAs bound to an Argonaute (AGO) protein in miRNA-induced silencing complexes (miRISCs) recognize their mRNA targets by base-pairing to partially complementary binding sites, which are predominantly located in the 3' UTR of the mRNA. AGO proteins interact with a GW182 protein, which in turn interacts with cytoplasmic poly(A)-binding protein (PABPC) and with the cytoplasmic deadenylase complexes PAN2-PAN3 and CCR4-NOT. The GW182 proteins consist of an amino-terminal AGO-binding domain (ABD) and a silencing domain (SD). miRNAs repress translation, but the precise molecular mechanism for this remains unclear. The emerging consensus is that miRNAs inhibit translation initiation by interfering with the activity and/or assembly of the eIF4F complex. The eIF4F complex consists of the cap-binding protein eIF4E, the adaptor protein eIF4G and the DEAD box RNA helicase eIF4A. eIF4G serves as a scaffold for protein-protein interactions that are essential for the recruitment of the 43S pre-initiation complex and for translation initiation. The cap structure is shown as a black circle. DDX6, DEAD box protein 6; PAM2, PABP-interacting motif 2. Adapted from (*50*).

3 | mRNA turnover in the cytoplasm

Cytoplasmic mRNA decay pathways

Translation is intrinsically connected to messenger RNA decay, a natural end of its cellular life cycle. Degrading mRNA into nucleotides is essential to ensure the balance between its synthesis and turnover; accordingly, affecting this balance by inhibiting various stages of cytoplasmic mRNA decay results in dysfunctional cellular phenotypes. Furthermore, differential mRNA destabilization forms an additional layer of gene expression control. Cytoplasmic RNA decay machinery is also a key outcome of diverse co-translational quality control mechanisms, ensuring a rapid elimination of aberrant messengers (*52*). These aspects of mRNA metabolism in the cytoplasm are further detailed below.

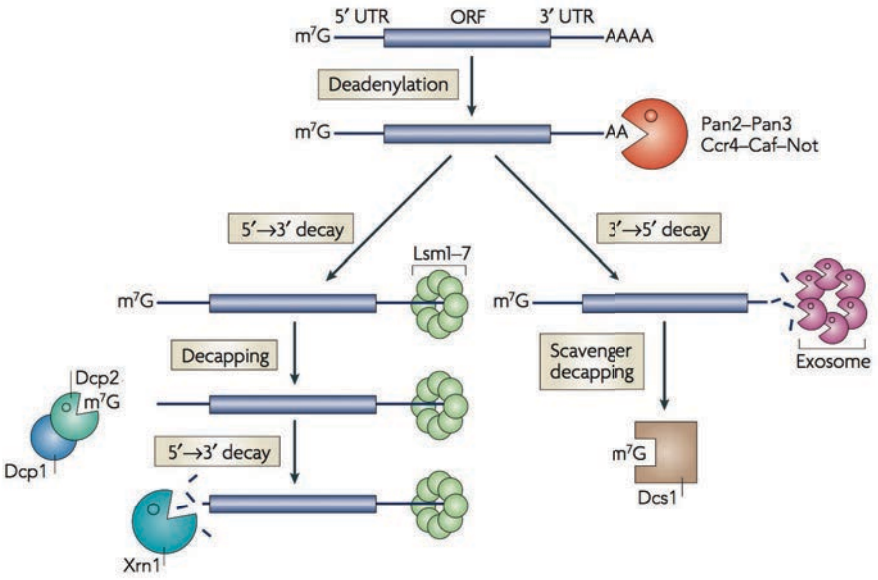

Figure 11. The mechanisms of cytoplasmic mRNA degradation.

Most mRNAs undergo decay by the deadenylation-dependent pathway. The poly(A) tail is removed by a deadenylase activity, shown here as either Ccr4-Caf-Not or Pan2-Pan3. Following deadenylation, two mechanisms can degrade the mRNA: either decapping followed by 5'-3' decay or 3'-5' decay. In the decapping pathway, the Lsm1-7 complex associates with the 3' end of the mRNA transcript and induces decapping by the Dcp1-Dcp2 complex. This leaves the mRNA susceptible to decay by the 5'-3' exoribonuclease Xrn1. Alternatively, the deadenylated mRNA can be degraded in the 3'-5' direction by the exosome, with the remaining cap structure being hydrolyzed by the scavenger-decapping enzyme Dcs1. Adapted from (*53*).

The canonical mRNA decay in the cytoplasm follows two general pathways—3' to 5' and 5' to 3' RNA degradation—that are both initiated by a gradual shortening of the 3' poly(A) tail (Figure 11). This process, known as deadenylation, is mediated by two cytoplasmic deadenylase complexes, Pan2-Pan3 and Ccr4-Not. Pan2 is the catalytic subunit of the Pan2-Pan3 assembly and has been reported to initiate deadenylation. The activity of Pan2 is stimulated by the presence of Pab1 bound to the poly(A) tail, and the dimer-forming Pan3 subunit was shown to directly interact with Pab1. Accordingly, in the cells where Pan2 is the sole active deadenylase, mRNAs with 20-25 A residues accumulate, reflecting the 3' extremity previously bound by the last Pab1 molecule. In wild-type cells, the initial Pan2-Pan3 mediated poly(A) tail shortening is thought to happen rapidly, since longer than 70-nucleotide poly(A) extensions can only be detected upon *PAN2* deletion (*54*).

Subsequently, the Ccr4-Not complex continues the deadenylation of an mRNA. It contains nine subunits, two of them having catalytic A-specific 3'-5' exoribonuclease activities: Ccr4 and Pop2 (also known as Caf1). However *in vivo*, the Ccr4 subunit shows the major deadenylase activity, Pop2 having predominantly structural role. In accordance, if *pop2*Δ cells accumulate deadenylation intermediates *in vivo*, this phenotype is not recapitulated by the presence of Pop2 catalytically inactive mutant. The Not1 subunit of the complex is a large scaffold protein, recruiting the rest of Ccr4-Not components—Not2-5, Caf40 and Caf130—whose functions remain largely unknown (*54*). An exception is the Not4 subunit, an E3 ubiquitin ligase that has been implicated in stalled ribosome-originated nascent peptide degradation and global translational repression (*55, 56*). Recently, Not2, Not3 and Not5 subunits have been shown to contribute to mRNA decapping via an interaction with the decapping activator Pat1 (*57*), as well as to accompany certain mRNAs from the nucleus to the cytoplasm, modulating its translational capacity (*58*).

In contrast to Pan2-Pan3 complex, Ccr4-Not activity is inhibited by the poly(A)-bound Pab1. This dependence suggests that Pab1 gradually dissociates from the poly(A) tail to make it accessible for Ccr4-mediated degradation; it also creates an intricate interplay between translation (as Pab1 is important for translation initiation, as mentioned above) and deadenylation. Consistently, mRNAs that undergo inefficient translation initiation (due to a poor AUG context, inhibitory secondary structures within 5' UTR, etc.) are targeted for accelerated deadenylation, suggesting

that these defects might somehow destabilize Pab1 binding. Ccr4-Not complex-mediated deadenylation has been also linked to translation termination. The release factor eRF3 interacts directly with Pab1, and the disruption of this interaction leads to the inhibition of the Ccr4-dependent poly(A) tail shortening. It was proposed that termination-coupled eRF3-Pab1 interaction transiently detaches Pab1 from the poly(A) sequence, providing an access for the Ccr4-Not complex (*52*). Interestingly, in certain metazoan organisms and plants, a third cytoplasmic deadenylase has been described. Known as poly(A)-specific ribonuclease (PARN), it is distinct in its activation by the 5' cap structure of an mRNA and has a role in diverse developmental processes (*53*).

After a critically short poly(A) length is attained (10-15 nucleotides, with no Pab1 associated), the mRNA is targeted for the subsequent degradation by two RNA decay pathways. In *S. cerevisiae*, the major mRNA turnover in the cytoplasm is mediated by the 5' to 3' RNA degradation. It begins with the removal of the 5' m7GpppN cap by the decapping enzyme complex Dcp1-Dcp2. In this complex, the Nudix pyrophosphatase family protein Dcp2 displays the catalytic activity, while Dcp1, belonging to the EVH family, is an essential cofactor of Dcp2 (*52*). In mammalian cells, Dcp1 and Dcp2 show weak interaction and require an additional scaffold protein Edc4 (also known as Hedls or Ge-1) to bridge the activation of Dcp2 (*53*). The cleavage of the cap results in 5'-monophosphorylated RNA and m7Gpp; the latter structure is further metabolized by its conversion to m7Gppp by nucleoside diphosphate kinase (Ynk1 in budding yeast) and subsequent cleavage by decapping scavenger enzyme Dcs1, yielding m7Gp (Figure 12) (*59*). Interestingly, along the main decapping activity of Dcp2, another cap cleaving enzyme has been discovered in metazoan organisms. Nudt16 also belongs to the Nudix family and has been shown to participate in various nuclear and cytoplasmic events; however, its target specificity and redundancy with Dcp2 remain elusive (*60*).

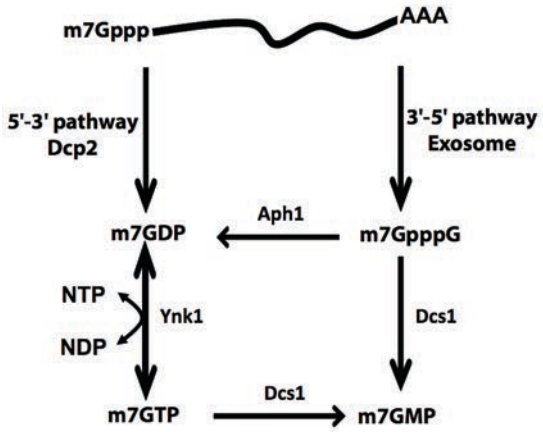

Figure 12. The schematic of cap (di-)nucleotide elimination.

Enzymes involved in the generation of cap (di-)nucleotide in the 5'-3' and 3'-5' mRNA decay pathways are indicated as well as the resulting products. At the 5' end, the cleavage of the cap structure by Dcp2 results in cap nucleotide, m7GDP, which is then phosphorylated by nucleoside diphosphate kinase, Ynk1, to yield the substrate for Dcs1. From the 3' extremity, exosomal degradation results in cap dinucleotide, m7GpppN, which can be directly metabolized by Dcs1 or cleaved by dinucleoside triphosphate hydrolase, Aph1, to yield m7GDP. The further metabolism of cap nucleotide monophosphate, m7GMP, remains unknown. Adapted from (*59*).

The activity of Dcp2 is modulated by a multitude of additional factors, known as decapping activators (or enhancers). These factors form a network of structural and functional interactions, connecting the 3' adenylation status of an mRNA to its 5' decapping. Moreover, some of these proteins act as translational repressors, providing a supplemental link between mRNA translation and decay. Importantly, all decapping activators were found to localize in cytoplasmic mRNP granules, the above-mentioned P-bodies and stress granules, and some of them are essential to its assembly (*52*).

An interesting case of 3' deadenylation repercussion to decapping activation is mediated by the Lsm1-7-Pat1 complex. Lsm1-7 complex, a heptameric ring of Sm-like proteins and similar to bacterial Hfq RNA chaperone, was found to predominantly bind deadenylated or oligoadenylated RNA. Its Lsm1 subunit confers the cytoplasmic localization of the complex, as the essential core Lsm2-7 subunits bind Lsm8 to form the nuclear Lsm2-8 complex, implicated in the metabolism of snRNA. Lsm2 and Lsm3 from the cytoplasmic Lsm1-7 complex recruit a large multidomain factor Pat1, and this association is thought to activate Pat1 binding to Dcp2, leading to decapping activation (*61*). Along Pat1, additional decapping

activators directly interact with Dcp2. Universally conserved factors Edc3 and Scd6, both containing Sm-like domain, compete for the same helical leucine-rich motif (HLM) of Dcp2, and this interaction has been proposed to prevent the Dcp1-Dcp2 complex to attain a catalytically inactive conformation. Yeast specific RNA-binding proteins Edc1 and Edc2 activate Dcp1-Dcp2 complex via Dcp1 binding. The multifunctional DEAD-box helicase Dhh1 also binds Dcp2 and stimulates its decapping activity (*60*). Interestingly, these factors form a highly interactive network, and recent structural studies have revealed that Pat1 and Edc3 use the same Phe-Asp-Phe binding site of Dhh1 C-terminal domain, suggesting the dynamic potential of mRNP remodeling by decapping activators (*62*).

Dhh1, Pat1 and Scd6 have been equally identified as inhibitors of translation. However, the key question remains whether these factors actually inhibit translation and kinetically favor eIF4F replacement by Dcp1-Dcp2 complex, or primarily act on decapping and subsequently compromise translation. Mechanistically, Scd6 has been shown to directly bind eIF4G, which in turn leads to the inhibition of 43S PIC recruitment. Dhh1 and Pat1 were also reported to inhibit translation initiation prior to 48S assembly (*52*), however later studies proposed that Dhh1 associates with slowly translating ribosomes and activates decapping (*60*). This is consistent with the finding that decapping occurs on polysomes, i.e. on translation-engaged mRNA (*63*). Moreover, as decapping activators together with the Dcp1-Dcp2 complex are the integral parts of cytoplasmic mRNP granules, the actual model proposes that sequestration of translationally inactive mRNA into P-bodies or stress-granules serves as a storage pathway, enabling its later re-entrance into translational pool (Figure 13) (*52*).

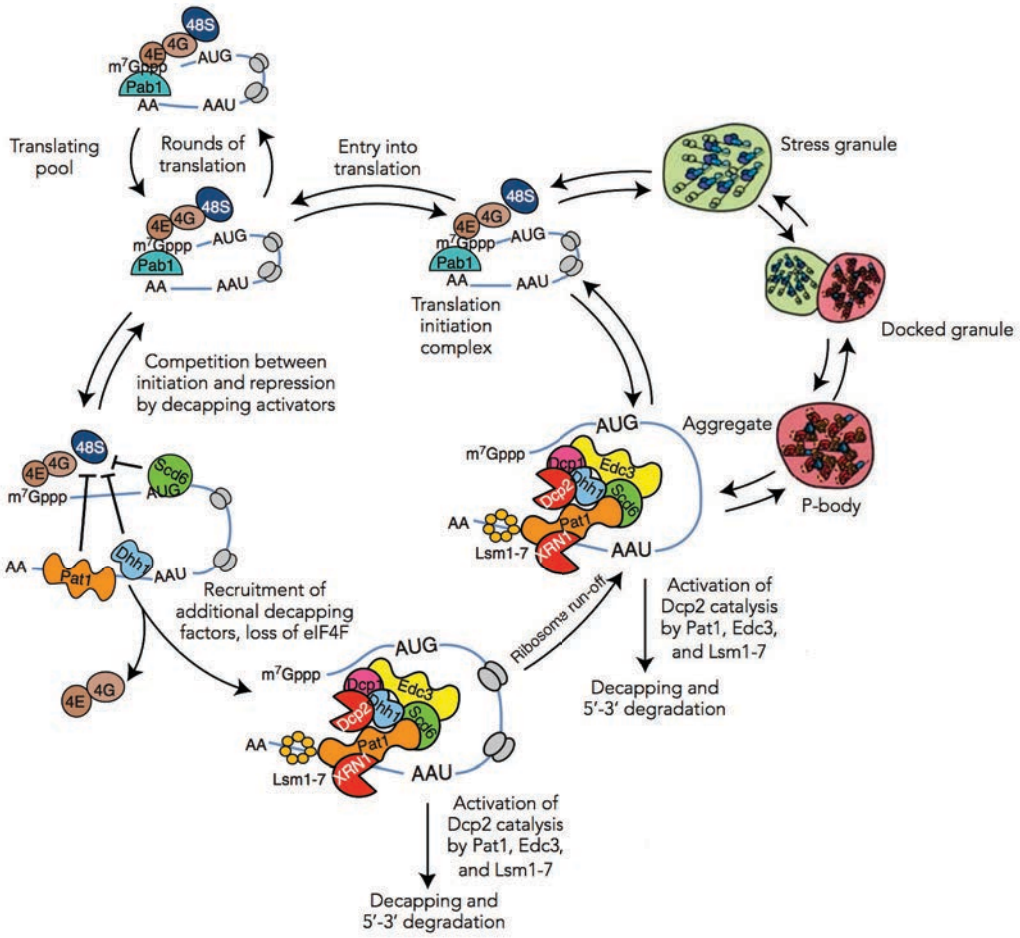

Figure 13. The model of the mRNA cycle.

The dynamic movement of mRNA between polysomes, P-bodies, and stress granules, and the possible mRNP transitions between the different states of the mRNA is depicted. Starting from top left (clockwise), mRNAs associate with translation factors and constitute the polysome-associated translating pool. Then, after initial association with translational repressors and decapping activators, mRNP loses translation factors and recruits additional decapping activators. From this point, mRNA decapping might occur, targeting the transcript for 5'-3' decay. Alternatively, decapping activators-bound mRNP might undergo ribosomal run-off and aggregate into P-bodies. Together with external stimuli-induced stress granules, P-bodies assure the storage of translationally inactive mRNPs, until translation can resume again. Adapted from (*46*).

As mentioned above, m7GpppN cap cleavage results in 5'-monophosphorylated mRNA that is the preferential substrate of the major cytoplasmic exoribonuclease Xrn1. This enzyme has a processive 5' to 3' exonucleolytic activity and is capable of degrading structured RNA due to its interdomain organization. Structural analysis of Xrn1 has revealed that it requires at least 4-nucleotide long 5' overhang for efficient degradation and cannot

accommodate 5' capped, triphosphorylated or hydroxylated RNA into its active site. Xrn1 is a highly conserved factor, implicated in RNA turnover throughout all eukaryotes, and along mRNA decay also participates in rRNA processing, tRNA turnover, and aberrant mRNA degradation (*19*). In budding yeast, Xrn1 targets a particular class of ncRNA, termed XUTs (Xrn1-sensitive unstable transcripts) (*64*), however recent data suggest that these ncRNAs undergo translation-associated quality control-mediated decapping and subsequent degradation by Xrn1 (*3*). It has been also shown that Xrn1 degrades mRNA co-translationally, as multiple 5' decay intermediates, sequentially protected by ribosomes, can be detected by high-throughput RNA sequencing techniques (*65*).

Xrn1 directly interacts with the decapping and decapping enhancing machinery, in accordance with a tight link between mRNA 5' cap removal and its irreversible targeting to degradation. Lsm1-7-Pat1 complex has been found to associate with Xrn1 via Pat1, relating mRNA 3' adenylation status to the cytoplasmic 5' to 3' decay pathway, as evoked previously. Xrn1 also binds to Dcp1-Dcp2 complex, addressing the unprotected 5'-monophosphate to degradation; this interaction is direct in *S. cerevisiae* and *D. melanogaster*, but requires the Edc4 scaffold in human cells (*19*). Xrn1 is activated by a cofactor protein Dcs1 *in vivo*, and this function is independent of Dcs1 catalytic decapping scavenger activity. Dcs1 also binds Xrn1 *in vitro* and stimulates its RNA-binding as well as catalytic activity (*66*). Xrn1 is an integral part of cytoplasmic mRNP granules, and has been found both in P-bodies and stress granules. Importantly, the disruption of the cytoplasmic 5' to 3' decay pathway affects the formation these mRNP granules, as *xrn1*Δ or *dcp2*Δ mutants accumulate large constitutive cytoplasmic foci, interpreted as the saturation of the cellular RNA turnover system with decay intermediates (*52*). Recently, an unexpected yet controversial nuclear function of Xrn1 has been proposed, attributing to this enzyme a key role in buffering mRNA levels in the cytoplasm and influencing transcription (*67, 68*).

In addition to the cytoplasmic 5' to 3' decay pathway, deadenylated mRNAs can enter the 3' to 5' decay pathway. It is mediated by the cytoplasmic RNA exosome, associated with its cofactor Ski complex. In budding yeast, 3'-5' decay pathway has been described to have a minor role in bulk mRNA turnover; accordingly, its activity is inhibited by 3' specific RNA-binding proteins, such as Pab1

or Lsm1-7. However, the inactivation of the 5'-3' decay pathway is not lethal, suggesting that these two RNA degradation mechanisms partially overlap. As mentioned above, both nuclear and cytoplasmic RNA exosomes share the same core composition, with the exception of nucleus specific catalytic subunit Rrp6; the cytoplasmic Exo10 contains thus the sole catalytic subunit Dis3. Similarly to the requirement for the TRAMP complex of the nuclear exosome, its cytoplasmic counterpart relies on an adaptor complex to access structured 3' ends of RNA (*52*).

The Ski2-Ski3-Ski8 heterotetramer, termed Ski complex, contains a large multi TPR motif scaffold protein Ski3 that recruits two Ski8 subunits and the catalytic Ski2 subunit. This organization is thought to activate Ski2, a DExH family ATP-dependent RNA helicase, that—in parallel to the nuclear Mtr4 from the TRAMP complex—unwinds RNA at the 3' end and guides it to the exosome core (*69*). The Ski complex associates with the cytoplasmic exosome via an adaptor protein Ski7, bridging the interaction between Ski3 and Ski4, a core exosome subunit. Ski7 is essential for exosome function *in vivo*, however it is present only in a subset of *Saccharomyces* species. Interestingly, the paralog of budding yeast Ski7 is Hbs1, a translational GTPase implicated in ribosome recycling, and it has been proposed that the latter factor, conserved throughout eukaryotes, might be involved in both processes (*70*). The Ski complex and cytoplasmic exosome-mediated mRNA decay results in 3' decay intermediate m7GpppN dinucleotide that is further processed into m7Gp by Dcs1 (*59*).

Along the canonical 5'-3' and 3'-5' mRNA decay pathways in the cytoplasm, the existence of non-canonical, endonucleolytic cleavage-induced decay pathway has been proposed in the literature. Although no cytoplasmic endoribonuclease has been identified in *S. cerevisiae*, its vacuolar T2-type ribonuclease Rny1 can be released to the cytoplasm upon oxidative stress, where it primarily targets tRNA and reduces the global translation level of the cell. It has been speculated that Rny1 might also cleave mRNA in the cytoplasm, although experimental evidence is still missing (*52*). In contrast, a recent study has shown that Rny1-mediated tRNA cleavage causes specific ribosome stalling upon oxidative stress, which in turn might lead to accelerated mRNA decay (*65*). It was also reported that upon nitrogen starvation-induced autophagy, bulk RNA is primarily targeted to vacuole where Rny1 efficiently degrades it (*71*). Similarly to Rny1, the mitochondrial nuclease Nuc1,

implicated in the programmed cell death, exits the mitochondria in response to oxidative stress and is further targeted to the nucleus. Its participation in cleavage-initiated decay has been equally proposed, supported by a negative genetic interaction between Nuc1 and cytoplasmic RNA decay factors. However, analogously to Rny1, no experimental evidence of Nuc1 mediated mRNA cleavage has yet been reported (*52*).

Regulation of cytoplasmic mRNA decay

Although cytoplasmic decay pathways act late in the mRNA life cycle, they are also subject to differential regulation. Mechanistically, this regulation depends on *cis*-elements within mRNA sequence, *trans*-interacting factors that recruit RNA decay machinery, as well as post-translational modifications of RNA degradation factors, modulating its activity (*52*). These control pathways target nearly all stages of cytoplasmic mRNA decay, and might be induced intrinsically (as a part of cellular gene expression program) or extrinsically (triggered by environmental cues) (*72*).

Enhanced deadenylation appears to be primarily induced by *cis*-elements, located within 3' UTR of mRNA. In *S. cerevisiae*, six Pumilio-homology family proteins Puf1-6 have been shown to control more than 1000 mRNAs (i.e. 10% of the transcriptome) via selective binding (cf. the above-mentioned control of the ASH1 mRNA by Puf6). Specifically, Puf1, Puf3, Puf4 and Puf5 are reported to induce deadenylation of bound mRNA, and, at least in the case of Puf5, to directly recruit Ccr4-Not complex via its Pop2 subunit. Similarly, the conserved SAM domain protein Vts1 has been shown to interact with a distinct stem-loop structure present in the 3' UTR of a subset of mRNAs and to enhance deadenylation via Ccr4-Not complex binding (*52*). In metazoan organisms, 3' UTR AU-rich elements (AREs) are targeted by a variety of RNA-binding proteins and constitute a potent regulatory pathway of cytoplasmic mRNA decay. AREs are involved in a highly complex network of interactions, as diverse ARE-binding proteins (ARE-BPs) might have a stabilizing or destabilizing effect and, furthermore, themselves are targets of post-translational modifications, integrating into various signaling pathways. Among the most studied examples of destabilizing ARE-BPs is tristetraprolin protein (TTP), directly recruiting deadenylase complexes in human cells via Not1, whose

deadenylase-interacting capacity is modulated by the p38 MAPK cascade (*73*). Importantly, the inhibition of deadenylation has been reported during various stress responses, independently of stress granule formation. The exact mechanism of this phenomenon remains unclear, although it has been associated with the stress-related translational repression (*52*).

Subsequently to enhanced deadenylation, *cis*-element binding factors participate in activating 5' decapping of mRNAs. In contrast, deadenylation-independent decapping has been also described. An interesting case is the degradation of EDC1 mRNA, coding for the above-mentioned decapping activator Edc1. Its 3' UTR contains a poly(U) stretch that is thought to sequester the downstream poly(A) tail, making it inaccessible for cytoplasmic deadenylases. Accordingly, EDC1 mRNA is mainly targeted to degradation by deadenylation-independent decapping and Xrn1. Mechanistically distinct degradation pathway targets RPS28B mRNA. It contains a stem-loop structure located in its 3' UTR and recruits the protein it codes for, Rps28b, to mediate the deadenylation-independent decapping. This process relies on decapping enhancer Edc3, as Rps28b possesses a *Saccharomyces* specific domain, responsible for Edc3 binding (*52*). A recent report has provided some additional data, attributing the stem-loop binding capacity to Edc3 and its activity-modulating role to Rps28b; however, it remains an intriguing example of feedback loop-type mechanism for regulating cytoplasmic mRNA decay (*74*). Interestingly, the rate of cytoplasmic decapping has been shown to be independent of P-body formation, as particular mutants defective in cytoplasmic foci assembly (*edc3*Δ *lsm4*Δ*c-ter* or *edc3*Δ *pat1*Δ) have comparable decay rates with cells capable to form P-bodies, at least in the case of model substrate mRNAs (*52*).

Recently, an important contribution to understanding the control of eukaryotic mRNA decay has been provided by a study focused on mRNA codon optimality. Genome-wide measurements of mRNA decay rates appeared to highly correlate with the codon composition of respective mRNAs. Furthermore, the importance of this interdependence has been demonstrated by replacing optimal codons in intrinsically stable mRNAs to synonymous suboptimal codons (without modifying the peptide sequence of the protein), which resulted in significant increase in mRNA decay rates. Similarly, intrinsically unstable mRNAs can be stabilized several-fold by optimizing its coding sequence (*75*). Importantly, these data suggest that translation is the key

process determining mRNA half-life, and connect translation elongation rate with mRNA decay, potentially via enhanced decapping (*76*).

Multiple factors involved in the decapping step of mRNA decay are targeted by post-translational modifications. The catalytic subunit of the decapping complex, Dcp2, undergoes phosphorylation by Ste20 kinase, and this modification is essential for Dcp2 localization within cytoplasmic stress granules. However, Dcp2 phosphorylation does not alter its catalytic activity, suggesting an alternative, signaling-related role for this modification. Decapping activator Pat1 is a substrate for the cyclic AMP (cAMP) dependent protein kinase (PKA) mediated phosphorylation. In contrast to phosphorylated Dcp2, Pat1 modification prevents the assembly of P-bodies. PKA activity is downregulated when yeast cells enter stationary phase, and maintaining low PKA activity is essential for cell survival after quiescence. Importantly, *pat1*Δ mutants show severely reduced survival rates, in accordance with nonphosphorylatable Pat1 mutant suppressor effect on constitutively active PKA-harboring cell survival. Entry to the cellular quiescence also relies on the specific protection of a subset of mRNAs from 5'-3' decay pathway. TORC1 inactivation upon nutrient limitation, associated with the stationary phase entry and initiation of the G_0 program, triggers protein kinase Rim15, which in turn targets two paralog proteins, Igo1 and Igo2. Phosphorylated Igo1/2 then interact with Dhh1, limiting its decapping activator and translational repressor function, licensing the expression of G_0-specific mRNAs (*52*).

The cytoplasmic 5'-3' exonuclease Xrn1 equally undergoes post-translational modifications. A recent study has reported that Snf1, the general glucose sensing-dependent protein kinase, targets Xrn1, together with Dhh1, Ccr4 and other RNA decay factors. Xrn1 phosphorylation is important for glucose-induced mRNA decay and, accordingly, constitutively active Snf1 in *reg3*Δ cells requires the presence of Xrn1, as *reg3*Δ and *xrn1*Δ mutations are synthetic lethal (*77*). An intricate link connects Xrn1 activity with general amino acid control. Through a cascade of enzymatic reactions, the sulfate assimilation pathway, required for methionine biosynthesis, yields a byproduct, adenosine 5',3' bisphosphate (pAp). Met16, responsible for producing pAp, is activated by Gcn4, the general amino acid biosynthesis factor (Figure 14). Importantly, pAp selectively inhibits the activity of Xrn1 and Rat1, dampening the cellular 5'-3' RNA decay pathway under amino acid

starvation conditions. pAp is further metabolized in the cytoplasm by Met22, yielding AMP and phosphate. The latter enzyme is a target of selective inhibition by lithium and sodium ions, connecting the 5'-3' exonucleolytic activity with the osmotic stress response and lithium toxicity (*78, 79*).

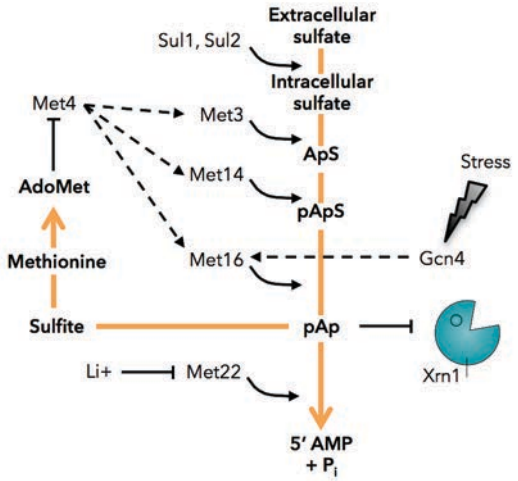

Figure 14. The sulfate assimilation pathway shares steps with the methionine biosynthesis pathway.

Sulfate enters the cell via the Sul1 and Sul2 transporters. Activities of Met3, Met14, and Met16 contribute to sulfate assimilation (marked by orange arrows) by producing sulfite. Sulfite is used in the production of methionine. Control of sulfate assimilation by methionine is ensured by subsequent production of AdoMet, which in turn represses the activity of Met4. Met4 participates in the transcriptional activation of *MET3*, *MET14*, and *MET16* genes (dashed arrows). Reduction of pApS by Met16 releases adenosine 3',5' bisphosphate (pAp), which is metabolized into AMP and inorganic phosphate by Met22. pAp inhibits the activity of the 5'-3' exonuclease Xrn1. Alternatively, pAp can accumulate because of the inhibition of Met22 in the presence of lithium ions. Independently of Met4, Gcn4 production activates *MET16* expression (dashed arrows). Adapted from (*78*).

The activity of Xrn1 in the cytoplasm has been associated with the metabolism of diverse carbon sources in budding yeast. It is important for cellular respiration, as *xrn1*Δ as well as *xrn1* catalytic mutant fail to utilize non-fermentable carbon sources (*66*). As evoked previously, Xrn1 is activated by Dcs1, decapping scavenger enzyme, which has been also associated with nutrient response signaling. Dcs1 forms a heterodimer with its paralog Dcs2, induced upon glucose-limited conditions via cAMP-PKA pathway. Dcs2 locates within cytoplasmic P-bodies, and its association with Dcs1 upon yeast entry into diauxie reportedly suppresses the catalytic activity of Dcs1 (*80, 81*). Finally, a recent study has shown that Xrn1 co-localizes with

membrane-associated cytoplasmic foci, termed eisosomes, in post-diauxic yeast cells, suggesting a link between the spatial isolation of this key ribonuclease and the entry into quiescence program (*82*).

Quality control pathways of aberrant mRNAs

Eukaryotic mRNA undergoes a thorough surveillance throughout its life cycle, and mRNA transcription- as well as maturation-related nuclear quality control pathways have been addressed in section 1 of this part of the chapter. In the cytoplasm, ribosome plays a key role in assessing the quality of the information encoded within an mRNA. Therefore, distinct cellular machineries have evolved to sense inefficiently elongating or terminating ribosomes and, subsequently, to target the associated aberrant mRNA to degradation (*52*).

Various signals might lead to the stalls in translation elongation, among which mRNA truncation (e.g. due to a missplicing event), absence of a stop codon (e.g. due to a nonstop mutation), presence of a strong secondary structure within the coding sequence, chemical RNA damage (e.g. nucleotide base oxidation) resulting in defective decoding, specific aminoacyl-tRNA insufficiency, homopolymeric (e.g. poly(A) stretches coding for polylysines) and GC-rich sequences within the ORF, etc. have been reported. Initially, two distinct pathways have been identified respective to the origin of the ribosome stall, namely nonstop decay (NSD; targeting truncated or stop codonless mRNA, with ribosome reaching the very 3' end of the transcript) and no-go decay (NGD; primarily triggered by internal stalls during elongation, before reaching the stop codon). However, later studies have provided an increasing amount of evidence for the functional interrelation of these pathways. Among many examples, the polylysine stretch, produced during the poly(A) tail translation in the absence of a stop codon, causes the ribosome to stall due to the electrostatic interaction between positive lysine charges and a negatively charged rRNA backbone within the ribosomal peptide exit tunnel (*83*).

Two important mechanistic implications underlie these mRNA quality control pathways, involved in rescuing stalled ribosomes. Firstly, as ribosomes get stalled on sense codons (e.g. at a strong secondary structure or on a stretch of rare codons), a non-canonical termination and recycling mechanism is needed to dissociate these

ribosomes from aberrant mRNAs, in order to avoid non-functional mRNP accumulation. Secondly, to prevent the potential damage of an abortive translation product, the nascent peptide has to be degraded together with the aberrant mRNA. Various studies have contributed to identify a list of factors, acting in a timely manner to ensure the coordinated stalled ribosome rescue and accompanying decay of the aberrant transcript together with the nascent peptide chain (Figure 15) (*83*).

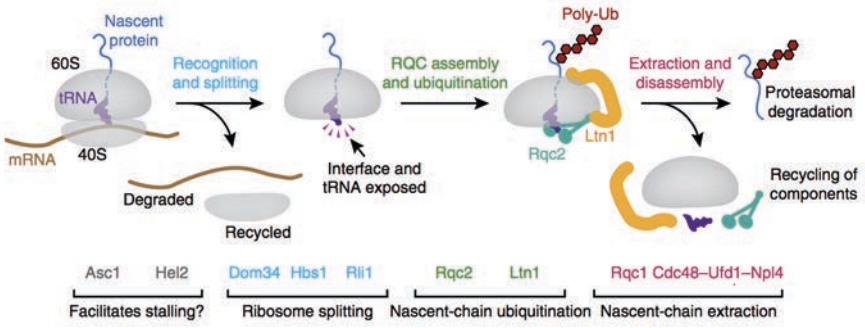

Figure 15. Primary steps and factors of ribosome-associated quality control.

A stalled ribosome is recognized (potentially assisted by Hel2 and Asc1) and acted upon by ribosome-recycling factors Dom34-Hbs1-Rli1 that split the ribosomal subunits. Removal of the 40S subunit exposes the peptidyl-tRNA and intersubunit interface on the 60S subunit. These cues are recognized by RQC components, Ltn1 and Rqc2, whose assembly on the 60S subunit permits ubiquitination of nascent polypeptides by Ltn1. The polyubiquitinated complex is then disassembled by the Rqc1-recruited Cdc48-Ufd1-Npl4, thus allowing recycling of the factors and degradation of the nascent chain. Adapted from (*83*).

Although the exact mechanism of stalled ribosome sensing remains unclear, two factors have been implicated at the early stages of the NGD pathway. Both Hel2, an E3 ubiquitin ligase, and Asc1 (known as RACK1 in metazoans), core component of the 40S ribosomal subunit, facilitate ribosome stalling on rare codons. Accordingly, *hel2*Δ and *asc1*Δ mutants show increased ribosome passage through stall-inducing elements, however without compromising the further stages of the quality control pathway (*83*). Interestingly, the activity of Hel2 has been linked to proteasome-unrelated K63 polyubiquitination, probably implicated in stall signaling events. This activity is NGD specific, as inhibiting it does not cause any observable effect on NSD substrates (*84*). The principal difference between NSD and NGD lies in the differential A site occupancy of the stalled ribosome (i.e. empty in NSD versus

containing sense codon in NGD), and it might cause alternative routes within these two generally converging pathways (*83*).

Following the detection of a stalled ribosome, two specific dissociation factors Dom34 (also known as Pelota in metazoans) and Hbs1 mediate its splitting into subunits. These factors are evolutionarily conserved homologues of canonical translation termination factors eRF1 and eRF3, and share significant respective structural similarity. Dom34, like its counterpart eRF1, adopts tRNA-like structural conformation that allows its accommodation into ribosomal A site. However, Dom34 lacks the essential GGQ motif of eRF1, important for stop codon recognition, in accordance with its binding to termination incompetent ribosomes. Hbs1, similarly to eRF3, is a translational GTPase, and its GTP hydrolyzing activity is necessary for Hbs1 function *in vivo* and *in vitro*. Following the GTP hydrolysis, Hbs1 dissociate from Dom34 and is subsequently replaced by ATP-bound Rli1, to proceed with the later steps of ribosome recycling (Figure 16). *In vitro*, Dom34-Hbs1 complex efficiently splits ribosomes with short RNA sequences downstream the ribosomal A site, suggesting that it primarily targets nonstop substrates *in vivo*. This effect depends on Hbs1 steric positioning on the ribosome, and might account for apparently low dissociation efficiency of certain NGD substrates (*83*).

The hallmark of NGD and NSD pathways is the endonucleolytic cleavage of the stalled ribosome-associated mRNA. Multiple studies have reported this cleavage to take place upstream of the stall site, however the enzyme responsible for this activity is still unknown. Dom34-Hbs1 complex stimulates the cleavage yet is not required for it to occur. This suggests that ribosome dissociation is dispensable for the cleavage event, although the exact order of ribosome splitting in relation to mRNA cleavage remains unclear. The cleavage results in two fragments, so called 5' and 3' decay intermediates, which are rapidly degraded by the cytoplasmic exosome and Xrn1, respectively, and thus remain mostly undetectable in wild-type cells. Importantly, every cleavage event generates a poly(A) tailless 5' decay intermediate that becomes a substrate for reiterated NSD cleavage in response to stalls of upstream ribosomes. In contrast, some 3' decay intermediates are detectable in the presence of functional cytoplasmic RNA degradation machinery, suggesting that a ribosome might remain associated with an aberrant mRNA fragment, escaping the splitting by Dom34-Hbs1 (*85, 86*).

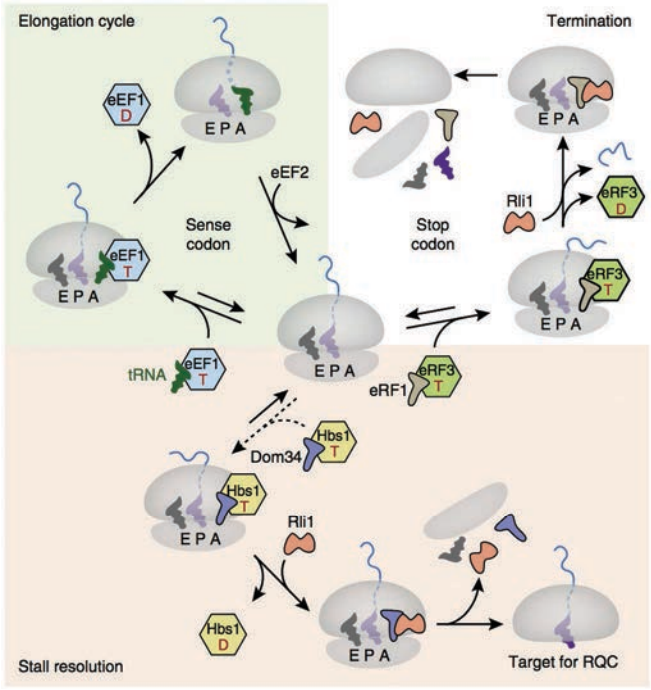

Figure 16. The working model for recognition of a stalled ribosome by recycling factors.

Top left (green background), a simplified translation elongation cycle is shown. A translating ribosome in the nonrotated state engages the tRNA-eEF1A-GTP ternary complex in response to a sense codon in the A site. Codon recognition by the tRNA triggers GTP hydrolysis by eEF1A, release of the latter from the ribosome and accommodation of the tRNA to catalyze peptide-bond formation. The ribosome is then translocated by one codon via the action of eEF2 to complete the cycle. Top right (white background), when a stop codon enters the A site, it is recognized by an eRF1-eRF3-GTP complex that functions analogously to the elongation complex. Upon accommodation of eRF1, the ATPase Rli1 is recruited, and peptidyl-tRNA is hydrolyzed, thus releasing the nascent protein. The ribosomal subunits are separated by the action of the eRF1-Rli1 complex. Bottom (pink background), failure to be engaged by either the eEF1 or eRF1 complex permits "default" engagement by the Dom34-Hbs1-GTP complex, which does not exhibit codon specificity. These factors act similarly to the homologous eRF1–eRF3 complex, with the exception that Dom34 does not catalyze peptidyl-tRNA hydrolysis. Thus, subunit separation results in a 60S-peptidyl-tRNA complex that is targeted by the RQC. T, GTP; D, GDP; E, exit tunnel. Adapted from (*83*).

As evoked previously, the canonical GGQ motif, present in eRF1, is absent in Dom34 and, accordingly, Dom34 is incapable to promote the nascent peptidyl-tRNA hydrolysis. After the ribosomal subunit dissociation, peptidyl-tRNA remains bound to the 60S subunit and the RQC (ribosome quality-control) complex is recruited. This complex contains multiple conserved factors, among which Ltn1,

Rqc1, Rqc2, and Cdc48-Npl4-Udf1 subcomplex have been identified. Rqc2 (NEMF in metazoans) interacts with the tRNA, located in the P site of 60S subunit, and prevents the reassociation of 40S subunit. This step is crucial for the subsequent binding of Ltn1 (also known as Listerin), as it shares the interaction surface of 60S with 40S. Ltn1 is a RING domain E3 ubiquitin ligase, polyubiquitinylating the nascent peptide associated with the 60S and thus targeting it for proteasomal degradation. During the later steps of the pathway, Rqc1 is thought to facilitate the binding of Cdc48 to the ubiquitinylated peptide and its subsequent extraction from the 60S. Strikingly, Rqc2 has been shown to participate in the ejection of the nascent peptide via untemplated peptide chain elongation. This phenomenon, termed CAT (C-terminal alanine-threonine) tailing, selectively adds alanine and threonine residues to the C-terminus of the stalled peptide, and is essentially dependent on Rqc2-mediated accommodation of aminoacyl-tRNA into the A site of 60S. Furthermore, CAT tails have been shown to induce the aggregation of aberrant peptides if not efficiently targeted for proteasomal degradation and also, in parallel, to trigger the general stress response via Hsf1 activation (*83*).

Along the stalls in translation elongation, a pathway targeting inefficient translation termination has also been described. It is termed nonsense-mediated decay (NMD), as it has been initially discovered while studying mRNAs containing a premature stop codon (PTC). Later data have broadened this observation, as it appears that a significant amount of messengers without any apparent PTC are equally targeted by NMD. Importantly, a current view of NMD exceeds the unique role in degradation of aberrant, PTC-containing mRNA, and integrates an intricate mechanism dedicated to controlling gene expression (*87*).

The recognition of inefficient translation termination complexes remains elusive, despite the multitude of experimental evidences associated with this event. In the case of PTC, its distance from the poly(A) tail-bound Pab1 is proposed among the key features triggering NMD. Consistently with this observation, longer 3' UTR sequences have been shown to destabilize the mRNA. In metazoans, a dedicated complex is deposited in the vicinity of splicing sites of an mRNA, termed exon junction complex (EJC), and accompanies the mRNP from the nucleus to the cytoplasm. During the pioneering translation round, this complex is displaced by the elongating ribosome. Importantly, most intron containing mRNAs have its stop

codon localized within the last exon, positioning all EJCs upstream the translation termination signal. Accordingly, EJC plays an important role in distinguishing PTCs, as its subunits have been shown to interact with the core NMD machinery. However, intronless mRNAs also undergo NMD-mediated destabilization in metazoans, and EJC-devoid organisms (as is the case of *S. cerevisiae*) have fully functional NMD pathway (*88*).

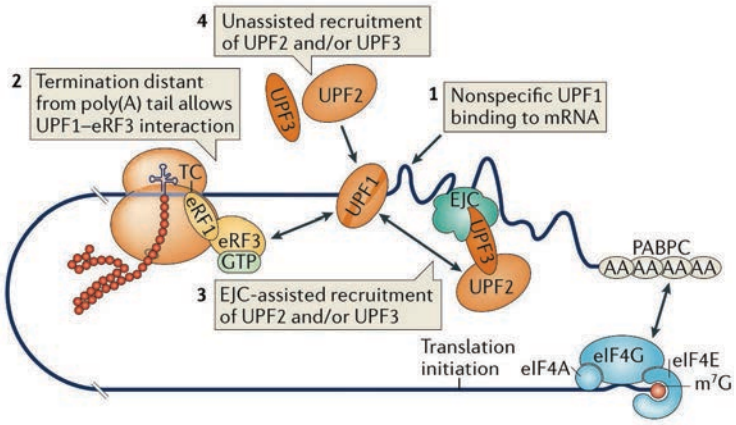

Figure 17. The model of the recruitment of Upf1, Upf2 and Upf3 during inefficient translation termination.

Upf1 can bind mRNA nonspecifically (1) but NMD is enabled when ribosome-associated eRF3 interacts with Upf1, thereby recruiting Upf2 and/or Upf3 (2). Recruitment of Upf2 and/or Upf3 is assisted by an exon–junction complex (EJC) bound to the 3' UTR in metazoans (3) but it can also occur independently (4). Some mRNAs may escape NMD for one or more rounds of translation, owing to the inefficient recruitment of Upf1, Upf2 and/or Upf3 to the terminating ribosome. Adapted from (*87*).

The universally conserved core complex of NMD factors contains Upf1, Upf2 and Upf3 proteins (Figure 17). Upf1, the ATP-dependent RNA helicase, has been termed the master regulator of NMD, as it mediates the PTC-bound ribosome recognition and recruits Upf2 and Upf3, to trigger the decay of associated mRNA. Biochemical data suggest that Upf1 directly binds 40S via Rps26, as well as Upf2 and Dcp2. A current model proposes that Upf1 accompanies the translation of all mRNAs, via a transient interaction with the 40S subunit. In the case of inefficient translation termination, interaction with Upf2-Upf3 stabilizes Upf1 binding to the ribosome and promotes the translation termination, as Upf1 has been shown to directly interact with the eRF1-eRF3 complex. In parallel, Upf2-Upf3-activated

Upf1 recruits the Dcp1-Dcp2 decapping complex, targeting the mRNA to deadenylation-independent decapping and Xrn1-mediated degradation (*88*).

In metazoan cells, additional factors are involved in mediating NMD. The previously mentioned EJC directly interacts with Upf3, and it has been proposed that an EJC downstream of a PTC provides the Upf2-Upf3 complex for Upf1 activation. Along the core NMD factors, Smg1 and Smg5-9 proteins also act in the pathway. Upf1 phosphorylation by Smg1 is associated with launching the downstream events of NMD. Among other effects, phosphorylated Upf1 has an increased binding affinity to Smg6, the NMD-specific endonuclease, primarily initiating the decay of NMD targets in metazoans. In human cells, Upf1-interacting Smg5-Smg7 dimer also directs target mRNAs to enhanced deadenylation, via Smg7-Pop2 interaction, recruiting the Ccr4-Not deadenylase complex. Consistently with the variety of actors involved in the NMD pathway, a significant part of human transcriptome has been reported to rely on NMD-mediated regulation (*87*).

Surprisingly, very little data exist on the fate of the truncated peptide produced during the NMD target translation. It has been shown that budding yeast Upf1 possesses ubiquitin ligase activity and contributes to the proteasomal degradation of the product translated from a PTC-containing transcript. However, the destabilization of NMD-related peptides appears to be context specific and so far lacks a generalized mechanism (*52*). It has been thus proposed that controlling gene expression via NMD-mediated mRNA destabilization allows the differential penetrance of target protein production (*87*).

PART II | Intermolecular RNA-RNA interaction: a core mechanism of gene regulation

A single stranded RNA molecule has an intrinsic ability to engage in intra- and intermolecular interactions via base pairing, and this feature stands among fundamental molecular mechanisms of all living systems. The most prominent examples include the structural and functional organization of essential RNPs, such as ribosome, spliceosome, and telomerase, among many others (*89*). Intermolecular RNA-RNA interactions primarily act as structural elements for scaffold assemblies or catalytic activities, and recent genome-wide data support the widespread presence of secondary and tertiary RNA structures in diverse organisms (*90-92*), although cells actively control the extent of these interactions (*93*). On the other hand, intermolecular RNA-RNA interactions cover a great variety of mechanistic implications, from the decoding principle via mRNA-tRNA pairing inside the ribosome to the miRNA-mediated translational repression (*89*). Moreover, latest data continue to enlarge the already extended list of roles for ncRNA-dependent gene regulation (*94, 95*). The Part II of this chapter aims (1) to review the origins of the interaction-competent RNAs, and then (2) to focus on the long-range intermolecular RNA-RNA interactions, particularly following the evidences and molecular outcomes of mRNA-mRNA pairing.

1 | The origins of interaction-competent RNA species

Various studies have suggested diverse classifications of RNA, mostly ncRNA beyond the canonical mRNA-rRNA-tRNA triplet, based on the length, function, stability, subcellular localization, relation to adjacent DNA elements, and other properties of the transcript (*95*). In the scope of RNA-RNA interactions, we propose the criterion of the origin of interaction-competent RNA to distinguish between *in cis* and *in trans* produced interacting RNA (Table 1), in contrast to generally suggested *cis*- and *trans*-acting ncRNA opposition (*96*).

Table 1. The origins of interaction competent RNA.

RNA species	*Function*
***In trans* produced interacting RNA**	
rRNA	mRNA translation
tRNA	mRNA decoding and cognate amino acid supply
snRNA	pre-mRNA maturation
snoRNA	rRNA, tRNA, snRNA modification Alternative splicing mRNA destabilization
miRNA	mRNA destabilization and translational repression miRNA maturation
piRNA	Transposable element silencing
lncRNA/mRNA	mRNA destabilization mRNA stabilization Alternative splicing
***In cis* produced interacting RNA**	
siRNA	Heterochromatin formation Transposable element silencing Antiviral response Transcription termination
lncRNA/mRNA	Chromatin remodeling Alternative splicing mRNA stabilization Translational activation

In trans produced interacting RNA

We define *in trans* produced interacting RNA as a pair of partially or entirely complementary RNAs that are transcribed from the spatially separated genomic loci and thus their complementarity is not inherently dependent of their genomic origin. The best-described examples of this class of interacting RNA are short RNA species such as snoRNA, miRNA, piRNA (Piwi-interacting RNA) and to some extent siRNA (see bellow). Importantly, tRNA and snRNA also fall into the category of *in trans* produced interacting RNA; their fundamental role in mRNA decoding and pre-mRNA processing, respectively, has been addressed in the Part I of this chapter (see sections 1 and 2).

snoRNAs are a class of ncRNAs having a well defined secondary structure, further subdivided in C/D box and H/ACA box families (Figure 18). These short RNAs have diverse genomic origins, and are mostly encoded within introns of ribosome metabolism-associated mRNAs (e.g. in metazoans), although some originate from other elements within mRNAs or are transcribed from dedicated Pol II or Pol III promoters (e.g. in fungi or plants). The canonical function of C/D box and H/ACA box snoRNA is to guide the post-transcriptional methylation and pseudouridylation, respectively, of rRNA and snRNA. The base pairing between snoRNA and its target indicates the exact position of the modification that is further mediated by the snoRNA-associated protein assembly. However, several non-canonical functions of snoRNA have been described. U3 snoRNA contains the characteristic C/D box element but does not guide RNA methylation; instead, it is responsible for guiding pre-rRNA cleavage important for 18S rRNA maturation (*97*). A subclass of snoRNAs has been termed orphan snoRNAs as they show no sequence complementarity with the canonical ncRNA targets. Interestingly, mRNA targets can be bioinformatically predicted for certain orphan snoRNAs, and several cases have been experimentally validated. SNORD115 (also known as snoRNA HBII-52) targets serotonin 5C receptor mRNA in human brain cells and induces its alternative splicing; this snoRNA-pre-mRNA interaction is of clinical importance as the absence of SNORD115 contributes to the traits of Prader-Willi syndrome (*98*). The mouse homologue of SNORD115 controls the alternative splicing of five mRNA targets; surprisingly, it acts in its processed form, after loosing the C/D box element (*99*). Recently, human SNORD27, which is active in guiding rRNA methylation, was also shown to be responsible for the alternative splicing of transcription factor E2F7 pre-mRNA, probably via direct competition with U1 snRNP. SNORD27 knockdown leads to exon inclusion within additional mRNAs, suggesting its broader scope for promoting alternative splicing (*100*). Furthermore, latest dsRNA crosslink-based genome-wide data revealed that another orphan C/D box snoRNA in human cells, SNORD83B, acts in downregulating the steady-state levels of at least three mRNA, albeit via an unknown mechanism (*92*). The growing list of non-canonical snoRNA functions supports the highly versatile potential of nuclear RNA-RNA pairing.

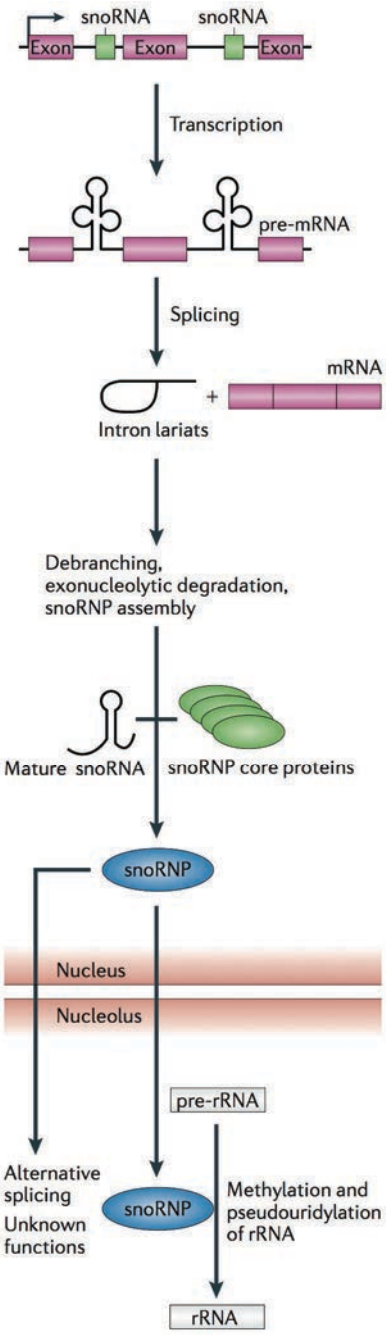

Figure 18. The biogenesis and effector machinery of snoRNA.

Small nucleolar RNAs are predominantly located in introns. Following splicing, debranching and trimming, mature snoRNAs are either exported, in which case they function in ribosomal RNA (rRNA) processing, or remain in the nucleus, where they are involved in alternative splicing and additional yet unknown functions. snoRNP, small nucleolar ribonucleoprotein. Adapted from (*101*).

The above-mentioned miRNAs (see section 2 of Part I) constitute a highly potent reservoir of so-called fine-tuning regulators of metazoan and plant gene expression. Similarly to snoRNAs, these short RNAs are also produced from mRNA

or ncRNA introns, or from dedicated Pol II or Pol III promoters; distinctively, multiple miRNAs can originate from a single polycistronic transcript (Figure 19). miRNAs undergo a cascade of nuclear and cytoplasmic processing events, and act in the cytoplasm associated with RISC complex. In animal cells, the status of complementarity between miRNA and mRNA defines the primary effect of miRISC on interacting mRNA. If the complete (or near complete) complementarity within the miRNA-mRNA interaction leads to mRNA cleavage by a member of AGO family proteins, the partial complementarity induces translational repression and mRNA destabilization (*102*). It has been long debated on the mechanistic details of partially complementary miRISC-mediated repression, and the current model proposes that translational repression occurs rapidly and primarily targets the early stages of translation (*103, 104*). However, subsequent mRNA destabilization has a dominant effect on steady-state gene expression levels (*105*). Interestingly, a recent report proposes an integrated view of miRNA-induced translational repression and mRNA decay, as miRNA-dependent co-translational mRNA decapping and 5'-3' degradation has been detected (*106*). In plant cells, where miRISC-induced mRNA cleavage predominates over translational repression, miRNA biogenesis pathway has a key role in determining the miRNA-mediated effect. Independently of the complementarity status of the miRNA-mRNA duplex, *Arabidopsis thaliana* Dicer-like 1 (DCL1) interacting proteins DRB1 and DRB2 direct the respective miRISCs to mRNA cleavage or translational repression route, representing an alternative mechanism within the miRNA-mRNA interaction pathway (*107*).

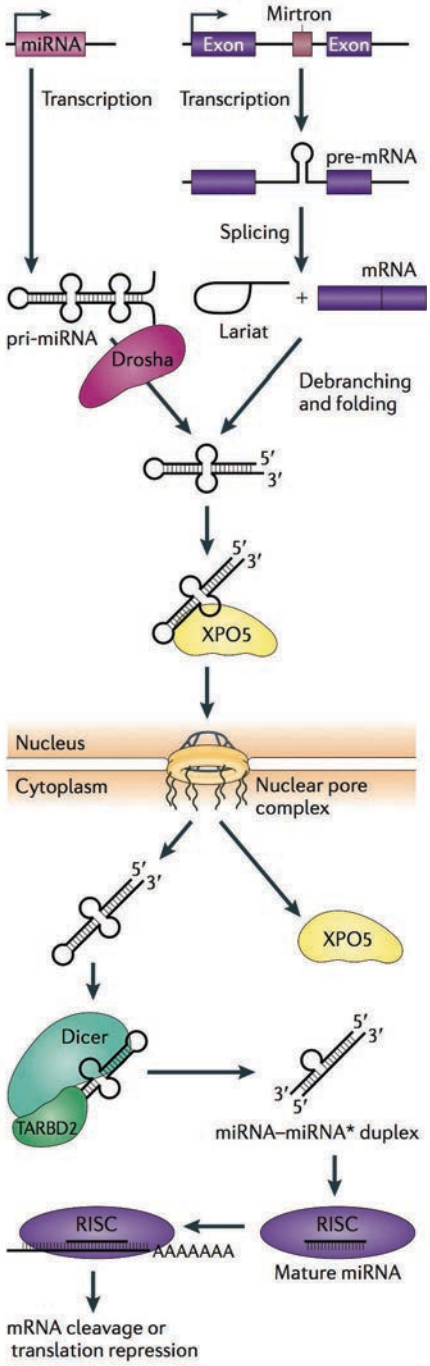

Figure 19. The biogenesis and effector machinery of miRNA.

miRNAs are transcribed as individual units (primary miRNA, pri-miRNA) or together with host genes (mirtrons). Following processing by the Drosha complex or the lariat-debranching enzyme, respectively, precursor miRNAs (pre-miRNAs) are exported from the nucleus by exportin 5 (XPO5). Further processing by Dicer and TAR RNA-binding protein 2 (TARBP2) generates mature miRNAs, which are loaded into the RNA-induced silencing complex (RISC). Adapted from (*101*).

Recent data continue to accumulate evidence for multiple non-canonical miRNA interactions. In accordance, various miRNAs have been detected in the

nucleus, suggesting their extended effect beyond the cytoplasmic translational repression. Of notice, mouse miRNA miR-709 interacts with the precursor form of miR-15a/16-1 in the nucleus, blocking the processing of miR-15a/16-1 towards its mature form and thus contributing to the upregulation of its targets. Similarly, *let-7* miRNA of *C. elegans* binds its own precursor in the nucleus, promoting the processing of a primary transcript and creating a positive feedback loop for *let-7* production. Additionally, various miRNAs, in their mature of precursor form, interact with a repertoire of ncRNAs, resulting in the effects on primary miRNA processing or mature miRNA sequestration. The latter interaction has been termed "the sponge phenomenon", describing the ncRNA competition with the target mRNA(s) for miRNA-binding. Importantly, miRNA-interacting ncRNAs, also known as molecular decoys, can contain multiple binding sites for the same or different miRNA, serving as potent miRNA reservoirs. This is the case of H19 ncRNA that sponges *let-7* miRNA in human cells; moreover, several mRNAs also compete for *let-7* binding, notably high mobility group AT-hook 2 (HMGA2) and transforming growth factor beta receptor III (TGFBR3) mRNAs, implicated in lung cancer progression. A separate class of miRNA decoys contains circular RNAs (circRNAs), primarily originating from a non-canonical splicing event, joining 5' donor and 3' acceptor sites. These RNAs show high stability due to their resistance to exonucleolytic degradation pathways and thus effectively sponge miRNAs, as is the case of human ciRS-7/CDR1as circRNA, containing more than 70 miR-7 binding sites (*108*).

An interesting case of mRNA interacting small RNA is piRNA, produced *in trans* in respect to their primary targets, although, similarly to siRNAs, piRNAs can also act *in cis* (discussed bellow). piRNAs are metazoan germline specific short RNA species, essential for genome protection from active transposable elements (TEs). These RNAs are produced from distinct genomic loci, termed piRNA clusters, containing fragments of TE sequences (Figure 20). Initiated as long Pol II transcripts, piRNAs are maturated in the cytoplasm into 24-35 nucleotide-long fragments, antisense to TE mRNAs. Loaded within the effector complexes, piRISC, piRNA mediate cognate TE mRNA cleavage. Importantly, mRNA cleavage intermediates are not targeted by canonical cytoplasmic RNA degradation machinery but rather participate in the generation of secondary piRNA, within the so-called ping-pong

cycle. In addition to the post-transcriptional repression of TE mRNA in the cytoplasm, most animals also silence transposon-producing loci transcriptionally. This pathway relies on *in trans* produced piRNA loading into Piwi protein containing RISC (Piwi-piRISC) that further recognizes nascent TE transcripts in the nucleus and recruits the general heterochromatin-inducing machinery (*109*). Intriguingly, piRNAs can also act *in cis* to activate the transcription of their own cluster, and this activation depends on generally repressive chromatin mark, histone H3 Lys9 trimethylation (H3K9me3), present in the locus. It was shown that accessory chromatin-interacting factors (such as Rhi and Cuff in *Drosophila*), specific for germline cells, determine the activating role of piRNAs (*110, 111*).

In addition to well defined *in trans* produced small interacting RNA classes, several cases of mRNA-mRNA or mRNA-lncRNA (long non-coding RNA) interactions have been described. Interestingly, multiple examples report the involvement of Staufen1 (STAU1) protein, human ortholog of *Drosophila* Staufen, implicated in mRNA localization during developmental processes. STAU1 contains a dsRNA-binding domain, and has been first identified to bind the intramolecular double-stranded structure within the 3' UTR of ADP-ribosylation factor 1 (ARF1) mRNA. STAU1 directly recruits the NMD factor Upf1 and independently from Upf2-Upf3 targets respective mRNA to cytoplasmic decay; this process has been termed STAU1-mediated decay (StMD) (*112*). STAU1 knockdown leads to downregulation of a set of mRNAs, many of which do not contain predictable secondary structures. It has been shown later that mRNAs containing Alu elements within their 3' UTRs can interact with homologous Alu element-carrying lncRNAs (*113*) or mRNAs (*114*) to recruit STAU1 and elicit StMD (Figure 21). Global mapping of STAU1 targets revealed that this factor primarily targets intramolecular dsRNA regions within 3' UTRs (*115*). However, STAU1 has been found to mediate the interaction of several lncRNA-mRNA pairs. Human terminal differentiation-induced ncRNA (TINCR) interacts with a set of mRNAs essential for epidermal differentiation, via a 25-nucleotide "TINCR-box" motif. In contrast to Alu element-bound STAU1, TINCR-associated STAU1 stabilizes its targets and ensures proper differentiation (*116*).

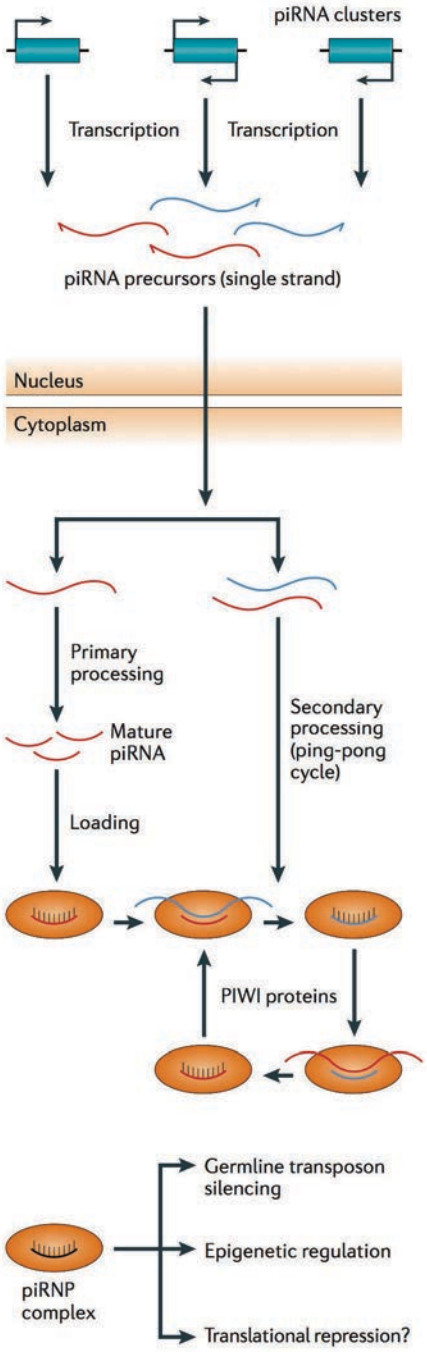

Figure 20. The biogenesis and effector machinery of piRNA.

PIWI-interacting RNAs (piRNAs) are mainly expressed as ssRNA from mono- or bidirectional clusters. Subsequently, additional piRNAs are produced through a PIWI-protein-catalyzed amplification loop (called the "ping-pong cycle") via sense and antisense intermediates. The PIWI ribonucleoprotein (piRNP) complex functions in transposon repression through target degradation and epigenetic silencing. Adapted from (*101*).

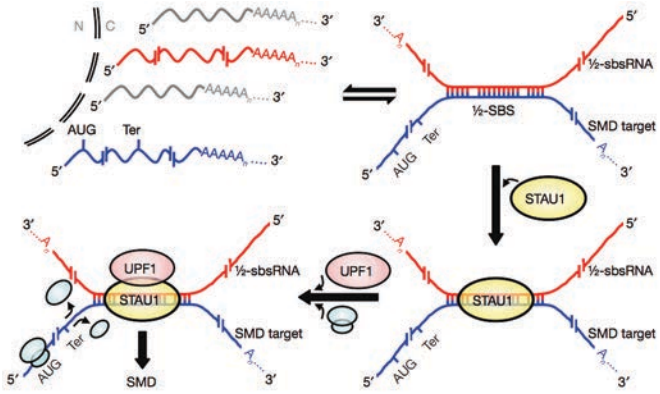

Figure 21. The model of STAU1 mediated mRNA decay (StMD).

An Alu-element-containing 1/2-SBS RNA that is polyadenylated and largely cytoplasmic (red) base-pairs with a partially complementary Alu element, that is, a half-STAU1 binding site (1/2-SBS), within the 3' UTR of a particular mRNA (blue) to trigger StMD. Base-pairing forms a functional SBS. The STAU1-bound SBS triggers StMD in a Upf1-dependent mechanism when translation terminates sufficiently upstream of the SBS so that translating ribosomes (blue ovals) do not remove bound STAU1. The 1/2-SBS RNA is not destroyed in the process. C, cytoplasm; N, nucleus; Ter, termination codon. Adapted from (*113*).

Latest genome-wide data probing RNA-RNA interactions in human, mouse and budding yeast cells reveal the high extent of intermolecular RNA pairing *in vivo* (*90-92*). These studies confirm many well-known interactions between rRNA, rRNA-snoRNA, snRNA-snoRNA, etc., as well as uncover unknown mRNA-mRNA and mRNA-lncRNA interactions. Interestingly, all individually confirmed mRNA-mRNA interactions of human tymosin β_4 (TMSB4X) transcript fall into *in trans* produced RNA category, suggesting that these interactions are prevalent between mRNAs that do not share apparent sequence complementarity. Moreover, interacting mRNA pairs show a probability to cluster within the same gene ontology (GO) term, as observed in human lymphoblastoid and embryonic stem (ES) cells (*90*). However, these data sets have been acquired using chemically assisted dsRNA crosslink and subsequent proximity ligation-based techniques, known to underrepresent perfect (or near perfect) RNA-RNA interactions (e.g. miRNA-mRNA) (*92*). It is thus possible that further technical improvements of genome-wide data acquisition will reveal additional classes of *in trans* produced interacting RNA.

In cis produced interacting RNA

We define *in cis* produced interacting RNA as a pair of partially or entirely complementary RNAs that are transcribed from the same genomic locus and thus their complementarity is inherently dependent of their genomic origin. The most widely studied example of this type of RNA interaction remains siRNA in plants and various fungi (*111*). In parallel, this category of intermolecular RNA-RNA interaction-competent transcripts includes the plethora of various types of sense-antisense RNA pairs that have emerged over the past decade, due to extensive high-throughput RNA sequencing from diverse organisms.

siRNA are 20-25 nucleotide-long RNA species, implicated in antiviral response in various organisms (essential for plants, insects, nematodes, and recently detected in mammalian ES cells (*117*)). siRNA-loaded RISC induces post-transcriptional repression of cognate viral RNAs, similarly to above-discussed miRNA action. However, extensive studies in fission yeast *Schizosaccharomyces pombe* as well as in plants identified a key role for endogenous siRNA in heterochromatin maintenance (*111*). In *S. pombe*, siRNAs are essential for pericentromeric heterochromatin assembly (Figure 22). They originate from long dsRNA-forming sense and antisense transcripts, covering the centromeric region, that are further processed by Dicer endonuclease (Dcr1) to yield siRNA. In addition to these so-called primary siRNAs, a small population of Dcr1-independent siRNA, termed primal siRNA, is produced from ssRNA substrates via nuclear RNA surveillance pathways (*118*). siRNA are loaded to the nuclear RNA-induced transcriptional silencing (RITS) complex, containing an AGO family protein (Ago1) essential for siRNA recruitment. RITS complex interacts with chromatin and with the RNA-dependent RNA polymerase complex (RDRC), the latter mediating the amplification of primary siRNA population. The key step in the current siRNA-mediated transcriptional silencing model, known as *nascent transcript* model, is the siRNA and nascent pericentromeric ncRNA pairing-induced RITS complex recruitment, which later triggers the local deposition of repressive chromatin marks, H3K9me2/3, via the associating CLRC complex. Moreover, RITS complex has an affinity for the H3K9me2/3 containing chromatin, creating the self-reinforcing epigenetic loop for siRNA production (*111*).

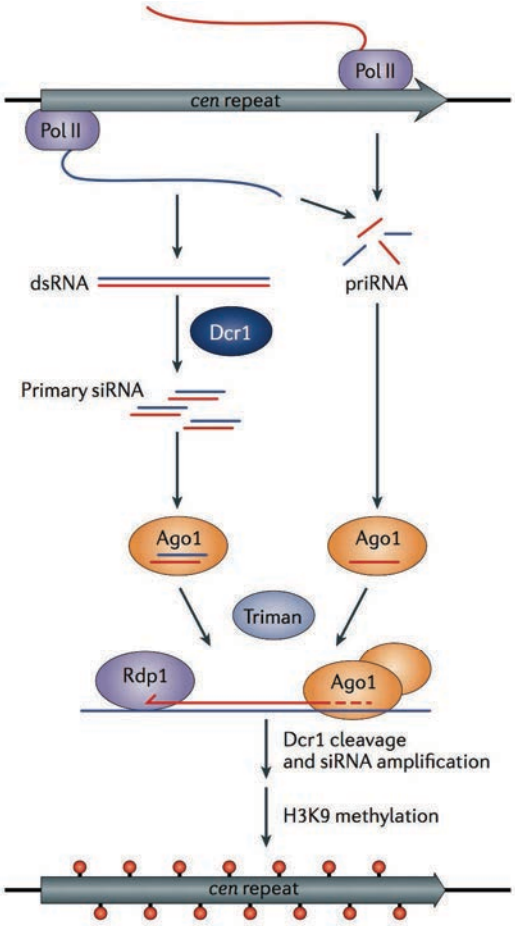

Figure 22. The biogenesis and effector machinery of heterochromatic siRNA.

In *S. pombe*, non-coding transcripts generated from pericentromeric DNA repeats (cen) act as the precursors for heterochromatic siRNAs. Heterochromatic small RNAs arise from Dicer 1 (Dcr1) cleavage of dsRNA formed by the base-pairing of complementary pericentromeric transcripts. The RNA-dependent RNA polymerase Rdp1 amplifies the siRNA pool by generating additional Dcr1 substrates. A small population of Dcr1-dependent, Rdp1-independent heterochromatic siRNAs, known as primary siRNAs, is generated, which is thought to arise from the base-pairing and Dcr1 processing of bidirectionally synthesized transcripts. Additionally, a class of Dcr1-independent heterochromatic siRNAs, known as primal RNAs (priRNAs), is synthetized. All these sub-populations of siRNAs are capable of directing H3K9 methylation in an Ago1-dependent manner, yielding transcriptionally silent heterochromatin. Adapted from (*111*).

Importantly, while siRNA-induced heterochromatin assembly is highly efficient *in cis*, numerous attempts to target *in trans*-produced siRNA for silencing a genomic locus of choice have been generally unsuccessful. Recently, a conserved RNA polymerase associated factor 1 complex (Paf1C) has been identified to suppress the *trans* siRNA-mediated heterochromatin formation. Accordingly, various Paf1C mutants readily allow the silencing of *trans* siRNA-targeted endogenous loci.

Mechanistically, these results emphasize the importance of efficient mRNA 3'-end formation and nascent transcript release, as mRNA retention at the chromatin opens the kinetic way for its targeting by the RNAi machinery (*119*).

In plants, transposable elements and repetitive DNA sequences are silenced transcriptionally by siRNA-guided heterochromatin formation and DNA methylation. In contrast to *S. pombe*, *A. thaliana* has multiple Dicer, AGO and RDRC homologues, each specific for a particular RNAi-based pathway (*120*). However, similarly to fission yeast, plant siRNA-bound Ago4 targets the nascent cognate sequence-containing transcript in the nucleus and triggers the local DNA cytosine methylation. Additionally, H3K9me nucleosomes and methylated DNA show affinity to the accessory factors of plant specific RNA Pol IV and Pol V, respectively, directing siRNA precursor transcription and creating the above-mentioned self-reinforcing epigenetic loop (*111*). Importantly, due to the potent RNAi machinery, plant cells actively protect endogenous mRNAs from entering siRNA-producing pathway, leading to transcriptional or post-transcriptional repression. The cellular RNA decay factors—nuclear as well as cytoplasmic—play the key role in ensuring the timely degradation of endogenous transcripts; accordingly, various RNA decay mutants accumulate siRNA corresponding to endogenous coding loci (*121*). Recently, it has been shown that decapped Pol II transcripts are readily transformed into dsRNA leading to siRNA production, while capped RNAs yield siRNAs only upon pairing with an antisense transcript (*122*). This connection between the RNAi and RNA decay pathways underlines an intricate crosstalk within RNA processing mechanisms, suggesting its substantial regulatory potential.

An intriguing case of interplay between RNAi and RNA maturation pathways has been reported in *Cryptococcus neoformans*, a pathogenic budding yeast. *C. neoformans* has a functional RNAi pathway that is essential to protect its genome from TE activation. In contrast to siRNA biogenesis in fission yeast, cryptococcal siRNA production is associated to the nuclear splicing machinery (Figure 23). In fact, transposon mRNAs are enriched in inefficient splicing sites and get stalled on spliceosomes. This stalling then triggers the activity of the spliceosome-bound SCANR complex that contains RNA-dependent RNA polymerase subunit. Inefficiently spliced transposon mRNA is subsequently transformed into dsRNA substrate for Dicer, yielding TE-directed siRNA (*123*).

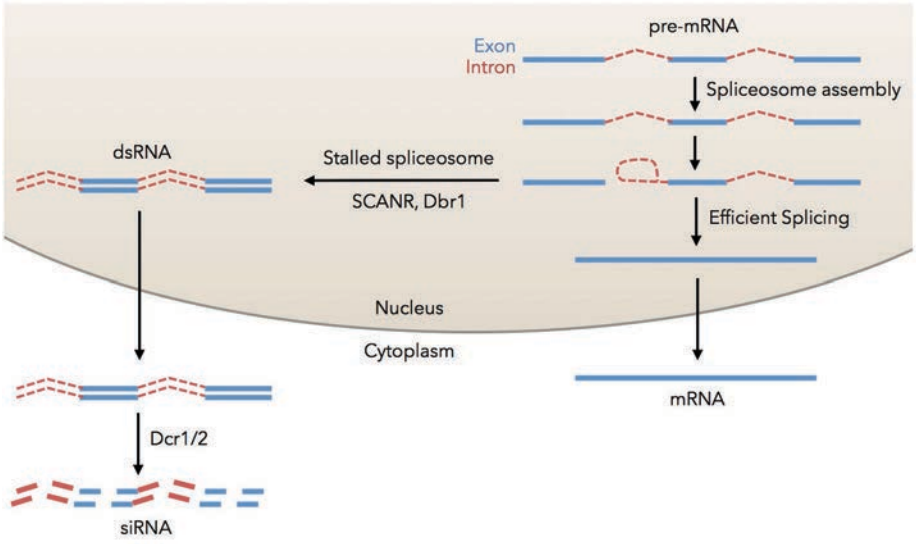

Figure 23. The kinetic competition model for SCANR-mediated genome defense.

The utilization of particular mRNA precursors by SCANR is influenced by their splicing efficiency due to a kinetic competition between splicing and dsRNA synthesis: transcripts that are inefficiently spliced, such as foreign genetic elements, exhibit increased spliceosome association, thereby facilitating their conversion to dsRNA, and, ultimately, siRNA. In this hypothetical example, a lariat intermediate is produced by stalled splicing of a transcript's first intron; downstream introns remain incompletely spliced. This intermediate is acted upon by SCANR and debranching enzyme Dbr1 to produce a dsRNA substrate for Dcr1/2. Adapted from (*123*).

In mammalian cells, siRNA-mediated transcriptional gene silencing has been confirmed significantly later than in other organisms (*124*), probably due to a delayed evidence for the existence of endogenous siRNA (discussed below). However, exogenous siRNA can trigger chromatin modification and DNA methylation-associated transcriptional repression in human cells (*125*). This effect follows the previously described nascent transcript model, as transcription *in cis* (even at very low level) is essential for siRNA-mediated heterochromatin assembly, underlining the importance of RNA-RNA interactions for this mechanism (*126*). Interestingly, siRNA-induced local deposition of repressive chromatin marks within coding regions induces Pol II pausing, resulting in increased rates of alternative splicing (*127*). Recent studies have also revealed a crosstalk between Pol II transcription termination and siRNA in human cells. Nascent RNA hybrids with homologous DNA strand (also known as R-loops), appearing upon Pol II pausing on GC-rich terminator

sequences, have been suggested to trigger siRNA production and local deposition of repressive chromatin marks, further contributing to Pol II release (*128*).

2 | The nuclear fate of sense-antisense RNA pairs

In addition to *in cis* produced siRNA, the major reservoir of interaction-competent long RNA species lies in the widespread antisense transcription. The existence of overlapping transcription units is readily detectable in organisms containing compact genomes (e.g. budding or fission yeast), but has been equally confirmed in metazoans and plants (*129-131*). In *S. cerevisiae*, the convergent orientation between neighboring coding genes is highly prevalent and the transcription of these genes generates 3' overlapping mRNAs (*9*). Similarly, animal genomes contain a significant portion of overlapping coding transcripts, and some convergent gene pairs show non-random retention during evolution (*132-134*). Moreover, various non-coding transcription units complement the coding landscape of the cell, creating a phenomenon known as pervasive transcription (*18*). Over the past decade, extensive high-throughput RNA-sequencing experiments revealed the universal presence of antisense transcription units in diverse organisms (Figure 24), and functional roles continue to emerge for these interaction-competent RNAs (*135, 136*). The regulatory outcomes of sense-antisense (comprising both coding and non-coding) RNA pairing are reviewed in following sections, distinguishing nuclear and cytoplasmic events.

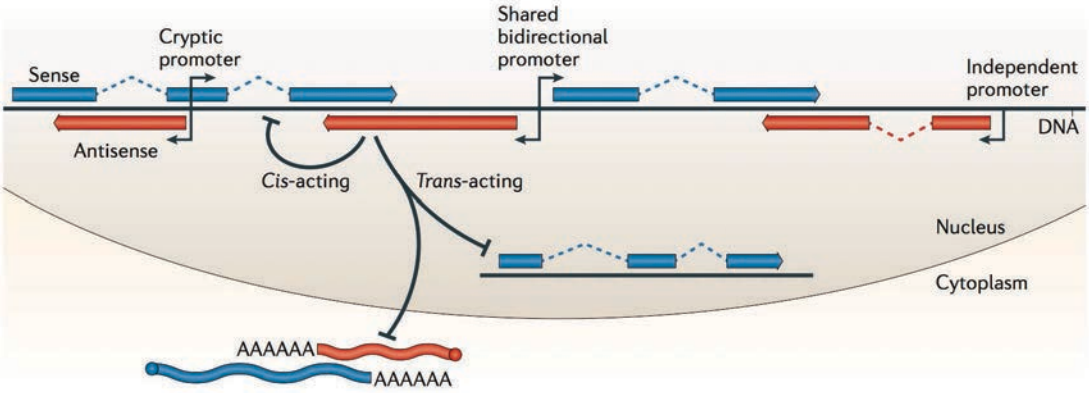

Figure 24. The variety of antisense transcription.

Antisense transcription can initiate from a dedicated promoter, share a bi-directional promoter with a divergent transcription unit, or arise from a cryptic promoter within another gene. In regard to sense transcription units, it might be tail-to-tail, head-to-head oriented or overlap internally. Mechanistically, antisense transcription can act *in cis* or *in trans*, the later effect localized to the nucleus (e.g. regulating alternative splicing) or to the cytoplasm (e.g. regulating RNA stability). Adapted from (*136*).

Antisense transcription mediated chromatin remodeling

RNA Pol II is associated with a powerful machinery of chromatin modifications, and distinct histone marks have been detected for promoter, enhancer, gene body, and terminator DNA regions. Accordingly, overlapping transcription units might interfere with respective chromatin modifications and result in the biased expression from only one DNA strand. The increasing amount of evidence suggests that the act of transcription *per se* accounts for an important contribution in antisense transcription mediated gene regulation (*137*). In *S. cerevisiae*, various physiological processes have been shown to depend on antisense transcription induced chromatin remodeling. The thoroughly studied *GAL1-GAL10* locus is activated when galactose becomes the sole available carbon source and is efficiently repressed when glucose is available. This locus produces various antisense ncRNAs that are crucial for maintaining a sensitive threshold for *GAL* locus activation (*138*). Mechanistically, H3K4 and H3K36 methylation by Pol II associated histone methyltransferases Set1 and Set2, respectively, leads to the recruitment of histone deacetylase complexes Set3C and Rdp3C, further mediating repressive chromatin states. The changes in chromatin marks during different states of *GAL1-GAL10* locus (*139, 140*) have been

later generalized genome-wide (*141*), implementing the model of co-transcriptional chromatin remodeling (*135-137*).

In metazoan organisms, antisense transcription has been also linked with repressive histone marks and DNA methylation. It is the case of human hemoglobin α1 (*HBA1*), tumor suppressor locus *CDKN2B–CDKN2A*, mouse insulin-like growth factor 2 receptor (*Igf2r*), among many others, that are targeted by antisense transcription-induced transcriptional repression (*136*). Numerous antisense ncRNAs have been found associated with mammalian Polycomb repressive complex 2 (PRC2), the major actor in repressive histone mark deposition, suggesting that antisense transcripts might recruit the chromatin remodeling machinery *in cis* (*142*). However, this model has been challenged lately by multiple genome-wide observations that witness the low specificity of PRC2-RNA interactions and incites to propose a more intricate relationship between ncRNA molecule itself and downstream chromatin modifications (*111*).

An interesting aspect of antisense transcription is the mechanistic configuration of RNA Pol II complexes, moving in convergent directions. In the case of two strong promoters in sense-antisense orientation, Pol II collision has been observed in budding yeast (*143, 144*). Indeed, most experimental evidence comes from studies of artificial convergent transcription initiated at strongly induced *GAL* promoters and the genome-wide prevalence of endogenous Pol II collision remains largely unexplored. It has been suggested that collided polymerases are targeted by ubiquitin-mediated proteasomal degradation, as in cells with inefficient Pol II ubiquitinylation paused polymerases can be detected within convergent genes (*143*). However, the most prevalent situation of convergent transcription is thought to happen at distinct time points for sense and antisense directions (*136*). It might lead to the existence of overlapping transcripts in the nucleus or in the cytoplasm (discussed below), although latest single molecule fluorescence *in situ* hybridization (smFISH) experiments suggest a frequent bi-modal sense-antisense RNA expression pattern. In *S. cerevisiae*, multiple antisense ncRNA appear to be expressed in a mutually exclusive manner in relation to sense RNA at a single cell level (*137*).

As evoked previously, *S. cerevisiae*, together with the members of *Saccharomyces* sensu stricto clade (but also including some distant yeast species), has lost RNAi during evolution. It has been shown that the absence of dsRNA-mediated genome

defense pathway favored the acquisition of dsRNA viruses, conferring the advantageous killer phenotype (*145*). Importantly, a recent study has reported that upon RNAi loss, various budding yeast species have undergone a significant expansion of antisense non-coding transcription, as compared to the RNAi competent counterpart, *S. castellii* (*146*). Furthermore, an artificial reconstitution of RNAi in *S. cerevisiae* has permitted to identify a set of ATP-dependent chromatin remodelers as active suppressors of overlapping non-coding transcription (*147*). These results support the importance of concerted expression within the overlapping transcription units, and chromatin dynamics plays a key role in this regulation.

Small RNA mediated chromatin remodeling

The presence of overlapping sense-antisense RNA pairs within the cells competent for RNAi—in contrast to *S. cerevisiae*—might lead to the potential dsRNA recognition by Dicer type of enzymes, producing siRNA and triggering downstream events. The examples from diverse organisms converge to the model suggesting that overlapping transcripts form dsRNA structures in the nucleus, although it is not the exclusive outcome of sense-antisense transcription. In human HEK923 cells, nucleus-localized Dicer has an important function in limiting endogenous dsRNA accumulation. Upon Dicer knockdown, dsRNA accumulate in the nucleus and later trigger interferon-mediated inflammatory cascades in the cytoplasm, eventually leading to apoptosis. The study reporting Dicer association with transcribing Pol II proposes that upon convergent transcription-dependent dsRNA formation, Dicer produces *in cis* acting siRNAs leading to the local deposition of repressive H3K9me2 marks and subsequent limiting of convergent transcription (*148*). However, another study found that HEK293 cells produce a relatively small amount of endogenous siRNAs, corresponding to less that 2% of coding genes and of which only a half could account on overlapping sense-antisense transcripts. These results suggest that the efficiency of siRNA production from convergent transcription units might depend on cell type specificity (*149*).

Indeed, recent data provide evidence that endogenous siRNAs originating from convergent loci are stably produced in mouse germ cells. The siRNA signature has been first detected in mouse oocytes and early preimplantation embryos and,

importantly, not only *in cis* synthesized sense-antisense RNA pairs but also *in trans* produced gene-pseudogene interacting RNAs are sources of endogenous siRNAs. These short RNA species target transposons and other coding/non-coding genes for transcriptional silencing; even though it shares some targets with piRNA, endogenous siRNA has been shown to be structurally and functionally distinct from piRNA (*150, 151*). Later on, mouse male germ cells have been also shown to contain high levels of endogenous siRNAs (*152*). Interestingly, the occurrence of endogenous siRNA in mouse oocytes coincides with the downregulated activity of miRNA biogenesis pathway during this particular cellular context (*153, 154*). The explanation for this phenomenon came from the study that identified the oocyte specific Dicer isoform, lacking the N-terminal DExD helicase domain present in the somatic isoform. This N-truncated Dicer shows increased affinity for perfectly complementary dsRNA substrates in contrast to mismatch-containing canonical pre-miRNA stem-loops (*155*). The high level of siRNA production in germline and ES cells has been associated to the downregulation of interferon-induced inflammatory response in these cell types (*148*).

Isolated examples of endogenous convergent loci-derived siRNA have been reported in diverse organisms (*156*), including mouse brain cells (*157*), zebrafish embryos (*158*), and various human cell types (*124*). These short RNA species have been also characterized in *Drosophila* somatic cells, mirroring the function of germline piRNA (*159, 160*). In fission yeast, convergent transcription induces transient heterochromatin formation within the overlapping regions. This phenomenon is RNAi-dependent, as sense-antisense-derived siRNAs are loaded to the RITS complex, resulting in the deposition of H3K9me3 chromatin marks. Interestingly, *S. pombe* methylated H3K9 recognizing factor Swi6 (ortholog of metazoan heterochromatin protein 1, HP1) binds nuclear cohesin and limits the overlapping transcription. This interaction occurs during the G1-S phase of the cell cycle; after mitosis, cohesin is released from intergenic regions between convergent genes and the subsequent transcription produces overlapping RNA again (*161*). Moreover, multiple factors essential for RNAi (including Dcr1, Ago1, Rdp1 among others) are in convergent genomic orientation in *S. pombe*. Accordingly, the expression of these factors decreases during the G1-S phase, autoregulating the pathway in which they are involved (*162*).

An interesting example of convergent RNA processing into siRNA and subsequent chromatin remodeling comes from the studies of filamentous fungus *Neurospora crassa*. Among multiple types of small RNA species, this organism harbors Dicer-independent siRNA (disiRNA), an abundant population of approximately 22 nucleotide-long RNAs, predominantly arising from convergent transcription units (*163*). It has been found that disiRNAs induce a dynamic DNA methylation at corresponding loci, which are further enriched for repressive H3K9me3 marks. The exact mechanism of disiRNA-mediated DNA methylation is unknown but has been shown to differ significantly from the well-described RIP DNA methylation pathway in *N. crassa* (*164*). The fact that methylation of disiRNA-producing loci is dependent on transcription suggests the mechanism related to the nascent transcript model, previously described in *S. pombe*. However, disiRNA biosynthesis is independent of Dicer, emphasizing the parallel variety of siRNA-mediated transcriptional repression in diverse fungi.

Antisense RNA mediated alternative splicing

Along the antisense transcription- and endogenous siRNA-induced chromatin remodeling, post-transcriptional RNA-RNA interactions might take place in the nucleus. Particularly, *in cis* produced overlapping RNAs have been reported to promote alternative splicing (*135, 136*). The earliest example is the human B cell thyroid hormone receptor-α (TRα) mRNA, containing two alternatively spliced isoforms, TRα1 and TRα2, the ratio of which depended on the levels of antisense transcript RevErbAα (*165*). Another case study reported the alternative splicing of Zeb2 mRNA, induced upon epithelial-mesenchymal transition (EMT) in human cells. ZEB2 codes for a transcriptional repressor of E-cadherin and its synthesis is upregulated upon EMT induction. This upregulation is mediated via the retention of a large intron within the 3' UTR of Zeb2 mRNA, containing an IRES element and thus promoting Zeb2 translation. Importantly, an antisense transcript, overlapping the included 3' UTR exon, is essential for this alternative splicing event (*166*). This mechanism of splicing site masking by complementary RNA is reminiscent of the widely used antisense oligonucleotide-mediated alternative splicing technique, which is of great experimental and therapeutic importance (*167*).

Two additional examples, associating antisense RNA with alternative splicing, suggest a scaffold role for ncRNA. Fas, a member of tumor necrosis factor receptor superfamily, plays a major role in apoptosis signaling and is active in its membrane-bound form (mFas). The skipping of exon 6 in Fas mRNA results in a soluble form of the receptor (sFas), incapable of triggering apoptosis; sFas expression is significantly increased in various types of leukemia. An antisense ncRNA to Fas, FAS-AS1, suppresses the exon 6 skipping, however FAS-AS1 promoter DNA is often hypermethylated in cancer cells, leading to decreased FAS-AS1 production and subsequent increased Fas exon 6 skipping. Mechanistically, FAS-AS1 has been found to associate with RNA binding motif protein 5 (RBM5), a factor interacting with the splicing machinery (*168*). Similarly, fibroblast growth factor receptor 2 (FGFR2) contains the optional exon IIIb in epithelial-state cells, and looses it upon EMT. The alternative splicing of FGFR2 depends on the chromatin state of the locus but also on the epithelial specific antisense ncRNA, asFGFR2. This antisense transcript acts in recruiting PRC2, which in turn deposes splicing-modulating chromatin marks. Surprisingly, asFGFR2 has a comparable effect on FGFR2 splicing if expressed *in trans* and, furthermore, this regulation is not dependent on RNAi machinery (*169*). Altogether, these two examples underline the mediator role of antisense transcription in regulating alternative splicing, which is not mutually exclusive with the above-mentioned cases of direct RNA-RNA interaction-mediated exon inclusion.

Genome-wide studies have also confirmed the global effect of sense-antisense transcription units on alternative splicing. Analysis of 176 human lymphoblastoid cell lines yielded a significant correlation between the presence of overlapping RNA regions and exon retention within these regions. For example, certain 3' proximal exons of mutS homolog 6 (MSH6) mRNA were alternatively skipped/included upon the presence of overlapping F-box protein 11 (FBXO11) mRNA; of notice, this is a case of convergent mRNA-mRNA pair-dependent alternative splicing. Interestingly, the same study has found that alternative splicing coincides with antisense transcription in another 12 metazoan species. Mechanistically, Pol II occupancy increases within the overlapping regions between sense-antisense transcription units, and this observation has been proposed to corroborate the endogenous siRNA-mediated alternative splicing model (see section 1 or Part II). However, no detailed mechanism has been proposed to explain these genome-wide correlations (*170*).

A similar study in *A. thaliana*, analyzing exclusively sense-antisense coding gene pairs, has also revealed a significant bias for spliced (including alternatively spliced) mRNAs among the large population (approximately 2000 transcripts) of convergent and overlapping messengers. Interestingly, this report suggests siRNA-independent outcomes of convergent mRNA interactions, based on the highly prevalent co-expression of both sense and antisense transcripts (*171*). However, a detailed study of three overlapping convergent gene pairs identified the unexpected existence of alternative polyadenylation sites, limiting the extent of sense-antisense overlap (*172*). The high prevalence of alternative polyadenylation in plants has also been confirmed genome-wide, with potential interconnections to antisense transcription (*173*).

Recently, the nuclear maturation of cancer-related lncRNA MALAT1 has been reported to depend on an antisense ncRNA. MALAT1 (also known as NEAT2) is an abundant nuclear transcript, involved in metastatic progression in various types of cancers. Among other functions, MALAT1 regulates the alternative splicing of B-MYB, an oncogenic transcription factor, potentially recruiting accessory splicing factors (*174*). Interestingly, MALAT1 undergoes a non-canonical nuclear maturation. It contains a tRNA-like structure at its 3' end that is removed by the nuclear RNase P; this 3' end byproduct is a stable tRNA-like fragment, called mascRNA, accumulating in the cytoplasm. After the cleavage, the 3' end of MALAT1 folds into a triple helix structure, acting as a stabilizing element. Importantly, an antisense ncRNA to MALAT1, termed TALAM1, is essential for this 3' end cleavage. The absence of TALAM1 decreases significantly MALAT1 levels, and this effect is mutually observable, as MALAT1 acts in stabilizing TALAM1 as well (*175*). Although the exact mechanism of this co-regulation is unclear, it broadens the scope of antisense RNA-mediated nuclear events.

3 | The cytoplasmic outcomes of sense-antisense RNA interaction

Although overlapping sense-antisense RNA can interact in the nucleus as discussed above, most Pol II transcripts are readily exported into the cytoplasm. As evoked previously, various studies have shown that inefficient transcription termination, nascent transcript release, or defective splicing might trigger nuclear RNAi response (*119, 123*). In contrast, the canonical co-transcriptional mRNP assembly favors rapid mRNA exclusion from the nucleus, and non-coding Pol II transcripts generally also follow this rule (*137*). Moreover, if the convergent transcription is thought to proceed at different time points from sense and antisense orientation, the highest probability of co-existence (and, potentially, interaction) of overlapping RNA remains in the cytoplasm (*136*). Further sections thus focus on the cytoplasmic fate of sense-antisense RNA pairs, reviewing the latest examples of RNA-RNA interaction-mediated remodeling of mRNA stability and translation.

Remodeling RNA stability by RNA-RNA interactions

The majority of the examples reporting RNA-RNA interaction-induced transcript stabilization come from the study of mammalian systems, with a few cases from other metazoans. In rat hepatocytes, inducible nitric oxide synthase (iNOS) responds to the interleukin 1β (IL-1β) signaling, leading to nitric oxide production. iNOS induction correlates with the increase of iNOS mRNA levels and the production of antisense ncRNA, iNOS-AS. This pair of convergent transcripts interacts in the cytoplasm and recruits the AU-rich element binding protein HuR, via the ARE located in the 3' UTR of both RNA (*176*). HuR, also associated to hnRNP L, is a well-known RNA binding protein implicated in the stabilization of mRNA in the cytoplasm, counteracting numerous destabilizing ARE-BPs (*73*). iNOS-AS is of a therapeutic interest as it is conserved in humans and expressed in hepatocarcinoma tissues. Artificially downregulating cytoplasmic levels of iNOS-AS effectively destabilizes iNOS mRNA (*177*). Similarly, the fibroblast growth factor receptor 3 (FGFR3) mRNA is stabilized in the cytoplasm by its 3' overlapping antisense non-coding transcript FGFR3-AS1. This interaction is also clinically relevant as downregulating FGFR3-AS1 reduced FGFR3 mRNA levels, resulting in

proliferation arrest of osteosarcoma cells (*178*). A noteworthy case of mutual stabilization by RNA-RNA interaction is a gastric cancer related pair DHPS-WDR83 (coding for deoxyhypusine synthase and WD repeat domain 83 protein, respectively), of which both transcripts are overlapping mRNAs (*179*).

Expression patterns of human cerebellar degeneration-related protein 1 (*CDR1*) constitute an intricate example of multilayered gene regulation. An antisense ncRNA to CDR1, termed CDR1as, is important to stabilize CDR1 mRNA levels in the cytoplasm, independently of chromatin remodeling. However, as evoked previously, CDR1as undergoes a non-canonical back-splicing event and functions as a circRNA; its increased stability is in turn directed towards CDR1 mRNA stabilization. CDR1as contains a near-perfect binding site for miR-671 and thus its nuclear turnover is initiated by the Ago2-mediated cleavage (*180*). Furthermore, CDR1as contains a large number of binding sites for miR-7 (although not perfectly complementary and thus incapable of inducing cleavage), and it was subsequently renamed ciRS-7 (for circular RNA sponge for miR-7). Intriguingly, CDR1/ciRS-7 competes for miR-7 binding with CDR1 mRNA, this way additionally contributing for CDR1 expression (*181*).

Mechanistic outcomes of the cytoplasmic RNA-RNA interaction in metazoan cells can also interfere with miRNA pathway. Several studies have reported sense-antisense RNA pairing-mediated transcript stabilization due to inaccessible miRNA binding sites. The expression of human β-secretase-1 (*BACE1*), implicated in Alzheimer's disease pathophysiology, is positively regulated by its antisense transcript, BACE1-AS (*182*). This antisense lncRNA competes with miR-485-5p for binding the same region within BACE1 mRNA and, concomitantly, BACE1-AS overexpression or miR-485-5p knockdown have a comparable effect in stabilizing BACE1 mRNA (*183*). Similarly, mouse sirtuin 1 (Sirt1) mRNA is stabilized in the cytoplasm by its convergent and 3' overlapping lncRNA, Sirt1 AS. The interaction between these RNAs masks the miR-34a binding site within Sirt1 3' UTR, inhibiting its targeting by the miRISC complex. Functional stabilization of Sirt1 transcript in the cytoplasm has been shown to promote myoblast proliferation and delay differentiation (*184*).

In contrast to miRNA site masking, dsRNA structures arising from sense-antisense RNA pairing might also trigger cytoplasmic RNAi pathway. The first

example of the cytoplasmic siRNA-mediated cleavage of overlapping RNAs came from a study in *Arabidopsis*. Upon high salt stress, similar to RCD one 5 (SRO5) transcription is induced, producing an mRNA that is convergent and overlapping with Δ^1-pyrroline-5-carboxylate dehydrogenase (D5CDH) mRNA. The increase in SRO5 mRNA levels corresponds with the decrease of D5CDH mRNA, repressing D5CDH activity and contributing to the intracellular increase of proline, important for high salt response. The detailed analysis of SRO5-D5CDH mRNA-mRNA interaction revealed that firstly, Dicer-like 2 (DCL2) cleaves this duplex and produces 24 nucleotide-long endogenous siRNA. Later on, D5CDH mRNA is transformed into dsRNA by RDR6 and processed into secondary 21 nucleotide-long siRNA by DCL1. Both populations of siRNA are loaded to the RISC complex and effectively degrade D5CDH transcript (*185*). A recent example of modulating transcript stability in *A. thaliana* within a convergent mRNA-mRNA pair includes NFYA5-NERF, nuclear transcription factor Y subunit A-5 and NFYA5 enhancing ring finger, respectively. Particularly, the 3' overlapping region between two mRNAs shares a target site with miR169, creating an intricate interplay between various small RNA-dependent pathways (*186*). Similarly, a pair of convergent mRNAs has been reported to control neuronal excitability in *Drosophila*. Two ion channel genes, *pickpocket 29* (*ppk29*) and *seizure* (*sei*), give rise to 3' overlapping mRNAs. Interestingly, it is the potassium channeling activity of *sei* that is important for neuronal excitability, while *ppk29* mRNA is essential in regulating *sei* mRNA levels independently from ppk29 protein. This mechanism depends on endogenous siRNA-producing Dicer-2, suggesting that *sei* mRNA decay is mediated by the cytoplasmic siRISC (*187*).

The outcomes of cytoplasmic sense-antisense RNA interaction in non-metazoan organisms have been recently addressed. Although transcriptome analysis of various fungi confirmed the universally widespread antisense transcription (*188*), only few examples exist that target cytoplasmic RNA-RNA interaction. In one study, several hundreds of convergent RNAs have been identified in basidiomycete fungus *Ustilago maydis*, predominantly in diploid teliospores, the pathogenic form *U. maydis* that infects maize. Interestingly, these RNAs are not present in haploid fungus and correlate with the increased stability of their convergent counterparts; more than a half of these antisense transcripts are mRNAs. Several antisense transcripts have been associated with the specific life cycle form of this fungus and, furthermore,

showed conservation in a close-related *U. hordei*. This conservation is of a particular interest, as *U. maydis* has lost the RNAi pathway (similarly to *S. cerevisiae*) while *U. hordei* contains all canonical RNAi factors. However, the conserved antisense transcripts accumulated in dormant teliospores of both species, and, in accordance, Dicer as well as RNA-dependent RNA polymerase were not expressed in *U. hordei* at this specific stage. The authors propose the model where timely antisense transcription in dormant *U. maydis* teliospores contributes to mRNA storage in the cytoplasm, for subsequent translational activation during germination (*189*).

A recent report has addressed the differential stability of various classes of antisense ncRNA in *S. cerevisiae*. Budding yeast harbors a large panel of stable unannotated transcripts (SUTs), as well as Xrn1-sensitive unstable transcripts (XUTs), both non-coding RNA species. These RNAs are often antisense to coding transcription units and frequently map to the same genomic region, suggesting they might be isoforms. Importantly, XUTs are degraded in the cytoplasm by Xrn1, and thus are efficiently transported from the nucleus; it is unknown to what extent SUTs occur in the cytoplasm. The genome-wide 5' end profiling of SUTs and XUTs has revealed that a fraction of these transcripts are actually isoforms, with XUTs containing 3' extensions. Upon RNAi reconstitution in *S. cerevisiae* (expressing Dicer and Argonaute from *S. castellii*), mRNA-overlapping SUTs and XUTs produce readily detectable siRNAs, suggesting that both ncRNA species exist in the cytoplasm and also in the same cell as their convergent mRNAs. Interestingly, the cytoplasmic decay of some XUTs depends on the NMD machinery, indicating that these ncRNAs are targeted by the translation-associated quality control. The authors of this study propose a mechanism where the differential overlap status between mRNA and SUT/XUT determines the decay rate of the latter. 3' extension containing XUTs are potential targets of RNA helicases Mtr4 and Dbp2, as shown for the ARG1 mRNA interacting XUT1678. Once unwound by these helicases, XUT1678 can undergo translation and subsequent decay by NMD. Accordingly, genome-wide ribosome profiling data in *S. cerevisiae* suggest that ribosomes associate preferentially with the 5' portion of ncRNA. In contrast, perfectly complementary to ARG1 mRNA, SUT768 forms a duplex that is not accessible for Mtr4 and Dbp2, protecting it from translation and decay (*190*).

Remodeling mRNA translation by RNA-RNA interactions

Sense-antisense RNA interaction in the cytoplasm—when at least one transcript from the pair is coding—might interfere with translational processes. Convergent transcription frequently results in 3' overlapping RNAs and antisense RNA can potentially extend to the ORF region of sense RNA. Alternatively, antisense transcription might initiate within the gene body of sense RNA (Figure 24), and create 5' or even fully overlapping RNAs. In all cases, sense-antisense RNA pairing might influence translation. Indeed, an early example of human tumor necrosis factor receptor superfamily member 17 (TNFRSF17, also known as BCMA) mRNA translation confirmed a role of an antisense RNA: its overexpression significantly reduced TNFRSF17 protein production, without altering mRNA levels (*191*). Similarly, human hematopoietic transcription factor PU.1 (or SPI1) translation is negatively regulated by an antisense ncRNA. In the cytoplasm, this pair of sense-antisense transcripts interacts, and a selective knockdown of the antisense RNA increases the protein output of PU.1, without influencing the levels of mRNA. Interestingly, PU.1 antisense ncRNA has been detected in cell lines that do not express PU.1 sense, suggesting that the translational repression of this B-lineage specific factor is an additional, fail-safe mechanism to ensure its timely expression (*192*).

In contrast to antisense RNA mediated translational repression, sense-antisense RNA interaction has been also reported to activate translation. In the case of mouse ubiquitin carboxyterminal hydrolase L1 (Uchl1) mRNA, its translation is modulated by a 5' overlapping antisense lncRNA, AS Uchl1. Importantly, this antisense non-coding RNA contains an inverted SINE (short interspersed nuclear element), termed SINEB2, that is essential for regulating Uchl1 mRNA translation. AS Uchl1 is a predominantly nuclear transcript; however, upon various stress-stimuli (e.g. mTORC1 inhibition by rapamycin), it relocates to the cytoplasm where it upregulates Uchl1 mRNA translation. The 5' complementarity between sense and antisense Uchl1 transcripts is required for this effect and similarly to the above-described examples, AS Uchl1 increases Uchl1 protein levels without altering Uchl1 mRNA expression pattern. Of notice, AS Uchl1 acts in the context when cap-dependent translation is inhibited (i.e. upon rapamycin treatment), suggesting that

antisense RNA-mediated translation upregulation acts to ensure important cap-independent protein synthesis. The case of Uchl1 is not unique, as analogously oriented ubiquitously expressed transcript (Uxt) and its SINEB2-containing AS Uxt also follow the same translational activation mechanism (*193*). The efficiency to upregulate translation by this RNA-RNA interaction has lead to investigations on the therapeutic use of SINE-containing antisense RNA (*194, 195*).

Several additional examples confirm the antisense RNA-induced translational activation. In humans, RNA binding motif protein 15 (RBM15), a regulator of megakaryocyte (MK) differentiation, is activated by the transcription factor RUNX1. Mechanistically, RUNX1 induces the transcription of RBM15 but also of its antisense lncRNA AS-RBM15, overlapping the 5' UTR of RBM15. These two RNAs interact in the cytoplasm, yielding significantly increased RBM15 protein production. Interestingly, the overlapping region alone of AS-RBM15 is sufficient to activate RBM15 translation, suggesting an alternative mechanism to SINE-dependent effect, discussed above (*196*). A similar case has been reported in rice plant *Oryza sativa*, where the translation of a protein essential for phosphate metabolism (PHO1;2) is regulated by a cognate antisense ncRNA (termed *cis*-$NAT_{PHO1;2}$). PHO1;2 mRNA is constitutively expressed in roots and leaves, while *cis*-$NAT_{PHO1;2}$ is induced in these organs by phosphate insufficiency. Critically, PHO1;2 mRNA levels remain stable during differential phosphate supply, and only upon *cis*-$NAT_{PHO1;2}$ expression PHO1;2 protein production is activated (*197*).

To our knowledge, no global analysis on antisense RNA-dependent translational regulation has been performed up to date. In contrast to a widespread use of RNA sequencing techniques, employed to detect subtle changes in cellular transcriptomes, quantitative proteomics is currently emerging as a powerful tool to follow translational dynamics and will undoubtedly provide the missing data on the interplay between regulatory circuits of RNA-RNA interaction and translation.

The diversity of outcomes from sense-antisense RNA interaction, addressed in the second part of this chapter, is illustrated in Figure 25.

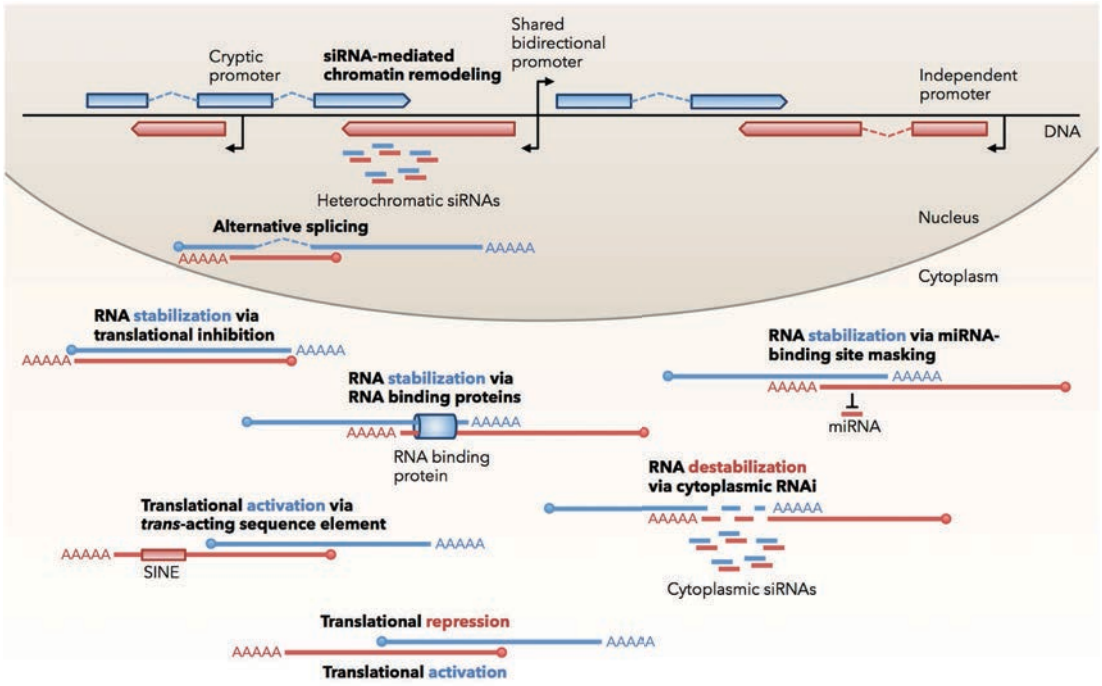

Figure 25. The diversity of outcomes from sense-antisense RNA interactions.

In the nucleus, *in cis* produced interacting sense-antisense RNAs might induce transcriptional silencing via siRNA-mediated chromatin remodeling. In parallel, RNA-RNA interaction might promote alternative splicing. In the cytoplasm, RNA-RNA interaction might lead to RNA stabilization (in blue) or destabilization (in red). The stabilizing effect depends on RNA-binding proteins, masked miRNA binding sites, or inhibited translation; the destabilizing effect is triggered by cytoplasmic RNAi. RNA-RNA interaction might also affect translation, activating it (in blue) or repressing it (in red). A sole mechanism for translational activation has been described, relying on a repetitive sequence element (SINE) embedded in the *trans*-acting RNA. Mechanisms of other examples reporting translational activation or repression by RNA-RNA interaction remain unknown.

PART III | Objectives of the study

The model organism *Saccharomyces cerevisiae* has a compact genome, resulting in the frequent overlap between coding and non-coding transcription units, as evoked in Part II of this chapter. Several cases of transcriptional repression have been reported for non-coding antisense RNA-containing genes (*198, 199*), however the majority of antisense transcription events do not cause significant downregulation of the sense RNA (*200*). Moreover, very few examples exist that have addressed the interplay between convergent coding transcripts in budding yeast (*201*), and no data is available on the cytoplasmic outcome of potential mRNA-mRNA interactions. Importantly, *S. cerevisiae* is naturally devoid of the RNAi pathway, positioning this organism as a particularly attractive genetically tractable model to study long-range RNA interactions independently of siRNA production. As discussed previously, an increasing amount of evidence accumulates suggesting the relevance of RNAi-independent outcomes of RNA-RNA interaction in organisms that harbor fully functional RNAi pathway. Finally, RNAi machinery can be artificially reconstituted in *S. cerevisiae*, providing the possibility to capture the *in vivo* signature of intermolecular RNA-RNA duplexes (*51*).

Our previous results, coming from the physiological regulation studies of the cytoplasmic 5'-3' RNA decay pathway, revealed a list of proteins whose expression was strictly dependent on the functionality of this pathway. The cells defective for the 5'-3' cytoplasmic RNA degradation show multiple phenotypes (*19*), including an impaired respiratory growth and mitochondrial abnormalities (ref. *66* and unpublished data). The absence of the major cytoplasmic 5'-3' exoribonuclease Xrn1, as well as its cofactor Dcs1, compromises the expression of multiple mitochondrion-related factors, as is the case of the mitochondrial outer membrane porin Por1 (also known as VDAC in metazoans) (*66*).

At the beginning of this study we have rapidly detected the key elements in the pattern of Por1 expression. Importantly, while the cytoplasmic 5'-3' RNA decay activity was crucial for the production of Por1 protein, POR1 mRNA remained insensitive to the absence of Xrn1, in contrast to the prevailing transcriptional repression model (*64*). In parallel, we have identified an antisense RNA to POR1, OCA2 mRNA, readily reactive to the changes in the activity of Xrn1, and

overlapping with POR1 mRNA. These initial observations have guided this study in two major directions.

Firstly, we aimed to understand the molecular mechanism of *POR1* expression, particularly in relation to its convergent genic orientation in respect of *OCA2*. Our objectives debuted with the functional characterization of POR1 and OCA2 mRNAs, seeking to understand at which stage of Por1 expression the absence of 5'-3' cytoplasmic RNA decay intervened. We addressed directly the role of OCA2 transcript in the mechanism of Por1 expression, and we further hypothesized that the apparent post-transcriptional effect of this antisense RNA might be mediated by the interaction between two partially complementary RNAs. As the evidences of the POR1 and OCA2 mRNA-mRNA interaction in the cytoplasm accumulated, we continued to define the molecular details and prerequisites for this interaction. Finally, we drew our attention to the translational outcomes of the interaction between overlapping mRNAs in the cytoplasm.

Secondly, we aimed to use our insights from the detailed study of POR1 expression and address the generalities of the convergent gene expression in *S. cerevisiae*. We have postulated that an important number of convergent and overlapping mRNAs might interact in the cytoplasm, mirroring the situation of POR1 and OCA2. We thus chose to use the reconstituted RNAi approach to identify the genome-wide prevalence of mRNA-mRNA interactions. Our objectives initiated with the validation of the functional RNAi pathway—siRNA production—in the case of POR1 and OCA2 mRNA-mRNA interaction. Further on, we assessed siRNA production within convergent loci by small RNA sequencing from RNAi-competent cells, having functional or mutated (*xrn1*Δ) 5'-3' cytoplasmic RNA decay pathway. We focused our analysis on the role of Xrn1 in surveilling the cytoplasmic mRNA-mRNA interactions, as we hypothesized on the key effect of this factor in controlling the pool of overlapping and potentially interacting RNA.

These two major focuses of our study, at individual gene level and genome-wide, were directed to obtain some broader insights on long-range intermolecular RNA-RNA interactions and its role in gene regulation. Using a model organism naturally devoid of RNAi pathway we aimed to assess how prevalent overlapping mRNA-mRNA interactions are and what functional outcomes it could have. Mechanistically, we sought to evaluate the possibility of an alternative fate of

overlapping coding transcription units, namely the cytoplasmic interaction of convergent mRNA pairs, independently of—or, at least, additionally to—transcriptional repression, currently preconized in the field. As our initial observations favored the post-transcriptional lead in Xrn1-mediated regulation of Por1 expression, we questioned the possibility that mRNA-mRNA interactions might function in remodeling RNA stability or translation in the cytoplasm, as discussed in Part II of this chapter. However, only a part of the references cited therein was available at the time when this study began, and no data on potential post-transcriptional regulation of sense-antisense RNA pairs in fungi was present. Finally, as our study continued to address the pleiotropic effects of a deficient cytoplasmic 5'-3' RNA decay pathway, we hypothesized of a new role of Xrn1, implicated in the post-transcriptional gene regulation.

CHAPTER II | RESULTS

Outline

The results of this work are presented following two major axes of experiments. Firstly, we analyzed in details the *POR1* gene expression, in relation to the presence of its convergent gene *OCA2*, resulting in the production of an overlapping mRNA pair. We then assessed the possibility of a cytoplasmic interaction between these two mRNAs, evaluating the molecular details and prerequisites of this interaction. We have further detected a cleavage of POR1 mRNA, characteristic for the no-go decay pathway, which confirmed the translation-associated cytoplasmic effect of OCA2 transcript on POR1 mRNA. Finally, we addressed the ribosome-stalling potential of interacting mRNAs, as well as the remodeling of mRNA stability by 3' overlapping interactions.

The second part of this chapter is focused on the results we obtained during the analysis of small RNA sequencing data from RNAi-competent *S. cerevisiae* cells. Initially, we validated the presence of siRNA signal corresponding to the overlap region between POR1 and OCA2 mRNAs, and extended this observation genome-wide. We then examined the role of Xrn1 in controlling the global mRNA-mRNA interaction landscape of the cell. The results presented in this part were obtained in collaboration with the group of Antonin Morillon (UMR3244, Institut Curie, Paris).

PART I | The study of *POR1-OCA2* locus expression

1 | Por1 is downregulated in *xrn1*Δ cells

As presented in the Objectives of the study section (Chapter I – Part III), the previous results from our laboratory demonstrated a severe respiratory growth defect of *S. cerevisiae xrn1*Δ cells (*66*). Accordingly, *xrn1*Δ cells cannot grow on non-fermentable carbon sources (such as glycerol, ethanol, lactate, etc.) and show mitochondrial abnormalities. A two-dimensional gel electrophoresis coupled to mass spectroscopy permitted to identify a list of proteins downregulated in *xrn1*Δ cells, in the context of glycerol as the sole carbon source. The majority of the factors identified is important for the respiratory function of the cell and/or is associated to mitochondria. A representative example of this list is the mitochondrial outer membrane voltage-dependent anion channel (VDAC), also known as porin, Por1. In the growth medium containing glycerol, Por1 protein production decreases several-fold in *xrn1*Δ cells, as compared to isogenic wild-type cells (*66*). Interestingly, in the fermentation medium containing glucose (YPD), the defective 5'-3' RNA decay pathway in the cytoplasm also leads to the decreased production of Por1. As shown in Figure 26, western blot analysis of cellular extracts from WT and *xrn1*Δ cells demonstrate a significant decrease of Por1 protein in the absence of Xrn1, as compared to a loading standard Pgk1.

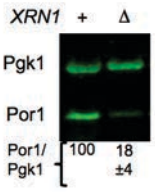

Figure 26. The expression of Por1 protein.

Western blot of Por1 protein in wild-type and *xrn1*Δ strains. Pgk1 was used as a loading control. Quantifications of Por1 relative to Pgk1 are indicated below the blot. All quantifications (here and below) are indicated with standard errors calculated from at least three independent experiments, unless stated otherwise.

2 | POR1 mRNA does not vary in quantity in *xrn1*Δ cells

The *xrn1*Δ mutant cells have been associated with the antisense transcription-mediated transcriptional gene repression (*64*). To verify if the compromised production of Por1 protein in the absence of Xrn1 is due to a decreased level of steady state POR1 mRNA, we analyzed total RNA extracted from isogenic WT and *xrn1*Δ cells. As shown in Figure 27, probing a northern blot with POR1 mRNA specific oligonucleotide revealed no significant difference in POR1 mRNA quantity between the two strains, suggesting that *POR1* locus does not undergo transcriptional repression in *xrn1*Δ cells.

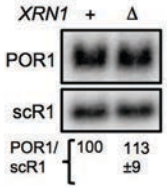

Figure 27. The expression of POR1 mRNA.

Northern blot showing steady-state levels of POR1 mRNA in wild-type and *xrn1*Δ strains. The scR1 RNA served as a loading control. Quantifications of POR1 relative to scR1 are indicated below the blot.

3 | POR1 mRNA is functional in *xrn1*Δ cells

As Xrn1 is the major cytoplasmic exoribonuclease, its mutants accumulate various RNA decay intermediates (*65*). This enzyme is selectively active on decapped, 5'-monophosphorylated mRNA (*19*), and we reasoned that the population of POR1 mRNA might be predominantly decapped in *xrn1*Δ cells, thus inherently defective for cap-dependent translation initiation (see section 2 of Chapter I, Part I). To analyze potential differences in capped to decapped POR1 mRNA ratio between WT and *xrn1*Δ situations, we treated total RNA extracted from these two strains by Xrn1 *in vitro*. As shown in Figure 28, this treatment efficiently eliminates 18S and 25S rRNA species that are 5'-monophosphorylated *in vivo*. However, the Xrn1-resistant fraction of POR1 mRNA, as revealed by northern probing with POR1 specific oligonucleotide, remains comparable in WT and *xrn1*Δ cells, suggesting that similar quantity of capped POR1 mRNA exists in these two contexts.

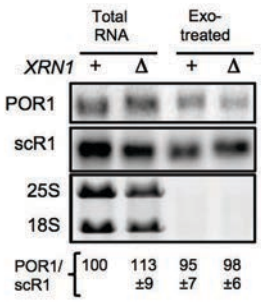

Figure 28. POR1 mRNA is similarly sensitive to 5'-3' exoribonuclease treatment in vitro in WT and *xrn1*Δ cells.

Northern blot showing steady-state levels of POR1 mRNA in wild-type and *xrn1*Δ strains. Total RNA and total RNA treated with Xrn1 *in vitro* (Exo-treated) to deplete 5'-monophosphorylated RNA (i.e., 5'-decapped RNA) are shown. The 25S and 18S rRNAs have 5'-monophosphate group and serve as a positive control for exonucleolytic digestion. The scR1 RNA (a Pol III transcript insensitive to exoribonuclease treatment) served as a loading control. Quantifications of POR1 relative to scR1, before or after 5'-3' exoribonuclease treatment, are indicated below the blot.

We have also considered that *xrn1* mutant cells could stabilize various 3' RNA decay intermediates, as canonically, both 5'-3' and 3'-5' mRNA decay pathways are initiated by 3' poly(A) tail shortening (see section 3 of Chapter I, Part I). It is thus possible that deadenylated albeit still capped POR1 mRNA could predominate in the absence of Xrn1, lowering the translation efficiency of POR1 mRNA. To determine the deadenylation status of POR1 mRNA in WT and *xrn1*Δ conditions, we have applied the widely used technique based on the 3' proximal RNA cleavage by RNase H, via a DNA oligonucleotide guide (*202*). This cleavage (1) generates the homogenous population of 5' extremities of a particular RNA and (2) creates short RNA species differing solely in their 3' adenylation status and thus easily analyzable in high-resolution polyacrylamide gels. However, we could not apply this technique for POR1 mRNA due to a very inefficient initial hybridization step of the DNA oligonucleotide for RNase H cleavage. We suspected that OCA2 mRNA (see below), overlapping POR1 mRNA within the region of oligonucleotide binding, competes for POR1 mRNA binding (data not shown).

We further chose to use an alternative technique to determine the adenylation status of POR1 mRNA. We applied the PCR-based ePAT (extension poly(A) test) method (see Figure 29A) (*203*) and found that inactivating Xrn1 does not significantly

alter the poly(A) tail length of POR1 mRNA, as shown in Figure 29B. We thus conclude that POR1 mRNA is translation-competent in *xrn1*Δ cells, as it contains the canonical prerequisites *in cis*—the 5' cap and the 3' poly(A) tail—at comparable levels to WT cells.

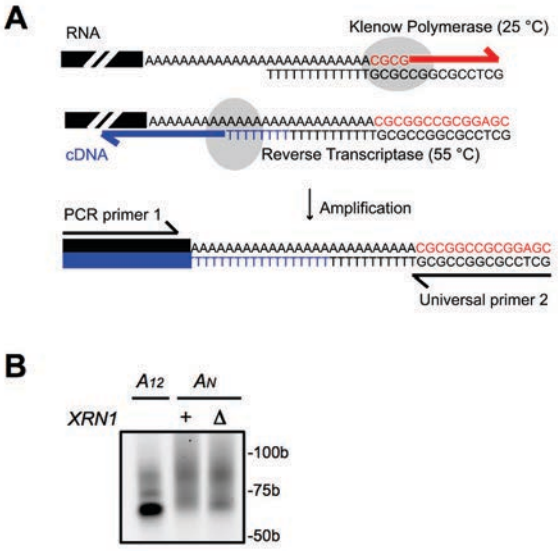

Figure 29. POR1 mRNA is similarly polyadenylated in WT and *xrn1*Δ cells.

(A) The schematic of the extension poly(A) test (ePAT) (*203*). Poly(A) containing RNAs are 3' extended by Klenow DNA polymerase, using an anchored oligo(dT) primer. After the inactivation of Klenow polymerase, extended RNAs are reverse transcribed at higher temperature, limiting the hybridization of the anchored oligo(dT) primer to the 3' extremity of the poly(A) tail. Finally, the cDNA is amplified with a universal and gene-specific primer by PCR.

(B) The analysis of the length of poly(A) tail (A_N) of POR1 mRNA by ePAT from WT and *xrn1*Δ cells. On the left, the A_{12} containing POR1 amplicon is shown. On the right, DNA size marker is shown.

4 | OCA2 mRNA, antisense to POR1, accumulates in *xrn1*Δ cells

As described previously, POR1 mRNA is expressed at similar steady-state levels in wild-type and *xrn1* mutant cells. Seeking to understand the mechanism of influence of a defective cytoplasmic 5'-3' RNA decay pathway to the expression of Por1 protein, we considered that *xrn1*Δ cells could accumulate an RNA, potentially controlling *POR1* expression post-transcriptionally. Similarly to the class of Xrn1-sensitive unstable transcripts (XUTs), ncRNA species detected in *xrn1*Δ cells that act in transcriptional gene repression (*64*), we hypothesized the existence of a ncRNA,

capable to repress Por1 translation. However, extensive probing in the vicinity of *POR1* locus yielded no detectable ncRNA (data not shown). In contrast, we have identified a coding RNA, OCA2 (oxidant-induced cell-cycle arrest), partially overlapping POR1 mRNA due to the convergent orientation of *POR1* and *OCA2* genes (see Figure 30A). Interestingly, OCA2 mRNA is significantly stabilized in *xrn1*Δ cells, as revealed by a northern blot shown in Figure 30B. Moreover, this transcript is highly sensitive to the treatment by Xrn1 *in vitro*, suggesting that decapped OCA2 mRNA—a 5' RNA decay intermediate—accumulates in *xrn1*Δ cells.

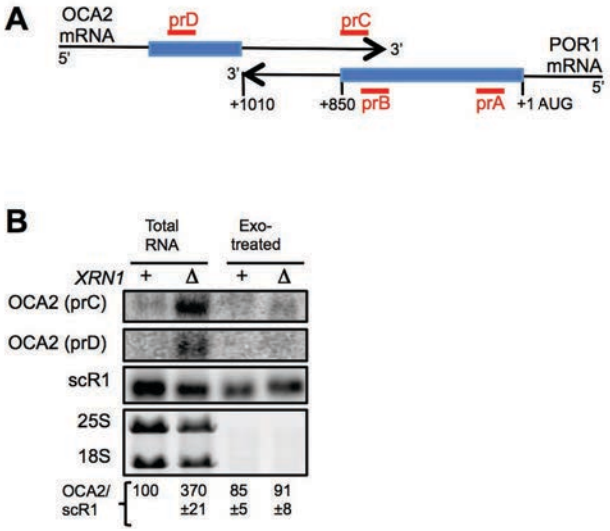

Figure 30. OCA2 mRNA accumulates in *xrn1*Δ cells.

(A) Schematic of POR1 and OCA2 mRNAs showing the extent of overlap in their 3' regions. Filled boxes indicate open reading frames, and black arrows represent 3' UTRs. Probes prA, prB, prC, and prD used in northern blot analysis are indicated in red.

(B) Northern blot showing steady-state levels of OCA2 mRNA in wild-type and *xrn1*Δ strains. Total RNA and total RNA treated with Xrn1 *in vitro* (Exo-treated) to deplete 5'-monophosphorylated RNA (i.e., 5'-decapped RNA) are shown. The 25S and 18S rRNAs have 5'-monophosphate group and serve as a positive control for exonucleolytic digestion. The scR1 RNA (a Pol III transcript insensitive to exoribonuclease treatment) served as a loading control. Quantifications of OCA2 (using probe prD) relative to scR1, before or after 5'-3' exoribonuclease treatment, are indicated below the blot.

Oca2 is a putative tyrosine phosphatase, implicated in a network of genetic and physical interactions with other proteins of Oca family, namely Oca1, Oca3 (also known as Siw14), Oca4 and Oca6 (*204*). The biological function of Oca2 (as well as

other Oca proteins) is unknown although it has been associated with oxidative stress response (*205*).

5 | OCA2 mRNA downregulates Por1 production *in trans*

The stabilization of an antisense transcript to POR1 mRNA in the absence of Xrn1 incited us to examine the role of OCA2 mRNA on the expression of Por1 protein. For this, we have constructed a plasmid (pLB001) expressing OCA2 under a strong (tetracycline repressible) $tetO_7$ promoter (*206*). The presence of OCA2 mRNA expressed *in trans* contributed to an additional downregulation of Por1 protein, in both WT and *xrn1*Δ cells, as compared to an empty vector control (Figure 31B). Importantly, the ectopic expression of OCA2 mRNA did not modify the steady-state levels of POR1 mRNA, as demonstrated by a northern blot analysis (Figure 31A).

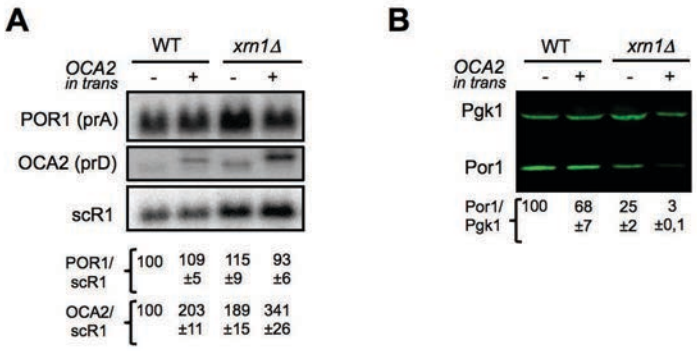

Figure 31. OCA2 mRNA downregulates Por1 production *in trans*.

(A) Northern blot analysis of POR1 mRNA steady-state levels upon expression (+) or not (-) of OCA2 mRNA *in trans*. Differences in OCA2 mRNA size are due to OCA2 expression from a plasmid-borne $tetO_7$ promoter that extends the 5' UTR sequence by 132 nucleotides. The scR1 RNA served as a loading control. Quantifications of RNAs relative to scR1 are indicated below the blot.

(B) Western blot analysis performed as in (A) upon expression (+) or not (-) of OCA2 mRNA *in trans*. Pgk1 was used as a loading control. Quantifications of Por1 relative to Pgk1 are indicated below the blot.

6 | 3' OCA2 fragment is sufficient to downregulate Por1 production *in trans*

As the overexpression of OCA2 mRNA *in trans* contributes to downregulating Por1 protein, we further wondered if this effect could be unambiguously attributed to the partial complementarity between POR1 and OCA2 mRNAs, rather than an indirect contribution of Oca2 protein. For this purpose, we constructed a plasmid (pLB002, derivative of pLB001) expressing the 3' portion of OCA2 mRNA—complementary to the 3' of POR1 mRNA—under the $tetO_7$ promoter (see Figure 32A for schematic). As shown in Figure 32B, western blot analysis of protein extracts from WT and *xrn1*Δ cells revealed that non-coding 3' OCA2 fragment *in trans* also decreases the production of Por1 protein, as compared to an empty vector control, similarly to the full length OCA2. This observation confirmed that a convergent and overlapping transcript to POR1 mRNA, OCA2, controlled the production of Por1 protein and, furthermore, this effect was valid *in trans*, depending on the complementarity between two RNAs rather than the coding capacity of OCA2 mRNA.

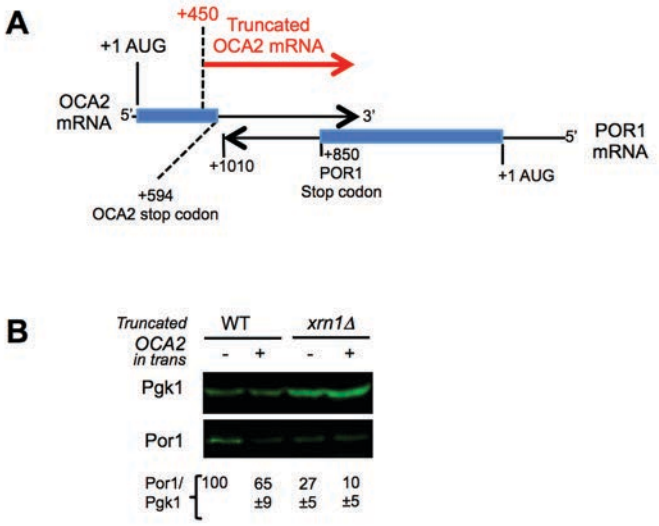

Figure 32. Truncated OCA2 transcript downregulates Por1 production *in trans*.

(A) Schematic representation of POR1 and OCA2 mRNAs. A red arrow defines the sequence of the truncated version of OCA2 mRNA expressed *in trans*. Filled boxes indicate open reading frames and black arrows represent 3' untranslated regions of OCA2 and POR1 genes.

(B) Western blot of Por1 protein in wild-type and *xrn1*Δ strains in the presence (+) or absence (-) of a truncated OCA2 mRNA expressed *in trans*. Pgk1 was used as a loading control. Quantifications of Por1 relative to Pgk1 are indicated below the blot.

7 | POR1 mRNA associates with polysomes in *xrn1*Δ cells

Accordingly to the observation discussed above, we reasoned that OCA2 mRNA, accumulating in the absence of Xrn1, could interfere with the translation of POR1 mRNA, as a 3' complementary RNA (coding as well as non-coding) *in trans* affected the production of Por1 protein. To determine the translational status of POR1 mRNA in wild-type and *xrn1* mutant cells, we fractionated cellular extracts from these two strains on sucrose gradient by ultracentrifugation, collected the ribosome-free, monosomal and polysomal fractions, and analyzed its content by northern blots. As shown in Figure 33, we found POR1 mRNA associated with polyribosomes in WT cells, as expected for a long and well-translated messenger, similarly to a control mRNA PGK1. However, we have also detected POR1 mRNA associated with polyribosomes in *xrn1*Δ cells, in contrast to a decreased protein output of this condition. Similarly to WT cells, PGK1 mRNA was found in the polysomal fraction in the absence of Xrn1. This observation suggests that POR1 mRNA is similarly engaged to the translation process in both WT and *xrn1* mutant cells. Interestingly, OCA2 mRNA was detected in the polysomal fraction in *xrn1*Δ cells, independently of its predominantly 5'-monophosphate status (see section 4 of this chapter).

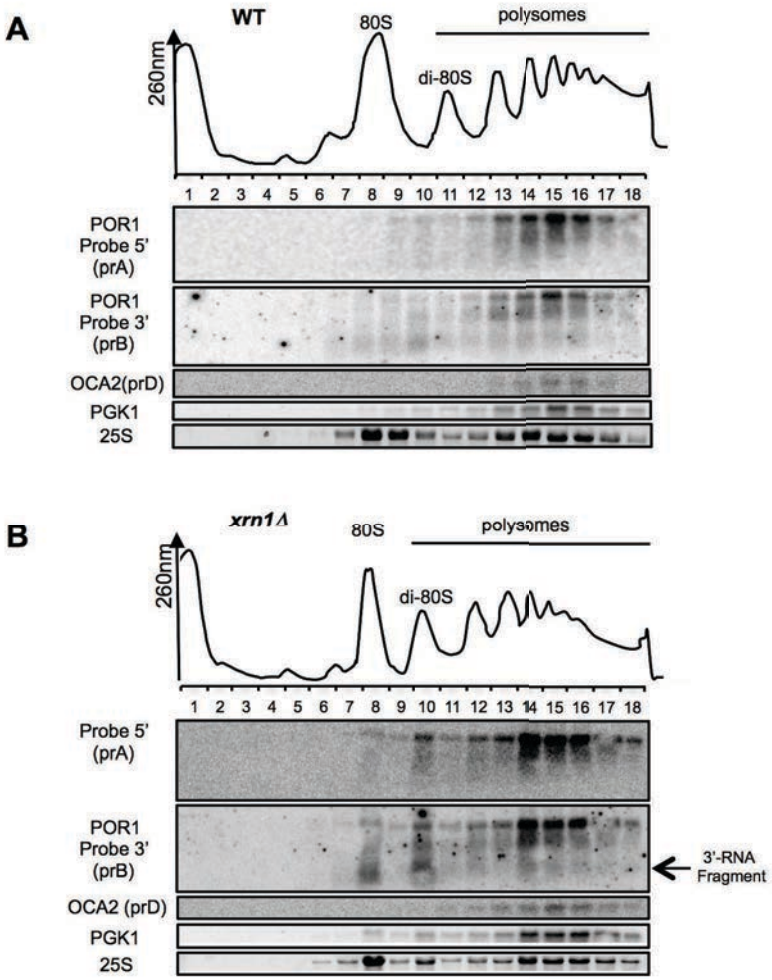

Figure 33. POR1 and OCA2 mRNAs associate with polyribosomes.

(A) Northern blot analysis was performed on RNA isolated from sucrose-gradient fractions of wild-type cell lysates. Monosomes (80S), disomes (di-80S) and polyribosomes (polysomes) are indicated above the profiles. Fraction numbers are indicated below the profiles and correspond to lanes of the panels shown below.

(B) The same analysis as in (A), performed on RNA isolated from *xrn1*Δ cell lysates. A 3' RNA fragment of POR1 is detected with probe prB in monosome and disome fractions (lanes 8 and 10).

8 | OCA2 mRNA induces 3' proximal cleavage of POR1 mRNA

During the analysis of polysome-bound POR1 mRNA, we also detected short RNA species, in the fraction of 80S and two ribosomes, while probing with the 3' end specific oligonucleotide for POR1 mRNA. This fragment was only detected in *xrn1*Δ cells (Figure 33). Seeking to confirm that OCA2 mRNA, stabilized in the absence of Xrn1, is connected to the production of this short RNA species, we overexpressed OCA2 mRNA *in trans* in WT and *xrn1*Δ cells and sequentially probed total RNA with

different POR1-specific oligonucleotide probes (Figure 34). The 5' proximal probe revealed only a single band of full-length POR1 mRNA; however, probing with different 3' proximal probes, the short fragment gradually appeared selectively in *xrn1*Δ cells upon OCA2 overexpression. A low signal of short RNA species was detected when probed 200 nucleotides upstream of POR1 stop codon, and increased approaching 3' end of POR1 mRNA. As this fragment, mapping to the 3' proximal part of POR1 mRNA, is solely detectable in the absence of Xrn1, it is unlikely to correspond to a exonucleolytic decay intermediate; we thus postulated that POR1 mRNA-overlapping OCA2 transcript induces the endonucleolytic cleavage of POR1 mRNA.

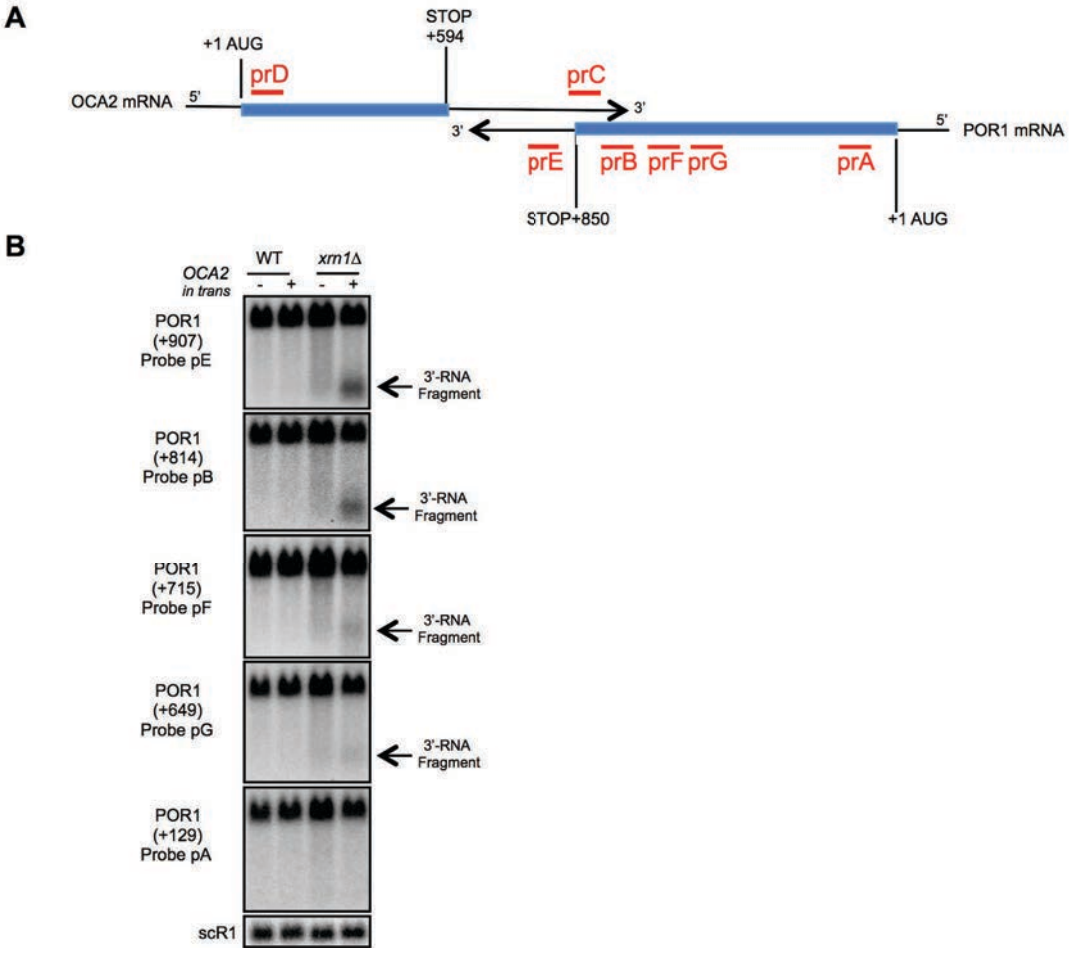

Figure 34. Mapping of the 3' POR1 fragment.

(A) Schematic of POR1 and OCA2 mRNAs. Transcription of POR1 is convergent to OCA2 transcription and transcripts overlap in their 3' regions. Filled boxes indicate open reading frames and black arrows represent 3' untranslated regions. Probes prA, prB, prC, prD, prE, prG and prF used in northern blot analysis are indicated in red.

(B) Detection of the 3' RNA fragment of POR1 mRNA using different probes, when expression (+) of OCA2 mRNA is produced in trans or not (-). Coordinates of the first 5' nucleotides of POR1 mRNA detected by each probe are indicated in brackets and are relative to the AUG of POR1 (A as the +1). The scR1 RNA served as a loading control.

9 | POR1 mRNA is targeted by no-go decay

OCA2 mRNA extends into the 3' portion of *POR1* ORF, as revealed by genome-wide transcript isoform profiling in *S. cerevisiae* (*9*) and our northern blot analysis (see Figure 30). The detection of a short RNA fragment in *xrn1*Δ cells, corresponding to the 3' end of POR1 mRNA, incited us to compare this situation to the no-go decay (NGD) pathway. In fact, the study that reported this pathway for the first time, analyzed a synthetic reporter mRNA, termed PGK1-SL, with a strong stem-loop structure inserted within the coding sequence of PGK1 mRNA (*207*). Upon translation of this transcript, ribosomes stall upstream of the stem-loop and trigger the endonucleolytic cleavage of the aberrant mRNA. As discussed in the Chapter I, this cleavage results in two decay intermediates—namely 5' and 3' NGD fragments—that are rapidly degraded in the cytoplasm by the exosome and Xrn1, respectively. The non-canonical ribosome release factors Dom34 and Hbs1 stimulate the endonucleolytic cleavage, although are not required for it to take place (*208*). We thus speculated that in case of an interaction between POR1 and OCA2 mRNAs, RNA duplex might form extending into the coding region of *POR1* and, mirroring the intramolecular secondary structure of PGK1-SL, induce NGD cleavage (see Figure 35A for schematic).

To validate this hypothesis, we extracted total RNA from isogenic WT, *xrn1*Δ, *dom34*Δ and *xrn1*Δ *dom34*Δ cells, and analyzed the RNAs corresponding to the 3' end of POR1 by northern blot. As shown in Figure 35B, we detected short RNA species in the *xrn1* mutant while probing with the POR1 3' end-specific oligonucleotide, similarly to the result observed during the northern analysis of polysome fractions (see section 7 of this chapter). In accordance with previous observations in the literature, the deletion of *DOM34* decreased (more than two-fold) the efficiency of POR1 RNA 3' fragment production in *xrn1*Δ cells.

In parallel to the analysis of POR1 mRNA in the presence or absence of Dom34, we also quantified Por1 protein in the above-mentioned contexts by western

blot. As shown in Figure 35C, Por1 protein remained at similar levels in wild-type and *dom34*Δ cells, and it was equally downregulated in both *xrn1*Δ and *xrn1*Δ *dom34*Δ mutants, suggesting that the deficient rescue of stalled ribosomes does not affect the aberrant protein output, as reported earlier (*208*).

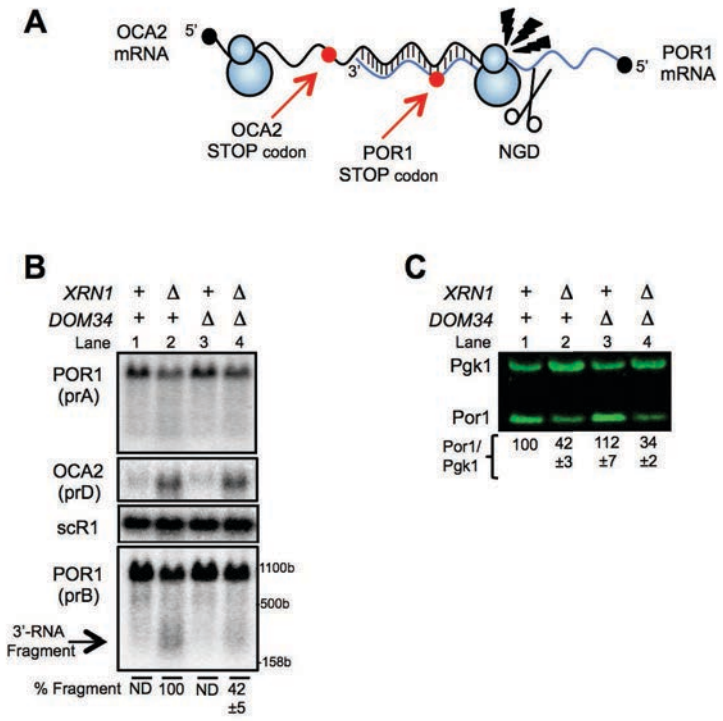

Figure 35. POR1 mRNA is targeted in by NGD in *xrn1*Δ cells.

(A) Model of NGD of the POR1 mRNA in the presence of the OCA2 transcript. The translational stop codon of POR1 mRNA is indicated by a red circle. The ribosome stall due to the presence of OCA2-POR1 mRNA duplex is symbolized by lightning flashes and the subsequent endonucleolytic cleavage is indicated by scissors. Bars indicate the potential interactions between the two RNA sequences.

(B) Detection of the 3' RNA fragment of POR1 in wild-type, *xrn1*Δ, or/and *dom34*Δ strains by northern blot analysis. The presence of the 3' RNA fragment of POR1 is indicated by an arrow. The % fragment is the relative ratio of fragment to full-length POR1 standardized to a percentage of 100% for *xrn1* mutant. The scR1 RNA served as a loading control. ND, not determined.

(C) Western blot of Por1 protein in wild-type, *xrn1*Δ, or/and *dom34*Δ strains. Pgk1 was used as a loading control. Quantifications of Por1 relative to Pgk1 are indicated below the blot.

The short fragment production corresponding to the 3' end of POR1 mRNA was also readily detected in *xrn1*Δ cells upon OCA2 mRNA overexpression *in trans* (Figure 36). Similarly to the observations described above, combining *dom34* and *xrn1*

mutations—but not *ski2* mutation, inactivating cytoplasmic 3' RNA decay pathway—decreased (approximately three-fold) the production of the POR1 3' end-corresponding fragment. In addition, the deletion of *HBS1* in the context of mutated *xrn1* also decreased the NGD signal of POR1 mRNA upon OCA2 mRNA overexpression, mirroring the absence of Dom34 (Figure 37).

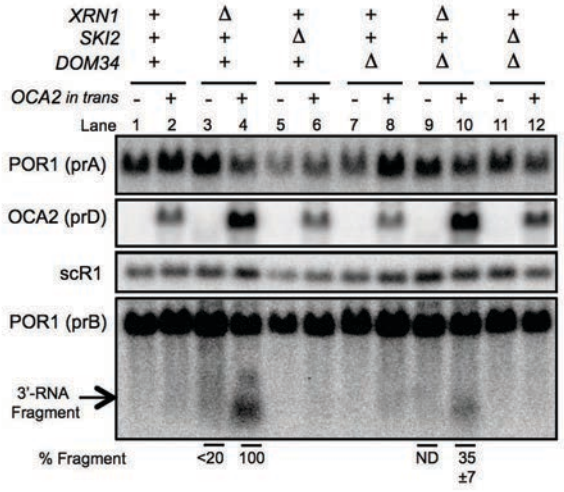

Figure 36. The production of the POR1 3' NGD fragment is dependent on Dom34.

Northern blot analysis in *xrn1*Δ, *ski2*Δ, and *dom34*Δ genetic contexts. Analysis of the production of the POR1 3' RNA fragment upon expression (+) or not (-) of OCA2 mRNA *in trans*. The presence of the 3' RNA fragment of POR1 is indicated by an arrow. The % fragment is the relative ratio of fragment to full-length POR1 standardized to a percentage of 100% for *xrn1* mutant. The scR1 RNA served as a loading control. ND, not determined.

At this point, we have concluded that POR1 3' UTR- and 3' proximal portion of the ORF-overlapping mRNA OCA2 induces the NGD cleavage of POR1 mRNA, detectable in the absence of Xrn1. This observation was important for our study, as it showed that OCA2 mRNA is implicated in a cytoplasmic and translation-associated event, leading to a compromised production of Por1 protein.

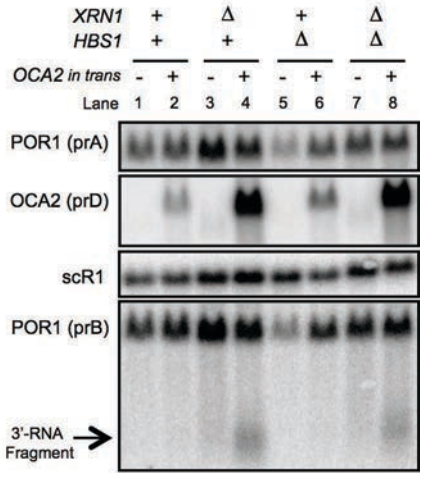

Figure 37. The production of the POR1 3' NGD fragment is dependent on Hbs1.

Northern blot analysis in *xrn1*Δ and *hbs1*Δ genetic contexts. Analysis of the production of the POR1 3' RNA fragment upon expression (+) or not (-) of OCA2 mRNA *in trans*. The presence of the 3' RNA fragment of POR1 is indicated by an arrow. The scR1 RNA served as a loading control.

10 | Physiological inhibition of Xrn1 is sufficient to detect the no-go decay cleavage of POR1 mRNA

The targeting of POR1 mRNA for NGD pathway is intrinsically connected to a dysfunctional 5'-3' cytoplasmic RNA decay pathway: firstly, OCA2 mRNA accumulates in the absence of Xrn1 and, secondly, 3' NGD intermediates can only be detected in Xrn1-deficient cells. However, we wanted to address the fate of interacting POR1 and OCA2 mRNAs in the presence of Xrn1. For this purpose, we chose to inhibit the 5'-3' RNA decay pathway in wild-type cells by a temporary presence of lithium ions in the growth medium (*209*). In the absence of sulfur-containing amino acids in the medium, methionine and cysteine biosynthesis is induced, yielding a by-product, adenosine 5',3' bisphosphate (pAp). Met22, which is selectively inhibited by lithium ions, further metabolizes the pAp; however, upon Met22 inhibition, pAp accumulates and further inhibits the activity of Xrn1 (*78, 79*).

After one-hour treatment with 100 mM LiCl of WT yeast cells, 3' NGD fragment of POR1 mRNA could be detected upon OCA2 mRNA overexpression *in trans*, as revealed by northern blot analysis (Figure 38). As expected, in *xrn1*Δ cells, OCA2 overexpression *in trans* triggered the NGD of POR1 mRNA in both presence and absence of lithium ions. This result suggests that the decreased activity of the cytoplasmic 5'-3' RNA decay pathway, but not Xrn1 itself, is responsible for the

accumulation of NGD-inducing antisense transcript. Additionally, it also shows that post-transcriptional events can be triggered in WT cells, upon physiological downregulation of the cytoplasmic 5'-3' RNA decay pathway.

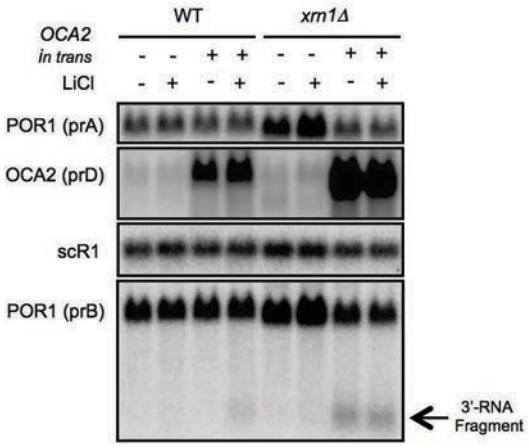

Figure 38. The inhibition of Xrn1 by lithium ions permits the detection of NGD cleavage of POR1 mRNA.

Northern blot analysis of mRNAs in the presence (+) or in the absence (-) of lithium in culture media, and with the expression (+) of OCA2 mRNA *in trans* or not (-). A 3' RNA fragment of POR1 is detected (indicated by an arrow), similar to what is observed for POR1 mRNAs in the *xrn1* mutant. The scR1 RNA served as a loading control.

11 | Sequence complementarity between POR1 and OCA2 mRNAs is essential for no-go decay cleavage

We further aimed to determine the prerequisites of interaction between two convergent and overlapping mRNAs and we chose to use NGD cleavage as the signal of a functional output from such interaction. We wondered if a differential complementarity status between the 3' UTR and the 3' proximal ORF region of two overlapping mRNAs lead to a distinct NGD signal. For this purpose, we firstly constructed a yeast strain that harbored isolated *POR1* and *OCA2* loci, abrogating thus the inherent complementarity between the respective mRNAs. Technically, we inserted a drug resistance cassette between *POR1* and *OCA2* genes, at the same time modifying the 3' UTR of POR1 mRNA but leaving *POR1* ORF intact. We then constructed two plasmids (containing the same $tetO_7$ promoter as pLB001, overexpressing OCA2; designated RNA1 in Figure 39A), permitting to overexpress differentially overlapping mRNAs *in trans*. One plasmid yielded an RNA (designated

RNA2) complementary to the mutated 3' UTR of POR1 mRNA but not the 3' proximal part of POR1 ORF; the second plasmid yielded an RNA (designated RNA3) complementary to both mutated 3' UTR and 3' proximal ORF region of POR1 mRNA (and thus complementary to only 3' proximal part of wild-type POR1 mRNA; see Figure 39A for a schematic). We then combined WT and *xrn1*Δ strains carrying wild-type or mutant 3' UTR alleles of *POR1* with the three RNAs *in trans* (or with an empty vector control) and assessed the appearance of NGD cleavage in these contexts by northern blot. As shown in Figure 39B, RNA1 (OCA2 mRNA) overexpression in wild-type POR1 allele-carrying *xrn1*Δ cells lead to a detectable 3' NGD cleavage of POR1 mRNA, as expected. Importantly, RNA3 overexpression in mutant 3' UTR POR1 allele-harboring *xrn1*Δ cells also caused NGD of POR1 mRNA. Among the situations tested, only these two produced a detectable 3' NGD intermediate, suggesting that the full complementarity—within the 3' proximal part of ORF and the 3' UTR—of interacting RNAs is required for triggering NGD.

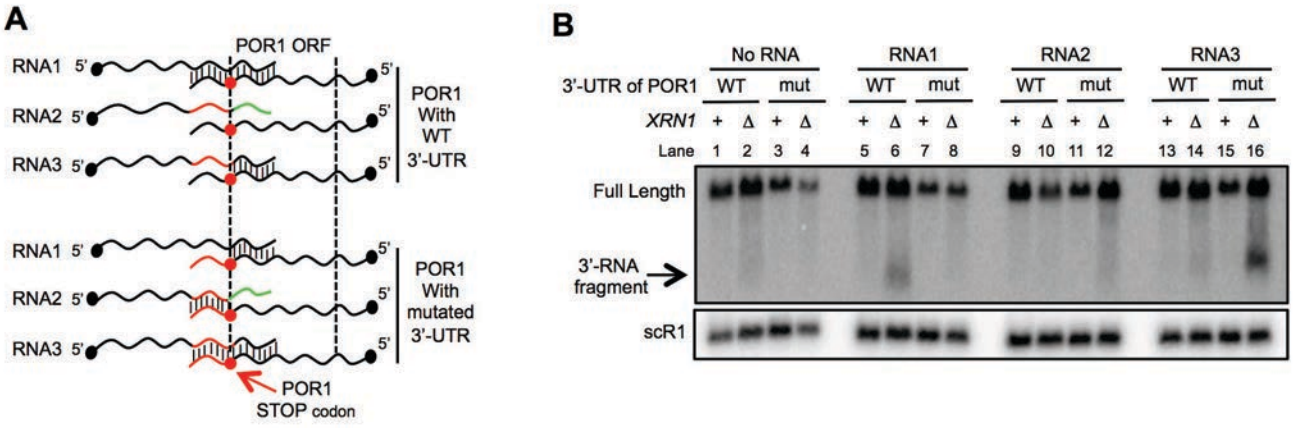

Figure 39. The sequence complementarity is important for mRNA-mRNA interactions.

(A) Schematic of RNA1, RNA2, and RNA3 expressed *in trans* and having putative full, partial, or no complementarity with the wild-type or mutated 3' UTR region of POR1 mRNAs. Bars indicate the potential interactions between sequences of the two RNAs. The POR1 ORF is delimited by dashed lines. The translational stop codon is indicated by a red circle.

(B) Northern blot analysis of POR1 mRNA with wild-type and mutated 3' UTR in the presence of RNA1, RNA2, or RNA3 expressed *in trans*. Probe prB is designed for the detection of the 3' RNA fragment that can be produced by both wild-type and modified POR1 mRNAs. The production of a 3' RNA fragment of POR1, indicated by an arrow, is analyzed in the presence or absence (no RNA) of RNA1, RNA2, or RNA3 expressed *in trans*. WT and mut specify the nature of

the 3' UTR of POR1 mRNA analyzed, wild-type 3' UTR, and mutated 3' UTR, respectively. The scR1 RNA served as a loading control.

We also examined the expression of differentially overlapping transcripts (RNA1, RNA2 and RNA3) used in the above-discussed experiment. Similarly to the results revealed in Figure 36, RNA1 and RNA2 were highly expressed in the four contexts tested, as expected for a transcript under the $tetO_7$ promoter (data not shown). Surprisingly, RNA3 was present in all conditions tested but the one where its expression triggered NGD of POR1 3' UTR mutant mRNA (Figure 40B, lane 4). We ruled out the possibility that RNA3 was not expressed in *xrn1Δ POR1* 3' UTR mutant cells, as the 3' NGD fragment was clearly detected in this context, in contrast to the empty vector control (see Figure 39B, lanes 4 and 16). We thus suspected that RNA3 construct was expressed at a sufficient level to trigger NGD of the complementary mRNA but at the same time was highly destabilized.

It is possible that POR1 3' UTR mutant mRNA extends into the 3' proximal portion of RNA3 ORF, similarly to RNA3 extending into the ORF of POR1. As described in Chapter IV (Materials and methods), *POR1* 3' UTR mutant harbors the 3' UTR sequence of *RPS12*; RNA3 was constructed modifying *RAD17* gene, convergent to *RPS12*, replacing the *RPS12* ORF-overlapping sequence within the 3' UTR of *RAD17* with the 3' proximal region of *POR1*. In support of our hypothesis, RPS12 mRNA was detected to extend into the ORF of *RAD17* in genome-wide mRNA isoform profiling experiments (*9*). We thus speculated that POR1 3' UTR mutant mRNA and RNA3 mutually trigger NGD, leading to an unexpected destabilization of RNA3. To test this possibility we modified RNA3 by inserting an upstream stop codon within its ORF (designated RNA4 and produced from pLB124 plasmid), resulting in a truncated but overlap-free coding sequence (see Figure 40A for schematic). We then examined the fate of POR1 3' UTR mutant mRNA, RNA3 and RNA4 in WT and *xrn1Δ* contexts by northern blot. As shown in Figure 40C, protecting RNA4 ORF from overlapping POR1 3' UTR mutant mRNA resulted in its stabilization, as expected. However, in contrast to RNA3, RNA4 *in trans* failed to trigger NGD of POR1 3' UTR mutant mRNA (Figure 40C, lanes 3 and 4), suggesting that mRNA-mRNA interaction-proximal translation is mutually necessary for NGD induction.

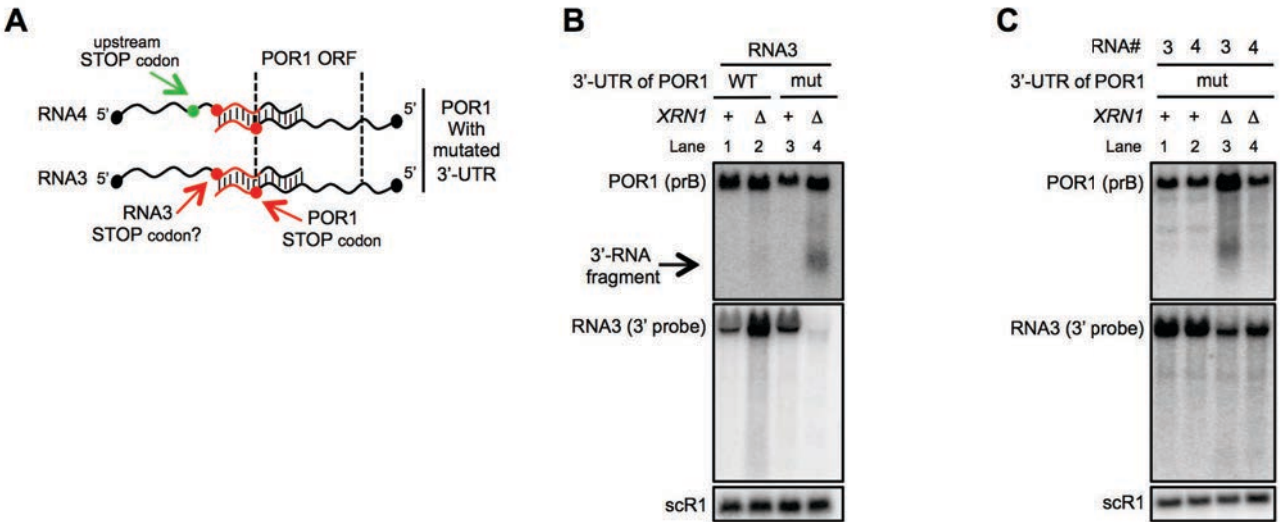

Figure 40. mRNA-mRNA overlap-proximal translation is mutually necessary for NGD.

(A) Schematic of RNA3 and RNA4 expressed *in trans* and having putative full complementarity with the mutated 3' UTR region of POR1 mRNA. RNA4 has a stop codon inserted upstream of the complementarity region with POR1 3' UTR mutant (indicated by a green circle). Bars indicate the potential interactions between sequences of the two RNAs. The POR1 ORF is delimited by dashed lines. The endogenous translational stop codon is indicated by a red circle.

(B) Northern blot analysis of POR1 mRNA with wild-type and mutated 3' UTR in the presence of RNA3 expressed *in trans*. RNA3 was detected with a probe specific for its 3' extremity.

(C) Northern blot analysis of POR1 3' UTR mutant mRNA in the presence of RNA3 or RNA4 expressed *in trans*. RNA3 was detected with a probe specific for its 3' extremity. RNA3 and RNA4 were detected with a probe specific for its 3' extremity. WT and mut specify the nature of the 3' UTR of POR1 mRNA analyzed, wild-type 3' UTR, and mutated 3' UTR, respectively. The scR1 RNA served as a loading control.

12 | Por1 translation termination is compromised by overlapping OCA2 mRNA

As described above, an overlap between two convergent mRNAs, extending from the 3' UTR to the 3' proximal ORF region, triggers a detectable NGD cleavage, indicating of a preceding ribosomal stall. We thus reasoned that elongating ribosomes get stalled (or, at least, significantly slowed down) at the region of interaction between two RNAs and cannot efficiently reach the termination codon. To test this hypothesis we have constructed a reporter plasmid (pLB102) containing

the entire sequence of POR1-OCA2 locus (including the endogenous promoters) and replaced the 5' proximal part of POR1 ORF in the plasmid by c-myc-tagged GFP sequence; this fusion construct contained the intact transcription termination region of OCA2 mRNA. Mirroring the situation of endogenous Por1, c-myc-tagged GFP protein was downregulated in the absence of Xrn1, as revealed by a western blot (Figure 41B, lanes 1 and 3). Further on, we modified this construct (pLB104) by inserting a stop codon at the end of GFP sequence (and thus upstream of the 3' proximal region of POR1, overlapping with OCA2 mRNA), creating a pair of overlapping mRNAs containing the respective translation stop codons outside of the overlap region (see Figure 41A for a schematic). In contrast to the former construct, the latter yielded a comparable amount of c-myc-tagged GFP protein in both WT and *xrn1*Δ cells (Figure 41B, lanes 2 and 4), suggesting that creating a translation termination event upstream of the interaction region between two overlapping RNAs anticipates the ribosome stall and suppresses the subsequent translational defect.

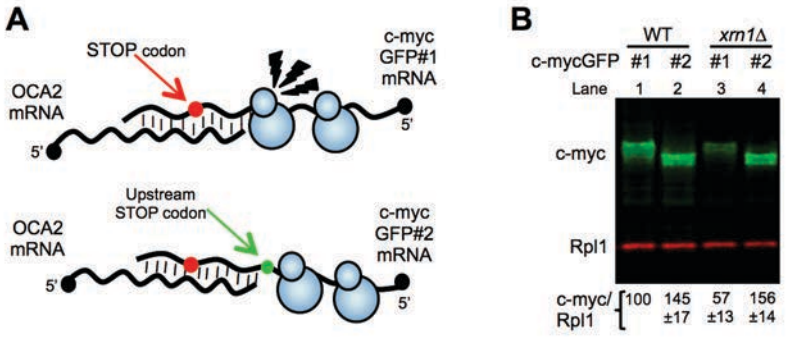

Figure 41. Por1 translation termination is compromised by OCA2 mRNA.

(A) Schematic of c-mycGFP no. 1 mRNA reporter expressed from a c-myc-tagged GFP integrated in the POR1 gene. The overlap between the OCA2 and POR1 mRNAs is conserved in these constructs, but c-mycGFP no. 2 mRNA differs from c-mycGFP no. 1 mRNA only by an additional stop codon (indicated by a green circle) inserted upstream of the RNA duplex containing the normal stop codon (red circle). Lightning flashes represent the clash between elongating ribosomes and the RNA duplex formed on c-mycGFP no. 1.

(B) Western blot analysis of c-mycGFP no. 1 (#1) and c-mycGFP no. 2 (#2) tagged proteins in wild-type and *xrn1*Δ strains. Rpl1 was used as a loading control, and quantifications of c-myc relative to Rpl1 are indicated below the blot.

13 | Translational stall at overlapping RNA region triggers no-go decay

The results described in the previous section revealed that 3' proximal ORF region-overlapping interaction between two convergent mRNAs lead to ribosome stalling during translation and associated NGD cleavage. We thus tested if the reporters used in the section 12 of this chapter—identical for the overlap status between convergent mRNAs but differing in the translation termination position in respect of this overlap—triggered NGD. For this purpose, we constructed isogenic *por1*Δ and *por1*Δ *xrn1*Δ strains (to ensure the detection of plasmid-borne mRNAs only), transformed it with the two c-myc-tagged GFP reporter constructs and ectopically OCA2 mRNA overexpressing plasmids (pLB001), and probed for the 3' NGD fragment by northern blotting. As shown in Figure 42, OCA2 mRNA *in trans* triggered the NGD cleavage of ribosome stall-inducing c-myc GFP#1 reporter but not of the upstream termination-inducing c-myc GFP#2 reporter, detectable in *xrn1*Δ cells. This result underlines that overlapping mRNA-induced NGD cleavage is translation dependent and relies on the elongating ribosome stalling.

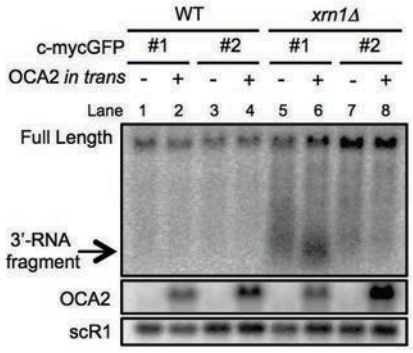

Figure 42. Translational stalling is essential for overlapping RNA-induced NGD cleavage.

Northern blot analysis of c-mycGFP no. 1 (#1) and c-mycGFP no. 2 mRNAs (#2) upon expression (+) or not (-) of OCA2 mRNA *in trans* in the wild-type or in the *xrn1* mutant. The probe prB is designed for the detection of a 3' RNA fragment (indicated by an arrow) produced by either mRNA no. 1 or no. 2. The scR1 RNA served as a loading control.

14 | Por1 is not targeted by RQC in *xrn1*Δ cells

As discussed above, the detection of OCA2 mRNA-induced POR1 mRNA cleavage, detectable in *xrn1* mutant cells, revealed the implication of the NGD pathway in the interacting RNA-related events. However, as visible in Figure 39B, even when the 3' NGD fragment of POR1 mRNA is detected, the full-length transcript remains at comparable levels in both WT and *xrn1*Δ cells, suggesting that targeting POR1 mRNA for NGD is not sufficient to decrease its steady-state level. In contrast, the steady-state quantity of Por1 protein differs significantly between these two contexts, prompting us to consider downstream events, following RNA interaction-induced ribosome stalling. One possibility is that interacting POR1 and OCA2 mRNAs limit the termination of translating ribosomes, targeting POR1 mRNA for predominantly "one-round" translation and thus accumulating slowly translating ribosomes (which is in agreement with POR1 mRNA presence in the polysomal fraction of *xrn1*Δ cells). Alternatively, aberrant mRNA-associated nascent peptides can be targeted for proteasomal degradation, as it is the case for various NGD and NSD substrates (*210, 211*).

Ltn1 is the ubiquitin ligase component of the RQC complex, responsible for nascent peptide polyubiquitination upon ribosome stalling (*212*). We thus chose to assess the role of Ltn1 for the downregulation of Por1 protein in the context of defective 5'-3' cytoplasmic RNA decay pathway. For this, we have constructed isogenic *ltn1*Δ and *ltn1*Δ *xrn1*Δ strains and quantified Por1 protein in these contexts by western blotting. As shown in Figure 43A, we did not detect any significant difference in Por1 protein quantity comparing *xrn1*Δ and *ltn1*Δ *xrn1*Δ cell extracts; interestingly, Por1 protein was also similarly expressed in WT and *ltn1*Δ cells. We have also examined the role of Ltn1 for the production of c-myc-tagged GFP reporter used in the sections 12-13 of this chapter. As the yeast anti-porin antibody used throughout this study appeared to target a C-terminal fragment of Por1 (data not shown), we hypothesized that translationally stalled nascent Por1 chains might be inefficiently detected by anti-Por1 antibody and thus chose to use the N-terminally tagged reporter protein. However, similarly to the endogenous Por1 protein, c-myc-tagged GFP reporter remained comparably downregulated in both *xrn1*Δ and *ltn1*Δ

*xrn1*Δ cells (Figure 43B), suggesting that the ubiquitination by Ltn1 is not the major pathway of Por1 protein decay, in both wild-type and *xrn1*Δ cells.

Another ubiquitin ligase, Hel2, is known to act upstream of the stalled ribosome dissociation and RQC recruitment (*83*). This factor does not ubiquitinate the nascent polypeptide itself (*84*), although it facilitates the early ribosome stalling during NGD, and its deletion causes the passage through rare codons (*210*). We thus speculated that Hel2 might participate in the OCA2 mRNA-induced translational stall on POR1 mRNA and chose to assess the effect of this factor on Por1 protein expression. For this purpose we constructed isogenic *hel2*Δ and *hel2*Δ *xrn1*Δ strains and quantified Por1 protein in these contexts by western blot. As shown in Figure 43C, similarly to the effect of *LTN1* deletion, the absence of Hel2 did not influence the production of Por1 protein in both wild-type and *xrn1*Δ cells, suggesting that Hel2 does not contribute to the ribosome stalling on POR1 mRNA.

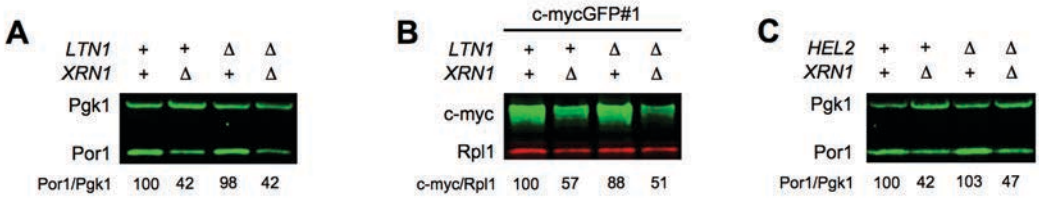

Figure 43. Por1 protein is not targeted by the RQC in *xrn1*Δ cells.

(A) Western blot of Por1 protein in wild-type, *xrn1*Δ, or/and *ltn1*Δ strains. Pgk1 was used as a loading control. Quantifications (from two independent experiments) of Por1 relative to Pgk1 are indicated below the blot.

(B) Western blot analysis of c-mycGFP no. 1 (#1) tagged proteins in wild-type, *xrn1*Δ, or/and *ltn1*Δ strains. Rpl1 was used as a loading control, and quantifications (from two independent experiments) of c-myc relative to Rpl1 are indicated below the blot.

(C) Western blot of Por1 protein in wild-type, *xrn1*Δ, or/and *hel2*Δ strains. Pgk1 was used as a loading control. Quantifications (from two independent experiments) of Por1 relative to Pgk1 are indicated below the blot.

15 | OCA2 mRNA is required for Por1 production in WT cells

In accordance to our results suggesting that OCA2 mRNA accumulation caused by the defective cytoplasmic 5'-3' RNA decay pathway leads to a decrease in Por1 protein production, we hypothesized that eliminating OCA2 transcript would

suppress the effect of *xrn1* mutation on Por1 downregulation. For this, we constructed isogenic *oca2*Δ and *oca2*Δ *xrn1*Δ strains (deleting the promoter and 5' proximal ORF region of *OCA2* but preserving the transcription termination region of *POR1*) and compared the levels of Por1 protein in these and other contexts (see Figure 44A). Surprisingly, the deletion of *OCA2* caused a strong decrease in Por1 production, independently of *XRN1* allele.

To get more details on the expression of *POR1* in the absence of *OCA2*, we prepared total RNA extracts from *oca2*Δ and *oca2*Δ *xrn1*Δ cells and analyzed POR1 mRNA by northern blot. As shown in Figure 44B, in the absence of *OCA2*, the steady-state quantity of POR1 mRNA decreased approximately two-fold as compared to in the presence of *OCA2*. The combination of *oca2*Δ and *xrn1*Δ mutations increased the steady-state level of POR1 mRNA in respect of *oca2*Δ situation; however, this increase was not sufficient to compensate the production of Por1 protein, as illustrated in Figure 44A.

As evoked previously, *xrn1* mutants accumulate various decay intermediates, among which—decapped mRNAs. To assess the 5' status of POR1 mRNA in the absence of *OCA2*, we treated total RNAs from the strains described above with Xrn1 *in vitro* and analyzed POR1 mRNA by northern blotting. As shown in Figure 44B, Xrn1-resistant quantity of POR1 mRNA remained comparable between *oca2*Δ and *oca2*Δ *xrn1*Δ strains, suggesting that POR1 mRNA contains an increased population of 5'-monophophorylated (thus decapped) transcripts in the absence of *OCA2* and *XRN1*.

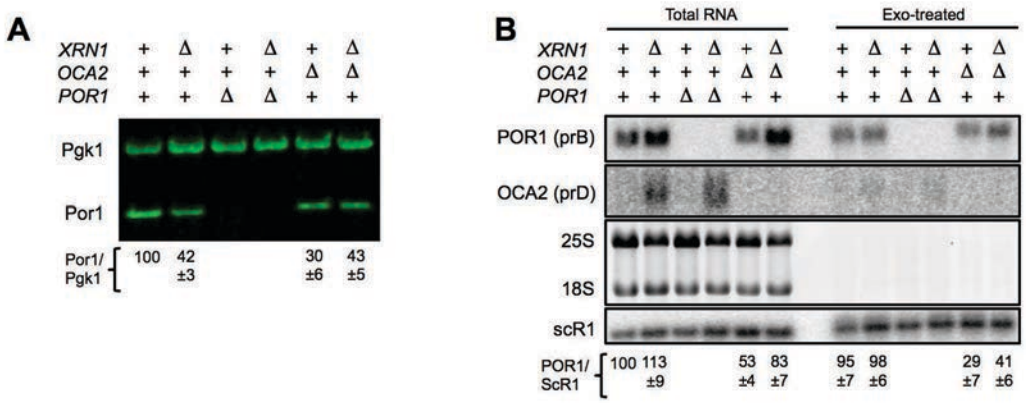

Figure 44. OCA2 mRNA is required for Por1 production.

(A) Western blot of Por1 protein in wild-type, *por1*Δ, *oca2*Δ and/or *xrn1*Δ strains. Pgk1 was used as a loading control. Quantifications of Por1 relative to Pgk1 are indicated below the blot.

(B) Northern blot showing steady-state levels of POR1 and OCA2 mRNAs in wild-type, *por1*Δ, *oca2*Δ and/or *xrn1*Δ strains. Total RNA and total RNA treated with Xrn1 *in vitro* (Exo-treated) to deplete 5'-monophosphorylated RNA (i.e., 5'-decapped RNA) are shown. The 25S and 18S rRNAs have 5'-monophosphate group and serve as a positive control for exonucleolytic digestion. The scR1 RNA (a Pol III transcript insensitive to exoribonuclease treatment) served as a loading control. Quantifications of POR1 relative to scR1, before or after 5'-3' exoribonuclease treatment, are indicated below the blot.

16 | OCA2 mRNA stabilizes POR1 mRNA *in trans*

As the deletion of *OCA2* significantly decreased the steady-state levels of POR1 mRNA, we reasoned that OCA2 mRNA might influence the stability of the overlapping transcript, POR1. To test this possibility, we overexpressed OCA2 mRNA *in trans* in WT and *xrn1*Δ cells and determined the decay rates of POR1 mRNA in these contexts. Technically, we treated yeast cultures with thiolutin, rapidly inhibiting RNA Pol II, and extracted total RNA from yeast cells at indicated time-points after thiolutin addition. As shown in Figure 45, OCA2 mRNA overexpression *in trans* stabilized POR1 mRNA in both wild-type and *xrn1*Δ cells, as compared to an empty vector control. This suggests that even when the 5'-3' cytoplasmic RNA decay pathway is defective, an interaction between convergent and overlapping mRNAs confers an additional stabilizing effect, prompting us to question the implication of the cytoplasmic exosome in this phenomenon.

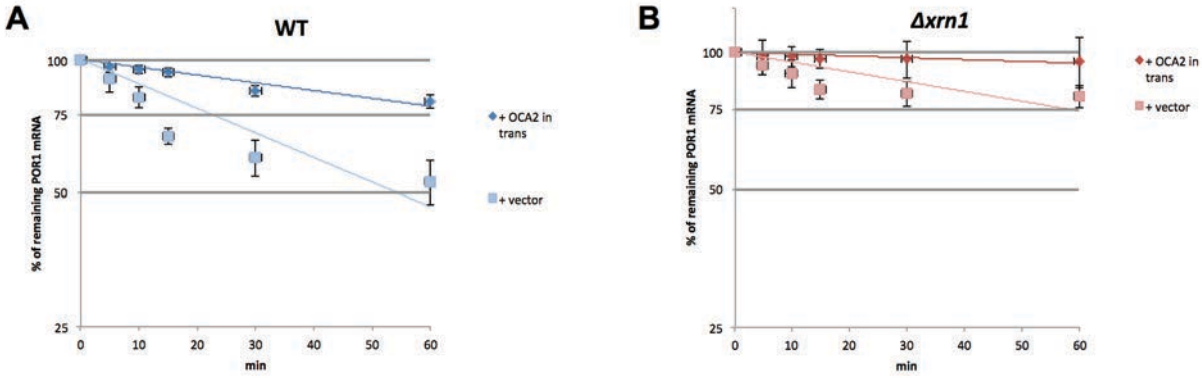

Figure 45. OCA2 mRNA stabilizes POR1 mRNA *in trans*.

(A) Degradation of POR1 mRNA in wild-type cells, upon expression (+) or not (-) of OCA2 mRNA *in trans*, was monitored after transcriptional shut-off with thiolutin, and RNA was isolated at the indicated time points. Half-life measurements of POR1 transcripts were carried out by quantification of the mRNA remaining at each time point by northern blot, after correcting for loading differences using scR1 RNA. The % of remaining mRNA axis is in logarithmic scale.

(B) Degradation of POR1 mRNA in *xrn1*Δ cells, analyzed as in (A).

17 | Cytoplasmic exosome mutation suppresses the necessity of OCA2 mRNA for Por1 production

Seeking to determine if the interaction between POR1 and OCA2 mRNAs confers a stabilizing effect for POR1 mRNA, we complemented the *oca2*Δ strain with the deletion of *SKI2*, an essential cofactor of the cytoplasmic exosome. We then compared the steady-state levels of POR1 mRNA and Por1 protein in the double mutant *oca2*Δ *ski2*Δ and other strains analyzed previously. As detected by northern and western blot analysis (Figure 46A-B), the inactivation of the cytoplasmic exosome increased the quantity of POR1 mRNA and, accordingly, of Por1 protein in *oca2*Δ *ski2*Δ cells as compared to *oca2* mutant. Further on, we also analyzed the decay rates of POR1 mRNA in WT, *oca2*Δ and *oca2*Δ *ski2*Δ cells following RNA Pol II inhibition by thiolutin, as described above (Figure 46E). In wild-type cells, POR1 mRNA had a half-life of 49 minutes, positioning it as a relatively stable transcript (*75*). However, in the absence of *OCA2*, POR1 mRNA remained at its initial decreased level even after thiolutin treatment (for at least an hour), suggesting additional prerequisites for POR1 messenger decay. Importantly, in the *oca2 ski2* double mutant, the exponential POR1 mRNA decay pattern was recovered, similarly to wild-type situation. These results point out that in wild-type condition, the interaction between convergent and

overlapping POR1 and OCA2 mRNAs is important to limit the destabilization of POR1 mRNA by the cytoplasmic exosome (Figure 46C-D). Moreover, as POR1 mRNA levels decrease but remaine highly stable in *oca2*Δ cells (Figure 46E), the absence of OCA2 mRNA changes the fate of POR1 messenger, and strikingly the absence of 3'-5' exonucleolytic attack suppresses this phenomenon.

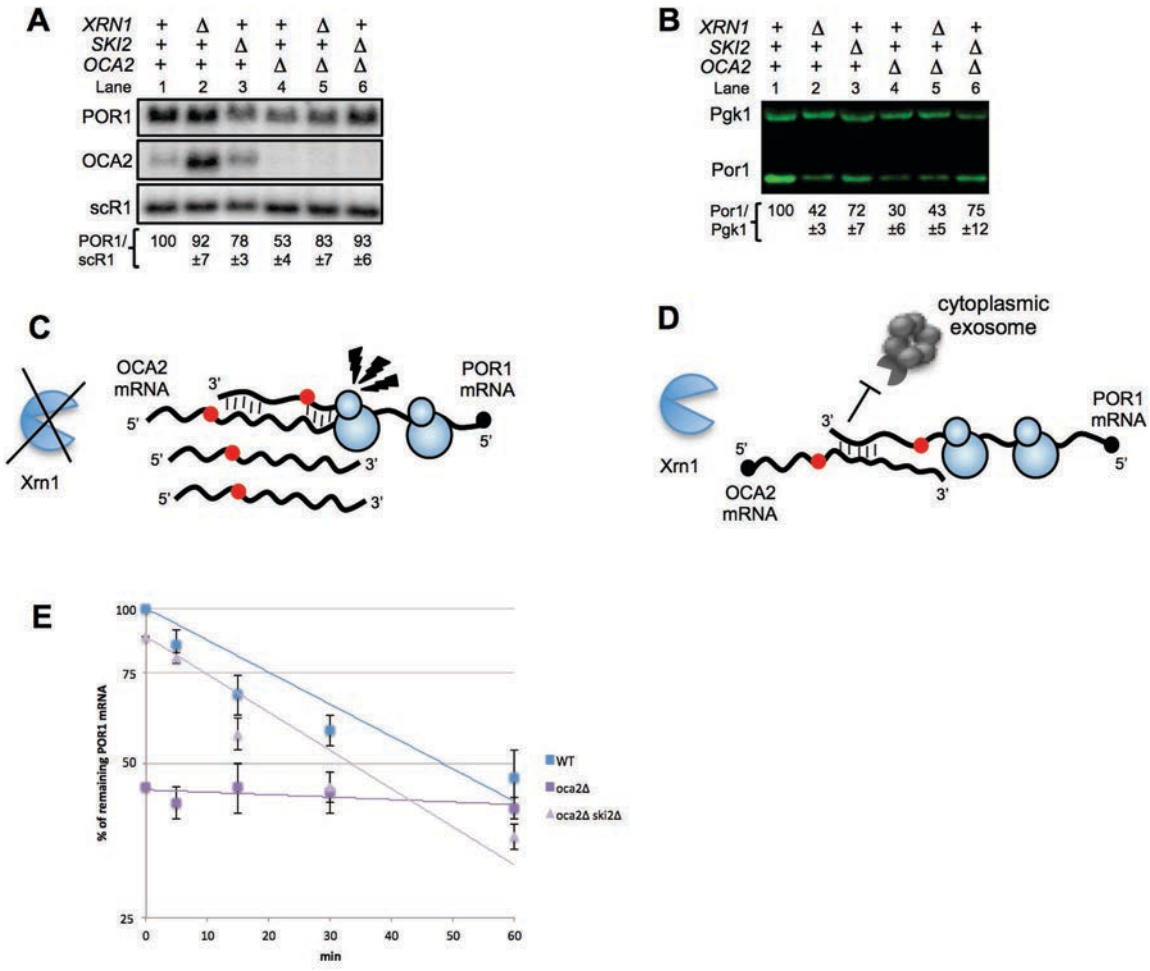

Figure 46. Inactivating cytoplasmic exosome restores the Por1 expression in the absence of OCA2 mRNA.

(A) Northern blot analysis of POR1 and OCA2 mRNAs in different *xrn1*Δ, *ski2*Δ, and *oca2*Δ genetic contexts. The scR1 RNA served as a loading control. Quantifications of POR1 relative to scR1 are indicated below the blot.

(B) Western blot analysis of Por1 protein in different *xrn1*Δ, *ski2*Δ, and *oca2*Δ genetic contexts. Pgk1 was used as a loading control. Quantifications of Por1 relative to Pgk1 are indicated below the blot.

(C) Model of the translational regulation of POR1 mRNA in the *xrn1* mutant. The ribosome stall due to the presence of OCA2-POR1 mRNA duplex is symbolized by lightning flashes.

(D) Model of the functional stabilization of POR1 mRNA by OCA2 mRNA in wild-type cells. Interaction with the OCA2 mRNA protects the POR1 mRNA 3' end against the exosome-mediated 3'-5' degradation. Translational stop codons are indicated by red circles. Bars indicate the potential interactions between sequences of the two RNAs.

(E) Degradation of POR1 mRNA in wild-type, *oca2*Δ and *oca2*Δ *ski2*Δ cells was monitored after transcriptional shut-off with thiolutin, and RNA was isolated at the indicated time points. Half-life measurements of POR1 transcripts were carried out by quantification of the mRNA remaining at each time point by northern blot, after correcting for loading differences using scR1 RNA. The % of remaining mRNA axis is in logarithmic scale.

18 | Cytoplasmic deadenylase mutation suppresses the necessity of OCA2 mRNA for Por1 production

As described in the Part I of Chapter I (see section 3), both 5'-3' and 3'-5' cytoplasmic decay pathways are initiated by a common step of 3' poly(A) tail shortening. To address the role of deadenylation prior to the 3'-5' exosome attack of POR1 mRNA in the *oca2*Δ context, we deleted *CCR4*, coding for the catalytic subunit of the major cytoplasmic deadenylase complex Ccr4-Not, in wild-type and *oca2*Δ cells. We then determined the steady-state quantity of POR1 mRNA and Por1 protein by northern and western blotting. As shown in Figure 47A, the absence of Ccr4 also stabilized POR1 mRNA in the *oca2*Δ context, as compared to the *OCA2* deletion alone; similarly, the quantity of Por1 protein increased in the double *oca2*Δ *ccr4*Δ mutant as compared to the simple *oca2*Δ mutant (Figure 47B). Additionally, we determined the decay kinetics of POR1 mRNA in *oca2*Δ *ccr4*Δ cells, and compared it to the results described above from *OCA2* and *oca2*Δ cells (Figure 47C). Following increased steady-state levels of POR1 mRNA in the double *oca2 ccr4* mutant, the decay rate of POR1 messenger was significantly reduced in this context. These results confirm that, as expected, 3'-5' deadenylation step precedes the cytoplasmic exosome attack of POR1 mRNA, essential for the functional stability of POR1 messenger in the absence of interacting OCA2 mRNA.

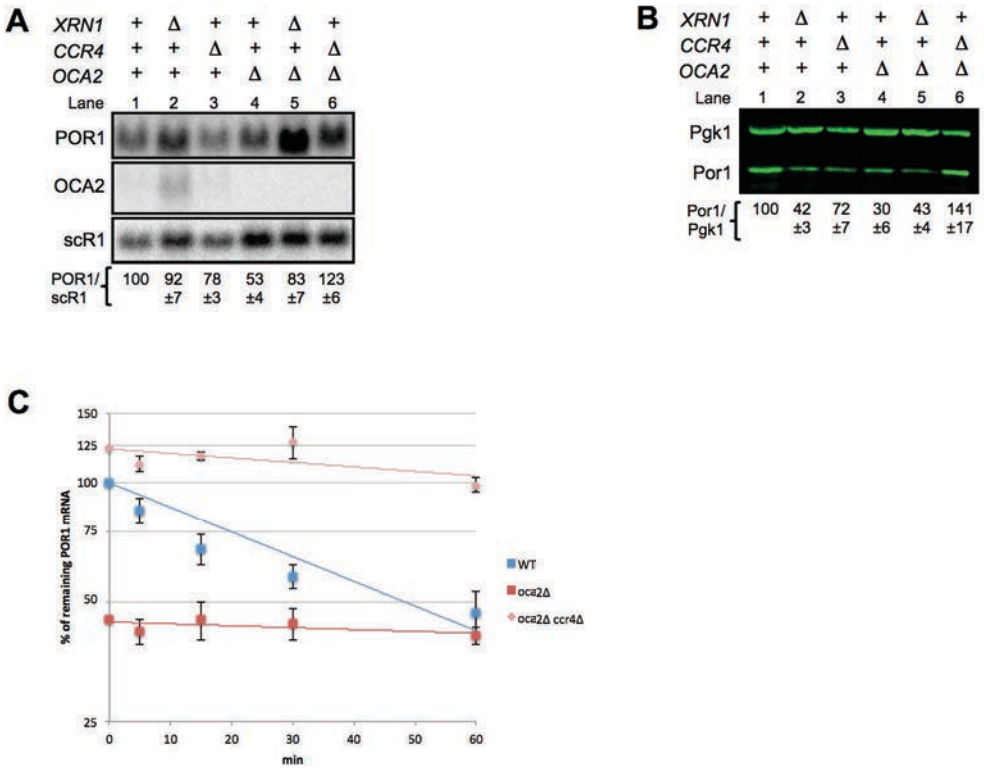

Figure 47. Inactivating cytoplasmic deadenylase restores the Por1 expression in the absence of OCA2 mRNA.

(A) Northern blot analysis of POR1 and OCA2 mRNAs in different *xrn1*Δ, *ccr4*Δ, and *oca2*Δ genetic contexts. The scR1 RNA served as a loading control. Quantifications of POR1 relative to scR1 are indicated below the blot.

(B) Western blot analysis of Por1 protein in different *xrn1*Δ, *ccr4*Δ, and *oca2*Δ genetic contexts. Pgk1 was used as a loading control. Quantifications of Por1 relative to Pgk1 are indicated below the blot.

(C) Degradation of POR1 mRNA in wild-type, *oca2*Δ and *oca2*Δ *ccr4*Δ cells was monitored after transcriptional shut-off with thiolutin, and RNA was isolated at the indicated time points. Half-life measurements of POR1 transcripts were carried out by quantification of the mRNA remaining at each time point by northern blot, after correcting for loading differences using scR1 RNA. The % of remaining mRNA axis is in logarithmic scale.

19 | Defective ribosome recycling does not suppress the necessity of OCA2 mRNA for Por1 production

A recent study has reported that in the absence of Dom34 a number of mRNAs accumulate ribosomes on their 3' UTRs, as revealed by ribosome profiling experiments (*213*). These 3' UTR-associated ribosomes were described to be

translationally inactive, as their footprints lack the characteristic three-nucleotide phasing. Interestingly, this study identified POR1 mRNA among the transcripts showing the most significant enrichment of ribosomes on its 3' UTR upon *DOM34* deletion. This information led us to hypothesize that in the absence of Dom34 the ribosomes reportedly associated with the 3' UTR of POR1 mRNA could protect POR1 mRNA from the 3'-5' exonucleolytic attack and thus alleviate the defect of Por1 protein production in *oca2*Δ cells. A similar mechanism has been already proposed for stop codon-less mRNAs (NSD substrates), the stability of which is enhanced by the presence of 3' end-associated ribosomes, inefficiently dissociated in *dom34* mutants (*208*). To address the effect of Dom34 for the compromised expression of Por1 protein in the *oca2*Δ context, we constructed the *oca2*Δ *dom34*Δ double mutant and quantified Por1 protein in isogenic WT, *oca2*Δ, *dom34*Δ and *oca2*Δ *dom34*Δ strains by western blot. However, as shown in Figure 48, the defective ribosome recycling did not suppress the downregulation of Por1 protein in the absence of Oca2, suggesting that the presence of ribosomes on 3' UTRs is not directly linked to the mRNA 3' end accessibility for the degradation by the cytoplasmic exosome.

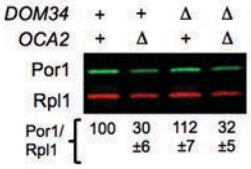

Figure 48. *DOM34* deletion does not restore the Por1 expression in the absence of OCA2 mRNA.

Western blot of Por1 protein in wild-type, *oca2*Δ, or/and *dom34*Δ strains. Rpl1 was used as a loading control. Quantifications of Por1 relative to Rpl1 are indicated below the blot.

20 | POR1 and OCA2 mRNAs differentially respond to carbon source changes

As evoked previously, Por1 protein plays an important role in mitochondrion- and respiration-related processes of the cell. The earlier study from our laboratory has revealed that the expression of *POR1* is upregulated at mRNA as well as protein level when cells start to metabolize non-fermentable carbon sources (*66*). We thus

wanted to address the functional crosstalk between the interaction of POR1 and OCA2 mRNAs and the remodeling of carbon metabolism. Firstly, we cultured wild-type yeast cells for an extended time lapse (7 days) during which cells are known to shift from fermentative to respiratory metabolism (diauxic shift) (*82*). We collected cells during the exponential growth phase and then subsequently every 24 hours. As shown in Figure 49A, northern blot analysis of POR1 mRNA from these cells confirmed its upregulation within the first 24 hours of the experiment, when glucose concentration becomes limiting and cells shift to respiration. Interestingly, OCA2 mRNA expression decreased within the same timeframe. During the following days, POR1 mRNA expression gradually decreased, as did the level of OCA2 mRNA. In contrast, the level of Por1 protein initially followed the upregulation of POR1 mRNA but remained stable afterwards (Figure 49B), suggesting its high stability in post diauxic shift cells.

Further on, we chose to zoom on POR1 and OCA2 mRNAs at the beginning of the switch to respiratory metabolism. For this purpose, we transferred exponentially growing cells (in the presence of glucose) to the medium where glycerol was the sole carbon source. We harvested cells at time points indicated in Figure 49C and analyzed the expression of POR1 and OCA2 mRNAs in isogenic WT, *xrn1*Δ and *oca2*Δ strains. Similarly in these three contexts, the expression of POR1 mRNA gradually increased after the shift to glycerol-containing medium. Interestingly, OCA2 mRNA remained stable during the first two hours after switching to respiratory conditions, suggesting that the decrease in its quantity happens in subsequent stages of respiratory metabolism.

S. cerevisiae is a well-known "specialist" organism, using exclusively glucose as a preferred carbon source, even when other carbon sources are present. Accordingly, it has evolved to rapidly remodel its carbon metabolism in case glucose becomes newly available. This trait resumes a specific trade-off between the advantageous growth when glucose is abundant and the delayed adaptation upon carbon source change (*214*). We thus chose to analyze the expression of POR1-OCA2 mRNA pair upon glucose repletion. We pre-incubated isogenic WT, *xrn1*Δ and *oca2*Δ strains in glycerol-containing medium for 3 hours and subsequently collected cell samples after glucose addition. As shown in Figure 49D, similarly to the steady-state levels in the presence of glucose, the quantity of POR1 mRNA was lower in *oca2*Δ cells after 3

hours in respiration-requiring medium, as compared to WT or *xrn1*Δ cells. Remarkably, within 30 to 60 minutes after glucose repletion, all POR1 mRNA was degraded in the three strains tested, and gradually resumed its quantity within 3 to 4 hours after glucose addition. Importantly, in accordance to our previous steady-state results in YPD medium, POR1 mRNA failed to reach the wild-type level in the absence of *OCA2* during the metabolic shift from respiration to fermentation. We also probed OCA2 mRNA during our analysis, and the usage of RNA decay mutant revealed the striking dynamics of *OCA2* expression. In wild-type cells, OCA2 mRNA was barely detectible during the passage from glucose to glycerol in the growth medium; however, in *xrn1* mutant, we observed a massive accumulation of OCA2 transcript, increasing with time after glucose repletion. Intriguingly, after a prolonged exposition of the *OCA2* probe-hybridized northern blot we could detect the increase in length of the OCA2 mRNA. This result prompted us to hypothesize on the existence on distinct 3' isoforms of OCA2 mRNA, as supported by the genome-wide mRNA isoprofiling experiments (Figure 49E).

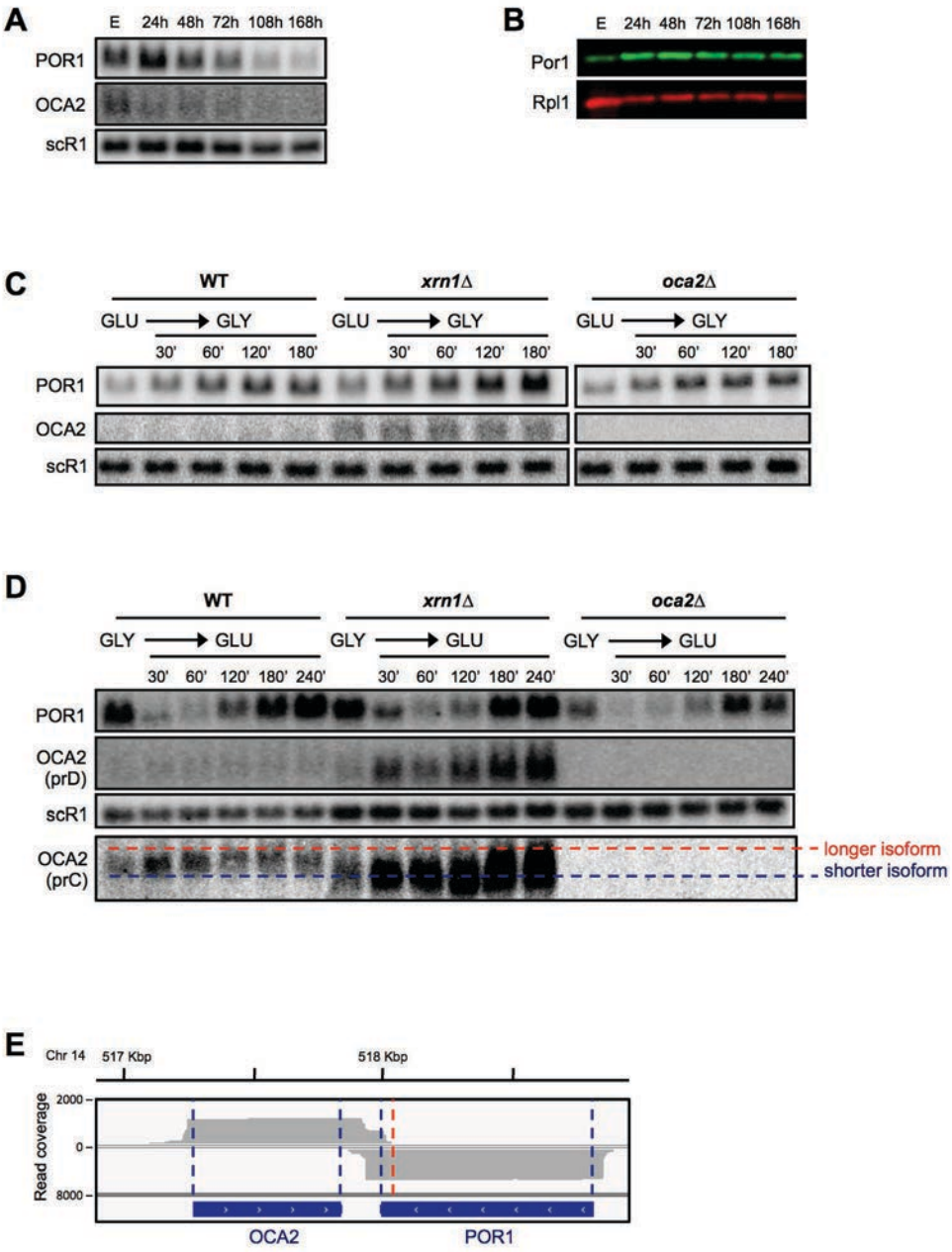

Figure 49. The expression of POR1 and OCA2 mRNAs upon carbon source changes.

(A) Northern blot analysis of POR1 and OCA2 mRNAs in wild-type cells during exponential growth phase (E) or 24, 48, 72, 108, and 168 hours after inoculation. The scR1 RNA served as a loading control. Probes prB (POR1) and prD (OCA2) were used.

(B) Western blot analysis of Por1 protein in wild-type cells, from the cultures used for northern analysis in (A). Rpl1 was used as a loading control.

(C) Northern blot analysis of POR1 and OCA2 mRNAs in wild-type, *xrn1*Δ and *oca2*Δ cells in YPD medium (GLU) or after a transfer in YPG medium (GLY), at indicated time points. The scR1 RNA served as a loading control. Probes prB (POR1) and prD (OCA2) were used.

(D) Northern blot analysis of POR1 and OCA2 mRNAs in wild-type, *xrn1*Δ and *oca2*Δ cells after 3 hours in YPG medium (GLY) and after glucose repletion (GLU), at indicated time points. The scR1 RNA served as a loading control. Probes prB (POR1), prC and prD (OCA2) were used. Blue and red dashed lines indicate putative shorter and longer isoforms of OCA2 mRNA, respectively.

(E) POR1 and OCA2 mRNA isoform profiling data, retrieved from Pelechano *et al.* (*9*). Blue dashed lines delimit ORFs of POR1 and OCA2. Red dashed line marks the putative longer isoform of OCA2 transcript. Chromosome 14 coordinates are indicated above.

PART II | The study of convergent mRNA interaction genome-wide

1 | RNAi reconstitution in *S. cerevisiae*

As commented in Chapter I, RNA interference, a gene-silencing pathway triggered by double-stranded RNA, has been lost in the model budding yeast *S. cerevisiae*, but the introduction of Dicer (Dcr1) and Argonaute (Ago1) from *S. castellii* leads to the production of artificial siRNAs. RNAi-competent *S. cerevisae* thus allows the identification of small RNAs (19-23 nucleotides), corresponding to naturally formed RNA duplexes *in vivo*. Importantly, in *S. castellii*, where more than a fourth of total number of ORFs is in convergent orientation, the majority of convergent mRNAs overlaps and approximately a half of these overlapping transcripts produces endogenous siRNAs (*51*). We therefore used this RNAi tool to determine the density of siRNAs originating from convergent gene loci produced in the presence or absence of Dcr1/Ago1 in *S. cerevisiae*. The wild-type strains harboring chromosomally integrated *DCR1* and *AGO1* genes from *S. castellii* were provided by David Bartel's laboratory that initially reported reconstituted RNAi in *S. cerevisiae* (*51*). We then constructed the isogenic *xrn1*Δ strains, used for further analysis described below.

2 | RNAi-competent cells produce POR1-OCA2 corresponding siRNA *in vivo*

We first chose to validate the functional reconstitution of RNAi in *S. cerevisiae* by examining the siRNA production corresponding to *POR1-OCA2* locus. We extracted total RNA from isogenic WT and *xrn1*Δ cells, positive (expressing Dcr1/Ago1) or negative for RNAi, and probed for small RNA by northern blot. As shown in Figure 50, approximately 23 nucleotide-long RNA species were detected specifically in RNAi-positive *xrn1*Δ cells, corresponding to the OCA2 strand within the region of POR1 and OCA2 mRNA overlap. As a positive control, Ty1 specific siRNAs were detected in RNAi-positive wild-type and, to a greater extent, *xrn1*Δ cells, similarly to the report by Drinnenberg *et al.* (*51*). These results indicated that reconstituted RNAi

was functional in both WT and *xrn1*Δ *S. cerevisiae* cells and incited us to proceed to small RNA sequencing experiments.

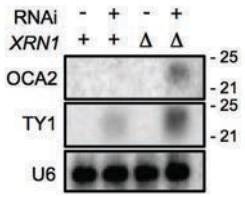

Figure 50. Small RNAs originating from interacting POR1 and OCA2 mRNAs are produced in RNAi-competent cells.

Northern blot analysis of OCA2 mRNA-originating small RNAs in wild-type and *xrn1*Δ, RNAi-positive or negative cells. Ty1 transcript-derived small RNA was used as a positive control, as described (*51*). U6 snRNA was used as a loading control. On the right, ssRNA size marker is shown.

Small RNA sequencing libraries were constructed using size-selected (10- to 40-nucleotide) RNAs, extracted from RNAi-positive or negative wild-type and *xrn1*Δ cells. 19- to 23-nucleotide reads obtained by single-end sequencing were uniquely mapped to *S. cerevisiae* reference genome. In accordance to the siRNA detection by northern blot originating from POR1-OCA2 mRNA duplex, these small RNA were also detected by RNA sequencing. Figure 51 shows the snapshot of siRNA signal densities in RNAi-positive WT and *xrn1*Δ cells, corresponding to *POR1-OCA2* locus. Small RNAs mapping to the intergenic region between *POR1* and *OCA2* were readily detected in Dcr1/Ago1 expressing wild-type cells. As expected, siRNA signal was significantly stronger in Dcr1/Ago1 expressing *xrn1* mutant cells. Moreover, in the absence of Xrn1, small RNA density was detected within the 3' proximal ORF region of *POR1*, confirming our previous results that showed POR1-OCA2 mRNA interaction extending within the coding part of POR1 mRNA.

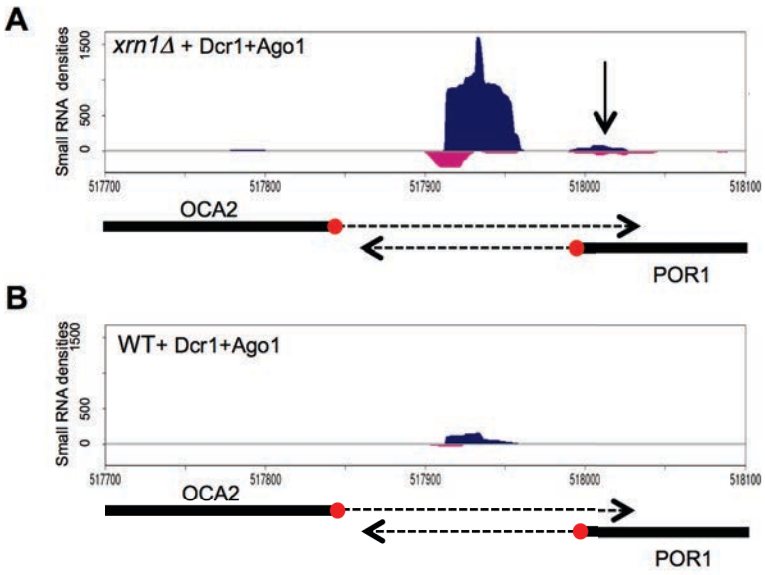

Figure 51. siRNA originating from interacting POR1 and OCA2 mRNAs are detected by small RNA sequencing in RNAi-competent cells.

(A) Snapshot of the small RNA accumulation of 19–23 nucleotides within the *OCA2-POR1* region in wild-type cells, in the presence of Dcr1/Ago1. The arrow shows reads corresponding to the *POR1* ORF.

(B) Snapshot as in (A), in the *xrn1* mutant. Small RNA reads originating from Watson strand are shown in blue, and small RNA reads originating from Crick strand are shown in red. Read densities are in linear scale. Dashed arrows represent the non-translated regions of the OCA2 and POR1 mRNAs, and filled boxes indicate ORFs. Chromosome 14 coordinates are indicated below (from 517700 to 518100).

3 | siRNA production is selectively enriched among convergent mRNAs in RNAi competent cells

We extended our analysis of the *POR1-OCA2* locus to other convergent mRNA pairs. At a global scale, we analyzed 4013 mRNAs, consisting of 2557 annotated mRNAs with no obvious antisense (coding or non-coding) RNA, which we called solo mRNAs, and 1456 convergent and overlapping mRNAs (see Table S1 in Sinturel *et al.* (*215*)). Among convergent mRNAs, we defined 1142 mRNAs overlapping in their 3' UTR only and 315 overlapping within their ORF. A significant amount of mRNAs was not included in our analysis due the presence of overlapping non-coding transcription; dubious ORF-containing messengers were equally not considered. For each analyzed mRNA, we calculated the density of small RNAs produced in the presence or absence of Dcr1/Ago1 in both wild-type and *xrn1*

mutant strains. For convergent mRNAs, only the region of overlap between two transcripts was considered for siRNA density calculations; for solo mRNAs, siRNAs originating from the entire transcript were included in the analysis. Figure 52B reveals that solo mRNAs were poor sources of small RNAs in the presence of Dcr1/Ago1 and were not affected by the absence of Xrn1. In contrast, convergent mRNAs, overlapping either in their 3' UTR region or also overlapping within their ORFs, were sensitive to the presence of Dcr1/Ago1 and produced an additional enrichment in small RNAs in the *xrn1* mutant (Figure 52B). In the RNAi-negative cells, solo mRNAs and both subclasses of convergent mRNAs yielded similar small RNA densities, independently of *XRN1* allele (Figure 52A), and originated mostly from abundant RNA (rRNA, tRNA, mRNA, etc.) decay intermediates, as previously reported (*51*).

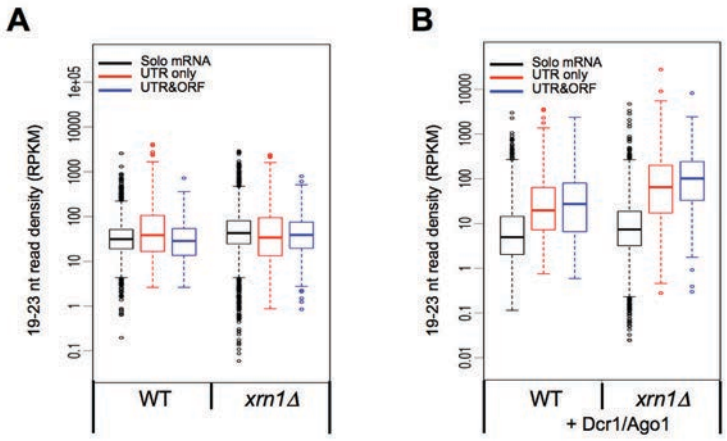

Figure 52. Genome-wide production of siRNA is enriched at convergent gene loci.

(A) Boxplots showing the average density of 19–23 nucleotide tags within 3' overlapping convergent genes and non-overlapping genes in the absence of Dcr1/Ago1 in the indicated strains. 4013 mRNAs were defined (see Chapter IV). 1456 3' overlapping convergent mRNAs and 2557 mRNAs having no antisense (called solo mRNAs) were analyzed. Among convergent mRNAs, we defined 1142 mRNAs overlapping in their 3' UTR only (UTR only) and 315 overlapping within their ORF (ORF and UTR). Read densities are in rpkm (reads per kilobase of transcript per million mapped reads).

(B) Boxplots showing the average density of 19–23 nucleotide tags within 3' overlapping convergent genes and non-overlapping genes as in (A) except for the presence of Dcr1/Ago1.

4 | siRNA signal corresponding to convergent mRNA is amplified in RNAi competent *xrn1*Δ cells

We pursued the analysis of small RNA sequencing data by representing each mRNA, either in the wild-type or in the *xrn1* mutant, by following the distribution of its small RNA density ratio (in the presence of Dcr1/Ago1 versus in the absence of Dcr1/Ago1) plotted against the average of its small RNA density (in the presence and in the absence of Dcr1/Ago1), a representation known as an MA plot (Figure 53A-B; see Table S2 in Sinturel *et al.* (*215*)). We analyzed 1923 solo mRNAs and 1087 convergent mRNAs after filtering out the 1003 mRNAs (25% of the 4013 mRNA set defined for Figure 52), showing the lowest density average in the four strains ($WT_{+Dcr1+Ago1}$ + WT + *xrn1*$\Delta_{+Dcr1+Ago1}$ + *xrn1*Δ; see Figure 53C). The MA plot showed that a third of analyzed convergent mRNAs produced a ≥2-fold enrichment of small RNA levels in RNAi-competent wild-type cells (Figure 53A). Remarkably, in the absence of Xrn1, we detected a massive enrichment of mRNA-mRNA interactions involving 678 mRNAs (Figure 53B), which corresponds to a two-fold increase as compared to RNAi-competent wild-type cells.

In Figure 53D, we represented the numbers and percentages of mRNAs producing a ≥2-fold enrichment of siRNA signal upon RNAi reconstitution (retrieved from MA plot analysis) by Venn diagrams. Interestingly, 324 convergent mRNAs (29.8%) showed a significant enrichment of siRNA signal in both wild-type and *xrn1*Δ RNAi-competent cells, as compared to 54 solo mRNAs (2.8%). Moreover, approximately a half of analyzed convergent mRNAs (48.3%) produced a ≥2-fold increase in siRNA signal originating from the both transcripts of a convergent mRNA pair in the absence of Xrn1, and a significant part of these pairs (16.2%) yielded an enrichment of siRNA signal independently of *XRN1* allele.

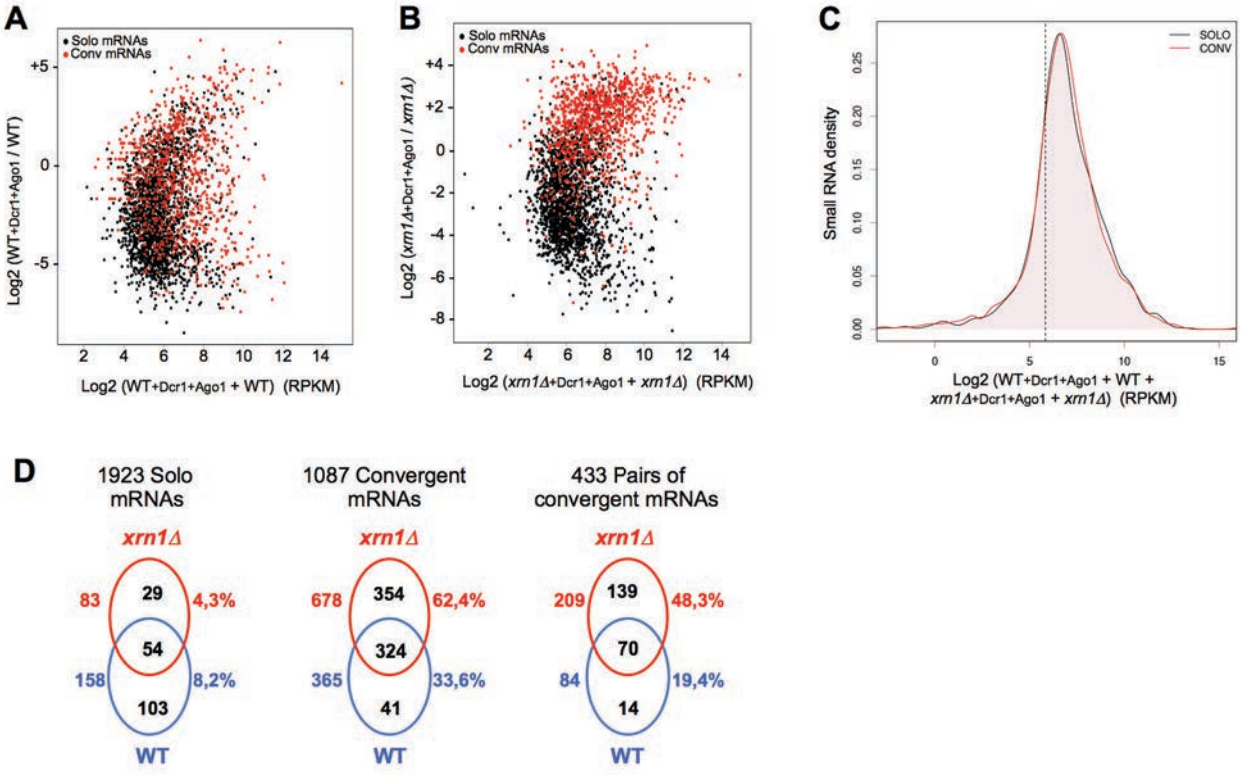

Figure 53. Genome-wide production of siRNA at convergent gene loci is amplified in *xrn1*Δ cells.

(A) MA plot representation of 19- to 23-nt small RNA density ratio in wild-type strains with or without Dcr1/Ago1 (labeled Log_2 [$WT_{+Dcr1+Ago1}$/WT]) as a function of 19- to 23-nt small RNA density average in the same strains (labeled Log_2 [$WT_{+Dcr1+Ago1}$ + WT]). In this MA plot, 1923 solo mRNAs (black dots) and 1087 convergent mRNAs (conv mRNAs indicated by red dots) are represented after filtering out 1003 mRNAs, corresponding to 25% of the 4013 mRNAs showing the lowest small RNA density, as shown in (C).

(B) MA plot representation performed as in (A), in the xrn1 mutant strain, with the same subset of mRNAs.

(C) 25% of the mRNA set showing the lowest density average in the four strains ([$WT_{+Dcr1+Ago1}$ + WT + *xrn1*$\Delta_{+Dcr1+Ago1}$ + *xrn1*Δ]/4) were filtered out (1003 mRNAs below the threshold represented by a dashed line). Read densities are in rpkm (reads per kilobase of transcript per million mapped reads).

(D) Venn diagram representation of numbers and percentages of solo mRNAs, convergent mRNAs, or pairs having both convergent mRNAs showing ≥2-fold small RNA enrichment in the presence of Dcr1/Ago1 in the wild-type and in the *xrn1* mutant (data obtained from the MA plot analyses).

5 | siRNA signal is specific for the overlap region between convergent mRNAs

Seeking to analyze the specificity of the siRNA signal originating from convergent mRNAs in RNAi-competent cells, we represented small RNA densities obtained from RNAi-positive or negative wild-type and *xrn1*Δ cells in the form of a metagene. We divided the convergent mRNAs used for MA plot analysis (Figure 53A-B) into two subgroups (overlapping in their 3' UTR only or overlapping within their ORF) defined for boxplot analysis (Figure 52) and calculated metagene siRNA densities for 70 mRNA pairs, showing a ≥2-fold enrichment of small RNA signal upon RNAi reconstitution in both WT and *xrn1*Δ cells, and for 209 mRNA pairs, showing a significant enrichment in the absence of Xrn1 (as illustrated in Venn diagrams of Figure 53D). In both subgroups of convergent mRNA pairs, the siRNA signal produced upon Dcr1/Ago1 reintroduction was limited to the region of overlap between sense and antisense mRNAs, irrespective of the presence or absence of Xrn1; moreover, this signal was highly dependent on RNAi reconstitution, as RNAi-negative cells failed to accumulate small RNA densities above background level (Figure 54).

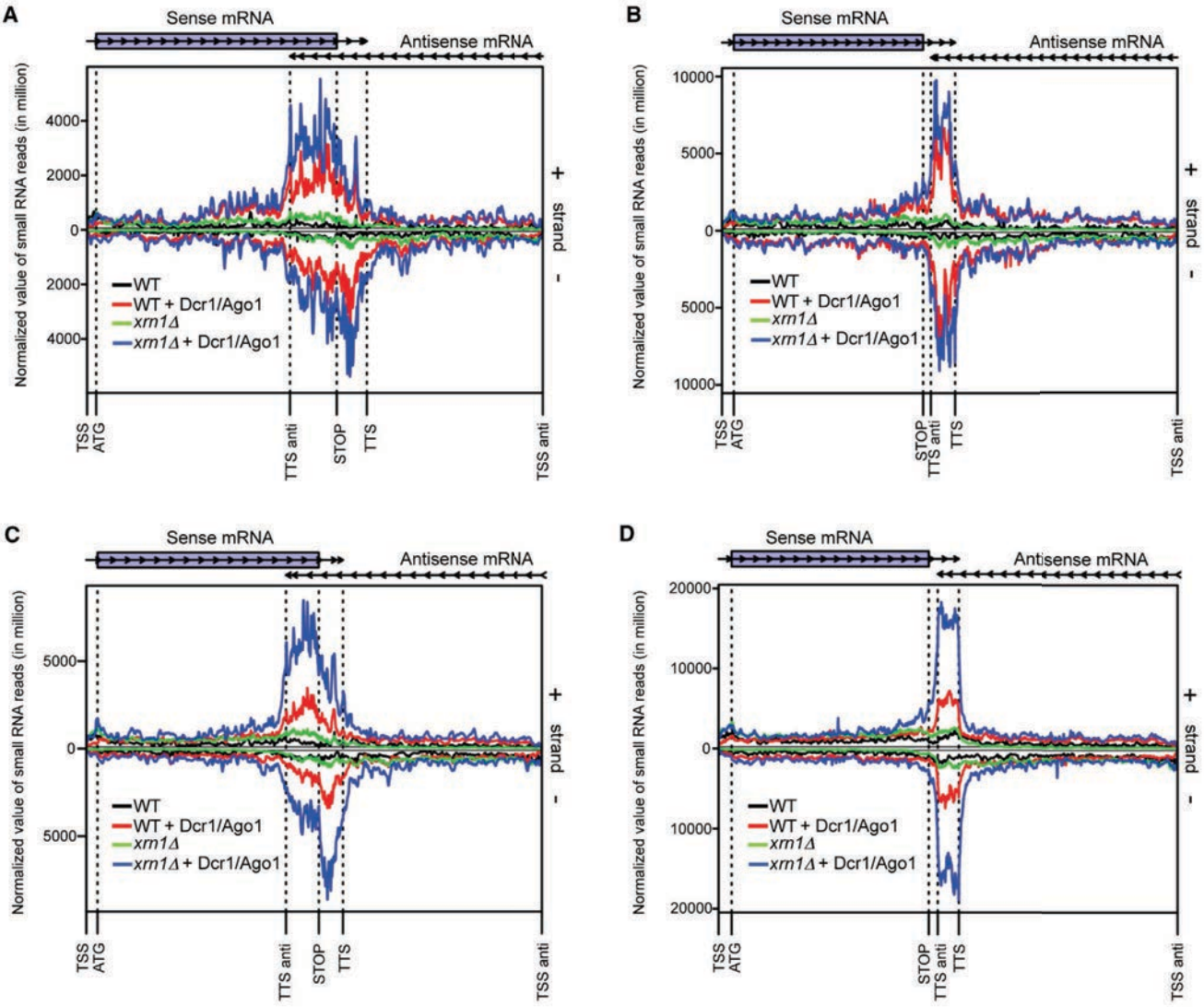

Figure 54. Genome-wide production of siRNA at convergent loci is restricted to the region of overlap between sense-antisense mRNAs.

Analysis was performed for mRNAs showing ≥2-fold small RNA enrichment in the presence versus the absence of Dcr1/Ago1 in the indicated strains (see Tables S3 and S4 from Sinturel *et al.* (*215*) for details on selected mRNAs).

(A) Small RNA meta-signal along mRNAs overlapping within their ORFs (ORF-UTR) and showing small RNA enrichment in both wild-type and *xrn1*Δ conditions.

(B) Same as (A) for mRNAs overlapping in their 3' UTR only (UTR only).

(C) Small RNA meta-signal along mRNAs overlapping within their ORFs (ORF-UTR) and showing small RNA enrichment in the *xrn1* mutant only.

(D) Same as (C) for mRNAs overlapping in their 3' UTR only (UTR only). Above each panel, arrows represent sense and antisense mRNAs, and the ORF of the sense mRNA is indicated by a blue box. Different boundaries delimiting regions of interest are defined: sense mRNA transcription start site (TSS), translation start (ATG) and stop codon (STOP), and transcription termination site (TTS); antisense mRNA transcription termination site (TTS anti) and transcription start site (TSS anti). x and

y axis correspond, respectively, to virtual nucleotides and to the mean of the normalized small RNA signal in rpkm at each virtual nucleotide (see Chapter IV).

6 | Validation of new pairs of messenger-interacting messenger RNA (mimRNA)

To validate the existence of other NGD-targeted messenger-interacting messenger RNAs (mimRNAs), two pairs of convergent and 3' overlapping mRNAs, SRB7-PMT7 and SNM1-PEX29, which show small RNA enrichment in the presence of Dcr1/Ago1, were selected for further analysis. Similarly to the expression pattern of POR1 and OCA2 mRNAs, SRB7 mRNA extends within the 3' proximal ORF region of PMT7 mRNA; remarkably, both SNM1 and PEX29 mRNAs extended within respective 3' proximal portions of ORF (Figure 55A-B). RNA duplexes between the 3' UTR and ORF regions of these two RNAs in wild-type cells were observed and, furthermore, were strongly increased in Xrn1-deficient cells (Figure 55C-D). Seeking to analyze the interaction of SRB7 and PEX29 mRNAs with their respective mimRNAs PMT7 and SNM1, we constructed two plasmids, placing *SRB7* (pLB132) and *PEX29* (pLB131) under control of $tetO_7$ promoter (analogously to *OCA2* in pLB001). We then extracted total RNA from SRB7 or PEX29 mRNA overexpressing wild-type and *xrn1*Δ cells and probed for PMT7 or SNM1 3' NGD intermediates by northern blot. As shown in Figure 55E-F, in the absence of Xrn1, we observed an accumulation of 3' RNA fragments corresponding to the PMT7 and SNM1 mRNAs that was dependent on the overexpression of SRB7 and PEX29 mRNAs, respectively, as compared to an empty vector control. These results, mirroring the OCA2 mRNA-induced NGD cleavage of POR1 mRNA, suggest that additional mRNA-mRNA interactions occur *in vivo* and are subjected to Xrn1 surveillance.

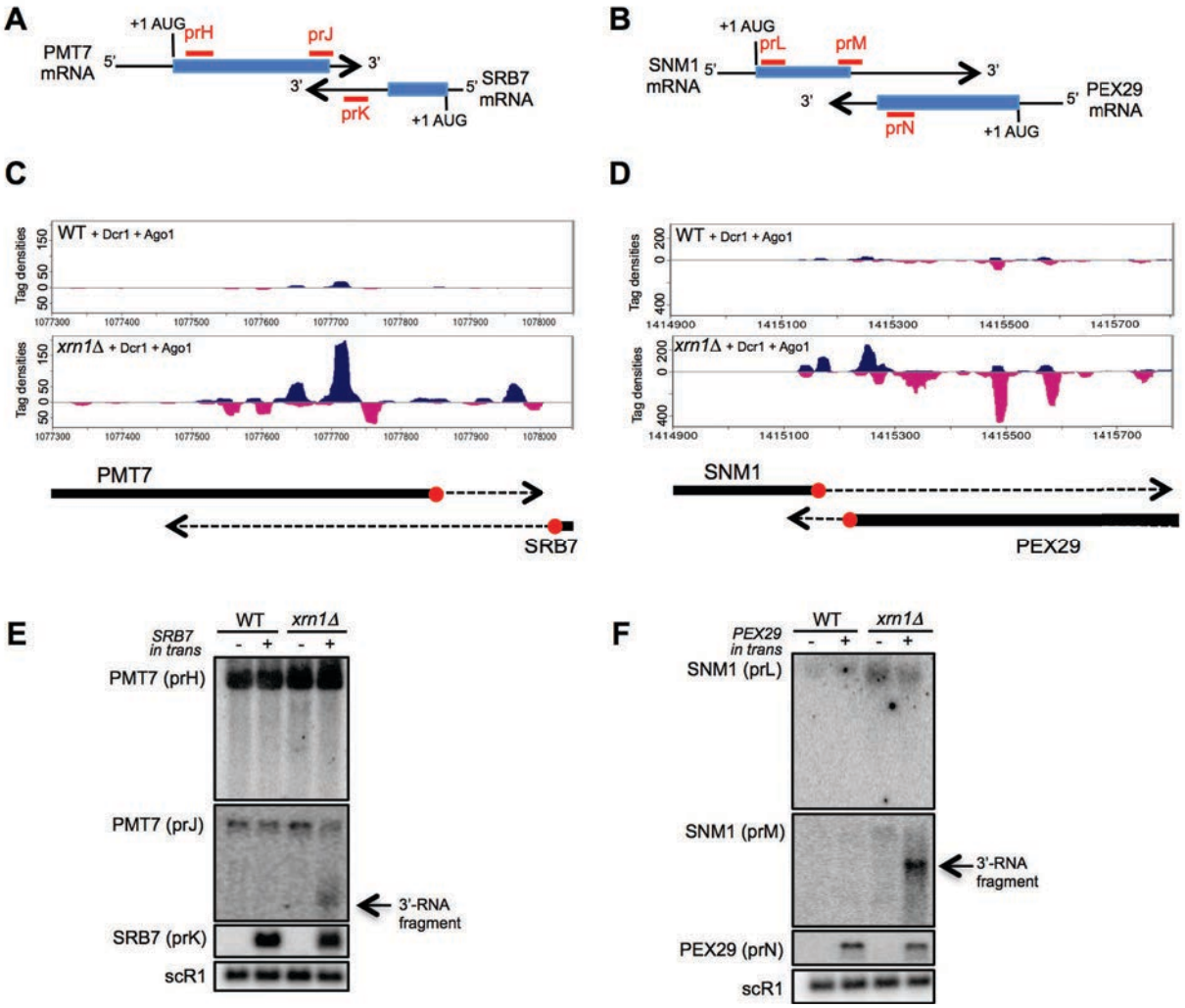

Figure 55. Validation of new messenger-interacting messenger RNA (mimRNA) pairs.

(A) Schematic of the convergent PMT7 and SRB7 mRNA pair (YDR307W and YDR308C, respectively). These convergent mRNAs overlap and were shown to produce small RNA enrichment in the presence of Dcr1/Ago1. Filled boxes indicate open reading frames, and black arrows represent 3' untranslated regions of mRNAs. Specific probes prH, prJ, and prK used in northern blots are indicated in red.

(B) Schematic of convergent SNM1 and PEX29 mRNA pairs (YDR478W and YDR479C, respectively). Specific probes prL, prM, and prN used in northern blots are indicated in red.

(C) Snapshot of the small RNA accumulation of 19–23 nucleotides within the *PMT7-SRB7* region in wild-type and *xrn1* mutant strains in the presence of Dcr1/Ago1. Small RNA reads originating from Watson strand are shown in blue, and small RNA reads originating from Crick strand are shown in red. Read densities are in linear scale. Chromosome 4 coordinates are indicated below. Dashed arrows represent the UTR regions of PMT7 and SRB7 mRNAs, and filled boxes indicate ORFs. Translational stop codons are indicated by red circles.

(D) Snapshot as in (C) of the small RNA accumulation of 19–23 nucleotides within the *SNM1-PEX29* region.

(E) Northern blot analysis to detect the 3' RNA fragment of PMT7 mRNA when SRB7 mRNA is expressed *in trans* (+) or not (-).

(F) Northern blot analysis to detect the 3' RNA fragment of SNM1 mRNA when PEX29 mRNA is expressed *in trans* (+) or not (-).

7 | Cytoplasmic 5' RNA decay factors are downregulated in *xrn1*Δ cells

As discussed above, we detected multiple interactions between sense-antisense mRNA pairs genome-wide, and thus we were particularly intrigued by the fact that the key factors in the cytoplasmic 5'-3' RNA decay pathway—Xrn1, along the members of Dcp1-Dcp2 complex—share the convergent orientation in yeast genome. Similarly to the auto-regulation of nuclear RNAi factors in *S. pombe* that are predominantly convergent genes (*162*), we speculated that 5'-3' RNA decay factors might also be subjected to auto-regulation via mRNA-mRNA interactions. To explore this idea, we decided to use translational fusion reporters of *XRN1*, *DCP2* and *DCP1* genes (Figure 56A). Technically, we constructed three plasmids, containing the entire region of the convergent gene pair (*XRN1-BUD13*, *DCP2-NCS2*, and *DCP1-SPT20*), and replaced the ORF sequence of the gene of interest by the Flag-tagged GFP ORF. We then transformed isogenic WT and *xrn1*Δ strains with these plasmids (pLB177, pLB178 and pLB179, respectively) and examined the expression of the Flag-GFP reporter by western blot. Of notice, the *XRN1::Flag-GFP* construction allowed us to address the expression pattern of *XRN1* reporter in the absence of Xrn1. As shown in Figure 56B, all three reporters were downregulated in *xrn1*Δ cells, similarly to our observation of the decreased production of Por1 protein in an analogous context. Importantly, this effect was not a general translational defect of *xrn1* mutant as another convergent gene reporter, *DED1::Flag-GFP*, was comparably produced in both WT and *xrn1*Δ cells. These findings suggest that cytoplasmic 5'-3' RNA decay factors might undergo a negative feed-back regulation rather than auto-regulation.

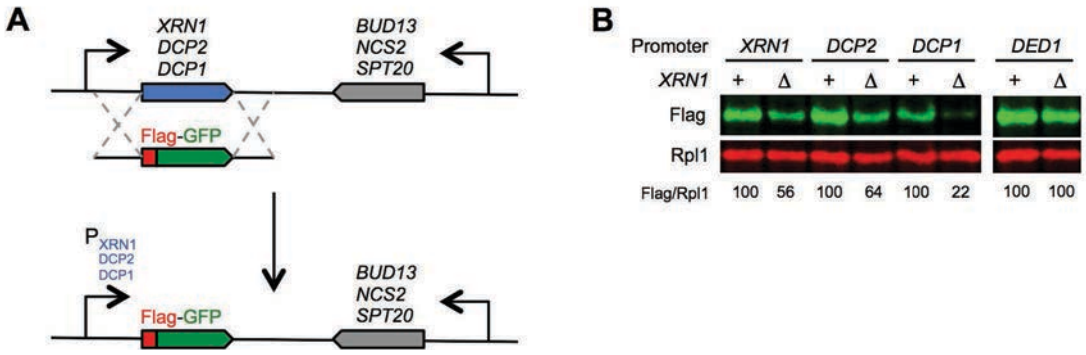

Figure 56. The expression of translational fusion reporters of *XRN1*, *DCP2* and *DCP1* genes.

(A) The schematic of translational fusion reporters, used to determine the expression of cytoplasmic 5'-3' RNA decay factors (Xrn1, Dcp2 and Dcp1). The genomic regions of *XRN1-BUD13*, *DCP2-NCS2*, and *DCP1-SPT20* were cloned in centromeric plasmids (upper panel). Then, the ORFs of the genes of interest were replaced by Flag-tagged GFP sequence, preserving natural promoter, 5' and 3' UTR sequences, as well as the convergent genes (lower panel).

(B) Western blot analysis of Flag-tagged GFP reporter, expressed under *XRN1*, *DCP2*, or *DCP1* promoter, in wild-type and *xrn1*Δ cells. The same reporter under *DED1* promoter and in convergent orientation with *GEP3* was used as a control. Rpl1 served as a loading control. Quantifications (from two independent experiments) of Flag to Rpl1 ratio are indicated below the blots.

CHAPTER III | DISCUSSION AND PERSPECTIVES

Outline

The findings of this work are discussed following the two major axes that also structured the presentation of results (Chapter II). Firstly, the observations of the detailed study of *POR1* and *OCA2* expression are contextualized and compared with the currently available data from the literature. The post-transcriptional nature of the interaction between POR1 and OCA2 mRNAs is emphasized, followed by the focus on two distinct outcomes of this interaction—remodeled stability and translation of POR1 mRNA. The second part of the chapter aims to summarize the observations of the analysis of small RNA sequencing data from RNAi-competent *S. cerevisiae*. The role of Xrn1, a conserved cytoplasmic 5'-3' exoribonuclease, is discussed, underlining its implication in controlling the extent of sense-antisense mRNA interactions. A particular attention is drawn to the potential physiological implications of mRNA-mRNA interaction genome-wide, evoking the interplay between transcriptional and post-transcriptional control mechanisms. Finally, the conclusions of this work are presented.

PART I | Interaction of POR1 and OCA2 mRNAs in the cytoplasm

1 | Post-transcriptional outcomes of convergent transcription

The major part of this work has been focused on the expression of *POR1* gene in *Saccharomyces cerevisiae*, the transcription of which produces an mRNA overlapping with its convergent partner, OCA2 mRNA. A number of results, presented in the Chapter II, permitted us to favor the hypothesis that an antisense mRNA—OCA2—affects post-transcriptionally the expression of a sense mRNA, POR1. Firstly, in the context where OCA2 mRNA is highly stabilized (*xrn1* mutant), the quantity of POR1 mRNA remains equivalent to the wild-type level while the production of Por1 protein is compromised. Secondly, OCA2 mRNA—as well as its non-coding portion complementary to POR1 mRNA—affects the expression of Por1 protein *in trans*. Thirdly, the accumulation of OCA2 mRNA leads to a detection of no-go decay intermediates of POR1 mRNA, dependent on NGD factors Dom34 and Hbs1, pointing out to the cytoplasmic and translation-associated processes involving these two overlapping mRNAs.

These results incited us to explore the idea of a base-pairing dependent interaction between sense-antisense mRNA pairs in the cytoplasm. Several points of evidence support this hypothesis. Importantly, the sequence complementarity between two mRNAs is essential for triggering NGD, a signature of the interaction between overlapping transcripts. In accordance to the direct link of NGD to translation, preventing ribosomes from entering the region of overlap between two mRNAs suppresses the translational defect and the accompanying detection of NGD intermediate. Furthermore, in RNAi-competent *S. cerevisiae* cells, siRNAs corresponding to the region of overlap between POR1 and OCA2 mRNAs are readily detected *in vivo*, contributing to the evidence that two partially complementary messengers can exist within the same cell and interact in the cytoplasm. In parallel, several results support the functional importance of the interaction between POR1 and OCA2 mRNA pair, as in the absence of *OCA2*, the stability—and, concomitantly, the expression—of POR1 mRNA is compromised.

These results converge to an alternative, post-transcriptional model for the antisense transcription mediated gene regulation. Indeed, our initial results could not be explained by the Pol II collision model (*144*), or the repressive chromatin remodeling (*64*) that addressed the outcomes of convergent transcription; as described above, the repression at the protein level of Por1 was not preceded by the transcriptional downregulation of POR1 mRNA. Rather, our findings echoed the examples of antisense RNA-mediated translational repression coming from studies of metazoan organisms. As evoked in the Chapter I, antisense RNAs to TNFRSF17 (*191*) or PU.1 (*192*) negatively affected the translation of their sense counterparts, without changing the steady-state levels of the sense mRNA. However, in these two cases the antisense RNAs were non-coding transcripts, and their genetic disposition exhibited important differences when compared to POR1 and OCA2 mRNAs (both TNFRSF17 AS and PU.1 AS originate from the coding region of TNRRSF17 and PU.1, respectively, and are not pairs of convergent transcription units as defined for the coding genes of *S. cerevisiae*). At the molecular level, PU.1 AS RNA was suggested to inhibit the translation initiation of PU.1 mRNA, as the overlap between two RNAs extended within the 5' UTR of PU.1; this was in contrast to the 3' proximal overlap between POR1 and OCA2 mRNAs and the comparable presence of POR1 mRNA in the polysomal fraction, in both WT and OCA2 accumulating *xrn1*Δ contexts.

Interestingly, while this work was in progress, one study addressed the fate of sense-antisense mRNA pairs in *S. cerevisiae* within the framework of transcriptional interference (*201*), a phenomenon usually referring to the antisense transcription mediated chromatin remodeling. Analyzing the available RNA-seq and NET-seq (native elongating transcript sequencing) data, Wang *et al.* found that 15% of convergent mRNA pairs showed anti-correlated expression pattern (termed "anti-regulation"), proposing transcriptional interference as the working model to explain this result. The further detailed examination of selected convergent gene pairs revealed that antisense mRNA could downregulate the sense transcript ectopically and, moreover, this effect was strictly dependent of the 3' UTR complementarity between the two mRNAs. Importantly, the pairs analyzed by Wang *et al.* belong to the category of 3' UTR and ORF overlapping transcripts, as defined in our analysis described in the Chapter II, and thus could potentially interfere with each other's translation, as detected for POR1 and OCA2 mRNAs. However, the authors of the

study preconized the transcriptional repression model, commenting that their findings reported the first case of ectopic transcriptional interference in *S. cerevisiae*, albeit with an unknown mechanism.

We have observed the downregulation of Por1 protein in the context of the defective cytoplasmic 5'-3' RNA decay pathway and associated it with the accumulation of OCA2 mRNA. We have reasoned that OCA2 transcript, stabilized in *xrn1*Δ cells, corresponds to a 5' decay intermediate as it is mainly 5'-monophosphorylated. Interestingly, several studies have recently attributed a nuclear role for Xrn1, traditionally described as a canonical cytoplasmic RNA decay factor. Conflicting reports exist whether Xrn1 contribution to transcriptional regulation is direct (*67*) or indirect (*68*), however the general line of evidences point out the Xrn1-mediated crosstalk between mRNA decay and synthesis rates. Importantly, at least in the case of OCA2 transcription, several observations lead us to exclude the derepression of mRNA synthesis in the absence of Xrn1. Firstly, Pol II occupancy over *OCA2* locus did not increase upon *XRN1* deletion, as revealed by Pol II-associated chromatin immunoprecipitation and DNA sequencing (ChIP-seq) analysis (*64*). Secondly, in wild-type and *xrn1*Δ or *xrn1* catalytic mutant cells, *OCA2* transcription rate remained unchanged, as deduced from the *OCA2* corresponding NET-seq signal (personal communication from A. Morillon). Thirdly, inhibiting Xrn1 by lithium ions (via accumulation of pAp, direct inhibitor of Xrn1) was sufficient to detect OCA2 mRNA-induced NGD cleavage of POR1 mRNA in wild-type cells, arguing that the defective catalytic 5'-3' RNA decay activity in the cytoplasm is essential for stabilizing OCA2 mRNA post-transcriptionally. In sum, our analysis of *POR1* and *OCA2* expression revealed a new type of post-transcriptional outcomes from the transcription of convergent gene loci.

2 | Remodeling mRNA stability by sense-antisense mRNA interaction

Our initial analysis has revealed that, in contrast to OCA2 mRNA accumulating in *xrn1*Δ cells, POR1 mRNA quantity remains stable in both wild-type and *xrn1* mutant contexts. Importantly, this feature is dependent on OCA2 mRNA, as in *oca2*Δ cells, POR1 mRNA level decreases approximately two-fold as compared

to wild-type. We have demonstrated that the interaction between POR1 and OCA2 mRNAs has a stability-remodeling effect on POR1 mRNA. Ectopically expressing OCA2 mRNA leads to a decreased decay rate of POR1 mRNA—in both WT and *xrn1*Δ cells, suggesting that 3'-5' RNA decay pathway is involved in this phenomenon. Indeed, in the absence of OCA2 mRNA, POR1 mRNA undergoes increased attack by the cytoplasmic deadenylase and exosome complexes, as deduced from the recovered steady-state levels and decay rates of POR1 mRNA in *oca2*Δ *ccr4*Δ and *oca2*Δ *ski2*Δ cells, in comparison to *oca2* mutant. These findings can be related to the reports from metazoan organisms, discussed in the Chapter I. Similarly to OCA2 mRNA, non-coding antisense RNAs iNOS AS (*176*) and FGFR3 AS1 (*178*) stabilize their respective sense counterparts, iNOS and FGFR3 mRNAs. Interestingly, a human sense-antisense RNA pair where both transcripts are coding, DHPS-WDR83, mutually act on each-other's stability (*179*); mechanistically, it is the closest example to the stabilizing interaction between POR1 and OCA2 mRNAs, suggesting that convergent gene orientation might account for the downstream stabilization of the respective transcripts in RNAi competent organisms. Importantly, our results provide the molecular details for RNA interaction-mediated inhibition of the cytoplasmic exosome attack, as previously hypothesized by the study addressing bidirectional stabilization of DHPS and WDR83 mRNAs (*179*).

In the case of iNOS mRNA stabilization by its antisense transcript iNOS AS, a trans-acting factor, HuR, is recruited to the AU-rich element present within 3' UTRs of both RNAs (*176*). HuR is a well-known competitor of destabilizing ARE-BPs, increasing thus the stability of its targets (*73*). In *S. cerevisiae*, two DExD/H family RNA helicases have been associated with the stability of sense-antisense transcript pairs. In the case of mRNA overlapping antisense XUTs, protruding 3' extensions of mRNA-XUT duplexes are recognized by Mtr4 and Dbp2 helicases, and the unwound form of the ncRNA is then targeted by the NMD pathway (*190*). It is thus tempting to speculate that 3' overlapping mRNA pairs can recruit dsRNA specific factors that would in turn contribute to remodeling its stability and/or translation. RNA helicase activity-containing proteins appear as most promising candidates in regulating mRNA-mRNA interaction *in trans*. However, at least in the case of Mtr4 and Dbp2, several evidences argue against the implication of these helicases in the convergent mRNA related processes. Importantly, Mtr4 and Dbp2 are

predominantly nuclear factors, and their activity is greatly enhanced by the 3' overhangs in the dsRNA substrate (*216, 217*). In contrast, cytoplasmic mRNA-mRNA interactions yield 3' overlapping and thus 5' protruding duplexes, unfavorable for recruiting these helicases.

Unlike OCA2 mRNA, accumulating in *xrn1*Δ cells in the form of a decapped 5' decay intermediate, POR1 mRNA contains 5' and 3' functional prerequisites in both wild-type and *xrn1* mutant contexts. This feature, similarly to the functional stabilization of POR1 mRNA discussed above, is also dependent on the presence of OCA2 mRNA. The deletion of *XRN1* in *oca2*Δ cells increases the quantity of steady-state POR1 mRNA, in contrast to the comparison between POR1 mRNA levels in *XRN1* and *xrn1* cells. This suggests that in the absence of OCA2 mRNA, POR1 mRNA becomes also a target of Xrn1, presupposing that POR1 messenger is accessible to the cytoplasmic 5'-3' RNA decay pathway. Indeed, after a treatment by Xrn1 *in vitro*, the resistant fraction of POR1 mRNA becomes comparable between *oca2*Δ and *oca2*Δ *xrn1*Δ mutants, witnessing the stabilization of a decapped 5' decay intermediate in the absence of Xrn1. This observation is compatible with similarly defective production of Por1 protein in both *oca2*Δ and *oca2*Δ *xrn1*Δ contexts.

The differential accessibility of POR1 mRNA to the cytoplasmic 5'-3' RNA decay pathway depending on the presence of OCA2 mRNA prompted us hypothesize that the single-stranded versus double-stranded 3' end state might be sensed by a decapping-related factor. Interestingly, the heptameric Lsm1-7 complex is known to bind oligoadenylated, preferentially single-stranded mRNA 3' ends and trigger the 5' decapping via its interaction with Pat1 and other decapping activators (*61*). It is plausible that the interaction between POR1 and OCA2 mRNAs not only prevents the cytoplasmic exosome attack at the 3' end but also limits the association of the Lsm1-7 complex, additionally stabilizing POR1 messenger at the 5' extremity.

In contrast to an apparent stabilization of POR1 mRNA by OCA2 mRNA at the 5' end, OCA2 mRNA remains in a 5'-3' decay pathway accessible form, as it is highly stabilized in *xrn1* mutant. This suggests that OCA2 transcript undergoes deadenylation independent decapping, even when POR1 mRNA is highly abundant; in accordance, we have not detected an increase in the steady-state OCA2 mRNA levels by inactivating *CCR4*, a major cytoplasmic deadenylase, or deleting *SKI2*, compromising the downstream attack by the cytoplasmic exosome. In addition,

removing POR1 mRNA did not change the pattern of OCA2 accumulation in the absence of Xrn1. Interestingly, OCA2 mRNA contains a three-codon upstream ORF (uORF) overlapping the main start codon, a feature that has been previously associated with an increased sensitivity to the NMD pathway (*218*). Indeed, in NMD-defective (*upf1*Δ) cells, OCA2 mRNA shows a two-fold increase in the steady-state quantity, as detected by a genome-wide study (*190*). It is thus possible that targeting OCA2 mRNA for the NMD-induced decapping is an inherent property of *OCA2* expression.

In parallel to the functional stabilization of POR1 mRNA by OCA2 mRNA in wild-type cells, increased accumulation of OCA2 mRNA in the absence of Xrn1 triggers the NGD of POR1 mRNA. To our knowledge, this is the first report of a trans-acting RNA induced NGD cleavage, complementing the large panel of signals leading to the ribosomal stall and aberrant mRNA decay *in cis* (*83*). Importantly, we have described the prerequisites of sequence complementarity between overlapping transcripts, essential for triggering NGD. The analysis of OCA2 and POR1 mRNAs revealed that the complementarity extending to the 3' proximal region of *POR1* ORF is essential for targeting POR1 transcript to the NGD pathway, conformingly to its intrinsic dependence of translation. However, the sole complementarity within the coding region of POR1 mRNA between two transcripts is not sufficient to trigger NGD, arguing that the overlap between the 3' UTRs is equally necessary for the interaction. It is possible that 3' UTRs, unoccupied by ribosomes, are more prone to engage in the RNA-RNA interaction and serve as a landing pad for subsequently invading the 3' proximal coding region.

Interestingly, the coding potential of the interacting RNA partners appears to be important for triggering NGD, as revealed by differential outcomes of the interaction between POR1 3' UTR mutant mRNA and RNA3 or RNA4 (Figure 40). POR1 3' UTR mutant mRNA and RNA3 mirror the complete complementarity between POR1 and OCA2 mRNAs, and thus trigger NGD. RNA4, similarly to RNA3, has the region of overlap with its partner mRNA intact, but it exhibits a displaced translational stop codon, limiting the ribosome access to the region of complementarity with POR1 3' UTR mutant. Surprisingly, this disposition does not lead to NGD of the interacting partner; it also stabilizes RNA4, ruling out the possibility that this transcript could be targeted by the NMD pathway, although it

formally contains a PTC. Indeed, RNA4 preserves a relatively long ORF upstream its PTC-prolonged 3' UTR, a feature associated with the evasion from the NMD (*219*). We interpret this result suggesting that an artificial lengthening of the 3' UTR of RNA3 (i.e. creating RNA4 by un upstream PTC) favors a structural remodeling of the ribosome devoid 3' UTR of RNA4, subsequently lowering its interaction potential towards the complementary mRNA. Additional experiments are needed to address this possibility, however it raises a fundamental question regarding convergent transcription derived RNA-RNA interactions, covering by definition the 3' untranslated regions and possibly extending within the 3' proximal coding sequences. In addition to convergent coding gene pairs, a multitude of non-coding transcripts contains 3' overlaps with mRNAs (*64, 190*). It is thus highly plausible that, having all mRNA-like prerequisites, these antisense ncRNAs could interact with their sense mRNAs in the cytoplasm, similarly to OCA2 and POR1 mRNAs. Importantly, the absence of the coding potential and, concomitantly, low ribosome coverage might favor the secondary structure formation in these ncRNA, counteracting their intermolecular RNA-RNA interaction.

3 | Remodeling mRNA translation by sense-antisense mRNA interaction

We have demonstrated that the interaction between POR1 and OCA2 mRNAs has an impact on POR1 mRNA translation. Initially, we have associated the accumulation of OCA2 mRNA in the absence of Xrn1 with the decreased production of Por1 protein; importantly, this effect is valid not only *in cis*, but also *in trans*, as ectopic expression of OCA2 mRNA additionally downregulates Por1 protein. Further on, we have detected that the interaction between the POR1-OCA2 mRNA pair triggers the NGD of POR1 messenger, confirming the translational mechanism of the phenomenon. Summing up our findings, we have termed this interaction "translational interference", underlining the fact that the region of overlap between two messengers (extending into the coding part of at least one of the interacting partners) might interfere with elongating ribosomes. In accordance, preventing ribosomes from entering the region of interaction between two mRNAs (by displacing the translation termination codon upstream the overlapping sequence)

suppresses the translational interference and avoids identifying the interacting mRNA as an aberrant transcript, subsequently targeted for NGD (*207*).

Similarly to synthetic reporters described in the literature (PGK1 mRNA containing a stem-loop structure within its ORF (*207*), termed PGK1-SL, or GFP mRNA followed by a stretch of rare arginine codons (*208*), designated GFP-Rare, among others), the production of the 3' NGD intermediate of POR1 mRNA is dependent on ribosome rescuing factors, Dom34 and Hbs1; the deletion of either of these factors decrease the steady-state levels of the 3' NGD fragment. In accordance to a widely proposed model, impairing stalled ribosome dissociation favors the coverage of an aberrant mRNA with blocked ribosomes, severely decreasing the number of cycles the messenger is fully translated. In result, the cleavage efficiency is lowered, as translation remains inactive. In parallel, *dom34* mutants show similarly deficient protein output from an aberrant mRNA (*208*), and, accordingly, we have detected equally downregulated levels of Por1 protein in both *xrn1*Δ and *xrn1*Δ *dom34*Δ cells.

As evoked previously, OCA2 mRNA-induced NGD targets a relatively low proportion of POR1 mRNA, as compared to a several-fold downregulation of Por1 protein in OCA2 mRNA accumulating *xrn1*Δ cells. It is plausible that stalled ribosomes on POR1 mRNA are not efficiently recognized and rescued by Dom34-Hbs1 complex, explaining its modest endonucleolytic cleavage. In fact, Dom34-Hbs1 efficiently recognizes ribosomes stalled on NSD substrates—stop codon-less mRNAs—exhibiting an empty ribosomal A site at the 3' extremity of an aberrant transcript (*208*). However, NGD substrates (e.g. mRNAs containing rare codons) have been already reported to possess lower affinity to Dom34-Hbs1 complex (*220*). It has been speculated that the access of Dom34-Hbs1 to the ribosomal A site is restricted when the ribosome is stalled on a sense codon (*83*), as would be the case of interacting POR1 and OCA2 mRNAs. Moreover, in reconstituted yeast and mammalian translation systems *in vitro*, Dom34-Hbs1-Rli1 efficiently dissociates ribosomes stalled at the 3' extremity of model substrates, and increasing length of the 3' protruding end downstream of the stalled ribosome leads to decreasing efficiency of stalled ribosome rescue, an effect dependent on Hbs1 (*221, 222*). Analyzing the structural data of ribosome-interacting Dom34-Hbs1, the N-terminal domain of

Hbs1 appears to localize in the vicinity of mRNA exit tunnel of 80S, suggesting that the presence of protruding mRNA could interfere with Hbs1 recruitment (*222*).

Accordingly to a low efficiency of POR1 mRNA targeting by Dom34-Hbs1 complex, Por1 nascent protein is not affected by the impaired RQC complex: deleting *LTN1* failed to augment the level of Por1 protein in *xrn1*Δ cells. Indeed, RQC complex is selectively recruited on peptidyl-tRNA-containing 60S subunits, as Ltn1, the E3 ubiquitin ligase of the RQC complex, shares the interaction surface on 60S with 40S ribosomal subunit (*83*). We have not either detected a peptidyl-tRNA form of Por1 nascent chain, similarly as no peptidyl-tRNA intermediate from NGD substrates have been reported to our knowledge. However, canonical NGD substrates (e.g. containing rare codons) are actively targeted by RQC complex (*210*), suggesting the existence of an alternative mechanism for stalled ribosome dissociation. Intriguingly, a recent study has reported that canonical translation termination factors eRF1-eRF3 also generate substrates for RQC-mediated proteasomal degradation (*223*).

The translational interference that we observed upon OCA2 and POR1 mRNA interaction was also unaffected by *HEL2* deletion. Hel2, another RING domain E3 ubiquitin ligase implicated in co-translational protein surveillance, is essential for stimulating ribosome stalling on NGD substrates (although dispensable for the decay of non-stop nascent peptides) (*84, 210*). Even though Hel2 acts upstream of the stalled ribosome dissociation, it is evident that ribosomes stalled at OCA2 mRNA-interacting POR1 messenger follow an alternative, Hel2- and Ltn1-independent route in downregulating Por1 production. Asc1 (known as RACK1 in metazoans), a core component of 40S ribosomal subunit, has been equally implicated in stimulating ribosome stalling on NDG substrates, acting upstream of Dom34-Hbs1 (*210*). Recently, Asc1 has been shown to significantly stimulate the endonucleolytic cleavage of NSD substrates in yeast (*224*). It is thus possible that this factor might also affect the ribosome stalling on POR1 mRNA in *xrn1* mutant. However, as to date we have not identified a factor that could suppress the defect of Por1 production in *xrn1*Δ cells, we preconize the model where elongating ribosomes slow down or stall upon the region of interaction between POR1 and OCA2 mRNAs. The Dom34-Hbs1 complex dissociates a fraction of stalled ribosomes, albeit inefficiently as the ribosomal A site is occupied by a non-rare cognate tRNA, and the 3' end of POR1

mRNA counteracts the Hbs1 recruitment. Accordingly, we postulate that the main outcome of translational interference between POR1 and OCA2 mRNAs is a decreased rate of POR1 mRNA translation, as the translation termination at the stop codon, embedded within the mRNA-mRNA duplex, becomes rate-limiting.

Although recycling factors Dom34-Hbs1-Rli1 dissociate stalled ribosomes on artificial aberrant mRNA reporters, endogenous targets of these complexes remain largely unknown. Recently, two studies have addressed the genome-wide effect of *dom34* and *rli1* mutations, using ribosome profiling technique (*213, 225*). The absence of Dom34 efficiently stabilizes the unspliced HAC1 mRNA, a well-known actor of unfolded protein response that undergoes non-canonical cytoplasmic splicing upon endoplasmic reticulum stress. This messenger is the sole evident target of Dom34 *in vivo*; surprisingly, the *dom34*Δ cells also accumulate ribosomes on the 3' UTRs of a subset of mRNAs (*213*). Similarly, ribosome accumulation on 3' UTRs is even more wide-spread in Rli1-depleted cells, accordingly to the role of Rli1 in a canonical ribosome recycling (*225*). We have been particularly intrigued to learn that POR1 mRNA has the highest *dom34*Δ to WT ratio of the 3' UTR ribosome density, as revealed by Guydosh and Green. These 3' UTR ribosomes are reported to be translationally inactive in the *dom34* mutant (in contrast to ribosomes capable to resume translation in Rli1-depleted cells), and we speculated that ribosome presence within the 3' UTR of POR1 mRNA might protect it from the 3'-5' exonucleolytic attack. However, deleting *DOM34* does not suppress the downregulation of Por1 protein in *oca2*Δ cells, the context where POR1 mRNA is particularly sensitive to the cytoplasmic exosome-mediated decay. It is possible that the ribosome passage through the stop codon in *dom34*Δ cells is of low efficiency, and thus insufficient to significantly stabilize POR1 mRNA. Alternatively, we cannot rule out that the interaction between POR1 and OCA2 mRNAs is essential for the Dom34-dependent ribosome association to 3' UTRs. Indeed, a significant portion of mRNAs, detected to contain ribosomes on their 3' UTRs in the absence of Dom34, result from convergent—coding as well as non-coding—loci (82% of transcripts showing ≥10 rpkm signal of ribosome density on their 3' UTRs) (Figure 57). Moreover, the authors of the study show that ribosomes on 3' UTRs are dependent on the 3' UTR sequence, and altering this sequence abolishes the detection of associated ribosomes in *dom34*Δ cells. Interestingly, these alterations were constructed

by replacing an antisense overlap-containing 3' UTR with an antisense transcription-devoid sequence, further supporting our hypothesis. It is thus tempting to speculate that translational interference between sense-antisense mRNA pairs might induce ribosome passage through the stop codon, possibly including frame-shifting (as it is the case of ribosomes translating 3' UTRs in the *rli1* mutant).

Gene	Signal in *dom34* mutant, rpkm	Genomic context	Gene	Signal in *dom34* mutant, rpkm	Genomic context
CPR5	330,5	XUT	*NOP56*	18,3	CONV
POR1	140,5	CONV	*BMH1*	16,0	SOLO
RPL18A	109,2	SOLO	*CCT6*	15,7	SOLO
RIB4	78,5	CONV	*RPO21*	14,6	CONV
DPM1	71,3	CUT/NUT	*CHA1*	14,5	SOLO
SGE1	62,9	SUT/XUT	*DDP1*	14,3	CONV
MFA2	57,9	NUT	*URA7*	14,3	CONV
VMA4	48,0	CUT	*SMX2*	14,2	CONV
SSA1	42,9	SOLO	*BAT1*	14,2	XUT/NUT
APA1	41,4	CONV	*RPT2*	14,0	XUT
RPL35A	39,6	CONV	*YNL208W*	13,7	XUT
TRX1	38,2	XUT/NUT	*DRS2*	13,4	CONV
RPS19A	36,7	XUT	*MGT1*	13,3	CONV
PCS60	35,7	CONV	*TFC7*	12,9	CUT/NUT
HUB1	28,1	CONV	*HAS1*	12,7	CONV
RPS14B	25,9	CONV	*RPE1*	12,6	CONV
RRS1	24,9	SOLO	*LAT1*	12,3	XUT
ACB1	24,6	SUT/XUT	*RNR2*	12,1	XUT/NUT
RPL42A	24,6	NUT	*ZUO1*	11,9	NUT
VOA1	24,2	CONV	*LYS4*	11,9	SOLO
RPL14B	24,2	SUT/XUT	*YIP5*	11,4	CONV
DUG1	23,2	SOLO	*OST5*	11,3	CONV
ERG10	21,1	SUT	*YLR179C*	10,8	XUT
SNU13	18,9	CONV	*SPT15*	10,1	CONV
ASP1	18,5	SOLO	*YPR1*	10,0	CONV

Figure 57. Dom34 rescues ribosomes from 3' UTRs of convergent mRNAs.

Fifty genes, mRNAs of which show ≥10 rpkm ribosome protected fragment density within the 3' UTR in *dom34*Δ cells, are listed. The genomic context of these genes is also indicated: coding for mRNAs without no detectible antisense transcription (SOLO, light blue); coding for convergent mRNAs (CONV, red); and coding for mRNAs with non-coding antisense transcription (light red), as described in (*190*). Non-coding transcription units are abbreviated as follows: XUT, Xrn1-sensitive unstable transcripts; CUT, cryptic unstable transcript; SUT, stable unannotated transcript; NUT, Nrd1-unterminated transcript. Data retrieved from (*213*).

In accordance to majorly 5'-monophosphorylated form of OCA2 mRNA in *xrn1*Δ cells, we postulated that this transcript is a stabilized 5' decay intermediate, inherently deficient for translation initiation. We also speculated that particularly this form of mRNA, showing supposedly lower ribosome occupancy, would be more prone to engage in RNA-RNA interactions. However, in wild-type as well as *xrn1*Δ

cells, we have detected OCA2 mRNA associated with polysomes, independently of its 5' status. This observation is compatible with the recently discovered wide-spread co-translational mRNA decay, suggesting that OCA2 mRNA undergoes decapping in both *XRN1* and *xrn1* contexts while still associated to ribosomes. In parallel, we have speculated that the presence of OCA2 mRNA in polysomal fractions might be explained by its interaction with POR1 mRNA, which in turn would indirectly cause OCA2 mRNA detection in the bottom of density gradient, independently of its translation. Further experimental data is needed to address the both possibilities in more details.

PART II | Genome-wide interaction of sense-antisense mRNAs in the cytoplasm

1 | siRNA signature of overlapping mRNAs in RNAi-competent *Saccharomyces cerevisiae*

Our experiments combining wild-type and 3' UTR mutant POR1 mRNAs with their differentially overlapping counterparts RNA1, RNA2 and RNA3 illustrate that the phenomena discussed in the Part I of this chapter extend outside the example of interaction between POR1 and OCA2 mRNAs. We have thus chosen the reconstituted RNAi approach to examine the dsRNA prevalence originating from the convergent gene loci genome-wide. Initially, we have confirmed that upon *S. castellii* Dicer and Argonaute introduction in *S. cerevisiae* cells, siRNAs corresponding to the intergenic region between *POR1* and *OCA2* can be detected by small RNA northern blot and small RNA sequencing. We have noticed that siRNAs originating from the OCA2 strand are significantly more prevalent than those of the complementary POR1 strand. We have attributed this bias to the intrinsic properties of the Argonaute protein, which is known to selectively stabilize one of the two Dicer-produced siRNA strands, although the details of this selection remain largely unknown (*226*).

The mapping of siRNA reads from RNAi-competent *S. cerevisiae* cells has revealed that convergent mRNA pairs yield a significantly stronger siRNA signal as compared to messengers with no apparent antisense transcript. Importantly, this effect was dependent on the RNAi factors, as *S. cerevisiae* strains without Dcr1/Ago1 failed to show the convergent versus solo mRNA enrichment pattern. Our analysis of 3' UTR only and 3' UTR and ORF overlapping mRNA pairs has also revealed that the latter category of transcripts shows an additional enrichment in siRNA signal. It is thus possible that messengers that undergo translational interference might engage in more stable interactions; in parallel, the accompanying NGD cleavage could contribute in generating dsRNA substrates for Dicer and account for an increased RNAi signal. Accordingly, we have detected the NGD cleavage originating from two

3' UTR and ORF overlapping mRNA pairs selected from our analysis, suggesting that translational interference is a recurrent phenomenon *in vivo*.

In previously reported analyses, overlapping mRNAs originating from convergent gene loci accounted for 20% of total coding transcripts in *S. cerevisiae* (*201*), or 21% in *S. castellii* (*51*). In our set of 4013 mRNAs, convergent and overlapping (for at least 23 nucleotides to allow for Dicer cleavage) messengers constitute 36% of the total mRNAs. Our preselected pool of mRNAs is devoid of overlapping coding and non-coding RNA pairs, which were presumably included in above-mentioned references, underlining the prevalence of the overlapping mRNA-mRNA pairs. Interestingly, in *S. castellii*—containing an endogenous RNAi pathway—43% of overlapping mRNA pairs yield siRNA signal (*51*), similarly to 34% of convergent mRNAs showing higher than two-fold enrichment in siRNA signal upon RNAi reconstitution in *S. cerevisiae*. Moreover, the deletion *AGO1* or *DCR1* in *S. castellii* does not significantly change its transcriptomic landscape (with the two exceptions of highly stabilized Y' element RNA and inverted Ty-like repeat-containing RNA), nor yield a detectable growth defect (*51*); similarly, Dcr1/Ago1 containing *S. cerevisiae* appear to grow normally. These findings incite us to think that while siRNA originating from overlapping mRNAs serve as a valuable signature mark for mRNA-mRNA interactions, the mechanistic implications of these interactions extend outside the RNAi machinery and thus are relevant in both RNAi positive and negative organisms.

We have detected that 19% of convergent and overlapping mRNAs show higher than two-fold enrichment in siRNA signal originating from the both transcripts of the pair in RNAi-positive *S. cerevisiae* cells. Remarkably, a recent study in human iESCs has reported a functional clustering of interacting mRNA-mRNA pairs, although these interactions are not specified to be originating from convergent gene loci (*90*). We have thus hypothesized that an implication in functionally related cellular processes might accompany interacting convergent mRNA pairs. We have determined the gene ontology (GO) terms of 84 convergent mRNA pairs (relative to Figure 53D) yielding the highest siRNA signal in RNAi-competent cells. As shown in Table 2, this subset of mRNAs is modestly but significantly enriched in nucleus- and organelle-associated genes, suggesting that these cellular compartments might be sensitive to perturbations in mRNA-mRNA interactions. Of notice, from the two

pairs of interacting mRNAs, for which we validated the interaction by observing the translational interference-dependent NGD cleavage, each contains a transcript coding an endomembrane located protein. Pex29 is a peroxin, associated to the ER membrane and involved in the biogenesis of peroxisomes (*227*), and Pmt7 is a putative mannosyltransferase, similar to Pmt1 and located in the ER (*228*). Interestingly, it has been reported that organellar co-translationally translocated proteins are particularly sensitive to impaired NGD and NSD pathways, presumably due to their inaccessibility for cytoplasmic proteasome (*229*). It is thus tempting to speculate that a functional link exists between mRNA-mRNA interaction-induced translational interference and organellar metabolism.

Table 2. Gene Ontology (GO) information on convergent mRNAs showing strong mRNA-mRNA interactions.

	GO ID	GO term	Cluster frequency	P-value of cluster	Number of Paired Genes in cluster	Background frequency
Biological process	6139	nucleobase-containing compound metabolic process	69 of 168 genes, 41.1%	0.00142	28 paired genes with both genes having this GO term	1638 of 6602 genes, 24.8%
	34641	cellular nitrogen compound metabolic process	73 of 168 genes, 43.5%	0.00195	36 paired genes with both genes having this GO term	1788 of 6602 genes, 27.1%
	36211	protein modification process	35 of 168 genes, 20.8%	0.00636	10 paired genes with both genes having this GO term	641 of 6602 genes, 9.7%
	43412	macromolecule modification	38 of 168 genes, 22.6%	0.00738	12 paired genes with both genes having this GO term	730 of 6602 genes, 11.1%
Cellular component	44428	nuclear part	51 of 168 genes, 30.4%	0.00067	16 paired genes with both genes having this GO term	1076 of 6602 genes, 16.3%
	44422	organelle part	98 of 168 genes, 58.3%	0.00178	46 paired genes with both genes having this GO term	2755 of 6602 genes, 41.7%
	5634	nucleus	80 of 168 genes, 47.6%	0.00282	40 paired genes with both genes having this GO term	2110 of 6602 genes, 32.0%
	32991	macromolecular complex	79 of 168 genes, 47.0%	0.00732	34 paired genes with both genes having this GO term	2124 of 6602 genes, 32.2%

2 | Xrn1 as a key factor controlling the extent of sense-antisense mRNA interaction

We have extended our analysis of convergent gene loci-originating siRNAs to RNAi-competent *xrn1*Δ cells and detected a massive increase in siRNA signal in this condition, specifically corresponding to convergent mRNAs. We interpret this result as a major outcome of stabilized dsRNA substrates for Dicer cleavage *in vivo*, rather than of an increased stability of Dicer products, if we consider that guide siRNA protection by Argonaute and complementary siRNA decay by the cytoplasmic exosome follow similar dynamics in WT and *xrn1*Δ cells. These observations are important for our model, where Xrn1 plays a key role in controlling the extent of mRNA-mRNA interactions. In accordance to multiple and pleotropic phenotypes of *xrn1* mutants (*19*), we propose that this evolutionary conserved exoribonuclease is essential in limiting the potentially deleterious interactions between overlapping transcripts in the cytoplasm.

Interestingly, converging observations have been reported from studies performed in distant model organisms. In plants, the importance of cytoplasmic degradation machinery has been shown by the fact that a severe growth defect of double *xrn4 ski2* mutants can be rescued by the inactivation of *RDR6* gene (coding the RNA-dependent RNA polymerase, implicated in the biogenesis of endogenous siRNA) (*230*). Indeed, the defective turnover of cytoplasmic mRNA triggers a massive and uncontrolled post-transcriptional gene silencing (PTGS) of coding genes in *A. thaliana*, resulting early growth arrest. Additionally, as evoked in the Chapter I (see section 1 of Part II), decapped mRNAs—targets of Xrn4, plant homolog of Xrn1, in the cytoplasm—are more readily transformed into dsRNA, substrate for Dicer cleavage, than cap-containing messengers (*122*), underlining the importance of a timely elimination of decay intermediates from the cellular RNA pool.

In mammals, a different mechanism prevents the recognition of dsRNA structures as viral genome elements and the subsequent interferon response. Double-stranded RNA-specific adenosine deaminases (ADARs) modify dsRNA structures by changing adenosine to inosine and disrupting the perfect duplexes (*231*). The ubiquitously expressed ADAR1 member of the mammalian ADAR family has been shown to modify multiple sites within mRNAs, predominantly in 3' UTRs. ADAR1

is an essential protein and *Adar1* homozygous mutations result in embryonic lethality, showing activated interferon-dependent inflammatory pathways. Remarkably, this severe effect of ADAR1 loss can be completely suppressed by *Ihif1* mutation, inactivating the major cellular RIG-I-like dsRNA receptor (*232*). Even though mostly intramolecular dsRNA structures have been described as ADAR1 targets in mammalian cells, recent findings reveal the existence of a broad network of intermolecular RNA-RNA interactions in various organisms (*90-92*). It is thus highly plausible that ADARs might also act on intermolecular RNA duplexes, preventing its recognition by dsRNA receptors. In this case, we envisage that the upstream action of cytoplasmic RNA decay machinery contributes to the tight control of the interaction-competent RNA pool.

However, we notice that if the loss of Xrn1 significantly amplifies the outcomes of sense-antisense mRNA interactions, these interactions readily exist in wild-type cells. Our analysis of small RNA sequencing data in the form of metagene representation underlines that a significant proportion of convergent mRNA pairs yield a comparable siRNA signal in both WT and *xrn1*Δ cells. Remarkably, the metagene analysis also reveals how specific the siRNA signal is within the region of overlap between sense-antisense mRNA pairs. We thus propose that a deficient cytoplasmic 5'-3' RNA decay pathway simply accentuates the pre-existing interactions. Our findings on the functional interdependence between POR1 and OCA2 mRNAs in wild-type cells further support this idea, suggesting an intricate balance between the timely elimination of potentially deleterious interaction partners and the retention of the beneficial RNA-RNA pairing.

3 | Physiological implications of sense-antisense mRNA interaction

As evoked above, the interaction between convergent mRNA pairs might remodel the corresponding mRNP particles, and 3' UTR associating factors are the evident targets for a modified affinity of RNA binding. The differential status of the 5' and 3' extremities of POR1 mRNA (as discussed in the Part I of this chapter) in response to the presence or absence of OCA2 mRNA prompted us to speculate that interaction-engaged 3' UTR of POR1 mRNA might be less accessible for 3' end associating decapping activators, such as Lsm1-7-Pat1 complex. Continuing with this

idea, we hypothesized that a modified target RNA binding of these factors might be a general feature of sense-antisense mRNA pairs. To explore this possibility, we have analyzed the cross-link and immunoprecipitation (CLIP)-derived RNA sequencing data from Mitchel *et al.* that determined the localization of Pat1, Dhh1, Lsm1 and Sbp1 binding on mRNAs genome-wide (*233*). We have calculated the percentage of mRNAs that interact with these factors and annotated these results according to whether the transcripts belonged to solo or convergent mRNA categories, as defined by our small RNA sequencing analysis (Figure 58A). Remarkably, we have detected a reduced association of all four RNA binding factors with mRNAs originating from convergent loci, as compared to antisense-devoid messengers. Moreover, those convergent mRNAs that show higher than two-fold enrichment in siRNA signal in RNAi-competent *S. cerevisiae* additionally avoid binding these four factors. This result is specific for Pat1, Dhh1, Lsm1 and Sbp1, as Slf1, an RNA interacting protein not known to favor 3' ends, binding (*234*) or m^6A modification (*235*) equally affects the subsets of solo and convergent mRNAs.

We have pursued our analysis in determining the GO terms of convergent mRNAs, showing the reduced binding to 3' end associating decapping activators. Interestingly, we have detected that genes corresponding to protein modification ($P = 3\times10^{-5}$) and cellular response to DNA damage categories ($P = 2\times10^{-3}$) are enriched among our subset of mRNAs. As discussed in the Chapter I (see section 3 of Part I), all four factors—Pat1, Lsm1, Dhh1 and Sbp1—are components of cytoplasmic foci (predominantly P-bodies); moreover, Pat1, Lsm1 and Dhh1 have been found to form cytoplasmic foci in response to DNA damage stress (*236*). The current model proposes that cytoplasmic mRNP foci contain translationally inactive mRNAs that can return to the polysome-associated mRNA pool upon the removal of foci-inducing stimulus (*46*). One of the many intriguing examples is the transcriptional upregulation of glucose metabolism genes upon glucose starvation, the mRNAs of which are rapidly sequestrated within the cytoplasmic stress granules, postponing its translation to the upcoming glucose repletion. In parallel, a group of heat-shock protein (Hsp) genes are also upregulated in response to glucose starvation, although their corresponding mRNAs remain in the translationally active polysome-associated pool (*44*). These findings incite us to speculate that upon DNA damage (or other external stress stimuli), sense-antisense mRNA interaction prevents the binding

of 3' specific decapping activators to stress response-related messengers and maintains these mRNAs translationally active, facilitating the expression of stress response genes (Figure 58B).

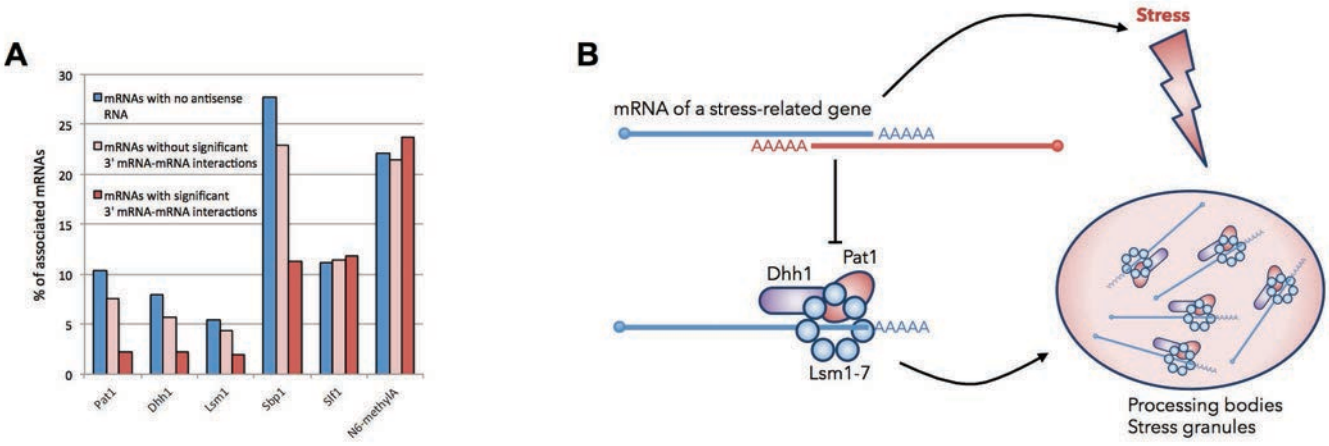

Figure 58. Pat1, Dhh1, Lsm1 and Sbp1 show reduced binding to mRNAs originating from convergent gene loci.

(A) A diagram showing the percentage of mRNAs binding to a specific RNA-binding protein (Pat1, Dhh1, Lsm1, Sbp1, Slf1) or the percentage of m^6A-containing mRNAs. The distribution of mRNAs was divided in three groups: mRNAs with no antisense RNA (blue bars), convergent mRNAs without significant siRNA signal in genome-wide analysis (light red bars), and convergent mRNAs showing higher than two-fold enrichment in siRNA signal in genome-wide analysis (dark red bars). The percentages shown were calculated within each group separately. Data retrieved from (*233*) (Pat1, Dhh1, Lsm1, Sbp1), (*234*) (Slf1), (*235*) (m^6A RNA).

(B) A model illustrating the putative role of mRNA-mRNA interactions in stress response. 3' overlapping and interacting mRNAs are poor substrates for specific RNA-binding factors, such as decapping activators Dhh1, Pat1, Lsm1-7. Upon stress (e.g. DNA damage stress), these factors form cytoplasmic foci, relocating the associated mRNAs to the translationally inactive stress-granules. However, convergent mRNAs, due to their poor binding to 3' associating decapping activators, remain in the translationally active mRNA pool, contributing to the stress-response pathway.

Our analysis of POR1 and OCA2 mRNAs in the context of remodeled carbon source metabolism also indicates that sense-antisense mRNA interactions are potentially implicated in cellular physiology-related processes. In the absence of OCA2 mRNA, POR1 mRNA exhibits lower steady-state levels, reduced activation upon respiratory growth conditions, and defective restoration of its levels after glucose-induced decay, as compared to wild-type situation. We suggest that this effect is mediated by an interaction between the two convergent transcripts, affecting

POR1 mRNA stability at its 3' and 5' ends, as discussed in the Part I of this chapter. Additionally, our observation that distinct isoforms of OCA2 mRNA exist during different stages of carbon source metabolism opens the possibility to propose a molecular model, explaining translational interference versus functional stabilization outcomes of POR1-OCA2 mRNA interaction. It is plausible that the shorter isoform of OCA2 mRNA does not extend to the coding portion of POR1 mRNA and thus solely protects POR1 messenger from the attack of cytoplasmic exosome. In contrast, the longer 3' end of OCA2 mRNA interacts with the 3' proximal region of POR1 ORF, causing translational interference, as described in *xrn1*Δ cells. These findings incite us to speculate that a mechanism connecting convergent mRNA poly(A) site selection with particular physiological conditions of the cell could exist.

Upon the exhaustion of glucose in the growth medium, *S. cerevisiae* activate the respiratory metabolism program and start using ethanol as a carbon source. This growth phase is termed diauxic shift, and Xrn1 has been documented to localize within P-bodies at this stage. Interestingly, in post-diauxic cells, Xrn1 has been found associated to eisosomes, equally distributed in the vicinity of cellular cortex, independently of other P-body components (*82*). We have hypothesized that this subcellular redistribution of Xrn1 might affect its activity in the bulk cytoplasmic mRNA decay. However, in post-diauxic cells (approximately at 48 hours after inoculation and beyond), OCA2 mRNA, an Xrn1 target, remains at low levels, similarly to during diauxic shift (24 hours post-inoculation), suggesting that its decay is independent of the dynamics of Xrn1 localization.

Our analysis has revealed that POR1 mRNA is a target of glucose-induced RNA decay, a pathway destabilizing a subset of respiration-related messengers upon respiration-to-fermentation transition (*237*). To our knowledge, POR1 mRNA has not been previously identified to respond to this pathway. Interestingly, POR1 mRNA, after being completely degraded, returns to initial steady-state levels within 3 to 4 hours after glucose repletion, and thus behaves as a non-canonical target of glucose-induced RNA decay. Moreover, POR1 messenger undergoes comparable degradation in both WT and *xrn1*Δ cells, suggesting that, differently from canonical glucose-induced RNA decay targets, it does not exclusively rely on the cytoplasmic 5'-3' RNA decay pathway. We have also detected that in *oca2*Δ cells, POR1 mRNA readily disappears within 30 minutes after respiration-to-fermentation transition, as

opposed to its enhanced stability in YPD medium after Pol II inhibition by thiolutin. This suggests that remodeling carbon source metabolism affects POR1 mRNP, making it accessible to RNA decay pathways.

Recently, glucose-induced decay of ADH2 mRNA has been shown to depend on its promoter sequence, Snf1-mediated signaling cascade and transcription factor Adr1 (*238*). *POR1* promoter also contains Adr1 consensus sequence and has been reported to bind this factor *in vivo* (*239*). As evoked previously, upon glucose starvation, a subset of Hsp mRNAs remain in the polysome-associated RNA pool when translation is globally repressed and ribosomes run off from mRNAs. Surprisingly, this phenomenon is defined by the promoter sequences of *HSP* genes, and is dependent on Hsf1 binding to heat-shock elements (HSE) embedded within these promoters (*44*). It is thus tempting to propose that Adr1, similarly to Hsf1, might direct the remodeling of the subcellular localization of POR1 mRNA, licensing its glucose-induced decay. Alternatively, POR1 mRNA stability has been associated with the RNA binding protein Ngr1 (*240, 241*), which could constitute an additional level of control for *POR1* expression. Further investigations are needed to establish a link between the transcriptional program of respiration-to-fermentation transition, mRNA-mRNA interactions, and RNA binding proteins.

An interesting question in the analysis of physiological significance of sense-antisense mRNA interactions remains the evolutionary conservation of convergent gene pairs. In metazoan organisms, 3' overlapping mRNA producing gene pairs (in both UTR-UTR and UTR-ORF configuration) are non-randomly retained, as opposed to general counterselection against overlapping coding genes (*133*). The above-mentioned human DHPS-WDR83 sense-antisense mRNA pair retains is genomic disposition since tetrapoda, and the mutually stabilizing effect of this mRNA-mRNA interaction has also been observed in mouse cell cultures (*179*). The convergent orientation of *POR1* and *OCA2* genes is conserved in post-WGD (whole genome duplication) yeast species (*242*), suggesting that POR1-OCA2 mRNA interaction might have acted as a positive factor for its selection. Interestingly, although *OCA* family genes are yeast specific and thus not conserved in metazoans, *VDAC* genes (*POR1* homologs) are in convergent orientation in various metazoan species. In *C. elegans*, *vdac-1* is convergent with *glt-6*; in humans, mice and zebrafish, among three *VDAC* genes, *VDAC2* and *VDAC3* are convergent, and, moreover, their

antisense counterparts are also conserved within these organisms. However, additional analyses are required to determine if *POR1* homologs produce overlapping transcripts to their convergent counterparts, and which mechanisms, regarding mRNA-mRNA interactions, are common to *S. cerevisiae* and other model organisms.

The convergent orientation of the main cytoplasmic 5'-3' RNA decay factors—Xrn1 and Dcp2—is conserved within various budding yeast species, both pre- and post-WGD; the convergent orientation of Dcp1 is only present in post-WGD yeasts (*242*). Interestingly, *xrn-1* is a convergent gene in *C. elegans*, and *DCP2* has an antisense gene in *C. elegans*, *D. melanogaster*, *M. musculus* and *H. sapiens*. We have thus hypothesized that convergent orientation of genes coding cytoplasmic 5'-3' RNA decay factors might be relevant for their auto-regulation by mRNA-mRNA interaction. As revealed by *XRN1*, *DCP2* and *DCP1* translational fusion reporters, the impaired cytoplasmic 5'-3' exoribonucleolytic activity results in a decreased production of reporter proteins, suggesting that these 5'-3' RNA decay factors engage in a negative feed-back control. This result appears counterintuitive if we consider that overlapping mRNA accumulation might stabilize mRNAs coding for RNA decay factors and subsequently upregulate their expression. However, antisense mRNAs, stabilized in *xrn1* mutant context, might also induce translational interference, as is the case for OCA2 mRNA, explaining the downregulation of the reporter proteins. In parallel, it is possible that the modified ration of 5' capped to uncapped mRNA in *xrn1*Δ cells might be sensed as a signal to reduce the cytoplasmic decapping, resulting in the downregulated expression of *DCP2* and *DCP1* reporters. In sum, the significance of the convergent orientation of cytoplasmic 5'-3' RNA decay factors remains an intriguing subject for a more detailed study.

PART III | Conclusions

The analysis of a convergent gene pair in *Saccharomyces cerevisiae*, *POR1* and *OCA2*, lead us to identify a novel post-transcriptional mechanism of gene expression control. The transcription of the *POR1-OCA2* locus produces two partially overlapping mRNAs, and we showed that these messengers interact in the cytoplasm. This interaction is dependent on the sequence complementarity between two transcripts, and is essential for the correct expression of *POR1*. We demonstrated that in the absence of *OCA2*, POR1 mRNA is significantly destabilized, due to an increased cytoplasmic exonucleolytic attack. However, we also showed that the uncontrolled accumulation of OCA2 mRNA leads to a deleterious effect on the production of Por1 protein. OCA2 transcript covers the 3' UTR and might extend within the coding sequence of POR1 mRNA. In this case, OCA2 mRNA induces ribosome stalling on POR1 mRNA and subsequently triggers no-go decay pathway. In sum, our findings reveal that the cytoplasmic interaction between POR1 and OCA2 mRNAs remodels the stability and translation of POR1 mRNA.

Seeking to generalize these principles, we used a reconstituted RNAi approach in *S. cerevisiae* and showed that mRNAs originating from convergent gene loci form RNA duplexes *in vivo* and yield a robust and overlap-specific small RNA signal. We identified Xrn1, a highly conserved major cytoplasmic exoribonuclease, as a key factor controlling the extent of sense-antisense mRNA interactions. We successfully validated the interaction between two additional mRNA pairs, confirming that our model outreaches the case of POR1 and OCA2 mRNAs.

Our findings open the perspective to consider post-transcriptional, cytoplasmic outcomes of convergent and overlapping transcription, which is highly prevalent within eukaryotic genomes. We anticipate that cytoplasmic mRNA-mRNA interaction might play a major role in mRNP remodeling and influence a broad range of cellular processes.

CHAPTER IV | MATERIALS AND METHODS

1 | Yeast media

Saccharomyces cerevisiae strains were grown in YPD medium or in synthetic minimum medium (SD). Minimal medium was completed for auxotrophy, and tryptophan, leucine, histidine, or uracil were omitted to keep selection for plasmids when necessary. 200 µg/mL G418 (American Bioanalytical), 100 µg/mL ClonNat (Werner Bioagents) and 200 µg/mL Hygromycin B (Sigma-Aldrich) were added in YPD medium plates to select for KanMX4, NatMX6 and HphMX4 cassettes, respectively. For lithium inhibition, cells were pre-grown in SD medium without methionine, cysteine and tryptophan and then transferred to the same medium with 100 mM LiCl (for 1 hour). For the mRNA decay analyses, thiolutin (15 mg/mL) was added to the cultures in order to block the transcription (generous gift from Pfizer). For carbon source shift analyses, cells were pre-cultured in YPD until early log phase, and then the medium was changed for YPG (containing 2% glycerol). Alternatively, cells where pre-incubated in YPG for 3 hours, and then 2% of glucose (final concentration) was added to the medium. The cells were harvested at the time points indicated in Figure 45, Figure 46, Figure 47, and Figure 49.

2 | Yeast strains

S. cerevisiae strains used in this work are listed in Table 3. The *dom34*Δ strains were generated by the one-step gene replacement using PCR fragment of the NatMX6 cassette amplified from plasmid pFA6a-NatMX6 (*243*) with primers OLB495 and OLB496 and transformed into strains YRP840 (WT), YRP884 (*xrn1*Δ), YRP1192 (*ski2*Δ) and YLB037 (*oca2*Δ, described below) respectively. Correct integration was confirmed by PCR with primers OLB497 and OLB498. In each case one strain was selected for analysis, YLB057, YLB058, YLB059 and YLB106. The *ura3* version of YRP884 (*xrn1*Δ) strain, YLB068, was constructed by replacing *URA3*-disrupted *XRN1* gene with HphMX4-containing cassette, amplified from pAG32 (*244*) using primers OLB372 and OLB373; the correct integration was confirmed by PCR with primers OLB379 and OLB380. To construct mutant *POR1* 3'UTR strain,

PCR amplified NatMX6 cassette using primers OLB492 and OLB493, and *RPS12* 3'UTR containing fragment, PCR amplified with primers OLB482 and OLB485, were PCR- fused using primers OLB490 and OLB491 and transformed into YRP840 (WT) and YRP884 (*xrn1*Δ), resulting in YLB043 and YLB044, respectively. The correct integration was confirmed by PCR using primers OLB488 and OLB489. To construct strains YLB066 (*ccr4*Δ) and YLB067 (*xrn1*Δ *ccr4*Δ), *CCR4* ORF in strains YRP840 (WT) and YRP884 (*xrn1*Δ) was replaced by the NatMX6 cassette, amplified by PCR with primers OLB525 and OLB526. Correct integration was confirmed by PCR with primers OLB527 and OLB528. Strains YLB078 (*ltn1*Δ) and YLB079 (*ltn1*Δ *xrn1*Δ) were constructed by replacing *LTN1* ORF in YRP840 (WT) and YLB068 (*xrn1*Δ) strains with NatMX6 cassette, using PCR primers OLB607 and OLB608; the correct integration was confirmed by PCR with primers OLB609 and OLB610. Strains carrying *HEL2* deletion (YLB080 and YLB081) were constructed by amplifying *hel2*Δ::KanMX4 region from Y03625 strain (Euroscarf) by PCR with primers OLB628 and OLB629, and recombining it into YRP840 (WT) and YLB068 (*xrn1*Δ) strains, respectively. The constructions were verified by PCR with primers OLB630 and OLB631. To obtain the *oca2*Δ strains, the *OCA2* ORF from -2 to +300 was replaced by the KanMX4 cassette, amplified by PCR from plasmid pFA6a-KanMX4, with primers OLB363 and OLB364 and transformed into YRP840 (WT), YRP884 (*xrn1*Δ), YRP1192 (*ski2*Δ), and YLB066 (*ccr4*Δ), respectively. Correct integration was confirmed by PCR with primers OLB360 and OLB361. In each case one strain was selected for analysis, YLB037, YLB038, YLB047, and YLB076.

Table 3. *Saccharomyces cerevisiae* strains used in this work.

Strain	Alias	Genotype	Reference
YRP840	WT	MATa *his4-539 leu2-3,112 trp1 ura3-52 cup1::LEU2/PGK1pG/MFA2pG*	(*245*)
YRP884	*xrn1*Δ	MATa *trp1 leu2-3,112 lys2-201 ura3-52 cup1::LEU2/PGK1pG/MFA2pG* ***xrn1::URA3***	(*245*)
YRP1192	*ski2*Δ	MATa *his4-539 leu2-3,112 trp1 ura3-52* ***ski2::LEU2***	(*246*)
YLB057	*dom34*Δ	As YRP840 but with ***dom34::NatMX6***	
YLB058	*xrn1*Δ *dom34*Δ	As YRP884 but with ***dom34::NatMX6***	
YLB059	*ski2*Δ *dom34*Δ	As YRP1192 but with ***dom34::NatMX6***	
YLB068	*xrn1*Δ	As YRP884 but with ***xrn1::HphMX4***	
YLB043	*POR1 3' UTR-mut*	As YRP840 but with ***3' UTR POR1::RPS12/NatMX6***	

YLB044	*xrn1*Δ *POR1 3' UTR-mut*	As YRP884 but with ***3' UTR POR1::RPS12/NatMX6***	
YBL066	*ccr4*Δ	As YRP840 but with ***ccr4::NatMX6***	
YBL067	*xrn1*Δ *ccr4*Δ	As YRP884 but with ***ccr4::NatMX6***	
YLB037	*oca2*Δ	As YRP840 but with ***oca2::KanMX4***	
YLB038	*xrn1*Δ *oca2*Δ	As YRP844 but with ***oca2::KanMX4***	
YLB047	*ski2*Δ *oca2*Δ	As YRP1192 but with ***oca2::KanMX4***	
YLB076	*ccr4*Δ *oca2*Δ	As YBL066 but with ***oca2::KanMX4***	
YLB078	*ltn1*Δ	As YRP840 but with ***ltn1::NatMX6***	
YLB079	*ltn1*Δ *xrn1*Δ	As YLB068 but with ***ltn1::NatMX6***	
YLB080	*hel2*Δ	As YRP840 but with ***hel2::KanMX4***	
YLB081	*hel2*Δ *xrn1*Δ	As YLB068 but with ***hel2::KanMX4***	
YLB106	*oca2*Δ *dom34*Δ	As YLB037 but with ***dom34::NatMX6***	
YAM1725	WT for RNAi	MATα *leu2-3,112 trp1-1 can1-100 ura3::EGFP(S65T)-KanMX6 ade2-1 his3-11,15*	(*51*)
YAM1730	WT for RNAi + Dcr1 Ago1	MATα ***LEU2::pTEF-Dcr1 TRP1::pTEF-Ago1*** *can1-100 ura3::EGFP(S65T)-KanMX6 ade2-1 his3-11,15*	(*51*)
YAM1982	*xrn1*Δ for RNAi	As YAM1725 but with ***xrn1::His3MX6***	
YAM2271	*xrn1*Δ for RNAi + Dcr1 Ago1	As YAM1730 but with ***xrn1::His3MX6***	

3 | Oligonucleotides

Oligonucleotides used in this work are listed in Table 4.

Table 4. Oligonucleotides used in this work.

Name	Alias	Application	Sequence 5' to 3'
OLB495		*DOM34* deletion by NatMX6	CATTCGTTGCTGCATCGTTGTCATTTTGTTCAATTATCGCATTCCTATCATAGCAAAAATCGGATCCCCGGGTTAATTAA
OLB496		*DOM34* deletion by NatMX6	CGATTTATTATAGGGTTGCAAATTTTATGTGTACATTACTTTTTTCTTACATAGTAAATCGAATTCGAGCTCGTTTAAAC
OLB497		*DOM34* deletion verification	GCGTCATCTTCTAACACCG
OLB498		*DOM34* deletion verification	GTGAACAGGTTCAGACAACTTCAAAGC
OLB607		*LTN1* deletion by NatMX6	TTCTTAATAGTTCAAATCTGCTAAGCCATCAAAAAAAGTTCAAGCAATAGTTGGTTCTTACAGCTGAAGCTTCGTACGC
OLB608		*LTN1* deletion by NatMX6	AATAAAAGGAACTTTGTTTAAAAAATGTAGTACATTTATATGAAATTTATATGCGATAGTGCATAGGCCACTAGTGGATCTG
OLB609		*LTN1* deletion verification	ACCCAGATTATGCCCCAACA
OLB610		*LTN1* deletion verification	AATATCCGGGTGATGGGCTG

OLB628		*hel2*Δ::KanMX4 amplification	AACGCGAAGGGGAAATAATGC
OLB629		*hel2*Δ::KanMX4 amplification	AAGGGCATCTGAACGACTGC
OLB630		*HEL2* deletion verification	TGTCCACAGGCCTTTTTCAG
OLB631		*HEL2* deletion verification	GTTTCATACTCGCGGTCACTC
OLB372		*XRN1* deletion by HphMX4	GTTTATTTTCTAAAGGATACTGTCTTCTTCCGTACTTATAATCGGGTTCACGATCTGTTTAGCTTGCCTTG
OLB373		*XRN1* deletion by HphMX4	CAATCCCCATTTGTTATAAGCTTTTTCTTAACAAGATCAACGATTAAATACCTCGTTTTCGACACTGGATG
OLB379		*XRN1* deletion verification	CTATTCTCACGATTAATGGTC
OLB380		*XRN1* deletion verification	ATGGGAGACGTGCAAAAGC
OLB525		*CCR4* deletion by NatMX6	CAGCATAAGGGACACCAGCAAGGGAACTCCGACTGACGTTATCCCTGCAAACTACCGCTACGGATCCCCGGGTTAATTAA
OLB526		*CCR4* deletion by NatMX6	CAATTTTATAATGAGGTAGTGTACAGAGAGGAGGGAGGGAGTGGGATGAAAGTGTGCGGTGAATTCGAGCTCGTTTAAAC
OLB527		*CCR4* deletion verification	TGTTATCGCCCGTACTTGCG
OLB528		*CCR4* deletion verification	GATAAGCTGTCGCCGTAACG
OLB363		kanMX integration in 5'-*OCA2* region	ATATCAACAAGTGAAGGCAAAATACCAATAGAGTGCACGTTTCTTGATGCCGACACTGGATGGCGGCGTT
OLB364		kanMX integration in 5'-*OCA2* region	ATCCCTGTAACAACTTTCTTATGATTCCCACAACTACACCAACACGGTGCTTGCCGATATCAAGCTTGCCTCGTC
OLB361		*OCA2* deletion verification	TCAGATTGGTTTCGAACATTG
OLB360		*OCA2* deletion verification	CCCAGATTTTTCAGTAATTTTTG
OLB362		Construction of pLB001	ATAGCTCTTGGTGCTAGGATCGGATCCCCCAGATTTTTCAGTAATTTTTG
OLB376		Construction of pLB001 and pLB002	GACTGTTGGCCGCGGCCGCTTGTTAAGACCTGGCGTCAC
OLB374		Construction of pLB002	GACATTGATCGGATCCGTTTATCACAATGTTCGAAACCAA
OLB456		Construction of pLB102	GACATGGATCCCACAGAACGGTGTGGAAACAA
OLB457		Construction of pLB102	GAAATGCGGCCGCATCGACCGTGAAGGTACGC
OLB611		Construction of pLB131	GACATATCGATAGGATGGCGTCTAGGTGAATAC
OLB612		Construction of pLB131	GAAATGCGGCCGCCCAAGAAGAGAAAGTTGAACTCAC
OLB613		Construction of pLB132	GACATATCGATCCAACCGTACGTGCTTAATTAC
OLB614		Construction of pLB132	GAAATGCGGCCGCCAGCAGAATTCAATCATACCTTATC
OLB512		Construction of pLB105	GACATGGATCCCTGACCAGCGATACAGACGG
OLB513		Construction of pLB105	GAAATGCGGCCGCAATGGTCAGAATGGGCGCTT

OLB508		Construction of pLB104	TCAAGCGTCGAAGGACAAAGAC
OLB509		Construction of pLB104	GGATCCCCCGAATTGATCC
OLB510		Construction of pLB104	CAATTCGGGGGATCC
OLB511		Construction of pLB104	TCCTTCGACGCTTGAGTGATGTTCATTAGTACTTTTTC
OLB520		Construction of pLB108	AACCTCTCAACAATGAGCAAGGGCGAGGAGCTG
OLB521		Construction of pLB108	ATTAGATATATACGTTTACTTGTACAGCTCGTCCATG
OLB518		Construction of pLB108	ACGTATATATCTAATATATATATGTTCAC
OLB519		Construction of pLB108 & pLB110	CATTGTTGAGAGGTTGGTTTG
OLB542		Construction of pLB110	AACCTCTCAACAATG CGTACGCTGCAGGTCGAC
OLB543		Construction of pLB110	CTCCTCGCCCTTGCT GCTAGTGGATCCGTTCAAGTC
OLB541		Construction of pLB110	AGCAAGGGCGAGGAGC
OLB568		Construction of pLB114	GACGAGCTGTACAAGCAGGTCGGTGCTAAGGCTAC
OLB303		Construction of pLB114 & pLB115	CAAGCAACCAATATGATTATGAGAACC
OLB569		Construction of pLB115	GACGAGCTGTACAAGTAACAGGTCGGTGCTAAGGCTAC
OLB567		Construction of pLB114 & pLB115	CTTGTACAGCTCGTCCATGC
OLB306		Construction of pLB114 & pLB115	CTGGTTCTCATAATCATATTGGTTGC
OLB574		Construction of pLB124	ACCCAAGGATTAACCCGGTATAG
OLB575		Construction of pLB124	CTATACCGGGTTAATCCTTGGGT
OLB737		Construction of pLB168	GACATACTAGTATGAGCAAGGGCGAGGAG
OLB738		Construction of pLB168	GATACCATATGTTACTTGTACAGCTCGTCC
OLB741		Construction of pLB172	GAAATGCGGCCGCCCACTACCCGTGAAAACCCT
OLB742		Construction of pLB172	GACATGGATCCTCGAGACTGAGCTCCTTGAGA
OLB743		Construction of pLB173	GAAATGCGGCCGCGGACAACCGCTTCTATCCGT
OLB744		Construction of pLB173	GACATACTAGTTGAAACTCTTTCTTTCCTTCCACT
OLB745		Construction of pLB174	GAAATGCGGCCGCACCCATAACACCCTTGGTGC
OLB746		Construction of pLB174	GACATGGATCCCCGCTTATGTTTGATTCGACCC
OLB768		Construction of pLB177	TATCAGTACGGTATGGATTATAAAGATGACGATGACAAGG
OLB769		Construction of pLB177	TTTTAGTCGTATGTTTTACTTGTACAGCTCGTCCATG
OLB770		Construction of pLB177	CATACCGTACTGATATATATTTGTTGC

OLB771		Construction of pLB177	AACATACGACTAAAAACGAAGTATATTC
OLB772		Construction of pLB178	TCGCATTAGACAATGGATTATAAAGATGACGATGACAAGG
OLB773		Construction of pLB178	ATAACACTTATTCTTTTACTTGTACAGCTCGTCCATG
OLB774		Construction of pLB178	CATTGTCTAATGCGAAGGTACTT
OLB775		Construction of pLB178	AAGAATAAGTGTTATACGTTTTATAGACACAC
OLB776		Construction of pLB179	AAGCTGTGCACAATGGATTATAAAGATGACGATGACAAGG
OLB777		Construction of pLB179	ATCTCACCTCTGTGCTTACTTGTACAGCTCGTCCATG
OLB778		Construction of pLB179	CATTGTGCACAGCTTTTCTTGTGA
OLB779		Construction of pLB179	GCACAGAGGTGAGATGCCCA
OLB322	prA	probe in 5' region of POR1 mRNA	CGTCTTTGACAGGCTGTTTAGCCTTC
OLB301	prB	probe in 3' region of POR1 mRNA	CAAAGACCAACCTAGCTTGTGAAC
OLB304	prC	probe in 3' region of OCA2 mRNA	GTTCACAAGCTAGGTTGGTCTTTG
OLB323	prD	probe in 5' region of OCA2 mRNA	CTTTCTTATGATTCCCACAACTACAC
OLB502	prE	probe in 3' region of POR1 mRNA	CTAGTGGCGAATTCGATGTTGAC
OLB501	prF	probe in 3' region of POR1 mRNA	GCCAAAGTGACAATACCGGAATC
OLB302	prG	probe in 3' region of POR1 mRNA	CAAGAACGAGCACATATATGGTATATAG
OLB652	prH	probe in 5' region of PMT7 mRNA	CCCTGTAATCTCAAGTCCTTCATG
OLB583	prJ	probe in 3' region of PMT7 mRNA	GGCAATCTAACGATTCTAAAAGACAC
OLB588	prK	probe in 3' region of SRB7 mRNA	GTCCTCTGTTCGCAAATTTCTTTG
OLB651	prL	probe in 5' region of SNM1 mRNA	CCTGGTCTTTATTCATTTTTGGGATCC
OLB587	prM	probe in 3' region of SNM1 mRNA	CTATATGTTCACTGATGATACAGTCAAC
OLB583	prN	probe in 3' region of PEX29 mRNA	GAGAGTGTCTGTTAGCTCTTCG
OLB473		probe in 3' region of RNA3/RNA4	GTTATAAAATTACCAAATGCTGAATGAAG
OLB378	scR1	detection of scR1 mRNA	GTCTAGCCGCGAGGAAGG
OLB25S	25S	detection of 25S rRNA	ATCCGCTAAGGAGTGTGTAACAACTCACC
OLBPGK1	PGK1	detection of PGK1 mRNA	TAAGATGGCCAAGAATGGTCTGGTTGGG
OLB891	U6	detection of U6 snRNA	CTGATCATCTCTGTATTG
AMO1509		detection of Ty1 siRNA	CCGTTAGACGTTTCAGCTTCCAAAACAGAAGAATGTGAGAAGGCTTCCACTAAG
OLB492		POR1 mut 3' UTR generation	CGGATCCCCGGGTTAATTAA

OLB493		POR1 mut 3' UTR generation	CATTAATATATTTGTATAAATAAAACAACAAGTCTAAA GATCAAGTAGAAAAATGGAAACGAATTCGAGCTCGTTT AAAC
OLB482		POR1 mut 3' UTR generation	GCTTTGAAGTTGTCTGAACCTGTTCACAAGCTAGGTT GGTCTTTGTCCTTCGACGCTTGAACGGTGTGATGTTC ATTAGTACTTTTTC
OLB485		POR1 mut 3' UTR generation	TTAATTAACCCGGGGATCCGCAATCGTAATAGTTTTAA AAAATCAACTATA
OLB490		POR1 mut 3' UTR generation	GTGTCGGTTCCTCTTTCGATGCTTTGAAGTTGTCTGA ACCTG
OLB491		POR1 mut 3' UTR generation	TTACAAAGAATATTTTCTGTCATTAATATATTTGTATA AATAAAACAACAAG
OLB488		POR1 mut 3' UTR verification	GTTAAGACCTGGCGTCACTC
OLB489		POR1 mut 3' UTR verification	GAAGTGAGTCTATTCTCACTGCC
OLB440		universal primer for ePAT	GCGAGCTCCGCGGCCGCGTTTTTTTTTTTT
OLB441		universal primer for ePAT	GCGAGCTCCGCGGCCGCGTTTTTTTTTTTTVN
OLB442		universal primer for ePAT	GAGCTCCGCGGCCGCG
OLB305		POR1 mRNA specific primer for ePAT	CTATATACCATATATGTGCTCGTTCTTG

4 | Plasmids

Plasmids used in this work are listed in Table 5 and were constructed using standard molecular biology procedures. Briefly, to construct pLB001, pLB002, pLB102, pLB131 and pLB132, DNA fragments containing *OCA2*, truncated fragment of *OCA2*, *RAD17*, *PEX29* or *SRB7* were amplified by PCR from yeast genomic DNA using primers OLB362-OLB376, OLB374-OLB376, OLB456-OLB457, OLB611-OLB612, or OLB613-614, respectively, and cloned into pCM185 (*206*) via *BamH*I-*Not*I or *Cla*I-*Not*I sites. The entire region of *POR1-OCA2* was amplified using primers OLB512-OLB513 and cloned into pRS316 (*247*) via *BamH*I-*Not*I sites, resulting in pLB105. The regions of *XRN1-BUD13*, *DCP2-NCS2* and *DCP1-SPT20* were amplified using primers OLB741-OLB742, OLB743-OLB744, and OLB745-OLB746, respectively and cloned into pRS313 or pRS316 (*247*) via *BamH*I-*Not*I or *Spe*I-*Not*I sites, resulting in pLB172, pLB173, pLB174 respectively. The Flag-GFP plasmid pLB168 was created by amplifying GFP sequence from pYM27 (*248*) and cloning it into p415ADH1-Flag plasmid, a generous gift from J. Banroques. Other plasmids used in this study were constructed using CloneEZ (GenScript) recombination-based cloning kit. The sequence

complementary to *RPS12* in the 3'-UTR of *RAD17* in pLB102 was replaced by that of POR1 recombining two PCR-generated fragments, amplified using OLB508-OLB509 from pLB001 and OLB510-OLB511 from pLB102, resulting in pLB104. The pLB104 was modified using Quick Change site-directed mutagenesis technique with primers OLB574-OLB575, yielding pLB124. To construct pLB108, GFP-containing PCR product amplified using OLB520-OLB521 from pYM27 (*248*) was recombined with the pLB105 fragment, PCR-generated using OLB518-OLB519. The GFP in pLB108 was tagged with 9 c-myc at N-terminus, recombining the OLB542-OLB543 PCR product from pYM21 (*248*) with the PCR-linearized pLB108 (using OLB519-OLB541), resulting in pLB110. Plasmids pLB114 and pLB115 were constructed recombining PCR products generated using OLB568-OLB303 and OLB569-OLB303, respectively, from pLB105 with the PCR fragment amplified using OLB567-OLB306 from pLB110. Plasmids pLB177, pLB178, and pLB179 were constructed by PCR-linearizing pLB172, pLB173, and pLB174 using primer pairs OLB770-OLB771, OLB774-OLB775, and OLB778-OLB779, and recombining it with Flag-GFP containing PCR products (from pLB168) amplified with OLB768-OLB769, OLB772-OLB773, and OLB776-OLB777, respectively. *Escherichia coli* TOP10 (Invitrogen) thermocompetent cells were used for cloning; all the plasmids were verified by sequencing (Eurofins Genomics).

Table 5. Plasmids used in this work.

Plasmid Name	Short Description	Marker	Reference
pCM185	centromeric vector with a tetracycline-regulatable promoter	*TRP1*	(*206*)
pRS316	centromeric yeast/*E. coli* shuttling vector with MCS	*URA3*	(*247*)
pRS313	centromeric yeast/*E. coli* shuttling vector with MCS	*HIS3*	(*247*)
pYM27	GFP Tag		(*248*)
pYM21	9Myc Tag		(*248*)
pFA6a-NatMX6	NatMX6 confers resistance to nourseothricin		(*243*)
pFA6a-KanMX4	KanMX4 confers resistance to Geneticin		(*243*)
pAG32	HphMX4 confers resistance to hygromycin B		(*244*)
pLB001	*OCA2* under *CYC1* promoter (pCM185 derivative) used for production of RNA1	*TRP1*	

pLB002	*OCA2* truncated version (pLB001 with *BamH*I-*Bts*BI deletion)	*TRP1*	
pLB131	*PEX29* under *CYC1* promoter (pCM185 derivative)	*TRP1*	
pLB132	*SRB7* under *CYC1* promoter (pCM185 derivative)	*TRP1*	
pLB102	used for production of RNA2 *in trans* (pCM185 derivative)	*TRP1*	
pLB104	used for production of RNA3 *in trans* (pCM185 derivative)	*TRP1*	
pLB124	used for production of RNA4 *in trans* (pCM185 derivative)	*TRP1*	
pLB105	*POR1-OCA2* region cloned into pRS316	*URA3*	
pLB108	GFP integrated in *POR1* gene (pRS316 derivative)	*URA3*	
pLB110	c-myc-tagged GFP integrated in *POR1* gene (pRS316 derivative)	*URA3*	
pLB114	Construction for c-mycGFP #1 (pRS316 derivative)	*URA3*	
pLB115	Construction for c-mycGFP #2 (pRS316 derivative)	*URA3*	
pLB168	Flag-tagged GFP under *ADH1* promoter (p415ADH1 derivative)	*LEU2*	
pLB172	*XRN1-BUD13* region cloned in pRS313	*HIS3*	
pLB173	*DCP2-NCS2* region cloned in pRS316	*URA3*	
pLB174	*DCP1-SPT20* region cloned in pRS316	*URA3*	
pLB177	*XRN1* gene replaced by Flag-tagged GFP (pRS313 derivative)	*HIS3*	
pLB178	*DCP2* gene replaced by Flag-tagged GFP (pRS316 derivative)	*URA3*	
pLB179	*DCP1* gene replaced by Flag-tagged GFP (pRS316 derivative)	*URA3*	

5 | Northern blot analysis

Total RNA was extracted according to the hot phenol extraction method. The cell pellet corresponding to 8 OD_{600} units was resuspended in AE buffer (50 mM NaOAc at pH 5.3 and 10 mM EDTA) with 1% SDS, extracted with phenol at 65°C for 10 min, and precipitated with 100% ethanol. 10 µg of total RNA was resolved on 1.2% agarose gel, transferred onto positively charged nylon membrane and UV-crosslinked. Blots were hybridized with [γ-^{32}P]-ATP labeled oligonucleotide probes in Roti-Hybrid-Quick (Carl Roth) buffer. Scans were performed with Typhoon scanner (Amersham Bioscience) and quantified using ImageJ software.

6 | 5'-phosphorylated RNA removal *in vitro*

5'-monophosphorylated RNA was degraded in vitro by Xrn1 treatment. 10 µg of total RNA was incubated with 15 ng/µl Xrn1 (final concentration) from *Kluyveromyces lactis* (a generous gift from Liang Tong) in the reaction buffer (50 mM Tris-HCl pH 7.9, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT with 20 units of RNase inhibitor) for 90 min at 37ºC. The reaction was stopped by the addition of 12.5 mM EDTA (final concentration) and the enzyme was inactivated by heat (5 min at 65ºC). The reaction mixture was resolved on an agarose gel, and Xrn1-resistant RNA was analyzed by northern blot.

7 | Extension poly(A) test (ePAT)

Extension poly(A) test was performed essentially as described by Jänicke *et al.* (*203*). Maxima H Minus RT (Thermo Fisher) was used to synthesize cDNA and Phusion polymerase (Thermo Fisher) was used for PCR. Universal primers OLB440, OLB441 and OLB442 have been described (*203*). POR1 mRNA was analyzed with primer OLB305.

8 | Polysome analysis

Polysome analysis was performed as described by Hu *et al.* (*63*). 200 mL of cell culture grown to mid-log phase were treated with cycloheximide to a final concentration of 100 µg/mL for 1 min in an ice-water bath and collected by centrifugation. Cell pellets were lysed in Polysome lysis buffer (10 mM Tris-HCl, pH 7.4, 100 mM NaCl, 30 mM $MgCl_2$, 1 mM DTT, 100 µg/mL cycloheximide) by bead bashing. Gradients were made by layering and freezing 15–45% w/w sucrose solutions in Polysome buffer (50 mM Tris-Acetate pH 7.0, 50 mM NH_4Cl, 12 mM $MgCl_2$, 1 mM DTT). 20 OD_{260} units of cell lysate were loaded onto each gradient, thawed overnight at 4ºC prior loading. Gradients were centrifuged at 37000 rpm for 3 h at 4ºC in a Beckman SW-41Ti rotor and fractionated using a Brandel Fractionation System and an Isco UA-6 ultraviolet detector. RNA was precipitated

by isopropanol, resuspended in AE buffer and extracted as described for Northern analysis.

9 | Western blot analysis

Crude extracts were prepared from yeast cells by bead bashing. The cell pellet corresponding to 8 OD_{600} units was resuspended in 5 M urea solution, boiled and vortexed with equal volume of glass beads. The extracts were then buffered with Lysis buffer (125 mM Tris-HCl, pH 6.8, 2% SDS) and clarified by centrifugation. Equivalent OD_{280} of extracts were loaded onto a NuPage (Invitrogen) Bis-Tris 12% polyacrylamide gel. After electrophoresis in MOPS SDS running buffer, proteins were electrotransferred onto a nitrocellulose membrane. The membrane was then blocked with Tris-buffer saline containing 0,05% Tween 20 (TBS-T) and 50 mg/mL BSA. After washing with TBS-T, the membrane was probed with a mouse anti-Por1 antibody, mouse anti-c-myc antibody or mouse anti-Flag antibody (all from Invitrogen), and left shaking overnight. The membranes were washed 3 x 10 min with TBS-T and blotted with anti-mouse IgG conjugated to IR dye 800 (LI-COR Bioscience-US) for 1h, protected from light. The membranes were washed 3 x 10 min min with TBS-T, scanned with Li-COR Odyssey Imager and quantified using Image Studio™ software. The Western blots were probed with a mouse anti-Pgk1 antibody (Invitrogen) or rabbit anti-Rpl1 antibody (generous gift from François Lacroute) as a loading control.

10 | RNA-sequencing

Small RNA-seq libraries were constructed according to the Small RNA Sample Preparation Guide (Illumina) using 10- to 40-nt small RNAs purified from total RNA on 15% TBE-urea PAGE. Size selection was validated by qualitative analysis of a sample of the purified small RNA fraction on a Small RNA chip in a 2100 bioanalyzer (Agilent). Single-end sequencing (40 nt) of libraries was performed on a Genome Analyzer IIx (Illumina). After removal of adaptor sequences using the Cutadapt tool (*249*), reads were uniquely mapped to the reference genome (*S. cerevisiae* S288C, 12/12/2011, retrieved from SGD; www.yeastgenome.org) using

version 0.12.7 of Bowtie (*250*) with a tolerance of three mismatches. All subsequent bioinformatic analyses were performed using uniquely mapped 19- to 23-nt small RNA reads.

11 | Small RNA analysis

4013 mRNAs (after having excluded mRNAs with annotated antisense ncRNAs and dubious ORFs) were defined (see Table S1 in Sinturel *et al.* (*215*)). 1456 convergent and 3'-overlapping mRNAs and 2557 mRNAs having no antisense (called solo mRNAs) were analyzed. Among convergent mRNAs, we defined 1142 mRNAs overlapping in their 3' UTR only and 315 additionally overlapping within their ORF. Only ≥23-nt overlaps between two convergent mRNAs were considered. For convergent mRNAs, 19- to 23-nt small RNA reads were counted within the region of overlap, on the considered strand. For solo mRNAs, reads were counted within the full-length transcripts. For each gene class, total read counts were expressed in rpkm (reads per kilobase of transcript per million mapped reads). Annotations of UTRs were retrieved from the reference genome R63 and imported into our reference genome using Blastn alignments. For MA plot analysis, 1923 solo mRNAs and 1087 convergent mRNAs are represented after filtering out the 1003 mRNAs (25% of the 4013 mRNA set, see Table S1 in Sinturel *et al.* (*215*)), showing the lowest density average in the four strains ($WT_{+Dcr1+Ago1}$ + WT + $xrn1\Delta_{+Dcr1+Ago1}$ + *xrn1*Δ; Figure 53C; see Table S2 in Sinturel *et al.* (*215*)).

12 | Metagene representations

Metagenes of small RNA signal were generated using R scripts. Boundaries (TSS, ATG, STOP, TTS, TTS anti, and TSS anti) defined in Figure 54 delimited five regions of interest: region 1 from TSS to ATG (Figure 54A-D); region 2 from ATG to STOP (Figure 54B and D) or from ATG to TTS anti (Figure 54A and C); region 3 from STOP to TSS anti (Figure 54B and D) or from TSS anti to STOP (Figure 54A and C); region 4 from TTS anti to TTS (Figure 54B and D) or from STOP to TTS (Figure 54A and C); and region 5 from TTS to TSS anti (Figure 54A-D). We first determined the mean length *L* of each region for each of the four subsets

of genes analyzed in Figure 54A-D. For regions 1–5 determined for Figure 54A, the mean length *L* is 69, 1464, 353, 228, and 1341 nt, respectively; for regions 1–5 for Figure 54B, the mean length *L* is 101, 1614, 66, 207, and 1896 nt, respectively; for regions 1–5 for Figure 54C, the mean length *L* is 76, 1280, 223, 164, and 1361 nt, respectively; and for regions 1–5 for Figure 54D, the mean length *L* is 97, 1465, 67, 157, and 1625 nt, respectively. Addition of the five *L* values for each subset of genes gave the length in virtual nucleotides of the corresponding metagene. For each subset of genes, metagene was computed for each strain as follows. Signal for regions 1–5 was computed in rpkm for each *i* gene (*i* comprised between 1 and *n*). Each region was scaled to its corresponding *L* value. Accordingly, signal associated to each nucleotide was scaled upon polynomial interpolation, so that each *j* virtual nucleotide of the scaled region was associated to a scaled signal $S_{i,j}$. Scaled regions 1–5 were then concatenated, and $S_{i,j}$ values of all virtual nucleotides were summed to give T_i, corresponding to the whole signal for the *i* gene. To ensure that each gene contributed equally to the final metagene, the $S_{i,j}$ value of the *j* virtual nucleotide of the *i* gene was divided by its associated T_i value, giving a normalized $S^*_{i,j}$ signal for each position. Finally, the metasignal at each *j* virtual nucleotide was computed as $M_j = \left(\sum_{i=1}^{n} S^*_{i,j}\right) \times T/n$, where *T* corresponds to the sum of the T_i values of the *n* genes of the subset.

CHAPTER V | REFERENCES

1. S. Sainsbury, C. Bernecky, P. Cramer, Structural basis of transcription initiation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16**, 129-143 (2015).
2. H. Neil *et al.*, Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature* **457**, 1038-1042 (2009).
3. C. Malabat, F. Feuerbach, L. Ma, C. Saveanu, A. Jacquier, Quality control of transcription start site selection by nonsense-mediated-mRNA decay. *Elife* **4**, (2015).
4. J. P. Hsin, J. L. Manley, The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev* **26**, 2119-2137 (2012).
5. I. Jonkers, J. T. Lis, Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16**, 167-177 (2015).
6. Q. Zhou, T. Li, D. H. Price, RNA polymerase II elongation control. *Annu Rev Biochem* **81**, 119-143 (2012).
7. O. Porrua, D. Libri, Transcription termination and the control of the transcriptome: why, where and how to stop. *Nat Rev Mol Cell Biol* **16**, 190-202 (2015).
8. M. Schmid *et al.*, The Nuclear PolyA-Binding Protein Nab2p Is Essential for mRNA Production. *Cell Rep* **12**, 128-139 (2015).
9. V. Pelechano, W. Wei, L. M. Steinmetz, Extensive transcriptional heterogeneity revealed by isoform profiling. *Nature* **497**, 127-131 (2013).
10. A. Ghosh, C. D. Lima, Enzymology of RNA cap synthesis. *Wiley Interdiscip Rev RNA* **1**, 152-172 (2010).
11. E. Tutucci, F. Stutz, Keeping mRNPs in check during assembly and nuclear export. *Nat Rev Mol Cell Biol* **12**, 377-384 (2011).
12. H. Le Hir, J. Saulière, Z. Wang, The exon junction complex as a node of post-transcriptional networks. *Nat Rev Mol Cell Biol* **17**, 41-54 (2016).
13. A. Köhler, E. Hurt, Exporting RNA from the nucleus to the cytoplasm. *Nat Rev Mol Cell Biol* **8**, 761-773 (2007).
14. C. A. Niño, L. Hérissant, A. Babour, C. Dargemont, mRNA nuclear export in yeast. *Chem Rev* **113**, 8523-8545 (2013).
15. A. Bonnet, B. Palancade, Regulation of mRNA trafficking by nuclear pore complexes. *Genes (Basel)* **5**, 767-791 (2014).
16. V. O. Wickramasinghe, R. A. Laskey, Control of mammalian gene expression by selective mRNA export. *Nat Rev Mol Cell Biol* **16**, 431-442 (2015).
17. C. Kilchert, S. Wittmann, L. Vasiljeva, The regulation and functions of the nuclear RNA exosome complex. *Nat Rev Mol Cell Biol* **17**, 227-239 (2016).
18. T. H. Jensen, A. Jacquier, D. Libri, Dealing with pervasive transcription. *Mol Cell* **52**, 473-484 (2013).
19. V. K. Nagarajan, C. I. Jones, S. F. Newbury, P. J. Green, XRN 5'→3' exoribonucleases: structure, mechanisms and functions. *Biochim Biophys Acta* **1829**, 590-603 (2013).
20. X. Jiao *et al.*, Identification of a quality-control mechanism for mRNA 5'-end capping. *Nature* **467**, 608-611 (2010).
21. S. Xiang *et al.*, Structure and function of the 5'-->3' exoribonuclease Rat1 and its activating partner Rai1. *Nature* **458**, 784-788 (2009).

22. J. H. Chang *et al.*, Dxo1 is a new type of eukaryotic enzyme with both decapping and 5'-3' exoribonuclease activity. *Nat Struct Mol Biol* **19**, 1011-1017 (2012).
23. A. Lebreton, R. Tomecki, A. Dziembowski, B. Séraphin, Endonucleolytic RNA cleavage by a eukaryotic exosome. *Nature* **456**, 993-996 (2008).
24. R. K. Gudipati *et al.*, Extensive degradation of RNA precursors by the exosome in wild-type cells. *Mol Cell* **48**, 409-421 (2012).
25. J. LaCava *et al.*, RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell* **121**, 713-724 (2005).
26. F. Wyers *et al.*, Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* **121**, 725-737 (2005).
27. M. Thoms *et al.*, The Exosome Is Recruited to RNA Substrates through Specific Adaptor Proteins. *Cell* **162**, 1029-1038 (2015).
28. A. Lardenois *et al.*, Execution of the meiotic noncoding RNA expression program and the onset of gametogenesis in yeast require the conserved exosome subunit Rrp6. *Proc Natl Acad Sci U S A* **108**, 1058-1063 (2011).
29. R. J. Jackson, C. U. Hellen, T. V. Pestova, The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* **11**, 113-127 (2010).
30. A. G. Hinnebusch, J. R. Lorsch, The mechanism of eukaryotic translation initiation: new insights and challenges. *Cold Spring Harb Perspect Biol* **4**, (2012).
31. T. E. Dever, R. Green, The elongation, termination, and recycling phases of translation in eukaryotes. *Cold Spring Harb Perspect Biol* **4**, a013706 (2012).
32. C. Schmidt *et al.*, Structure of the hypusinylated eukaryotic translation factor eIF-5A bound to the ribosome. *Nucleic Acids Res* **44**, 1944-1951 (2016).
33. E. Gutierrez *et al.*, eIF5A promotes translation of polyproline motifs. *Mol Cell* **51**, 35-45 (2013).
34. L. K. Doerfel *et al.*, EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. *Science* **339**, 85-88 (2013).
35. T. E. Dever, T. G. Kinzy, G. D. Pavitt, Mechanism and Regulation of Protein Synthesis in Saccharomyces cerevisiae. *Genetics* **203**, 65-107 (2016).
36. D. E. Weinberg *et al.*, Improved Ribosome-Footprint and mRNA Measurements Provide Insights into Dynamics and Regulation of Yeast Translation. *Cell Rep* **14**, 1787-1799 (2016).
37. F. Gebauer, M. W. Hentze, Molecular mechanisms of translational control. *Nat Rev Mol Cell Biol* **5**, 827-835 (2004).
38. P. P. Roux, I. Topisirovic, Regulation of mRNA translation by signaling pathways. *Cold Spring Harb Perspect Biol* **4**, (2012).
39. J. Kong, P. Lasko, Translational control in cellular and developmental processes. *Nat Rev Genet* **13**, 383-394 (2012).
40. K. D. Meyer *et al.*, 5' UTR m(6)A Promotes Cap-Independent Translation. *Cell* **163**, 999-1010 (2015).
41. J. Zhou *et al.*, Dynamic m(6)A mRNA methylation directs translational control of heat shock response. *Nature* **526**, 591-594 (2015).
42. E. Lécuyer *et al.*, Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* **131**, 174-187 (2007).

43. A. R. Buxbaum, G. Haimovich, R. H. Singer, In the right place at the right time: visualizing and understanding mRNA localization. *Nat Rev Mol Cell Biol* **16**, 95-109 (2015).
44. B. M. Zid, E. K. O'Shea, Promoter sequences direct cytoplasmic localization and translation of mRNAs during starvation in yeast. *Nature* **514**, 117-121 (2014).
45. C. E. Simpson, M. P. Ashe, Adaptation to stress in yeast: to translate or not? *Biochem Soc Trans* **40**, 794-799 (2012).
46. C. J. Decker, R. Parker, P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. *Cold Spring Harb Perspect Biol* **4**, a012286 (2012).
47. Y. Lin, D. S. Protter, M. K. Rosen, R. Parker, Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol Cell* **60**, 208-219 (2015).
48. J. R. Buchan, R. M. Kolaitis, J. P. Taylor, R. Parker, Eukaryotic stress granules are cleared by autophagy and Cdc48/VCP function. *Cell* **153**, 1461-1474 (2013).
49. S. Jain *et al.*, ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell* **164**, 487-498 (2016).
50. S. Jonas, E. Izaurralde, Towards a molecular understanding of microRNA-mediated gene silencing. *Nat Rev Genet* **16**, 421-433 (2015).
51. I. A. Drinnenberg *et al.*, RNAi in budding yeast. *Science* **326**, 544-550 (2009).
52. R. Parker, RNA degradation in Saccharomyces cerevisae. *Genetics* **191**, 671-702 (2012).
53. N. L. Garneau, J. Wilusz, C. J. Wilusz, The highways and byways of mRNA decay. *Nat Rev Mol Cell Biol* **8**, 113-126 (2007).
54. E. Wahle, G. S. Winkler, RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes. *Biochim Biophys Acta* **1829**, 561-570 (2013).
55. R. Matsuda, K. Ikeuchi, S. Nomura, T. Inada, Protein quality control systems associated with no-go and nonstop mRNA surveillance in yeast. *Genes Cells* **19**, 1-12 (2014).
56. S. Preissler *et al.*, Not4-dependent translational repression is important for cellular protein homeostasis in yeast. *EMBO J* **34**, 1905-1924 (2015).
57. N. Alhusaini, J. Coller, The deadenylase components Not2p, Not3p, and Not5p promote mRNA decapping. *RNA* **22**, 709-721 (2016).
58. I. Gupta *et al.*, Translational Capacity of a Cell Is Determined during Transcription Elongation via the Ccr4-Not Complex. *Cell Rep*, (2016).
59. V. Taverniti, B. Séraphin, Elimination of cap structures generated by mRNA decay involves the new scavenger mRNA decapping enzyme Aph1/FHIT together with DcpS. *Nucleic Acids Res* **43**, 482-492 (2015).
60. M. Arribas-Layton, D. Wu, J. Lykke-Andersen, H. Song, Structural and functional control of the eukaryotic mRNA decapping machinery. *Biochim Biophys Acta* **1829**, 580-589 (2013).
61. H. Sharif, E. Conti, Architecture of the Lsm1-7-Pat1 complex: a conserved assembly in eukaryotic mRNA turnover. *Cell Rep* **5**, 283-291 (2013).
62. H. Sharif *et al.*, Structural analysis of the yeast Dhh1-Pat1 complex reveals how Dhh1 engages Pat1, Edc3 and RNA in mutually exclusive interactions. *Nucleic Acids Res* **41**, 8377-8390 (2013).

63. W. Hu, T. J. Sweet, S. Chamnongpol, K. E. Baker, J. Coller, Co-translational mRNA decay in Saccharomyces cerevisiae. *Nature* **461**, 225-229 (2009).
64. E. L. van Dijk *et al.*, XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. *Nature* **475**, 114-117 (2011).
65. V. Pelechano, W. Wei, L. M. Steinmetz, Widespread Co-translational RNA Decay Reveals Ribosome Dynamics. *Cell* **161**, 1400-1412 (2015).
66. F. Sinturel, D. Bréchemier-Baey, M. Kiledjian, C. Condon, L. Bénard, Activation of 5'-3' exoribonuclease Xrn1 by cofactor Dcs1 is essential for mitochondrial function in yeast. *Proc Natl Acad Sci U S A* **109**, 8264-8269 (2012).
67. G. Haimovich *et al.*, Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis. *Cell* **153**, 1000-1011 (2013).
68. M. Sun *et al.*, Global analysis of eukaryotic mRNA degradation reveals Xrn1-dependent buffering of transcript levels. *Mol Cell* **52**, 52-62 (2013).
69. F. Halbach, P. Reichelt, M. Rode, E. Conti, The yeast ski complex: crystal structure and RNA channeling to the exosome complex. *Cell* **154**, 814-826 (2013).
70. E. Kowalinski, A. Schuller, R. Green, E. Conti, Saccharomyces cerevisiae Ski7 Is a GTP-Binding Protein Adopting the Characteristic Conformation of Active Translational GTPases. *Structure* **23**, 1336-1343 (2015).
71. H. Huang *et al.*, Bulk RNA degradation by nitrogen starvation-induced autophagy in yeast. *EMBO J* **34**, 154-168 (2015).
72. S. Huch, T. Nissan, Interrelations between translation and general mRNA degradation in yeast. *Wiley Interdiscip Rev RNA* **5**, 747-763 (2014).
73. D. R. Schoenberg, L. E. Maquat, Regulation of cytoplasmic mRNA decay. *Nat Rev Genet* **13**, 246-259 (2012).
74. F. He, C. Li, B. Roy, A. Jacobson, Yeast Edc3 targets RPS28B mRNA for decapping by binding to a 3' untranslated region decay-inducing regulatory element. *Mol Cell Biol* **34**, 1438-1451 (2014).
75. V. Presnyak *et al.*, Codon optimality is a major determinant of mRNA stability. *Cell* **160**, 1111-1124 (2015).
76. T. Sweet, C. Kovalak, J. Coller, The DEAD-box protein Dhh1 promotes decapping by slowing ribosome movement. *PLoS Biol* **10**, e1001342 (2012).
77. K. A. Braun *et al.*, Phosphoproteomic analysis identifies proteins involved in transcription-coupled mRNA decay as targets of Snf1 signaling. *Sci Signal* **7**, ra64 (2014).
78. L. Benard, Inhibition of 5' to 3' mRNA degradation under stress conditions in Saccharomyces cerevisiae: from GCN4 to MET16. *RNA* **10**, 458-468 (2004).
79. A. L. Todeschini, C. Condon, L. Bénard, Sodium-induced GCN4 expression controls the accumulation of the 5' to 3' RNA degradation inhibitor, 3'-phosphoadenosine 5'-phosphate. *J Biol Chem* **281**, 3276-3282 (2006).
80. N. Malys, K. Carroll, J. Miyan, D. Tollervey, J. E. McCarthy, The 'scavenger' m7GpppX pyrophosphatase activity of Dcs1 modulates nutrient-induced responses in yeast. *Nucleic Acids Res* **32**, 3590-3600 (2004).
81. N. Malys, J. E. McCarthy, Dcs2, a novel stress-induced modulator of m7GpppX pyrophosphatase activity that locates to P bodies. *J Mol Biol* **363**, 370-382 (2006).
82. T. Grousl, M. Opekarová, V. Stradalova, J. Hasek, J. Malinsky, Evolutionarily conserved 5'-3' exoribonuclease Xrn1 accumulates at plasma

membrane-associated eisosomes in post-diauxic yeast. *PLoS One* **10**, e0122770 (2015).
83. O. Brandman, R. S. Hegde, Ribosome-associated protein quality control. *Nat Struct Mol Biol* **23**, 7-15 (2016).
84. K. Saito, W. Horikawa, K. Ito, Inhibiting K63 polyubiquitination abolishes no-go type stalled translation surveillance in Saccharomyces cerevisiae. *PLoS Genet* **11**, e1005197 (2015).
85. M. Graille, B. Séraphin, Surveillance pathways rescuing eukaryotic ribosomes lost in translation. *Nat Rev Mol Cell Biol* **13**, 727-735 (2012).
86. C. J. Shoemaker, R. Green, Translation drives mRNA quality control. *Nat Struct Mol Biol* **19**, 594-601 (2012).
87. S. Lykke-Andersen, T. H. Jensen, Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nat Rev Mol Cell Biol* **16**, 665-677 (2015).
88. F. He, A. Jacobson, Nonsense-Mediated mRNA Decay: Degradation of Defective Transcripts Is Only Part of the Story. *Annu Rev Genet* **49**, 339-366 (2015).
89. T. R. Cech, J. A. Steitz, The noncoding RNA revolution-trashing old rules to forge new ones. *Cell* **157**, 77-94 (2014).
90. J. G. Aw *et al.*, In Vivo Mapping of Eukaryotic RNA Interactomes Reveals Principles of Higher-Order Organization and Regulation. *Mol Cell* **62**, 603-617 (2016).
91. Z. Lu *et al.*, RNA Duplex Map in Living Cells Reveals Higher-Order Transcriptome Structure. *Cell* **165**, 1267-1279 (2016).
92. E. Sharma, T. Sterne-Weiler, D. O'Hanlon, B. J. Blencowe, Global Mapping of Human RNA-RNA Interactions. *Mol Cell* **62**, 618-626 (2016).
93. S. Rouskin, M. Zubradt, S. Washietl, M. Kellis, J. S. Weissman, Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature* **505**, 701-705 (2014).
94. J. J. Quinn, H. Y. Chang, Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet* **17**, 47-62 (2016).
95. G. St Laurent, C. Wahlestedt, P. Kapranov, The Landscape of long noncoding RNA classification. *Trends Genet* **31**, 239-251 (2015).
96. M. Guttman, J. L. Rinn, Modular regulatory principles of large non-coding RNAs. *Nature* **482**, 339-346 (2012).
97. F. Dupuis-Sandoval, M. Poirier, M. S. Scott, The emerging landscape of small nucleolar RNAs in cell biology. *Wiley Interdiscip Rev RNA* **6**, 381-397 (2015).
98. S. Kishore, S. Stamm, The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. *Science* **311**, 230-232 (2006).
99. S. Kishore *et al.*, The snoRNA MBII-52 (SNORD 115) is processed into smaller RNAs and regulates alternative splicing. *Hum Mol Genet* **19**, 1153-1164 (2010).
100. M. Falaleeva *et al.*, Dual function of C/D box small nucleolar RNAs in rRNA modification and alternative pre-mRNA splicing. *Proc Natl Acad Sci U S A* **113**, E1625-1634 (2016).
101. M. Esteller, Non-coding RNAs in human disease. *Nat Rev Genet* **12**, 861-874 (2011).

102. M. Ha, V. N. Kim, Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol* **15**, 509-524 (2014).
103. A. A. Bazzini, M. T. Lee, A. J. Giraldez, Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish. *Science* **336**, 233-237 (2012).
104. S. Djuranovic, A. Nahvi, R. Green, miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. *Science* **336**, 237-240 (2012).
105. S. W. Eichhorn *et al.*, mRNA destabilization is the dominant effect of mammalian microRNAs by the time substantial repression ensues. *Mol Cell* **56**, 104-115 (2014).
106. T. T. Tat, P. A. Maroney, S. Chamnongpol, J. Coller, T. W. Nilsen, Cotranslational microRNA mediated messenger RNA destabilization. *Elife* **5**, (2016).
107. R. S. Reis, G. Hart-Smith, A. L. Eamens, M. R. Wilkins, P. M. Waterhouse, Gene regulation by translational inhibition is determined by Dicer partnering proteins. *Nat Plants* **1**, 14027 (2015).
108. S. Guil, M. Esteller, RNA-RNA interactions in gene regulation: the coding and noncoding players. *Trends Biochem Sci* **40**, 248-256 (2015).
109. M. J. Luteijn, R. F. Ketting, PIWI-interacting RNAs: from generation to transgenerational epigenetics. *Nat Rev Genet* **14**, 523-534 (2013).
110. B. Czech, G. J. Hannon, One Loop to Rule Them All: The Ping-Pong Cycle and piRNA-Guided Silencing. *Trends Biochem Sci*, (2016).
111. D. Holoch, D. Moazed, RNA-mediated epigenetic regulation of gene expression. *Nat Rev Genet* **16**, 71-84 (2015).
112. Y. K. Kim, L. Furic, L. Desgroseillers, L. E. Maquat, Mammalian Staufen1 recruits Upf1 to specific mRNA 3'UTRs so as to elicit mRNA decay. *Cell* **120**, 195-208 (2005).
113. C. Gong, L. E. Maquat, lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* **470**, 284-288 (2011).
114. C. Gong, Y. Tang, L. E. Maquat, mRNA-mRNA duplexes that autoelicit Staufen1-mediated mRNA decay. *Nat Struct Mol Biol* **20**, 1214-1220 (2013).
115. Y. Sugimoto *et al.*, hiCLIP reveals the in vivo atlas of mRNA secondary structures recognized by Staufen 1. *Nature* **519**, 491-494 (2015).
116. M. Kretz *et al.*, Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature* **493**, 231-235 (2013).
117. M. L. Li, K. F. Weng, S. R. Shih, G. Brewer, The evolving world of small RNAs from RNA viruses. *Wiley Interdiscip Rev RNA*, (2016).
118. M. Halic, D. Moazed, Dicer-independent primal RNAs trigger RNAi and heterochromatin formation. *Cell* **140**, 504-516 (2010).
119. K. M. Kowalik *et al.*, The Paf1 complex represses small-RNA-mediated epigenetic gene silencing. *Nature* **520**, 248-252 (2015).
120. F. Borges, R. A. Martienssen, The expanding world of small RNAs in plants. *Nat Rev Mol Cell Biol* **16**, 727-741 (2015).
121. L. Liu, X. Chen, RNA Quality Control as a Key to Suppressing RNA Silencing of Endogenous Genes in Plants. *Mol Plant*, (2016).
122. J. S. Parent *et al.*, Post-transcriptional gene silencing triggered by sense transgenes involves uncapped antisense RNA and differs from silencing

intentionally triggered by antisense transgenes. *Nucleic Acids Res* **43**, 8464-8475 (2015).
123. P. A. Dumesic *et al.*, Stalled spliceosomes are a signal for RNAi-mediated genome defense. *Cell* **152**, 957-968 (2013).
124. M. S. Weinberg, K. V. Morris, Transcriptional gene silencing in humans. *Nucleic Acids Res*, (2016).
125. K. V. Morris, S. W. Chan, S. E. Jacobsen, D. J. Looney, Small interfering RNA-induced transcriptional gene silencing in human cells. *Science* **305**, 1289-1292 (2004).
126. J. Han, D. Kim, K. V. Morris, Promoter-associated RNA is required for RNA-directed transcriptional gene silencing in human cells. *Proc Natl Acad Sci U S A* **104**, 12422-12427 (2007).
127. M. Alló *et al.*, Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat Struct Mol Biol* **16**, 717-724 (2009).
128. K. Skourti-Stathaki, K. Kamieniarz-Gdula, N. J. Proudfoot, R-loops induce repressive chromatin marks over mammalian gene terminators. *Nature* **516**, 436-439 (2014).
129. S. Katayama *et al.*, Antisense transcription in the mammalian transcriptome. *Science* **309**, 1564-1566 (2005).
130. X. J. Wang, T. Gaasterland, N. H. Chua, Genome-wide prediction and identification of cis-natural antisense transcripts in Arabidopsis thaliana. *Genome Biol* **6**, R30 (2005).
131. Y. Zhang, X. S. Liu, Q. R. Liu, L. Wei, Genome-wide in silico identification and analysis of cis natural antisense transcripts (cis-NATs) in ten species. *Nucleic Acids Res* **34**, 3465-3475 (2006).
132. C. R. Sanna, W. H. Li, L. Zhang, Overlapping genes in the human and mouse genomes. *BMC Genomics* **9**, 169 (2008).
133. G. Soldà *et al.*, Non-random retention of protein-coding overlapping genes in Metazoa. *BMC Genomics* **9**, 174 (2008).
134. E. J. Wood, K. Chin-Inmanu, H. Jia, L. Lipovich, Sense-antisense gene pairs: sequence, transcription, and structure are not conserved between human and mouse. *Front Genet* **4**, 183 (2013).
135. M. A. Faghihi, C. Wahlestedt, Regulatory roles of natural antisense transcripts. *Nat Rev Mol Cell Biol* **10**, 637-643 (2009).
136. V. Pelechano, L. M. Steinmetz, Gene regulation by antisense transcription. *Nat Rev Genet* **14**, 880-893 (2013).
137. M. Castelnuovo, F. Stutz, Role of chromatin, environmental changes and single cell heterogeneity in non-coding transcription and gene regulation. *Curr Opin Cell Biol* **34**, 16-22 (2015).
138. T. L. Lenstra, A. Coulon, C. C. Chow, D. R. Larson, Single-Molecule Imaging Reveals a Switch between Spurious and Functional ncRNA Transcription. *Mol Cell* **60**, 597-610 (2015).
139. J. Houseley, L. Rubbi, M. Grunstein, D. Tollervey, M. Vogelauer, A ncRNA modulates histone modification and mRNA induction in the yeast GAL gene cluster. *Mol Cell* **32**, 685-695 (2008).
140. M. Pinskaya, S. Gourvennec, A. Morillon, H3 lysine 4 di- and tri-methylation deposited by cryptic transcription attenuates promoter activation. *EMBO J* **28**, 1697-1707 (2009).

141. T. Kim, Z. Xu, S. Clauder-Münster, L. M. Steinmetz, S. Buratowski, Set3 HDAC mediates effects of overlapping noncoding transcription on gene induction kinetics. *Cell* **150**, 1158-1169 (2012).
142. M. Magistri, M. A. Faghihi, G. St Laurent, C. Wahlestedt, Regulation of chromatin structure by long noncoding RNAs: focus on natural antisense transcripts. *Trends Genet* **28**, 389-396 (2012).
143. D. J. Hobson, W. Wei, L. M. Steinmetz, J. Q. Svejstrup, RNA polymerase II collision interrupts convergent transcription. *Mol Cell* **48**, 365-374 (2012).
144. E. M. Prescott, N. J. Proudfoot, Transcriptional collision between convergent genes in budding yeast. *Proc Natl Acad Sci U S A* **99**, 8796-8801 (2002).
145. I. A. Drinnenberg, G. R. Fink, D. P. Bartel, Compatibility with killer explains the rise of RNAi-deficient fungi. *Science* **333**, 1592 (2011).
146. E. A. Alcid, T. Tsukiyama, Expansion of antisense lncRNA transcriptomes in budding yeast species since the loss of RNAi. *Nat Struct Mol Biol* **23**, 450-455 (2016).
147. E. A. Alcid, T. Tsukiyama, ATP-dependent chromatin remodeling shapes the long noncoding RNA landscape. *Genes Dev* **28**, 2348-2360 (2014).
148. E. White, M. Schlackow, K. Kamieniarz-Gdula, N. J. Proudfoot, M. Gullerova, Human nuclear Dicer restricts the deleterious accumulation of endogenous double-stranded RNA. *Nat Struct Mol Biol* **21**, 552-559 (2014).
149. A. Werner *et al.*, Contribution of natural antisense transcription to an endogenous siRNA signature in human cells. *BMC Genomics* **15**, 19 (2014).
150. O. H. Tam *et al.*, Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* **453**, 534-538 (2008).
151. T. Watanabe *et al.*, Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* **453**, 539-543 (2008).
152. R. Song *et al.*, Male germ cells express abundant endogenous siRNAs. *Proc Natl Acad Sci U S A* **108**, 13159-13164 (2011).
153. J. Ma *et al.*, MicroRNA activity is suppressed in mouse oocytes. *Curr Biol* **20**, 265-270 (2010).
154. N. Suh *et al.*, MicroRNA function is globally suppressed in mouse oocytes and early embryos. *Curr Biol* **20**, 271-277 (2010).
155. M. Flemr *et al.*, A retrotransposon-driven dicer isoform directs endogenous small interfering RNA production in mouse oocytes. *Cell* **155**, 807-816 (2013).
156. X. Zhang *et al.*, Mechanisms of small RNA generation from cis-NATs in response to environmental and developmental cues. *Mol Plant* **6**, 704-715 (2013).
157. K. H. Ling *et al.*, Derivation of an endogenous small RNA from double-stranded Sox4 sense and natural antisense transcripts in the mouse brain. *Genomics* **107**, 88-99 (2016).
158. O. E. Andrews, D. J. Cha, C. Wei, J. G. Patton, RNAi-mediated gene silencing in zebrafish triggered by convergent transcription. *Sci Rep* **4**, 5222 (2014).
159. B. Czech *et al.*, An endogenous small interfering RNA pathway in Drosophila. *Nature* **453**, 798-802 (2008).
160. M. Ghildiyal *et al.*, Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. *Science* **320**, 1077-1081 (2008).

161. M. Gullerova, N. J. Proudfoot, Cohesin complex promotes transcriptional termination between convergent genes in S. pombe. *Cell* **132**, 983-995 (2008).
162. M. Gullerova, D. Moazed, N. J. Proudfoot, Autoregulation of convergent RNAi genes in fission yeast. *Genes Dev* **25**, 556-568 (2011).
163. H. C. Lee *et al.*, Diverse pathways generate microRNA-like RNAs and Dicer-independent small interfering RNAs in fungi. *Mol Cell* **38**, 803-814 (2010).
164. Y. Dang, L. Li, W. Guo, Z. Xue, Y. Liu, Convergent transcription induces dynamic DNA methylation at disiRNA loci. *PLoS Genet* **9**, e1003761 (2013).
165. M. L. Hastings, C. Milcarek, K. Martincic, M. L. Peterson, S. H. Munroe, Expression of the thyroid hormone receptor gene, erbAalpha, in B lymphocytes: alternative mRNA processing is independent of differentiation but correlates with antisense RNA levels. *Nucleic Acids Res* **25**, 4296-4300 (1997).
166. M. Beltran *et al.*, A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. *Genes Dev* **22**, 756-769 (2008).
167. M. M. Scotti, M. S. Swanson, RNA mis-splicing in disease. *Nat Rev Genet* **17**, 19-32 (2016).
168. L. Sehgal *et al.*, FAS-antisense 1 lncRNA and production of soluble versus membrane Fas in B-cell lymphoma. *Leukemia* **28**, 2376-2387 (2014).
169. I. Gonzalez *et al.*, A lncRNA regulates alternative splicing via establishment of a splicing-specific chromatin signature. *Nat Struct Mol Biol* **22**, 370-376 (2015).
170. A. S. Morrissy, M. Griffith, M. A. Marra, Extensive relationship between antisense transcription and alternative splicing in the human genome. *Genome Res* **21**, 1203-1212 (2011).
171. C. H. Jen, I. Michalopoulos, D. R. Westhead, P. Meyer, Natural antisense transcripts with coding capacity in Arabidopsis may have a regulatory role that is not linked to double-stranded RNA degradation. *Genome Biol* **6**, R51 (2005).
172. E. Zubko, A. Kunova, P. Meyer, Sense and antisense transcripts of convergent gene pairs in Arabidopsis thaliana can share a common polyadenylation region. *PLoS One* **6**, e16769 (2011).
173. X. Wu *et al.*, Genome-wide landscape of polyadenylation in Arabidopsis provides evidence for extensive alternative polyadenylation. *Proc Natl Acad Sci U S A* **108**, 12533-12538 (2011).
174. V. Tripathi *et al.*, Long noncoding RNA MALAT1 controls cell cycle progression by regulating the expression of oncogenic transcription factor B-MYB. *PLoS Genet* **9**, e1003368 (2013).
175. X. Zong *et al.*, Natural antisense RNA promotes 3' end processing and maturation of MALAT1 lncRNA. *Nucleic Acids Res* **44**, 2898-2908 (2016).
176. K. Matsui *et al.*, Natural antisense transcript stabilizes inducible nitric oxide synthase messenger RNA in rat hepatocytes. *Hepatology* **47**, 686-697 (2008).
177. E. Yoshigai *et al.*, Natural antisense transcript-targeted regulation of inducible nitric oxide synthase mRNA levels. *Nitric Oxide* **30**, 9-16 (2013).
178. J. Sun *et al.*, Long noncoding RNA FGFR3-AS1 promotes osteosarcoma growth through regulating its natural antisense transcript FGFR3. *Mol Biol Rep* **43**, 427-436 (2016).
179. W. Y. Su *et al.*, Bidirectional regulation between WDR83 and its natural antisense transcript DHPS in gastric cancer. *Cell Res* **22**, 1374-1389 (2012).

180. T. B. Hansen *et al.*, miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO J* **30**, 4414-4422 (2011).
181. T. B. Hansen *et al.*, Natural RNA circles function as efficient microRNA sponges. *Nature* **495**, 384-388 (2013).
182. M. A. Faghihi *et al.*, Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat Med* **14**, 723-730 (2008).
183. M. A. Faghihi *et al.*, Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biol* **11**, R56 (2010).
184. G. Q. Wang *et al.*, Sirt1 AS lncRNA interacts with its mRNA to inhibit muscle formation by attenuating function of miR-34a. *Sci Rep* **6**, 21865 (2016).
185. O. Borsani, J. Zhu, P. E. Verslues, R. Sunkar, J. K. Zhu, Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. *Cell* **123**, 1279-1291 (2005).
186. W. Gao, W. Liu, M. Zhao, W. X. Li, NERF encodes a RING E3 ligase important for drought resistance and enhances the expression of its antisense gene NFYA5 in Arabidopsis. *Nucleic Acids Res* **43**, 607-617 (2015).
187. X. Zheng, V. Valakh, A. Diantonio, Y. Ben-Shahar, Natural antisense transcripts regulate the neuronal stress response and excitability. *Elife* **3**, e01849 (2014).
188. M. E. Donaldson, B. J. Saville, Natural antisense transcripts in fungi. *Mol Microbiol* **85**, 405-417 (2012).
189. M. E. Donaldson, B. J. Saville, Ustilago maydis natural antisense transcript expression alters mRNA stability and pathogenesis. *Mol Microbiol* **89**, 29-51 (2013).
190. M. Wery *et al.*, Nonsense-Mediated Decay Restricts LncRNA Levels in Yeast Unless Blocked by Double-Stranded RNA Structure. *Mol Cell* **61**, 379-392 (2016).
191. A. Hatzoglou *et al.*, Natural antisense RNA inhibits the expression of BCMA, a tumour necrosis factor receptor homologue. *BMC Mol Biol* **3**, 4 (2002).
192. A. K. Ebralidze *et al.*, PU.1 expression is modulated by the balance of functional sense and antisense RNAs regulated by a shared cis-regulatory element. *Genes Dev* **22**, 2085-2092 (2008).
193. C. Carrieri *et al.*, Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature* **491**, 454-457 (2012).
194. A. Indrieri *et al.*, Synthetic long non-coding RNAs [SINEUPs] rescue defective gene expression in vivo. *Sci Rep* **6**, 27315 (2016).
195. S. Zucchelli *et al.*, SINEUPs: A new class of natural and synthetic antisense long non-coding RNAs that activate translation. *RNA Biol* **12**, 771-779 (2015).
196. N. T. Tran *et al.*, The AS-RBM15 lncRNA enhances RBM15 protein translation during megakaryocyte differentiation. *EMBO Rep* **17**, 887-900 (2016).
197. M. Jabnoune *et al.*, A rice cis-natural antisense RNA acts as a translational enhancer for its cognate mRNA and contributes to phosphate homeostasis and plant fitness. *Plant Cell* **25**, 4166-4182 (2013).

198. J. Camblong, N. Iglesias, C. Fickentscher, G. Dieppois, F. Stutz, Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in S. cerevisiae. *Cell* **131**, 706-717 (2007).
199. C. F. Hongay, P. L. Grisafi, T. Galitski, G. R. Fink, Antisense transcription controls cell fate in Saccharomyces cerevisiae. *Cell* **127**, 735-745 (2006).
200. M. Castelnuovo *et al.*, Role of histone modifications and early termination in pervasive transcription and antisense-mediated gene silencing in yeast. *Nucleic Acids Res* **42**, 4348-4362 (2014).
201. L. Wang *et al.*, 3' Untranslated regions mediate transcriptional interference between convergent genes both locally and ectopically in Saccharomyces cerevisiae. *PLoS Genet* **10**, e1004021 (2014).
202. T. W. Nilsen, Measuring the length of poly(A) tails. *Cold Spring Harb Protoc* **2015**, 413-418 (2015).
203. A. Jänicke, J. Vancuylenberg, P. R. Boag, A. Traven, T. H. Beilharz, ePAT: a simple method to tag adenylated RNA to measure poly(A)-tail length and other 3' RACE applications. *RNA* **18**, 1289-1295 (2012).
204. C. Romá-Mateo *et al.*, Phylogenetic and genetic linkage between novel atypical dual-specificity phosphatases from non-metazoan organisms. *Mol Genet Genomics* **285**, 341-354 (2011).
205. N. Alic, V. J. Higgins, I. W. Dawes, Identification of a Saccharomyces cerevisiae gene that is required for G1 arrest in response to the lipid oxidation product linoleic acid hydroperoxide. *Mol Biol Cell* **12**, 1801-1810 (2001).
206. E. Garí, L. Piedrafita, M. Aldea, E. Herrero, A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in Saccharomyces cerevisiae. *Yeast* **13**, 837-848 (1997).
207. M. K. Doma, R. Parker, Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* **440**, 561-564 (2006).
208. T. Tsuboi *et al.*, Dom34:hbs1 plays a general role in quality-control systems by dissociation of a stalled ribosome at the 3' end of aberrant mRNA. *Mol Cell* **46**, 518-529 (2012).
209. B. Dichtl, A. Stevens, D. Tollervey, Lithium toxicity in yeast is due to the inhibition of RNA processing enzymes. *EMBO J* **16**, 7184-7195 (1997).
210. O. Brandman *et al.*, A ribosome-bound quality control complex triggers degradation of nascent peptides and signals translation stress. *Cell* **151**, 1042-1054 (2012).
211. Q. Defenouillère *et al.*, Cdc48-associated complex bound to 60S particles is required for the clearance of aberrant translation products. *Proc Natl Acad Sci U S A* **110**, 5046-5051 (2013).
212. M. H. Bengtson, C. A. Joazeiro, Role of a ribosome-associated E3 ubiquitin ligase in protein quality control. *Nature* **467**, 470-473 (2010).
213. N. R. Guydosh, R. Green, Dom34 rescues ribosomes in 3' untranslated regions. *Cell* **156**, 950-962 (2014).
214. J. I. Roop, K. C. Chang, R. B. Brem, Polygenic evolution of a sugar specialization trade-off in yeast. *Nature* **530**, 336-339 (2016).
215. F. Sinturel *et al.*, Cytoplasmic Control of Sense-Antisense mRNA Pairs. *Cell Rep* **12**, 1853-1864 (2015).
216. J. Bernstein, D. N. Patterson, G. M. Wilson, E. A. Toth, Characterization of the essential activities of Saccharomyces cerevisiae Mtr4p, a 3'->5' helicase partner of the nuclear exosome. *J Biol Chem* **283**, 4930-4942 (2008).

217. W. K. Ma *et al.*, Recruitment, Duplex Unwinding and Protein-Mediated Inhibition of the Dead-Box RNA Helicase Dbp2 at Actively Transcribed Chromatin. *J Mol Biol* **428**, 1091-1106 (2016).
218. J. A. Arribere, W. V. Gilbert, Roles for transcript leaders in translation and mRNA decay revealed by transcript leader sequencing. *Genome Res* **23**, 977-987 (2013).
219. L. Decourty *et al.*, Long open reading frame transcripts escape nonsense-mediated mRNA decay in yeast. *Cell Rep* **6**, 593-598 (2014).
220. L. Chen *et al.*, Structure of the Dom34-Hbs1 complex and implications for no-go decay. *Nat Struct Mol Biol* **17**, 1233-1240 (2010).
221. V. P. Pisareva, M. A. Skabkin, C. U. Hellen, T. V. Pestova, A. V. Pisarev, Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. *EMBO J* **30**, 1804-1817 (2011).
222. C. J. Shoemaker, R. Green, Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. *Proc Natl Acad Sci U S A* **108**, E1392-1398 (2011).
223. N. Shcherbik, T. A. Chernova, Y. O. Chernoff, D. G. Pestov, Distinct types of translation termination generate substrates for ribosome-associated quality control. *Nucleic Acids Res*, (2016).
224. K. Ikeuchi, T. Inada, Ribosome-associated Asc1/RACK1 is required for endonucleolytic cleavage induced by stalled ribosome at the 3' end of nonstop mRNA. *Sci Rep* **6**, 28234 (2016).
225. D. J. Young, N. R. Guydosh, F. Zhang, A. G. Hinnebusch, R. Green, Rli1/ABCE1 Recycles Terminating Ribosomes and Controls Translation Reinitiation in 3'UTRs In Vivo. *Cell* **162**, 872-884 (2015).
226. H. A. Meijer, E. M. Smith, M. Bushell, Regulation of miRNA strand selection: follow the leader? *Biochem Soc Trans* **42**, 1135-1140 (2014).
227. F. J. Vizeacoumar, J. C. Torres-Guzman, Y. Y. Tam, J. D. Aitchison, R. A. Rachubinski, YHR150w and YDR479c encode peroxisomal integral membrane proteins involved in the regulation of peroxisome number, size, and distribution in Saccharomyces cerevisiae. *J Cell Biol* **161**, 321-332 (2003).
228. M. Gentzsch, W. Tanner, The PMT gene family: protein O-glycosylation in Saccharomyces cerevisiae is vital. *EMBO J* **15**, 5752-5759 (1996).
229. T. Izawa *et al.*, Roles of dom34:hbs1 in nonstop protein clearance from translocators for normal organelle protein influx. *Cell Rep* **2**, 447-453 (2012).
230. X. Zhang *et al.*, Plant biology. Suppression of endogenous gene silencing by bidirectional cytoplasmic RNA decay in Arabidopsis. *Science* **348**, 120-123 (2015).
231. K. Nishikura, A-to-I editing of coding and non-coding RNAs by ADARs. *Nat Rev Mol Cell Biol* **17**, 83-96 (2016).
232. B. J. Liddicoat *et al.*, RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself. *Science* **349**, 1115-1120 (2015).
233. S. F. Mitchell, S. Jain, M. She, R. Parker, Global analysis of yeast mRNPs. *Nat Struct Mol Biol* **20**, 127-133 (2013).
234. L. Schenk, D. M. Meinel, K. Strässer, A. P. Gerber, La-motif-dependent mRNA association with Slf1 promotes copper detoxification in yeast. *RNA* **18**, 449-461 (2012).

235. S. Schwartz *et al.*, High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis. *Cell* **155**, 1409-1421 (2013).
236. J. M. Tkach *et al.*, Dissecting DNA damage response pathways by analysing protein localization and abundance changes during DNA replication stress. *Nat Cell Biol* **14**, 966-976 (2012).
237. S. Prieto, B. J. de la Cruz, I. E. Scheffler, Glucose-regulated turnover of mRNA and the influence of poly(A) tail length on half-life. *J Biol Chem* **275**, 14155-14166 (2000).
238. K. A. Braun, K. M. Dombek, E. T. Young, Snf1-Dependent Transcription Confers Glucose-Induced Decay upon the mRNA Product. *Mol Cell Biol* **36**, 628-644 (2016).
239. C. Tachibana *et al.*, Combined global localization analysis and transcriptome data identify genes that are directly coregulated by Adr1 and Cat8. *Mol Cell Biol* **25**, 2138-2146 (2005).
240. L. M. Buu, L. T. Jang, F. J. Lee, The yeast RNA-binding protein Rbp1p modifies the stability of mitochondrial porin mRNA. *J Biol Chem* **279**, 453-462 (2004).
241. L. C. Chang, F. J. Lee, The RNA helicase Dhh1p cooperates with Rbp1p to promote porin mRNA decay via its non-conserved C-terminal domain. *Nucleic Acids Res* **40**, 1331-1344 (2012).
242. K. P. Byrne, K. H. Wolfe, The Yeast Gene Order Browser: combining curated homology and syntenic context reveals gene fate in polyploid species. *Genome Res* **15**, 1456-1461 (2005).
243. P. Hentges, B. Van Driessche, L. Tafforeau, J. Vandenhaute, A. M. Carr, Three novel antibiotic marker cassettes for gene disruption and marker switching in Schizosaccharomyces pombe. *Yeast* **22**, 1013-1019 (2005).
244. A. L. Goldstein, J. H. McCusker, Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. *Yeast* **15**, 1541-1553 (1999).
245. L. Hatfield, C. A. Beelman, A. Stevens, R. Parker, Mutations in trans-acting factors affecting mRNA decapping in Saccharomyces cerevisiae. *Mol Cell Biol* **16**, 5830-5838 (1996).
246. J. S. Anderson, R. P. Parker, The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. *EMBO J* **17**, 1497-1506 (1998).
247. R. S. Sikorski, P. Hieter, A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics* **122**, 19-27 (1989).
248. C. Janke *et al.*, A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* **21**, 947-962 (2004).
249. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011* **17**, (2011).
250. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25 (2009).

CHAPTER VI | ANNEXES

1 | List of tables

CHAPTER I

CHAPTER III

CHAPTER IV

2 | List of figures

CHAPTER I