JUSTUS-LIEBIG-
UNIVERSITÄT
GIESSEN

Analysis of Cbf5-dependent mRNA pseudouridylation and its function in the mRNA life cycle

Dissertation

to obtain the academic degree
doctor rerum naturalium (Dr. rer. nat.)
submitted to the Department of Biology and Chemistry
of Justus Liebig University Giessen
by

Kristin Hühn

This thesis was carried out at the Institute of Biochemistry (Department of Biology and Chemistry) of the Justus Liebig University in Giessen by supervision of Prof. Dr. Katja Sträßer.

Day of submission: 13.01.2023

1^{nd} Reviewer: Prof. Dr. Katja Sträßer
Department of Biology and Chemistry
Institute of Biochemistry
Justus Liebig University Giessen

2^{nd} Reviewer: Apl. Prof. Dr. Elena Evguenieva-Hackenberg
Department of Biology and Chemistry
Institute of Microbiology and Molecular biology
Justus Liebig University Giessen

Contents

Summary

Chemical nucleotide modifications of (ribonucleic acids) RNAs take center stage of the current research. The isomerization of a uridine to a pseudouridine, known as pseudouridylation, is the most abundant RNA modification. One pseudouridine synthase that catalyzes pseudouridylation of ribosomal RNAs (rRNAs) is Cbf5 in *Saccharomyces cerevisiae*. Cbf5 harbors the catalytic activity of the H/ACA small nucleolar ribonucleoprotein particle (H/ACA snoRNP), which also contains the proteins Gar1, Nhp2 and Nop10 and one guide H/ACA small nucleolar RNA (H/ACA snoRNA). In 2014/15, transcriptome-wide mapping of pseudouridylation sites showed that Cbf5 also pseudouridylates residues of messenger RNAs (mRNAs). So far, little is known about the prevalence, mechanism and function of Cbf5-dependent mRNA pseudouridylation and the role of Cbf5 in the mRNA life cycle. Previous results revealed an RNA-independent interaction between Cbf5 and TREX. TREX (TRanscription and EXport) is a conserved nuclear protein complex, which has a central role in coordination of mRNA synthesis and nuclear mRNA export.

In this study it was shown that Cbf5 also functions in transcription and nuclear mRNA export. Fluorescence *in situ* hybridizations (FISHs) demonstrated an mRNA export defect when Cbf5 is depleted. In addition, Cbf5 influences transcription elongation of the RNA polymerase II (RNAPII) and it also controls transcription of genes involved in metabolism and energy derivation shown by RNAPII chromatin immunoprecipitation sequencing (ChIPseq). Cbf5-dependent pseudouridylation in the 5' untranslated region (UTR) of the transcription activator *GCN4* regulates its protein amount that might influence transcription of genes involved in amino acid metabolism. Moreover, a point mutation at the Cbf5-dependent pseudouridylation site of *ADE13* regulates mRNA synthesis of this gene.

Additionally, it was shown that also Gar1 and Nhp2 of the H/ACA snoRNP interact with TREX in an RNA-independent manner. Chromatin immunoprecipitations (ChIPs) and RNA immunoprecipitations (RIPs) demonstrate that Cbf5 is present at all genes and mRNAs, independent of whether the mRNA has a pseudouridylation site. Furthermore, it was shown that Cbf5 occupancy at the genes is transcription- and RNA-dependent. Primer extension after N-cyclohexyl-N9-(2-morpholinoethyl)-carbodiimide metho-p-toluenesulphonate (CMC) labeling of nascent and poly(A) mRNA suggests that Cbf5-dependent pseudouridylation occurs post-transcriptionally.

To summarize, it was shown for the first time that Cbf5 has a role in mRNA synthesis and nuclear mRNA export just as the TREX complex. Whereas the role of Cbf5 in transcription is pseudouridylation-dependent, FISHs demonstrated that the nuclear mRNA export function of Cbf5 is pseudouridylation-independent.

Zusammenfassung

Chemische Nukleotidmodifikationen von Ribonukleinsäuren (RNAs) stehen im Mittelpunkt der aktuellen Forschung. Die Isomerisierung von einem Uridin zu einem Pseudouridin, bekannt als Pseudouridylierung, ist die häufigste RNA-Modifikation. Eine Pseudouridin-Synthase, die die Pseudouridylierung von ribosomalen RNAs (rRNAs) katalysiert, ist Cbf5 in *Saccharomyces cerevisiae*. Cbf5 besitzt die katalytische Aktivität des kleinen nukleolären Ribonukleoproteinpartikels H/ACA (H/ACA snoRNP), welches auch die Proteine Gar1, Nhp2 und Nop10 und die kleine, nukleoläre H/ACA Ribonukleinsäure (H/ACA snoRNA) enthält. 2014/15 zeigte eine transkriptomweite Kartierung von Pseudouridylierungen, dass Cbf5 auch Boten RNAs (mRNAs) pseudouridyliert. Wenig ist bisher bekannt über die Prävalenz, den Mechanismus und die Funktion von Cbf5-abhängigen mRNA-Pseudouridylierungen und die Rolle von Cbf5 im mRNA-Lebenszyklus. Bisherige Resultate zeigen eine RNA-unabhängige Interaktion zwischen Cbf5 und TREX. TREX (TRanskription and EXport) ist ein konservierter Kernproteinkomplex, der eine zentrale Rolle bei der Koordination der mRNA-Synthese und des mRNA-Kernporenexports hat.
In dieser Studie wurde gezeigt, dass Cbf5 auch eine Funktion während der Transkription und dem mRNA Kernporenexport besitzt. Fluoreszenz-*in-situ*-Hybridisierungen (FISHs) verdeutlichen einen mRNA Export Defekt, wenn Cbf5 depletiert ist. Zudem wurde mittels RNAPII-Chromatin-Immunpräzipitation und anschließender Sequenzierung (ChIPseq) nachgewiesen, dass Cbf5 die Transkriptionselongation der RNAPII beeinflusst und die Transkription von Genen, die eine Rolle beim Metabolismus und der Energieerzeugung besitzen, reguliert. Cbf5-abhängige Pseudouridylierungen an der 5' nicht translatierten Region (UTR) des Transkriptionsaktivators *GCN4* beeinflussen dessen Proteinmenge, welche die Transkription von Genen, die am Aminosäurestoffwechsel beteiligt sind, regulieren könnten. Darüber hinaus wurde mittels einer Punktmutation an der Cbf5-abhängigen Pseudouridylierungsstelle von *ADE13* gezeigt, dass Cbf5 eine Rolle bei der Transkription besitzt.
Zusätzlich wurde nachgewiesen, dass auch Gar1 und Nhp2 des H/ACA snoRNPs RNA-unabhängig mit TREX interagieren. Chromatin-Immunpräzipitationen (ChIPs) und RNA-Immunpräzipitationen (RIPs) zeigen, dass sich Cbf5 an allen untersuchten Genen und mRNAs befindet unabhängig davon, ob die mRNA eine Pseudouridylierungsstelle besitzt. Zudem wurde nachgewiesen, dass die Cbf5 Okkupanz an den Genen transkriptions- und RNA-abhängig ist. Die Primerverlängerung nach N-Cyclohexyl-N9-(2-morpholinoethyl)-carbodiimidmetho-p-toluolsulfonat (CMC)-Mar- kierung von naszierender und poly(A)-mRNA deutet darauf hin, dass eine Cbf5-abhängige Pseudouridylierung posttranskriptional erfolgt.

Zusammenfassend wurde erstmals gezeigt, dass Cbf5 eine Funktion bei der mRNA Synthese und dem mRNA Kernporenexport besitzt. Während die Rolle von Cbf5 bei der Transkription pseudouridylierungsabhängig ist, zeigen FISHs, dass die Funktion von Cbf5 beim mRNA Kernporenexport pseudouridylierungsunabhängig ist.

1 Introduction

1.1 mRNA Biogenesis and Export

In 1952, Hershey and Chase identified that the deoxyribonucleic acid (DNA) is the hereditary material [Hershey and Chase, 1952]. Six years later, Francis Crick described the transfer of information encoded in the DNA sequence into proteins as the central dogma of molecular biology [Crick, 1970]. The transcription of genes into a RNA sequence is the first step of this pathway and therefore one of the most important processes in all living cells. Eukaryotic cells contain a nucleus, where the genetic information is stored and transcribed by up to five different RNA polymerases into RNA. One of them is RNAPII, which transcribes among others protein-coding genes into mRNAs. The transcribed nascent mRNA undergoes many processing steps such as capping, splicing, polyadenylation and is finally packed into an export-competent messenger ribonucleoprotein particle (mRNP). This mRNP is exported to the cytoplasm, where the mRNA is translated into a protein. Only the highly regulated interplay of these processes ensures efficient and accurate expression of the genetic information [Maniatis and Reed, 2002/ Cole and Scarcelli, 2006].

1.1.1 mRNA synthesis

The process of transcription can be divided into transcription initiation, transcript elongation, transcription termination and recycling of the RNAPII. A key element in guiding the RNAPII through the transcription cycle and in coupling the transcription process to mRNA processing events is the C-terminal domain (CTD) of Rpb1, the largest subunit of RNAPII. The CTD consists of highly conserved heptapeptide $(Y_1 - S_2 - P_3 - T_4 - S_5 - P_6 - S_7)_n$ repeats, which differ between species only in the number of the repeats. In *S. cerevisiae* 26 of these repeats are present, whereas 52 heptapeptides are existing in humans [Corden and Ingles, 1992/ Chapman et al., 2008]. CTD kinases and phosphatases catalyze the phosphorylation or dephosphorylation of the heptapeptide serines, tyrosines and threonines during the transcription cycle [Phatnani and Greenleaf, 2006/ Chapman et al., 2008]. According to the phosphorylation state of these amino acids of the CTD, proteins involved in transcription, chromatin remodeling or mRNA processing are recruited to the CTD. Threonines and serines can also be glycosylated and in addition, an isomerization of prolines was observed [Phatnani and Greenleaf, 2006].
For eukaryotic transcription initiation, the preinititation complex (PIC) is recruited to the promoter. PIC consists of RNAPII containing an unphosphorylated CTD, the mediator complex and general transcription factors (GTFs) [Muhlbacher et al., 2014]. The mediator

complex, which binds to the unphosphorylated CTD of RNAPII, stimulates basal transcription [Figure 1]. It also relays regulatory signals from activators or repressors by acting as a bridge between upstream regulatory factors and RNAPII [Lariviere et al., 2006/ Cai et al., 2010]. GTFs are a set of the five transcription factors (TFs) TFIIB, TFIID, TFIIE, TFIIF and TFIIH.

Initially, the TATA-binding protein (TBP), which is a subunit of the multi-protein complex TFIID, recognizes a special AT-rich sequence at the promoter. Upon binding to the TATA-sequence or the TATA-like sequence, the DNA bent by 90° [Kim et al., 1993/ Tsai and Sigler, 2000]. TFIIB stabilizes the TFIID-DNA complex and recruits the TFIIF complex [Bushnell et al., 2004/ Kostrewa et al., 2009]. Consecutive binding of TFIIE and TFIIH complete PIC assembly and transcription initiation starts. TFIIE and TFIIH stimulate DNA unwinding and promoter melting. The transcription bubble is forming, which consists of a 15 nucleotide (nt) single-stranded DNA [Holstege et al., 1996/ Kim et al., 2000]. When single-stranded DNA is available in the active site of the promoter complex, incorporation of the first nucleotides can occur. Kin28, a subunit of the TFIIH, phosphorylates the CTD S_5 and S_7 residues [Figure 1] [Akhtar et al., 2009/Hengartner et al., 1998]. S_5 phosphorylation leads to dissociation of the mediator complex and to promoter clearance [Sogaard and Svejstrup, 2007]. RNAPII leaves the promoter, when the newly synthesized RNA has reached a length of 10 nt [Margeat et al., 2006] and transcription elongation starts.

In *S. cerevisiae*, transition from initiation to productive elongation is characterized by phosphorylation of S_2 and gradual S_5 dephosphorylation of the CTD by Rtr1 or Ssu72. S_2 phosphorylation is carried out by Bur1 and Ctk1 of the CTDK-I complex [Figure 1] [Mosley et al., 2009/ Cho et al., 2001/ Werner-Allen et al., 2011]. During this process, transcription elongation factors like Spt4, Spt5 and Spt6 are recruited to the RNAPII complex [Figure 1] [Martinez-Rucobo et al., 2011]. Spt6 is recruited to the phosphorylated S_2 of the CTD [Yoh et al., 2007]. It is involved in nucleosome remodeling to enable transcription elongation through chromatin. Furthermore, it has an important role in RNAPII processivity [Narain et al., 2021]. Spt5 also binds directly to RNAPII and its NGN domain associates with Spt4 to form the Spt4/5 heterodimer [Klein et al., 2011/ Martinez-Rucobo et al., 2011]. The main function of Spt4/5 is to stimulate transcription elongation mediated by the NGN domain [Hirtreiter et al., 2010]. With increasing S_2 phosphorylation of the CTD during transcription elongation, RNA cleavage factors are recruited [Mandel et al., 2008].

At the end of the transcription cycle, RNAPII has to release the RNA transcript and dissociates from the DNA template to terminate the transcription. Termination is predominantly mediated through the poly(A)-dependent pathway. The 3' end of each transcribed gene

contains a highly conserved poly(A)-signal sequence (AAUAAA) upstream and a G/U-rich sequence downstream. While RNAPII runs over the poly(A) signal, transcription decelerates and is finally terminated [Kuehner et al., 2011]. Termination factors, which are a part of the two multi-subunit complexes cleavage factor I (CFI) and cleavage and polyadenylation factor (CPF) are involved in this process [Mandel et al., 2008].

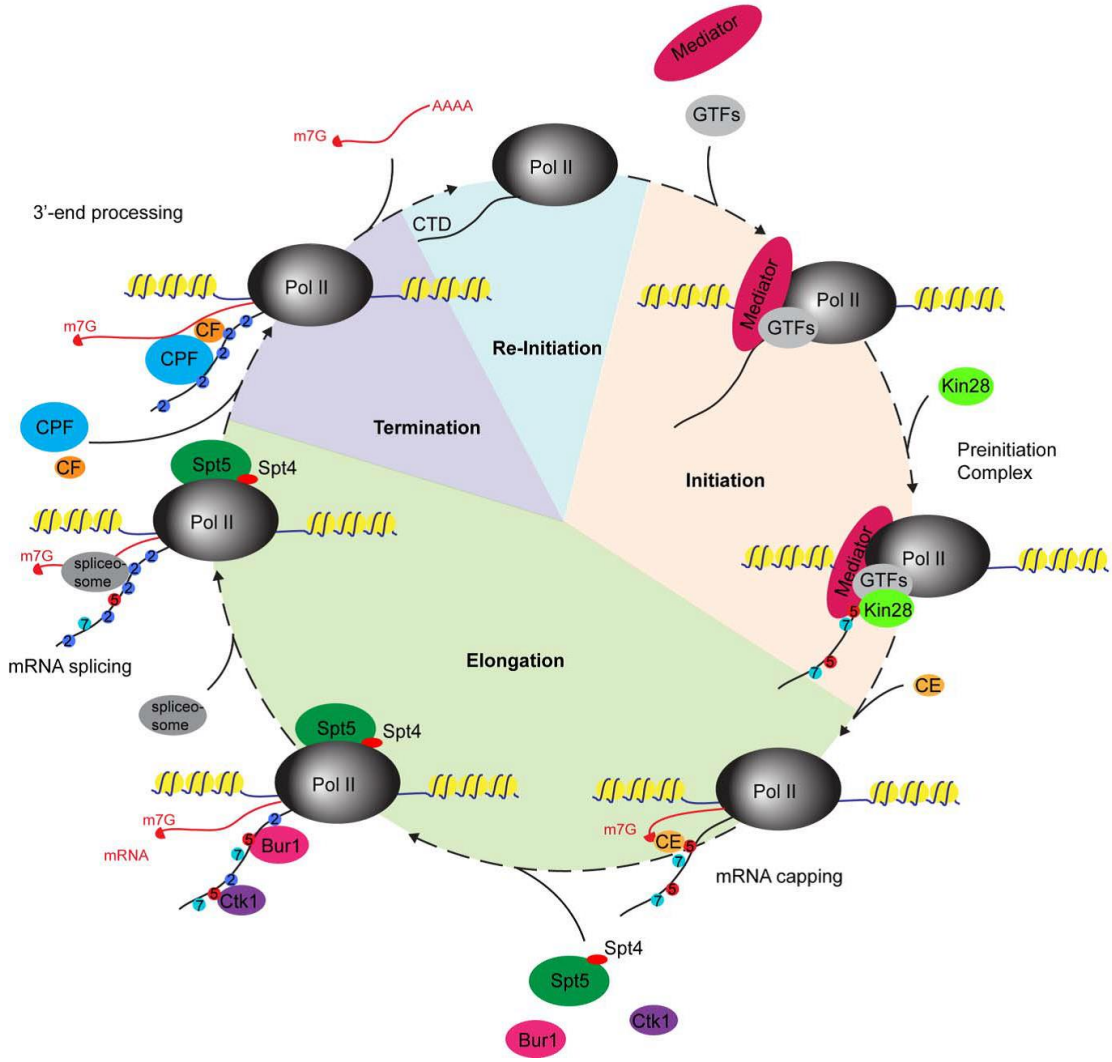

Figure 1: Schematic overview of the RNAPII transcription cycle in *S. cerevisiae*. The phosphorylation state of the C-terminal domain (CTD) of the RNA polymerase II (Pol II) and the proteins bound to it during the transcription cycle are shown. DNA is depicted as a blue line with histones in yellow and RNA as a red line. Pol II and its CTD are drawn in black. CTD phosphorylation S_7 (7), S_5 (5) and S_2 (2) is depicted as blue, red and light blue dots, respectively. GTFs: general transcription factors; CE: capping enzyme; CF: cleavage factor; CPF: cleavage and polyadenylation factor; m7G: guanine-N7 methyl cap; AAAA: poly(A) tail [Schreick, 2013]

These termination factors direct endonucleolytic cleavage of the upstream pre-mRNA and degradation of the downstream transcript by the 5'- 3' exonuclease Rat1 [Kim et al., 2004]. The termination process has been described by two different models, the "anti-terminator model" and the "torpedo model". The "anti-terminator model" describes that binding of RNA 3'-processing factors or the passage of the poly(A) signal induce structural rearrangements in the RNAPII elongation complex, which elicit termination [Logan et al., 1987]. The "torpedo model", on the other hand, stresses the role of the 5'- 3' exonuclease Rat1. Rat1 catches up with the still transcribing polymerase during the degradation of the downstream transcript after endonucleolytic cleavage. This model postulates that Rat1 destabilizes RNAPII on the DNA by its collision with this exonuclease [Kim et al., 2004]. The real situation is probably best described by a mixture of both models [Luo et al., 2006].
After transcription is terminated, RNAPII is displaced from the DNA template and has to be recycled in order to participate in a new round of transcription. Recycling requires not only the removal of transcription elongation or termination factors but also the dephosphorylation of the CTD by phosphatases. In *S. cerevisiae*, Fcp1 dephosphorylates S2, whereas S5 is dephosphorylated by Ssu72 [Krishnamurthy et al., 2004/ Hausmann et al., 2004].

1.1.2 mRNA processing and mRNP assembly

The different stages of transcription are closely linked to corresponding processing steps and the packaging of the nascent RNA transcript into an mRNP. The first steps of processing already start co-transcriptionally. After transcription initiation, the S_5-phosphorylated CTD together with the C-terminal repeats of the transcription elongation factor Spt5 recruit the RNA 5'-capping machinery, which adds the 5' guanine-N7 methyl cap (m7G) to nascent mRNAs [Lidschreiber et al., 2013]. This 5' cap structure protects the mRNA against 5' to 3' degradation and enhances mRNA translation initiation [Sonenberg and Hinnebusch, 2009/ Schwer et al., 1998].
The capping reaction comprises three steps. Firstly, in *S. cerevisiae* the RNA 5' triphosphatase Cet1 removes the γ-phosphate from the 5' triphosphate, secondly, the guanyltransferase Ceg1 adds a guanine monophosphate to the 5'-end diphosphate, and thirdly, the methyltransferase Abd1 adds a methyl group to the N7 atom of the terminal guanine group [Rodriguez et al., 1999/ Shibagaki et al., 1992/ Mao et al., 1995]. In the nucleus, the cap is bound by the cap-binding complex (CBC). CBC consists of the small subunit Cbp20 and the large subunit Cbp80. CBC promotes several downstream processes like splicing, polyadenylation and mRNA export as Cbp20 and Cbp80 are able to recruit proteins involved in these processes [Nojima et al., 2007/ Pabis et al., 2013/ Flaherty et al., 1997].

The second step of mRNA processing is the removal of introns of the pre-mRNA by splicing. This process is mediated by the spliceosome, a large ribonucleoprotein complex consisting of the five small nuclear RNAs (snRNAs) U1, U2, U4, U5, and U6 and about 170 associated proteins. These components assemble in a step-wise manner and undergo many dynamic rearrangements during the splicing reaction [Herzel et al., 2017/ Wahl et al., 2009]. Removal of an intron and ligation of the flanking exons is achieved by two successive transesterification reactions. Sequences containing an intron are characterized by short consensus sequences at the 5' and 3' intron-exon junction and a branch point inside the intron that serve as recognition sites for spliceosome components [Will and Luehrmann, 2011]. Although only 5% of *S. cerevisiae* genes contain introns [Parenteau et al., 2008], splicing is an important process, as intron-containing genes account for more than one fourth of total transcripts [Ares et al., 1999].

After correct splicing, the last processing step of the nascent mRNA is the 3' end processing. CPF is recruited to the cleavage site at the 3' end of the gene [Tian and Graber, 2012]. The CPF complex consists of more than 20 proteins. The recruitment of these proteins is mediated through interaction with the RNA and the S_2-phosphorylated CTD of RNAPII [Casanal et al., 2017]. After the mRNA is cleaved, poly(A) polymerase, which is also a subunit of CPF, elongates the poly(A) tail. Several poly(A)-binding proteins (e.g. Nab2, Pab1) stabilize the poly(A) tail [Amrani et al., 1997/ Mandel et al., 2008]. After initial polyadenylation, the yeast poly(A) nuclease complex (PAN) shortens the tail to a length of 60-80 nt [Dunn et al., 2005]. Binding of Nab2 to the mature mRNA regulates poly(A) tail length by preventing readenylation [Viphakone et al., 2008].

Many other proteins bind to the nascent mRNA during and after transcription to form the mRNP. These proteins are involved in mRNA export, mRNA localization and mRNA translation [Meinel and Straesser, 2015]. The formation of an mRNP will be only successful when the involved proteins are recruited to the RNA or DNA at the right time. To enable precise protein-RNA/DNA interaction, three different recruitment platforms are described.

One recruitment platform is the CTD of Rpb1, the composition and function of which have been described previously in chapters 1.1.1. and 1.1.2..

The general transcription elongation factor Spt5 that binds directly to RNAPII [Klein et al., 2011], serves as another recruitment platform. Similar to the CTD of Rbp1, Spt5 contains a repetitive C-terminal region (CTR). In *S. cerevisiae*, the CTR consists of 15 hexapeptides with the consensus sequence $(S_1 - T/A_2 - W_3 - G_4 - G_5 - A/Q_6)_n$ [Swanson et al., 1991]. These regions are phosphorylated on their S_1 residues during transcription elongation [Lindstrom et al., 2003]. As already mentioned, Spt5 is the recruiter of the 5' capping machinery

[Lidschreiber et al., 2013]. S_1 phosphorylation by the kinase Bur1 leads also to the recruitment of the 3' RNA cleavage factor CFI and the PAF complex. This complex is involved in several processes like chromatin remodeling and 3' end processing [Mayer et al., 2012/ Liu et al., 2009].
The RNA itself also provides a platform to recruit proteins, which recognizes either special sequence motifs or structural elements of the RNA. Those proteins mostly contain an RNA-binding domain to interact with the RNA such as an RNA recognition motif (RRM) or an arginine/glycine-rich (RGG/RG) domain [Maris et al., 2005/ Ozdilek et al., 2017]. Whereas CPF recognizes the A-rich sequence of the RNA, Nab2 binds to the mRNA body and poly(A) tail of the maturated 3'-end of the mRNA [Mandel et al., 2008]. Subcomponents of the spliceosome also recognize motifs within the RNA by base-pairing of spliceosomal RNA components with the mRNA [Will and Luehrmann, 2011]. Moreover, some mRNA binding proteins mediate the recruitment of other proteins once they are bound to the mRNA. For example, Yra1, Nab2 and Npl3 act as adaptor proteins to recruit the export receptor Mex67-Mtr2 to the mRNA.
These three recruitment platforms do not act independently of each other. There are some proteins, that are recruited simultaneously by two or all three of the recruitment mechanisms [Mayer et al., 2012]. For example, the CPF complex is recruited to the S_2-phosphorylated CTD of Rpb1, whereas CPF recruitment is also RNA-dependent and requires the A-rich sequence of the mRNA [Mandel et al., 2008].
Taken together, all three platforms: the CTD of the RNAPII, the CTR of Spt5 and the mRNA itself act together to recruit mRNA binding proteins for efficient mRNA processing and mRNP assembly. Disruption of any of these tightly coupled steps leads to the degradation of the mRNP [Mitchell et al., 2013].

1.1.3 Nuclear mRNA export

After successful processing, the mRNP needs to be transported from the nucleus into the cytoplasm. mRNA export is mediated by the highly conserved export receptor Mex67-Mtr2 [Santos-Rosa et al., 1998/ Katahira et al., 1999]. Mex67-Mtr2 binds to the mRNA but only with a very low affinity [Segref et al., 1997/ Katahira et al., 1999]. It was shown that Mex67-Mtr2 is recruited to mRNAs by direct interaction with mRNA-binding proteins. These proteins serve as export adaptors and increase the efficiency of mRNA export. One export adaptor is the TREX component Yra1. Yra1 interacts with Mex67-Mtr2 to initiate mRNA export [Straesser and Hurt, 2000]. Other export adaptors are the TREX protein Hpr1, Nab2 or the SR-protein Npl3 [Figure 2]. Hpr1 binds directly to the ubiquitin-associated-binding

domain of Mex67, when it is ubiquitylated by the ubiquitin ligase Rsp5 [Gwizdek et al., 2006]. Npl3 needs to be dephosphorylated to interact with Mex67-Mtr2 [Gilbert and Guthrie, 2004]. Nab2 also serves as an export adaptor by direct interaction with Mex67-Mtr2 and the export adaptor Yra1 to form a ternary complex [Hector et al. 2002/ Iglesias et al., 2010]. A direct interaction of export adaptor proteins with Mlp1 of the nuclear pore complex (NPC) is necessary for mRNA export [Figure 2]. This interaction facilitates a final quality check and many remodeling steps of the mRNP [Fasken et al. 2008]. Incorrectly processed mRNAs will be degraded by the nuclear exosome, while export-competent mRNPs are transported through the NPC. The NPC is structured as a cylindrical pore comprised of eight spokes to form a central channel. With a size of around 50 MDa, it is one of the largest protein complexes in *S. cerevisiae* [Alber et al., 2007]. The NPC can be divided into three segments: the nuclear face with the basket, the central channel, and the cytoplasmic face with the fibrils [Bui et al., 2013]. Approximately 30 different nucleoporins constitute one NPC [Oeffinger and Zenklusen, 2012/ Alber et al., 2007]. Specialized nucleoporins create a scaffold to bind proteins. These nucleoporins contain phenylalanine/glycine repeats (FG-repeats), which appear to be largely unfolded. The FG-repeats serve as binding sites for the mRNA export receptor Mex67-Mtr2 to facilitate the transport of the export competent mRNP through the NPC [Straesser et al., 2000]. At the cytoplasmic side, additional remodeling steps of the mRNP take place. These remodeling steps are induced by the ATP-dependent DEAD-box helicase Dbp5, which binds to the mRNP in the nucleus and shuttles together with it through the NPC [Noble et al. 2011]. In the cytoplasm, Dbp5 is activated by Gle1 and inositol hexaphosphate (IP6) in an ATP-dependent manner, which then remodels the mRNP [Figure 2]. The mRNP remodeling releases the export factors, to prevent the re-import of the mRNA [Tran et al., 2007].

1.1.4 TREX couples transcription to nuclear mRNA export

TREX has a critical part in regulating mRNA biogenesis and nuclear mRNA export. In *S. cerevisiae*, TREX is composed of the heteropentameric THO complex, (consisting of Hpr1, Tho2, Thp2, Mft1 and Tex1) and the proteins Sub2, Yra1, Gbp2 and Hrb1 [Figure 2] [Chavez et al., 2000/Straesser et al., 2002]. TREX is recruited to the transcription site by direct interaction of THO with the S_2 and S_5 di-phosphorylated CTD as well as an RNA-dependent recruitment of several THO subunits to nascent mRNA [Meinel et al., 2013/ Abruzzi et al., 2004]. It is also known that the Prp19 complex, a protein complex involved in splicing, is needed for TREX occupancy at transcribed genes [Chanarat et al., 2012]. THO is also necessary to load the TREX complex proteins Sub2 and Yra1 co-transcriptionally onto the nascent RNA transcript [Straeser and Hurt, 2001].
THO was initially found to facilitate transcription elongation especially through long, GC-rich or repeat-containing genes [Chavez et al., 2001]. Additionally, THO protects the cell from transcription-dependent hyper-recombination [Chavez et al., 2000]. During hyper-recombination, a DNA/RNA hybrid of the nascent mRNA is formed. Those R-loops result in an elongation block for the next polymerase. They can also cause replication impairment or single strand breaks in the single-stranded DNA [Huertas and Aguilera, 2003/Domiguez-Sanchez et al., 2011].
Hpr1, a component of THO, is required for 3' end formation [Chavez et al., 2001] and also serves as an export adaptor of Mex67-Mtr2 [Gwizdek et al., 2006]. Thus, Hpr1 couples transcription with nuclear mRNA export. Another component functioning in nuclear mRNA export is Yra1. This mRNA export adaptor, which is recruited to the CTD of RNAPII [MacKellar and Greenleaf, 2011], interacts with the 3'-end processing factor Pcf11 [Johnson et al., 2009]. After recruitment to the 3' processing machinery, Yra1 interacts with Sub2 [Straeser and Hurt, 2001]. In exchange for Sub2, Yra1 recruits Mex67-Mtr2 to facilitate nuclear mRNA export [Straesser and Hurt, 2000]. Studies with temperature sensitive mutants of *yra1* and *sub2* revealed a similar or even worse hyper-recombination phenotype as for THO subunit deletion mutants [Jimeno et al., 2002].
Besides Yra1 and Sub2, Gbp2 and Hrb1 associate stably with TREX. Hrb1 and Gbp1 belong to the family of serine/arginine-rich (SR) protein-like factors and are co-transcriptionally recruited to the nascent mRNA transcript via TREX [Hurt et al., 2004]. Both proteins are key surveillance factors for the nuclear export of correctly spliced mRNAs in *S. cerevisiae*. Their absence leads to the nuclear export of unspliced pre-mRNAs into the cytoplasm [Hackmann et al., 2014]. Additionally, Gbp2 and Hrb1 interact with the CTDK-1 complex [Hurt et al., 2004]. A synthetic growth defect can be also observed in a *GBP2* and *CTK1* deletion strain,

which provides a link between Gbp2 and the transcription machinery [Xie et al., 2021b]. Recently, cryo-electron microscopy structures of the yeast and human THO complex and *UAP56/DDX39*/ Sub2 and *ALYREF* (yeast Yra1) were published, which reveals molecular insights of the THO complex assembly and the interaction with *UAP56/DDX39*/Sub2 and *ALYREF* [Xie et al., 2021a/ Pühringer et al., 2020]. Furthermore, cross-linking mass spectrometry (XL-MS) of the THO complex suggests that TREX serves as a landing site of Gbp2 to facilitate its loading onto the mRNP [Xie et al., 2021a].

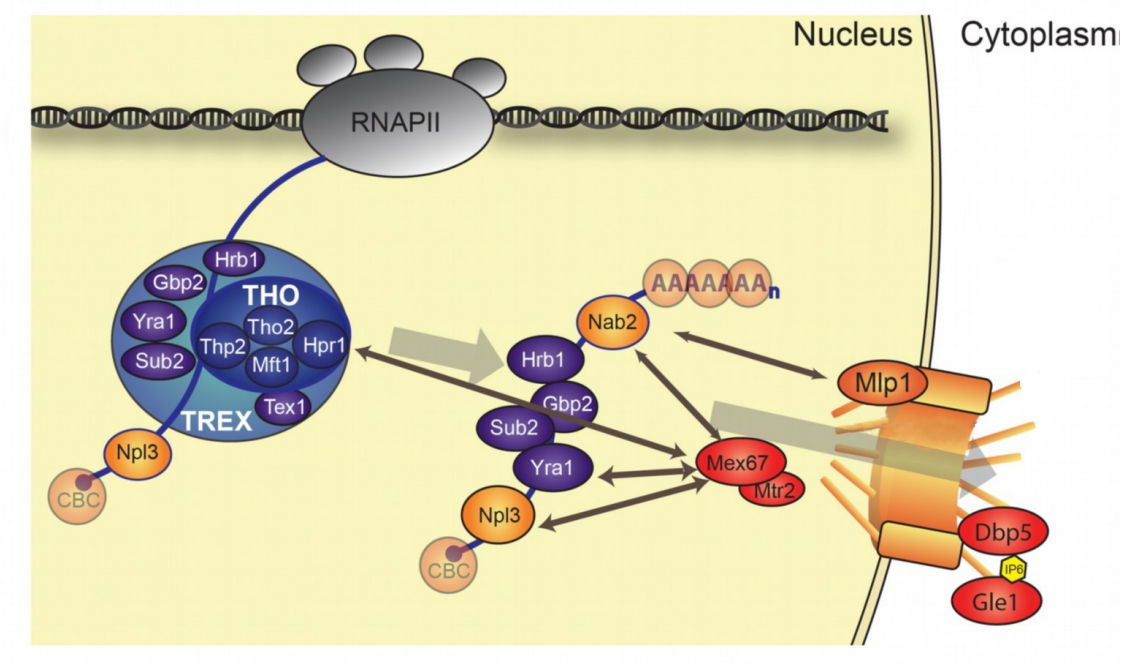

Figure 2: TREX couples transcription and export. While the RNAPII transcribes nascent mRNA, the TREX complex is recruited co-transcriptionally to the pre-mRNA via interaction with the CTD of the RNAPII. The TREX proteins Sub2, Yra1, Hrb1 and Gbp2 are a part of the mRNP. Yra1, Npl3, Nab2 and Hpr1 recruit the export receptor Mex67-Mtr2 to the mRNP. The mRNP undergoes a final quality check, before passing the NPC. The mRNP is remodeled at the cytoplasmic site by Dbp5, which is activated through the molecule inositol-hexaphosphat (IP_6) and the protein Gle1 [Figure is modified after Chanarat et al., 2012].

Taken together, TREX couples transcription to nuclear mRNA export by facilitating transcription elongation, enabling co-transcriptional mRNP assembly and packaging of the mRNA into an export-competent mRNP complex. Due to these functions, the TREX complex could be the central factor of an mRNP packaging station, that coordinates the individual steps of mRNP formation [Meinel and Straesser, 2015].

1.2 Pseudouridylation

1.2.1 Pseudouridine - the fifth RNA nucleotide

So far, over 160 different RNA modifications have been identified, one of them being pseudouridine (Ψ) [Boccaletto et al., 2018]. In 1951, pseudouridine was discovered by Cohn and Volkin as the first RNA modification [Cohn and Volkin, 1951]. It is found in all three phylogenetic domains of life. As it is also the most abundant post-transcriptional RNA modification, pseudouridine is called the "fifth RNA nucleotide" [Davis and Allen, 1957]. In *S. cerevisiae* 1.2% of all nucleotides in rRNAs are pseudouridines [Maden, 1990]. In addition to rRNAs, transfer RNAs (tRNAs), snRNAs and mRNAs also contain pseudouridines [Carlile et al., 2014].
Pseudouridine is a uridine (U) isomer (5-Ribosyluracil). A pseudouridine synthase catalyzes the site-specific isomerization of a uridine residue that is already part of an RNA chain. Initially, the nitrogen-carbon (N_1 - C_1) bond between the uracil and the ribose is broken, due to deprotonation of the ribose by the pseudouridine synthase [Veerareddygari et al., 2016]. Uracil rotates 180° along the N_3 - C_6 axis, allowing the formation of a new C_5 - C_1 bond, which links the uracil to the ribose [Figure 4] [Goldwasser and Heinrikson, 1966].

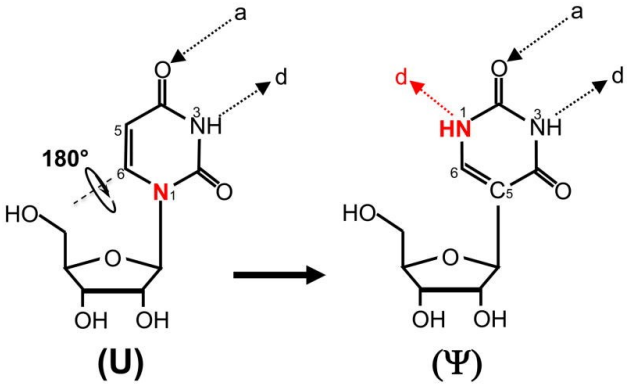

Figure 3: Schematic reaction of pseudouridylation. The structures of uridine (U) and pseudouridine (Ψ) are shown. Ψ is formed of U through an isomerization reaction, where the nitrogen - carbon bond linking the ribose to the uracil is broken to form a carbon-carbon (C_5 - C_1) bond. The nitrogen at position 1 in U and the additional hydrogen bond in Ψ are marked in red. Hydrogen bond acceptor (a) and hydrogen bond donor (d) are also indicated [Ge and Yu, 2013]

.

The C_5 - C_1 glycosidic bond of pseudouridine increases the rotational freedom, which results in a higher conformational flexibility of pseudouridines compared to uridines [Lane et al., 1995]. The additional amino-group of a Ψ functions as another hydrogen donor, which offers new base-pairing possibilities [Fernandez et al., 2013]. Furthermore, the amino-group can

form a hydrogen bond to the phosphate-group of the pseudouridine ribose in the presence of water. This hydrogen bond facilitates the formation of a rigid RNA backbone [Lane et al., 1995]. Crystal structures of tRNAs demonstrate a conformational change of the structure in presence of pseudouridine [Arnez and Steitz, 1994]. Moreover, the additional H bond of Ψ comparing to U can enhance base stacking and thus influence RNA folding [Davis, 1995]. These changed chemical properties of a pseudouridine, its abundance and the existence of Ψ in functionally important regions of the RNA suggest that pseudouridines have a crucial role in the RNA life cycle.

1.2.2 Pseudouridine synthases

The isomerization of a uridine to a pseudouridine is catalyzed by pseudouridine synthases. In eukaryotes, a distinction is made between guide RNA-independent and guide RNA-dependent pseudouridine synthases. Both types of pseudouridine synthases share a conserved aspartate residue in the catalytical domain of the enzyme [Veerareddygari et al., 2016]. Guide RNA-independent pseudouridine synthases (Pus) recognize a specific consensus sequence around the pseudouridylation-site of a target RNA. Pseudouridylation occurs without the help of any other proteins, RNAs or co-factors. In *S. cerevisiae*, nine different RNA-independent pseudouridine synthases (Pus1-Pus9) have been described [Rintala-Dempsey and Kothe, 2017]. Each pseudouridine synthase recognizes either few specific or sometimes multiple, target sequences to modify rRNAs, tRNAs, snRNAs or mRNAs [Motorin et al., 1998]. By contrast, one guide RNA-dependent pseudouridine synthase exists in eukaryotes and archaea. This pseudouridine synthase needs a small nucleolar RNA (snoRNA) complementary to the sequence to be modified for target selection [Ni et al., 1997]. The guide RNA and the pseudouridine synthase are a part of the H/ACA snoRNP, which also consists of three other proteins. In *S. cerevisiae*, H/ACA snoRNPs consist of the proteins Cbf5, Gar1, Nop10 and Nhp2 and one of multiple different snoRNAs [Watkins et al., 1998]. Pseudouridines of rRNAs are mostly modified by H/ACA snoRNPs in *S. cerevisiae*. Besides, snRNAs and mRNAs also contain Cbf5-dependent pseudouridylation sites (further informations about structure and function of H/ACA snoRNPs see chapter 1.2.5.) [Carlile et al., 2014/ Ni et al., 1997].

1.2.3 Occurrence and function of pseudouridylation

Occurrence and function of rRNA pseudouridylation

As already mentioned, 1.2% of all nucleotides in rRNAs are pseudouridines [Maden, 1990]. Many rRNA pseudouridylation sites are situated close to the peptidyl transferase center

of the ribosomes, the decoding center and the sites where ribosomal subunits interact. A temperature-sensitive *S. cerevisiae* strain containing a mutated *cbf5* displays pronounced defects in 18S and 25S rRNA synthesis at restrictive temperatures [Cadwell et al., 1997]. Deletion of five H/ACA snoRNAs, that are responsible for the formation of pseudouridines at the peptidyl transferase center of rRNAs, results in translation defects and severe growth impairment [King et al., 2003]. Moreover, Jack and colleagues reported that hypo-pseudouridylated rRNAs show a decreased affinity for tRNAs at the A- and P-site of the ribosome, resulting in decreased translational precision [Jack et al., 2011]. Furthermore, mutated pseudouridylation sites at the decoding center of 18S rRNAs impairs translation and strongly delays pre-rRNA processing [Liang et al., 2009]. These examples show that rRNA pseudouridylation has a strong impact on ribosome activity by altering rRNA folding and interaction [Decatur and Fournier, 2002]. Besides, pseudouridylation of rRNAs has an important role in rRNA biogenesis and processing [Cadwell et al., 1997].

Occurrence and function of tRNA pseudouridylation

tRNA pseudouridylation also has a role in translation. *In vitro* experiments with unpseudouridylated tRNAs in *Escherichia coli* (*E. coli*) show a decreased translation rate and translation precision [Harrington et al., 1993]. In *S. cerevisiae*, Pus1 deletion increases CGC misreading by $tRNA^{His}$ or misreading of UAG stop codons by a $tRNA^{Gln}$ [Khonsari and Klassen, 2020]. tRNA pseudouridylation sites were mainly found at the anti-codon loop, the D-loop and at the highly-conserved position 55 of tRNAs [Becker et al., 1997/ Motorin et al., 1998] and are catalyzed by Pus1, Pus3, Pus4, Pus7 or Pus8 [Carlile et al., 2014]. A Cbf5-dependent tRNA pseudouridylation is not known yet. The changed chemical properties due to a pseudouridylation result in conformational change of the tRNA structure [Arnez and Steitz, 1994]. Based on this fact, tRNA pseudouridylation increases tRNA stability and enhances codon-anticodon base pairing, which has a positive impact on translation efficiency [Harrington et al., 1993/ Khonsari and Klassen, 2020].

Occurrence and function of snRNA pseudouridylation

Similar to rRNAs and tRNAs, snRNA pseudouridylation sites are located at functional important regions of the corresponding snRNAs. Thus, they are mainly found at the branch point recognition region of the U2 snRNAs or at the 5' end of the U5 snRNAs [Yang et al., 2005]. Experiments in oocytes of *Xenopus laevis* show that snRNA pseudouridylation at these functionally important regions have a role in small nuclear ribonucleoprotein (snRNP) biogenesis and splicing. Many of these modified nucleotides are conserved across species [Massenet et al., 1998]. In *S. cerevisiae*, snRNA pseudouridylation was identified in three snRNPs (U1,U2,U5) [Carlile et al., 2014]. Ψ at position 35 of the U2 snRNA is required for pre-mRNA splicing in *S. cerevisiae* [Yang et al., 2005]. Furthermore, Newby and Greenbaum showed that this pseudouridylation site of the U2 snRNA is important for maintaining the correct structure of the branch-point adenosine [Newby and Greenbaum, 2002].
snRNA pseudouridylation also has a role in cell differentiation. Filamentous growth of yeast cells is induced due to a Pus1-dependent pseudouridylation of the U5 snRNA [Basak and Query, 2014].
Additionally, other studies showed that snRNA pseudouridylation were induced during stress conditions. Wu and colleagues published that the U2 snRNA of *S. cerevisiae* has two stress-induced pseudouridylation sites Ψ56 and Ψ93. Both pseudouridylation events were induced during nitrogen starvation, whereas only Ψ93 can be observed during heat shock. They could also show that this stress-induced pseudouridylation of the U2 snRNA reduces pre-mRNA splicing [Wu et al., 2011].
Taken together, snRNA pseudouridylation has an important role in cell differentiation, snRNP biogenesis and pre-mRNA splicing [Yang et al., 2005/ Wu et al., 2011]/ Basak and Query, 2014].

1.2.4 mRNA pseudouridylation

Since 1951, it is known that rRNAs and tRNAs contain pseudouridines [Cohn and Volkin, 1951]. Shortly thereafter, the existence of snRNA pseudouridylation was verified. For a long time scientists dealt with the question whether mRNAs are also pseudouridylated. The low abundance of mRNAs compared to rRNAs or tRNAs has made it extremely difficult to answer this question. High-throughput sequencing enabled the identification of mRNA pseudouridylation for the first time.

Identification of mRNA pseudouridylation

In 2014/2015, the first transcriptome-wide data on mRNA pseudouridylation were published by four independent groups [Schwartz et al., 2014/ Carlile et al., 2014/ Lovejoy et al., 2014/

Li et al., 2015a]. The methods (Pseudo-seq, Ψ-seq, PSI-seq, and CeU-seq) used in these publications rely on selective labeling of pseudouridines by the chemical compound CMC. Under physiological conditions, CMC acylates G at the N1 position and U/Ψ at the N3 position of the nucleotides. Exclusively, the Ψ-CMC adduct is stable to alkaline hydrolysis. The other acetylated nucleotides are susceptible to hydrolysis at a pH 10.4. In addition, CMC efficiently blocks reverse transcription (RT) one base downstream of the pseudouridine [Bakin and Ofengand, 1993]. This discontinuation can be detected by next generation sequencing [Figure 4].

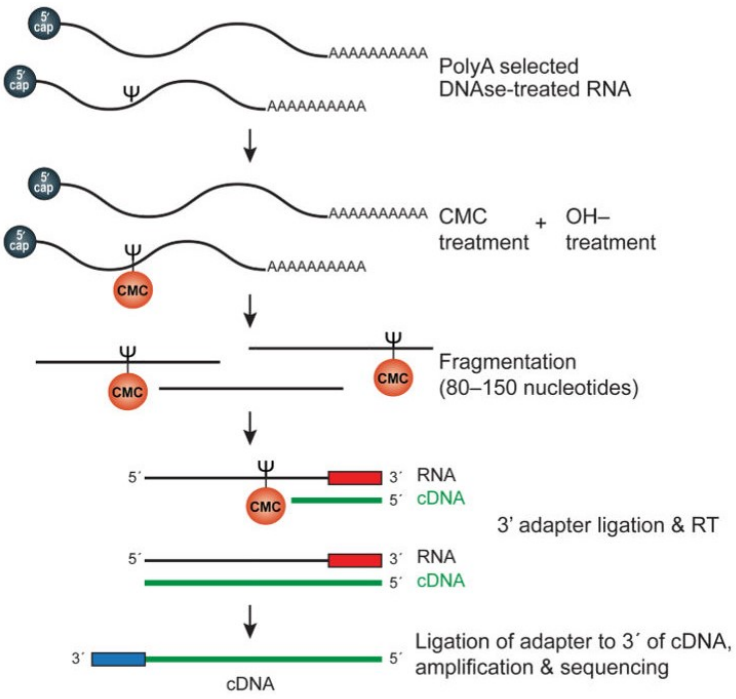

Figure 4: Scheme of the CMC-based pseudouridylation-sequencing method: Poly(A) mRNA is treated with CMC, which covalently binds to U, Ψ and G. The following incubation at an alkaline pH leads to hydrolysis of U-CMC/ G-CMC adducts. RNA is then fragmented to a size range of 80-150 nt, followed by adapter ligation to the 3' end. Reverse transcription (RT) terminates at the pseudouridylation site. A second adapter is ligated to the 3' end of the cDNA, libraries are amplified and sequenced [Schwartz et al. 2014].

Previously unknown pseudouridylation sites in rRNAs, tRNAs, snRNAs, long non-coding (lncRNA) and snoRNAs were detected with each of the four methods based on CMC labeling of RNAs [Figure 4]. Furthermore, pseudouridylation sites were identified in about 2000 human mRNAs and around 850 mRNAs in *S. cerevisiae* [Schwartz et al., 2014/ Carlile

et al., 2014/ Lovejoy et al., 2014/ Li et al., 2015a]. These mRNA pseudouridylation sites are located throughout the coding regions and were also found in 5' untranslated regions (UTR) and the 3' UTRs [Carlile et al., 2014/ Lovejoy et al., 2014].
However, only a small subset of the detected pseudouridylation-sites were identified by all four methods. Only one of 402 mRNA pseudouridylation sites in the coding region of *S. cerevisiae* was consistently found applying Pseudo-seq, Ψ-seq and PSI-seq. 38 yeast mRNA pseudouridylation sites were identified by at least two of these methods [Zaringhalam and Papavasiliou, 2016]. Reasons for these differences might be variability in the CMC labeling protocol, differences in the RNAseq library preparations, the sequencing approach or data processing. Furthermore, there are also chemical limitations of CMC, as Ψ-specific CMC-modification may not be 100% efficient [Zaringhalam and Papavasiliou, 2016]. Nevertheless, these methods allowed to detect the existence of mRNA pseudouridylation sites for the first time. In addition to standard growth conditions, two studies also mapped pseudouridylation sites under stress conditions [Schwartz et al. 2014/ Carlile et al., 2014]. snoRNA-independent pseudouridylation by the pseudouridine synthase Pus7 is induced during heat shock in *S. cerevisiae*. Pseudouridylated mRNAs are expressed at 25% higher levels in wild-type strains in comparison to *pus7* deletion strains during heat shock. The authors concluded that pseudouridylation may contribute to RNA stability [Schwartz et al. 2014]. By contrast, only few pseudouridylation events were induced during cold shock or at stationary phase [Schwartz et al. 2014]. The Gilbert group identified several hundred pseudouridylation events during post-diauxic shift in *S. cerevisiae*. 42% of these pseudouridylation sites are regulated, as they were not detected under exponential growth conditions [Carlile et al., 2014]. These results show that pseudouridylation of mRNAs may have a function under different environmental or stress conditions.
During the last eight years, methods to detect pseudouridylation are still evolving. In 2017, a quantitative PCR (qPCR)-based method to detect pseudouridylation sites was established, which also uses CMC-labeled RNA. Pseudouridylation sites were identified based on different melting temperatures of qPCR products [Lei and Yi, 2017]. Another method uses a modified carbodiimide that permits chemical coupling of biotin to pseudouridine residues to enrich Ψ-containing RNA fragments for sequencing [Li et al., 2015b]. Furthermore, RBS-Seq was established to simultaneously map the RNA modifications N1-methyladenosine, 5-methyl cytosine and pseudouridine through bisulfite treatment of RNA [Fleming et al., 2019]. In 2020, HydraPsiSeq for quantitative mapping of pseudouridylation was published. This method is based on cleavage at all uridine sites by a combination of hydrazine and aniline cleavage whereas Ψ, C, G and A show no or only a low hydrazine reactivity [Marchand et al., 2020].

Taken together, many methods to detect mRNA pseudouridylation sites were established during the last years. It remains to be seen, which method will become the standard in the future.

Function of mRNA pseudouridylation

So far, little is known about the physiological function of pseudouridines in mRNAs.
High-throughput sequencing results indicate that mRNA pseudouridylation is dynamic and inducible in response to various stress conditions such as heat stress or nutrient deficiency [Carlile et al., 2014/ Schwartz et al., 2014]. However, a detailed function of individual pseudouridylation events under stress conditions is not known, yet.
So far, it was suspected that Pus7-dependent pseudouridylation may stabilize transcripts during heat shock [Schwartz et al., 2014]. Nevertheless, this assumption is consistent with previous reports by Kariko et al., who published that *in vitro* transcribed pseudouridylated mRNA shows a higher translation rate and stability in human cells. Besides, modified mRNAs were less immunogenic than unpseudouridylated mRNAs [Kariko et al., 2008/ Kariko et al., 2012]. For this reason, N1-methyl-pseudouridine is incorporated into the mRNA coding for the transmembrane S protein of the COVID-19 mRNA vaccines comirnaty© and spikevax©. N1-methyl-Ψ was substituted for all uridines throughout the mRNA sequence, including the uridines in the stop codons. N1-methyl-pseudouridine was used instead of pseudouridine, as N1-methyl-Ψ increases the protection against nucleases [Polack et al., 2020/ Corbett et al., 2020]. Another study by Nakamoto and colleagues showed that mRNA pseudouridylation has a destabilizing effect on mRNAs. They published that mRNAs of *Toxoplasma gondii* containing Pus1-dependent pseudouridylation sites are less stable than *pus1* mutants [Nakamoto et al., 2017]. Taken together, it appears that Ψ can act in two opposite directions, enhancing mRNA stability in some instances and reducing mRNA stability in others, depending on the organism, gene and conditions.
In addition to the described function in mRNA stability, mRNA pseudouridylation has an impact on translation. The Weissmann group published that *in vitro* transcribed pseudouridylated mRNA shows a higher translation rate [Kariko et al., 2008]. Pseudouridylation in a sense codon can repress translation, whereas pseudouridylation can also convert a nonsense codon into a sense codon [Karijolich and Yu, 2011]. Additionally, a stop codon that contains a pseudouridine results in translational read-through. This observation indicates that pseudouridylation of mRNAs may have different effects on mRNA translation. The modified stop codon is recognized by a specific aminoacylated tRNA instead of the release factor, leading to continued translation. Mass spectrometric analyses showed that ΨAA and ΨAG direct serine and threonine incorporation, while ΨGA directs tyrosine and phenylala-

nine incorporation [Karijolich and Yu, 2011]. Another group crystallized ribosomes bound to the pseudouridylated stop codon ΨAG and its "cognate" anticodon of $tRNA^{Ser}$. They showed that unusual base pairing occurs between two purines during decoding, suggesting that the catalytic center of the ribosome is rather flexible [Fernandez et al., 2013].
More recently, it was reported that Ψ might affect the coding specificity during translation. Using an *E.coli* translation system, it was shown that a pseudouridine could alter how ribosomes or codons interact with cognate and near-cognate tRNAs, leading to amino acid substitution. The authors suggested that this amino acid substitution mechanism could be a valuable source for adaptation under stress conditions [Eyler et al., 2019].
To summarize, mRNA pseudouridylation plays a role in mRNA stability and translation. Other functions of mRNA pseudouridylation are still unknown and need to be unraveled.

1.2.5 The chromatin-binding-factor 5 (Cbf5) of *S. cerevisiae*

Discovery and structure of Cbf5

In 1993, Cbf5 was first isolated using yeast centromere DNA affinity column chromatography (CEN). This is a method to purify putative centromere and kinetochore binding proteins of chromatin extracts. Cbf5 was discovered as a major low-affinity CEN binding protein that is situated at the centromere during mitosis/meiosis [Jiang et al., 1993]. Other studies showed that the homologue of Cbf5 in rats, Nap57, localizes in the nucleolus, as Nap57 was co-immunoprecipitated with a previously identified nucleolar protein. Furthermore, immunofluorescence and immunogold electron microscopy with Nap57 specific antibodies showed that Nap57 localizes to the dense fibrillar components of the nucleolus, to coiled bodies and to the nucleoplasm [Meier and Blobel, 1994]. In 1997, Cadwell and colleagues published that Cbf5 is involved in rRNA biosynthesis and interacts genetically with the RNA polymerase I transcription factor Rrn3 [Cadwell et al., 1997].

Cbf5 consists of 483 amino acids (aa) with a high content of charged aa, and has a molecular weight of 54.7 kDa. It is composed of a conserved catalytic domain (*CAT*) and a *PUA* domain to interact with RNA. As many other centromere-binding proteins, it has a centromere-binding domain *KKE/D*, containing the repeating *KKE/D* sequence near the C-terminus [Jiang et al., 1993] [Figure 5]. Furthermore, Cbf5 has sequence homology to TruB of *E. coli* and to the human protein dyskerin in the CAT and PUA domain [Koonin, 1996/ Meier and Blobel, 1994].

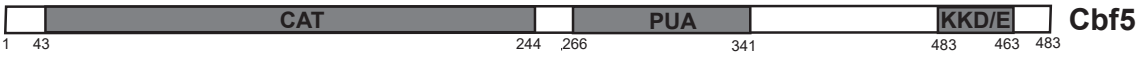

Figure 5: Domain structure of *CBF5* The localization of the *CAT, PUA* and *KKE/D* domain (grey boxes) within *CBF5* are illustrated. The numbers below represent the numbers of amino acids of the domains.

Cbf5 is a part of H/ACA snoRNPs

snoRNPs are ribonucleoproteins consisting of one snoRNA and several proteins. In 1979, snoRNPs were described as 110S ribonucleoprotein particles containing a 45S RNA [Fujisawa et al., 1979]. They function as endonucleases that cleave the pre-rRNA or as sequence-specific guides to modify rRNA transcripts [Tollervey and Kiss, 1997]. Based on conserved sequence motifs, nearly all known snoRNPs belong to the group of H/ACA snoRNPs or C/D snoRNPs. These snoRNPs are highly conserved and abundant in eukaryotes as well as in archaea [Balakin et al., 1996]. SnoRNPs containing the box C and the D motif direct the site-specific $2'$O-methylation of RNAs [Cavaille et al., 1996]. The majority of the H/ACA snoRNPs function as guide RNAs in the site-specific pseudouridylation of rRNAs, snRNAs and mRNAs [Ni et al. 1997]. There are only two known H/ACA snoRNAs (snR30 and snR10) that are required for the 18S rRNA cleavage [Tollervey 1987]. H/ACA snoRNAs of eukaryotes always consist of two hairpins, whereas they can have one to three hairpins in archaea.

The 133 nucleotide long H/ACA snoRNA folds into the conserved "hairpin-hinge-hairpin-tail" structure, which contains the ANANNA consensus sequence (H-Box) in the hinge region and ACA triplets (ACA box) in the tail region [Figure 6] [Ganot et al., 1997]. These motifs are important for nucleolar localization and stabilization of the H/ACA snoRNA as well as for proper pseudouridylation [Bortolin et al., 1999/ Lange et al., 1999]. The hairpins contain a single-stranded, internal loop, called the pseudouridylation pocket. 3-10 nucleotides of the pseudouridylation pocket are complementary to the substrate RNA to form an RNA duplex around the target uridine [Figure 6A] [Ganot et al., 1997]. The recruitment and the complementarity of the substrate RNA to the pseudouridylation pocket of the H/ACA

snoRNP results in a conformational change of the H/ACA snoRNA. The upper stems approach the CAT domain of Cbf5 and the lower parts of the stem loops undergo a rotation. This conformational change facilitates the proper positioning of the target uridine in the pseudouridylation-pocket [Duan et al., 2009/ Liang et al., 2009]. Two coaxially stacked helices are formed by the interaction between the pseudouridylation pocket of the H/ACA RNA and the substrate RNA. Moreover, two unusual three-way junctions are created during RNA duplex formation, stabilizing the otherwise flexible pseudouridylation pocket and placing the uridine that will be modified in a favorable position in the active site of Cbf5 [Jin et al., 2007/ Wu and Feigon, 2007]. At least eight base pairs between the guide H/ACA snoRNA and the substrate RNA are required for pseudouridylation. Pseudouridylation still occurs when unpaired nucleotides of different length are added to the upper part of the pseudouridylation pocket, which reveals the flexibility of the pseudouridylation pocket to accommodate a variety of substrate RNAs without perfectly matching residues [De Zoysa et al., 2018]. Another study showed that a minimum of 3 bp on each side of the target uridine are required for pseudouridylation. However, the initial velocity of pseudouridylation is reduced in these mutants, the formation of the RNA duplex can be observed [Kelly et al., 2019].

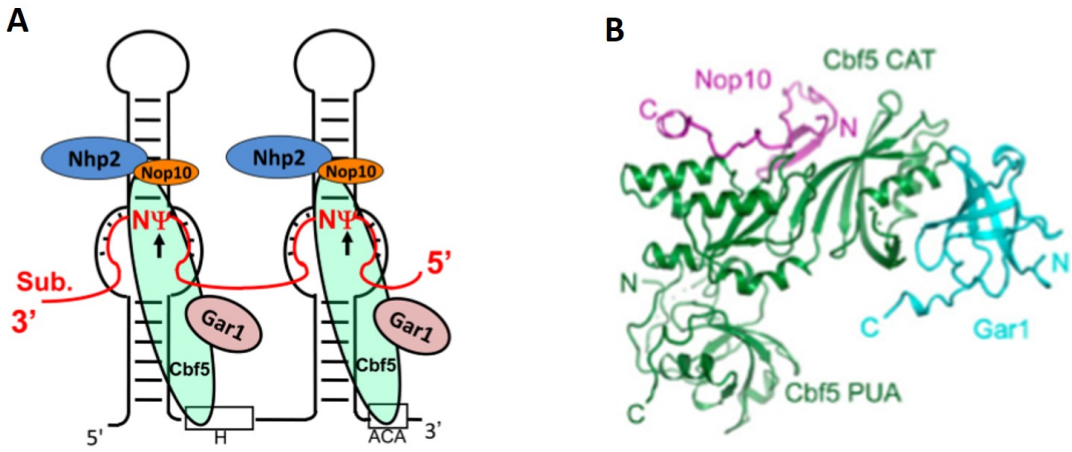

Figure 6: Structure of the H/ACA snoRNP: **A** shows the schematic structure of the H/ACA snoRNP. The color-coded ovals depict the four core proteins Cbf5, Nhp2, Nop10 and Gar1. The black line represents the H/ACA snoRNA in the typical "hairpin-hinge-hairpin-tail" structure. A substrate RNA (red line) is paired with the H/ACA snoRNA in the pseudouridylation pockets. Ψ (red) is the target nucleotide converted from uridine and N (red) represents any nucleotide. Boxes H and ACA of the guide RNA are also shown [Ge and Yu 2013]. **B** shows the crystal structure of the *S. cerevisiae* Cbf5–Nop10–Gar1 complex as a ribbon representation. Cbf5 is shown in green, Nop10 in magenta, and Gar1 in cyan. The N- and C-termini of each protein and the CAT and PUA domain are labeled [Li et al. 2011]

.

H/ACA snoRNPs of *S. cerevisiae* also contain the proteins Cbf5, Nop10, Nhp2 and Gar1 [Figure 6] [Watkins et al., 1998/ Henras et al., 1998]. The PUA domain of Cbf5 binds to the conserved sequence of the H box and the ACA box. The catalytic domain, carrying the catalytically essential aspartate residue of Cbf5, interacts with the pseudouridylation pocket of the H/ACA snoRNA and the uridine of the target RNA to form a reactive state [Li and Ye, 2006/ Duan et al., 2009]. Nop10 and Nhp2 interact with each other and the upper stem of the H/ACA snoRNA hairpin, thereby anchoring the snoRNA to the substrate RNA. In addition, Nop10 stabilizes the CAT domain of Cbf5 by direct interaction with Cbf5 and functions as a platform to bind Nhp2 [Figure 6] [Li et al., 2011]. Nhp2 consists of a conserved RNA-binding domain that binds RNA unspecifically in the absence of the other proteins, while specific binding to the snoRNA seems to be governed by formation of the Nhp2-Nop10-Cbf5 subcomplex [Wang and Meier, 2004/ Henras et al., 2001]. By contrast, Gar1 interacts only with Cbf5 and the substrate RNA to facilitate the formation of a catalysis competent conformation [Figure 6] [Li et al., 2011]. Furthermore, fluorescence correlation spectroscopy experiments showed that Gar1 destabilizes the product complex, which promotes substrate release [Yang et al., 2012]. Gar1 contains several glycine-arginine rich (RGG) domains, both at its N- and C-terminus, which enhance RNA folding, resulting in higher activity of the H/ACA snoRNP [Trucks et al., 2021].

Cbf5, Nop10 and the H/ACA snoRNA are essential to form a snoRNP *in vitro* that catalyzes the isomerization of uridine to pseudouridine. *In vivo* experiments show that Cbf5, Gar1 and Nop10 can form a complex in absence of Nhp2 and the H/ACA snoRNA. A depletion of one protein of the H/ACA snoRNP inhibits the expression of Cbf5, Gar1, or Nop10, whereas Nhp2 levels are little affected [Henras et al., 2004]. So far, only a crystal structure of a ternary Nop10-Cbf5-Gar1 subcomplex, comprising only the core domains of Cbf5 and Gar1, has been solved [Li et al., 2011]. Furthermore, a nuclear magnetic resonance (NMR) structure of Nhp2 in *S. cerevisiae* is published [Koo et al., 2011]. The crystal structure revealed that the eukaryotic complex displays considerable structural similarity to archaeal RNPs [Li et al., 2011/ Duan et al., 2009]. Despite these results of the structure, dynamics and function of the H/ACA snoRNP, the exact function of eukaryotic-specific features in H/ACA snoRNA-guided pseudouridylation remains elusive.

Assembly of the H/ACA snoRNP

The initial assembly model for the H/ACA RNP is based on interactions identified *in vitro*. These interactions suggested that Nop10 interacts with Cbf5, followed by the association of Nhp2, which binds to Nop10. Thus, the trimeric core RNP is generated [Henras et al.,

2001/ Wang and Meier, 2004]. The assembly factor Shq1 binds to the PUA domain and to the C-terminal end (CTE) of Cbf5, preventing interaction with RNAs [Walbott et al., 2011/ Li et al., 2011]. Afterwards, all four proteins co-translocate into the nucleus [Yang et al., 2002/ Darzacq et al., 2006]. The immature RNP is recruited to the H/ACA RNA transcription site. The H/ACA snoRNA is transcribed and the immature H/ACA snoRNP is formed containing the H/ACA snoRNA, Cbf5, Nhp2 and Nop10 as well as the assembly factor Naf1, which binds to Cbf5 [Ballarino et al., 2005]. A mature complex is marked by the displacement of the assembly factor Naf1 by Gar1. Naf1 and Gar1 compete for binding to Cbf5. Once Gar1 associates with the mature H/ACA snoRNP, the complex localizes to the sites of function [Yang et al., 2002/ Darzacq et al., 2006].

Function of Cbf5

The main function of Cbf5 as a part of the H/ACA snoRNP is the isomerization of uridines to pseudouridines. Cbf5-dependent pseudouridylation sites are detected in rRNAs and mRNAs [Schattner et al., 2004/ Schwartz et al., 2014]. So far, it is not known whether tRNAs and snRNAs contain Cbf5-dependent pseudouridylation sites. The conserved pseudouridines of U2 snRNA are only generated by H/ACA snoRNPs in vertebrates, whereas the guide RNA-independent pseudouridine synthase Pus7 catalyzes the formation of these pseudouridines in *S. cerevisiae* [Wu et al., 2011/ Donmez et al., 2004]. 44 different pseudouridylation-sites exist in rRNAs. Using snoGPS, a program for computational screening of genomic sequences for H/ACA guide snoRNAs, 41 of them were found to be catalyzed by H/ACA snoRNPs [Schattner et al., 2004]. These Cbf5-dependent rRNA pseudouridylation events have a strong impact on ribosome activity by altering rRNA folding and interaction, which in turn influences translation precision and efficiency [King et al., 2003/ Jack et al., 2011]. Besides, pseudouridylation of rRNAs have an important role in rRNA biogenesis and processing [Liang et al., 2009]. About 37 Cbf5-dependent mRNA pseudouridylation were identified performing Ψ-seq in a *cbf5* depletion mutant [Schwartz et al., 2014]. Until now, a function of Cbf5-dependent mRNA pseudouridylation is not known.

Cbf5 also has pseudouridylation-independent functions. Viability of a *cbf5* deletion can be rescued by expressing a catalytically inactive *cbf5* variant [Zebarjadian et al., 1999]. In 2009, it was published that the H/ACA snoRNP containing the snoRNA snR30 (U17) may act as a guide RNA for 18S rRNA processing that is not related to pseudouridylation. The 3' hairpin of snR30 contains sequences that act as conserved 18S rRNA recognition elements in the bulge of the hairpin, which base-pair with the pre-rRNA. In the upper stem of the 3' hairpin, there is a putative binding site for a protein that might be involved in the nucleolytic processing of the pre-rRNA 35S into the 18S rRNA [Fayet-Lebaron et al., 2009].

Another example of a putative pseudouridylation-independent function of Cbf5, is its role in chromosome segregation. In 1993, Cbf5 was discovered as a centromere-binding protein [Jiang et al., 1993]. Jiang et al. could also show that most cells of a temperature-sensitive *S. cerevisiae* strain containing a mutated *cbf5*, arrest in G_1/S phase of the cell cycle. A yeast strain with a C-terminally truncated *cbf5* delays at the G_2/M phase of the cell cycle [Jiang et al., 1993]. Two years later, the same group published that overexpression of *mck1*, which encodes a meiosis and centromere regulatory kinase, suppresses the temperature-sensitive phenotype of *cbf5* [Jiang et al., 1995]. In the last 27 years, only one article about the function of the human homologue of Cbf5, dyskerin, in chromosome segregation was published. It was shown that dyskerin tethers pericentric chromatin to the spindle axis during mitosis [Snider et al., 2014]. Besides, dyskerin is a part of the human telomerase and regulates its activity. It plays a role in stabilization, maintenance and quality control of the telomerase snoRNA htr [Kroustallaki et al., 2019/ Shukla et al., 2016]. Furthermore, dyskerin is involved in regulating hematopoietic stem and progenitor cell (HSPC) differentiation into primary erythroblasts and also has a role in self-renewal and differentiation of mesenchymal stem cells [Richards et al., 2020/ Zhang et al., 2017]. It is not known whether Cbf5 has also a function in regulation of telomerase in *S. cerevisiae*

In sum, besides the main function of Cbf5 as a pseudouridine synthase of mRNAs and rRNAs, Cbf5 has a pseudouridylation-independent function in rRNA biogenesis and chromosome aggregation

1.3 Cbf5 interacts with TREX

A former master student of our lab performed mass-spectrometric analysis of the purified and cross-linked TREX complex to unravel the structure of TREX. In addition to crosslinks between the components of TREX, crosslinks between Cbf5 at the amino acid (aa) residue 419 and the proteins Tho2, Sub2, Hpr1 and Thp2 were identified [unpublished data Manuel Koschitza] [Figure7].

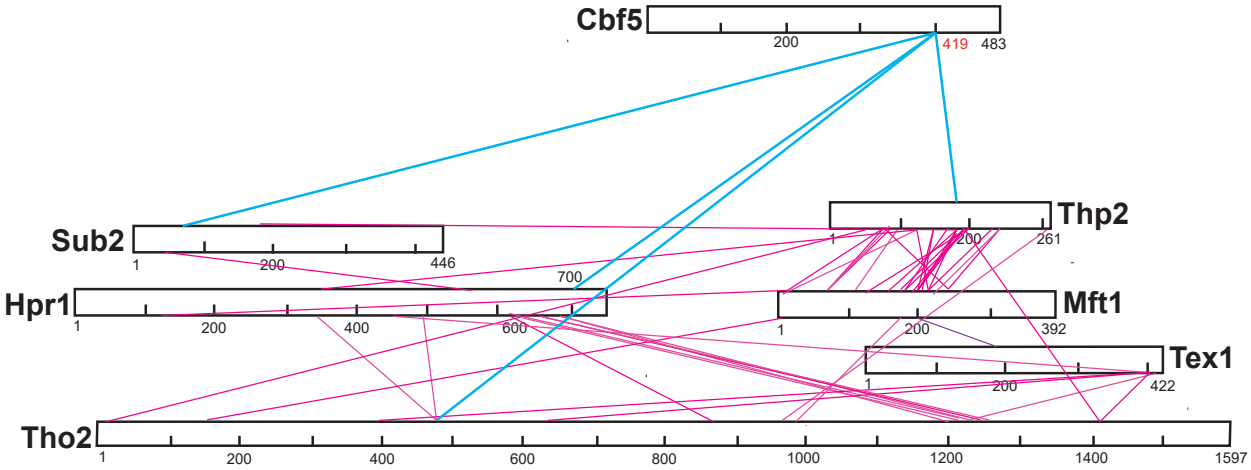

Figure 7: Crosslinks between Cbf5 and Sub2, Hpr1, Tho2 and Thp2 of TREX were identified. XL/MS with the crosslinker BS^3 was performed of the purified TREX complex to identify crosslinks between the proteins of TREX. The figure shows the identified crosslinks between proteins of TREX (pink lines). Crosslinks between Sub2, Hpr1, Tho2 and Thp2 of TREX and Cbf5 are shown as light blue lines. Proteins with the corresponding number of amino acids are represented by boxes. Purification and crosslinking of TREX was performed by Manuel Koschitza. Mass-spectrometry and data analysis were conducted by Alexander Leitner (ETH Zürich).

To examine whether Cbf5 interacts with TREX, I did tandem affinity purifications and subsequent Western Blot analyses in presence or absence of RNase A in my master thesis. These experiments showed that Cbf5 interacts with TREX in an RNA-independent manner, whereas the interaction of Cbf5 with other components of the mRNP, i.e. Npl3, Nab2 and Cbp80 was RNA-dependent [Gruellich, 2017]. Furthermore, methylation-tracking (M-track)-assays to detect an *in vivo* protein/protein interaction showed that Cbf5 interacts with Sub2 and Thp2. This experiment, performed by the group of Claudine Kraft, could verify that Cbf5 interacts with TREX [Gruellich, 2017]. The function of this RNA-independent interaction is not known, yet.

1.4 Aims and scope of this thesis

The overall aim of this thesis is to unravel the mechanism and function of Cbf5-dependent mRNA pseudouridylation and to elucidate the role of the Cbf5/TREX interaction.

In *S. cerevisiae*, Cbf5 is the enzyme of the H/ACA snoRNP, which catalyzes the isomerization of uridine to pseudouridine [Ni et al., 1997/ Watkins et al., 1998]. In 2014/2015, it was published that Cbf5-mediated mRNA pseudouridylation exists [Schwartz et al., 2014]. However, the function of Cbf5-dependent mRNA pseudouridylation is still unknown. Besides, it is not known whether Cbf5 has a pseudouridylation-independent function in the mRNA life cycle. In my master thesis I showed that Cbf5 interacts with TREX in an RNA-independent manner [Gruellich, 2017]. TREX has a variety of functions spanning from transcription to mRNA nuclear export. [Chavez et al., 2000/ Straesser and Hurt, 2000/ Straeser and Hurt, 2001].
The first aim of this thesis is to identify whether the whole H/ACA snoRNP interacts with TREX. Therefore, XL-MS of the purified H/ACA snoRNP and TREX and tandem-affinity purifications (TAPs) with subsequent Western Blots will be conducted.
Secondly, I want to investigate the recruitmernt mechanism of Cbf5 to the transcribed genes and mRNAs performing ChIPs and RIPs. In particular, I want to assay whether TREX has a role in Cbf5 recruitment.
Thirdly, it is still unknown when Cbf5-dependent pseudouridylation occurs. To study whether it occurs co-transcriptionally or post-transcriptionally primer extension after CMC labeling of purified nascent mRNA, poly(A) mRNA and total RNA will be performed.
The fourth aim is to elucidate the function of the Cbf5/TREX interaction. TREX has a role in transcription elongation and mRNA nuclear export. RNAPII ChIPs at the large gene *YLR454* and ChIPseq of RNAPII will be conducted to investigate whether Cbf5 also functions in mRNA transcription. Furthermore, FISHs will be performed to examine whether Cbf5 has a role in nuclear mRNA export.
As the function of mRNA pseudouridylation is still unknown, the last aim of this thesis is to investigate the role of Cbf5-dependent mRNA pseudouridylation in the mRNA life cycle. To study this, point mutations will be inserted at Cbf5-dependent pseudouridylation (PseudoU) sites of mRNAs. Thiolutin assays, *in vivo* transcription assays and cycloheximide assays of these PseudoU site mutants will be performed to investigate whether Cbf5-dependent pseudouridylation functions in transcription, mRNA- or protein stability.

2 Results

2.1 TREX interacts with the H/ACA snoRNP

H/ACA snoRNPs, which consist of a H/ACA snoRNA and the proteins Cbf5, Gar1, Nhp2 and Nop10, catalyze the formation of rRNA and mRNA pseudouridylation events [Watkins et al., 1998]. TREX is a highly conserved complex that couples mRNA transcription and mRNA nuclear export [Chavez et al., 2000 and Straesser et al., 2002]. In my master thesis I showed that Cbf5 interacts with TREX in an RNA-independent manner [Gruellich, 2017].

To identify further crosslinks between proteins of TREX and H/ACA snoRNPs, XL/MS of both complexes was performed. To this end, TREX and H/ACA snoRNPs were purified using tandem affinity purification (TAP), with a total yield of 107 μg H/ACA snoRNP and 73 μg TREX purified from 60 L and 40 L yeast culture at OD_{600}=3.5, respectively. Figure 8A/B shows SDS-polyacrylamide gels of the TAP eluates/lysates of proteins from *CBF5-FTpA* [Figure 8A] and *HPR1-FTpA* [Figure 8B]. Afterwards, cross-linking of 60 μg protein of each complex with 500 μM BS^3 was conducted. Mass-spectrometry of the purified and cross-linked complexes and data analysis was performed by Alexander Leitner (ETH Zürich). Figure 8C shows a scheme of the TREX and H/ACA snoRNP proteins with the corresponding number of amino acids and the identified crosslinks. Crosslinks between proteins of one complex are shown in purple, whereas crosslinks between complexes are drawn as blue lines. Crosslinks between all proteins of the H/ACA snoRNP except Gar1 and Nhp2 were identified. Crosslinks between Thp2 and Mft1 as well as the C-terminus of Hpr1 and Tho2 were observed within TREX. Furthermore, two crosslinks between Gar1 at aa residue 59 and the C-terminus of Tho2 could be identified. This finding indicates that Gar1 interacts with Tho2.

To verify the XL/MS result and to answer the question whether H/ACA snoRNPs interact with TREX, H/ACA snoRNPs of the TAP-tagged strains *GAR1-TAP HPR1-HA* and *NHP2-TAP* were purified in presence and absence of 0.1 mg/ml RNase A. Subsequent Western Blots were conducted to assess a co-purification of TREX components or other proteins of the mRNP. Figure 9A shows a polyacrylamide gel of the TAP lysates/eluates of proteins from *GAR1-TAP HPR1-HA* +/- RNase A. The purified protein *GAR1-CBP* of a size of around 28 kDa and the H/ACA snoRNP component Cbf5 of a size of around 65 kDa and Nhp2 (25 kDa) are visible on the gel after the first (TEV) and the second (EGTA) purification step. The allocation of the purified proteins is given according a TAP of Cbf5 and a subsequent liquid chromatography mass-spectrometry (LC–MS/MS) analysis of the purified

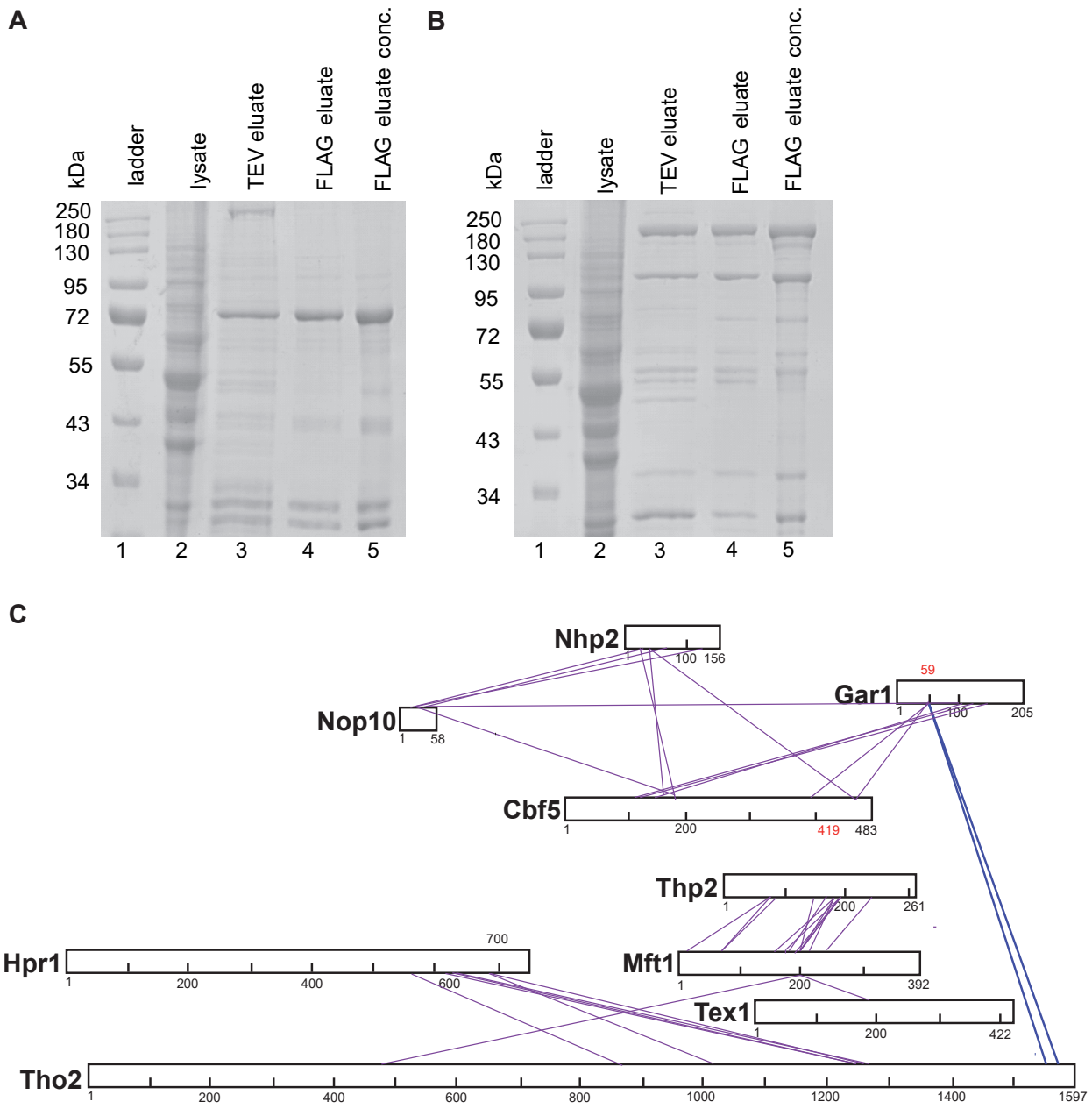

Figure 8: Crosslinks between Gar1 of the H/ACA snoRNP and Tho2 of TREX were identified. XL/MS with the crosslinker BS^3 was performed of the purified TREX and H/ACA snoRNP to identify crosslinks between these complexes. TAPs were performed to purify the H/ACA snoRNP and the TREX complex. **A**&**B:** Coomassie stained SDS-polyacrylamide gels of the purified and concentrated H/ACA snoRNP (A) and the TREX complex (B). 2 μl of the yeast cell lysate, 10 μl of the eluate after the first (TEV eluate) and second (FLAG eluate) purification step and 2.5 μl of the concentrated eluate (FLAG eluate conc.) were used to perform SDS-PAGES. **C:** Scheme of the identified crosslinks between both complexes (blue lines) as well as within one complex (purple lines). Proteins with the corresponding number of amino acids are represented by boxes. Mass spectrometry and data analysis was performed by Alexander Leitner (ETH Zürich).

proteins [Schwer et al., 2011]. A purification of proteins from *HPR1-HA* served as control. Western Blot analysis confirmed an RNA-independent co-purification of the TREX proteins *HPR1-HA* and Yra1. Other proteins of the mRNP as Npl3 or Cbp80 and also the ribosomal protein Rps8 co-purify in an RNA-dependent manner.
Figure 9B shows a Coomassie-stained gel of the TAP lysates/eluates of proteins from *NHP2-TAP* +/- RNase A. The purified protein *NHP2-CBP* of a size of around 27 kDa is visible on the polyacrylamide gel after the first and the second purification step, whereas Nop10 of a size of around 18 kDa is only visible on the gel after the first purification step. Cbf5 was not detected on the gel of the Nhp2 purification in presence of 0.1 mg/ml RNase A in the EGTA eluate. As the molecular weight of Gar1 and *NHP2-CBP* is approximately equal, they are visible as one band on the gel. Purification of proteins of an untagged wild-type (WT) strain served as control. Western Blot analysis of proteins co-purifying with *NHP2-CBP* demonstrated an RNA-independent interaction with Yra1, whereas Npl3, Cbp80 and Rps8 co-purified in an RNA-dependent manner.
To verify the co-purification of TREX proteins with the purified H/ACA snoRNP, a TAP of TREX proteins of the strain *HPR1-TAP GAR1-HA* +/- RNase A was performed [Figure 9C]. In this experiment the strain *GAR1-HA* served as control. All components of TREX are visible on the Coomassie-stained gel as it was described in Straesser et al., 2002. Western Blot analysis of the purified complex with an antibody against the HA-tag revealed an RNA-independent co-purification of *GAR1-HA*.

The purifications with a subsequent Western Blot analysis confirmed the XL/MS results. They show that Gar1 and Nhp2 of the H/ACA snoRNP also interact with proteins of TREX in an RNA-independent manner.

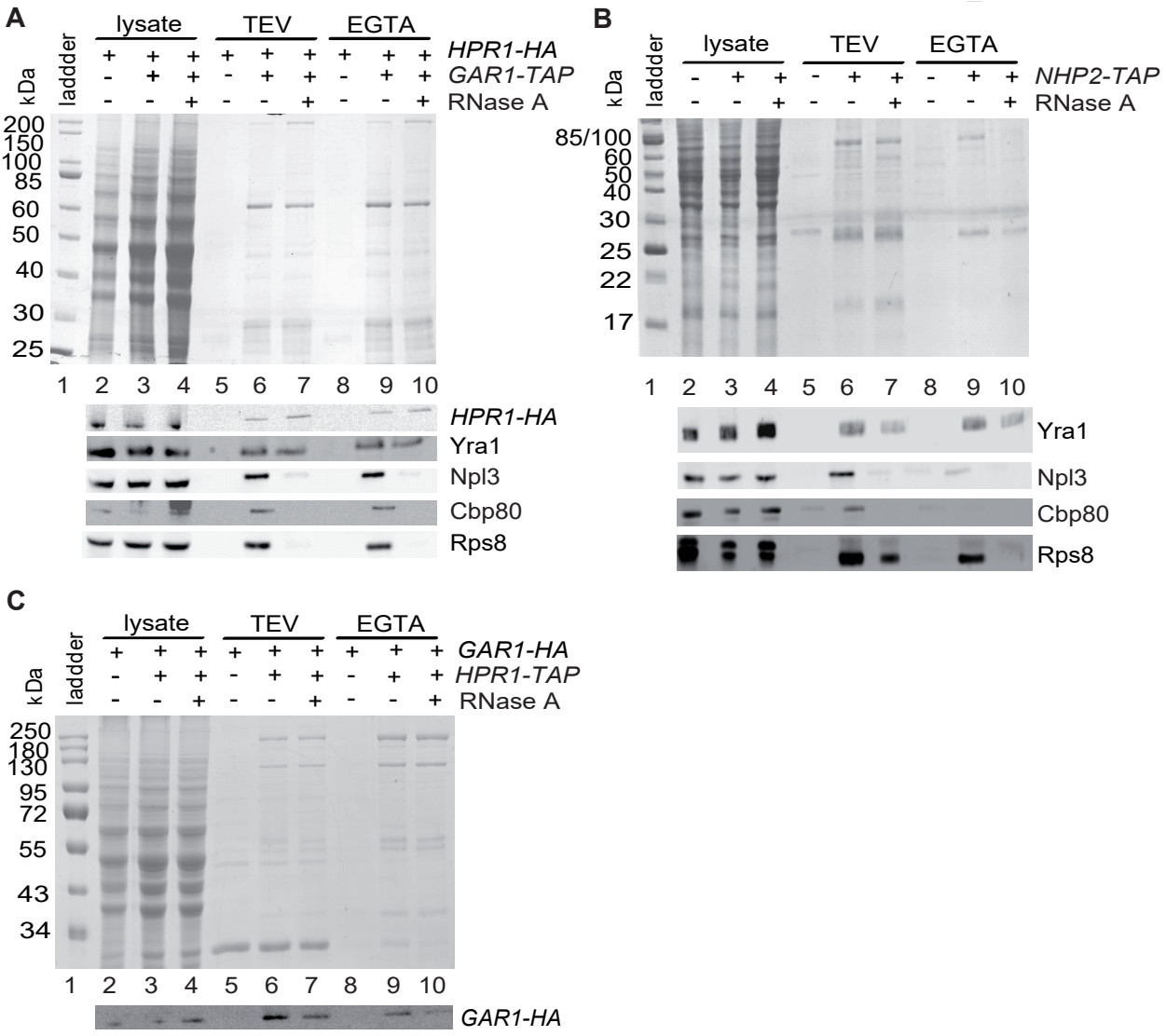

Figure 9: Gar1 and Nhp2 of the H/ACA snoRNP interact with Hpr1 and Yra1 of TREX. Tandem affinity purifications of TREX/H/ACA snoRNPS from **A:** *GAR1-TAP HPR1-HA* **B:** *NHP2-TAP HPR1-HA* and **C:** *HPR1-TAP GAR1-HA* in presence and absence of 0.1 mg/ml RNase A were conducted to detect an RNA-independent co-purification performing Western Blots. Three Coomassie stained SDS-polyacrylamide gels with the corresponding chemiluminescent signals on the nitrocellulose membranes of the Western Blots are shown to detect a co-purification. 2 μl of the yeast cell lysates and 15 μl of the eluates after the first (TEV) and second (FLAG) purification step were analyzed by SDS-PAGE. The purification and the Western Blot of figure C was performed by the bachelor student Jana Müller.

2.2 Cbf5 recruitment occurs transcription-dependent and RNA-dependent

The next aim of this thesis was to elucidate the function of the Cbf5/TREX interaction. As THO is needed to load the proteins Hrb1 and Gbp1 as well as Sub2 and Yra1 co-transcriptionally onto the nascent RNA transcript [Hurt et al., 2004/ Straeser and Hurt, 2001], I wanted to analyze whether THO/TREX plays the same role for Cbf5. In addition, this would lead to the question whether Cbf5 is recruited to transcribed genes/mRNAs.

2.2.1 The occupancy of Cbf5 at RNAPII transcribed genes is transcription-dependent

To determine the occupancy of Cbf5 at transcribed genes, endogenous Cbf5 was TAP-tagged and a ChIP was performed. Cbf5 occupancy was analyzed at three exemplary highly transcribed genes [Figure 10B] and at five genes containing a Cbf5-dependent pseudouridylation site [Figure 10C]. Primer extension after CMC treatment was performed in the *CBF5* depletion strain (*osTIR CBF5-AID*) to identify reliable Cbf5-dependent pseudouridylation sites based on the published Pseudo-seq data by Schwartz et al., 2014 [Appendix Figure 31/32]. The eight analyzed genes showed a highly significant increase in Cbf5 fold enrichment compared to the untagged strain. This result shows that Cbf5 is present at all tested transcribed genes, independent of whether the mRNA of this transcript contains a Cbf5-dependent pseudouridylation site.

To determine whether Cbf5 occupancy at transcribed genes is dependent on transcription, Cbf5 occupancy at the inducible *GAL1* gene was assessed. For this, the strain *CBF5-TAP* was grown in media with glucose (YPD) or galactose (YPG). *GAL1* is highly transcribed in YPG, but the gene is repressed in glucose-containing medium. Cbf5 is not recruited to *GAL1* under repressed conditions, whereas the occupancy of Cbf5 at *GAL1* grown in media containing galactose is significantly increased [Figure 10D]. This result shows that Cbf5 occupancy at genes is transcription-dependent.

RIP was performed to study whether Cbf5 is recruited to mRNAs. Cbf5 of the strain *CBF5-TAP* was co-immunoprecipitated from whole-cell extracts with IgG-sepharose. An untagged wild-type strain served as control. The presence of the three highly abundant mRNAs and the five mRNAs containing Cbf5-dependent pseudouridylation sites in the co-immunoprecipitate was assessed using reverse transcription (RT) followed by quantitative polymerase chain reaction (qPCR).

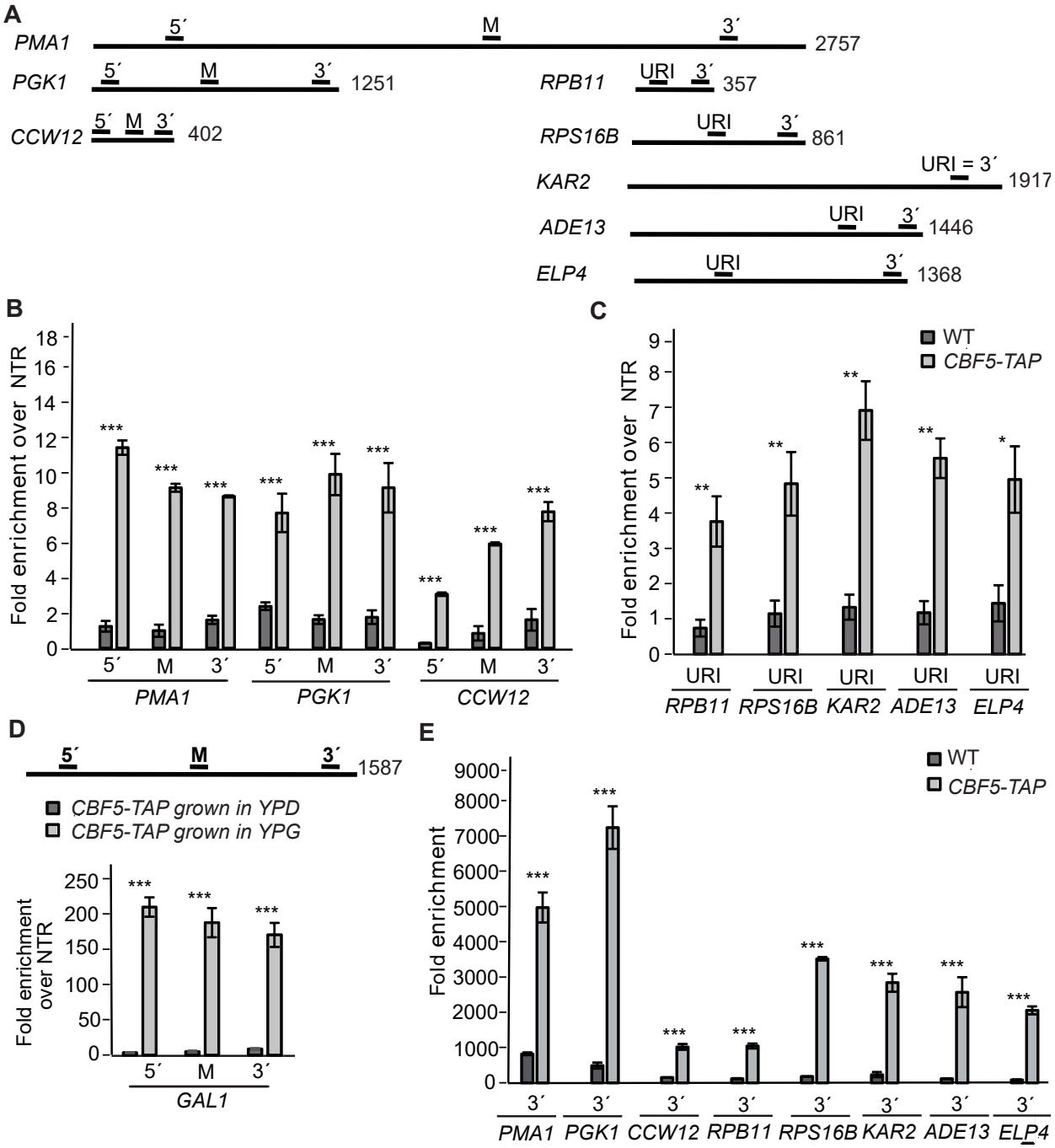

Figure 10: Cbf5 is present at transcribed genes in a transcription-dependent manner. **A:** Scheme of the tested genes. *PMA1, PGK1 and CCW12* are representatives of highly transcribed genes. *RPB11, RPS16B, KAR2, ADE13, ELP4* are genes containing a Cbf5-dependent pseudouridylation site. Open reading frames (ORFs) are represented by a solid line. Positions of primer pairs used for qPCR are indicated as bars above the solid lines. 5': 5' end of the ORF, M: middle of the ORF, 3': 3' end the ORF, URI: Cbf5-dependent pseudouridylation site **B&C:** Cbf5 enrichment at the highly transcribed genes (B) and genes containing a Cbf5-dependent pseudouridylation site (C) was determined by ChIP and subsequent qPCR. The occupancy of Cbf5 is increased in *CBF5-TAP*. The fold enrichment of Cbf5 at the tested genes was normalized to an untranscribed region (NTR). **D:** ChIP and qPCR of Cbf5 whose strain (*CBF5-TAP*) was grown in media with glucose (YPD)

or galactose (YPG). The occupancy of Cbf5 at *GAL1* increases in presence of galactose. The fold enrichment of Cbf5 at the tested genes was normalized to an untranscribed region.
E: Cbf5 enrichment at the tested mRNAs was determined by RIP and subsequent RT-qPCR. Cbf5 occupancy increases at all tested mRNAs. **B-E:** Results of three independent biological replicates are shown. Bars represent +/- standard deviation of students t-test. (***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$)

Figure 10E shows that all tested mRNAs are much more enriched when Cbf5 is TAP-tagged than in the untagged strain. The RIP indicates that Cbf5 is recruited to mRNAs with a broad specificity.
Taken together, Cbf5 is present at all tested genes and mRNAs independently of whether the transcript contains a Cbf5-dependent pseudouridylation site. The occupancy of Cbf5 at the transcribed genes is transcription-dependent.

2.2.2 Identification of the Cbf5 recruiter to transcribed genes and mRNAs

It was already shown that the H/ACA snoRNP interacts with TREX in an RNA-independent manner, whereas the interaction between the H/ACA snoRNP and Cbp80 of the cap-binding complex is RNA-dependent. Additionally, it is published that Cbf5 interacts genetically and physically with Cbp20 of CBC [Fortes et al., 1999/ Schwer et al., 2011]. Both complexes function as recruiters of proteins onto nascent RNA transcripts [Hurt et al., 2004/ Sen et al., 2019]. To investigate whether TREX or the CBC is needed for Cbf5 occupancy at transcribed genes, ChIPs of Cbf5 in *HPR1/THO2* or *CBP20/CBP80* deletion strains were performed. Cbf5 occupancy is unchanged in an *HPR1* deletion strain both at the highly transcribed genes [Figure 11A] as well as at genes containing a Cbf5-dependent pseudouridylation site [Figure 11B]. However, reduced Cbf5 enrichment was detected at *PMA1* 5' in the same strain. Cbf5 fold enrichment in the *THO2* deletion strain is decreased at *PMA1, PGK1, CCW12 3' and RPS16B* compared to the wild-type. The other tested genes showed nearly the same Cbf5 occupancy in the wild-type and the mutant [Figure 12A/B]. Cbf5 occupancy in a *CBP80* and *CBP20* deletion strain is decreased at *PGK1, CCW12 5', RPB16B and ADE13* (only in *Δcbp80*). *CBF5-TAP Δcbp80* showed a significantly increased Cbf5 occupancy at *CCW12 M and 3'* compared to the wild-type strain. The other tested genes showed nearly the same Cbf5 occupancy in the wild-type and the mutants [Figures 13A/B]. Furthermore, the occupancy of RNAPII was determined. As Cbf5 occupancy is transcription-dependent, reduced levels of RNAPII in *HPR1, THO2, CBP20 and CBP80* deletion strains compared to the wild-type could lead to reduced levels of Cbf5 at the genes even if the proteins had no direct role in Cbf5 recruitment to transcribed loci.

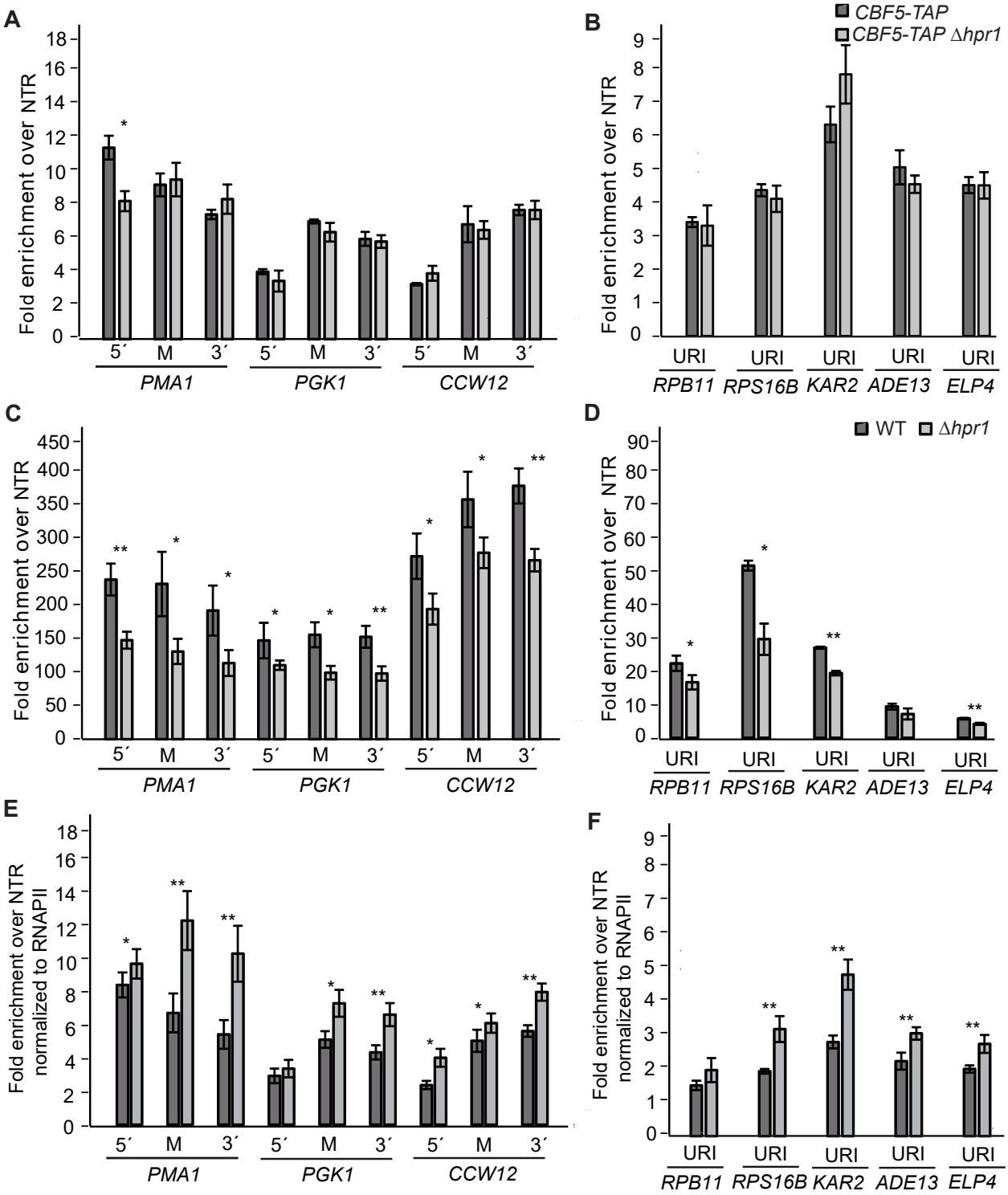

Figure 11: Hpr1 is not needed for Cbf5 occupancy at transcribed genes. The occupancy of Cbf5 (A&B) and RNAPII (C&D) at the tested genes was determined by ChIP. The fold enrichment was normalized to an untranscribed region (NTR). **A**&**B:** The occupancy of Cbf5 is unchanged in a *HPR1* deletion strain. **C**&**D:** The occupancy of RNAPII is decreased in a *HPR1* deletion strain. **E**&**F:** Cbf5 occupancy normalized to the occupancy of RNAPII is increased. Bars represent the mean +/- standard deviation of three independent biological replicates, p values of students t-test: **: $p < 0.01$; *: $p < 0.05$

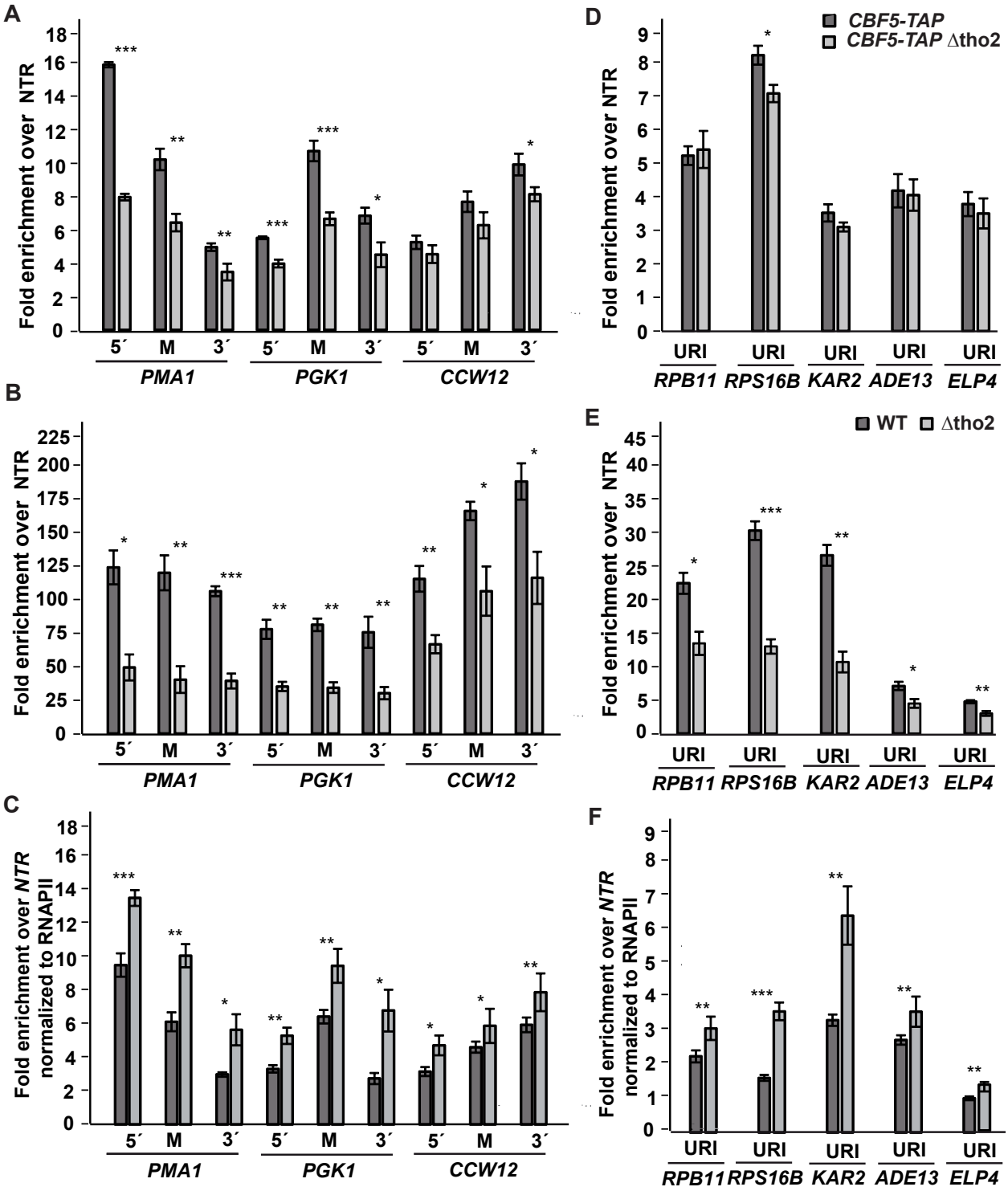

Figure 12: Tho2 is not needed for Cbf5 occupancy at transcribed genes. The occupancy of Cbf5 (A&B) and RNAPII (C&D) at the tested genes was determined by ChIP. The fold enrichment was normalized to an untranscribed region (NTR). **A**&**B:** Occupancy of Cbf5 is unaffected or decreased in a *THO2* deletion strain. **C**&**D:** The occupancy of the RNAPII is decreased in a *THO2* deletion strain. **E**&**F:** Cbf5 occupancy normalized to the occupancy of the RNAPII is increased. Bars represent the mean +/- standard deviation of three independent biological replicates, p values of students t-test: ***: $p < 0.001$ **: $p < 0.01$; *: $p < 0.05$

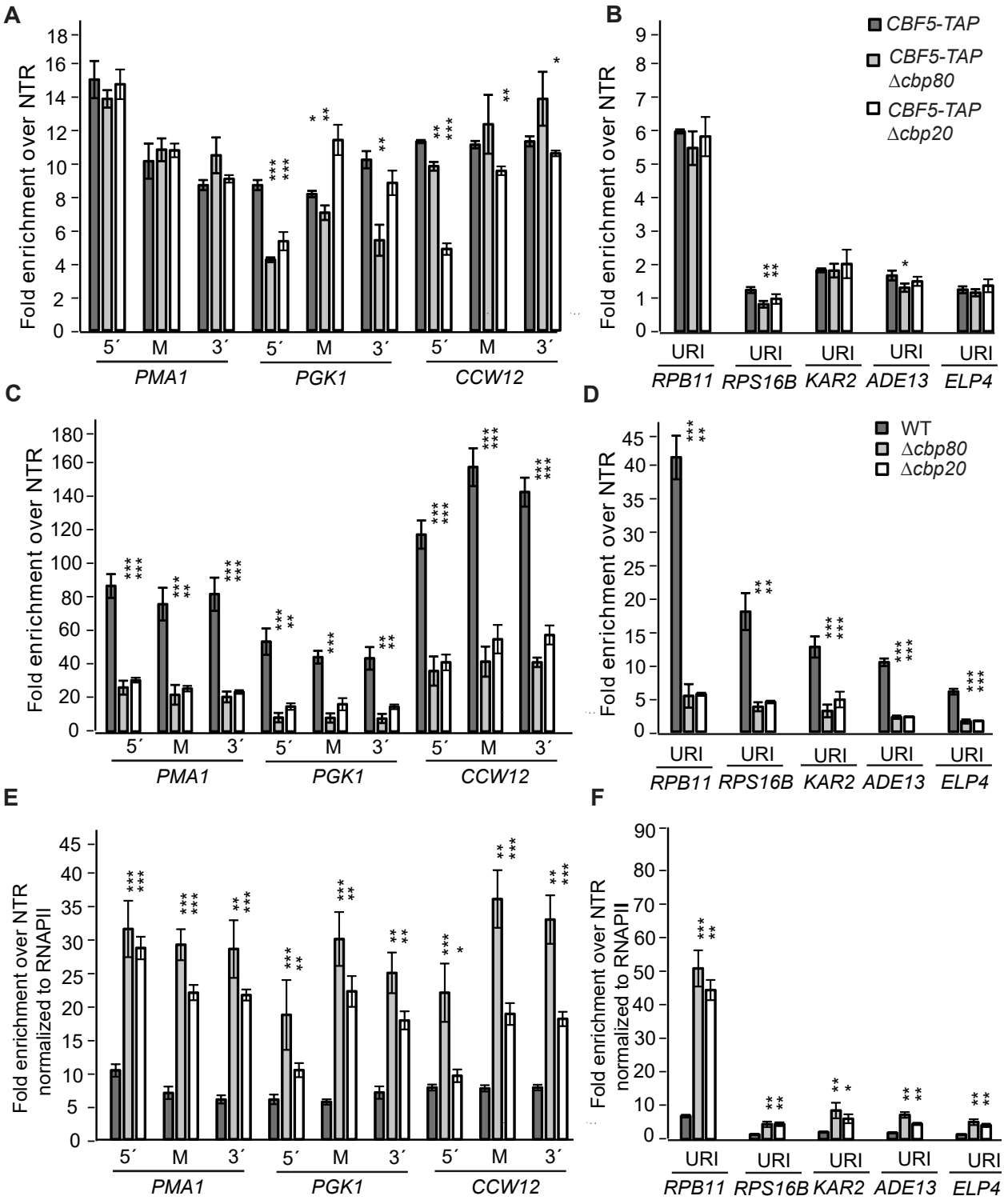

Figure 13: Cbp80 and Cbp20 are not needed for Cbf5 occupancy at transcribed genes. The occupancy of Cbf5 (A&B) and RNAPII (C&D) at the tested genes was determined by ChIP. The fold enrichment was normalized to an untranscribed region (NTR). **A**&**B:** The occupancy of Cbf5 is unaffected or decreased depending on the gene in a *CBP80* or *CBP20* deletion strain. **C**&**D:** The occupancy of the RNAPII is decreased in a *CBP80* or *CBP20* deletion strain. **E**&**F:** Cbf5 occupancy normalized to the occupancy of RNAPII is increased. Bars represent the mean +/- standard deviation of three independent biological replicates, p values of students t-test: ***: $p < 0.001$ **: $p < 0.01$; *: $p < 0.05$

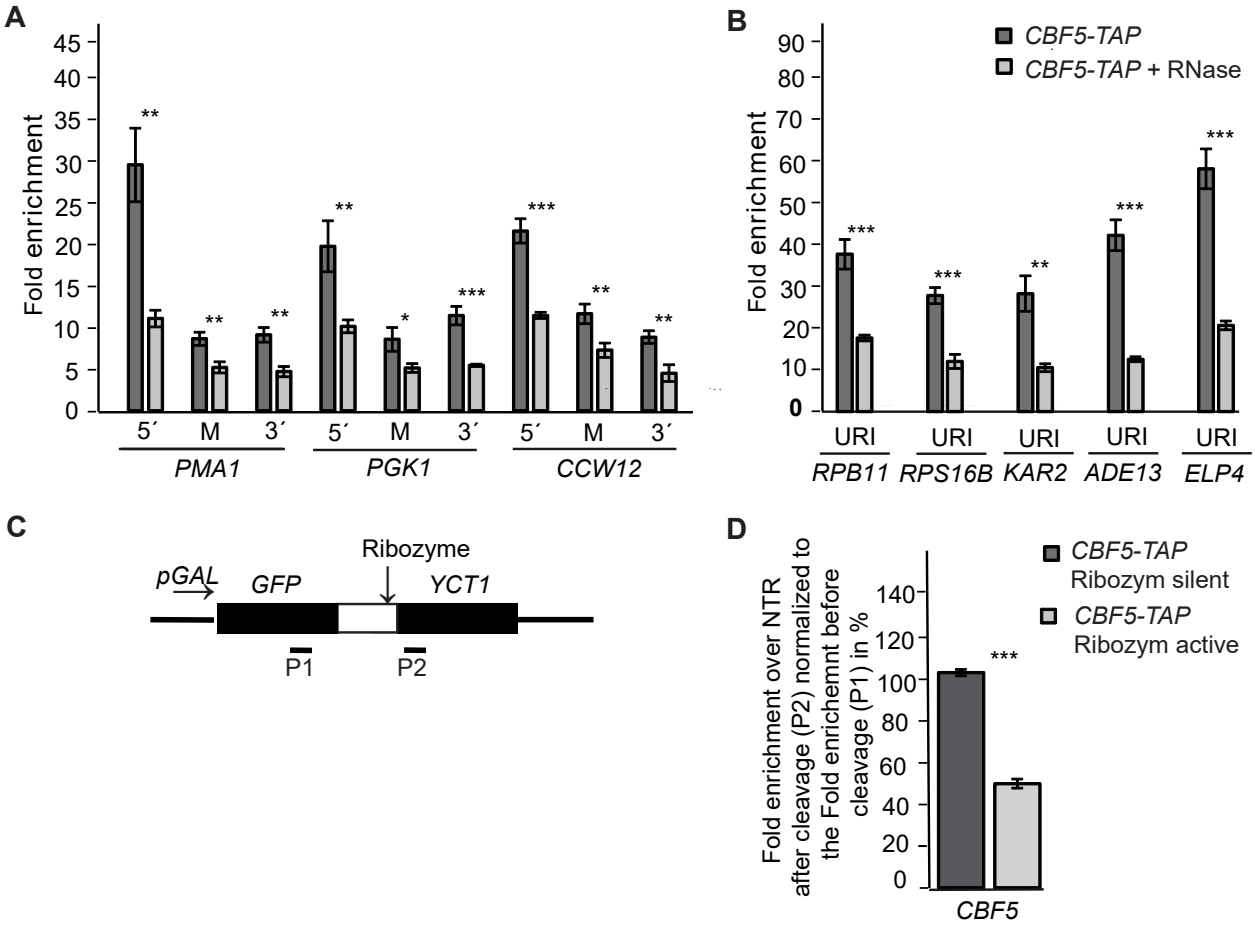

Figure 14: Cbf5 occupancy at transcribed genes is RNA-dependent. A&B: The occupancy of Cbf5 at the highly transcribed genes (A) and at the genes containing a Cbf5-dependent pseudouridylation site (B) was determined by ChIP in presence of 7.5 U RNase A and 300 U RNase T1. Cbf5 occupancy decreases in presence of RNase A and T1. **C:** Scheme of the ribozyme containing reporter, which consists of a sequence coding for the *GAL1* promoter (pGAL), *GFP* and the hepatitis λ ribozyme (wild-type or inactive mutant) inserted 5' of the nonessential gene *YCT1*. The positions of the primer (P1 and P2) used for qPCR are indicated as bars below the gene. **D:** To assess the dependence of the Cbf5 occupancy on RNA a hepatitis λ ribozyme ChIP assay of the silent and the active ribozyme was performed. Cbf5 occupancy in the ribozyme active construct is 50% decreased compared to the silent ribozyme. The fold enrichment of Cbf5 at the tested genes was normalized to an untranscribed region (NTR). Bars represent the mean +/- standard deviation of three independent biological replicates, p values of students t-test: ***: $p < 0.001$ **: $p < 0.01$

Therefore, an immunoprecipitation of the whole-cell lysate using an antibody against Rbp1 of RNAPII was conducted. The four tested deletion strains showed a significant decreased RNAPII occupancy compared to the wild-type. Whereas the RNAPII occupancy in *Δhpr1* is decreased about 25-30% [Figures 11B/C], the *THO2* deletion strain showed a 50-65% decreased RNAPII occupancy at transcribed genes [Figures 12B/C].

The highest reduction in RNAPII occupancy of about 75% was observed in the *CBP80* and *CBP20* deletion strain [Figures 13B/C]. Figures 11, 12, 13E/F show the Cbf5 occupancy normalized to the RNAPII level at the tested genes. The normalized Cbf5 occupancy is increased at almost all tested genes of the four different deletion strains. One exception is the normalized Cbf5 occupancy at *PGK 5' and RBP11* in the *HPR1* deletion strain, which is unaffected. These results demonstrate that neither THO/TREX nor CBC is needed for Cbf5 occupancy at transcribed genes.

Besides the CTD of Rbp1 [Chapman et al., 2008 Phatnani and Greenleaf, 2006] and the transcription elongation factor Spt5 [Klein et al., 2011], the RNA itself provides a platform to recruit proteins to transcribed genes/mRNA [Maris et al., 2005 Ozdilek et al., 2017]. To investigate whether Cbf5 occupancy is RNA-dependent, a Cbf5 ChIP in presence of RNase and a hepatitis λ ribozyme-based ChIP assay were performed.

Figures 14A/B show the result of a ChIP and subsequent qPCR analysis of Cbf5 in presence and absence of RNase A and T1. 7.5 U of RNase A and 300 U of RNase T1 were added to the whole-cell lysate and incubated for 30 min at room temperature prior to the immunoprecipitation. Furthermore, the cross-linking time was reduced from 20 to 5 min, as described in Abruzzi et al., 2004. The qPCR for the highly transcribed genes [Figure 14A] as well as for genes containing a pseudouridylation site [Figure 14B] showed a reduced co-immunoprecipitation with Cbf5 in presence of RNase A and T1.

To confirm this result, a hepatitis λ ribozyme ChIP assay was performed. Therefore, a *CBF5* TAP-tagged strain containing the silent (mutated) or active hepatitis λ ribozyme inserted 5' of the nonessential gene *YCT1* under the control of the *GAL1* promoter was used [Figure 14C]. This strain was grown in media containing galactose to induce transcription of the ribozyme-containing reporter. During synthesis of the *YCT1* mRNA, the internal ribozyme sequence folds into an enzymatically active RNA and initiates co-transcriptional self-cleavage. This cleavage event releases the 5' end of the nascent mRNA and any proteins bound to it, while RNAPII and the 3' end of the transcribed mRNA remain at the transcription site. ChIP with subsequent qPCR of the *CBF5-TAP* ribozyme silent/ active strain was performed. Cbf5 occupancy 5' (P1) and 3' (P2) of the ribozyme cleavage site was determined and compared to the Cbf5 occupancy of the silent ribozyme, which cannot initiate co-transcriptional self-cleavage. Cbf5 occupancy at P2 was normalized to the occupancy 5' of the cleavage site (P1) and was set to 100% for the inactive ribozyme construct, whereas the ratio of P2/P1 for the active ribozyme was calculated relative to the inactive ribozyme [Meinel et al., 2013].

Normalization of the result of the hepatitis λ ribozyme ChIP assay showed that Cbf5 occupancy in the active ribozyme construct is reduced by about 50% [Figure 14D]. This result confirmed that Cbf5 occupancy at transcribed genes is RNA-dependent.

2.3 Cbf5-dependent pseudouridylation occurs post-transcriptionally

So far, it was shown that Cbf5 is present at all tested transcribed genes and mRNAs independently of whether the mRNA contains a PseudoU site [Figure 10]. Besides, Cbf5 occupancy is transcription and RNA-dependent [Figures 10/14]. It is not known yet, when Cbf5-dependent pseudouridylation takes place. An RNA modification can occur during transcription or afterwards. To elucidate whether Cbf5-dependent pseudouridylation occurs co-transcriptionally or post-transcriptionally, nascent mRNA, poly(A) mRNA and total RNA were isolated. The relative amounts of the purified RNA types were analyzed performing RT-qPCR. Afterwards, primer extension after CMC treatment was carried out to quantify the pseudouridylation level of the different RNAs.
Nascent mRNA was isolated from transcribing RNAPII using immunoprecipitation (IP) with anti-FLAG M2 affinity beads of the protein lysate from the strain *RPB3-FTpA*. Subsequent isolation of nascent mRNA of the eluate was performed using the miRNeasy Kit. Figure 15A shows a polyacrylamide gel of two lysates/eluates. Western Blots with antibodies against the ribosomal protein Rpl8 and the mRNA-binding proteins Cbp80, Npl3 and Sub2 were performed to analyze the quality of the IPs [Figure 15B]. Rpl8 was not identified in the eluates of the IPs, whereas mRNA-binding proteins were detected. Sub2 was even enriched in the eluates compared to the lysates. These Western Blots demonstrated that it was possible to purify RNAPII, which is located with other mRNA binding-proteins at the nascent mRNA. The next step was the quantification of the amount of nascent mRNA and poly(A) mRNA of the three purified RNA types performing RT-qPCR. To this end, special primers were used for reverse transcription. To generate cDNA of nascent mRNA, RT primers annealing downstream of the poly(A) cleavage site were designed [Figure 15C]. This sequence does no longer exist in mature poly(A) mRNA. For reverse transcription of poly(A) mRNA, oligo dT_{19} primers were used. To generate cDNA of 25S rRNA, a primer annealing at the 3' end of the 25S rRNA was designed. Quantification of the amounts of nascent and mature *KAR2*, *RPS16B* and *ADE13* mRNA and of *25S rRNA* was performed using qPCR primers annealing at the 3' end of the mRNAs/rRNA gene.

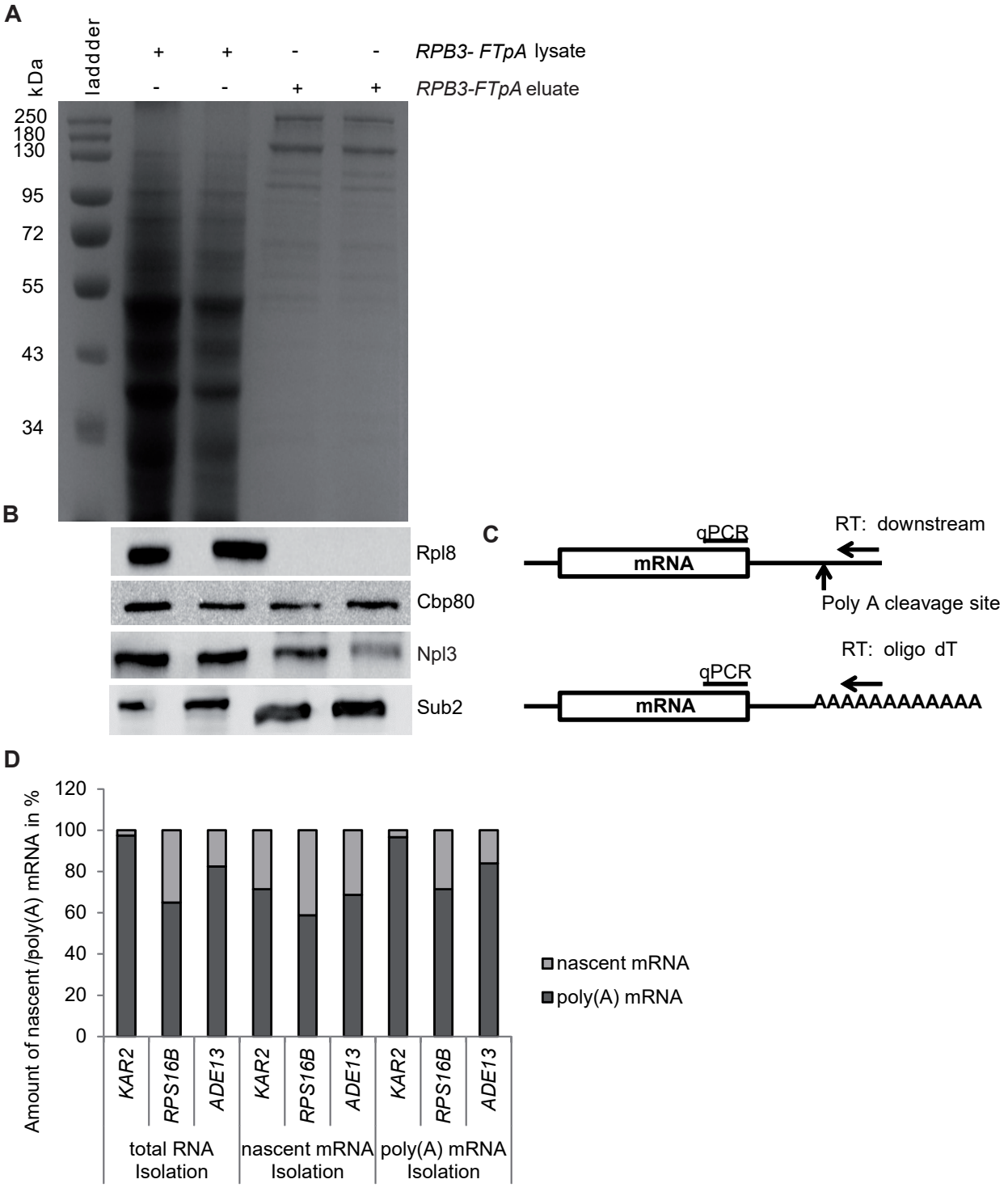

Figure 15: Isolation and quality control of total RNA, poly(A) mRNA and nascent mRNA. A&B: Coomassie-stained polyacrylamide gel and subsequent Western Blot analysis of the lysates and eluates of Rpb3 immunoprecipitations to purify nascent mRNA. 2 μl of the yeast cell lysates and 20 μl of the eluates were used to perform the SDS-PAGES. **C:** Scheme of an mRNA to mark the position of the primers used for the reverse transcription of nascent mRNA (RT: downstream) and poly(A) mRNA (RT: oligo dT) as well as for qPCR. The arrow labeled with poly(A) cleavage site shows the position of the poly(A) cleavage site.

D: Bar chart of the quantified amount of *KAR2, RPB16B, ADE13* nascent mRNA/ poly(A) mRNA after total RNA isolation, nascent mRNA isolation and poly(A) mRNA isolation.

Table 1: CT-values of the RT-qPCR of total RNA; nascent- and poly(A) mRNA to assess the quality of the purified RNA types

	total RNA isolation		**nascent mRNA isolation**		**poly(A) mRNA isolation**	
	CT (poly(A) mRNA)	**CT (nascent mRNA)**	**CT (poly(A) mRNA)**	**CT (nascent mRNA)**	**CT (poly(A) mRNA)**	**CT (nascent mRNA)**
KAR2	24.206	29.476	26.482	27.802	19.054	23.929
RPS16B	22.65	23.544	22.264	22.777	16.788	18.107
ADE13	26.731	28.957	26.957	28.088	20.587	22.97
25S rRNA	10.597		11.591		11.7	

Table 1 shows the result of the RT-qPCR of total RNA, nascent and poly(A) mRNA. Additionally, a bar chart, which exemplifies the share of *KAR2, RPB16B, ADE13* nascent - and poly(A) mRNA of each type of RNA isolation, was created [Figure 15D]. It was possible to enrich nascent mRNAs performing a Rpb3 immunoprecipitation compared to the amount of nascent mRNA of the total RNA isolation. Figure 15D shows that the enrichment of nascent mRNA varies. Whereas nascent mRNA of *KAR2* could be enriched around 25%, nascent mRNA of *RPS16B* was enriched around 6% [Figure 15D]. Table 1 also shows that the amount of poly(A) mRNA as well as nascent mRNA of the oligo dT beads purification is highly enriched compared to total RNA isolation. However, the ratio between nascent - and poly(A) mRNA of isolated poly(A) mRNA is similar to isolated total RNA. The amount of rRNAs of isolated poly(A) mRNA is merely reduced. RT-qPCR at 25S rRNA of the nascent mRNA and the poly(A) mRNA is decreased compared to total RNA. However, the amount of 25S rRNA is still high [Table 1].

Following quality control, primer extension after CMC treatment of total RNA, poly(A) and nascent mRNA was performed to quantify relative pseudouridylation levels of the different RNAs. Primer extension of dilution series at *KAR2, ADE13* and *RPS16B* were carried out to identify an RNA concentration in the linear range [Appendix Figure 33/34/35]. 2 μg CMC-treated RNA was used to perform the Primer extension of *KAR2* and 1 μg RNA was used for the Primer extension of *ADE13* and *RPS16B*.

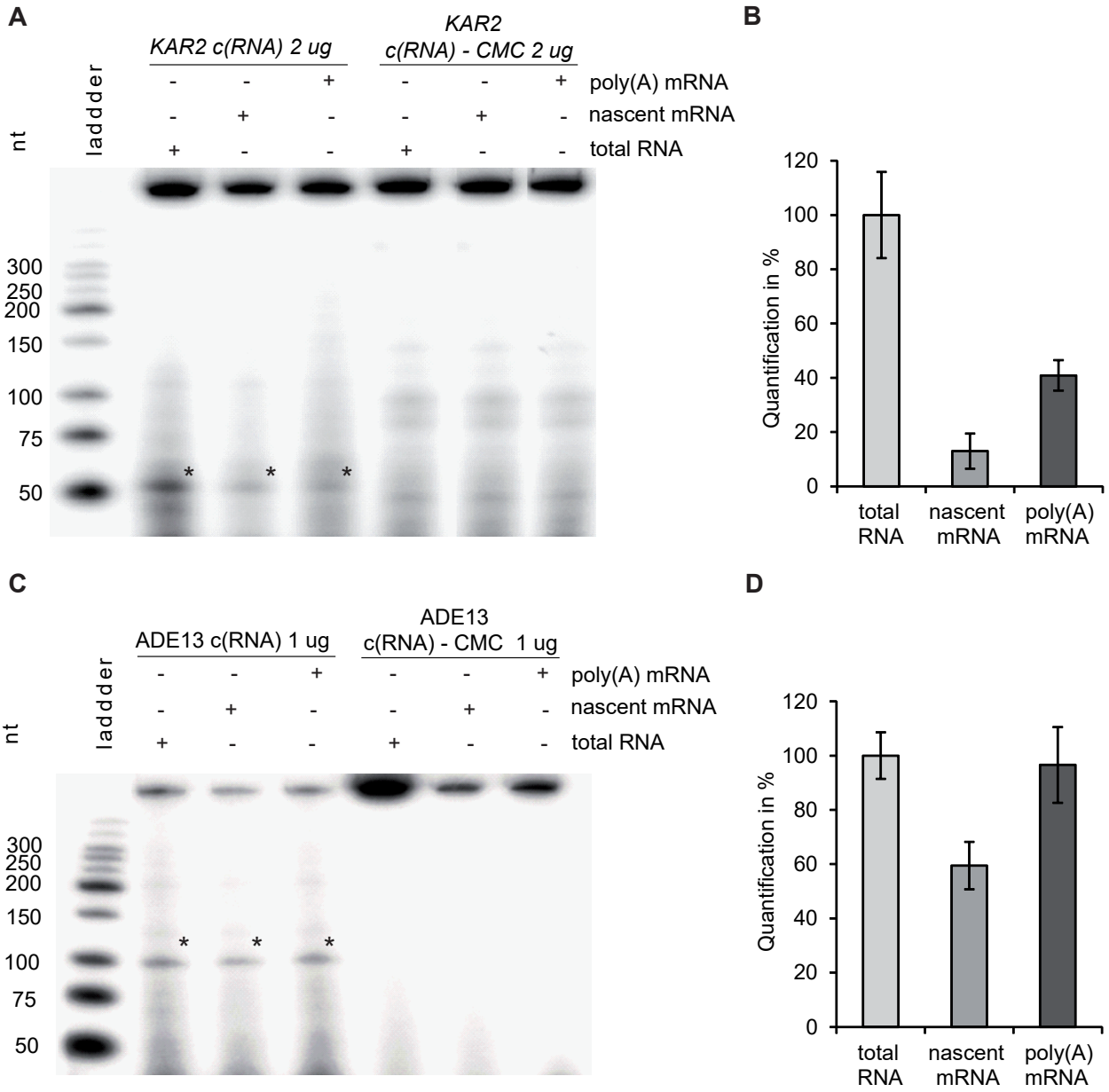

Figure 16: Cbf5-dependent pseudouridylation might occur post-transcriptionally. Primer Extension after CMC treatment of total RNA, poly(A) and nascent mRNA to assess the pseudouridylation rate of the different RNA types. **A**&**C**: Autoradiographs of the Primer Extension at the Cbf5-dependent pseudouridylation site of *KAR2* and *ADE13*. A concentration (c) of 2 μg/1 μg CMC-treated RNA or CMC-untreated (-CMC) RNA was used for the reverse transcription. Asterisks labeled bands show the radioactive signal of the cDNA after the discontinuation of the reverse transcription. **B**&**D**: Quantification of the discontinuation signal of the reverse transcription shows a lower level of pseudouridylation for nascent mRNA. Bars represent the mean +/- standard deviation of three independent biological replicates.

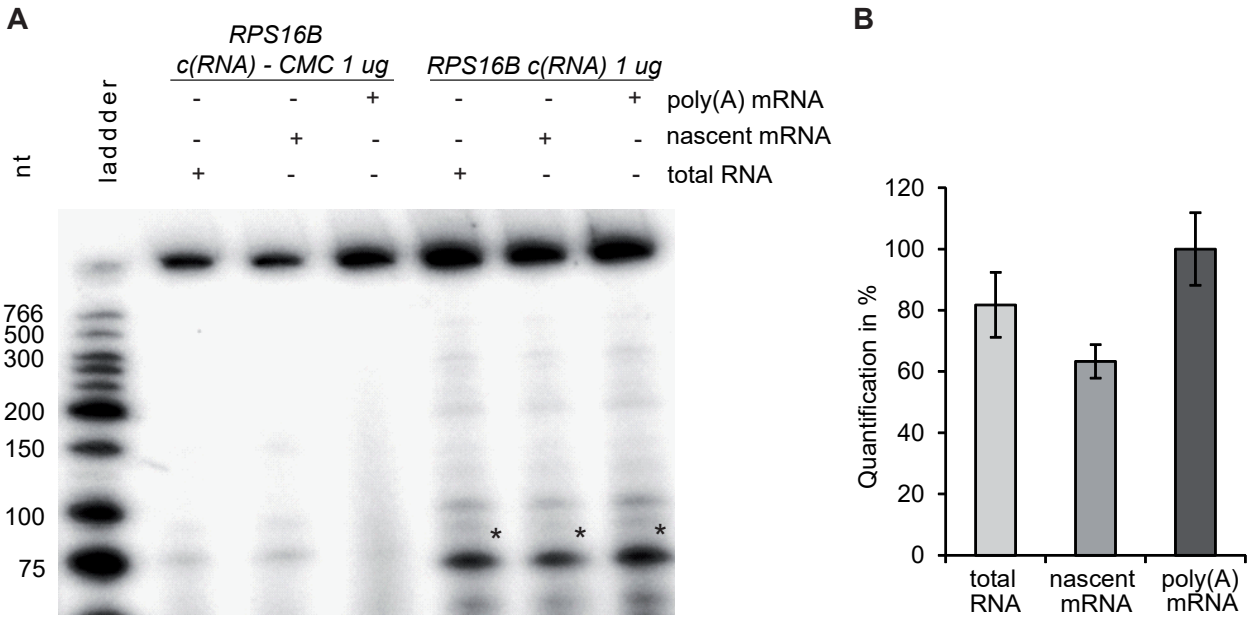

Figure 17: Cbf5-dependent pseudouridylation might occur post-transcriptionally. Primer Extension after CMC treatment of total RNA, poly(A) and nascent mRNA to assess the pseudouridylation rate of the different RNA types. **A**: Autoradiograph of the Primer Extension at the Cbf5-dependent pseudouridylation site of *RPS16B*. A concentration (c) of 1 μg CMC-treated RNA or CMC-untreated (-CMC) RNA was used for reverse transcription. Asterisks labeled bands show the radioactive signal of the cDNA after the discontinuation of the reverse transcription. **B**: Quantification of the discontinuation signal of the reverse transcription shows a lower level of pseudouridylation for nascent mRNA. Bars represent the mean +/- standard deviation of three independent biological replicates.

Figure 16A/C & Figure 17A show autoradiographs of the Primer extension of total RNA, nascent mRNA and poly(A) mRNA at the Cbf5-dependent pseudouridylation sites of *KAR2, RPS16B* and *ADE13*. Asterisks-labeled bands display radioactive signals of the cDNA after the discontinuation of reverse transcription at the modified site. These discontinuations of the RT represent the pseudouridylated share of the different RNA types at the defined position of the mRNA. CMC-untreated RNA served as negative control. Many other radioactive signals of a low intensity are existing in all lanes of the autoradiographs.
Figure 16B/D & Figure 17B show the quantified radioactive signals of the discontinuation of the reverse transcription using ImageJ. The quantification shows a lower level of pseudouridylation for nascent mRNA at the three tested mRNAs compared to total RNA and poly(A) mRNA. That indicates that Cbf5-dependent pseudouridylation might occur post-transcriptionally. However, pseudouridylation of isolated RNA after RNAPII IP is clearly detectable. It is around 60% at the isolated RNA of *ADE13* and *RPS16B*.

2.4 Cbf5 functions in mRNA transcription and nuclear mRNA export

Cbf5 is present at all tested transcribed genes and mRNAs independent of whether the mRNA contains a PseudoU site [Figure 14]. Besides Cbf5 interacts with TREX in an RNA-independent manner [Gruellich, 2017]. However, TREX is not needed for Cbf5 occupancy at transcribed genes [Figures 11/12]. To unravel the function of the Cbf5/TREX interaction and to understand the reason for the occupancy of Cbf5 at all tested transcribed genes and mRNAs, the next aim was to elucidate whether Cbf5 has a role in transcription and nuclear mRNA export just as TREX [Straesser et al., 2002].
For this, a *CBF5* depletion strain and a catalytically inactive *cbf5* mutant were generated. To generate a *CBF5* depletion strain, the auxin-inducible degron system was used. This is a depletion system of plants, where the plant hormone auxin induces degradation of the indole-3-acetic acid (IAA) family of transcription repressors by the Skp1-cullin 1-F-box containing (SCF) E3 ubiquitin ligase.

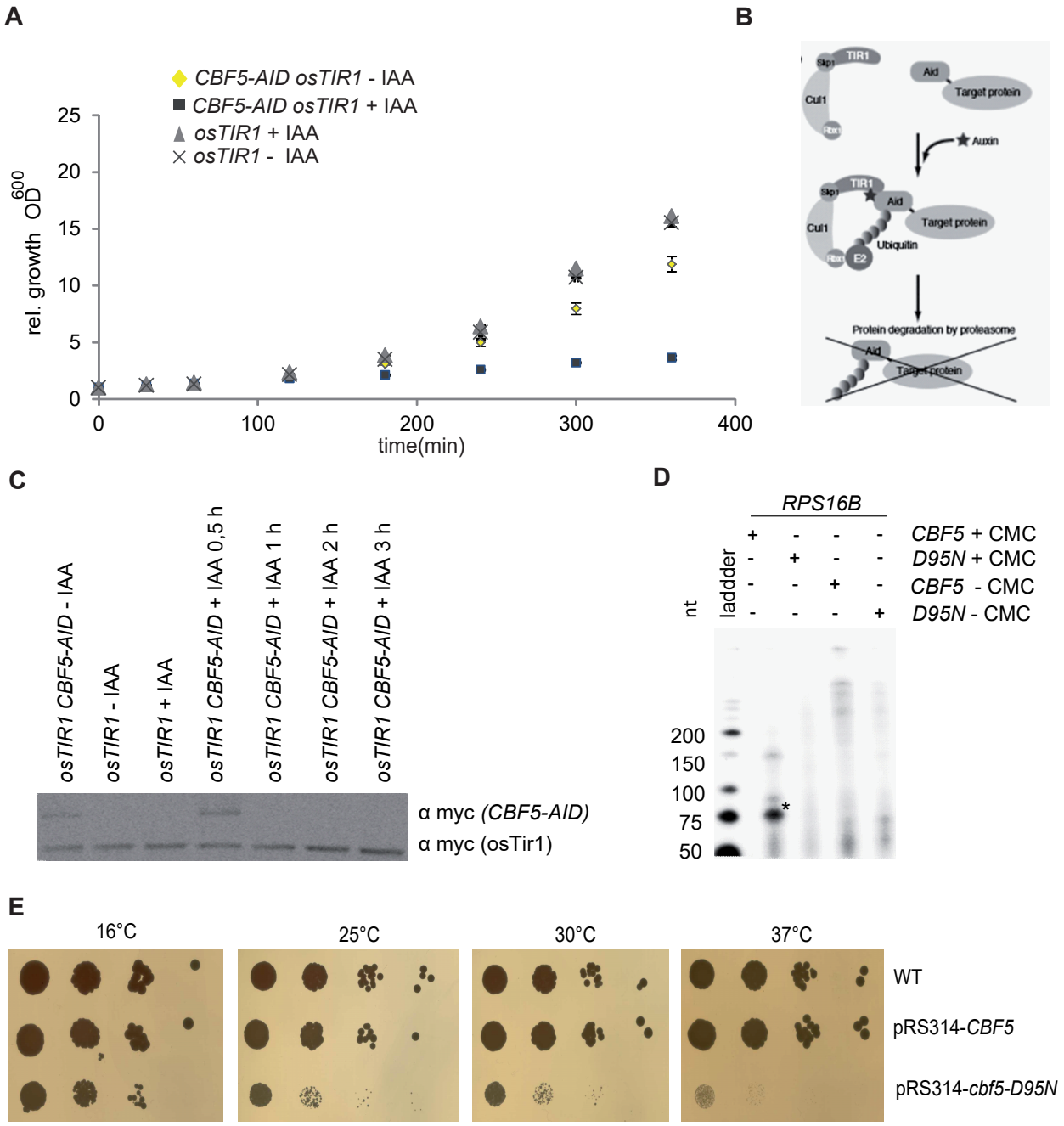

Figure 18: Characterization of the *CBF5* depletion strain and the catalytically inactive *cbf5* mutant A: *osTIR1 CBF5-AID* shows a strong growth defect in presence of auxin (IAA). Growth curve of the strains *osTIR1 CBF5-AID* and *osTIR1* +/- IAA at 30°C. Bars represent the mean +/- standard deviation of at least three independent biological replicates. **B:** Scheme of the auxin induced degradation pathway. In presence of auxin an *AID*-tagged protein interacts with Tir1, which leads to ubiquitylation of the tagged protein and subsequent degradation by the proteasome [Nishimura et al., 2009]. **C:** Western Blot of protein lysates from *osTIR1 CBF5-AID* and *osTIR1* +/- IAA shows that Cbf5 is depleted after 1 h IAA (c(IAA) = 1 mM) treatment. *osTIR1* serves as a loading control. **D:** Primer extension after CMC treatment at *RPS16B* shows that the Cbf5-dependent pseudouridylation does not exist in the catalytically inactive *cbf5* mutant (*D95N*), whose RNA is CMC-treated (+CMC).

CMC-untreated RNA (-CMC) serves as negative control. **E:** *cbf5* catalytically inactive mutant shows a strong growth defect. Ten-fold serial dilutions of the wild-type strain (WT) and yeast strains containing pRS314-*CBF5* and pRS314-*cbf5-D95N* were spotted on YPD plates and incubated 2-5 days at 16°C, 25°C, 30°C and 37°C.

In presence of auxin the interaction between a protein containing an auxin-inducible degron (*AID*) and the substrate recognition domain of *TIR1* is mediated. This interaction leads to a ubiquitylation of the target protein by recruitment of the SCF E3 ubiquitin ligase and finally proteasomal degradation [Figure 18B] [Nishimura et al., 2009/ Morawska and Ulrich, 2013]. The SCF pathway is highly conserved among eukaryotes and also exists in yeast. However, *S. cerevisiae* can not respond to auxin, as *TIR1* does not exist in the genome. By integrating *osTIR1* (*TIR1* of rice plant *Oryza sativa*) and *AID*-tagging of the target protein, the auxin-inducible degron system can be transplanted into non-plant cells for rapid degradation of proteins [Nishimura et al., 2009/ Morawska and Ulrich, 2013]. As Cbf5 is an essential protein, *CBF5* was *AID*-tagged for depletion. Figure 18A shows a growth curve of the strains *CBF5-AID osTIR1* and *osTIR1* in presence and absence of 1 mM IAA. *CBF5-AID osTIR1* has a strong growth defect in presence of IAA, suggesting that Cbf5 is efficiently degraded. The *osTIR1* strain, which does not contain the *AID* tag at *CBF5*, shows the same growth phenotype in presence and absence of auxin [Figure 18A]. *CBF5-AID osTIR1* in absence of auxin has a small growth defect compared to *osTIR*. The Western Blot in Figure 18C shows that Cbf5 can not be detected after 1 h IAA treatment. The antibody against myc detects *CBF5-AID* as well as *osTIR1*, as *MYC* was integrated upstream/ downstream of both genes. OsTir1 serves as a loading control.
Furthermore, a *cbf5* catalytically inactive mutant was generated. Therefore, a point mutation at the catalytic center of *CBF5* at aa residue 95 was inserted according to Ramamurthy et al., 1999. Dot spots incubated at 16°C, 25°C, 30°C and 37°C of the catalytically inactive *cbf5* mutant (pRS314-*cbf5-D95N*) show a strong growth defect at all temperatures compared to *CBF5* expressed from a plasmid in the *CBF5* shuffle strain (pRS314-*CBF5*) and the wild-type strain [Figure 18E]. Primer extension after CMC treatment at *RPS16B* shows that the Cbf5-dependent pseudouridylation does not exist in *CBF5* shuffle pRS314-*cbf5-D95N* (*D95N*), whereas the pseudouridylation was detected in *CBF5* shuffle pRS314-*CBF5* (*CBF5*) of CMC-treated RNA (asterisk-labeled band) [Figure 18D]. CMC-untreated RNA (-CMC) serves as negative control.

2.4.1 Cbf5 has a function in mRNA transcription elongation

To get a first indication whether Cbf5 has a role in transcription, the catalytically inactive *cbf5* mutant was examined for 6-AU sensitivity. 6-AU influences yeast growth, when other mutations of a strain affect transcription elongation [Exinger and Lacroute, 1992]. To investigate whether the strain *cbf5-D95N* has a growth defect on plates containing 6-AU, growth assays on SDC -Ura plates containing 25 μg/ml, 50 μg/ml 6-AU or the solvent dimethyl sulfoxide (DMSO) were performed [Figure 19B]. Interestingly, the catalytically inactive *cbf5* mutant grows faster on plates containing 6-AU than on plates with DMSO, where the mutant shows a strong growth defect. Dst1 is commonly used as a positive control for 6-AU sensitivity. The *DST1* deletion strain and the double mutant *CBF5* shuffle pRS314-*CBF5* *Δdst1* (*Δdst1* pRS314-*CBF5*) have a growth defect on plates containing 6-AU, as Dst1 is the transcription elongation factor TFIIS. Interestingly, the double mutant *CBF5* shuffle *pRS314-cbf5-D95N* *Δdst1* (*Δdst1* pRS314-*cbf5-D95N*) partially rescues the growth defect of the *DST1* deletion strain, indicating that Cbf5 has a role in transcription elongation.
To further investigate a potential function of Cbf5 in transcription elongation, RNAPII ChIPs of the catalytically inactive *cbf5* mutant +/- 6-AU at the 7887 base pairs (bp) long gene *YLR454* under the control of the *GAL1* promoter were performed. *YLR454* is the longest gene in *S. cerevisiae*. It is used to study the transcription rate and processivity of RNAPII [Mason and Struhl, 2005]. Transcription processivity is determined by measuring the level of RNAPII association across the *YLR454* coding region in galactose-grown cells. Yeast strains containing pRS314-*cbf5-D95N* or pRS314-*CBF5* were grown in medium lacking uracil with 2% raffinose to the early exponential phase, followed by a galactose-induction for 2.5 h and a 6-AU treatment for 30 min prior to formaldehyde-crosslinking [Mason and Struhl, 2005]. Afterwards, ChIPs with an antibody against the large subunit of RNAPII were performed. The occupancy of RNAPII was measured at seven positions within *YLR454* as shown in Figure 19B. Figure 19C shows that the RNAPII occupancy increases at the 3' end of *YLR454* in the catalytically inactive *cbf5* mutant, whereas the occupancy of RNAPII increases along the whole gene in the yeast strain containing pRS314-*cbf5-D95N* in presence of 6-AU compared to the corresponding wild-type. This result demonstrates that RNAPII processivity increases in the catalytically inactive *cbf5* mutant, suggesting that Cbf5 could have an active role in regulating transcription elongation.

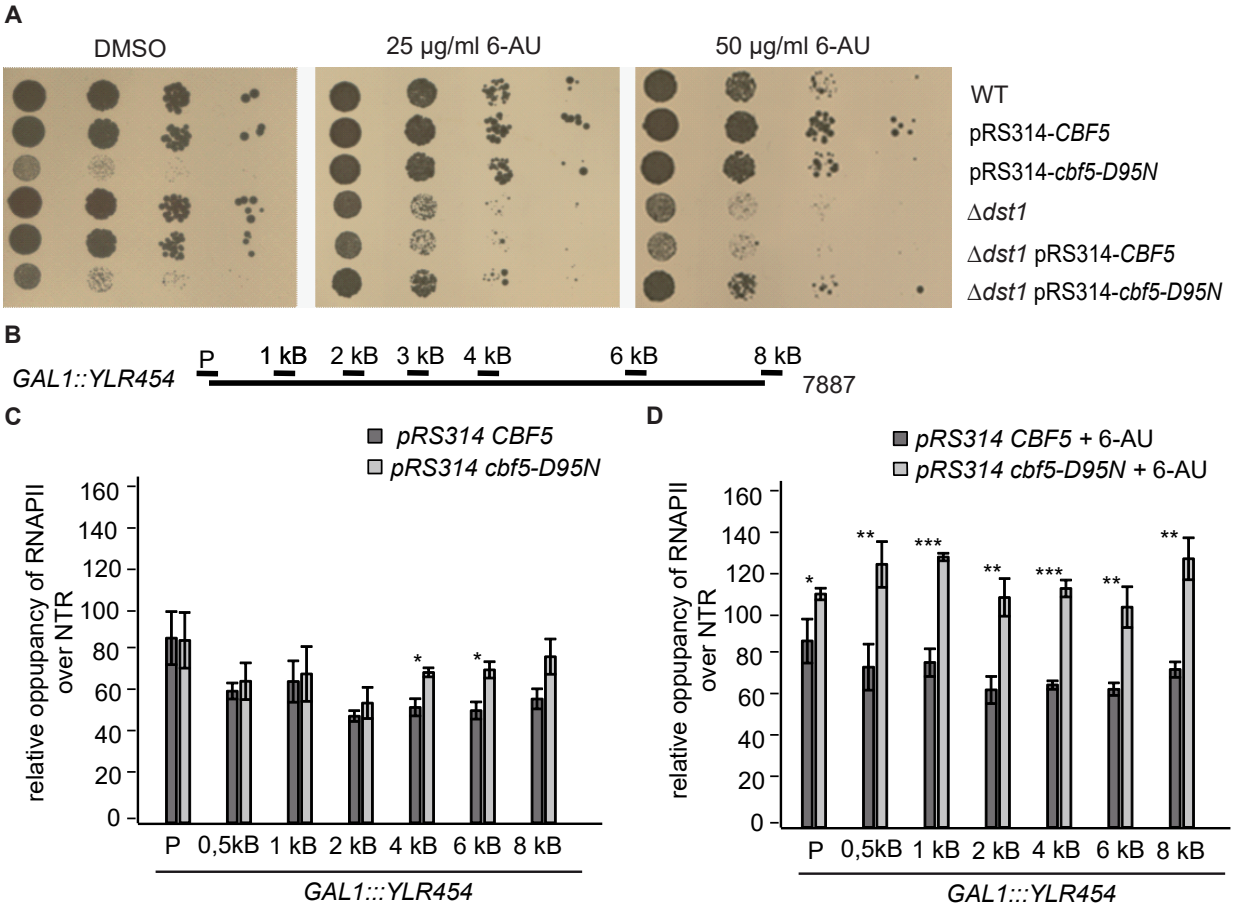

Figure 19: Cbf5 has a function in transcription-elongation A: The *cbf5* catalytically inactive mutant partially rescues 6-azauracil (6-AU) sensitivity of a *DST1* deletion strain. Ten-fold serial dilutions of wild-type (WT), *CBF5*-shuffle pRS314-*CBF5* (pRS314-*CBF5*), *CBF5*-shuffle pRS314-*cbf5-D95N* (pRS314-*cbf5-D95N*), *Δdst1*, *Δdst1* pRS314-*CBF5* and *Δdst1* pRS314-*cbf5-D95N* cells were spotted on SDC -URA plates containing DMSO, 25 μg/ml 6-AU or 50 μg/ml 6-AU and incubated for 2-3 days at 30°C **B:** Scheme of the 7887 base pairs (bp) long gene *GAL1::YLR454* represented as a solid line. Positions of the primers used for qPCR are indicated as bars above the solid line. **C/D:** RNAPII processivity increases in the catalytically inactive *cbf5* mutant. RNAPII occupancy within *YLR454* in yeast strains containing pRS314-*CBF5* and pRS314-*cbf5-D95N* grown in medium with galactose in absence (C) and presence of 50 μg/ml 6-AU (D) was determined by ChIP. The fold enrichment of RNAPII was normalized to an untranscribed region (NTR). Bars represent the mean +/- standard deviation of at least three independent biological replicates. p values of students t-test: ***: $p < 0.001$ **: $p < 0.01$; *: $p < 0.05$

2.4.2 Cbf5 regulates expression of genes involved in metabolism and energy generation

To unravel the function of Cbf5 in transcription, RNAPII ChIPseq of the *CBF5* depletion strain and the catalytically inactive *cbf5* mutant were performed. An antibody against Rbp1 of RNAPII was used to conduct the immunoprecipitation. Library preparation and sequencing were peformed by the group of Stefan Günther (Max Planck Institut für Herz und Lungenforschung, Bad Nauheim). Bioinformatics was done by Samira Reuscher (group of Kathi Zarnack, Buchmann Institute for Molecular Life Sciences, Goethe University Frankfurt). The volcano plot and the bar diagram of Figure 20A/C show that around 600 genes are significantly (adjusted p value < 0.05) up and down regulated in the RNAPII ChIPseq of the *CBF5* depletion strain. ChIPseq of the catalytically inactive *cbf5* mutant indicates that the occupancy of RNAPII significantly decreases at 66 genes, whereas only three genes were significantly upregulated [Figure 20B/C]. One of three genes that show an increased occupancy of RNAPII is *CBF5* itself. 45 genes show a significantly changed RNAPII occupancy in both ChIPseq experiments [Figure 20D]. The gene ontology (GO) term enrichment analysis shown in Figure 20E and Figure 21A/B display the type of genes that are regulated in the two different mutants. GO term analysis of the RNAPII ChIPseq in the catalytically inactive *cbf5* mutant shows that mainly genes involved in metabolic processes and energy generation are regulated. These genes are predominantly amino acid biosynthetic genes or carbohydrate biosynthetic genes [Figure 20E]. Similar results are obtained by GO term enrichment analysis of the upregulated genes of the *CBF5* depletion strain. Figure 21B displays that mainly genes involved in metabolism of carbohydrates or nucleic acids are regulated. Besides, genes involved in peptide synthesis, ribosome biogenesis or rRNA processing show a decreased RNAPII occupancy in the *CBF5* depletion strain [Figure 21B].
To validate the RNAPII ChIPseq results, ChIPs with a subsequent qPCR analysis in the *CBF5* depletion strain and the catalytically inactive *cbf5* mutant were performed. 15 genes showing a significantly changed RNAPII occupancy in at least one of the two ChIPseq data were chosen for single-gene validation. These genes are permeases or have a role in rRNA biogenesis or biosynthesis of amino acids, fatty acids, nucleotides or carbohydrates.

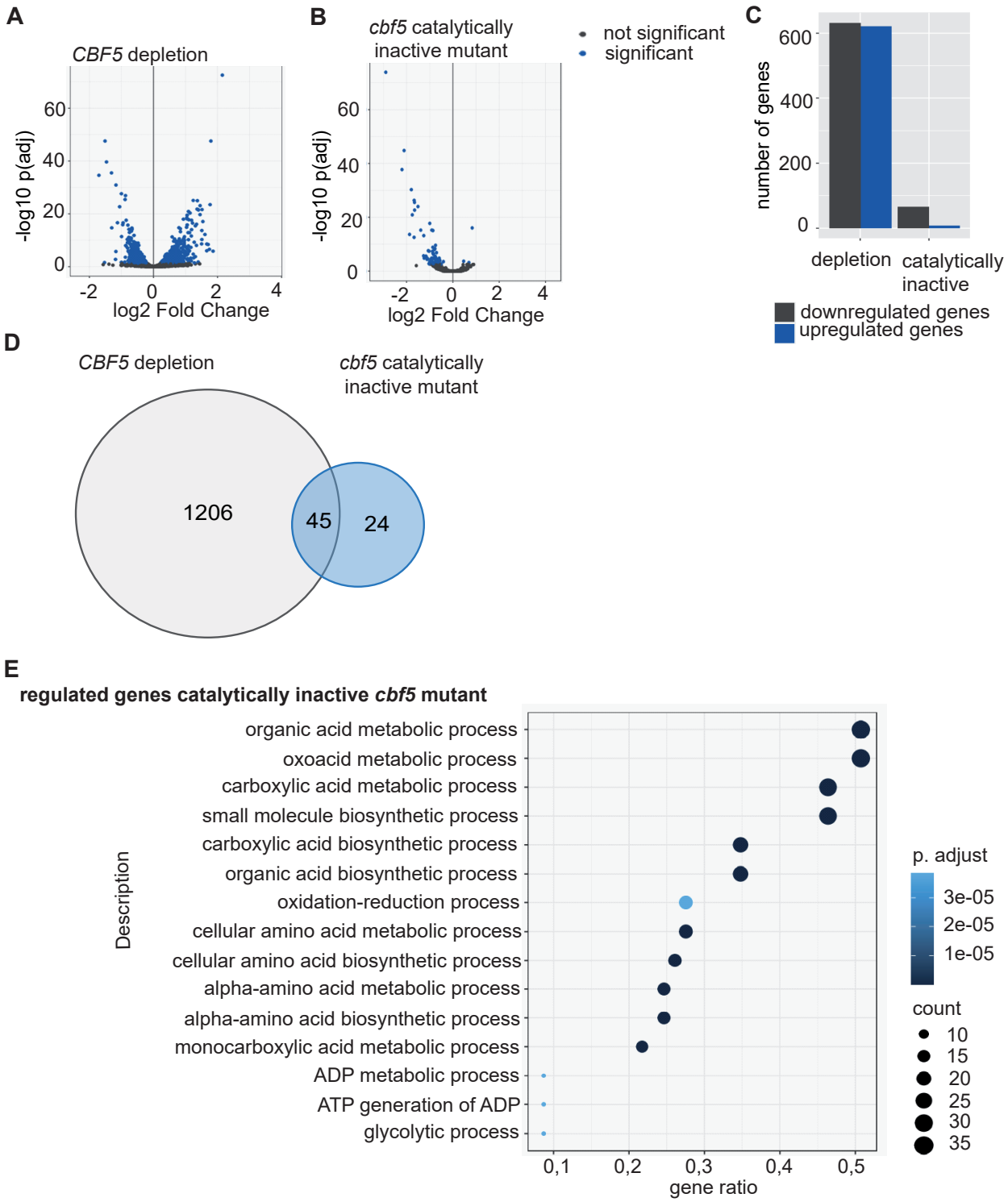

Figure 20: RNAPII ChIPseq data show that mainly genes involved in metabolism and energy generation are regulated. A/B: Volcano plots of the RNAPII ChIPseq signals in the *CBF5* depletion strain (A) and catalytically inactive mutant *cbf5* mutant (B) vs. wild-type show the amount of the significantly and not significantly up and down regulated genes. (adjusted p value < 0.05) **C:** Regulated genes in the different strains are shown in a bar diagram (adjusted p value < 0.05). **D:** Venn diagram displays that 45 genes show a changed RNAPII occupancy in both *cbf5* mutants. **E:** The gene ontology term enrichment analysis of the significantly

regulated genes of the RNAPII ChIPseq in the catalytically inactive *cbf5* mutant shows that genes involved in amino acid or carbohydrate metabolism and energy generation are regulated. Bioinformatics was performed by Samira Reuscher.

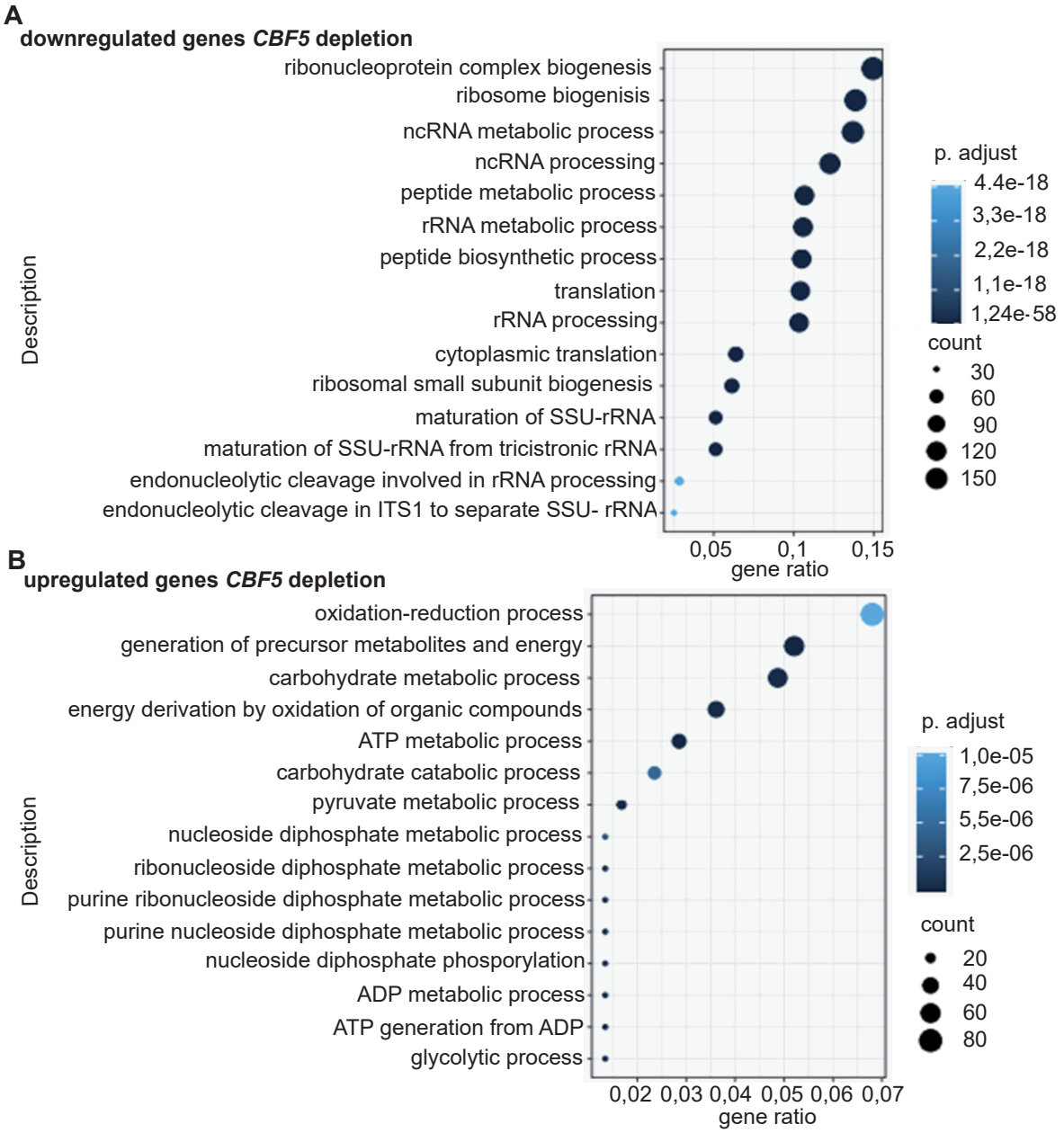

Figure 21: RNAPII ChIPseq data show that mainly genes involved in metabolism and energy generation are regulated. The gene ontology term enrichment analysis of the significantly (adjusted p value $< 0{,}05$) down- and upregulated genes of the RNAPII ChIPseq in the *CBF5* depletion strain indicates that genes involved in amino acid synthesis, ribosome biogenesis or rRNA processing are mainly downregulated (A), whereas genes that have a function in the metabolism of carbohydrates or nucleic acids were upregulated (B). Bioinformatics was performed by Samira Reuscher.

Table 2: log2 fold change of 15 selected genes of the RNAPII ChIPseq data of the *CBF5* depletion strain and the *cbf5* catalytically inactive mutant; adjusted p value $< 0{,}05$; n.s.: not significant

gene name	**log2 fold change RNAPII ChIPseq *cbf5* catalytically inactive mutant**	**log2 fold change RNAPII ChIPseq *CBF5* depletion strain**	**function of the gene**
ATR1	-2.14	-1.15	permease
MUP1	-0.79	-1.03	permease
MET17	-1.68	-1.76	amino acid biosynthesis
HIS4	-1.56	-0.72	amino acid biosynthesis
CYS3	-0.60	-0.74	amino acid biosynthesis
UTP22	n.s.	-0.69	rRNA biogenesis
SDA1	n.s.	-0.78	rRNA biogenesis
RPA190	n.s.	-0.87	rRNA biogenesis
OLE1	-1.74	n.s.	fatty acid biosynthesis
FAS1	-0.86	n.s.	fatty acid biosynthesis
ADE5/7	n.s.	-0.94	nucleotide biosynthesis
URA7	n.s.	-0.76	nucleotide biosynthesis
GUA1	n.s.	-0.58	nucleotide biosynthesis
PGK1	-0.68	0.68	carbohydrate biosynthesis
CIT2	-1.46	1.22	carbohydrate biosynthesis

Table 2 displays the log2 fold change of 15 genes of the RNAPII ChIPseq results in both *cbf5* mutants. Figure 22A/B show qPCR results of RNAPII ChIPs in these mutants. A comparison of both data of the catalytically inactive *cbf5* mutant displays that the single-gene analysis is similar to the log2 fold change of the ChIPseq at the 15 genes. The verification of the RNAPII ChIPseq in the *CBF5* depletion strain shows a conformance at 13 genes.

PGK1 and *CIT2* show a decreased RNAPII occupancy performing ChIP and a subsequent qPCR analysis, whereas the ChIPseq result displays an increased log2 fold change compared to the wild-type.

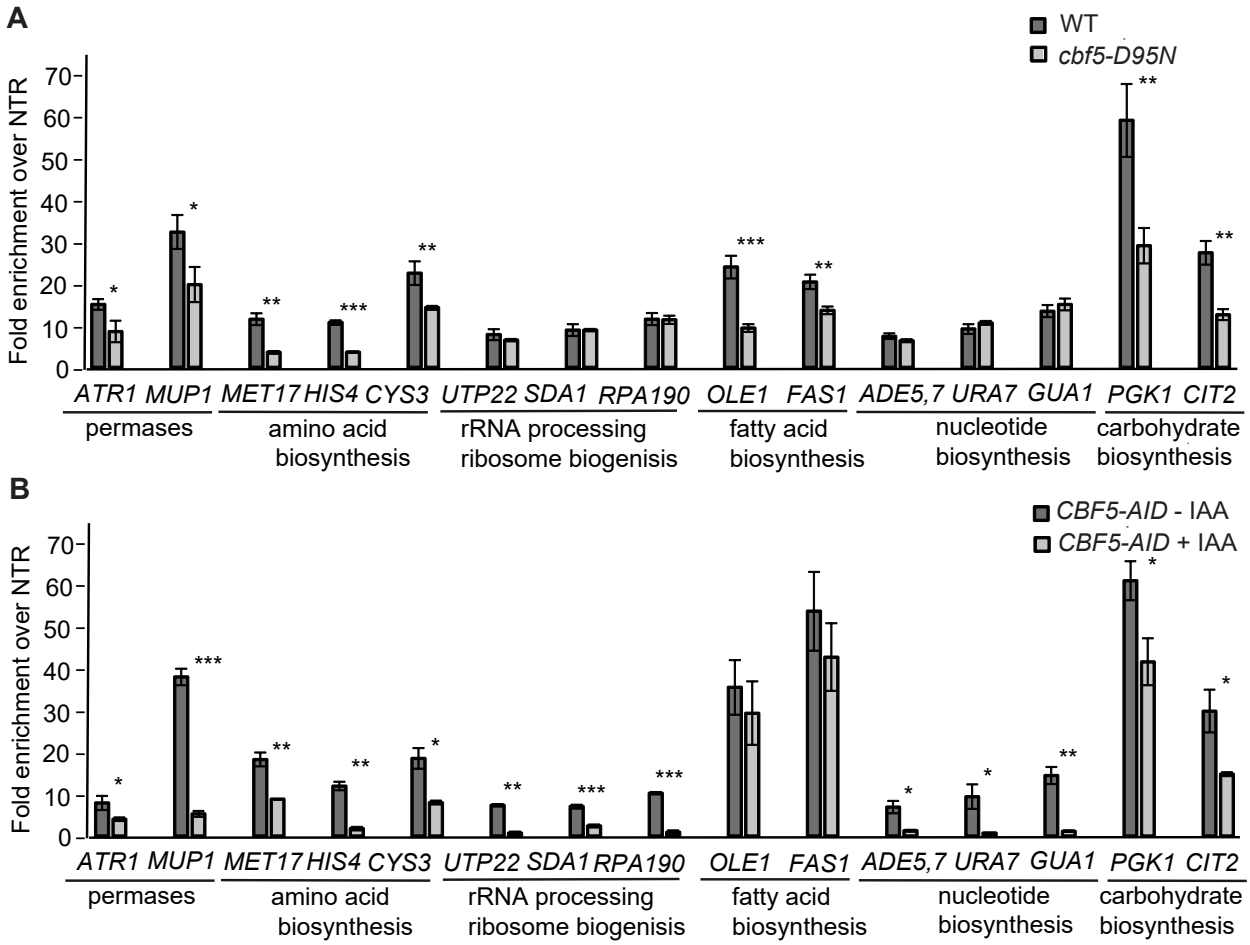

Figure 22: Verification of the ChIPseq data performing ChIP and a subsequent qPCR at 15 selected genes A/B: ChIP qPCR in the *cbf5* catalytically inactive mutant (A) at 15 selected gens shows a similar RNAPII fold enrichment as the log2 fold change of the ChIPseq data. The verification of the RNAPII ChIPseq in the *CBF5* depletion strain (B) shows a conformance at 13 genes. *PGK1* and *CIT2* show a decreased RNAPII occupancy, when Cbf5 is depleted, whereas the ChIPseq result displays an increased fold2 fold change compared to the wild-type. The fold enrichment was normalized to an untranscribed region (NTR). Primers annealing at the 3' end of each gene were used. Bars represent the mean +/- standard deviation of three independent biological replicates, p values of students t-test: ***: $p < 0.001$ **: $p < 0.01$; *: $p < 0.05$

2.4.3 Cbf5 regulates the protein amount of the transcription activator *GCN4*

So far, it was shown that Cbf5 has a role in transcription elongation. ChIPseq data indicated that Cbf5 regulates expression of genes involved in metabolism and energy generation. To further investigate these results, a comparison of the RNAPII ChIPseq results of the *cbf5* mutants [Figures 20/21] with other RNAPII ChIPseq data of transcription factor deletion strains in *S. cerevisiae* was performed [Kemmeren et al., 2014]. This comparison revealed that 93 genes that show a changed RNAPII occupancy in the *GCN4* deletion strain are affected in at least one of the *cbf5* mutants [Figure 23A], corresponding to 29% of all regulated genes of the published RNAPII ChIPseq in the *GCN4* deletion strain [Kemmeren et al., 2014]. Gcn4 is a transcription activator of amino acid biosynthetic genes. The expression of *GCN4* is tightly regulated at both the transcriptional and translational levels. The transcription activator responds to amino acid starvation [Natarajan et al., 2001/ Lucchini et al., 1984/ Hinnebusch and Natarajan, 2002].
To test, whether Cbf5 regulates the mRNA or protein amount of *GCN4*, RT-qPCR and quantitative Western Blots were performed in the *CBF5* depletion strain and the catalytically inactive *cbf5* mutant. To study this, the *cbf5* mutants were grown in minimal media containing 120 μM instead of 178 μM of each required aa, as *GCN4* is expressed under amino acid starvation. Oligo dT primers were used for reverse transcription and qPCR was performed with primers annealing at the 3' end of *GCN4*. Figure 23B shows that the fold enrichment of *GCN4* is unchanged in the *CBF5* depletion strain and the catalytically inactive *cbf5* mutant compared to the wild-type, indicating that Cbf5 does not regulate mRNA levels of *GCN4*. Furthermore, quantitative Western Blots of TAP-tagged Gcn4 in the *CBF5* depletion strain and the catalytically inactive *cbf5* mutant were performed. The Western Blot analysis of Cbf5 in *CBF5-AID* +/- IAA shows that Cbf5 is depleted after 1 h IAA treatment [Figure 23C]. Pgk1 serves as a loading control for normalization. Quantification of the Western Blot analysis showed that the protein amount of Gcn4 is reduced about 50% when Cbf5 is depleted [Figure 23C]. The protein level of Gcn4 in the catalytically inactive *cbf5* mutant is decreased about 95% [Figure 23C]. These results display, that Cbf5 regulates the protein amount of Gcn4 in a pseudouridylation-dependent manner.

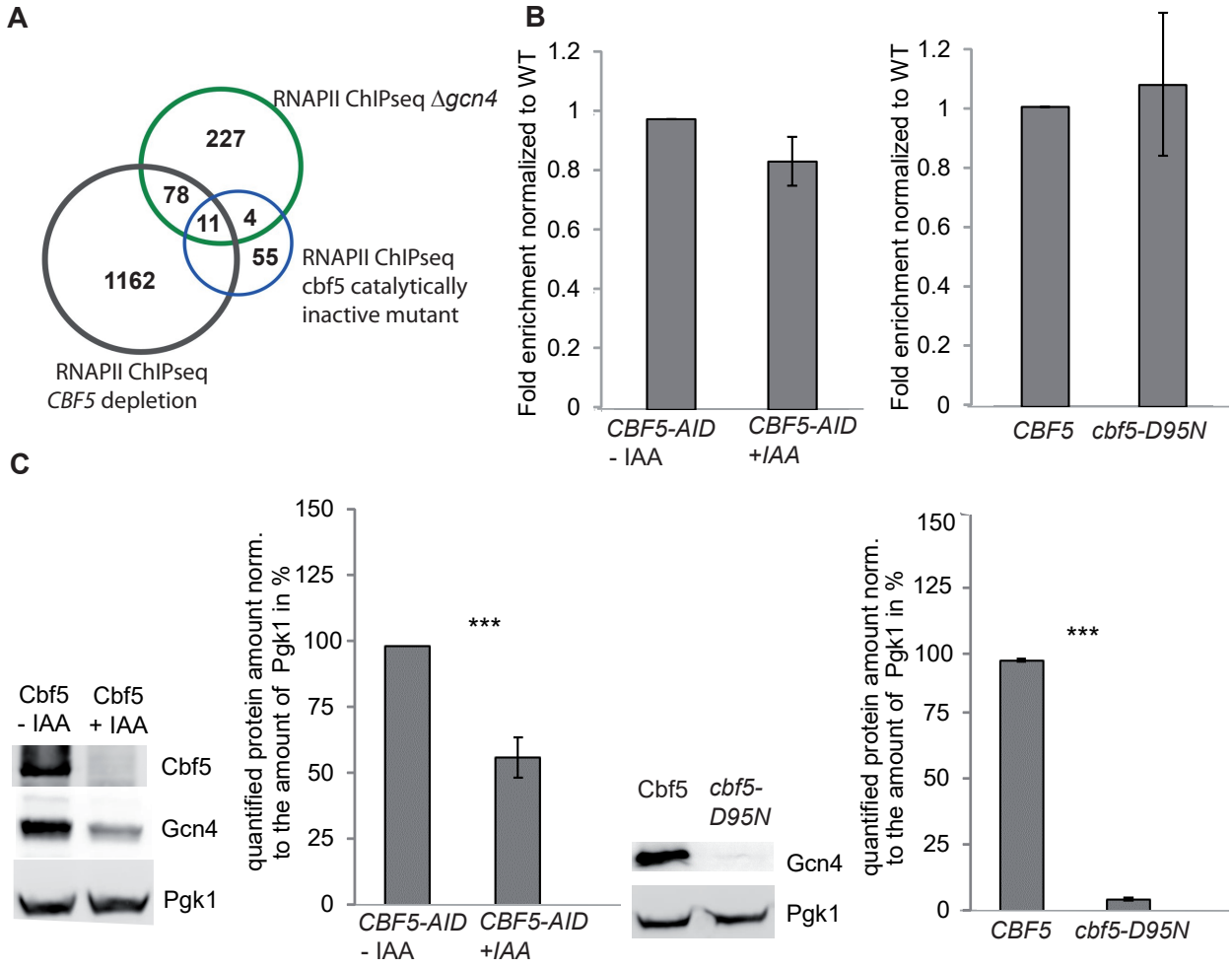

Figure 23: Cbf5 regulates the protein amount of Gcn4 A: Venn Diagram of RNAPII ChIPseq data display that 93 genes show a changed RNAPII occupancy in the *GCN4* deletion strain and in one of the *cbf5* mutants (*CBF5* depletion strain or *cbf5* catalytically inactive mutant). **B:** Total mRNA level of *GCN4* does not change in the *CBF5* depletion strain and the catalytically inactive *cbf5* mutant. mRNA levels were determined by RT-qPCR. **C:** Total protein level of Gcn4 is about 50% decreased in the *CBF5* depletion strain and about 95% decreased in the *cbf5* catalytically inactive mutant. Protein levels were determined by quantitative Western Blot of the whole cell extract and quantified by ImageJ. **B/C:** Bars represent the mean +/- standard deviation of three independent biological replicates, p values of students t-test: ***: $p < 0.001$

GCN4 contains four upstream open reading frames (uORFs) to regulate protein expression of *GCN4* under amino acid starvation [Hinnebusch and Natarajan, 2002]. It is known that N^6-methylation of the adenosine (m^6A) in the 5' UTR regulates translation of the mammalian functional Gcn4 homolog Atf4 [Zhou et al., 2018]. To investigate whether *GCN4* contains Cbf5-dependent pseudouridylations in the 5' UTR, Primer Extension after CMC treatment along the 5' UTR of *GCN4* was performed in the *cbf5* catalytically inactive mutant. For

this purpose, primers at a distance of 100-150 bp were designed along the 5' UTR [Figure 24A]. Figure 24B/C shows the result of the primer extension after CMC treatment with the primers P1 and P5. The asterisks-labeled bands indicate discontinuations of the reverse transcription at the CMC-treated RNA of *CBF5* shuffle pRS314-*CBF5*. These discontinuations were not detected on the autoradiograph of the primer extension of CMC-untreated RNA as well as of CMC-treated RNA in the catalytically inactive *cbf5* mutant. This result indicates that the 5' UTR of *GCN4* contains four Cbf5-dependent pseudouridylation sites. Primer extension at *RPS16B* serves as a positive control, as this gene contains a known Cbf5-dependent pseudouridylation site [Appendix Figure 31A]. The other Primer extension with the primers P2-P4 and P6-P7 show the same discontinuation of the reverse transcription in the catalytically inactive *cbf5* mutant and in the corresponding wild-type [Appendix Figure 41].

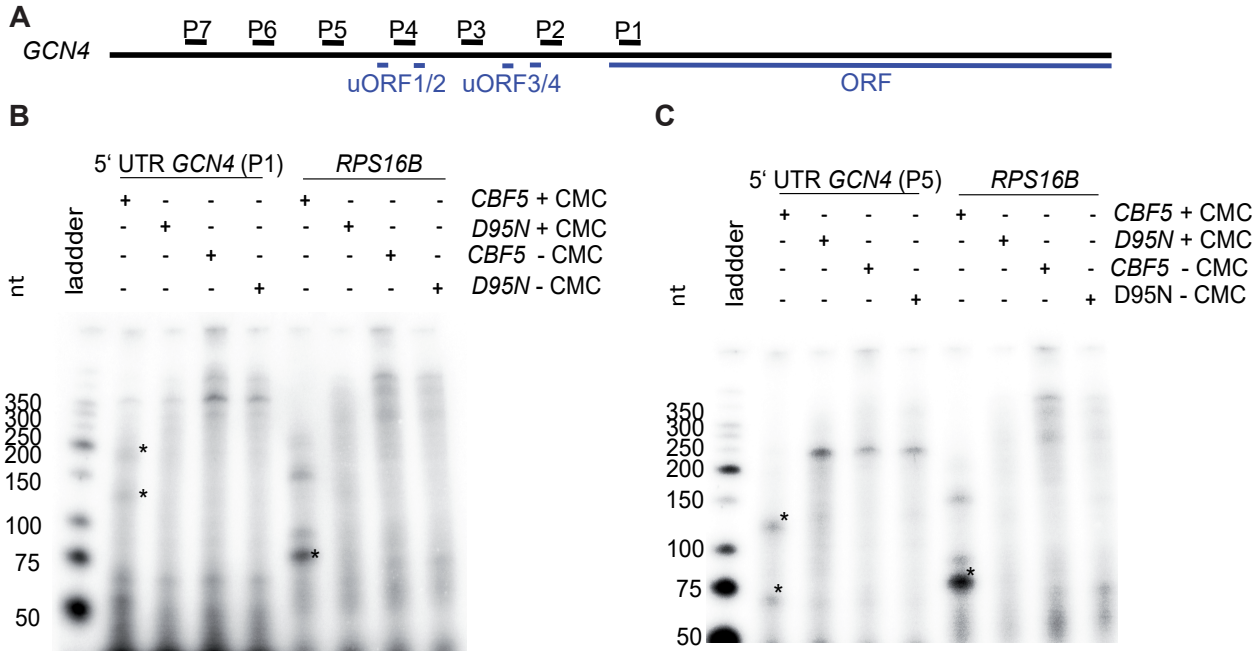

Figure 24: The 5' UTR of *GCN4* contains four Cbf5-dependent pseudouridylation-sites. A: Scheme of *GCN4* including the 5' untranslated region (5' UTR) is represented by a solid black line. The open reading frame (ORF) and the four untranslated open reading frames (uORFs 1-4) are marked by blue lines below. Positions of the primers used for the primer extension after CMC treatment are indicated by black bars above the solid line. **B/C:** Primer extension of CMC-treated (+CMC) and CMC-untreated (-CMC) RNA in the catalytically inactive *cbf5* mutant (*D95N*) and the corresponding wild-type (*CBF5*) at the 5' UTR of *GCN4* and at *RPS16B*. Asterisk-labeled bands show discontinuations of the reverse transcription. These are possible Cbf5-dependent pseudouridylation sites.

To summarize, Cbf5 has different functions in transcription. It influences transcription elongation [Figure 19] and it regulates the expression of genes involved in metabolism and energy generation [Figure 20/21]. Further experiments showed that Cbf5 regulates the protein amount of the transcription activator Gcn4 during amino acid starvation [Figure 23].

2.4.4 Cbf5 has a role in nuclear mRNA export

To unravel whether Cbf5 has a role in nuclear mRNA export FISHs in the *cbf5* catalytically inactive mutant and the *CBF5* depletion strain were performed. To visualize the localization of poly(A) mRNA, Cy3-labeled oligo dT_{50} primers were used. DAPI-staining of the DNA was conducted to visualize the nucleus. The microscopic images of Figure 25A show that the poly(A) mRNA is localized in the whole cell in the catalytic inactive *cbf5* mutant as well as in *CBF5* shuffle pRS314-*CBF5* and the wild-type strain. The yeast strain containing the point mutation at the catalytically center of *CBF5* (*cbf5-D95N*) displays no visible mRNA export defect. The quantification of the FISH verifies that a nuclear mRNA export defect can not be detected in the catalytically inactive *cbf5* mutant [Figure 25B]. FISHs in the *CBF5* depletion strain show that the poly(A) mRNA is localized as dots near the nuclear rim in some cells. Other cells show a poly(A) localization all over the cell. The Violin plot of the quantification displays that the *CBF5* depletion strain has a highly significant mRNA export defect. FISHs in wild-type cells in presence and absence of IAA show that auxin treatment does not induce an mRNA export defect.

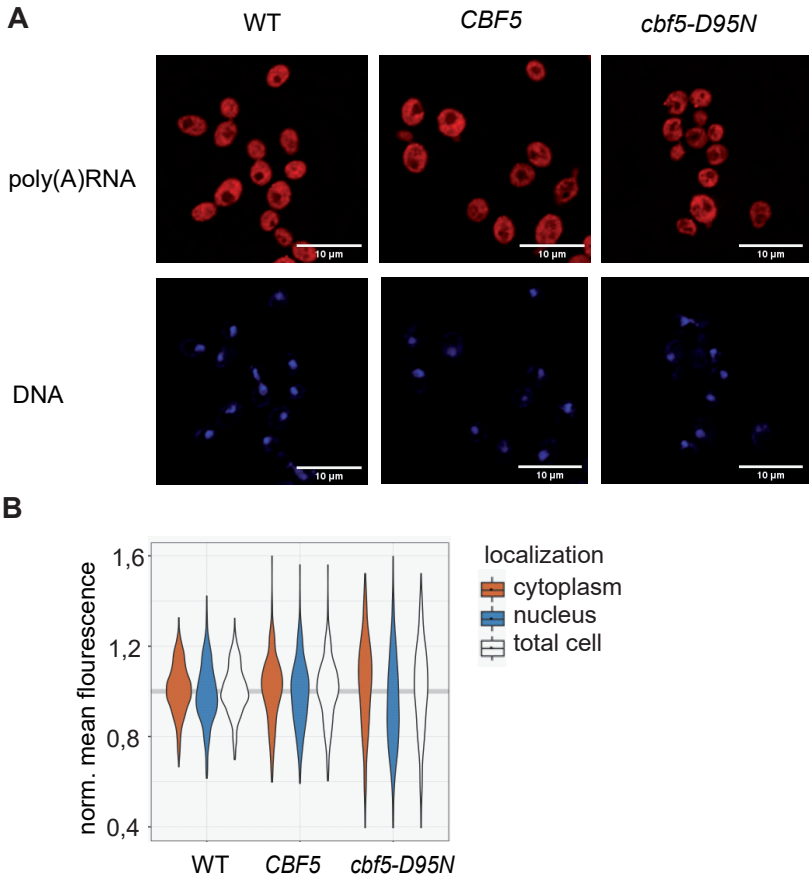

Figure 25: *cbf5* catalytically inactive mutant does not have a nuclear mRNA export defect A: FISHs were performed in the *cbf5* catalytically inactive mutant (*cbf5-D95N*), the *CBF5* shuffle strain, where *CBF5* is expressed from a plasmid (*CBF5*) and in the wild-type W303 (WT). Bulk mRNA was visualized with oligo dT_{50} primer coupled to a Cy3-fluorescent dye. DNA was stained with DAPI to locate the nucleus. **B:** Violin plot displays that the catalytically inactive *cbf5* mutant does not show a nuclear mRNA export defect. Quantification was conducted of 300 randomly selected cells using ImageJ. FISHs and the quantification were performed by Johanna Seidler.

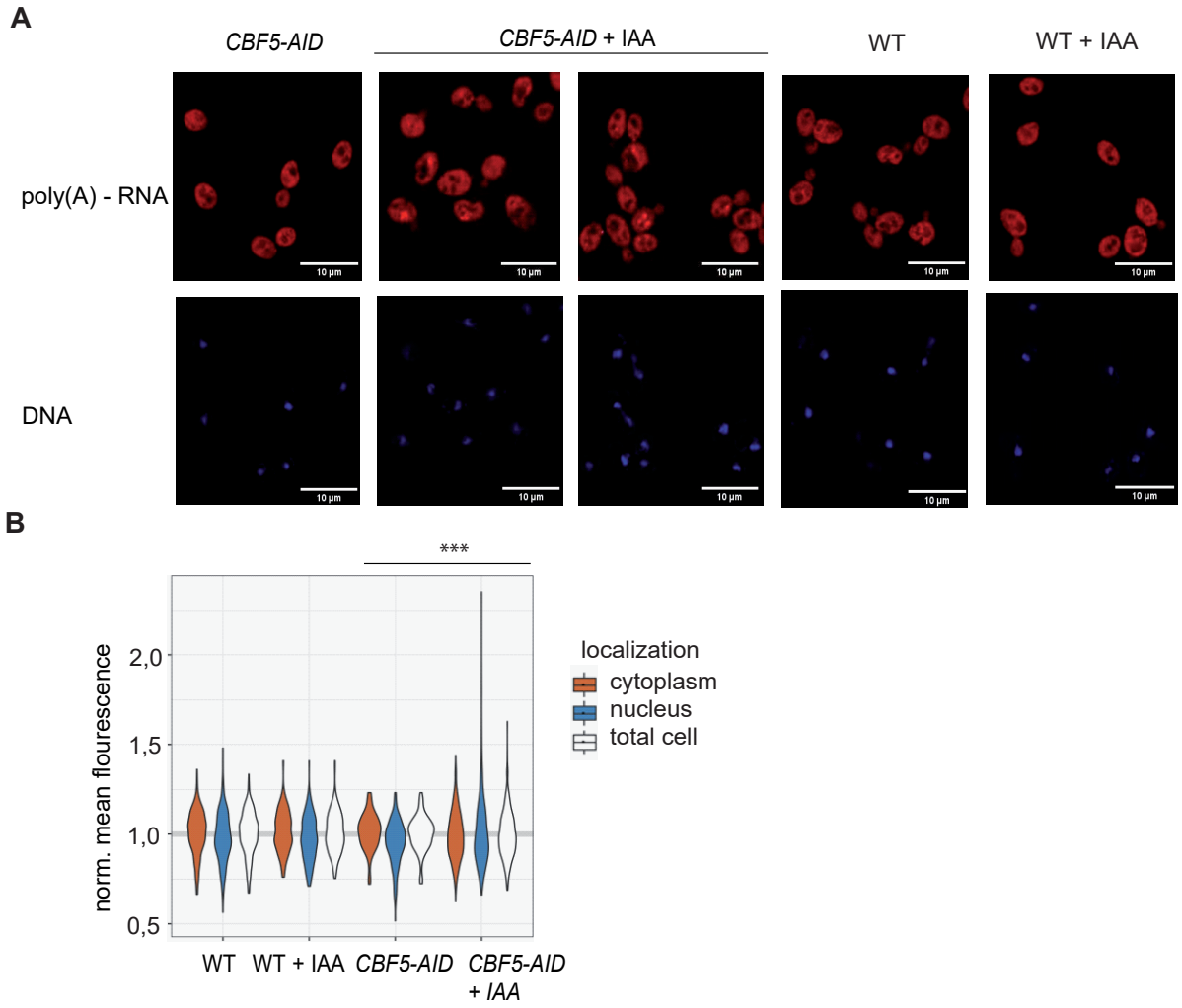

Figure 26: *CBF5* depletion strain shows a nuclear mRNA export defect A: FISHs were performed in the *CBF5* depletion strain (*CBF5-AID*) +/- auxin (IAA) and the wild-type strain (WT) +/- IAA. Bulk mRNA was visualized with oligo dT_{50} primer coupled to a Cy3-fluorescent dye. DNA was stained with DAPI to locate the nucleus. **B:** Violin plot displays that the *CBF5* depletion strain has a significant nuclear mRNA export defect. Quantification was conducted of 300 randomly selected cells using ImageJ. FISHs and the quantification were performed by Johanna Seidler.

2.5 Cbf5-dependent pseudouridylation has a role in transcription

Since 2014, it is known that mRNAs contain Cbf5-dependent pseudouridylation [Schwartz et al., 2014]. So far, nothing is known about the physiological function of these mRNA pseudouridylation. To study this, point mutations were inserted at the Cbf5-dependent pseudouridylation site of *ADE13* and *ELP4*.

At *ADE13* a silent point mutation at aa residue 350 (*ade13-S350S*) could be generated. The newly generated codon has a codon usage of 14.2%, whereas the codon usage of the unmutated codon is 23.5%. Other silent point mutations resulting in codons with a codon usage of 9.8% and 18.7% were generated at this position of the mRNA [*GenScript Codon Usage Frequency Table(chart) Tool*]. Quantitative Western Blots of whole cell lysates of these *ade13* PseudoU site mutants were performed to investigate whether different codon usages influence protein expression of *ADE13*. Figure 42 of the appendix shows that different codon usages of *ade13-S350* do not change the protein amount of Ade13 significantly. Further experiments were performed with *ADE13* shuffle pRS314-*ade13-S350S* were *ade13-S350* has a codon usage of 14.2%.

To generate point mutations at the pseudouridylation site in *ELP4*, amino acid exchanges have to be inserted. To reduce the possibility that an amino acid exchange induces protein misfolding or degradation, the PoPMuSiC v3.0 web server from dezyme was used [Dehouck et al., 2011]. This tool provides an *in silico* prediction of thermodynamic stability changes of the protein upon single-site mutation. Based on the prediction analysis, phenylalanine 126 was mutated to tyrosine and tryptophan. An *ELP4* deletion strain and an *ADE13* shuffle strain were transformed with the plasmids containing the point mutations at the Cbf5-dependent pseudouridylation sites. Primer extension after CMC treatment was performed to assess that Cbf5-dependent pseudouridylation does not exist in the PseudoU site mutants. The autoradiographs of Figure 40 in the appendix show that the discontinuation of the primer extension can not be detected when the Cbf5-dependent PseudoU site of *ELP4* and *ADE13* is mutated. CMC-untreated RNA serves as negative control. The primer extension at *ELP4* was also performed in the catalytically inactive *cbf5* mutant and the corresponding wild-type as a control [Appendix Figure 40B].

Figure 27 shows the analysis of the function of the Cbf5-dependent pseudouridylation at *ADE13*. At first RT-qPCR of total RNA at *ade13-S350S* and *ADE13* was performed. Oligo dT primers were used for reserve transcription and primers annealing at the 3' end of *ADE13* were utilized to quantify total mRNA amounts. To determine Ade13 protein levels of the *ade13* PseudoU site mutant and the corresponding wild-type quantitative Western Blots were performed. For this purpose, whole cell lysates of *ADE13-TAP* tagged yeast

strains and an antibody against protein A of the TAP tag was used. Pgk1 was used for normalization and the quantification was done with ImageJ. Figure 27A&B shows that the mRNA level of *ADE13* as well as its protein level is reduced in the *ADE13* shuffle strain containing a silent point mutation at *ade13-S350*. The mRNA amount is decreased to around 40% [Figure 27A] and the quantification of the Western Blot analysis shows that the protein level is reduced to about 50% [Figure 27B]. To further investigate the decreased *ADE13* mRNA and protein amount of *ADE13* shuffle pRS314-*ade13-S350S*, an *in vivo* transcription assay was performed. To assess mRNA synthesis of the *ade13* PseudoU site mutant *in vivo*, yeast strains containing *ADE13* or *ade13-S350S* under the control of the galactose inducible promoter *GAL10* were generated. The enrichment of *ADE13* was normalized to the amount of the permanently transcribed gene *PMA1*. Levels of *ADE13* mRNA were decreased in the PseudoU site mutant [Figure 27C], suggesting that the Cbf5-dependent pseudouridylation in *ADE13* might influence transcription of this gene. To study whether reduced *ADE13* mRNA levels were due to decreased mRNA stability in the *ade13* PseudoU site mutant, a thiolutin assay was performed. Cells were treated with 8 μg/mL thiolutin to block transcription. Samples were harvested 0, 10, 20, 40 and 60 min after the addition of thiolutin. The determined half-life of *ADE13* mRNA is 18.37 +/- 5 min and the mRNA half-life of *ade13-S350S* mRNA amounts 12.54 +/- 2 min. Although the measured half life was lower in the mutant, the difference was not statistically significant [Figure 27D].

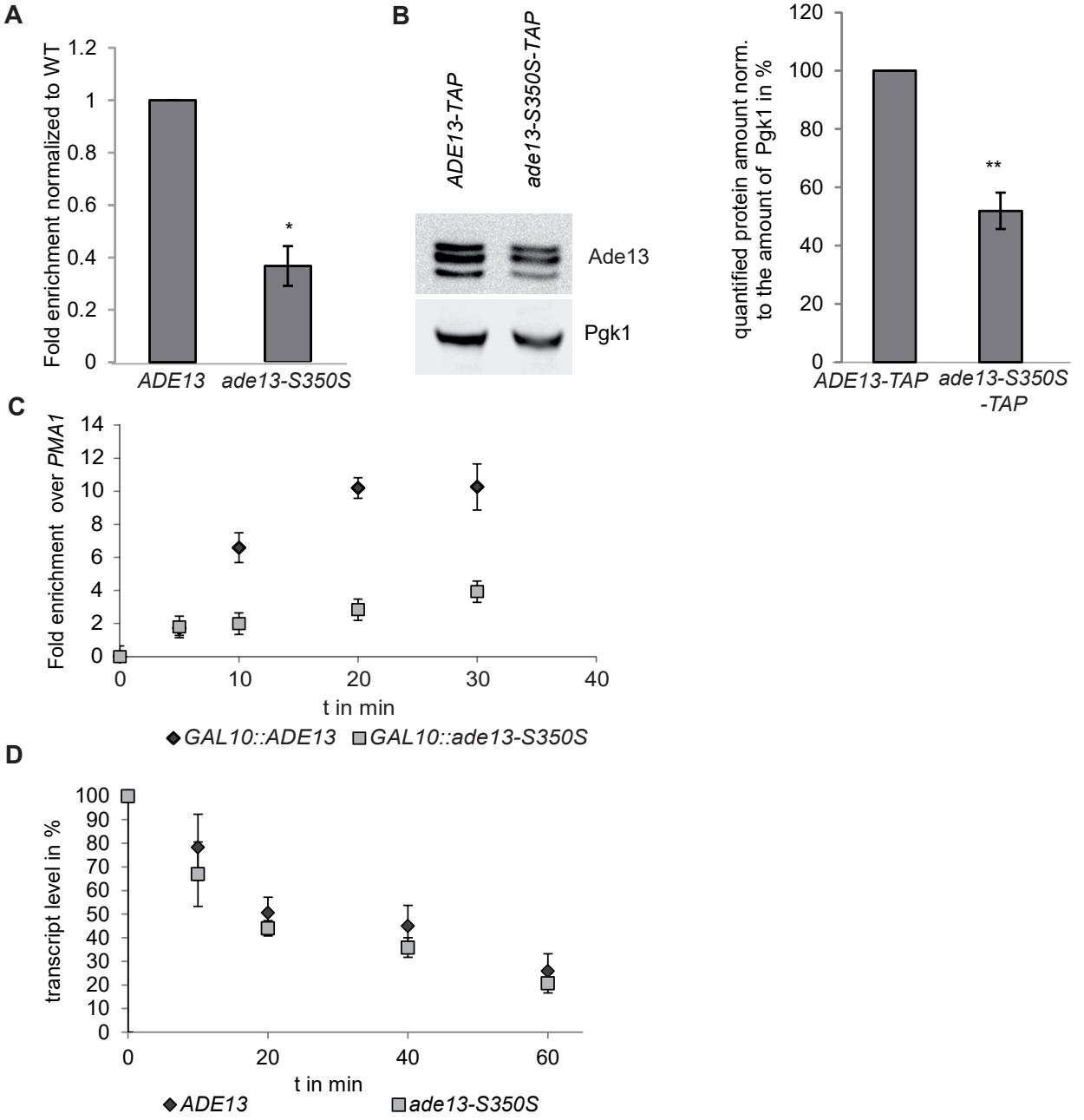

Figure 27: Mutation of the Cbf5-dependent pseudouridylation in *ADE13* influences its transcription. A: Total mRNA level of *ade13-S350S* is 60% decreased compared to *ADE13*. mRNA levels were determined by RT-qPCR. **B:** Ade13 protein level of *ADE13* shuffle pRS314-*ade13-S350S* is 40% decreased compared to the corresponding wild-type strain. Protein levels were determined by quantitative Western Blots of whole cell extracts and quantification was conducted by ImageJ. Pgk1 was used for normalization. **C:** mRNA synthesis of *ade13-S350S* is reduced compared to *ADE13*. *ADE13* and *ade13-S350S* were expressed under the control of the inducible *GAL10* promoter. The level of *ADE13* was assessed by RT-qPCR 0, 5, 10, 20 and 30 min after induction with galactose and normalized to the amount of the permanently transcribed *PMA1*. **D:** The mRNA half-live of *ADE13-S350S-TAP* ($ade13\text{-}S350S_{50\%}$ = 18.37 +/- 5 min) does not show a significant change compared to the wild-type ($ADE13_{50\%}$ = 12.54 +/- 2 min). Yeast cells were treated with 8 μg/mL thiolutin to block transcription.

The level of *ADE13* was assessed by RT-qPCR 0, 10, 20, 40 and 60 min after the addition of thiolutin. **A-D:** Bars represent the mean +/- standard deviation of at least three independent biological replicates. p values of students t-test: **: $p < 0.01$; *: $p < 0.05$

Figures 28/29 show the analysis of the function of the Cbf5-dependent pseudouridylation in *ELP4*. The experiments described above were also performed in the *elp4* PseudoU site mutants. RT-qPCR shows that the total mRNA amount of *ELP4* is reduced in *Δelp4* pRS314-*elp4-F126W*, however it is unchanged in *Δelp4* pRS314-*elp4-F126Y* [Figure 28A]. The *in vivo* transcription assay indicates that mRNA synthesis of *ELP4* is not affected when the PseudoU site of *ELP4* at the aa residue 126 is mutated. The thiolutin assay to determine mRNA stability displays that *elp4-F126W* mRNA half-life of 2.23 +/- 2 min is slightly reduced compared to the wild-type ($mRNA_{50}$ *ELP4* = 8.55 +/- 4 min), whereas the mRNA half-life of *elp4-F126Y* (*elp4-F126Y*$_{50\%}$ = 5.96 +/- 2 min) is not significantly changed. The quantitative Western Blot analysis demonstrates that the Elp4 protein amount is highly increased in both *elp4* PseudoU site mutants [Figure 29A], indicating that Cbf5-dependent pseudouridylation might influence protein expression. To investigate whether Elp4 protein stability is altered when the PseudoU site is mutated, a cycloheximide assay was performed. To stop mRNA translation 125 μg/ml cycloheximide was added to the yeast culture and samples were taken 0, 30, 60 and 90 min after the addition of cycloheximide. Quantitative Western Blot analysis of whole cell lysates of *ELP4-TAP* tagged yeast strains with an antibody against protein A was performed. The highly stable protein Pgk1 serves as a loading control. The cycloheximide assay shows that the protein stability of *ELP4-TAP* is increased in *Δelp4* pRS314-*elp4-F126Y* and *Δelp4* pRS314-*elp4-F126W* [Figure 29B]. *ELP4-TAP* of *Δelp4* pRS314-*elp4-F126Y* has a protein half-life of 137.59 +/- 23 min, the same protein of *Δelp4* pRS314-*elp4-F126W* has a half-life of 141.63 +/- 24 min, whereas the protein half-life of the unmutated *ELP4-TAP* amounts 73.00 +/- 14 min.
To sum up, Cbf5-dependent pseudouridylation in *ADE13* might decrease mRNA transcription of this gene, whereas mutations at the PseudoU site of *ELP4* might regulate its protein amount.

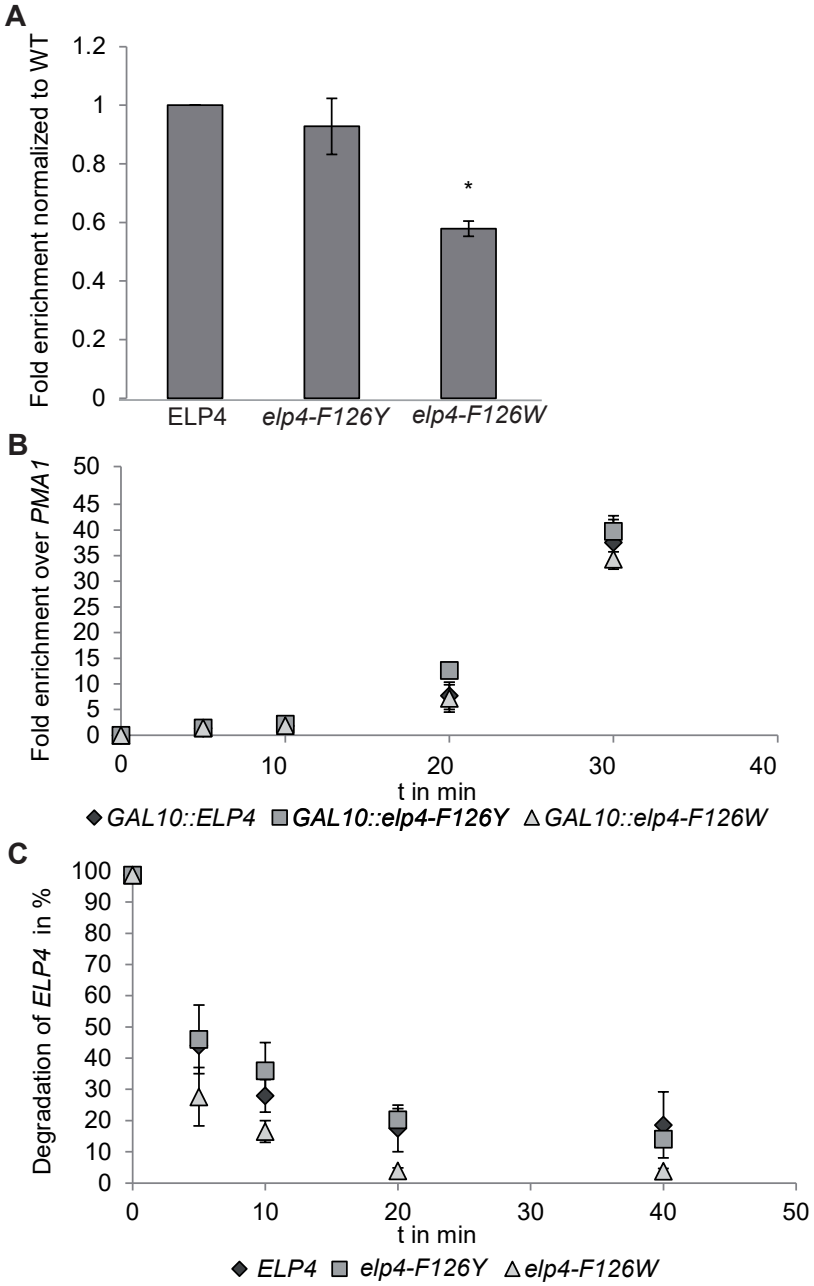

Figure 28: Mutation of the Cbf5-dependent pseudouridylation in *ELP4* does not influence its mRNA expression. A: Total mRNA level of *elp4-F126W* is 40% decreased compared to the wild-type, whereas the mRNA level of elp4-F126Y is similar to the wild-type. mRNA levels were determined by RT-qPCR. **B:** *elp4* PseudoU site mutants does not influence its mRNA levels. *ELP4* and *elp4-F126W/Y* were expressed under the control of the inducible *GAL10* promoter. The level of *ELP4* was assessed by RT-qPCR 0, 5, 10, 20 and 30 min after induction with galactose and normalized to the amount of the permanently transcribed *PMA1*. **C:** The mRNA half-life of *elp4-F126W* ($elp4\text{-}F126W_{50\%}$ = 2.23 +/- 2 min) is slightly reduced compared to the wild-type ($ELP4_{50\%}$ = 8.55 +/- 4 min), whereas the mRNA half-life of *elp4-F126Y* ($elp4\text{-}F126Y_{50\%}$ = 5.96 +/- 2 min) is not significantly changed. Yeast cells were treated with 8 μg/mL thiolutin to block transcription. The level of *ELP4* was assessed by RT-qPCR 0, 5, 10, 20 and 40 min after the addition of thiolutin. **A-D:** Bars represent the mean +/- standard deviation of at least three independent biological replicates. p values of students t-test: *: $p < 0.05$

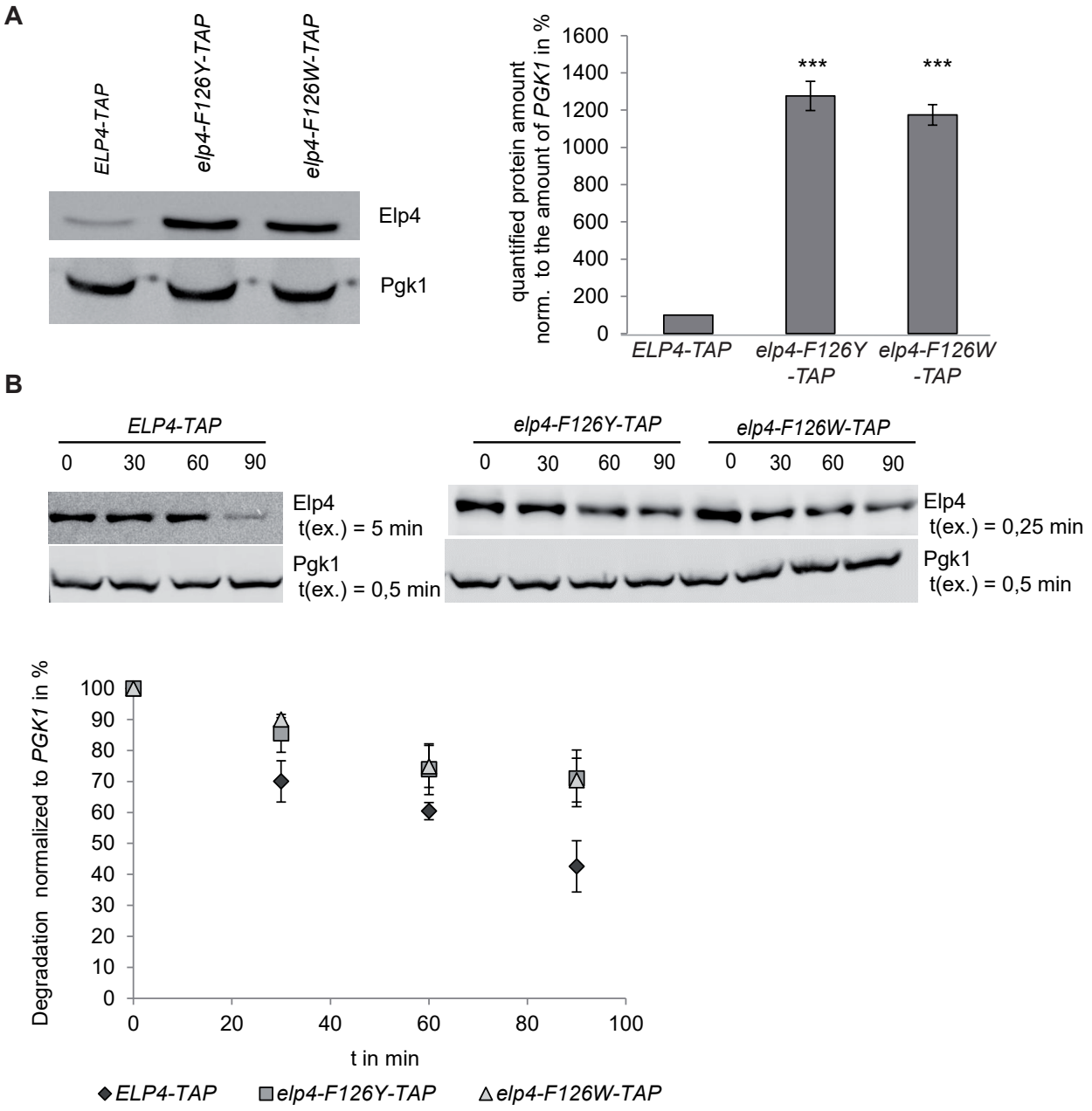

Figure 29: Mutation of the Cbf5-dependent pseudouridylation in *ELP4* regulates its protein amount. A: *ELP4-TAP* protein levels of *Δelp4* pRS314-*ELP4-F126Y-TAP* and *Δelp4* pRS314-*ELP4-F126W-TAP* are 12-14 fold increased compared to the corresponding wild-type strain. Protein levels were determined by quantitative Western Blots of whole cell extracts and quantification was conducted by ImageJ. Pgk1 was used for normalization. **B:** The protein half-live of *ELP4-TAP* is increased in *Δelp4* pRS314-*ELP4-F126Y-TAP* ($ELP4\text{-}F126Y\text{-}TAP_{50\%}$ = 137.59 +/- 23 min) and *Δelp4* pRS314-*ELP4-F126W-TAP* ($ELP4\text{-}F126YW\text{-}TAP_{50\%}$ = 141.63 +/- 24 min) compared to the corresponding wild-type strain ($ELP4\text{-}TAP_{50\%}$ = 73.00 +/- 14 min). The level of *ELP4-TAP* was assessed by quantitative Western Blot 0, 30, 60 and 90 min after the addition of 125 μg/ml cycloheximid and quantified by ImageJ. Pgk1 serves as a loading control. t(ex.) = exposure time **A-B:** Bars represent the mean +/- standard deviation of at least three independent biological replicates. p values of students t-test: ***: $p < 0.001$

3 Discussion

Cbf5 was originally described as a centromere-binding factor that has a role during mitosis and meioses [Jiang et al., 1993]. As a part of the H/ACA snoRNP, Cbf5 catalyzes pseudouridylation of uridine residues of rRNAs [Schattner et al., 2004/ Watkins et al., 1998] and it has also a function in rRNA processing [Fayet-Lebaron et al., 2009]. In 2014, it was published that Cbf5 modifies uridines of mRNAs [Schwartz et al., 2014]. The function of Cbf5-dependent mRNA pseudouridylation is unknown and remains to be understood. The aim of the present study was to unravel the function and mechanism of Cbf5-dependent mRNA pseudouridylation and to elucidate the role of Cbf5 in the mRNA life cycle.

3.1 TREX interacts with the H/ACA snoRNP

To unravel whether the H/ACA snoRNP or only Cbf5 interacts with TREX, two crosslinks between Gar1 at aa residue 59 and the C-terminus of Tho2 were identified, performing a XL-MS analysis of both complexes. The detected crosslinks between Cbf5 and the TREX components Tho2, Sub2, Hpr1 and Thp2 of former experiments could not be reproduced [Figure 7]. In general, only few crosslinks within one complex were identified. The identified crosslinks between the C-terminus of Hpr1 and Tho2 as well as between Thp2 and Mft1 are the most frequently identified crosslinks of a XL-MS analysis of TREX [Figure 8C], which were also found by Manuel Koschitza [Figure 7] and by Xie et al., 2021a. Crosslinks between proteins of the H/ACA snoRNP were mainly identified between Cbf5 and Gar1 as well as Nop10 and Nhp2 [Figure 8C], which is consistent to the published XL-MS results of the H/ACA snoRNP [Ding et al., 2016]. That shows that the XL-MS results of this study are reliable. Due to the low number of the identified crosslinks between both complexes, it was not possible to predict a structure of the H/ACA snoRNP/TREX interaction. This is exacerbated by the fact that the crosslinks at aa residue 419 of Cbf5 are located in a structural flexible region, whose structure is not solved, yet. However, these results indicate that the H/ACA snoRNP interacts with TREX.

TAPs of one complex and subsequent Western Blots of the other complex verified the XL-MS results. A physical interaction between Yra1 and Gar1, as well as Nhp2 was also identified applying TAP of the H/ACA snoRNP and a subsequent LC-MS/MS [Schwer et al., 2011]. The purification of Nhp2 in presence of RNase A and the subsequent Western Blot analysis showed that Yra1 interacts directly with Nhp2. The H/ACA snoRNP fall apart into its individual proteins during the RNase A treatment, as Nhp2 does not directly interacts with Cbf5 or Gar1 [Li et al., 2011/ Koo et al., 2011]. That explains the absence of Cbf5 on the

SDS-polyacrylamide gel of the eluate after RNase treatment [Figure 9B] and shows that the interaction between Nhp2 and Yra1 is not mediated by Cbf5.
To finally conclude, an RNA-independent interaction between Nhp2/Gar1 of the H/ACA snoRNP and Hpr1/Yra1 of TREX could be shown. Although, it was not investigated whether Nop10 interacts with TREX, it is highly likely that the whole H/ACA snoRNP interacts with TREX.

3.2 Cbf5 recruitment occurs transcription-dependent and RNA-dependent

The next aim of this thesis was to unravel the recruitment mechanism of Cbf5 to chromatin and RNAs. So far, it is only published that Cbf5 is associated with the H/ACA snoRNA transcribed genes *SNR10, SNR30*, which are also RNAPII-transcribed. However, Cbf5 is not present at the C/D snoRNA transcribed gene *SNR13* [Ballarino et al., 2005].

3.2.1 The occupancy of Cbf5 at RNAPII transcribed genes is transcription-dependent

The ChIP of the strain *CBF5-TAP* showed that Cbf5 is present at three exemplary highly transcribed genes and five genes containing a Cbf5-dependent pseudouridylation-site [Figure 10B/C]. Although the enrichment of Cbf5 was only determined at eight genes, it is highly likely that Cbf5 is present at all RNAPII transcribed genes independently of whether the genes contain a Cbf5-dependent pseudouridylation-site. Further, a RIP of the strain *CBF5-TAP* revealed that Cbf5 binds to mRNAs [Figure 10E]. RIP also displays that the pseudouridine-synthase does not distinguish between mRNAs containing a pseudouridylation and those that do not have one.
Cbf5 can bind DNA, as it contains the centromere-binding domain *KKE/D*. Furthermore, a centromere DNA affinity column chromatography showed that Cbf5 is situated at the centromere during mitoses and meioses [Jiang et al., 1993]. However, it is not possible to distinguish whether Cbf5 binds to DNA sequences and/or nascent mRNA performing a ChIP, as formaldehyde crosslinks N-terminal amino groups and side chains of cysteines, histidines, lysines, tryptophans and arginines [Metz et al., 2004] to amino and imino groups of nucleic acid bases [Chang and Loew, 1994/ Fraenkel-Conrat, 1954], allowing no distinction between RNA and DNA.
The ChIP result of the strain *CBF5-TAP* at the inducible *GAL1* gene [Figure 10D] is consistent with the finding that Cbf5 is only present at actively transcribed regions of the H/ACA

snoRNA transcribed genes *SNR10, SNR30*. Cbf5 recruitment upstream of these transcribed genes was not observed [Ballarino et al., 2005 Figure 1].
Taken together, Cbf5 is present at RNAPII transcribed genes and mRNAs independent of whether the genes contain a Cbf5-dependent pseudouridylation-site. It is interesting to determine the function of Cbf5 at whose transcribed genes that do not have a pseudouridylation-site.

3.2.2 Identification of the Cbf5 recruiter to transcribed genes

THO and CBC function as recruiter of proteins onto nascent RNA transcripts [Hurt et al., 2004/ Sen et al., 2019]. Furthermore, both complexes interact with Cbf5 of the H/ACA snoRNP [Gruellich, 2017/ Fortes et al., 1999/ Schwer et al., 2011]. However, in this study it was shown that neither THO nor CBC is needed for Cbf5 occupancy at transcribed genes. Interestingly, Cbf5 enrichment normalized to the occupancy of RNAPII in the deletion strains increase compared to the wild-type strains [Figures 11/12/13]. A Cbf5 ChIP in a *NPL3* deletion strain normalized to RNAPII levels showed nearly the same result. However, the occupancy of Cbf5 is not increasing at every gene [Appendix Figure 36]. An RNAPII normalized ChIP of Cbf5 in a *SUB2* overexpression strain showed an unchanged Cbf5 occupancy. The enrichment of Cbf5 is only decreased at *CCW12* M [Appendix Figure 39]. This result showed that an overexpression of a TREX component does not have the opposite effect to a *HPR1* deletion. It seems as Cbf5 fills unoccupied positions at transcribed genes or newly synthesized mRNAs in a rather unspecific manner. To prove this hypothesis and to investigate whether Cbf5 binds to unoccupied positions due its strong RNA/DNA binding affinity, electrophoretic mobility shift assays (EMSAs) can be performed of Cbf5 and *cbf5* mutants containing point mutations in the *PUA* domain or the *KKE/D* domain.
To further investigate, how Cbf5 is recruited to transcribed genes, a Cbf5 ChIP in presence of RNase A and RNase T1 and a hepatitis λ ribozyme ChIP assay showed that the occupancy of Cbf5 is RNA-dependent [Figure 14A/B]. The RNase ChIP was conducted according to Abruzzi et al., 2004. The approach used in this study bears a technical shortcoming: The digestion of the RNA is carried out after cross-linking. Thus, proteins that need RNA to be recruited or to be stabilized at the transcription site can be cross-linked to RNA or DNA and would therefore be in-sensitive during RNase digestion in the ChIP assay. However, a 50% reduced Cbf5 occupancy was observed. The hepatitis λ ChIP assay, which also shows that Cbf5 occupancy is reduced by 50% verified this result [Figure 14D]. When a protein binds to nascent RNA containing the hepatitis λ ribozyme or needs RNA to be recruited or stabilized at the transcription site, the occupancy of this protein should drop downstream

of the cleavage site. The disadvantage of this experiment is that the effect can be only observed at *YCT1* as only this gene contains the hepatitis λ ribozyme. However, the same result was obtained with two different strategies. As Cbf5 occupancy is only reduced by 50%, it is highly likely that Cbf5 is not only recruited through direct interaction with RNA to transcribed genes. Besides the RNA itself, the CTD of Rbp1 of RNAPII serves as a recruitment platform of proteins involved in mRNA transcription and mRNA processing [Phatnani and Greenleaf, 2006]. Cbf5 ChIPs in *rbp1-S2A and rpb1-S5A* show that S2 and S5 phosphorylation of the CTD is not needed for Cbf5 occupancy [Appendix Figure 37/38]. The general elongation factor Spt5, which consists of the CTR serves as another recruitment platform. As a physical interaction between Cbf5 and Spt5 was observed performing mass spectrometry of a *SPT5-FLAG* co-immunoprecipitation [Lindstrom et al., 2003], it is worth to investigate whether the CTR of Spt5 serves as the recruiter of Cbf5.
To sum up, it could be shown that the occupancy of Cbf5 is RNA-dependent. However, TREX or CBC are not needed for Cbf5 occupancy at transcribed genes.

3.3 Cbf5-dependent pseudouridylation occurs post-transcriptionally

An RNA modification can occur while RNA polymerase II transcribes the mRNA or afterwards. To elucidate whether Cbf5-dependent pseudouridylation occurs co-transcriptionally or post-transcriptionally, Primer extension after CMC-treatment of total RNA, nascent mRNA and poly(A) mRNA were performed.
The analysis of the amount of poly(A) mRNA and nascent mRNA of the oligo-dT purification shows that both RNA types are highly enriched compared to total RNA [Table 1]. Although nascent mRNA does not contain a poly(A) tail, newly transcribed RNA was purified using oligo-dT beads. To optimize poly(A) mRNA isolation, different protocols and oligo-dT beads of several companies were used. Unfortunately, all purifications highly enrich poly(A) mRNA as well as nascent mRNA.
The amount of 25S rRNA of the isolated nascent mRNA and poly(A) mRNA is only slightly decreased [Table1]. That shows, that it was not possible to eliminate rRNAs performing an IP of *RBP3-FtpA* or an oligo-dT bead purification.
Nevertheless, primer extension after CMC treatment of total RNA, poly(A) mRNA and nascent mRNA was performed. Besides the asterisks labeled bands, many other unspecific bands are shown on the autoradiographs [Figures 16A/C/17A], as the reverse transcription of CMC-untreated RNA terminates at different points. These unspecific cDNAs were also detected on the autoradiographs of CMC-treated RNA, potentially due to unspecific CMC treatment of uridines or guanosines. Quantification of the asterisks marked signals on the au-

toradiographs suggest that Cbf5-dependent pseudouridylation occurs post-transcriptionally. However, Cbf5-dependent pseudouridylation levels of nascent mRNA were measurable due to poly(A) mRNA contaminations of the purified nascent mRNA. It was also not possible to purify pure poly(A) mRNA using oligo dT-beads. Furthermore, the rRNA content of nascent mRNA and poly(A) mRNA is only slightly reduced. All these problems make it difficult to draw a definitive conclusion whether Cbf5-dependent pseudouridylation occurs post-transcriptionally, particularly as the results of a primer extension after CMC treatment are only semi-quantitative [Heiss and Kellner, 2017/ Zaringhalam and Papavasiliou, 2016].
Recently, it was published that mRNA pseudouridylation by the pseudouridine synthases Pus1, Pus7 and RPUSD4 occurs co-transcriptionally in human cells [Martinez et al., 2022]. They isolated chromatin-associated RNA using a biochemical cellular fractionation that enriches for intron-containing unspliced pre-mRNAs and detected pseudouridylation sites performing Pseudo-seq [Carlile et al., 2014]. That shows that two different methods to elucidate the occurrence of mRNA pseudouridylation receive two different results. However, it is possible that Cbf5-dependent pseudouridylation occurs post-transcriptionally in *S. cerevisiae*, while pseudouridylation by the pseudouridine synthases Pus1, Pus7 and RPUSD4 occur co-transcriptionally in human cells.
Taken together, primer extension after CMC treatment of total RNA, nascent- and poly(A) mRNA indicates that Cbf5-dependent pseudouridylation occurs predominantly post-transcriptionally. However, technical problems with the experiment especially during nascent - and poly(A) mRNA isolation leave some doubt on the result. Another method to isolate nascent- and poly(A) mRNA has to be conducted to verify whether Cbf5-dependent pseudouridylation occurs post-transcriptionally.

3.4 Cbf5 functions in mRNA transcription and nuclear mRNA export

So far, it was shown that Cbf5 is present at all tested transcribed genes [Figure 10] and it interacts with TREX in an RNA-independent manner [Gruellich, 2017]. A function of the Cbf5/TREX interaction is not known, yet. TREX is a nuclear complex, which couples transcription and nuclear mRNA export [Straesser et al., 2002]. The next aim was to unravel whether Cbf5 also functions in mRNA synthesis or nuclear mRNA export. To this end, a catalytically inactive *cbf5* mutant was generated. Interestingly, this strain is not lethal [Figure 18]. Viability of a *CBF5* depletion strain can be rescued by expressing a catalytically inactive *cbf5* mutant [Zebarjadian et al., 1999]. This result showed that Cbf5 might have a pseudouridylation-independent function, which is essential for viability of *S. cerevisiae*. That is probably related to the centromere-binding function of Cbf5, as a C-terminal deletion of *CBF5* including the KKE/D domain is lethal. However, it is possible to delete half of the KKE/D domain, as *cbf5-aa 1-443* is viable [Jiang et al., 1993]. Besides the 18S rRNA processing function of Cbf5 as a part of the H/ACA snoRNP might also be essential for viability [Fayet-Lebaron et al., 2009]. Furthermore, it is conceivable that Cbf5 has another unknown essential function.

3.4.1 Cbf5 has a function in mRNA transcription elongation

To study whether Cbf5 functions in transcription elongation, growth assays on plates containing 6-AU suggested that Cbf5 has a role in transcription elongation [Figure 18]. Growth assays on plates containing 6-AU is a common experiment in *S. cerevisiae* to get an indication of a potential role in transcription elongation. 6-AU leads to a reduction of intracellular GTP and UTP levels that influences yeast growth combined with mutations that affect transcription elongation [Exinger and Lacroute, 1992]. Interestingly, the catalytically inactive mutant of *cbf5* grows faster on plates containing 6-AU than on plates containing DMSO (negative control). Furthermore, this mutant partially rescues 6-AU sensitivity of the transcription elongation factor Dst1. It seems as Cbf5 has a role in transcription elongation in the opposite direction than Dst1, which has a growth defect on plates containing 6-AU.
Paul B. Mason and Kevin Struhl established experiments to determine RNAPII processivity (nucleotide additions per initiation event) and RNAPII elongation rate (nucleotide additions per min) to learn more about factors that influence transcription elongation in *S. cerevisiae* [Mason and Struhl, 2005]. In this study, ChIPs to identify RNAPII processivity verified that Cbf5 functions in transcription elongation [Figure 18C/D]. A previous study about the

deubiquitylase Ubp15 showed that 6-AU sensibility of a *Δdst1* strain can be rescued in a *Δdst1 Δubp15* double mutant. However, reduced RNAPII processivity in *Δdst1* could not be rescued in the double mutant, as RNAPII processivity is unchanged in *Δubp15* [Eyboulet et al., 2020]. It will be interesting, whether the catalytically inactive *cbf5* can rescue RNAPII processivity in *Δdst1*. Furthermore, ChIPs at the long gene could be conducted to assess whether the transcription rate of RNAPII is changed when Cbf5 is depleted or catalytically inactive. *In vivo* and *in vitro* transcription assays could be performed in the *cbf5* mutants to confirm that Cbf5 influences transcription elongation. Nevertheless, the results suggest that Cbf5 has a pseudouridylation-dependent function in mRNA transcription elongation.

3.4.2 Cbf5 regulates expression of genes involved in metabolism and energy generation

To further investigate the role of Cbf5 in transcription elongation, RNAPII ChIPseq of the *CBF5* depletion strain and the catalytically inactive *cbf5* mutant identified 1556 genes that show a changed RNAPII occupancy in at least one of the *cbf5* mutants [Figures 20]. Most of the regulated genes in the yeast strain containing the mutation *cbf5-D95N* show a decreased RNAPII occupancy, whereas nearly the same amount of genes were donwregulated and upregulated in the *CBF5* depletion strain [Figure 20C], indicating that Cbf5 regulates transcription in different ways. However, only four genes in the catalytically inactive *cbf5* mutant show an increased RNAPII occupancy [Figure 20C]. One of these upregulated genes is *CBF5*. That is probably a regulatory mechanism of the cell to compensate the strong growth defect in the catalytically inactive *cbf5* mutant. Furthermore, gene ontology term enrichment analysis display that many genes involved in energy generation or in metabolism of amino acids, carbohydrates or nucleic acids show a changed RNAPII occupancy [Figures 20E; 21A/B]. It seems as Cbf5 regulates the expression of transcription factors involved in transcription of these genes. Besides many genes that have a function in rRNA processing and ribosome biogenesis were downregulated in the *CBF5* depletion strain, as the H/ACA snoRNP containing the snoRNA snR30 has a pseudouridylation-independent role in 18S rRNA processing [Zebarjadian et al., 1999]. ChIP qPCR of both *cbf5* mutants confirmed this statement. Genes involved in ribosome biogenesis and rRNA processing show a decreased RNAPII occupancy in the *CBF5* depletion strain, whereas RNAPII occupancy in the catalytically inactive *cbf5* mutant is unchanged [Figure 22A].
To summarize, it was shown for the first time that Cbf5 might regulate the expression of genes involved in metabolism and energy generation.

3.4.3 Cbf5 regulates the protein amount of the transcription activator *GCN4*

To unravel whether Cbf5 regulates expression of transcription factors, a comparison of RNAPII ChIPseq results of this study with other RNAPII ChIPseq data of transcription factor deletion strains in *S. cerevisiae* was performed [Kemmeren et al., 2014]. This comparison revealed that 29% (93 genes) of all regulated genes of a RNAPII ChIPseq in a *GCN4* deletion strain show also a changed RNAPII occupancy in at least one the *cbf5* mutants [Figure 23A]. Further experiments showed that Cbf5 regulates the protein amount of the transcription activator Gcn4, whereas the mRNA level of *GCN4* is not affected by Cbf5 [Figure 23B/C]. Furthermore, it was shown that the 5' UTR of *GCN4* highly likely contains four Cbf5-dependent PseudoU sites [Figure 24B/C]. PseudoU sequencing of the 5' UTR of *GCN4* has to be performed to confirm the existence of these Cbf5-dependent pseudouridylation sites, as it is not possible to determine the exact position of a pseudouridylation applying primer extension after CMC labeling. Afterwards, point mutations at the detected Cbf5-dependent pseudouridylation site can be inserted to investigate whether translation of *GCN4* is pseudouridylation-dependent.

It is known that translation of *GCN4* is regulated through four uORFs in the 5' UTR of *GCN4*. Under normal conditions the methionyl initiator tRNA (Met-tRNA$_i$) is transfered to the 40S ribosomal subunit in a ternary complex with the elongation factor 2 (elF2) and GTP to form the 43S initiation complex [Sonenberg and Hinnebusch, 2009]. The initiation complex scans the 5' UTR of *GCN4* to search AUG of the uORF1. Upon base-pairing of the anticodon of Met-tRNA$_i$ with AUG, GTP of the ternary complex is hydrolyzed to GDP, releasing elF2-GDP. elF2-GDP is recycled to elF2-GTP by the elongation factor 2B (elF2B) to reform the ternary complex. Nearly all initiation complexes scanning downstream from uORF1 rebind the ternary complex and re-initiate translation at uORF2/3/4, after which they dissociate from the mRNA [Hinnebusch, 1997]. Translation of *GCN4* is repressed under normal conditions. During amino-acid starvation the protein kinase Gcn2 phosphorylates elF2 [Natarajan et al., 2001]. Phosphorylation of elF2 converts eIF2-GDP from a substrate to an inhibitor of eIF2B, impeding the formation of the ternary complex. Due to these delayed association of the 40S ribosomal subunit with the ternary complex, uORF 3/4 are often bypassed, which results in efficient translation of the *GCN4* open reading frame [Hinnebusch and Natarajan, 2002].

One potential Cbf5-dependent PseudoU site in the 5' UTR of *GCN4* is maybe a part of uORF4. It is the PseudoU site that was detected on the autoradiograph of the primer extension with the primer P1. PseudoU sequencing of the *cbf5* catalytically inactive mutant will determine, whether uORF4 contain a Cbf5-dependent pseudouridylation.

It was also shown that the functional mammalian homologue Atf4, whose translation is regulated in a similar way as the protein expression of *GCN4*, contains m^6A modifications in the 5' UTR of *ATF4* including one within uORF2 [Zhou et al., 2018]. Furthermore, it was shown that the m^6A modification in the uORF2 influences translation. Removal of m^6A of the uORF2 results in increased translation of the ORF and the uORF2 of *ATF4* [Zhou et al., 2018]. Zhou et al. proposes that m^6A impedes scanning ribosomes, thus promoting translation initiation of uORF2 and the ORF of *ATF4*. This regulation mechanism is independent of the translation regulation during amino acid starvation. Both regulation mechanism enable efficient translation of the *ATF4* ORF.
To investigate whether Cbf5-dependent pseudouridylation in-stead of m^6A modification influences ribosome scanning and thereby regulates translation of *GCN4* in *S. cerevisiae*, ribosome profiling can be performed. It is already known that mRNA pseudouridylation has an impact on translation. *In vitro* transcribed pseudouridylated mRNA shows a higher translation rate [Kariko et al., 2008]. Furthermore, it is published that mRNA pseudouridylation can modulate translatability [Eyler et al., 2019].
To summarize, it is highly likely that the 5' UTR of *GCN4* contains Cbf5-dependent pseudouridylation sites, which regulate the translation of this gene. RNAPII ChIPseq in the *CBF5* depletion strain and the catalytically inactive *cbf5* mutant indicates that many genes involved in amino acid or peptide metabolism show a changed RNAPII occupancy. Cbf5 might regulate the expression of these genes, as it controls translation of Gcn4, the transcription activator of amino acid biosynthetic gene. Further research is needed to understand this potential pseudouridylation-dependent regulation mechanism. Nevertheless, that is the first indication of a function of a pseudouridylation site in the 5' UTR of a mRNA.

3.5 Cbf5 has a role in nuclear mRNA export

It was shown that Cbf5 has a role in transcription elongation just as TREX. To investigate, whether Cbf5 also functions in nuclear mRNA export FISHs were performed, which showed a significant nuclear mRNA export defect in the *CBF5* depletion strain [Figure 26]. That export defect was not observed in the catalytically inactive *cbf5* mutant [Figure 25], which demonstrates that Cbf5 has a pseudouridylation-independent function in nuclear mRNA export. The export defect in the *CBF5* depletion strain could not be observed in every cell. That may be due to the fact that Cbf5 is not equally degraded in every cell at the same time during depletion. The poly(A) mRNA in the cells that show a nuclear mRNA export defect is located as dots around the nuclear rim. To localize poly(A) mRNA more precisely, co-localization with proteins of the nuclear membrane e.g. nucleoporins can be

performed. Furthermore, it will be interesting, whether all or only specific mRNAs were not exported to the cytoplasm in a *CBF5* depletion strain. To study this, single molecule FISH of specific mRNAs can be conducted. Taken together, it was shown for the first time that a pseudouridine synthase can alter nuclear mRNA export. So far, it was only published that the human m^6A- binding protein *YTHDC1* mediates nuclear mRNA export of m^6A modified RNA [Roundtree et al., 2017]. Furthermore, this pseudouridylation-independent function of Cbf5 in nuclear mRNA export is probably essential for viability of *S. cerevisiae*. To prove this hypothesis, a *cbf5* mutant, which does not have a role in nuclear mRNA export, has to be generated.

3.6 Cbf5-dependent pseudouridylation has a role in transcription

The last aim of this thesis was to elucidate the function of Cbf5-dependent pseudouridylation on mRNA. It is known that mRNA pseudouridylation generated by stand-alone pseudouridine synthases have an impact on mRNA splicing [Martinez et al., 2022], mRNA translation [Karijolich and Yu, 2011/ Fernandez et al., 2013] and mRNA stability [Kariko et al., 2008/ Nakamoto et al., 2017]. Some mRNA pseudouridylation were regulated under stress conditions like heat stress [Schwartz et al., 2014] or glucose starvation [Carlile et al., 2014]. The physiological function of Cbf5-dependent pseudouridylation was studied at *ADE13* and *ELP4*.

Ade13 is an adenylosuccinate lyase, which catalyzes two steps in the *de novo* purine nucleotide biosynthetic pathway [Guetsova et al., 1997]. An *ade13* PseudoU site mutant reduces its mRNA and protein levels [Figure 27A/B], as the *ade13-S350S* transcription rate is reduced in *ADE13* shuffle pRS314-*ade13-S350S* compared to the wild-type strain [Figure 27C]. To perform the *in vivo* transcription assay, oligo dT primers instead of a primer amplifying nascent *ADE13* mRNA were used to quantify the amount of *ADE13* after galactose induction. As oligo dT primers amplify mRNAs containing a poly(A) tail, not exclusively newly synthesized mRNAs were quantified during RT-qPCR. However, as the thiolutin assay of Figure 27D shows that stability of the *ADE13* mRNA is unchanged, a repetition of the *in vivo* transcription assay with a primer amplifying only nascent *ADE13* mRNA will be a good control. Nevertheless, these experiments suggest that the Cbf5-dependent PseudoU site of *ADE13* has a function in transcription. Additionally, the ChIPseq result of the *CBF5* depletion strain shows that RNAPII occupancy at *ADE13* decreases (log2 fold change = -0.48) when Cbf5 is depleted. This result supports that Cbf5-dependent pseudouridylation at *ADE13* regulates its transcription. To confirm this result an *in vivo* transcription assay in the *CBF5* depletion strain compared to the wild-type strain will be performed. Further

research is needed to unravel how Cbf5 regulates transcription of *ADE13*. To investigate whether transcription elongation of *ADE13* is influenced by a Cbf5-dependent pseudouridylation, growth assays on SDC -Ura plates in presence and absence of 6-AU can be performed [Riles et al., 2004].
Many other genes involved in purine nucleotide biosynthesis show a decreased RNAPII occupancy in the *CBF5* depletion strain [Figure 21]. It is not known yet, whether these genes also contain a Cbf5-dependent PseudoU site. To unravel whether Cbf5-dependent pseudouridylation at other genes involved in purine nucleotide biosynthesis regulates transcription, the generated ChIPseq data and PseudoU-seq data have to be compared. For this purpose, reliable PseudoU-seq data of a *CBF5* depletion strain or a *cbf5* catalytically inactive mutant with the corresponding wild-types are needed.

Elp4 is a subunit of the hexameric RecA-like ATPase Elp456 Elongator subcomplex, which is a major histone acetyltransferase component of the RNAPII responsible for transcriptional elongation [Wittschieben et al., 1999]. The analysis of the function of the Cbf5-dependent pseudouridylation in *ELP4* shows that the protein amount as well as the protein stability is increased in the *elp4* PseudoU site mutants [Figure 29A/B]. The amino acid exchange of phenylalanine to tryptophan or tyrosine has a destabilizing effect according the prediction analysis using the PoPMuSiC v3.0 web server [Dehouck et al., 2011], indicating that the amino acid exchanges do not cause increased protein stability of the *elp4* PseudoU site mutants. Whereas the protein half-life is increased around 2 fold ($ELP4\text{-}F126Y\text{-}TAP_{50\%}$ = 137.59 +/- 23 min; $ELP4\text{-}F126W\text{-}TAP_{50\%}$ = 141.63 +/- 24 min; $ELP4\text{-}TAP_{50\%}$ = 73.00 +/- 14 min), the protein amount increases around 12-14 fold in the *elp4* PseudoU site mutants compared to the wild-type [Figure 29A]. That indicates that a Cbf5-dependent *ELP4* pseudouridylation might have another function. As it is known that mRNA pseudouridylation has a role in mRNA translation [Karijolich and Yu, 2011/ Fernandez et al., 2013], it is worth studying whether a Cbf5-dependent *ELP4* pseudouridylation functions in mRNA translation. Therefore, a sucrose density gradient analysis of translation complexes of the *elp4* PseudoU site mutants can be performed.

Taken together, a previously unknown function of mRNA pseudouridylation was identified. It was shown for the first time that Cbf5-dependent pseudouridylation might have a role in transcription. Furthermore, the PseudoU site in *ELP4* regulates its protein amount. Further research is needed to unravel, whether mRNA pseudouridylation of *ELP4* influences mRNA translation.

3.7 Conclusion

XL-MS analysis of purified and crosslinked TREX could identify several crosslinks between Cbf5 and TREX components [unpublished data, Manuel Koschitza 2015]. During the research for my master thesis, I confirmed that Cbf5 indeed interacts with TREX [Figure 30] [Gruellich, 2017]. In this study, I showed that also Gar1 and Nhp2 of the H/ACA snoRNP interact with Cbf5 in an RNA-independent manner. Therefore, it is highly likely that the whole H/ACA snoRNP interacts with TREX. ChIPs showed that the occupancy of Cbf5 to all tested genes is transcription- and RNA-dependent. However, TREX is not needed for Cbf5 occupancy at transcribed genes. RIPs verified that Cbf5 binds to all mRNAs, independent of whether the mRNA contains a Cbf5-dependent pseudouridylation [Figure 30]. Primer extension after CMC-labeling of nascent and poly(A) mRNA indicates that Cbf5-dependent pseudouridylation occurs post-transcriptionally.
Further analyses to identify the function of the Cbf5/TREX interaction revealed that Cbf5 affects mRNA synthesis in different ways. On the one hand Cbf5 influences RNAPII transcription elongation. On the other hand Cbf5 regulates transcription of genes involved in metabolism and energy derivation [Figure 30]. Cbf5-dependent pseudouridylation-sites in the 5' UTR of *GCN4* were detected, suggesting that these pseudouridylation-sites regulate translation of the transcription activator *GCN4*, which controls transcription of genes involved in amino acid metabolism. Moreover, the Cbf5-dependent pseudouridylation site of *ADE13*, whose protein is involved in the purine nucleotide biosynthetic pathway, influences its transcription. These results verify the ChIPseq data suggesting that Cbf5 regulates transcription of genes involved in metabolism of amino acids and nucleic acids. In addition, FISH experiments showed that Cbf5 has a pseudouridylation-independent function in nuclear mRNA export [Figure 30].
To summarize, novel insights into the recruitment mechanism and function of Cbf5 and Cbf5-dependent pseudouridylation in the mRNA life cycle were gained. It was shown for the first time that Cbf5 functions in transcription and nuclear mRNA export just as the TREX complex. However, it is not known whether these functions of Cbf5 are linked to the Cbf5/TREX interaction. For this purpose, a *cbf5* mutant that does not interact with TREX anymore, has to be generated. This mutant will be helpful for detailed analysis of the Cbf5 function in transcription and nuclear mRNA export.
Furthermore, this study suggests the first function of a mRNA pseudouridylation in the 5' UTR of a gene. It will be interesting to unravel the mechanism of the Cbf5-dependent regulation of *GCN4* translation.

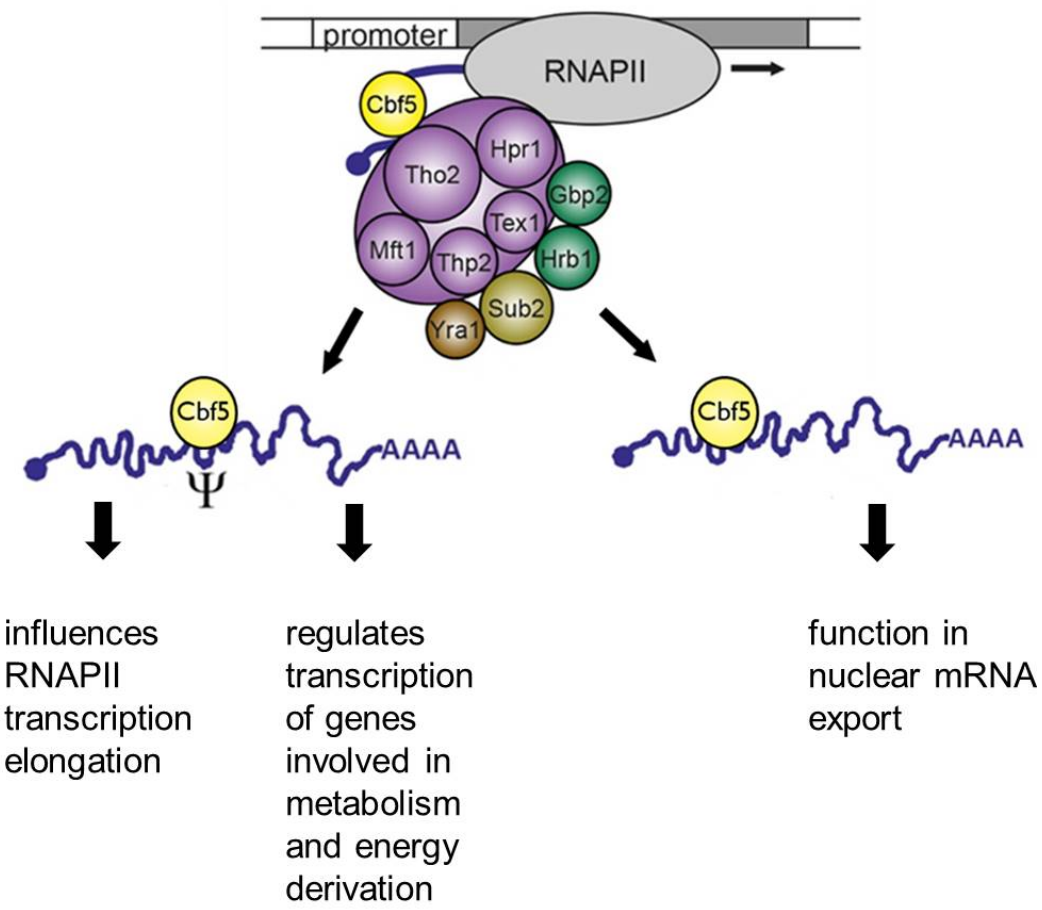

Figure 30: Cbf5 interacts with TREX and functions in mRNA transcription and nuclear mRNA export Cbf5 interacts with TREX in an RNA-independent manner. Cbf5 binds transcription-dependent to mRNAs (shown as blue lines), that contain a Cbf5-dependent pseudouridylation (marked with Ψ) and also to unpseudouridylated mRNAs. Cbf5 influences RNAPII transcription elongation and regulates transcription of genes involved in metabolism and energy derivation. Furthermore, Cbf5 has a pseudouridylation-independent function in nuclear mRNA export [Figure was created by Katja Sträßer and was slighty modified].

4 Material

4.1 Chemicals

The used chemicals and consumables with their respective supplier are listed in Table 4.

Table 3: List of the used chemicals with their respective supplier

Name of the equipment	**Supplier**
2-Propanol	Carl Roth
4-Thiouracil	Sigma-Aldrich
5-Fluoroorotic acid (5-FOA)	Apollo Scientific Ltd
6-Azuracil (6-AU)	Sigma-Aldrich
Acetic acid	VWR Chemicals
Acrylamide (29:1) 40 %	AppliChem
Agar Bacteriology grade	AppliChem
Agarose	Applichem
Ammonium persulfate (APS)	VWR Chemicals
Ampicillin	Applichem
ANTI-FLAG M2 Affinity Gel	Sigma-Aldrich
BactoTM Peptone	BD Biosciences
BactoTM Yeast extract	BD Biosciences
Bicine	Fluka
Bovine serum albumin (BSA)	Carl Roth
Bromophenol blue	Applichem
Calcium chloride ($CaCl_2$)	Fluka
Calmodulin Affinity resin	Agilent Technologies
Chloroform	Merck
Coomassie Brilliant Blue G-250	Applichem
Coomassie Brilliant Blue R-250	Applichem
D-Galactose	Applichem
D-Glucose Monohydrate	Sigma-Aldrich
D-Raffinose	Roth
Dextran sulfate	Sigma-Aldrich
Dimethyl sulfoxide (DMSO)	Grüssing

Disodium hydrogenphosphate (Na_2HPO_4)	Sigma-Aldrich
Dithiothreitol (DTT)	Sigma-Aldrich
dNTPs (dATP, dTTP, dCTP, dGTP)	Thermo Fisher Scientific
Dynabeads^{TM} M-280 Tosylactivated	Invitrogen
Dynabeads^{TM} Protein G	Invitrogen
ECL Solution	Applichem
E.coli t-RNA	Roche diagnostics
Ethanol	Fisher Chemical
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich
Ethyleneglycol-bis(aminoethylether)tetraacetic acid (EGTA)	Merck
Ficoll 400	Carl Roth
FLAG Peptide	Sigma-Aldrich
Formaldehyde	ORG Laborchemie
Formamide	Merck
Gel loading dye, purple (6x)	NEB
Gel loading dye, purple (6x) w/o SDS	NEB
Genetecin (G418)	ThermoFisher (Gibco)
Glycerol	Carl Roth
Glycine	Labochem international
Herring Sperm DNA	Invitrogen
HDGreen^{TM} DNA stain	Intas
HEPES	Carl Roth
Hydrochloric acid (HCl)	Carl Roth
IGEPAL CA-630	Sigma-Aldrich
IgG Sepharose 6 Fast Flow	GE Healthcare
Indol-3-acetid acid (IAA)	Sigma-aldrich
L-Arginine-HCl	Biomol GmbH
L-Aspartic acid	Sigma-Aldrich
Leupeptin (Hemisulfate)	Carl Roth
L-Histidine	Sigma-Aldrich
L-Isoleucine	Sigma-Aldrich
Lithium acetate ($C_2H_3LiO_2$)	Carl Roth

Lithium chloride (LiCl)	Merck
L-Leucine	Sigma-Aldrich
L-Lysine Monohydrochloride	Sigma-Aldrich
L-Methionine	Sigma-Aldrich
L-Phenylalanine	Sigma-Aldrich
L-Threonine	Sigma-Aldrich
L-Tryptophan	Sigma-Aldrich
L-Tyrosine	Sigma-Aldrich
L-Valine	Biomol GmbH
Magnesium chloride ($MgCl_2$)	Merck
Methanol	Merck-Millipore
Octylphenoxypolyethoxyethanol (IGEPAL CA-630; NP-40)	Sigma-Aldrich
Pepstatin A	Applichem GmbH
Phenylmethane sulfonyl fluoride (PMSF)	Carl Roth
Phosphoric acid	Carl Roth
Polyethylene glycol (PEG) 3800/4000	Carl Roth
Polyethylene glycol (PEG) 8000	Fluka
Polysorbate 20 (Tween 20)	Merck
Polyvinylpyrrolidone (PVP)	Sigma-Aldrich
Ponceau S	Serva
Potassium acetat (CH_3CO_2K)	Applichem
Potassium chloride (KCl)	ORG Laborchemie
Potassium hydroxide (KOH)	Merck
Powdered milk, fat free, blotting grade	Carl Roth
Rothi-Mount FluorCare DAPI	Carl Roth
Salmon sperm DNA (SSD)	Applichem
Sodium acetate ($C_2H_3NaO_2$)	Merck
Sodium carbonate (Na_2CO_3)	Merck
Sodium chloride (NaCl)	Merck
Sodium citrate	Carl Roth

Sodium deoxycholate	Sigma-Aldrich
Sodium dihydrogenphosphate (NaH_2PO_4)	Sigma-Aldrich
Sodium dodecyl sulfate (SDS)	Serva
Sodium hydroxide (NaOH)	Merck
Sulfosalicylic acid	Merck
Tetramethylethylendiamin (TEMED)	Carl Roth
Thiolutin	Biomol
Trichloroacetic acid (TCA)	Merck
Tris(hydroxymethyl)aminomethane (Tris)	Applichem
Triton X-100	Applichem
TRIzol reagent	Thermo Fisher
Tryptone BioChemica	Applichem
Uracil	Sigma-Aldrich
Urea	Applichem
Yeast nitrogen base, w/o amino acids	Formedium

4.2 Equipment

The used equipment with their respective supplier is listed in Table 5.

Table 4: List of the used equipment with their respective supplier

Name of the equipment	**Supplier**
Agarose gel electrophoresis system	Southern biological
AM100, micro scale	Mettler-Toledo
Apollo, liquid nitrogen container	Cryotherm
Avanti JXN-26 with JLA-8.1, JA-10	Beckman Coulter
Bioruptor UCL 200	Diagenode
ChemoCam Imager ECL HR 16-3200	Intas
CME microscope	Leica
CO8000 Cell Density Meter	WPA
Delta Vision Ultra fluorescence microscope	Cytiva
Dissection microscope MSM 400	Singer Instruments
Electrophoresis Power Supply Consort E835	Neolab
Electrophoresis Power Supply EPS 3500	Pharmacia Biotech
Eppendorf centrifuge 5424	Eppendorf
Eppendorf centrifuge 5430R	Eppendorf
Eppendorf centrifuge 5415D	Eppendorf
FastPrep-24^{TM} 5G	MP Biomedicals
Freezer/Mill 6870D	Spex SamplePrep
Frezzer Forma 900 Series	Thermo Scientific
Gel iX20 Transilluminator/gel docu	Intas
Hera safe, laminar flow cabinet	Thermo Fisher Scientific
HeraFreeze HFU T Series	Thermo Fisher Scientific
Heidolph shaker duomax 1030	Neolab
IKA KS 4000 ic control, shaking incubator	IKA Labortechnik
IKAMAG RCT, magnetic stirrer	IKA Labortechnik
Incubators	Memmert
Innova 44 shaking incubator	New Brunswick
LED bluelight transilluminator	Nippon genetics
Megafuge 40R, 75003180R	Thermo Scientific

Milli-Q integral water purification system	Merck
Mini-Protean Tetra Electrophoresis Cell	Bio-Rad Laboratories
Multitron Pro / Labotron shaker	Infors HT
Nano Drop ND - 1000 Spectrometer	Thermo Scientific
Optima XPN-80 Ultracentrifuge with 70 Ti	Beckman Coulter
PeqStar XS Thermocycler	Peqlab
Personal Thermocycler	Biometra
Pipetboy acu	IBS Integra Biosciences
Pipettes 20, 100, 200, 1000	Brand
Pipettes 2, 5000	Gilson
PM2000, scale	Mettler-Toledo
pH 211 Microprocessor pH meter	HANNA instruments
Pulverisette	Fritsch
Rotator	neoLab
StepOnePlus Real Time PCR System	Applied Biosystems
Sunrise Microplate Absorbance Reader	Tecan Group
SW22, shaking waterbath	Julabo
T3 Thermocycler	Biometra
Thermomixer compact	Eppendorf
Thermomixer 5436	Eppendorf
Trans-Blot Turbo Transfer System	Bio-Rad Laboratories
Typhoon FLA 9500	GE Healthcare
Vakulan CVC 3000	Vacuubrand
VF2, vortex mixer	IKA Labortechnik
VX-150, autoclave	Systec
WT 12, tumbling shaker	Biometra

4.3 Media and buffers

Buffers and media were prepared with MilliQ water (Merck) and autoclaved at 120°C for 20 min. Heat sensitive solutions and buffers were sterile-filtered (0.22 μm) before use. Specific buffers for experiments are mentioned in the chapter methods. All the other buffers were listed in the Table 6.

Table 5: List of the used buffers and their composition

Buffer	**Composition**
Coomassie stain solution	0.25% (w/v) Coomassie Brilliant Blue R-250, 30% (v/v) ethanol, 10% (v/v) acetic acid
1 x PBS	137 mM NaCl, 2.7 mM KCl, 20 mM NaH_2PO_4, 10 mM Na_2HPO_4 (pH 7.5)
Protease inhibitor (100x)	8 ng/ml Leupeptin, 137 ng/ml Pepstatin A, 17 ng/ml PMSF, 0.33 mg/ml Benzamidine, solved in 100% EtOH (p.a.)
4 x SDS sample buffer	0.2 M TRIS-HCl (pH 6.8), 40% (v/v) glycerol, 8% (w/v) SDS, some mg of bromophenol blue, 0.1 M DTT
SDS-PAGE running buffer	25 mM TRIS, 0.1% (w/v) SDS, 0.19 mM glycine
4 x separating SDS gel buffer	3 M TRIS, 0.4% (w/v) SDS, pH 8.8 (HCl)
Semi-Dry-Blotting-buffer	25 mM TRIS, 192 mM glycine, 20% (w/v) methanol
4 x stacking SDS-gel buffer	0.5 M TRIS, 0.4% (w/v) SDS, pH 6.8 (HCl)
1 x TAE	40 mM TRIS, 2 mM EDTA, pH 8.0; 20 mM acetic acid
1 x TBST	137 mM NaCl, 2.7 mM KCl, 12.5 mM TRIS-HCl 0.1 % (v/v) Tween-20
1 x TBE	100 mM TRIS; 100 mM H_3BO_3; 2,5 mM EDTA
1 x TE	1 mM EDTA, 10 mM TRIS-HCl, pH 8.0

Table 6: List of the used media and their composition

Media	Composition
LB - Bacteria full medium	1% (w/v) tryptone; 0.5% (w/v) yeast extract; 0.5 % (w/v) NaCl; (2 % (w/v) agar for plates)
SDC - Synthetic complete dropout medium	0.67% (w/v) yeast nitrogen base; 0.06% (w/v) complete synthetic mix of aa; drop out as required; 2% (w/v) glucose; (2% (w/v) agar for plates)
SFC - Synthetic complete dropout medium with raffinose	0.67% (w/v) yeast nitrogen base; 0.06% (w/v) complete synthetic mix of aa; drop out as required; 2% (w/v) raffinose; (2% (w/v) agar for plates)
YPD - Yeast full medium	2% (w/v) peptone; 2% (w/v) glucose; 1% (w/v) yeast extract; (2% (w/v) agar for plates)
YPG - Yeast full medium with galactose	2% (w/v) peptone; 2% (w/v) galactose; 1% (w/v) yeast extract; (2% (w/v) agar for plates)

4.4 Radioactivity

$\gamma-^{32}$P ATP (370 MBq/ml, 10 mCi/ml) was obtained from Hartmann Analytik.

4.5 Organism

E.coli strains and cultivation

E. coli strains supplemented with the corresponding antibiotic were grown in LB medium at 37°C and 180 rpm. They were used for molecular cloning.

Table 7: List of the used *E. coli* strains

Strain	Genotype	Reference
$DH5\alpha$	F^- *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG* $\theta 80dlacZ\Delta M15\Delta(lacZYAargG)U169$ $hsdR17_K(r_K^-m_B^+)$ $\lambda-$	Woodcock et al. 1989

S. cerevisiae strains and cultiviation

If not mentioned otherwise, cultures were grown in YPD at 30°C and shaking at 230 rpm.

Table 8: List of the used *S. cerevisiae* strains

Strain	**Genotype**	**Reference**
RS453	MATa;*ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1;can1-100; GAL+*	Straesser and Hurt, 2000
W303	MATa;*ura3-1; trp1-1; his3-11,15; leu2-3,112; ade2-1,can1-100; GAL+*	Thomas and Rothstein, 1989
W303 *CBF5-TAP*	MATa;*CBF5-TAP::TRP1; ura3-1;trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+*	this study
W303 *Δhpr1*	MATa;*HPR1::KanMX4; ura3-1; trp1-1; his3-11,15;leu2-3,112; ade2-1; can1-100; GAL+*	Chavez et al., 2000
W303 *CBF5-TAP Δhpr1*	MATa;*CBF5-TAP::TRP1; HPR1::KanMX4;* ura3-1;trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+	this study
W303 *Δtho2*	MATa; *THO2::KanMX4; ura3-1; trp1-1; his3-11,15;leu2-3,112; ade2-1; can1-100; GAL+*	Piruat and Aguilera, 1998
W303 *CBF5-TAP Δtho2*	MATa; *CBF5-TAP::TRP1; THO2::KanMX4; ura3-1; trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+*	this study
W303 *Δcbp20*	MATa; *CBP20::KanMX4; ura3-1; trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+*	this study
W303 *Δcbp80*	MATa; *CBP80::KanMX4; ura3-1; trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+*	this study
W303 *CBF5-TAP Δcbp20*	MATa; *CBF5-TAP::TRP1; CBP20::KanMX4; ura3-1; trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+*	this study
W303 *CBF5-TAP Δcbp80*	MATa; *CBF5-TAP::TRP1; CBP80::KanMX4; ura3-1; trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+*	this study

BY *CBF5-TAP* Ribozyme active	MATa; *CBF5-TAP::URA3; Ribo-YCT1::KanMX; his3-1; leu2-0; met15-0; ura3-0*	this study
BY *CBF5-TAP* Ribozyme silent	MATa; *CBF5-TAP::URA3; RiboMut-YCT1::KanMX; his3-1; leu2-0; met15-0; ura3-0*	this study
W303 *RPB3-FTpA*	MATa; *RPB3-FtpA::HIS3; ura3-1; trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+*	this study
W303 *CBF5-FTpA*	MATa; *CBF5-FtpA::HIS3; ura3-1; trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+*	this study
W303 *HPR1-FTpA*	MATa; *HPR1-FtpA::HIS3; ura3-1; trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+*	this study
RS453 *HPR1-HA*	MATa; *HPR1-HA::HIS3; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	this study
RS453 *GAR1-HA*	MATa; *GAR1-HA::HIS3; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	this study
RS453 *HPR1-TAP GAR1-HA*	MATa; *HPR1-TAP::TRP1; GAR1-HA::HIS3; ade2-1; his3-11,15; ura3-52;leu2-3,112; trp1-1; can1-100; GAL+*	this study
RS453 *GAR1-TAP HPR1-HA*	MATa; *GAR1-TAP::TRP1; HPR1-HA::HIS3; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; can1-100; GAL+*	this study
W303 *CBF5* shuffle	MATa; *CBF5::KanMX; pRS316-CBF5; ura3-1; trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+*	this study
W303 *Δdst1*	MATa; *dst1::URA3; ura3-1; trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+*	Minocha et al., 2018
W303 *CBF5* shuffle *Δdst1*	MATa; *CBF5::KanMX; pRS316-CBF5; dst1::URA3; ura3-1; trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+*	this study
W303 *osTIR*	MATa; *OSTIR::URA3; ura3-1; trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+*	this study

W303 *osTIR* *CBF5-AID*	MATa; *OSTIR::URA3; CBF5-AID-9myc::HYG; ura3-1; trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+*	this study
W303 *CBF5* shuffle *GCN4-TAP*	MATa; *CBF5::KanMX; pRS316-CBF5; GCN4-TAP::TRP1; ura3-1; trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+*	this study
W303 *osTIR* *CBF5-AID* *GCN4-TAP*	MATa; *OSTIR::URA3; CBF5-AID-9myc::HYG; GCN4-TAP::TRP1; ura3-1; trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+*	this study
W303 *osTIR* *CBF5-AID* *HPR1-TAP*	MATa; *OSTIR::URA3; CBF5-AID-9myc::HYG; HPR1-TAP::TRP1; ura3-1; trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+*	this study
W303 *Δelp4*	MATa; *ELP4::KanMX4; ura3-1; trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+*	this study
W303 *ADE13* shuffle	MATa; *ADE13::KanMX4; pRS316-ADE13 ura3-1; trp1-1; his3-11,15; leu2-3,112; ade2-1; can1-100; GAL+*	this study

4.6 Oligonucleotides

Used oligonucleotides are listed in the Tables 10 to 16 and were synthesized by Biolegio (Njimegen, Netherlands).

Table 9: List of the oligonucleotide sequences used for genomic tagging

Name	**Sequence 5' - 3'**
CBF5-TAP tag frw	TGAAGACGGTGATTCTGAGGAAAAGAAATCTAAGAAAT CTAAGAAATCCATGGAAAAGAGAGG
CBF5-TAP tag rev	AGAAAGCTGTTAAATATATAGGATGAGATGGAGTGATG AGAATCATACGACTCACTATAGGG
CBF5-FtpA tag frw	TCTGAAGACGGTGATTCTGAGGAAAAGAAATCTAAGAAA TCTAAGAAAGATCACGACGGTGACTAC
CBF5-FtpA tag rev	AAAAGAAAGCTGTTAAATATATAGGATGAGATGGAGTGA TGAGAATCAATCGATGAATTCGAGCTCG
HPR1-HA tag frw	GCTACTTCGAACATTTCTAATGGTTCATCTACCCAAGATA GAAACGTACGCTGCAGGTCGAC
HPR1-HA tag rev	ATGAATTTCTTATCAGTTTAAAATTTCTATTAAGAGGATA ATTTAATCGATGAATTCGAGCTCG
RPB3-FtpA tag frw	GCATCTCAAATGGGTAATACTGGATCAGGAGGGTATGATA ATGCTTGGGATCACGACGGTGACTAC
RPB3-FtpA tag rev	GGTTCGTTCACTTGTTTTTTTTCCTCTATTACGCCCACTT GAGAACTACATCGATGAATTCGAGCTCG
GAR1-TAP tag frw	GGATCTCGTGGCGGATCTCGTGGTGGTTTCAGAGGAGGT AGAAGAATGGAAAAGAGAAG
GAR1-TAP tag rev	AAGATAAAGGATGGTTTTTTTTTTTATTTTTTTCTAGCTA GATTATCATACGACTCACTATAGGG
GAR1-HA tag frw	GGATCTCGTGGCGGATCTCGTGGTGGTTTCAGAGGAGGT AGAAGACGTACGCTGCAGGTCGAC
GAR1-HA tag rev	AAGATAAAGGATGGTTTTTTTTTTTATTTTTTTCTAGCTA GATTAATCGATGAATTCGAGCTCG
CBF5-AID tag frw	GAAGACGGTGATTCTGAGGAAAAGAAATCTAAGAAATCT AAGAAACCTAAAGATCCAGCCAAACC
CBF5-AID tag rev	AGAAAGCTGTTAAATATATAGGATGAGATGGAGTGATGAG AATCACAGTATAGCGACCAGCATTC

GCN4-TAP tag frw	TTGGAAAATGAGGTTGCCAGATTAAAGAAATTAGTTGGC GAACGCATGGAAAAGAGAAG
GCN4-TAP tag rev	CACGAGAATGAAATAAAAAATATAAAATAAAAGGTAAATG AAATCATCATACGACTCACTATAGGG

Table 10: List of the oligonucleotide sequences used for deletion of genes

Name	**Sequence 5' - 3'**
CBP20 500 bp 5' UTR frw	GATGACACAAATGATGAGAAATAGTC
CBP20 500 bp 5' UTR rev	CGCAAAATATTTAGTGATTACTATACCC
CBP80 500 bp 5' UTR frw	GCACCATCTTTTGACGATTCTTCGAG
CBP80 500 bp 5' UTR rev	GATATGCATTACTTTCCCGAACCGC
CBF5 300 bp 5' UTR frw	GAGGCAGCTGCATTGTAAAC
CBF5 300 bp 5' UTR rev	GATTGCAGCACAAGATTCAGG
DST1 300 bp 5' UTR frw	GGATGGTGGGTTGTTCTGTACC
DST1 300 bp 5' UTR rev	GCCTCCTTAAACTCCTGGTGTG
ELP4 frw	GATAGTTTAGAGATGAGGTGCAGAGAACAACATTG TATAACAAATTCGGCTCCCAAATATCGCATGTACCA GCTTGCCTCGTCCCCG
ELP4 rev	AAAAAAATAATACATTCACCACGCTTGTTAAAAAG CATGCCGTATATTTCCCATAAATTGAACCATATTCC GGCGTTAGTATCGAATCGACAG
ADE13 frw	CCAAGGACCAAAGCTTTTCCAGTAAGTTGCTACAA CTAAAGTGAAATAAACTATCCGAATTATATAAAAGA GCTTGCCTCGTCCCCG
ADE13 rev	GAATTGTTATTTTTTTCATTAGTAACATATAATAGG GTAATAAACATATTATTTGTACATGTAAAGGCTTCG GCGTTAGTATCGAATCGACAG

Table 11: List of the oligonucleotide sequences used for Gibson assembly

Name	**Sequence 5' - 3'**
pRS314-*CBF5* vector frw	CGCCGATTATTGACTAGTGGATCCCCCGGG
pRS314-*CBF5* vector rev	ACTACATCTGGAGAGCTCCAATTCGCCCTATAGT
pRS314-*CBF5* insert frw	CGAATTGGAGCTCTCCAGATGTAGTTAAGTTCGGC CAA
pRS314-*CBF5* insert rev	GGGATCCACTAGTCAATAATCGGCGATATTAGCTTA AAATCAAATGAGAA
pRS315-*CBF5* vector frw	CGCCGATTATTGCGATACCGTCGACCTCGAG
pRS315-*CBF5* vector rev	ACTACATCTGGAGGGGGGATCCACTAGTTCTAGAG CG
pRS315-*CBF5* insert frw	TAGTGGATCCCCCCTCCAGATGTAGTTAAGTTCGG CCAAA
pRS315-*CBF5* insert rev	GTCGACGGTATCGCAATAATCGGCGATATTAGCTT AAAATCAAATGAGCCAAA
pRS316-*CBF5* vector frw	CGAATTGGAGCTCTCCAGATGTAGTTAAGTTCGGC CAA
pRS316-*CBF5* vector rev	ACTACATCTGGAGTAGTTCTAGAGCGGCCGCCAC
pRS316-*CBF5* insert frw	GCTCTAGAACTACTCCAGATGTAGTTAAGTTCGGC CAAATATTG
pRS316-*CBF5* insert rev	TAAGCTTGATATCCAATAATCGGCGATATTAGCTTA AAATCAAATGAGAA
pointmutation *cbf5-D95N* frw	CTCTGGTACATTGAATCCAAAAGTTACAGG
pointmutation *cbf5-D95N* rev	CCTGTAACTTTTGGATTCAATGTACCAGAG
pRS314-*ELP4-TAP* insert frw	CTCTAGAACTAGTGGATCCCATGTCATTTCGTAAA AGAGGTGAAATACTGA
pRS314-*ELP4-TAP* insert rev	TTCTTTTTCCATCTTCTCTTATAGTCTAAAGATAT CTTGGTCTTCTTTGCAGG
pRS314-*ELP4-TAP* vector frw	CCAAGATATCTTTAGACTATAAGAGAAGATGGAAA AAGAATTTCATAGCCG
pRS314-*ELP4-TAP* vector rev	CCTCTTTTACGAAATGACATGGGATCCACTAGTTC TAGAGCGG

pRS314-*ELP4* insert rev	CCGGAATTAGCTTGGCTCGACTCAATAGTCTAAAG ATATCTTGGTCTTCTTTGC
pRS314-*ELP4* vector frw	CCAAGATATCTTTAGACTATTGAGTCGAGCCAAGC TAATTCCGG
pointmutation *elp4-F126Y* frw	ACTCTATTCTCGGTAAACTGTATGCCGCACAG
pointmutation *elp4-F126Y* rev	ACGATACCCTGTGCGGCATACAGTTTACCGAG
pointmutation *elp4-F126W* frw	ACTCTATTCTCGGTAAACTGTGGGCCGCACAG
pointmutation *elp4-F126W* rev	ACGATACCCTGTGCGGCACCCAGTTTACCGAG
GAL1010::ELP4 vector frw	CCGTACTTCAATATAATGTCATTTCGTAAAAGAGG TGAAA
GAL1010::ELP4 vector rev	TTTAACGTCAAGGAGTTCACTGGCCGTCGTTTTAC A
GAL1010::ELP4 insert frw	ACGACGGCCAGTGAACTCCTTGACGTTAAAGTAT AGAGG
GAL1010::ELP4 insert rev	TTTACGAAATGACATTATATTGAAGTACGGATTAG AAGCCGC
pRS314-*ADE13-TAP* insert frw	TAGTGGATCCCATGCCTGACTATGACAATTACACTA CG
pRS314-*ADE13-TAP* insert rev	CCATCTTCTCTTAACATTTAACTTGACTTGTTCATC GTTTAGGTAC
pRS314-*ADE13-TAP* vector frw	TTAAATGTTTAGAAGAGAAGATGGAAAAAGAATTT CATAGCC
pRS314-*ADE13-TAP* vector rev	TAGTCAGGCATGGGATCCACTAGTTCTAGAGCGG
pRS314-*ADE13* insert rev	CCGGAATTAGCTTGGCTCGACTCAAACATTTAACT TGACTTGTTCATCGTTTAGG
pRS314-*ADE13* vector frw	AAGTTAAATGTTTGAGTCGAGCCAAGCTAATTCCG G
pRS316-*ADE13* insert frw	CTCCAAATTGATCAAGGCAGCTTTTGTTCCCTTTA GTGAGG

pRS316-*ADE13* insert rev	TAAAGGGAACAAAAGCTGCCTTGATCAATTTGGAGTCACCA
pRS316-*ADE13* vector frw	ACTATAGGGCGAATTGCTAAGCCTGCCACTGAATAGG
pRS316-*ADE13* vector rev	TTCAGTGGCAGGCTTAGCAATTCGCCCTATAGTGAGTCG
pointmutation *ade13-S350S* frw	GATATTCTATTAAGTACTTTGTTGAACATCTCATCCGG
pointmutation *ade13-S350S* rev	TCAACAAAGTACTTAATAGAATATCTGCGGTTAAAAATGC
GAL1010::ADE13 vector frw	CCGTACTTCAATATAATGCCTGACTATGACAATTACAC
GAL1010::ADE13 insert rev	GTCATAGTCAGGCATTATATTGAAGTACGGATTAGAAGCCGC

Table 12: List of the oligonucleotide sequences used for colony PCR

Name	**Sequence 5' - 3'**
KANMX4 rev	GCGGCCATCAAAATGTATGG
KANMX4 frw	ATTGTTGATGCGCTGGCAG
CBP20 5' UTR frw	GGATTCCTAAATCCTTGAGGAGAACTT
CBP80 5' UTR frw	GCTTAAAGGACGTTTACAGTAGATGAAT
osTIR frw	GGCGCATGCAACTTCTTTTC
osTIR rev	ACTGAGAGTGCACCACGC
CBF5 5' UTR 500 bp frw	TCAGCAAGGAAAGGAGACG
CBF5 300 bp 3' UTR frw	GTAGTTCAGCAATTGGCCCA
pRS314/pRS315/pRS316 rev	TGCTGGCCTTTTGCTCAC
DST1 3' UTR rev	CGAACGAGTGACTCTGGAACTAT
ELP4 5' UTR frw	GGGCGATAATCAAGGAGAATATCA
ELP4 3' UTR rev	CTTTTTCTACCTTCCAGCTGCC
ADE13 5' UTR frw	GATGATGGCTTAGAACCGGATG
ELP4 frw	GGTGTGAGGAGTCTTGTCAAG
ADE13 frw	GTTTAGTGACGCCGTTCAA
ADE13 3' frw	GCCTATCTGGGAAGAATTAGATTC
GAL10 rev	ATTAGCTCTACCACAGTGTGTG

Table 13: List of the oligonucleotide sequences used for qPCR

Name	**Sequence 5' - 3'**
YER frw	TGCGTACAAAAAGTGTCAAGAGATT
YER rev	ATGCGCAAGAAGGTGCCTAT
PMA1 5' frw	CGATGACGCTGCATCTGAA
PMA1 5' rev	CCGTGATTAGATTGTAGTTCTTCGATT
PMA1 M frw	AAATCTTGGGTGTTATGCCATGT
PMA1 M rev	CCAAGTGTCTAGCTTCGCTAACAG
PMA1 3' frw	CGTCTTCGCTGTCGACATCA
PMA1 3' rev	TTTTCAGACCACCAACCGAAT
PGK1 5' frw	CCAACCATCAAGTACGTTT
PGK1 5' rev	CCCAAGTGAGAAGCCAAGACA
PGK1 M frw	AAGGCTTTGGAGAACCCAAC
PGK1 M rev	CGACCTTGTCCAACAAGTTGTC
PGK1 3' frw	TGACAAGATCTCCCATGTCTCTACTG
PGK1 3' rev	TGGCAATTCCTTACCTTCCAA
CCW12 5' frw	AGTGTCGCTTCTATCGCCGC
CCW12 5' rev	TTGGCTGACAGTAGCAGTGG
CCW12 M frw	CTGTCTCCCCAGCTTTGGTT
CCW12 M rev	GGCACCAGGTGGTGTATTGA
CCW12 3' frw	TGAAGCTCCAAAGAACACCACC
CCW12 3' rev	AGCAGCAAGCACCAGTGTAAG
GAL1 5' frw	AGCTAAAGGTAAAACCGAGAAGTCAC
GAL1 5' rev	AATGGATGGATTTTTCTCGTTCA
GAL1 M frw	ATCATGCTCTATACGTTGAGTTC
GAL1 M rev	TGCGGAAATTTAAACGGAGTAGC
GAL1 3' frw	GGTGGTTGTACTGTTCACTTGGTT
GAL1 3' rev	TCATTGGCAAGGGCTTCTTT
KAR2 URI frw	GAAACCGCCATTGCTGAAGA
KAR2 URI rev	GAAGTAATTGGATAAGCGACCTTG
ADE13 3' frw	GAGGGATGAATTTTTCAAGCCT
ADE13 3' rev	CTCAACTTGTTGTGGAGCTC
ADE13 URI frw	CCAAGTGCATTTTTAACCGC
ADE13 URI rev	CACAACTAAACCGGATGAGAT
ELP4 3' frw	GGGGAAGATTCAACATGGGC

ELP4 3' rev	CTGCCATTCTTAAATGCCCATTC
ELP4 URI frw	CGACAGAATTTCACTCTATTCTCG
ELP4 URI rev	CTGAAATTCTATTATGAACGATACCC
RPS16B 3' frw	GGTGGTGGTCATGTTTCTCAAG
RPS16B 3' rev	GGTGGTAAGCGACTAAACCC
RPS16B URI frw	AACATCGATATCAGAGTTAGAGTCACTG
RPS16B URI rev	CAATAGCTTGTCTGATGGCG
RBP11 3' frw	GCTGAGAATACAGACCACCGAGGG
RBP11 3' rev	GTCTTCAGGGCGCCAAGCTTG
RBP11 URI frw	ACCTGATTCGTGCAGAACTTC
RBP11 URI rev	GGGATGCTCAACCTTGTAAGC
GFP 3' (P1) frw	CCAATTGGTGATGGTCCAGTCT
GFP 3' (P1) rev	GTCTTTTCGTTTGGATCTTTGGAT
YCT1 5' (+100 bp) (P2) frw	TCGCTGATGTTACGCTAGCATT
YCT1 5' (+100 bp) (P2) rev	CAGGCGTGATTTCTGGAACTG
25S rRNA frw	GAATAAGGCGTCCTTGTGGC
25S rRNA rev	CAAGCACCCAAGGCCTTT
ATR1 frw	GAGGTGCTCAGTACCTTGGG
ATR1 rev	CCCTGCGGCCTTTTATGAAGC
MUP1 frw	GGGCCCAATTATTGCCAAGATGG
MUP1 rev	CTCCAGAACCCATCTTCACCAAGCAC
MET17 frw	GTTGGTGATGCCAAGACCTTAGTC
MET17 rev	TCCTTGGTAACACCAGATGCCAAC
HIS4 frw	GTGCCAACACTGCAACCTTCCAAAAG
HIS4 rev	CCCTCCTTCTTGGCAACGCAC
CYS3 frw	GTTCACATTGGCCGAATCCC
CYS3 rev	GGCCTCCTTTGGGATACCAC
ADE5,7 frw	CTTCAACATGGGTGTTGGTATGG
ADE5,7 rev	CTTGCAAAGAACCAAGCTCCC
URA7 frw	GGGTTTAGTTGCTGCTTCCGCT
URA7 rev	CGCCGGCCTCAAGATCGTAC
GUA1 frw	CATGACTGCTGACTGGTTCCC
GUA1 rev	TGACCCTAGCAACCCCATCGAC

OLE1 frw	CTGGTATTGTTCACGACGTATCTGG
OLE1 rev	GTAGACACCACCACTGAAAGCC
FAS1 frw	CGGAAAGTACATTCCAAACTTGACTGC
FAS1 rev	CCTTGATAGGTTCGGAGCCAG
RPA190 frw	CATGTCAATTCTTGACCAAGGCTG
RPA190 rev	CCACCACAATTCTAGCGGAAGG
SDA1 frw	CATGCGTTCTACTACCAACAGAG
SDA1 rev	GGTGATGTGTGCACGTAACAC
UTP22 frw	GCAACGTTAAACCAGTGGATGATG
UTP22 rev	CCATGTCATTCCCGAATGCAG
YLR454w 194/+35 (promoter) frw	CTGGGGTAATTAATCAGCGAAGCGATG
YLR454w 194/+35 (promoter) rev	CACTTGTACAGTAGAACATTAATCGGAAAC
YLR454w +392/+605 (0.5 kb) frw	CGCAATTAGTCAACAACGATATCACGATTG
YLR454w +392/+605 (0.5 kb) rev	CCTACTTGAAGTCCATCCTTCAGAGG
YLR454w +945/+1147 (1 kb) frw	CAATACCAACAGGTTCAGAAATGAGATGC
YLR454w +945/+1147 (1 kb) rev	GAGAGAACAAATTGGTTTCGCCAAATATCG
YLR454w +1986/+2199 (2 kb) frw	CATATCATCCACCCTAGGTGCTAGGTCGG
YLR454w +1986/+2199 (2 kb) rev	GAGCTGACCAGACCTAACCATAGTAGCGTG
YLR454w +4069/+4268 (4 kb) frw	AGATATTACTCGTTGTTCGTGCCCAG
YLR454w +4069/+4268 (4 kb) rev	TCCCAAAACCCTAGTTTAACAGAAGGATC
YLR454w +5904/+6074 (6 kb) frw	CGTACTGTTGAAATGGAACGAGGACGC
YLR454w +5904/+6074 (6 kb) rev	ATCGCTTCCATACTCGTTGTATCATCAGTC

YLR454w +7701/+7850 (8 kb) frw	GAGGGTCACA GATCTATTAC TTGCCC
YLR454w +7701/+7850 (8 kb) rev	GTTGTGAGTTGCTTCAGTGGTGAAGTG
GCN4 3' frw	CGTTCTCGTGCGAGAAAGTTG
GCN4 3' rev	CTGGCAACCTCATTTTCCAAGTG

Table 14: List of the oligonucleotide sequences used for reverse transcription

Name	**Sequence 5' - 3'**
oligo dT(16)	AAAAAAAAAAAAAAAA
ADE13 RT nascent	GGAAGAAGATGAAAAATATATCCCAG
KAR2 RT nascent	CTGAGTACCTTAACCCCAG
RPS16B RT nascent	GGGAATGATTCTAAATTAAC

Table 15: List of the used oligonucleotide sequences used for primer extension

Name	**Sequence 5' - 3'**
ADE13	GCCATAAAAGGTAGTTCACCC
KAR2	ATAATCAGCGGCACCAG
RPS16B	TAGCTTGTCTGATGGCG
RBP11	TCAACCTTGTAAGCGGC
ELP4	GACAATGACATGAGTGTCGC
MRPS12	GCAGATACCACGTTCCC
YPL225W 5' UTR P1	ACGTAACTTGGAACCGG
GCN4 5' UTR P2	GGCGAGTAAACCTGGATAATTTGACAG
GCN4 5' UTR P3	CTCTTCAAAAAACTGACAGTTTTCAAAAAAAGTAAAGG
GCN4 5' UTR P4	CAATAAATTTAACACATAATTCTCTTAATAATTTTCTAAT AATAATCTACTTTAAAAAC
GCN4 5' UTR P5	GCGAAGTAGATGAGTGAGCTGTG
GCN4 5' UTR P6	GGGGAGAGTAACCTGTGTTGTGAG
GCN4 5' UTR P7	GCGAGCCGAAACTTCTTTAAAGAAGAC

4.7 Plasmids

A list with plasmids used in this study can be found in the table below.

Table 16: List of the used plasmids

Name	Description	Reference
pYM15	for C - terminal HA - tagging with *HIS3* marker	Euroscarf
pBS1479	for C - terminal TAP - tagging with *TRP1* marker	Puig et al., 2001
pBS1539	for C - terminal TAP - tagging with *URA3* marker	Puig et al., 2001
pFA6a-FTpA-HIS3	for C - terminal FtpA - tagging with *HIS3* marker	Kressler et al., 2012
pRS314	pBlueScript based yeast centromere vector with TRP1 marker	Sikorski and Hieter, 1989
pRS314-*CBF5*	ORF + 500 bp of 5' and 300 bp of 3' UTR of *CBF5* was cloned into pRS314	this study
pRS314-*cbf5-D95N*	insertion of pointmutation *cbf5-D95N*	this study
pRS314-*ADE13*	ORF + 500 bp of 5' and 300 bp of 3' UTR of *ADE13* was cloned into pRS314	this study
pRS314-*GAL10::ADE13*	*GAL10* promotor was cloned upstream of *ADE13* into pRS314-*ADE13*	this study
pRS314-*ade13-S350S*	insertion of pointmutation *ade13-S350S*	this study
pRS314-*GAL10::ade13-S350S*	*GAL10* promotor was cloned upstream of *ade13-S350S* into pRS314-*ade13-S350S*	this study
pRS314-*ELP4*	ORF + 500 bp of 5' and 300 bp of 3' UTR of *ELP4* was cloned into pRS314	this study
pRS314-*elp4-F126Y*	insertion of pointmutation *elp4-F126Y*	this study
pRS314-*GAL10::elp4-F126Y*	*GAL10* promotor was cloned upstream of *elp4-F126Y* into pRS314-*elp4-F126Y*	this study
pRS314-*elp4-F126W*	insertion of pointmutation *elp4-F126W*	this study
pRS314-*GAL10::elp4-F126W*	*GAL10* promotor was cloned upstream of *elp4-F126W* into pRS314-*elp4-F126W*	this study

pRS314-TAP	pBlueScript based yeast centromere vector with TRP1 marker and TAP-tag	Roether and Straesser, 2007
pRS314-*ADE13-TAP*	ORF + 500 bp of 5' *ADE13* was cloned into pRS314-TAP	this study
pRS314-*ADE13-S350S-TAP*	insertion of pointmutation *ade13-S350S*	this study
pRS314-*ELP4-TAP*	ORF + 500 bp of 5' *ELP4* was cloned into pRS314-TAP	this study
pRS314-*ELP4-F126Y-TAP*	insertion of pointmutation *ELP4-F126Y-TAP*	this study
pRS314-*ELP4-F126W-TAP*	insertion of pointmutation *ELP4-F126W-TAP*	this study
pRS315	pBlueScript based yeast centromere vector with LEU2 marker	Sikorski and Hieter, 1989
pRS315-*CBF5*	ORF + 500 bp of 5' and 300 bp of 3' UTR of CBF5 was cloned into pRS315	this study
pRS315-*cbf5-D95N*	insertion of pointmutation *cbf5-D95N*	this study
pRS316	pBlueScript based yeast centromere vector with URA3 marker	Sikorski and Hieter, 1989
pRS316-*CBF5*	ORF + 500 bp of 5' and 300 bp of 3' UTR of CBF5 was cloned into pRS316	this study
pRS316-*ADE13*	ORF + 500 bp of 5' and 300 bp of 3' UTR of ADE13 was cloned into pRS316	this study

4.8 Antibodies

Used primary and secondary antibodies are listed in the tables below.

Table 17: List of the used primary antibodies

Name	**Source**	**Dilution**	**Supplier**
anti-Cbf5	rabbit, polyclonal	WB: 1:1.000	Dragon lab, Quebec, Kanada
anti-Cbp80	rabbit, polyclonal	WB: 1:20.000	Görlich lab, Göttingen, Germany
anti-HA	rat, monoclonal	WB: 1:1.000	Roche, Basel, Switzerland
anti-HA	rabbit, monoclonal	WB: 1:1.000	R&D Systems, Minneapolis, USA
anti-Npl3	rabbit, polyclonal	WB: 1:5.000	Guthrie lab, California, San Francisco
anti-myc	rabbit, polyclonal	WB: 1:5.000	Upstate Cell Signaling Solution, New York, USA
Peroxidase-anti-peroxidase (PAP)	rabbit, polyclonal	WB: 1:5.000	Sigma, Taufkirchen, Germany
anti-Sub2	rabbit, polyclonal	WB: 1:10.000	Straesser lab, Gießen, Germany
anti-Yra1	rabbit, polyclonal	WB: 1:2.000	Straesser lab, Gießen, Germany
anti-Rpl6	rabbit, polyclonal	WB: 1:5.000	G. Dieci lab, Parma, Italy
anti-Rps8	rabbit, polyclonal	WB: 1:5.000	G. Dieci lab, Parma, Italy
8WG16	rabbit, monoclonal	IP: 1:50	BioLegend, San Diego, Kalifornia

Table 18: List of the used secondary antibodies

Name	**Source**	**Dilution**	**Supplier**
anti-rabbit-HRPO	goat, monoclonal	1:3.000	Bio-Rad, Hercules, Kalifornia
anti-rat- HRPO	goat, monoclonal	1:5.000	Sigma, Taufkirchen, Germany

5 Methods

5.1 Cultivation of *S. cerevisiae*

Yeast strains were cultivated in either full media (YPD) or synthetic complete (SC) medium, lacking one specific amino acid or nucleotide on a solid agar plate or in a liquid culture. Unless otherwise specified in a protcol, yeast cultures were cultivated at 30°C and liquid cultures were shaken at 250 rpm. Cell density of the yeast culture was determined with a spectrophotometer at a wavelength of 600 nm. One OD at 600 nm corresponds to app. 2.5×10^7 cells.

5.1.1 Dot spot

To test for potential growth defects, mutant strains were analyzed using dot spots. Therefore, a half loop of freshly grown cells was mixed in 1 mL of H_2O. The cell density was adjusted using a spectrometer. The solutions were diluted 4 times, each 10 fold and 5 - 10 μl of these dilutions were spotted on the respective media plates. The air dried plates were incubated at 16°C, 30°C, 37°C and 20°C for up to 5 days.

5.1.2 Growth curve

An overnight culture was diluted to an OD_{600} of 0.2 and was grown for 2 hours before the measurement of the growth curve started. The OD_{600} was measured every 30/60 minutes. To keep cells in mid-log growth phase, cells were always diluted before reaching an OD_{600} of 0.8.

5.2 Molecular cloning methods

5.2.1 Preparation of genomic DNA of *S. cerevisiae*

10 ml of an overnight culture with an OD_{600} nm > 1 were centrifuged (3 min, 3600 rpm). The pellet was washed with 10 ml H_2O. Afterwards, the cells were resuspended in 500 μl H_2O and 200 μl TSNTE buffer, (composition see table) 300 μl glass beads and 300 μl phenol:chloroform:isoamylalcohol (25:24:1) were added. The mixture was vortexed for 5 min. After a centrifugation step for 5 min at 13,000 rpm, the upper phase was transferred to a new reaction tube. An equal volume of chloroform was added, vortexed and centrifuged for 5 min at 13,000 rpm. The upper phase was transferred to a new reaction tube. Genomic DNA was precipitated and resuspended in 100 μl H_2O (see the section "DNA precipitation").

Table 19: Composition of the TSNTE-buffer

TSNTE-buffer
2% Triton X - 100
2% SDS
100 mM NaCl
10 mM TRIS - HCl, pH 8.0
1 mM EDTA

5.2.2 Polymerase chain reaction

All PCRs for cloning and genome integration were carried out using Phusion DNA polymerase (New England Biolabs (NEB) or Q5 High Fidelty DNA Polymerase (NEB). A typical PCR reaction per tube is shown in the table below. The annealing temperature depends on the melting temperature of the primers and the time of the elongation depends on the length of the PCR product.

Table 20: Composition of PCR

Component	**Amount**
each primer [10 μM]	2.5 μl
template [10 - 100 ng]	1 μl
dNTPs [2.5 mM]	4 μl
5 $\times$ Phusion or Q5 buffer	10 μl
DMSO	1.5 μl
Polymerase	1 μl
H_2O	27.5 μl
total volume	50 μl

For verification of gene disruptions, colony PCR was performed in 20 μl using Taq DNA polymerase (prepared by Heike Büngen) with one primer annealing in the sequence of the auxotrophy marker used for disruption and one primer annealing in the flanking genomic region. A small amount of freshly grown yeast cells were added to all components of the colony PCR excluding the polymerase. The PCR reaction tube was placed to a cold PCR cycler and the cells were lysed 10 min at 95°C. Afterwards, the Taq polmerase was added to the PCR reaction.
A typical colony PCR reaction per tube is shown in the table below.

Table 21: Composition of colony PCR

Component	**Amount**
each primer [100 μM]	0.2 μl
dNTPs [2.5 mM]	2 μl
10 × Taq puffer	2 μl
Taq polymerase	2 μl
H_2O	13.6 μl
total volume	20 μl

5.2.3 Agarose gel elektrophorese

1% agarose was solved in 1x TAE to prepare an agarose gel. HD Green Plus DNA Stain (Intas) was added to the agarose and the agarose gel was prepared. The gel was covered with 1 × TAE. PCR samples and the 1 kb DNA ladder (NEB) or 1 kb plus DNA ladder (NEB) were mixed with 6 × DNA Loading Dye Purple (NEB) and loaded on the gel. A voltage of 200 V was applied for approximately 30 min to separate DNA fragments according to their size. DNA bands were visualized by UV-illumination.

5.2.4 DNA/RNA precipitation

DNA/RNA was precipitated by addition of 2.5 volumes of 100% cold ethanol and 1/10 volume of 3 M sodiumacetat ($C_2H_3NaO_2$) (pH 5.2). The mixture was incubated for at least 1 h at -20°C. After centrifugation (14,000 rpm 4°C, 20 min) the pellet was washed with 70% ethanol, dried and resuspended in an appropriate volume of water.

5.2.5 Transformation of *S. cerevisiae*

For genomic integration of DNA or plasmid transformation, an overnight yeast culture was diluted to an OD_{600} of 0.2 in 50 ml YPD. The culture was harvested at an OD_{600} of 0.6 - 0.8 (3600 rpm, 3 min). The pellet was washed once with 10 ml H_2O and 0.5 ml transformation solution I (composition see table). The yeast cells were resuspended in 250 μl transformation solution I. For each transformation 50 μl of the yeast cell suspension was mixed with 300 μl transformation solution II (composition see table), 500ng plasmid DNA or 300 μl precipitated PCR product for genomic integration and 5 μl single stranded carrier DNA (2mg/ml). The transformation reaction was incubated for 30 min at room temperature on a turning wheel. 35 μl DMSO was added. Afterwards, the transformation was heat-shocked for 10 min at 42°C, followed by 3 min incubation on ice. 1 ml H_2O was added to dilute the solution. The

yeast cells were pelleted 3 min at 3600 rpm, solved in 200 μl H_2O and spread on a selective agar. For genome integrations, the pelleted cells were solved in 1 mL YPD and incubated for 1 - 3 h at 30°C on a turning wheel prior to spreading on selective media plates. Cells were typically incubated for 2 - 4 days at 30°C.

Table 22: Composition of the transformation solution I & II

transformation solution I	**transformation solution II**
1 × TE	1 × TE
100 mM $C_2H_3LiO_2$	100 mM $C_2H_3LiO_2$
	40% PEG - 4000

5.2.6 Transformation of *E. coli*

A 60 μl aliquot of competent *E. coli* DH5α cells were thawed on ice and added to 1 μl of the ligation product or 2 μl of the Gibson reaction. The chemically competent cells with DNA were kept on ice for 15 min, heat-shocked at 42°C for 1 min and after another minute on ice resuspended in 450 μl LB media. The transformants were incubated on the shaker at 37°C for 30 min to form resistance. Finally, cells were spread on LB plates containing respective antibiotics and incubated at 37°C overnight.

5.2.7 Plasmid-minipreperation of *E. coli*

Single colonies were picked, transferred to 3 mL LB media containing respective antibiotics and incubated on the shaker at 37°C overnight. The cells were harvested to extract the plasmid DNA from the cells. Therefore, the NucleoSpin Plasmid (NoLid)-kit (Macherey-Nagel) according to the manufacturer's instructions was used.

5.2.8 Restriction and ligation

The restriction digest was performed with the appropriate restriction enzymes (NEB) and the corresponding restriction buffer. 0.25 μl of each restriction enzyme was added to a restriction digestion of a volume of 50 μl. The restriction digestion was incubated at the appropriate temperature for 1 h. The digestion was stopped by incubation for 10 min at 80°C. An agarose gelelektrophorese in a 1% agarose gel was performed to verify the success of the restriction digestion (see section "Agarose gel elektrophorese"). If the digestion was successful, another agarose gelelektrophorese of the whole restriction digestion was made. The resulting bands of the PCR products were excised under UV light. The DNA was

purified using the PCR NucleoSpin Gel and PCR Clean-up-kit (Macherey-Nagel) according to the manufacturer's instruction. Restricted plasmids and inserts were ligated in a volume ratio of 1:3 using T4 DNA ligase (NEB) at 16°C overnight. 1 μl of the reaction was used for a *E. coli* transformation

5.2.9 Gibson assembly

Gibson assembly was performed essentially as described in Gibson et al., 2009. The DNA fragments of the vector and the inserts were PCR amplified. PCR primers were used to get PCR fragments with an overlapping region of 20 - 25 bp to each other. The PCR products were purified using the PCR NucleoSpin Gel and PCR Clean-up-kit (Macherey-Nagel) according to the manufacturer's instruction. 100 ng vector and a volume raio of 1:1 of vector to insert was mixed in 5 μl and added to the Gibson assembly master mix (composition see table). The Gibson reaction was incubated for 1 h at 50°C. Afterwards, a DpnI digest was performed for 3 h at 37°C. 2 μl of the reaction was used for a *E. coli* transformation

Table 23: Composition of the Gibson assembly master mix

Gibson assembly master mix
1 x Isothermal reaction buffer (NEB)
4 U/μl T5 exonuclease (NEB)
4 U/μl Taq DNA ligase (NEB)
25 U/μl Phusion DNA polymerase (NEB)

5.2.10 Genomic tagging of *S. cerevisiae*

The genomic integration of any kind of protein tag is obtained by homologous recombination. Special primers are needed for this method. The forward primer consists of 40 - 50 nucleotides of the yeast genomic sequence, flanking the integration site before the stop codon plus the 5' end of the sequence containing the protein tag and an auxotrophy marker. The reverse primer consists of 40 - 50 nucleotides of the 3' UTR of the yeast genomic sequence region, the stop codon and the 3' end of the sequence containing the protein tag and an auxotrophy marker. These primers were used to amplify the protein tag according to the PCR protocol (see the section "Polymerase chain reaction"), The PCR product was precipitated (see the section "DNA precipitation") and transformed into the yeast cells to achieve integration into the genome by homologous recombination (see the section "Transformation of *S. cerevisiae*"). Transformants were tested for integration of the tag by colony PCR (see the section "Colony

PCR") or Western Blot analysis (see the section "Western Blot analysis"). To this end, a whole cell extract was prepared of freshly grown cells. A white loop of cells was suspended in 100 μl 1 $\times$ SDS sample buffer and approximately 30 μl of glass beads. The suspension was lysed using the FastPrep-24 5G device (20 s at 6 m/s) 15 μl of the lysate was loaded on a SDS-polyacrylamide gel for Western Blot analysis.

5.2.11 Deletion of a gene in *S. cerevisiae*

The deletion of a gene in *S. cerevisiae* is obtained by homologous recombination. The coding sequence of the respective gene has to be replaced by an auxotrophy marker or an antibiotic resistance gene. Special primers are used, that carry homologous sequences to the promoter and the 3'-downstream region of the gene to be deleted. A PCR reaction of a volume of 300 μl was performed. A plasmid containing the auxotrophy or the antibiotic resistance gene for gene replacement was used as a template. The following steps are described in the section "Genomic tagging of *S. cerevisiae*".

5.3 Biochemical methods

5.3.1 SDS polyacrylamide gel electrophoresis

SDS-PAGE was carried out using a Mini Protean II system (BioRad). Samples for SDS-PAGE were mixed with SDS sample buffer (final concentration 1 $\times$), incubated at 95°C for 5 min and centrifuged at 13,000 rpm for 3 min. After sample loading, the SDS-PAGE was performed in 1 $\times$ SDS-PAGE running buffer at 200 - 250 V until the dye front reached the bottom of the separating gel. 7 μl Unstained Protein standard Broad Range 10 - 200 kDA (NEB) or the Color Protein standard Broad Range (NEB) was used as marker. The gel was disassembled and stained in heated Coomassie stain solution for 20 min. The gel was destained in heated 10 % acetic acid serval times until the background of the gel is becoming colorless. A typical 10% SDS-polyacrylamid gel contains the following mixture.

Table 24: Composition of a SDS polyacrylamide gel

component	10% separating gel	4.5% stacking gel
4 × separating gel buffer	1.25 ml	-
4 × stacking gel buffer	-	1 ml
30% acrylamide/bisacrylamide (29:1)	1.66 ml	0.33 ml
H_2O	2.1 ml	1.8 ml
TEMED (100%)	5 μl	5 μl
10% APS	50 μl	15 μl

5.3.2 Urea polyacrylamide gel electrophoresis

Urea PAGE was carried out using a Mini Protean II system (BioRad). Samples for Urea PAGE were mixed with 2 × formamide dye, incubated at 95°C for 3 min and centrifuged at 13,000 rpm for 3 min. For preparing an urea polyacrylamide gel the following components were heated for 5 - 6 s in the microwave.

Table 25: Composition of an urea polyacrylamide gel

Urea polyacrylamid gel
8 M urea
2.8 ml 37% acrylamid/bisacrylamid (29:1)
1.5 ml 10 × TBE buffer
3.5 ml H_2O

8 μl TEMED and 80 μl 10% APS were added before pouring the gel using a Mini Protein II system (BioRad). The pockets were cleaned to remove extra urea and the urea gel was incubated at 180 V for 20 - 30 min before the samples were pipetted into the pockets of the urea polyacrylamide gel. After loading, the urea PAGE was performed in 1 × TBE buffer at 200 V for 45 min.

5.3.3 Western Blot analysis

A polyacrylamide gel after SDS PAGE can be also used to perform a Western Blot analysis to detect specific proteins. Separated proteins of a polyacrylamide gel were transferred onto a nitrocellulose membrane (Porablot, Macherey & Nagel) using the semi-dry blotting system (Biorad) for 45 min at 25 V. The transfer of proteins was verified by Ponceau S staining. After the transfer, the membrane was blocked for 1 h with blocking solution (5% milk-powder solved in 1x TBST). The primary antibody dissolved in blocking solution was added

and incubated overnight at 4°C. Excess of the first antibody was removed by washing the membrane 3 × for 10 min with TBST at room temperature on a shaker. The secondary antibody coupled to horseradish peroxidase was applied and incubated for 1.5 h. After washing with 1x TBST for 3 × each 10 min, the visualization of the immunodetected proteins was performed using the CheLuminate-HRP PicoDetect ECL kit (Applichem). Pictures of the blot were taken with the ChemoCam Imager (Intas).

5.3.4 Quantitative Western Blot analysis

Applying quantitative Western Blots, an overnight yeast culture was diluted to an OD_{600} of 0.2 in 25 ml YPD. 5 OD_{600} units of an OD_{600} of 0.6 - 0.8 were harvested. (3600 rpm, 3 min) The pellet was washed once with 10 ml PBS and resuspended in cold H_2O. 150 μl pre-treatment solution (composition see table below) were added and the cells were alkaline lysed for 20 min on ice. Proteins were precipitated adding 150 μl 55% trichloroacetid acid (TCA). After an incubation time for 20 min on ice the samples were centrifuged for 20 min at 13000 rpm (4°C). The protein pellet is solved in 85 μl of 1 × SDS sample buffer and 20 μl 1 M TRIS. The samples were incubated for 10 min at 95°C. A SDS- PAGE and a subsequent Western Blot were applyed as it is described in the section "SDS polyacrylamide gelelectrophoresis" and "Western Blot analysis". The Western Blot signals were qantified using ImageJ.

Table 26: Composition of the pre-treatment solution

pre-treatment solution
7,5% β-Mercaptoethanol
1,85 M NaOH

5.3.5 Determination of protein stability

To determine protein stability, an overnight yeast culture was diluted to an OD_{600} of 0.2 in 30 ml YPD. 2.5 OD_{600} units yeast culture per time point of an OD_{600} of 0.6 - 0.8 were harvested (3600 rpm, 3 min) and resuspended in 1 ml fresh 30°C pre-warmed YPD per time point. The yeast cells were equilibrated for 5 min in a 30°C water-bath. 950 μl yeast cell suspension was pipetted to a reaction tube containing 50 μl 20 × stop-mix (composition see table below). These tubes were vortexed and stored in ice. Immediately after this step, a final concentration of 125 μg/mL cycloheximide was added to the cell culture. A sample of 950 μl cell culture was taken after 30 min, 60 min and 90 min and added to the stop solution. When all samples have been collected, the samples were centrifuged at 6500 g for

30 s. The pellet was resuspended in 100 μl cold H_2O. Afterwards, 100 μl 0.2 M NaOH was added and mixed by vortexing. After an incubation time of 5 min at room temperature, the samples were centrifuged 30 s at 18000 g. The pellet was solved in 100 μl SDS sample buffer. The samples were incubated for 5 min at 95°C. A SDS - PAGE and a subsequent Western Blot analysis were applied as it is described in the section "SDS polyacrylamide gelelectrophoresis" and "Western Blot analysis". The Western Blot signals were qantified using ImageJ.

Table 27: Composition of the stop solution

stop solution
200 mM sodium azide
5 mg/ml bovine serum albumin

5.3.6 Tandem affinity purification

Tandem affinity purification (TAP) was performed essentially as described in Straesser et al., 2002. Yeast strains containing a TAP-tag fused to the protein to be purified were used. The TAP-tag consists of two protein A domains, the TEV cleavage site and the calmodulin binding peptide.
Yeast cells were grown in 2 L YPD at 30°C to an OD_{600nm} of 3 - 3.6. Cells were harvested (4000 rpm; 5 min) and the pellet was resuspended in 2.5 ml LB-buffer. The cell suspension was dropped in a vessel containing liquid nitrogen to get flash frozen cells. The frozen cells were lysed using a freezer mill 6870D (SPEX SamplePrep) and stored at -80°C until use.
10 ml TAP-buffer + 1 × protease inhibitor and 1 mM DTT were added to the frozen lysate. The lysate was thawed at room temperature on a turning-wheel. The thawed lysate was cleared from cellular debris by centrifugation for 10 min at 4000 rpm (4°C) and 60 min at 40000 rpm (4°C). The top fatty phase was removed by aspiration and the clear lysate was collected. An aliquot of 30 μl was taken for further analysis.
400 μl Fast Flow IgG Sepharose beads (GE) were washed 3 × in cold TAP-buffer (2 min, 1800 rpm, 4°C) and were added to the lysate. After an incubation step for 1.5 h on a turning wheel at 4°C, the IgG beads were centrifuged and transferred to a mobicol. Beads were washed with 15 ml TAP-buffer + 0.5 DTT by gravity flow. Bound proteins were eluted by TEV cleavage for 1 h on a turning wheel at 16°C in 120 μl TAP-buffer containing 4 μl (2 mg/ml) TEV protease (prepared by Heike Büngen). An aliquot of the TEV eluate was taken for further analysis. The TEV-eluate was bound to 400 μl prewashed (3 × in TAP-buffer + 1 mM DTT and 2 mM $CaCl_2$) Calmodulin Sepharose Beads (Agilent Technologies) for 1 h at

4°C in 150 μl TAP-buffer supplemented with 1 mM DTT and 4 mM $CaCl_2$. Calmodulin beads were washed with 5 ml TAP-buffer + 1 mM DTT and 2 mM $CaCl_2$ by gravity flow. 600 μl TAP elution buffer (10 mM TRIS HCl pH 8.0; 5 mM EGTA) was added to the mobicol to elute the purified protein complex. Calmodulin beads were incubated for 15 min at 37°C at a thermoshaker. Proteins were precipitated using 10% trichloroacetic acid (TCA) for 20 min on ice and 20 min centrifugation at 13000 rpm and 4°C. The protein pellet was resuspended in 60 μl 1 × sample buffer and neutralized using 1 M TRIS.
SDS-PAGE was carried out to identify the qualitiy of the purification (see section "SDS polyacrylamide gelelectrophoresis").

Table 28: Composition of the TAP-buffer

TAP-buffer
50 mM TRIS HCl pH 7.5
100 mM NaCl
1.5 mM $MgCl_2$
0.15% NP-40

5.3.7 Cross-linking mass spectrometry

Isolation and cross-linking of proteins

Yeast strains containing a FTpA-tag fused to the protein to be purified were used. The FTpA-tag consists of two protein A domains, the TEV cleavage site and the FLAG tag.
Harvesting and cell lysis was performed essentially as described in the chapter "tandem affinity purification". Yeast cells of 24 L YPD OD_{600nm} 3.6 were harvested to purify one protein complex for XL/MS. 5 ml LB-buffer + 1 × protease inhibitor and 1 mM DTT (composition see table below) were added to the frozen lysate. The thawed lysate was cleared from cellular debris by centrifugation for 10 min at 4000 rpm (4°C). The supernatant of 1.5 thawed lysates was put to one ultracentrifuge tube. After ultracentrifugation at 40000 rpm for 60 min (4°C) the clear lysate of two ultracentrifuge tubes was collected in one 50 ml falcon.
600 μl Fast Flow IgG Sepharose beads (GE) were added to the cleared lysates. The samples were incubated on a turning wheel at 4°C over night. IgG beads were washed 3 times with 10 ml wash buffer (LB-buffer + 0.5 mM DTT) 3 min at 1800 rpm. The second wash step was performed with 10 ml wash buffer by gravity flow. To this end, the beads of two 50 ml falcons were collected into one mobicol. TEV cleavage was performed in 100 μl wash buffer containing 15 μl (2 mg/ml) TEV protease at 16°C for 1 h at a turning wheel. 150 μl TEV-

eluate was bound to 150 μl prewashed (3 × in TAP-buffer + 1 mM DTT) anti-FLAG M2 affinity beads (Sigma aldrich) for 1 h at 4°C. FLAG beads were washed with 5 ml XL-buffer (compositon see table below) + 0.15% NP-40 by gravity flow. Proteins were eluted in 100 μl XL-buffer containing 100 μg/ml FLAG peptide (Sigma aldrich) 15 min on a turning wheel at RT. The purified proteins were concentrated using a Vivaspin concentrator (Sartorius) according to the manufacturer's instructions. The protein concentration was determined according to Bradford (Applichem).
XL-reactions were performed in 100 μl. The appropriate concentration of BS^3 H12/D12 was determined performing different crosslinking tests. 60 - 70 μl of the purified proteins were incubated with the appropriate concentration of BS^3 H12/D12 and 50 μl XL-buffer for 30 min at 37°C 500 rpm. The reaction was stopped adding 50 mM NH_4HCO_3. The cross-linked proteins were stored at -80°C until XL/MS was performed.

Table 29: Composition of the LB-buffer and XL-buffer

LB-buffer	**XL-buffer**
200 mM KCl	150 mM KCl
500 mM HEPES pH 7.5	20 mM HEPES pH 8.4
1.5 mM $MgCl_2$	
0.15% NP-40	

Sample processing for mass spectrometry

After quenching or gel filtration, cross-linked samples were evaporated to dryness in a vacuum centrifuge and processed according to Leitner et al., 2014. In brief, disulfide bonds were reduced with tris(2-carboxyethyl)phosphine and free thiol groups in cysteines were alkylated with iodoacetamide. Proteins were digested in a two-step procedure, first with endoproteinase Lys-C (Wako; 1:100, 37°C for 2-3 h), then with trypsin (Promega; 1:50, 37°C overnight). Samples were purified by solid-phase extraction (SepPak tC18 cartidges, Waters) and fractionated by peptide-level size-exclusion chromatography (SEC; Superdex Peptide PC 3.2/300, GE) on an ÄKTA micro FPLC system.

Liquid chromatography-tandem mass spectrometry

SEC fractions and additionally some unfractionated samples were analyzed by nanoflow LC-MS/MS. The cross-linked samples were initially analyzed on an Orbitrap Elite system and later reanalyzed on an Orbitrap Fusion Lumos instrument (both Thermo Fisher). The Orbitrap Elite instrument was connected to an Easy nLC-1000 HPLC system (Thermo Fisher) with a 150 mm × 75 μm Acclaim PepMap RSLC C18 column (2 μm particle size, Thermo

Fisher) and mobile phases A = water/acetonitrile/formic acid (98:2:0.15, v/v/v) and B = acetonitrile/water/formic acid (98:2:0.15, v/v/v). Peptides were separated at a flow rate of 300 nL/min with a gradient of 9 to 35% B in 60 min. The mass spectrometer was operated in data-dependent acquisition mode with detection of precursor ions in the Orbitrap at 120000 resolution and detection of fragment ions in the linear ion trap at normal resolution. For each sequencing cycle, the 10 most abundant precursor ions (top n mode) were selected for sequencing if they had a charge state of +3 or higher and were fragmented at 35% normalized collision energy. Dynamic exclusion was activated for 30 s after one sequencing event. The Orbitrap Fusion Lumos mass spectrometer was connected to an Easy nLC-1200 HPLC instrument (Thermo Fisher) with a 250 mm × 75 μm Acclaim PepMap RSLC C18 column (2 μm particle size, Thermo Fisher) and mobile phases A = water/acetonitrile/formic acid (98:2:0.15, v/v/v) and B = acetonitrile/water/formic acid (80:20:0.15, v/v/v). Peptides were separated at a flow rate of 300 nL/min with a gradient of 11 to 40% B in 60 min. The mass spectrometer was operated in data-dependent acquisition mode with detection of precursor ions in the Orbitrap at 120,000 resolution and detection of fragment ions in the linear ion trap at rapid resolution (low-resolution MS/MS) or in the Orbitrap at 30,000 resolution (high-resolution MS/MS). For each sequencing cycle, the most abundant precursor ions were selected for sequencing with a cycle time of 3 s (top speed mode) if they had a charge state of +3 or higher and were fragmented at 35% normalized collision energy. Dynamic exclusion was activated for 30 s after one sequencing event.

Identification of cross-linked peptides

MS/MS data acquired in the native Thermo .raw format were converted into the mzXML format by msconvert/ProteoWizard [Chambers et al., 2012] and searched using xQuest [Walzthoeni et al., 2012], version 2.1.4, against a custom database of target proteins and identified contaminants. The false discovery rate (FDR) was controlled by a target/decoy search using reversed and shuffled sequences as decoys. xQuest search parameters included: enzyme = trypsin, maximum number of missed cleavages = 2, fixed modifications = carbamidomethylation on Cys, initial MS1 error tolerance = +/-15 ppm, MS2 error tolerance = 0.2 Da for "common" fragment ions and 0.3 Da for "xlink" fragment ions (ion trap) or = +/-20 ppm (Orbitrap), respectively. Cross-linking specificity was set to K and N terminus. Primary search results (target and decoy hits) were then further filtered according to the experimentally observed mass accuracy, a TIC value of >0.14 and a ΔS value of <0.9.

All remaining candidate hits were further curated and only assignments for which a minimum of four bond cleavages overall or three consecutive bond cleavages were assigned for each peptide were retained. The FDR was adjusted to 5% at the cross-linked peptide pair level after the manual curation steps.

5.4 Molecular biology methods

5.4.1 Total RNA extraction

Experiments with RNA require general precautions to provide an RNase-free enviromnent. Surfaces and pipettes were cleaned with ethanol. All buffers were prepared with DEPC-treated water. Total RNA was extracted from yeast cells growing in YPD at 30°C to an OD_{600nm} of 0.8. The pellet was resuspended in 500 μl AE buffer (composition see table below). 55 μl 10% SDS and 500 μl phenol/chloroform/isoamyl ethanol (25:24:1) (PCI) were added. The suspension was incubated for 20 min at 65°C, 1000 rpm. Afterwards, the suspension was cooled on ice for 2 min and centrifugated at 14000 rpm for 10 min. The upper phase was taken, mixed with an equal volume of PCI and vortexed for 30 s at maximum intensity. After centrifugation (14000 rpm; 5 min), the upper phase was taken. The nucleic acid solution was mixed with 400 μl chloroform, vortexed and centrifuged at 14000 rpm for 5 min. The upper phase was taken and a DNA/RNA precipitation was performed. (see the section "RNA/DNA precipitation")
To eliminate DNA a DNase digest was conducted. For this purpose, the precipitated nucleic acid was solved in 39 μl DEPC-trated water, 5 μl 10 × DNase buffer (Thermo Fisher), 5 U DNase I (Thermo Fisher) and 1 μl RNase Inhibitor (Thermo Fisher) and was incubated 1 h at 37°C. A phenol/chloroform extraction was performed to eliminate the DNase. The RNA concentration was measured with the Nanodrop and the RNA was stored at -80°C until further use.

Table 30: Composition of the AE-buffer

AE-buffer
50 mM $C_2H_3NaO_2$
10 mM EDTA
10 mM Tris HCl pH 7.5

5.4.2 Isolation of poly(A) mRNA

Poly(A) mRNA was isolated of DNase-treated total RNA using oligo$(dT)_{25}$ Dynabeads (Thermo Fisher) according to the manufacturer's instruction. The amount of poly(A) mRNA was measured with the Nanodrop. Afterwards, poly(A) mRNA was stored at -80°C until further use.

5.4.3 Isolation of nascent mRNA

The *S. cereviseae* strain RPB3-FTpA was used to purify nascent RNA. Yeast cells were grown in 2 L YPD at 30°C to an OD_{600nm} of 0.8. Cells were harvested (4000 rpm 5 min) and the pellet was resuspended in 2.5 ml lysis buffer. The cell suspension was dropped in a vessel containing liquid nitrogen to get flash frozen cells. Frozen cells were lysed using a freezer mill 6870D (SPEX SamplePrep) and stored at -80°C until use.
Lysed cells were resuspended in 5 ml lysis-buffer (composition see table below) + 1 × protease inhibitor; + 50 U/ml RNase inhibitor (Thermo Fisher). 660 U DNase I (Thermo Fisher) was added to the lysate and the DNase digest was incubated for 20 min on ice. The digested lysate was centrifuged 10 min, 16000 g at 4°C. The cleared lysate was collected, input samples were taken and the lysate was bound to 400 μl prewashed (2 × in lysis buffer) anti-FLAG M2 affinity beads (Sigma aldrich) for 2.5 h at 4°C on a turning wheel. Beads were washed 4 × with 10 ml wash buffer (composition see table below). Washed beads were transferred to a mobicol and centrifuged 2 min at 1000 g to remove the washing buffer. Proteins were eluted 2 × in 300 μl lysis buffer containig 1 mg/ml FLAG peptide (Sigma aldrich) for 30 min on ice. The mobil was inverted every 5 minutes. A sample of the eluate was taken to perform a Western Blot. Nascent mRNA of the remaining sample was purified using the miRNeasy Kit (Qiagen) according to the manufacturer's instruction. The concentration of the nascent mRNA was measured with the Nanodrop. The RNA stored at -80°C until further use.

Table 31: Composition of the lysis buffer and the wash buffer

lysis buffer	**wash buffer**
20 mM HEPES pH 7.4	20 mM HEPES pH 7.4
110 mM CH_3CO_2K	110 mM CH_3CO_2K
0.5% Triton-X100	0.5% Triton-X100
0.1% Tween 20	0.1% Tween 20
10 mM $MnCl_2$	1mM EDTA

5.4.4 Reverse transcription

A reverse transcription was performed of 100 ng DNase treated RNA using Super Script II Reverse Transcriptase (Thermo Fisher) according to the manufacturer's instruction. The cDNA was precipitated and solved in 100 μl water. The cDNA was stored at -20°C until further use.

5.4.5 Quantitative real-time PCR

qPCR is a method to quantify the amount of DNA by usage of a sensitive fluorescent dye. In this study SYBR green, a component of the Power Sybr Green PCR Master Mix (Applied Biosystems) and the StepOnePlus Real Time PCR system (Applied Biosystems) was used as described in the manual.
2.5 μl of the diluted DNA or cDNA and 7.5 μl of the reaction mix containing 0.1 μl (100 pmol/μl of each primer, 5 μl 2 × SYBR Green qPCR Master Mix (Thermo Fisher) and 2.3 μl of ddH_2O were pipetted into a Micro Amp Optical 96 well plate (Applied Biosystems). To determine primer efficiencies (E), a standard curve was performed. The specificity and dimer formation of the primers were tested by melting curve analysis of the individual amplification products. A non-transcribed region on Chromosome V (174131-174200) (*YER*) served as a negative control to normalize for background. Each tested sample or controls (standard curve, water control) were performed as triplicates. A typical qPCR run is shown in table 33, a list of qPCR primers is given in table 14.

Table 32: Program for qPCR

Temperature	**Time**
95°C	10 min
95°C	15 s
60°C	60 s
	steps 2 - 3 were repeated 45 ×
95°C	15 s
60°C	60 s
60 - 95°C	continuously 0.3°C
95°C	15 s

CT values were calculated by the StepOne^{TM} Software (version 1.5.1) (Applied Biosystems). The occupancy was calculated according to the comparative CT method as the enrichment

of the tested gene relative to the non-transcribed gene *YER*. The following formula was used as described in Livak and Schmittgen, 2001 (E = primer efficiency):

$$\frac{\left(E^{CT\ IP-CT\ IN}\right)_{YER}}{\left(E^{CT\ IP-CT\ IN}\right)}$$

5.4.6 Determination of mRNA stability

To determine mRNA stability of selected transcripts, yeast cells were grown at 30°C until they reached an OD_{600} of 0.5. Immediately after harvesting 5 mL culture, a final concentration of 8 μg/mL thiolutin was added to the cell culture. 4 × 5 ml cell culture were harvested within 1 h. Exact time points depends on the half-life of the selected transcripts. Cells were added to 2.5 ml methanol on dry-ice to stop the metabolism. Samples were centrifuged 2 min at 4000 rpm (4°C), washed once with PBS and flash frozen in liquid nitrogen. Afterwards, yeast cells were lysed, the RNA was extracted, DNase treated and precipitated. (see the sections "Total RNA extraction & DNA/RNA precipitation"). Total RNA was reverse transcribed using Super Script II as it was described in the chapter 5.4.4.. The resulting cDNA was analyzed by qPCR (see chapter 5.4.5.).

5.4.7 *In vivo* transcription assay

To determine the transcription rate, yeast strains containing the selected transcripts under the control of the galactose promoter, were used. These yeast strains were grown in media containing raffinose until they reached an OD_{600} of 0.6. 5 ml yeast culture was harvested. Immediately after this step, 1 % galactose was added to the yeast cell suspension to induce transcription of genes, which were regulated by a galactose promoter. 5 ml yeast cells were harvested 5, 10, 20 and 30 min after the addition of galactose. The cells were added to 2.5 ml methanol on dry-ice to stop the metabolism immediately. Samples were centrifuged 2 min at 4000 rpm (4°C), washed once with PBS and flash frozen in liquid nitrogen. Afterwards, yeast cells were lysed, RNA was extracted, DNase treated and precipitated. (see the sections "Total RNA extraction & DNA/RNA precipitation"). Total RNA was reverse transcribed using Super script II as it was described in the chapter 5.4.4.. The resulting cDNA was analyzed by qPCR (see chapter 5.4.5.).

5.4.8 Chromatin-Immunoprecipitation

ChIP was performed to analyze the association of proteins to their target genes. ChIP was performed as described in Meinel et al., 2013. *S. cerevisiae* strains were grown in 100 ml of the appropriate medium from OD_{600} 0.2 to 0.8 and cross-linked with 1% formaldehyde for 20 min at room temperature on a shaker. The process is stopped by addition of 0.25 M glycine. Cells were incubated for 5 min at room temperature with mild agitation. Cross-linked cells were harvested by centrifugation (3600 rpm, 3 min, 4°C), washed 2 × with cold 1 × PBS and flash frozen in liquid nitrogen. Pellets were stored at - 80°C until further use.
Pellets were thawed on ice and resuspended in 800 μl FA-Lysis buffer (low salt) + 1 × protease inhibitor and 800 μl glass beads. Cells were lysed using the FastPrep 24G device two times for 45 sec (6.5 m/s) with a 2 min break on ice in between. Afterwards, the cell lysate was sonicated 3 times for 15 min with a 5 min break on ice in between using a Bioruptor (Diagenode) to fragmentate the chromatin to a size of 250-500 bp . The lysate was cleared by centrifugation two times 5 min and then 10 min at 13000 rpm at 4°C. DNA concentration (adsorption at 280 nm) of each sample was measured using the Nano Drop ND-1000 spectrometer (Thermofisher Scientific). All lysates were diluted to an equal DNA concentration. An aliquot of 10 μl was taken as the input (IN) sample. The remaining sample was incubated with 15 μl prewashed (3 × in FA-Lysis buffer low salt) magnetic beads (Invitrogen) coupled to IgG (for TAP-tagged proteins) or a specific antibody against the protein of interest. IgG coupled magnetic beads were incubated for 2.5 h at room temperature on a turning wheel. Uncoupled antibodies were incubated 1.5 h at room temperature. Afterwards, 15 μl magnetic beads were added and incubated for one hour on the turning wheel. After an incubation time of 2.5 h, the beads were collected and washed with 800 μl of the following buffers (2 × FA–Lysis low salt, 2 × FA-Lysis high salt, 2 × TLEND and 1 × TE). Washed beads were resuspended in 130 μl elution buffer and incubated for 20 min at 65°C on a thermoskaker. Afterwards, the beads were collected and the supernatant (IP sample) was transferred to a 0.5 ml tube. The IN samples and IP samples were mixed with 80 μl 1 × TE and 10 μl Proteinase K (10 mg/mL). Additionally, 80 μl ChIP elution puffer was added to the IN samples. The samples were incubated 2 h at 37°C, followed by 12 - 16 h incubation at 65°C. Samples were purified using the Nucleo Spin Gel and PCR clean up Kit (Macherey & Nagel) according to the manufacturer's instructions. The purified DNA is eluted in 140 μl elution puffer (Macherey & Nagel). qPCR (see the section "Quantitative real-time PCR") was performed to quantify the amount of the DNA.

Table 33: Composition of the ChIP buffers

FA - Lysis (low salt)	**FA - Lysis (high salt)**	**TLEND**	**Elution Buffer**
50 mM HEPES	50 mM HEPES	10 mM Tris HCl pH 8.0	50 mM Tris HCl pH 7.5
1 mM EDTA	1 mM EDTA	1 mM EDTA	10 mM EDTA
0.1% SDS	0.1% SDS	0.25 M LiCl	1% SDS
0.1% sodium deoxycholate	0.1% sodium deoxycholate	0.5% sodium deoxycholate	
1% Triton - X100 (Triton - X100)	1% Triton - X100	0.5% NP 40	
150 mM NaCl	500 mM NaCl		

5.4.9 Chromatin-Immunoprecipitation sequencing

ChIP was performed as described above (see the section "Chromatin-Immunoprecipitation") with the following modifications. To optimize the wash step, it was washed 6 × with FA–Lysis low salt. Purified DNA samples were solved in water in stead of elution buffer (Macherey & Nagel).

Libary preparation

DNA samples were quantified by Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific). 10 ng of DNA was used as input for TruSeq ChIP Library Preparation Kit (Illumina) with following modifications. Instead of gel-based size selection before final PCR step, libraries were size selected by SPRI-bead based approach after final PCR with 18 cycles. In detail, samples were 1st cleaned up by 1x bead:DNA ratio to eliminate residuals from PCR reaction, followed by 2-sided-bead cleanup step with initially 0.6x bead:DNA ratio to exclude larger fragments. Supernatant was transferred to new tube and incubated with additional beads in 0.2x bead:DNA ratio for eliminating smaller fragments, like adapter and primer dimers. Bound DNA samples were washed with 80 % ethanol, dried and resuspended in TE buffer. Library integrity was verified with LabChip Gx Touch 24 (Perkin Elmer). Sequencing was performed on the NextSeq2000 instrument (Illumina) using P3 flowcell with 1x72bp single end setup.

Data analysis

After mapping to the genome, reads were counted with Python package HTSeq-count (version 0.13.5) [Anders et al., 2015]. Gene assignment was based on annotation data by *Saccha-*

romyces cerevisiae Genome Database (genome version R64-1-1). For differential gene expression analysis, R/Bioconductor package DeSeq2 (version 1.28.1) was used [Love et al., 2014]. Log2 transformed fold changes were modified applying the shrinkage estimator "apeglm". Genes with adjusted p-value < 0.05 were considered as significantly differently expressed genes and further classified according to Gene Ontology Biological Processes. Classification was performed in R using the package clusterProfiler (version 3.16.0) [Yu et al., 2012].

5.4.10 RNA Immunoprecipitation

RIP was performed to analyze the association of proteins to RNAs. TAP-tagged *S. cerevisiae* strains were grown in 400 ml YDP to an OD_{600} of 0.8, harvested and stored at -80°C until further use.

Frozen pellets were thawed at ice, resuspended in 1 ml RNA IP-buffer (composition see table below) + 20 μl 100 × protease inhibitor and lysed using the Fast-Prep machine (3 × 20 s 6 m/s). The lysate was cleared by centrifugation 5 min at 4000 rpm 4°C and 10 min at 13000 rpm 4°C. 900 μl of the cleared lysate was incubated with 660 units DNase I (Thermo Fisher) for 30 min on ice. Two aliquots of 10 μl and 7.5 μl were taken as the input (IN) samples for SDS page and RNA isolation. 40 μl IgG-coupled magnetic dynabeads were added and the samples were incubated for 3 h at 4°C. Beads were washed 8 times with RNA-IP-buffer. For RNA-analysis, 1 ml TRIzol was added to the washed beads. RNA was isolated and precipitated according to the manufacturer's instruction. The RNA of the input and IP samples were analyzed performing reverse transcription with subsequent qPCR to quantify the amount of the immunoprecipitated RNA (see the section "reverse transcription" and "Quantitative real-time PCR").

For protein analysis, an acetone precipitation of the phenol- and interphase was performed. Therefore, 5 × volume of acetone was added to the phenol- and interphase and incubated over night at -20°C. After centrifugation (15.000 rpm, 1 h, 4°C), the precipitated proteins were washed 2 × with acetone, dried and resuspended in 60 μl of 1 × SDS loading dye. SDS PAGE of the input and IP samples was performed to identify the quality of the immunoprecipitation (see section "SDS Polyacrylamid Gel Electrophoresis").

Table 34: Composition of RNA IP-buffer

RNA IP-Buffer
25 mM TRIS HCl pH 7.5
150 mM NaCl
2 mM $MgCl_2$
0.2% Triton-X-100

5.4.11 Fluorescence *in situ* hybridization with oligo dT

In situ hybridization against poly(A) + RNA was done according to Amberg et al., 1992. Cells were grown at 30°C in YPD medium to mid-log phase. An aliquot of 10 mL cell culture was fixed with 4% formaldehyde, washed after fixation and spheroplasted with 100 T zymolyase. After adhering spheroplasts to a poly-lysine-coated slide, cells were prehybridized for 1 hour at 37°C in prehybridization buffer in a humid chamber. To hybridize with oligo dT 50-Cy3 0.75 μl of 1 pmol/μl probe was added and incubated at 37°C over night in a humid chamber. After hybridization, cells were washed, mounted with ROTI Mount FluorCare DAPI and covered with a coverslip. Cells were inspected with a Delta Vision Ultra fluorescence microscope (Cytiva). The quantification was conducted of 300 randomly selected cells. The average intensity of the nucleus and the cytoplasm was determined using ImageJ. Violin blots were generated by means of a R-script.

Table 35: Composition of the FISH buffers

prehybridization buffer	**wash buffer FISH**	**20x SSC pH 7.0**
50% formamide	1.2 M sorbitol	3 M NaCl
10% dextran sulphate	100 mM KPO_4	300 mM sodium citrate
125 μ/mL *E. coli* tRNA tRNA		
500 μ/mL H.S. DNA		
4 × SSC		
0.02% polyvinyl pyrrolidone		
0.02% BSA		
0.02% Ficoll-40		

5.4.12 [$\gamma-^{32}$P] ATP labeling of oligonucleotides

5 μl 10 × T4 PNK buffer, 2 μl of the purified 1:10 diluted oligonucleotides, 3 μl $\gamma-^{32}$P ATP (7000 Ci/mmol, 150 mCi/m) and 38 μl nuclease-free water were mixed. 2 μl (10 U/ml) T4 polynucleotide kinase was added and incubated for 30 min at 37°C. The kinase was heat-inactivated for 20 min at 65°C. Unincorporated $\gamma-^{32}$P ATP was removed by column chromatography using a "Mini quick spin oligo column" (Roche) according to the manufacturer's instruction.

5.4.13 Primer extension after CMC treatment

Primer extension after CMC labeling was performed to detect the Pseudouridylation site of a RNA. *S. cerevisiae* strains were grown in 25 ml YDP to an OD_{600} of 0.8, harvested and stored at -80°C until further use.
Yeast cells were lysed, the RNA was extracted, DNase treated and precipitated. (see the sections "Total RNA extractration & DNA/RNA precipitation") To mark uridines, pseudouridines and also some guanosines, the RNA was treated with 200 mM CMC in 75 μl BEU buffer (composition see table below) for 30 min at 40°C. In parallel, the same sample was incubated with BEU buffer lacking CMC as a negative control. Afterwards, CMC-treated RNA was precipitated. The RNA is resuspended in 40 μl $NaCO_3$ buffer (composition see table below) to remove CMC from non-pseudouridines. The solution is incubated at 50°C for 2 h and subsequently precipitated. The dried pellet is solved in 30 - 50 μl nuclease - free water. The concentration of the RNA was determined by NanoDrop.

Table 36: Composition of the BEU buffer Na_2CO_3 buffer

BEU buffer	**$NaCO_3$ buffer**
7 M urea	50 mM $NaCO_3$ pH 10.2
4 mM EDTA	2 mM EDTA
50 mM bicine pH 8.5	

A reverse transcription followed. An equal concentration of 1-2 μg total RNA was mixed with 10 mM dNTPs, 0.5 μl [$\gamma-^{32}$P] ATP labeled primer and nuclease free water to a total volume of 10 μl. The solution was incubated 5 min at 70°C for denaturation. 2 μl 10 × M-MulV buffer, 1.2 μl M-MulV reverse transcriptase (200 U/μ), 0.2 μl RNase inhibitor (40 U/μl) and 6.6 μl nuclease - free water were added. The reaction was incubated for 1 h at 42°C. cDNA was precipitated (see the section "DNA/RNA precipitation"). The pellet was resuspended

in 4.5 μl RNase A (0.4 mg/ml) and incubated for 3 min at room temperature to dissolve the RNA strand out of the duplex. 4.5 μl 2 $\times$ Formamide DNA loading dye was added and an urea PAGE was performed at 200 V for 45 min (see section "Urea polyacrylamid eletrophoresis"). The urea polyacrylamid gel was exposed overnight to a storage phosphor screen and analyzed by a phosphorimager (Typhoon 9400, GE Healthcare).

List of Figures

List of Tables

References

Abruzzi, K. C., S. Lacadieand, and M. Rosbash (2004). "Biochemical analysis of TREX complex recruitment to intronless and intron-containing yeast genes". In: *The EMBO journal* 23, pp. 2620–2631.

Akhtar, M. S., M. Heidemann, J. R. Tietjen, D. W. Zhang, R. D. Chapman, D. Eick, and A. Z. Ansari (2009). "TFIIH kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II". In: *Molecular cell* 34, pp. 387–393.

Alber, F., S. Dokudovskaya, L. M. Veenhoff, W. Zhang, J. Kipper, D. Devos, A. Suprapto, O. Karni-Schmidt, R. Williams, B. T. Chait, A. Sali, and M. P. Rout (2007). "The molecular architecture of the nuclear pore complex." In: *Nature* 450, pp. 695–701.

Amberg, D. C., A. L. Goldstein, and C. N. Cole (1992). "Isolation and characterization of RATI: an essential gene of *Saccharomyces cerevisiae* required for the efficient nucleo cytoplasmic trafficking of mRNA". In: *Genes & Development* 6, pp. 1173–1189.

Amrani, N., M. Minet, M. Le Gouar, F. Lacroute, and F. Wyers (1997). "Yeast Pab1 interacts with Rna15 and participates in the control of the poly(A) tail length in vitro". In: *Molecular Cell Biology* 17, pp. 3694–3701.

Anders, S., P. H. Pyl, and W. Huber (2015). "HTSeq–a Python framework to work with high-throughput sequencing data". In: *Bioinformatics* 31, pp. 166–169.

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

Arnez, J. G. and T. A. Steitz (1994). "Crystal structure of unmodified tRNA(Gln) complexed with glutaminyl-tRNA synthetase and ATP suggests a possible role for pseudo-uridines in stabilization of RNA structure". In: *Biochemistry* 33, pp. 7560–7567.

Bakin, A. and J. Ofengand (1993). "Four newly located pseudouridylate residues in *Escherichia coli* 23S ribosomal RNA are all at the peptidyltransferase center: analysis by the application of a new sequencing technique". In: *Biochemistry* 32, pp. 9754–9762.

Balakin, A. G., L. Smith, and M. J. Fournier (1996). "The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions". In: *Cell* 6, pp. 823–834.

Ballarino, M., M. Morlando, F. Pagano, A. Fatica, and I. Bozzoni (2005). "The cotranscriptional assembly of snoRNPs controls the biosynthesis of H/ACA snoRNAs in *Saccharomyces cerevisiae*". In: *Molecular and Cellular Biology* 25, pp. 5396–5403.

Basak, A. and C. C. Query (2014). "A pseudouridine residue in the spliceosome core is part of the filamentous growth program in yeast". In: *Cell reports* 8, pp. 966–973.

Becker, H. F., Y. Motorin, R. J. Planta, and H. Grosjean (1997). "The yeast gene *YNL292w* encodes a pseudouridine synthase (Pus4) catalyzing the formation of psi55 in both mitochondrial and cytoplasmic tRNAs". In: *Nucleid acid research* 25, pp. 4493–4499.

Boccaletto, P., M. A. Machnicka, E. Purta, P. Piatkowski, B. Baginski, T. K. Wirecki, V. de Crecy-Lagard, R. Ross, P. A. Limbach, A. Kotter, M. Helm, and J. M. Bujnicki (2018). "MODOMICS a database of RNA modification pathways. 2017 update". In: *Nucleic Acids Research* 46, pp. 303–307.

Bortolin, M. L., P. Ganot, and T. Kiss (1999). "Elements essential for accumulation and function of small nucleolar RNAs directing site-specific pseudouridylation of ribosomal RNAs". In: *EMBO Journal* 18, pp. 457–469.

Bui, K. H., A. von Appen, A. L. DiGuilio, A. Ori, L. Sparks, M. Mackmull, T. Bock, W. Hagen, A. A. s Pons, J. S. Glavy, and M. Beck (2013). "Integrated structural analysis of the human nuclear pore complex scaffold". In: *Cell* 155, pp. 1233–1243.

Bushnell, D. A., K. D. Westover, R. E. Davis, and R. D. Kornberg (2004). "Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms". In: *Science* 303, pp. 983–988.

Cadwell, C., H. J. Yoon, Y. Zebarjadian, and J. Carbon (1997). "The yeast and nucleolar protein Cbf5p is involved in rRNA biosynthesis and interacts genetically with the RNA". In: *Molecular and Cellular Biology* 7, pp. 6175–6183.

Cai, G., T. Imasaki, K. Yamada, F. Cardelli, Y. Takagi, and F. J. Asturias (2010). "Mediator head module structure and functional interactions". In: *Nature Structural & Molecular Biology* 17, pp. 273–279.

Carlile, T. M., M. F. Rojas-Duran, B. Zinshteyn, H. Shin, K. M. Bartoli, and W. V. Gilbert (2014). "Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells". In: *Nature* 515.7525, pp. 143–146.

Casanal, A., A. Kumar, C. H. Hill, A. D. Easter, P. Emsley, G. Degliesposti, Y. Gordiyenko, B. Santhanam, J. Wolf, K. Wiederhold, G. L. Dornan, M. Skehel, C. V. Robinson, and L. A. Passmore (2017). "Architecture of eukaryotic mRNA 3-end processing machinery". In: *Science* 358, pp. 1056–1059.

Cavaille, J., M. Nicoloso, and J. P. Bachellerie (1996). "Targeted ribose methylation of RNA in vivo directed by tailored antisense RNA guides". In: *Nature* 383, pp. 732–735.

Chambers, M. C., B. Maclean, R. Burke, D. Amodei, D. L. Ruderman, S. Neumann, L. Gatto, B. Fischer, B. Pratt, and J. et al. Egertson (2012). "A cross-platform toolkit for mass spectrometry and proteomics". In: *Nature Biotechnology* 30, pp. 918–920.

Chanarat, S., C. Burkert-Kautzsch, D. M. Meinel, and K. Straesser (2012). "Prp19C and TREX: interacting to promote transcription elongation and mRNA export". In: *Transcription* 3, pp. 8–12.

Chang, Y. T. and G. H. Loew (1994). "Reaction mechanisms of formaldehyde with endocyclic imino groups of nucleic acid bases". In: *Journal of the American Chemical Society* 116, pp. 3548–3555.

Chapman, R. D., M. Heidemann, C. Hintermair, and D. Eick (2008). "Molecular evolution of the RNA polymerase II CTD". In: *Trends in Genetics* 24, pp. 289–296.

Chavez, S., T. Beilharz, A. G. Rondon, H. Erdjument-Bromage, P. Tempst, J. Q. Svejstrup, T. Lithgow, and A. Aguilera (2000). "A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in *Saccharomyces cerevisiae*." In: *EMBO Journal* 19, pp. 5824–5834.

Chavez, S., M. Garcia-Rubio, F. Prado, and A. Aguilera (2001). "Hpr1 is preferentially required for transcription of either long or G+C-Rich DNA sequences in *Saccharomyces cerevisiae*". In: *Molecular and Cellular Biology* 20, pp. 7054–7064.

Cho, E. J., M. S. Kobor, M. Kim, J. Greenblatt, and S. Buratowski (2001). "Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser2 of the RNA polymerase II C-terminal domain". In: *Genes and development* 15, pp. 3319–3329.

Cohn, W. E. and E. Volkin (1951). "Nucleoside-5'-phosphates from ribonucleic acid". In: *Nature* 167, pp. 483–484.

Cole, C. N. and J. J. Scarcelli (2006). "Transport of messenger RNA from the nucleus to the cytoplasm." In: *Current Opinion in Cell Biology* 18, pp. 299–306.

Corbett, K. S., D. K. Edwards, S. R. Leist, O. M. Abiona, S. Boyoglu-Barnum, R. A. Gillespie, S. Himansu, A. Schaefer, C. T. Ziwawo, A. T. DiPiazza, K. H. Dinnon, K. M. Morabito, R. S. B. Mascola, R. S. Baric, A. Carfi, and B. Graham (2020). "SARS-CoV-2 mRNA vaccine design enabled by prototype pathogen preparedness". In: *Nature* 586, pp. 567–571.

Corden, J. L. and C. J. Ingles (1992). "Carboxy-terminal domain of the largest subunit of eukaryotic RNA Polymerase II." In: *Transcriptional Regulation* 6, pp. 81–108.

Crick, F. (1970). "Central dogma of Molecular Biology". In: *Nature* 227, pp. 561–563.

Darzacq, X., N. Kittur, S. Roy, Y. Shav-Tal, R. H. Singer, and U. T. Meier (2006). "Stepwise RNP assembly at the site of H/ACA RNA transcription". In: *Journal of Cell Biology* 173, pp. 207–218.

Davis, D. R. (1995). "Stabilization of RNA stacking by pseudouridine". In: *Nucleic acids research* 23, pp. 5020–5026.

Davis, F. F. and F. W. Allen (1957). "Ribonucleic acids from yeast which contain a fifth nucleotide". In: *The Journal of biological chemistry* 227, pp. 907–915.

De Zoysa, M. D., G. Wu, R. Katz, and Y. T. Yu (2018). "Guide-substrate base-pairing requirement for box H/ACA RNA-guided RNA pseudouridylation". In: *RNA* 24, pp. 1106–1117.

Decatur, W. A. and M. J. Fournier (2002). "rRNA modifications and ribosome function". In: *Trends in biochemical sciences* 27, pp. 344–351.

Dehouck, Y., J. M. Kwasigroch, D. Gilis, and M. Rooman (2011). "PoPMuSiC 2.1: a web server for the estimation of protein stability changes upon mutation and sequence optimality". In: *BMC Bioinformatics* 12, pp. 1–12.

Ding, Y., S. Fan, S. Li, B. Feng, N. Gao, K. Ye, S. He, and M. Dong (2016). "Increasing the depth of mass-spectrometry-based structural analysis of protein complexes through the use of multiple cross-linkers". In: *Analytical Chemistry* 88, pp. 4461–4469.

Domiguez-Sanchez, M. S., S. Barroso, B. Gomez-Gonzalez, R. Luna, A. Aguilera, and C. E. Pearson (2011). "Genome instability and transcription elongation impairment in human cells depleted of THO/TREX". In: *PLoS Genetics* 7, pp. 1–17.

Donmez, G., K. Hartmuth, and R. Luhrmann (2004). "Modified nucleotides at the 5' end of human U2 snRNA are required for spliceosomal E-complex formation". In: *RNA* 10, pp. 1925–1933.

Duan, J., L. Li, J. Lu, W. Wang, and K. Ye (2009). "Structural mechanism of substrate RNA recruitment in H/ACA RNA-guided pseudouridine synthase". In: *Molecular Cell* 34, pp. 427–439.

Dunn, E. F., C. M. Hammell, C. A. Hodge, and C. N. Cole (2005). "Yeast poly(A)-binding protein, Pab1, and PAN, a poly(A) nuclease complex recruited by Pab1, connect mRNA biogenesis to export". In: *Genes and development* 19, pp. 90–103.

Exinger, F. and F. Lacroute (1992). "6-azauracil inhibition of GTP biosynthesis in *Saccharomyces cerevisiae*". In: *Curr. Genet* 22, pp. 9–11.

Eyboulet, F., C. Jeronimo, J. Cote, and F. Robert (2020). "The deubiquitylase Ubp15 couples transcription to mRNA export". In: *eLife* 9, pp. 1–21.

Eyler, D. E., M. K. Franco, Z. Batool, M. Z. Wu, M. L. Dubuke, M. Dobosz-Bartoszek, J. D. Jones, S. Polikanov, B. Roy, and S. Koutmou (2019). "Pseudouridinylation of mRNA coding sequences alters translation." In: *PNAS* 116, pp. 23068–23074.

Fasken, M. B., M. Stewart, and A. H. Corbett (2008). "Functional significance of the interaction between the mRNA-binding protein, Nab2, and the nuclear poreassociated protein, Mlp1, in mRNA export." In: *The Journal of Biological Chemistry* 283, pp. 27130–27143.

Fayet-Lebaron, E., V. Atzorn, Y. Henry, and T. Kiss (2009). "18S rRNA process-ing requires base pairings of snR30 H/ACA snoRNA to eukaryote-specific 18S sequences." In: *The EMBO journal* 28, pp. 1260–1270.

Fernandez, I., N. C. Leong, A. C. Kelley, G. Wu, Y. Yu, and V. Ramakrishnan (2013). "Unusual base pairing during the decoding of a stop codon by the ribosome". In: *Nature* 500, pp. 107–110.

Flaherty, S. M., P. Fortes, E. Izaurralde, I. W. Mattaj, and G. M. Gilmartin (1997). "Participation of the nuclear cap binding complex in pre-mRNA 3' processing". In: *PNAS* 94, pp. 11893–11898.

Fleming, A. M., A. Alenko, J. P. Kitt, A. M. Orendt, P. F. Flynn, J. M. Harris, and C. J. Burrows (2019). "Structural elucidation of bisulfite adducts to pseudouridine that result in deletion signatures during reverse transcription of RNA". In: *Journal of the American Chemical Society* 141, pp. 16450–16460.

Fortes, P., J. Kufel, M. Fornerod, M. Polycarpou Schwarz, D. La Fontaine, D. Tollervey, and I. W. Mattaj (1999). "Genetic and physical interactions involving the yeast nuclear cap-binding complex". In: *Molecular and Cellular Biology* 19, pp. 6543–6553.

Fraenkel-Conrat, H. (1954). "Reaction of nucleic acid with formaldehyde". In: *Biochim. Biophys. Acta* 15, pp. 307–309.

Fujisawa, T., K. Imai, Y. Tanaka, and K. Ogata (1979). "Studies on the protein components of 110S and total ribonucleoprotein particles of rat liver nucleoli". In: *The journal of Biochemistry* 8, pp. 277–286.

Ganot, P., M. L. Bortolin, and T. Kiss (1997). "Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs". In: *Cell* 89, pp. 799–809.

Ge, J. and Y. Yu (2013). "RNA pseudouridylation: new insights into an old modification". In: *Trends in Biochemical Sciences* 38, pp. 210–218.

GenScript. *GenScript Codon Usage Frequency Table(chart) Tool.* URL: `https://www.genscript.com/tools/codon-frequency-table`.

Gibson, D. G., L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison, and H. O. Smith (2009). "Enzymatic assembly of DNA molecules up to several hundred kilobases". In: *Nature methods* 6, pp. 343–345.

Gilbert, W. and C. Guthrie (2004). "The Glc7p Nuclear Phosphatase Promotes mRNA Export by Facilitating Association of Mex67p with mRNA". In: *Molecular Cell* 13.2, pp. 201–212.

Goldwasser, E. and R. L. Heinrikson (1966). *The biochemistry of pseudouridine IN: Progress in nucleic acid research and mol. biol.* Academic Press.

Gruellich, K. (2017). "Analysis of the potential interaction between a Pseudouridine Synthase with nuclear mRNA binding proteins". MA thesis. Institute of Biochemistry; Justus-Liebig-University, Giessen.

Guetsova, M. L., K. Lecoq, and B. Daignan-Fornier (1997). "The isolation and characterization of *Saccharomyces cerevisiae* mutants that constitutively express purine biosynthetic genes". In: *Genetics* 147, pp. 383–397.

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

Hackmann, A., H. Wu, U. M. Schneider, K. Meyer, K. Jung, and H. Krebber (2014). "Quality control of spliced mRNAs requires the shuttling SR proteins Gbp2 and Hrb1". In: *Nature communications* 5, p. 3125.

Harrington, K. M., I. A. Nazarenko, D. B. Dix, R. C. Thompson, and O. C. Uhlenbeck (1993). "*In vitro* analysis of translational rate and accuracy with an unmodified tRNA." In: *Biochemistry* 32, pp. 7617–7622.

Hausmann, S., B. Schwer, and S. Shuman (2004). "An encephalitozoon cuniculi ortholog of the RNA polymerase II carboxyl-terminal domain (CTD) serine phosphatase Fcp1". In: *Biochemistry* 43, pp. 7111–7120.

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

Heiss, M. and S. Kellner (2017). "Detection of nucleic acid modifications bychemical reagents". In: *RNA biology* 14, pp. 1166–1174.

Hengartner, C. J., V. E. Myer, S. M. Liao, C. J. Wilson, S. S. Koh, and R. R. Young (1998). "Temporal regulation of RNA polymerase II by Srb10 and Kin28 cyclin dependent kinases". In: *Molecular cell* 2, pp. 43–53.

Henras, A., Y. Henry, C. Bousquet-Antonelli, J. Noaillac-Depeyre, J. P. Gelugne, and M. Caizergues-Ferrer (1998). "Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs". In: *The EMBO journal* 17, pp. 7078–7090.

Henras, A., C. Dez, J. Noaillac-Depeyre, Y. Henry, and M. Caizergues-Ferrer (2001). "Accumulation of H/ACA snoRNPs depends on the integrity of the conserved central domain of the RNA-binding protein Nhp2". In: *Nucleid Acids Research* 29, pp. 2733–2746.

Henras, A. K., R. Capeyrou, Y. Henry, and M. Caizergues-Ferrer (2004). "Cbf5p, the putative pseudouridine synthase of H/ACA-type snoRNPs, can form a complex with Gar1p and Nop10p in absence of Nhp2p and box H/ACA snoRNAs". In: *RNA* 10, pp. 1704–1712.

Hershey, A. D. and M. Chase (1952). "Independet functions of viral protein and nucleic acid in growth of bacteriophage". In: *The Journal of General Physiology* 36, pp. 39–56.

Herzel, L., D. S. M. Ottoz, T. Alpert, and K. M. Neugebauer (2017). "Splicing and transcription touch base: co-transcriptional spliceosome assembly and function". In: *Nature reviews Molecular cell biology* 18, pp. 637–650.

Hinnebusch, A. G. (1997). "Translational Regulation of Yeast *GCN4*: A window on factors that control initiator-tRNA binding to the ribosome". In: *Journal of biological chemistry* 272, pp. 21661–21664.

Hinnebusch, A. G. and K. Natarajan (2002). "Gcn4p, a master regulator of gene expression, is controlled at multiple levels by diverse signals of starvation and stress". In: *Eukaryot Cell* 1, pp. 22–32.

Hirtreiter, A., G. E. Damsma, A. C. Cheung, D. Klose, D. Grohmann, E. Vojnic, A. C. Martin, P. Cramer, and F. Werner (2010). "Spt4/5 stimulates transcription elongation through the RNA polymerase clamp coiled-coil motif." In: *Nucleic Acids Research* 38, pp. 4040–4051.

Holstege, F. C., P. C. van der Vliet, and H. T. Timmers (1996). "Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH". In: *EMBO Journal* 15, pp. 1666–1677.

Huertas, P. and A. Aguilera (2003). "Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination." In: *Molecular Cell* 12, pp. 711–721.

Hurt, E., M. Luo, S. Roether, R. Reed, and K. Straesser (2004). "Cotranscriptional recruitment of the serine-arginine-rich (SR)-like proteins Gbp2 and Hrb1 to nascent mRNA via the TREX complex." In: *PNAS* 101, pp. 1858–1862.

Iglesias, N., E. Tutucci, C. Gwizdek, P. Vinciguerra, E. Dach, A. H. von Corbett, C. Dargemont, and F. Stutz (2010). "Ubiquitin mediated mRNP dynamics and surveillance prior to budding yeast mRNA export." In: *Genes & Development* 24, pp. 1927–1938.

Jack, K., C. Bellodi, D. M. Landry, R. O. Niederer, A. Meskauskas, S. Musalgaonkar, N. Kopmar, O. Krasnykh, A. M. Dean, S. R. Thompson, D. Ruggero, and J. D. Dinman (2011). "rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells". In: *Molecular Cell* 44, pp. 660–666.

Jerabek-Willemsen, M., C. J. Wienken, D. Braun, P. Baaske, and S. Duhr (2011). "Molecular Interaction Studies Using Microscale Thermophoresis". In: *Assay Drug Dev Technol.* 9, pp. 342–353.

Jiang, W., J. Middleton, H. Yoon, C. Fouquet, and J. Carbon (1993). "An essential yeast protein, Cbf5p, binds In vitro to centromeres and microtubules". In: *Molecular and Cellular Biology* 13.8, pp. 4884–4893.

Jiang, W., M. Y. Lim, H. J. Yoon, J. Thorner, G. S. Martin, and J. Carbon (1995). "Overexpression of the yeast *MCK1* protein kinase suppresses conditional mutations in centromere-binding protein genes *CBF2* and *CBF5*". In: *Molecular Genetics and Genomics* 246, pp. 360–366.

Jimeno, S., A. G. Rondon, R. Luna, and A. Aguilera (2002). "The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability". In: *EMBO* 21, pp. 3526–3535.

Jin, H., P. Loria, and P. B. Moore (2007). "Solution structure of an rRNA substrate bound to the pseudouridylation pocket of a box H/ACA snoRNA". In: *Molecular Cell* 26, pp. 205–215.

Johnson, S. A., G. Cubberley, and D. L. Bentley (2009). "Cotranscriptional recruitment of the mRNA export factor Yra1 by direct interaction with the 3' end processing factor Pcf11." In: *Molecular Cell* 33, pp. 215–226.

Karijolich, J. and Y. Yu (2011). "Converting nonsense codons into sense codons by targeted pseudouridylation". In: *Nature* 474, pp. 395–398.

Kariko, K., H. Muramatsu, F. A. Welsh, J. Ludwig, H. Kato, S. Akira, and D. Weissman (2008). "Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability". In: *Molecular Therapy : the Journal of the American Society of Gene Therapy* 16, pp. 1833–1840.

Kariko, K., H. Muramatsu, J. M. Keller, and D. Weissman (2012). "Increased erythropoiesis in mice injected with submicrogram quantities of pseudouridine-containing mRNA encoding erythropoietin". In: *Molecular Therapy : the Journal of the American Society of Gene Therapy* 20, pp. 948–953.

Katahira, J., K. Straesser, A. Podtelejnikov, M. Mann, J. U. Jung, and E. Hurt (1999). "The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human." In: *EMBO Journal* 18, pp. 2593–2609.

Kelly, E. K., D. P. Czekay, and U. Kothe (2019). "Base-pairing interactions between substrate RNA and H/ACA guide RNA modulate the kinetics of pseudouridylation, but not the

affinity of substrate binding by H/ACA small nucleolar ribonucleoproteins". In: *RNA* 25, pp. 1393–1404.

Kemmeren, P., K. Sameith, L. A. L. van de Pasch, J. J. Benschop, T. L. Lenstra, T. Margaritis, E. O Duibhir, E. Apweiler, S. van Wageningen, C. W. Ko, S. van Heesch, M. M. Kashani, G. Ampatziadis-Michailidis, M. O. Brok, N. A. C. H. Brabers, A. J. Miles, D. Bouwmeester, S. R. van Hooff, H. van Bakel, E. Sluiters, L. V. Bakker, B. Snel, P. Lijnzaad, D. van Leenen, M. J. A. Groot Koerkamp, and F. C. P. Holstege (2014). "Large-scale genetic perturbations reveal regulatory networks and an abundance of gene-specific repressors". In: *Cell* 3, pp. 740–752.

Khonsari, B. and R. Klassen (2020). "Impact of Pus1 pseudouridine synthase on specific decoding events in *Saccharomyces cerevisiae*". In: *Biomolecules* 10, pp. 1–16.

Kim, M., N. J. Krogan, L. Vasiljeva, O. J. Rando, E. Nedea, J. F. Greenblatt, and S. Buratowski (2004). "The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II". In: *Nature* 432, pp. 517–22.

Kim, T. K., R. H. Ebright, and D. Reinberg (2000). "Mechanism of ATP-dependent promoter melting by transcription factor IIH". In: *Science* 288, pp. 1418–1422.

Kim, Y., J. H. Geiger, S. Hahn, and P. B. Sigler (1993). "Crystal structure of a yeast TBP/TATA-box complex". In: *Nature* 365, pp. 512–520.

King, T. H., B. Liu, R. R. McCully, and M. J. Fournier (2003). "Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center". In: *Molecular Cell* 11, pp. 425–435.

Klein, B. J., D. Bose, K. J. Baker, Z. M. Yusoff, X. Zhang, and K. S. Murakami (2011). "RNA polymerase and transcription elongation factor Spt4/5 complex structure". In: *PNAS* 108, pp. 546–550.

Koo, B. K., C. J. Park, C. F. Fernandez, N. Chim, Y. Ding, G. Chanfreau, and J. Feigon (2011). "Structure of H/ACA RNP protein Nhp2p reveals cis/trans isomerization of a conserved proline at the RNA and Nop10 binding interface". In: *Journal of Molecular Biology* 411, pp. 927–942.

Koonin, E. V. (1996). "Pseudouridine synthases: four families of enzymes containing a putative uridine-binding motif also conserved in dUTPases and dCTP deaminase". In: *Nucleic Acids Research* 24, pp. 2411–2414.

Kostrewa, D., M. E. Zeller, K. J. Armache, M. Seizl, K. Leike, M. Thomm, and P. Cramer (2009). "RNA polymerase II-TFIIB structure and mechanism of transcription initiation". In: *Nature* 462, pp. 323–330.

Krishnamurthy, S., X. He, M. Reyes-Reyes, C. Moore, and M. Hampsey (2004). "Ssu72 is an RNA polymerase II CTD phosphatase". In: *Mol Cell* 14, pp. 387–394.

Kroustallaki, P., L. Lirussi, S. Carracedo, P. You, Q. Y. Esbensen, A. Goetz, L. Jobert, L. Alsoe, P. Saetrom, S. Gagos, and H. Nilsen (2019). "*SMUG1* promotes telomere maintenance through telomerase RNA processing". In: *Cell Reports* 28, pp. 1690–1702.

Kuehner, J. N., E. L. Pearson, and C. Moore (2011). "Unravelling the means to an end: RNA polymerase II transcription termination." In: *Nature Review Molecular Cell Biology* 12, pp. 283–294.

Lane, B. G., J. Ofengand, and M. Gray (1995). "Pseudouridine and O-methylated nucleosides. Significance of their selective occurrence in rRNA domains that function in ribosome-catalyzed synthesis of the peptide bonds in proteins". In: *Biochimie* 77, pp. 7–15.

Lange, T. S., M. Ezrokhi, F. Amaldi, and S. A. Gerbi (1999). "Box H and box ACA are nucleolar localization elements of U17 small nucleolar RNA". In: *Molecular biology of the cell* 10, pp. 3877–3890.

Lariviere, L., S. Geiger, S. Hoeppner, S. Roether, K. Straesser, and P. Cramer (2006). "Structure and TBP binding of the Mediator head subcomplex Med8-Med18-Med20". In: *Nat Struct Mol Biol* 13, pp. 895–901.

Lei, Z. and C. Yi (2017). "A radiolabeling-free, qPCR-based method for locus-specific pseudouridine detection". In: *Angewandte Chemie International Edition* 5, pp. 1–6.

Leitner, A., T. Walzthoeni, and R. Aebersold (2014). "Lysine-specific chemical cross-linking of protein complexes and identification of cross-linking sites using LC-MS/MS and the xQuest/xProphet software pipeline." In: *Nature Protocols* 9, pp. 120–137.

Li, L. and K. Ye (2006). "Crystal structure of an H/ACA box ribonucleoprotein particle". In: *Nature* 443, pp. 302–307.

Li, S., J. Duan, D. Li, B. Yang, M. Dong, and K. Ye (2011). "Reconstitution and structural analysis of the yeast box H/ACA RNA-guided pseudouridine synthase". In: *Genes & Development* 25, pp. 2409–2421. ISSN: 0890-9369.

Li, X., S. Ma, and C. Yi (2015a). "Pseudouridine Chemical Labeling and Profiling". In: *Methods in Enzymology* 560, pp. 247–272.

Li, Y., P. Zhu, S. Ma, J. Song, J. Bai, F. Sun, and C. Yi (2015b). "Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome". In: *Nature Chemical Biology* 11, pp. 592–597.

Liang, X., Q. Lui, and M. J. Fournier (2009). "Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing". In: *RNA* 15, pp. 1716–1728.

Lidschreiber, M., K. Leike, and P. Cramer (2013). "Cap completion and C-terminal repeat domain kinase recruitment underlie the initiation-elongation transition of RNA polymerase II". In: *Molecular and cellular biology* 33, pp. 3805–3816.

Lindstrom, D. L., S. L. Squazzo, N. Muster, T. A. Burckin, K. C. Wachter, C. A. Emigh, J. A. McCleery, J. R. Yates, and G. A. Hartzog (2003). "Dual roles for Spt5 in pre mRNA processing and transcription elongation revealed by identification of Spt5- associated proteins." In: *Molecular Cell Biology* 23, pp. 1368–1378.

Liu, Y., L. Warfield, C. Zhang, J. Luo, J. Allen, W. H. Lang, J. Ranish, K. M. Shokat, and S. Hahn (2009). "Phosphorylation of the transcription elongation factor Spt5 by yeast Bur1 kinase stimulates recruitment of the PAF complex". In: *Molecular and cellular biology* 29, pp. 4852–4863.

Livak, K. J. and T. D. Schmittgen (2001). "Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method." In: *Methods* 24, pp. 402–408.

Logan, J., E. Falck-Pedersen, J. E. Darnell, and T. Shenk (1987). "A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse beta maj-globin gene". In: *PNAS* 84, pp. 8306–8310.

Love, M. I., W. Huber, and S. Anders (2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2". In: *Genome biology* 15, pp. 1–21.

Lovejoy, A. F., D. P. Riordan, and P. Brown (2014). "Transcriptome-wide mapping of pseudouridines: pseudouridine synthases modify specific mRNAs in *S. cerevisiae*". In: *PLOS ONE* 9, e110799.

Lucchini, G., A. G. Hinnebusch, C. Chen, and G. R. Fink (1984). "Positive regulatory interactions of the *HIS4* gene of *Saccharomyces cerevisiae*." In: *Mol Cell Biol* 4, pp. 1326–1333.

Luo, W., A. W. Johnson, and D. L. Bentley (2006). "The role of Rat1 in coupling mRNA 3'-end processing to transcription termination: Implications for a unified allosteric-torpedo model". In: *Genes and Development* 20, pp. 954–965.

MacKellar, A. L. and A. L. Greenleaf (2011). "Cotranscriptional association of mRNA export factor Yra1 with C-terminal domain of RNA polymerase II." In: *Journal of Biological Chemistry* 286, pp. 36385–36395.

Maden, B. E. (1990). "The numerous modified nucleotides in eukaryotic ribosomal RNA." In: *Prog Nucleic Acid Res Mol Biol.* 39, pp. 241–303.

Mandel, C. R., Y. Bai, and L. Tong (2008). "Protein factors in pre-mRNA 3' -end processing." In: *Cellular and Molecular Life Sciences* 65, pp. 1099–1122.

Maniatis, T. and R. Reed (2002). "An extensive network of coupling among gene expression machines". In: *Nature* 416, pp. 499–506.

Mao, X., B. Schwer, and S. Shuman (1995). "Yeast mRNA cap methyltransferase is a 50-kilodalton protein encoded by an essential gene". In: *Molecular and cellular biology* 15, pp. 4167–4174.

Marchand, V., F. Pichot, P. Neybecker, L. Ayadi, V. Bourguignon-Igel, L. Wacheul, D. L. J. Lafontaine, A. Pinzano, M. Helm, and Y. Motorin (2020). "HydraPsiSeq: a method for systematic and quantitative mapping of pseudouridines in RNA". In: *Nucleic Acids Research* 48, e110 1–14.

Margeat, E., A. N. Kapanidis, P. Tinnefeld, Y. Wang, J. Mukhopadhyay, R. H. Ebright, and S. Weiss (2006). "Direct observation of abortive initiation and promoter escape within single immobilized transcription complexes". In: *Biophysical journal* 90, pp. 1419–1431.

Maris, C., C. Dominguez, and F. H. Allain (2005). "The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression". In: *FEBS Journal* 272, pp. 2118–2131.

Martinez, N. M., A. Su, M. C. Burns, J. K. Nussbacher, C. Schaening, S. Sathe, G. W. Yeo, and W. V. Gilbert (2022). "Pseudouridine synthases modify human pre-mRNA co-transcriptionally and affect pre-mRNA processing". In: *Molecular Cell* 82, pp. 645–659.

Martinez-Rucobo, F. W., S. Sainsbury, A. C. Cheung, and P. Cramer (2011). "Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity". In: *The EMBO journal* 30, pp. 1302–1310.

Mason, P. M. and K. Struhl (2005). "Distinction and relationship between elongation rate and processivity of RNA polymerse II *in vivo*". In: *Molecular Cell* 17, pp. 831–840.

Massenet, S., A. Mougin, and C. Branlant (1998). "Posttranscriptional modifications in the U small nuclear RNAs". In: *Modification and Editing of RNA* 18, pp. 201–227.

Mayer, A., A. Schreieck, M. Lidschreiber, K. Leike, D. E. Martin, and P. Cramer (2012). "The *SPT5* C-terminal region recruits yeast 3' RNA cleavage factor I". In: *Molecular Cell Biology* 32, pp. 1321–1331.

Meier, U. T. and G. Blobel (1994). "NAP57, a mammalian nucleolar protein with a putative homolog in yeast and bacteria". In: *The Journal of Cell Biology* 127, pp. 1505–1514.

Meinel, D. M. and K. Straesser (2015). "Co-transcriptional mRNP formation is coordinated within a molecular mRNP packaging station in *S. cerevisiae*". In: *Bioessays* 37, pp. 666–677.

Meinel, D. M., C. Burkert-Kautzsch, A. Kieser, E. O Duibhir, M. Siebert, A. Mayer, P. Cramer, J. Soeding, F. C. P. Holstege, and K. Straesser (2013). "Recruitment of TREX to

the transcription machinery by its direct binding to the phospho-CTD of RNA polymerase II". In: *PLOS Genetics* 9, e1003914.

Metz, B., G. F. A. Kersten, P. Hoogerhout, H. F. Brugghe, H. A. M. Timmermans, A. de Jong, H. Meiring, J. ten Hove, W. E. Hennink, D. J. A. Crommelin, and W. Jiskoot (2004). "Identification of formaldehyde-induced modifications in proteins: reactions with model peptides". In: *Journal Biol. Chem.* 279, pp. 6235–6243.

Mitchell, S. F., S. Jain, M. She, and R. Parker (2013). "Global analysis of yeast mRNPs." In: *Structural Molecular Biology* 20, pp. 127–133.

Morawska, M. and H. D. Ulrich (2013). "An expanded tool kit for the auxin-inducible degronsystem in budding yeast". In: *Yeast* 30, pp. 341–351.

Mosley, A. L., S. G. Pattenden, M. Carey, S. Venkatesh, J. M. Gilmore, L. Florens, J. L. Workman, and M. P. Washburn (2009). "Rtr1 is a CTD phosphatase that regulates RNA polymerase II during thetransition from serine 5 to serine 2 phosphorylation". In: *Molecular cell* 34, pp. 168–178.

Motorin, Y., G. Keith, C. Simon, D. Foiret, G. Simos, E. Hurt, and E. Grosjean (1998). "The yeast tRNA:pseudouridine synthase Pus1p displays a multisite substrate specificity". In: *RNA* 4, pp. 856–869.

Muhlbacher, W., S. Sainsbury, M. Hemann, M. Hantsche, S. Neyer, F. Herzog, and P. Cramer (2014). "Conserved architecture of the core RNA polymerase II initiation complex". In: *Nature communication* 5, pp. 4310–4315.

Nakamoto, M. A., A. F. Lovejoy, A. M. Cygan, and J. C. Boothroyd (2017). "mRNA pseudouridylation affects RNA metabolism in the parasite Toxoplasma gondii". In: *RNA* 23, pp. 1834–1849.

Narain, A., P. Bhandare, B. Adhikari, S. Backes, M. Eilers, L. Doelken, A. Schlosser, F. Erhard, A. Baluapuri, and E. Wolf (2021). "Targeted protein degradation reveals a direct role of Spt5 in RNAPII elongation and termination". In: *Mol Cell* 81, pp. 3110–3127.

Natarajan, K., M. R. Meyer, B. M. Jackson, D. Slade, C. Roberts, A. G. Hinnebusch, and M. Marton (2001). "Transcriptional profiling shows that Gcn4p is a master regulator of gene expression during amino acid starvation in yeast". In: *Mol Cell Biol* 21, pp. 4347–4368.

Newby, M. I. and N. L. Greenbaum (2002). "Sculpting of the spliceosomal branch site recognition motif by a conserved pseudouridine". In: *Nature Structural Biology* 9, pp. 958–965.

Ni, J., A. L. Tien, and M. J. Fournier (1997). "Small nucleolar RNAs direct sitespecific synthesis of pseudouridine in ribosomal RNA". In: *Cell* 89, pp. 565–573.

Nishimura, K., T. Fukagawa, H. Takisawa, T. Kakimoto, and M. Kanemaki (2009). "An auxin-based degron system for the rapid depletionof proteins in nonplant cells". In: *Nature Methods* 6, pp. 917–923.

Noble, K. N., E. J. Tran, A. R. Alcazar-Roman, C. A. Hodge, C. N. Cole, and R. S. Wente (2011). "The Dbp5 cycle at the nuclear pore complex during mRNA export II: nucleotide cycling and mRNP remodeling by Dbp5 are controlled by Nup159 and Gle1". In: *Genes and Development* 25, pp. 1065–1077.

Nojima, T., T. Hirose, H. Kimura, and M. Hagiwara (2007). "The interaction between cap-binding complex and RNA export factor is required for intronless mRNA export". In: *The Journal of biological chemistry* 282, pp. 15645–15651.

Oeffinger, M. and D. Zenklusen (2012). "To the pore and through the pore: A story of mRNA export kinetics". In: *Biochimica et Biophysica Acta (BBA)* 1819, pp. 494–506.

Ozdilek, B. A., V. F. Thompson, N. S. Ahmed, C. I. White, R. T. Batey, and J. C. Schwartz (2017). "Intrinsically disordered RGG/RG domains mediate degenerate specificity in RNA binding". In: *Nucleic acids research* 45, pp. 7984–7996.

Pabis, M., N. Neufeld, M. C. Steiner, T. Bojic, Y. Shav-Tal, and K. M. Neugebauer (2013). "The nuclear cap-binding complex interacts with the U4 U6 U5 tri-snRNP and promotes spliceosome assembly in mammalian cells". In: *RNA* 19, pp. 1054–1063.

Parenteau, J., M. Durand, S. Veronneau, A. Lacombe, G. Morin, V. Guerin, B. Cecez, J. Gervais-Bird, C. Koh, D. Brunelle, R. J. Wellinger, B. Chabot, and S. A. Elela (2008). "Deletion of many yeast introns reveals a minority of genes that require splicing for function". In: *Molecular Biology of the Cell* 19, pp. 1932–1941.

Phatnani, H. P. and A. E. Greenleaf (2006). "Phosphorylation and functions of the RNA polymerase II CTD". In: *Genes & Development* 20, pp. 2922–2936.

Polack, F. P., S. J. Thomas, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, J. L. Perez, S. Uenal, D. B. Tresnan, S. Mather, P. R. Dormitzer, U. Sahin, K. U. Jansen, and W. C. Gruber (2020). "Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine". In: *The New England Journal of Medicine* 383, pp. 2603–2615.

Puig, O., F. Caspary, G. Rigaut, R. Rutz, E. Bouveret, E. Bragado-Nilsson, M. Wilm, and B. Séraphin (2001). "The tandem affinity purification (TAP) method: a general procedure of protein complex purification". In: *Methods* 24, pp. 218–229.

Pühringer, T., U. Hohmann, L. Fin, B. Pacheco-Fiallos, U. Schellhaas, J. Brennecke, and C. Plaschka (2020). "Structure of the human core transcription-export complex reveals a hub for multivalent interactions". In: *eLife* 9, pp. 1–21.

Ramamurthy, V., S. L. Swann, J. L. Paulson, C. J. Spedaliere, and E. G. Mueller (1999). "Critical aspartic acid residues in pseudouridine synthases". In: *Journal Biol Chem.* 274, pp. 22225–22230.

Richards, L. A., A. Kumari, K. Knezevic, J. A. Thoms, G. von Jonquieres, C. E. Napier, Z. Ali, R. OBrien, J. Marks-Bluth, M. F. Maritz, H. A. Pickett, J. Morris, J. E. Pimanda, and K. L. MacKenzie (2020). "*DKC1* is a transcriptional target of GATA1 and drives upregulation of telomerase activity in normal human erythroblasts". In: *Haematologica* 105, pp. 1517–1526.

Riles, L., R. J. Shaw, M. Johnston, and D. Reines (2004). "Large-scale screening of yeast mutants for sensitivity to the IMP dehydrogenase inhibitor 6-azauracil". In: *Yeast* 21, pp. 241–248.

Rintala-Dempsey, A. C. and Ute Kothe (2017). "Eukaryotic stand-alone pseudouridine synthases - RNA modifying enzymes and emerging regulators of gene expression?" In: *RNA biology* 14, pp. 1185–1196.

Rodriguez, C. R., T. Takagi, E. J. Cho, and S. Buratowski (1999). "A *Saccharomyces cerevisiae* RNA 5'-triphosphatase related to mRNA capping enzyme". In: *Nucleic acids research* 27, pp. 2181–2188.

Roundtree, I. A., G. Luo, Z. Zhang, X. Wang, T. Zhou, Y. Cui, J. Sha, X. Huang, L. Guerrero, P. Xie, E. He, B. Shen, and C. He (2017). "YTHDC1 mediates nuclear export of N 6-methyladenosine methylated mRNAs". In: *eLIFE* 6, pp. 1–28.

Santos-Rosa, H., H. Moreno, G. Simos, A. Segref, B. Fahrenkrog, N. Pante, and E. Hurt (1998). "Nuclear mRNA export requires complex formation between Mex67p and Mtr2p at the nuclear pores". In: *Molecular and cellular biology* 18, pp. 6826–6838.

Schattner, P., W. A. Decatur, C. A. Davis, M. l. Ares, M. J. Fournier, and T. M. Lowe (2004). "Genome-wide searching for pseudouridylation guide snoRNAs: analysis of the *Saccharomyces cerevisiae* genome". In: *Nucleic acids research* 32.14, pp. 4281–4296.

Schreick, A. E. (2013). "Role of the RNA polymerase II C-terminal domain in transcription termination and function of Spt5 in 3' RNA processing factor recruitment". PhD thesis. Ludwig-Maximilians-University, Munich.

Schwartz, S. S., D. A. Berstein, M. R. Mumbach, M. Jovanovic, R. H. Herbst, X. L. Brian, J. M. Engreitz, M. Guttmann, R. Satiija, E. S. Lander, and et al. (2014). "Transcriptome - wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA". In: *Cell* 159, pp. 148–162.

Schwer, B., X. Mao, and S. Shuman (1998). "Accelerated mRNA decay in conditional mutants of yeast mRNA capping enzyme". In: *Nucleic acids research* 26, pp. 2050–2057.

Schwer, B., H. Erdjument-Bromage, and S. Shuman (2011). "Composition of yeast snRNPs and snoRNPs in the absence of trimethylguanosine caps reveals nuclear cap binding protein as a gained U1 component implicated in the cold-sensitivity of tgs1". In: *Nucleic Acids Research* 39, pp. 6715–6728.

Segref, A., K. Sharma, V. Doye, A. Hellwig, J. Huber, R. Luehrmann, and E. Hurt (1997). "Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)+ RNA and nuclear pores". In: *EMBO journal* 16, pp. 3256–3271.

Sen, R., P. Barman, A. Kaja, J. Ferdoush, S. Lahudkar, A. Roy, and S. R. Bhaumik (2019). "Distinct functions of the Cap-binding complex in stimulation of nuclear mRNA Export". In: *Mol Cell Biol* 39, pp. 1–20.

Shibagaki, Y., N. Itoh, H. Yamada, S. Nagata, and K. Mizumoto (1992). "mRNA capping enzyme: Isolation and characterization of the gene encoding mRNA guanylytransferase subunit from *Saccharomyces cerevisiae*". In: *Journal of Biological Chemistry* 267, pp. 9521–9528.

Shukla, S., J. C. Schmidt, K. C. Goldfarb, T. R. Cech, and R. Parker (2016). "Inhibition of telomerase RNA decay rescues telomerase deficiency caused by dyskerin or PARN defects". In: *Nature Structural and Molecular Biology* 23, pp. 286–292.

Sikorski, R. S. and P. Hieter (1989). "A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*". In: *Genetics* 122, pp. 19–27.

Snider, C. E., A. D. Stephens, J. G. Kirkland, O. Hamdani, R. T. Kamakaka, and B. Kerry (2014). "Dyskerin, tRNA genes, and condensin tether pericentric chromatin to the spindle axis in mitosis". In: *Journal of Cell Biology* 207, pp. 189–199.

Sogaard, T. M. and J. Q. Svejstrup (2007). "Hyperphosphorylation of the C-terminal repeat domain of RNA polymerase II facilitates dissociation of its complex with mediator". In: *The Journal of biological chemistry* 282, pp. 14113–14120.

Sonenberg, N. and A. G. Hinnebusch (2009). "Regulation of translation initiation in eukaryotes: mechanisms and biological targets". In: *Cell* 136, pp. 731–745.

Straeser, K. and E. Hurt (2001). "Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p". In: *Nature* 11, pp. 648–652.

Straesser, K. and E. Hurt (2000). "Yra1p, a conserved nuclear RNA - binding protein, interacts directly with Mex67p and is required for mRNA export". In: *EMBO* 19, pp. 410–420.

Straesser, K., J. Bassler, and E. Hurt (2000). "Binding of the Mex67p/Mtr2p heterodimer to Fxfg, Glfg, and Fg Repeat nucleoporins is essential for nuclear mRNA export." In: *The Journal of Cell Biology* 150, pp. 695–706.

Straesser, K., S. Masuda, P. Mason, J. Pfannstiel, M. Oppizzi, S. Rodriguez-Navarro, A. G. Rondon, A. Aguilera, K. Struhl, R. Reed, and E. Hurt (2002). "TREX is a conserved complex coupling transcription with messenger RNA export". In: *Nature* 417, pp. 304–307.

Swanson, M. S., E. A. Malone, and F. Winston (1991). "*SPT5*, an essential gene important for normal transcription in *Saccharomyces cerevisiae*, encodes an acidic nuclear protein with a carboxy-terminal repeat". In: *Molecular and cellular biology* 11, pp. 3009–3019.

Tian, B. and J. H. Graber (2012). "Signals for pre-mRNA cleavage and polyadenylation". In: *Wiley interdisciplinary reviews* 3, pp. 385–396.

Tollervey, D. (1987). "A yeast small nuclear RNA is required for normal processing of pre-ribosomal RNA". In: *EMBO Journal* 6, pp. 4169–4175.

Tollervey, D. and T. Kiss (1997). "Function and synthesis of small nucleolar RNAs". In: *Current Opinion in Cell Biology* 9, pp. 337–342.

Tran, E. J., Y. Zhou, A. H. Corbett, and S. R. Wente (2007). "The DEAD-box protein Dbp5 controls mRNA export by triggering specific RNA:protein remodeling events". In: *Mol Cell* 28, pp. 850–859.

Trucks, S., G. Hanspach, and M. Hengesbach (2021). "Eukaryote specific RNA and protein features facilitate assembly and catalysis of H/ACA snoRNPs". In: *Nucleic Acids Research* 49, pp. 4629–4642.

Tsai, F. T. and P. B. Sigler (2000). "Structural basis of preinitiation complex assembly onhuman pol II promoters". In: *EMBO* 19, pp. 25–36.

Veerareddygari, G. R., S. K. Singh, and E. G. Mueller (2016). "The pseudouridine synthases proceed through a glycal intermediate". In: *Journal of the American Chemical Society* 138, pp. 7852–7855.

Viphakone, N., F. Voisinet-Hakil, and L. Minvielle-Sebastia (2008). "Molecular dissection of mRNA poly(A) tail length control in yeast". In: *Nucleic Acids Research* 36, pp. 2418–2433.

Wahl, M. C., C. L. Will, and R. Luehrmann (2009). "The spliceosome: design principles of a dynamic RNP machine". In: *Cell* 136, pp. 701–718.

Walbott, H., R. Machado-Pinilla, D. Liger, M. Blaud, S. Rety, P. N. Grozdanov, K. Godin, H. van Tilbeurgh, G. Varani, U. T. Meier, and N. Leulliot (2011). "The H/ACA RNP assembly factor Shq1 functions as an RNA mimic". In: *Genes and development* 25, pp. 2398–2408.

Walzthoeni, T., M. Claassen, A. Leitner, F. Herzog, S. Bohn, F. Forster, M. Beck, and R. Aebersold (2012). "False discovery rate estimation for cross-linked peptides identified by mass spectrometry." In: *Nature methods* 9, pp. 901–903.

Wang, C. and U. T. Meier (2004). "Architecture and assembly of mammalian H/ACA small nucleolar and telomerase ribonucleoproteins". In: *EMBO* 23, pp. 1857–1867.

Watkins, N. J., A. Gottschalk, G. Neubauer, B. Kastner, P. Fabrizio, M. Mann, and R. Luehrmann (1998). "Cbf5p, a potential pseudouridine synthase and Nhp2p, a putative and RNA-binding protein are present together with Gar1p in all H BOX/ACA-motif snoRNPs constitute a common bipartite and structure". In: *RNA* 4, pp. 1549–1568.

Werner-Allen, J. W., C. Lee, P. Liu, N. I. Nicely, S. Wang, A. L. Greenleaf, and P. Zhou (2011). "cis-proline-mediated Ser(P)5 dephosphorylation by the RNA polymerase II C-terminal domain phosphatase Ssu72". In: *The Journal of biological chemistry* 286, pp. 5717–5726.

Will, C. L. and R. Luehrmann (2011). "Spliceosome structure and function". In: *Cold Spring Harbor Perspectives in Biology* 3, a003707.

Wittschieben, B. O., G. Otero, T. de Bizemont, J. Fellows, H. Erdjument-Bromage, R. Ohba, Y. Li, C. D. Allis, P. Tempst, and J. Q. Svejstrup (1999). "A novel histone acetyltransferase is an integral subunit of elongating RNA polymerase II holoenzyme". In: *Mol Cell* 4, pp. 123–128.

Wu, G., M. Xiao, C. Yang, and Y. T. Yu (2011). "U2 snRNA is inducibly pseudouridylated at novel sites by Pus7p and snR81 RNP". In: *EMBO Journal* 30, pp. 79–89.

Wu, H. and J. Feigon (2007). "H/ACA small nucleolar RNA pseudouridylation pockets bind substrate RNA to form three-way junctions that position the target U for modification". In: *PNAS* 104, pp. 6655–6660.

Xie, Y., B. P. Clarke, Y. J. Kim, A. L. Ivey, P. S. Hill, Y. Shi, and Y. Ren (2021a). "Cryo-EM structure of the yeast TREX complex and coordination with the SR-likeprotein Gbp2". In: *eLife* 10, pp. 1–18.

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

Yang, C., D. S. McPheeters, and Y. T. Yu (2005). "Psi35 in the branch site recognition region of U2 small nuclear RNA is important for pre-mRNA splicing in *Saccharomyces cerevisiae*." In: *J Biol Chem.* 280, pp. 6655–6662.

Yang, P. K., G. Rotondo, T. Porras, P. Legrain, and G. Chanfreau (2002). "The Shq1p.Naf1p complex is required for box H/ACA small nucleolar ribonucleoprotein particle biogenesis". In: *The journal of Biological Chemistry* 277, pp. 45235–45242.

Yang, X., J. Duan, S. Li, P. Wang, S. Ma, K. Ye, and X. S. Zhao (2012). "Kinetic and thermodynamic characterization of the reaction pathwayof box H/ACA RNA-guided pseudouridine formation". In: *Nucleic Acids Research* 40, pp. 10925–10936.

Yoh, S. M., H. Cho, L. Pickle, R. M. Evans, and K. A. Jones (2007). "The Spt6 *SH2* domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export". In: *Genes and development* 21, pp. 160–174.

Yu, G., L. Wang, Y. Han, and Q. He (2012). "clusterProfiler: an R package for comparing biological themes among gene clusters". In: *Omics : a journal of integrative biology* 16, pp. 284–287.

Zaringhalam, M. and F. N. Papavasiliou (2016). "Pseudouridylation meets next-generation sequencing". In: *Methods* 107, pp. 63–72.

Zebarjadian, Y., T. King, M. J. Fournier, L. Clarke, and J. Carbon (1999). "Point mutations in yeast *CBF5* can abolish *in vivo* pseudouridylation of rRNA". In: *Molecular Cell Biology* 19, pp. 7461–7472.

Zhang, Y., C. Xu, D. Gu, M. Wu, B. Yan, Z. Xu, Y. Wang, and H. Liu (2017). "H/ACA Box small nucleolar RNA 7A promotes the self-renewal of humanumbilical cord mesenchymal stem cells". In: *Stem Cells* 35, pp. 222–235.

Zhou, J., J. Wan, X. E. Shu, Y. Mao, X. M. Liu, X. Yuan, X. Zhang, M. E. Hess, J. C. Bruening, and S. B. Qian (2018). "N6 -methyladenosineguides mRNA alternative translation during integrated stress response". In: *Mol. Cell* 69, pp. 1–29.

Appendix

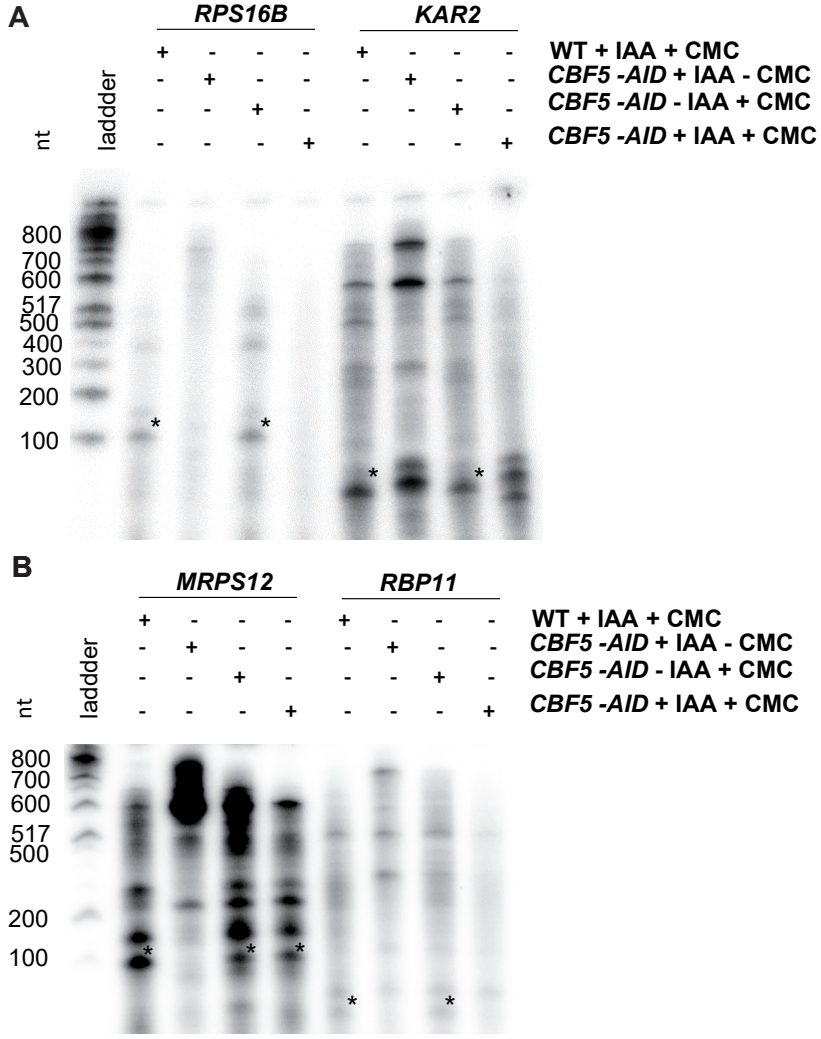

Figure 31: Identification of Cbf5-dependent pseudouridylation at *RPS16B, KAR2 and RBP11*, whereas the pseudouridylation at *MRPS12* is not Cbf5-dependent Primer extension after CMC treatment of CMC-treated (+ CMC) and CMC-untreated RNA (- CMC) in the wild-type strain (WT) and the depleted (*CBF5-AID* +IAA) or undepleted (*CBF5-AID* -IAA) *CBF5* depletion strain to identify Cbf5-dependent pseudouridylation sites. Figures show autoradiographs of the Primer Extension at *KAR2*/*RPS16B* (**A**) and *MRPS12*/*RPB11* (**B**). The asterisks labeled bands show pseudouridylation-sites as a radioactive signal of the cDNA after the discontinuation of the reverse transcription. nt: nucleotides

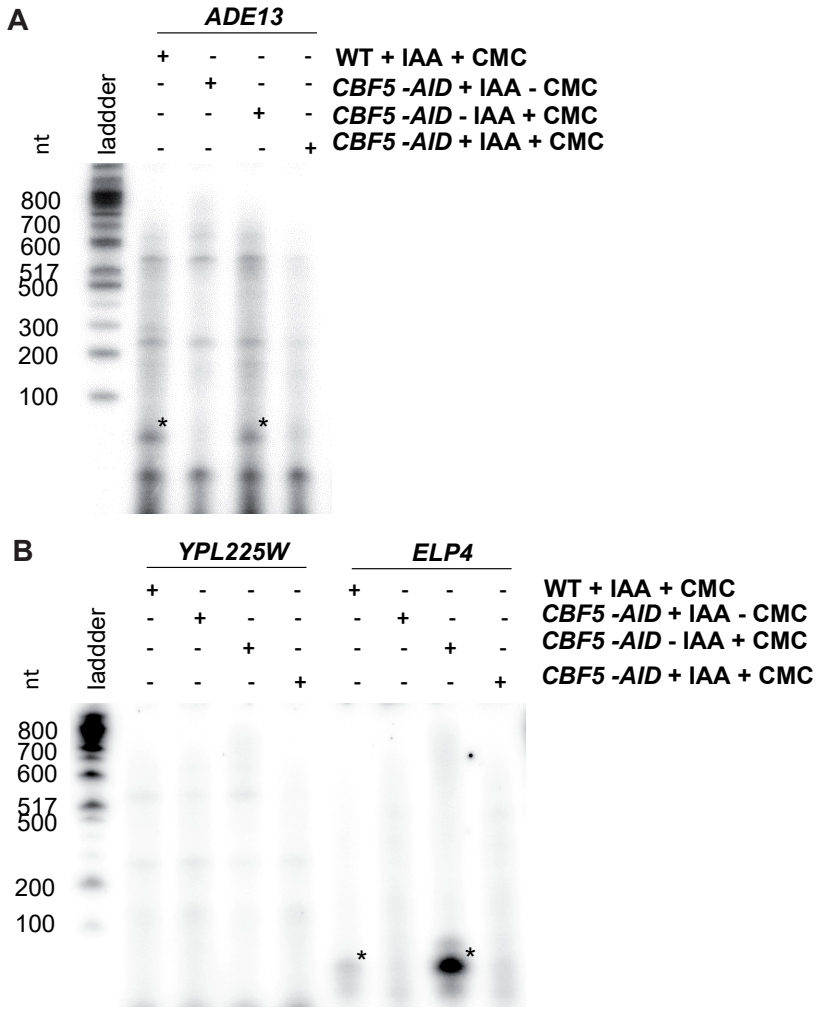

Figure 32: Identification of Cbf5-dependent pseudouridylation at *ADE13 and ELP4*, whereas *YPL225W* does not have a pseudouridylation site. Primer extension after CMC treatment of CMC-treated (+ CMC) and CMC-untreated RNA (- CMC) in the wild-type strain (WT) and the depleted (*CBF5-AID* +IAA) or undepleted (*CBF5-AID* -IAA) *CBF5* depletion strain to identify Cbf5-dependent pseudouridylation. Figures show autoradiographs of the Primer Extension at *ADE13* (**A**) and *YPL225W*/*ELP4* (**B**). The asterisks labeled bands show pseudouridylation-sites as a radioactive signal of the cDNA after the discontinuation of the reverse transcription. nt: nucleotides

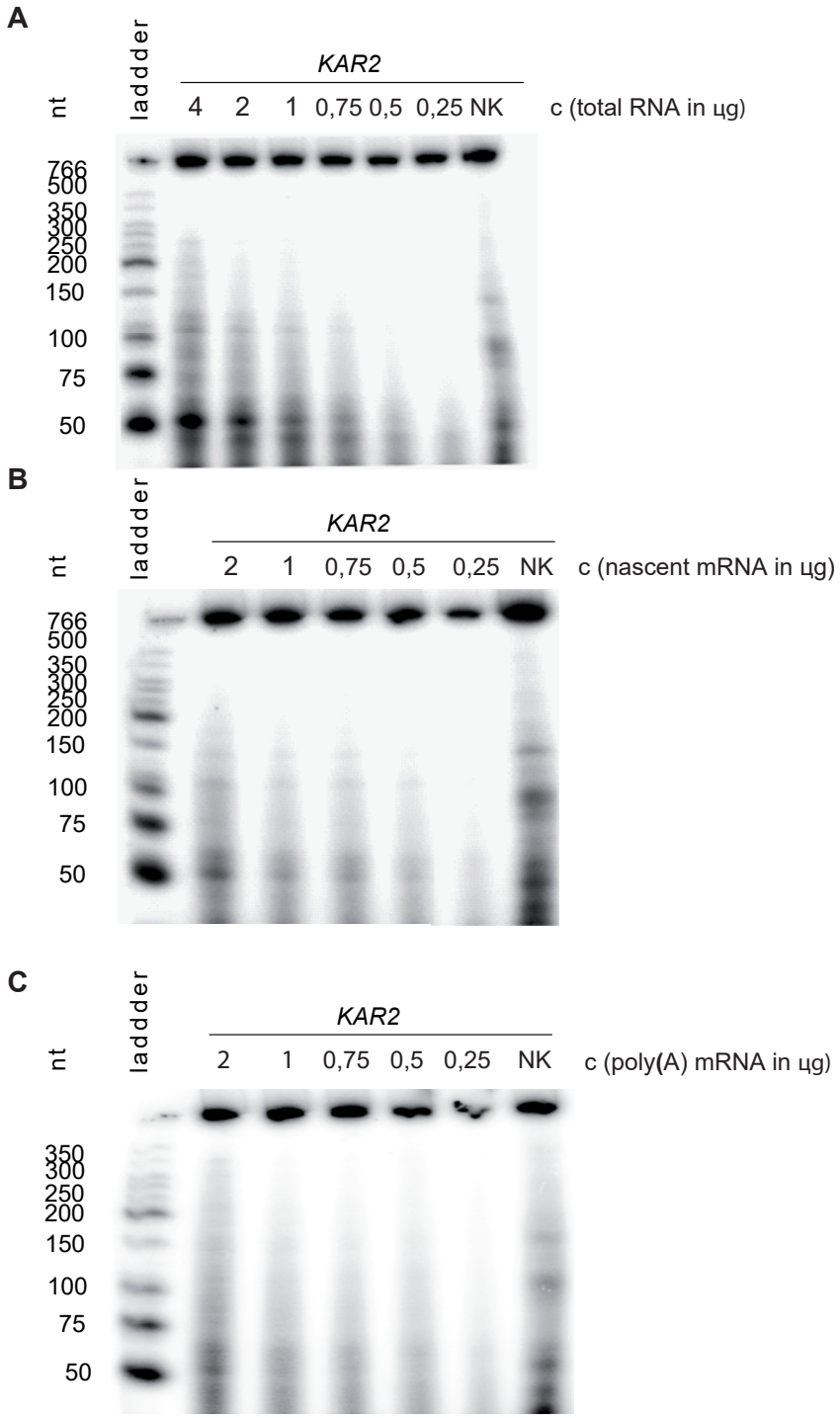

Figure 33: Primer extension after CMC treatment of different concentrations of total RNA, nascent mRNA and poly(A) mRNA to identify an RNA concentration in the exponential growth phase that was used for following experiments. Figures show autoradiographs of the Primer Extension at *KAR2* of total RNA (**A**), nascent mRNA (**B**) and poly(A) mRNA (**C**). nt: nucleotides

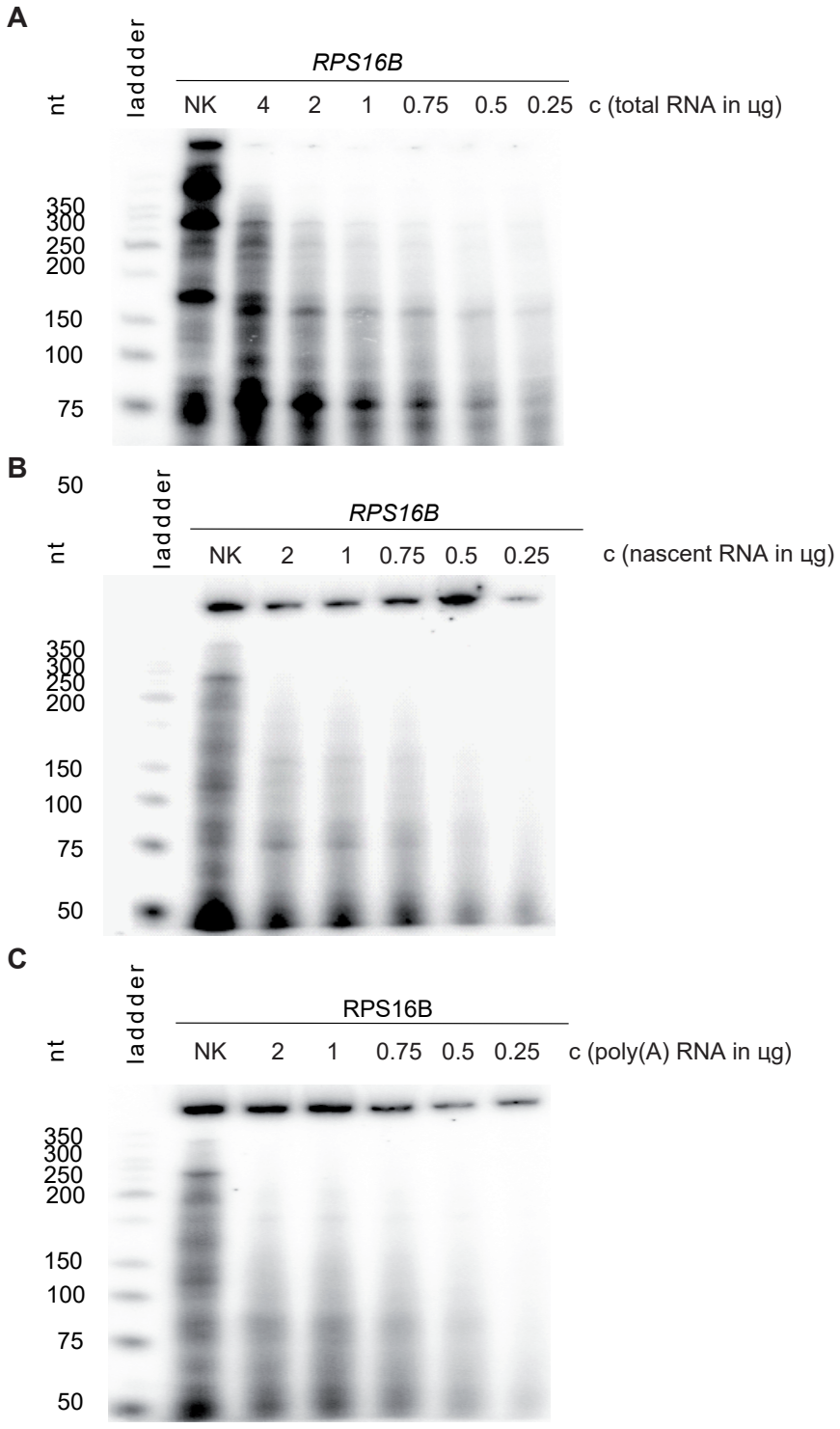

Figure 34: Primer extension after CMC treatment of different concentrations of total RNA, nascent mRNA and poly(A) mRNA to identify an RNA concentration in the exponential growth phase that was used for following experiments. Figures show autoradiographs of the Primer Extension at *RPS16B* of total RNA (**A**), nascent mRNA (**B**) and poly(A) mRNA (**C**). nt: nucleotides

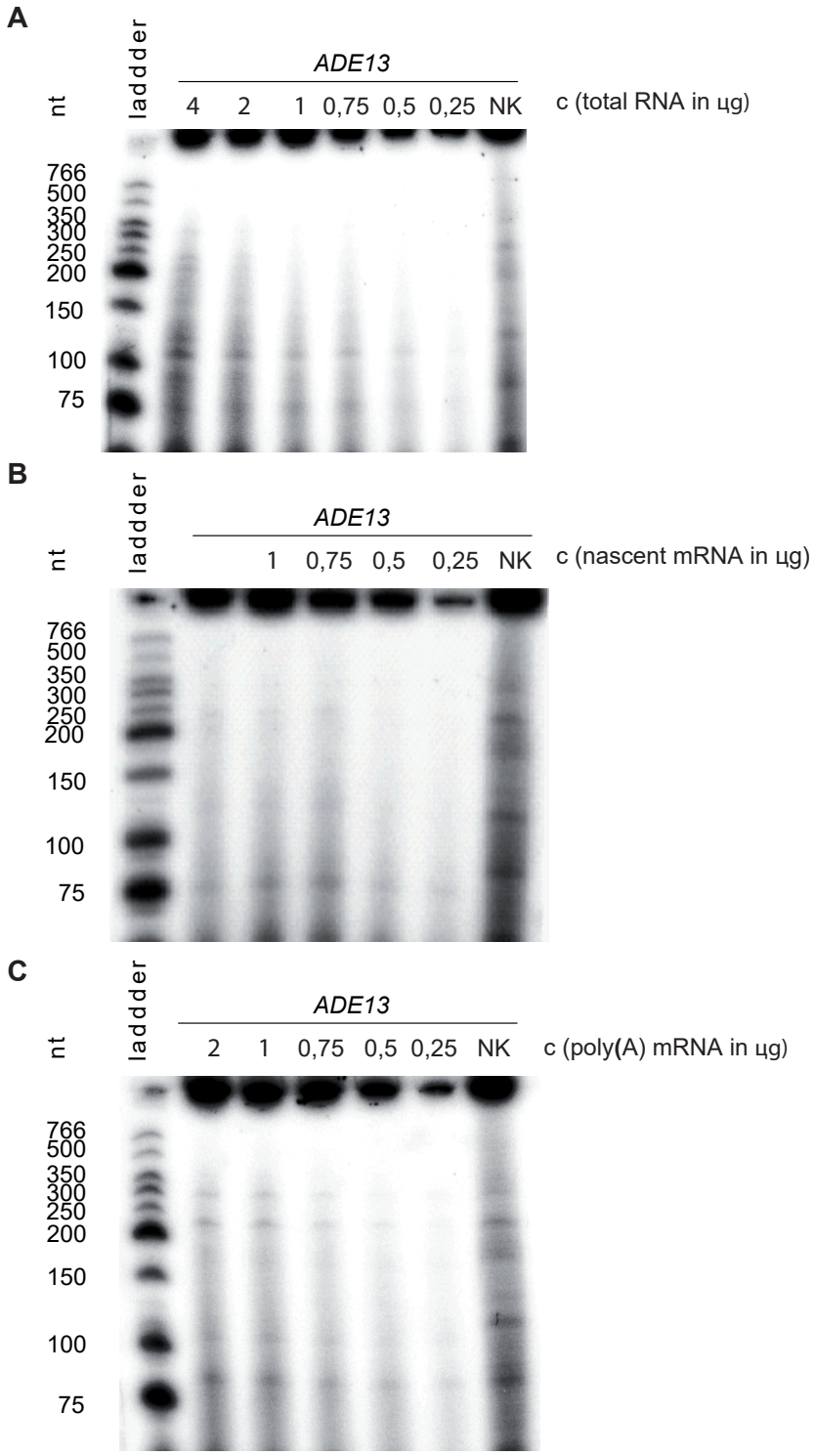

Figure 35: Primer extension after CMC treatment of different concentrations of total RNA, nascent mRNA and poly(A) mRNA to identify an RNA concentration in the exponential growth phase that was used for following experiments. Figures show autoradiographs of the Primer Extension at *ADE13* of total RNA (**A**), nascent mRNA (**B**) and poly(A) mRNA (**C**). nt: nucleotides

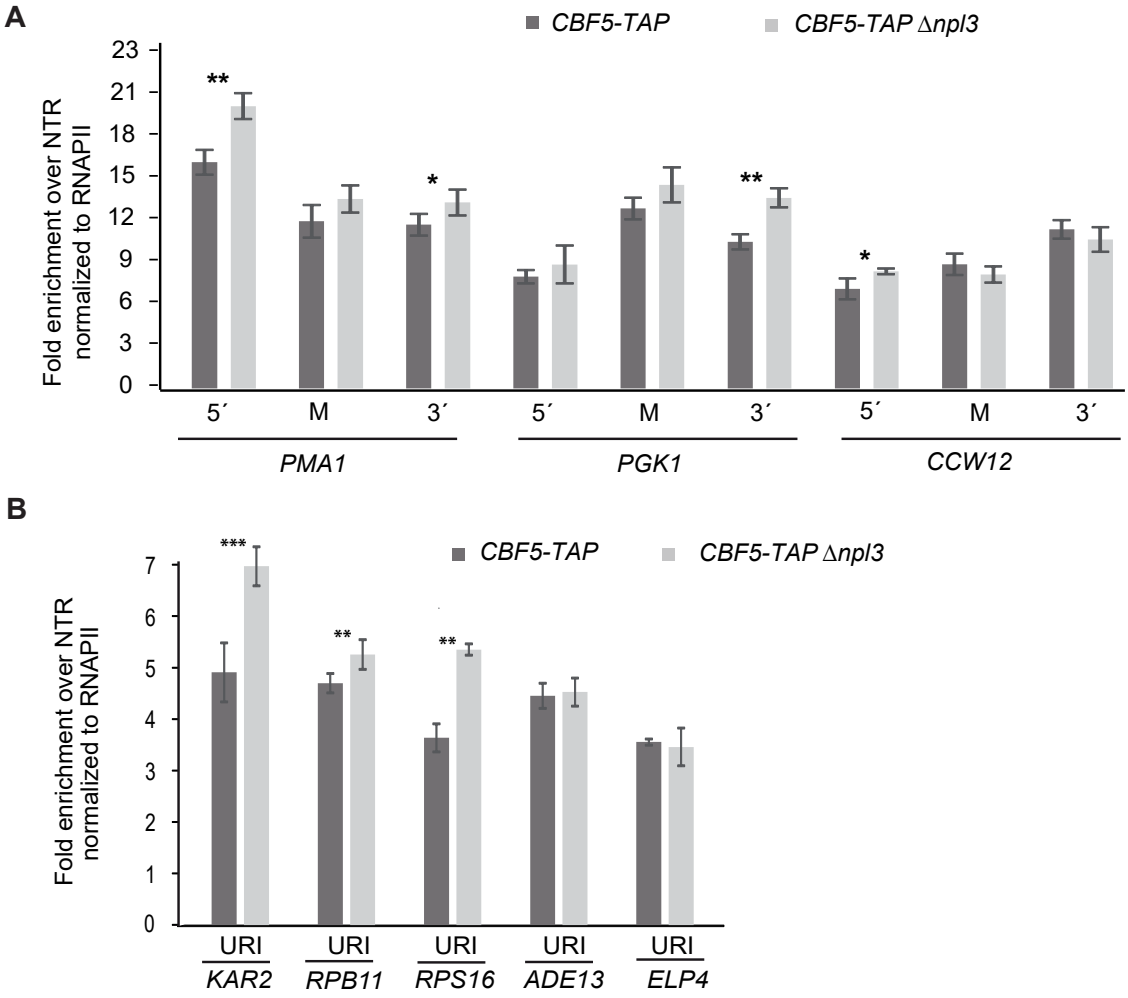

Figure 36: Npl3 is not needed for Cbf5 occupancy at transcribed genes. The occupancy of Cbf5 and RNAPII at the tested genes was determined by ChIP. Fold enrichment was normalized to an untranscribed region (NTR). The same primer pairs for ChIP analysis were used as described in figure 10. Figure A&B show that the normalized Cbf5 occupancy to the occupancy of RNAPII is unchanged or increased in the *Npl3* deletion strain. Bars represent the mean +/- standard deviation of three independent biological replicates, p values of students t-test: ***: $p < 0.001$ **: $p < 0.01$; *: $p < 0.05$

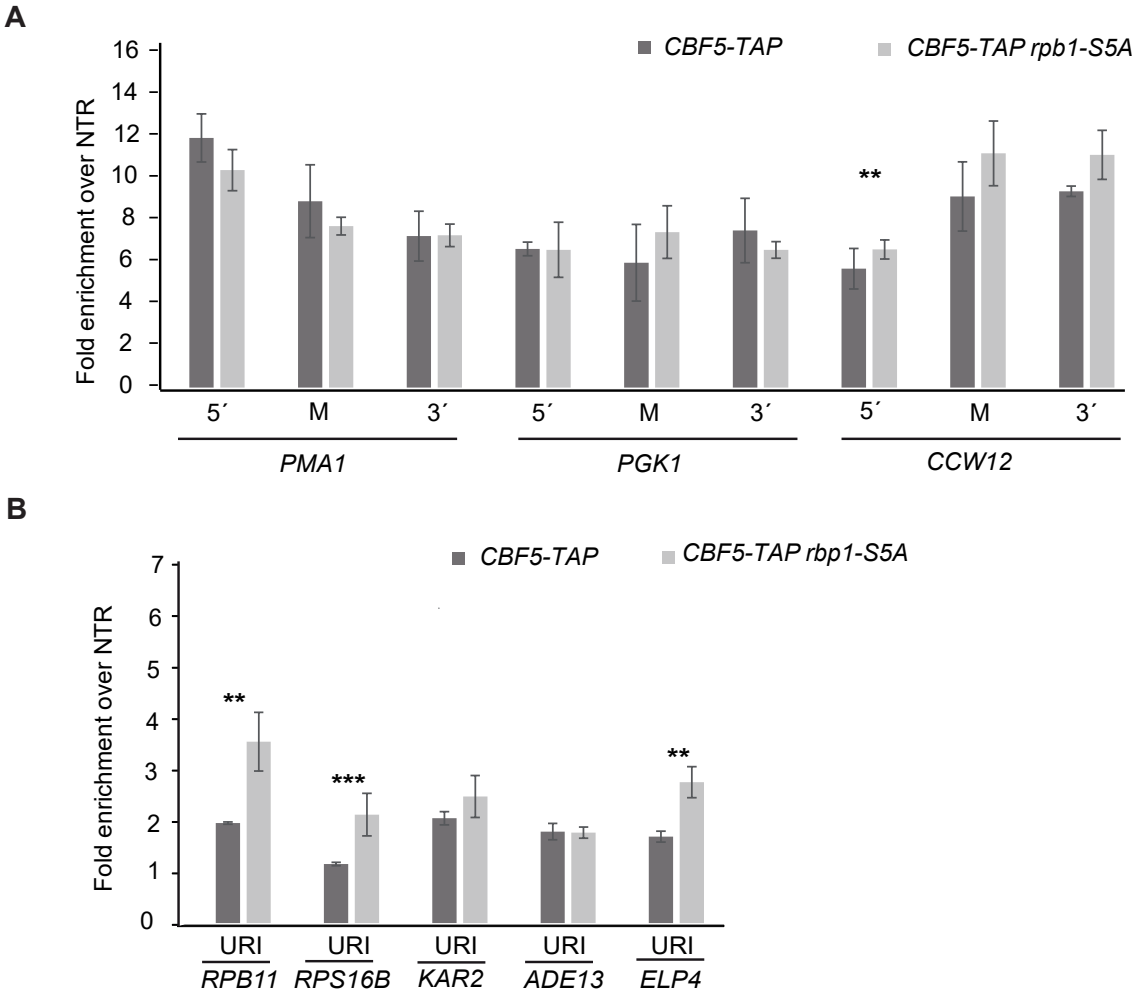

Figure 37: Phosporylation of serine 5 of the C-terminal domain of *RBP1* is not needed for Cbf5 occupancy at transcribed genes. The occupancy of Cbf5 at the tested genes was determined by ChIP. Fold enrichment was normalized to an untranscribed region (NTR). The same primer pairs for ChIP analysis were used as described in figure 10. Figure A&B show that the occupancy of Cbf5 is unchanged or increased when serine 5 of the C-terminal domain of *RBP1* is phosphorylated. Bars represent the mean +/- standard deviation of three independent biological replicates, p values of students t-test: ***: $p < 0.001$ **: $p < 0.01$; *: $p < 0.05$

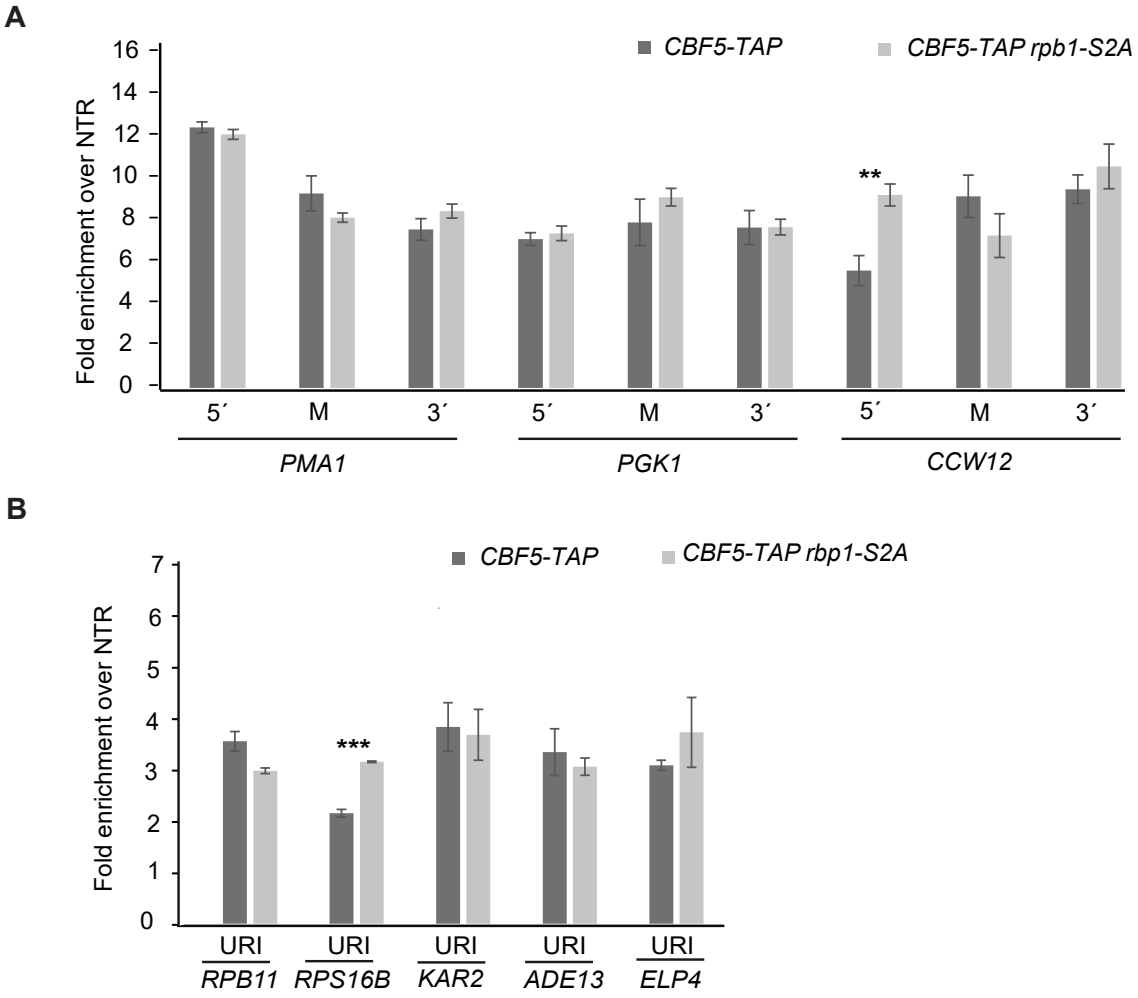

Figure 38: Phosporylation of serine 2 of the C-terminal domain of *RBP1* is not needed for Cbf5 occupancy at transcribed genes. The occupancy of Cbf5 at the tested genes was determined by ChIP. Fold enrichment was normalized to an untranscribed region (NTR). The same primer pairs for ChIP analysis were used as described in figure 10. Figure A&B show that the occupancy of Cbf5 is unchanged or increased when serine 2 of the C-terminal domain of *RBP1* is phosphorylated.. Bars represent the mean +/- standard deviation of three independent biological replicates, p values of students t-test: ***: $p < 0.001$ **: $p < 0.01$; *: $p < 0.05$

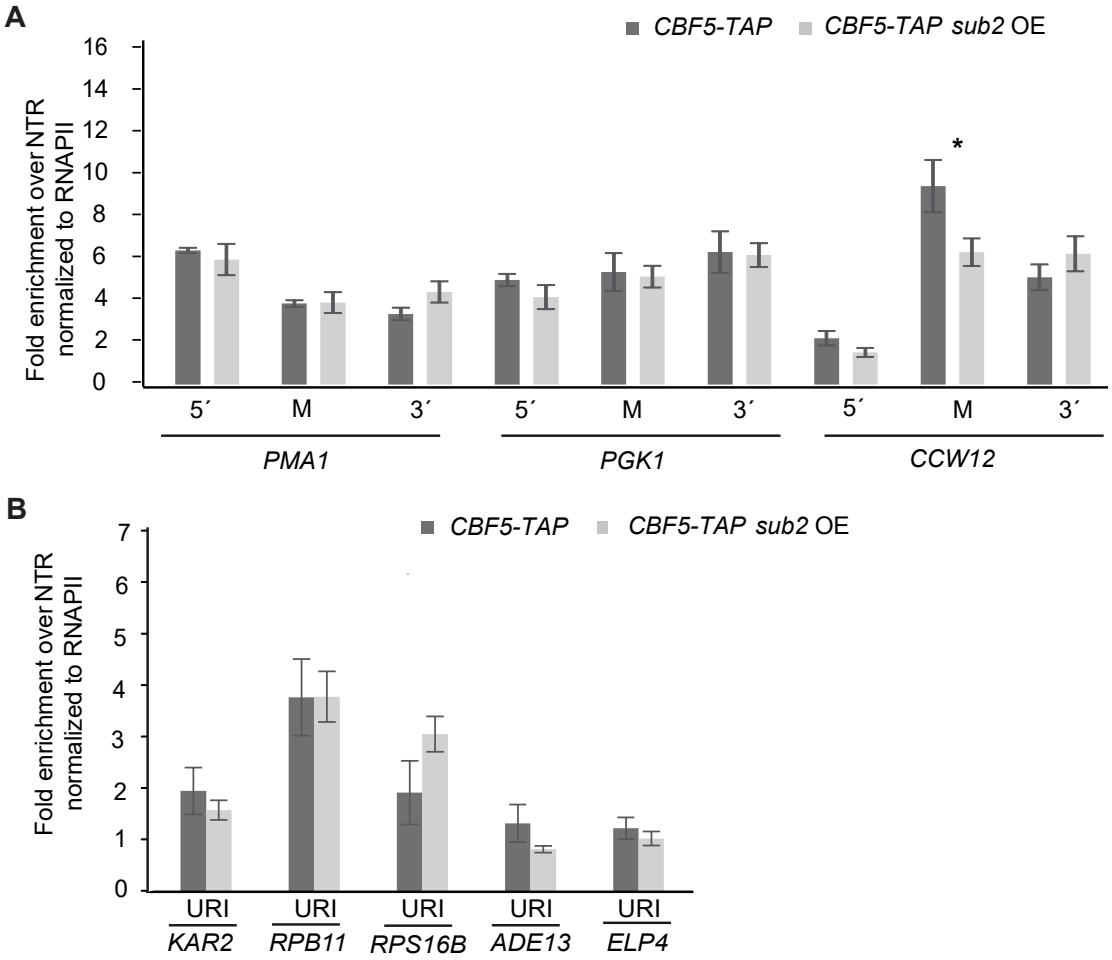

Figure 39: *sub2* overexpression does not show a decreased Cbf5 occupancy at transcribed genes. The occupancy of Cbf5 and RNAPII at the tested genes was determined by ChIP. Fold enrichment was normalized to an untranscribed region (NTR). The same primer pairs for ChIP analysis were used as described in figure 10. Figure A&B show that the normalized Cbf5 occupancy to the occupancy of RNAPII is mostly unchanged in a *sub2* overexpression (OE) strain. Bars represent the mean +/- standard deviation of three independent biological replicates, p values of students t-test: ***: $p < 0.001$ **: $p < 0.01$; *: $p < 0.05$

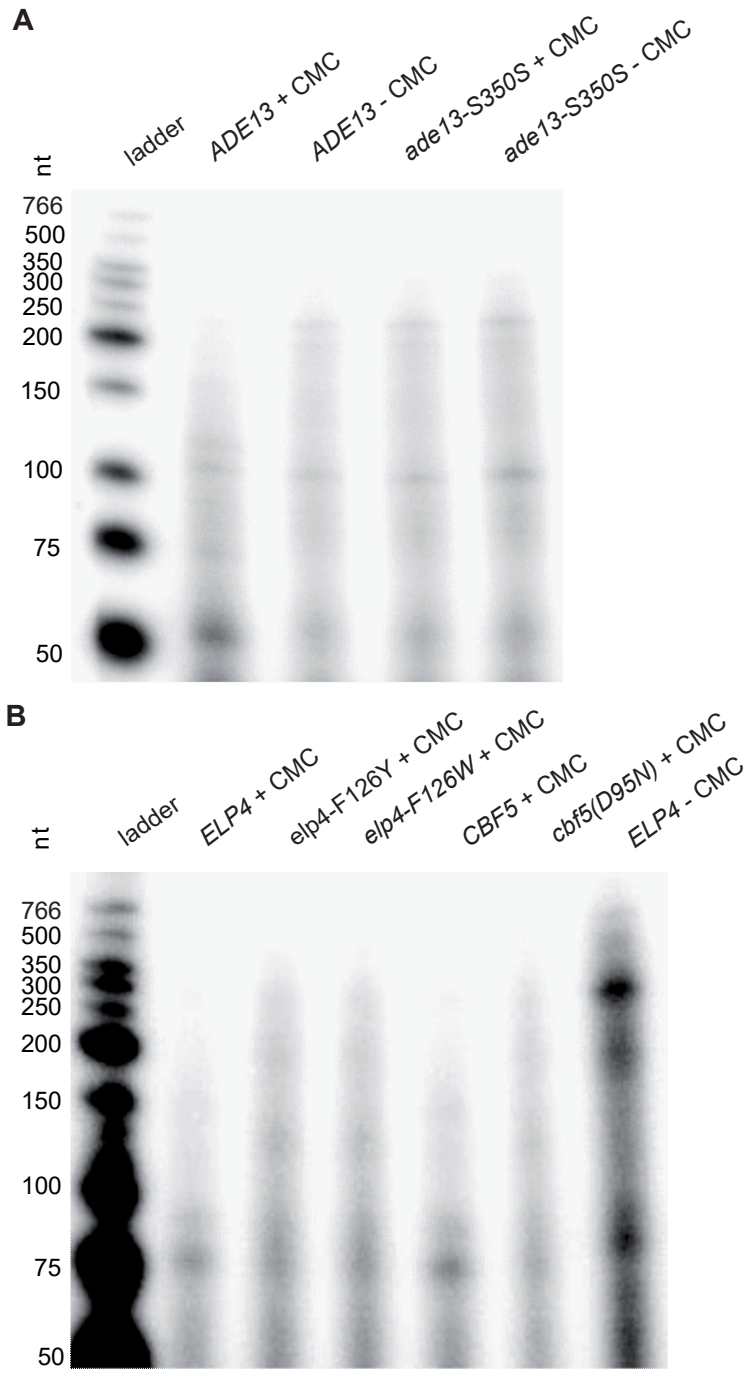

Figure 40: Cbf5-dependent pseudouridylation does not exist in *elp4* and *ade13* PseudoU site mutants. Primer extension of CMC-treated (+CMC) and CMC-untreated (-CMC) RNA in the PseudoU site mutants of *ade13* (*ade13-S350S*) (A) and *elp4* (*elp4-F126Y*; *elp4-F126W*) (B) as well as in the catalytically inactive *cbf5* mutant (*D95N*) and the corresponding wild-type (*ADE13*; *ELP4*; *CBF5*). nt: nucleotides

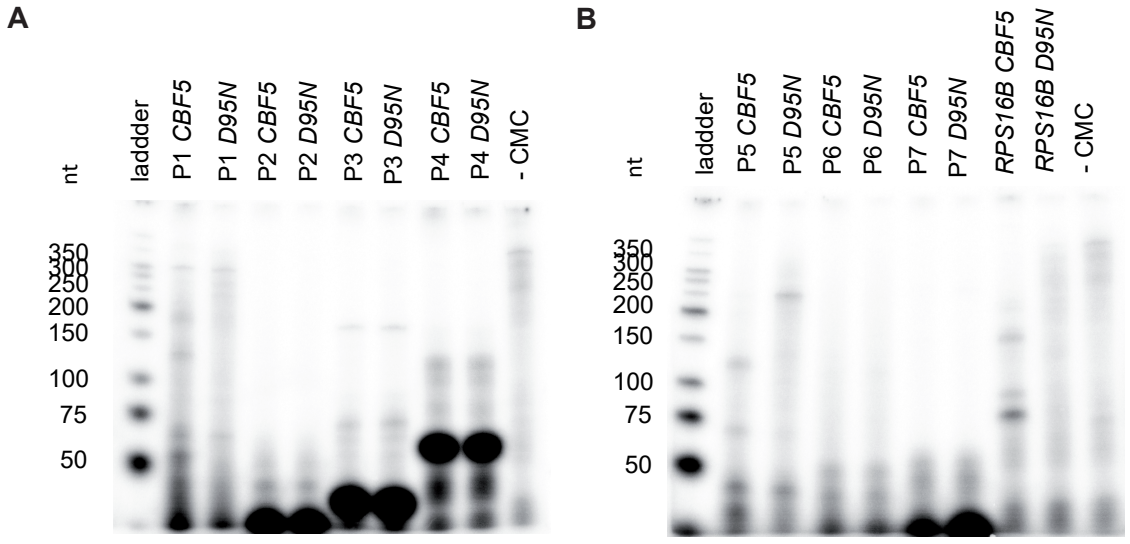

Figure 41: Primer extension of CMC-treated (+CMC) and CMC-untreated (-CMC) RNA in the catalytically inactive *cbf5* mutant (*D95N*) and the corresponding wild-type (*CBF5*) along the 5' UTR of *GCN4* and at *RPS16B*. nt: nucleotides

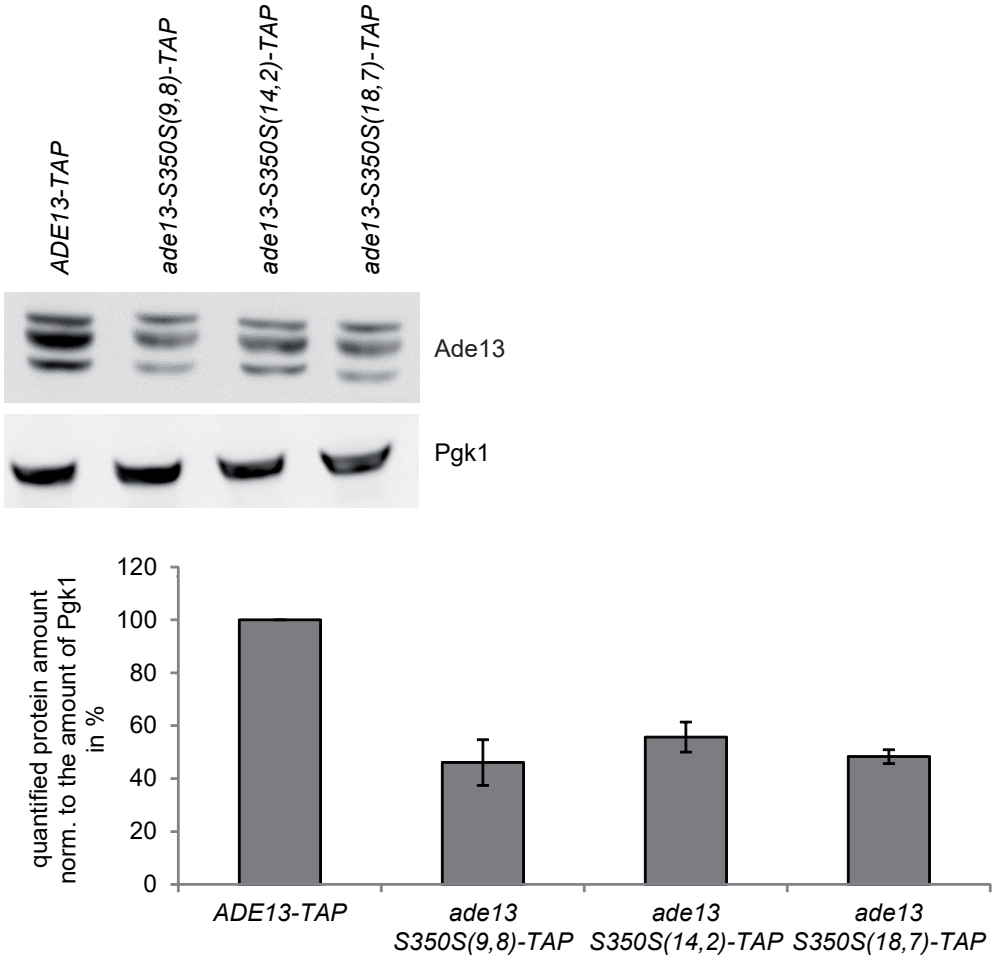

Figure 42: Different codon usages of *ade13-S350S* do not change protein expression of *ADE13*. Protein levels were determined by quantitative Western Blot analysis of whole cell extracts. Qantification was conducted by ImageJ. Pgk1 was used for normalization.

Publications

I contributed to the following publication:

Npl3 functions in mRNP assembly by recruitment of mRNP components to the transcription site and their transfer onto the mRNA
Philipp Keil, Alexander Wulf, Nitin Kachariya, Samira Reuscher, Kristin Hühn, Ivan Silbern, Janine Altmüller, Kathi Zarnack, Michael Sattler, Henning Urlaub, Katja Sträßer

Nucleid acids research

Danksagung

Nun ist es an der Zeit einmal Inne zu halten, um die vergangenen Jahre *revue* passieren zu lassen. Es war eine sehr prägende Zeit mit vielen Höhen und Tiefen, die ich nicht missen möchte. Ich möchte mich an dieser Stelle bei allen Menschen bedanken, die mich auf diesem Weg unterstützt haben.

Besonderen Dank gebührt meiner Doktormutter Prof. Dr. Katja Sträßer, die mir die Möglichkeit gegeben hat, dieses spannende Projekt außerhalb Ihrer Forschungsschwerpunktes zu bearbeiten. Speziell die fachlichen Diskussionen und Anregungen, als auch die aufbauenden Worte in so manchen Momenten waren sehr hilfreich zur Realisierung dieser Arbeit. Danke, für das entgegengebrachte Vertrauen und die Möglichkeit meine eigenen Ideen verwirklichen zu dürfen. Besonders schätze ich den offenen Umgang zwischen uns.

Apl. Prof Dr. Elena Evguenieva-Hackenberg danke ich für die Übernahme des Zweitgutachtens sowie für die angenehme Zusammenarbeit in dem PseudoU Nebenprojekt. Auch möchte ich mich bei den anderen Mitgliedern des Prüfungskomitees für Ihre Zeit bedanken.

Meinen Kollaborationspartnern Dr. Stefan Günther, Dr. Kathi Zarnack und Samira Reuscher als auch Dr. Alexander Leitner danke ich für die Durchführung der Sequenzierungen als auch Massenspektrometrie und die dazugehörige Datenauswertung. Ihr habt einen wertvollen Beitrag zu dieser Arbeit geleistet. Es war eine angenehme Zusammenarbeit mit euch.

Ein besonders großes Dankeschön geht an alle ehemaligen und derzeitigen Mitarbeiter der Biochemie für die angenehme Arbeitsatmosphäre, die Hilfsbereitschaft sowie die fachliche als auch mentale Unterstützung. Ihr ward ein Grund weswegen ich gern zur Arbeit gegangen bin. Jacqueline danke ich für die Unterstützung bei der Durchführung der Dot Blot assays. Bei Cornelia möchte ich mich für die fachlichen Diskussionen und den ein oder anderen Denkanstoß bedanken. Ein großes Dankeschön geht an dieser Stelle an Laura und Johanna für die schöne gemeinsame Zeit als Arbeitskollegen und Freunde. Seien es wissenschaftliche, methodische oder auch private Probleme, bei euch habe ich immer ein offenes Ohr gefunden! Außerdem bedanke ich mich bei Laura für die Hilfe bei ChIPseq und bei Johanna für die Durchführung der FISH Experimente und der dazugehörigen Quantifizierung. Nicht zuletzt möchte ich mich bei Nadine, Cornelia, Wolle, Corinna, Laura und Johanna für die Korrektur dieser Arbeit bedanken.
Außerdem bedanke ich mich bei meinen Studenten Anna-Lena, Jana und Rieka, die mir so

manche Arbeit abgenommen haben. Es hat mir großen Spaß gemacht, euch die ein oder andere Methode beizubringen. Den Medien HIWIs Isabel, Verena, Nils und Vivien danke ich fürs Ansetzen von unzähligen Litern Medium und Puffer und Petra danke ich fürs Spülen und Autoklavieren der Glaswaren.

Zudem bedanke ich mich bei allen Freunden und Familienmitgliedern, die mich stets ermutigt haben weiterzumachen. Danke für euer Verständnis, wenn ich mal wieder keine Zeit hatte. Ein großes Dankeschön geht auch an meine Eltern. Den Ehrgeiz und das Durchhaltevermögen, welches ihr mir auf den Weg gegeben habt, waren sehr hilfreich in den letzten Jahren.
Das größte Dankeschön geht an meinen Ehemann Dominic, der dieses Projekt über die gesamte Zeit mitgetragen oder auch manchmal ertragen musste. Ich danke dir für dein Einfühlvermögen, dein Verständnis und den ein oder anderen Rat aus eigener Erfahrung. Du warst der Fels in der Brandung, der immer darauf geachtet hat, dass es mir gut geht. Danke, dass du mir vor allem in den letzten Monaten den Rücken freigehalten hast, diese Arbeit zu beenden. Ich bin froh, dass es dich gibt!

Eidstaatliche Erklärung

Ich erkläre: Ich habe die vorgelegte Dissertation selbstständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Ich stimme einer evtl. Überprüfung meiner Dissertation durch eine Antiplagiat-Software zu. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten.

Giessen, den