Investigating the function of Fal1/eIF4AIII
in *S. pombe*

by

ANDREA COLELLA

Supervised by Dr. Saverio Brogna

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

School of Biosciences
College of Life and Enviromental Sciences
The University of Birmingham
September 2019

Summary

The Exon Junction Complex (EJC) is a protein complex formed by a core of three proteins, eIF4AIII, MAGO, Y14, which is cotranscriptionally deposited on spliced mRNAs in mammalian cells, typically 24 nucleotides before the exon-exon boundaries. EJC deposition is believed to occur in evolutionarily divergent organisms, possibly all those that express the three core proteins. However, recent observations suggest that EJC factors are also associated with mRNAs derived from intron-less genes.

My PhD project aimed to clarify whether the EJC components are recruited on mRNAs independently of splicing in *Schizosaccharomyces pombe*. The data I have gathered indicate that Fal1, the ortholog of eIF4AIII in *S. pombe,* is associated with Pol II transcribed genes independently of the presence of an intron. To investigate whether Fal1 has a direct role in transcription or pre-mRNA processing, I engineered an Auxin-Inducible Degron (AID) system to rapidly deplete Fal1. However, I found that the application of the system causes a decrease in the Fal1 protein levels even in the absence of auxin, accompanied by severe growth defects. Therefore, I have discussed the limiting factors which might prevent the use of the AID system in *S. pombe*.

I have also investigated whether an EJC is formed in fission yeast. My data suggest that a mammalian-like EJC might not be formed in *S. pombe,* and that Fal1 can exist independently of MAGO and Y14, even though these two proteins are likely to form a heterodimer also in this organism, which in mammalian cells is sufficient to stabilize the association of eIF4AIII with splice junctions.

Acknowledgements

Abbreviations used in this thesis

AA	Anchor-Away
AID	Auxin-Inducible degron
BPS	Branch-point sequence
BSA	Bovine serum albumin
CBC	Cap-binding complex
CBP	Cap-binding protein
CDS	Coding sequence
CTD	C-terminal domain
ChIP	Chromatin Immunoprecipitation
DEPC	Diethylpyrocarbonate
DSE	Downstream sequence element
DSR	Determinant of selective removal
EDTA	Ethylenediaminetetraacetic acid
eIF	Eukaryotic translation initiation factor
EJC	Exon Junction Complex
GFP	Green Fluorescent Protein
KO	Knock-Out
mRNP	Messenger ribonucleoprotein particle
NMD	Nonsense-mediated mRNA decay
NPC	Nuclear pore complex
ORF	Open reading frame
PABP	Poly(A) binding protein

PBS	Phosphate buffered saline
PEG	Polyethylene glycol
PTC	Premature termination codon
PMSF	Phenylmethaneshulfonyl fluoride
RNA Pol (I, II, III)	RNA polymerases (I, II, III)
Rpb	DNA-directed RNA polymerase II subunit
SDS	Sodium dodecyl sulphate
SS	Splice site
Tris	Tris (hydroxymethyl) aminomethane
TSS	Transcription start site
UPF	Up-frameshift
UTR	Untranslated region

Table of Contents

Chapter 1: Introduction

1.1 - The process of gene expression in eukaryotes

Gene expression is a fundamental biological mechanism that converts the genetic information deposited in the DNA sequence of protein-coding genes into messenger RNAs (mRNAs), which are then translated into the several thousands of proteins that make and operate a living organism. In eukaryotes, protein-coding genes are transcribed by RNA Polymerase II (RNA Pol II) into transcripts that are extensively modified, mostly co-transcriptionally, by capping, splicing and polyadenylation (Bentley, 2014, Moore and Proudfoot, 2009). Contrary to prokaryotes, where the mRNA is immediately available to the ribosome for its translation, in eukaryotic cells the mRNA has to be transported from the nucleus to the cytoplasm, passing through nuclear pore complexes (NPCs), where it will be translated by the ribosome, into a protein (Figure 1.0). Due to its complexity, regulation of gene expression can occur at different stages, such as chromatin remodeling, transcription, mRNA processing, nuclear export and translation.
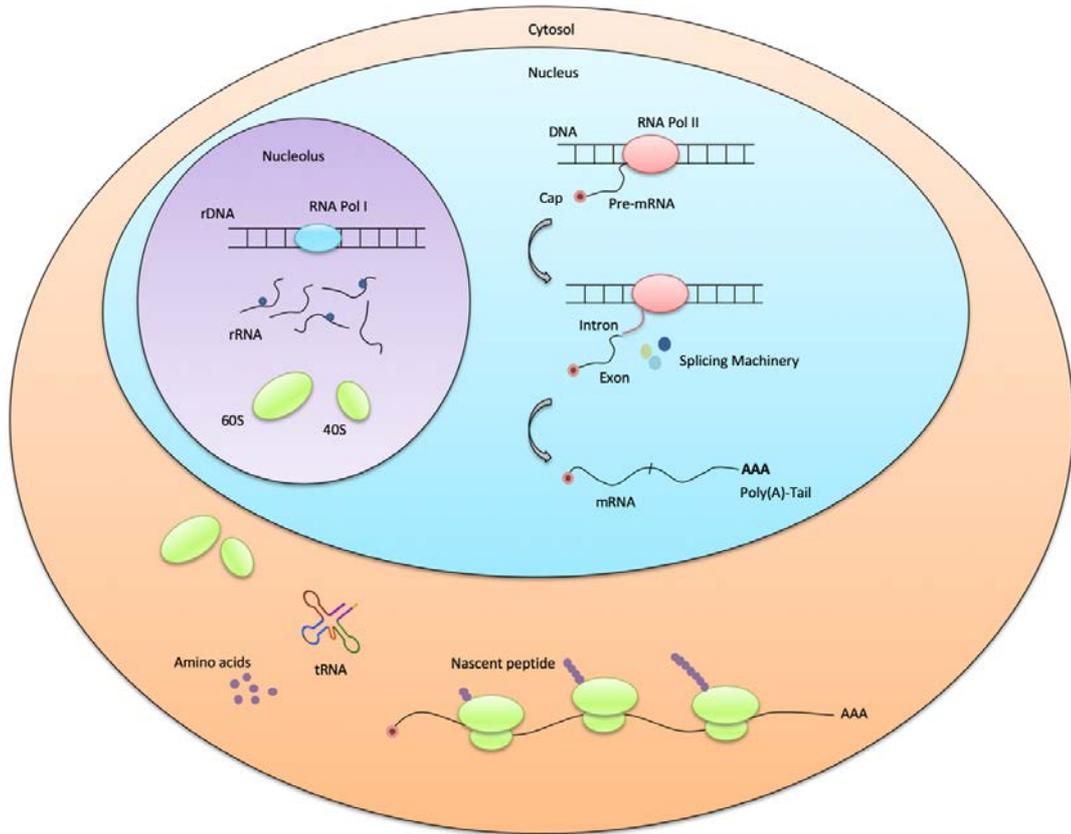
Figure 1.0 - The process of gene expression in eukaryotes. Protein-coding genes are transcribed by RNA Polymerase II (RNA Pol II) in the nucleus where the pre-mRNA is co-transcriptionally processed via capping, splicing and polyadenylation. RNA Polymerase I (RNA Pol I) transcribes the ribosomal DNA (rDNA) in the nucleolus, producing the rRNAs, which together with ribosomal proteins form the ribosome. The mRNA, exported to the cytoplasm is used as a template by the ribosomes to synthesize the proteins.

1.2 - Transcription

Transcription is the first step of gene expression, through which the information stored in the DNA is copied into RNA by RNA polymerases. Protein-coding genes are transcribed by RNA Pol II, which is a multiprotein complex, composed of several subunits varying from 8 to 12 according to the organism. RNA Pol II from *Saccharomyces cerevisiae* is formed by 12 subunits, denoted from Rpb1 to Rpb12, in order of decreasing molecular weight. The subunits Rpb1, -2, -3, -6 and 11 are the counterpart of the bacterial subunits β', β, α^I, α^{II}, ω, respectively (Ebright, 2000). Additionally, the subunits Rpb5, -6, -8, -10 and 12 are also present in RNA Pol I and III, confirming the high level of homology between the three enzymes. The subunits Rpb4 and Rpb7 form instead a heterodimer, which associates reversibly with RNA Pol II; in *S. cerevisiae*, these two subunits are present in the complex during the early stationary phase, while in exponential growth, they are present in substoichiometric quantity, suggesting that their recruitment occurs under poor growth conditions (Choder, 2004, Craighead et al., 2002).

The largest subunit of RNA Pol II, Rbp1, is characterized by a unique C-terminal domain (CTD), formed by repeats of the heptad $Y_1S_2P_3T_4S_5P_6S_7$; the number of these repeats can vary in different organisms, according to their complexity; for example, the CTD in human consists of 52 repeats, in *D. melanogaster* of 45, in *S. pombe* of 29 and in *S. cerevisiae* of 27. The CTD can be phosphorylated on the heptad residues during different stages of transcription and, as I will explain in the next paragraphs, its phosphorylation changes between transcription stages and is linked to pre-mRNA processing (Hsin and Manley, 2012).

1.3 - Pre-mRNA processing is coupled to transcription

The maturation of the pre-mRNA typically involves three major covalent modifications: 5'-end capping, splicing and 3'-end cleavage and polyadenylation (Bentley, 2014, Proudfoot, 2000). Intriguingly each of these modifications is followed by the deposition of specific factors on the transcript, which determine downstream later processes of gene expression. As a consequence, transcripts are never in the form of naked RNA and are always associated with several proteins, forming messenger ribonucleoprotein complexes (mRNP) (Singh et al., 2015).

The CTD of the Rpb1 subunit plays a pivotal role in the transition through different stages of transcription and pre-mRNA processing. In vivo, RNA Pol II is present in two different forms: RNAPIIA, with a non-phosphorylated CTD, and RNAPII0, which is highly phosphorylated (Weeks et al., 1993). Using polytene chromosome of *D. melanogaster*, it was initially observed that RNAPIIA associated with inactive genes and promoters, whereas, the RNAPII0 was found at the sites of active transcription (O'Brien et al., 1994, Weeks et al., 1993). The phosphorylation status of RNA Pol II and its biological meaning have become clear thanks to Chromatin Immunoprecipation (ChIP) technique which allows identifying protein-DNA interactions in vivo (Collas, 2010, Gilmour and Lis, 1984, Gilmour and Lis, 1985)

At different stages of transcription, RNA Pol II shows a distinctive pattern of phosphorylation, which permits recruiting factors involved in the histone modifications, in the transition from initiation to terminations and pre-mRNA processing (Buratowski, 2009). Phosphorylation occurs on the heptad repeats of the CTD of Rpb1. Ser5 and Ser2 represent the most studied phosphorylation sites. In mammals Ser*5* is phosphorylated by the cyclin-dependent kinase 7, CDK7, a subunit of the transcription factor IIH (TF2H); this residue is

highly phosphorylated in RNA Pol II loaded within the first few hundred nucleotides of the gene (Komarnitsky et al., 2000). This phosphorylation appears to play a critical role in transcription initiation and the recruitment of enzymes needed for the pre-mRNA capping (Hsin and Manley, 2012). Conversely, Ser5 is dephosphorylated in downstream regions (Phatnani and Greenleaf, 2006). Instead, the phosphorylation on Ser2 by p-TEFb (which is composed by CDK9 and cyclin T) in mammals, is fundamental to stimulate the transcription elongation and for the recruitment of the splicing and poly(A) machinery (Ahn et al., 2004). As a consequence, the phosphorylation levels of Ser2 increases progressively as RNA Pol II moves along the coding sequence of the gene (Buratowski, 2009).

In the next paragraphs of this section, I will explain more in detail the mechanism of pre-mRNA processing and how this is connected with transcription.

1.3.1 - 5' pre-mRNA capping

When the 5'-end of the transcript emerges through the RNA Pol II exit channel, typically as its length reaches 20-30 nucleotides, it is promptly modified by the addition of a 7-methylguanosine (m^7G), a guanine nucleotide attached to the 5'-end of the nascent transcript by an unconventional 5' to 5' triphosphate bridge, which is also modified by methylation of the base at position-7 (Moore, 2005, Noe Gonzalez et al., 2018). These chemical modifications require three enzymatic activities: an RNA triphosphatase, which cleaves the 5' triphosphate of the pre-mRNA to a diphosphate, a guanylyl-transferase which adds the guanosine monophosphate (GMP) to the 5', and the methyltransferase which methylates the guanine (Furuichi et al., 1976, Shatkin, 1976). The capping enzymes are recruited to the 5' end of the nascent mRNAs through interactions with the CTD of the Rbp1 subunit (Buratowski, 2003). The association of the capping enzymes to the CTD is stimulated by the

phosphorylation of Ser5 (Martinez-Rucobo et al., 2015). The CTD appears essential for the mechanism of capping; for the mRNAs derived by chimeric gene constructs in which a gene normally transcribed by RNA Pol II is under the control of an RNA Pol I or Pol III promoter, capping is prevented or partially repressed (Smale and Tjian, 1985). The enzymatic activities required for the capping are evolutionarily conserved (Yue et al., 1997). However, the organization of the different subunits and the modalities by which they bind the phosphorylated CTD can vary according to the species (Pei and Shuman, 2002). For example, in *S. cerevisiae*, the three enzymatic activities are on different polypeptides; the methyltransferase and the guanylyl-transferase components bind the CTD of RNA Pol II independently of each other, while the RNA triphosphatase does not bind the CTD directly but is associated to the guanylyl-transferase. In metazoans, the capping enzymes are formed by a bifunctional RNA triphosphatase/guanylyl-transferase and a methyltransferase. The guanylyl-transferase domain, located at the C-terminal, binds the CTD of RNA Pol II, but the triphosphatase does not (Pei and Shuman, 2002). The cap is then recognized by a nuclear cap-binding complex (nCBC), consisting of a cap-binding subunit 20 (CBP20), which exploits its ribonucleoprotein (RNP) domain to bind the cap, and an auxiliary cap-binding protein 80 (CBP80), whose function is to stabilize the interaction between CBP20 on the cap (Topisirovic et al., 2011). Besides, CBP80 appears to interacts with REF (RNA and export factor binding protein), and with the CTD of RNA Pol II (Topisirovic et al., 2011)

The nCBC not only stabilizes the transcript, but it appears to play a pivotal role in numerous aspects of gene expression, stimulating pre-mRNA splicing and nuclear export (Topisirovic et al., 2011). In the cytoplasm the heterodimer CBP80 and CBP20 is believed to persist on the mRNA until the first round of translation; then it is replaced by the eukaryotic Translation

Initiation Factor 4E (eIF4E), which remains associated to the transcript for the remaining rounds of translation (Maquat et al., 2010).

1.3.2 - pre-mRNA splicing

Before being exported to the cytosol, introns are removed from the mRNA via splicing. In *S. cerevisiae* splicing appears to occur on the nascent-mRNAs, before RNA Pol II has completed the transcription of the gene. This notion derives from single-molecule RNA-sequencing (SMIT) experiments, which indicate that splicing happens when Pol II is only a few hundred nucleotides downstream the intron, indicating that as soon introns emerge from RNA Pol II they are rapidly excised (Oesterreich et al., 2016). However, this conclusion contrasts with previous observations that reported that splicing can happen co-transcriptionally in yeast only for those genes which have a long second exon, while for the majority of the genes introns are removed predominantly post-transcriptionally (Tardiff et al., 2006). The SMIT experiments in *S. pombe* confirm that also in fission yeast, splicing might occur early in transcription when RNA Pol II is just downstream of the 3' splice site (Oesterreich et al., 2016).

In mammalian cells, the sequencing of the newly synthesized transcripts from human B-cells, via 4-thio-uridine (4sU) labeling, has shown that the frequency of co-transcriptional splicing events is approximately 75% (Windhager et al., 2012). A similar result was obtained through sequencing of total RNA from human brain (Ameur et al., 2011). That said, some human long-non coding RNAs (lncRNAs) or alternative exons, are spliced post-transcriptionally or not spliced at all (Tilgner et al., 2012).

The CTD of the Rpb1 subunit also plays a crucial role in splicing; during transcription, the phosphorylated form of Ser2 allows the recruitment of splicing factors involved in the

definition of the splice sites and in the assembly of the spliceosomal machinery (Corden and Patturajan, 1997, Hsin and Manley, 2012).

Splicing requires regulatory cis-acting RNA elements for its functioning (Xiao et al., 2007). Each intron is characterized by a set of distinctive sequences which are: the 5' splice site also known as donor site, generally defined by a dinucleotide GU, the 3' splice site, characterized by a dinucleotide AG, and a branch point adenosine, typically found 20-25 bases upstream of the 3' splice site (Wahl et al., 2009)

Exploiting these consensus motifs, splicing proceeds through two consecutive reactions of transesterification: in the first, the ribose 2'OH group of an adenosine located at the branch point, attacks the 5'phosphate of the G located in the 5'splice site, causing the formation of a lariat intermediate and a free 5' exon. In the second reaction of transesterification, the 3'-OH of the last residue of the free exon attacks with the 5' phosphate of the first residue of the second exon, thus joining the two exons and releasing the lariat structure (Figure 1.1).
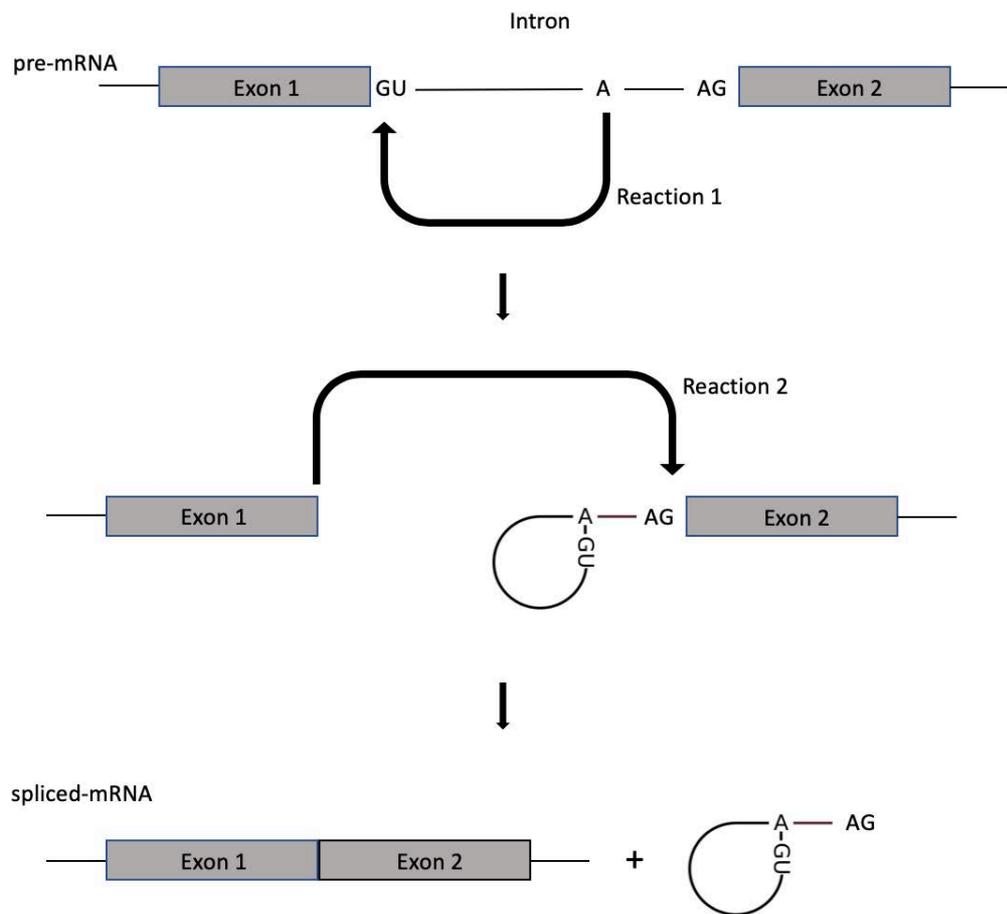
Figure 1.1 - Splicing occurs through two transesterification reactions. In the first reaction of transesterification, the 2'OH group of an adenosine's ribose located at the branch point attacks the 5'phosphate of the G located in the 5'SS causing the formation of a lariat intermediate and a free 5' exon. In the second reaction of transesterification, the 3' OH of the last residue of the free exon attacks the 5' phosphate of the first residue of the second exon, causing the release of the lariat structure and juxtaposition of the two exons.

To allow intron removal and ligation of the two flanking exons, splicing requires five small nuclear uridine-rich RNAs (U1, U2, U4, U5, U6 snRNAs), which are associated with a number of additional proteins, varying from 6 to 10 for each snRNA, forming the small nuclear ribonucleoprotein particles (snRNP). The snRNPs together with additional non-snRNPs factors, which assist in the splicing reactions, form the spliceosome machinery (Kramer, 1996). The two reactions of transesterification require the formation of several splicing complexes, formed through the association and dissociation of the snRNPs during the different stages of splicing (Figure 1.2). The first spliceosome component to bind the transcript is the U1 snRNP; U1 associates at the 5'splice site of the intron, via base pairing, and at the same time some auxiliary factors, mark the BPS and the 3' splice site, to form the inactive E complex. The recruitment of the U2 snRNP, at the BPS through base pairing, marks the transition from the E complex to the A-complex, or pre-spliceosome. With the recruitment of a pre-assembled trimeric complex formed by U4/U6 and U5 snRNPs, the pre-spliceosomal complex, is converted in the pre-catalytical spliceosome, or complex B. The further dissociation of U1 and U4 is required to form an active splicesome or complex B*; during the first reaction of transesterification, the complex B* is converted in complex C, which leads to the second reaction of transesterification (Wahl et al., 2009).
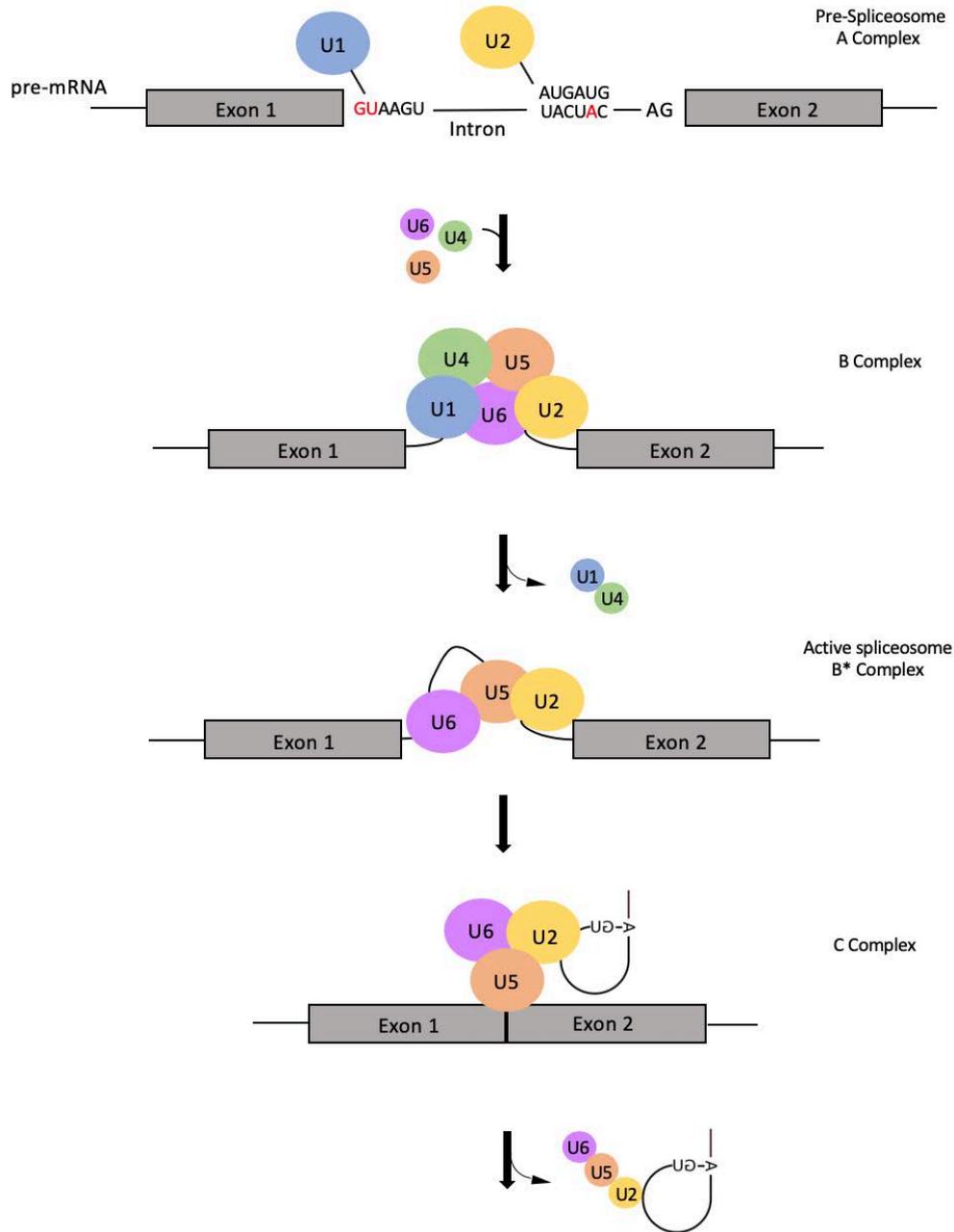
Figure 1.2 - Assembly of spliceosome complexes. snRNA U1 and U2 are recruited on the 5' splice site and BPS respectively, where they form a stable bond via base pairing (A Complex). A trimeric complex, formed by U4/U6 and U5, is then recruited on the Complex A, to generate the B Complex, which is still in an inactive form until the release of U1 and U4 (active spliceosome or B* Complex). With the formation of an active spliceosome, the first transesterification is performed, followed by a second reaction driven by the assembly of a C complex.

1.3.3 - 3' end processing

Cleavage and polyadenylation of the 3'-end of the nascent transcript is an essential step in the maturation of the pre-mRNA. Specific sequences on the transcript dictate the site of cleavage and polyadenylation. In humans, the core upstream element, also known as poly(A) signal, is characterized by an AAUAAA or AUUAAA sequence, and it is found in almost 80% of the polyadenylated transcripts (Proudfoot, 1991). The remaining 20% are characterized by a non-canonical hexamer sequence, occurring mainly in those transcripts modified by alternative cleavage and polyadenylation (Nunes et al., 2010, Tian et al., 2005).

Downstream of the poly(A) signal, two other sequences can be detected; a cleavage site, consisting of a CA dinucleotide, where the endonucleolytic cleavage occurs, and a U-rich or GU-rich motif, also known as downstream sequence element (DSE). These cis-elements are recognized by trans-acting factors; the cleavage and polyadenylation specificity factor (CPSF), contacts the poly(A) signal through its 160 kDa subunit (CPSF-160), whose binding specificity seems to be guaranteed also by CPSF-100, CPSF-73 and CPSF-30 (Proudfoot et al., 2002). The binding of CPSF to the poly(A) signal is strengthened by the Cleavage Stimulation Factor (CstF), which consists of CstF-77, CstF-64 and CstF-50. While CstF-64 contacts the DSE, CstF-77 makes a bridge between CPSF-160 and Cstf-64 and Cstf-50 (Fong and Bentley, 2001). The binding of the two complexes on their respective elements is cooperative since when CstF is bound to the DSE, it enhances the affinity of CPSF for the poly(A) signal and vice-versa (Proudfoot et al., 2002).

The 3'-end of the transcript is defined by the cleavage reaction, which is perfomed by the CPSF-73 endonuclease (Hill et al., 2019, Mandel et al., 2006). *In vitro* experiments demonstrated that Ysh1, the yeast homolog of CPSF-73, is incorporated into an CPF core

complex of eight subunits, which increases the formation of an active catalytic complex, required for the cleavage reaction (Hill et al., 2019).

After the cleavage, the Poly(A)-polymerase (PAP) catalyses the polyadenylation of the transcript (Bienroth et al., 1993, Wahle, 1995, Wahle et al., 1991).

Following cleavage, polyadenylation and release of the mature mRNA from the transcription site, RNA Pol II remains engaged to the DNA and continues transcribing the downstream region for up to several kb before it is released. How RNA Pol II is removed from the DNA template after cleavage and polyadenylation, is still not completely understood. One possibility represented by the "allosteric model", is that RNA Pol II dissociates as a result of a conformational change when it reaches the poly(A) signal, or by loss of elongation factors and subsequent decrease in the processivity (Loya and Reines, 2016). Another model, known as "torpedo" proposes instead that the endonucleolytic cleavage at the poly-(A) site generates an entry site for a 5'-3' exonuclease, XRN2 in Human or Rat1 *in S. cerevisiae*, which reaches the uncapped transcript behind RNA Pol II, degrading the RNA and dislodging the polymerase (Eaton and West, 2018, West et al., 2004). However, new recent observations suggest that the two models could be not mutually exclusive. Cortazar and colleagues reported that downstream the poly-(A) site, RNA Pol II decelerates because of dephosphorylation of the Transcription Elongation Factor 5 (Spt5), which is promoted by a protein phosphatase (PNUTS-PP1) (Cortazar et al., 2019). Through this change in speed, XRN2 can reach RNA Pol II and terminate the process. The model has been defined as "sitting duck torpedo" (Cortazar et al., 2019). Similarly, Eaton and colleagues proposed that allosteric changes, occurring beyond the poly-(A) site and responsible for an RNA Pol II slowing down, allow XRN2 to reach the uncapped transcript to promote termination (Eaton et al., 2020).

1.4 - The Exon Junction Complex (EJC)

1.4.1 - Splicing alters the composition of the mRNPs by deposition of the Exon Junction Complex

Pre-mRNA splicing is thought to alter the composition of the newly synthesized mRNP, by the deposition of a specific set of proteins which form the Exon Junction Complex (EJC) (Le Hir et al., 2001b, Le Hir et al., 2000). This deposition is thought to act as a signature for those transcripts that have been processed by the splicing machinery and it is believed to regulate several downstream biological mechanisms such as nuclear export, mRNA localisation and mRNA decay (Le Hir et al., 2016). In light of this, the EJC seems to explain how an intron, which is removed upon splicing, can still affect translation and nonsense-mediated RNA decay (NMD) in the cytoplasm (Brogna et al., 2016, Le Hir et al., 2003, Shaul, 2017).

The EJC was discovered in 2000 by Le Hir and colleagues who reported that following in vitro splicing in HeLa nuclear extracts, the spliced mRNAs were resistant to RNase H digestion at a specific position, typically 20-24 nucleotides upstream of the exon-exon junction. The same observation was also confirmed in vivo, by injecting intron-containing gene reporters into the oocyte nuclei of *Xenopus*: following dissection, isolation and RNase H treatment of the nuclei, the same pattern of protection was observed, suggesting that the phenomenon was evolutionarily conserved (Le Hir et al., 2000). The authors concluded that the protection was caused by the presence of a specific set of proteins bound to the transcripts at this position.

This initial observation led to further understanding of the factors associated in this particular position. Since the binding was in close vicinity of the exon-exon junctions, the complex which they were forming was named EJC.

Subsequent studies revealed that the EJC is a highly stable hetero-tetramer composed of four pivotal factors, eIF4AIII (DDX48), Y14 (RRM8A), MAGO and Barentz (BTZ or MNL51, or CASC3), and it is surrounded by numerous peripheral components, which affect the destiny of the mRNA during the different stages of gene expression (Bono et al., 2004, Chan et al., 2004, Le Hir et al., 2001b). MAGO and its partner Y14, were initially shown to be crucial factors for the localisation of the *oskar* mRNA to the posterior pole of *D. melanogaster* during oogenesis, thus providing a link between splicing and mRNA localisation (Hachet and Ephrussi, 2001, Hachet and Ephrussi, 2004, Le Hir et al., 2001a). Further discoveries demonstrated that EJC factors were involved in nonsense-mediated mRNA decay (NMD) a surveillance pathway which degrades defective mRNAs harboring premature termination codons (PTC) (Conti and Izaurralde, 2005, Lejeune and Maquat, 2005, Maquat, 2005, Palacios et al., 2004). Two other peripheral factors, Aly/REF and UAP56, enhanced the export of the spliced mRNAs from the nucleus to the cytoplasm (Le Hir et al., 2016).

1.4.2 - eIF4AIII is the core component of the EJC

The anchor point of all the EJC factors is represented by eIF4AIII, which directly contacts the spliced mRNAs, 24 nucleotides before each exon-exon junction in a sequence-independent fashion, and associates with MAGO, Y14, and BTZ (Chan et al., 2004, Zhang and Krainer, 2007). eIF4AIII belongs to the Eukaryotic Translation Initiation Factor family 4A, a group of ATP dependent RNA helicases comprising also eIF4AI and eIF4AII (Cordin et al., 2006). Structurally, eIF4AIII is an Asp-Glu-Ala-Asp DEAD-box helicase: members of this family use ATP hydrolysis to rearrange secondary structures of RNA, in order to facilitate translation initiation, pre-rRNA processing, or to clear away factors bound to nucleic acids (Bono et al., 2006, Le Hir and Andersen, 2008). Like other members of the DEAD-box family, eIF4AIII

has a standard modular structure, characterized by two domains, RecA1 (located at N-terminal) and RecA2 (located at C-terminal), separated by a short linker of 11 amino acids (Figure 1.3A). Biochemically, each RecA contains a parallel β sheet formed by a variable number of β strands, with loops on both the amino and carboxyl-terminal (Caruthers et al., 2000, Cordin et al., 2006). Within the two RecA domains, numerous short and highly conserved motifs, involved in RNA and ATP binding, can be identified (Figure 1.3B). Motif I, also known as Walker A, is characterized by a typical GKT (glycine-lysine-threonine) sequence, and it is involved in ATP binding and hydrolysis. The presence of the lysine allows it to establish interactions with the γ-phosphate of the ATP; if it is substituted with an Arginine the ATP binding, but not its hydrolysis, can occur, while when substituted with an Alanine the ATP binding is compromised (Sharma et al., 2013). The threonine instead coordinates the interaction with Mg^{2+} ions. Motif II, also known as Walker B, contains the DEAD-box sequence and it is involved in the ATP binding and hydrolysis (Walker et al., 1982). In this domain, the glutamate regulates the catalytic activity and the aspartate binds the Mg^{2+} ions. Motif III or serine acetyltransferase (SAT) domain, coordinates with Motif V the interactions between the ATP and the RNA binding motifs. Motif IV contains a conserved residue of phenylalanine, followed by a cysteine, and it is involved in the interaction with RNA. Motif VI is involved in ATP binding and hydrolysis, whose function is explicated by two evolutionary conserved residues of arginine. The Motif QxxR, where glutamine is separated by non-conserved amino acids from an arginine, is a typical signature of DEAD-box helicases and plays a crucial role in RNA binding; the conserved Q and R residues help to orientate the helicase to the RNA surface. Motif Ia and Ib, contribute to RNA binding, interacting with the sugar-phosphate backbone of the RNA (Linder and Jankowsky, 2011, Rocak and Linder, 2004).

The binding of the ATP and RNA to eIF4AIII causes a conformational change of the protein; whereas in absence of ATP, the RecA1 and RecA2 domains do not show a clear orientation and the structure is characterized by an "open" conformation, when ATP and RNA are bound, the two ReC domains move in close vicinity (closed conformation), thus forming a stable anchor point on the mRNA (Bono et al., 2006). Despite eIF4AIII contains all the domains of a typical DEAD box-RNA helicase, it does not function as a canonical translocating helicase, but it is generally recognized as a clamp or a "place holder" for the recruitment of other RNA binding proteins (Bono et al., 2006, Le Hir et al., 2016). Supporting this view, eIF4AIII appears to be functionally distinct from the other two members of the same family, eIF4AI and eIF4AII, both of which are required to melt secondary structures in the 5'-UTRs, facilitating the correct attachment and positioning of the 40S ribosomal subunit. Moreover, in light of the functional difference with eIF4AI and eIF4AII, eIF4AIII primarily localises in the nucleus differently to the others two isoforms of the same family, which accumulate in the cytoplasm (Chan et al., 2004).
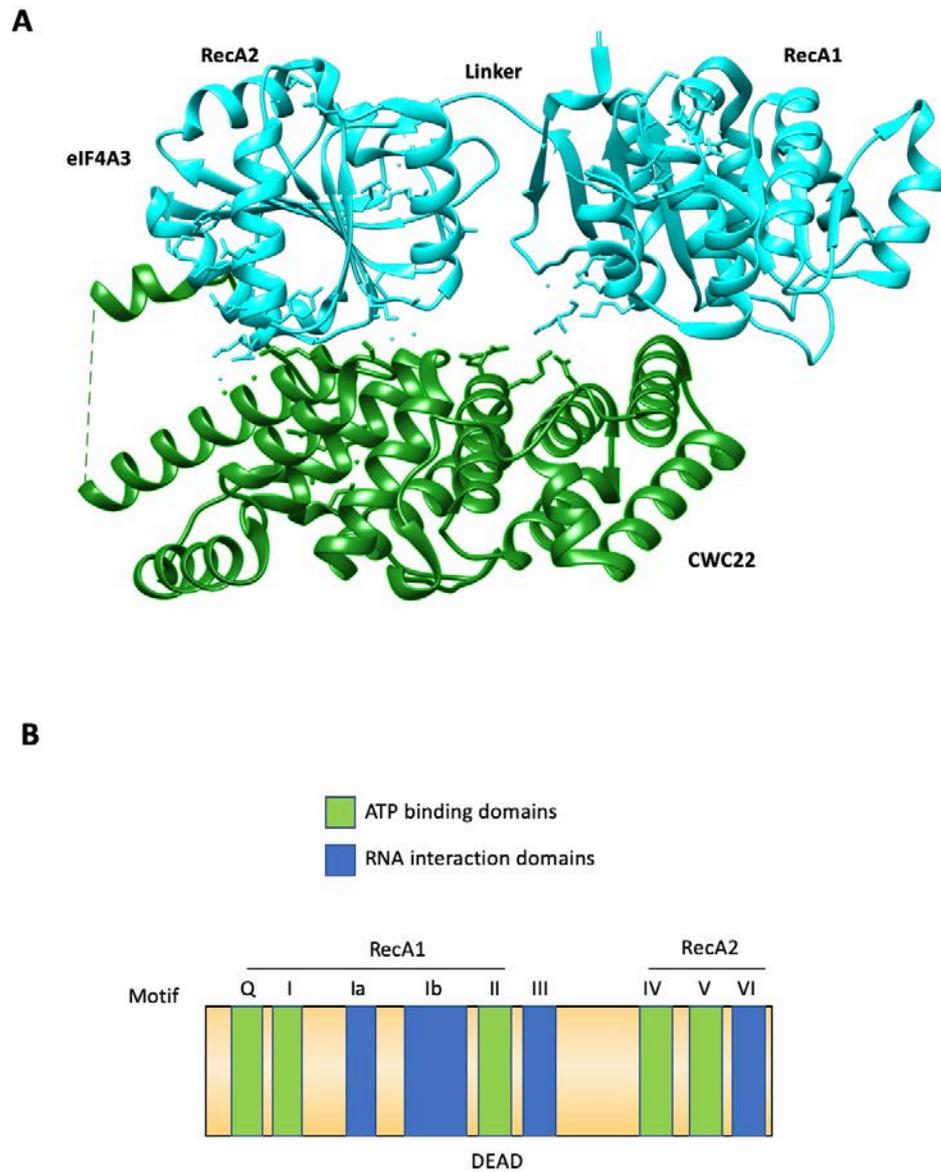
Figure 1.3 - Crystal structure of the human eIF4AIII associated with CWC22. (A) The amino-terminal of eIF4AIII corresponds to the RecA1 domain and the carboxyl-terminal to the RecA2 domain (both are represented in cyan); CWC22, a partner of eIF4AIII in mammalian cells, establishing contacts with RecA2 is in green (Images elaborated using Chimera, PDB: 4C9B ((Buchwald et al., 2013)). (B) Conserved sequence motifs of a typical DEAD-box helicase, involved in ATP binding and RNA binding.

1.4.3 - MAGO and Y14 form a stable heterodimer

MAGO and Y14, which are also known as MAGO NASHI and Tsunagi/RBM8 respectively, are evolutionarily conserved proteins across eukaryotes and they form an obligatory heterodimer in all the organisms in which their association has been biochemically investigated (Gong et al., 2014). Heterodimerisation is required for their biological function (Fribourg et al., 2003, Lau et al., 2003). To keep the structural interactions throughout evolution, two kinds of strategies were evolved: conservation of the same amino-acid residues involved in the heterodimerisation, or coordinated sequence variations in both the proteins (Gong et al., 2014). Supporting this idea, a stable heterodimer of MAGO and Y14 has been confirmed in at least 11 different species, from Human to the plant *Physalis floridiana* (Gong et al., 2014). Y14 does not contact directly the RNA; using an electrophoretic mobility shift assay, neither MAGO or Y14 bind the RNA (Shi and Xu, 2003). This observation is consistent with the crystal structure of the EJC, where eIF4AIII is anchored on the mRNA, and MAGO and Y14 are perpendicular to eIF4A3 (Bono et al., 2006) (Figure 1.4). Whilst eIF4AIII has been identified in all the eukaryotic groups, MAGO and Y14 are absent in organisms with low intron-density such as *S. cerevisiae* or the alga *Cyanidioschyzon merolae* (Bannerman et al., 2018). However, both proteins are present in the fission yeast S. *pombe*, and in the budding yeast *Y. lipolytica,* where the percentage of genes containing introns is approximately 45% and 37% respectively (Boisramé et al., 2019, Kuhn and Käufer, 2003). Y14 and MAGO are involved in NMD in mammalian cells (Gehring et al., 2003). However, in *D. melanoagaster*, where the heterodimer is essential for cell viability, MAGO and Y14 are required for mRNA localisation, but not for NMD (Gatfield et al., 2003, Hachet and Ephrussi, 2001, Hachet and Ephrussi, 2004). MAGO and Y14 are also not required for NMD in *S. pombe* (Wen and Brogna, 2010).
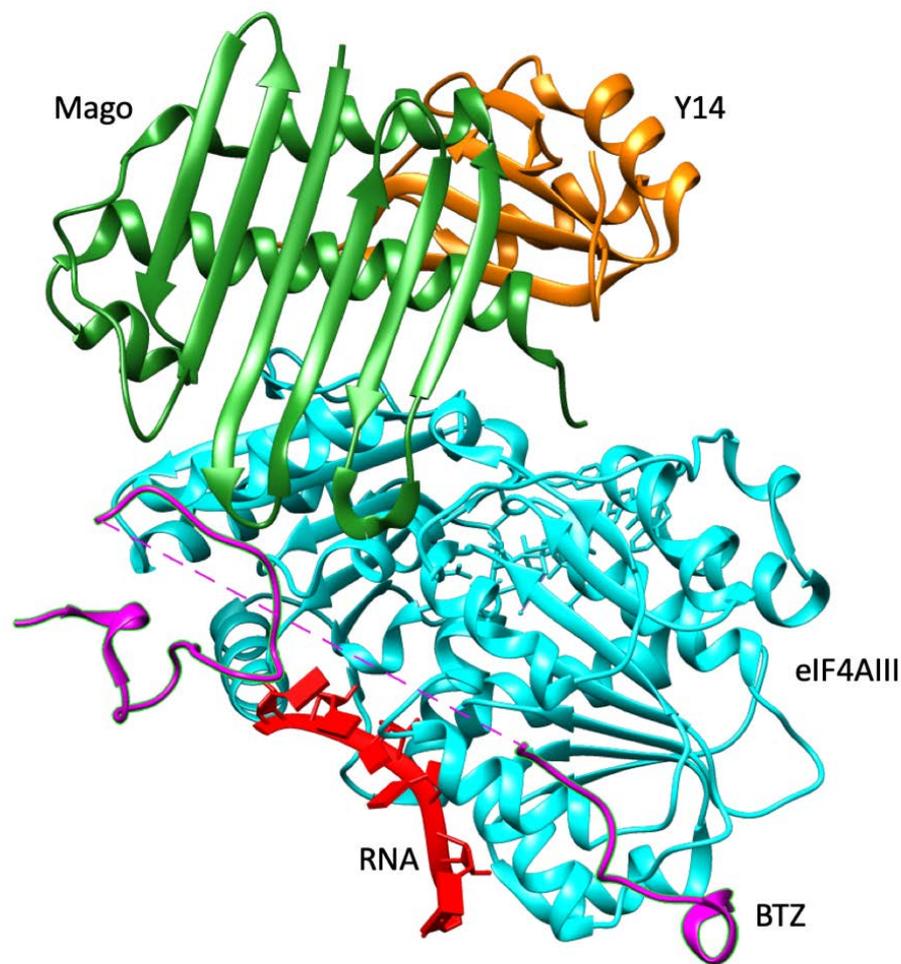
Figure 1.4 - Crystal structure of the EJC at 2.2 Å resolution. The EJC forms an L-shaped structure with MAGO (green) and Y14 (orange) almost perpendicular to eIF4AIII (cyan) and BTZ (magenta), located instead at the base of the structure. In the presence of ATP, eIF4AIII binds the RNA by using both the RecA1 and RecA2 domains. BTZ does not assume a globular conformation, but it is extended along with the two RecA domains of eIF4AIII and also contacts MAGO. The MAGO and Y14 heterodimer interact mainly with the RecA2 domain of eIF4AIII (Bono et al., 2006). Figure generated with Chimera (PDB accession number: 2J0S).

1.4.4 - Mechanisms of the EJC deposition and assembly

EJC deposition depends on splicing *in vivo*, and the complex can be assembled in vitro in the presence of RNA and the other core EJC components (Ballut et al., 2005). In mammalian cells, the assembly and positioning of the EJC require an eIF4G-like protein known as CWC22. CWC22 is a spliceosome associated protein, which associates *in vivo* with eIF4AIII, but not with MAGO and Y14 (Steckelberg et al., 2012). The association between eIF4AIII and CWC22 is mediated by its MIF4G domain (Middle domain of Eukaryotic Initiation Factor 4G) (Steckelberg et al., 2012).

The biological function of CWC22 in the deposition of the EJC is to escort eIF4AIII to the proximity of the splicing machinery and to prevent premature binding with the pre-mRNA; this model derives from an *in vitro* experiment by Barbosa and colleagues, which showed that in absence of CWC22, eIF4AIII does not immunoprecipitate with spliced mRNA (Barbosa et al., 2012). However, contrary to mammalian cells, in *D. melanogaster,* CWC22 is not required for association of eIF4AIII on the nascent mRNA; in fact, the conditional depletion of CWC22 does not interfere with the association of eIF4AIII and Y14 on the nascent transcripts (Choudhury et al., 2016).

EJC assembly and deposition on the spliced mRNAs follows a precise hierarchical pathway in mammalian cells; when an intron-containing transcript is incubated in a splicing nuclear extract, the EJC components precipitate with fully spliced mRNAs, but a portion of eIF4AIII, MAGO and Y14 associates with pre-mRNA, thus suggesting that the deposition occurs in the early stage of the splicing reaction (Gehring et al., 2009a). This observation seems to be supported by the fact that the EJC components have been purified within an active spliceosomal C-complex (Bessonov et al., 2008). Additionally, when the 3'-splice site is mutated, thus preventing the exon-exon ligation, eIF4AIII, MAGO and Y14, but not BTZ, can

still precipitate on splicing intermediates, suggesting not only that three major components of the EJC are recruited before the exon ligation, but also that BTZ associates with the EJC after dissociation of the spliceosome (Gehring et al., 2009a).

The binding of the EJC to exon-exon junctions also requires a "closed" conformation of eIF4AIII; the factor can oscillate from an open to a closed conformation, according to whether the helicase is bound with ATP (closed) or not bound (open). Introducing point mutations in the pocket of eIF4AIII required for ATP binding, MAGO and Y14, but not BTZ, do not precipitate with eIF4AIII, suggesting that the recruitment of the heterodimer can occur only when eIF4AIII is stably associated with the RNA; conversely, BTZ can be loaded on the EJC regardless of the conformation state of eIF4AIII (Bono et al., 2006, Gehring et al., 2009a). Additionally, the closed conformation of eIF4AIII on the mRNA is kept by the interaction with MAGO and Y14, which are believed to inhibit the hydrolysis of the ATP, thus permitting the formation of a stable grip on the transcript, with the remarkable ability to translocate from the nucleus to the cytoplasm, clamping on the same position (Ballut et al., 2005).

1.4.5 - Mechanism of EJC disassembly

EJC disassembly occurs during translation of the mRNA in the cytoplasm (Gehring et al., 2009b). Despite it being initially hypothesized that the helicase activity of the ribosome was sufficient to displace the EJC from the transcript, later studies identified Partner of Y14 and MAGO (PYM), as a pivotal protein for EJC disassembly (Gehring et al., 2009b, Park and Muench, 2007). The interaction of PYM with the 40S ribosomal subunit and the heterodimer MAGO-Y14, via its N-terminal and C-terminal domains respectively, is believed to trigger the disassembly of the EJC and the recycling of the factors from the cytoplasm to the nucleus

(Bono et al., 2004, Bono and Gehring, 2011). Moreover, the interaction between PYM and the EJC factors, is probably restricted to the first round of translation, when the first ribosome scans the mRNA; a crucial observation supporting this hypothesis is that on sucrose gradient fractionation, Y14 sediments in monosomal but not polysomal fractions (Dostie and Dreyfuss, 2002, Gehring et al., 2009b).

1.4.6 - Canonical and non-canonical deposition of the EJC

As previously stated, the EJCs are deposited 20-24 nucleotides upstream of each exon-exon junction (Le Hir et al., 2000). Initial studies about the function, the assembly and the deposition of the EJCs were performed *in vitro*, using protein expressed exogenously or gene reporters transfected into the cells. However, *in vivo* approaches have revealed a greater level of complexity in the EJC deposition.

Using the RNA:protein immunoprecipitation in tandem sequencing (RIPiT-seq), a high-throughput approach to map the sites of interaction of a specific protein on the mRNP, Singh and colleagues confirmed that in mammalian cells the majority of the EJC factors were positioned at the canonical positions (cEJC), 24 nucleotides before each exon-exon junction (Singh et al., 2012). However, it was also observed that almost 20% of spliced mRNAs were not associated with any EJCs, and more interestingly that EJCs were also positioned in non-canonical positions (nEJC), typically localised in the exon regions of spliced mRNAs (Singh et al., 2012, Singh et al., 2014). Another study based on the use of the ultraviolet Cross-linking and Immunoprecipitation sequencing (CLIP-seq), has shown that eIF4AIII is associated both at canonical and not canonical positions, in variable quantities and at purine-rich motifs: however, eIF4III is not detected on mRNAs derived from intron-less genes

(Saulière et al., 2012). Interestingly in the CLIP-seq analysis from Sauliere and colleagues, not all the exon-exon junctions contained an EJC.

Contrary to the general notion that the EJC is recruited only on spliced mRNAs, in *D. melanogaster*, the recruitment of EJC components can occur on the nascent transcripts independently of splicing (Choudhury et al., 2016). Moreover, it was also reported that the association of MAGO and Y14 on the nascent-RNA occurred independently of eIF4AIII, and the depletion of neither MAGO or Y14 altered the loading of eIF4AIII, suggesting that in *D. melanogaster* it is not likely that the three components form a stable complex (Choudhury et al., 2016).

Another observation arguing against the loading of the EJC occurring only on spliced mRNAs comes from a recent study from Akhtar and colleagues; using ChIP-seq analysis they reported that the EJC factors accumulated at the transcription start site (TSS) of any gene, regardless the presence of an intron, both in mammalian cells and in S2 cells of *D. melanogaster* (Akhtar et al., 2019). The ChIP signal of the three components was sensitive to an RNase treatment, suggesting that their recruitment was occurring on the RNA co-transcriptionally (Akhtar et al., 2019).

Consistent with these studies, Viphakone and colleagues, have recently shown in their CLIP-seq analysis that, in mammalian cells, eIF4AIII and BTZ, associate also with intronless-transcripts (Viphakone et al., 2019).

Additionally, eIF4AIII has been found associated at the 5'-end of CBC-bound mRNAs in Hela cells, and the recruitment was independent of the presence of an intron (Choe et al., 2014).

1.5 - Functions of the EJC

1.5.1 - Role of the EJC in mRNA surveillance

Typical error-containing mRNAs are those that harbour **P**remature Translational **T**ermination **C**odons (PTCs) (Brogna and Wen, 2009, Hentze and Kulozik, 1999, Maquat, 2004). PTCs can arise for a variety of reasons; for instance a single nucleotide substitution can change a sense codon into a termination one (TAA, TAG, TGA), or a deletion or insertion can alter the original sequence of the CDS, leading to the formation of a termination codon (Chang et al., 2007). Another source of PTCs derives from chromosomal rearrangements: an example is the V(D)J recombination mechanism in mammalian cells, adopted by B and T-cells to increment the repertoire of T-cell receptors (TCRs) and immune-globulin (IgG); two times out of three it leads to an unfruitful rearrangement, with subsequent generation of transcripts harboring a PTC (Chang et al., 2007, Nicholson et al., 2010).

PTC-containing transcripts are considered the natural target of nonsense-mediated mRNA decay (NMD), whose function is to eliminate defective transcripts, preventing their translation and the subsequent production of deleterious truncated peptides (Maquat, 1995, Zhang and Maquat, 1997). The first evidence of NMD dates back to 1979 when two different groups simultaneously observed that genes harboring PTCs, associated with unexpectedly low levels of mRNA both in yeast and in human cells (Chang and Kan, 1979, Losson and Lacroute, 1979).

Although NMD is recognized as an evolutionarily conserved surveillance pathway across eukaryotes, it remains still nebulous how the cell discriminates between a normal stop codon and a PTC (Brogna et al., 2016, Karousis and Mühlemann, 2019).

The up-frameshift (UPF) protein UPF1, originally described in *S. cerevisiae*, is the core component of the NMD machinery (Culbertson et al., 1980). UPF1 is an ATP-dependent RNA helicase, and *in vitro* can translocate through single and double-stranded nucleic acids to remodel the mRNPs (Fiorini et al., 2015). The protein appears involved in several biological mechanisms according to its localization. In the nucleus, there are evidences of its involvement in DNA replication and genome stability, release of the nascent RNA from chromosomal active sites and viral RNA nuclear export (Ajamian et al., 2015, Azzalin and Lingner, 2006, Singh et al., 2019).

In the cytoplasm, UPF1 plays an active role in the degradation of the PTC-containing transcripts via NMD. Depletion of UPF1 via RNA interference in higher eukaryotes, or knock out in yeast, rescues aberrant mRNAs from degradation, suggesting that UPF1 is essential for NMD (Brogna et al., 2016, Brogna and Wen, 2009, Karousis and Mühlemann, 2019). Two up-frameshift proteins, UPF2 and UPF3, also play a role in NMD, and their depletion results in the stabilization of the PTC-containing transcripts (Conti and Izaurralde, 2005).

Despite a huge effort to identify the mechanisms and the factors involved in NMD, a central question about this process remains still open: how a PTC is recognized and discriminated from a normal stop codon remains unsatisfactorily understood (Brogna et al., 2016)

A first hypothesis regarding the mechanism of NMD was proposed by Losson and Lacroute, who observed that when nonsense mutations were introduced in the URA3 gene of *S. cerevisiae*, the levels of the URA3 mRNA were lower when the PTCs were close to the 5' end of the gene (Losson and Lacroute, 1979). To explain this phenomenon, they conjectured the existence of a discriminatory signal, the Downstream Sequence Element (DSE), located downstream of the PTC and occurring more likely in the CDS (Brogna et al., 2016, Brogna and Wen, 2009) Therefore, if the PTC associated earlier within a transcript, higher was the

probability for a DSE to occur downstream. The model was also explored mechanistically: in the presence of a PTC-containing transcript, the ribosome stalled on the PTC, allowing the recruitment of a surveillance complex, formed by UPF1, UPF2, and UPF3, which scanned the transcript looking for the DSE. The interaction between UPF1 and a specific factor located on the DSEs, known as HRP1, triggered the mRNA decay (González et al., 2000, Brogna et al., 2016). Despite these interesting considerations, the DSE model has been largely abandoned since the putative DSE is not characterized by a specific consensus sequence (Maquat and Li, 2001).

However, the idea that a signal located downstream of the PTC can still influence the discrimination from a normal stop codon, is still largely accepted. In particular, in mammalian cells, it was observed that a PTC induced NMD only when was located upstream of an intron, suggesting that the discrimination of a PTC from a normal stop codon was driven by a splice junction or some factors associated in their proximity (Zhang et al., 1998, Carter et al., 1996).

Since the splice junctions are highly degenerate sequences, it has been hypothesized that the EJC, which are normally deposited 24 nucleotides upstream of the exon-exon boundaries, could play a pivotal role in the recognition of a PTC (Le Hir et al., 2001b).

According to the model, the ribosome displaces the EJC, during the first round of translation, and the encounter with a stop codon, it causes the translation termination. However, in the presence of a PTC, the interaction between factors associated on the ribosome, such as UPF1, eRF1 and eRF3, and the downstream EJC components, is responsible for the NMD activation. More specifically, UPF1 interacts with UPF2, which acts as a bridge, bringing together UPF1 and UPF3, which was previously loaded into the EJC in the nucleus. The interaction triggers, by a still vague and unclear mechanism, the phosphorylation of UPF1 by SMG-1, and the

degradation of the transcript by exo-and endonuclease landed on the complex (Figure 1.5) (Kashima et al., 2006, Buchwald et al., 2010, Brogna et al., 2016).
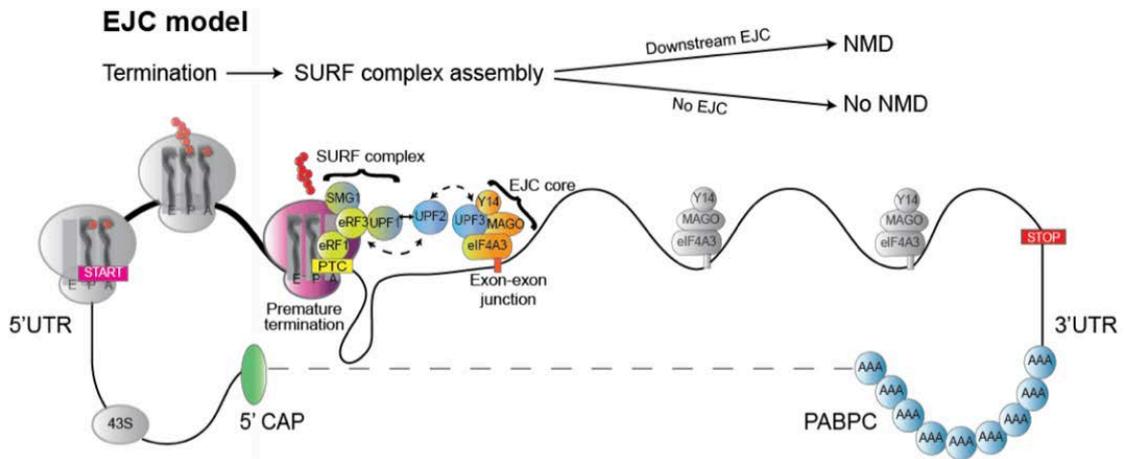
Figure 1.5 - NMD pathway triggered by EJC. During translation, the ribosome scans the transcript displacing the EJC. However, if the mRNA contains a PTC, the ribosome stalls in its proximity and a surveillance complex, known as SURF, formed by UPF1, SMG1 and the release factors 1 and 3 (eRF1 and eRF3), is recruited. The SURF complex interacts with the downstream EJC via UPF2, which acts as a bridge between UPF1 and UPF3. The interaction triggers the activation of the UPF1 helicase activity and the mRNA degradation (Figure taken from (Brogna et al., 2016)).

However, despite these observations supporting a link between EJC and the NMD, the EJC model has been argued by Wen and Brogna, which observed that an intron enhances NMD regardless of whether it is positioned downstream or upstream of a PTC (Wen and Brogna, 2010). In fact, when fission yeast is transformed with an endogenous PTC containing intron-less gene, the reduction in the mRNA level is less prominent compared to the effect observed from the same reporter in which an intron is artificially inserted, suggesting that an intron strongly enhances the NMD activity; however, this occurs regardless of whether the intron being located before or after the PTC. This finding appears in stark contrast to the EJC model since in the case the intron is located upstream of a PTC, the EJC would be removed by the translating ribosome before the PTC recognition, and NMD would not be triggered (Wen and Brogna, 2010).

With regard to eIF4AIII, its role in NMD has been investigated in mammalian cells in particular, in which its depletion leads to stabilization of PTC containing transcripts (Shibuya et al., 2004). A similar effect was also observed when cells are depleted of Y14 (Singh and Lykke-Andersen, 2003).

In mammalian cells the role of eIF4AIII in NMD has also been recently evaluated by the use of specific inhibitors; for example Pateamine A (Pat A), a natural drug derived from the sea sponge *Mycale,* appears to bind specifically members of the eukaryotic translation initiation factor 4A, such as eIF4AI, eIF4AII, and eIF4AIII stabilizing their conformation in an "*off-state*" and leading to the suppression of their biological function; when Pat A is used on HeLa or HEK293T cells, the PTC-containing mRNAs are stabilized (Dang et al., 2009). The same effect was also reported for a new compound, known as 1,4- diacylpiperazine, which is also considered a selective inhibitor of eIF4AIII (Ito et al., 2018, Ito et al., 2017, Iwatani-Yoshihara et al., 2017, Mazloomian et al., 2019, Mizojiri et al., 2018).

1.5.2 - Role of EJC in splicing

Further investigations relative to the function of the EJC factors have demonstrated that the EJC is required for the splicing of specific transcripts, both in *D. melanogaster* and in yeast (Ashton-Beaucage et al., 2010, Marayati et al., 2016). For example, in eukaryotes, the RAS/MAPK signaling pathway is a fundamental cascade, linked to several biological processes, such as cell division, differentiation, and survival; in flies the conditional depletion of eIF4AIII and MAGO, correlates with a reduction in the MAPK mRNA levels and with the formation of new mRNA isoforms derived from the same gene, suggesting that the depletion of EJC factors could trigger alternative splicing events (Ashton-Beaucage et al., 2010).

In *S. pombe*, it has been reported that Fal1, the homolog of eIF4AIII, interacts with Red5, a key component of the MTREC, a multiprotein complex involved in the targeting and degradation of the meiotic transcripts during vegetative growth. Deletion of Fal1 in fission yeast, associates with an accumulation of different splicing isoforms of meiotic mRNAs (Marayati et al., 2016).

How splicing itself is affected upon depletion of EJC components remains unclear. Ashton-Becauge and colleagues have proposed that when splicing occurs, the EJCs are firstly loaded on the exon-exon junctions separated by long introns, which are typically spliced more efficiently; their presence seems to facilitate the splicing of the small surrounded exons, thus permitting that the process occurs in a specific direction and impeding alternative splicing events. Instead, when the EJC components are depleted, the hierarchy of the splicing events is compromised and the exons are randomly juxtaposed (Ashton-Beaucage and Therrien, 2011). Another possibility is that EJC affects splicing indirectly, primarily as a consequence of regulating RNA Pol II elongation rate (Wang et al., 2014). It has been reported that RNA Pol

II elongation rate has a strong effect on alternative splicing; for example a faster RNA Pol II correlates with an increase in intron-retaining, probably due to less time for the recognition of the 5'SS and 3'SS (de la Mata et al., 2003, Wang et al., 2014). Wang and colleagues observed that when mammalian cells are depleted of eIF4AIII, the expression of some target genes increase faster compared to the WT, upon block and release of transcription with the elongation inhibitor 5,6-dichoro-1-β-D-ribofuranosylbenzimidazole (DRB); according to their interpretation, the deletion of eIF4AIII correlates with variation in RNA Pol II progression, which triggers the inclusion or the skipping of alternative exons (Wang et al., 2014).

1.5.3 - Role of EJC in mRNA export

After synthesis and processing in the nucleus, the mRNPs are exported to the cytoplasm through nuclear pore complexes (NPCs). The most important mRNA export factor is the Tap-p15 heterodimer, which is conserved from yeast to human (Reed, 2003). The Tap (NXF1) protein contains four domains; a non-canonical RNP-type RNA binding domain (RRM) at the N-terminus, a domain needed for the heterodimerisation with p15/NXT1, four leucine-rich repeats (LRR) and a ubiquitin-associated domain (UBA) that allows interaction with the nucleoporins, the constituents of the NPCs which coordinate the flow of macromolecules through the NPCs (Erkmann and Kutay, 2004). The Tap-p15 heterodimer interacts with a member of the Transcription/Export complex (TREX), such as Aly-REF, which is recruited on the nascent mRNA during its synthesis and pre-mRNA processing, and functions as an adaptor to facilitate the interaction between the nuclear export receptor Tap-p15 and the mRNA. Once the Tap-p15 gains access to the mRNP, it facilitates the translocation through the NPCs (Katahira, 2012).

Since the EJC dynamically changes its composition as the mRNA move from the nucleus to the cytoplasm, it has been reported that the EJC provides a binding platform also for those factors involved in mRNA nuclear export (Gatfield and Izaurralde, 2002, Le Hir et al., 2001b, Reed, 2003, Viphakone et al., 2019). In particular, several studies report that eIF4AIII interacts with Tap and Aly-REF, suggesting a role of the EJC factor in enhancing the export efficiency of the spliced-mRNAs. In agreement with this hypothesis, knock-down of eIF4AIII leads to moderate retention of poly(A) transcripts in the nucleus, probably due to its function in facilitating the deposition of the Aly-REF factors on the spliced mRNAs, and interestingly also on intron-less gene transcripts (Viphakone et al., 2019).

1.6 - Functions of eIF4AIII outside the EJC

Fal1, the ortholog of eIF4AIII in *S. cerevisiae*, is predominantly localised in the nucleolus and its depletion leads to a reduction in 18S rRNA levels (Kressler et al., 1997). Specifically, Kressler and colleagues observed that depletion of Fal1 is associated with defects in specific endonucleolytic cleavage reactions needed for the biogenesis of 18S rRNA. It has been proposed that the RNA helicase activity of eIF4AIII is required to unwind secondary structures of the rRNA, thus facilitating its maturation (Kressler et al., 1997). Despite the role of eIF4AIII in ribosomal biogenesis being overlooked for several years, recently, the role it plays in ribosomal RNA biogenesis has been confirmed in numerous studies also in mammalian cells (Alexandrov et al., 2011, Badertscher et al., 2015, Mao, 2016, Mao et al., 2016, Tafforeau et al., 2013, Wild et al., 2010, Zhang et al., 2011). Therefore, the function appears evolutionarily conserved across eukaryotes, since the *D. melanogaster* eIF4AIII can complement the phenotype of the Fal1 deletion in *S. cerevisiae* (Alexandrov et al., 2011).

A second function attributed to eIF4AIII, which appears independent of the EJC, is the ability to melt secondary structures located at the 5'-UTR of spliced and unspliced transcripts, thus providing a role for eIF4AIII in translation initiation (Choe et al., 2014). In fact, Choe and colleagues have reported that eIF4AIIII associates at the 5'-UTR of any transcript via a direct interaction with CTIF (CBC- dependent translation initiation factor), an MIF4G domain-containing protein which directly interacts with CBP80, and whose depletion blocks translation in vitro (Kim et al., 2009); the presence of eIF4AIII at the 5'-UTR appears crucial to remove secondary structures artificially inserted into a reporter plasmid, which might prevent the attachment and the movement of the small ribosomal subunit to the translation initiation codon (Choe et al., 2014).

However, this notion seems in contrast to what Li and colleagues reported more than one decade ago; they demonstrated that eIF4AIII, as eIF4AI and II, has an ATP-dependent RNA helicase activity *in vitro*, but it fails to substitute eIF4AI in an *in vitro* ribosome binding assay, and differently from the other two members of the same family it does not stimulate translation*,* proposing that eIF4AIII is not involved in translation initiation. (Li et al., 1999). Generally, most of the evidence available in the literature, are in agreement with this finding (Chan et al., 2004, Ferraiuolo et al., 2004, Shibuya et al., 2004).

1.7 - *S. pombe* as a model organism

The fission yeast *Schizosaccharomyces pombe* is an ascomycete unicellular fungus, used as a model organism in molecular and cell biology. Wild type cells have a rod shape, with a diameter varying from 3-4 μm and a length of 12-15 μm (Forsburg, 2003). Due to a short generation time, ranging from 2 to 5 hours, depending on the strain, the temperature and the media, *S. pombe* is relatively easy to grow and maintain. The small genome size (13.8 Mb), which consists of only 3 chromosomes and contains approximately 4950 protein-coding genes and 450 non-protein coding RNAs, have made *S. pombe* a popular model system, to study gene expression and cell cycle regulation (Wood et al., 2002, Yanagida, 2002). For example, Paul Nurse, who won the Nobel prize in Physiology and Medicine in 2001, has used fission yeast to identify genes and mechanisms involved in cell cycle progression (Nurse, 1990). Moreover, almost 50 genes involved in human diseases, are highly conserved in *S. pombe*, thus making this organism suitable for the understanding of the molecular basis of human pathologies (Wood et al., 2002).

S. pombe has two forms of propagation, sporulation and binary fission and it can be found either in haploid or diploid state (Forsburg, 2003). To evaluate whether the EJC factors have a direct role in transcription and pre-mRNA processing, we have decided to use *S. pombe* as model organism, for its ability to preferentially grow in haploid condition (each chromosome present in a single copy); therefore any change in the EJC loci or protein levels, will be associated with a unique cellular or molecular phenotype and not covered or masked by other genetic backups. Because EJCs are highly conserved across eukaryotes, we expect that any consideration coming from our study can be reasonably extended to multicellular organisms.

Another important aspect that makes *S. pombe* particularly convenient for gene expression studies, is the fact 43% of its genes contain introns, while most *S. cerevisiae* genes have no

introns. This wider complexity is therefore particularly important when biological mechanisms connected to splicing are investigated. For example, most of the EJC factors are retained in fission yeast, but not in *S. cerevisiae* where only eIF4AIII/Fal1 has been identified (Kressler et al., 1997, Wen and Brogna, 2010).

Lastly, *S. pombe* is a powerful model organism for studying eukaryotic gene expression, for the ability to easily manipulate its genome and generate deletion mutants and tagged proteins: due to a powerful homologous recombination machinery, a fragment of exogenous DNA can be transformed and integrated into the desired locus.

Furthermore, the repertoire of genetic tools available in fission yeast is vast and continually expanded; for example, CRISPR-Cas9, Auxin-Inducible degron (AID) system and Anchor Away (AA) have been developed and applied for this organism (Ding et al., 2014, Kanke et al., 2011, Rodríguez-López et al., 2016).

1.8 - Aim of this study

My work aims to increase our understanding of the function of eIF4AIII in *S. pombe*, by using a new experimental approach to rapidly deplete eIF4AIII and investigate its function.

The RNA-helicase eIF4AIII is a highly conserved protein across eukaryotes. Previous studies reported that eIF4AIII might regulate transcription and pre-mRNA processing (Akhtar et al., 2019, Ashton-Beaucage and Therrien, 2011, Ashton-Beaucage et al., 2010, Marayati et al., 2016, Wang et al., 2014). However, a clear limitation coming from these conclusions is that to study the function of eIF4AIII, its depletion has been accomplished using siRNA mediated knock-down, which normally requires several hours or days; during this time frame cells may adapt, thus preventing discrimination between direct and indirect effects. (Akhtar et al., 2019, Ashton-Beaucage and Therrien, 2011, Ashton-Beaucage et al., 2010, Wang et al., 2014). Furthermore, other studies have based their conclusions from molecular phenotypes deriving from the deletion of *eIF4AIII,* which causes severe growth defects (Marayati et al., 2016).

In this project three research lines have been explored:

1) Development and optimisation of a cutting-edge technology known as Auxin-Inducible Degron (AID) system, to induce rapid depletion of Fal1/eIF4AIII and then assess the effects that this has on transcription and pre-mRNA processing. The results of these experiments are discussed in Chapter 3.

2) Investigating the binding of eIF4AIII and other putative EJC components at specific transcription active loci using ChIP. The results are described in Chapter 4.

3) Investigation whether eIF4AIII might form a stable complex with MAGO and Y14 in *S. pombe*. The results are described in Chapter 5.

Chapter 2: Materials and Methods

2.1 - *S. pombe*

2.1.1 - *S. pombe* strains

A list of the *S. pombe* strains used in this study is reported in Appendix II

2.1.2 - *S. pombe* long term storage

The *S. pombe* strains were stored at the temperature of -80°C. To prepare the yeast stock, cells were plated on YES (Yeast Extract with Supplements) or EMM (Edinburgh Minimal Media), (see Appendix I for details), according to the yeast genotype, and incubated overnight at 30°C. Then, cells were scraped from the plate using a sterile tip, and resuspended in 1mL of YES or EMM liquid media, containing 667 μL of 80% glycerol, and stored at the temperature of -80°C.

2.1.3 - *S. pombe* cell culturing

Two kinds of media were used for cell culturing: YES and EMM, according to the genotype of the yeast strain. For most of the experiments, a fresh single colony of *S. pombe* was inoculated in 5 mL of YES or EMM and grown overnight at the temperature of 30°C in a shaking incubator; the pre-culture was subsequently scaled up, to achieve the desired final volume, according to the experiment. For cell selection, G418 disulfate (Melford), hygromycin B (Melford) and Nourseothricin-dihydrogen sulfate (Melford) were added to YES or EMM plates at a final concentration of 200 μg/ml, 200 μg/ml and 100 μg/ml respectively. For experiments based on the rapid degradation induced by the AID, the plant

auxin 1-Naphthaleneacetic acid (NAA) (Sigma) powder was dissolved in 70% ethanol and diluted to the required concentration (1 mM). For experiment based on promoter shut-off, Thiamine powder (Sigma) was dissolved in sterile distilled water and diluted to the required concentration (20 μM).

2.1.4 - Counting *S. pombe* cells: method 1

The number of cells was determined using a hemocytometer. In brief, a coverslip was positioned over the counting chamber. 10 μL of cell suspension was firstly diluted in 90 μL of sterile H$_2$O; then 10 μL of the diluted cell suspension was gently pipetted into the space generated between the coverslip and the counting chamber. The chamber was placed under an optical microscope and cells in the red squares of the grid were counted (Figure 2.0). In case the cells touched the sides of the square, only cells located on two arbitrary borders were counted. To determine the total number of cells for mL, the following formula was applied:

$$number\ of\ \frac{cells}{ml} = number\ of\ cells\ counted\ \times \frac{1}{2} \times dilution\ factor \times 10000$$
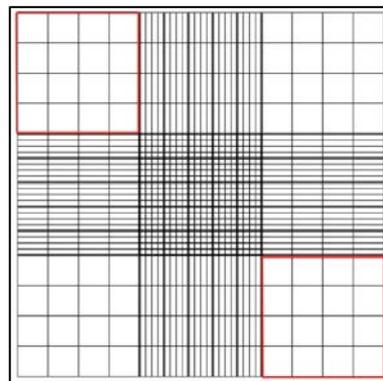
Figure 2.0 – Grid of the hemocytometer.

2.1.5 - Counting *S. pombe* cells: method 2

The cell density was determined using a spectrophotometer (Amersham Pharmacia's GeneQuant pro). 100 μL of cell culture was diluted in 900 μL of sterile H_2O in a cuvette (Sarstedt), and the absorbance at 600 was measured. The OD_{600} read by the instrument was multiplied by the dilution factor (x10). For all the experiments performed in this study, cultures were maintained in mid-exponential growth with an OD_{600}=0.4/05, which corresponds to approximately $1x10^7$ cell/mL.

2.1.6 - Spot growth assay

The functionality and the viability of the *S. pombe* yeast strains were estimated using the spot growth assay. A fresh single colony of *S. pombe* was inoculated in 5 mL of YES or EMM liquid media (according to the yeast genotype) and grown until OD_{600}=0.5. Then, ten-fold serial dilutions of the cell culture were prepared in sterile H_2O. 10 μL of the diluted samples, containing approximately 10^5, 10^4, 10^3, 10^2 and 10 cells, were spotted on YES and EMM agar plates and incubated for 72 hours at 30°C. For some experiments, the EMM plates were supplemented with 1-Naphthaleneacetic acid (final concentration of 1mM) and Thiamine (final concentration of 20 μM).

2.1.7 - Growth assay in liquid media

The growth rate of *S. pombe* was estimated by growth assay in liquid media. A fresh single colony of *S. pombe* was inoculated in 5 mL of EMM or YES liquid media (according to the yeast genotype) and grown overnight until OD_{600}=0.5. The cell culture was diluted in 20 mL of liquid media and synchronized at OD_{600}=0.03. Using a hemocytometer, the number of cells per mL was determined every two hours. To avoid the saturation of the cell culture, the media

was refreshed every 8 hours, removing 10 mL of cell culture and adding 10 mL of fresh media.

2.1.8 - *S. pombe* microscopy

Images of *S. pombe* were acquired using confocal microscopy (Nikon A1R of the Birmingham Advanced Light Microscopy (BALM) facility) or a 2D array scanning laser confocal miscroscope (Infinity 3, VisiTech) (University of Leicester).

Cells were analyzed in the vegetative growth phase for all the experiments.

To prepare the sample, a single fresh colony of *S. pombe* strain was inoculated in 5 mL of YES or EMM liquid media and grown overnight until OD=0.5. Next day, 1 mL of cell culture was transferred into a 1.5 mL tube and centrifuged at 15700 g for 1 minute at RT. The supernatant was removed and cells were washed twice in 1mL of sterile H_2O. Then cells were resuspended in 200 μL of sterile H_2O, of which 10 μL were gently pipetted onto a glass microscope slide (Thermo Scientific), covered with a coverslip.

2.2 - *S. pombe* DNA techniques

For *S. pombe* DNA extraction, method 1 (2.2.1) has been used. Alternatively, method 2 (2.2.2), was adopeted when GeneJet Genomic DNA purification kit was not available.

2.2.1 - Genomic DNA extraction: method 1

The genomic DNA from *S. pombe* was extracted using the GeneJet Genomic DNA purification kit (Thermo Fisher Scientific). A fresh single colony of *S. pombe* was grown in 5 mL of YES or EMM (according to the yeast genotype) until OD_{600}=0.5. Cells were harvested in a microfuge at 800 x g for 5 minutes. The supernatant was discarded and the pellet was

resuspended in 1 mL of Yeast Lysis Buffer (not provided with the kit) consisting of 5 mg/mL Zymolyase 20T (AMS Biotechnology), 0.1 M EDTA pH 8.0, and 1 M sorbitol. The cell suspension was incubated for 1 hour at 37°C. Cells were harvested at 2700 x g for 10 minutes, and the pellet was resuspended in 180 μL of digestion solution (provided with the kit), containing 20 μL of Proteinase K (20 mg/mL). The suspension was incubated for 1 hour at 56°C. To degrade the RNA, the lysate was treated with 20 μL of RNase A (10 mg/mL), and incubated for 30 minutes at room temperature. After incubation, 200 μL of Lysis Solution (provided with the kit) was added to the mixture, which was vortexed for a few seconds to obtain a homogeneous solution. To precipitate the DNA, 400 μL of 50% of ethanol was added to the lysate, which was then transferred into a GeneJET Genomic DNA Purification column (provided with the kit). The column previously inserted into a collection tube, was centrifuged at 5900 x g for 1 minute and the flow-through was discarded. The column was firstly washed with 500 μL of Wash buffer I and then with 500 μL of Wash buffer II. After the final washing, the column was transferred into a sterile 1.5 mL Eppendorf tube. To elute the DNA 50 μL of Elution buffer was pipetted in the middle of the column; the sample was firstly incubated for 1 minute at room temperature and then centrifuged at maximum speed for 2 minutes. The DNA concentration was evaluated by nanodrop (ND-1000 spectrometer) and the quality was checked by running a 1% agarose gel.

2.2.2 - Genomic DNA extraction: method 2

To extract the genomic DNA from *S. pombe* (only to use as a template in a PCR reaction), an alternative method to the kit protocol previously described has been also adopted (Murray et al., 2016). Cells were grown in 10 mL of YES or EMM (according to the yeast genotype) until OD_{600}=0.5. The cell culture was harvested at maximum speed for 1 minute, washed with

1 mL of sterile H$_2$O, and resuspended in 1 mL of CSE buffer (1.2 M Sorbitol, 40mM EDTA pH 8.0, 1mg/mL of Zymolase 20T). The mixture was incubated at 37°C for 15 minutes. After incubation, the cell suspension was harvested at maximum speed for 1 minute and resuspended in 450 µL of 5x TE buffer. To the lysate 50 µL of 10% SDS was added, mixed by inversion and incubated for 5 minutes at room temperature. Upon addition of 150 µL of 5M CH$_3$COOK, the sample was incubated on ice for 10 minutes. The suspension was pelleted at maximum speed for 10 minutes and the supernatant was transferred in a sterile 1.5 mL Eppendorf tube, to which 1 volume of isopropanol was added. The DNA was recovered by centrifugation at maximum speed for 10 minutes, washed twice with 75% ethanol and air-dried at room temperature for 5-10 minutes. The pellet was resuspended in 50 µL of 1xTE buffer, containing 1 µL of 10mg/mL RNase A, and incubated for 15 minutes at 37°C. Typically, 1 µL of genomic DNA was used as a template in the PCR reaction.

2.2.3 - Agarose gel electrophoresis and gel extraction

Plasmids and PCR products were mixed with 6X Gel Loading Dye (NEB) and resolved on a 1% agarose gel and run in 1X TAE buffer (40 mM Tris base, 20 mM acetic acid and 1 mM EDTA), containing 0.5 µg/mL of ethidium bromide, at a constant voltage of 90 V. The size of the band of interest was identified by referring to the 100 bp or 1kb DNA ladders (NEB). In presence of multiple bands, the fragment of interest was extracted using the Monarch DNA extraction Kit (NEB) or the QIAquick Gel extraction kit (QIAGEN). Briefly, the desired DNA band was cut under the UV light; according to its weight, three volumes of Dissolving Buffer (provided with the kit) were added into a 1.5 mL Eppendorf tube containing the band of interest, which was melted at 55°C for approximately 5-7 minutes. The suspension was loaded into the column (provided with the kit), and centrifuged at maximum speed for 1

minute. The flow-through was discarded and the DNA was washed twice with 700 µL of Wash buffer (provided with the kit). The column was positioned into a sterile 1.5 Eppendorf tube, to which 50 µL of Elution buffer was applied for the elution. The DNA was eluted by centrifugation at 15700 x *g* for 1 minute at room temperature.

In the presence of a single band, the DNA was purified using a QIAquick PCR purification kit; briefly, the PCR reaction solution (approximately 50 µL), was mixed with 5 volumes of Binding Buffer (provided with the kit). The mixture was transferred into a QIAquick Spin column, and centrifuged for 1 minute at maximum speed. The flow-through was discarded and the column was firstly washed with 750 µL of Wash Buffer I and then with the same volume of Wash Buffer II. Lastly, the DNA was eluted by applying 50 µL of Elution Buffer and centrifuging at maximum speed for 1 minute at room temperature.

2.2.4 - Standard PCR

Standard PCR was performed in a reaction volume of 50 µL (1X Q5 Reaction Buffer, 200 µM dNTPs, 0.5 µM Forward Primer, 0.5 µM Reverse Primer, 0.02 U/µL Q5 high fidelity DNA polymerase, nuclease-free water and template DNA). The sequences of the primers used in this study are reported in Appendix III. Primers were designed using Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/) and bought from Integrated DNA Technologies (ITD) or Sigma. PCR was executed in a thermocycler (Biorad, Mycycler™) and samples were resolved on 1% agarose gel.

2.2.5 - Single Colony PCR

Confirmation of endogenous gene tag or gene knock-out in *S. pombe* was also performed by single colony PCR. For the colony PCR, the DNA was amplified using the GoTaq DNA

polymerase (Promega) in a volume of 25 µL. The colony PCR reaction set up is shown in table 1.

Reagent	Volume
GoTaq polymerase	0.125 µL
5x Go Taq buffer	5 µL
FW Primer (10 µM)	1 µL
RV Primer (10 µM)	1 µL
dNTP (100 mM)	0.2 µL
MgCL (25 mM)	1.5 µL
ddH$_2$O	16.175 µL
Template	Colony

Table 1 - Colony PCR reaction set up

PCR was executed in a thermocycler (Biorad, Mycler^{TM}), using the following setting: pre-denaturation (to break the cell wall and release the nucleic acids) at 95°C for 10 minutes; 30 cycles of denaturation at 95°C for 30 seconds; annealing at 52°C for 30 seconds; extension at72°C for 90 seconds; final extension at 72°C for 5 minutes. The annealing temperature was calculated using the New England BioLabs TM Calculator (http://tmcalculator.neb.com/#!/). Amplified fragments were resolved on 1% agarose gel.

2.2.6 - DNA sequencing

DNA fragments were sequenced by either GATC biotech (Konstanz, Germany) or Source Bioscience (Nottingham, UK).

2.2.7 - *S. pombe* transformation

To transform plasmids and DNA fragments in *S. pombe*, a fast transformation protocol was adopted (Gietz and Woods, 2002). A fresh single colony was inoculated into 5 mL of YES or EMM liquid media (according to the yeast genotype), and grown overnight until OD_{600}=0.5. Cells were harvested at 1000 x g for 1 minute at room temperature, and the pellet was washed with 1 mL of 0.1 M Lithium Acetate. After a centrifugation at 1000 x g for 1 minute, the pellet was resuspended in a solution consisting of: 120 μL of PEG (50% w/v), 18 μL of 1M Lithium Acetate, 10 μL of 10 mg/mL salmon sperm DNA, the DNA fragment/plasmid to transform (typically 1 μg of DNA) and sterile H_2O until a final volume of 180 μL. Cells were incubated for 1 hour at room temperature followed by a heat shock at 42°C for 15 minutes. After the heat shock, cells were harvested at 800 x g for 30 seconds at room temperature, resuspended in 150 μL of sterile H_2O and plated on a YES or EMM agar plate (according to the yeast genotype). The plate was incubated at 30°C for 24 hours, after which it was replica plated on a fresh plate containing the desired antibiotic. The selective plate was incubated for 2-3 days at 30°C until colony formation.

2.2.8 - Endogenous gene tagging- Two-step PCR approach

To create strains in which epitopes are added endogenously at the C-terminal or N-terminal of a target protein, we have adopted a two-step PRC approach/PCR sewing (Krawchuk and Wahls, 1999). This method is characterized by the use of pFA6a plasmid, and its derivates, which can contain different kinds of epitopes, along with an antibiotic resistance gene. Initially, two PCR reactions are set up to produce two DNA fragments of approximately 500 bp (Figure 2.1). In the first PCR reaction, a FW primer (FW1) anneals to the coding sequence of the target gene, approximately 500 bp before the stop codon; the reverse primer (RV1) is a hybrid, which anneals at the end of the target gene (excluding the stop codon), and contains a short sequence of homology with pFA6a plasmid (approximately 25-30 bp). In the second PCR reaction, another fragment of 500 bp is amplified; in this case the FW primer (FW2) is a hybrid, containing a sequence of homology with pFA6a plasmid (approximately 25-30 bp) and it begins after the stop codon of the target gene, while the RV primer (RV2) is positioned 500 bp after the stop codon (Figure 2.1). The two PCR products are then purified using the QIAquick PCR Purification Kit (Qiagen) or the PCR and DNA Cleanup Kit (Monarch), as previously described. After the purification, the two PCR products are used in a third PCR reaction with pFA6a plasmid as a template. The previous PCR products will anneal with pFA6a plasmid through the short sequences of homology, and the two external primers F1 and RV2, will amplify a DNA fragment formed by a portion of the gene of interest in frame with the tag, a terminator sequence and a selection marker (Figure 2.1). The fragment is transformed in S. *pombe* and integrated into the genome via homologous recombination.
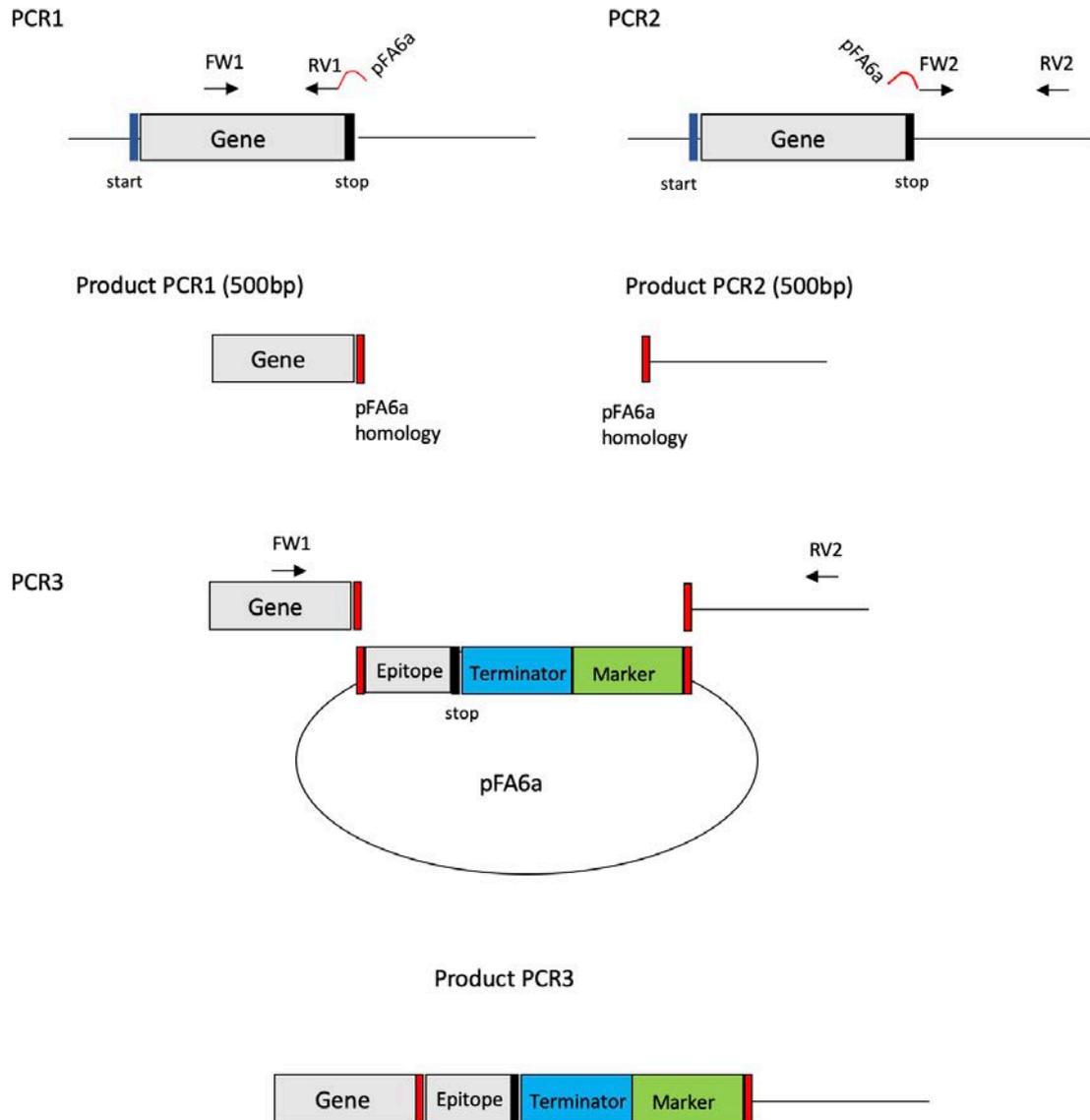
Figure 2.1 – Two-step PCR approach/PCR sewing. Two DNA fragments of approximately 500 bp, having a sequence of homology with pFA6a plasmid, are amplified in the first step of the two step-PCR approach. The two PCR products are then purified and used in a second PCR reaction with pFA6a plasmid as a template. By using external primers FW1 and RV2, a PCR product containing part of the gene of interest in frame with the epitope, a terminator and a selectable marker is amplified and subsequently transformed in *S. pombe* where it will be integrated into the desired locus via homologous recombination.

2.2.9 - Chromatin Immunoprecipitation (ChIP)

2.2.9.1 - Crosslinking and Sonication

A fresh single colony of *S. pombe* was inoculated in 10 mL of YES liquid media and grown overnight at 30°C in a shaking incubator. The next day, the cell culture was transferred in 90 mL of YES, diluted until $OD_{600}=0.1$ and grown until $OD_{600}=0.5$ ($1x\ 10^7$ cells/mL). Cells were fixed for 5 minutes at room temperature by adding 2.7 mL of 37% formaldehyde (Sigma) so that the final concentration was 1%. To stop the reaction, 20 mL of 2.5 M glycine was added to the cell culture and incubated for 10 minutes at room temperature on a shaker. The cell culture was transferred in two 50 mL Falcon tubes and centrifuged at 2700 x *g* for 3 minutes at 4°C. The pellet was resuspended in 40 mL of cold 1X PBS, and combined into a single 50 mL Falcon tube. Cells were harvested at 2700 x *g* for 3 minutes at 4°C, and two more washing in cold 1X PBS were made at the same condition. Pellet was resuspended in 750 μL of cold FA Lysis Buffer (100 mM HEPES-KOH, pH.7.5, 300 mM NaCl, 2mM EDTA, 2% Triton X-100, 0.2% Na Deoxycholate) containing an EDTA-free protease inhibitors cocktail tablet (Roche, 1 tablet for 10 ml of FA Lysis Buffer). The cell suspension was transferred into a 2 mL screw-cap tube and pelleted at 3300 x *g* for 2 minutes at 4°C. The supernatant was removed and the pellet was resuspended in 750 μL FA lysis Buffer with EDTA-free protease inhibitor cocktail tablet, to which 400 μL of Zirconia beads (0.7mm of diameter, Biospec) was added. Cells were broken using a cell homogenizer (Bertin Technologies, Precellys 24); to break almost 90% of the cells, 10 cycles of 30 seconds each at 2700 x *g* were applied, alternate with 2 minutes in ice. The bottom of the screw cup tube was pierced twice with a hot 25 G needle and each tube was immediately transferred in the barrel of a syringe previously inserted into a 15 mL falcon tube. The lysate was collected at 110 x *g* for 1 minute

at 4°C; the residual lysate was collected by adding 500 μL of FA lysis Buffer, and centrifuging at the same conditions. The cell lysate was transferred in a 1.5 mL microfuge tube and pelleted at 15700 x g for 25 minutes at 4°C. Pellet was resuspended in 600 μL of FA lysis Buffer containing a mini EDTA-free protease inhibitor cocktail tablet. Then cell lysate (approximately 750 μL) was transferred in a 15 mL tube (Sumitomo): 250 μL of FA lysis Buffer with mini EDTA-free protease inhibitor cocktail tablet was added to reach the final volume of 1 mL. To increase sonication efficiency 20 μL of 10% SDS and 20 μL of 100 mM PMSF was added to the mixture and gently resuspended, avoiding bubbles. Samples were sonicated for 15 cycles at 4°C using the Bioruptor (Diagenode).

2.2.9.2 - Sonication efficiency checking

After sonication, the sample was transferred to a 1.5 mL microfuge tube, centrifuged at 15700 x g for 25 minutes at 4°C. The supernatant was collected and transferred to a fresh 1.5 mL Eppendorf tube: to check the sonication efficiency, 50 μL of sonicated chromatin were mixed with 150 μL of TES buffer (10 mM Tris, pH 7.5; 10 mM EDTA, pH 8; 0.5% SDS), and incubated overnight at 65° C. The next day, 200 μL of TE buffer (10 mM Tris, pH 7.5; 10 mM EDTA, pH 8) were added to reach the final volume of 400 μL. To digest RNA, 1 μL of RNase A (10 mg/mL) was added and incubated at 37°C for 1 hour. To digest proteins, the mixture was incubated with 10 μL of Proteinase K (20 mg/mL), for 2 hours at 55°C. The chromatin was extracted using the QIAquick PCR purification kit, as previously described. The eluted DNA (20 μL) was mixed with 4 μL of DNA gel loading Dye (6x) and resolved on a 1% agarose gel.

2.2.9.3 - Immunoprecipitation (IP)

A 50 μL aliquot of Dynabeads Protein G (Thermofisher) was transferred in a 1.5 mL DNA low binding tube (Eppendorf) and washed three times with 1 mL of cold 1xPBS containing 5% BSA. Between each washing, the tube containing the beads was rotated on a magnetic rack 5 times. After the final washing, the beads were resuspended in 500 μL 1x PBS/ 5% BSA and transferred into a fresh 1.5 mL DNA low binding tube. Depending on the target, 3-10 μg of antibody were added to the beads, and gently resuspended. The mixture was incubated on a rotor for 1 hour at room temperature. To remove the unbound antibody, the beads were washed 3 times with 1 mL of 1xPBS /5% BSA. After the final washing, the supernatant was removed and the beads were resuspended in the appropriate amount of sonicated chromatin and transferred in a fresh 1.5 mL DNA low binding tube. The coated beads and the chromatin were incubated overnight at 4 °C on a rotor.

2.2.9.4 - Beads Washing and ChIP DNA elution

The next day, each tube containing an IP reaction was positioned into a magnetic rack. The supernatant was removed and the beads were washed for 5 minutes at room temperature on a rotor using the following buffer:

- 2 times with Wash Buffer I (50mM HEPES-KOH pH 7.5, 150 mM NaCl, 1mM EDTA pH 8.0, 1% Triton X-100, 0.1% Sodium deoxycholate, 0.1% SDS)
- 2 times with Wash Buffer II (50mM HEPES-KOH pH 7.5, 500 mM NaCl, 1mM EDTA pH 8.0, 1% Triton X-100, 0.1% Sodium deoxycholate, 0.1% SDS)
- 2 times with Wash Buffer III (10mM Tris-HCl pH 8.0, 1mM EDTA pH 8.0, 0.25mM LiCl, 0.5% Igepal CA, 1% Sodium deoxycholate)

- 2 times with TE (10mM Tris-HCl pH 8.0, 1mM EDTA pH 8.0)

Between each washing, beads were briefly centrifuged and they were rotated on the magnetic rack for 5 times. After the final washing, the beads were resuspended in 100 μL of Elution Buffer (50mM Tris-HCl pH 7.5, 10 mM EDTA, 1% SDS) and incubated for 10 minutes at 65°C and occasionally vortexed. Then supernatant (elution) was recovered and transferred to a fresh 1.5 mL DNA low binding tube. The elution was performed a second time using the same conditions.

2.2.9.5 - De-crosslinking and DNA extraction

To the input, 150 μL of elution buffer were added and incubated at 65°C overnight to allow de-crosslinking. IP sample was de-crosslinked in parallel using the same condition. Next day, 200 μL of TE buffer was added to the Input and IP, and briefly vortexed. To degrade RNA, 1 μL of RNase A (10 mg/ml) was added to the Input and to the IP which were subsequently incubated at 37 °C for 30 minutes. To remove proteins from the DNA, 10 μL Proteinase K (20 mg/ml) was added and samples were incubated at 55 °C for two hours. DNA was then extracted using the QIAquick PCR purification kit, as previously described.

2.2.9.6 - Quantitative Real-Time PCR (qRT-PCR)

After DNA extraction Input and IP were diluted in sterile distilled water: input was diluted to 200 pg/ μL; IP was diluted 1:50 in sterile H_2O. qRT-PCR of diluted Input and IP was performed using an ABI-Prism 7000 real-time PCR thermocycler. PCR reactions were prepared in technical triplicate using a 96 well plate: each reaction contains 10 μL of 2x SyGreen Blue Mix Hi-ROX (PCR-biosystems), 4 μL of diluted Input/IP, 3 μL of 0.4 μM of

each forward and reverse primer. The qRT-PCR was run using the following setting: 3 minutes at 95°C, followed by 40 cycles of 15 seconds at 95°C and 30 s at 60°C.

2.2.9.7 - Calculation of Fold enrichment

ChIP-qPCR enrichments are normalized to the input and expressed as fold enrichment relative to a negative locus (intergenic), which is not expressed. The fold enrichment is calculated with the following formula;

$$\Delta CT1 = \text{Average CT (IP gene)} - \text{Average CT (IP intergenic)}$$

$$\Delta CT2 = \text{Average CT (Input gene)} - \text{Average CT(Input intergenic)}$$

$$\Delta\Delta CT = \Delta CT1 - \Delta CT2$$

$$\text{Fold enrichment} = 2\text{-}^\wedge \Delta\Delta CT$$

2.2.9.8 - RNase sensitivity

To check whether a protein is associated with nascent RNA, an RNase sensitivity step was introduced before the DNA elution. After two washes in buffer 3, beads were washed three times more with 1xTE buffer, to remove excess detergent. Then the beads were resuspended in 500 µL of TE buffer and transferred to a fresh 1.5 mL DNA low binding tube. 5 µL of RNase A, DNAse and protease-free (10 mg/mL), was added to the mixture, gently resuspended and incubated for 30 minutes at room temperature on a rotor. Then the microfuge was positioned into the magnetic rack, the supernatant was removed and the beads were washed two times more with 1 mL of TE buffer. Elution was performed as previously described.

2.3 - *S. pombe* protein analysis techniques

2.3.1 - Protein extraction from *S. pombe* cell cultures

The protein extraction was performed using a rapid method based on NaOH (Matsuo et al., 2006). A fresh single colony of *S. pombe* was inoculated in 2 mL of YES or EMM (according to the yeast genotype), and grown at 30°C until an OD_{600}=0.5. The 2 mL culture was poured in 3 mL of pre-warmed media so that the final volume was 5 mL and grown until an OD_{600}= 0.5. Cells were harvested by centrifugation at 2700 x g for 5 minutes at room temperature; the supernatant was completely removed and cells were washed with 1 mL of sterile H_2O. The pellet, obtained by centrifugation at the same conditions, was resuspended in 300 µL of sterile H_2O, to which one volume of 0.6M NaOH (~300 µL) was added so that the final concentration of NaOH was 0.3M. The suspension was gently mixed by pipetting and the sample was incubated for 10 minutes at room temperature. The cell lysate was pelleted by centrifugation at 4000 x g for 5 minutes at 4°C. The NaOH was completely removed and the proteins were resuspended in 70 µL of 1X SDS loading buffer (4% SDS, 0.04% bromophenol blue, 20% glycerol, 60mM Tris HCl, pH 6.8, 4% mercaptoethanol, 0.01% and 1 mM phenylmethanesulfonylfluoride (PMSF)). The samples were boiled for 5 minutes and centrifuged at 15700 x g at 4°C. The samples were stored at -20°C, or immediately processed.

2.3.2 - SDS-PAGE and Western Blotting

<u>S</u>odium <u>d</u>odecyl <u>s</u>ulphate <u>p</u>olyacrylamide <u>g</u>el electrophoresis (SDS-PAGE) gels were prepared as described in Molecular Cloning 4th edition (Green and Sambrook, 2012). 20 µL of the protein extract in 1X SDS loading buffer were loaded in each well. The SDS-PAGE gels were run in 1x SDS running buffer (0.025 M Tris-HCl, 0.192 M Glycine, 0.1% (w/v) SDS,

pH 8.3), using a Bio-Rad Mini-PROTEAN Tetra Cell electrophoresis chamber, at a constant voltage of 100V for 90 minutes. Proteins were transferred onto a nitrocellulose membrane (Amersham™ Potran™, 0.2 μm, GE healthcare), using a Bio-Rad Mini PROTEAN II electrophoresis apparatus, at 350 mA for 2 hours in a cold room in 1X transfer buffer (190 mM glycine, 25 mM Tris-HCl, 20% (v/v) Methanol). The membrane was blocked with 10 mL of 1x TBS (25mM Tris-HCl, 137 mM NaCl), containing 5% milk. The primary antibodies used in this study (reported in Appendix V), were diluted according to the company, in 2.5% skimmed milk in TBST (25mM Tris-HCl, 137 mM NaCl, with 0.1% Tween 20), and incubated overnight in a cold room on a shaker. The next day, the membrane was washed 3 times for 10 minutes, in 30 mL of 1xTBST (25mM Tris-HCl, 137 mM NaCl, with 0.1% Tween 20), and incubated for 1 hour at room temperature, with a secondary horseradish peroxidase (HRP) conjugated antibody (Sigma), using a standard dilution of 1:10000 in TBST. After incubation with the secondary antibody, the membrane was washed 6 times for 5 minutes in 1xTBST. To develop the signal of the protein of interest, 1 mL of a mix formed by Buffer A and Buffer B of the ECL Western Blotting Detection Reagent (GE Healthcare) was applied to the membrane: the membrane was wrapped in a cling film and the signal was detected using the ChemiDoc imaging system. For the secondary antibody conjugated with fluorescent dyes, some protocol modifications were made. Blocking of the membrane was performed in 1xTBS containing 2.5% dried milk. The membrane was washed using the TBST-high salt (TBS supplemented with NaCl, to the final concentration of 500mM,). Additionally, the last washing before the detection of the signal was carried out in 1xTBS to eliminate residual Tween-20 that could interfere with the detection. The detection of the signal was performed using the Odyssey Imaging Software, LI-COR.

2.3.4 - Western Blot quantification analysis

Quantification of the band of the protein of interest was performed using ImageJ. The western blot image was imported in ImageJ. The intensity of the band of interest was quantified by selecting the area around the band. Then this value was divided by the intensity derived by the quantification of α-tubulin.

2.3.5 - Bioinformatics analysis

Primary protein sequences were downloaded from Pombase, FlyBase or Uniprot. Alignment, visualization, and editing were performed using Jalview (http://www.jalview.org/). Visualization and editing of the 3D structures were performed using UCSF Chimera. Homology modelling of 3D protein structures was performed using Swiss-Model (https://swissmodel.expasy.org).

2.4 - *S. pombe* RNA analysis techniques

2.4.1 - RNA extraction from *S. pombe*

To extract RNA from *S. pombe,* 15 mL of the cell culture was grown in YES or EMM, according to the genotype, until OD_{600}=0.5. Cells were harvested at 4000 x *g* for 5 minutes at 4°C. The supernatant was removed and the pellet was resuspended in 1mL of cold sterile H_2O, and transferred into a sterile 1.5 mL Eppendorf tube. Cells were harvested for 10 seconds at maximum speed at 4°C. The supernatant was removed and the pellet was resuspended in 400 μL of AE buffer (50mM NaOAc pH 5.3, 10mM EDTA), 40 μL of 10%

SDS and 400 μL of acid phenol. The tube was vortexed for 20 seconds, and incubated at 65°C into a heat block, for 10 minutes. After incubation, the sample was frozen at -80°C for 10 minutes; then it was defrosted on the bench and spun at maximum speed at room temperature for 10 minutes. The aqueous upper phase was transferred into a cold 1.5 mL Eppendorf tube containing 400 μL of acid phenol: chlorophorm: isoamyl (25:24:1), vortexed for 30 seconds and spun at maximum speed for 10 minutes at room temperature. The upper phase was then transferred into a cold 1.5 mL Eppendorf tube containing 400 μL chloroform, vortexed for 30 seconds and spun at maximum speed for 5 minutes at room temperature. Lastly, the upper phase was transferred to a fresh Eppendorf tube containing 1 mL of 100% EtOH and 40 μL of 7.5 M NH4Ac. The tube was inverted for a few times and RNA was precipitated for 2 hours at -80°C. The RNA was pelleted at maximum speed for 10 minutes at 4°C; EtOH was completely removed and the pellet was washed with 500 μL of 70% EtOH. After centrifugation at the same condition, the pellet was air-dried at room temperature for 20 minutes and resuspended in 20 μL of ddH$_2$O. The concentration of 1 μL of purified RNA was checked via nanodrop (ND-1000 spectrometer), and the remaining RNA was stored at - 20°C for downstream applications.

2.4.2 - Agarose gel electrophoresis of RNA

To check the quality of RNA, 4-5 μg of RNA (approximately 1 μL), were diluted in 10 μL of DEPC treated H$_2$O. Then 2 μL of 6x DNA gel Loading Dye (NEB) was added to the sample. The RNA samples were loaded in a 1% agarose gel, which was made by using DEPC treated H$_2$O. RNA was run at 70 V and bands were analyzed by the use of UV.

2.4.3 - Northern Blot

All the solutions were made in DEPC-treated dH_2O. The samples were run on 1% agarose gel, prepared as shown in Table 2.

Reagents	Quantity
Agarose	1.5 g
10x MOPS	15 mL
Formaldehyde (37%)	7.5 mL
DEPC H_2O	127.5 mL

Table 2 - Agarose formaldehyde gel composition for Northern Blot analysis

The RNA samples for Northern Blot analysis were prepared on ice, by mixing 5 μg of RNA, 1 μL of Ethidium bromide, 5 μL of denaturing buffer and DEPC-dH_2O up to 10 μL. Then the RNA samples were denatured at 70°C for 5 minutes and quickly placed on ice. The electrophoresis chamber was filled with approximately 600 mL of 1xMOPS buffer. 10 μL of RNA samples were loaded in each lane in parallel with 5 μL of 6xDNA loading buffer (2.5 mg/mL bromophenol blue, 2.5 mg/mL xylene cyanol, 39% glycerol, 10mM EDTA pH 8.0), used to check the electrophoresis run. The gel was run at 80V until the bromophenol blue reached 2/3 of the gel. After the run, the fractionated RNA was blotted overnight into a neutral nylon Amersham Hybond-N membrane (GE Healthcare), and UV-cross-linked to immobilize the RNA. After the cross-linking, the membrane was placed into a glass hybridization tube, to which 13 mL of RNA Hybridization Buffer (PerfectHyb, Sigma) was added. The pre-hybridization was performed for 45 minutes at 42°C. The probes were prepared by mixing 1 μL of 10 μM of a specific primer, 1 μL of T4 Polynucleotide Kinase

(NEB), 1.5 µL of T4 Polynucleotide Kinase reaction buffer, 1 µL of Adenosine triphosphate, labeled on the gamma phosphate group with ^{32}P, and DEPC-dH$_2$O to a final volume of 15 µL. The reaction was incubated at 37°C for 30 minutes and inactivated at 65°C for 20 minutes. The labeled probes were added to the pre-hybridized membrane and incubated for 5 hours at 42°C. Then the membrane was washed 3x5 minutes with 100 mL of hot 2xSSC containing 0.1% SDS. After the final washing, the membrane was wrapped in a cling film and the signal was visualised using the Quantity One Software.

2.4.4 - Polysome Profiling

2.4.4.1 - Preparation of sucrose gradients

All the solutions were prepared in DEPC-treated sterile H$_2$O, and all the steps were carried out in ice or at 4°C to avoid RNA degradation. The sucrose gradient used in this study varied from 10 to 50 %, and it was prepared using a Biocomp Master Base Unit (Wolf Laboratories), by mixing sucrose solutions at 10% and 50% (prepared in 10mM Tris-acetate pH 7.4, 70mM ammonium acetate, 4mM magnesium acetate and 100 µg/mL cycloheximide). In brief, using a cannula attached to a syringe, the 10% solution was firstly dispensed into an SW41 centrifuge tube, until the solution level reached almost half of the tube. Then the 50% solution was dispensed by inserting a clean cannula at the bottom of the tube and slowly filled until half of the tube. Then, the tube was positioned into the rack of the Biocomp Master Base Unit, and the gradient was made using the 1-step long 10-50% sucrose program. The gradients were kept at 4°C and used just before the loading of the samples.

2.4.4.2 - Cell lysis

An *S. pombe* cell culture was grown in 50 mL of YES until OD_{600}=0.5. Then the cells were treated with 100 μg/mL cycloheximide for 5 minutes at 30°C and the culture was immediately poured into a 50 mL centrifuge tube containing ice. The cells were harvested at 1000 x *g* and resuspended in a cold wash buffer (20 mM HEPES pH 7.4, 2 mM magnesium acetate, 2 mM magnesium chloride, 100 mM potassium acetate, 100 μg/mL cycloheximide). The cells were pelleted at the same condition and resuspended in 600 μL of lysis buffer (wash buffer with 1 mM dithiothreitol (DTT), 1 mM phenylmethanesulfonyl fluoride (PMSF), 1x cOmpleteTM Mini EDTA-free Protease Inhibitor Cocktail (Roche) and either 10 mM Ribonucleoside Vanadyl Complex (RVC) (New England Biolabs) or 20 units of RiboLock (Thermo Fisher Scientific) as an RNase inhibitor). The cell suspension was then transferred into a clean 2 mL-screw cap tube, containing approximately 1 mL of either acid-washed glass or Zirconia silica beads (0.5 mm diameter, Biospec products, distributed by Cole-Palmer). To break the cells, a tissue homogenizer (Precellys 24, Bertin instruments) was used: a cell lysate was obtained by applying 3x15s beads shaking pulse at 5500 rpm. Between each pulse, the sample was cooled on ice for two minutes. To recover the lysate, the 2 mL screw-cap tube was pierced on the bottom using a 25G needle; then the screw cap was positioned into a 5 mL syringe, located itself into a 15 mL tube. The 15 mL tube was then centrifuged at 1000 x g for 5 minutes at 4°C and the lysate was transferred into a clean 1.5 mL Eppendorf tube. The cell debris was pelleted by centrifugation at maximum speed for 5 minutes at 4°C; then the supernatant was transferred into a new 1.5 mL tube and centrifuged at the same conditions for a second clearing.

2.4.4.3 - Polysome fractionation

To evaluate the volume of the cell lysate to load into the sucrose gradient, the absorbance abs_{260} of 1 μL of lysate was measured by the use of a nanodrop (ND-1000 spectrometer). The volume corresponding to 20 abs_{260} was loaded into the gradient, and the tubes were positioned into the bucket of a Beckman SW41 rotor. Centrifugation was carried out at 1200 x g at 4°C for 150 minutes. After sucrose fractionation, the sample was recovered from the bottom by the action of a peristaltic pump (P-1, Pharmacia), attached to a capillary tube. By the use of a flow-through UV spectrophotometer (ISCO UA-6 detector), the absorbance at 254 was constantly measured and recorded using a chart recorder (PicoLog Data Logger). Each fraction within the sucrose gradient was recovered using a fraction collector (FRAC100, Pharmacia). Samples were stored at -20°C or processed immediately for downstream applications.

2.4.4.4 Protein extraction from polysomal fractions

The proteins were extracted from the polysomal fractions using the Methanol/Chloroform method. A 150 μL aliquot of polysomal fraction, was mixed with 600 μL of Methanol, 150 μL of Chloroform and 450 μL of sterile H_2O. The sample was centrifuged at maximum speed for 5 minutes at room temperature. After centrifugation, a white layer of proteins formed between the bottom organic phase and the upper aqueous phase: the aqueous phase was removed, and 650 μL of methanol was added to the sample. The sample was centrifuged at 15700 x g for 5 minutes at RT. The supernatant was completely removed and the pellet was air-dried for 5-10 minutes. The precipitated proteins were resuspended in 35 μL of 1X SDS loading buffer (4% SDS, 0.04% bromophenol blue, 20% glycerol, 60mM Tris HCl, pH 6.8, 4% mercaptoethanol, 0.01% and 1 mM phenylmethanesulfonylfluoride (PMSF)), and boiled

for 5 minutes. The protein extracts were run on 10-12% SDS polyacrylamide gel and analysed by Western Blot.

2.5 - *Escherichia coli* methods

2.5.1 - *E. coli* Strains

The *E. coli* strains used to amplify plasmids were DH5 alpha, with genotype: F⁻ endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20 φ80dlacZΔM15Δ(lacZYA-argF) U169, hsdR17($r_K^-m_K^+$), λ⁻. The plasmids used in this project can be found in Appendix IV.

2.5.2 - *E. coli* media and growth condition

Luria-Bertani (LB) liquid media and LB agar plates were used to grow *E. coli.* Typically, a single colony of *E.* coli was inoculated in 5 ml and grown at 37°C overnight using a shaking incubator at 220 rpm. The subsequent culture was used to scale up to the desired volume. The ampicillin was added to LB liquid media to a concentration of 100 μg/mL. For plated cells, they were grown overnight on LB agar plates at 37°C overnight. Plates were supplemented with ampicillin at the same concentration.

2.5.3 - Bacterial Transformation

1 μL (1 ng) of plasmid was added to 50 μL of DH5 alpha competent cells, gently mixed and incubated on ice for 30 minutes. After incubation, cells were heat-shocked at 42°C for 30

seconds using a water-bath. Cells were immediately transferred on ice for 2 minutes. To recover the cells after the heat-shock, 500 μL of NZY media (details of which is reported in Appendix I) was added and cells were incubated at 37 °C for 1 hour in a thermo-shaker at 250 rpm. 100 μL of cells was spread on LB agar plates containing 100 μg/mL Ampicillin. Cells were grown overnight at 37°C.

2.5.4 - Cloning by Gibson assembly

To clone a fragment of DNA into a vector, the Gibson Assembly Cloning technique was adopted in this study. The fragment of DNA was amplified via PCR using primers forward and reverse with at least 25 bp of homology to the region of insertion on the plasmid (Figure 2.2). The plasmid was amplified via PCR, with primer forward and reverse with the same 25 bp of homology of the fragment. The two products were then purified using the QIAquick PCR purification kit (Qiagen) or the PCR and DNA cleanup kit (Monarch). The assembly reaction was performed using the NEBuilder HIFI DNA assembly Master Mix (NEB). The reaction was set up by combining 3 μL of 2x Gibson Assembly Master Mix, 0.5 pmols of each DNA fragment and 6 μL of sterile water. The mix was incubated in a thermocycler at 50°C for 15 minutes. After incubation 10 μl of the mix was used to transform the DH5 alpha competent cells. A schematic representation of the assembly is shown in Figure 2.2
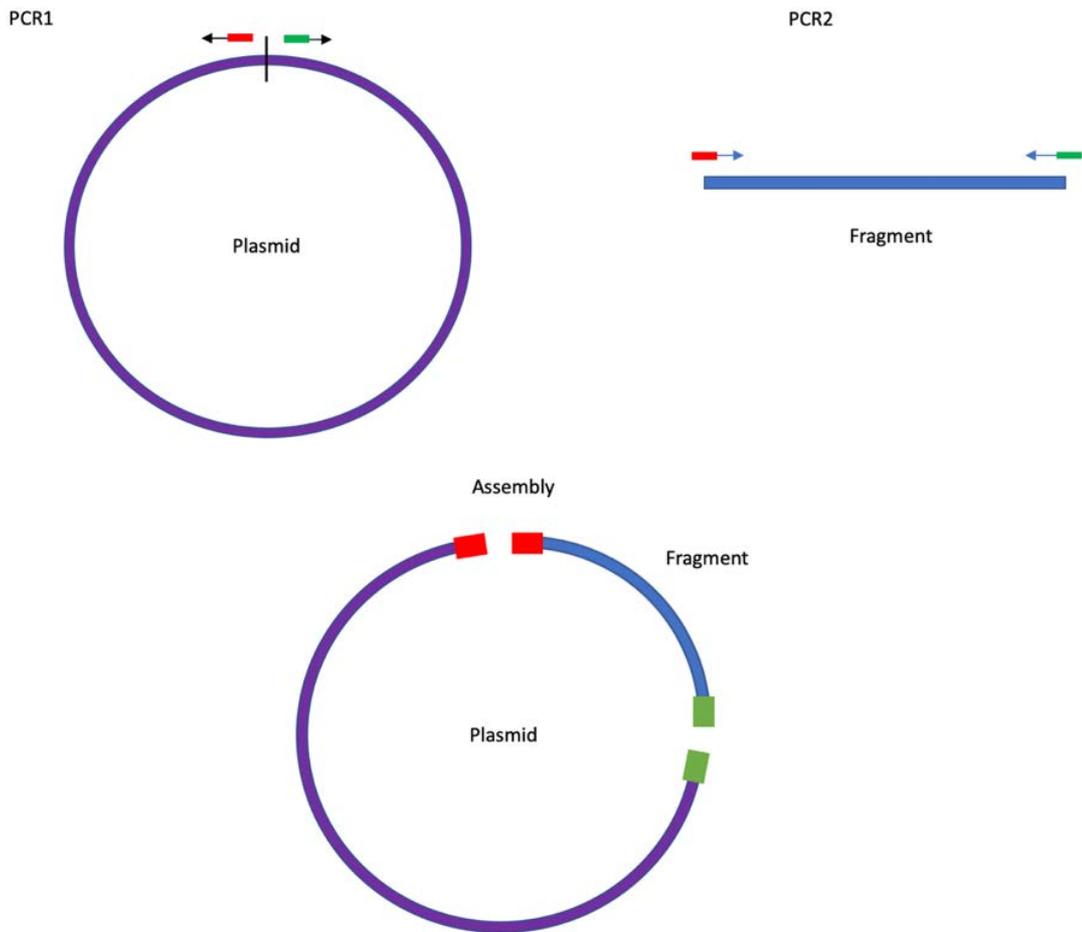
Figure 2.2 - Gibson assembly schematic representation. A DNA fragment to insert into a plasmid of interest are both amplified by PCR, using hybrid primers having at least 25 nucleotides of homologous sequence. The PCR amplified products are purified using a Monarch PCR Purification Kit (or a QIAquick PCR Purification Kit). The Gibson assembly is then performed by combining the purified fragment and the vector backbone with the Gibson Assembly Master Mix.

2.5.5 - Small scale plasmid preparation (Method 1)-Boiling prep

A single colony of DH5 alpha was inoculated in 2-3 mL of LB liquid media and grown overnight at 35.5°C. After 12 hours, the culture was poured into a 1.5 mL tube and centrifuged for 5 minutes at 300 g. The supernatant was discarded and the pellet was resuspended in 400 μL of STET buffer (8% sucrose, 0.1% Triton-X 100, 50 mM EDTA pH8, 50 mM Tris pH 8) and 30 μL of lysozyme (Sigma). The mixture was firstly incubated at room temperature for 5 minutes and then boiled for further 3 minutes. Cells were pelleted at maximum speed for 10 minutes and the supernatant was transferred into a sterile 1.5 mL tube, to which 60 μL of RNase A (10mg/mL) and 50 μL of NaAc pH 5.2 was added. The sample was incubated at room temperature for 5 minutes, after which 1 volume of isopropanol, was added and centrifuged at maximum speed for 10 minutes. The supernatant was discarded and the pellet was washed with 650 μL of 75% EtOH and air-dried. The DNA was dissolved in 50 μL of 1xTE buffer and plasmid concentration was evaluated by nanodrop (ND-1000 spectrometer).

2.5.6 - Small scale plasmid preparation (Method 2)

A second method was adopted at the beginning of this study, before Boiling Prep (2.5.5) was optimized. A single colony of DH5 alpha was inoculated overnight in 2 mL of LB liquid media at 37°C using a shaking incubator at 220 rpm. After 12 hours, the culture was poured in a 1.5 ml tube and centrifuged for 5 minutes at 300 g. Plasmid DNA was purified using an Invitrogen mini-prep kit. The final DNA pellet was resuspended in 50 μL of TE buffer and the concentration of the plasmid was evaluated by nanodrop (ND-1000 spectrometer).

2.5.8 - Plasmid diagnostic digestion

Plasmids were cut using restriction enzymes from NEB. The digestion conditions varied according to the restriction enzyme: (https://www.neb.com/tools-and-resources/usage-guidelines/nebuffer-performance-chart-with-restriction-enzymes). Typically, 1 µg of DNA plasmid was digested. The reactions were set up as shown in Table 3. The buffer was chosen according to the performance chart. After reaction set-up, the mix was incubated at 37°C, for an hour, and products were resolved on a 1% agarose gel.

Reagent	Quantity
10x NEBuffer	5µL
Restriction enzyme	1µL
Plasmid DNA	1µg
ddH$_2$O	Up to 50

Table - 3 Diagnostic digestion set-up conditions

Chapter 3: Application of the Auxin-Inducible Degron (AID) system for the rapid depletion of Fal1 in *S. pombe*

3.1 - Introduction

Numerous methods have been developed to control the expression of ectopic genes in *S. pombe*. The *nmt1* (*no message* in *thiamine*) gene of *S. pombe* is typically highly expressed but it is subjected to transcription repression when the media is supplemented with thiamine (Vitamin B1) (Maundrell, 1990). The promoter of this gene, Pnmt1, can be exploited to substitute the regulatory elements of heterologous essential genes, and to evaluate the subsequent phenotype upon thiamine repression (Basi et al., 1993).

However, the strong activity of the Pnmt1 makes this tool unsuitable to repress genes which are already expressed at low levels because even in the presence of thiamine, the original Pnmt1 retains a low level of transcription activity (Watson et al., 2013).

To overcome this issue, two weaker versions of the *nmt1* promoter, the Pnmt41 and the Pnmt81, derived by nucleotide changes in the TATA box, have been developed, and represent the most common systems to "shut-off" an essential gene in *S. pombe*. However, these versions still require a considerable temporal window, from 6 to 21 hours (depending on the half-life of the protein), to cause a substantial reduction of the levels of the target protein (Basi et al., 1993, Maundrell, 1993). Since the cell cycle is completed within three hours for most of the strains that I used in this project, these promoters represent a clear drawback to answer advanced questions relative to gene expression (Watt et al., 2008). A truncated form of the Pnmt1 has been also characterized; the promoter is repressed in a shorter time, but it

requires a temperature shift rather than thiamine repression, thus causing relevant changes in the gene expression profile due to stress conditions (Kumar and Singh, 2006, Watt et al., 2008). A faster alternative to regulate the expression of a target gene at transcriptional levels, is represented by the *urg1* promoter: Watt and colleagues have shown that upon addition of uracil, the promoter is activated in less than 30 minutes, and a similar rapid inactivation is observed upon removal of uracil (Watt et al., 2008). However, several studies have argued about the utility of the urg1 system, due to a pivotal caveat: despite a substantial reduction in the levels of mRNA, the basal level of the protein regulated by the *urg1* promoter is still too high to determine a null-like phenotype (Watson et al., 2013).

Another strategy available in *S. pombe* is to sequester the protein of interest or delocalize it in another cellular compartment. For example, the target can be endogenously tagged with the hormone-binding domain (HBD) of the estrogen receptor (ER). The tag attracts the Hsp90 complex, which binds the HBD, rendering the protein inactive: however, upon addition of estradiol to the media, the hormone binds the HBD domain, causing a change of conformation, with the release of the Hsp90 and subsequent activation of the protein of interest (Bøe et al., 2008). Despite the system appearing conceptually attractive, the key limitation of this tool is the cytoplasmic localisation of the Hsp90 complex, which might limit the application only to those proteins which function in the cytoplasm (Pai et al., 2012).

Similar to the HBD system, another way to sequester a protein in another cellular compartment is represented by the Anchor-Away (AA) system originally developed in *S. cerevisiae* to deplete nuclear proteins (Haruki et al., 2008). The anchor away system is based on the use of rapamycin which allows the interaction between two protein domains: the FRB domain, of the TOR1 protein, and the FKBP12 domain, present in the FRP1 factor (Otsubo and Yamamato, 2008, Weisman, 2004). Therefore, the application of this method requires

that the protein of interest is endogenously tagged with the FRB domain, while the FKB12 domain will be added to a highly expressed protein that rapidly translocates from the nucleus to the cytoplasm, such as a ribosomal protein. In the presence of rapamycin, the protein of interest will be quickly transported to the cytosol, permitting the analysis of the consequent phenotype. Since translocation of the target occurs rapidly, it has been observed that after 30-60 minutes from addition of rapamycin, the nucleus appears depleted of the protein of interest. However, several limitations are associated with the AA system: firstly, since the protein sequestration occurs in the cytoplasm, it can be mainly applied to those proteins which have a nuclear function. Secondly, the sequestration of the target protein from the nucleus to the cytoplasm may cause the translocation of interacting partners, thus impeding the correct discrimination of the protein function. Lastly, it appears to be protein-specific, and not applicable to all the nuclear proteins (Haruki et al., 2008).

The Auxin-Inducible Degron (AID) is another tool originating from plants, whose activation permits the rapid degradation of a protein of interest upon addition of the plant hormone auxin (Nishimura et al., 2009). The correct growth and development of a plant can require a spatio-temporal regulation of many proteins, regulated as in other organisms, by the release and action of different hormones. In the case of auxin, whose name derives from the Greek word "auxien", which means to grow, it specifically targets transcriptional repressors containing a typical IAA17 epitope. In brief, auxin binds the pocket of a perceptive F-box protein, the Transport Inhibitor Response 1 (TIR1), acting as a molecular glue for the insertion in the pocket of an AUX/IAA transcriptional repressor. TIR1 is part of an SCF-type E3 ubiquitin-protein ligase complex, which comprises Skp1, Cullin and RBX1. The last two factors, transfer the ubiquitin from an E1/E2 enzyme system to the Aux/IAA transcriptional repressor, thus causing its polyubiquitination and degradation by the proteasome (Gray et al., 2001,

Kepinski and Leyser, 2005). In absence of auxin, the transcriptional repressor is not degraded, and its binding to the promoter of target genes prevents the activation of signaling pathways required for the growth and development of the plant (Chapman and Estelle, 2009).

Application of the AID tool in non-plant organisms requires that the gene of interest is tagged with the IAA17 epitope, normally referred to as AID, and the transplantation of the TIR1 element from plant to the desired organism, typically inserted in a silent area of the genome. The proteins forming the SCF complex (Skp1, Cul1, and Rbx1) are all highly conserved across eukaryotes, therefore they are not required to be transplanted for the engineering of the AID system (Figure 3.0) (Natsume et al., 2016). Normally, a successful application of the AID system allows the degradation of a target protein after 30 minutes by the addition of auxin (Natsume et al., 2016, Nishimura et al., 2009, Tanaka et al., 2015).
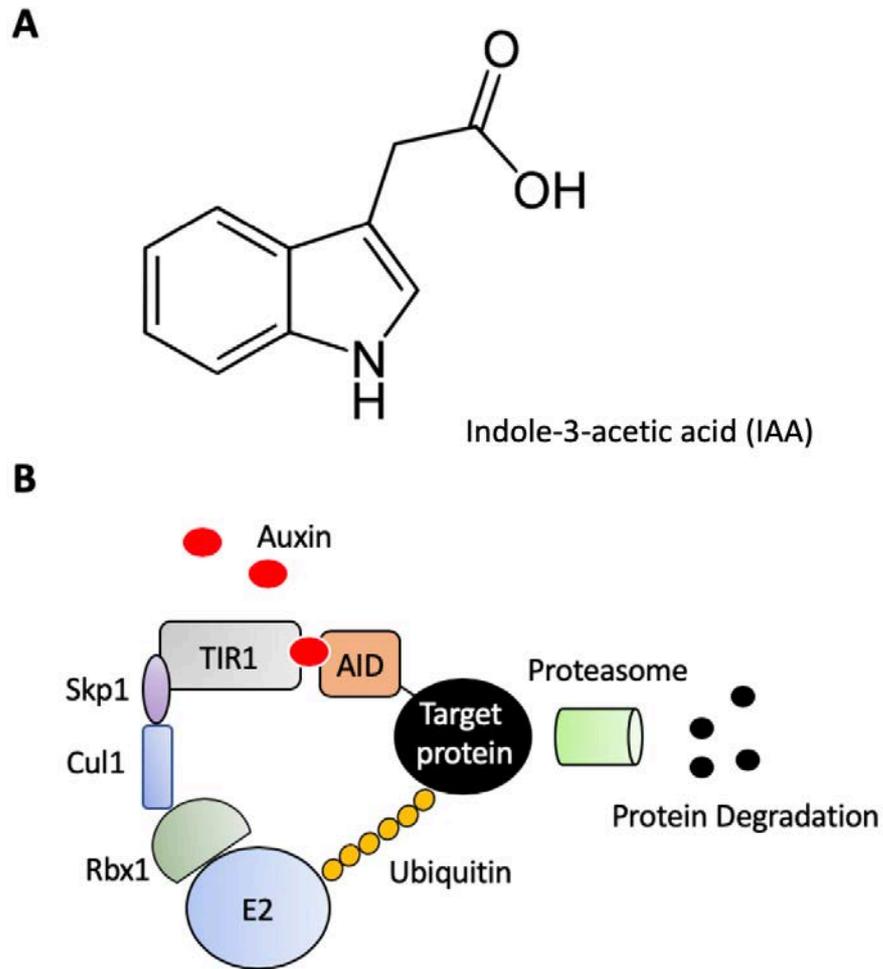
Figure 3.0 - Chemical structure of the Indole-3-acetic acid (IAA) and the mechanism of action of the Auxin-Inducible degron system (AID). (A) The Indole-3-acetic acid is the most basic and abundant auxin in plants, which is formed by an aromatic ring (indole) and a carboxylic acid group. (B) Schematic representation showing the mechanisms of action of the Auxin-Indicuble degron tool: auxin binds TIR1 (present in SCF-TIR1 complex), which induces the interaction of TIR1 with the IAA17(AID)-tagged protein. SCF-TIR1 recruits E2, which causes the ubiquitination of the target protein and the subsequent degradation by the proteasome.

Another system that allows the rapid degradation of the target protein is known as degradation tag (dTAG) (Nabet et al., 2018). This tool requires the fusion of the protein of interest with the FKBP12 protein, which contains a point mutation (F36V), causing a cavity in FKBP12. The "hole" is recognized by a bifunctional degrader (dTAG), which also contains a binding site for an E3 ligase. The formation of the ternary complex, formed by FKB12^{F36V}, dTAG and E3, causes the rapid ubiquitination and degradation of the target protein (Nabet et al., 2018).

Rapid degradation of the protein of interest can be also achieved using a Heat-Inducible Degron system. In *S. cerevisiae*, it was reported that a Dihydrofolate Reductase (DHFR), harboring a destabilizing arginine residue at the N-terminal, is normally expressed at 23°C, but it is ubiquitinated and degraded by the proteasome when the temperature is shifted to 37°C. Thus, the arg-DHFR element can be fused with the target protein to trigger its rapid degradation after temperature shift (Dohmen et al., 1994). This system has been further modified: Iwamoto and colleagues reported that a DHFR mutant protein from *E. coli* (ecDHFR) can confer instability to a fused protein, which is rapidly degraded. However, ecDHFR can be stabilized by a small high-affinity ligand know as Trimethoprim (TMP). As a consequence of this mechanism, in absence of TMP, the protein levels of the target are hardly detectable (Iwamoto et al., 2010).

In this chapter we present the application of a system, which we termed "OFF-AID system", for the rapid depletion of Fal1, the homolog of eIF4AIII in *S. pombe,* which exploits both the *nmt81* promoter and the AID system, acting therefore at two levels of regulation: gene expression and protein turnover. The final goal is to rapidly deplete Fal1 from fission yeast and evaluate whether it has a direct function in transcription and pre-mRNA processing, by extracting and sequencing the nascent-RNA (Churchman and Weissman, 2011).

Some of the experiments presented in this Chapter were done under the supervision and in collaboration with Dr. Kayoko Tanaka, a Principal Investigator at the University of Leicester, who also made some of the figures below: Figures 3.2B, 3.5B, 3.7B, and 3.10A.

3.2 - Results

3.2.1 - Fal1 is conserved in *S. pombe*

Before embarking on the application of the AID system for the rapid depletion of Fal1, I first evaluated whether Fal1 is conserved in *S. pombe.* With this analysis, I aimed to investigate whether future prospectives coming from the study of Fal1 might be applied in higher eukaryotes. To this aim, the Fal1 protein sequence obtained from *PomBase* was aligned with the protein sequences of eIF4AIII across common model eukaryotes. The alignment was performed using Jalview, software for multiple sequence alignment, editing, visualization and analysis (Clamp et al., 1998, Clamp et al., 2004).

The result of this investigation confirmed that eIF4AIII is conserved across eukaryotes and that Fal1 shows high levels of similarity with eIF4AIII of Human (79%), *D. melanogaster* (78%), *S. cerevisiae* (65%), *C. elegans* (78%), *Arabidopsis* (72%) and *Xenopus* (76%) (Figure 3.1). Additionally, the domains for the binding/hydrolysis of the ATP and the RNA binding appear particularly highly conserved among species, including *S. pombe*, except for two single amino acid substitutions; a phenylalanine which substitutes a threonine in the domain Ib, which is important to establish interactions with the sugar-phosphate backbone of the RNA, and a threonine which substitutes a lysine in the Motif Q, which regulates the hydrolysis of the ATP (Figure 3.2). Despite these two changes, this analysis reveals that Fal1 might have the same, or at least comparable, ATP and RNA binding properties as in other eukaryotes and that *S. pombe* is a suitable model organism to study the *in vivo* function of eIF4AIII.
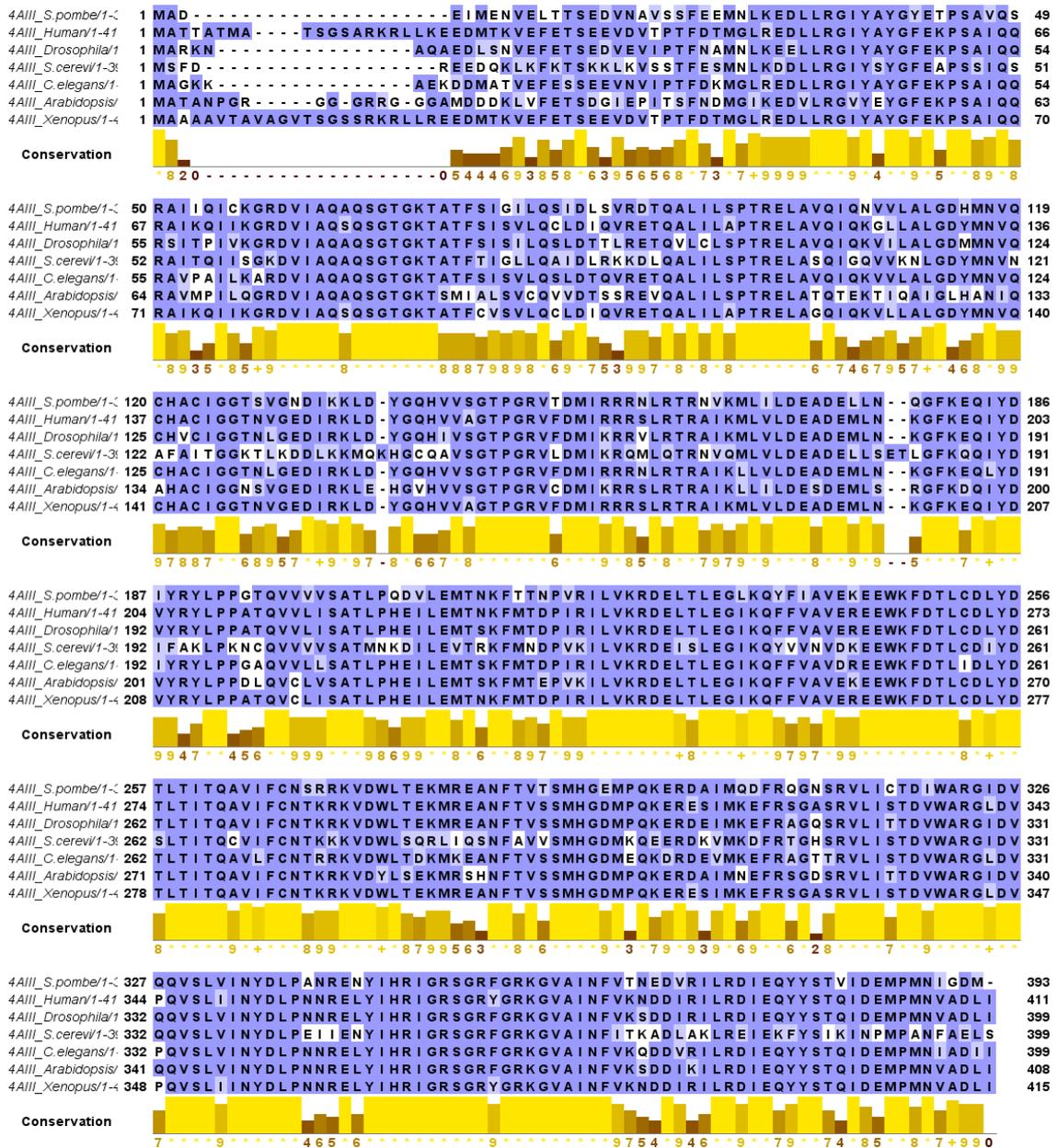
Figure 3.1 - Fal1 amino acid sequence shows high similarity with the eukaryotic counterparts. Fal1 shows a high percentage of identity with its counterparts in Human (78%), *D. melanogaster* (79%), *C. cerevisiae* (65%), *C. elegans* (78%), *Arabidopsis* (72%) and Xenopus (76%). Alignment and visualization are performed using Jalview. Identical residues are indicated in dark blue; gaps are indicated in white; residues with a positive score in the Blosum62 matrix are indicated in light blue.
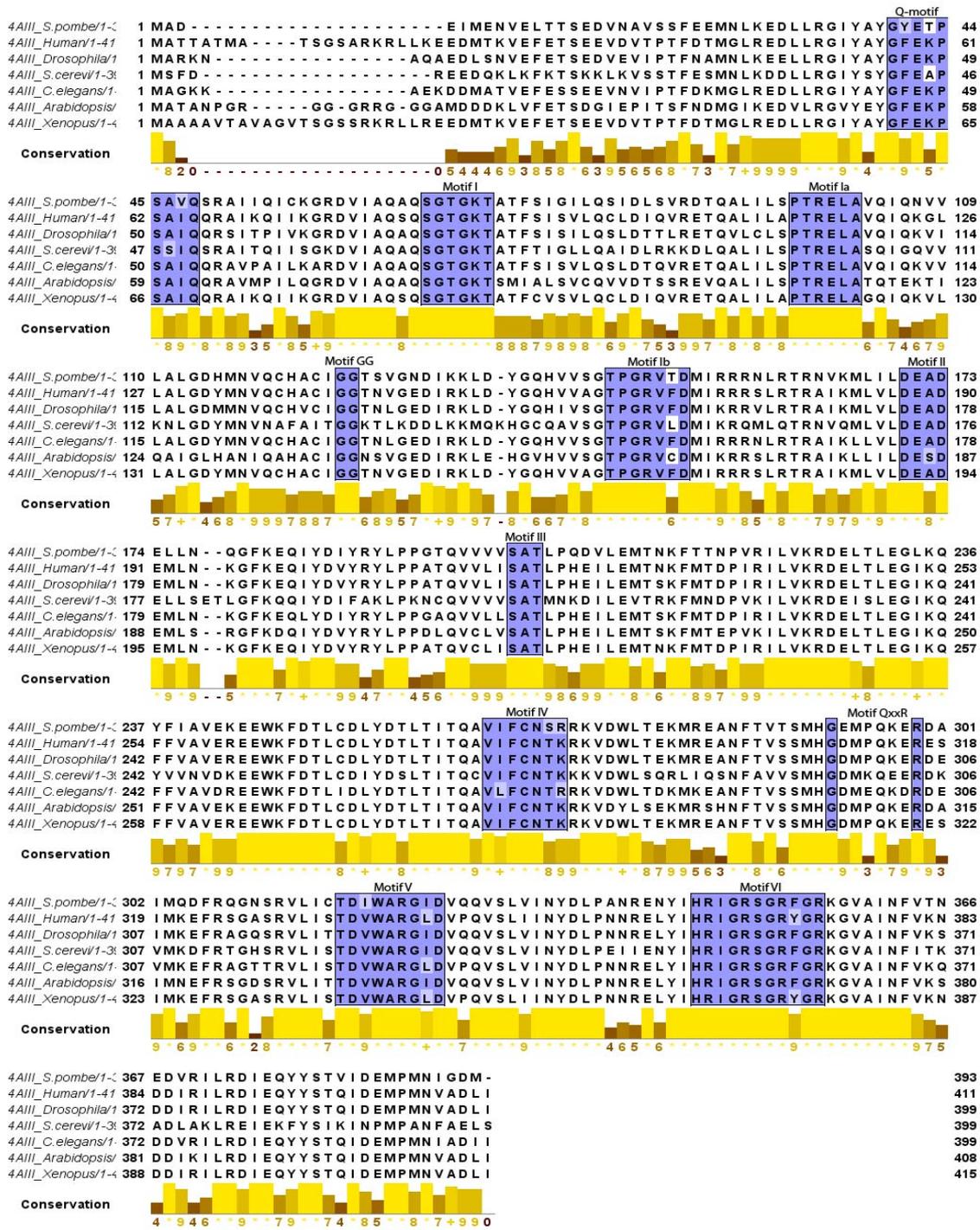
Figure 3.2 - Fal1 protein motifs are highly conserved across eukaryotes. Motifs involved in the ATP binding/hydrolysis and interaction with the substrate are aligned using Jalview and conservation grade is evaluated using the amino acids substitution matrix Blosum62.

3.2.2 - Fal1 is required for normal cell growth in *S. pombe*

The deletion of *eIF4AIII*, or reduction of its protein levels, correlates with inviability or severe growth defects in all the organisms in which its function has been investigated (Choudhury et al., 2016, Favaro et al., 2014, Haremaki et al., 2010, Kittler et al., 2004, Kressler et al., 1997, Marayati et al., 2016).

In *S. cerevisiae,* tetrad analysis performed after replacing the *fal1* ORF with a gene cassette conferring resistance for G418 has shown that deletion of *fal1* is lethal for budding yeast (Kressler et al., 1997). In *D. melanogaster* the depletion of eIF4AIII associates with an aberrant organization of the cytoskeleton, reduction of nuclei size and disaggregation of the follicular epithelium, a clear indication of cell death (Palacios et al., 2004). Choudhury and colleagues also reported growth defects in the salivary glands (Choudhury et al., 2016).

The knock-down of eIF4AIII in the frog *Xenopus laevis* results in whole body paralysis and defects in the neural crest (Haremaki et al., 2010).

In mice, eIF4AIII haploinsufficiency also causes neural crest defects, and in human, a noncoding expansion in the 5'-UTR of the gene leads to acrofacial dysostosis, defined as Richieri-Costa Pereira syndrome (RCPS) (Favaro et al., 2014, Miller et al., 2017). Notably, the depletion of eIF4AIII in zebrafish causes craniofacial alterations similar to RCPS (Favaro et al., 2014). Lastly, the depletion of eIF4AIII in HeLa cells, causes a mitotic arrest and alteration of the morphology of the spindle (Kittler et al., 2004). Overall these observations suggest that eIF4AIII plays a crucial role in growth and development across eukaryotes.

Previous attempts to generate a *fal1* deletion strain in fission yeast were not successful, suggesting that the gene could be essential for *S. pombe* (Wen and Brogna, 2010). Meanwhile, Marayati and colleagues reported that gene deletion results in a severe growth defect but nonetheless, the *fal1* deletion is not lethal (Marayati et al., 2016). To clarify this situation

we performed gene knock-out of *fal1* in diploid cells, followed by a tetrad dissection, which was done in collaboration with Dr. Kayoko Tanaka. The *fal1* gene disruption in heterozygous diploids (potentially *fal1+/fal1:ClonNAT*) was verified by PCR (Figure 3.3A).

Then, the heterozygous diploid strains were induced to sporulation and asci were dissected using a micromanipulator; the four spores derived by its dissection were aligned vertically on a YE+adenine (YEA) plate, and after a few days, replica plated on YEA supplemented with the selective antibiotic (clonNAT). We found that the spores derived from each ascus were able to generate colonies on the YEA plate, suggesting that the *fal1* gene was not essential; additionally, on the selective plate, just two out of four colonies containing the antibiotic resistance survived, supporting the expected *fal1+/fal1::ClonNAT* phenotype of the parental diploid cells (Figure 3.3B).

Importantly, we also noted that the colonies derived from the *fal1* deleted spores, which were ClonNAT resistant, were substantially smaller; they emerged from the plates after several days of delay compared to the WT, suggesting that *fal1* gene deletion greatly reduces cell viability (Figure 3.3B).

To confirm whether the deletion of *fal1* affects cell growth, the growth of the haploid *fal1* deletion mutant was compared to the WT strain by growth spot assay. The plates were incubated at 30°C and 32°C and survival was evaluated after 72 hours. Interestingly we found that the deletion of *fal1* associated with a severe growth defect in *S. pombe* (Figure 3.3 C), confirming the previous observation made by Marayati and colleagues (Marayati et al., 2016).

A
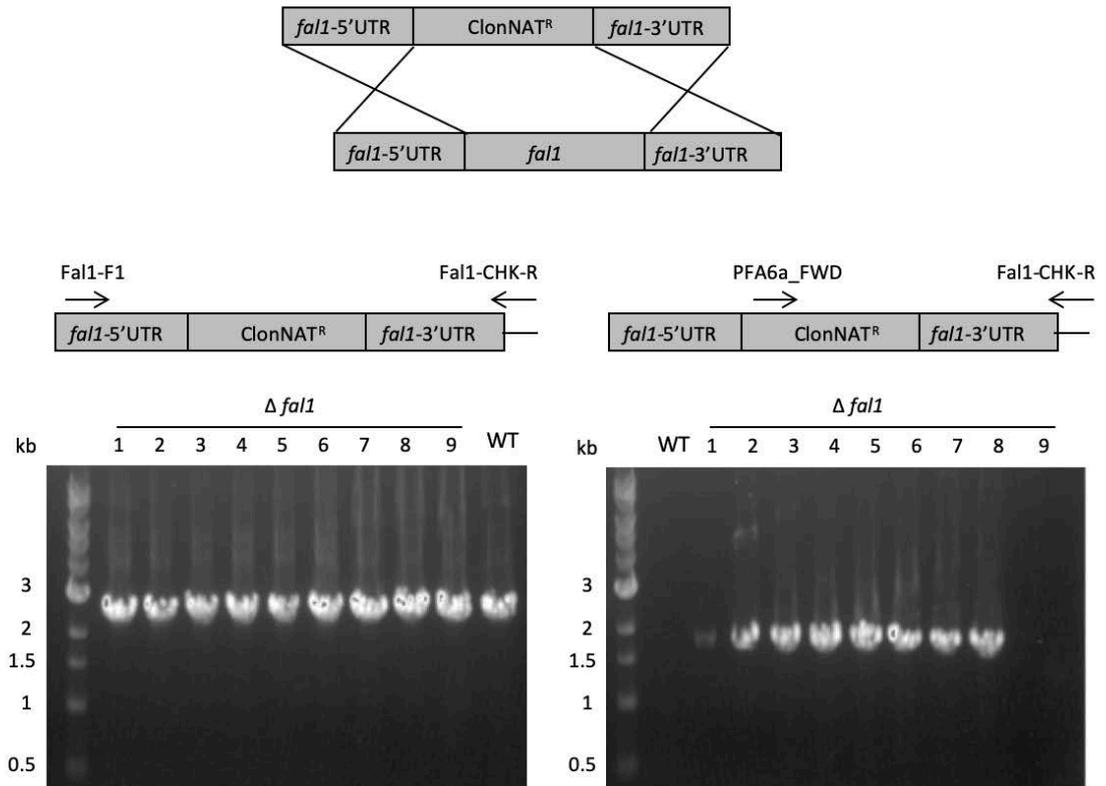
B
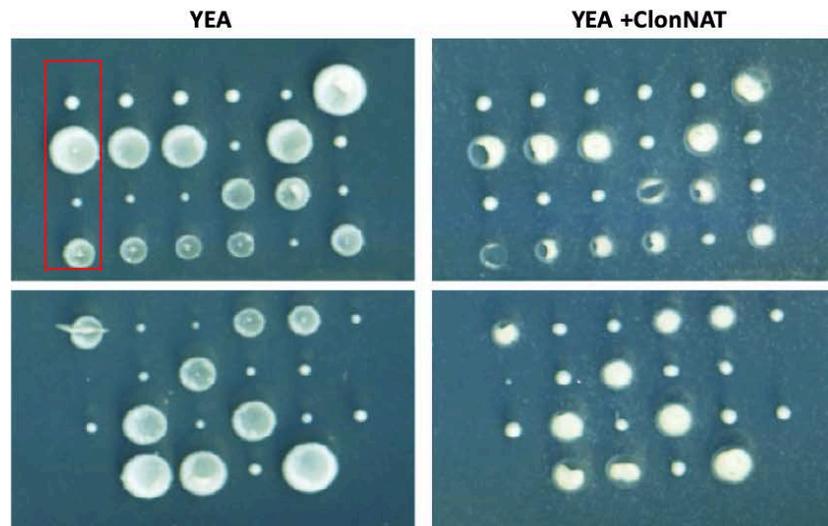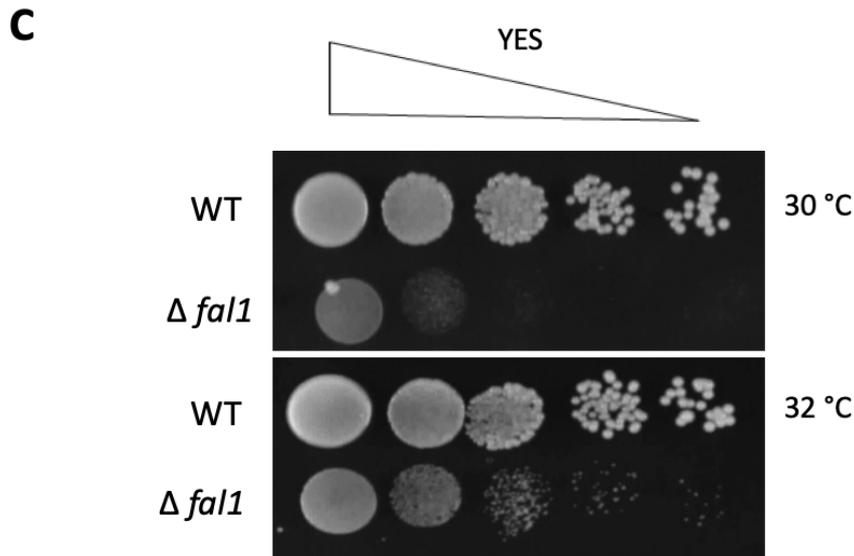
Figure 3.3 - *fal1* knock-out and tetrad dissection. (A) Verification of the deletion of *fal1* by PCR in diploid heterozygous yeast cells. Deletion of *fal1* is confirmed by genome PCR, using primers Fal-F1 and Fal-1-CHK-R, where the second anneals outside the region of recombination. An expected band of 2.3 kb is amplified in all the colonies analyzed. A second PCR is also performed using the primer pair pFA6a-FWD and Fal1-CHK-RV, where the first anneals in the ClonNAT cassette; an expected band of approximately 1.7 kb is amplified for 8 colonies, while no amplification is detected from the WT genomic DNA. A schematic representation of the gene deletion is reported in Appendix VIII.

(B) Diploid heterozygous for *fal1*+; *Δfal*1::ClonNAT are analyzed by tetrad dissection. The products of each tetrad are aligned vertically (red box) on a YEA plate, and replica plated on YEA plates containing the ClonNAT antibiotic. The tetrad dissection analysis shows the *fal1* deleted spores survive on the YEA+ClonNAT plate, suggesting that the gene is not essential. (C) The growth of the haploid *fal1* deletion mutant is compared to the WT strain, by spot growth assay on YES plates. Diluted spots contain approximately 10^5, 10^4, 10^3, 10^2, 10 cells. Plates are incubated at 30°C and 32°C and the growth is evaluated after 72 hours.

The tetrad dissection in Figure.3.2B was performed by Dr. Kayoko Tanaka.

3.2.3 - Fal1 is nuclear in *S. pombe*

As previously explained, the main goal of this chapter is to generate a system which allows the depletion of Fal1 in a rapid and conditional manner. Therefore, it is pivotal to evaluate the levels of Fal1 at different times, by fluorescence and/or WB.

Due to the lack of antibodies against Fal1 in *S. pombe*, we first generated a yeast strain in which Fal1 was endogenously tagged at the C-terminus with a double tag, the GFP-2xFLAG epitope, whose integration in the genomic locus was verified by PCR (Figure 3.4A).

The expression of the endogenous tag was confirmed by WB using an anti-FLAG antibody (M2 Sigma): Fal1, GFP and x2 FLAG have a molecular weight of 44.71, 25 and 2.2 kDa respectively, therefore a band of approximately 72 kDa was expected and detected by immunoblotting (Figure 3.4B). Since *fal1* gene knock-out affects cell viability, to assess that the protein was still functional after the addition of the tag, we performed a spot growth assay, comparing the viability of the Fal1-GFP-2xFLAG strain to the WT (non-tagged). Cells were spotted in serial dilution on a YES plate, which was incubated at 30°C for 72 hours. We found that the Fal1-GFP-2xFLAG strain had no growth defects, indicating that the tag did not affect the functionality of the protein, and it could be used for further applications (Figure 3.4C).

A

B

C

YES 30 C

WT
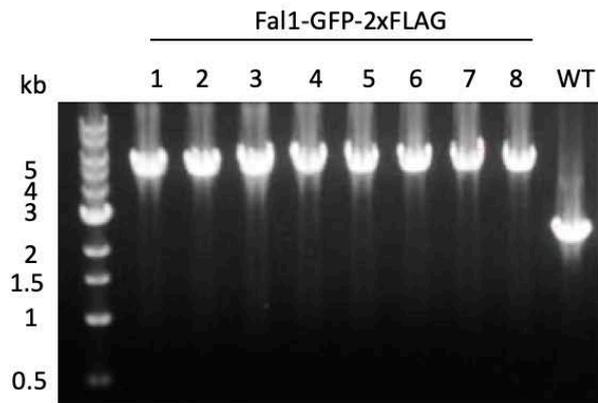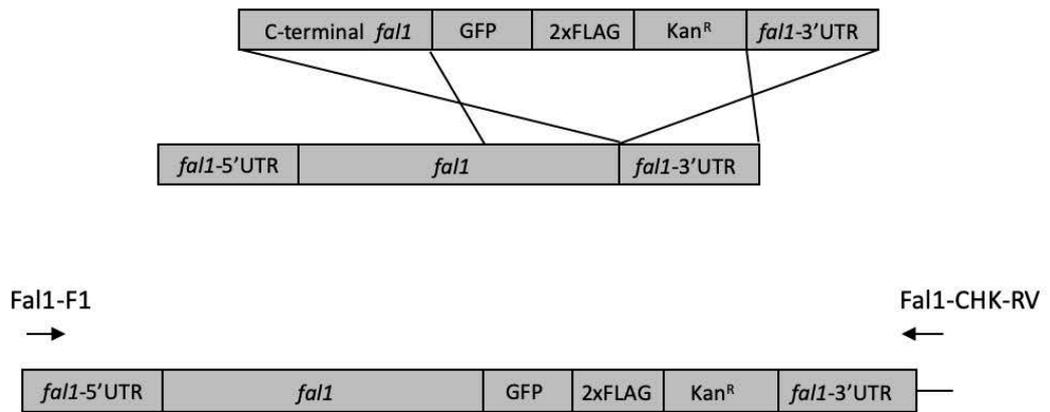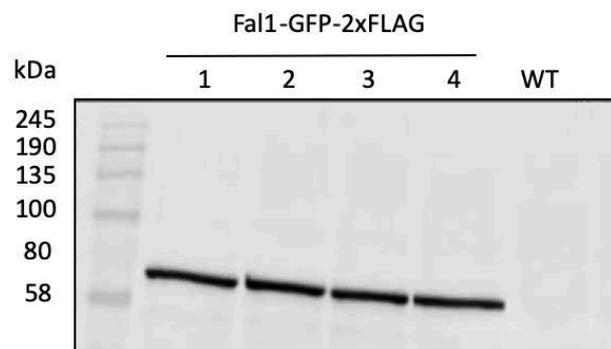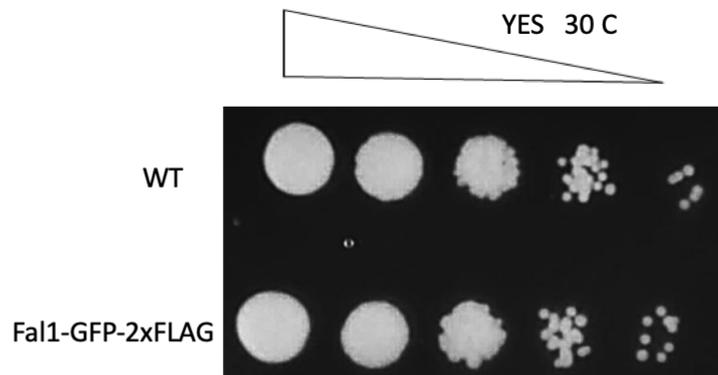
Fal1-GFP-2xFLAG

Figure 3.4 - Endogenous GFP-2xFLAG tag of Fal1 does not affect cell viability. (A) Verification of the GFP-2xFLAG tag of Fal1 by PCR: integration of the tag is confirmed by PCR using as template the genomic DNA extracted from eight selected colonies. Using the primer pair Fal1-F1 and Fal1-CHK-RV, where the second is designed outside the region of recombination, expected bands of approximately 5 kb and 2.3 kb for the tag and the WT strains respectively are amplified. A schematic representation of the gene tagging is reported in Appendix VIII.

(B) Expression of the tag is confirmed by WB, using an anti-FLAG antibody (M2, Sigma); a band of 72 kDa is detected for four colonies.

Note that the WB is qualitative and aimed to confirm the presence of the FLAG-tag in the positive colonies; therefore, the membrane was not probed with a loading control antibody.

(C) The growth of the Fal1-GFP-2xFLAG is compared to the WT, by spot growth assay on a YES agar plate. Diluted spots contain approximately 10^5, 10^4, 10^3, 10^2, 10 cells. Plates are incubated at 30°C and survival is evaluated after 72 hours.

Having tagged Fal1 with GFP, we exploited this tag to determine its subcellular localisation in *S. pombe*. Previous studies reported that eIF4AIII was predominantly in the nucleus of a wide range of organisms from flies to mammals (Ferraiuolo et al., 2004, Palacios et al., 2004). In *S. cerevisiae*, Fal1 is associated but not restricted to the nucleolar region, the site of ribosome biogenesis (Kressler et al., 1997). Nucleolar localisation of eIF4AIII was also reported in *Arabidopsis thaliana* (Pendle et al., 2005).

To evaluate the localisation of Fal1 in *S. pombe,* the Fal1-GFP-2xFLAG strain was grown until mid-log phase and the GFP signal was visualized by confocal microscopy. We found that in *S. pombe,* as in other eukaryotes, Fal1 is predominantly nuclear (Figure 3.5A). Additionally, a cell live imaging experiment showed that Fal1 persistently localised to the nuclear compartment during nuclear division, indicating that the main function of this protein is carried out in the nucleus (Figure 3.5B).
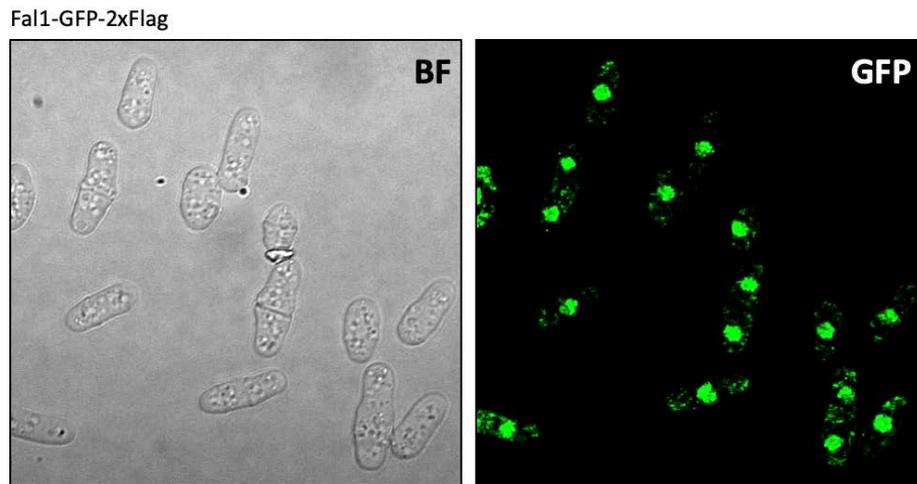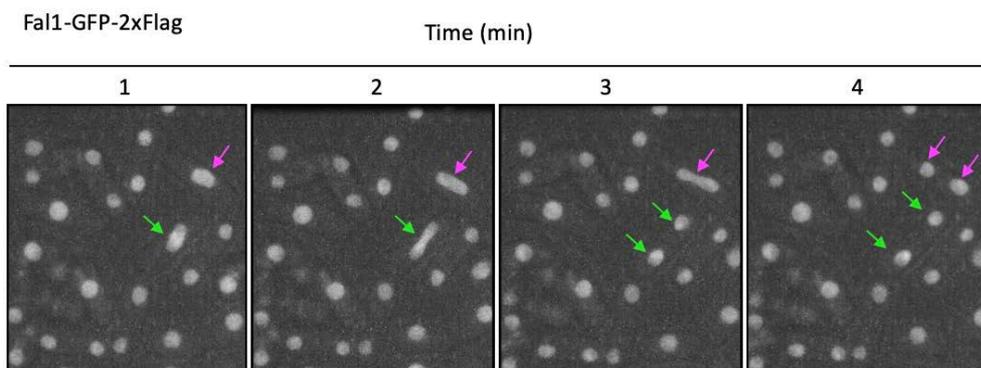
Figure 3.5 - Fal1 is localised to the nucleus of *S. pombe*. (A) The Fal1-GFP-2xFLAG strain was visualized by confocal Microscopy (Nikon A1R); the signal derived by the GFP tag is predominantly nuclear. (B) Cell live imaging of Fal1-GFP-2xFLAG: images were acquired every 60 seconds for 30 minutes using a 2D array scanning laser confocal microscopy (Infinity 3, VisiTech). The figures are representative of a time frame of 0-4 minutes. Two nuclei in fission are indicated with green and magenta arrows.

Figure B was generated by Dr. Kayoko Tanaka.

3.2.4 - Swapping the Fal1 promoter with the Pnmt81

Before tagging Fal1 with AID, the physiological promoter of *fal1* was substituted with the weakest version of the nmt promoters, Pnmt81, which is effectively repressed by thiamine (Basi et al., 1993, Maundrell, 1993, Tamm, 2012). Having previously tagged Fal1 with the GFP-2xFLAG epitope, we decided to change the promoter in this strain, so that the repression of Fal1 could be monitored both by fluorescence and immuno-blotting. The shut-off system was made through the "PCR sewing" method, using the N-terminal tag approach and the integration in the genomic locus was verified by PCR (Figure 3.6A).

The functionality of the Pnmt81 promoter was tested via WB using an anti-flag antibody (Sigma, M2), detecting an expected band of 72 kDa for two positive colonies (Figure 3.6B). Moreover, we also verified that the expression of Fal1 under the Pnmt81 promoter did not affect its nuclear localisation (Figure 3.6B).

A

B
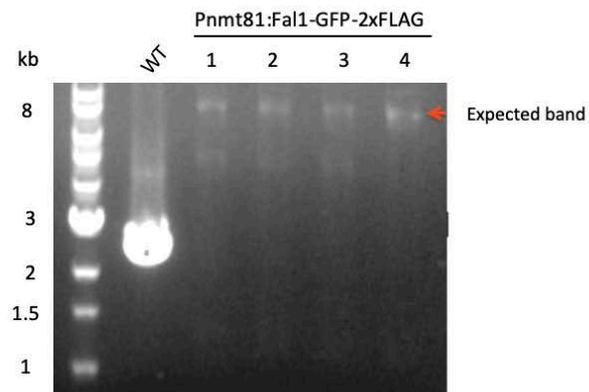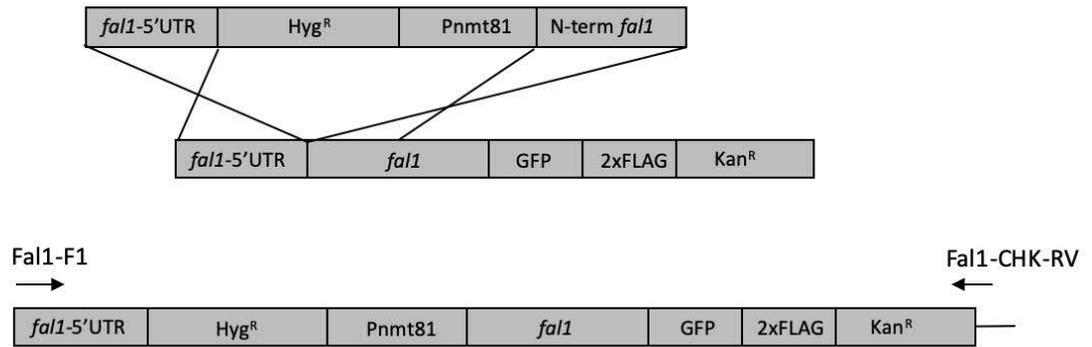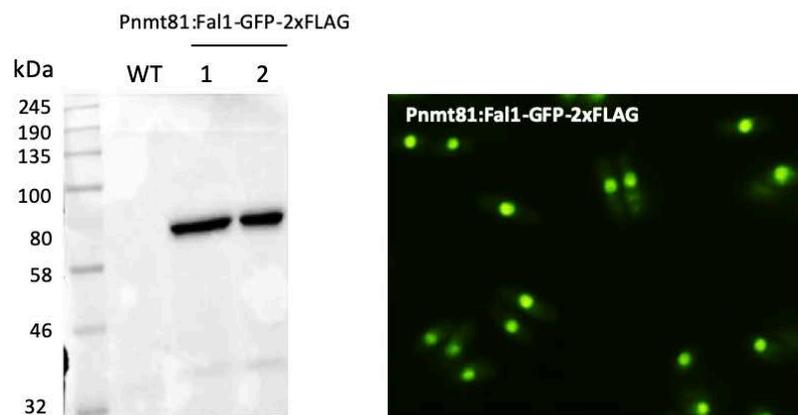
Figure 3.6 - *fal1* promoter swapping with the nmt81 promoter. (A) The physiological promoter of *fal1* is substituted with the Pnmt81. A DNA fragment containing the hygromycin resistant cassette upstream the Pnmt81, and sequence of homology with the 5'-UTR and the N-terminal region of *fal1*, is transformed in the Fal1-GFP-2xFLAG strain and integrated into the genome via homologous recombination. The integration of the fragment is verified by PCR on the genomic DNA of the select colonies: using primer Fal1-F1, as forward, and Fal1-CHK-rev as reverse, an expected band of approximately 8 kb is amplified in four colonies. The same primer pair amplifies a band of 2.3 from the WT genomic DNA. A schematic representation of the promoter swapping is reported in Appendix VIII.

(B) Western blot analysis and localisation analysis of Fal1 under the nmt81 promoter: WB using anti-flag antibody (M2, Sigma) reveals that the Pnmt81 is functional, able to express the Fal1-GFP-2xFLAG protein in two positive colonies. The expression of Fal1 under Pnmt81 does not affect the protein nuclear localisation.

Note that the WB is qualitative and aimed to confirm the presence of the FLAG-tag in the positive colonies; therefore, the membrane was not probed with a loading control antibody.

3.2.5 - Pnmt81 "shut-off" causes a depletion of Fal1 within 12 hours

To evaluate the temporal window required to deplete Fal1 upon repression of the nmt81 promoter, we performed a time course WB experiment; the Pnmt81:Fal1-GFP-2xFLAG strain was grown in EMM liquid media, supplemented with thiamine and aliquots of the cell culture were collected at 0, 6, and 12 hours. The protein levels of Fal1 were evaluated by probing with an anti-Flag antibody. We found that the protein levels of Fal1 were strongly reduced after 12 hours of thiamine treatment suggesting that the Pnmt81 promoter was responsive to thiamine repression (Figure 3.7A). We also monitored and compared the GFP signal of the Pnmt81:Fal1-GFP-2xFLAG strain to the signal of the parental strain (in which Fal1-GFP-2xFLAG is expressed by the physiological promoter) during the temporal window 0-12 hours. The two strains were grown in EMM with thiamine and a small aliquot of the cell culture was used to visualize the GFP fluorescence every three hours. Coherently with the WB analysis, we found that the GFP signal was almost undetectable after 12 hours of thiamine repression; as expected no effect was observed for GFP levels of the parental strain. Our result indicated that the shut-off system specifically depleted cells of Fal1 in response to thiamine addition (Figure 3.7B).

A

Pnmt81:Fal1-GFP-2xFLAG

| kDa | 0 | 6 | 12 | Time (hour) |

245
135
190
100
80

◄ Fal1-GFP-2xFLAG

58

46

58

◄ α-tubulin

46

B

| | Fal1-GFP-2xFLAG | Pnmt81:Fal1-GFP-2xFLAG |
| | EMM + Thiamine | EMM + Thiamine |

0 h

3 h

6 h

9 h

12 h
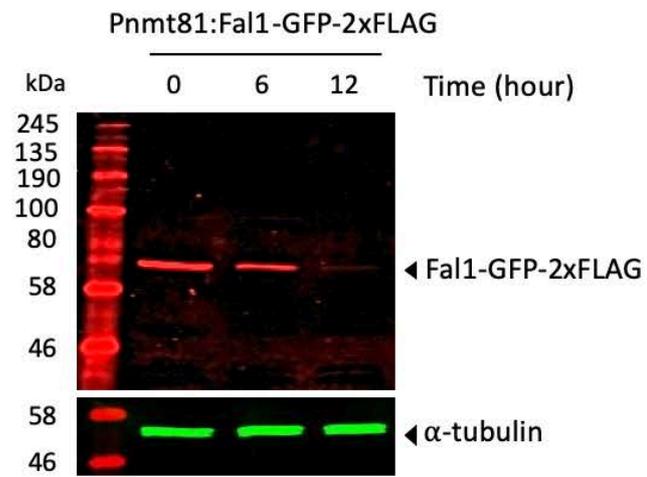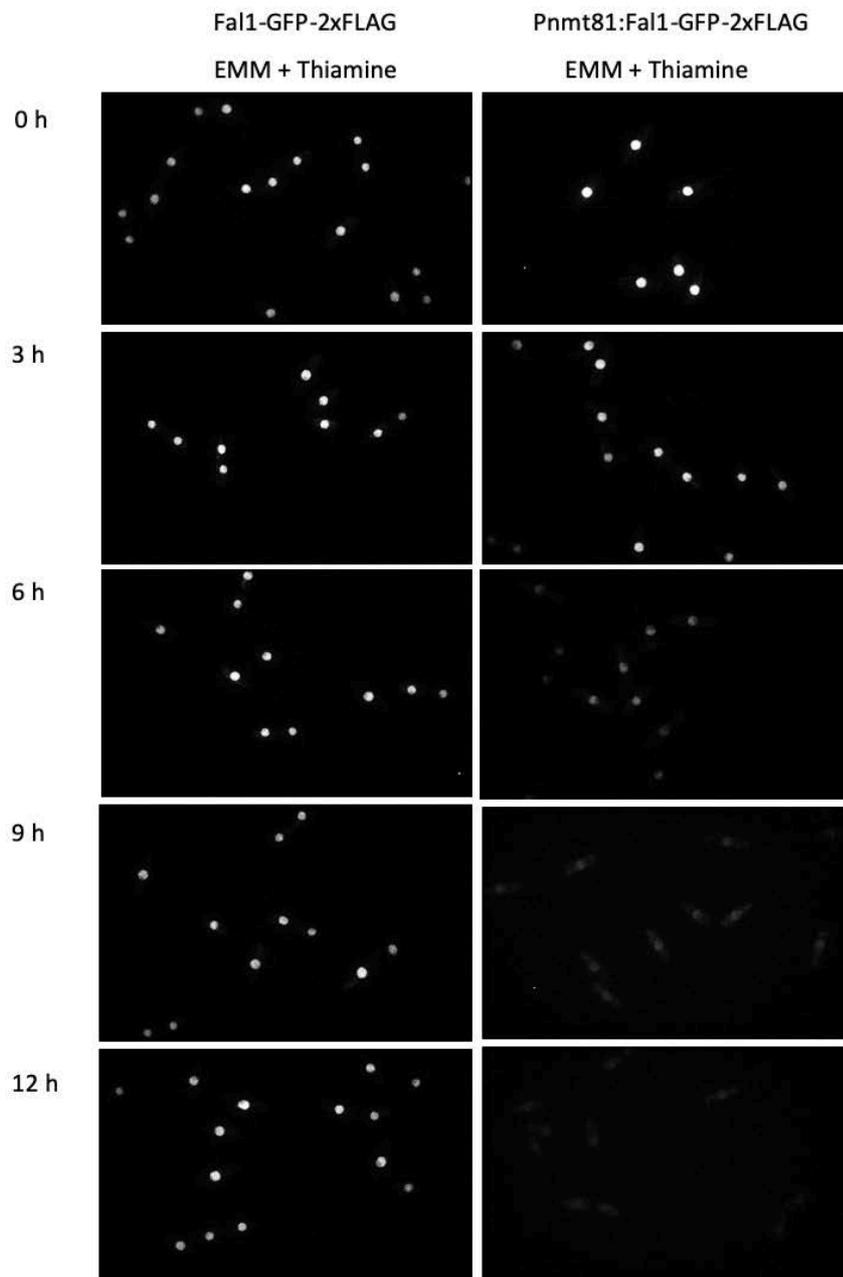
Figure 3.7 - Evaluation of the Fal1 protein levels upon thiamine repression. (A) WB analysis of the Pnmt81:Fal1-GFP-2xFLAG, grown in EMM liquid media supplemented with thiamine. Cells are collected at 0, 6, 12 hours; the protein levels of Fal1 are evaluated probing the membrane with anti-flag antibody. The antibody against α-tubulin is used as a loading control. (B) The GFP-signal relative to the Pnmt81:Fal1-GFP-2xFlag strain is compared to the signal of the parental strain (no repressible promoter). The two strains are grown in EMM liquid media supplemented with thiamine, and images of the GFP fluorescence are taken every three hours. The images are acquired using a 2D array scanning laser confocal microscope (Infinity 3, VisiTech).

Figure B was generated by Dr. Kayoko Tanaka

Since the deletion of Fal1 strongly affected cell viability, we predicted that the repression of the Pnmt81 promoter should cause a reduction of cell viability. Therefore, to validate the efficiency of the repressible system, we first performed a spot growth assay comparing the growth of the conditional mutant, Pnmt81:Fal1-GFP-2xFLAG, to that of the parental strain, Fal1-GFP-2xFLAG strain, which was previously shown to grow as WT (Figure 3.8A). Not surprisingly, we found that the addition of thiamine impaired cell growth for the shut-off strain, but did not affect the viability of the parental (Figure. 3.8A).

To evaluate the minimal temporal window required to induce growth defects, we performed a growth assay of the shut-off strain, in EMM liquid media. Cells were set at approximately OD_{600}=0.03/0.04 in liquid media, with or without thiamine, and the cell densities were counted every two hours using a hemocytometer. After eight hours, the cell cultures were diluted to fresh media to avoid saturation of the cell culture. It was observed that up to about 8 hours the two cultures behaved similarly (Figure 3.8B).

Notably, we found that after 12 hours from the beginning of the time course, cells cultured in presence of thiamine started to exhibit growth defects (Figure 3.8B). The time of appearance of the growth phenotype, at about 12 hours after addition of thiamine, coincided with almost complete disappearance of Fal1 protein, judged by WB and fluorescence microscopy (Figure 3.7A and Figure 3.7B).
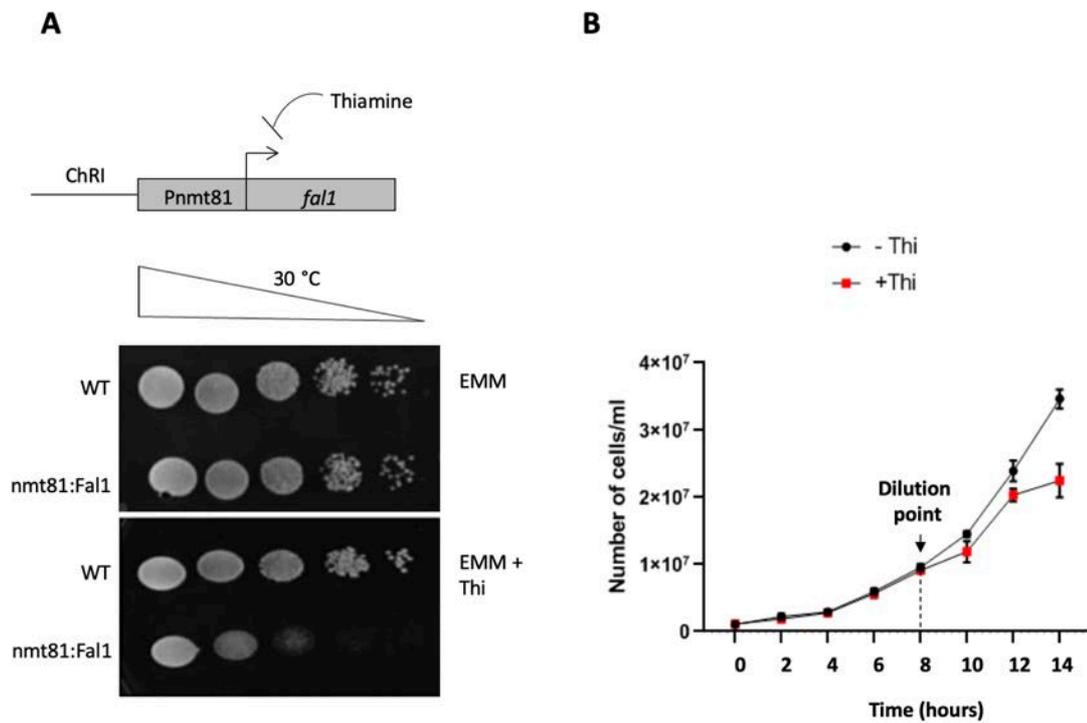
Figure 3.8 - Evaluation of the cell phenotype upon transcriptional repression of *fal1*.
(A) Spot growth assay comparing the growth of the strain in which *fal1* is under the control of the nmt81 promoter, to Fal1-GFP-2xFLAG, which we previously showed to grow as WT. Diluted spots contain approximately 10^5, 10^4, 10^3, 10^2, 10 cells. The addition of thiamine to the EMM agar plate affects cell growth of the shut-off strain while no effect is observed for the parental.

(B) Growth assay in EMM liquid media of the Pnmt81:Fal1-GFP-2xFLAG strain, cultured with and without thiamine; cells are synchronized at approximately OD=0.03/0.04 and counted every two hours using a hemocytometer. After 8 hours, the cell cultures are diluted (dilution point; the number of cells counted after the dilution point were multiplied by the dilution factor, to keep continuity in the graph). Cell growth is monitored for 14 hours; growth defects emerge after 12 hours from the beginning of the time course. Error bars represent mean ± SEM (n=2).

3.2.6 - Engineering of the AID system for the rapid depletion of Fal1

The "shut-off" method described in paragraph 3.2.5, represents a powerful system to conditionally regulate the protein levels of Fal1; however, it took 12 hours to deplete the target protein and to observe a cellular phenotype. Considering that in fission yeast the cell cycle is completed within 2-3 hours, a considerable disadvantage of this tool is represented by the inability to discriminate primary effects from the secondary ones, which could emerge as a consequence of cell adaptation.

To overcome this issue, we have applied the AID system to rapidly reduce Fal1 protein levels; the AID system was engineered on top of the "shut-off strain, thus allowing the promoter repression and the protein degradation by addition of thiamine and auxin respectively.

The plasmid containing the AID epitope (pAW8ENdeI-cHAIAA7-spo5DSR) was a gift from Professor Tony Carr, from the University of Sussex, and the map is reported in Appendix IV. Watson and colleagues reported that the AID system poorly functioned in *S. pombe,* due to the inability to completely deplete the target protein (Watson et al., 2013). To improve the system, they introduced a DSR element (Determinant of Selective Removal), a sequence naturally occurring in the 3'-UTR of the meiotic mRNAs of *S. pombe*. The DSR allows the selective elimination of meiotic mRNAs during the mitotic growth (Harigaya et al., 2006). During the vegetative growth, the mRNAs harboring the DSR are recognized and bound by the Meiotic mRNA Interception 1 protein (Mmi1), which interacts with components of the nuclear exosome causing the degradation of the transcripts (Watson et al., 2013, Yamashita et al., 2012). However, during meiosis, Mmi1 interacts with meiRNA, a lncRNA that retains Mmi1 on the sme2 locus, thus permitting the translation of the meiotic mRNAs (Ding et al., 2013) (Figure 3.9A).

Based on previous observations from Watson and colleagues, we decided to introduce the DSR in our AID system, to cause a reduction in the Fal1 mRNA levels.

Therefore, we substituted the 3-UTR of Fal1, with the 3-UTR of the meiotic gene *spo5* (SPBC29A10.02). A schematic representation of the engineering of the AID system is reported in Figure 3.9B.

The integration in the genomic locus was verified by PCR (Figure 3.9C). The amplified fragment was sent for sequencing to verify that no point mutations occurred in the junction between the GFP and the AID epitopes; 100% of homology with the expected sequence was obtained for a region of almost 700 bp starting from the GFP-seq2 primer, which annealed in the GFP sequence epitope (Appendix VI). Then, the AID tagged strain was crossed with a TIR1 expressing strain, bought from the National BioResource Project-Yeast (details in Appendix II), to generate the full degron tool.

A

Mitotic Growth

RNA exosome

Mmi1

DSR

Meiotic Transcript

AAAA

Meiosis

Mmi1

meiRNA

sme2

Meiotic Transcript

DSR

AAAA

B

| GFP | AID | DSR | ClonNAT^R | fal1-3'UTR |

| Pnmt81 | fal1 | GFP | 2xFLAG | Kan^R | fal1-3'UTR |

| Pnmt81 | fal1 | GFP | AID | DSR | clonNAT^R | fal1-3'UTR |

C

Fal1-GFP-F1

Fal1-CHK-RV

| fal1 | GFP | AID | DSR | ClonNAT^R | fal1-3'UTR |

Pnmt81:Fal1-GFP-AID:DSR

kb   1   2   3   4   ◄ colonies

4
3
2
1.5
1

0.5

96
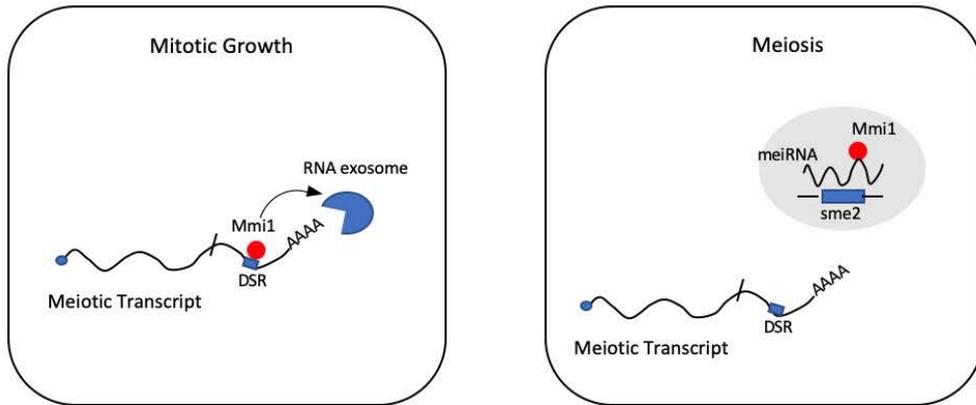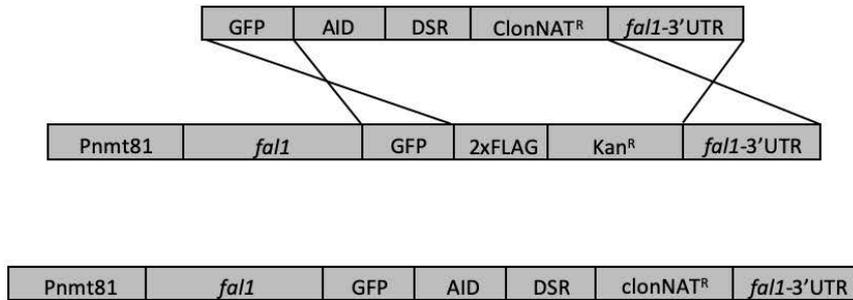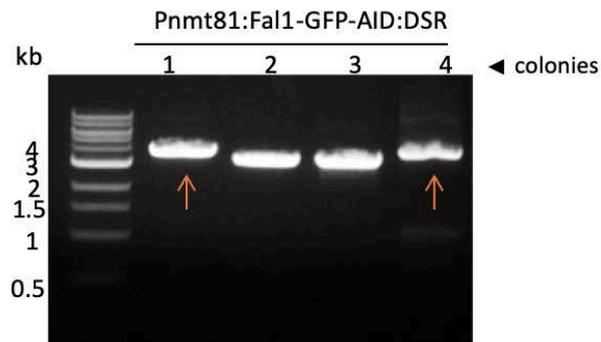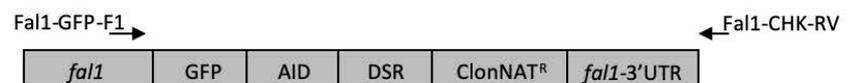
Figure 3.9 - Engineering of the AID system for the rapid depletion of Fal1. (A) Role of the DSR element in the selective elimination of the meiotic transcripts. During mitotic growth, the meiotic transcripts containing a DSR element, are recognized by Mmi1 and degraded by the nuclear exosome. During meiosis, Mmi1 is sequestered at the sme2 locus, by interaction with a lncRNA. (B) Schematic representation of the method used to tag Fal1 with the AID epitope: a DNA fragment containing the GFP tag, the AID epitope (IAA17), the DSR element, the marker conferring resistance for the antibiotic Nourseothricin sulfate (ClonNAT), and a sequence of homology with the 3'-UTR of Fal1 is generated using overlapping PCRs, transformed in the nmt81:Fal1 strain and integrated in the genomic locus via homologous recombination. A detailed schematic representation of the gene tagging with the AID, is reported in Appendix VIII. (C) Integration in the genomic locus is verified by PCR on the genomic DNA of four selected colonies: using primer Fal1-GFP-F1 as forward, which anneals in the Fal1 gene, and Fal1-CHK-RV as reverse, which is designed outside the region of recombination, an expected band of 3.8 kb is amplified in two colonies, indicated by arrows.

3.2.7 - Characterisation of the AID system

To evaluate the protein levels of Fal1 upon addition of thiamine and auxin, we first monitored the GFP signal of the Pnmt81:Fal1-AID:DSR (strain in which *fal1* was under the control of the Pnmt81 promoter and contains the AID epitope and the DSR element for transcription down-regulation) comparing to the Fal1 parental strain. The two strains were inoculated in EMM with thiamine and 1-Naphthaleneacetic acid (NAA), a synthetic analog of auxin, and images of the GFP fluorescence were taken at time 0-90 minutes using confocal microscopy. We found that the GFP nuclear signal of the Pnmt81: Fal1-AID strain reduced after 90 minutes from thiamine and auxin treatment, while as expected, the parental strain showed no effect (Figure.3.10A).

This result suggested that the AID system was working upon addition of thiamine and auxin; however, we also noted that the intensity of the GFP signal for the Pnmt81:Fal1-AID strain was substantially reduced even before the addition of thiamine and NAA, compared to the parental (Figure.3.10A).

This effect was probably due to:

- destabilization of the protein caused by the AID tag, as reported by Eaton and colleagues (Eaton et al., 2018);

- a protein degradation caused by a basal level of interaction between Fal1-AID and the TIR1 (Mendoza-Ochoa et al., 2019, Morawska and Ulrich, 2013, Natsume et al., 2016);

- or a reduction in the mRNA levels due to the DSR element (Watson et al., 2013).

To confirm the situation by WB analysis, the Pnmt8:Fal1-AID:DSR strain was grown in EMM liquid media, supplemented with thiamine and NAA; cells were collected at times

0, 30, 60, 90 and 120 minutes. The protein levels of Fal1 were evaluated by WB, probing with an anti-GFP antibody (AHP975, Bio-rad). The presence of the AID tag increased the molecular weight of the Fal1-AID strain so that a band just below 100 kDa was expected (Figure 3.10B).

This experiment confirmed that the intensity of the band of interest was much weaker when compared to the parental strain at time zero (Figure 3.10B). Actually, almost double of the protein extract typically loaded (30 µL) had to be used to detect the Fal1-GFP-AID protein.

We also found that the level of Fal1-AID-GFP was almost undetectable after 60 minutes from the addition of thiamine and NAA (Figure 3.10B).

A

Fal1-GFP-2xFLAG   Clone 1   Clone 2

T=0

T=90 min
+NAA
+ Thi

T=180 min
+NAA
+Thi

B

Pnmt81: Fal1-GFP-AID

Time (min)	0	30	60	90	120	wt	Fal1-GFP-2xFLAG	kDa

Fal1-GFP-AID ►

α-tubulin ►

kDa
245
190
135
100
80
58
58
46

100
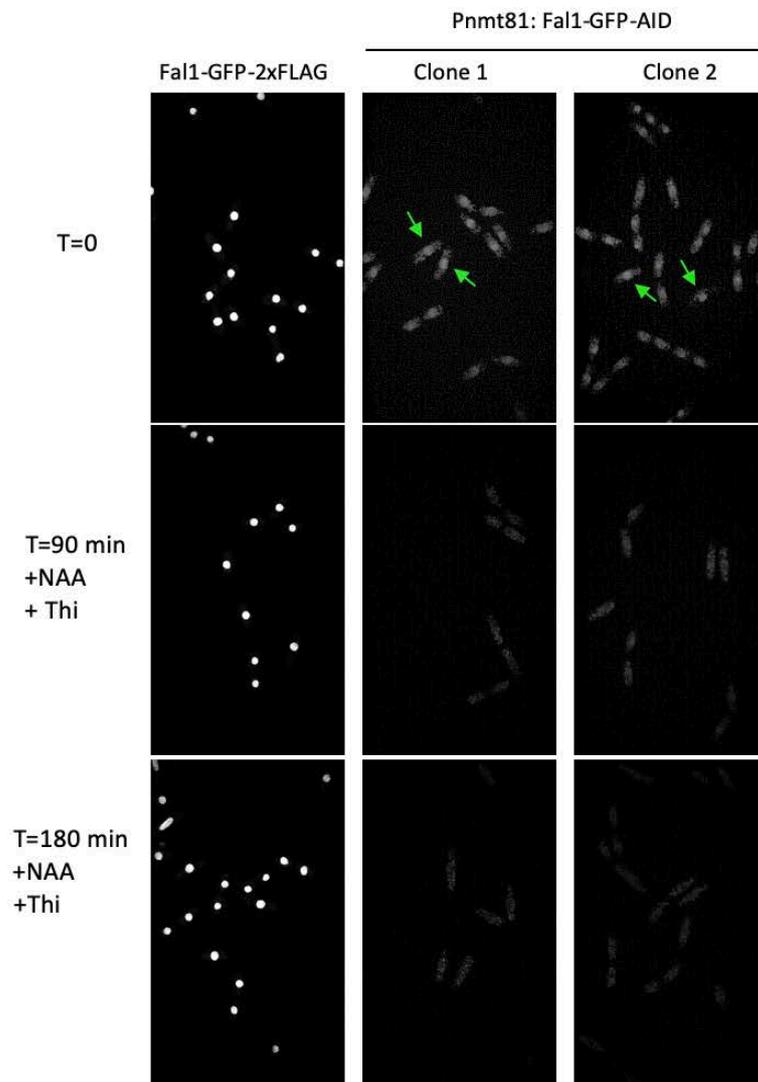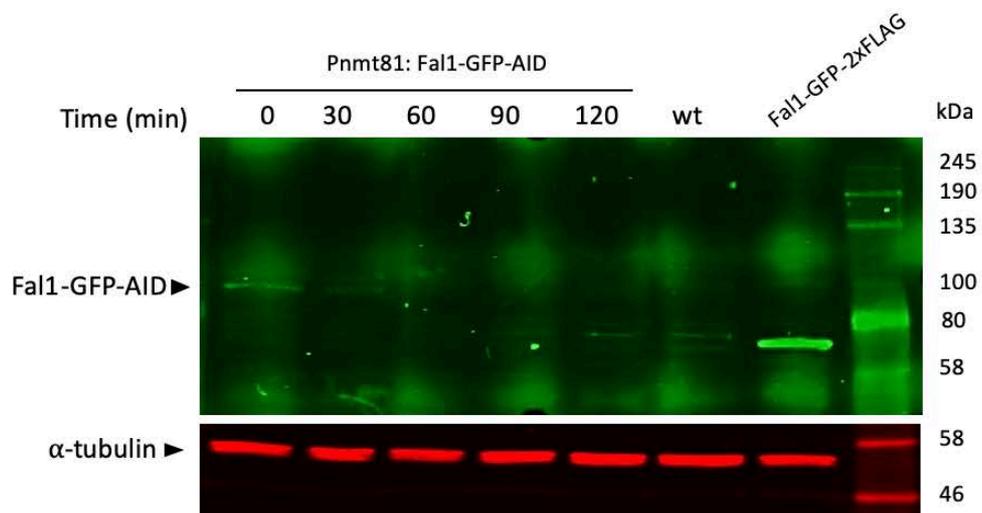
Figure 3.10 - Evaluation of the protein levels of Fal1 upon addition of thiamine and auxin. (A) The GFP signal of two degron clones is compared to the signal of the Fal1 parental strain, upon thiamine and NAA (a synthetic analog of auxin) treatment. Images of the GFP fluorescence are taken at time 0-90 minutes using a 2D array scanning laser confocal microscope (Infinity 3, VisiTech). Example of nuclear GFP signals are indicated by arrow. Compared to the Fal1-GFP-2xFLAG strain, Pnmt81:Fal1-GFP-AID cells produce much weaker GFP signals.

(B)Time-course experiments to detect the protein levels of the degron strain by WB; cells are collected at 0, 30, 60, 90, 120 minutes upon addition of thiamine and NAA. Protein extracts from WT and Fal1-GFP-2xFLAG strains are used as negative and positive control respectively. The membrane is incubated with anti-GFP antibody. Antibody against α-tubulin is used as a loading control.

Figure in A was generated by Dr. Kayoko Tanaka

To test the efficiency of the AID system from a phenotypical point of view, and to verify whether the reduced protein levels observed in Figure 3.10A and Figure 3.10B affected cell viability, we performed an extensive analysis of the cell growth phenotype under different conditions.

Firstly, we performed spot growth assay comparing the TIR1 strain used to generate the degron system, to the Fal1-GFP-2xFLAG strain, which we had previously shown to grow as WT. We could confirm that the presence of the plant element TIR1 in the yeast background did not affect cell viability (Figure 3.11A).

Then, we compared the growth of the degron strain (Pnmt81:Fal1-GFP-AID, which also contains the DSR element) which lacked TIR1, with the parental strain Fal1-GFP-2xFLAG. Interestingly, we found that the two strains grew identically, indicating that the AID epitope and the DSR element did not affect cell viability (Figure 3.11B).

Lastly, we compared the growth of the degron strain containing TIR1 (Pnmt81:Fal1-GFP-AID: DSR + TIR1), to the Fal1-GFP-2xFLAg strain. Cells were spotted on EMM, EMM supplemented with thiamine, EMM supplemented with NAA, and EMM supplement with thiamine and NAA; surprisingly we found that the presence of TIR1 and the AID-tagged protein in the same background strongly affected cell viability even in absence of thiamine and NAA. Additionally, the presence of thiamine and NAA to the media only slightly affected the cell growth, which was already compromised on EMM only (Figure 3.11C). Overall, the results presented in this paragraph indicate that despite the quick decrease in the protein levels of Fal1 observed upon treatment with NAA and thiamine, our application of the "OFF-AID system" strongly affects cell viability, making the strain unsuitable for further experiments.

A

Fal1-GFP-2xFLAG

TIR1

30°C

EMM

B

Fal1-GFP-2xFLAG

Pnmt81: Fal1-GFP-AID:DSR
(not containing TIR1)

30°C

EMM

C

Fal1-GFP-2xFLAG

Pnmt81: Fal1-GFP-AID:DSR +
TIR1

Fal1-GFP-2xFLAG

Pnmt81: Fal1-GFP-AID:DSR +
TIR1

Fal1-GFP-2xFLAG

Pnmt81: Fal1-GFP-AID:DSR +
TIR1

Fal1-GFP-2xFLAG

Pnmt81: Fal1-GFP-AID:DSR +
TIR1

30°C

EMM

EMM+ Thia

EMM+ NAA

EMM+ Thi+ NAA
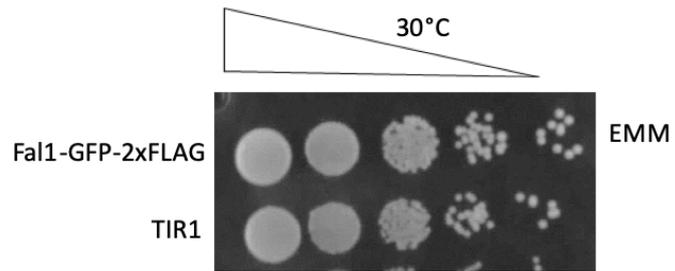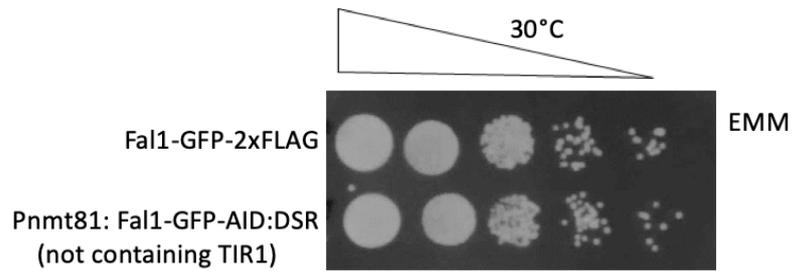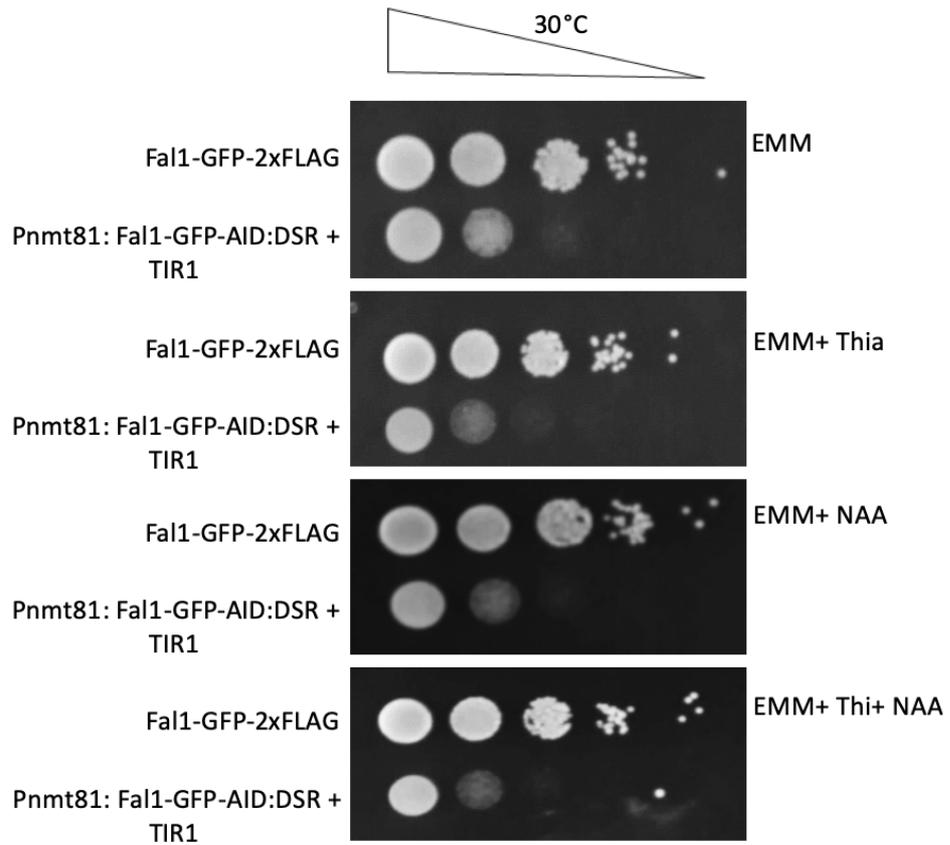
Figure 3.11 - Characterization of the growth phenotype of the AID system. (A) Evaluation of the growth of the TIR strain on EMM agar plate; the growth of the TIR1 and the Fal1-GFP-2xFlag (which grows as WT), is compared by spot growth assay. Diluted spots contain approximately 10^5, 10^4, 10^3, 10^2, 10 cells.

(B) Evaluation of the growth of the Fal1 degron strain prior conjugation with the TIR1 element: Diluted spots contain approximately 10^5, 10^4, 10^3, 10^2, 10 cells.

(C) Evaluation of the growth of the Fal1 degron containing the TIR1 element: cells are spotted on EMM, EMM supplemented with thiamine, EMM supplemented with auxin, and EMM supplemented with both the drugs. Diluted spots contain approximately 10^5, 10^4, 10^3, 10^2, 10 cells.

3.2.8 – Generation of a tool to be used to assess the role of Fal1 on the RNA Pol II elongating complex.

The spot growth assay from Figure 3.11 reveals that growth defects occurred only when TIR1 was introduced in the Pnmt81:Fal1-AID:DSR tagged strain (Figure 3.11C), indicating that the DSR element *per se* was not responsible for the phenotype (Figure 3.11B).

Therefore, we predicted that we still could use the Pnmt81-Fal1-AID strain, not containing TIR1, where the DSR element would reduce the time required for the depletion of Fal1.

Since the original aim of this project was to isolate and sequence the nascent-RNA upon rapid depletion of Fal1, we decided to use the Pnmt81:Fal1-AID:DSR strain, not containing TIR1, for this purpose. The purification of the nascent RNA requires an immunoprecipitation (IP) of RNA Pol II elongating complex, consisting of DNA, RNA Pol II, and the nascent RNA still attached to it (Churchman and Weissman, 2011). To accomplish this goal, with technical expertise of Dr. Kayoko Tanaka, the Pnmt81:Fal1-AID:DSR strain, not containing TIR1, was crossed with the strain in which the third subunit of RNA Pol II, Rpb3, is tagged at the N-terminal with a single FLAG-tag, that could be used to purify the elongating complex (Figure 3.12A). After crossing, the resultant strains of at least 10 tetrads were analyzed by colony PCR, to check the presence of the FLAG-Rpb3. By using primer pair Rpb3-FW (annealing to the 5'-UTR of Rpb3) and Rpb3-RV (annealing to the Rpb3 coding sequence), the presence of the FLAG epitope, could be detected and we successfully isolated several clones harbouring both Pnmt81-Fa1-GFP-AID:DSR and FLAG-Rpb3 (Figure 3.12B). To check the time frame required to deplete Fal1, we collected cells at 0, 3, 6, 9 and 12 hours after addition of thiamine and we performed WB using anti-GFP antibody. Interestingly we found that the level of Fal1-AID was lower compared to the parental strain expressing Fal1-GFP-2xFLAG only (Figure 3.12C). This was probably due to the presence of the DSR element (Watson et al., 2013).

However, the level was much higher compared to Fal1-AID containing TIR1 (Figure 3.12C and Figure 3.10B), confirming again that the main issue relative to the AID system derived from TIR1.

Moreover, we noted that in only 3 hours of thiamine repression, the levels of Fal1 drastically reduced and in 6 hours, the signal was almost undetectable, in agreement with our initial prediction (Figure 3.12C).

To check whether a growth phenotype was associated with the depletion of Fal1, a spot growth assay was performed on EMM plates in presence and absence of thiamine, comparing the growth of the Fal1-AID;FLAG-Rpb3 strain, with the FLAG-Rpb3 and the Fal1-GFP-2xFLAG (WT) strains; interestingly we found that the Fal1-AID after crossing with the FLAG-Rpb3, was growing slightly less than the WT on EMM, but a severe growth defect was observed in presence of thiamine, indicating that the tool could be exploited for further experiments (Figure 3.12D)

A

B

Rpb3-FW →   ← Rpb3-RV

| FLAG | Rpb3 |

C
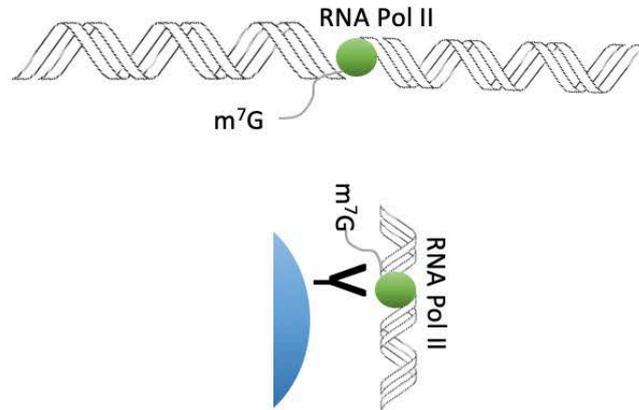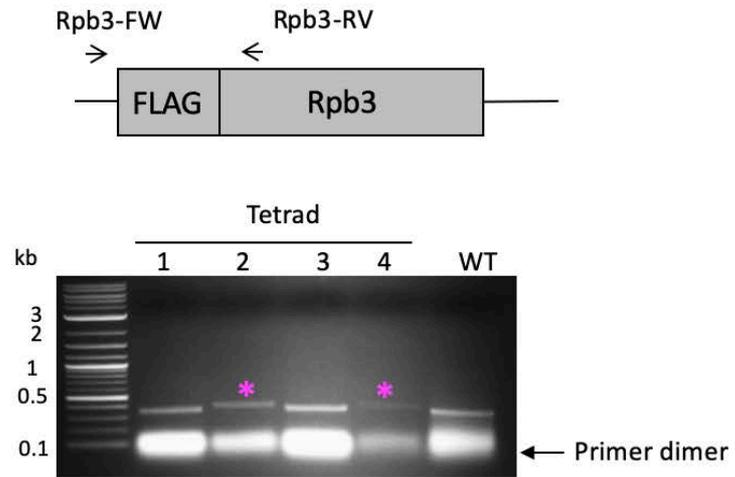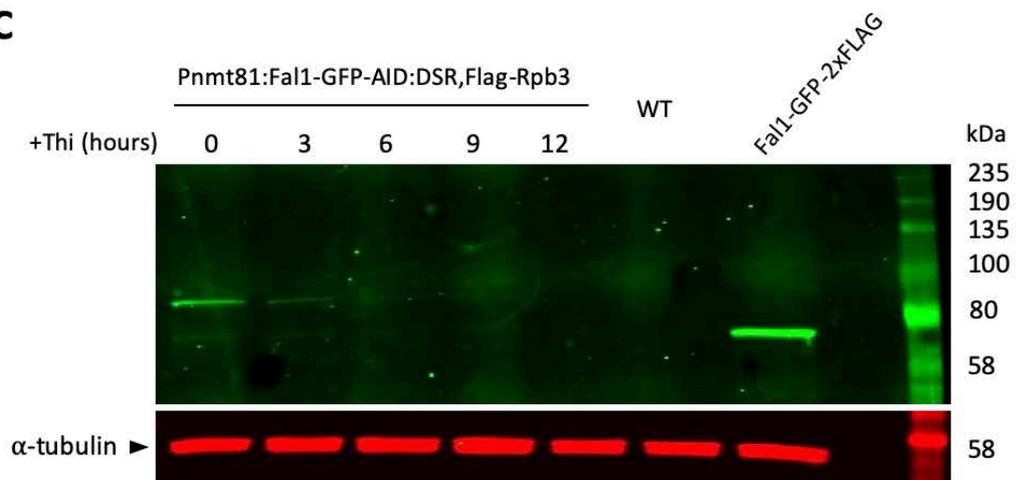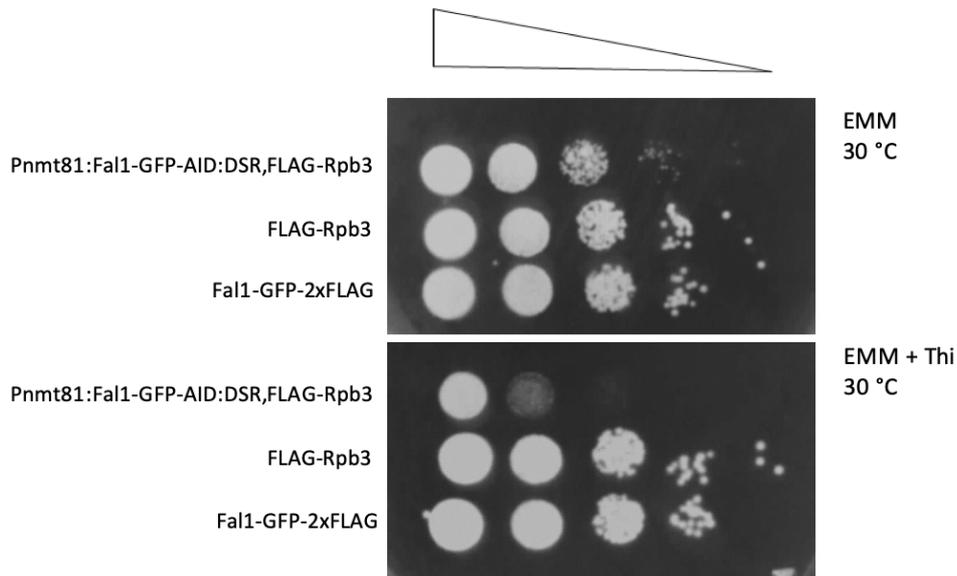
Figure 3.12 - Creation of a tool for isolation of nascent RNA upon depletion of Fal1.
(A) Schematic representation of the nascent RNA purification: RNA Pol II elongating complex is purified via immunoprecipitation (IP), by exploiting an antibody which binds a tagged subunit of RNA Pol II. (B) Generation of Pnmt81-Fal1-GFP-AID:DSR strain harbouring FLAG-Rpb3 but lacking TIR. The successful introduction of the FLAG-Rbp3 tag results in an increase in the size of the PCR products. In this instance, colony 2 and 4 are positive for the FLAG-Rbp3 tagging (indicated by a Magenta asterisk)

(C) The temporal window to deplete Fal1 in the newly generated strain is assessed by WB: cells are collected at time 0, 3, 6, 9, 12 hours and Fal1 is detected using anti-GFP antibody. Protein extracts from WT and Fal1-GFP-2xFLAG are used as negative and positive control respectively. The membrane is also probed with anti-α-tubulin, as a loading control.

(D) Spot growth assay comparing the Fal1-AID:Flag-Rpb3 (not containing TIR1) strain to the FLAG-Rpb3 and Fal1-GFP-2xFLAG strains on EMM plates with and without thiamine.

3.3 - Discussion

In the first part of this chapter, we have shown a characterization of Fal1 in terms of sequence conservation between divergent organisms, cellular phenotype derived by its deletion, and localisation. Our data indicate that Fal1 shares a high percentage of identity with its orthologues. Moreover, the domains for the binding/hydrolysis of the ATP and the RNA binding appear highly conserved in Fal1, thus suggesting an evolutionarily conserved function, which is kept in fission yeast (Figure 3.2).

Using a gene knock-out approach and subsequent tetrad dissection we have shown that the *fal1* gene is not essential but it is required for the normal cell growth of *S. pombe;* severe growth defects emerge upon *fal1* deletion, implying its role in an essential biological function (Figure 3.3B and Figure 3.3C).

Another important feature which appears conserved is that Fal1/eIF4AIII is found predominantly in the nucleus as divergent organisms such as flies, mammalian cells, plants and yeast (Ferraiuolo et al., 2004, Koroleva et al., 2009, Kressler et al., 1997, Marayati et al., 2016, Palacios et al., 2004, Choudhury et al., 2016). By tagging endogenous Fal1 with the GFP epitope, we observed that also in *S. pombe*, Fal1 is distributed in the nucleus, where it remains constantly associated during nuclear division, suggesting that the function of Fal1 is mainly nuclear (Figure 3.5A and Figure 3.5B).

The fact that *fal1* deletion affects cell viability makes the knock-out strain inappropriate for any experiment relative to gene expression, since it does not allow to discriminate direct and indirect effects that can arise as a consequence of a cell adaptation. To overcome this issue, we have generated an OFF-AID system, in which Fal1 is regulated by transcriptional repression (OFF), and protein turn-over (AID). Additionally, the 3'-UTR of the gene has been

replaced by an mRNA destabilizing element (DSR), which has been shown to improve the efficiency of the AID system in *S. pombe* (Watson et al., 2013).

The shut-off system (OFF) which we first created, substituting the physiological promoter with a repressible one, the Pnmt81, appears to function both at phenotypical and protein levels: upon thiamine repression, the Fal1 protein levels decrease within 12 hours from the beginning of the treatment (Figure 3.7A; Figure 3.7B). Accompanying, growth defects also emerge around 12 hours (Figure 3.8B). At a glance, therefore, this shut-off tool represents a huge step forward to study the function of Fal1. However, cell adaptations could potentially still occur during the 12 hours time frame required for Fal1 depletion. Therefore, to overcome this issue, the shut-off strain has been further improved by the addition of the AID and the DSR elements. By adding these two components to the previous tool, the depletion of Fal1 is accomplished within almost 60 minutes from thiamine and auxin treatment (Figure 3.10B). However, we observed a substantial reduction of the protein compared to the parental strain, which is independent of auxin incubation (Figure 3.10A; Figure 3.10B; Figure 3.11C). The reduced levels of Fal1-AID affect cell viability, which is compromised before the addition of auxin and thiamine (Figure 3.11C).

To understand the reason behind this phenotype we have analyzed the cellular phenotypes of the strains used to generate the degron system; particularly, the strain containing the AID-DSR element, but not TIR1, can grow on EMM as the parental, without showing obvious growth defect (Figure 3.11B). Moreover, the strain containing only the TIR1 element does not exhibit phenotypical defects when compared to the WT (Figure 3.11A). This suggests that growth defects arise when TIR1 and the AID-tagged protein coexist in the same background. One possibility to explain this phenomenon, is that Fal1 is degraded before the addition of auxin, because of an intrinsic basal affinity between TIR1 and AID, characterized by the

entrance of the IAA17 epitope in the pocket of TIR1 even in absence of the auxin. This explanation appears consistent with numerous recent observations (Mendoza-Ochoa et al., 2019, Natsume et al., 2016, Watson et al., 2013, Yesbolatova et al., 2019).

Overall our work has revealed that a key limitation for the application of the degron system in fission yeast is represented by the constitutive expression of the TIR1 element.

To overcome this problem, Yesbolatova and colleagues have recently developed a TIR1 antagonist known as auxinole, which binds TIR1 preventing its association with the AID-fused protein; then when the degradation of the protein of interest is required, the auxinole is washed off and cells are resuspended into a fresh media containing auxin (Yesbolatova et al., 2019). The auxinole, which is commercially available, will be tested in the future to verify whether it will be able to abolish the TIR1 leakage effect, which prevents us to use the degron tool for further experiments.

Another approach to test in the future is to change the TIR1 promoter with the *urg1* promoter, which has been shown to be activated in almost 30 minutes upon addition of uracil; the system in this case will be characterized by a "block and release" of TIR1 expression and subsequent degradation of Fal1 by addition of auxin (Watt et al., 2008).

Since we had observed that the growth defects occurred only when the TIR1 element was introduced in the AID tagged strain (Figure 3.11C), we decided to cross the Pnmt81:Fal1-AID-DSR strain, not containing TIR1, with a strain in which the third subunit of the RNA Pol II was FLAG-tagged at the N-terminal, to generate a tool for isolation of nascent-mRNA upon depletion of Fal1 (Churchman and Weissman, 2011). In fact, we predicted that the presence of the DSR, would have reduced the mRNA levels of Fal1, thus causing a faster depletion of Fal1, but without affecting cell viability. As we predicted, the protein levels of Fal1 drastically reduced in almost three hours upon addition of thiamine and the signal is almost

undetectable in six hours. The protein levels of Fal1 at time zero appear lower compared to the parental strain, probably due to the presence of the DSR (Figure 3.12C). Interestingly we found that the newly generated strain harbouring both Fal1-GFP-AID:DSR and FLAG-Rbp3, is sensitive to thiamine, but it shows a minimal variation in cell growth, when compared to the WT (Figure 3.12D).

To give an explanation to this phenomenon, we have hypothesized that the presence of a double tag for Fal1 and Rbp3, could negatively affect a putative interaction between the two factors, which was previously speculated for other organisms, thus causing a minimal reduction in cell viability (Akhtar et al., 2019, Singh et al., 2019, Wang et al., 2014).

This new tool appears extremely promising since it allows to obtain fast depletion of Fal1 in the background of a tagged Pol II subunit. The double tag can be exploited for a co-immunoprecipitation (co-IP) experiment, to evaluate whether the two proteins interact and whether the interaction is RNA sensitive, as observed in *D. melanogaster* (Singh et al., 2019).

In case of interaction, since previous studies suggested a role for eIF4AIII in the regulation of Pol II elongation and pausing, we will examine the changes in Pol II loading through ChIP-seq analysis, before and after rapid degradation of Fal1 (Akhtar et al., 2019, Wang et al., 2014). The change in Pol II density after fast depletion of Fal1, will be further investigated using a high-resolution approach, which permits to isolate the nascent RNA via isolation of the ternary complex formed by DNA-RNA-Pol II, and to capture the Pol II position at single-nucleotide resolution (Churchman and Weissman, 2011). Because we will assess nascent-RNA upon rapid depletion of eIF4AIII, we should be able to discriminate whether the effect on Pol II elongation is direct or indirect. Several studies also report a role of eIF4AIII in the regulation of splicing (Akhtar et al., 2019, Ashton-Beaucage and Therrien, 2011, Ashton-Beaucage et al., 2010, Marayati et al., 2016, Wang et al., 2014). However, we suspect that this

effect can be indirect, mainly due to the sick phenotype of the cells observed upon eIF4AIII depletion. To clarify this situation, we will purify the newly transcribed RNA with the 4-thiouridine labeling to examine the role of Fal1 on mRNA processing at high-resolution (Mata and Wise, 2017, Rutkowski and Dölken, 2017).

Chapter 4: Fal1 associates with transcriptionally active gene loci regardless of the presence of an intron

4.1 - Introduction

For several years it has been reported that splicing alters the composition of the mRNPs, primarily, as discussed in Chapter 1, by deposition of the EJC, typically at 20-24 nucleotides upstream of each exon-exon junction, in a sequence-independent fashion (Le Hir et al., 2001b).

However a new observation suggests that the EJC factors might associate co-transcriptionally also on mRNAs derived from intron-less genes (Akhtar et al., 2019, Choudhury et al., 2016, Viphakone et al., 2019)

The model of deposition of the EJC has been called into question for the first time by Choudhury and colleagues, who found that eIF4AIII, MAGO and Y14 are associated on the polytene chromosomes of *D. melanogaster*, both on intron-containing as well as intron-less genes. Furthermore, the association was dependent on nascent transcripts, since the signals were RNase sensitive (Choudhury et al., 2016). This breakthrough has been recently confirmed in mammalian cells and *D. melanogaster*, by the use of high-throughput approaches (Akhtar et al., 2019, Viphakone et al., 2019).

To date, in light of these new pieces of evidence, the mechanisms underlying EJC deposition appear still controversial since in stark contrast to the traditional model that the EJC is deposited only on spliced mRNA (Le Hir et al., 2000, Le Hir et al., 2016).

This chapter aims to clarify this by using ChIP, whether the deposition of the core EJC occurs regardless the presence of an intron in *S. pombe,* which appears a suitable model organism considering that more than 40% of the genes contain introns.

4.2 - Results

4.2.1 - Standardization of ChIP in *S. pombe*

Chromatin immunoprecipitation (ChIP) is a powerful molecular biology technique used to evaluate whether a protein interacts with DNA, either directly or through association with the nascent RNA (Abruzzi et al., 2004, Carey et al., 2009).

To use the ChIP assay in fission yeast, I have optimized all aspects of the technique, starting from the sonication. The quality and reproducibility of the ChIP assay depends on the size of the sonicated DNA (Carey et al., 2009, Pillai and Chellappan, 2009). I have optimised the sonication procedure to generate DNA fragments which mainly accumulate between 800 to 200 bp (Figure 4.0A and Materials and Methods).

The IP step was initially optimized using a strain in which the third subunit of the RNA Pol II, Rpb3, was tagged at the N-terminal with a single FLAG-tag, confirmed by PCR on genomic DNA (Appendix VI). The occupancy of Rpb3 was evaluated by qPCR, on two specific target genes; the intron-less gene *fba1* (1077 nt) encodes for fructose-bisphosphate aldolase, a highly transcribed gene that generates an average of 520 molecules of RNA per cell during the vegetative growth (Marguerat et al., 2012). The intron-containing gene, *gpd3,* encodes for glyceraldehyde-3-phosphate dehydrogenase, which generates an average of 110 molecules of RNA per cell during the vegetative growth phase (Marguerat et al., 2012). For these genes,

several primer pairs were designed to evaluate the level of enrichment and the pattern of Rpb3 along the coding sequence (Figure 4.0B). The occupancy of Rpb3 along the CDS of these genes was normalized to the input and expressed as fold enrichment relative to a negative locus (intergenic), as described in Material and Methods (2.2.9.7)

I found that on *fba1,* Rpb3 showed an enrichment of almost 20-fold relative to the intergenic; the pattern of Rpb3 along the CDS of *fba1*, appeared to be even along the gene (Figure 4.0B) and similar to that reported by Larochelle and colleagues*,* suggesting that the ChIP assay was functioning appropriately (Larochelle et al., 2018).

On *gpd3*, Rpb3 showed enrichment of almost 15-fold relative to the intergenic, in the middle of the CDS, therefore slightly lower when compared to the enrichment observed on *fba1,* which might reflect the difference in terms of gene expression between the two genes. Interestingly, on the first part of *gpd3*, the enrichment of Rpb3 was particularly modest, varying from 1.02 to 1.43-fold; from the melting curve analysis, no apparent issues were revealed, therefore we hypothesized that the poor signal could be due to the conformation of the chromatin in that particular region which might reduce the accessibility of the FLAG-antibody (Figure 4.0B).

To further evaluate the robustness of our ChIP protocol, we also performed the ChIP assay using an anti-phospho-Ser2 antibody (ab5095), which binds the phosphorylated form of the Ser2 of the CTD. As described in Chapter 1, the phosphorylation level of Ser2 is low at the beginning of the CDS and it increases constantly throughout the gene body (Buratowski, 2003, Buratowski, 2009).

I found that for *fba1,* the phospho-Ser2 signal was modest at the beginning of the gene (20-fold enrichment relative to the intergenic) and firmly increased in the downstream regions (90, and 120-fold enrichment relative to the intergenic) (Figure 4.0C). Notably, the pattern of

phospho-Ser2 on *fba1* was in line with the profile reported by Larochelle and colleagues for the same gene in *S. pombe* (Larochelle et al., 2018).

Additionally, I also tested the phospho-Ser2 signal on a gene expressed at low levels during the vegetative growth, *arg11* (2658 nt), which encodes for the N-acetyl-gamma-glutamyl-phosphate reductase. I confirmed not only the expected pattern of phospho-Ser2 along the coding sequence, but I also noted a correlation between the occupancy of the RNA Pol II and the level of expression; differently from *fba1*, whose level of expression is 520 molecules of RNA per cell during the the vegetative growth, the RNA levels for *arg1*1 is only 8.5, and the subsequent phospho-Ser2 signal varied from 2-fold, at the beginning of the gene, to 5-fold at the end of the CDS. Lastly, on *gpd3,* the phospho-Ser2 signal was varying from 2.44 to 30-fold enrichment relative to the intergenic region (Figure 4.0C). Due to the presence of a small enrichment detected in the first region of *gpd3,* I decided to exclude it from further analysis. Overall, considering the expected patterns of Rpb3 and phospho-Ser2*,* I concluded that the ChIP protocol was working in my hands and could be used to further examine whether Fal1 and other EJC proteins are associated with gene loci.

A
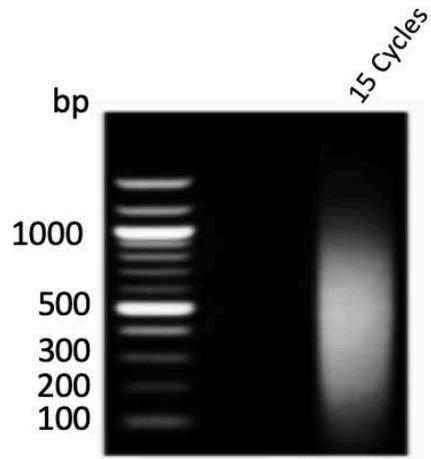
B
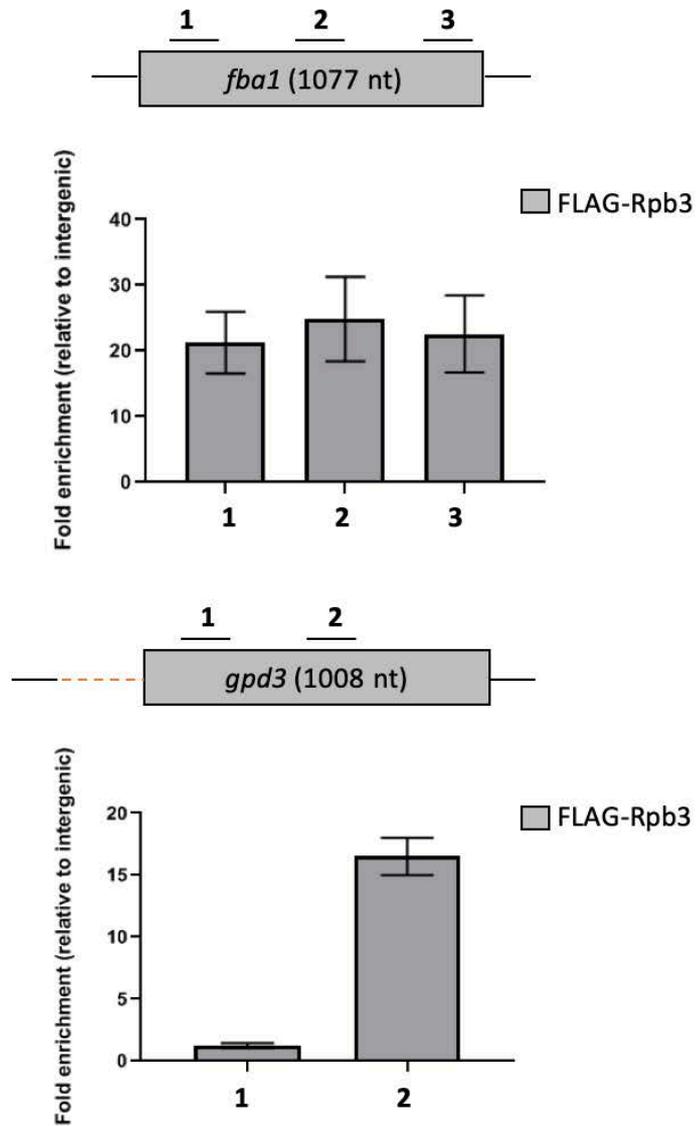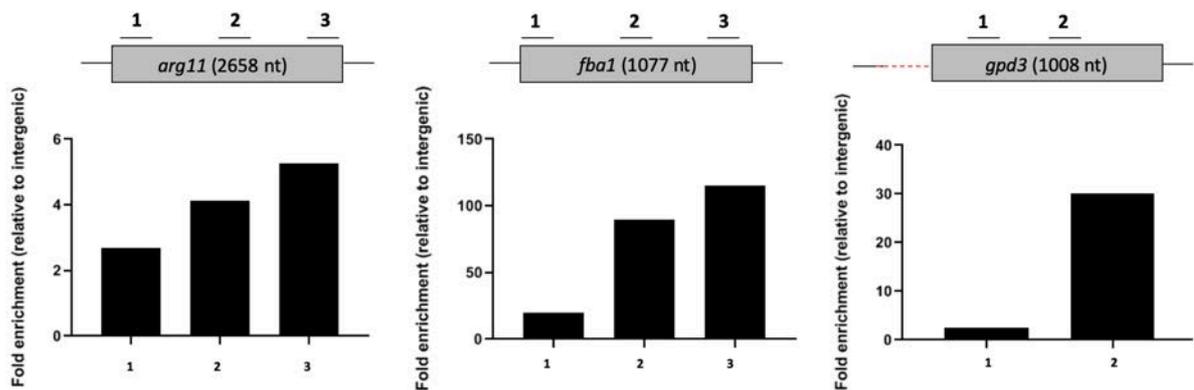
Figure 4.0 - Standardization of ChIP assay in *S. pombe*. (A) The fragmentation of the chromatin is performed using the Diagenode Bioruptor sonicator as described in more details in the Material and Methods; 15 cycles of sonication, of 20 seconds each, allow to generate a pattern of digested fragments accumulated between 200 and 800 bp.

(B) ChIP-qPCR of the Rpb3 subunit of the RNA Pol II, FLAG-tagged at the N-terminal. The occupancy of the Rpb3 subunit is evaluated on two highly expressed genes, *fba1* (1077 nt) and *gpd3* (1008 nt); the numbers reported on the top of the genes correspond to the position of each primer pair, and the red segment in the *gpd3* gene represents an intron located in the 5'-UTR. The signal derived by the qPCR-ChIP analysis is normalized to the input and expressed as fold change relative to a negative locus (intergenic) of *S. pombe.* The error bars represent mean ± SEM (n=2). (C) ChIP-qPCR using the anti-phopsho Ser2 antibody (ab5095) against the phosphorylated form of the Ser2 of the Rpb1 subunit; occupancy is evaluated on a low expressed gene (*arg11*) and two highly expressed ones (*fba1* and *gpd3*), and expressed as fold change relative to a negative locus. Figure in C derives from a single experiment.

4.2.2 - Fal1 associates at transcriptionally active sites regardless of the presence of an intron

To assess the association of Fal1 at transcriptionally active sites, I performed the ChIP assay with a strain in which Fal1 was tagged at the C-terminal with a GFP-2xFLAG epitope, which as discussed earlier, does not affect the functionality of the protein (Chapter 3, Figure 3.4C). The IP step was performed using an anti-FLAG antibody (Sigma, M2), which is highly specific for the FLAG-tag strain as demonstrated by immunoblotting (Figure 4.1A).

The occupancy of Fal1 was evaluated along two RNA Pol II transcribed genes: *fba1* (intron-less) and *gpd3* (intron-containing). Interestingly, I found that Fal1 is associated with *gpd3,* giving a 5 to 6-fold enrichment relative to the intergenic region, hence indicating that Fal1 associates with transcriptionally active sites also in *S. pombe* (Figure 4.1B).

Notably, I also found an association of Fal1 along the CDS of the intron-less gene *fba1;* the enrichment varied from an average of 4-fold in the first region of the gene to almost 6-fold at the end of the CDS (Figure 4.1B). The Fal1 ChIP signal appears to be specific as there is little enrichment in the WT (untagged) strain, which was around 2-fold or below compared to the intergenic region; the enrichment in the untagged strain was set background for the next ChIP assays (Figure 4.1B).

Additionally, the occupancy on *gpd3*, appears to be similar to that on *fba1,* suggesting that in fission yeast, Fal1 does not bind preferentially to intron-containing genes.

The association of Fal1 was also tested on a second intron-less locus, *pma1,* confirming that the binding of Fal1 can occur regardless of an intron (Appendix VI). Overall the experiments demonstrated that Fal1 is specifically associated with transcriptionally active Pol II transcribed genes in *S. pombe,* and the recruitment is intron independent. To evaluate the

global binding profile of Fal1 in *S. pombe,* we collected IP chromatin for ChIP-seq analysis, and the sequencing will be performed in the immediate future.

A
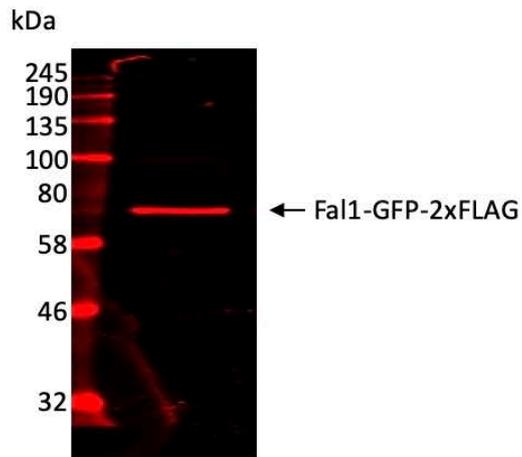
B
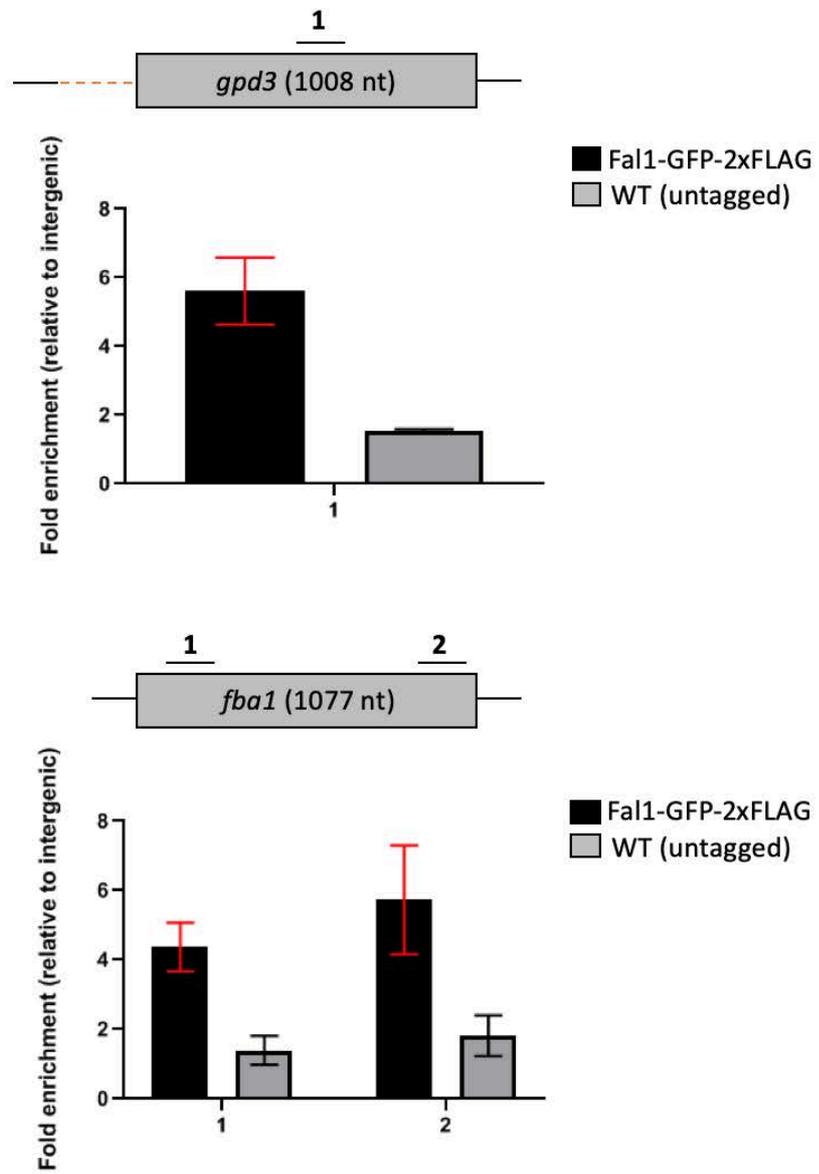
Figure 4.1 - Fal1 associates with RNA Pol II transcribed loci regardless of the presence of an intron. (A) The specificity of the FLAG antibody (Sigma, M2) used for the ChIP assay, is verified by immunoblotting of Fal1-GFP-2xFLAG, detecting a unique band of approximately 72 kDa. Note that the WB is qualitative and aimed to evaluate the specificity of the FLAG antibody; therefore, the membrane was not probed with a loading control antibody.

(B) ChIP-qPCR of Fal1-GFP-2xFLAG and WT (untagged) strain. The occupancy is evaluated on *fba1* (intron-less gene) and *gpd3* (intron-containing gene); the signals are normalized to the input and expressed as fold enrichment relative to a negative locus (not expressed). Error bars represent mean ± SD of independent experiments (n= 4 on *fba1*); (n= 2 on *gpd3*).

4.2.3 - The binding of Fal1 at gene loci is partially sensitive to RNase

Since Fal1 is an ATP dependent-RNA helicase, we tested whether its association with the chromatin was mediated by interaction with the nascent-RNA. To this aim, an RNase step was introduced in the ChIP assay (details in Materials and Methods). The principle behind the RNase treatment is conceptually straightforward; if the target protein is bound to the nascent RNA, the corresponding DNA sequences will be lost during the RNase digestion and the derived signal detected by ChIP-qPCR will be lower compared to the non-treated control. An illustration of the RNase treatment is reported in Figure 4.2A.

I found that the RNase treatment caused a significant reduction (p-value < 0.05) of the Fal1 occupancy on *gpd3* in three independent experiments, of approximately 24%, 47% and 59%, suggesting that Fal1 might be co-transcriptionally recruited to intron-containing transcripts (Figure 4.2B). However, the only partial sensitivity to RNase could also mean that Fal1 can establish additional contacts with the chromatin or with the transcription machinery.

On *fba1,* I found a reduction in the first region of the gene only in two experiments out of three performed (almost 31% and 75% reduction). Conversely, in the last region of the *fba1* gene, I found a reduction in the three experiments performed, of almost 58%, 70% and 30%, but not significant from a statistical point of view (Figure 4.2B). However, also for the second region of *fba1*, the signal of the RNase treated sample appeared partially reduced compared to the control (an average of 52% of reduction) (Figure 4.2B).

A

RNA Pol II DNA

RNase

Fal1

Nascent-mRNA

RNA Pol II DNA

Fal1

B

1

gpd3 (1008 nt)

■ - RNase
▨ + RNase

*
- 43 %

Fold enrichment (relative to intergenic)

1

2   1

fba1 (1077 nt)

■ - RNase
▨ + RNase

NS
- 52%

NS

Fold enrichment (relative to intergenic)

1   2
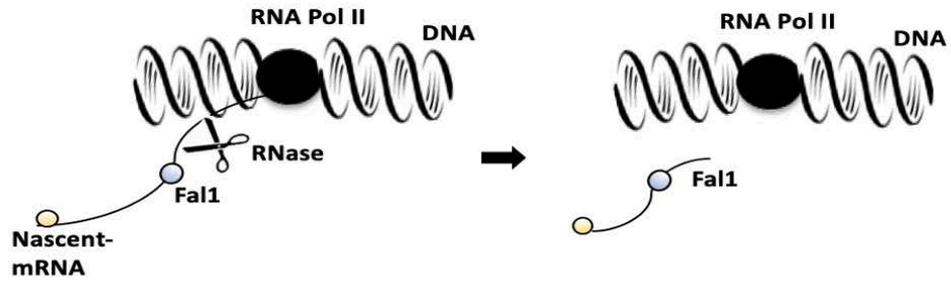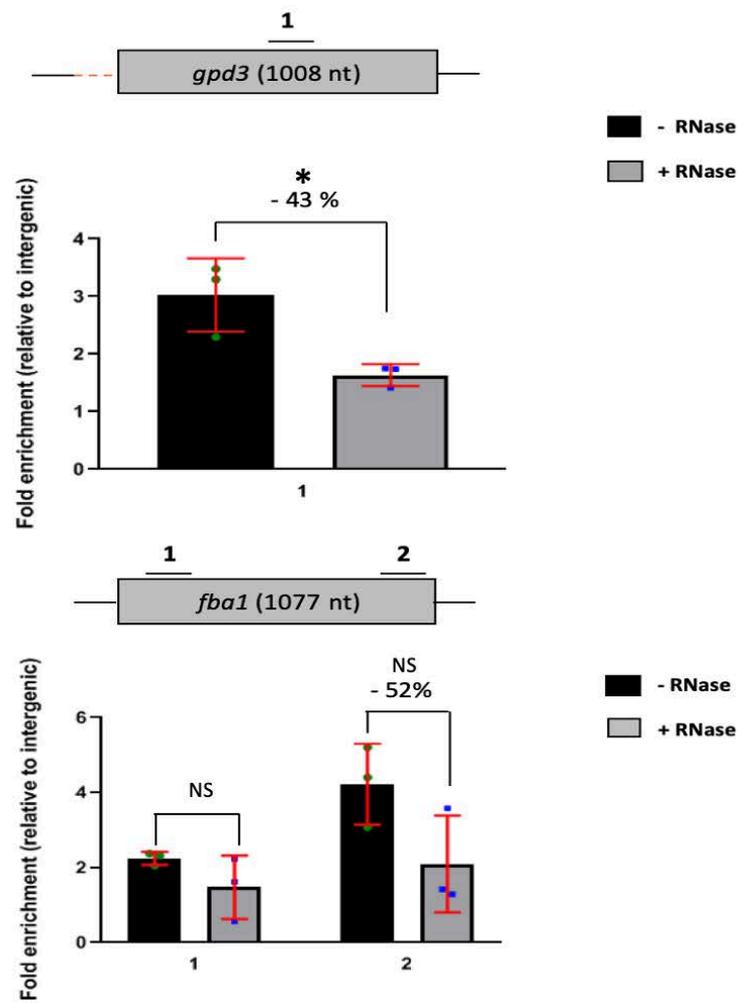
Figure 4.2 – The association of Fal1 with gene loci is partially sensitive to RNase (A) Illustration of the ChIP-RNase treatment: RNase A treatment is performed to evaluate whether Fal1 associates with the nascent-RNA. If the protein is bound to the nascent-RNA, the DNA sequences in its proximity are lost during the treatment, causing a reduction in the ChIP signal compared to the not-treated control. (B) The enrichment of Fal1 is evaluated on *gpd3* and *fba1*, with or without RNase treatment in the ChIP assay. The Fal1 occupancy is normalized to the input and expressed as fold enrichment relative to a negative locus. The average reduction of the Fal1 signal after the RNase treatment is reported in the graph.

The error bars represent the mean ± SD of independent experiments (n=3).

Statistical analysis was performed using unpaired Student t-test: NS (not significant); *p < 0.05, **p < 0.01, ***p < 0.001.

Since the interaction of the EJC components with the transcription machinery has been recently observed in two different studies in *D. melanogaster* and the profile of eIF4AIII on the polytene chromosomes of the same organism, remarkably overlaps with that of RNA Pol II*, I predicted that the partial sensitivity to RNase treatment, could be due to an interaction with the polymerase (Akhtar et al., 2019, Choudhury et al., 2016, Singh et al., 2019).

To test this hypothesis, the binding profile of Fal1 was compared to that of Rpb3, along the CDS of the *fba1* gene. Notably, I found that the pattern of Fal1 and Rpb3 were similar throughout the CDS of the gene (Figure 4.3). For both Fal1 and Rpb3, the enrichment was lower at the beginning of the gene, then it peaked in the middle of the CDS, and lastly decreased in the proximity of the end of the gene (Figure 4.3).

Overall, the data reported in this paragraph indicate that Fal1 is partially sensitive to the RNase treatment and it might interact with the transcription machinery.
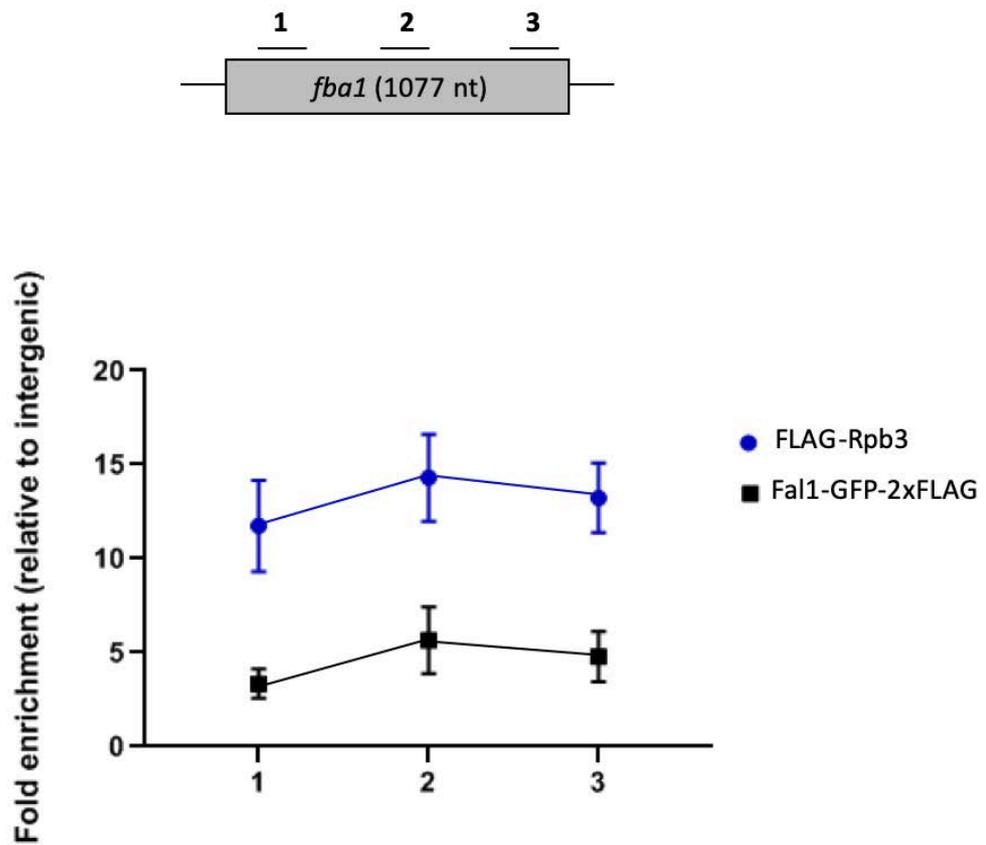
Figure 4.3 – The binding pattern of Fal1 is similar to that of Rpb3. Evaluation of the binding profile of Fal1 and Rpb3 along the CDS of the *fba1* gene (1077 nt). The enrichment of Fal1 and Rpb3 is normalized to the input and expressed as fold enrichment relative to a negative locus. The error bars represent the mean ± SEM of independent experiments (n=3).

4.2.4 - Y14 and MAGO are mostly found in the nucleus of *S. pombe*

According to the EJC model, Fal1 together with MAGO and Y14, makes a stable grip with the RNA (Le Hir et al., 2016). Having found that Fal1 is recruited at active gene loci in an intron-independent fashion, I also investigated whether other members of the EJC exhibited a similar binding mode. To test this hypothesis, I generated yeast strains in which Y14 and MAGO were endogenously tagged at the C-terminal with the GFP-2xFLAG epitope (Figure 4.4A). The expression of the endogenous GFP-2xFLAG tagged Y14 and MAGO was evaluated by western blot using an anti-FLAG antibody; expected bands of approximately 43.2 kDa and 46.5 were detected for Y14 and MAGO respectively (Figure 4.4B). Since the localisation of MAGO and Y14 has never been reported in *S. pombe,* I exploited the GFP-tag to determinate their subcellular distribution. Using confocal microscopy, I found that in *S. pombe,* as in other eukaryotes, the two factors mainly localise in the nucleus, which suggests that their function is predominantly nuclear in fission yeast (Figure 4.4C) (Choudhury et al., 2016, Kataoka et al., 2011, Koroleva et al., 2009, Le Hir et al., 2001a, Pendle et al., 2005). The images were acquired with technical assistance of Dr. Alessandro di Maio, microscopy manager at UoB.

A

B

C

Y14-GFP-2xFLAG
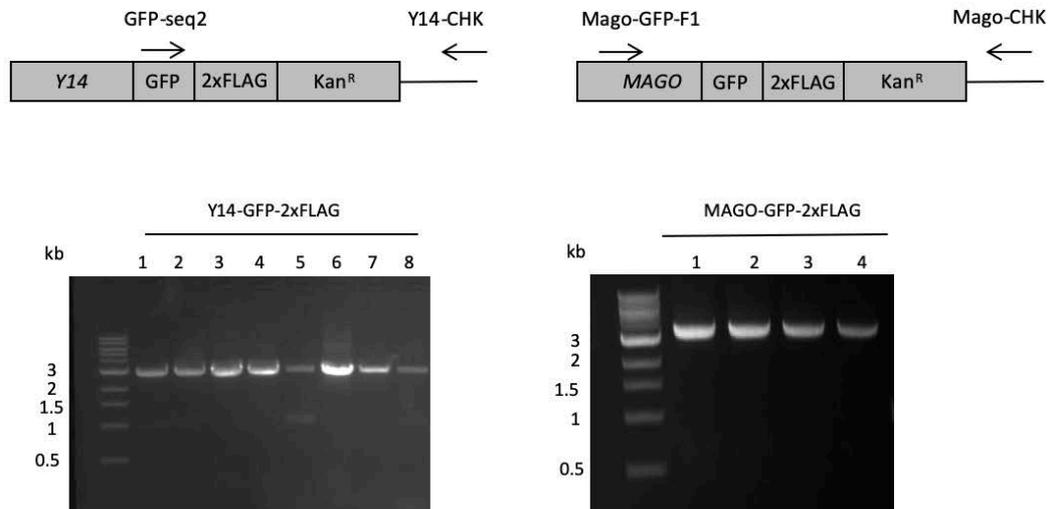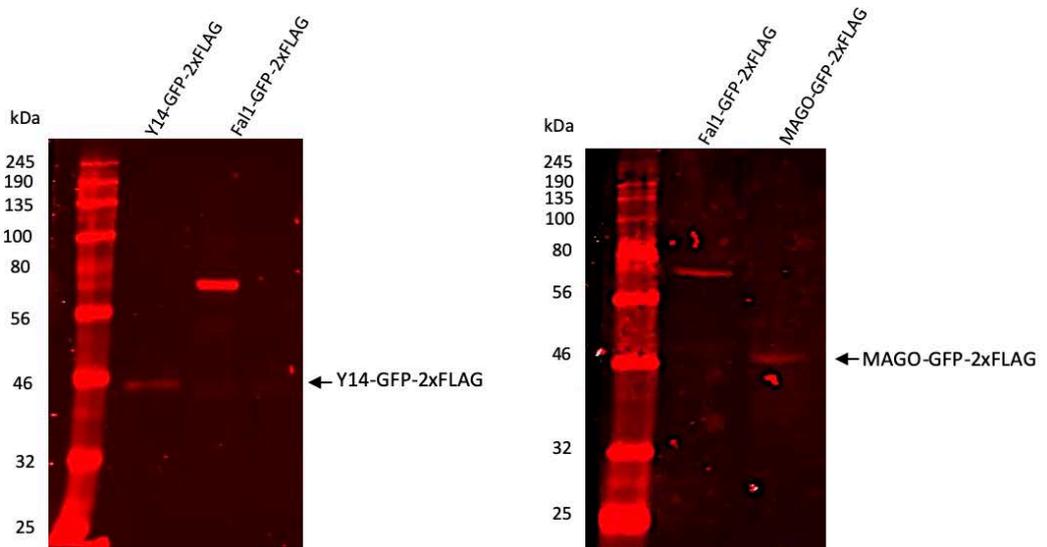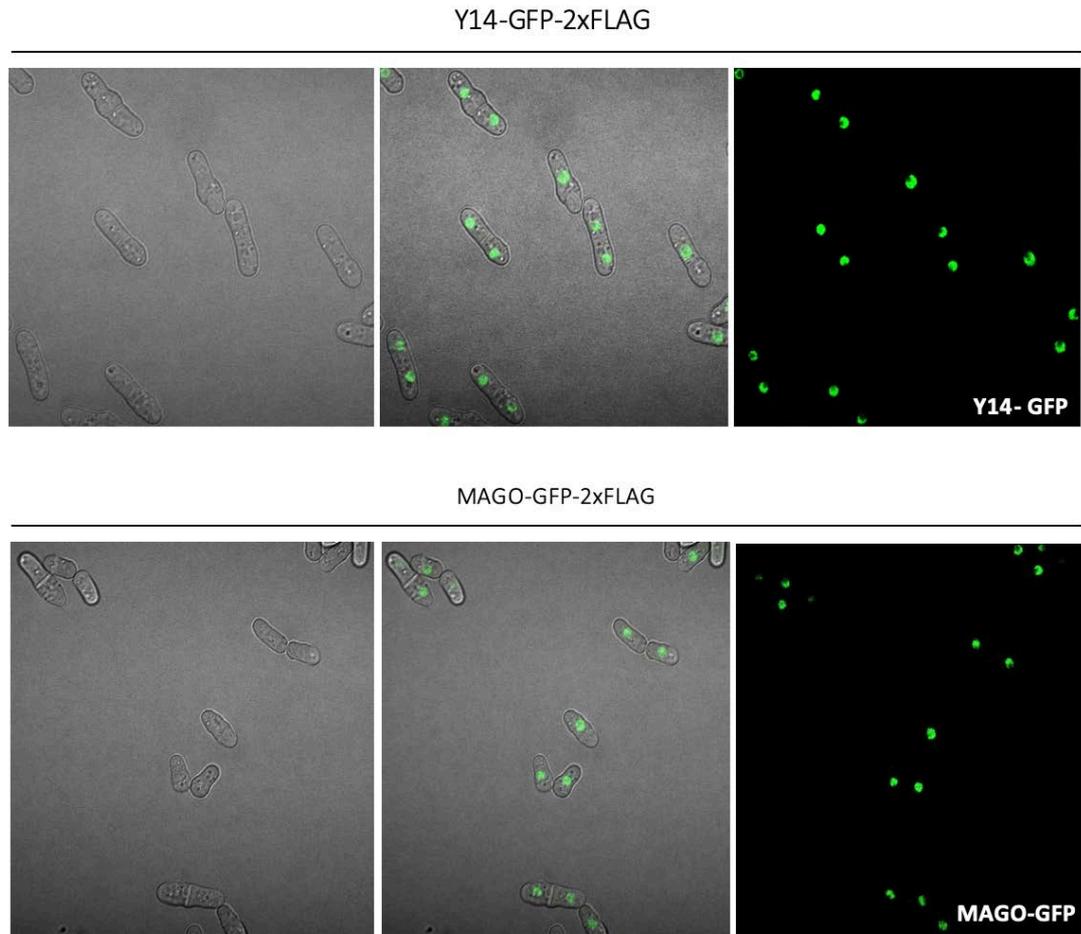
MAGO-GFP-2xFLAG

 Figure 4.4 - MAGO and Y14 are mostly found in the nucleus of *S. pombe.* (A) The integration of the GFP-2xFLAG tag is verified by PCR as illustrated.

(B) The expression of Y14 and MAGO, endogenously tagged with the GFP-2xFLAG epitope, is detected by WB, using anti-FLAG antibody (M2, Sigma); expected bands of approximately 43.2 kDa and 46.5 are detected for Y14 and MAGO respectively. Note that the WB is qualitative and aimed to confirm the presence of the FLAG-tag in the positive colonies; therefore, the membrane was not probed with a loading control antibody.

(C) The GFP signal of MAGO and Y14 was visualized by confocal microscopy (Nikon A1R).

4.2.5 - Y14 might not stably associate with transcription active sites

To test whether Y14 exhibited the same binding pattern of Fal1 on the target genes, a ChIP assay was performed using the newly created strain. Surprisingly, I found a modest enrichment on *gpd3* (almost 2-fold relative to the intergenic), only in one experiment out of three performed; in the other two experiments, the enrichment (varying from 1 to 1.5-fold) was similar to the ChIP signal derived from previous experiments with a WT-untagged strain (Figure 4.5). Also, for *fba1* a modest signal (almost 3-fold relative to the intergenic), was observed in one experiment out of three performed (Figure 4.5). The occupancy of Y14 was also tested on a second highly expressed intron-less gene, *pma1,* observing similar results (Appendix VI). These data suggest that Y14 might not be associated or at least not stably, along the CDS of *fba1* and *gpd3*.

I also performed preliminary ChIP of MAGO, whose occupancy was tested on *fba1,* and *gpd3*, but, I did not observe an obvious enrichment above the intergenic, in agreement with the fact the MAGO and Y14 might form a heterodimer in *S. pombe* (Appendix VI).
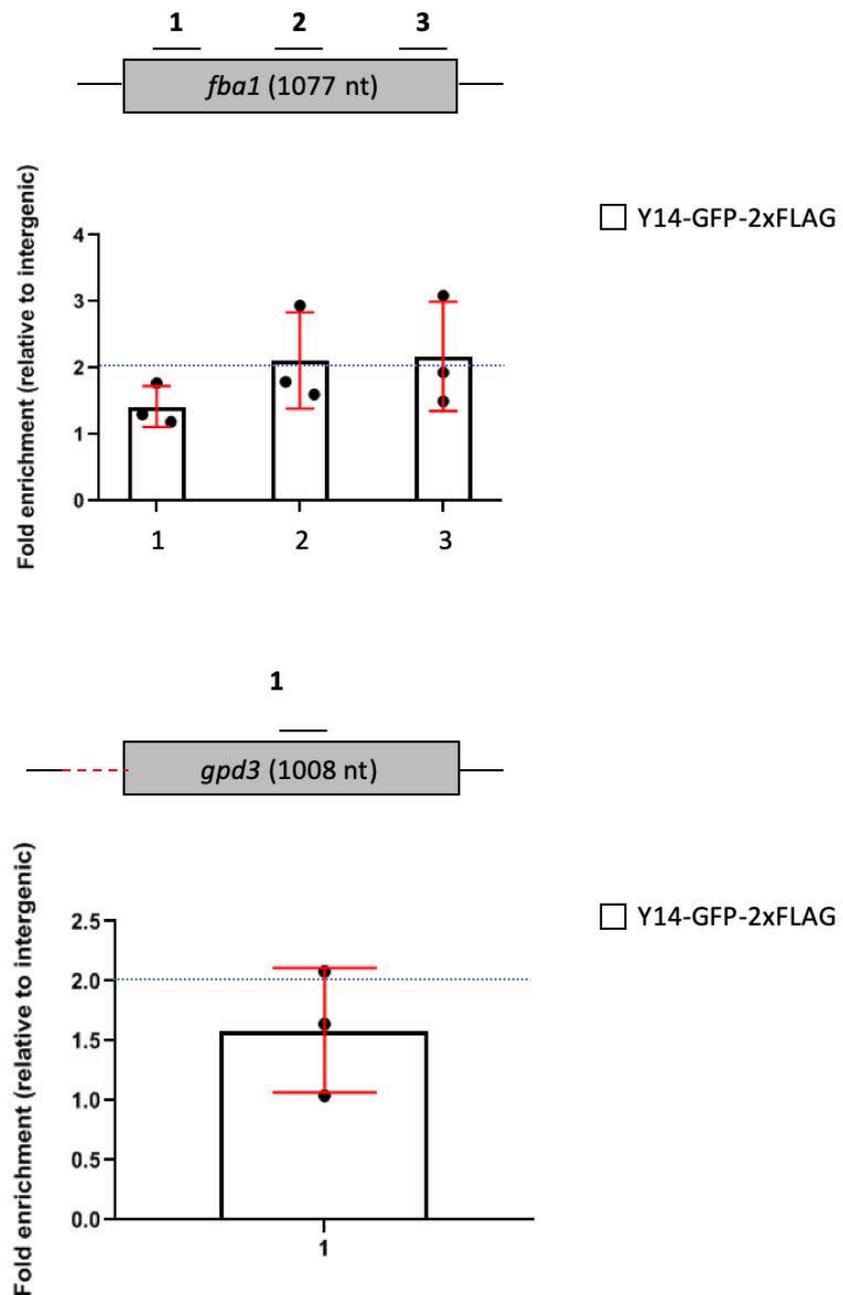
Figure 4.5 - Y14 may not associates or at least not stably with transcription active sites. ChIP-qPCR of Y14 on *fba1* (intron-less gene) and *gpd3* (intron-containing gene); the Y14 signal is normalized to the input and expressed as fold enrichment relative to a negative locus (not expressed). The blue dashed line indicates a level of background previous reported for an untagged-WT strain. Error bars represent mean $\pm$ SD of independent experiments (n=3).

4.3 - Discussion

According to the EJC model, eIF4AIII is deposited 20-24 nucleotides upstream of each exon-exon junction, in a sequence-independent fashion (Le Hir et al., 2000, Le Hir et al., 2016). In mammalian cells a non-canonical deposition of eIF4AIII was revealed using CLIP-seq experiments; eIF4AIII appears prevalently associated with the canonical position, but it also marks different exonic regions of spliced mRNAs. However, its binding seems to occur only on the mRNAs derived from intron-containing genes (Saulière et al., 2012, Singh et al., 2012).

Contrary to this common view, Choudhury and colleagues have reported that in *D. melanogaster* eIF4AIII, and other EJC components, are recruited on the nascent RNA regardless of the presence of an intron (Choudhury et al., 2016). This binding fashion has been recently confirmed in mammalian cells and *D. melanogaster* (Akhtar et al., 2019, Viphakone et al., 2019).

With this project, I aimed to investigate whether eIF4AIII associates with transcriptionally active sites and whether the association depends on the presence of an intron *in S. pombe.* To this end I have performed ChIP of Fal1, the homolog of eIF4AIII in *S. pombe;* its occupancy has been evaluated on two RNA Pol II transcribed genes; *fba1*, which is intron-less, and *gpd3*, which contains an intron in the 5'-UTR. The data indicate that Fal1 is associated along the CDS of *fba1* and *gpd3,* without showing a preferential association for the latter (Figure 4.1). The association on these genes is probably mediated via nascent RNA since when an RNase treatment is introduced in the ChIP assay, it causes a partial reduction of the Fal1 occupancy on *gpd3* (an average of 43%) and *fba1* (an average of 52%, despite the decrease appears reproducible only in the last region of the gene) (Figure 4.2). However, on *fba1* the reduction of Fal1 occupancy is not statistically significant. This can be either due to technical

considerations such as the low number of replicates (N=3) used for the statistical analysis or to a biological mechanism which suggests that Fal1 might associate with RNA Pol II but its deposition of Fal1 occurs only on the mRNA derived by intron-containing genes. It is also important to note that in the ChIP assay, the fixation of protein-chromatin interactions with formaldehyde, could reduce the accessibility of the RNase to the nascent-mRNAs, thus causing a substantial inter-experimental variability (Abruzzi et al., 2004, De et al., 2011). Due to this limitation, it could be interesting to clarify whether Fal1 binds single-exon transcripts using a more direct approach such as Cross-Linking Immunoprecipitation (CLIP) (Ule et al., 2003).

The partial sensitivity to RNase could also reflect a putative interaction of Fal1 with the transcription machinery. For example, in *D. melanogaster* eIF4AIII interacts with RNA Pol II, and the interaction is sensitive to RNase; a similar feature is also shared by MAGO (Akhtar et al., 2019, Singh et al., 2019). Additionally, in *D. melanogaster,* the signal of eIF4AIII, but not that of MAGO and Y14, remarkably overlaps with that of RNA Pol II (Choudhury et al., 2016). From this evidence, Choudhury and colleagues conjectured that eIF4AIII might have a global role in gene expression, and Akthar and colleagues concluded that it may affect the elongation of RNA Pol II, as previously anticipated in mammalian cells (Akhtar et al., 2019, Choudhury et al., 2016, Wang et al., 2014). In support of this hypothesis, I found that Fal1 and the Rpb3, exhibit a similar binding pattern along the *fba1* gene (Figure 4.3). Whether Fal1 has a direct role in the elongation of RNA Pol II needs to be evaluated in the future, possibly following rapid depletion of Fal1 (as discussed in Chapter 3).

To test whether other EJC proteins exhibit a Fal1-like binding profile, Y14 and MAGO have been endogenously tagged with the same GFP-2xFLAG epitope, which has revealed that the two factors accumulate mainly in the nuclear compartment (Figure 5.4), a subcellular

distribution that is in agreement with that observed in other eukaryotes (Choudhury et al., 2016, Kataoka et al., 2011, Koroleva et al., 2009, Le Hir et al., 2001a, Pendle et al., 2005).

However, ChIP experiments of Y14 indicate that the protein might not be or at least not stably associated on the tested genes, since the enrichment was modest and similar to the signal derived from the WT-untagged strain (Figure 5.5). Some preliminary ChIP of MAGO suggest a similar behavior (Appendix VI). This data indicates that eIF4AIII might be associated at some gene loci independently of the heterodimer MAGO and Y14, which is consistent with what has been recently observed in *D. melanogaster* (Choudhury et al., 2016). It is also possible that the association of Y14 on these gene loci is transitory and not stable enough to be captured using the time of fixation of our ChIP protocol. Therefore, it will be crucial in the future to extend the fixation time, and to evaluate the occupancy of Y14 on more gene loci.

Overall the main conclusion of this study is that Fal1 associates with transcriptionally active loci regardless the presence of an intron, in agreement with recent findings reported in *D. melanogaste*r and mammalian cells (Akhtar et al., 2019, Choudhury et al., 2016, Viphakone et al., 2019).

Chapter 5: The main function of Fal1 is independent of MAGO and Y14 in *S. pombe*

5. 1 - Introduction

The current EJC model predicts that the tight association of eIF4AIII with a splice junction depends on the interactions eIF4AIII makes with the MAGO-Y14 heterodimer (Bono et al., 2006). The observation that Y14 is not, or at least not stably associated on spliced mRNA, as discussed in Chapter 4, led me to better explore whether a mammalian-like EJC is formed in *S. pombe.* Here I describe some evidence which argue against the formation of a mammalian-like EJC in fission yeast. Despite MAGO and Y14 being conserved and likely to form a heterodimer in *S. pombe*, their deletion does not affect eIF4AIII stability, nuclear localisation and cell viability; moreover eIF4AIII, unlike MAGO and Y14, associates with active translating mRNA.

5.2 - Results

5.2.1 - MAGO and Y14 are conserved in *S. pombe*

Previous studies concluded that a mammalian-like EJC is unlikely to be formed in *S. pombe* (Marayati et al., 2016, Wen and Brogna, 2010). One possibility to explain this phenomenon is that in *S. pombe,* MAGO and Y14 have diverged too much from higher eukaryotes, and hence structural differences might prevent an interaction with Fal1/eIF4AIII, and therefore also the formation of an EJC.

To explore this hypothesis, I have aligned the protein sequence of MAGO from *S. pombe,* with its orthologs across different model organisms. I found that the sequence of MAGO is highly conserved, having 63% of homology with its counterpart in Human, 61% with *D. melanogaster*, 62% with *C. elegans*, 63 % with *Arabidopsis Thaliana* and 62% with *Xenopus* (Figure 5.0A). Since the primary sequence of MAGO appeared conserved in *S. pombe,* I next investigated whether the predicted secondary structures showed the same level of similarity across the examined organisms. Using the primary amino acid sequence, the secondary structure of MAGO has been predicted based on the similarity with known secondary structures deposited in the database (Cuff et al., 1998). Then each prediction is visualized in Jalview (Clamp et al., 1998, Clamp et al., 2004).

From this analysis it emerged that the MAGO local segments exhibited similar conformations across eukaryotes, presenting almost identical patterns of α-helices (green arrows) and β-sheets (red tubes) in the same positions (Figure 5.0B). Lastly, the 3D modeling from the primary amino acid sequence, using SWISS-MODEL, a protein homology-modeling server, predicted that the tertiary structure of MAGO deriving from *S. pombe*, had a high degree of structural similarity with its Human and *D. melanogaster* orthologs (Figure 5.0C).

A

```
Mago_S.pombe/1-1    1 M------SDFYVRYYSGHHGRFGHEFLEFDYHSDGLARYANNSYRNDSLIRKEMFVSELVLK   57
Mago_Human/1-14     1 ME------SDFYLRYYVGHKGKFGHEFLEFEFRPDGKLRYANNSYKNDVMIRKEAYVHKSVME  58
Mago_Drosophila/1   1 MST-----EDFYLRYYVGHKGKFGHEFLEFEFRPDGKLRYANNSYKNDTMIRKEAFVHQSVME  59
Mago_C.elegans/1·   1 MSGEEEKAADFYVRYYVGHKGKFGHEFLEFEFRPNGSLRYANNSYKNDTMIRKEATVSESVLS  64
Mago_Arabidopsis/   1 MAAE--EASEFYLRYYVGHKGKFGHEFLEFEFREDGKLRYANNSYKNDTIIRKEVFLTPAVLK  62
Mago_Xenopus/1-1    1 MG-----SDFYLRYYVGHKGKFGHEFLEFEFRPDGKLRYANNSYKNDVMIRKEAYVHKSVME  58
Conservation
                       *'''''0''''49'''''''''''''''''''99848'37''''''''9''69'''779435'94
```

```
Mago_S.pombe/1-1   58 EVQRIVDDSEIIKESDESWPPENKDGKQELEIRMNGKHIMFETCKLGSLADVQNSDDPEGLKVF  121
Mago_Human/1-14    59 ELKRIIDDSEITKEDDALWPPPDRVGRQELEIVIGDEHISFTTSKIGSLIDVNQSKDPEGLRVF  122
Mago_Drosophila/1  60 ELKRIIIDSEIMQEDDLPWPPPDRVGRQELEIVIGDEHISFTTSKTGSLVDVNRSKDPEGLRCF  123
Mago_C.elegans/1·  65 ELKRIIEDSEIMQEDDNWPEPDRVGKQELEIILYKNEHISFTTGKIGALADVNNSKDPDGLRSF  128
Mago_Arabidopsis/  63 ECKRIVSESEILKEDDNNWPEPDRVGKQELEIVLGNEHISFATSKIGSLVDVQSSNDPEGLRIF  126
Mago_Xenopus/1-1   59 ELKRIIDDSEVTKEDDALWPPPDRVGRQELEIVIGDEHISFTTSKIGSLIDVNQSKDPEGLRVF  122
Conservation
                       '8'7''939''977'7'34''5589'4'9''''''47457'6'4'7'8'7''965'5''9'95
```

```
Mago_S.pombe/1-1  122 YYLIQDLKALCFSLISLNFKLRPVKN                                        147
Mago_Human/1-14   123 YYLVQDLKCLVFSLIGLHFKIKPI--                                        146
Mago_Drosophila/1 124 YYLVQDLKCLVFSLIGLHFKIKPI--                                        147
Mago_C.elegans/1· 129 YYLVQDLKCLVFSLIGLHFKIKPI--                                        152
Mago_Arabidopsis/ 127 YYLVQDLKCLVFSLISLHFKIKPI--                                        150
Mago_Xenopus/1-1  123 YYLVQDLKCLVFSLIGLHFKIKPI--                                        146
Conservation
                       ''''9'''9'9''''8'5''+9'9--
```

B

```
Mago_S.pombe/1-1    1 M------SDFYVRYYSGHHGRFGHEFLEFDYHSDGLARYANNSYRNDSLIRKEMFVSELVL   56
Mago_Human/1-14     1 ME------SDFYLRYYVGHKGKFGHEFLEFEFRPDGKLRYANNSYKNDVMIRKEAYVHKSVM  57
Mago_Drosophila/1   1 MST-----EDFYLRYYVGHKGKFGHEFLEFEFRPDGKLRYANNSYKNDTMIRKEAFVHQSVM  58
Mago_C.elegans/1·   1 MSGEEEKAADFYVRYYVGHKGKFGHEFLEFEFRPNGSLRYANNSYKNDTMIRKEATVSESVL  63
Mago_Arabidopsis/   1 MAAE--EASEFYLRYYVGHKGKFGHEFLEFEFREDGKLRYANNSYKNDTIIRKEVFLTPAVL  61
Mago_Xenopus/1-1    1 MG-----SDFYLRYYVGHKGKFGHEFLEFEFRPDGKLRYANNSYKNDVMIRKEAYVHKSVM  57
jnetpred
jnetpred
jnetpred
jnetpred
jnetpred
jnetpred
```

```
Mago_S.pombe/1-1   57 KEVQRIVDDSEIIKESDESWPPENKDGKQELEIRMNGKHIMFETCKLGSLADVQNSDDPEGLK  119
Mago_Human/1-14    58 EELKRIIDDSEITKEDDALWPPPDRVGRQELEIVIGDEHISFTTSKIGSLIDVNQSKDPEGLR  120
Mago_Drosophila/1  59 EELKRIIIDSEIMQEDDLPWPPPDRVGRQELEIVIGDEHISFTTSKTGSLVDVNRSKDPEGLR  121
Mago_C.elegans/1·  64 SELKRIIEDSEIMQEDDNWPEPDKIGRQELEIILYKNEHISFTTGKIGALADVNNSKDPDGLR  126
Mago_Arabidopsis/  62 KECKRIVSESEILKEDDNNWPEPDRVGRQELEIVLGNEHISFATSKIGSLVDVQSSNDPEGLR  124
Mago_Xenopus/1-1   58 EELKRIIDDSEVTKEDDALWPPPDRVGRQELEIVIGDEHISFTTSKIGSLIDVNQSKDPEGLR  120
jnetpred
jnetpred
jnetpred
jnetpred
jnetpred
jnetpred
```

```
Mago_S.pombe/1-1  120 VFYYLIQDLKALCFSLISLNFKLRPVKN                                      147
Mago_Human/1-14   121 VFYYLVQDLKCLVFSLIGLHFKIKPI--                                      146
Mago_Drosophila/1 122 CFYYLVQDLKCLVFSLIGLHFKIKPI--                                      147
Mago_C.elegans/1· 127 SFYYLVQDLKCLVFSLIGLHFKIKPI--                                      152
Mago_Arabidopsis/ 125 IFYYLVQDLKCLVFSLISLHFKIKPI--                                      150
Mago_Xenopus/1-1  121 VFYYLVQDLKCLVFSLIGLHFKIKPI--                                      146
jnetpred
jnetpred
jnetpred
jnetpred
jnetpred
jnetpred
```

C

Human   *D. melanogaster*   *S. Pombe*

Figure - 5.0 MAGO is conserved across eukaryotes. (A) The MAGO amino-acid sequence from *S. pombe* shows a high percentage of identity with its counterparts from Human (63%), in *D. melanogaster* (61%), in *C. elegans* (62%), in *Arabidopsis* (62%) and in *Xenopus* (63%). Alignment and visualization are performed using Jalview. Identical residues are indicated in dark blue; gaps are indicated in white; residues with a positive score in the Blosum62 matrix are indicated in light blue. (B) Secondary structure prediction, alignment and visualization of the MAGO sequences across eukaryotes. The prediction of the secondary structure is performed using Jpred and it is visualized in Jalview; beta-sheets and alpha-helices are represented by green arrows and red tubes respectively. The local conformations show a high degree of similarity across the analyzed sequences. (C) The prediction of the tertiary structures of MAGO from Human*, D. melanogaster,* and *S. pombe* is performed using SWISS-MODEL (Schwede et al., 2003).

I also evaluated the conservation grade of Y14 across eukaryotes; the protein sequences from the same organisms were aligned as previously described and visualized in Jalview. I found that Y14 showed 55% of homology with its counterpart in Human, 47% in *D. melanogaster*, 48% in C. *elegans,* 57% in *Arabidopsis* and 51% in *Xenopus* (Figure 5.1A). However, despite the high level of similarity, I also noted that in fission yeast, Y14 lacked the N-terminal domain, which in higher eukaryotes seems to be fundamental to establish interactions with MAGO (Fribourg et al., 2003). As expected, the predicted secondary structures of Y14 exhibited high variability at the N-terminal domain and became progressively more conserved in the central region where identical patterns of α-helices and β-sheets were observed (Figure. 5.1B). The 3D modeling predicted that the tertiary structure of Y14 deriving from *S. pombe*, had a high degree of structural similarity with its Human and *D. melanogaster* orthologs, despite a shorter N-terminal portion (Figure 5.0C). Overall the data presented in this paragraph reveal that MAGO and Y14 are conserved in *S. pombe;* however, the lack of the N-terminal region of Y14, lead us to better explore its relationship with MAGO, whose results are presented in the next paragraphs.

A

```
Y14_S.pombe/1-121       1 M-------------------------------------------------------- 1
Y14_Human/1-174         1 MADVLD---LHEAGGEDFAMDEDGD---------ESIHKL-------KEKAKKRK 36
Y14_D.melanogaster/1-   1 MADVLD---IDN--AEEFEVDEDGD---------QGIVRL-------KEKAKHRK 34
Y14_C.elegans/1-142     1 MSDNEV--------------------------EMEDV------VANAERSR 19
Y14_Arabidopsis/1-201   1 MANIESEAVDFEP-EEDDLMDEDGTAIDGADVSPRAGHPRLKSAIAGANGESAKKTK 56
Y14_Xenopus/1-174       1 MADVLD----LHEAGGEDFAMDEDGD---------ESIHKL-------KEKAKKRK 36

Conservation
                         ' 2 2 0 0 1 - - - - - - - - - - - - - - - - - - - - - 0 1 0 2 2 - - - - - - - 1 0 1 3 2 3 1 3

Y14_S.pombe/1-121       2 -------------------------RPAKSVEGYIIIVTGVHPEATEEQ 25
Y14_Human/1-174        37 GRGFGSEEGSRARMR---EDYDSVE---QDGDEPGPQRSVEGWILFVTGVHEEATEED 88
Y14_D.melanogaster/1-  35 GRGFGSDSNTREAIH----SYERVRNEDDDELEPGPQRSVEGWILFVTSIHEEAQEDE 88
Y14_C.elegans/1-142    20 GRGLAQSR-NRERIT-----YDVVD-EESSATSGAPQRSVEGWIVFVTNIHEEATEDD 70
Y14_Arabidopsis/1-201  57 GRGFREEKDSDRQRRLSSRDFESLG---SDGG-RGPQRSIEGWIILVSGVHEEAQEDD 110
Y14_Xenopus/1-174      37 GRGFGADEGTRARIR---EDYDSVE---QDGDEPGPQRSVEGWILFVTGVHEEATEED 88

Conservation
                         3 3 3 1 0 0 1 1 - 2 2 0 0 0 1 - - - - - 2 3 0 2 0 - - - 1 2 0 0 - - 0 4 ' 6 9 ' 9 - ' * 9 8 8 7 9 ' 5 ' 8 9 7

Y14_S.pombe/1-121      26 VEDLFADFGPVKNLHLNLDRRTGYVKGYALIEYATLEQAQKAVDEKNL-SLLDEKLEV 82
Y14_Human/1-174        89 IHDKFAEYGEIKNIHLNLDRRTGYLKGYTLVEYETYKEAQAAMEGLNGQDLMGQPISV 146
Y14_D.melanogaster/1-  89 IQEKFCDYGEIKNIHLNLDRRTGFSKGYALVEYETHKQALAAKEALNGSEIMGQTIQV 146
Y14_C.elegans/1-142    71 VHDKFSEYGKIKNIHLNLDRRTGFLKGYALVEYETQKEANEAIDQSNDTDLLGQNVKV 128
Y14_Arabidopsis/1-201 111 LTNAFGDFGEIKNLHLNLDRRTGFVKGYALIEYEKYEEAQKAISAMNGAELLTQNVSV 168
Y14_Xenopus/1-174      89 IHDKFGEFGEIKNIHLNLDRRTGYLKGYALVEYETYKEALAAMEGLNGQDLMGQPISV 146

Conservation
                         9 4 7 4 ' 7 9 9 ' 3 9 ' ' * ' ' 9 5 ' 8 ' 9 ' 4 7 4 7 8 ' 6 3 ' 5 6 4 3 ' 3 0 6 * 9 5 8 4 9 4 '

Y14_S.pombe/1-121      83 DFAFLEPPERAPRPS---ISTR-SRSQSPEVQHRDRDVAMAEP 121
Y14_Human/1-174       147 DWCFVRGPPKGKRR-----GGR-RRSRSPDRRRR- 174
Y14_D.melanogaster/1- 147 DWCFVKGPKRVKKS-----DKR-R--------R- 165
Y14_C.elegans/1-142   129 DWCFVKGKKTSGK-------------R- 142
Y14_Arabidopsis/1-201 169 DWAFSRGPSAESYRRKNSRSGRSQRSRSPRRRY- 201
Y14_Xenopus/1-174     147 DWGFVRGPPKGKRR-----SGR-RRSRSPERRRR- 174

Conservation
                         ' 9 9 ' 5 7 7 4 2 4 3 3 6 0 - - - - - 1 1 3 - 1 - - - - - - - - - - 3
```

B

```
Y14_S.pombe/1-121       1 M------------------------------------------------- 1
Y14_Human/1-174         1 MADVLD----LHEAG GEDF AM DEDGD            ESIHKL      KEKAKKRKGRGF GSEEGS 46
Y14_D.melanogaster/1-   1 MADVLD----IDN--AEEFEVDEDGD            QGIVRL      KEKAKHRKGRGF GSDSNT 44
Y14_C.elegans/1-142     1 MSDNEV          EMEDV       VANAERSRGRGL AQSR-N 28
Y14_Arabidopsis/1-201   1 MANIESEAVDFEP-EEDDLMDEDGTAIDGADVSPRAGHPRLKSAIAGANGESAKKTKGRGF REEKDS 66
Y14_Xenopus/1-174       1 MADVLD----LHEAG GEDF AM DEDGD            ESIHKL      KEKAKKRKGRGF GADEGT 46
jnetpred
jnetpred
jnetpred
jnetpred
jnetpred
jnetpred

Y14_S.pombe/1-121       2 -------------------------RPAKSVEGY IIIVT GVHPEAT EEQVEDLF ADFGP VKNLHLNL DR 45
Y14_Human/1-174        47 RARMR---EDYDSVE---QDGDEPGPQRSVEGW ILFVT GVHEEAT EEDIHDKF AEYGE IKNIHLNL DR 108
Y14_D.melanogaster/1-  45 REAIH----SYERVRNEDDDELEPGPQRSVEGW ILFVT SIHEEAQ EDEIQEKF CDYGE IKNIHLNL DR 108
Y14_C.elegans/1-142    29 RERIT-----YDVVD-EESSATSGAPQRSVEGW IVFVT NIHEEAT EDDVHDKF SEYGK IKNIHLNL DR 90
Y14_Arabidopsis/1-201  67 DRQRRLSSRDFESLG---SDGG-RGPQRSIEGW IILVS GVHEEAQ EDDLTNAF GDFGE IKNLHLNL DR 130
Y14_Xenopus/1-174      47 RARIR---EDYDSVE---QDGDEPGPQRSVEGW ILFVT GVHEEAT EEDIHDKF GEFGE IKNIHLNL DR 108
jnetpred
jnetpred
jnetpred
jnetpred
jnetpred
jnetpred

Y14_S.pombe/1-121      46 RTGYVKGY ALIEYAT LEQAQKAVDEKNL-SLLD EKLEVDF AFLEPPERAPRPS---ISTR-SRSQSPE 108
Y14_Human/1-174       109 RTGYLKGY TLVEYET YKEAQAAMEGL NGQDLMGQPISVDW CFVRGPPKGKRR-----GGR-RRSRSPD 170
Y14_D.melanogaster/1- 109 RTGFSKGY ALVEYET HKQALAAKEAL NGSEIMGQTIQVDW CFVKGPKRVKKS-----DKR-R------ 164
Y14_C.elegans/1-142    91 RTGFLKGY ALVEYET QKEANEAIDQS NDTDLLGQNVKVDW CFVKGKKTSGK---------------- 141
Y14_Arabidopsis/1-201 131 RTGFVKGY ALIEYEK YEEAQKAISAMNGAELLTQNVSVDW AFSRGPSAESYRRKNSRSGRSQRSRSPR 198
Y14_Xenopus/1-174     109 RTGFLKGY ALVEYET YKEALAAMEGL NGQDLMGQPISVDW GFVRGPPKGKRR-----SGR-RRSRSPE 170
jnetpred
jnetpred
jnetpred
jnetpred
jnetpred
jnetpred

Y14_S.pombe/1-121     109 VQHRDRDVAMAEP 121
Y14_Human/1-174       171 RRRR-------- 174
Y14_D.melanogaster/1- 165 ---R-------- 165
Y14_C.elegans/1-142   142 ---R-------- 142
Y14_Arabidopsis/1-201 199 RRY--------- 201
Y14_Xenopus/1-174     171 RRRR-------- 174
jnetpred
jnetpred
jnetpred
jnetpred
jnetpred
jnetpred
```

142

Figure 5.1 - Y14 is conserved across eukaryotes. (A) Y14 amino-acid sequence from *S. pombe* shows a high percentage of identity with its counterparts in Human (55%), in *D. melanogaster* (47%), in *C. elegans* (48%), in *Arabidopsis* (57%) and in *Xenopus* (51%). Alignment and visualization are performed using Jalview. Identical residues are indicated in dark blue; gaps are indicated in white; residues with a positive score in the Blosum62 matrix are indicated in light blue. (B) Secondary structure prediction, alignment and visualization of the Y14 sequences across eukaryotes. The prediction of the secondary structure is performed using Jpred and it is visualized in Jalview; beta-sheets and alpha-helices are represented by green arrows and red tubes respectively. (C) The prediction of the tertiary structures of Y14 from Human*, D. melanogaster, and S. pombe* is performed using SWISS-MODEL (Schwede et al., 2003)

5.2.2 – Deletion of *mago* and *y14* do not affect cell viability of *S. pombe*

In higher eukaryotes eIF4AIII associates with MAGO and Y14, forming a stable complex which is tightly locked on the mRNA (Bono et al., 2006). Due to the nature of the complex, it is also possible that *mago* and *y14* deletion mutants could exhibit a severe growth defect as that described for *Fal1* in Chapter 3. Previous studies reported that the deletion of *mago* and *y14* in *S. pombe* did not affect cell viability (Marayati et al., 2016). To test whether this is the case, I decided to generate my own strains and evaluate whether the deletion of *mago* and *y14* affected cell viability in *S. pombe.* I firstly generated deletion strains of *mago* and *y14* (Figure 5.2A); then I performed a spot growth assay on YES plates, comparing the growth of the mutants to the WT. As expected, I confirmed that the deletion of *mago* and *y14* did not affect cell viability; this result suggests that the sick phenotype deriving from the deletion of *fal1* does not depend by the disruption of a putative EJC in *S. pombe* (Figure 5.2B). To exclude the presence of MAGO and Y14 paralogs that might potentially rescue growth defects in *S. pombe*, the amino acid sequence of the two factors was aligned with the Basic Local Alignment Search Toll (BLAST), searching for putative homologs within the same organism. The sequence of MAGO did not align with any protein in the *S. pombe* database (data not shown). For Y14, my search reported 11 hits, which were characterized by a low percentage of identity (Appendix VI). This analysis confirms that *mago* and *y14* might be unique within the *S. pombe* genome, and their deletion, does not impair cell viability. Moreover, the severe growth defect derived by the deletion of *fal1* indicates that it could play an essential function independently of the EJC components.
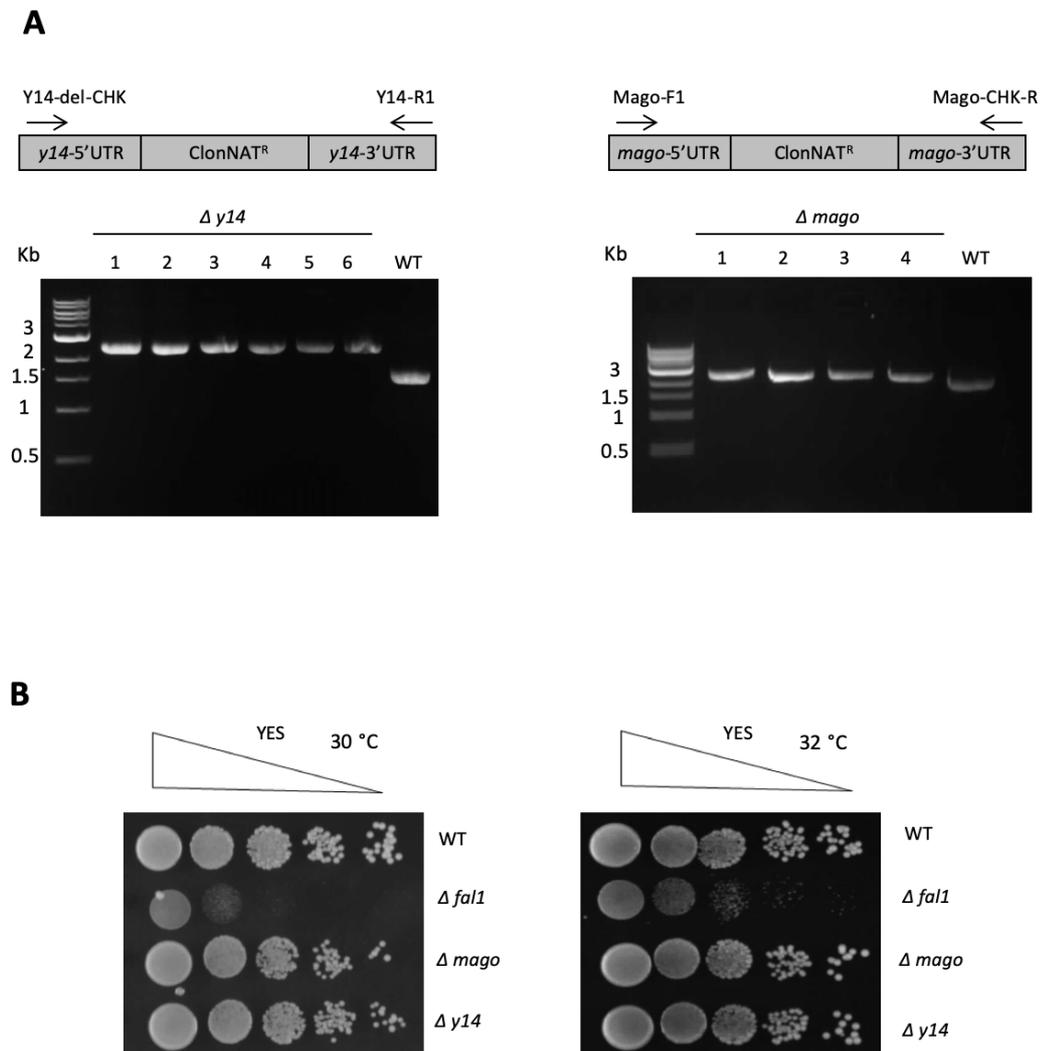
Figure 5.2 – The deletion of *mago* and *y14* do not affect cell viability. (A) Confirmation of the deletion of *mago* and *y14* by PCR: deletion of *y14* is verified by PCR on the DNA extracted from selected colonies, using the primer pair Y14-del-CHK and Y14-R1, with the first anneals outside the region of recombination, amplifying expected bands of approximately 2.3 kb and 1.5 kb for the KO and the WT respectively. Deletion of *mago* is verified via PCR on the DNA extracted from selected colonies, using the primer pair Mago-F1 and Mago-CHK, where the latter anneals outside the region of recombination, amplifying expected bands of approximately 2.4 kb and 2 kb for the KO and the WT respectively. (B) The growth of *mago* and *y14* deletion mutants is compared to the *fal1* deletion and WT strains, by growth spot assay. Dilute spots contain approximately 10^5, 10^4, 10^3, 10^2, 10 cells. The YES plates are incubated at 30°C and 32°C and survival is evaluated after 72 hours.

5.2.3 - MAGO and Y14 fractionate separately from Fal1 on a sucrose gradient

According to the EJC model, the EJC factors remain associated with spliced mRNAs until they are removed by the ribosome during the first round of translation. Key evidence supporting this view is that in mammalian cells, Y14 is present in the monosomal, but not in the polysomal fractions (Dostie and Dreyfuss, 2002).

To gain further insight into the nature of the EJC in *S. pombe,* I have analyzed with the technical help of Dr. Marija Petric, a fellow PhD student in the group, the distribution of the EJC proteins in polysomal fractions of fission yeast cytoplasmic extracts. A schematic representation of the procedure is reported in figure 5.3A.

However, to avoid possible issues relative to the functionality of MAGO and Y14, which were previously tagged with a GFP-2xFLAG epitope, we decided to perform the polysome profiling with new strains in which MAGO and Y14 were endogenously tagged at the C-terminal with a 5xFLAG, that being smaller should have no impact on the protein functionality (creation of the strains, validation and WB analysis is reported in additional results of Appendix VI). Notably, we found that Y14 and MAGO did not sediment with Fal1 by sucrose gradient fractionation of fission yeast cytoplasmic extracts (Figure 5.3B). In fact, while MAGO and Y14 stayed on the light fraction (lane 11 and 12), corresponding to free proteins or small molecular weight complexes, Fal1 was detected both in the monosomal and polysomal fractions, suggesting that it is not removed by the translating ribosome. The sedimentation profile of Y14 and MAGO also suggests that they are not bound to mRNAs in the cytoplasm, contrary to what the EJC model predicts for mammalian cells, but in agreement to more recent observation in *D. melanogaster* (Choudhury et al., 2016).

Additionally, the different pattern of sedimentation of the EJC factors, suggests that it is not likely that a stable EJC is formed in fission yeast.
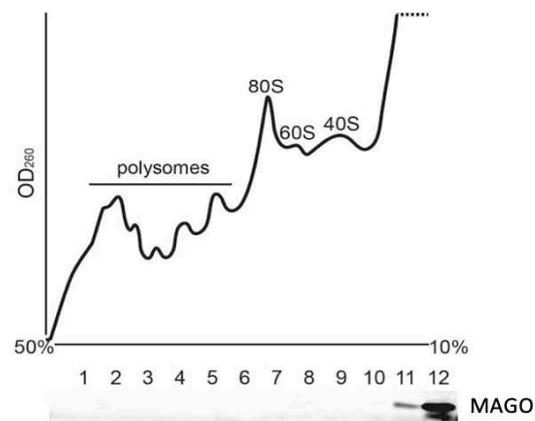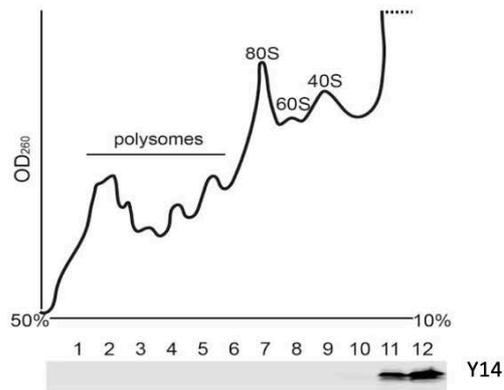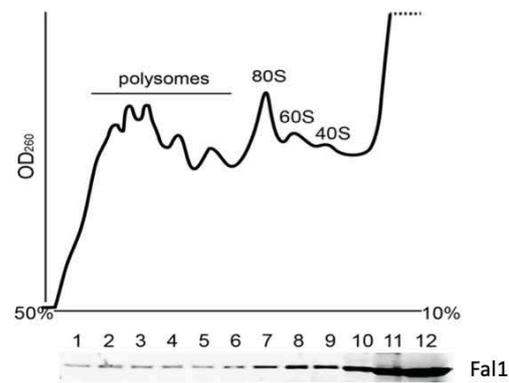
Figure 5.3 –MAGO and Y14 fractionate separately from Fal1 on sucrose gradient. Polysome profiling and western blotting analysis of the sucrose gradient fractions, show the distribution of eIF4AIII, Y14 and MAGO.

5.2.4 - MAGO and Y14 are required for the stability of each other in *S. pombe*

Next, I examined whether MAGO and Y14 can form a heterodimer in *S. pombe*. The two components form an obligatory heterodimer in almost 12 species in which the interaction has been studied (Gong et al., 2014); according to previous studies they not only interact, but they are also required for the stability of each other. (Boisramé et al., 2019, Choudhury et al., 2016, Le Hir et al., 2001a, Miller, 2017). However, in fission yeast the interaction and interdependency between MAGO and Y14 have not yet been investigated.

To gain further insight into this aspect, I firstly inspected whether the residues of interaction between MAGO and Y14, known from previous studies in mammalian cells, were conserved in *S. pombe*. Based on the crystal structure of the mammalian EJC, Bono and colleagues reported the conserved residues of MAGO which were required for its interaction with Y14 (Bono et al., 2006). By aligning and visualizing the MAGO sequences from *S. pombe, D. melanogaster* and Human, I found that most of these residues were highly conserved (Figure 5.4A); however, in *S. pombe,* a conserved histidine in position 53 (HIS53), is replaced with a serine. In order to investigate whether this amino acid substitution can be accommodated in the mammalian 3D structure of the EJC (PDB accession number: 2J0S), without major rearrangements of the protein, I have used USCF Chimera, a package for molecular visualisation, which provides several tools for exploring and editing 3D structures (Pettersen et al., 2004). Interestingly, the substitution of HIS53 in the mammalian 3D structure of the EJC, with the serine rotamer (the conformation of the side-chain of an amino acid) having the highest probability (40% of probability according to the Dunbrack rotamers library of 2010; Table in Appendix VI) does not appear to cause clashes or unfavorable interactions in the

structure (Figure 5.4B). Overall the strong conservation of these amino acids and the apparent

neutrality of the SER53, suggest that MAGO might interact with Y14 also in fission yeast.

A


```
Mago_S.pombe/1-147       1 MS - - DF YVRYYSGHHGRFGHEFLEFDYHSDGLARYANNSNYRNDSLIR    46
Mago_D.melanogaster/1-147 1 MSTEDFYLRYYVGHKGKFGHEFLEFEFRPDGKLRYANNSNYKNDTMIR    48
Mago_Human/1-146         1 ME - SDFYLRYYVGHKGKFGHEFLEFEFRPDGKLRYANNSNYKNDVMIR    47

Conservation

                          * 6 - - * * * 9 * * 6 * 9 * 9 * * * * * * * * * * * 9 9 8 7 * * 6 7 * * * * * * 9 * * 6 9 * *

Mago_S.pombe/1-147       47 KEMFVSELVLKEVQRIVDDSEIIKESDESWPPENKDGKQELEIRMNGK    94
Mago_D.melanogaster/1-147 49 KEAFVHQSVMEELKRIIIDSEIMQEDDLPWPPPDRVGRQELEIVIGDE    96
Mago_Human/1-146         48 KEAYVHKSVMEELKRIIDDSEITKEDDALWPPPDRVGRQELEIVIGDE    95

Conservation

                          * * 8 9 * 4 6 5 * 9 7 * 9 7 * * 9 4 * * * * 7 7 * 7 * 3 4 * * * 5 8 9 5 * 9 * * * * * 4 9 7 5 7

Mago_S.pombe/1-147       95 HIMFETCKLGSLADVQNSDDPEGLKVFYYLIQDLKALCFSLISLNFKL   142
Mago_D.melanogaster/1-147 97 HISFTTSKTGSLVDVNRSKDPEGLRCFYYLVQDLKCLVFSLIGLHFKI   144
Mago_Human/1-146         96 HISFTTSKIGSLIDVNQSKDPEGLRVFYYLVQDLKCLVFSLIGLHFKI   143

Conservation

                          * * 6 6 * 7 7 * * * 7 * * 9 7 * 6 * * * * * 9 9 * * * 9 * * * * 9 9 * * 8 * 5 * * +

Mago_S.pombe/1-147      143 RPVKN                                                147
Mago_D.melanogaster/1-147 145 KPI - -                                              147
Mago_Human/1-146        144 KPI - -                                              146

Conservation
```

B

Figure 5.4 - Residues of interaction of MAGO with Y14 are conserved in *S. pombe*. (A) Alignment and visualization of the MAGO amino acid sequences from *S. pombe, D. melanogaster* and Human. The amino acid residues of MAGO, required for the interaction with Y14 (derived from (Bono et al., 2006)), are grouped and visualized in Jalview. Identical residues are indicated in dark blue; gaps are indicated in white; residues with a positive score in the Blosum62 matrix are indicated in light blue. (B) Most of the residues are conserved except for a histidine in position 53 (HIS53) of *S. pombe,* which is replaced by a serine. The substitution of the HIS53 with the most likely serine rotamer (40% of probability according to Dunbrack rotamers library of 2010) using USCF Chimera, does not cause clashes or unfavorable interactions in the mammalian EJC structure.

Contrary to the highly conserved features of MAGO, the protein sequence of Y14 in *S. pombe* lacks an extensive N-terminal region where seven residues of interaction with MAGO were previously identified in mammalian cells (Fribourg et al., 2003) (Figure 5.5A); however the lack of the N-terminal region seems to be not critical for a putative interaction between MAGO and Y14, since also Y14 from the budding yeast *Yarrowia lipolytica* lacks a 50 amino-acid N-terminal extension present in the human orthologue, it can still interact and stabilize MAGO (Boisramé et al., 2019).

The central domain of Y14 is characterized by 13 amino acids involved in the interaction with MAGO, which are highly conserved, except for phenylalanine in position 76 (PHE 76), which is replaced with an isoleucine. Interestingly, the substitution of the PHE 76, in the mammalian 3D structure of the EJC, with the most likely isoleucine rotamer (68% of probability according to the Dunbrack rotamers library of 2010; table in Appendix VI), seems to cause a clash with the valine in position 56 of MAGO (VAL56), and the cysteine in position 149 (CYS149) of Y14 (Figure 5.5B).

A

Y14_S.pombe/1-121	1 M -	1
Y14_D.melanogaster/1-165	1 M A D V L D I D N A - - E E F E V D E D G D Q G I V R L K E K A K H R K G R G F G S D S N T R E A I H -	49
Y14_Human/1-174	1 M A D V L D L H E A G G E D F A M D E D G D E S I H K L K E K A K K R K G R G F G S E E G S R A R M R E	52

Conservation

Y14_S.pombe/1-121	2 - - - - - - - - - - - - - - R P A K S V E G Y I I I V T G V H P E A T E E Q V E D L F A D F G P V K N	38
Y14_D.melanogaster/1-165	50 S Y E R V R N E D D D E L E P G P Q R S V E G W I L F V T S I H E E A Q E D E I Q E K F C D Y G E I K N	101
Y14_Human/1-174	53 D Y D S V E Q - - - D G D E P G P Q R S V E G W I L F V T G V H E E A T E E D I H D K F A E Y G E I K N	101

Conservation

- - - - - - - - - - - - - 4 ' 6 9 ' ' ' ' ' + ' + 8 ' ' 8 9 ' 5 ' ' 8 ' 9 7 9 5 9 6 ' 9 9 9 ' 5 9 ' '

Y14_S.pombe/1-121	39 L H L N L D R R T G Y V K G Y A L I E Y A T L E Q A Q K A V D E K N - L S L L D E K L E V D F A F L E P	89
Y14_D.melanogaster/1-165	102 I H L N L D R R T G F S K G Y A L V E Y E T H K Q A L A A K E A L N G A E I M G Q T I Q V D W C F V K G	153
Y14_Human/1-174	102 I H L N L D R R T G Y L K G Y T L V E Y E T Y K E A Q A A M E G L N G Q D L M G Q P I S V D W C F V R G	153

Conservation

+ ' ' ' ' ' ' ' ' ' 9 5 ' ' ' 8 ' 9 ' 4 ' 5 7 8 ' 7 5 ' 5 9 4 6 ' - 5 6 + 9 5 8 4 + 6 ' ' 9 9 ' 9 7 7

Y14_S.pombe/1-121	90 P E R A P R P S I S T R S R S Q S P E V Q H R D R D V A M A E P	121
Y14_D.melanogaster/1-165	154 P K R V K K S E K - - R R - - - - - - - - - - R - - - - - - - - -	165
Y14_Human/1-174	154 P P K G K R R G G - - R R R S R S P D R R R R - - - - - - - - -	174

Conservation

' 3 9 9 4 9 4 4 4 ' 6 ' ' ' ' ' ' ' ' '

B

PHE 76   ILE 76   CYS 149   VAL 56   ILE 76

154

Figure 5.5 - Residues of interaction of Y14 with MAGO are not fully conserved in *S. pombe*. (A) Alignment and visualization of the Y14 sequences from *S. pombe, D. melanogaster* and Human. The amino acid residues of Y14, required for the interaction with MAGO (derived from(Bono et al., 2006, Fribourg et al., 2003)), are grouped and visualized in Jalview. Identical residues are indicated in dark blue; gaps are indicated in white; residues with a positive score in the Blosum62 matrix are indicated in light blue. In *S. pombe,* Y14 lacks an N-terminal portion, containing 7 conserved residues involved in the interaction with MAGO. In the central domain, all the residues are conserved except for phenylalanine in position 76 (PHE76) of *S. pombe,* which is substituted with an isoleucine. (B) The substitution of the PHE76 with the most likely isoleucine rotamer (68% of probability according to Dunbrack rotamers library of 2010) using USCF Chimera, causes unfavorable interactions with the VAL56 of MAGO and CYS149 of Y14 (reported as yellow line).

To further investigate whether MAGO and Y14 could form a stable complex in *S. pombe* I decided to explore experimentally the nature of the interaction, asking whether MAGO and Y14 are interdependent in *S. pombe,* as in other eukaryotes (Gong et al., 2014). In fact, it has been reported that the stability of MAGO and Y14 depends on their heterodimerisation; therefore evaluating the protein levels of a component in absence of its partner, could be an indirect approach to evaluate whether the two factors form a stable complex (Choudhury et al., 2016).

To this aim, I firstly generated deletion of *mago* in a strain in which Y14 was previously tagged with a 5xFLAG epitope at the C-terminal (Figure 5.6B). Next, I performed western blot analysis using an anti-FLAG antibody, comparing the protein levels of Y14-5xFLAG with those in which *mago* was deleted. Interestingly, I found that the deletion of *mago* caused a reduction of the protein levels of Y14, in agreement with previous observation in mammalian cells and *D. melanogaster* (Choudhury et al., 2016, Miller, 2017) (Figure 5.6B).
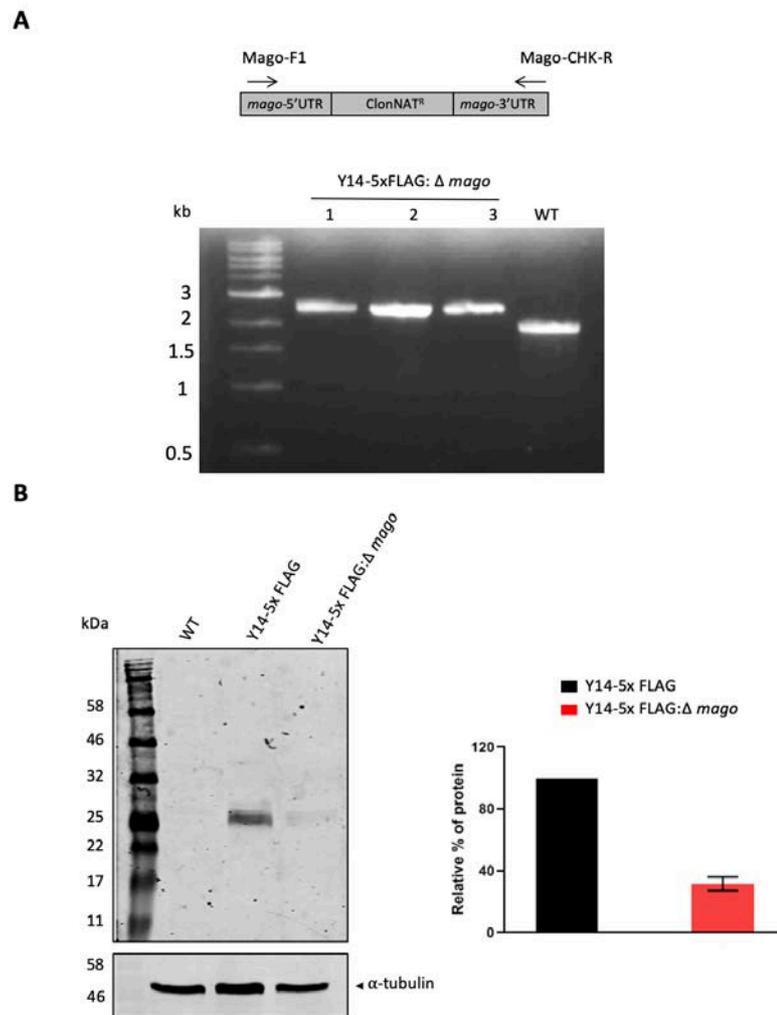
Figure - 5.6 MAGO is required for the stability of Y14. (A) Confirmation of the deletion of *mago* in a strain in which Y14 is tagged at the C-terminal with a 5xFLAG epitope: deletion of *mago* is verified by PCR on DNA extracted from three colonies, using the primer pair Mago-F1 and Mago-CHK-R, where the second anneals outside the region of recombination, amplifying expected bands of approximately 2.4 kb and 2 kb for the KO and the WT respectively. (B) WB analysis that compares the protein levels of Y14-5xLAG with Y14-5xFLAG:Δ*mago*. Quantification of the bands of interest is normalized to the loading control (α-tubulin), and expressed as % of protein relative to the Y14-5xFLAG strain, which is set at 100%. Error bars are mean ± SEM (n=2). Quantification was performed using ImageJ.

To evaluate whether the two factors were interdependent, I also generated a deletion of *y14* in a strain in which MAGO was previously tagged with a 5xFLAG epitope at the C-terminal (Figure 5.7A). The WB experiment confirmed that deletion of *y14* caused a reduction of the protein levels of MAGO of approximately the same magnitude (Figure 5.7B). Overall the results reported in this paragraph indicate that MAGO and Y14 are required for their reciprocal stability in *S. pombe;* therefore they likely form a stable complex as in other organisms (Gong et al., 2014). Furthermore, the few variations in the amino acids which participate in the interactions between MAGO and Y14, could be interpreted as a strategy of fission yeast, to preserve the heterodimerisation as suggested by Gong and colleagues (Gong et al., 2014). Additionally, the lack of the N-terminal portion of Y14, containing some amino acids which are fundamental for its interaction with MAGO in mammalian cells, might be not crucial in yeast as recently observed also by Boisrame and colleagues (Boisramé et al., 2019).
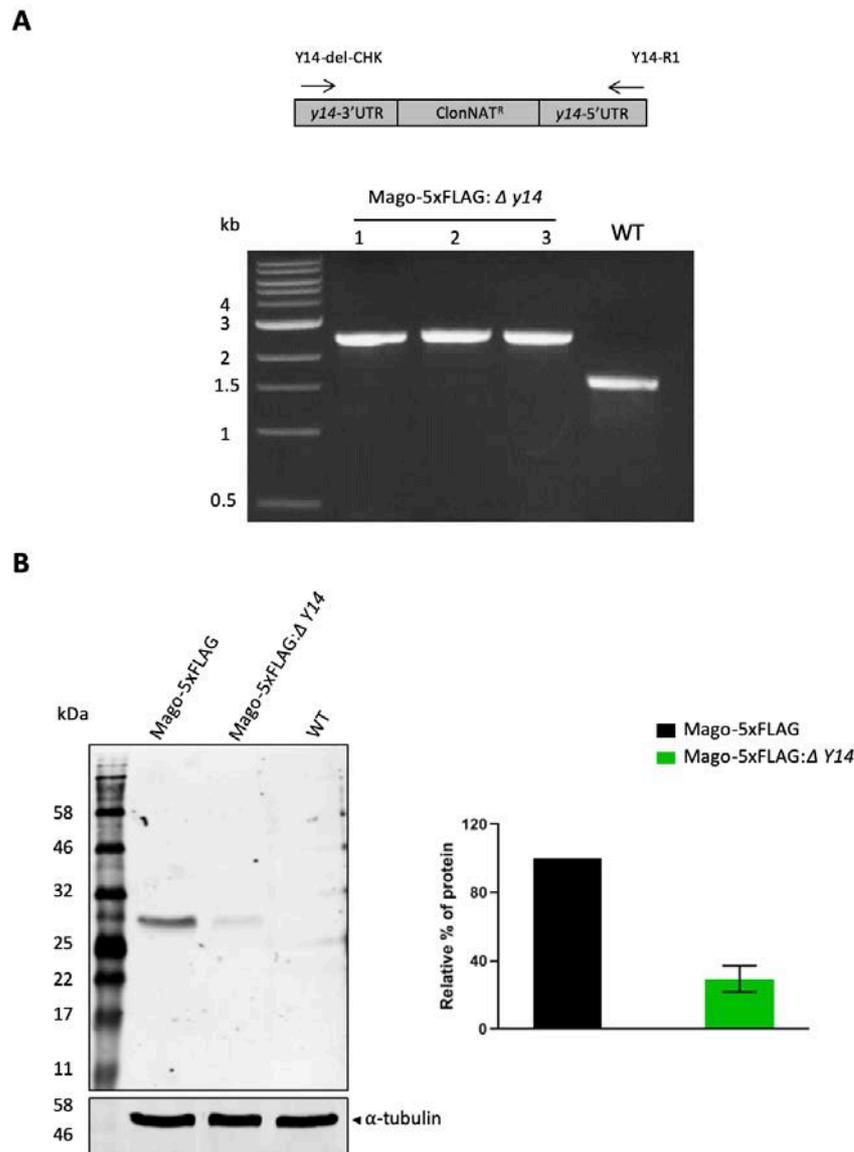
Figure 5.7 Y14 is required for the stability of MAGO. (A) Confirmation of the deletion of *y14* in a strain in which MAGO is tagged at the C-terminal with a 5xFLAG epitope: deletion of *y14* is verified by PCR on the DNA extracted from three colonies, using the primer pair Y14-del-CHK and Y14-R1, with the first annealing outside the region of recombination, amplifying expected bands of approximately 2.3 kb and 1.5 kb for the KO and the WT respectively. (B) Western blot analysis comparing the protein levels of MAGO-5xFLAG with MAGO-5xFLAG:ΔY14. Quantification of the bands of interest is normalized to the loading control (α-tubulin), and expressed as % of protein, relative to the MAGO-5xFLAG strain, which is set at 100%. Error bars are mean ± SEM (n=2). Quantification has been performed using ImageJ.

159

5.2.5 - *y14* deletion causes a depletion of MAGO nuclear signal and vice-versa

As shown in the previous paragraph, the deletion of *mago* reduces the levels of Y14 and vice-versa, suggesting that also in *S. pombe,* the two proteins are required for the stability of each other. To confirm this aspect from a qualitative point of view, I also evaluated the GFP signal of MAGO upon deletion of *y14*, and the GFP signal of Y14 upon deletion of *mago.*

To accomplish this goal, I deleted *y14* and *mago* in the strains in which their respective partners were tagged at the C-terminal with a GFP-2xFLAG epitope.

Then I visualized the GFP signal of MAGO, with and without the deletion of *y14* by confocal microscopy. As previously shown in Chapter 4, MAGO mainly localised in the nuclear compartment of *S. pombe;* interestingly, upon deletion of *y14*, the GFP signal decreased, and the remaining appeared fragmented (Figure 5.8).

To evaluate the inter-dependency of the mechanism, I visualized the GFP signal of Y14, with and without the deletion of *mago* by confocal microscopy; in agreement with the previous observation, the Y14-GFP nuclear accumulation decreased when its partner was deleted (Figure 5.9).

MAGO-GFP-2xFLAG
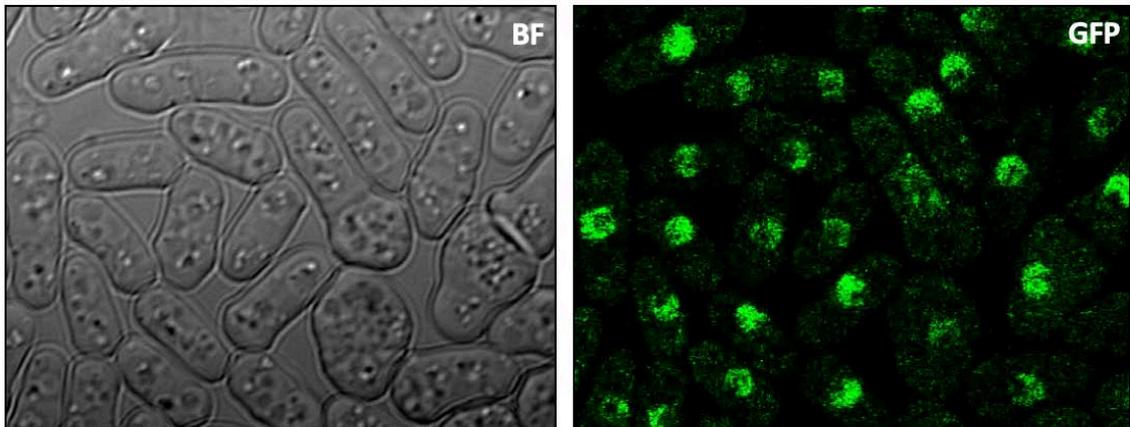
MAGO-GFP-2xFLAG: *Δy14*
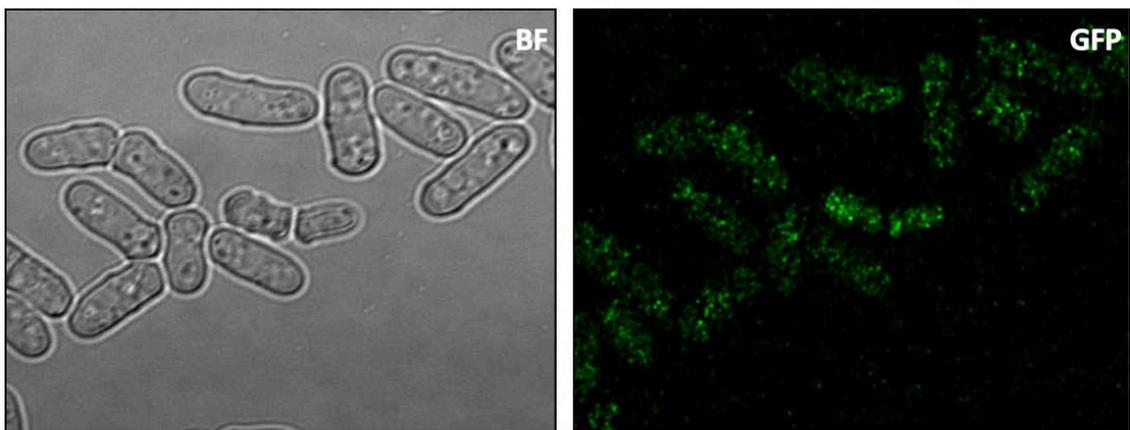
Figure 5.8 – *y14* deletion causes a depletion of MAGO nuclear signal. *Y14* deletion is performed in a strain in which MAGO was tagged at the C-terminal with a GFP-2xFLAG epitope. The GFP signal of MAGO with and without deletion of *y14* was visualized by confocal microscopy (Nikon A1R).

Y14-GFP-2xFLAG
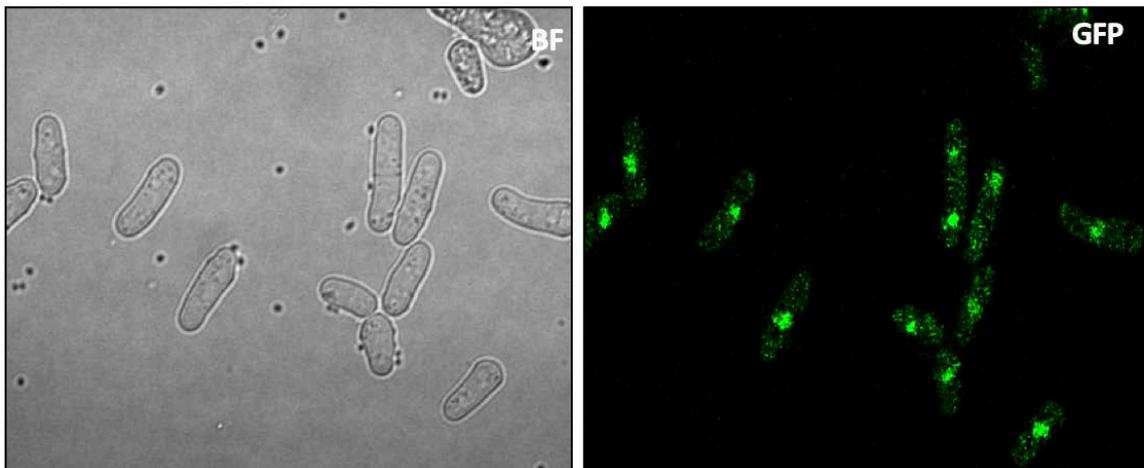
Y14-GFP-2xFLAG: *Δmago*
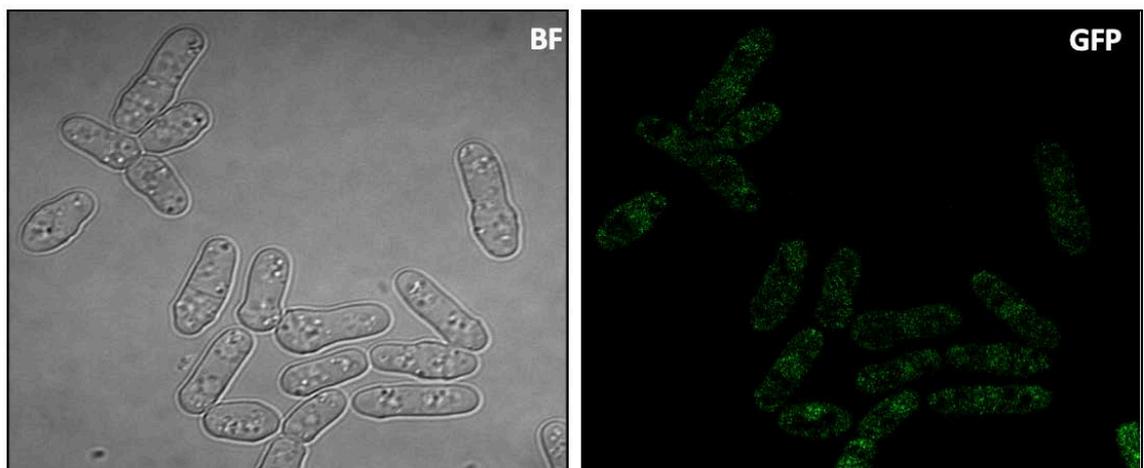
Figure 5.9 - *mago* deletion causes a depletion of Y14 nuclear signal. *mago* deletion is performed in a strain in which Y14 was tagged at the C-terminal with a GFP-2xFLAG epitope. The GFP signal of Y14 with and without deletion of *mago* was visualized by confocal microscopy (Nikon A1R).

5.2.6 - Fal1 exists independently of MAGO-Y14 in *S. pombe*

As shown in the previous paragraphs, the deletion of *fal1* but not *mago* and *y14*, severelly affected cell viability, indicating that the Fal1 main function is independent of MAGO and Y14 in *S. pombe*. To investigate whether there is an inter-dependency between Fal1 and MAGO/Y14, I studied whether protein level, nuclear localisation and the functionality of Fal1 were affected upon deletion of *mago* and *y14*.

To this aim, I deleted *mago* and *y14* in a strain in which Fal1 was tagged at the C-terminal with the GFP-2xFLAG epitope (Figure 5.10). Next, I evaluated by confocal microscopy, whether the nuclear localisation of Fal1 was affected upon deletion of *mago* and *y14*; not surprisingly, I found that the deletions of *mago* and *y14* did not affect the nuclear localisation of Fal1 (Figure 5.11).

To evaluate whether *mago* and *y14* deletions affected the protein levels of Fal1, I performed WB analysis comparing the protein levels of Fal1-GFP-2xFLAG with those in which *mago* and *y14* were deleted. I found that their deletion did not cause a significant reduction of the protein levels of Fal1 (Figure 5.12A). Moreover, the deletions of *mago* and *y14* did not alter the functionality of Fal1, since the tagged strain did not show any growth defect (Figure 5.12B). Overall, the results reported in this paragraph indicate that Fal1 exists in *S. pombe* regardless of MAGO and Y14.
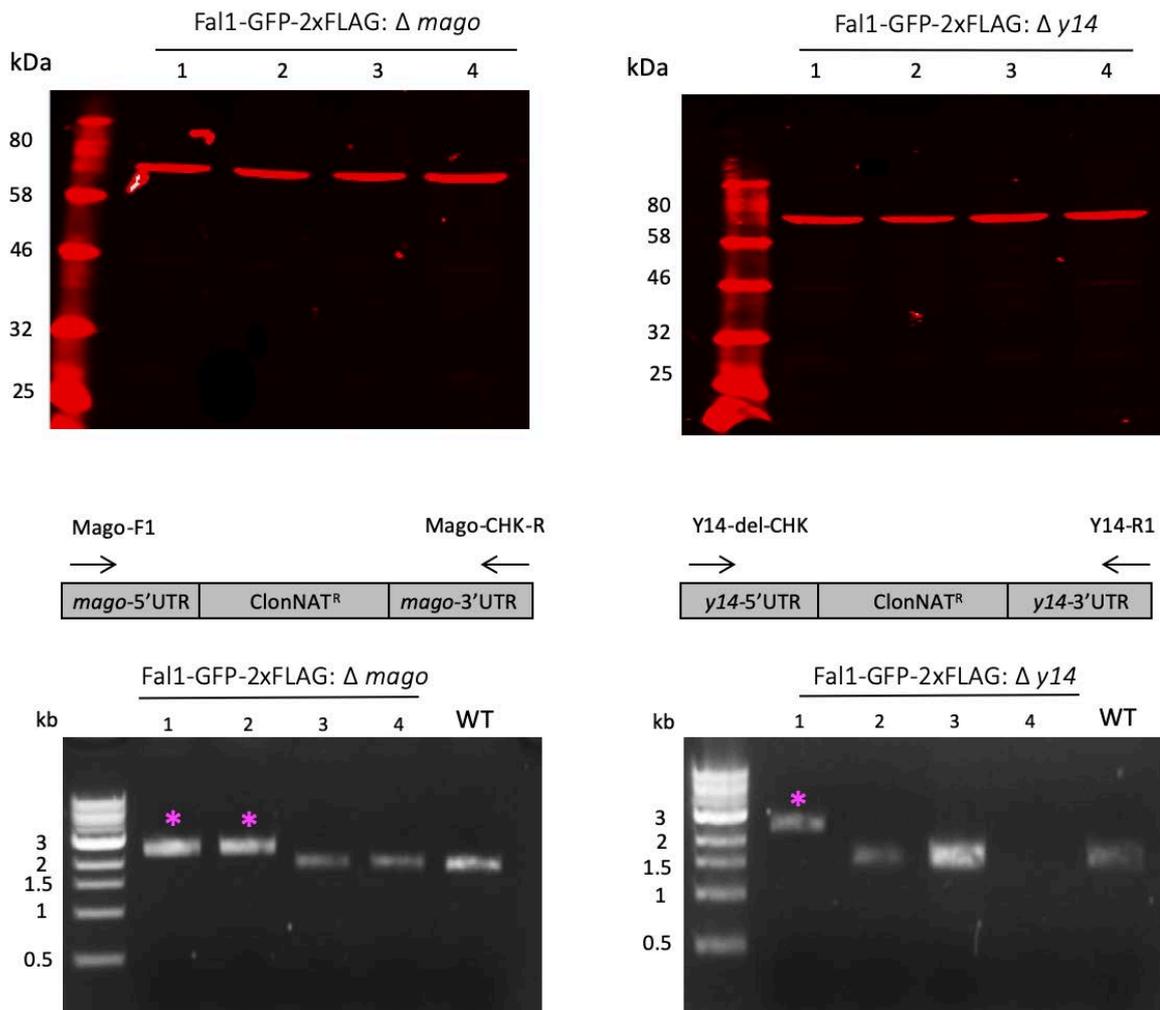
Figure 5.10 - Deletion of *mago* and *y14* from Fal1-GFP-2xFLAG tagged strain. Confirmation by PCR of the deletion of *mago* and *y14* in a strain in which Fal1 was tagged at the C-terminal with a GFP-2xFLAG epitope is performed with the primer pairs annotated in the schematic representation. Positive colonies for the deletion are indicated with a Magenta asterisk. Note that the WB is qualitative and aimed to confirm the presence of the FLAG-tag in the positive colonies; therefore, the membrane was not probed with a loading control antibody.
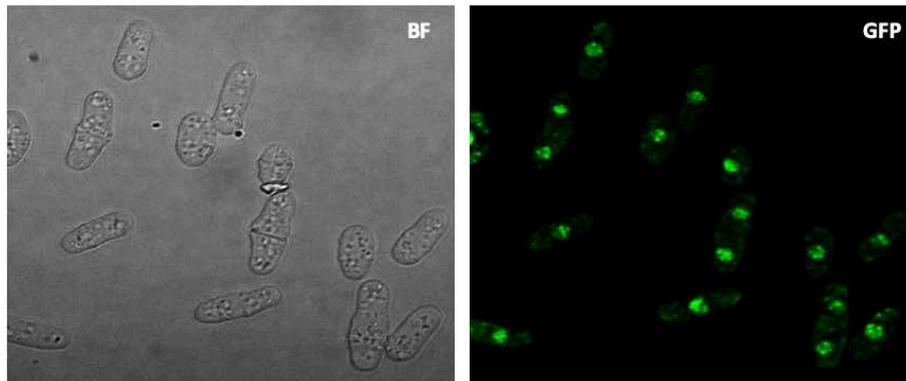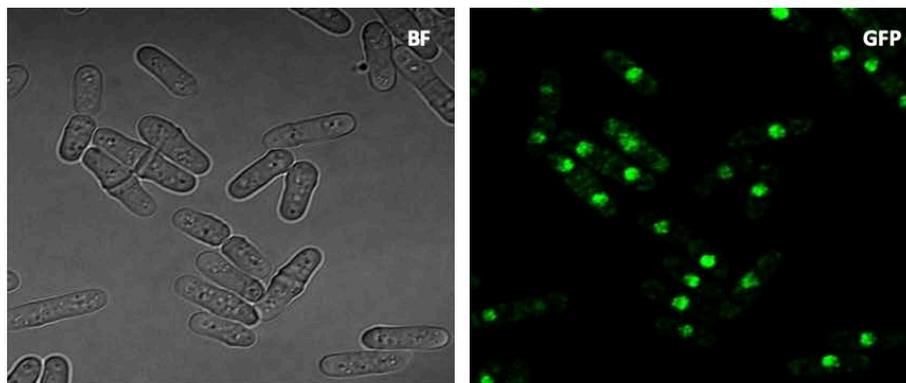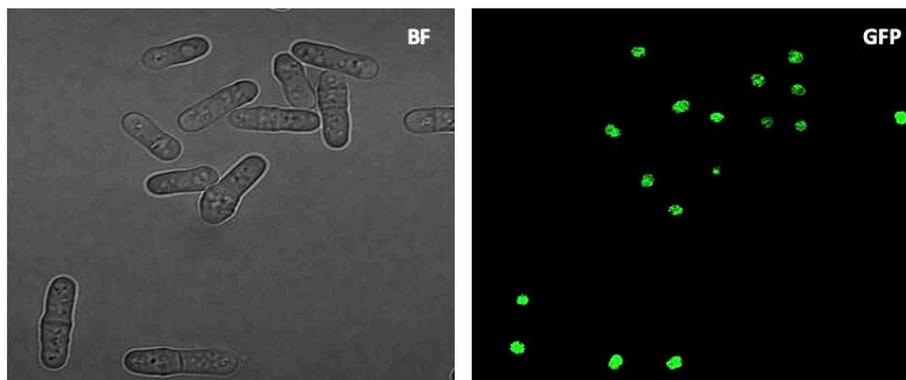
Figure 5.11 - Deletion of *mago* and *y14* do not affect Fal1 nuclear localisation. The GFP signal of Fal1, with and without deletion of *mago* and *y14*, was visualized by confocal microscopy (Nikon A1R). The deletion of *mago* and *y14* does not affect the Fal1 nuclear localisation.
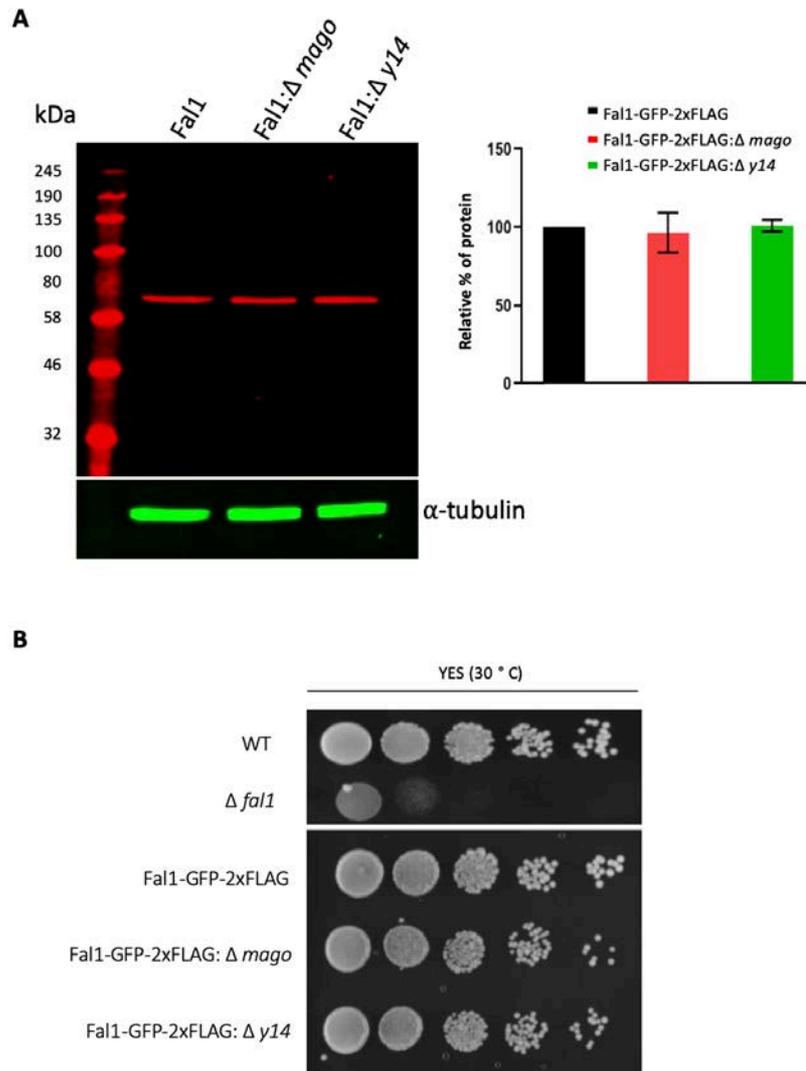
Figure 5.12 - Deletion of *mago* and *y14* do not affect Fal1 protein stability and functionality. (A) WB analysis of the protein levels of Fal1-GFP-2xFLAG, Fal1-GFP-2xFLAG:Δ*mago* and Fal1-GFP-2xFLAG:Δ*y14*. The quantification of the bands of interest is normalized to the loading control (α-tubulin), and expressed as % of protein relative to the Fal1-GFP-2xFLAG strain, which is set at 100%. Error bars represent mean ± SD (n=3). Quantification has been performed using ImageJ.

(B) Spot growth assay comparing the growth of Fal1-GFP-2xFLAG, Fal1-GFP-2xFLAG: Δ*mago* and Fal1-GFP-2xFLAG:Δ*y14*. Diluted spots contain approximately 10^5, 10^4, 10^3, 10^2, 10 cells. Plates are incubated at 30°C and survival is evaluated after 72 hours.

5.2.7 - Depletion of Fal1 might be associated with abnormalities in rRNA processing

The observation that the deletion of *fal1*, but not *mago* and *y14*, affected cell viability, lead me to hypothesize that the protein could play a critical biological function independently of MAGO and Y14. In *S. cerevisiae,* Fal1 localises also in the nucleolus and its depletion leads to a reduction of the steady-state 18S rRNA levels, and as a consequence a decrease of the 40S ribosomal subunit (Kressler et al., 1997). The function appears evolutionarily conserved, since the Human and *D. melanogaster* orthologs of eIF4AIII, can complement the phenotype of the *Fal1* deletion in *S. cerevisiae* (Alexandrov et al., 2011). Moreover, the knock-down of eIF4AIII correlates with defects in ribosomal biogenesis also in mammalian cells (Alexandrov et al., 2011, Badertscher et al., 2015, Mao, 2016, Tafforeau et al., 2013, Wild et al., 2010, Zhang et al., 2011)

To evaluate whether the role of Fal1 in the biogenesis of the 18S rRNA is conserved in *S. pombe*, I performed, with help of Dr. Pawel Grzechnik, Principal Investigator at the University of Birmingham, a preliminary Northern Blot analysis, upon Fal1 thiamine-repression using the strain previously characterized in Chapter 3; interestingly, I found that Fal1 down-regulation associated with a reduction of the steady-state 18S rRNA levels in fission yeast (Figure 5.13). However, the molecular mechanism behind this reduction needs to be better investigated in the future.

A

nmt81:Fal1

| - | + | Thi |

28S
18S

LMW

B

nmt81:Fal1

| - | + | Thi |

18S

7SL
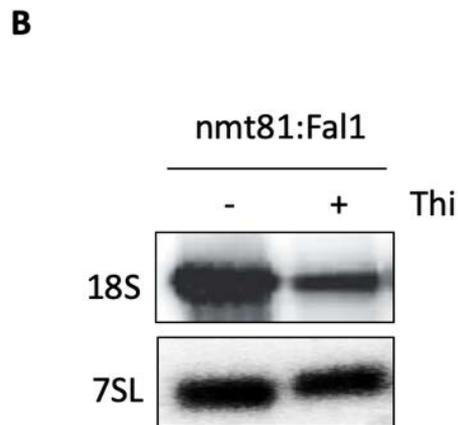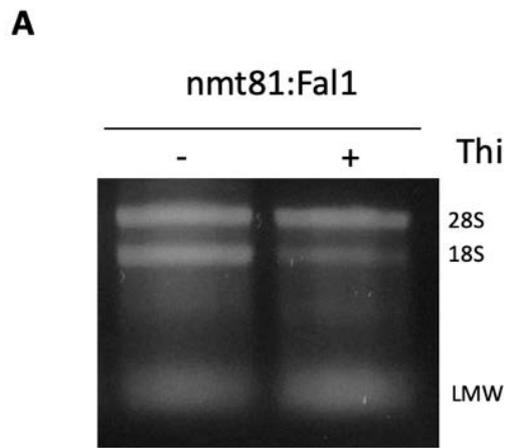
Figure 5.13 - Depletion of Fal1 is associated with the reduction of the steady-state 18S rRNA levels. (A) Ethidium bromide-stained 1% non-denaturing agarose gel of total RNA (5 μg) isolated from yeast cells in which Fal1 was repressed by thiamine for 13 hours, or not repressed (t=0) (B) Northern blot analysis of the 18S rRNA levels, upon depletion of Fal1. The RNA of the signal recognition particle 7SL was used as a loading control.

5.2 - Discussion

The results reported in this chapter are in agreement with previous studies that have argued against the formation of a mammalian-like EJC complex in fission yeast (Marayati et al., 2016, Wen and Brogna, 2010).

To study whether an EJC is formed in *S. pombe,* I firstly reasoned in terms of conservation asking whether MAGO and Y14 are conserved in fission yeast. For instance, it is possible that differences in the protein structures compared to other eukaryotic counterparts may affect the formation of a putative EJC in fission yeast. Instead, I found that MAGO is highly conserved across eukaryotes in terms of primary, secondary and tertiary structure (Figure 5.0). Also Y14 appears conserved*,* despite lacking a long N-terminal segment in *S. pombe*, which in the mammalian counterpart contains numerous residues of interaction with MAGO (Figure 5.1) (Fribourg et al., 2003). However, despite the absence of the N-terminal segment, my data suggest that it is likely that MAGO and Y14 form a stable heterodimer in *S. pombe;* in support of it, I found that the deletion of *y14* reduced the protein levels of MAGO and vice-versa, suggesting that in fission yeast the two factors are required for the stability of each other (Figure 5.6 and Figure 5.7). MAGO and Y14 are mainly localised in the nucleus of *S. pombe* as reported in Chapter 4; however, the deletion of *mago* or *y14,* causes a reduction of the nuclear signals of the respective partners. In mammalian cells and *D. melanogaster* the inter-dependency of MAGO and Y14 is the result of a change in the protein stability of the two factors, rather than a variation in the mRNA levels (Choudhury et al., 2016, Miller, 2017). Therefore, it will be essential in the future to verify the molecular mechanism of this phenomenon in fission yeast.

Despite the fact that MAGO and Y14 are conserved and might form a stable complex in *S. pombe,* an observation which argues against the formation of an EJC in fission yeast, is that

the deletion of *mago* and *y14* does not impair cell viability, contrary to *fal1* whose deletion is associated with severe growth defects. However, despite appearing reasonable that the main function of Fal1 is independent of MAGO and Y14, we cannot fully rule out that an EJC is formed in *S. pombe.*

However, another piece of evidence which argues against the formation of an EJC, is that MAGO and Y14 fractionate separately from Fal1 on sucrose gradient fractionation. The EJC model conjectures that the EJC components remain associated with spliced mRNAs until they are displaced by the ribosome during the first round of translation (Dostie and Dreyfuss, 2002). Contrarily, I found not only that MAGO and Y14 remain on top of the gradient, corresponding to free proteins or small molecular weight complexes, but also that Fal1 associated both with monosomal and polysomal fractions, as previously reported in *D. melanogaster* (Choudhury et al., 2016). This distribution indicates that MAGO and Y14 might dissociate from the mRNA before the translation initiation and it is not likely that they form a stable complex with Fal1 on translating mRNA in fission yeast. The lack of association with translating ribosomes also suggests that MAGO and Y14 are not involved in translation as in mammalian cells (Nott et al., 2004). Moreover, since MAGO and Y14 are found only in the light fraction, it is likely that they disengage from the transcripts either in the nucleus or as soon as they pass through the nuclear pore complexes, before the mRNA is bound by the ribosome. This mode of binding suggests that the two proteins might escort the transcript to the cytoplasm, perhaps to determine its cytoplasmic localisation, as previously reported in *D. melanogaster* (Hachet and Ephrussi, 2004, Palacios, 2002). This hypothesis will be tested through a poly(A) RNAs *in situ* hybridization, to verify whether the deletion of MAGO and Y14 triggers the accumulation of mature transcripts in the nucleus.

The fact that the deletion of *fal1* affects cell viability, unlike *mago* and *y14* (Figure 5.2), indicates that Fal1 is not inter-dependent with MAGO and Y14. My data indicate that deletions of *mago* and *y14* do not affect neither the protein levels nor the nuclear distribution (Figure 5.11 and Figure 5.12) indicating that Fal1 exists in *S. pombe* independently of other EJC components, as reported in mammals and *D. melanogaster* (Alexandrov et al., 2011, Choudhury et al., 2016, Miller, 2017). Moreover, a preliminary analysis it also suggests that the severe growth defect derived by the deletion of *fal1* is not due to the EJC disassembly, but perhaps to a reduction in the steady-state levels of the 18S rRNA (Figure 5.13), which is a phenotype observed in other organisms (Alexandrov et al., 2011, Badertscher et al., 2015, Kressler et al., 1997, Mao, 2016, Tafforeau et al., 2013, Wild et al., 2010, Zhang et al., 2011). However, the molecular mechanism by which this defect occurs needs to be explored in the future. In conclusion, my observations argue against the formation of an EJC. This is in agreement with previous studies in *S. pombe* (Marayati et al., 2016, Wen and Brogna, 2010), as well as in Drosophila (Choudhury et al., 2016).

Chapter 6: Conclusions and Discussion

6.1 - Summary of this project

Although a widely held consensus is that the EJC components are deposited on spliced mRNAs, 20-24 nucleotides upstream of the exon-exon junctions, new compelling evidence suggests that the EJC factors might bind mRNAs regardless of the presence of an intron in the corresponding genes, as discussed in the introduction. The following have driven the work of this PhD project: (I) evaluating whether eIF4AIII and other members of the EJC are associated at transcriptional active loci independently of splicing; (II) understanding the direct function of eIF4AIII at this transcriptionally active site by engineering of the Auxin-inducible Degron (AID) system for the rapid depletion of eIF4AIII; (III) clarifying whether a mammalian-like EJC is formed in fission yeast.

6.2 - Fal1 is associated with active gene loci regardless of the presence of an intron

Recent observations show that the EJC factors might be associated co-transcriptionally on the mRNA independently of introns, both in mammalian cells and *D. melanogaster* (Akhtar et al., 2019, Choe et al., 2014, Choudhury et al., 2016, Viphakone et al., 2019). Therefore, the question at the start of this project was whether in *S. pombe* the EJC factors also associate at active gene loci transcribed by RNA Pol II, and if so, whether the association is splicing independent. This question has partly been answered by ChIP assay using two genes carrying either an intron or intronless as targets. The main conclusion of this work is that Fal1 is recruited on both the intron-containing and intron-less gene, and interestingly without showing a preferential association for the latter. My data indicate that the binding of Fal1 to

172

chromatin is partially sensitive to the RNase treatment thus in part this reflects association with the nascent transcript. This observation is consistent with eIF4AIII being found associated with RNA Pol II transcription sites in Drosophila polytene chromosomes (Choudhury et al., 2016). It is possible that Fal1 might interact with the transcription machinery, a hypothesis that seems supported by the fact that in *D. melanogaster* eIF4AIII co-precipitates with RNA Pol II and the interaction is sensitive to RNAse (Singh et al., 2019). It could be essential in the future to evaluate whether Fal1 interacts with RNA Pol II in *S. pombe,* and if so, whether the interaction is RNA dependent. This hypothesis will be tested by Co-IP using an available tool in which Fal1 and Rpb3 are both tagged endogenously in the same background.

According to the EJC model, eIF4AIII acts as an anchor point for the binding of MAGO and Y14 (Bono et al., 2006, Fribourg et al., 2003). Therefore, I decided to assess whether these components exhibited a binding pattern similar to Fal1. Despite a strong nuclear accumulation, my data suggest that Y14 might not associate or at least not stably on the regions in which Fal1 is reproducibly precipitated. Preliminary results indicate a similar behavior for MAGO. This observation suggests either that Fal1 can associate on the chromatin independently of Y14 and MAGO, meaning that an EJC is not constitutively assembled in fission yeast, or that the interaction between the EJC components is transitory and therefore not stable enough to be captured during the fixation time of our ChIP protocol (Choudhury et al., 2016).

6.3 - is there an EJC in *S. pombe* ?

The EJC model predicts that eIF4AIII, MAGO and Y14 are engaged in a stable grip on the RNA (Bono et al., 2006). Since I have shown that the deletion of *fal1* severely affects cell

viability, I have also evaluated whether the deletion of its partners correlated with a similar phenotype. My results indicate instead, that the deletion of *mago* and *y14* does not cause any obvious growth defect, in agreement with previous findings (Marayati et al., 2016). This observation suggests that an EJC may not be formed in fission yeast (Marayati et al., 2016, Wen and Brogna, 2010). Notably, the silent phenotype deriving from the deletion of *mago* and *y14* in fission yeast*,* is divergent from that observed in other eukaryotes, which instead might depend on that disruption of an EJC; for example, in *D. melanogaster,* the depletion of MAGO and Y14 is accompanied by abnormalities in the dorsal-ventral region, and in *C. elegans* a single point mutation of MAGO causes a prominent masculinization during the sex determination (Gong et al., 2014). Also, in *Arabidopsis thaliana*, the knock-down of MAGO and Y14 is associated with a lethal phenotype (Gong et al., 2014, Park et al., 2009) and in mouse haploinsufficiency for MAGO or Y14 cause abnormalities in the cortical development and microcephaly (Mao et al., 2016).

According to the consensus, the EJC components remain associated with spliced mRNAs until they are displaced by the ribosome during the first round of translation. The piece of evidence which supports this view is that Y14 sediments in the monosomal fractions, but not in the polysomal ones in mammalian cells (Dostie and Dreyfuss, 2002). Contrarily, I have found that MAGO and Y14 remain on top of the sucrose gradient, corresponding to free proteins or small molecular weight complexes, a pattern that has been also observed in *D. melanogaster* (Choudhury et al., 2016). This distribution indicates that MAGO and Y14 might dissociate from the mRNA before the translation initiation and that it is not likely that they form a stable complex with Fal1, which instead persists on translating mRNA.

The fact that *mago* and *y14* deletion mutants grow similarly to the WT strain, also suggests that Fal1 is not inter-dependent from MAGO and Y14. In fact, my data indicate that the

deletion of *mago* and *y14* do not affect its protein levels and the nuclear localisation. A conserved function of eIF4AIII, from human to yeast, is the ability to regulate the biogenesis of the 18S rRNA (Alexandrov et al., 2011, Badertscher et al., 2015, Kressler et al., 1997, Mao, 2016, Mao et al., 2016, Tafforeau et al., 2013, Wild et al., 2010, Zhang et al., 2011). My preliminary results suggest that the function is conserved in *S. pombe,* since depletion of Fal1 correlates with a reduction in the steady-state levels of the 18S rRNA. However, the mechanisms behind this process have not been yet investigated. Mechanistically, one possibility is that Fal1, as in *S. cerevisiae* and Human, could exploit its helicase activity to favor some of the endonucleolytic cleavage required for the maturation of the 18S rRNA (Alexandrov et al., 2011, Kressler et al., 1997). However, It may possible that the reduction of the 18S rRNA is an indirect effect; for example, Fal1 might regulate the expression of other genes which are required for the metabolism of the rRNAs or the expression of ribosomal proteins, as recently observed by Mao and colleagues (Mao et al., 2016). Therefore, this mechanism requires to be confirmed in the future.

Since Fal1 is not dependent on MAGO and Y14, I also evaluated whether there is an inter-dependency between the latter in *S. pombe*. Both quantitative and qualitative experiments indicate that the two proteins are required for the stability of each other in fission yeast. Since it has been demonstrated that the stability of the two components depends on their heterodimerisation, it is likely that they form an obligatory interaction in *S. pombe* as in other eukaryotes (Boisramé et al., 2019, Choudhury et al., 2016, Gong et al., 2014, Le Hir et al., 2001a, Miller, 2017). However, it is crucial to evaluate whether this inter-dependency is a result of a change in the stability of the proteins, as observed in other organisms, or due to a change in the mRNA levels, caused for example by a negative feedback on gene expression (Choudhury et al., 2016, Miller, 2017).

6.4 - The AID system causes a decrease in the Fal1 protein levels before the addition of auxin

My PhD research aimed to implement the AID degron system to rapidly deplete eIF4AIII so to allow discrimination between direct and indirect functions of this protein in gene expression as detailed earlier. I have applied an OFF-AID-DSR system for the rapid depletion of Fal1, the homolog of eIF4AIII in fission yeast, which is characterized by the presence of repressible promoter (OFF), the degron system (AID), and an mRNA destabilizing element (DSR). However, my results indicate a substantial reduction of the protein at basal levels, which reflects unfortunately on cell viability, already compromised before the addition of auxin and thiamine.

Although the AID system did not work as expected, since a cellular phenotype is observed even in absence of auxin and thiamine, my data are still relevant because they indicate that TIR1 could be the limiting factor for the application of the AID system in fission yeast. Based on genetic evidence this investigation concludes that growth defects arise only when the AID tagged protein and TIR1 coexist in the same background. Supporting my observation, recent studies pinpointed that an important caveat for the application of the AID system in budding yeast and mammalian cells, is represented by the constitutive expression of TIR1 under the control of a strong promoter: high-levels of expression of TIR1 cause an auxin-independent degradation of the target (Mendoza-Ochoa et al., 2019, Natsume et al., 2016, Yesbolatova et al., 2019). To overcome this issue in view of future applications, two strategies can be adopted; the first is represented by the use of a commercially available TIR1 inhibitor, auxinole, which blocks the association between TIR1 and the AID domain of the target protein, thus suppressing basal degradation (Yesbolatova et al., 2019). A second possibility is to conditionally regulate the expression of TIR1; for instance, the actual promoter of TIR1,

the *adh1* promoter, can be substituted with the rapid uracil-regulatable promoter, which is repressed or activated in almost 30 minutes (Watt et al., 2008).

6.5 - Conclusion and future directions

My data demonstrate that Fal1 can associate at both an intron-containing and an intron-less gene in *S. pombe*. To evaluate the global binding profile of Fal1, ChIP-seq will need to be performed in the future. My experimental evidences also indicate that the binding of Fal1 on the gene loci is partially sensitive to RNase, suggesting that Fal1 can establish interactions with the nascent-RNA or the transcription machinery. However, it could be interesting to clarify whether Fal1 binds the single-exon transcripts using a more direct approach such as Cross-Linking Immunoprecipitation combined with high-throughput sequencing (CLIP-seq) (Ule et al., 2003).

To assess the role of Fal1 on transcription and pre-mRNA processing, a strain that allows its rapid depletion, in a background with tagged RNA Pol II has been generated. Using this tool, we will purify and sequence the newly transcribed RNA with the 4-thiouridine labeling, to evaluate whether Fal1 depletion can affect directly the RNA synthesis and splicing (Mata and Wise, 2017, Rutkowski and Dölken, 2017). Moreover, to investigate whether the effects on transcription are depedent from changes in the RNA Pol II loading or rate, we will isolate the nascent-RNA via purification of the ternary complex formed by DNA-RNA-RNA Pol II, and the position of the polymerase will be captured at single-nucleotide resolution (Churchman and Weissman, 2011). Using this approach, we will evaluate the effect of Fal1 rapid depletion on the density of RNA Pol II across the whole genome of *S. pombe.*

References

ABRUZZI, K. C., LACADIE, S. & ROSBASH, M. 2004. Biochemical analysis of TREX complex recruitment to intronless and intron-containing yeast genes. *The EMBO journal,* 23**,** 2620-2631.

AHN, S. H., KIM, M. & BURATOWSKI, S. 2004. Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. *Molecular cell,* 13**,** 67-76.

AJAMIAN, L., ABEL, K., RAO, S., VYBOH, K., GARCÍA-DE-GRACIA, F., SOTO-RIFO, R., KULOZIK, A. E., GEHRING, N. H. & MOULAND, A. J. 2015. HIV-1 recruits UPF1 but excludes UPF2 to promote nucleocytoplasmic export of the genomic RNA. *Biomolecules,* 5**,** 2808-2839.

AKHTAR, J., KREIM, N., MARINI, F., MOHANA, G., BRÜNE, D., BINDER, H. & ROIGNANT, J.-Y. 2019. Promoter-proximal pausing mediated by the exon junction complex regulates splicing. *Nature communications,* 10**,** 521.

ALEXANDROV, A., COLOGNORI, D. & STEITZ, J. A. 2011. Human eIF4AIII interacts with an eIF4G-like partner, NOM1, revealing an evolutionarily conserved function outside the exon junction complex. *Genes & development,* 25**,** 1078-1090.

AMEUR, A., ZAGHLOOL, A., HALVARDSON, J., WETTERBOM, A., GYLLENSTEN, U., CAVELIER, L. & FEUK, L. 2011. Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. *Nature structural & molecular biology,* 18**,** 1435.

ASHTON-BEAUCAGE, D. & THERRIEN, M. 2011. The exon junction complex: a splicing factor for long intron containing transcripts? *Fly (Austin),* 5**,** 224-33.

ASHTON-BEAUCAGE, D., UDELL, C. M., LAVOIE, H., BARIL, C., LEFRANÇOIS, M., CHAGNON, P., GENDRON, P., CARON-LIZOTTE, O., BONNEIL, É. & THIBAULT, P. 2010. The exon junction complex controls the splicing of MAPK and other long intron-containing transcripts in Drosophila. *Cell,* 143**,** 251-262.

AZZALIN, C. M. & LINGNER, J. 2006. The human RNA surveillance factor UPF1 is required for S phase progression and genome stability. *Current Biology,* 16**,** 433-439.

BADERTSCHER, L., WILD, T., MONTELLESE, C., ALEXANDER, L. T., BAMMERT, L., SARAZOVA, M., STEBLER, M., CSUCS, G., MAYER, T. U. & ZAMBONI, N. 2015. Genome-wide RNAi screening identifies protein modules required for 40S subunit synthesis in human cells. *Cell reports,* 13**,** 2879-2891.

BALLUT, L., MARCHADIER, B., BAGUET, A., TOMASETTO, C., SÉRAPHIN, B. & LE HIR, H. 2005. The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity. *Nature structural & molecular biology,* 12**,** 861.

BANNERMAN, B. P., KRAMER, S., DORRELL, R. G. & CARRINGTON, M. 2018. Multispecies reconstructions uncover widespread conservation, and lineage-specific elaborations in eukaryotic mRNA metabolism. *PloS one,* 13**,** e0192633.

BARBOSA, I., HAQUE, N., FIORINI, F., BARRANDON, C., TOMASETTO, C., BLANCHETTE, M. & LE HIR, H. 2012. Human CWC22 escorts the helicase eIF4AIII to spliceosomes and

promotes exon junction complex assembly. *Nature structural & molecular biology,* 19**,** 983.

BASI, G., SCHMID, E. & MAUNDRELL, K. 1993. TATA box mutations in the Schizosaccharomyces pombe nmt1 promoter affect transcription efficiency but not the transcription start point or thiamine repressibility. *Gene,* 123**,** 131-136.

BENTLEY, D. L. 2014. Coupling mRNA processing with transcription in time and space. *Nat Rev Genet,* 15**,** 163-75.

BESSONOV, S., ANOKHINA, M., WILL, C. L., URLAUB, H. & LÜHRMANN, R. 2008. Isolation of an active step I spliceosome and composition of its RNP core. *Nature,* 452**,** 846.

BIENROTH, S., KELLER, W. & WAHLE, E. 1993. Assembly of a processive messenger RNA polyadenylation complex. *The EMBO journal,* 12**,** 585-594.

BØE, C. A., GARCIA, I., PAI, C.-C., SHAROM, J. R., SKJØLBERG, H. C., BOYE, E., KEARSEY, S., MACNEILL, S. A., TYERS, M. D. & GRALLERT, B. 2008. Rapid regulation of protein activity in fission yeast. *BMC cell biology,* 9**,** 23.

BOISRAMÉ, A., DEVILLERS, H., ONÉSIME, D., BRUNEL, F., POUCH, J., PIOT, M. & NEUVÉGLISE, C. 2019. Exon junction complex components Y14 and Mago still play a role in budding yeast. *Scientific reports,* 9**,** 849.

BONO, F., EBERT, J., LORENTZEN, E. & CONTI, E. 2006. The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA. *Cell,* 126**,** 713-725.

BONO, F., EBERT, J., UNTERHOLZNER, L., GÜTTLER, T., IZAURRALDE, E. & CONTI, E. 2004. Molecular insights into the interaction of PYM with the Mago–Y14 core of the exon junction complex. *EMBO reports,* 5**,** 304-310.

BONO, F. & GEHRING, N. H. 2011. Assembly, disassembly and recycling: the dynamics of exon junction complexes. *RNA biology,* 8**,** 24-29.

BROGNA, S., MCLEOD, T. & PETRIC, M. 2016. The meaning of NMD: translate or perish. *Trends in genetics,* 32**,** 395-407.

BROGNA, S. & WEN, J. 2009. Nonsense-mediated mRNA decay (NMD) mechanisms. *Nature structural & molecular biology,* 16**,** 107.

BUCHWALD, G., EBERT, J., BASQUIN, C., SAULIERE, J., JAYACHANDRAN, U., BONO, F., LE HIR, H. & CONTI, E. 2010. Insights into the recruitment of the NMD machinery from the crystal structure of a core EJC-UPF3b complex. *Proceedings of the National Academy of Sciences,* 107**,** 10050-10055.

BUCHWALD, G., SCHÜSSLER, S., BASQUIN, C., LE HIR, H. & CONTI, E. 2013. Crystal structure of the human eIF4AIII–CWC22 complex shows how a DEAD-box protein is inhibited by a MIF4G domain. *Proceedings of the National Academy of Sciences,* 110**,** E4611-E4618.

BURATOWSKI, S. 2003. The CTD code. *Nature Structural & Molecular Biology,* 10**,** 679.

BURATOWSKI, S. 2009. Progression through the RNA polymerase II CTD cycle. *Mol Cell,* 36**,** 541-6.

CAREY, M. F., PETERSON, C. L. & SMALE, S. T. 2009. Chromatin immunoprecipitation (chip). *Cold Spring Harbor Protocols,* 2009**,** pdb. prot5279.

CARTER, M., LI, S. & WILKINSON, M. 1996. A splicing-dependent regulatory mechanism that detects translation signals. *The EMBO journal,* 15**,** 5965-5975.

CARUTHERS, J. M., JOHNSON, E. R. & MCKAY, D. B. 2000. Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. *Proceedings of the National Academy of Sciences,* 97**,** 13080-13085.

CHAN, C. C., DOSTIE, J., DIEM, M. D., FENG, W., MANN, M., RAPPSILBER, J. & DREYFUSS, G. 2004. eIF4A3 is a novel component of the exon junction complex. *Rna,* 10**,** 200-9.

CHANG, J. C. & KAN, Y. W. 1979. beta 0 thalassemia, a nonsense mutation in man. *Proceedings of the National Academy of Sciences of the United States of America,* 76**,** 2886-2889.

CHANG, Y.-F., IMAM, J. S. & WILKINSON, M. F. 2007. The nonsense-mediated decay RNA surveillance pathway. *Annu. Rev. Biochem.,* 76**,** 51-74.

CHAPMAN, E. J. & ESTELLE, M. 2009. Mechanism of auxin-regulated gene expression in plants. *Annual review of genetics,* 43**,** 265-285.

CHODER, M. 2004. Rpb4 and Rpb7: subunits of RNA polymerase II and beyond. *Trends in biochemical sciences,* 29**,** 674-681.

CHOE, J., RYU, I., PARK, O. H., PARK, J., CHO, H., YOO, J. S., CHI, S. W., KIM, M. K., SONG, H. K. & KIM, Y. K. 2014. eIF4AIII enhances translation of nuclear cap-binding complex–bound mRNAs by promoting disruption of secondary structures in 5'UTR. *Proceedings of the National Academy of Sciences,* 111**,** E4577-E4586.

CHOUDHURY, S. R., SINGH, A. K., MCLEOD, T., BLANCHETTE, M., JANG, B., BADENHORST, P., KANHERE, A. & BROGNA, S. 2016. Exon junction complex proteins bind nascent transcripts independently of pre-mRNA splicing in Drosophila melanogaster. *Elife,* 5**,** e19881.

CHURCHMAN, L. S. & WEISSMAN, J. S. 2011. Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature,* 469**,** 368.

CLAMP, M., CUFF, J. & BARTON, G. 1998. JalView–analysis and manipulation of multiple sequence alignments. *EMBnet News,* 5**,** 16-21.

CLAMP, M., CUFF, J., SEARLE, S. M. & BARTON, G. J. 2004. The jalview java alignment editor. *Bioinformatics,* 20**,** 426-427.

COLLAS, P. 2010. The current state of chromatin immunoprecipitation. *Molecular biotechnology,* 45**,** 87-100.

CONTI, E. & IZAURRALDE, E. 2005. Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. *Current opinion in cell biology,* 17**,** 316-325.

CORDEN, J. L. & PATTURAJAN, M. 1997. A CTD function linking transcription to splicing. *Trends in biochemical sciences,* 22**,** 413-416.

CORDIN, O., BANROQUES, J., TANNER, N. K. & LINDER, P. 2006. The DEAD-box protein family of RNA helicases. *Gene,* 367**,** 17-37.

CORTAZAR, M. A., SHERIDAN, R. M., ERICKSON, B., FONG, N., GLOVER-CUTTER, K., BRANNAN, K. & BENTLEY, D. L. 2019. Control of RNA Pol II Speed by PNUTS-PP1 and Spt5 Dephosphorylation Facilitates Termination by a "Sitting Duck Torpedo" Mechanism. *Molecular cell,* 76**,** 896-908. e4.

CRAIGHEAD, J. L., CHANG, W.-H. & ASTURIAS, F. J. 2002. Structure of yeast RNA polymerase II in solution: implications for enzyme regulation and interaction with promoter DNA. *Structure,* 10**,** 1117-1125.

CUFF, J. A., CLAMP, M. E., SIDDIQUI, A. S., FINLAY, M. & BARTON, G. J. 1998. JPred: a consensus secondary structure prediction server. *Bioinformatics (Oxford, England),* 14**,** 892-893.

CULBERTSON, M. R., UNDERBRINK, K. M. & FINK, G. R. 1980. Frameshift suppression in Saccharomyces cerevisiae. II. Genetic properties of group II suppressors. *Genetics,* 95**,** 833-853.

DANG, Y., LOW, W.-K., XU, J., GEHRING, N. H., DIETZ, H. C., ROMO, D. & LIU, J. O. 2009. Inhibition of nonsense-mediated mRNA decay by the natural product pateamine A through eukaryotic initiation factor 4AIII. *Journal of Biological Chemistry,* 284**,** 23613-23621.

DE LA MATA, M., ALONSO, C. R., KADENER, S., FEDEDA, J. P., BLAUSTEIN, M., PELISCH, F., CRAMER, P., BENTLEY, D. & KORNBLIHTT, A. R. 2003. A slow RNA polymerase II affects alternative splicing in vivo. *Molecular cell,* 12**,** 525-532.

DE, S., VARSALLY, W., FALCIANI, F. & BROGNA, S. 2011. Ribosomal proteins' association with transcription sites peaks at tRNA genes in Schizosaccharomyces pombe. *Rna,* 17**,** 1713-1726.

DING, D.-Q., HARAGUCHI, T. & HIRAOKA, Y. 2013. The role of chromosomal retention of noncoding RNA in meiosis. *Chromosome research,* 21**,** 665-672.

DING, L., LAOR, D., WEISMAN, R. & FORSBURG, S. L. 2014. Rapid regulation of nuclear proteins by rapamycin-induced translocation in fission yeast. *Yeast,* 31**,** 253-264.

DOHMEN, R. J., WU, P. & VARSHAVSKY, A. 1994. Heat-inducible degron: a method for constructing temperature-sensitive mutants. *Science,* 263**,** 1273-1276.

DOSTIE, J. & DREYFUSS, G. 2002. Translation is required to remove Y14 from mRNAs in the cytoplasm. *Current Biology,* 12**,** 1060-1067.

EATON, J. D., DAVIDSON, L., BAUER, D. L., NATSUME, T., KANEMAKI, M. T. & WEST, S. 2018. Xrn2 accelerates termination by RNA polymerase II, which is underpinned by CPSF73 activity. *Genes & development,* 32**,** 127-139.

EATON, J. D., FRANCIS, L., DAVIDSON, L. & WEST, S. 2020. A unified allosteric/torpedo mechanism for transcriptional termination on human protein-coding genes. *Genes & Development,* 34**,** 132-145.

EATON, J. D. & WEST, S. 2018. An end in sight? Xrn2 and transcriptional termination by RNA polymerase II. *Transcription,* 9**,** 321-326.

EBRIGHT, R. H. 2000. RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. *Journal of molecular biology,* 304**,** 687-698.

ERKMANN, J. A. & KUTAY, U. 2004. Nuclear export of mRNA: from the site of transcription to the cytoplasm. *Experimental cell research,* 296**,** 12-20.

FAVARO, F. P., ALVIZI, L., ZECHI-CEIDE, R. M., BERTOLA, D., FELIX, T. M., DE SOUZA, J., RASKIN, S., TWIGG, S. R., WEINER, A. M. & ARMAS, P. 2014. A noncoding expansion in EIF4A3 causes Richieri-Costa-Pereira syndrome, a craniofacial disorder associated with limb defects. *The American Journal of Human Genetics,* 94**,** 120-128.

FERRAIUOLO, M. A., LEE, C.-S., LER, L. W., HSU, J. L., COSTA-MATTIOLI, M., LUO, M.-J., REED, R. & SONENBERG, N. 2004. A nuclear translation-like factor eIF4AIII is recruited to the mRNA during splicing and functions in nonsense-mediated decay. *Proceedings of the National Academy of Sciences,* 101**,** 4118-4123.

FIORINI, F., BAGCHI, D., LE HIR, H. & CROQUETTE, V. 2015. Human Upf1 is a highly processive RNA helicase and translocase with RNP remodelling activities. *Nature communications,* 6**,** 1-10.

FONG, N. & BENTLEY, D. L. 2001. Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. *Genes & Development,* 15**,** 1783-1795.

FORSBURG, S. L. 2003. S. pombe strain maintenance and media. *Current protocols in molecular biology,* 64**,** 13.15. 1-13.15. 5.

FRIBOURG, S., GATFIELD, D., IZAURRALDE, E. & CONTI, E. 2003. A novel mode of RBD-protein recognition in the Y14–Mago complex. *Nature Structural & Molecular Biology,* 10**,** 433.

FURUICHI, Y., MUTHUKRISHNAN, S., TOMASZ, J. & SHATKIN, A. 1976. Mechanism of formation of reovirus mRNA 5'-terminal blocked and methylated sequence, m7GpppGmpC. *Journal of Biological Chemistry,* 251**,** 5043-5053.

GATFIELD, D. & IZAURRALDE, E. 2002. REF1/Aly and the additional exon junction complex proteins are dispensable for nuclear mRNA export. *J Cell Biol,* 159**,** 579-588.

GATFIELD, D., UNTERHOLZNER, L., CICCARELLI, F. D., BORK, P. & IZAURRALDE, E. 2003. Nonsense-mediated mRNA decay in Drosophila: at the intersection of the yeast and mammalian pathways. *The EMBO journal,* 22**,** 3960-3970.

GEHRING, N. H., LAMPRINAKI, S., HENTZE, M. W. & KULOZIK, A. E. 2009a. The hierarchy of exon-junction complex assembly by the spliceosome explains key features of mammalian nonsense-mediated mRNA decay. *PLoS biology,* 7**,** e1000120-e1000120.

GEHRING, N. H., LAMPRINAKI, S., KULOZIK, A. E. & HENTZE, M. W. 2009b. Disassembly of exon junction complexes by PYM. *Cell,* 137**,** 536-548.

GEHRING, N. H., NEU-YILIK, G., SCHELL, T., HENTZE, M. W. & KULOZIK, A. E. 2003. Y14 and hUpf3b form an NMD-activating complex. *Mol Cell,* 11**,** 939-49.

GIETZ, R. D. & WOODS, R. A. 2002. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods in enzymology.* Elsevier.

GILMOUR, D. S. & LIS, J. T. 1984. Detecting protein-DNA interactions in vivo: distribution of RNA polymerase on specific bacterial genes. *Proceedings of the National Academy of Sciences,* 81**,** 4275-4279.

GILMOUR, D. S. & LIS, J. T. 1985. In vivo interactions of RNA polymerase II with genes of Drosophila melanogaster. *Molecular and cellular biology,* 5**,** 2009-2018.

GONG, P., ZHAO, M. & HE, C. 2014. Slow co-evolution of the MAGO and Y14 protein families is required for the maintenance of their obligate heterodimerization mode. *PLoS One,* 9**,** e84842.

GONZÁLEZ, C. I., RUIZ-ECHEVARRÍA, M. J., VASUDEVAN, S., HENRY, M. F. & PELTZ, S. W. 2000. The yeast hnRNP-like protein Hrp1/Nab4 marks a transcript for nonsense-mediated mRNA decay. *Molecular cell,* 5**,** 489-499.

GRAY, W. M., KEPINSKI, S., ROUSE, D., LEYSER, O. & ESTELLE, M. 2001. Auxin regulates SCF TIR1-dependent degradation of AUX/IAA proteins. *Nature,* 414**,** 271.

GREEN, M. R. & SAMBROOK, J. 2012. Molecular cloning. *A Laboratory Manual 4th*.

HACHET, O. & EPHRUSSI, A. 2001. Drosophila Y14 shuttles to the posterior of the oocyte and is required for oskar mRNA transport. *Current Biology,* 11**,** 1666-1674.

HACHET, O. & EPHRUSSI, A. 2004. Splicing of oskar RNA in the nucleus is coupled to its cytoplasmic localization. *Nature,* 428**,** 959.

HAREMAKI, T., SRIDHARAN, J., DVORA, S. & WEINSTEIN, D. C. 2010. Regulation of vertebrate embryogenesis by the exon junction complex core component Eif4a3. *Developmental Dynamics,* 239**,** 1977-1987.

HARIGAYA, Y., TANAKA, H., YAMANAKA, S., TANAKA, K., WATANABE, Y., TSUTSUMI, C., CHIKASHIGE, Y., HIRAOKA, Y., YAMASHITA, A. & YAMAMOTO, M. 2006. Selective elimination of messenger RNA prevents an incidence of untimely meiosis. *Nature,* 442**,** 45-50.

HARUKI, H., NISHIKAWA, J. & LAEMMLI, U. K. 2008. The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes. *Molecular cell,* 31**,** 925-932.

HENTZE, M. W. & KULOZIK, A. E. 1999. A perfect message: RNA surveillance and nonsense-mediated decay. *Cell,* 96**,** 307-310.

HILL, C. H., BOREIKAITĖ, V., KUMAR, A., CASAÑAL, A., KUBÍK, P., DEGLIESPOSTI, G., MASLEN, S., MARIANI, A., VON LOEFFELHOLZ, O. & GIRBIG, M. 2019. Activation of the endonuclease that defines mRNA 3' ends requires incorporation into an 8-subunit core cleavage and polyadenylation factor complex. *Molecular cell,* 73**,** 1217-1231. e11.

HSIN, J.-P. & MANLEY, J. L. 2012. The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes & development,* 26**,** 2119-2137.

ITO, M., IWATANI-YOSHIHARA, M., TANAKA, T., CARY, D. R., KAMADA, Y., ISHIBASHI, Y., OKI, H., SOGABE, S., NAKAO, S. & MORISHITA, D. 2018. Discovery of the first selective and cell-active eIF4A3 inhibitors. AACR.

ITO, M., TANAKA, T., CARY, D. R., IWATANI-YOSHIHARA, M., KAMADA, Y., KAWAMOTO, T., APARICIO, S., NAKANISHI, A. & IMAEDA, Y. 2017. Discovery of novel 1, 4-diacylpiperazines as selective and cell-active eIF4A3 inhibitors. *Journal of medicinal chemistry,* 60**,** 3335-3351.

IWAMOTO, M., BJÖRKLUND, T., LUNDBERG, C., KIRIK, D. & WANDLESS, T. J. 2010. A general chemical method to regulate protein stability in the mammalian central nervous system. *Chemistry & biology,* 17**,** 981-988.

IWATANI-YOSHIHARA, M., ITO, M., ISHIBASHI, Y., OKI, H., TANAKA, T., MORISHITA, D., ITO, T., KIMURA, H., IMAEDA, Y. & APARICIO, S. 2017. Discovery and characterization of a eukaryotic initiation factor 4a-3-selective inhibitor that suppresses nonsense-mediated mrna decay. *ACS chemical biology,* 12**,** 1760-1768.

KANKE, M., NISHIMURA, K., KANEMAKI, M., KAKIMOTO, T., TAKAHASHI, T. S., NAKAGAWA, T. & MASUKATA, H. 2011. Auxin-inducible protein depletion system in fission yeast. *BMC cell biology,* 12**,** 8.

KAROUSIS, E. D. & MÜHLEMANN, O. 2019. Nonsense-mediated mRNA decay begins where translation ends. *Cold Spring Harbor perspectives in biology,* 11**,** a032862.

KASHIMA, I., YAMASHITA, A., IZUMI, N., KATAOKA, N., MORISHITA, R., HOSHINO, S., OHNO, M., DREYFUSS, G. & OHNO, S. 2006. Binding of a novel SMG-1–Upf1–eRF1–eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes & development,* 20**,** 355-367.

KATAHIRA, J. 2012. mRNA export and the TREX complex. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms,* 1819**,** 507-513.

KATAOKA, N., DIEM, M. D., YOSHIDA, M., HATAI, C., DOBASHI, I., DREYFUSS, G., HAGIWARA, M. & OHNO, M. 2011. Specific Y14 domains mediate its nucleo-cytoplasmic shuttling and association with spliced mRNA. *Scientific reports,* 1**,** 92.

KEPINSKI, S. & LEYSER, O. 2005. The Arabidopsis F-box protein TIR1 is an auxin receptor. *Nature,* 435**,** 446.

KIM, K. M., CHO, H., CHOI, K., KIM, J., KIM, B.-W., KO, Y.-G., JANG, S. K. & KIM, Y. K. 2009. A new MIF4G domain-containing protein, CTIF, directs nuclear cap-binding protein CBP80/20-dependent translation. *Genes & development,* 23**,** 2033-2045.

KITTLER, R., PUTZ, G., PELLETIER, L., POSER, I., HENINGER, A. K., DRECHSEL, D., FISCHER, S., KONSTANTINOVA, I., HABERMANN, B., GRABNER, H., YASPO, M. L., HIMMELBAUER, H., KORN, B., NEUGEBAUER, K., PISABARRO, M. T. & BUCHHOLZ, F. 2004. An endoribonuclease-prepared siRNA screen in human cells identifies genes essential for cell division. *Nature,* 432**,** 1036-40.

KOMARNITSKY, P., CHO, E.-J. & BURATOWSKI, S. 2000. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes & development,* 14**,** 2452-2460.

KOROLEVA, O., BROWN, J. W. & SHAW, P. E. 2009. Localization of eIF4A-III in the nucleolus and splicing speckles is an indicator of plant stress. *Plant signaling & behavior,* 4**,** 1148-1151.

KRAMER, A. 1996. The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annu Rev Biochem,* 65**,** 367-409.

KRAWCHUK, M. D. & WAHLS, W. P. 1999. High-efficiency gene targeting in Schizosaccharomyces pombe using a modular, PCR-based approach with long tracts of flanking homology. *Yeast,* 15**,** 1419-1427.

KRESSLER, D., DE LA CRUZ, J., ROJO, M. & LINDER, P. 1997. Fal1p is an essential DEAD-box protein involved in 40S-ribosomal-subunit biogenesis in Saccharomyces cerevisiae. *Mol Cell Biol,* 17**,** 7283-94.

KUHN, A. N. & KÄUFER, N. F. 2003. Pre-mRNA splicing in Schizosaccharomyces pombe. *Current genetics,* 42**,** 241-251.

KUMAR, R. & SINGH, J. 2006. A truncated derivative of nmt1 promoter exhibits temperature-dependent induction of gene expression in Schizosaccharomyces pombe. *Yeast,* 23**,** 55-65.

LAROCHELLE, M., ROBERT, M.-A., HÉBERT, J.-N., LIU, X., MATTEAU, D., RODRIGUE, S., TIAN, B., JACQUES, P.-É. & BACHAND, F. 2018. Common mechanism of transcription termination at coding and noncoding RNA genes in fission yeast. *Nature communications,* 9**,** 4364.

LAU, C.-K., DIEM, M. D., DREYFUSS, G. & VAN DUYNE, G. D. 2003. Structure of the Y14-Magoh core of the exon junction complex. *Current Biology,* 13**,** 933-941.

LE HIR, H. & ANDERSEN, G. R. 2008. Structural insights into the exon junction complex. *Current opinion in structural biology,* 18**,** 112-119.

LE HIR, H., GATFIELD, D., BRAUN, I. C., FORLER, D. & IZAURRALDE, E. 2001a. The protein Mago provides a link between splicing and mRNA localization. *EMBO reports,* 2**,** 1119-1124.

LE HIR, H., GATFIELD, D., IZAURRALDE, E. & MOORE, M. J. 2001b. The exon–exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *The EMBO journal,* 20**,** 4987-4997.

LE HIR, H., IZAURRALDE, E., MAQUAT, L. E. & MOORE, M. J. 2000. The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon–exon junctions. *The EMBO journal,* 19**,** 6860-6869.

LE HIR, H., NOTT, A. & MOORE, M. J. 2003. How introns influence and enhance eukaryotic gene expression. *Trends in biochemical sciences,* 28**,** 215-220.

LE HIR, H., SAULIÈRE, J. & WANG, Z. 2016. The exon junction complex as a node of post-transcriptional networks. *Nature reviews Molecular cell biology,* 17**,** 41.

LEJEUNE, F. & MAQUAT, L. E. 2005. Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. *Current opinion in cell biology,* 17**,** 309-315.

LI, Q., IMATAKA, H., MORINO, S., ROGERS, G. W., JR., RICHTER-COOK, N. J., MERRICK, W. C. & SONENBERG, N. 1999. Eukaryotic translation initiation factor 4AIII (eIF4AIII) is functionally distinct from eIF4AI and eIF4AII. *Mol Cell Biol,* 19**,** 7336-46.

LINDER, P. & JANKOWSKY, E. 2011. From unwinding to clamping — the DEAD box RNA helicase family. *Nature Reviews Molecular Cell Biology,* 12**,** 505.

LOSSON, R. & LACROUTE, F. 1979. Interference of nonsense mutations with eukaryotic messenger RNA stability. *Proceedings of the National Academy of Sciences of the United States of America,* 76**,** 5134-5137.

LOYA, T. J. & REINES, D. 2016. Recent advances in understanding transcription termination by RNA polymerase II. *F1000Research,* 5.

MANDEL, C. R., KANEKO, S., ZHANG, H., GEBAUER, D., VETHANTHAM, V., MANLEY, J. L. & TONG, L. 2006. Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature,* 444**,** 953-956.

MAO, H. 2016. *Posttranscriptional Regulation of Embryonic Neurogenesis by the Exon Junction Complex.* Duke University.

MAO, H., MCMAHON, J. J., TSAI, Y.-H., WANG, Z. & SILVER, D. L. 2016. Haploinsufficiency for core exon junction complex components disrupts embryonic neurogenesis and causes p53-mediated microcephaly. *PLoS genetics,* 12**,** e1006282.

MAQUAT, L. E. 1995. When cells stop making sense: effects of nonsense codons on RNA metabolism in vertebrate cells. *Rna,* 1**,** 453-465.

MAQUAT, L. E. 2004. Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nature reviews Molecular cell biology,* 5**,** 89.

MAQUAT, L. E. 2005. Nonsense-mediated mRNA decay in mammals. *Journal of cell science,* 118**,** 1773-1776.

MAQUAT, L. E. & LI, X. 2001. Mammalian heat shock p70 and histone H4 transcripts, which derive from naturally intronless genes, are immune to nonsense-mediated decay. *Rna,* 7**,** 445-56.

MAQUAT, L. E., TARN, W. Y. & ISKEN, O. 2010. The pioneer round of translation: features and functions. *Cell,* 142**,** 368-74.

MARAYATI, B. F., HOSKINS, V., BOGER, R. W., TUCKER, J. F., FISHMAN, E. S., BRAY, A. S. & ZHANG, K. 2016. The fission yeast MTREC and EJC orthologs ensure the maturation of meiotic transcripts during meiosis. *RNA,* 22**,** 1349-1359.

MARGUERAT, S., SCHMIDT, A., CODLIN, S., CHEN, W., AEBERSOLD, R. & BÄHLER, J. 2012. Quantitative analysis of fission yeast transcriptomes and proteomes in proliferating and quiescent cells. *Cell,* 151**,** 671-683.

MARTINEZ-RUCOBO, F. W., KOHLER, R., VAN DE WATERBEEMD, M., HECK, A. J., HEMANN, M., HERZOG, F., STARK, H. & CRAMER, P. 2015. Molecular basis of transcription-coupled pre-mRNA capping. *Molecular cell,* 58**,** 1079-1089.

MATA, J. & WISE, J. A. 2017. 4-thiouridine labeling to analyze mRNA turnover in Schizosaccharomyces pombe. *Cold Spring Harbor Protocols,* 2017**,** pdb. prot091645.

MATSUO, Y., ASAKAWA, K., TODA, T. & KATAYAMA, S. 2006. A rapid method for protein extraction from fission yeast. *Bioscience, biotechnology, and biochemistry,* 70**,** 1992-1994.

MAUNDRELL, K. 1990. nmt1 of fission yeast. A highly transcribed gene completely repressed by thiamine. *Journal of Biological Chemistry,* 265**,** 10857-10864.

MAUNDRELL, K. 1993. Thiamine-repressible expression vectors pREP and pRIP for fission yeast. *Gene,* 123**,** 127-130.

MAZLOOMIAN, A., ARAKI, S., OHORI, M., EL-NAGGAR, A. M., YAP, D., BASHASHATI, A., NAKAO, S., SORENSEN, P. H., NAKANISHI, A. & SHAH, S. 2019. Pharmacological systems analysis defines EIF4A3 functions in cell-cycle and RNA stress granule formation. *Communications biology,* 2**,** 165.

MENDOZA-OCHOA, G. I., BARRASS, J. D., TERLOUW, B. R., MAUDLIN, I. E., DE LUCAS, S., SANI, E., ASLANZADEH, V., REID, J. A. & BEGGS, J. D. 2019. A fast and tuneable auxin-inducible degron for depletion of target proteins in budding yeast. *Yeast,* 36**,** 75-81.

MILLER, E. E. 2017. *The Discovery of EJC Independent Roles for EIF4A3 in Mitosis, Microtubules, and Neural Crest Development.* Duke University.

MILLER, E. E., KOBAYASHI, G. S., MUSSO, C. M., ALLEN, M., ISHIY, F. A., DE CAIRES, L. C., GOULART, E., GRIESI-OLIVEIRA, K., ZECHI-CEIDE, R. M. & RICHIERI-COSTA, A. 2017. EIF4A3 deficient human iPSCs and mouse models demonstrate neural crest defects that underlie Richieri-Costa-Pereira syndrome. *Human molecular genetics,* 26**,** 2177-2191.

MIZOJIRI, R., NAKATA, D., SATOH, Y., MORISHITA, D., SHIBATA, S., IWATANI-YOSHIHARA, M., KOSUGI, Y., KOSAKA, M., TAKEDA, J. & SASAKI, S. 2018. Design, synthesis and evaluations of novel class of orally bioavailable eIF4A3-selective inhibitors. AACR.

MOORE, M. J. 2005. From birth to death: the complex lives of eukaryotic mRNAs. *Science,* 309**,** 1514-8.

MOORE, M. J. & PROUDFOOT, N. J. 2009. Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell,* 136**,** 688-700.

MORAWSKA, M. & ULRICH, H. D. 2013. An expanded tool kit for the auxin-inducible degron system in budding yeast. *Yeast,* 30**,** 341-351.

MURRAY, J. M., WATSON, A. T. & CARR, A. M. 2016. Extraction of chromosomal DNA from Schizosaccharomyces pombe. *Cold Spring Harbor Protocols,* 2016**,** pdb. prot090985.

NABET, B., ROBERTS, J. M., BUCKLEY, D. L., PAULK, J., DASTJERDI, S., YANG, A., LEGGETT, A. L., ERB, M. A., LAWLOR, M. A. & SOUZA, A. 2018. The dTAG system for immediate and target-specific protein degradation. *Nature chemical biology,* 14**,** 431-441.

NATSUME, T., KIYOMITSU, T., SAGA, Y. & KANEMAKI, M. T. 2016. Rapid protein depletion in human cells by auxin-inducible degron tagging with short homology donors. *Cell reports,* 15**,** 210-218.

NICHOLSON, P., YEPISKOPOSYAN, H., METZE, S., OROZCO, R. Z., KLEINSCHMIDT, N. & MÜHLEMANN, O. 2010. Nonsense-mediated mRNA decay in human cells: mechanistic insights, functions beyond quality control and the double-life of NMD factors. *Cellular and molecular life sciences,* 67**,** 677-700.

NISHIMURA, K., FUKAGAWA, T., TAKISAWA, H., KAKIMOTO, T. & KANEMAKI, M. 2009. An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nature methods,* 6**,** 917.

NOE GONZALEZ, M., SATO, S., TOMOMORI-SATO, C., CONAWAY, J. W. & CONAWAY, R. C. 2018. CTD-dependent and -independent mechanisms govern co-transcriptional capping of Pol II transcripts. *Nat Commun,* 9**,** 3392.

NOTT, A., LE HIR, H. & MOORE, M. J. 2004. Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. *Genes & development,* 18**,** 210-222.

NUNES, N. M., LI, W., TIAN, B. & FURGER, A. 2010. A functional human Poly (A) site requires only a potent DSE and an A-rich upstream sequence. *The EMBO journal,* 29**,** 1523-1536.

NURSE, P. 1990. Universal control mechanism regulating onset of M-phase. *Nature,* 344**,** 503.

O'BRIEN, T., HARDIN, S., GREENLEAF, A. & LIS, J. T. 1994. Phosphorylation of RNA polymerase II C-terminal domain and transcriptional elongation. *Nature,* 370**,** 75-7.

OESTERREICH, F. C., HERZEL, L., STRAUBE, K., HUJER, K., HOWARD, J. & NEUGEBAUER, K. M. 2016. Splicing of Nascent RNA Coincides with Intron Exit from RNA Polymerase II. *Cell,* 165**,** 372-381.

OTSUBO, Y. & YAMAMATO, M. 2008. TOR signaling in fission yeast. *Critical reviews in biochemistry and molecular biology,* 43**,** 277-283.

PAI, C.-C., SCHNICK, J., MACNEILL, S. A. & KEARSEY, S. E. 2012. Conditional inactivation of replication proteins in fission yeast using hormone-binding domains. *Methods,* 57**,** 227-233.

PALACIOS, I. M. 2002. RNA processing: splicing and the cytoplasmic localisation of mRNA. *Current biology,* 12**,** R50-R52.

PALACIOS, I. M., GATFIELD, D., ST JOHNSTON, D. & IZAURRALDE, E. 2004. An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. *Nature,* 427**,** 753.

PARK, N.-I. & MUENCH, D. G. 2007. Biochemical and cellular characterization of the plant ortholog of PYM, a protein that interacts with the exon junction complex core proteins Mago and Y14. *Planta,* 225**,** 625-639.

PARK, N.-I., YEUNG, E. C. & MUENCH, D. G. 2009. Mago Nashi is involved in meristem organization, pollen formation, and seed development in Arabidopsis. *Plant Science,* 176**,** 461-469.

PEI, Y. & SHUMAN, S. 2002. Interactions between fission yeast mRNA capping enzymes and elongation factor Spt5. *Journal of Biological Chemistry,* 277**,** 19639-19648.

PENDLE, A. F., CLARK, G. P., BOON, R., LEWANDOWSKA, D., LAM, Y. W., ANDERSEN, J., MANN, M., LAMOND, A. I., BROWN, J. W. & SHAW, P. J. 2005. Proteomic analysis of the Arabidopsis nucleolus suggests novel nucleolar functions. *Molecular biology of the cell,* 16**,** 260-269.

PETTERSEN, E. F., GODDARD, T. D., HUANG, C. C., COUCH, G. S., GREENBLATT, D. M., MENG, E. C. & FERRIN, T. E. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry,* 25**,** 1605-1612.

PHATNANI, H. P. & GREENLEAF, A. L. 2006. Phosphorylation and functions of the RNA polymerase II CTD. *Genes & development,* 20**,** 2922-2936.

PILLAI, S. & CHELLAPPAN, S. P. 2009. ChIP on chip assays: genome-wide analysis of transcription factor binding and histone modifications. *Chromatin Protocols.* Springer.

PROUDFOOT, N. 1991. Poly(A) signals. *Cell,* 64**,** 671-4.

PROUDFOOT, N. 2000. Connecting transcription to messenger RNA processing. *Trends in biochemical sciences,* 25**,** 290-293.

PROUDFOOT, N. J., FURGER, A. & DYE, M. J. 2002. Integrating mRNA processing with transcription. *Cell,* 108**,** 501-12.

REED, R. 2003. Coupling transcription, splicing and mRNA export. *Current opinion in cell biology,* 15**,** 326-331.

ROCAK, S. & LINDER, P. 2004. DEAD-box proteins: the driving forces behind RNA metabolism. *Nature reviews Molecular cell biology,* 5**,** 232.

RODRÍGUEZ-LÓPEZ, M., COTOBAL, C., FERNÁNDEZ-SÁNCHEZ, O., BRAVO, N. B., OKTRIANI, R., ABENDROTH, H., UKA, D., HOTI, M., WANG, J. & ZARATIEGUI, M. 2016. A CRISPR/Cas9-based method and primer design tool for seamless genome editing in fission yeast. *Wellcome open research,* 1.

RUTKOWSKI, A. J. & DÖLKEN, L. 2017. High-resolution gene expression profiling of RNA synthesis, processing, and decay by metabolic labeling of newly transcribed RNA using 4-thiouridine. *Eukaryotic Transcriptional and Post-Transcriptional Gene Expression Regulation.* Springer.

SAULIÈRE, J., MURIGNEUX, V., WANG, Z., MARQUENET, E., BARBOSA, I., LE TONQUÈZE, O., AUDIC, Y., PAILLARD, L., CROLLIUS, H. R. & LE HIR, H. 2012. CLIP-seq of eIF4AIII reveals transcriptome-wide mapping of the human exon junction complex. *Nature structural & molecular biology,* 19**,** 1124.

SCHWEDE, T., KOPP, J., GUEX, N. & PEITSCH, M. C. 2003. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic acids research,* 31**,** 3381-3385.

SHARMA, D., SAY, A. F., LEDFORD, L. L., HUGHES, A. J., SEHORN, H. A., DWYER, D. S. & SEHORN, M. G. 2013. Role of the conserved lysine within the Walker A motif of human DMC1. *DNA repair,* 12**,** 53-62.

SHATKIN, A. 1976. Capping of eucaryotic mRNAs. *Cell,* 9**,** 645-653.

SHAUL, O. 2017. How introns enhance gene expression. *The international journal of biochemistry & cell biology,* 91**,** 145-155.

SHI, H. & XU, R.-M. 2003. Crystal structure of the Drosophila Mago nashi–Y14 complex. *Genes & development,* 17**,** 971-976.

SHIBUYA, T., TANGE, T. Ø., SONENBERG, N. & MOORE, M. J. 2004. eIF4AIII binds spliced mRNA in the exon junction complex and is essential for nonsense-mediated decay. *Nature structural & molecular biology,* 11**,** 346.

SINGH, A. K., CHOUDHURY, S. R., DE, S., ZHANG, J., KISSANE, S., DWIVEDI, V., RAMANATHAN, P., PETRIC, M., ORSINI, L. & HEBENSTREIT, D. 2019. The RNA helicase UPF1 associates with mRNAs co-transcriptionally and is required for the release of mRNAs from gene loci. *Elife,* 8**,** e41444.

SINGH, G., KUCUKURAL, A., CENIK, C., LESZYK, J. D., SHAFFER, S. A., WENG, Z. & MOORE, M. J. 2012. The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell,* 151**,** 750-764.

SINGH, G. & LYKKE-ANDERSEN, J. 2003. New insights into the formation of active nonsense-mediated decay complexes. *Trends in biochemical sciences,* 28**,** 464-466.

SINGH, G., PRATT, G., YEO, G. W. & MOORE, M. J. 2015. The clothes make the mRNA: past and present trends in mRNP fashion. *Annual review of biochemistry,* 84**,** 325-354.

SINGH, G., RICCI, E. P. & MOORE, M. J. 2014. RIPiT-Seq: a high-throughput approach for footprinting RNA: protein complexes. *Methods,* 65**,** 320-332.

SMALE, S. T. & TJIAN, R. 1985. Transcription of herpes simplex virus tk sequences under the control of wild-type and mutant human RNA polymerase I promoters. *Molecular and cellular biology,* 5**,** 352-362.

STECKELBERG, A.-L., BOEHM, V., GROMADZKA, A. M. & GEHRING, N. H. 2012. CWC22 connects pre-mRNA splicing and exon junction complex assembly. *Cell reports,* 2**,** 454-461.

TAFFOREAU, L., ZORBAS, C., LANGHENDRIES, J. L., MULLINEUX, S. T., STAMATOPOULOU, V., MULLIER, R., WACHEUL, L. & LAFONTAINE, D. L. 2013. The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. *Mol Cell,* 51**,** 539-51.

TAMM, T. 2012. A thiamine-regulatable epitope-tagged protein expression system in fission yeast. *Recombinant Gene Expression.* Springer.

TANAKA, S., MIYAZAWA-ONAMI, M., IIDA, T. & ARAKI, H. 2015. iAID: an improved auxin-inducible degron system for the construction of a 'tight'conditional mutant in the budding yeast Saccharomyces cerevisiae. *Yeast,* 32**,** 567-581.

TARDIFF, D. F., LACADIE, S. A. & ROSBASH, M. 2006. A genome-wide analysis indicates that yeast pre-mRNA splicing is predominantly posttranscriptional. *Molecular cell,* 24**,** 917-929.

TIAN, B., HU, J., ZHANG, H. & LUTZ, C. S. 2005. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic acids research,* 33**,** 201-212.

TILGNER, H., KNOWLES, D. G., JOHNSON, R., DAVIS, C. A., CHAKRABORTTY, S., DJEBALI, S., CURADO, J., SNYDER, M., GINGERAS, T. R. & GUIGÓ, R. 2012. Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. *Genome research,* 22**,** 1616-1625.

TOPISIROVIC, I., SVITKIN, Y. V., SONENBERG, N. & SHATKIN, A. J. 2011. Cap and cap-binding proteins in the control of gene expression. *Wiley Interdisciplinary Reviews: RNA,* 2**,** 277-298.

ULE, J., JENSEN, K. B., RUGGIU, M., MELE, A., ULE, A. & DARNELL, R. B. 2003. CLIP identifies Nova-regulated RNA networks in the brain. *Science,* 302**,** 1212-1215.

VIPHAKONE, N., SUDBERY, I., GRIFFITH, L., HEATH, C. G., SIMS, D. & WILSON, S. A. 2019. Co-transcriptional loading of RNA export factors shapes the human transcriptome. *Molecular cell*.

WAHL, M. C., WILL, C. L. & LÜHRMANN, R. 2009. The spliceosome: design principles of a dynamic RNP machine. *cell,* 136**,** 701-718.

WAHLE, E. 1995. Poly (A) tail length control is caused by termination of processive synthesis. *Journal of Biological Chemistry,* 270**,** 2800-2808.

WAHLE, E., MARTIN, G., SCHILTZ, E. T. & KELLER, W. 1991. Isolation and expression of cDNA clones encoding mammalian poly (A) polymerase. *The EMBO Journal,* 10**,** 4251-4257.

WALKER, J. E., SARASTE, M., RUNSWICK, M. J. & GAY, N. J. 1982. Distantly related sequences in the alpha-and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *The EMBO journal,* 1**,** 945-951.

WANG, Z., MURIGNEUX, V. & LE HIR, H. 2014. Transcriptome-wide modulation of splicing by the exon junction complex. *Genome biology,* 15**,** 551.

WATSON, A. T., DAIGAKU, Y., MOHEBI, S., ETHERIDGE, T. J., CHAHWAN, C., MURRAY, J. M. & CARR, A. M. 2013. Optimisation of the Schizosaccharomyces pombe urg1 expression system. *PLoS One,* 8**,** e83800.

WATT, S., MATA, J., LÓPEZ-MAURY, L., MARGUERAT, S., BURNS, G. & BÄHLER, J. 2008. urg1: a uracil-regulatable promoter system for fission yeast with short induction and repression times. *PloS one,* 3**,** e1428.

WEEKS, J. R., HARDIN, S. E., SHEN, J., LEE, J. M. & GREENLEAF, A. L. 1993. Locus-specific variation in phosphorylation state of RNA polymerase II in vivo: correlations with gene activity and transcript processing. *Genes Dev,* 7**,** 2329-44.

WEISMAN, R. 2004. The fission yeast TOR proteins and the rapamycin response: an unexpected tale. *TOR.* Springer.

WEN, J. & BROGNA, S. 2010. Splicing-dependent NMD does not require the EJC in Schizosaccharomyces pombe. *The EMBO journal,* 29**,** 1537-1551.

WEST, S., GROMAK, N. & PROUDFOOT, N. J. 2004. Human 5'→ 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature,* 432**,** 522.

WILD, T., HORVATH, P., WYLER, E., WIDMANN, B., BADERTSCHER, L., ZEMP, I., KOZAK, K., CSUCS, G., LUND, E. & KUTAY, U. 2010. A protein inventory of human ribosome biogenesis reveals an essential function of exportin 5 in 60S subunit export. *PLoS Biol,* 8**,** e1000522.

WINDHAGER, L., BONFERT, T., BURGER, K., RUZSICS, Z., KREBS, S., KAUFMANN, S., MALTERER, G., L'HERNAULT, A., SCHILHABEL, M. & SCHREIBER, S. 2012. Ultrashort and progressive 4sU-tagging reveals key characteristics of RNA processing at nucleotide resolution. *Genome research,* 22**,** 2031-2042.

WOOD, V., GWILLIAM, R., RAJANDREAM, M. A., LYNE, M., LYNE, R., STEWART, A., SGOUROS, J., PEAT, N., HAYLES, J., BAKER, S., BASHAM, D., BOWMAN, S., BROOKS, K., BROWN, D., BROWN, S., CHILLINGWORTH, T., CHURCHER, C., COLLINS, M., CONNOR, R., CRONIN, A., DAVIS, P., FELTWELL, T., FRASER, A., GENTLES, S., GOBLE, A., HAMLIN, N., HARRIS, D., HIDALGO, J., HODGSON, G., HOLROYD, S., HORNSBY, T., HOWARTH, S., HUCKLE, E. J., HUNT, S., JAGELS, K., JAMES, K., JONES, L., JONES, M., LEATHER, S., MCDONALD, S., MCLEAN, J., MOONEY, P., MOULE, S., MUNGALL, K., MURPHY, L.,

NIBLETT, D., ODELL, C., OLIVER, K., O'NEIL, S., PEARSON, D., QUAIL, M. A., RABBINOWITSCH, E., RUTHERFORD, K., RUTTER, S., SAUNDERS, D., SEEGER, K., SHARP, S., SKELTON, J., SIMMONDS, M., SQUARES, R., SQUARES, S., STEVENS, K., TAYLOR, K., TAYLOR, R. G., TIVEY, A., WALSH, S., WARREN, T., WHITEHEAD, S., WOODWARD, J., VOLCKAERT, G., AERT, R., ROBBEN, J., GRYMONPREZ, B., WELTJENS, I., VANSTREELS, E., RIEGER, M., SCHAFER, M., MULLER-AUER, S., GABEL, C., FUCHS, M., DUSTERHOFT, A., FRITZC, C., HOLZER, E., MOESTL, D., HILBERT, H., BORZYM, K., LANGER, I., BECK, A., LEHRACH, H., REINHARDT, R., POHL, T. M., EGER, P., ZIMMERMANN, W., WEDLER, H., WAMBUTT, R., PURNELLE, B., GOFFEAU, A., CADIEU, E., DREANO, S., GLOUX, S., et al. 2002. The genome sequence of Schizosaccharomyces pombe. *Nature,* 415**,** 871-80.

XIAO, X., WANG, Z., JANG, M. & BURGE, C. B. 2007. Coevolutionary networks of splicing cis-regulatory elements. *Proceedings of the National Academy of Sciences,* 104**,** 18583-18588.

YAMASHITA, A., SHICHINO, Y., TANAKA, H., HIRIART, E., TOUAT-TODESCHINI, L., VAVASSEUR, A., DING, D.-Q., HIRAOKA, Y., VERDEL, A. & YAMAMOTO, M. 2012. Hexanucleotide motifs mediate recruitment of the RNA elimination machinery to silent meiotic genes. *Open biology,* 2**,** 120014.

YANAGIDA, M. 2002. The model unicellular eukaryote, Schizosaccharomyces pombe. *Genome biology,* 3**,** comment2003. 1.

YESBOLATOVA, A., NATSUME, T., HAYASHI, K. I. & KANEMAKI, M. T. 2019. Generation of conditional auxin-inducible degron (AID) cells and tight control of degron-fused proteins using the degradation inhibitor auxinole. *Methods*.

YUE, Z., MALDONADO, E., PILLUTLA, R., CHO, H., REINBERG, D. & SHATKIN, A. J. 1997. Mammalian capping enzyme complements mutant Saccharomyces cerevisiae lacking mRNA guanylyltransferase and selectively binds the elongating form of RNA polymerase II. *Proceedings of the National Academy of Sciences,* 94**,** 12898-12903.

ZHANG, J. & MAQUAT, L. E. 1997. Evidence that translation reinitiation abrogates nonsense-mediated mRNA decay in mammalian cells. *The EMBO Journal,* 16**,** 826-833.

ZHANG, J., SUN, X., QIAN, Y., LADUCA, J. P. & MAQUAT, L. E. 1998. At least one intron is required for the nonsense-mediated decay of triosephosphate isomerase mRNA: a possible link between nuclear splicing and cytoplasmic translation. *Molecular and cellular biology,* 18**,** 5272-5283.

ZHANG, Y., FORYS, J. T., MICELI, A. P., GWINN, A. S. & WEBER, J. D. 2011. Identification of DHX33 as a mediator of rRNA synthesis and cell growth. *Mol Cell Biol,* 31**,** 4676-91.

ZHANG, Z. & KRAINER, A. R. 2007. Splicing remodels messenger ribonucleoprotein architecture via eIF4A3-dependent and -independent recruitment of exon junction complex components. *Proceedings of the National Academy of Sciences,* 104**,** 11574-11579.

Appendix

Appendix I - *S. pombe* Media

	Reagent	Quantity (for 1 L)	Final Concentration
YES	Glucose	30 g	3%
	Yeast extract	50 g	0.5%
	Adenine	225 mg	1.31 mM
	Uracile	225 mg	2.01 mM
	L-histidine	225 mg	1.45 mM
	L-leucine	225 mg	1.71 mM
	L-lysine	225 mg	1.23 mM
	Arginine	225 mg	1.3 mM
	Agar (for solid media)	20 g	2%
	ddH$_2$O	To 1 L	

	Reagent	Quantity (for 1L)	Final Concentration
EMM(G)	Glucose	20 g	2%
	Potassium hydrogen phtallate	3 g	14.7 mM
	Dibasic sodium phosphate	2.2 g	15.5 mM
	L-glutamic acid	3.38 g	2.3 mM
	Adenine	225 mg	1.31 mM
	uracil	225 mg	2.01 mM
	L-hystine	225 mg	1.45 mM
	L-leucine	225 mg	1.71 mM
	L-lysine	225 mg	1.23 mM
	Arginine	225 mg	1.3 mM
	50x salt stock	20 mL	1x
	1000x vitamin stock	1 mL	1x
	10000x mineral stock	0.1 mL	1x

Stock for use with S. *pombe* media

	Reagent	Quantity (for 1L)	Final Concentration
(1000x) Vitamin stock	Pantothenic acid	1 g	81.2 mM
	Nicotinic acid	10 g	81.2 mM
	Inositol	10 g	4.2 mM
	Biotin	10 mg	40.9μM

	Reagent	Quantity (for 1L)	Final Concentration
(1000x) Mineral stock	Boric acid	5 g	80.9 mM
	Magnesium sulphate	4 g	33.2 mM
	Zinc sulphate heptahydrate	4 g	13.9 mM
	Ferric chloride hexahydrate	2 g	7.4 mM
	Molybdic acid	0.4 mg	0.32 mM
	Potassium iodide	1 mg	6.02 mM
	Cupric sulfate pentahydrate	0.4 mg	1.6 mM
	Citric acid	10 mg	47.6 mM

E. coli media

	Reagent	Quantity (for 1L)	Final Concentration
LB (Luria-Bertani)	Yeast Extract	5 g	0.5%
	Bacto-tryptone	10 g	1%
	NaCl	10 g	1%
	Agar (for solid media)	15 g	1.5%
	ddH$_2$O	to 1 L	

	Reagent	Quantity (for 1L)	Final Concentration
NZY	Yeast Extract	5 g	0.5%
	Sodium chloride	5 g	85 mM
	Magnesium Chloride hexahydrate	2 g	10 mM
	NZ amine	10 g	1%
	ddH$_2$O	to 1 L	

Appendix II - *S. pombe* strains

S. *pombe* strains used in this study

Code	Description	Genotype	Originator
A4	ΔFal1	Otr1R(Sph1)::ura4+, pFal1-deletion:kanMX	Marayati et al., 2016
A13	Fal1-GFP-2xFlag	h⁻ ade6-M216 leu1.32,Fal1-GFP-2xFlag:: kanMX	This study
A14	Pnmt81:Fal1-GFP-2xFlag	h⁻ ade6-M216,leu1.32, Pnmt81:hygr: Fal1-GFP-2xFlag:: kanMx	This study
A19	Y14-GFP-2xFlag	h⁻ leu1-32,uraD4-D18, Y14-GFP-2xFLAG	This study
A21	ΔY14	h⁻ leu1-32 uraD4-D18, ΔY14:clonat	This study
A20	TIR1 strain (used for AID)	h- ade6::ade6+-Padh15-skp1-AtTIR1-2NLS Padh15-skp1-OsTIR1 ura4-D18	Masukata Hisao
A24	Fal1-degron (crossed with TIR1)	h⁻ ade6-M216 leu1.32, Pnmt81:hygr: Fal1-GFP-AID:clonat-TIR	This study
A35	TIR1 strain	h- ade6::ade6+-Padh15-skp1-AtTIR1-2NLS-9myc ura4-D18	Masukata Hisao

A36	Y14-5xFlag	h⁻ leu1-32 uraD4-D18, Y14-5xFlag	This study
A37	ΔMago	h⁻ leu1-32,uraD4-D18, ΔMago:clonat	This study
A38	Y14-GFP-2xFlag; ΔMago	h⁻ Leu1-32,uraD4-D18, Y14-GFP-2xFlag, ΔMago: clonat	This study
A41	Fal1-GFP-2xFlag; ΔMago	h⁻ ade6-M216,leu1.32, Fal1-GFP-2xFlag:: kanMX, ΔMago: clonat	This study
A45	WT	h⁻ leu1-32 uraD4-D18	Lab stock
A47	Mago-GFP-2xFlag	h⁻leu1-32,uraD4-D18, Mago-GFP-2xFlag; Knmx	This study
A49	Mago-GFP-2xFlag; ΔY14	h⁻ leu1-32,uraD4-D18, Mago-GFP -2xflag; knmx, ΔY14:clonat	This study
A54	Fal1-GFP-2xFlag; ΔY14	h⁻ ade6-M216,leu1.32, Fal1-GFP-2xFlag:: kanMX, ΔY14 :clonat	This study
A55	ΔFal1	H90/h90, ade6-M216/ade6-M210 leu1/leu1: ΔFal1:clonat	This study
A57	Mago-5xFlag	h⁻ leu1-32, uraD4-D18, Mago-5xFlag:knmx	This study
A58	Mago-5xFlag; ΔY14	h⁻ leu1-32 uraD4-D18, Mago-5xFlag:knmx, ΔY14:clonat	This study
A60	Y14-5xFlag; ΔMago	h⁻ leu1-32 uraD4-D18, Y14-5xFlag:knmx, ΔMago:clonat	This study
A71	nmt81:Fal1 with Rpb3-flag	h⁻ ade6-M216 leu1.32, mmt81:Fal1-GFP-AID:DSR:Clonat, rpb3-flag	This study
A89	Flag:rpb3	h⁻ ade6-M216 leu1.32 Flag:rpb3	This study

Appendix III - List of primers

List of primers used to create S. *pombe* strains

Code	Primer Name	Primer sequence
A1	4A3_CDS.F	GGTTGATTGGTTGACGGAGA
A2	4A3_Stop.R	GGGGATCCGTCGACCTGCAGCGTACGAAACC ATATCTCCTATATTCATAGGCA
A3	4A3_afterstop.F	GTTTAAACGAGCTCGAATTCATCGATATTTTC CATTGG AAAAAAAAAGCTG
A4	4A3_afterstop.R	TTAGTATCTCAAGCAACTAATCCTC
A5	4A3_Start.F	GGGGGGGATCCCGCGGACGAAATAATGGAG AA
A6	4A3_Stop.R	GGGGGGGATCCTTAAACCATATCTCCTATAT TC
A7	4A3.FW	CGAATATTGGTCAAGCGCGA
A8	4A3.RV	GCATCTCACCGTGCATACTG
A9	kanmx.FW	CGGTTGCATTCGATTCCTGT
A10	kanmx.RV	CCAGACTTGTTCAACAGGCC
A11	Fal1-F1	GTGGTGGGCAATTCTGGTTAGGAAACATTAA TTG
A12	Fal1-F2	TTAATTAACCCGGGGATCCGAATGGAAGTTG TAG TAGTTATGTTATGTCAGCCCAAC
A13	Fal1-R2	GTTTAAACGAGCTCGAATTCATTTTCCATTGG AAAAAAAAAGCTGGTTTTTTTTGGATTATTGT ACAG
A14	Fal1-R1	TTCGTGATGAAGTTTTCTGGTGGTCCAAGC
A15	Fal1-R1_nest	GCAAAGAAAGCCAATTCCTCA
A16	Fal1-seq1	GTGGAATCTATGCGTATGGCTATGAGACTC
A17	Fal1-F1	GTGGTGGGCAATTCTGGTTAGGAAACATTAA TTG
A18	Fal1-nmt-R1	GATGTGCCGCCTATACAAGCATGACATTG
A19	Fal1-nmt-F2	GTTTAAACGAGCTCGAATTCAATGGAAGTTG TAGTAGTTATGTTATGTCAGCCCAAC
A20	Fal1-nmt-R2	TATAGTCGCTTTGTTAAATCATGGCGGACG AAATAATGGAGAACGTGGAG
A21	Fal1-nmt-GFP-R2	GCATGGATGAACTATACAAAGCGGACGAAA TA ATGGAGAACGTGGAGCTTAC
A22	Fal1-R1	TTCGTGATGAAGTTTTCTGGTGGTCCAAGC
A23	GFP-seq2	CGGAAAACTTACCCTTAAATTTATTTGCACTA CTGG
A24	GFP-C-AID	CAGATTCAGCTCGACACTGCCCATCATCTTGT

		CATCGTCGTCCTTGTAGTCCTTG
A25	Spo5-DSR-marker	GTTAGACAAAGCCCCGGGATAACTTCGTATA TGG CGGATCCCCGGGTTAATTAAGGCG
A26	GFP-C-AID	CAGATTCAGCTCGACACTGCCCATCATCTTGT CATCGTCGTCCTTGTAGTCCTTG
A27	Fal1-CHK-R	CATTGCGCAAAGTCTATCAAGTCCTACTACT G
A28	Fal1-GFP-F1	CGAATATTGGTCAAGCGCGATGAACTTACTC
A29	GFP-seq3	GTCCACACAATCTGCCCTTT
A30	AID	CTAAAGATCCAGCCAAACCTCC
A31	Gar2-F1	CCCAGCCAACCCTCAACCATATGC
A32	Gar2-TMT	TTAATTAACCCGGGGATCCG GTCAAAGTGACTTTGTTCC CAGAAAATGG
A33	Gar2-R1	CATTTACTAGCTATTAGAAAGTGATCAACGA TGATTG
A34	Gar2-R2	GTTTAAACGAGCTCGAATTC GTATACTGGCTTTTAACTATATTG AGTTCATTTTTATGG
A35	GFP-C-AID_NEW	CAGATTCAGCTCGACACTGCCCATTTTGTATA GTTCATC CATGCCATGTGTAATCC
A36	Y14-GFP-F1	GCACCTCAATTTAGATCGACGTACAGGTTAT G
A37	Y14-GFP-F2	TTAATTAACCCGGGGATCCGGGGGCTCAGCCA TTGCTA CATCTCTATCTCGATGTTGAAC
A38	Y14-R2	GTTTAAACGAGCTCGAATTCATGGGCTTCTG AAAATAC ACAAAACTTATTTATGTTAG
A39	Y14-R1	GGCAGAAATTTGGGGCATAAATCATATAC
A40	Y14_CHK	CTCGTTGGAAAGCAGTCACA
A41	CBC20-FW	CCTTTAGGATCTTCTCTAGAAC
A42	CBC20-RV	CCTGTGAATCTTCCAATACCTC
A43	Y14_HA_R2	GTTTAAACGAGCTCGAATTCATCGATATGGG CTTCTGA AAATACACAAAACTTATTTATGTTAG
A44	Y14_HA_F2	GGGGATCCGTCGACCTGCAGCGTACGAGGGC TCAGCCATTGCTACATCTCTATCTCGATGTTG AAC
A45	Mago-GFP-F1	GGAGATGTTCGTCAGTGAGC
A46	Mago-GFP-F2	TTAATTAACCCGGGGATCCGGTTTTTAAC AGGTCGGAGCTTAAAATTTAGGGAG
A47	Mago-R2	GTTTAAACGAGCTCGAATTCATGCATTCCATT TTCTTATATATTAGG
A48	Mago-R1	TCCTTCATCGACTGCATCCA
A49	RV-marker-km	GGTATTCTGGGCCTCCATGTC
A50	Mago-CHK	CCAGGAGCAATGGAAAGGTA
A51	Mago-F2	TTAATTAACCCGGGGATCCGTTTCGAGCAAA

		TAGGAATATGATGG
A52	Mago R2-del	GTTTAAACGAGCTCGAATTCATGCATTCCATT TTCTTAT ATATTAGG
A53	Mago-F1	TGGATGACAGCTGTGAATGT
A54	Y14-GFP-F1	GCACCTCAATTTAGATCGACGTACAGGTTAT G
A55	Y14-GFP-F2	TTAATTAACCCGGGGATCCGGGGCTCAGCCA TTGCTAC ATCTCTATCTCGATGTTGAAC
A56	Y14 R2	GTTTAAACGAGCTCGAATTCATGGGCTTCTG AAAATACACAAAACTTATTTATGTTAG
A57	Y14 R1	GGCAGAAATTTGGGGCATAAATCATATAC
A58	Fal1-GFP-F2	TTAATTAACCCGGGGATCCGAACCATATCTC CTATAGGCATTTCATCTATAACAG
A59	Fal1-CHK-R	CATTGCGCAAAGTCTATCAAGTCCTACTACT G
A60	Y14-F2	TTAATTAACCCGGGGATCCGTATTCTAAATTT TAAAAATTGTAAGAAGGAGATTTC
A61	Y14-F1	GTGAAGATAGCGAATGTTGTTACCAGGAG
A62	Y14-del-CHk	AGGCACGAACAATACGATTCC
A63	AID	CTAAAGATCCAGCCAAACCTCC
A64	rbp3-rv	CCTACTGGAGGCGGTTCATC
A65	rpb3-fw	AGACCAAACGACACATTCTCA

List of primers used for ChIP-qRT-PCR analysis

Gene name & primer stock reference	Primer pair	Primer Sequence
Intergenic region (T67-T68) From Tina Mcleod	FW RV	GCGAAACCAGTATGGACGAT AACGGGCAAATGTAAAGACG
pma1 From Sandip De stocks	P1 FW P1 RV P2 FW P2 RV P3 FW P3 RV	CTCTAGAACATACGTTATTTAATCTCGA GTATTACCGACAATAGAAAAGGGG GTCTTCGTGATTGGGTCGAT GGGGTCACCATAGTGCTTGT ATCCCGTTTCCAAGAAGGTT GAGGATCGGAACAAGGCATA
fba1 (T77-T82) From Tina Mcleod	F1 FW F1 RV F2 FW	CCCTGCCATTAACGTCACCT CCAGCAAAGAAGTGAGCACC CTGGGATGTCTACCGTGAGC

	F2 RV	GTGTTGACCCAAAAGAGCGG
	F3 FW	CGACAAGCCCGTTTTCTTCG
	F3RV	AGGCGAATTGGGTATCAGTGT
gpd3 (JM80-JM83) From Jianming Wang	G1 F	CTCACTGGCAAGATCCAAGTTGTCG
	G1 R	ACCGTGGGTAGAGTCGTACTTGAAC
	G2 F	CAAGCGTGTCATCATCTCTGCTCCT
	G2 R	GTGCAAGAGGCGTTGGAGATAACC
arg11 This study	F1 FW	ACGGAATGGAGCGAAACACT
	F1 RV	TGGTTTCGCGAGTGGGTTTA
	F2 FW	TGGCCTCATGAAGCAACCAT
	F2 RV	GGCTACGGAAGATGTACGGG
	F3 FW	TGAAGGCGGTCTCATCGAAG
	F3RV	GTGCCTTCCAACTCACGAGA

Details of gene loci analysed by CHIP-qPRC

Gene	Systematic name	Locus	Length (nt) Transcription unit (CDS)	Product	RNA copy number per cell	Protein copy number per cell
pma1	SPAC1071.10c	Chr I 3876176 - 3871642	4535 (2760)	P-type proton ATPase, P3- type Pma1	190	438967
fba1	SPBC19C2.07	Chr II 1688336 - 1689599	1264 (1077)	fructose-bisphosphate aldolase Fba1	520	1376321
gpd3	SPBC354.12	Chr II 578063 - 580056	1994 (1008)	glyceraldehyde 3-phosphate dehydrogenase Gpd3	110	151540
Arg11	SPAC4G9.09C	Chr I 2272676-2270019	3004 (2658)	N-acetyl-gamma-glutamyl-phosphate reductase/acetylglutamate kinase	8.5	67667

List of primers used as probe for Northern-Blot

Primer	Sequence
18S - annealing in 18 rRNA (Reverse primer)	TGCACCACCACCCATAGAAT
7SL - annealing in 7SL (Reverse primer)	CCCGGTAGTGATGTGCATTG

Appendix IV - Maps of plasmids

Maps and description of the plasmids used to create *S. pombe* strains
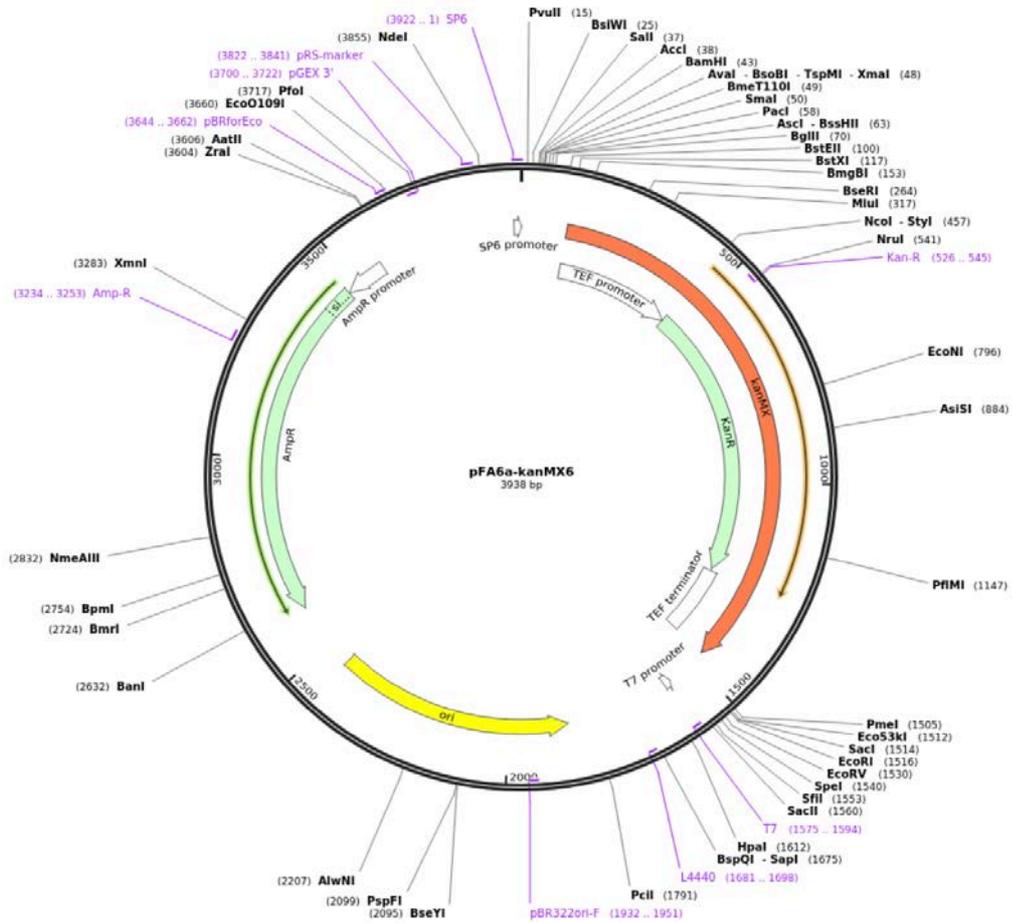
pFA6a-KanMX6 - Map of the pFA6a-KanMX6 plasmid: variant of this plasmid, containing GFP-2xFLAG upstream the KanMX6 gene cassette has been used to tag Fal1, MAGO and Y14.

pFA6a-GFP-2xFLAG-KanMX6 – a variant of pFA6a-KanMX6 plasmid: variant of this plasmid, containing GFP-2xFLAG upstream the KanMX6 gene cassette has been used to tag Fal1, MAGO and Y14. The nucleotide sequence shows the GFP sequence (in green), the 2xFLAG (in magenta), the stop codon (red). The remaning nucleotides represent the terminator and part of the KanMX6 cassette.

ACGCGGCCGCCAGCTGAAGCTTCGTACGCTGCAGGTCGACGGATCCCCGGGTTAATTAACAG
TAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTA
ATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATACGGAAAACTTACC
CTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTT
CACTTATGGTGTTCAATGCTTTTCAAGATACCCAGATCATATGAAACGGCATGACTTTTTCA
AGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTCAAAGATGACGGGAAC
TACAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAA
AGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACT
CACACAATGTATACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATT
AGACACAACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAAT
TGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTTCGA
AAGATCCCAACGAAAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTGCTGGGATT
ACACATGGCATGGATGAACTATACAAAGACTACAAGGACGACGATGACAAGGACTACAAGGA
CGACGATGACAAGTAGGGCGCGCCACTTCTAAATAAGCGAATTTCTTATGATTTATGATTTT
TATTATTAAATAAGTTATAAAAAAAATAAGTGTATACAAATTTTAAAGTGACTCTTAGGTTT
TAAAACGAAAATTCTTATTCTTGAGTAACTCTTTCCTGTAGGTCAGGTTGCTTTCTCAGGTA
TAGTATGAGGTCGCTCTTATTGACCACACCTCTACCGGCAGATCCGCTAGGGATAACAGGGT
AATATAGATCTGTTTAGCTTGCCTCGTCCCCGCCGGGTCACCCGGCCAGCGACATGGAGGCC
CAGAATACCCTCCTTGACAGTCTTGACGTGCGCAGCTCAGGGGCATGATGTGACTGTCGCCC
GTACATTTAGCCCATACATCCCCATGTATAATCATTTGCATCCATACATTTTGATGGCCGCA
CGGCGCGAAGCAAAAATTACGGCTCCTCGCTGCAGACCTGCGAGCAGGGAAACGCTCCCCTC
ACAGACGCGTTGAATTGTCCCCACGCCGCGCCCCTGTAGAGAAATATAAAAGGTTAGGATTT
GCCACTGAGGTTCTTCTTTCATATACTTCCTTTTAAAATCTTGCTAGGATACAGTTCTCACA
TCACATCCGAACATAAACAACCATGGGTAAGGAAAAGACTCACGTTTCGAGGCCGCGATTAA
ATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCA
GGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGG
CAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAAT
TTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACC
ACTGCGATCCCCGGCAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAA
TATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTC
CTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTG
GTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGA
AATGCATAAGCTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTTG
ATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATC
GCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATT
ACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTC
ATTTGATGCTCGATGAGTTTTTTCTAATCAGTACTGACAATAAAAAGATTCTTGTTTTCAAGA

ACTTGTCATTTGTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAATCAAATGTTAGCGTG
ATTTATATTTTTTTTCGCCTCGACATCATCTGCCCAGATGCGAAGTTAAGTGCGCAGAAAGT
AATATCATGCGTCAATCGTATGTGAATGCTGGTCGCTATACTGCTGTCGATTCGATACTAAC
GCCGCCATCCAGTTTAAACGAGCTCGAATTCATCGATGATATCAGATCCACTAGTGGCCTAT
GCGGCCGCGGATCTGCCGGTCTCCCTATAGTGAGTCGTATTAATTTCGATAAGCCAGGTTAA
CCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCG
CTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCAC
TCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGC
AAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGC
TCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA
GGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGAC
CCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAAT
GCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCAC
GAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCC
GGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGT
ATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACA
GTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTG
ATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGC
GCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGG
AACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGAT
CCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTG
ACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCC
ATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCC
CAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACC
AGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCT
ATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGT
TGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCG
GTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCC
TTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGC
AGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGT
ACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCA
ATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTC
TTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTC
GTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACA
GGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACT
CTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT
TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCA
CCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAG
GCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGG
AGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCA
GCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAG
AGTGCACCATATGGACATATTGTCGTTAGAACGCGGCTACAATTAATACATAACCTTATGTA
TCATACACATACGATTTAGGTGACACTATAG
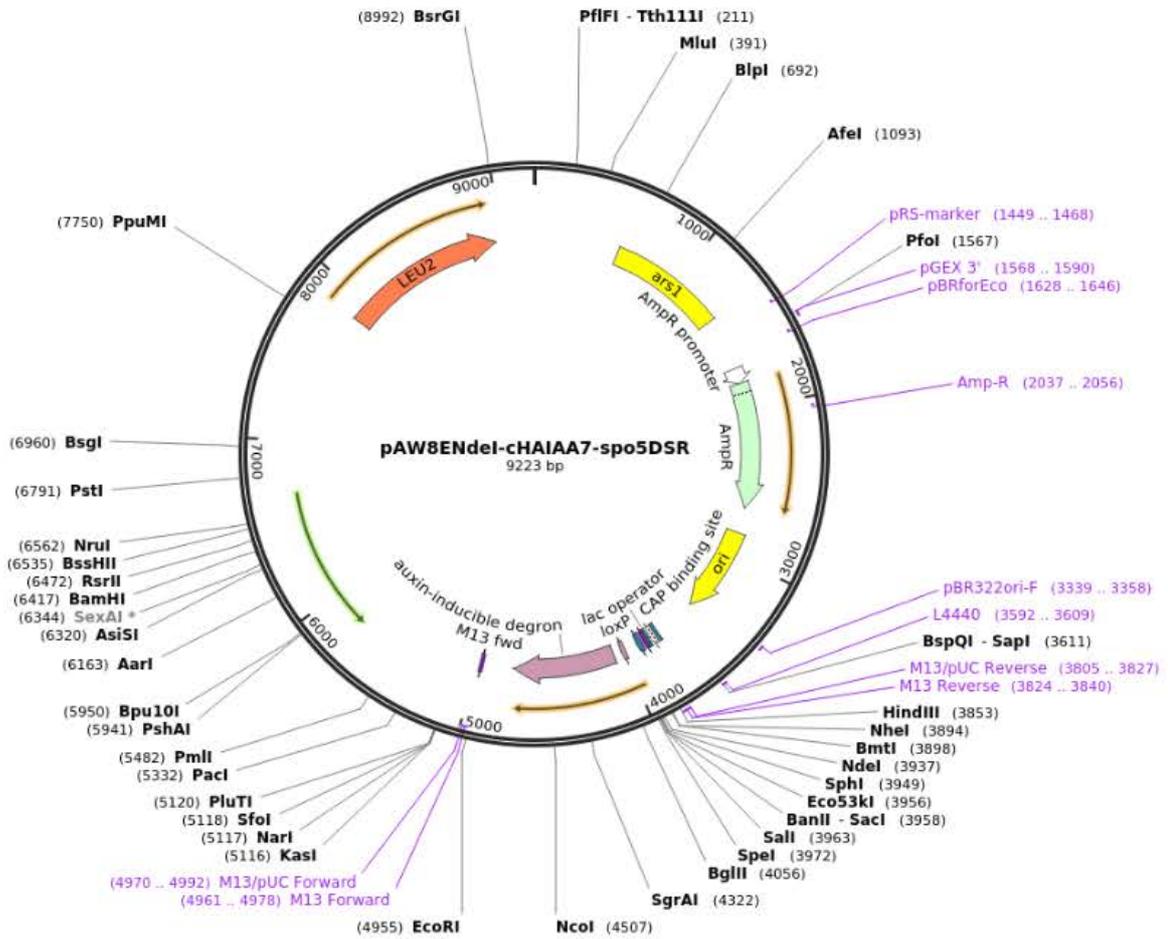
pAW8ENdeI-cHAIAA7-spo5DSR - Map of the pAW8ENdel-cHIAA7-spo5DSR plasmid: this plasmid has been used to tag Fal1 with the AID (IAA17) epitope.

The plasmid was a gift from Professor Anthony Carr, University of Sussex. The plasmid is mentioned in (Watson et al., 2013).

Brief description of the pAW8ENdeI-cHAIAA7-spo5DSR (Watson et al., 2013): it contains the AID sequence epitope in frame with the DSR element of *spo5* (*S. pombe* gene: SPBC29A10.02). The nucleotide sequence of the AID (in cyan), the spo5-DSR element (in magenta) and the stop codon (in red) is showed below.

The whole sequence of the plasmid can be found at this link:
https://www.addgene.org/49873/

ATGATGGGCAGTGTCGAGCTGAATCTGAGGGAGACTGAGCTGTGTCTTGGTCTTCCCGGTGG
AGATACAGTGGCTCCGGTAACCGGAAACAAGAGAGGGTTCTCAGAGACGGTTGATCTGAAGC
TAAATCTGAATAATGAGCCTGCAAACAAGGAAGGATCTACGACTCATGACGTCGTGACTTTT
GATTCCAAGGAGAAGAGTGCTTGTCCTAAAGATCCAGCCAAACCTCCGGCCAAGGCACAAGT
TGTGGGATGGCCACCGGTGAGATCATACCGGAAGAACGTGATGGTTTCCTGCCAAAAATCAA
GCGGTGGCCCGGAGGCGGCGGCGTTCGTGAAGGTATCAATGGACGGAGCACCGTACTTGAGG
AAAATCGATTTGAGGATGTATAAAGCTACGATGAGCTTTCTAATGCTTTGTCCAACATGTT
CAGCTCTTTTACCATGGGCAAACATGGAGGAGAAGAAGGAATGATAGACTTCATGAATGAGA
GGAAATTGATGGATTTGGTGAATAGCTGGGACTATGTTCCCTCTTATGAAGACAAAGACGGT
GATTGGATGCTCGTCGGCGACGTTCCTTGGCCAATGTTCGTCGATACATGCAAGCGTTTACG
TCTCATGAAAGGATCGGATGCCATTGGTCTCGCTCCGAGGGCGATGGAGAAGTGCAAGAGCA
GAGCTTGA
CCCGGGACTACGCCATATCATGCCCATCCCGCTACTTCTGCTGCTAATTAAATTTAAACATC
TTTCCAGCATTTACTTGAACATCAAACTCTAATATCGAAATCTTTTCAAACACTTATCACAT
ACTCAAACTTAAACCAACATAAAACCTGTTAGACAAAGCCCCGGGATAACTTCGTATATGG
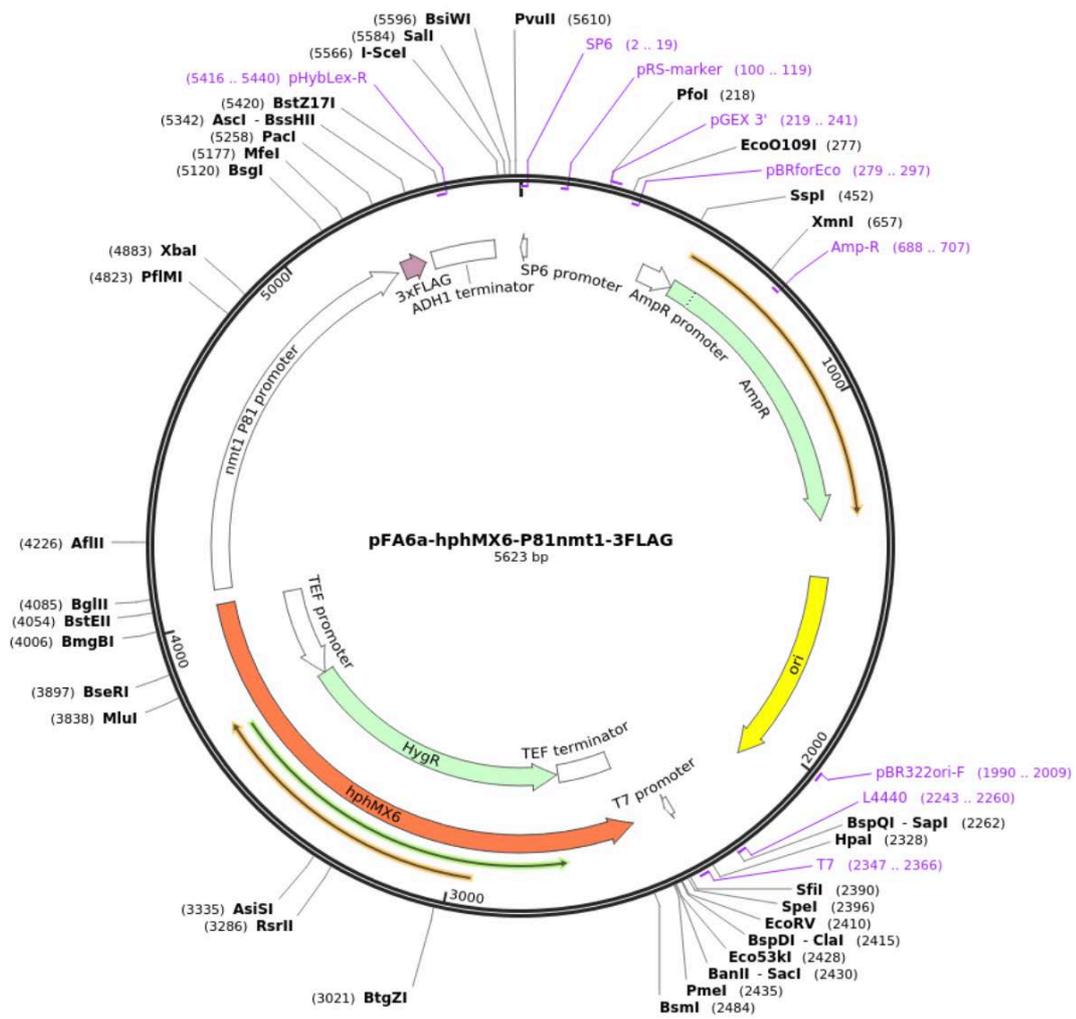
pFA6a-hphMx6-P81nmt1-3xFLAG - Map of the pFA6a-hphMx6-P81nmt1-3xFLAG plamisd: a gene cassette containing the Pnmt81 promoter dowstream the Hygromycin resistence gene, deriving from this plasmid, has been used to generate the Fal1 conditional knock-out, substituing the Fal1 promoter with the Pnmt81.
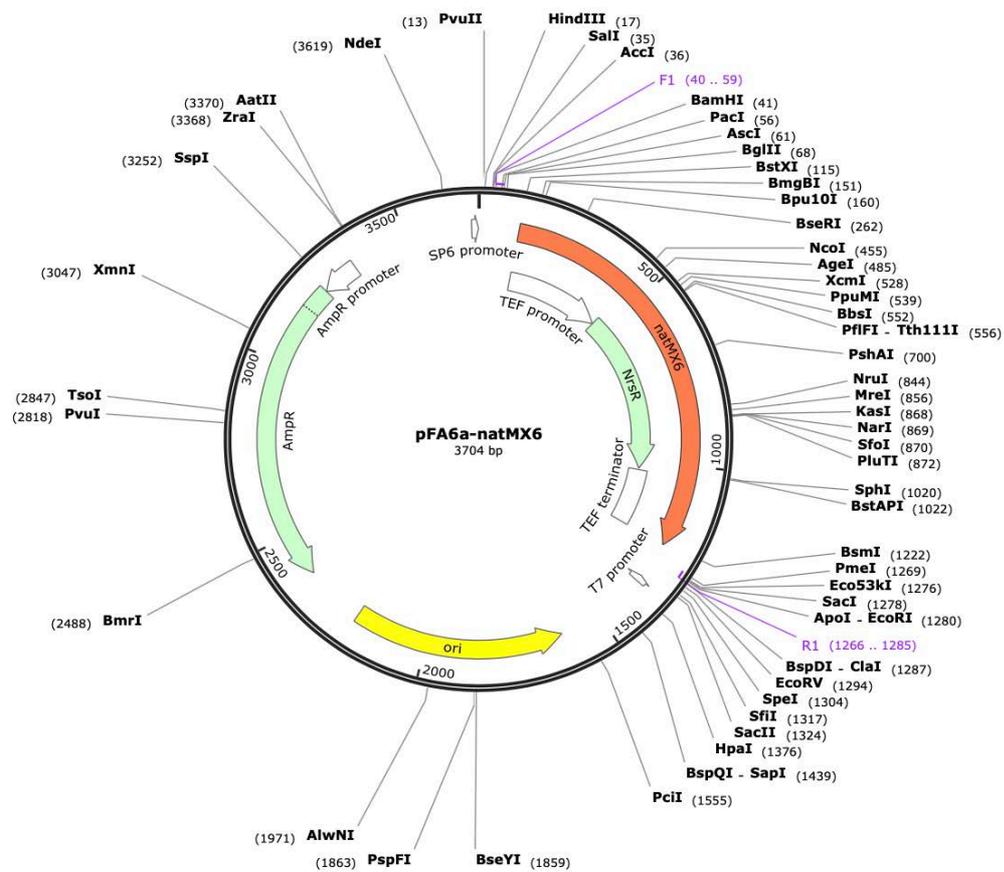
pFA6a-natMX6 - Map of the pFA6a-natMX plasmid: a gene cassette deriving from a

pFA6a-natMX has been used to generate the knock-out strains of Fal1, MAGO and Y14
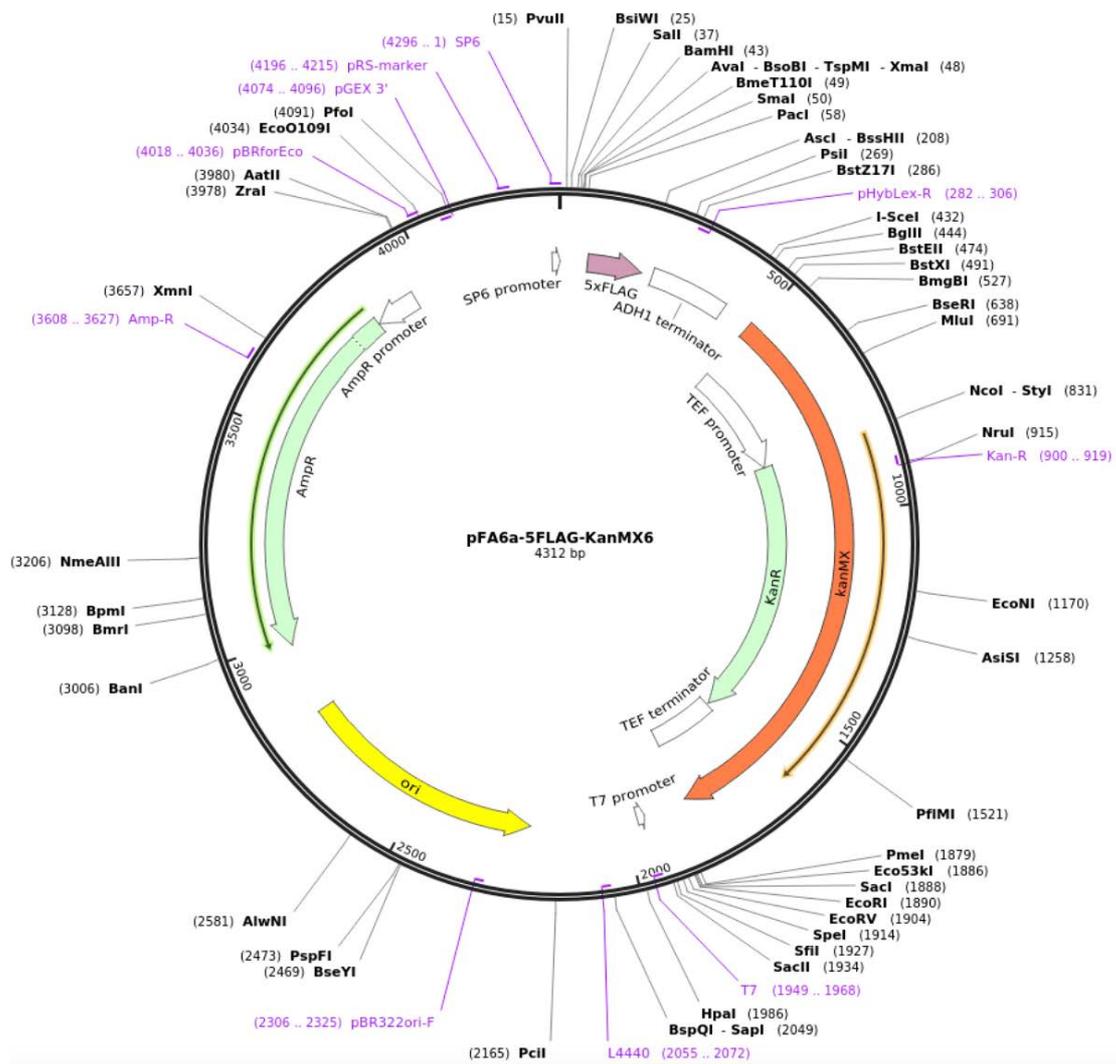
pFA6a-5xFlag-KanMX6 - Map of the pFA6a-5xFLAG-KanMX6 plasmid: this plasmid has been used to tag MAGO and Y14 with a 5xFLAG epitope.

Appendix V - List of antibodies

List of primary antibodies used for WB

Antibody	Reference	Dilution
anti-FLAG (mouse)	M2, monoclonal, Sigma	1:5000
anti-GFP (goat)	AHP75, polyclonal, Bio-Rad	1:5000

List of antibodies used for ChIP

Antibody	Reference
anti-Flag (mouse)	M2, monoclonal, Sigma
anti-phospho S2 (rabbit)	Ab5095, polyclonal, Abcam

Appendix VI - Additional Results
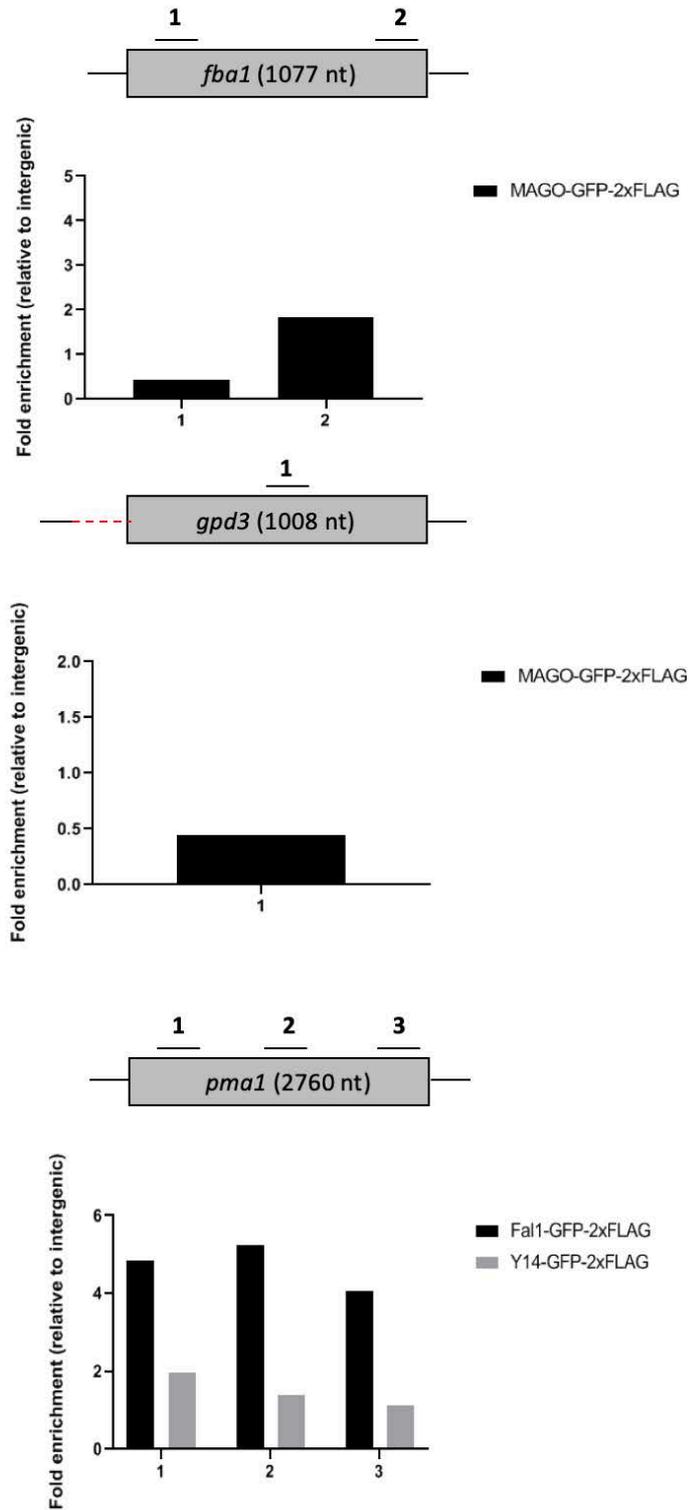
Additional ChIP experiments. ChIP of MAGO-GFP-2xFLAG on *fba1* and *gpd3*. ChIP of Fal1-GFP-2xFLAG and Y14-GFP-2xFLAG on *pma1*
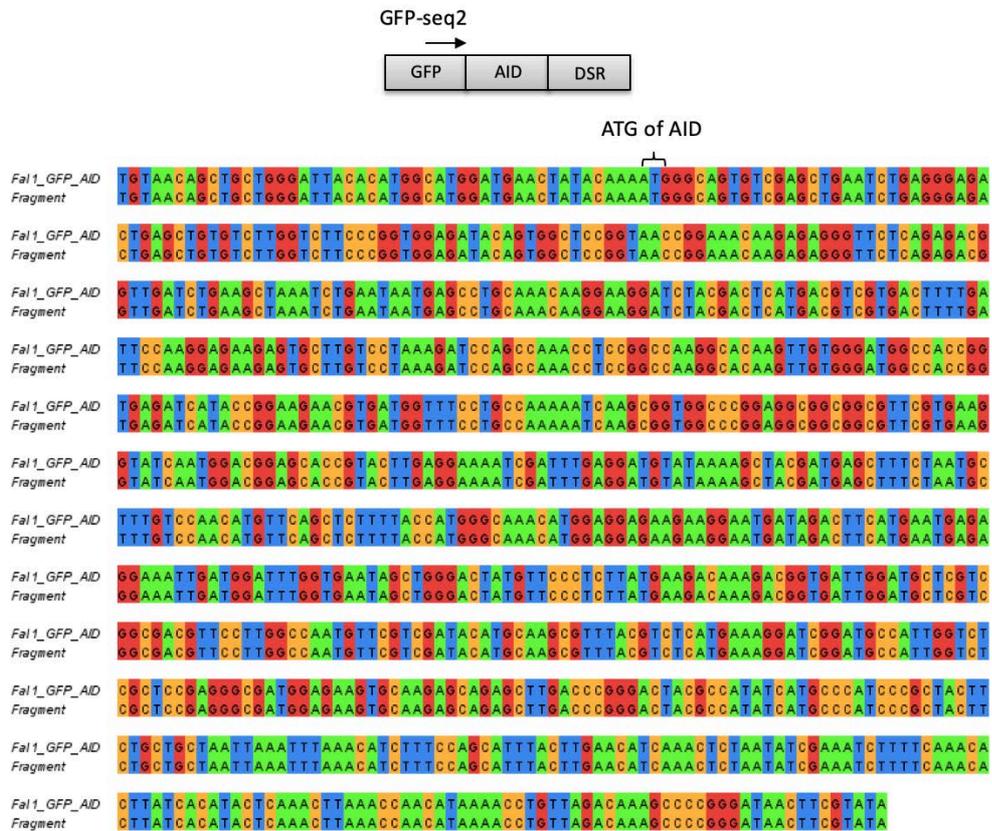
Verification of the correct integration of the AID tag in frame with the GFP epitope. The sequence of the AID fragment shows 100% of homology with the expected sequence and the GFP is in frame with the AID epitope.
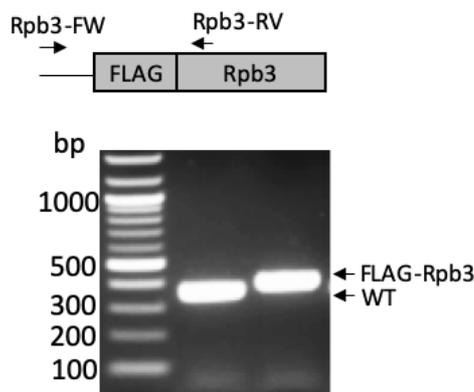
Verification of the Flag-Rpb3 strain by PCR on genomic DNA.
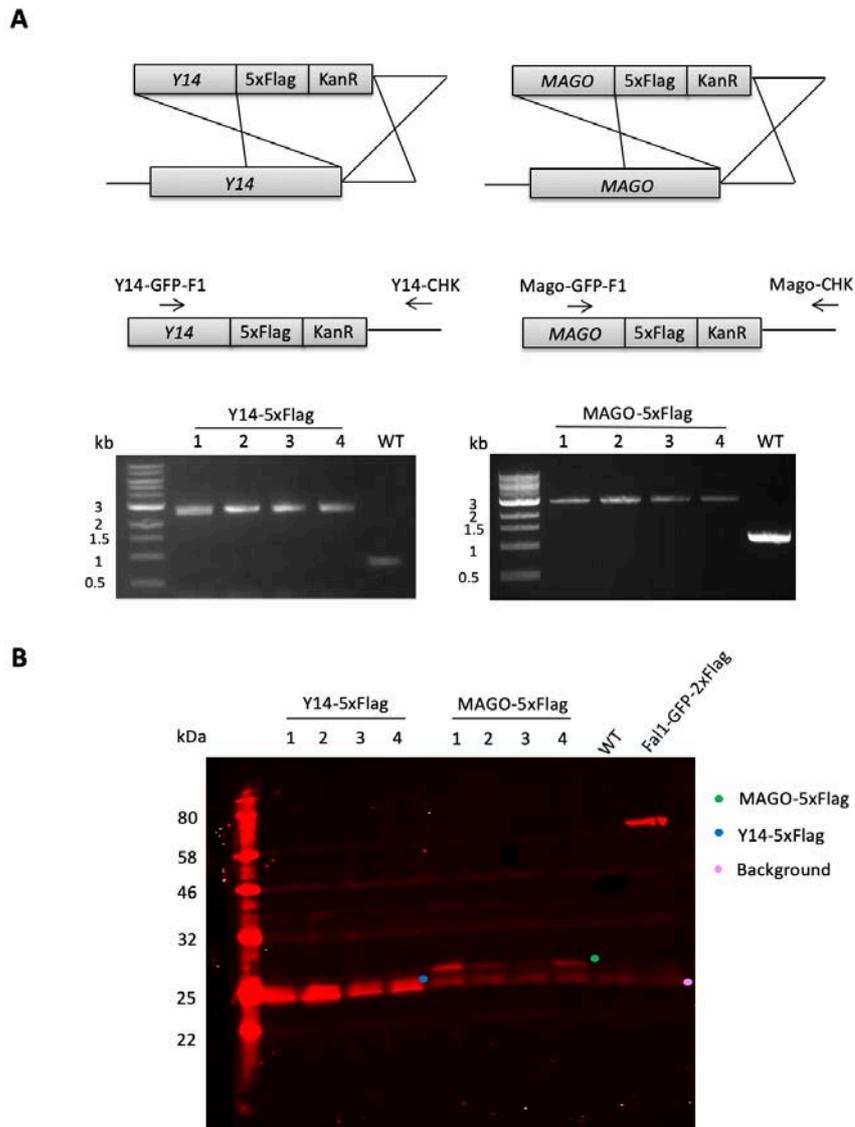
Creation and Validation of Y14-5xFLAG and MAGO-5xFLAG strains. (A)MAGO and Y14 are tagged at the C-terminal with a short 5xFLAG epitope. The integration of the tag is verified by PCR: for Y14, using the primer Y14-GFP-F1, as forward, and Y14-CHK as reverse (annealing outside the region of recombination), an expected band of almost 3 kb is successfully amplified in all the colonies analyzed. For MAGO, using the primer Mago-GFP-F1, as forward, and Mago-CHK, as reverse (annealing outside the region of recombination), an expected band of almost 3 kb was amplified in four colonies. (B) The expression of the 5xFLAG tagged Y14 and MAGO is verified by WB analysis using an anti-FLAG antibody (Sigma M2); bands of approximately 25 kDa and 28 kDa were detected for Y14 and MAGO respectively. A background line signal, also present in the WT-untagged strain is highlighted with a pink dot. Strains have been sequenced confirming the presence of the FLAG-tag.

A

Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
putative RNA-binding protein [Schizosaccharomyces pombe]	247	247	100%	2e-87	100.00%	NP_594439.1
unnamed protein product [Schizosaccharomyces pombe]	45.8	45.8	48%	8e-08	35.59%	BAA13806.1
translation initiation factor 3 [Schizosaccharomyces pombe]	45.4	45.4	46%	3e-07	37.50%	BAA25105.1
translation initiation factor eIF3g [Schizosaccharomyces pombe]	45.1	45.1	46%	4e-07	37.50%	NP_595727.1
putative RNA-binding protein [Schizosaccharomyces pombe]	43.9	43.9	61%	1e-06	30.67%	NP_596114.1
putative RNA-binding protein NIFK [Schizosaccharomyces pombe]	40.8	40.8	47%	1e-05	37.93%	NP_588551.1
Meu5 [Schizosaccharomyces pombe]	38.9	38.9	52%	5e-05	32.31%	BAB60882.1
Hypothetical protein YPR112c [Schizosaccharomyces pombe]	35.0	35.0	46%	0.001	25.00%	BAA87307.1
putative RNA-binding protein Mrd1 [Schizosaccharomyces pombe]	35.0	35.0	38%	0.001	29.79%	NP_595599.1
poly(A)-binding protein Crp79 [Schizosaccharomyces pombe]	33.9	33.9	47%	0.004	31.58%	NP_001018242.1
Crp79 [Schizosaccharomyces pombe]	33.9	33.9	47%	0.004	31.58%	AAL65912.1
RNA-binding protein Mug28 [Schizosaccharomyces pombe]	31.6	31.6	80%	0.024	30.00%	NP_593427.1

B

Dunbrack 2010 ILE rotamers

Chi 1	Chi 2	Probability
-60.4	170.3	0.687429
-55.4	-59.0	0.276404
-63.4	83.1	0.024766
-176.3	166.4	0.007430
-176.5	70.0	0.003449
60.6	171.3	0.000394
60.6	86.3	0.000074
-164.1	-80.1	0.000054
66.2	-71.1	0.000000

Dunbrack 2010 SER rotamers

Chi 1	Probability
66.8	0.403318
-63.6	0.309306
176.0	0.287376

(A) Percentage of identity of Y14 amino acid sequence of *S. pombe*, with putative paralogs. The Y14 amino acid sequence of *S. pombe* is aligned with putative paralogs using Protein Blast (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins). **(B) Rotamers of isoleucine and Serine according to the Dunbrack 2010 rotamers library.** Table of rotamers of Isoleucine and Serine. Chi1 and Chi2 represent the side chain torsion angles.

Appendix VII - Point by Point protocols

1. ChIP

1) A fresh patch of *S. pombe* strain is inoculated in 10 mL of YES media and grown overnight at 30°C. Next day, the cell culture is poured in 90 mL of YES and grown until OD_{595}=0.5 (1x 10^7 cells/ml) (Wait at least 2-3 division).

2) Cells are fixed for 5 minutes at RT by adding 2.7 mL of 37% Formaldehyde (Sigma).

3) To stop the reaction, 20 mL of 2.5 M glycine is added to the cells suspension and incubated for 10 minutes at RT on a rotor.

4) The cell culture is poured in two 50 mL Falcon tubes and centrifuged at 5000 rpm for 3 minutes at 4°C.

5) Pellet is resuspended in 40 mL of cold 1XPBS. Cells are centrifuged at 5000 rpm for 3 minutes at 4°C; other two washings with cold 1X PBS are perfoemd at the same condition.

6) Pellet is resuspend in 750 μL ice-cold FA Lysis Buffer (100 mM HEPES-KOH, pH.7.5, 300 mM NaCl, 2mM EDTA, 2% Triton X-100, 0.2% Na Deoxycholate) containing an EDTA-free protease inhibitors cocktail tablet (Roche, 1 tablet for 10 mL of FA Lysis Buffer); cell suspension is transferred in a 2 mL microfuge tube.

7) Cells are pelleted at 6000 rpm for 2 minutes at 4°C. to the pellet 750 μL FA lysis Buffer with EDTA-free protease inhibitor cocktail tablet and 400 μL of Zirconia beads, 0.7mm of diameter (Biospec), are added.

8) Cells are broken using a cell homogenizer (Bertin Technologies, Precellys 24) in 10 cycles: 30 seconds at 5500 rpm and 2 minutes in ice. The bottom of each screw cap tube is pierced two times with a red-hot 25 G needle and each tube is immediately

transferred in the barrel of a syringe and adapted in a 15 mL falcon tube. Lysate is collected at 1000 rpm for 1 minute at 4°C; remaining lysate is collected adding 500 μL of FA lysis Buffer with mini EDTA-free protease inhibitor cocktail tablet and centrifuging at the same conditions.

9) The cell lysate is transferred in a 1.5 ml microfuge tube and harvested at 13000 rpm for 25 minutes at 4°C. Pellet was resuspended in 600 μL of FA lysis Buffer with mini EDTA-free protease inhibitor cocktail tablet.

10) 750 μL of cell lysate transferred in a 15 mL tube (Sumitomo or in alternative a typical a 15mL tube): 250 μL of FA lysis Buffer with mini EDTA-free protease inhibitor cocktail tablet was added to reach the final volume of 1 mL. To allow the sonication <u>is essential</u> to add 20 μL of 10% SDS (<u>without SDS the chromatin is not sonicated</u>) and 20 μL of 100 mM PMSF is added to the mixture and gently resuspended. Samples are sonicated for 15 cycles, using Bioruptor (Diagenode).

11) After sonication the sample is transferred into a 1.5 mL tube and centrifuged 13000 rpm for 25 minutes at 4°C. The aqueous phase (containing the sonicated chromatin), is transferred into a 15 mL tube (in ice) and diluted with the same volume of FA lysis Buffer with mini EDTA-free protease inhibitor cocktail tablet, to reduce the concentration of SDS used in the sonication procedure. Keep the sampe in ice, during beads preparation.

12) From this sample, 50 μL of sonicated chromatin are collected as Input and stored at -20°C.

13) 50 μL of Dynabeads (Thermofisher), are transferred in a 1.5 mL DNA low binding tube (Eppendorf) and washed three times with 1 mL of cold 1xPBS containing 5% BSA. Typically, 10 μg of antibody (this depends about the target and the antibody)

was added and gently resuspended. The mixture is incubated on a rotor for 1 hour at room temperature. To remove unbound antibody, beads are washed three times with 1 mL 1xPBS containing 5% BSA.

14) After the final washing, supernatant is removed and beads are mixed with the appropriate amount of sonicated chromatin and transferred in a fresh 1.5 mL DNA Low- binding tube. Ab coated beads and the chromatin were then incubated overnight at 4 °C on a rotor. <u>Note</u>: To determine the amount of sonicated chromatin for each IP, 1µL of sonicated chromatin is measured by the nanodrop; usually 300 µL of sonicated chromatin with a concentration of 1 µg/µL are used for each IP. In case of multple reactions, the amount of chromatin is calculated proportionally.

15) Next day, the tube containing the IP reaction is positioned into a magnetic rack and it is rotated 10 times. The supernatant is removed and beads are washed for 5 minutes at room temperature on a rotor using the following buffer:

- o 2 times with Wash Buffer I (50mM HEPES-KOH pH 7.5, 150 mM NaCl, 1mM EDTA pH 8.0, 1% Triton X-100, 0.1% Sodium deoxycholate, 0.1% SDS)

- o 2 times with Wash Buffer II (50mM HEPES-KOH pH 7.5, 500 mM NaCl, 1mM EDTA pH 8.0, 1% Triton X-100, 0.1% Sodium deoxycholate, 0.1% SDS)

- o 2 times with Wash Buffer III (10mM Tris-HCl pH 8.0, 1mM EDTA pH 8.0, 0.25mM LiCl, 0.5% Igepal CA, 1% Sodium deoxycholate)

- o 2 times with TE (10mM Tris-HCl pH 8.0, 1mM EDTA pH 8.0)

   <u>Note</u>: between each washing, beads are briefly (not more than 3 seconds) centrifuged to avoid the attachment to the lid, and they are rotated on the

magnetic rack for 10 times. Every two washing the beads are also transferred into a clean 1.5 mL DNA Low-binding tube.

16) After the final washing, the beads are resuspended in 100 μL of Elution Buffer-EB (50mM Tris-HCl pH 7.5, 10mM EDTA, 1% SDS) and incubated for 10 minutes at 65°C and occasionally vortexed.

17) The supernatant (elution) is recovered by putting the tube on the magnetic rack, and it is transferred in a fresh 1.5 mL DNA low binding tube. Elution is performed a second time using same conditions.

18) To the input 150 μL of EB is added and incubated at 65°C overnight to allow de-crosslinking. IP sample is de-crosslinked in parallel using the same condition.

19) After de-crosslinling 5 μL Proteinase K (20 mg/mL) are added to the IP and Input and samples are incubated at 50 °C for two hours.

20) To remove RNA, 1 μL of RNase A (10 mg/mL) is added and the sample is incubated for 30 minutes at 37°C.

21) DNA is extracted using the QIAquick PCR purification kit, and eluted in 40 μL of Elution Buffer.

2. Mini-prep boiling methods

1) A single colony of *E.coli*, is inoculated in 2-3 mL of LB with Ampicillin (or other antibiotics according to the resistence) and grown for 13-14 hours at 35.5°C. (<u>Note</u>: it is essential to use this temperature). The culture is poured into a 1.5 mL Eppendorf tube and bacteria are harvested by centrifugation for 5 minutes at low speed (300g).

2) The supernatant is completely removed.

3) Add 400 µL STET buffer (8% sucrose, 5% Triton-x 100, 50 mM EDTA pH 8-8.5, 50 mM Tris pH 8-8.5. <u>Note</u>: do not autoclave STET buffer), 30 µL Lysozyme (Sigma, 10 mg/mL (w/v) in 10mM Tris pH 7.5, flash freeze aliquots in liquid nitrogen).

4) Vigorously vortex to mix and let sit for 5 seconds.

5) Boil the tube for 1-3 minutes.

6) Centrifuge for 10 minutes at maximum speed.

7) Prepare the following mix according to the number of preps: 10 µL RNseA (Roche, 2mg/ml (w/v) in 10mM Tris ph 7.5, 15 mM NaCl, flash freeze aliquots in liquid nitrogen) and 50 µL of 3M NaAc (pH5.2).

8) The pellet after centrifugation (step 6) will have a gluey texture. Remove with a toothpick and add 60 µL RNase/NaAc mix (prepared in step 7) and mix by inverting the tube. Wait 5 minutes.

9) Add 500 µL Isopropanol (or 1 Vol) and mix by inverting the tubes. Spin at maximum speed for 10 minutes.

10) Discard the supernatant, and wash pellet with 650 µLof 75% (v/v) EtOH.

11) Spin for 8 minutes, discard the supernatant, air dry the pellet.

12) Dissolve the pellet in 50µL of TE.

3. *S. pombe* transformation

1) Inoculate a fresh single colony into 5 mL of YES or EMM liquid media (according to the yeast genotype), and grown overnight until OD_{600}=0.5.

2) Harvest the cells by pelleting at 3000 rpm for 1 minute at room temperature.

3) Wash the pellet with 1 mL of 0.1 M Lithium Acetate and harvest again at 3000 rpm for 1 minute.

4) Resuspend the pellet in a solution consisting of: 120 μL of PEG (50% w/v), 18 μL of 1M Lithium Acetate, 10 μL of 10 mg/mL salmon sperm DNA, the DNA fragment/plasmid to transform (typically 1 μg of DNA) and sterile H_2O until a final volume of 180 μL.

5) Incubate the cells for 1 hour at room temperature followed by a heat shock at 42°C for 15 minutes.

6) Harvest the cells at 3000 rpm for 30 seconds at room temperature, and resuspend the pellet in 150 μL of sterile H_2O.

7) Plate on a YES or EMM agar plate (according to the yeast genotype).

8) Incubate the plate at 30°C for 24 hours.

9) Replica plate on a fresh plate containing the desired antibiotic. The selective plate is incubated for 2-3 days at 30°C until colony formation.

Appendix VIII - Schematic diagrams to generate yeast strains

(The following illustrations have been adapted from Dr. Kayoko Tanaka's illustrations).

Schematic diagram explaining how to tag a gene with the AID system
(primers sequences for this specific project are reported in Appendix II)
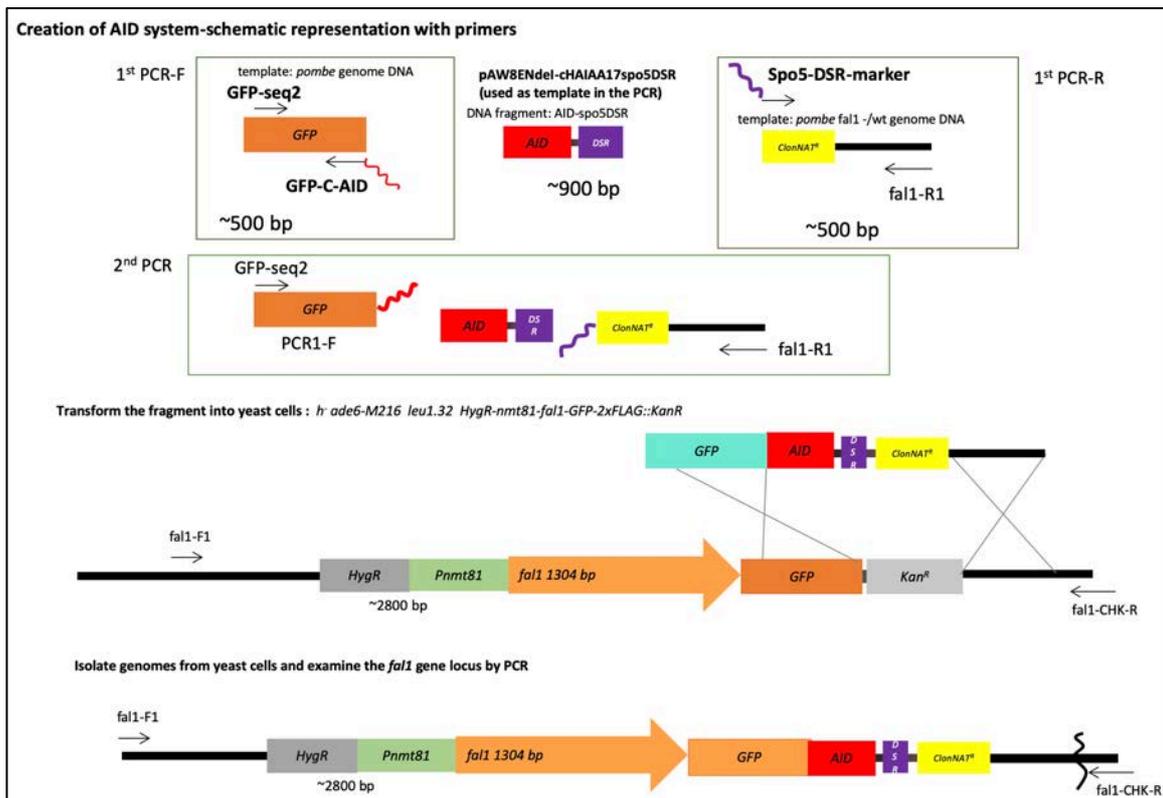
Schematic diagram explaining how to generate gene deletion. This plan can be used to generate gene deletion both in haploid and diploid *S. pombe* strains
(primers sequences for this specific project are reported in Appendix II)

fal1 gene disruption using "PCR sewing"
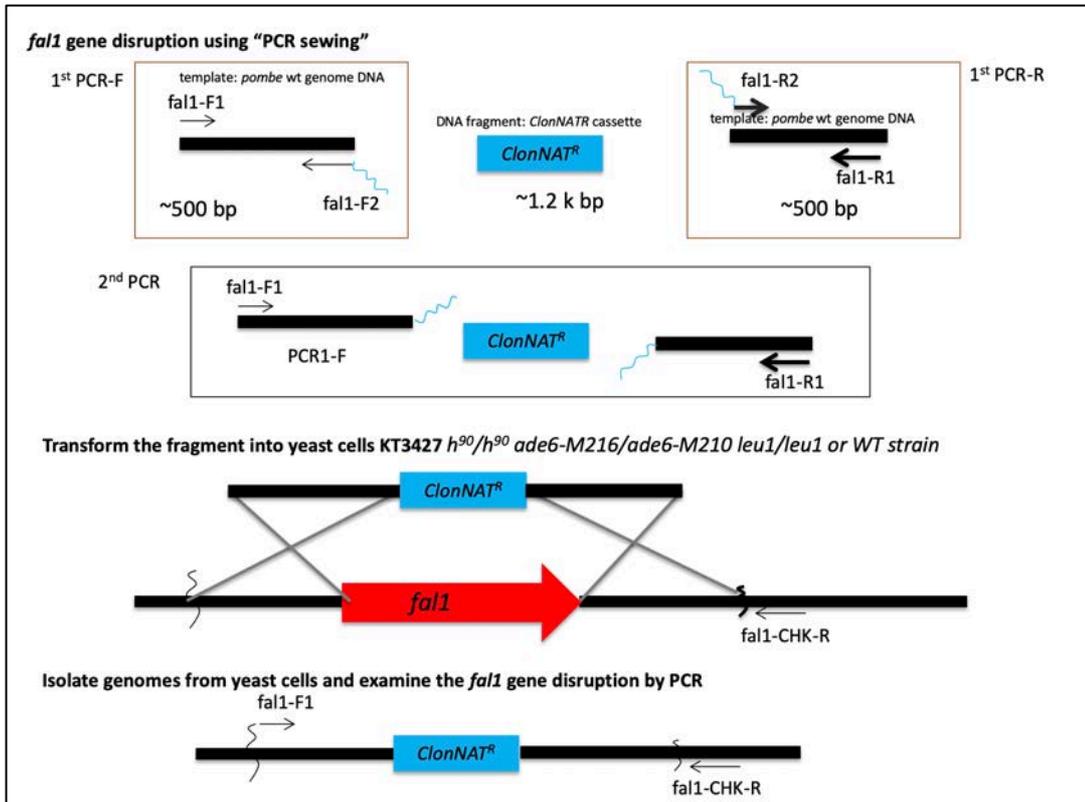
NB: the same approach was used to generate deletion of *mago*, y14 and

Schematic diagram explaining how to change the *fal1* promoter with the nmt versions (nmt1, nmt81, nmt41)
(primers sequences for this specific project are reported in Appendix II)
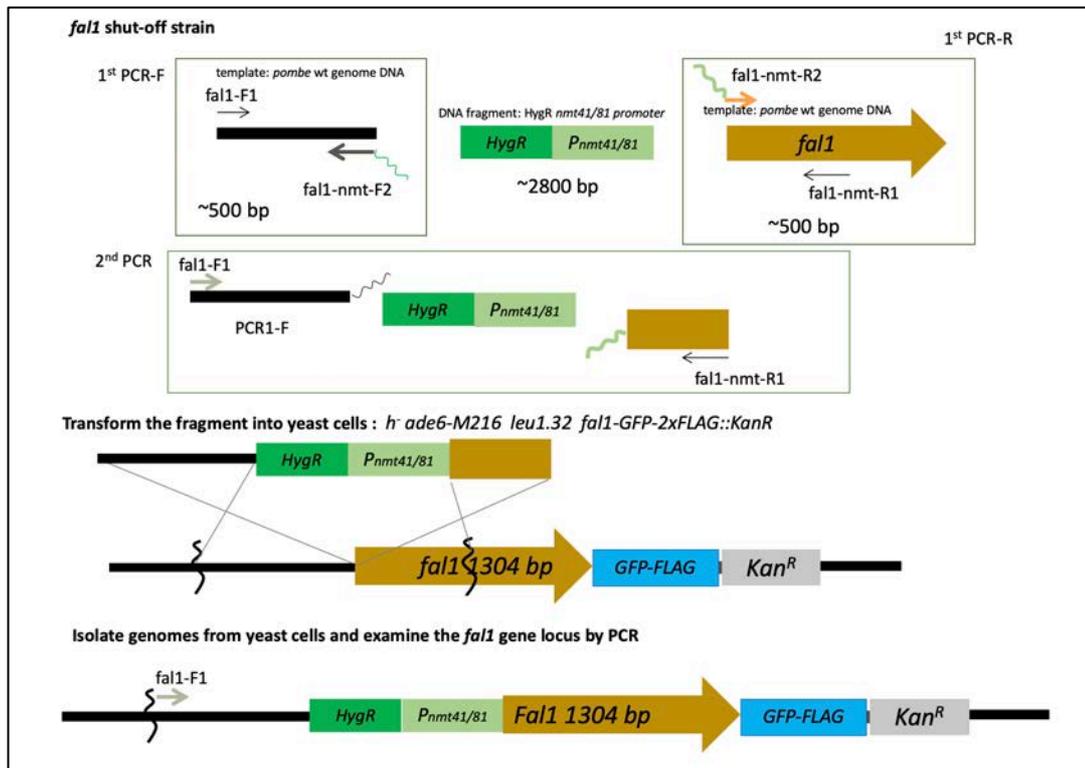
NB. This approach can be also used to tag a gene at the N-terminal region.

Schematic diagram explaining how to tag a gene with the GFP-2xFLAG epitope
(primers sequences for this specific project are reported in Appendix II)

fal1 gene tagging using "PCR sewing"
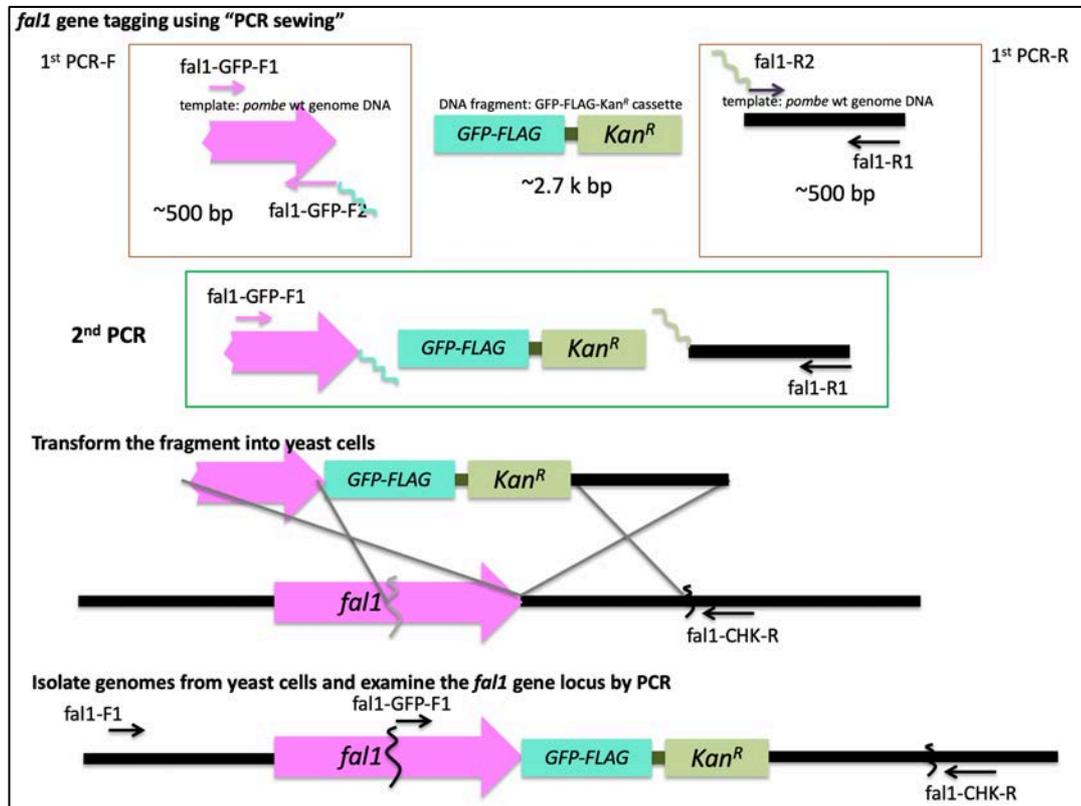
NB. The same plan can be used to tag a gene with any of the versions of the pFA6a plasmids (see 5xFLAG or 3xHA), without modifying the sequence of homology with the plasmid in the hybrid primers.

The same approach was used to tag MAGO and Y14 with the GFP-2xFLAG epitope and the 5xFLAG epitope.