JUSTUS-LIEBIG-
UNIVERSITÄT
GIESSEN

# Function of the RNA-binding activity of the nuclear mRNA-binding protein Npl3 in nuclear mRNA packaging

**Dissertation**

zur Erlangung des Doktorgrades der Naturwissenschaften

(Dr. rer. nat.)

am Fachbereich 08 für Biologie und Chemie

der Justus-Liebig-Universität Gießen

Vorgelegt von

**Philipp Keil**

Gießen, 2021

Die vorliegende Arbeit wurde am Institut für Biochemie (Fachbereich 08) der Justus-Liebig-Universität Gießen, unter Leitung von Prof. Dr. Katja Sträßer, erstellt.

Dissertation eingereicht am: 27.07.2021

Erstgutachter: Prof. Dr. Katja Sträßer
Fachbereich 08: Biologie und Chemie
Institut für Biochemie
Justus-Liebig-Universität Gießen

Zweitgutachter: Prof. Dr. Sandra Hake
Fachbereich 08: Biologie und Chemie
Institut für Genetik
Justus-Liebig-Universität Gießen

# List of content

# 1 Zusammenfassung

Ein wesentlicher Schritt der Genexpression ist die Bildung eines Boten-Ribonukleoprotein-Partikels (mRNP). Bereits während der Transkription binden RNA-Bindeproteine (RBPs) an die mRNA und bilden so ein mRNP. Dies ist essentiell für die Stabilität der mRNA und nur korrekt verpackte mRNPs werden aus dem Zellkern in das Zytoplasma transportiert. Der nukleäre Export von mRNPs wird über den mRNA-Exportrezeptor Mex67-Mtr2 vermittelt, welcher hauptsächlich über Adapterproteine wie das RNA-*annealing* Protein Yra1, das Poly(A)-Bindeprotein Nab2 und das SR-ähnliche Protein Npl3 zur mRNA rekrutiert wird.

Die generelle Funktion der Proteine, die an diesen Prozessen beteiligt sind, wurde bereits analysiert, indem das gesamte Protein deletiert beziehungsweise depletiert oder zumindest ganze Proteindomänen entfernt wurden. Dies führt allerdings dazu, dass mehrere Funktionen eines Proteins gleichzeitig unterbunden werden. Um gezielt die RNA-Bindefunktion von Proteinen ausschalten zu können, die an der nukleären mRNP-Bildung beteiligt sind, ohne andere Funktionen des Proteins zu beeinträchtigen, wurden in der vorliegenden Arbeit zunächst jene Aminosäuren, die unmittelbar an der RNA-Bindung beteiligt sind, identifiziert. Die Identifizierung erfolgte mittels UV-Licht-Quervernetzung (*crosslinking*) und anschließender Massenspektroskopie-Analyse. Dabei wurden in 16 verschiedenen Proteinen, die unmittelbar an der mRNP-Bildung beteiligt sind, 100 Aminosäuren identifiziert, die *in vivo* mit der RNA quervernetzt werden konnten, unter anderem in: Npl3, Nab2, Tho1, Mex67-Mtr2 sowie in Komponenten des TREX-Komplexes (Tho2, Hpr1, Mft1, Hrb1, Sub2 und Yra1).

Die Aminosäuren der putativen RNA-Bindestellen, die mittels MS identifiziert wurden, wurden zunächst mutiert und im Anschluss dahingehend untersucht, ob und in welchem Ausmaß sie an der RNA-Bindung beteiligt sind. Im Anschluss wurde die Bedeutung der verminderten RNA-Bindungsaktivität für die verschiedenen biologischen Funktionen der individuellen mRNP-Komponenten analysiert. In der vorliegenden Arbeit wurden drei RNA-Bindestellen von Npl3 genauer untersucht. Zwei davon liegen in klassischen RNA-bindenden Domänen, sogenannten *RNA recognition motifs* (RRMs), die dritte im nicht-strukturellen Linker-Bereich dazwischen. Npl3 ist unter anderem an der Transkription, 3‘-Prozessierung, mRNP-Bildung sowie dem mRNA-Export beteiligt. Nach Erzeugung der drei *npl3* Mutanten wurden die funktionellen Auswirkungen untersucht. Interessanterweise zeigen die drei *npl3* Mutanten mit reduzierter RNA-Bindung unterschiedliche Phänotypen. Des Weiteren konnte durch Analyse der *npl3-Linker* Mutante nachgewiesen werden, dass Npl3 am Transfer von nukleären mRNP-Komponenten auf die mRNA beteiligt ist.

Zusammenfassend wurden putative *in vivo* RNA-Bindestellen diverser nukleärer RBPs identifiziert. Darüber hinaus konnte gezeigt werden, dass die Inhibierung bestätigter RNA-Bindestellen in verschiedenen Bereichen von Npl3 spezifische und funktionell unterschiedliche Auswirkungen hat. Außerdem konnte gezeigt werden, dass Npl3 als initialer Faktor im Transfer von nukleären mRNP-Komponenten auf die mRNA fungiert.

# 2 Summary

An essential step of gene expression is the formation of a messenger ribonucleoprotein particle (mRNP). Already during transcription the nascent mRNA is bound by mRNA-binding proteins (RBPs) that package the mRNA into an mRNP. The formation of mRNPs is crucial for mRNA stability and only correctly packaged mRNPs are exported from the nucleus through the nuclear pore complexes to the cytoplasm. The nuclear export of mature mRNPs is mediated by the mRNA export receptor Mex67-Mtr2, which is mainly recruited by adaptor proteins like the RNA-annealing protein Yra1, the poly(A)-binding protein Nab2, and the SR-like protein Npl3. Besides nuclear events, bound RBPs often regulate cytoplasmic events such as localization, translation and degradation of the mRNA.

The functions of the proteins involved in these processes have largely been analyzed by depletion or deletion of either the whole protein or entire protein domains, which will often abrogate multiple activities of each protein at once. To determine specifically the functional significance of the RNA-binding activity of proteins involved in nuclear mRNP formation, we identified the amino acids that are in close proximity to RNA by UV-light crosslinking and subsequent mass spectrometric analyses. Approximately 100 amino acids that could be crosslinked to RNA *in vivo* were identified in various proteins involved in mRNP formation, namely in Npl3, Nab2, Tho1, Mex67-Mtr2 and in most components of the TREX complex (Tho2, Hpr1, Mft1, Hrb1, Sub2 and Yra1).

To selectively disrupt the interaction of RNA with the newly identified putative binding sites, selected amino acids that crosslinked to RNA were mutated and afterwards the RNA-binding activity was analyzed. After confirmation of reduced RNA-binding activity, the functional consequences of these mutations on mRNP assembly, mRNA export, mRNA expression and stability, and splicing were then analyzed in detail. In this study, we carried out mutational analysis of the protein Npl3. Npl3 is an SR-like protein involved in transcription, 3' end formation, mRNP assembly and nuclear mRNA export. Npl3 contains two RRM domains connected by a non-structured flexible linker, an SR/RGG motif at its C-terminus and an APQE motif in its N-terminus. Three *npl3* mutants were generated, one in the linker-region and one within each of the two RRM domains, and the functional consequences of these mutations were elucidated. Interestingly, reduction of mRNA-binding in the three different regions of Npl3 had different phenotypic outcomes, suggesting that the different functional activities of the protein are independent from each other. Moreover, analysis of the *npl3-Linker* mutant revealed a novel function of Npl3. Npl3 acts as an initial factor in the transfer of nuclear mRNP components to the mRNA.

In sum, putative *in vivo* RNA binding sites of nuclear mRNA binding proteins were identified. In addition, disruption of verified RNA binding sites in different regions of Npl3 has specific and surprisingly different functional consequences. Furthermore, a novel function of Npl3 in nuclear mRNP assembly was discovered.

# 3 Introduction

## 3.1 mRNA biogenesis and export

### 3.1.1 Gene expression

Eukaryotic cells are - in contrast to prokaryotic cells - divided in a nucleus and cytoplasm. The DNA - the genetic blueprint - is stored inside the nucleus and the protein-coding genes are transcribed by RNA polymerase II (RNAPII) into the messenger RNA (mRNA). The nascent synthesized pre-mRNAs undergo co-transcriptional processing including 5' capping, removing introns by splicing, 3' cleavage and polyadenylation. Already during transcription several RNA-binding proteins (RBPs) bind the newly synthesized mRNA and package it into an export-competent messenger ribonucleoprotein particle (mRNP). Only correctly packaged and processed mRNAs are exported from the nucleus to the cytoplasm. mRNPs undergo an additional remodeling step at the cytoplasmic face of the nuclear pore complex (NPC) before they are finally released into the cytoplasm and the mRNA is translated into the respective protein.

### 3.1.2 Recruitment of RBPs to transcribed genes

#### *3.1.2.1 RNAPII CTD as recruiting platform for processing and regulatory factors*

Gene expression is the sum of many processes from regulation of chromatin accessibility to the production of the final proteins, but has its starting point at the transcription of genes by RNAPII. RNAPII serves as one of three recruiting platforms for several factors and RBPs that orchestrate among others the processes of transcription, mRNP maturation and export. One of the unique characteristics of RNAPII among the other RNA polymerases is the carboxy-terminal domain (CTD) of its largest subunit Rpb1. The CTD consists of highly conserved repetitive heptapeptides ($Y_1S_2P_3T_4S_5P_6S_7$) with 26 repeats in *S. cerevisiae* and 52 repeats in humans (Corden *et al.*, 1985; Liu *et al.*, 2010). As RNAPII progresses through the various stages of the transcription cycle from initiation to termination, the heptapeptides of the CTD are intensely and reversibly modified. The tyrosine, threonine and serines can be phosphorylated, the threonine and serines can be glycosylated and the prolines can undergo isomerization. The differentially modified CTD serves as primary recruiting platform for RBPs and transcription and processing factors that act in mRNP biogenesis and chromatin remodeling. During the initiation phase, the $S_5$ and $S_7$ of the CTD repeats are phosphorylated by the transcription factor IIH-associated kinase Kin28. The $S_5$ phosphorylation is supported by the mediator-complex- associated kinase Srb10. In the mammalian system the respective kinases are Cdk7 and Cdk8 (Liao *et al.*, 1995; Rodriguez *et al.*, 2000; Zhang *et al.*, 2012). The

phosphorylation of $S_5$ is important for early chromatin remodeling by recruiting Set1, a histone methyltransferase of the COMPASS-complex that trimethylates H3 at position K4 (Ng *et al.*, 2003); for mRNA export by recruiting Sus1, a component of both the SAGA and THSC/TREX-2 complexes (Pascual-García *et al.*, 2008); and for mRNA 5' capping by recruiting Ceg1, a guanylyltransferase (Cho *et al.*, 1997). The placement of m7G cap on the nascent mRNA by the capping enzyme complex, which represents the first step of mRNP formation, stabilizes the mRNA by preventing its degradation by 5' to 3' exonucleases. The phosphorylation of $S_7$ is needed to prime the RNAPII CTD for recognition of the positive transcription elongation factor P-TEFb (Czudnochowski *et al.*, 2012).

After several abortive initiation cycles, until a transcript of a threshold length of approximately 10 nucleotides is synthesized, promoter clearance occurs. The transition from initiation to elongation is marked by a change in the phosphorylation pattern of the CTD. Bur1, a cyclin-dependent protein kinase (Cdk9 in mammals), is recruited to the $S_5$ phosphorylation marks and phosphorylates $S_2$ (Qiu *et al.*, 2009). Ctk1, the major $S_2$ kinase that substantially increases the phosphorylation of $S_2$ causes the priming of the CTD (Cho *et al.*, 2001; Jones *et al.*, 2004). Meanwhile, the predominant $S_5$ phosphorylation marks are removed by the CTD phosphatases Rtr1 and Ssu72, although a low level of $S_5$ phosphorylation remains over the open reading frame (Krishnamurthy *et al.*, 2004; Mosley *et al.*, 2009). In budding yeast this change occurs already behind the first nucleosome after promotor clearance (Mayer *et al.*, 2010). Due to the change in the phosphorylation pattern, initiation factors are exchanged for proteins required during elongation. The recruitment of elongation factors like Paf1, Spt4, Spt5, Spt6 and others occurs competitively on every gene independently of its length, type or expression (Mayer *et al.*, 2010). The recruitment of these factors is necessary for the processivity of transcription, histone modification and 3' end processing (Adkins and Tyler, 2006; Hartzog *et al.*, 1998; Jaehning, 2010; Martinez-Rucobo *et al.*, 2011). In terms of mRNP assembly and RNA processing the phosphorylated $S_2$ is known to recruit Prp40, a subunit of the spliceosome (Morris and Greenleaf, 2000) as well as the SR-like protein Npl3 that acts, among others, in transcription elongation, 3' end processing, splicing and mRNA export (Bucheli and Buratowski, 2005; Bucheli *et al.*, 2007; Dermody *et al.*, 2008; Lei *et al.*, 2001).

The next transition from elongation to termination is marked by increasing levels of phosphorylated $S_2$ and $Y_1$ of the CTD coupled with dephosphorylation of $S_5$ at the cleavage and polyadenylation site. The change in phosphorylation recruits 3' end processing factors like Pcf11, an essential component of the cleavage factor IA complex

that binds $S_2$ and promotes the release of RNAPII (Lunde *et al.*, 2010; Noble *et al.*, 2005; Zhang *et al.*, 2005). Several of the known termination and 3' processing factors preferentially associate with CTD phosphorylated at $S_2$ or $S_2$/$S_5$: Npl3, Rtt103, Rna14, Rna15, Ydh1, Yhh1, Pta1, and Pcf11 (Barillà *et al.*, 2001; Dichtl *et al.*, 2002; Kyburz *et al.*, 2003; Zhang *et al.*, 2012). Finally, the CTD becomes dephosphorylated again, which promotes the release of RNAPII from the DNA. Free RNAPII can then initiate the next cycle of transcription (Cho *et al.*, 1999; Steinmetz and Brow, 2003).

#### *3.1.2.2 Spt5 as a recruiting platform*

Spt5 is a general elongation factor that binds to the body of RNAPII (Klein *et al.*, 2011; Martinez-Rucobo *et al.*, 2011) and also acts as a recruitment platform for transcription factors and RBPs. Spt5 is the only known RNAPII-associated factor that is conserved in all three kingdoms of life (Grohmann and Werner, 2011). In a purification of Spt5 more than 90 yeast proteins involved in transcription elongation and termination, mRNP biogenesis and mRNA export have been identified as co-purifiers, demonstrating its general role as an additional platform for recruitment of mRNA binding proteins (Lindstrom *et al.*, 2003). For example, Spt5 interacts with the capping enzyme (Lidschreiber *et al.*, 2013; Pei and Shuman, 2002; Wen and Shatkin, 1999) and recruits the Paf1 complex (Liu *et al.*, 2009) as well as the pre-mRNA cleavage factor I (CFI) (Mayer *et al.*, 2012). Similar to Rpb1, Spt5 carries a repetitive C-terminal region (CTR) reminiscent of Rpb1-CTD. In *S. cerevisiae,* the CTR consists of 15 hexapeptides with the consensus sequence S,T/A,W,G,G,A/Q (Swanson *et al.*, 1991), while the human CTR has pentapeptides with the sequence G,S,R/Q,T,P (Yamada *et al.*, 2006). The recruitment of proteins like Paf1 can be mediated by the phosphorylation of the serines within the CTR repeats by the kinase Bur1 (Liu *et al.*, 2009; Zhou *et al.*, 2009).

#### *3.1.2.3 Direct mRNA interactors*

Independently of the recruiting platforms described above, the (pre-)mRNA itself is directly bound by several RBPs as it emerges from RNAPII. In comparison to DNA, the secondary and tertiary structure of RNA is much more diverse. The interaction between the RBPs and RNA occurs mainly via sequence specific motifs as well as structural elements of the RNA. RBPs contain distinct conserved RNA-binding domains that interact with RNA, for example RRM (RNA recognition motif; (Cléry *et al.*, 2008; Maris *et al.*, 2005)), RGG/RG domain (arginine/glycine-rich; (Ozdilek *et al.*, 2017)), KH domain (K homology; (Valverde *et al.*, 2008)) or ZnF (zinc finger domain; (Hall, 2005)). However, recent studies revealed that protein–RNA interactions do not always require canonical RBDs and that such unconventional RNA binding occurs more often than expected

(Hentze *et al.*, 2018 and references therein). In the case of small nuclear RNPs (snRNPs), subcomplexes of the spliceosome comprised of proteins and RNAs, the binding to specific RNA sites is mediated by base pairing with the target motif (Will and Lührmann, 2011). Moreover, some RBPs mediate the recruitment of other proteins or RNAs once they are bound to the mRNA. For example, Yra1, Nab2 and Npl3 act as adaptor proteins that recruit the heterodimeric export receptor Mex67-Mtr2 to the mRNA.

The co-transcriptional recruitment of RBPs to the pre-mRNA is not only crucial for correct RNA processing but also prevents the immediate degradation of the newly synthesized mRNA by nuclear surveillance machineries such as the highly conserved exosome complex. Furthermore, nuclear mRNP formation has an impact on cytoplasmic events that affect mRNA fate later on, and can influence translatability, stability and localization of mRNAs. For example, in *S. cerevisiae* the *ASH1* transcript that has a key function in regulating the mating type switch has to be asymmetric distributed inside the cell (Bobola *et al.*, 1996; Cosma, 2004). In *Drosophila melanogaster* oocytes the distribution of the *bicoid*, *oskar* and *nanos* mRNA to the corresponding poles is necessary for the appropriate spatial development of the body axis of the embryo; reviewed in (Becalska and Gavis, 2009). Apart from nuclear mRNA export, higher plants have evolved a *long-distance signaling* mechanism to realize inter-organ communication (Xia *et al.*, 2018). A prominent example in *Arabidopsis thaliana* is the *flowering locus T* mRNA that is transcribed in leaves, is transported to the distantly located meristematic shoot apex where it is translated to induce flowering (Huang *et al.*, 2005).

Taken together, all three platforms: The RNAPII, Spt5 and the mRNA itself act together to recruit RBPs for efficient mRNP assembly. Several nuclear RBPs are recruited simultaneously by all three platforms. Correct recruitment of RBPs at a given time or process is essential during mRNP formation as well as the later fate of an mRNA.

### 3.1.3 mRNA processing events and mRNP formation

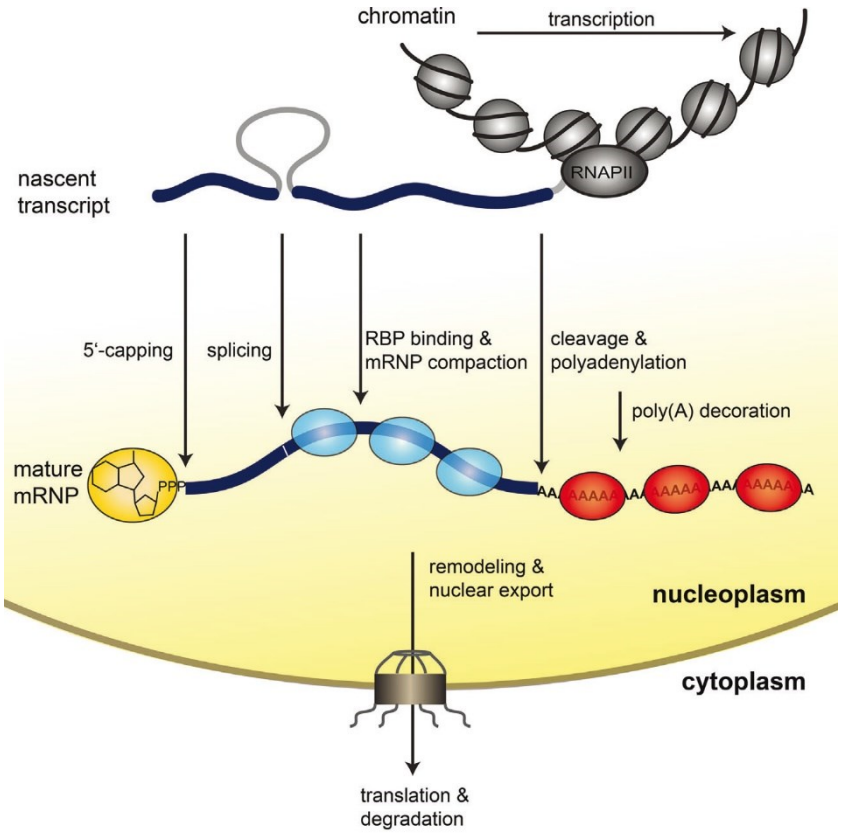



**Figure 1: mRNP biogenesis.** RNAPII transcribes protein-coding genes. Already during transcription, the nascent (pre-)mRNA is covered by different RNA binding proteins (RBPs) and proteins complexes. The RBPs regulate the processing of the mRNA (5' capping, splicing and 3' end formation). The mRNA is remodeled during these processes until an export-competent mRNP is formed. The mRNPs are exported from the nucleus through the nuclear pore complexes to the cytoplasm, where it is translated or degraded. Blue and red circles indicate different RBPs acting on the mRNA. (Meinel and Sträßer, 2015).

#### *3.1.3.1 5' capping*

Nascent mRNAs are processed co-transcriptionally and packaged into an mRNP by a variety of RBPs (Müller-McNicoll and Neugebauer, 2013). The first mRNA processing event is the 5' capping that occurs co-transcriptionally already after ~25 nucleotides emerge from the RNAPII. The cap, a N7-methylated guanosine (m7G) linked to the first nucleotide, protects the mRNA against 5' to 3' degradation by exonucleases such as Rat1 and Xrn1 (Bousquet-Antonelli *et al.*, 2000; He and Jacobson, 2001). To generate the cap structure, three enzymatic steps are performed. The RNA triphosphatase (Cet1) removes the γ-phosphate from the first nucleotide at the RNA 5′ end to generate a diphosphate RNA (Rodriguez *et al.*, 1999; Tsukamoto *et al.*, 1997). The RNA guanylyltransferase (Ceg1) transfers a guanine monophosphate from a guanosine triphosphate to the 5'-end diphosphate (Shibagaki *et al.*, 1992). Together, Cet1 and Ceg1 form the capping enzyme complex (Cho *et al.*, 1998). Ceg1 binds to phosphorylated $S_5$ on RNAPII CTD during transcription initiation (Ho and Shuman, 1999). Lastly, the

guanine-N7 methyltransferase (Abd1) adds a methyl group to the N7 of the guanosine (Mao *et al.*, 1995). Although the enzymatic procedure is conserved among all eukaryotes, the proteins involved vary between different organisms. In *S. cerevisiae,* three different enzymes (Cet1, Ceg1 and Abd1) are involved. In contrast, the RNA triphosphatase and the RNA guanylyltransferase is combined in a bifunctional protein in metazoans (Shuman, 2001). The cap is bound by the nuclear cap-binding complex (CBC) and the eukaryotic initiation factor eIF4E (Cdc33) in the cytoplasm. CBC consists of the small subunit Cbc2 (Cbp20) and the large subunit Sto1 (Cbp80). eIF4E is a general translation initiation factor that directs the ribosomes to the mRNA. The CBC and Cdc33 are the major binding factors of the cap that recruit RNA processing factors involved in pre-mRNA splicing, mediate processes like mRNA export and polyadenylation as well as the initiation of protein synthesis (Andersen *et al.*, 2013; Flaherty *et al.*, 1997; Nojima *et al.*, 2007; Pabis *et al.*, 2013).

#### *3.1.3.2 Splicing*

Splicing, the process that removes intragenic regions (introns) and ligates the expressed regions (exons) from the pre-mRNA also occurs co-transcriptionally. Splicing is a two-step transesterification reaction that is catalyzed by the spliceosome, a dynamic multimegadalton ribonucleoprotein (RNP) complex comprised of five snRNPs/snRNAs (U1, U2, U4, U5 and U6) and numerous proteins (Herzel *et al.*, 2017; Wahl *et al.*, 2009). In *S. cerevisiae*, only 5 % of all genes contain introns. By far, the majority of those genes contain only one intron. However, about 30 % of all mRNAs contain introns. This is due to the high expression level of genes coding for ribosomal proteins and other highly transcribed genes that contain an intron (Hooks *et al.*, 2014). This already indicates how important splicing is. In humans, almost every gene contains introns. In average human genes contain more than 7 introns per gene (Sakharkar *et al.*, 2004).

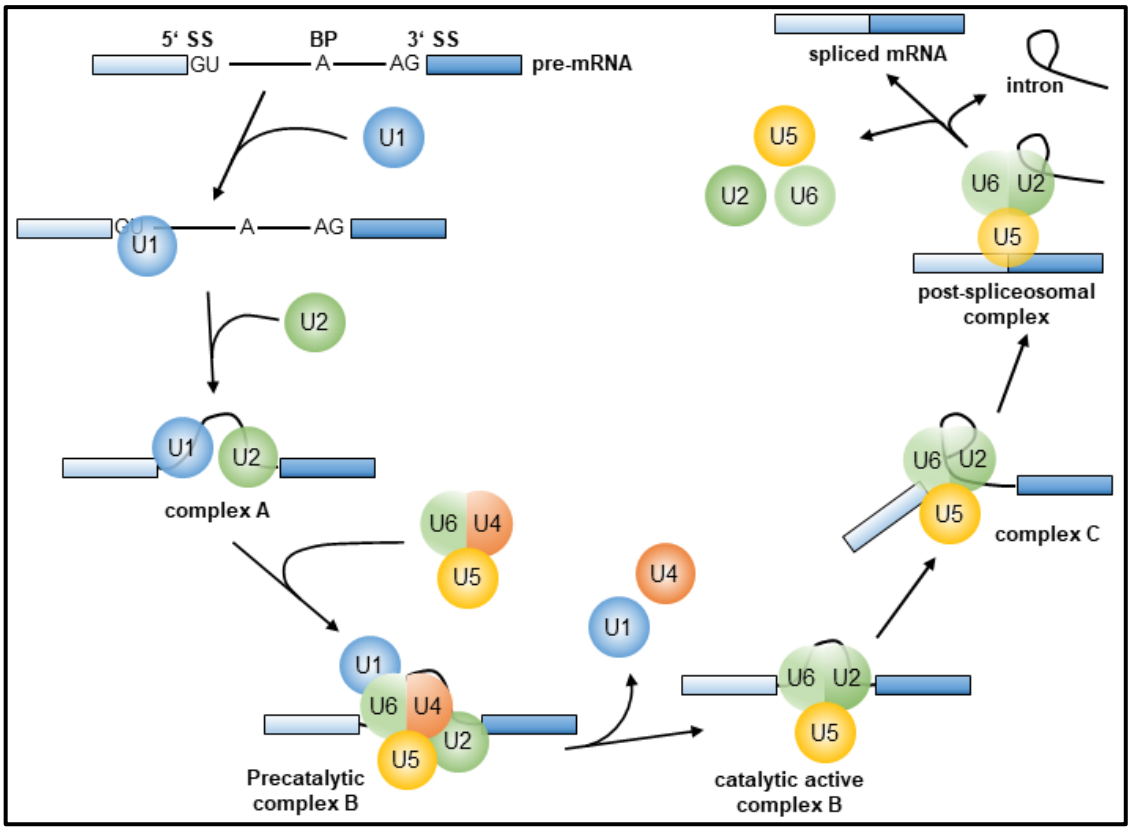



**Figure 2: Overview of splicing and the assembly of the spliceosome.** First, the U1 and U2 snRNPs bind to the 5' splice site and branch point, respectively and forming thereby the complex A. This results in recruitment of the U6/U4-U5 tri-snRNP. After remodeling and removing the S1 and S4 snRNP the catalytic actively complex B is formed. Two catalytic steps are performed to remove the intron and the exons are ligated together. After splicing occurred the spliced mRNA and the intron are released as well as the spliceosome components that are reused in another cycle of splicing. SS: splice site; BP: branch point. (Adapted from (Will and Lührmann, 2011))

Spliceable sequences are characterized by short consensus sequences at the 5' and 3' intron-exon junction and a branch point inside the intron that serve as a recognition site for the spliceosome components. Spliceosome assembly starts with U1 snRNP base pairing with the 5' splice site. Association of the U2 snRNP with the branch point then yields complex A. Next, the U4/U6-U5 tri-snRNP is recruited to generate the pre-catalytic complex B. After some conformational rearrangement of the complex B, U1 and U4 snRNPs are released and complex B activated. The first catalytic step of splicing consists of a nucleophilic attack of the 2' OH group at the adenosine of the branch point on the 5' splice site. This leads to pre-mRNA cleavage at the 5' splice site and ligation of the 5' end of the intron to the branch adenosine, resulting in a so-called lariat structure bound by complex C. Subsequently, the second catalytic splice step cleaves the pre-mRNA at the 3' splice site. This cleavage is mediated by an attack of the 3' splice site by the 3' OH group of the 5' exon. This leads to the ligation of the 5' and 3' ends of the two exons and a spliced mRNA product. The spliceosome, the spliced mRNA and the intron dissociate

and, after some remodelling, the released snRNPs are recycled for another round of splicing (Will and Lührmann, 2011) (Figure 2).

Even though relatively rare in *S. cerevisiae*, metazoans and other eukaryotic organisms with are more complicated exon-intron structure like fission yeast are able to splice in an alternative way (Douglass *et al.*, 2019; Fair and Pleiss, 2017). Due to alternative splicing, those organisms are able to generate different variants of mRNA from a single pre-mRNA species by including or excluding specific exonic or intronic elements. Thereby, they expand their proteome tremendously from a limited number of genes. In humans, alternative splicing is observed for almost every gene (Chen and Manley, 2009; Keren *et al.*, 2010; Nilsen and Graveley, 2010).

#### *3.1.3.3 3' end formation*

The last step of mRNA processing is the 3' end processing and polyadenylation. During 3' end formation, the newly transcribed RNA is released from RNAPII through cleavage and a poly(A) tail attached to the 3' end, which is then bound by poly(A)-binding proteins (PABs). These processes are mediated by the cleavage and polyadenylation factor (CPF), which is mainly recruited by canonical motifs inside the 3' UTR and the phosphorylated $S_2$ of the CTD of RNAPII (Moore and Proudfoot, 2009; Tian and Graber, 2012). The conserved CPF-complex consists of more than 20 different proteins and is organized in three major units, which each possess a different enzymatic activity. The nuclease module (Cft2) cleaves the nascent RNA, the polymerase module (Pab1) adds the poly(A) tail, and the phosphatase module regulates 3' end processing (Casañal *et al.*, 2017). Already during synthesis of the poly(A) tail, PABs like Nab2 and Pab1 bind the poly(A) tail and stabilize the RNA (Dunn *et al.*, 2005; Mandel *et al.*, 2008; Soucek *et al.*, 2012). Apart from the PAB binding, the length of the poly(A) tail is also crucial for the stability of the RNA. In *S. cerevisiae,* the poly(A) nuclease complex shortens the poly(A) tail to ~60 nucleotides (~250 nucleotides in humans; (Chan *et al.*, 2011; Moore and Proudfoot, 2009)). The termination of polyadenylation as well as the length of the poly(A) tail are controlled by Nab2 (Brockmann *et al.*, 2012; Kelly *et al.*, 2010).

Alternative polyadenylation (APA) is a widespread mechanism across all eukaryotic organisms from yeast to humans (Aparicio *et al.*, 2004; Derti *et al.*, 2012; Jan *et al.*, 2011; Ozsolak *et al.*, 2010; Smibert *et al.*, 2012; Wilkening *et al.*, 2013). APA generates mRNAs with different 3' untranslated regions, which serves as an additional layer of gene regulation. Inclusion or exclusion of RBP or microRNA binding sites due to APA can affect mRNA stability, translation, nuclear export, or localization of the mRNA (Geisberg *et al.*, 2014; Neve *et al.*, 2016; Spies *et al.*, 2013; Wang *et al.*, 2008). For example, a

recent study revealed a mechanism by which the localization of membrane proteins is regulated by APA (Berkovits and Mayr, 2015).

### 3.1.4 Nuclear mRNA export

In eukaryotes, the transcription of protein-coding genes into mRNA and the translation of mRNAs into the corresponding proteins are spatially separated and occur in the nucleus and cytoplasm, respectively. Since the mRNAs carry the information of the protein-coding genes they have to be exported from the nucleus through the nuclear pore complexes (NPCs) to become available to the ribosomes in the cytoplasm (Figure 3). For smaller molecules of up to 30-60 kDa it is possible to diffuse through the NPC passively (Knockenhauer and Schwartz, 2016; Mohr *et al.*, 2009; Paine and Feldherr, 1972; Paine *et al.*, 1975). Because an export-competent mRNP consists of an mRNA and a variety of bound proteins, export through the NPCs has to be actively managed. NPCs are multiprotein complexes that consist of ~30 different nucleoporins (Nups) and have a molecular mass of about 50 MDa in *S. cerevisiae* (Alber *et al.*, 2007). Vertebrate NPCs are estimated to have a size of ~112 MDa (Reichelt *et al.*, 1990). The ~30 Nups, which are organized in smaller defined subclasses, are highly conserved in all eukaryotes (Cronshaw *et al.*, 2002; Devos *et al.*, 2014; Rout *et al.*, 2000; Tamura *et al.*, 2010). Nups can be categorized into four groups: Transmembrane, core scaffold (inner and outer ring), linker and phenylalanine-glycine (FG-repeats)-containing Nups. The NPC has a cylindric structure with a central opening, which measures approximately 40 nm in diameter and is comprised of eight surrounding spokes (Bui *et al.*, 2013; Panté and Kann, 2002). Besides the central channel, NPCs consist of a nuclear face with a basket and the cytoplasmic face with fibrils. Numerous largely unfolded FG-Nups directly bind the core scaffold and occupy the surface of the central channel from the nuclear to the cytoplasmic face (Alber *et al.*, 2007; Rout *et al.*, 2000). The FG-repeats have a key role in the transport through NPCs as they serve as binding sites for most cargo complexes (Radu *et al.*, 1995; Terry and Wente, 2007). At the nucleoplasmic side of the NPC, proteins like Mlp1, Pml1 or Pml39 are involved in the association of mRNPs with the NPC. To this aim, Mlp1 interacts with the different export adaptor proteins Nab2, Gbp2, Hrb1 and Npl3 (for more information see below). Furthermore, these proteins are important to ensure that wrongly or incompletely processed mRNAs are retained in the nucleus (Fasken *et al.*, 2008; Galy *et al.*, 2004; Hackmann *et al.*, 2014; Palancade *et al.*, 2005; Vinciguerra *et al.*, 2005). Besides its vital role in transport, the NPC also has an impact on the regulation of gene expression by localizing actively transcribed genes to the nuclear periphery, a process postulated decades ago as *gene gating* (Blobel, 1985).

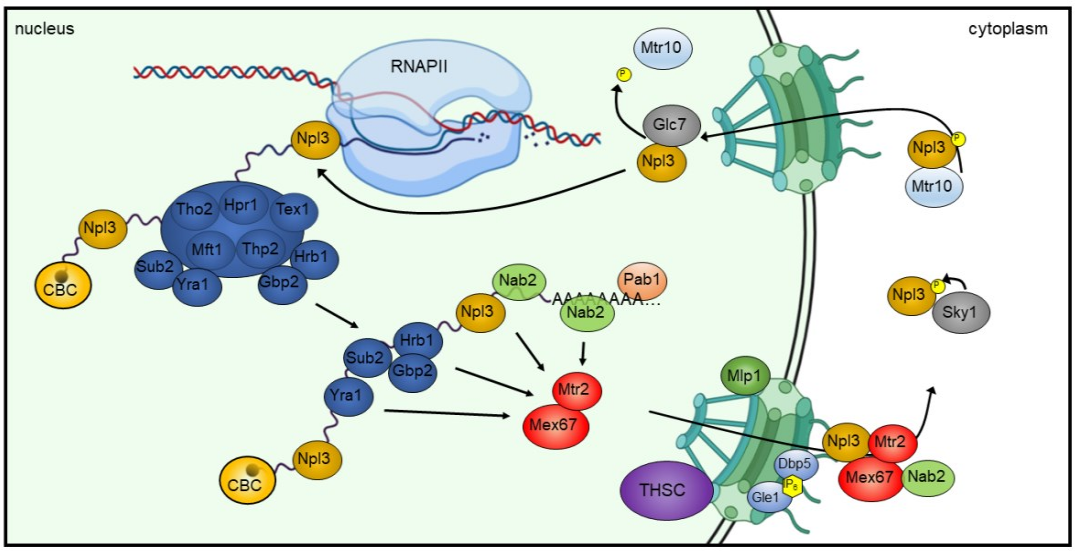



**Figure 3: Overview of nuclear mRNA export.** Npl3 and the THO/TREX complex is recruited to the mRNA co-transcriptionally. The subunits of TREX Sub2, Yra1, Hrb1 and Gbp2 are known to associate with the mRNA. Yra1, Hrb1, Nab2 and Npl3 serve as adaptor proteins that recruit the mRNA export-receptor Mex67-Mtr2 to the mRNP and thereby mediate the mRNA export. The THSC complex is located at the NPC and is able to recruit Mex67-Mtr2 as well. After the mRNP is transported to the cytoplasm, the DEAD-box helicase Dbp5 dissociates Mex67-Mtr2 and Nab2 from the mRNP at the cytoplasmic face of the NPC. Dbp5 is activated by $IP_6$ and Gle1 releases thereby the mRNA into the cytoplasm for translation. The shuttling protein Npl3 is released by phosphorylation by Sky1. The export factors are subsequently re-imported to the nucleus by importins (karyopherins) and are reused for a new cycle of mRNA export.

Due to the separation of nucleus and cytoplasm in eukaryotes, the evolution of proteins that bridge transcription and nuclear mRNA export became necessary. After successful assembly of an mRNP, correct 3' processing and release from RNAPII, the export-competent mRNP is ready to be exported. The export of the mRNPs is mediated by interaction with the heterodimeric export receptor Mex67-Mtr2 (Santos-Rosa *et al.*, 1998). Mex67 and Mtr2 are highly conserved from yeast to humans and termed NXF1-NXT1 or Tap-p15 (Katahira *et al.*, 1999). Although Mex67-Mtr2 are able to bind RNA, both bind with relatively low affinity (Katahira *et al.*, 1999). Therefore, both are thought to be recruited to mRNA mainly by adaptor proteins (Iglesias and Stutz, 2008; Kelly and Corbett, 2009; Köhler and Hurt, 2007).

Besides a variety of other proteins, the highly conserved TREX complex (transcription and export) is a major actor in this process and is recruited co-transcriptionally to the (pre-)mRNA in a length-dependent manner (Chávez *et al.*, 2000; Meinel *et al.*, 2013; Strässer *et al.*, 2002). The nonameric TREX complex is composed of the pentameric subcomplex THO (Hpr1, Tho2, Mft1, Thp2 and Tex1), the SR-like proteins Gbp2 and Hrb1, the DEAD-box RNA-helicase Sub2, and Yra1 (Chávez *et al.*, 2000; Strässer *et al.*, 2002). Sub2 and Yra1 form a heterodimer and recruit the mRNA export receptor Mex67-Mtr2 to the mRNA, which results in the export of the mRNP through the NPCs (Strässer and Hurt, 2000; Strässer and Hurt, 2001; Strässer *et al.*, 2002; Stutz *et al.*, 2000). It is

assumed that 3' end processing and the generation of an export-competent mRNP are linked by the CPF complex subunit Pcf11, which directly interacts with Yra1 (Johnson *et al.*, 2009). Moreover, it was shown that Mex67 is involved in the release of the CPF complex after 3' end processing (Qu *et al.*, 2009). The C-terminal region of Mex67 contains a nuclear transport factor 2-like domain (NTF2) that harbors two nuclear export signal motifs that are necessary for the interaction with the FG-repeat nucleoporins of the NPC (Grant *et al.*, 2002; Thakurta *et al.*, 2004). Mex67 is released from the mRNA at the cytoplasmic side of the NPC by the action of DEAD box helicase Dbp5, preventing the return of the mRNA to the nucleus. The ATPase activity of Dbp5 is activated by the cytoplasmic nucleoporin Gle1 and inositol hexakisphosphate (Adams and Wente, 2020; Alcázar-Román *et al.*, 2006; Lund and Guthrie, 2005; Tran *et al.*, 2007).

Besides Yra1, which binds only a subset of yeast mRNAs, there are other adaptor proteins that are involved in mRNA export (Hieronymus and Silver, 2003). It could be shown that recruitment of the THO subunit Hpr1 to a subset of mRNPs can be regulated by sumoylation via Siz1 and Siz2. However, the sumoylation of Hpr1 seems to have no effect on bulk mRNA export. The regulatory effect of the sumoylation of Hpr1 is restricted to acidic stress-induced genes (Bretes *et al.*, 2014).

Moreover, the THSC (TREX-2) complex, another conserved and multifunctional complex comprised of Sac3, Thp1, Sus1, Cdc31 and Sem1, is involved in mRNA export (Fischer *et al.*, 2004; Fischer *et al.*, 2002; Gallardo *et al.*, 2003; Rodríguez-Navarro *et al.*, 2004). Thp1 and Sac3 are responsible for attaching the mRNP to the NPC by interaction with the nucleoporins, Mex67-Mtr2 as well as the TREX component Yra1 (Fischer *et al.*, 2002). Even though the THSC complex is thought to be associated with the NPC and mainly involved in mRNA export, it was postulated that it is involved in transcription as well since the levels of mRNA drop in Δ*thp1* and Δ*sac3* mutant cells (Gallardo and Aguilera, 2001; Gallardo *et al.*, 2003). Furthermore, the THSC complex was shown to interact with the mediator complex and has an impact on the $S_5$ phosphorylation of the RNAPII CTD (Schneider *et al.*, 2015). Therefore, the THSC complex seems to be another component that links transcription and mRNA export.

Two other prominent examples of mRNA export adaptors are the poly(A)-binding protein Nab2 and the SR-like protein Npl3. Nab2 interacts directly with Mex67 and the export adaptor protein Yra1, which form a ternary complex (Iglesias *et al.*, 2010). Furthermore, overexpression of Nab2 is able to rescue the mRNA export defect of Δ*yra1* (TREX complex) and *Δthp1* (THSC complex) cells (Batisse *et al.*, 2009; Gallardo *et al.*, 2003). The interaction of those components and the fact that Nab2 is shuttling together with the

mRNA to the cytoplasm indicates that Nab2 is a key player in mRNA export. Like Mex67-Mtr2, Nab2 is released from the mRNA by the helicase Dbp5 and is transported back to the nucleus by the importer Kap104 (Aitchison *et al.*, 1996; Lee and Aitchison, 1999; Tran *et al.*, 2007).

## 3.2 The SR-like protein Npl3

In eukaryotes, the most abundant classes of RBPs are the heterogeneous nuclear ribonucleoproteins (hnRNPs) and in mammalian the serine-arginine-rich (SR-) proteins. Both groups of proteins are involved in several mRNA related-functions from transcription and pre-mRNA packaging over splicing and mRNA export to translation (Dreyfuss *et al.*, 2002; Shepard and Hertel, 2009). hnRNP proteins contain an RGG motif (closely repetitive RGG tripeptides and/or RG dipeptides) at their C-terminus that carries di-methylated arginines and one or more RRMs at their N-terminus (Boisvert *et al.*, 2005; Liu and Dreyfuss, 1995). SR-proteins are characterized by the presence of RRMs at their N-terminus and of a RS-rich region at their C-terminus, which are reversibly phosphorylated (Hertel and Graveley, 2005). In *S. cerevisiae*, the most abundant hnRNP is Npl3 (Ghaemmaghami *et al.*, 2003), which is also an SR-like protein. Hrb1 and Gbp2 are the only two other SR-like proteins characterized in *S. cerevisiae*. Npl3 contains two RRMs in the middle and an RS/RGG domain in its C-terminus. Furthermore, Npl3 contains APQE-repeats with unknown function in its N-terminus (Birney *et al.*, 1993; Deka *et al.*, 2008) (Figure 4).


**Figure 4: Schematic domain architecture of Npl3.** The RS/RGG domain is rich in arginine and serine residues (RS) and has several RG and RGG repeats.

Npl3 (nuclear protein localization 3; also known as nuclear polyadenylated RNA-binding 1 (Nab1), mRNA transport-defective 13 (Mtr13), mitochondrial targeting suppressor 1 (Mts1) or nucleolar protein 3 (Nop3)) was initially identified as a protein necessary for the localization of nuclear proteins and for processing pre-rRNAs (Bossie *et al.*, 1992; Russell and Tollervey, 1992). Shortly after, its role as an mRNA export factor was described (Kadowaki *et al.*, 1994; Lee *et al.*, 1996; Wilson *et al.*, 1994). Like the two TREX components Gbp2 and Hrb1, Npl3 was identified as a shuttling protein that travels together with the RNA to the cytoplasm and after release of the RNA back to the nucleus (Flach *et al.*, 1994; Häcker and Krebber, 2004; Lee *et al.*, 1996). However, its steady state location is nuclear (Wilson *et al.*, 1994). Furthermore, Npl3 was shown to be involved in chromatin modification, transcription elongation and termination, splicing, R-loop prevention and translation (Baierlein *et al.*, 2013; Bucheli and Buratowski, 2005;

Dermody *et al.*, 2008; Kress *et al.*, 2008; Moehle *et al.*, 2012; Pérez-Martínez *et al.*, 2020; Santos-Pereira *et al.*, 2013).

Npl3 is recruited co-transcriptionally by direct interaction with nascent mRNAs and RNAPII phosphorylated at $S_2$ of the CTD (Dermody *et al.*, 2008; Meinel *et al.*, 2013). Recently, it was shown that the cap binding complex contributes to the recruitment of Npl3 to actively transcribed genes (Sen *et al.*, 2019; Shen *et al.*, 2000). Transcript binding and release are regulated by a posttranslational phosphorylation cycle (Figure 3): Inside the nucleus, Npl3 is dephosphorylated by the phosphatase Glc7, enabling Npl3 to bind mRNAs and serve as an mRNA export adaptor for Mex67 to induce mRNA export (Gilbert and Guthrie, 2004; Lee *et al.*, 1996). Npl3 remains bound to the mRNA during export. In the cytoplasm, Npl3 is phosphorylated by the SR-protein kinase Sky1 at position S411 (Gilbert *et al.*, 2001; Siebel *et al.*, 1999). This phosphorylation induces its release from the mRNA and Npl3 is imported back to the nucleus by Mtr10 (Gilbert *et al.*, 2001; Pemberton *et al.*, 1997; Senger *et al.*, 1998). Besides the phosphorylation of Npl3 by Sky1, it has been shown that other kinases like the casein kinase 2 (CK2) phosphorylate Npl3 on other residues as well (Dermody *et al.*, 2008; Smolka *et al.*, 2007). The phosphorylation of those residues also promotes the dissociation of Npl3 from the mRNA and at the same time decreases the ability of Npl3 to compete with the termination factor Rna15 to regulate transcription termination (Dermody *et al.*, 2008). In addition, the activity of Npl3 is regulated by methylation of multiple arginines inside the RS/RGG domain by the methyltransferase Hmt1 (Henry and Silver, 1996; Siebel and Guthrie, 1996). The methylation of these arginines inhibits the phosphorylation by Sky1 and thereby negatively influences the nuclear import of Npl3 (Yun and Fu, 2000). Moreover, arginine methylation itself facilitates the nuclear export of Npl3 (Shen *et al.*, 1998). Besides its function in mRNA export as an adaptor protein for Mex67-Mtr2, Npl3 also functions as a nuclear export factor for the pre-60S ribosomal subunit (Hackmann *et al.*, 2011). However, during the export of ribosomal subunits Npl3 acts as an independent mediator by interacting with the 25S rRNA and Npl25 in contrast to the mRNA export with Mex67-Mtr2 (Hackmann *et al.*, 2011). Nevertheless, Mex67-Mtr2 also functions in the export of rRNAs (Yao *et al.*, 2007).

Since alternative splicing rarely occurs in *S. cerevisiae,* it is not surprising that classical SR-proteins that are mainly known for their function in splicing are missing in this organism and only three SR-like proteins are characterized (Blencowe *et al.*, 1999). However, only Npl3 promotes pre-mRNA splicing, while Gbp2 and Hrb1 are required for quality control of spliced mRNAs (Busch and Hertel, 2012; Hackmann *et al.*, 2014; Kress *et al.*, 2008). Npl3 interacts with the U1 and U2 snRNPs and is involved in the recruitment

of those snRNPs to the splice site (Kress *et al.*, 2008). The regulatory function of Npl3 in early spliceosome assembly is directly linked to the mono-ubiquitylation of histone H2B (H2Bub1) at position K124 (Moehle *et al.*, 2012); although there is no lysine at position 123, the ubiquitylated side chain is referred to as K123 in the literature. H2Bub1 is mediated by the Bre1-Rad6 heterodimer (Kim and Roeder, 2009). In this context, Npl3 was found to physically interact with the E3 ligase Bre1 (Moehle *et al.*, 2012). In higher eukaryotes, monoubiquitylation of H2B influences alternative splicing by decreasing the level of U1 and U2 snRNPs resulting in exon skipping events (Zhang *et al.*, 2013). Although there is little alternative splicing in budding yeast, H2Bub1 marks the intron/exon structure, which is required for efficient splicing (Shieh *et al.*, 2011). Moreover, the same histone modification affects the ubiquitylation of Swd2 (subunit of the COMPASS (Set1C) histone methyltransferase complex), which in turn results in a defective recruitment of the Mex67 adaptor proteins Yra1 and Nab2 and therefore in an mRNA export defect (Vitaliano-Prunier *et al.*, 2012).

Npl3 associates with the mRNA co-transcriptionally at the beginning of transcription and stimulates the elongation activity of RNAPII through interaction with its CTD (Dermody *et al.*, 2008; Lei *et al.*, 2001; Meinel *et al.*, 2013). Furthermore, Npl3 participates in transcription termination and 3' end formation (Bucheli and Buratowski, 2005), although the role of Npl3 in this process is controversial. Npl3 was shown to suppress transcriptional termination and the 3' end formation on cryptic or weak termination sites, containing UA-rich elements, by competing with other 3' end processing factors like Rna15 (Bucheli and Buratowski, 2005; Bucheli *et al.*, 2007). The competition between Npl3 and Rna15 exists no longer when Npl3 is phosphorylated by casein kinase 2 (CKA2; (Dermody *et al.*, 2008)). However, Npl3 was shown to promote transcriptional termination as well (Holmes *et al.*, 2015).

Although Npl3 is phosphorylated by Sky1 in the cytoplasm (Gilbert *et al.*, 2001; Siebel *et al.*, 1999), is thereby released from the mRNA and transported back to the nucleus, it was shown that at least some Npl3 molecules remain with the mRNA through all steps of translation. Npl3 supports monosome formation to initiate translation (Baierlein *et al.*, 2013) and can act as a translation suppressor as well (Windgassen *et al.*, 2004). The suppression of translation is probably mediated by the RGG domain of Npl3 since this domain was shown to be a binding-motif for eIF4G and act thereby as a general translation suppressor (Rajyaguru and Parker, 2012; Rajyaguru *et al.*, 2012). Moreover, Npl3 is involved in promoting accurate translation termination (Estrella *et al.*, 2009).

Moreover, Npl3 is involved in the maintenance of genome stability. Genome instability is defined as an accumulation of DNA double-strand breaks, hyperrecombination and increased levels of chromosome loss (Langie *et al.*, 2015). During replication fork progression, DNA-RNA hybrids that displace the non-template DNA strand, also known as R-loops, can form co-transcriptionally and induce transcription-associated recombination. In a ChIP-chip analysis, cells that lack Npl3 showed genome-wide, R-loop-dependent replication damage (Santos-Pereira *et al.*, 2013). Furthermore, Npl3 seems to be involved in the repair of DNA double strand breaks. Δ*npl3* cells are highly sensitive to double strand break-inducing chemicals and to the expression of EcoRI (McKinney *et al.*, 2013; Santos-Pereira *et al.*, 2013). Npl3 was shown to regulate the level of Exo1 that plays an important role during DNA repair (Colombo *et al.*, 2017). Recently, Npl3 was shown to associate with telomeres in an R-loop-dependent manner. At critically short telomeres, R-loops become stabilized and induce elongation of those telomeres by homology-directed repair. Npl3 seems to be involved in stabilizing those R-loops (Pérez-Martínez *et al.*, 2020).

Since Npl3 is involved in so many processes related to RNA metabolism, direct RNA binding is required. How the RNA recognition of Npl3 works is still not fully understood. Npl3 contains two RRM domains and an RGG domain which are all known to bind RNA. The structure of the two RRM domains was solved in free state by NMR (Deka *et al.*, 2008; Skrisovska and Allain, 2008). Although Npl3 has no clear preference for any RNA sequence motif according to cross-linking and analysis of cDNA (CRAC; (Holmes *et al.*, 2015)), crosslinking experiments showed that Npl3 prefers UA and UG-rich RNA sequences (Bucheli *et al.*, 2007). This coincides with its function in 3' end formation since polyadenylation signals are rich in UA and UG elements. Moreover, it seems that Npl3 binds five-fold more strongly to UG-rich RNA than to UA-rich RNAs of the same length, as shown by electrophoretic mobility shift assay (Deka *et al.*, 2008). It was proposed that this binding preference is mediated by RRM2 of Npl3 (Cléry *et al.*, 2013; Deka *et al.*, 2008).

## 3.3 Aims and scope of this study

Eukaryotic cells express a huge variety of RNA binding proteins (RBPs) that act in various cellular processes and in almost every aspect of RNA metabolism. RBPs cover nascent (pre-)mRNAs as soon as these emerge from RNA polymerase II and package the mRNA into a messenger ribonucleoprotein particle (mRNP). Hundreds of mRNA-binding proteins have been identified in *S. cerevisiae* and mammalian cells by proteome-wide approaches (Castello *et al.*, 2012; Caudron-Herger *et al.*, 2019; Klass *et al.*, 2013; Mitchell *et al.*, 2013). In *S. cerevisiae,* countless RBPs that are involved in mRNP assembly, nuclear mRNA export and mRNA processing are known. The principal functions of these RBPs have been described with the help of gene knockouts, deletion of entire protein domains as well as thermosensitive mutations. Moreover, RNA-binding motifs and bound RNAs have been identified by next-generation sequencing combined with methods like PAR-CLIP (photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation) for many RBPs in *S. cerevisiae* (Baejen *et al.*, 2014; Hafner *et al.*, 2010; Tuck and Tollervey, 2013). However, the converse information, the identification of the amino acid residues involved in mRNA binding for each RBP, is largely missing. RNP$^{XL}$, a proteomic approach that combines mass spectrometry with UV-crosslinking was developed a few years ago (Kramer *et al.*, 2014). The method allows the identification of the crosslinked amino acids within the proteins that are in close proximity to RNA.

The aim of this study is to identify the *in vivo* RNA interaction sites of RBPs that are involved in mRNA packaging and export (THO/TREX complex, Npl3, Sto1-Cbc2, Nab2, Tho1, THSC/TREX2 complex and Mex67-Mtr2) using the RNP$^{XL}$ pipeline. The identification of RNA-binding sites will be used to analyze how direct interactions between RBPs and mRNA influence the above-mentioned processes. The putative binding sites are selected according to their level of conservation and the corresponding amino acids are mutated to disrupt RNA binding.

The second aim of this Ph.D. project is to figure out the role of the mRNA-binding activity of Npl3 on the molecular level. Npl3 is a SR-like protein with multiple functions in mRNA processing, transcription, nuclear mRNA export, splicing and translation (see 3.2 for details). To reduce or inhibit the RNA-binding activity of the putative RNA-binding sites of Npl3, mutants will be generated at the identified crosslink positions. To confirm that RNA binding of the putative binding sites is impaired, RNA-immunoprecipitation and NMR analysis will be performed. To get a first indication whether the mutated binding sites have a significant impact on cell fitness, cell growth will be analyzed. Afterwards,

the functional consequences of the RNA-binding mutants will be investigated. Among others, the nuclear mRNA export of the mutant cells will be determined. Moreover, global expression levels and stability of transcripts will be investigated transcriptome-wide to determine whether this is impaired in the Npl3 RNA-binding mutants. The impact of binding site mutations on the recruitment of Npl3 and, as a consequence, of other RBPs to nuclear mRNPs will be assessed by tandem affinity purification (TAP) and subsequent quantitative Western blot (WB) and mass spectrometric analyses. Furthermore, the recruitment of Npl3 to the site of transcription will be tested by chromatin immunoprecipitation and RNA immunoprecipitation. In summary, the investigations will give new insights about the significance of mRNA binding of Npl3 and hopefully provide a more detailed view of the functions on Npl3.

# 4 Material

## 4.1 Chemicals and consumables

**Table 1: Chemicals and Consumables**

| Chemicals and Consumables | Supplier |
|---|---|
| 2-Propanol | Carl Roth |
| 4-Thiouracil | abcr GmbH |
| 5-Fluoroorotic acid (5-FOA) | Apollo Scientific Ltd |
| Acetic acid | VWR Chemicals |
| Acrylamide (29:1) 40 % | AppliChem GmbH |
| Adenine hemisulfate salt | Sigma-Aldrich |
| Agar Bacteriology grade | Applichem GmbH |
| Agarose | Applichem GmbH |
| Ammonium persulfate (APS) | VWR Chemicals |
| Ampicillin | Applichem GmbH |
| ANTI-FLAG® M2 Affinity Gel | Sigma-Aldrich |
| Bacto™ Peptone | BD Biosciences |
| Bacto™ Yeast extract | BD Biosciences |
| Benzamidine HCl | MP Biomedicals |
| Boric acid ($BH_3O_3$) | Applichem GmbH |
| 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 |
| Dextran sulfate | Sigma-Aldrich |
| D-Galactose | Applichem GmbH |
| D-Glucose Monohydrate | Sigma-Aldrich |
| Dimethyl sulfoxide (DMSO) | Grüssing GmbH |
| Dipotassium phosphate ($K_2HPO_4$) | Grüssing GmbH |
| Disodium phosphate ($Na_2HPO_4$) | Carl Roth |
| Dithiothreitol (DTT) | Sigma-Aldrich |
| dNTPs (dATP, dTTP, dCTP, dGTP) | Thermo Fisher Scientific |
| D-Sorbitol | Carl Roth |
| Dynabeads™ M-280 Tosylactivated | Invitrogen |
| Dynabeads™ Protein G | Invitrogen |
| *E. coli* tRNA | Roche diagnostics |
| ECL Solution | Applichem |
| 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 |
| HDGreen™ DNA stain | Intas |
| HEPES | Carl Roth |
| Herring Sperm DNA | ThermoFisher (Invitrogen) |
| Hydrochloric acid (HCl) | Carl Roth |
| IGEPAL CA-630 | Sigma-Aldrich |
| IgG Sepharose 6 Fast Flow | GE Healthcare |
| Imidazole | Merck |
| Isopropyl β-D-1-thiogalactopyranoside (IPTG) | Carl Roth |
| Kanamycin | Merck |
| L-Arginine-HCl | Biomol GmbH |
| L-Aspartic acid | Sigma-Aldrich |
| Leupeptin (Hemisulfate) | Carl Roth |
| L-Histidine | Sigma-Aldrich |
| L-Isoleucine | Sigma-Aldrich |
| Lithium acetate (LiOAc) | 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 |
| Magnesium sulfate ($MgSO_4$) | Carl Roth |
| Methanol | Merck-Millipore |
| Monopotassium phosphate ($KH_2PO_4$) | Carl Roth |
| Monosodium phosphate ($NaH_2PO_4$) | Merck |
| Nicotinamide adenine dinucleotide (NAD) | Sigma Aldrich |
| Octylphenoxypolyethoxyethanol (IGEPAL CA-630) | 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 |
| Polylysine | Sigma-Aldrich |
| Polysorbate 20 (Tween 20) | Merck |
| Polyvinylpyrrolidone (PVP) | Sigma-Aldrich |
| Ponceau S | Serva |
| Potassium chloride (KCl) | ORG Laborchemie |
| Potassium hydroxide (KOH) | Merck |
| Powdered milk, fat free, blotting grade | Carl Roth |
| Protino® Ni-NTA Agarose | Macherey-Nagel |
| Rothi®-Mount FluorCare DAPI | Carl Roth |
| Roti®-Aqua-Phenol | Carl Roth |
| Salmon sperm DNA (SSD) | Applichem GmbH |
| Sodium acetate (NaOAc) | Merck |
| Sodium chloride (NaCl) | Merck |
| Sodium citrate | Carl Roth |
| Sodium deoxycholate | Sigma-Aldrich |
| Sodium dodecyl sulfate (SDS) | Serva |
| Sodium hydroxide (NaOH) | Merck |
| Sulfosalicylic acid | Merck |

| | |
|---|---|
| Tetramethylethylendiamin (TEMED) | Carl Roth |
| Thiolutin | Sigma-Aldrich |
| Trichloroacetic acid (TCA) | Merck |
| Tris(hydroxymethyl)aminomethane (Tris) | Applichem GmbH |
| Triton X-100 | Applichem GmbH |
| Tryptone BioChemica | Applichem GmbH |
| Uracil | Sigma-Aldrich |
| Yeast nitrogen base, w/o amino acids | Formedium |

## 4.2 Equipment and devices

**Table 2: Equipment and devices**

| Name | Supplier |
|---|---|
| 70 Ti | Beckman Coulter |
| ÄKTA purifier | GE Healthcare |
| AM100, micro scale | Mettler-Toledo |
| Apollo®, liquid nitrogen container | Cryotherm |
| Avanti JXN-26 Centrifuge | Beckman Coulter |
| Axio observer, fluorescence microscope | Zeiss |
| Bioruptor UCD-200, Sonication System | Diagenode |
| BLX-254, UV-crosslinker | Vilber Lourmat |
| BLX-365, UV-crosslinker | Vilber Lourmat |
| ChemoCam Imager ECL HR 16-3200 | Intas |
| Duomax 1030, tumbling shaker | Heidolph Instruments |
| EPS 301, electrophoresis power supply | GE Healthcare |
| FastPrep-24$^{TM}$ 5G | MP Biomedicals |
| Freezer/Mill® 6870D | Spex®SamplePrep |
| Gel iX20, Transilluminator/gel docu | Intas |
| Hera safe, laminar flow cabinet | Thermo Fisher Scientific |
| HeraFreeze HFU T Series | Thermo Scientific |
| HT Multitron Pro shaking incubator | Infors |
| HXP 120 V, light source | Kübler Codix |
| IKA® KS 4000 ic control, shaking incubator | IKA Labortechnik |
| IKAMAG® RCT, magnetic stirrer | IKA Labortechnik |
| Incubator with HT Labotron, shaker | Aqua Lytic / Infors |
| Incubators | Memmert |
| Innova®44 shaking incubator | Eppendorf / New Brunswick |
| JLA-8.1, JA-25.50, JA-10 | Beckman Coulter |
| Lab phenomenal pH 1000L, pH meter | VWR |
| LED bluelight transilluminator | Nippon genetics |
| Megafuge 40R | Thermo Scientific, Heraeus |
| Milli-Q® integral water purification system | Merck |
| Mini-Protean® Tetra Electrophoresis Cell | Bio-Rad Laboratories |
| ND-1000, Spectrophotometer | NanoDrop |
| Optima XPN-80 Ultracentrifuge | Beckman Coulter |
| PeqStar XS Thermocycler | Peqlab |
| Pipetboy acu | IBS Integra Biosciences |
| PM2000, scale | Mettler-Toledo |
| Quant Studio 3, Real Time PCR System | Applied Biosystems, Thermo Fisher Scientific |
| RCT basic, magnetic stirrer | IKA Labortechnik |

| | |
|---|---|
| Research Pipettes 2.5, 10, 20, 100, 200, 1000, 5000 | Eppendorf |
| Rotator | neoLab |
| SBH130D, block heater | Stuart® |
| Sonifier 250 | Branson Ultrasonics™ |
| Sunrise Microplate Absorbance Reader | Tecan Group |
| Superdex 75 10/300 GL | GE Healthcare |
| SW22, shaking waterbath | Julabo |
| T3 Thermocycler | Biometra |
| Tabletop Centrifuge 5424, 5424R | Eppendorf |
| Tabletop Centrifuge 5430, 5430R | Eppendorf |
| Thermomixer 5436 | Eppendorf |
| Trans-Blot® Turbo Transfer System | Bio-Rad Laboratories |
| Typhoon FLA 9500 | GE Healthcare |
| Unichromat 1500 | Uniequip |
| Vakulan CVC 3000 | Vacuubrand |
| Vari-X-Link | $UVO_3$ Ltd. |
| VF2, vortex mixer | IKA Labortechnik |
| VX-150, autoclave | Systec |
| WT 12, tumbling shaker | Biometra |

## 4.3 Buffers, Media and Solutions

### 4.3.1 Water

The used water for any solutions and media was filtered/purified by the Milli-Q-synthesis System (Millipore) and autoclaved at 120°C for 20 min. Heat sensitive solutions and buffers were sterile filtered through a 0.22 µm filter.

### 4.3.2 Media

Lysogeny broth (LB) for 1 L
10 g peptone
5 g yeast extract
5 g NaCl
adjust to pH 7.2 (NaOH)
(15 g agar for plates)

SOC for 100 mL
2 g tryptone
0.5 g yeast extract
10 mM NaCl
0.5 mM KCl
10 mM $MgCl_2$
10 mM $MgSO_4$
adjust to pH 7.0 (NaOH)

Synthetic dropout medium (SDC) for 1 L
6.75 g yeast nitrogen base (w/o aa)
0.6 g complete synthetic amino acid mix (CSM)
20 g glucose
10 mL of each 100x aa stock except the required drop out*
adjust to pH 5.5 (NaOH)
(15 g agar for plates // 1 g 5-FOA if required)
*for 4-tU crosslinking only 6.7 mL 100x uracil stock was added to SDC-medium

Yeast complete medium (YPD) for 1 L
10 g yeast extract
20 g peptone
20 g glucose
adjust to pH 5.5 (HCl)
(15 g agar for plates)

### 4.3.3 Buffers and solutions

**Cloning**

50x TAE buffer
2 M TRIS
1 M NaOAc
50 mM EDTA (pH 8.0)
adjust to pH 8.0 (acetic acid)

5x Isothermal reaction buffer
25 % PEG-8000
500 mM TRIS (pH 7.5)
50 mM $MgCl_2$
50 mM DTT
1 mM of each dNTP
5 mM $NAD^+$

6 x Agarose loading dye
0.03 % Bromphenol blue
0.03 % Xylen cyanol
60 % Glycerin
60 mM EDTA
10 mM TRIS

Gibson assembly master mix
1 × Isothermal reaction buffer
4 U/µL T5 exonuclease
4 U/µL Taq DNA ligase
25 U/mL Phusion DNA polymerase

**SDS-PAGE**

4x Separating buffer
1.5 M TRIS (pH 8.8)
8 mM EDTA
0.6 % SDS

Separating gel (10 %)
3 mL acrylamide (40 %; 29:1)
3 mL 4x separating buffer
6 mL $H_2O$
100 µL 10 % APS
20 µL TEMED

6x SDS loading dye
7 mL stacking buffer
40 % glycerol
10 % SDS
0.5 M DTT
0.03 % bromphenol blue
1 % β-mercaptoethanol

10x Running buffer
250 mM TRIS
1.9 M Glycin
1 % SDS

4x Stacking buffer
0.5 M TRIS (pH 6.8)
8 mM EDTA
0.6 % SDS

Stacking gel (4 %)
400 µL acrylamid (40 %; 29:1)
1 mL 4x stacking buffer
2.6 mL $H_2O$
30 µL 10 % APS
10 µL TEMED

Hot-Coomassie
0.5 % Coomassie R250
25 % isopropanol
10 % acetic acid

Detain solution
10 % acetic acid

**Western blotting**

10x TBS-T (TRIS buffered saline + tween)
500 mM TRIS (pH 7.5)
1.5 M NaCl
1 % Tween 20

Semi dry Western blot buffer
25 mM TRIS
192 mM Glycin
20 % Methanol

**Yeast transformation**

Solution I
1x TE
100 mM LiOAc

10x TE
100 mM TRIS (pH 7.5)
10 mM EDTA (pH 8.0)

Solution II
1x TE
100 mM LiOAc
40 % PEG 3,800 (or PEG 4,000)

**Tandem affinity purification (TAP)**

Wash buffer
50 mM TRIS (pH 7.5)
1.5 mM $MgCl_2$
200 mM KCl
0.15 % NP 40
1 mM DTT
(1x protease inhibitor)

100x protease inhibitor (in 50 mL EtOH)
6.85 mg pepstatin A
1.42 mg leupeptin hemisulfat
850 mg PMSF
1.65 g benzamide HCl

**Fluorescence *in situ* hybridization (FISH)**

Prehybridisation buffer
50 % formamide
10 % dextran sulphate
125 µg/mL tRNA (*E. coli*)
500 µg/mL herring sperm DNA
4 x SSC
1 x Denhardt's solution

Wash buffer
1.2 M sorbitol
100 mM $KPO_4$ (pH 6.4)

20 x SSC (pH 7.0)
3 M NaCl
300 mM sodium citrate

50x Denhardt's solution
1 % polyvinylpyrrolidone (PVP)
1 % bovine serum albumin (BSA)
1 % ficoll-400

**Chromatin immunoprecipitation (ChIP)**

Low-salt buffer
50 mM HEPES (pH 7.5)
150 mM NaCl
1 mM EDTA
1 % Triton-X 100
0.1 % SDS
0.1 % sodium deoxycholate

TLEND
10 mM TRIS (pH 8.0)
0.25 M LiCl
1 mM EDTA
0.5 % NP-40
0.5 % sodium deoxycholate

10x TE
100 mM TRIS (pH 7.5)
10 mM EDTA

High-salt buffer
50 mM HEPES (pH 7.5)
500 mM NaCl
1 mM EDTA
1 % Triton-X 100
0.1 % SDS
0.1 % sodium deoxycholate

Elution buffer
50 mM TRIS (pH 7.5)
10 mM EDTA
1 % SDS

**RNA immunoprecipitation (RIP)**

RNA IP-buffer
25 mM TRIS (pH 7.5)
150 mM NaCl
2 mM MgCl2
0.2 % Triton X 100
500 µM DTT

## 4.4 Organisms

### 4.4.1 Yeast strains

**Table 3: Yeast strains**

| Yeast strain | Genotype | Reference |
|---|---|---|
| RS453 | MAT a, ade2-1, his3-11,15, ura3-52, leu2-3,112, trp1-1, can1-100, GAL+ | (Strässer and Hurt, 2000) |
| SUB2-TAP | MAT a, ade2-1, his3-11,15, ura3-52, leu2-3,112, trp1-1, can1-100, GAL+, SUB2-CBP-TEV-2x protA::TRP1-KL | (Strässer *et al.*, 2002) |
| HPR1-FTpA | MAT a, ade2-1, his3-11,15, ura3-52, leu2-3,112, trp1-1, can1-100, GAL+, HPR1-FLAG-TEV-2x protA::HIS3MX4 | This study |
| TAP-NPL3 | MAT a, ade2-1, his3-11,15, ura3-52, leu2-3,112, trp1-1, can1-100, GAL+, 2x protA-TEV-NPL3::TRP1-KL | This study |
| THO1-FTpA | MAT a, ade2-1, his3-11,15, ura3-52, leu2-3,112, trp1-1, can1-100, GAL+, THO1-FLAG-TEV-2x protA::HIS3MX4 | This study |
| NAB2-FTpA | MAT a, ade2-1, his3-11,15, ura3-52, leu2-3,112, trp1-1, can1-100, GAL+, NAB2-FLAG-TEV-2x protA::HISMX4 | This study |
| MTR2-FTpA | MAT a, ade2-1, his3-11,15, ura3-52, leu2-3,112, trp1-1, can1-100, GAL+, MTR2-FLAG-TEV-2x protA::HIS3MX4 | This study |
| CBC2-FTpA | MAT a, ade2-1, his3-11,15, ura3-52, leu2-3,112, trp1-1, can1-100, GAL+, CBC2-FLAG-TEV-2x protA::HIS3MX4 | This study |
| THP1-TAP | MAT a, ade2-1, his3-11,15, ura3-52, leu2-3,112, trp1-1, can1-100, GAL+, THP1-CBP-TEV-2x protA::TRP1-KL | This study |
| W303 | MATa, ura3-1, trp1-1, his3-11,15, leu2-3,112, ade2-1, can1-100, GAL+ | (Thomas and Rothstein, 1989) |
| npl3-shuffle | MATa, ura3-1, trp1-1, his3-11,15, leu2-3,112, ade2-1, can1-100, GAL+, NPL3::HIS3, pRS316-NPL3 | (Röther, 2007) |
| npl3-shuffle, CBC2-TAP | MATa, ura3-1, trp1-1, his3-11,15, leu2-3,112, ade2-1, can1-100, GAL+, NPL3::HIS3, pRS316-NPL3, CBC2-CBP-TEV-2x protA::TRP1-KL | This study |
| npl3-shuffle, CBC2-TAP, MEX67-HA | MATa, ura3-1, trp1-1, his3-11,15, leu2-3,112, ade2-1, can1-100, GAL+, NPL3::HIS3, pRS316-NPL3, CBC2-CBP-TEV-2x protA::TRP1-KL, MEX67-HA::KanMX4 | This study |
| npl3-shuffle, CBC2-TAP, HPR1-HA | MATa, ura3-1, trp1-1, his3-11,15, leu2-3,112, ade2-1, can1-100, GAL+, NPL3::HIS3, pRS316-NPL3, CBC2-CBP-TEV-2x protA::TRP1-KL, HPR1-HA::KanMX4 | This study |
| npl3-shuffle, CBC2-TAP, YRA1-HA | MATa, ura3-1, trp1-1, his3-11,15, leu2-3,112, ade2-1, can1-100, GAL+, NPL3::HIS3, pRS316-NPL3, CBC2-CBP-TEV-2x protA::TRP1-KL, YRA1-HA::KanMX4 | This study |
| npl3-shuffle, MEX67-HA | MATa, ura3-1, trp1-1, his3-11,15, leu2-3,112, ade2-1, can1-100, GAL+, NPL3::HIS3, pRS316-NPL3, MEX67-HA::KanMX4 | This study |
| npl3-shuffle, HPR1-HA | MATa, ura3-1, trp1-1, his3-11,15, leu2-3,112, ade2-1, can1-100, GAL+, NPL3::HIS3, pRS316-NPL3, HPR1-HA::KanMX4 | This study |
| npl3-shuffle, SUB2-FTpA | MATa, ura3-1, trp1-1, his3-11,15, leu2-3,112, ade2-1, can1-100, GAL+, NPL3::HIS3, pRS316-NPL3, SUB2-FLAG-TEV-2x protA::TRP1-KL | This study |
| npl3-shuffle, THO1-FTpA | MATa, ura3-1, trp1-1, his3-11,15, leu2-3,112, ade2-1, can1-100, GAL+, NPL3::HIS3, pRS316-NPL3, THO1- FLAG -TEV-2x protA::TRP1-KL | This study |
| npl3-shuffle, MEX67-FTpA | MATa, ura3-1, trp1-1, his3-11,15, leu2-3,112, ade2-1, can1-100, GAL+, NPL3::HIS3, pRS316-NPL3, MEX67- FLAG -TEV-2x protA::TRP1-KL | This study |
| npl3-shuffle, HPR1-FTpA | MATa, ura3-1, trp1-1, his3-11,15, leu2-3,112, ade2-1, can1-100, GAL+, NPL3::HIS3, pRS316-NPL3, HPR1- FLAG -TEV-2x protA::TRP1-KL | This study |
| npl3-shuffle, YRA1-TAP | MATa, ura3-1, trp1-1, his3-11,15, leu2-3,112, ade2-1, can1-100, GAL+, NPL3::HIS3, pRS316-NPL3, YRA1-CBP-TEV-2x protA::TRP1-KL | This study |

| | | |
|---|---|---|
| npl3-shuffle, NAB2-FTpA | MATa, ura3-1, trp1-1, his3-11,15, leu2-3,112, ade2-1, can1-100, GAL+, NPL3::HIS3, pRS316-NPL3 | This study |
| Δ*dhh1* | MATa, ura3-1, trp1-1, his3-11,15, leu2-3,112, ade2-1, can1-100, GAL+, DHH1::KanMX4 | This study |

### 4.4.2 *E. coli* strains

For any kind of cloning DH5α cells were used. BL21 Star (DE3) cells were used for heterologous protein expression.

**Table 4: *E. coli* strains**

| *E. coli* strain | Genotype | Reference |
|---|---|---|
| DH5α | F– endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20 φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(rK– mK+), λ– | (Taylor *et al.*, 1993) |
| BL21 Star (DE3) | F– ompT gal dcm lon hsdSB(rB–mB–) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB+]K-12(λS) rne131 | (Wood, 1966) |

## 4.5 Plasmids

**Table 5: Plasmids**

| Plasmid | Description | Reference |
|---|---|---|
| pBS1479 | For genomic C-terminal TAP-tagging (CBP-TEV-2x protein A), *TRP1-KL* | (Puig *et al.*, 2001) |
| pFA6a-FTpA-HIS3MX4 | For genomic C-terminal FTpA-tagging (FLAG-TEV-2x protein A), *HIS3MX4* | (Kressler *et al.*, 2012) |
| pRS316-*NPL3* | pBlueScript based yeast centromere vector with *URA3* marker and ORF + 500 bp of 5' and 300 bp of 3' UTR of *NPL3* | This study |
| pRS315 | pBlueScript based yeast centromere vector with *LEU2* marker | (Sikorski and Hieter, 1989) |
| pRS315-*NPL3* | ORF + 500 bp of 5' and 300 bp of 3' UTR of *NPL3* was cloned into pRS315 | This study |
| pRS315-*npl3-F160Y,F162Y,F165Y* | *npl3-F160Y,F162Y,F165Y* | This study |
| pRS315-*npl3-F160Y,F162Y* | *npl3-F160Y,F162Y* | This study |
| pRS315-*npl3-F160Y,F165Y* | *npl3-F160Y,F165Y* | This study |
| pRS315-*npl3-F162Y,F165Y* | *npl3-F162Y,F165Y* | This study |
| pRS315-*npl3- F160Y* | *npl3-F160Y* | This study |
| pRS315-*npl3-F162Y* | *npl3-F162Y (RRM1)* | This study |
| pRS315-*npl3- F165Y* | *npl3-F165Y* | This study |
| pRS315-*npl3-P196D,A197D* | *npl3-P196D,A197D (Linker)* | This study |
| pRS315-*npl3-P196D* | *npl3-P196D* | This study |
| pRS315-*npl3-A197D* | *npl3-A197D* | This study |
| pRS315-*npl3-F245I* | *npl3-F245I (RRM2)* | This study |
| pRS315-*npl3-F245Y* | *npl3-F245Y* | This study |
| pRS315-*npl3-RRM1-Linker* | *npl3-F162Y,P196D,A197D* | This study |
| pRS315-*npl3-RRM1-RRM2* | *npl3-F162Y,F245I* | This study |
| pRS315-*npl3-Linker-RRM2* | *npl3-P196D,A197D,F245I* | This study |
| pRS315-*npl3-RRM1-Linker-RRM2* | *npl3-F162Y,P196D,A197D,F245I* | This study |
| pRS315-*npl3-L219S* | *npl3-L219S* | This study |
| pRS315-*npl3-P196L,A197D* | *npl3-P196L,A197D* | This study |

| | | |
|---|---|---|
| pRS315-*npl3-P196D,A197K* | *npl3-P196D,A197K* | This study |
| pRS315-*npl3-P196L,A197K* | *npl3-P196L,A197K* | This study |
| pRS315-*TAP-NPL3* | *2x prot A-TEV-CBP-NPL3* | This study |
| pRS315-*TAP-npl3-RRM1* | *2x prot A-TEV-CBP-npl3-F162Y* | This study |
| pRS315-*TAP-npl3-Linker* | *2x prot A-TEV-CBP-npl3-P196D,A197D* | This study |
| pRS315-*TAP-npl3-RRM2* | *2x prot A-TEV-CBP-*n*pl3-F245I* | This study |
| pRS425 | *high copy number plasmid, 2µ, with LEU2 maker* | (Sikorski and Hieter, 1989) |
| pRS425-*NPL3* | *ORF + 500 bp of 5' and 300 bp of 3' UTR of NPL3 was cloned into pRS425* | This study |
| pRS425-*npl3-F162Y* | *npl3-F162Y (RRM1)* | This study |
| pRS425-*npl3-P196D,A197D* | *npl3-P196D,A197D (Linker)* | This study |
| pRS425-*npl3-F245I* | *npl3-F245I (RRM2)* | This study |
| pT7-*His$_6$-TEV-NPL3*$^{120-280}$ | Truncated version of *NPL3* for recombinant expression in *E. coli.* used for NMR | This study |
| pT7-*His$_6$-TEV-npl3-RRM1*$^{120-280}$ | Truncated version of *npl3-RRM1* mutant for recombinant expression in *E. coli.* used for NMR | This study |
| pT7-*His$_6$-TEV-npl3-Linker*$^{120-280}$ | Truncated version of *npl3-Linker* mutant for recombinant expression in *E. coli.* used for NMR | This study |
| pT7-*His$_6$-TEV-npl3-RRM2*$^{120-280}$ | Truncated version of *npl3-RRM2* mutant for recombinant expression in *E. coli.* used for NMR | This study |

## 4.6 Oligos

**Table 6: Oligos used for FISH and RNA-Binding assays**

| Name | Sequence | Assay |
|---|---|---|
| Oligo($dT_{50}$)-Cy3 | 50 x T coupled with Cy3 fluorescent dye | FISH |
| UAA-motif | AGCACCGUAAAGA | NMR/ITC |
| UGG-motif | AGCACCGUGGAGA | NMR/ITC |

**Table 7: Oligonucleotides used for cloning**

| Name | Sequence |
|---|---|
| pRS315 fwd | GGCCAGTGAATTGTAATACGACTCA |
| pRS315 rev | CCCTCACTAAAGGGAACAAAAGCTG |
| pRS315 (rev comp) fwd | CAGCTTTTGTTCCCTTTAGTGAGGG |
| pRS315 (rev comp) rev | TGAGTCGTATTACAATTCACTGGCC |
| npl3-F160Y-F162Y-F165Y fwd | AAGATCTTGAACGGCTACGCGTATGTTGAATATGAAGAAGCAGAATCC |
| npl3-F160Y-F162Y-F165Y rev | GGATTCTGCTTCTTCATATTCAACATACGCGTAGCCGTTCAAGATCTT |
| npl3-F160Y fwd | AAGATCTTGAACGGCTACGCGTTTGTTGAATTTG |
| npl3-F160Y rev | CAAATTCAACAAACGCGTAGCCGTTCAAGATCTT |
| npl3-F162Y fwd | TTGAACGGCTTCGCGTATGTTGAATTTGAAGAAG |
| npl3-F162Y rev | CTTCTTCAAATTCAACATACGCGAAGCCGTTCAA |
| npl3-F165Y fwd | TTCGCGTTTGTTGAATATGAAGAAGCAGAATCC |
| npl3-F165Y rev | GGATTCTGCTTCTTCATATTCAACAAACGCGAA |
| npl3-F160Y-F162Y fwd | AAGATCTTGAACGGCTACGCGTATGTTGAATTTGAAGAAG |
| npl3-F160Y-F162Y rev | CTTCTTCAAATTCAACATACGCGTAGCCGTTCAAGATCTT |
| npl3-F160Y-F165Y fwd | AAGATCTTGAACGGCTACGCGTTTGTTGAATATGAAGAAGCAGAATCC |
| npl3-F160Y-F165Y rev | GGATTCTGCTTCTTCATATTCAACAAACGCGTAGCCGTTCAAGATCTT |
| npl3-F162Y-F165Y fwd | CTTGAACGGCTTCGCGTATGTTGAATATGAAGAAGCAGAATCC |
| npl3-F162Y-F165Y rev | GGATTCTGCTTCTTCATATTCAACATACGCGAAGCCGTTCAAG |
| npl3-P196D-A197D fwd | TACTCTAAATTGGATGACAAGAGATACCGT |
| npl3-P196D-A197D rev | ACGGTATCTCTTGTCATCCAATTTAGAGTA |
| npl3-P196D fwd | TTACTCTAAATTGGATGCCAAGAGATACC |
| npl3-P196D rev | GGTATCTCTTGGCATCCAATTTAGAGTAA |
| npl3-A197D fwd | TCTAAATTGCCTGACAAGAGATACCGT |
| npl3-A197D rev | ACGGTATCTCTTGTCAGGCAATTTAGA |
| npl3-P196D-A197K fwd | TACTCTAAATTGGATAAGAAGAGATACCGT |
| npl3-P196D-A197K rev | ACGGTATCTCTTCTTATCCAATTTAGAGTA |
| npl3-P196L-A197D fwd | TACTCTAAATTGCTTGACAAGAGATACCGT |

| | |
|---|---|
| npl3-P196L-A197D rev | ACGGTATCTCTTGTCAAGCAATTTAGAGTA |
| npl3-P196L-A197K fwd | TACTCTAAATTGCTTAAGAAGAGATACCGT |
| npl3-P196L-A197K rev | ACGGTATCTCTTCTTAAGCAATTTAGAGTA |
| npl3-F245I fwd | GTACCGGTGCTCTAGAAATTCCTAGTGAAGAAATCTTGG |
| npl3-F245I rev | CAAGATTTCTTCACTAGGAATTTCTAGAGCACCGGTACC |
| npl3-F245Y fwd | GTACCGGTGCTCTAGAATACCCTAGTGAAGAAATCTTGG |
| npl3-F245Y rev | CAAGATTTCTTCACTAGGGTATTCTAGAGCACCGGTACC |
| pNOP-TAP fwd | AACCAGGTAAGCCATTTATATAGTTGAG |
| pNOP-TAP rev | CTTCAGACATGGATCCAAGTGC |
| NPL3 (for pNOP-TAP) fwd | ACTTGGATCCATGTCTGAAGCTCAAGAAACTCAC |
| NPL3 (for pNOP-TAP) rev | ATATAAATGGCTTACCTGGTTGGTGATC |
| pRS425 fwd | TTAACATCCTGGGATCCACTAGTTCTAGAGCGG |
| pRS425 rev | GATAGAGACGTCGACCTCGAGGGG |
| NPL3 (for pRS425) fwd | TCGAGGTCGACGTCTCTATCAATATGCAAATGCTCGG |
| NPL3 (for pRS425) rev | TAGTGGATCCCAGGATGTTAAATGTTATCATGGGATTATTCTATCTACTG |
| pT7-His6-TEV RRMs fwd | AATCAGATAAGGATCCGGCTGCTAACA |
| pT7-His6-TEV RRMs rev | TCGAGAGTTCGCCCTGAAAATACAGGTTTTCGG |
| npl3-RRMs (for pT7-His6…) fwd | TTTTCAGGGCGAACTCTCGAACACCAGATTG |
| npl3-RRMs (for pT7-His6…) rev | AGCCGGATCCTTATCTGATTGGTGGAGGATTGTCATCTC |
| pT7-His6-TEV FL fwd | CCAACCAGGTAAGGATCCGGCTGCTAACAAAG |
| pT7-His6-TEV FL rev | AGCTTCAGACATGCCCTGAAAATACAGGTTTTCGG |
| *NPL3* (for pT7-His6…) fwd | TATTTTCAGGGCATGTCTGAAGCTCAAGAAACTCAC |
| *NPL3* (for pT7-His6…) rev | GCAGCCGGATCCTTACCTGGTTGGTGATCTTTCAC |
| npl3-C211S fwd | CTTACCAGAAGGTTCTTCATGGCAAGATCTTAAAGATTTAGCC |
| npl3-C211S rev | CTTGCCATGAAGAACCTTCTGGTAAGTTTTTCATGG |
| npl3-D135C fwd | CCTTTCCCATTGTGCGTTCAAGAATCCGAGTTGAATG |
| npl3-D135C rev | TCGGATTCTTGAACGCACAATGGGAAAGGTCTAACAAACAATCTGG |
| npl3-E176C fwd | GCCAAAGCCATTTGTGAAGTTCACGGTAAGAGTTTTGC |
| npl3-E176C rev | ACCGTGAACTTCACAAATGGCTTTGGCAGCGGATTC |
| npl3-N185C fwd | TAAGAGTTTTGCTTGCCAACCTTTGGAAGTTGTTTACTC |
| npl3-N185C rev | CTTCCAAAGGTTGGCAAGCAAAACTCTTACCGTGAACTTC |
| npl3-E226C fwd | GAAAATAGTTTATGTACTACTTTTTCTAGCGTCAATACC |
| npl3-E226C rev | TAGAAAAAGTAGTACATAAACTATTTTCCCTGGC |
| npl3-D236C fwd | GTCAATACCAGATGTTTTGATGGTACCGGTGCTCTAG |
| npl3-D236C rev | GGTACCATCAAAACATCTGGTATTGACGCTAGAAAAAG |

**Table 8: Oligonucleotides used for genomic integration of affinity tags**

| Name | Sequence |
|---|---|
| HPR1-FTpA fwd | GCAGCTACTTCGAACATTTCTAATGGTTCATCTACCCAAGATATGA AAGATCACGACGGTGACTAC |
| HPR1-FTpA rev | TGCATGAATTTCTTATCAGTTTAAAATTTCTATTAAGAGGATAATTT AATCGATGAATTCGAGCTCG |
| THO1-FTpA fwd | TCCAGAGTAAGTAAAAACAGGAGAGGCAACCGCTCTGGTTACAGA AGAGATCACGACGGTGACTAC |
| THO1-FTpA rev | GAACCGAAACTAGAATGAAAAACTCCACCAAAACGGCTTGAGCCT TTAATCGATGAATTCGAGCTCG |
| NAB2-FTpA fwd | CCTCCGCAAACCAGTTTTACGCACCAAGAACAAGATACGGAAATG AACGATCACGACGGTGACTAC |
| NAB2-FTpA rev | AAGGGTCACAGGAACATGAATTTCGTTCCGTGATTTTAATAGTAAT CAATCGATGAATTCGAGCTCG |
| MTR2-FTpA fwd | TTTAACTATAACATGGTTTACAAACCCGAGGATTCTCTGCTAAAAA TTGATCACGACGGTGACTAC |
| MTR2-FTpA rev | TTTCTATATAATTTTGTTTTTCCCTGGTAGAGCGGAATCTTCCCACT AATCGATGAATTCGAGCTCG |

| | |
|---|---|
| CBC2-FTpA fwd | AGACCAGGTTTCGATGAAGAAAGAGAAGATGATAACTACGTACCT CAGGATCACGACGGTGACTAC |
| CBC2-FTpA rev | ATATATCTGTGTGTAGAATCTTTCTCAGATATAAATTGATTGATTCT AATCGATGAATTCGAGCTCG |
| THP1-FTpA fwd | AATGAACGAATCACCAAGATGTTTCCTGCCCATTCTCACGTTCTTT GGGATCACGACGGTGACTAC |
| THP1-FTpA rev | TGATATGTAGATATATATAGGAAATAGAAGAGAAGGAGCGACATTA TGATCGATGAA TTCGAGCTCG |
| CBC2-TAP fw | AGACCAGGTTTCGATGAAGAAAGAGAAGATGATAACTACGTACCT CAGTCCATGGAAAAGAGAAG |
| CBC2-TAP rev | TATATATATCTGTGTGTAGAATCTTTCTCAGATATAAATTGATTGAT TCTATACGACTCACTATAGGG |
| MEX67-HA fwd | TAAACTGTATATTTTTTGTGATACTGTGCGGCTGAAACAGGGAAC AATATCATTAATCGATGAATTCGAGCTCG |
| MEX67-HA rev | AAAGGGTTTTCAGAGTAGCATGAATGGCATCCCTAGAGAAGCATT TGTGCAGTTCCGTACGCTGCAGGTCGAC |
| HPR1-HA fwd | GCTACTTCGAACATTTCTAATGGTTCATCTACCCAAGATATGAAA CGTACGCTGCAGGTCGAC |
| HPR1-HA rev | ATGAATTTCTTATCAGTTTAAAATTTCTATTAAGAGGATAATTTAAT CGATGAATTCGAGCTCG |
| SUB2-FTpA fwd | GCTGAATTCCCAGAAGAAGGCATTGATCCGTCCACTTATTTGAAT AATGATCACGACGGTGACTAC |
| SUB2-FTpA rev | AAAATCTTTATATAATCTATATAAAAACGTATCTTTTTTCCTTTATT AATCGATGAATTCGAGCTCG |
| MEX67-FTpA fwd | TTTCAGAGTAGCATGAATGGCATCCCTAGAGAAGCATTTGTGCA GTTCGATCACGACGGTGACTAC |
| MEX67-FTpA rev | TATATTTTTTGTGATACTGTGCGGCTGAAACAGGGAACAATATCA TTAATCGATGAATTCGAGCTCG |
| YRA1-HA fwd | AAAATTAAATTTAATAAAACCAAATTAAATCAAACAAAAAATTGAC AATTAATTA ATCGATGAATTCGAGCTCG |
| YRA1-HA rev | TAAGAAAAGTCTTGAAGATCTGGACAAGGAAATGGCGGACTATT TCGAAAAGAAA CGTACGCTGCAGGTCGAC |

**Table 9: Oligonucleotides used for qPCR**

| Name | Sequence |
|---|---|
| NTR fwd (Mayer *et al.*, 2010) | TGCGTACAAAAAGTGTCAAGAGATT |
| NTR rev (Mayer *et al.*, 2010) | ATGCGCAAGAAGGTGCCTAT |
| PMA1 5' fwd | GTTTTTCGTCGGTCCAATTCA |
| PMA1 5' rev | AACCGGCAGCCAAAATAGC |
| PMA1 M fwd | AAATCTTGGGTGTTATGCCATGT |
| PMA1 M rev | CCAAGTGTCTAGCTTCGCTAACAG |
| PMA1 3' fwd | CAGAGCTGCTGGTCCATTCTG |
| PMA1 3' rev | GAAGACGGCACCAGCCAAT |
| CCW12 5' fwd | ACTGTCGCTTCTATCGCCGC |
| CCW12 5' rev | TTGGCTGACAGTAGCAGTGG |
| CCW12 M fwd | CTGTCTCCCCAGCTTTGGTT |
| CCW12 M rev | GGCACCAGGTGGTGTATTGA |
| CCW12 3' fwd | TGAAGCTCCAAAGAACACCACC |
| CCW12 3' rev | AGCAGCAGCACCAGTGTAAG |
| YEF3 5' fwd | GCTGACCCAACTGAAGTTCC |
| YEF3 5' rev | GGGACCATGATGGACAAAGT |
| YEF3 M fwd | GGTATCTGTGGTCCAAACG |
| YEF3 M rev | GGGTTGGGAAACCATCAA |
| YEF3 3' fwd | TCTGGTCACAACTGGGTTAGTG |
| YEF3 3' rev | GCAATCTTGTTACCCATAGCATCGA |

## 4.7 Enzymes

**Table 10: Enzymes**

| Enzyme | Supplier |
|---|---|
| Phusion® High-Fidelity DNA Polymerase | NEB |
| RNase A | Thermo Fisher Scientific |
| DNase I | Thermo Fisher Scientific |
| Proteinase K | Sigma Aldrich |
| Tobacco etch virus (TEV)-protease | self-made |
| Zymolyase 100T | Carl Roth |
| Zymolyase 20T | Carl Roth |
| Taq DNA Polymerase | self-made |
| T5 Exconuclease | NEB |
| Taq DNA Ligase | NEB |
| Restriction Enzymes | NEB |

## 4.8 Antibodies

**Table 11: Antibodies**

| Name | Source | Dilution | Supplier |
|---|---|---|---|
| anti-Mex67 | rabbit | 1:5,000 | (Strässer *et al*., 2000) |
| anti-HA-HRP | rabbit | 1:1,000 | R&D Systems |
| anti-Nab2 | mouse | 1:5,000 | Swanson lab (3F2) |
| anti-Npl3 | rabbit | 1:5,000 | Tracy Kress lab |
| anti-Sto1 | rabbit | 1:20,000 | Dirk Görlich lab |
| anti-Sub2 | rabbit | 1:10,000 | (Strässer *et al*., 2002) |
| anti-Tho1 | rabbit | 1:5,000 | Pineda lab |
| anti-Pgk1 | mouse, monoclonal | 1:5,000 | Abcam |
| Peroxidase anti-Peroxidase (PAP) | rabbit, monoclonal | 1:5,000 | Sigma-Aldrich |
| RNAPII (8WG16) | mouse | 1:250 | Covance |
| ChromPure rabbit IgG for ChIP | Rabbit | | Jackson IR Laboratories |
| anti-rabbit-HRP | goat, monoclonal | 1:3,000 | Biorad; #170-6515 |
| anti-mouse-HRP | goat, monoclonal | 1:3,000 | Biorad; #170-6516 |
| anti-rabbit-Alexa488 | Goat | 1:200 | Invitrogen |

# 5 Methods

## 5.1 Standard methods

If not further explained in the sections below, standard molecular cloning techniques like growth of bacteria, DNA isolation and DNA analysis on agarose gels in 1x TAE were performed according to (Sambrook and Russell, 2001) or as described in manufactures' manuals. For small-scale plasmid preparation from *E. coli,* the NucleoSpin® Plasmid (NoLid)-kit or for mid-scale the NucleoSnap Plasmid Midi-kit (Macherey-Nagel) was used. Purifying and gel extraction of DNA samples were carried out by NucleoSpin® Gel and PCR Clean-up-kit (Macherey-Nagel). All plasmids cloned for this study have been confirmed by sequencing (Microsynth). For visualizing DNA or RNA on agarose gels, Intas HDGreen™ was used.

## 5.2 Cloning

### 5.2.1 Polymerase chain reaction (PCR)

Depending on the required amplicon, Phusion High-Fidelity DNA polymerase (cloning and genomic integration) or self-made Taq DNA polymerase (colony PCR) were used to perform the PCR reaction. A typical PCR reaction for Phusion High-Fidelity DNA polymerase is shown in Table 12. For integration in yeast cells, a PCR reaction with a volume of 300 µL was performed.

**Table 12: Standard PCR reaction**

| Component | Amount [µL] | Final concentration | Stock concentration |
|---|---|---|---|
| dNTPs | 4 | 200 µM | 2.5 mM |
| Buffer | 10 | 1x | 5x |
| Primer fwd. | 0.25 | 500 nM | 100 µM |
| Primer rev. | 0.25 | 500 nM | 100 µM |
| Template (plasmid) | 0.5 | ≤ 10 ng | Varies |
| Water | 34.5 | | |
| Phusion DNA Polymerase | 0.5 | 1 U | 2 U/µL |
| | 50 µL | | |

| Temperature | Time | |
|---|---|---|
| 98°C | 30 sec | |
| 98°C | 10 sec | x32 |
| 54°C | 30 sec | |
| 72°C | 30 sec/kb | |
| 72°C | 5-10 min | |
| 4°C | ∞ | |

### 5.2.2 Gibson Assembly

Gibson assembly was used for any kind of cloning or introducing point mutations (adapted to (Gibson *et al.*, 2009)). Linear vector backbone as well as the inserts were generated by PCR. The primers were designed to yield PCR fragments with an overlapping region of about 20 - 25 bp to each other. PCR fragments were DpnI digested before Gibson assembly. For a typical assembly reaction, 50 ng of vector and a 1:1 ratio of insert to vector was mixed with 15 µL of Gibson assembly master mix and filled up to 20 µL. The mixture was incubated for 30 min at 50°C. For transformation in *E. coli,* 2 µL of the mixture was used.

### 5.2.3 Transformation of *E. coli*

Competent cells were made according to the manufacturer's manual of the Mix & Go *E. coli* Transformation Kit (Zymo Research Corp.) and stored at -80°C until needed. For each transformation 50 µL competent cells were thawed on ice. 2 µL of a Gibson assembly mix or 0.5 µL of a plasmid-miniprep were added to the cells and mixed by flicking several times. After incubation on ice for 10 min, the cells were heat shocked at 42°C for 45 sec. Immediately after the heat shock, the cells were placed on ice for 1 min. For recovery 400 µL SOC medium was added and incubated for 60 min at 37°C on a shaker at 200 rpm. Afterwards the cells were pelleted at 10,000 g for 20 sec and pellet was resuspended in 150 µL water. The cell suspension was spread out on a selective LB-plate and incubated overnight at 37°C.

### 5.2.4 Colony PCR for *E. coli*

To screen for successful plasmid cloning, single freshly grown colonies were picked and suspended in 20 µL water. 5 µL cell suspension was used for a PCR as shown in table 14 and analyzed on a 1 % agarose gel in TAE. For two positive clones a 3 mL LB culture with appropriate antibiotic was inoculated with the residual 15 µL cell suspension.

**Table 13: *E. coli* colony PCR**

| Component | Amount [µL] | Final concentration | Stock conc. |
|---|---|---|---|
| dNTPs | 1.6 | 200 µM | 2.5 mM |
| Buffer | 2 | 1x | 10x |
| Primer fwd. | 0.1 | 500 nM | 100 µM |
| Primer rev. | 0.1 | 500 nM | 100 µM |
| Template (cell suspension) | 5 | | |
| Water | 10.8 | | |
| Taq DNA Polymerase | 0.4 | | |
| | 20 µL | | |

| Temperature | Time | |
|---|---|---|
| 95°C | 5 min | |
| 95°C | 30 sec | x32 |
| 48°C | 45 sec | |
| 68°C | 1 min/kb | |
| 68°C | 5-10 min | |
| 4°C | ∞ | |

## 5.3 Transformation in *S. cerevisiae*

For transformation of linearized DNA or plasmid in *S. cerevisiae,* an overnight pre-culture was diluted in 50 mL appropriate medium to $OD_{600}$ = 0.2. When reaching mid-log phase ($OD_{600}$ = 0.6 - 0.8) after shaking at 200 rpm at 30°C, the culture was harvested at 2,800 g for 3 min and washed once with 10 mL water. The cells were washed once in 500 µL solution I, transferred to a 2 mL tube and resuspended in 250 µL of solution I. For each transformation 50 µL of cell suspension were mixed with 5 µl single-stranded carrier DNA (2 mg/mL), 300 µL solution II and 500 ng plasmid DNA or with a EtOH-precipitated 300 µL-PCR reaction in case of genomic integration. As a negative control a mixture without DNA was performed. After 30 min incubation on a rotating wheel at RT, the transformation mixture was heat shocked for 10 min at 42°C and immediately afterwards incubated on ice for 3 min. To remove the PEG solution, 1 mL water was added, and the cells were pelleted at 1,200 g for 3 min at RT in a tabletop centrifuge. For genomic integration 1 mL YPD was added to the cells. Cells were recovered for 2 hours at 30°C while shaking before being pelleted and resupended in 150 µL water to spread out on selective media. When transforming plasmids the cells were directly spread out on selective media plates. Plates were incubated for 2 - 4 days at 30°C.

### 5.3.1 Yeast colony PCR

To screen for positive genomic integration of protein tags or genomic deletions, a small amount of freshly grown yeast colony was picked and suspended in 15 µL zymolyase 20T solution (2.5 mg/mL) to digest the cell wall. The mixture was incubated for 20 min at RT, moved to 37°C for 5 min and afterwards boiled for 5 min at 95°C. The cell suspension was diluted by adding 60 µL water. Afterwards a PCR reaction with 5 µL of cell suspension as template was performed as shown in Table 13 and analyzed on a 1 % agarose gel in TAE.

### 5.3.2 Generating yeast strains with *npl3* point mutations

Strains that carry a point mutation in *NPL3* were generated by transforming a pRS315 plasmid with mutated ORF + 500 bp of 5' and 300 bp of 3' UTR of *NPL3* into the *npl3-*

*shuffle* strain and shuffling out the pRS316-*NPL3* by streaking cells two times over 5'-FOA plates.

## 5.4 Ethanol (EtOH) precipitation

To precipitate proteins and DNA, 2.5 - 3 volumes of 100 % EtOH and 1/10 volume 3 M NaOAc (pH 5.2) were added to the sample and mixed by inverting the tube. After an incubation of 20 min at -20°C, a centrifugation step at 4°C at 12,000 g for 20 min followed. After washing the pellet once with 70 % EtOH, the pellet was dried and dissolved in 15 µL pure water or the required buffer/volume.

## 5.5 SDS-PAGE

Sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) was done according to (Laemmli, 1970). The gels were cast according to (Sambrook and Russell, 2001), using a Mini-Protean II system (Biorad). After electrophoretic separation the proteins were either transferred to a membrane for subsequent Western blot or stained with a modified Fairbanks-Coomassie staining method (Fairbanks *et al.*, 1971; Wong *et al.*, 2000). For staining, the gel was covered with Coomassie solution, heated up in a microwave and incubated for 15 min on a rocker. To clear the background, the Coomassie solution was discarded, destain solution was added, heated up and incubated on a rocker. This was repeated two to three times until the background was completely destained.

## 5.6 Dot spots

Freshly grown yeast cells were picked with a small loop and suspended in 1 mL of water. To ensure that all spotted strains have the same number of cells the $OD_{600}$ was measured and diluted to 0.15. A 10-fold serial dilution was made four times for each strain. 5 µL of each strain and dilution were spotted on respective media plate, air dried and incubated at 16°C, 25°C, 30°C and 37°C for up to 7 days.

## 5.7 Growth curve in liquid media

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. To keep cells in mid-log growth phase, cells were always diluted before reaching an $OD_{600}$ of 1. For incubation at 37°C, a shaking water bath was used.

## 5.8 Tandem Affinity Purification (TAP)

Tandem affinity purification (TAP) was performed according to (Puig *et al.*, 2001; Rigaut *et al.*, 1999). The TAP tags consisted of two minimal protein A domains followed by a TEV cleavage side and either a calmodulin binding peptide (CBP) or a FLAG-tag. TAP was performed to purify the fused protein and interacting proteins from yeast. A cell pellet from a 2 L yeast culture grown to $OD_{600}$ = 3.0 - 3.5 was used. The cell pellet was resuspended in 2.5 mL TAP-buffer after centrifugation at 7,000 g for 4 min at RT. The resulting cell suspension was dropped into a vessel containing liquid nitrogen to get flash frozen cell beads. These deep-frozen cell beads were lysed using a freezer mill 6870D (SPEX SamplePrep). The lysate was stored at -80°C until use.

Before thawing the lysate 10 mL TAP-buffer + 1x protease inhibitor and 1 mM DTT were added. After thawing, the lysate was centrifuged at 3,500 g for 12 min at 4°C and the precleared lysate at 165,000 g for 1 hour at 4°C. During ultracentrifugation 500 µL IgG-sepharose 6 fast flow affinity resin was equilibrated with TAP-buffer three times by addition of 10 mL buffer, centrifugation at 700 g for 3 min and removal of the supernatant. The fatty top phase that resulted after ultracentrifugation was removed by aspiration. The clear supernatant was transferred to a 50 mL tube, IgG sepharose added, and the mix incubated for 1.5 hours on a rotating wheel at 4°C. In case that an RNase A treatment was performed, 100 µg/mL RNase A was added to the lysate and incubated at RT for 20 min on a rotating wheel before the IgG sepharose was added. The supernatant was removed by vacuum pump after IgG beads were centrifuged at 700 g for 3 min at 4°C. The IgG sepharose was washed once with 10 mL TAP-buffer before transferring to a Mobicol (Mobitec). A 10 mL syringe was plugged on top of the Mobicol to wash the IgG sepharose with an additional 10 mL TAP-buffer by gravity flow. The bound protein complexes were released by adding 140 µL TAP-buffer + 7.5 µL TEV-protease and incubating at 16°C for 1 hour on a rotating wheel. The TEV-eluate was spun into a fresh 2 mL tube. To further increase the purity, one of the following affinity purifications were performed in a fresh MobiCol.

CBP purification

Per sample, 600 µL calmodulin slurry was equilibrated in TAP-buffer with a final concentration of 2 mM $CaCl_2$. The equilibrated calmodulin slurry and additional TAP-buffer (equal volume as TEV-eluate) supplemented with 4 mM $CaCl_2$ were added to the TEV-eluate and incubated for 1 hour at 4°C on a rotating wheel. After washing with 10 mL TAP-buffer (+2 mM $CaCl_2$), the proteins were eluted by adding elution buffer and incubated at 37°C for 15 min. The calmodulin-eluate was precipitated with 10 % TCA for

10 min on ice, followed by centrifugation at 13,500 g for 10 min at 4°C. The protein pellet was resuspended in 1x SDS-sample loading buffer and pH neutralized with approximately 2 µL 1 M Tris base (until SDS-buffer turned from yellow to blue again).

FLAG-tag purification

The TEV-eluate was incubated for 1 hour at 4°C on a rotating wheel with 50 µL anti-FLAG M2 affinity gel after equilibrating in TAP-buffer. After washing with 10 mL TAP-buffer the proteins were eluted by adding 100 µg/mL FLAG peptide and incubated on a rotating wheel for 15 min at RT.

## 5.9 Identification of protein-RNA crosslinking sites

To prepare mRNPs for identification of RNA-binding sites, a PCR-amplified TAP- or FTpA-tag was genomically integrated. The PCR construct was transformed into yeast strain RS453 and successful integration verified by Western blot. For 4-tU labeling and *in vivo* cross-linking, yeast cells were grown in SDC-URA supplemented with 120 µM Uracil to $OD_{600}$ = 0.8 before 4-thiouracil was added to a final concentration of 500 µM. Cells were pelleted at 7,000 g for 4 min when reaching $OD_{600} \approx 3$. Cells were cross-linked by exposure to 365 nm UV light in a petri dish (dissolved in 20 mL 1 x PBS) on a water bath with crushed ice using a Bio-Link BLX-365-UV-Crosslinker. Cells were pelleted again, suspended in 2.5 mL TAP-buffer (50 mM TRIS, pH 7.5, 1.5 mM $MgCl_2$, 200 mM KCl, 0.15 % NP 40, 1 mM DTT) and flash frozen drop-by-drop in liquid nitrogen. Cell beads were ground using a freezer mill 6870D (SPEX SamplePrep). The lysate was cleared at 165,000 g for 1 hour at 4°C. The supernatant was incubated with IgG bead-slurry (IgG Sepharose 6 Fast Flow, GE Healthcare) for 1.5 h at 4°C on a turning wheel. After washing TEV-protease was added to the beads and incubated for 1 h at 16°C for elution. 100 µg of affinity-purified eluate were further processed to enrich for protein-RNA heteroconjugates as described in (Kramer *et al.*, 2014; Sonnleitner *et al.*, 2017). The enriched samples were subjected to the same self-packed C18 LC-system described above with a modified multi-step gradient at a flow rate of 0.3 - 0.4 µL/min. The specific mass spectrometric settings for protein-RNA crosslinking identification are described in detail in (Kramer *et al.*, 2014; Sonnleitner *et al.*, 2017). Spectra corresponding to peptide-RNA crosslinked species were analyzed and manually validated using the OpenMS module $RNP^{XL}$ (Kramer *et al.*, 2014).

## 5.10 Identification and quantification of affinity-purified proteins using MS

For identification and quantification of affinity-purified proteins using MS, the TAP purifications of Cbc2 and Npl3 were essentially carried out as described above. For

RNase A treatment, whole-cell protein extracts were incubated with RNase A at a final concentration of 100 µg/mL for 20 min at room temperature before IgG Sepharose was added. The TEV-eluates were analyzed by MS and quantitative Western blot. For MS analysis samples were processed according to a modified protocol (Shevchenko *et al.*, 1996) to ensure equal sample quality and comparability. Briefly, proteins contained in the eluates were subjected to SDS-PAGE analysis on a 4 - 12 % gradient gel (NuPage, Thermo Fisher) according to the manufacturer's specifications. Next, individual lanes of eluates were cut into 23 pieces and proteins were alkylated with 55 mM iodacetamide before digestion with trypsin (12 ng/µL) in-gel overnight at 37°C. Extracted peptides were dried by vacuum centrifugation and resuspended in 2 % acetonitrile, 0.05 % TFA before being subjected to LC-MS analysis using a self-packed C18 column (Dr. Maisch GmbH, 1.9 µM pore size). All samples were analyzed on a 58-minute linear gradient, ranging from 2 % - 80 % acetonitrile in the mobile phase. Acquired spectra from a QExactive HF mass spectrometer (Thermo Scientific) were searched for corresponding peptides that were also quantified using the LFQ algorithm built into MaxQuant v. 1.6.5 (Cox and Mann, 2008).

## 5.11 Quantitative Western blot

To quantify the total amount of a protein, cells were grown in liquid media. When reaching mid-log phase ($OD_{600}$ = 0.6), 10 mL culture were pelleted at 2,800 g for 3 min at 4°C. The supernatant was discarded and the pellet resuspended in 150 µL 1x SDS-loading buffer. The cell suspension was transferred to a 2 mL tube with screw cap and lysed by FastPrep-24 5G device. For quantification, equal amounts of total protein were separated by SDS-PAGE, semi-dry blotted (Towbin *et al.*, 1979) and proteins detected with the corresponding primary antibody, a horse radish peroxidase-coupled secondary antibody and CheLuminate-HRP ECLsolution (Applichem). Chemiluminescence signals were imaged using a ChemoCam Imager (Intas) and quantified with FIJI.

## 5.12 Determination of mRNA stability

To determine the half-life of selected transcripts, yeast cells were grown at 30°C until they reached an $OD_{600}$ of 0.5. Immediately after harvesting 10 mL culture, a final concentration of 8 µg/mL thiolutin was added to the cell culture. After 15, 30 and 90 min, additional 10 mL of the yeast culture were harvested. The cells were flash frozen in liquid nitrogen after harvesting. Afterwards the RNA was extracted according to 5.15. The total RNA was reverse transcribed using super script III according to the manufacturer's protocol. The resulting cDNA was analyzed by qPCR.

## 5.13 Fluorescence *in situ* hybridization (FISH) with oligo d(T)

*In situ* hybridization against poly(A)+ RNA was done according to (Amberg *et al.*, 1992). Briefly, cells were grown at 30°C in YPD medium to mid-log phase before being shifted to 37°C for 1 hour. An aliquot of 10 mL cell culture was removed before temperature shift. Cells were immediately fixed with 4 % formaldehyde, washed after fixation, and spheroplasted with 100T zymolyase. After adhering spheroplasts to a poly-lysine-coated slide, the cells were prehybridized for 1 hour at 37°C in prehybridization buffer (50 % formamide, 10 % dextran sulphate, 125 µg/mL of *E. coli* tRNA, 500 µg/mL H.S. DNA, 4x SSC, 0.02 % polyvinyl pyrrolidone, 0.02 % BSA, 0.02 % Ficoll-40) in a humid chamber. To hybridize with oligo $d(T)_{50}$-Cy3 0.75 µl of 1 pmol/µl probe was added and incubated at 37°C O/N in humid chamber. After hybridization, cells were washed, mounted with ROTI®Mount FluorCare DAPI, and covered with a coverslip. Cells were inspected with an Axio observer fluorescence microscope (Zeiss) connected to a CCD camera.

## 5.14 Chromatin Immunoprecipitation (ChIP)

To investigate the relative occupancy of proteins on target genes, chromatin immunoprecipitation (ChIP) was performed. An adapted ChIP protocol according to (Aparicio *et al.*, 2005) was used. An overnight pre-culture was diluted in 100 mL of an appropriate medium and crosslinked with 1 % formaldehyde for 20 min on a rocker at RT when reaching an $OD_{600}$ = 0.8. By adding 12.5 mL glycine (3 M) and incubating for 10 min, the cross-linking reaction was stopped. The cells were harvested at 2,800 g for 3 min and washed three times in 1x PBS. The washed cell pellet was flash frozen in liquid nitrogen and stored at -80°C until use. The pellet was thawed on ice, resuspended in 800 µL low-salt buffer and transferred to a screw cap tube with 300 µL glass beads. To lyse the cells the FastPrep 24G device was used two times for 45 sec (6.5 m/s setting) with a 2 min break on ice in between. The lysate was sonicated in a Bioruptor three times, each 15 min (alternating 30 seconds at high energy setting on and 30 seconds off) with 5 min breaks on ice to shear chromatin in ~250 bp fragments. The lysate was cleared by two consecutive spins, 5 min at 18,000 g and 10 min at 18,000 g, each time transferring supernatant into a fresh tube. The DNA concentration of the lysates was measured at the NanoDrop and each sample was adjusted to the lowest measured sample. After adjusting the samples, 10 µL of each lysate were kept at 4°C as an input sample. For TAP- or FTpA-tagged proteins, 15 µL magnetic Dynabeads coupled with IgG were added to the remaining sample and incubated for 2.5 hours on a rotating wheel at RT. For RNAPII-ChIPs, 4 µL of α-RPB1 antibody (CTD repeat YSPTSPS;

8WG16; 1 mg/mL) were added and incubated for 1.5 hours on a rotating wheel at RT before 15 µL of Dynabeads coupled with protein G were added and incubated for one hour more. Subsequently, the beads were collected on a magnetic particle collector (MPC) and washed with 800 µl (2x low-salt buffer, 2x high-salt buffer, 2x TLEND, 1x TE). Between washing steps the beads were incubated for 2 min on a rotating wheel. For elution, the beads were resuspended in 130 µL ChIP-elution buffer and incubated on a thermoshaker at 65°C and 1,000 rpm for 20 min. After short incubation on the MPC, the supernatant was transferred to a fresh low binding tube and 80 µL of 1x TE and 10 µL proteinase K (10 mg/mL) were added. The input sample was mixed with 80 µL ChIP-elution buffer, 80 µL 1x TE and 10 µL proteinase K. To reverse the crosslinks, all samples were incubated for 2 hours at 37°C followed by 12 - 14 hours at 65°C. To purify the DNA the PCR NucleoSpin® Gel and PCR Clean-up-kit (Macherey-Nagel) was used according to manufacturers' manual and the DNA eluted in 140 µL 1x TE.

## 5.15 RNA extraction

For total RNA extraction, a cell pellet was resuspended in 1 mL Trizol reagent and mixed by pipetting. After incubating for 5 min at RT, 200 µL chloroform were added and the sample incubated an additional 3 min at RT. To phase separate the proteins and nucleic acids, the sample was centrifuged at 18,000 g for 20 min at 4°C. The upper aqueous phase (RNA) was transferred to a fresh tube and precipitated by adding 0.5 mL isopropanol and 2 µL glycogen (5 mg/mL). The mixture was incubated for 10 min on ice and pelleted at 13,500 g for 10 min. The pellet was washed once with 75 % EtOH, dried and dissolved in DEPC-treated RNase-free water. If needed, contaminating DNA was digested by DNase I in presence of RNase inhibitor for 30 min at 37°C. For transcriptome sequencing the purity and integrity of the RNA was analyzed via Bioanalyzer on an RNA Nano Chip (Agilent Technologies) according to the manufacturer's manual.

## 5.16 RNA Immunoprecipitation (RIP)

FTpA- or TAP-tagged *S. cerevisiae* strains were grown in 400 mL YDP to an $OD_{600}$ of 0.8, harvested and stored at -80°C. Pellets were thawed on ice, resuspended in 1 mL RNA IP-buffer + protease inhibitor and lysed using the FastPrep-24 5G device (3 times for 20 s at 6 m/s). The lysate was cleared by centrifugation for 5 min at 1,500 g and 10 min at 16,000 g at 4°C. 900 µl of the cleared lysate (input) was incubated with 660 units DNase I for 30 min on ice. 40 µl IgG-coupled Dynabeads M-280 were added and incubated for 3 h at 4°C on a turning wheel. The beads were washed 8 times with RNA-IP-buffer. For RNA extraction 1 mL TRIzol reagent was added to the beads. For

protein analysis an acetone precipitation of the phenol- and interphase was performed. The RNA of the input and IP samples were reverse transcribed and subsequently analyzed by quantitative PCR on an Applied Biosystems StepOnePlus cycler using Applied Biosystems Power SYBRGreen PCR Master Mix. As a control the RIP was also performed with an untagged strain (nc). PCR efficiencies (E) were determined with standard curves. Protein enrichment over the untagged strain was calculated according to $\frac{E^{(ct\_IP-ct\_input)nc}}{E^{(ct\_IP-ct\_input)}}$. Mean values were calculated of at least three biological replicates.

## 5.17 Transcriptome wide sequencing (RNA seq)

Total RNA was extracted as mentioned above. After RNA quantification with the Qubit RNA BR Assay Kit, ERCC RNA Spike-In Mix (both Invitrogen) was added according to the manufacturer's instructions. Libraries were prepared using the Illumina® TruSeq® mRNA stranded sample preparation Kit. After poly-A selection (using poly-T oligo-attached magnetic beads), mRNA was purified and fragmented using divalent cations under elevated temperature. The RNA fragments underwent reverse transcription using random primers. This was followed by second strand cDNA synthesis with DNA Polymerase I and RNase H. After end repair and A-tailing, indexing adapters were ligated. The products were then purified and amplified to create the final cDNA libraries. After library validation and quantification (Agilent 4200 tape station), equimolar amounts of library were pooled. The pool was quantified by using the Peqlab KAPA Library Quantification Kit and the Applied Biosystems 7900HT Sequence Detection System and subsequently sequenced on a NovaSeq6000 sequencer with PE100 read length aiming at 25M reads/sample.

## 5.18 RNA-seq data analysis of differential gene expression and splicing

For all genomic analyses, *Saccharomyces cerevisiae* S288c genome and gene annotation assembly (version R64-1-1) were downloaded from the EnsemblFungi database (https://fungi.ensembl.org/index.html). After initial quality control using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), RNA-seq reads were mapped using the splice-aware alignment tool STAR (version 2.7.2a) with the following parameters: --outFilterMultimapNmax 1 --outFilterMismatchNoverLmax 0.04. For visualization, bam files were converted to bigwig files using deepTools, including count per million (CPM) normalization.

For differential gene expression analysis, reads were counted into exons using HTSeq-count from the HTSeq Python package. Comparison between *NPL3* wild type and each of the three *npl3* mutants was performed with the R/Bioconductor package DESeq2

including shrinkage of logarithmic fold changes to disregard genes with low read counts. $Log_2$-transformed fold changes of intron-containing and intron-less genes were compared using Wilcoxon rank-sum test.

Analysis of intron retention events was implemented using the Bioconductor package GenomicAlignments. To quantify intron retention, the average number of reads overlapping the exon-intron (EI) 5' and 3' boundaries was counted for each intron. Similarly, reads overlapping the exon-exon (EE) junction were counted to evaluate spliced isoforms. The percentage of intron retention was calculated as

$$\%IR = \frac{EI}{EI + EE} * 100\ \%$$

Only genes with EI + EE > 100 were considered (n = 215). Differential intron retention was analyzed between NPL3 and each of the three *npl3* mutants. For each gene, ΔIR values were calculated based as

$$\Delta IR = \%IR(mutant) - \%IR(NPL3)$$

Wilcoxon rank-sum tests were performed to determine statistically significant changes in ΔIR distributions for all intron-containing genes between NPL3 and *npl3* mutants. All analyses were performed in R (version 4.0.2).

## 5.19 Protein expression and purification

$His_6$-tagged Npl3 tandem RRM domains (aa 120-280) wild-type (named $Npl3^{120\text{-}280}$) and the three mutants *npl3-RRM1*, *-Linker* and *-RRM2* (named *npl3*$^{120\text{-}280}$*-RRM1*, *npl3*$^{120\text{-}280}$*-Linker* and *npl3*$^{120\text{-}280}$*-RRM2*) were expressed in *E.coli BL21* (DE3) in minimal M9 media supplemented with $^{15}N$ $NH_4Cl$ and/or $^{13}C$ glucose as sole nitrogen and carbon source, respectively. Protein expression was induced at an $OD_{600}$ of 0.8 with 0.5 mM IPTG and cells grown at 22°C overnight. Cells were harvested, stored at -20°C and lysed by sonication. Proteins were purified by standard affinity chromatography with Ni-NTA sepharose. The N-terminal His-tag was removed by cleavage with TEV protease. Further purification was done using ion exchange and size exclusion chromatography. Purified samples were exchanged with NMR buffer containing 20 mM sodium phosphate (pH 6.4), 50 mM NaCl, 1 mM DTT. 10 % $D_2O$ was added to lock the magnetic field.

## 5.20 NMR experiments and analysis

Backbone chemical shift assignments were obtained using 3D triple resonance experiments: HNCACB, HNcoCACB, HNCO, HNcaCO and HcccoNH. For RNA binding studies, $^{1}H$-$^{15}N$ heteronuclear single quantum coherence (HSQC) and $HN^{\varepsilon}$-selective

heteronuclear in-phase single quantum coherence (HISQC) spectra were recorded in a titration series using 50 µM of protein with increasing concentration of RNA (synthetic RNA purchased from Dharmacon, USA, or Biolegio BV, Germany). Chemical shift perturbations (CSPs, **Δδ**) were calculated as: $\mathbf{\Delta\delta = [(\Delta\delta^{1}H)^{2} + (\Delta\delta\,^{15}N)^{2}/25]^{1/2}}$.

Dissociation constants ($K_D$) were derived from NMR titrations by fitting to the following equation: $\mathbf{\Delta\delta_{obs} = \Delta\delta_{max}\,\{([P]_t + [L]_t + K_D) - [\,([P]_t + [L]_t + K_D)^2 - 4[P]_t\,[L]_t\,]^{1/2}\} / 2[P]_t}$, where, $\Delta\delta_{obs}$ is the observed chemical shift difference relative to the free state, $\Delta\delta_{max}$ is the maximum shift change in saturation, $[P]_t$ and $[L]_t$ are the total protein and ligand concentrations, respectively, and $K_D$ is the dissociation constant. $\{^{1}H\}$-$^{15}N$ heteronuclear NOE (hetNOE) was determined from the ratio of signal intensities with and without saturation in HSQC based experiments with 3 s interscan delay. HetNOE for the protein alone using 200 to 300 µM and the protein-RNA complex were recorded with 90 µM protein and a 3-fold molar access of RNA. A series of $^{1}H$,$^{15}N$ HSQC experiments were recorded over 4 hours to monitor amide hydrogen-deuterium exchange.

NMR measurements were carried out with NMR samples in a Shigemi tube (Shigemi Inc, Japan) at 25°C on Bruker spectrometers operation at 500, 600, 800, 900 or 950 MHz proton Larmor frequency equipped with cryogenic probes. NMR spectra were processed with a shifted sine-bell window function and zero-filling before Fourier transformation using Bruker Topspin 3.5pl6 and NMRPipe. Proton chemical shifts were referenced against sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS), respectively. All spectra were analyzed with the CCPNMR analysis v2.5 software package.

## 5.21 Isothermal titration calorimetry (ITC)

ITC experiments were performed with a MicroCal PEAQ-ITC device (Malvern, UK). All protein samples were dialyzed against NMR buffer (20 mM sodium phosphate, pH 6.4, 50 mM NaCl, 1 mM TCEP). The ITC cell was filled with a 15 µM concentration of RNA oligo and protein was added from a syringe. Titrations were performed with 39 points of 1 µl injections with a 150 s interval at 25°C. All measurements were conducted in duplicates and analyzed by the Malvern's MicroCal PEAQ-ITC analysis software v1.0.0.1259. Binding curves were fitted to one-site binding mode and thermodynamic parameters were extracted.

# 6 Results

Already during transcription, RNA-binding proteins bind the nascent (pre-)mRNA and regulate its co- and post-transcriptional fate. Many of these proteins are recruited through a combination of protein-protein and protein-RNA interactions. The aim of the present study was to determine how the RNA-binding activities of individual RNA-binding sites within various nuclear mRNP-binding proteins contribute to protein recruitment and function. To this end, we first identified the amino acids involved in RNA contacts for several nuclear mRNP-binding proteins to be able to mutate those binding sites.

## 6.1 Identification of *in vivo* RNA-binding sites in nuclear mRNP components on amino acid level

To identify amino acids that interact with RNA *in vivo,* the $RNP^{XL}$ mass spectrometry method (Kramer *et al.*, 2014) was applied. To do so, the following mRNP components were genomically TAP-tagged in *S. cerevisiae*: The SR-like protein Npl3; the poly(A)-binding protein Nab2; the small subunit of the cap-binding complex, Cbc2; the TREX components Hpr1 and Sub2; Tho1; the THSC (TREX-2) component Thp1; and the mRNA exporter component Mtr2. In each of those strains, the RNA was labeled with 4-thiouracil (4-tU), crosslinked to bound proteins by 365 nm UV light *in vivo,* and the tagged proteins natively purified via the protein A tag. The TEV-eluate was then used to identify the amino acids crosslinked to the RNA via liquid chromatography mass spectroscopy and the $RNP^{XL}$ pipeline (Figure 5A).

The crosslink analysis resulted in varying numbers of newly identified RNA-interacting residues for different mRNP components (Figure 5B, Supplementary table 1). In total, 101 peptides that crosslink to RNA were identified in 16 different mRNP components. In 70 out of the 101 peptides, the amino acid involved in the crosslink could be precisely determined. By far, the highest number of amino acids that crosslinked to RNA were identified for Sub2 and Yra1 (24 % and 22 %), followed by Nab2 and Npl3 (10 % and 9 %; Supplementary table 1). For some proteins like Hpr1, Mex67 or Tho2, only very few RNA-interacting residues were detected. Overall, most identified crosslink positions fall within known RNA-binding domains like RRM, RGG-, helicase, or zinc-finger domains. However, some identified crosslinks are in protein regions that were so far unknown to contribute to RNA binding (Figure 5B). We postulate that RBPs bind multiple positions at the RNA within a single mRNP and that the different RNA binding sides contribute to the functions of each protein in different ways. Mutating the newly identified putative RNA-interacting residues in RBPs enables now to analyze how the different RNA-binding

activities influence the functional consequences on mRNP assembly, mRNA export, mRNA expression and stability, and splicing.

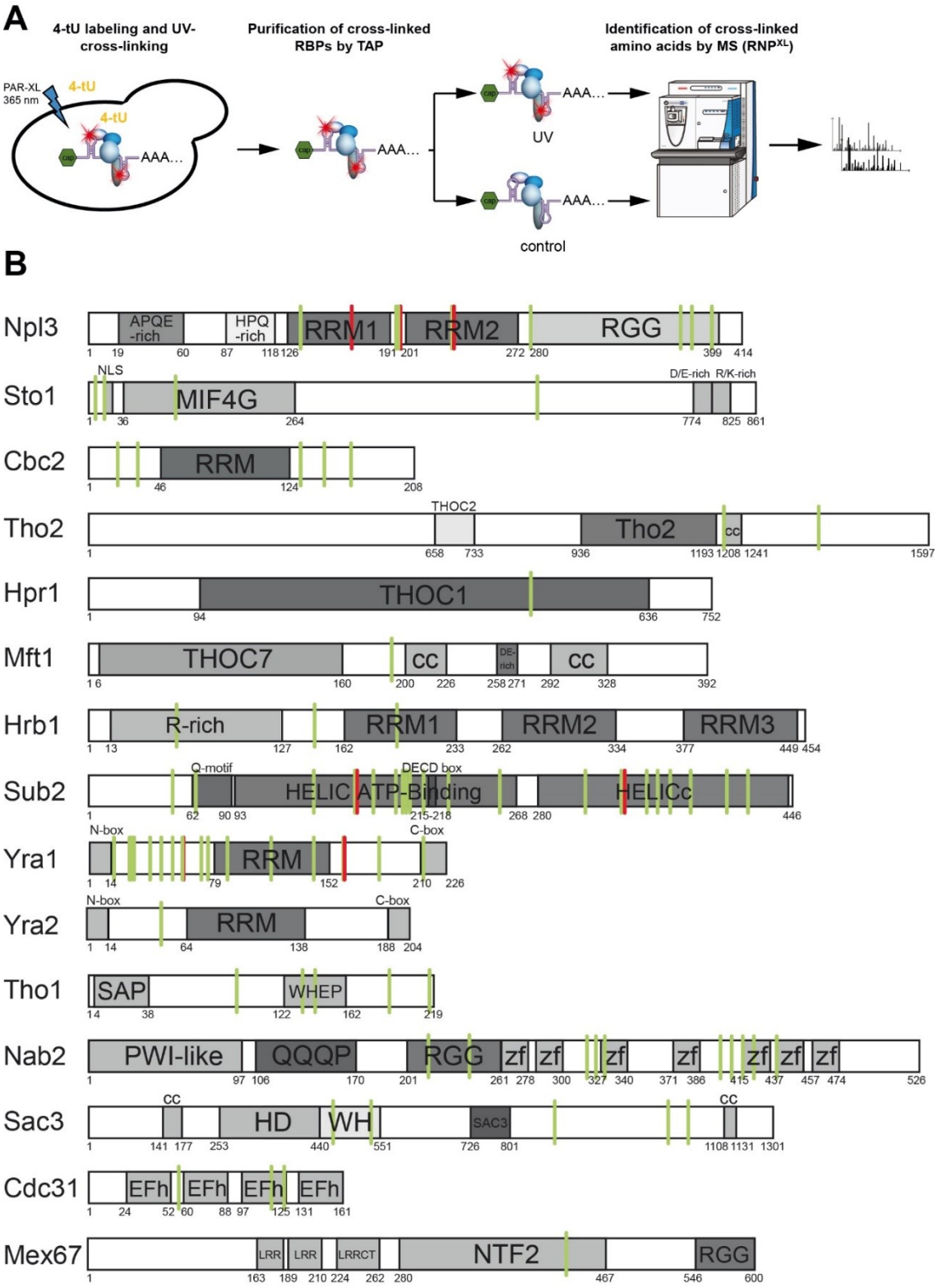



**Figure 5: Identification of RNA-protein crosslinks using the RNP^XL method. (A)** Schematic overview of the experimental workflow to identify amino acids that crosslink to RNA *in vivo*. The RNA was labelled with 4-thiouracil (4-tU) and crosslinked to proteins by 365 nm UV light (photoactivatable ribonucleoside enhanced crosslinking; PAR-XL) *in vivo*. TAP-tagged *S. cerevisiae* cells were lysed and proteins (protein-complexes)

were purified. The TEV-eluated samples (crosslinked and non-crosslinked control) were analyzed by mass spectroscopy and the RNP$^{XL}$ pipeline to identify the amino acids that crosslinked to RNA. **(B)** Position of identified peptides and amino acids that crosslink to RNA *in vivo* in nuclear mRNP components. Crosslinked peptides and amino acids that are newly identified are represented by green lines. Positions that were known by Kramer et al. 2014 are represented in red lines. Each protein is represented in its schematic domain organization and its appropriate amino acid positions. Sto1, Tho2, Hpr1, Sac3 and Mex67 are not drawn to scale. RRM: RNA recognition motif; RGG: arginine-glycine rich domain; NLS: nuclear localization signal; MIF4G: middle domain of eukaryotic initiation factor 4G; THOC2: THO complex, subunit 2, N-terminal; Tho2; THO complex, subunit 2, C-terminal; THOC1: THO complex subunit 1; THOC7: THO complex subunit 7; cc: coiled coil; Q-motif: characteristic for DEAD box helicases (GFXXPXPIQ); HELIC ATP binding: ATP-binding domain; HELICc: C-terminal helicase domain; N-box/C-box: highly conserved in REF proteins; SAP: SAF-A/B, Acinus and PIAS; WHEP-domain: discovered in TrpRS (W), HisRS (H), GluProRS (Lunde et al.); zf: C3H1-zinc finger; HD: superhelical domain; WH: winged helix domain; EFh; EF-hand; NTF2: nuclear transport factor 2. MS-analysis was performed by Alexander Wulf.

## 6.2 Mutagenesis of mRNA-binding sites of Npl3

The second major aim of this study was to figure out the functional relevance of the mRNA-binding activity of the newly identified binding sites. To do so, we focused on the SR-like protein Npl3 and mutated RNA-interacting residues that we had identified. Npl3 is involved in many aspects of mRNA metabolism like transcription, mRNA export, splicing, poly(A) tail processing and translation. Npl3 contains two canonical RNA recognition motifs (RRMs) connected by a flexible non-structured linker region in its centre, an RGG domain in its C-terminal region and an APQE-rich domain in its N-terminus (Figure 6A). The identified RNA-protein crosslinks for Npl3 are distributed over both RRMs and the RGG motif. Within RRM1, crosslinks were detected at the canonical phenylalanine in the RNP1 motif. A bit unexpected, two RNA-crosslinked amino acids (P196 and A197) were identified within the linker region between the two RRMs. Since these two crosslinks were identified several times (in this study as well as in an oligo(dT)-pulldown from Kramer et al. 2014), we speculated that interactions with these amino acids contribute to the RNA-binding capacity of Npl3.

A multiple sequence alignment was performed with *S. cerevisiae* Npl3 and 150 protein sequences of homologs across different yeast species using the ConSurf Server with standard settings (MAFFT, Homolog search algorithm: HMMER, Maximal and Minimal %ID between Sequences: 95 % and 35 %, respectively) to assess the conservation of the amino acids that we wanted to mutate (Figure 6B). A higher degree of conservation at the crosslinked amino acids suggests that the interaction with RNA at this position may also be conserved, a potential indicator of biological relevance. To reduce the possibility that mutagenesis induced protein misfolding or degradation, the PoPMuSiC v3.0 web server from dezyme was used, which gives an *in silico* prediction of protein thermodynamic stability changes upon single-site mutation. All generated strains that harbor a point mutation in *NPL3*, including those that were not used for further analysis are listed in (Supplementary table 2).

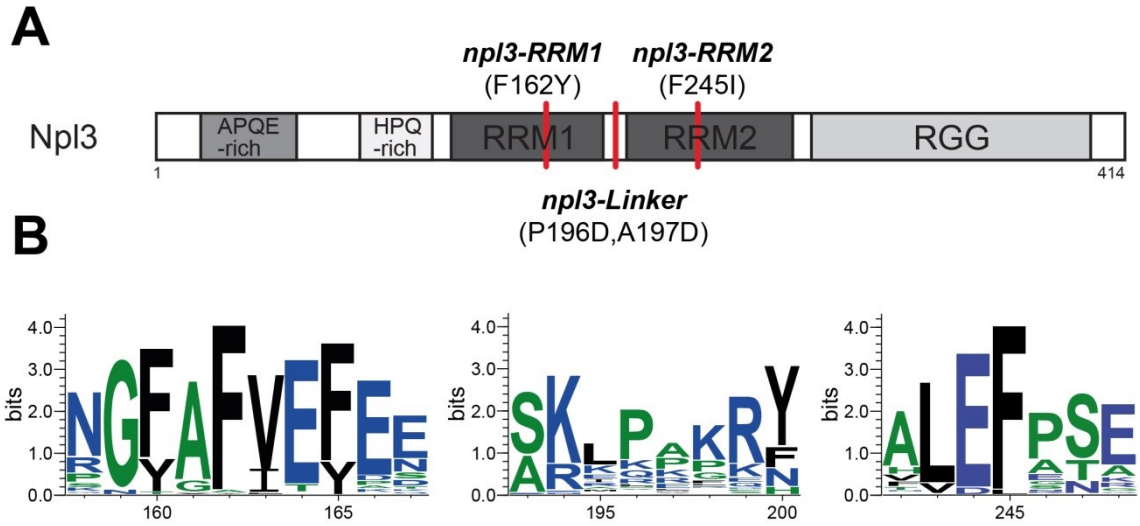



**Figure 6: Chosen RNA-crosslinks of Npl3 for mutagenesis and their conservation. (A)** Schematic overview of the domain architecture of Npl3. The red vertical lines represent the mutations that we introduced. RRM: RNA recognition motif. **(B)** Conservation of regions within Npl3 with newly identified crosslinks derived from multiple sequence alignment (MSA) depicted with WebLogo 3.6. Left panel: RNP1 of RRM1 (crosslink at position F160, F162, F265), middle panel: Linker-region between both RRMs (crosslink at position P196 and A197), right panel: pseudoRNP1 of RRM2.

## 6.3 Introduced mutations of mRNA-binding sites of Npl3 result in strong growth defects

In the absence of Npl3, cellular fitness is reduced and mutants display a slow growth phenotype. To get a first indication whether mutation of the binding sites affected the function of the protein, we determined growth rates by dot spot assay and liquid growth curves at different temperatures. The analysis of cell growth and expression level (6.4) in the newly generated mutants served as a pre-screen to exclude mutants that don't show growth defects, are lethal or have a lower expression level of Npl3. All three phenylalanines inside the RNP1 of RRM1 of Npl3 crosslinked and were mutated (F160, F162, F165; Figure 6B, left panel). Moreover, all possible combinations of those single mutants were generated. Afterwards, the growth properties of the different *npl3-RRM1* mutant variants were analyzed by a dot spot assay (Figure 7A). Every *RRM1-mutant* that carried the F162Y substitution was lethal at 37°C and showed a mild to strong growth defect at the other tested temperatures. The growth behavior did not differ between the double mutants (F160Y,F162Y; F160Y,F165Y; F162Y,F165Y). However, the cells that harbor the triple mutation (F160Y,F162Y,F165Y) grew slightly worse compared to single and double mutations that carry F162Y (Figure 7A). For further analysis, the single F162Y mutant was selected. P196 and A197 inside the linker region between the two RRM domains of Npl3 both crosslinked to RNA. Therefore, both were mutated into aspartic acid as single mutants; a combination of both was generated as well. The *npl3-P196D* and the *npl3-P196D,A197D* mutant both grew slowly at 16°C and at 37°C, although the phenotype of the double mutant is much stronger (Figure 7B). Interestingly,

the *npl3-A197D* mutation seems to enhance the effect of *npl3-P196D*, even though *npl3-A197D* alone shows no growth defect at all. Inside the RRM2 of Npl3, F245 of the pseudoRNP1 was mutated to a tyrosine or an isoleucine. Both RRM2 mutant variants cause a slight temperature-sensitive growth phenotype at 16°C, 25°C and 30°C and both are lethal at 37°C (Figure 7C). Since *npl3-F245I* shows a slightly better growth, this one was selected for further analysis (Figure 7C and 6E).

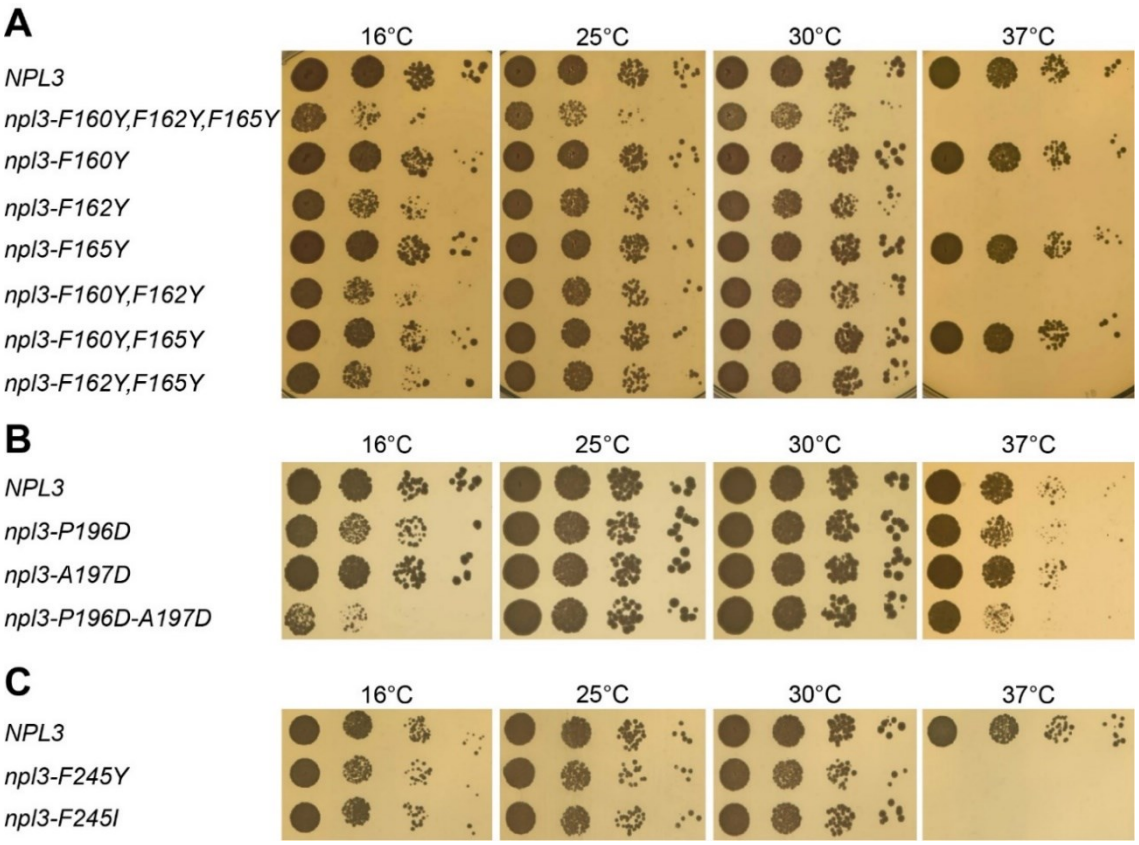



**Figure 7: Dot spot assay of cells carrying indicated NPL3 mutations in the appropriate crosslink position.** A 10-fold serial dilution series of wild-type cells and indicated mutants were spotted onto YPD plates and incubated for 2-3 days (25°C, 30°C, 37°C) or 7 days (16°C). Cells expressing mutated Npl3 variants inside RNP1 of RRM1 **(A)**, inside the linker region between both RRM domains **(B)** and inside the pseudo RNP1 of RRM2 **(C)**.

The finally selected *npl3* mutants were named according to their position within the protein: *npl3-RRM1* (F162Y), *npl3-Linker* (P196D,A197D) and *npl3-RRM2* (F245I). All three mutants showed growth defects in both assays (Figure 8 and Figure 9). Particularly, the growth defect of the *npl3* mutants is enhanced at lower and higher temperatures compared to 30°C. The *npl3-RRM1* mutant shows a strong growth defect at 16°C, a mild defect at 25°C, almost like wild type at 30°C and is lethal at 37°C. The *npl3-Linker* mutant shows a strong growth defect at 16°C and 37°C and very faint growth defect at 25°C and 37°C. The *npl3-RRM2* mutant is lethal at 37°C and growth at all other temperatures almost like wild type (Figure 8A). Furthermore, the different mutants were combined to screen for synergetic defects caused by genetic interaction of the different

single mutations. Any combination of *npl3-RRM2* results in lethality of the cells (Figure 8B). The *npl3-RRM1-Linker* mutant shows stronger growth defects than each of the two single mutations. It results in lethality at 16°C and 37°C and shows a very strong growth defect at 25°C and 30°C. However, the *npl3-RRM1-Linker* mutant as well as all other novel introduced mutants grow still better compared to the Δ*npl3* mutant (Figure 8A).

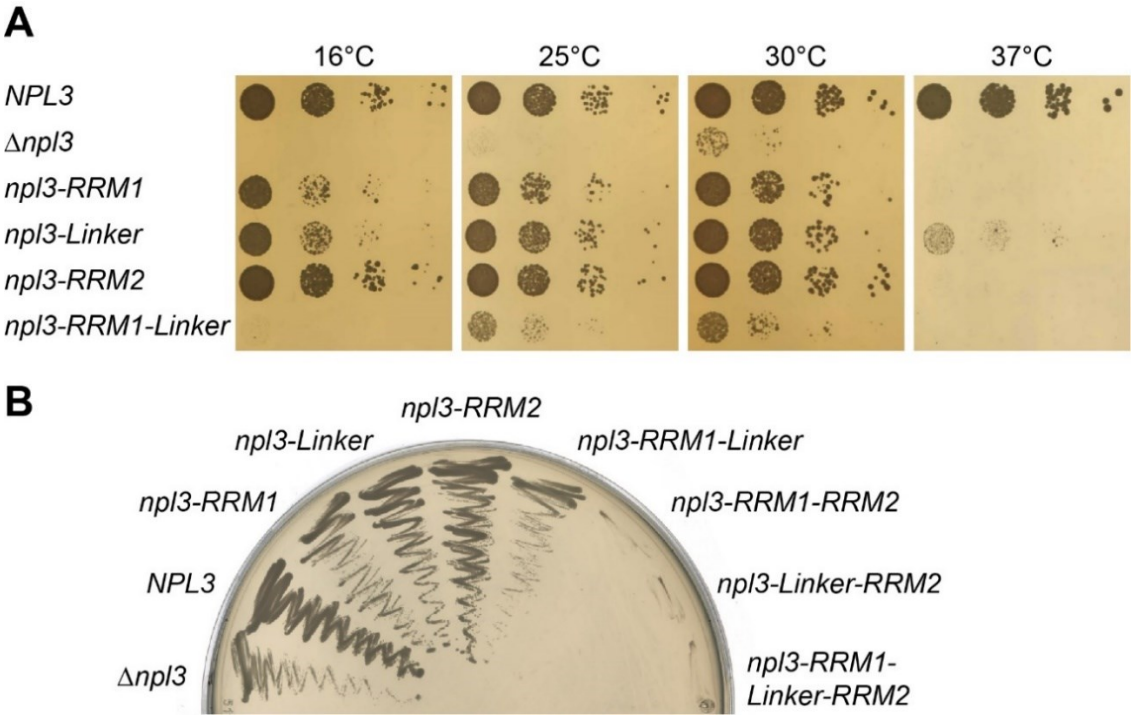



**Figure 8: Growth properties of novel Npl3 mutants. (A)** Dot spot assay with indicated Npl3 mutants at indicated temperatures. A 10-fold serial dilution series of wild-type cells and indicated mutants were spotted onto YPD plates and incubated for 2-3 days (25°C, 30°C, 37°C) or 7 days (16°C). **(B)** A shuffle strain with genomic deletion of *NPL3* carrying a URA3-plasmid (pRS316) encoding NPL3 was transformed with a LEU2-plasmid encoding the indicated Npl3 mutation. Yeast cells were streaked out on 5'-FOA plates to shuffle out the URA3-plasmid and incubated for 3 days at 30°C. Combination of any Npl3 mutant with *npl3-RRM2* leads to lethality.

The growth properties of all three mutants were also tested by liquid growth assay, which is more sensitive than dot spot assays. Although the growth defect of the three Npl3 mutants at 30°C is rather mild in the dot spot assay (Figure 8A), the defect is clearly visible in the liquid growth curves (Figure 9A, C, E). Wild-type *NPL3* cells have a doubling time of 1.52 h in YPD at 30°C; the three mutants have a doubling time of: *npl3-RRM1* 1.75 h, *npl3-Linker* 1.65 h and *npl3-RRM2* 1.73 h. Since the *npl3-RRM1* and *npl3*-RRM2 mutants are lethal at 37°C, this temperature was only tested for the *npl3-Linker* mutant (doubling time of wild-type *NPL3*: 1.83 h and *npl3-Linker*: 5.84 h at 37°C; Figure 9C). Moreover, cells were grown at 30°C and shifted to 37°C to determine at which time point the cells start to grow more slowly or die (Figure 9B, D, F). This was important to determine optimal conditions for subsequent experiments such as FISH, where cells were analyzed at 37°C. It takes approximately one hour until slower growth or a clear

growth arrest are detectable after shifting to 37°C. The *npl3-RRM2* mutant arrests more quickly and cells stop to grow after 0.6 to 1 hour (Figure 9F).

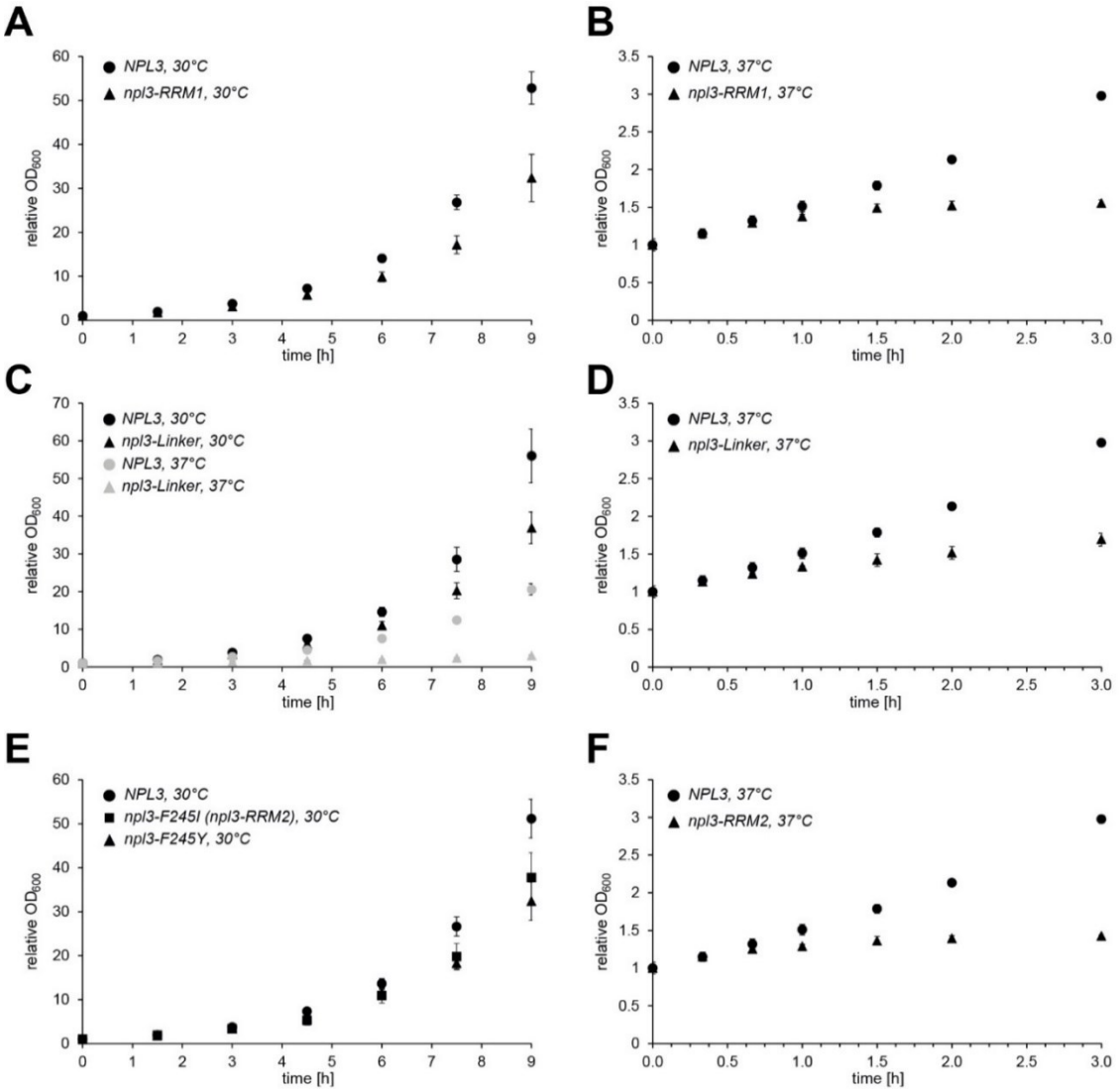



**Figure 9: Growth curves of the *npl3* mutants at 30°C and 37°C. (A, C and E)** Growth curves of various mutants at the indicated temperatures. **(B, D and F)** Growth curves of *npl3* mutant cells that were grown at 30°C and shifted to 37°C at time point 0. Mean ± standard deviation of three independent biological replicates.

Wild-type cells that carry a plasmid that overexpresses Npl3 wild-type or any of the mutated variants do not show any growth defect at 16°C, 25°C, or 30°C. However, overexpression of *npl3-RRM1* or *npl3-Linker* is lethal at 37°C, which appear to have a dominant negative effect at this temperature (Figure 10). Toxicity at 37°C is not observed for *NPL3* wild-type or *npl3-RRM2*.

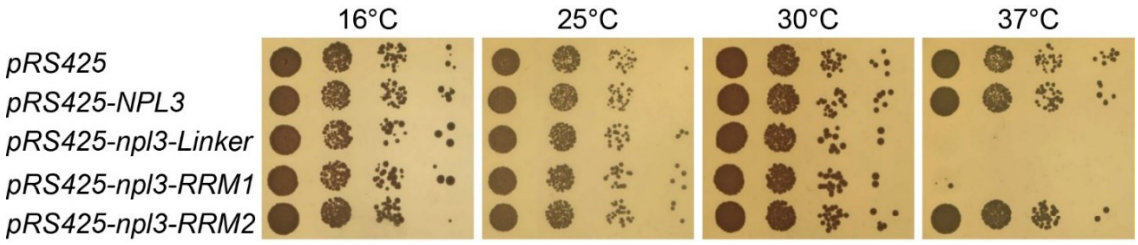



**Figure 10: Yeast cells overexpressing npl3-RRM1 or npl3-Linker show a dominant negative growth phenotype.** 10-fold serial dilutions of wild-type yeast cells transformed with a multi-copy plasmid (pRS425) encoding the indicated mutant of Npl3 were spotted onto SDC-leu plates and incubated for 2-3 days (25°C, 30°C, 37°C) or 7 days (16°C).

In summary, mutations in *NPL3* at positions that crosslink to RNA *in vivo* result in growth defects at the indicated temperatures. Furthermore, combination of different mutations exacerbate the phenotype.

## 6.4 Expression level of Npl3 in *npl3* mutants is like wild-type

To exclude that the growth defects are a result of a lower expression level of mutated Npl3, levels of wild-type and mutated Npl3 were compared by Western blotting of whole cell extracts. All mutants tested express the Npl3 protein at wild-type level (Figure 11). This shows that the growth defects are not caused by a change in protein level.

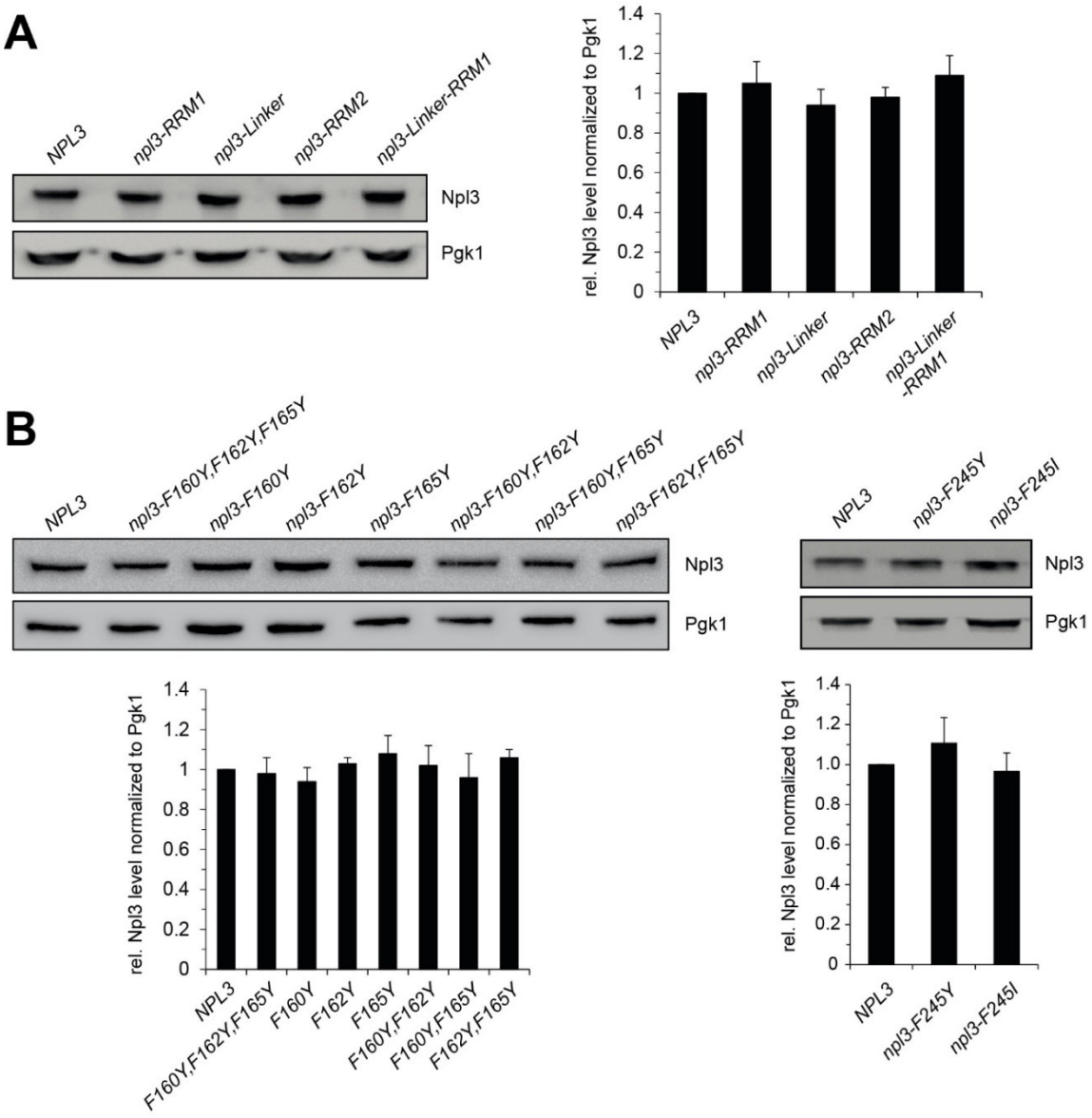



**Figure 11: *In vivo* expression level of Npl3 in different mutant variants.** Total Npl3 protein level of mutated variants is at wild-type level. Protein levels were determined by Western blot of whole cell extracts and quantified by FIJI (ImageJ) for at least three biological replicates. Pgk1 served as a loading control. **(A)** Npl3 mutants selected for further analysis. **(B)** Pre-selected variants with mutations in RRM1 (left panel) and RRM2 (right panel). Bars represent the mean ± standard deviation of three independent biological replicates.

## 6.5 RNA-binding ability of *npl3* mutants is reduced *in vivo*

To show that positions of Npl3 that crosslinked to RNA *in vivo* are truly involved in RNA binding the RNA-binding ability of the three Npl3 mutants *npl3-RRM1*, *npl3-Linker* and *npl3-RRM2* were analyzed. To determine RNA binding *in vivo,* RNA immunoprecipitation (RIP) experiments were performed with full length protein (Figure 12) and different NMR methods were used for *in vitro* analysis with truncated Npl3 versions.

For RIP, TAP-tagged Npl3 variants were immunoprecipitated with IgG-sepharose from whole cell extracts. The co-purified RNA was analyzed by reverse transcription and quantitative PCR (RT-qPCR). All three mutated Npl3 proteins showed a significant

reduction in the RNA-binding capacity for the tested transcripts (*PMA1*, *CCW12*, *YEF3*) compared to the wild-type (Figure 12), confirming that the identified crosslinks are involved in the RNA binding of Npl3.

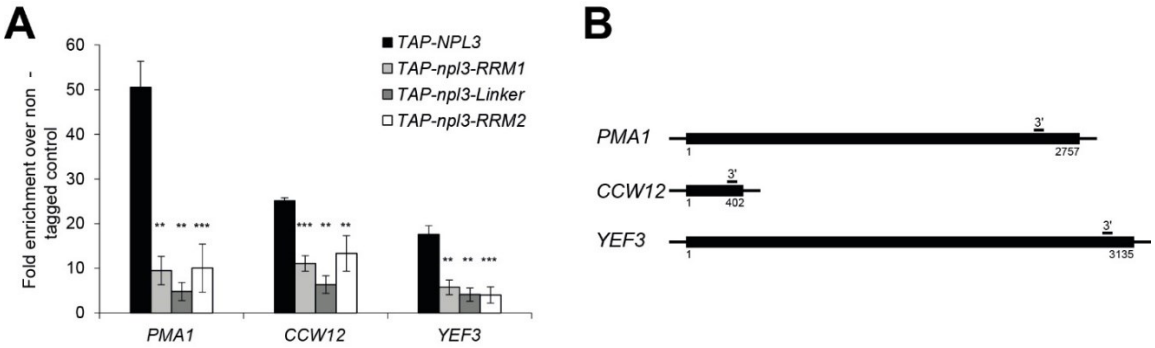



**Figure 12: RNA immunoprecipitation of TAP-tagged Npl3 mutants compared to wild-type Npl3. (A)** All three Npl3 mutants bind significantly less of the indicated transcripts *in vivo*. The level of mRNA was determined as enrichment over the level of a non-tagged Npl3 variant. Bars represent the mean ± standard deviation of three independent biological replicates, p-values of student's t-test: ** ≤ 0.01; *** ≤ 0.001. **(B)** Schematic of the transcripts with primer position used for the RT-qPCR in (A). Experiment performed by Kristin Hühn.

## 6.6 RNA-binding ability of *npl3* mutants is reduced *in vitro* and the *npl3-RRM1* and *npl3-Linker* mutants are structurally intact

To characterize the RNA-binding ability of Npl3 *in vitro*, nuclear magnetic resonance (NMR)-spectroscopy was used. At the same time, NMR-spectroscopy was used to describe the structure of Npl3 and verify protein integrity in case of the mutant variants. For all NMR-assays a truncated version harboring the two RRMs with its linker region (aa 120 - 280) of wild-type Npl3 and the three mutants was used (Figure 10A). Although the structures of both single RRM domains have been resolved previously (2OSQ and 2OSR by (Deka *et al.*, 2008); 2JVO and 2JVR by (Skrisovska and Allain, 2008)), the arrangement in RNA-bound vs. unbound state is so far poorly understood. Based on the structural model, a strong accumulation of positively charged amino acid side chains is visible at the binding surface of the two RRMs facing each other and the linker region (Figure 13B). To verify the structural integrity of the three Npl3 mutants and to thereby exclude that any effects of the mutants derive from unfolded or denatured protein, NMR fingerprint spectra of Npl3$^{120\text{-}280}$ were compared to npl3$^{120\text{-}280}$-RRM1, npl3$^{120\text{-}280}$-Linker and npl3$^{120\text{-}280}$-RRM2 (Figure 13C and D). Some distinct chemical shift perturbations (CSPs) are present in the linker region of npl3$^{120\text{-}280}$-Linker, suggesting structural rearrangement of this flexible region. For npl3$^{120\text{-}280}$-RRM1, only a few minor CSPs are noticeable (Figure 13C and D, left and middle panel). In case of npl3$^{120\text{-}280}$-RRM2 intense line-broadening for almost every residue in RRM2 domain is observed (Figure 13C and D, right panel). This indicates that the *npl3-RRM2* mutation strongly affects the structural integrity of the RRM2 domain. However, the RRM1 domain in npl3$^{120\text{-}280}$-RRM2 appears

to be intact. In sum, the NMR spectra show that the overall structural integrity is not affected by the *npl3-RRM1* or the *npl3-Linker* mutation, while the *npl3-RRM2* mutation results in an unfolded RRM2 domain.

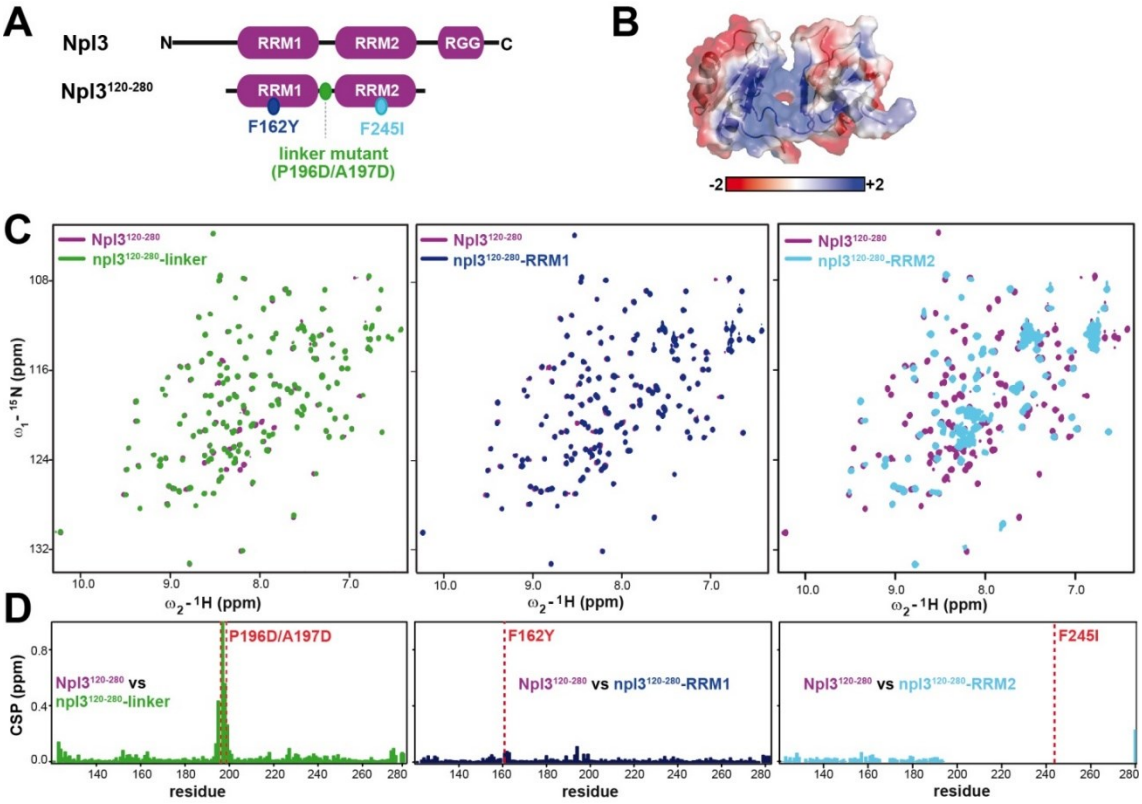



**Figure 13: Structural integrity of truncated Npl3 version by nuclear magnetic resonance (NMR)-spectroscopy. (A)** Schema of Npl3 with its domain organization of full length protein (upper) and truncated version used for NMR (lower). Mutant position of *npl3-RRM1*, *npl3-Linker* and *npl3-RRM2* are indicated. **(B)** Structural model of truncated Npl3 variant used for NMR shown with electrostatic surface potential. blue = positive charge, red = negative charge. **(C)** NMR $^{1}$H-$^{15}$N correlation spectra of Npl3$^{120-280}$ wild-type (purple) superimposed with Npl3$^{120-280}$-Linker (green; left), Npl3$^{120-280}$-RRM1 (dark-blue; middle) and Npl3$^{120-280}$-RRM2 (cyan; right). **(D)** Chemical shift perturbation of Npl3$^{120-280}$ mutants, same order as in (C). Experiment performed by Nitin Kachariya.

To assess the *in vitro* RNA-binding ability of the three *npl3* mutants compared to wild-type, NMR titrations and isothermal titration calorimetry (ITC) were performed using two different 13-mer RNA oligos with $^{15}$N-labelled Npl3$^{120-280}$. The sequences of the RNA oligos (5'-AGCACCGUAAAGA-3' and 5'-AGCACCGUGGAGA-3') are based on published data. In (Deka *et al.*, 2008) and (Cléry *et al.*, 2013), the binding of Npl3 was investigated by NMR and electrophoretic mobility shift assay (EMSA) with different RNA oligos. Wild-type Npl3$^{120-280}$ shows low RNA-binding affinity to the UAA RNA oligo in the NMR titration assay (Figure 14A). The NMR signals shift more with increasing concentration of RNA. A 4-fold molar surplus of RNA is necessary to achieve saturation (Figure 14). CSPs are clearly detectable for both RRM domains as well as for the linker region in the wild-type protein (Figure 14B). For RRM1 the CSPs can be assigned to the conserved RNP motifs. Despite the lack of conserved RNP motifs in RRM2, several

strong CSPs are observable as well. The linker region also shows some significant CSPs, with the strongest shifts at position K198 and K199 (Figure 14B). This indicates that the linker region is involved in the RNA binding as well. In ITC measurements, no binding was detectable for the RNA carrying the UAA motif (Figure 15E), indicating once more the low binding affinity to this motif. The $K_D$ value for UAA-RNA was calculated via NMR-titration to ~150 µM (Figure 14J and K). Compared to the UAA motif, the binding of the UGG-containing RNA is much stronger (Figure 14A). However, the spectral changes of the CSPs map to the same binding regions, even though there are some more detectable shifts for the UGG-oligo (Figure 14B). In ITC measurements binding could be determined for the UGG-containing RNA with a measured $K_D$ of 0.66 ± 0.04 µM (Figure 15A). Overall, the detected RNA binding regions correspond to the positively charged surface of the structural model with electrostatic surface potential (Figure 13B). Taken together, NMR-titrations and ITC measurements show clearly that both RRM domains as well as the linker region are involved in RNA binding of Npl3, with a higher affinity for the UGG motif than for the UAA motif.

To assess RNA binding of the *npl3$^{120-280}$-RRM1*, *npl3$^{120-280}$-Linker* and *npl3$^{120-280}$-RRM2* mutants *in vitro,* the mutant proteins were titrated against both RNA oligos. For the *npl3-Linker* mutant, almost no shifts or spectral changes were detected with the UAA-containing RNA, even at 4-fold molar excess. Some strongly reduced shifts are detectable using the UGG motif-containing RNA. In the ITC assay, the *npl3-Linker* mutant did not show any detectable binding for any of the used RNAs (Figure 15B and F). The NMR titration and ITC experiments of *npl3$^{120-280}$-RRM1* show strongly reduced RNA binding (Figure 14G and H, Figure 15C). Using the UGG-containing RNA in ITC experiments resulted in a $K_D$ value of 10.4 ± 0.2 µM for *npl3-RRM1* (Figure 15C), a ~15-fold reduction in RNA binding compared to wild type. The calculated $K_D$ value of the *npl3$^{120-280}$-RRM2* mutant is 4.8 ± 0.3 µM using the UGG RNA in the ITC experiments (Figure 15D), representing approximately 10-fold reduced RNA affinity compared to wild type. That the *npl3$^{120-280}$-RRM2* mutant shows a better binding compared to the other mutants was a bit surprising since this mutation results in an unfolded RRM2 domain. Mapping the chemical shift perturbations of the wild type and the three mutants onto the structure depicts once more that the binding of both RNAs is significantly reduced in all three mutants (Figure 14C, F and I).

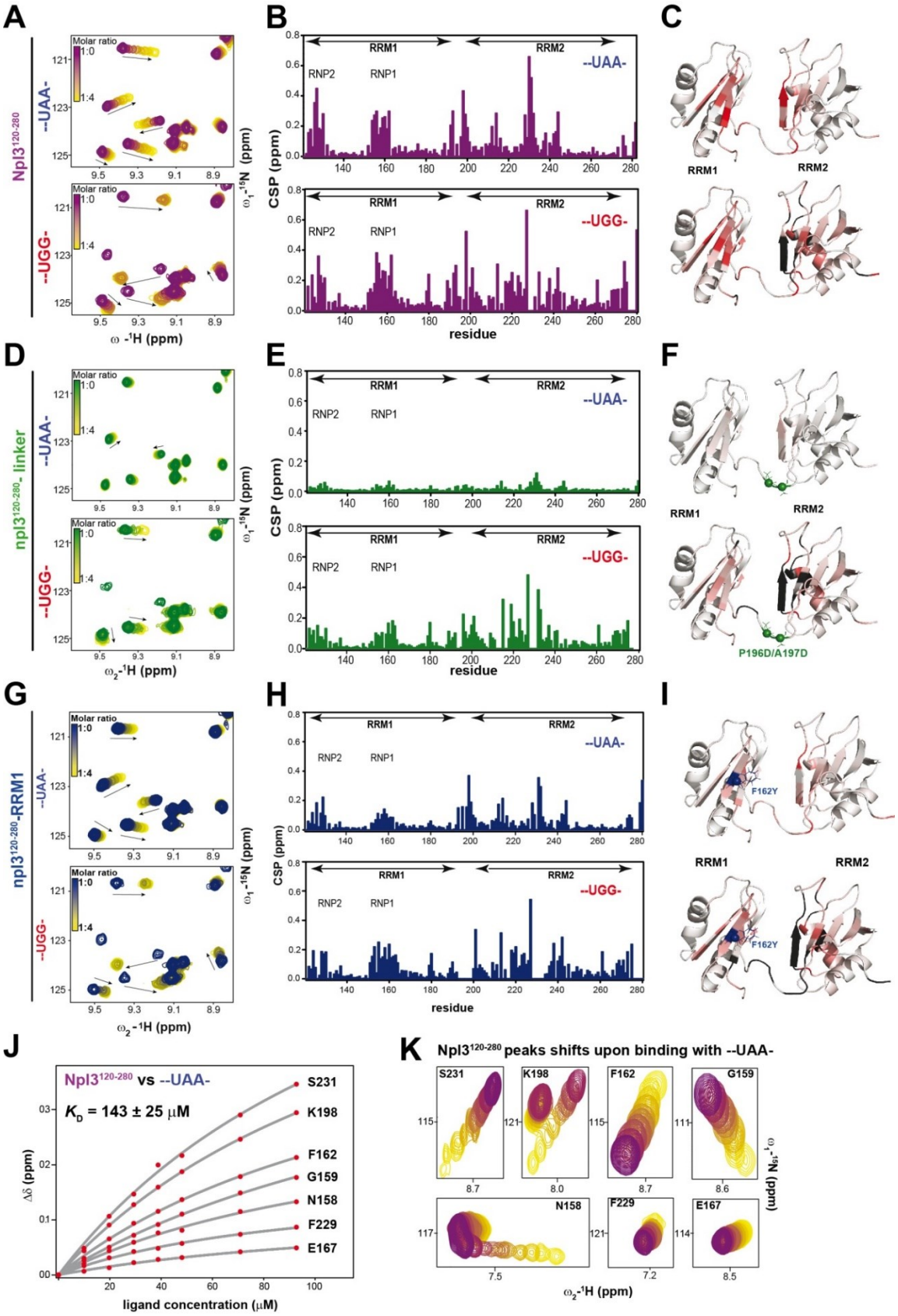



**Figure 14: NMR titration to determine the RNA binding affinity of npl3 mutants. (A, D, G)** Zoomed sections of NMR spectral superposition of Npl3 wild type, *npl3-Linker* and *npl3-RRM1* titrated against UAA containing (upper) and UGG containing (lower) RNA. Increasing chemical shift perturbations (CSP) are indicated by arrows. **(B, E, H)** Amino acid position and intensity of CSPs plotted for titration of UAA containing (upper) and UGG containing (lower) RNA to Npl3 wild type, *npl3-Linker* and *npl3-RRM1*. **(C, F, I)** Mapping

of RNA binding to the structure of Npl3, colored from gray to red with increasing CSP. Spheres indicate the position of the introduced mutations. **(J)** Estimated $K_D$ value from NMR CSPs of Npl3 wild type with UAA containing RNA. **(K)** Zoomed sections of NMR spectral superposition of Npl3 wild type titrated against UAA containing RNA used for estimation of $K_D$ value as shown in (J).

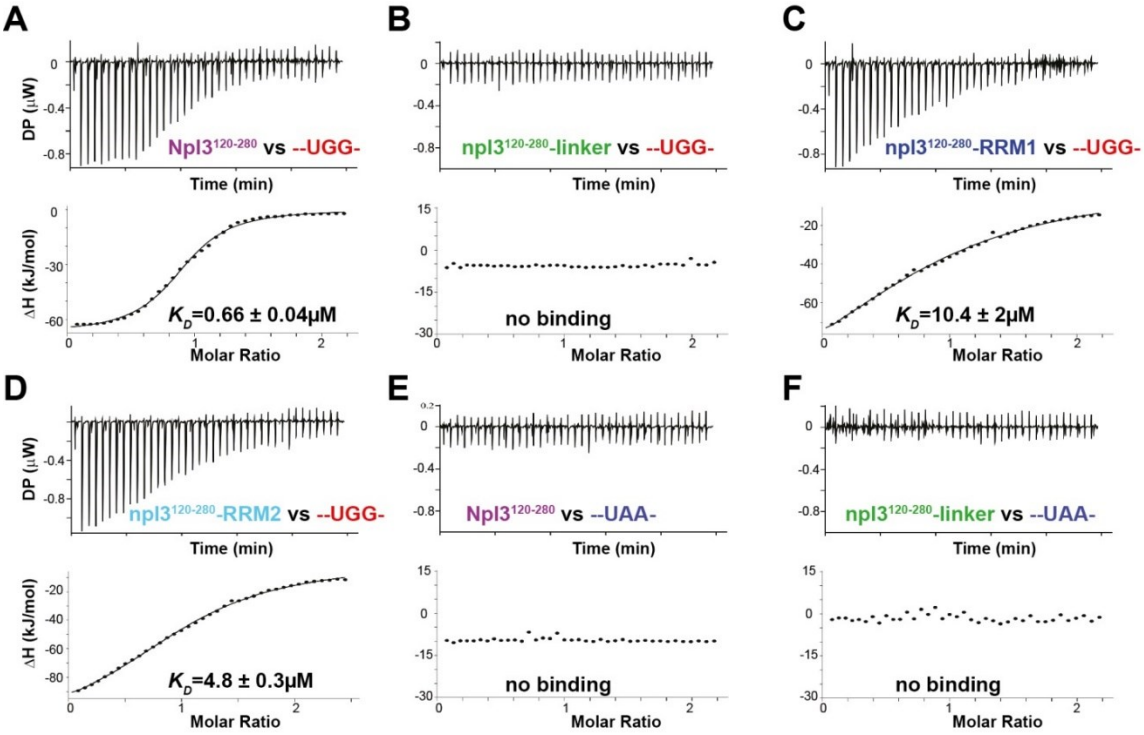



**Figure 15: Isothermal titration calorimetry (ITC) for RNA binding of npl3 mutants.** ITC for binding of UGG and UAA-containing RNA to Npl3 wild-type protein **(A and E)** and *npl3-RRM1* **(C)**, *npl3-Linker* **(B and F)** and *npl3-RRM2* **(D)** mutant proteins.

Both the *in vivo* and the *in vitro* assays provide clear evidence that the RNA-binding ability is strongly reduced in all three *npl3* mutants compared to wild type. The *npl3-Linker* mutant has the strongest mRNA binding defect. This suggests once more that the flexible region between the two RRM domains is truly and surprisingly strongly involved in the RNA-binding ability of Npl3.

## 6.7 RNA expression levels and splicing are globally affected in *npl3* mutants

To investigate whether any of the three *NPL3* mutants influence expression levels of mRNAs or affect splicing, the transcriptome was analyzed by next generation sequencing. RNA was extracted from *NPL3* wild-type and mutant yeast cultures and poly(A)-RNA was sequenced (RNA-seq). Before harvesting, yeast cells were either incubated at 30°C of shifted to 37°C for one hour. Differential expression analysis was performed using DESeq2 (doi: 10.1186/s13059-014-0550-8) and showed broad changes throughout the transcriptome in all mutants (adjusted P value < 0.05; Figure 16A - C). For all *npl3* mutants, abundance of more than 3,500 transcripts changed significantly when cells were grown at 30°C. Most of the differentially expressed

transcripts are shared between the different *npl3* mutants (Figure 16A, left panel). At 37°C, less transcripts show a change in expression level in all mutants than at 30°C, but the difference between mutants is more distinctive (Figure 16, right panel). Taken together, the loss of RNA-binding capacity of the *npl3* mutants results in a significant change to the transcriptome (Figure 16A and C). A more detailed analysis revealed that the levels of intron-containing transcripts drop significantly in all three mutants. In that regard, the *npl3-RRM1* and the *npl3-RRM2* mutant show the strongest effect at 37°C (Figure 16B). Since the lower levels of intron-containing transcripts indicate that splicing is affected in those mutants, the RNA-seq reads were quantified at the intron-exon transitions to analyze the splicing efficiency (Figure 16D). As expected, the *npl3* mutants show a defect in splicing, especially at 37°C. However, the *npl3-RRM1* and *npl3-RRM2* mutant show increased intron retention whereas the *npl3-Linker* mutant shows the opposite effect, a reduced level of intron retention compared to wild type at 37°C (Figure 16D). This indicates that the different domains of Npl3 affect splicing in different ways. Overall, the RNA-seq data of the *npl3* mutants reveal extensive variations of the mRNA levels transcriptome-wide as well as opposing splicing defects. Some effects are overlapping and some are different between the *npl3* mutants, indicating once more divergent functions of the RRM domains and the linker region of Npl3.

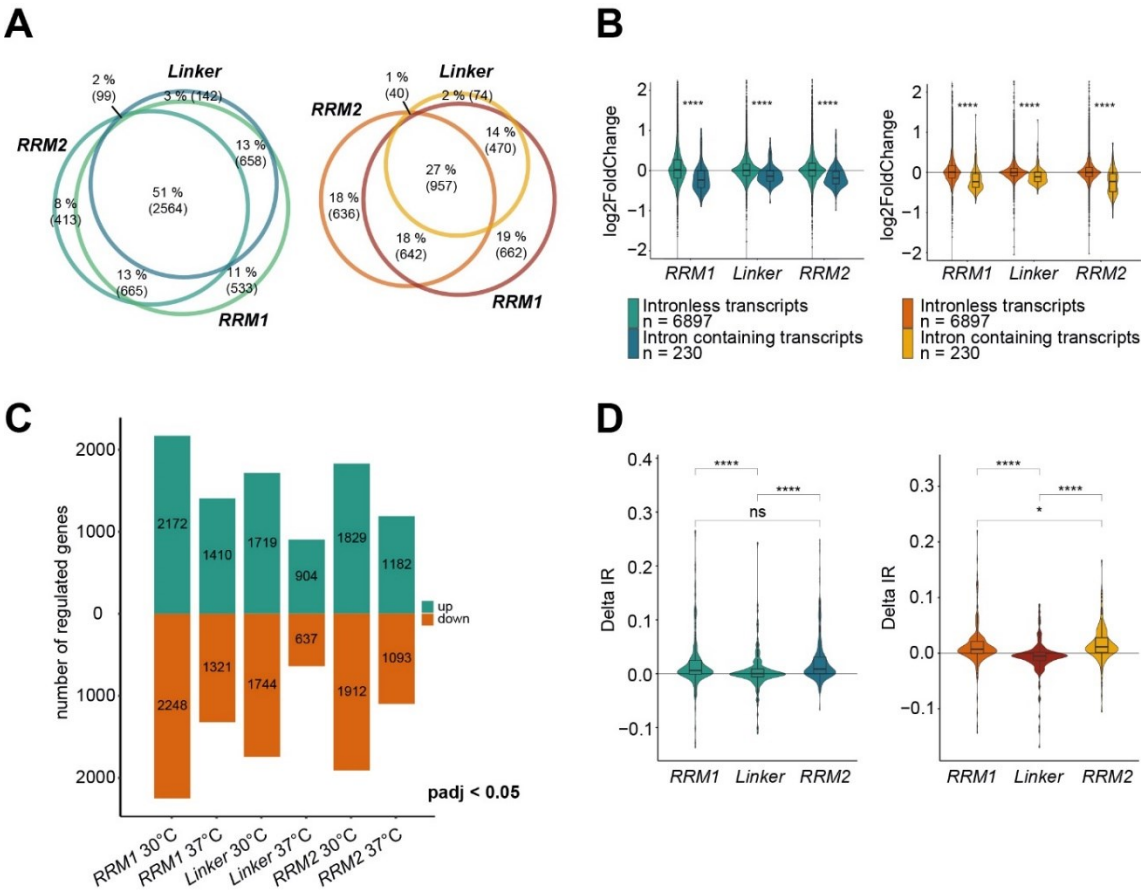



**Figure 16: RNA-seq data of *npl3* mutants reveal changes in transcript abundance and opposing splice defects. (A)** Venn diagrams display broad changes in transcript abundance in all *npl3* mutants at 30°C (left) and 37°C (right) compared to wild type. Adjusted P value < 0.05 **(B)** Intron-containing transcripts drop significantly in *npl3* mutants represented by violin boxplots. $log_2$-transformed fold changes in transcript abundance for intron-containing (n = 230) and intron-less (n = 6897) transcripts are compared between wild type and *npl3* mutants at 30°C (left) and 37°C (right). **** P value ≤ 0.0001, Wilcoxon rank-sum test. **(C)** Absolut numbers of up and down regulated transcripts of each *npl3* mutant at 30°C and 37°C. **(D)** Increased intron retention in the *npl3-RRM1* and *npl3-RRM2* mutants and slightly reduced intron retention in the *npl3-Linker* mutant depicted by violin boxplots. Intron retention (ΔIR; n = 215) for each *npl3* mutant compared to wild type at 30°C (left) and 37°C (right). * P value < 0.05, **** P value ≤ 0.0001, Wilcoxon rank-sum test. Sequencing was performed by Janine Altmüller, analysis of RNA-seq data by Samira Reuscher.

## 6.8 The half-life of selected transcripts is not effect in *npl3* mutants

To determine the half-life of selected transcripts in the *npl3* mutants, cells were treated with thiolutin to block transcription. After treatment total RNA was extracted, reverse transcribed and analyzed by qPCR. Three transcripts without intron (*PMA1, CCW12, YEF3)* and three with intron (*DBP2, ASC1, RPL28)* with different half-lives were selected. In the wild type, the determined half-lives agreed with published data (Geisberg *et al.*, 2014). None of the *npl3* mutants showed a significant change in the half-life in any of the tested transcripts (Figure 17).

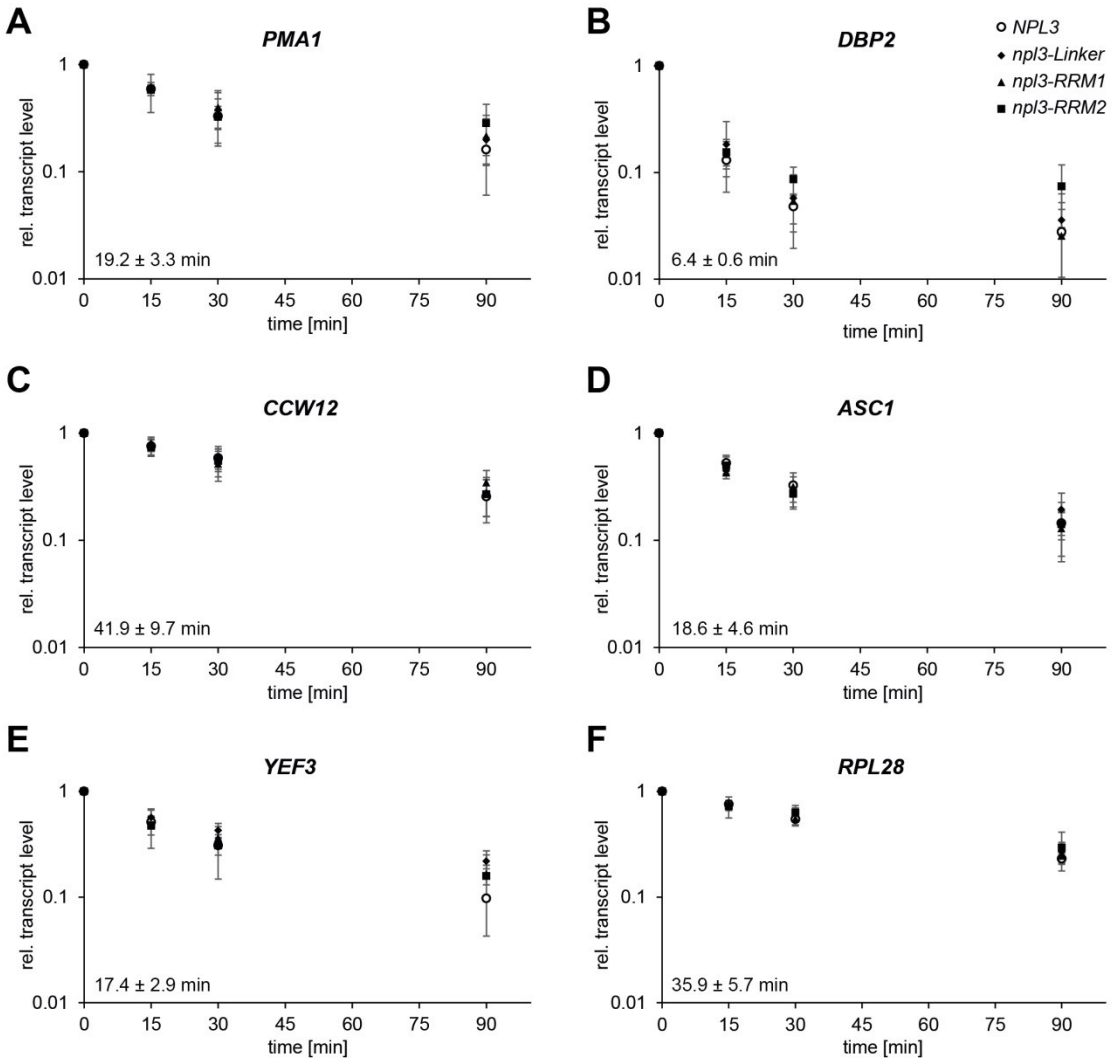



**Figure 17: Half-life determination of selected transcripts in *npl3* mutant cells.** The half-lives of selected transcripts were determined for *NPL3* wild type and *npl3* mutants. Yeast cells were treated with 8 µg/mL thiolutin to block transcription. Before treatment and 15, 30 and 90 min after treatment cells were harvested, total RNA was extracted, reverse transcribed and analyzed by qPCR. Three transcripts without intron (**A:** *PMA1*, **C:** *CCW12*, **E:** *YEF3*) and three with intron (**B:** *DBP2*, **D:** *ASC1*, **F:** *RPL28*) with different half-lives were selected. Represented is the mean ± standard deviation of three independent biological replicates. Because the Ct values of the time point 90 min after treatment were already or almost at background level in qPCR analysis, those values were not used for calculating the half-life.

## 6.9 Nuclear mRNA export is strongly defective in *npl3-Linker* mutant

To investigate the functional consequences of the *npl3* mutants, the mRNA export was visualized by fluorescence *in situ* hybridization (FISH) using an oligo($dT_{50}$) probe coupled to Cy3 against poly(A) tails. Despite the clear reduction in RNA binding ability evident in all three *npl3* mutants (Figure 12), they vary in their nuclear mRNA export phenotypes (Figure 18). The *npl3-RRM1* mutant displays no detectable mRNA export defect at 30°C and only a mild defect at 37°C. In contrast, the *npl3-Linker* cells show a strong mRNA export defect already at 30°C, which is even more pronounced at 37°C, to a level that almost no mRNA is detectable in the cytoplasm. The *npl3-RRM2* mutant has no mRNA export defect at 30°C and a minor defect at 37°C (Figure 18).

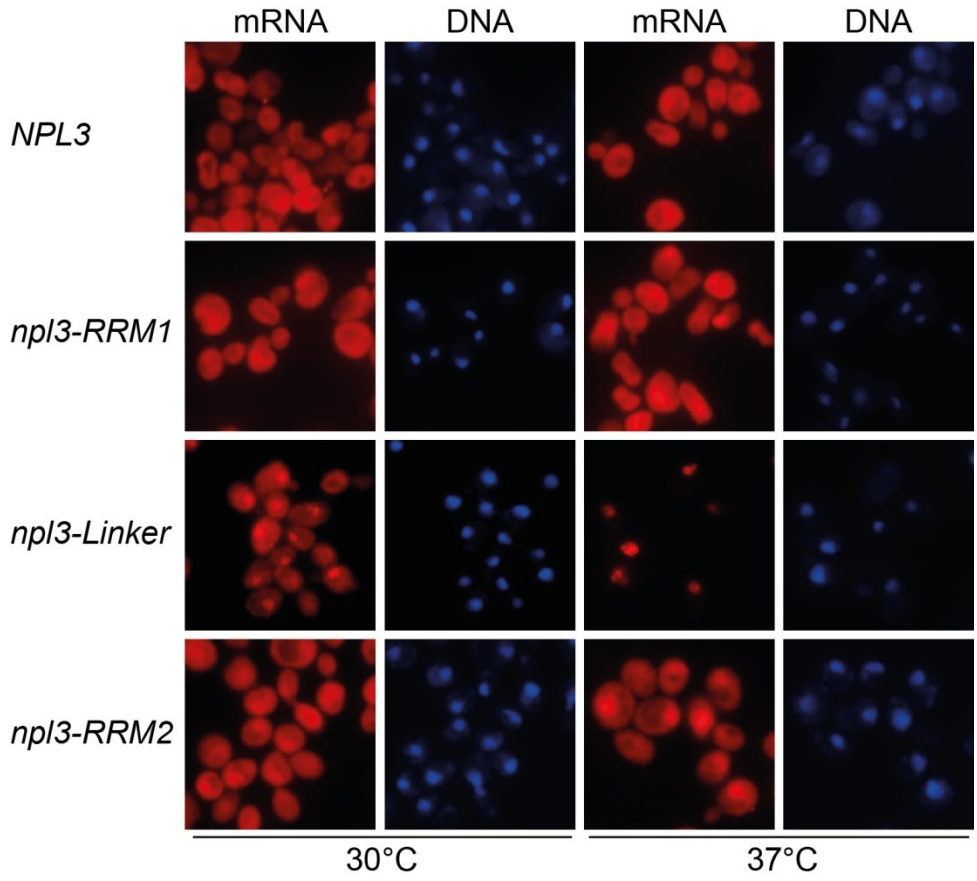



**Figure 18: Nuclear mRNA export of *npl3* mutants visualized by fluorescence *in situ* hybridization.** *npl3-RRM1* mutant cells show a minor export defect at 37°C, *npl3-Linker* cells show a strong mRNA export defect at 30°C and 37°C while *npl3-RRM2* cells show an intermediate mRNA export defect. Bulk mRNA was visualized by oligo($dT_{50}$) coupled to Cy3-fluorecent dye. DNA was stained with DAPI to locate the nucleus. Cells were either grown at 30°C or shifted to 37°C for one hour.

## 6.10 Localization of Npl3 protein in *npl3* mutant cells is not affected

Under native conditions the steady-state localization of Npl3 is in the cytoplasm. It is known from literature that some *npl3* mutants mislocalize to the cytoplasm and thereby generate disfunctions. To exclude that any of the phenotypes of the new *npl3* mutants are a result of mislocalized Npl3, an immunostain was performed. Visualizing the Npl3 protein in wild-type and mutant cells using anti-Npl3 antibody coupled to Alexa488-fluorescent dye displayed clear nuclear distribution in all cells at 30°C and 37°C (Figure 19).

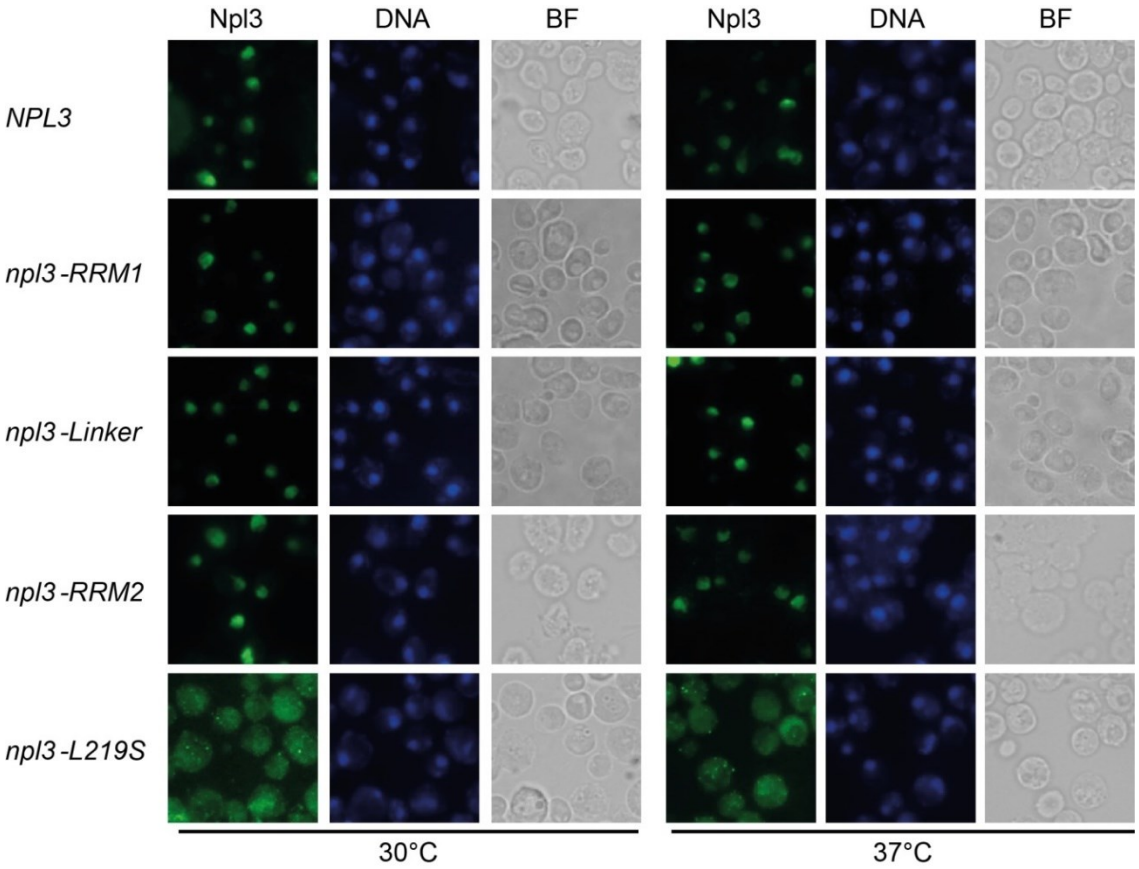



**Figure 19: Localization of Npl3 wild type and mutated variants was visualized in the cell by immunostaining.** None of the *npl3* RNA-binding mutants showed a mislocalization neither at 30°C nor at 37°C. *npl3-L219S* was used as mislocalized control cells. Npl3 protein was visualized with anti-Npl3 antibody coupled to Alexa488-fluorescent dye. DNA was stained with DAPI to locate the nucleus. Cells were either grown at 30°C (left panel) or shifted to 37°C (right panel) for one hour. BF: bright field.

## 6.11 mRNP composition is changed in *npl3* mutants

Since a nuclear mRNA export defect was verified for the *npl3* mutants, the composition of nuclear mRNPs in these mutants was analyzed. A purification until TEV-elution using genomically TAP-tagged Cbc2, the small subunit of the cap binding complex (CBC; for details see 3.1.3.1) was performed with subsequent quantitative determination of some co-purified proteins including Npl3, TREX components (Hpr1, Sub2, Yra1), the mRNA exporter subunit Mex67, Tho1 and the poly(A)-binding protein Nab2 by Western blot (Figure 20). Since all three *npl3* mutants show decreased RNA-binding affinity, it was expected that the level of Npl3 would be reduced in the nuclear mRNPs. This is the case for *npl3-Linker* and *npl3-RRM1* cells. However, the level of Npl3 in the *npl3-RRM2* mutant is increased in the nuclear mRNPs. Based on the NMR data, it seems this effect is due to dimer or oligomer formation by Npl3 induced by the *RRM2* mutation. The *npl3-RRM1* mutant exhibits only a reduced level of Hpr1 in nuclear mRNPs (Figure 20). For the *npl3-Linker* mutant, the levels of all tested THO/TREX components (Hpr1, Sub2 and Yra1), Tho1 and Mex67 decrease in the mRNP. This indicates that the accurate function of Npl3 is needed to recruit Hpr1, Sub2, Yra1, Tho1 and Mex67 to the nuclear

mRNPs or to keep them there. For the *npl3-RRM2* mutant, it was shown that the nuclear mRNP components Nab2 and Mex67 are increased (Figure 20). In sum, the three *npl3* mutants differ strongly in their phenotypes even though all three show reduced RNA binding.

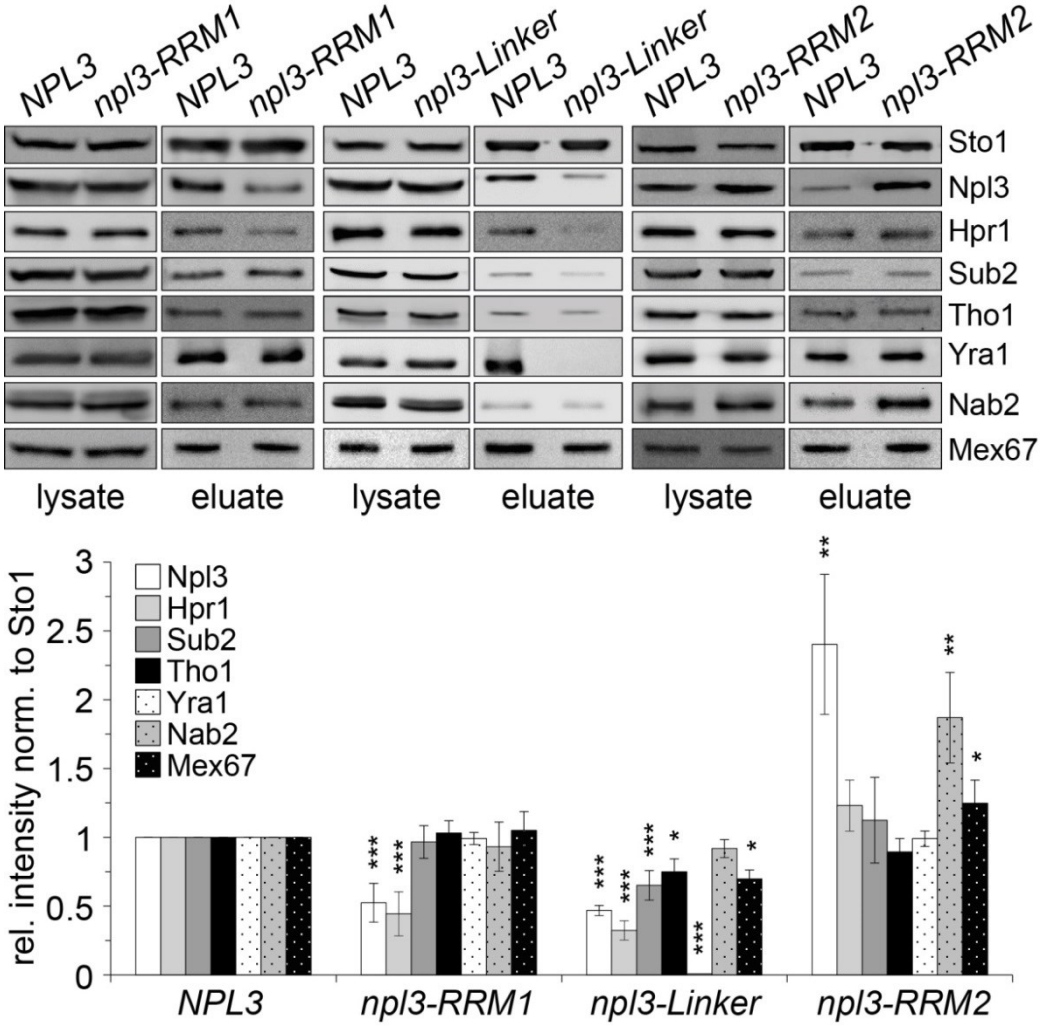



**Figure 20: Quantification of nuclear mRNP composition of npl3 mutants.** Nuclear mRNPs were purified by endogenously TAP-tagged Cbc2 until TEV-elution from wild-type, *npl3-RRM1*, *npl3-Linker* and *npl3-RRM2* cells and selected co-purifiers were determined by Western blot (upper) and quantified (lower) of at least three biological independent replicates. The quantified amounts were normalized to Sto1 and values for wild type were set to 1. * P value < 0.05, ** P value ≤ 0.01, *** P value ≤ 0.001, students t-test.

Besides the determination of the nuclear mRNP composition in the different *npl3* mutants, a purification of TAP-tagged Npl3 wild type and mutants itself was performed ± RNase A treatment. The purification was performed until TEV elution and afterwards analyzed by quantitiavtive Western blot of selected co-purified proteins including Sto1 (large subunit of CBC), the TREX components Sub2 and Yra1, the mRNA exporter subunit Mex67, Tho1 and the poly(A)-binding protein Nab2 (Figure 21). The *npl3-RRM1* mutant co-purifies significantly less Sto1 and Sub2 without RNase A treatment. Although changes for the other tested proteins are not significant, they tend

to be slightly reduced as well (Figure 21A). As expected, when treated with RNase A, co-purification of all tested proteins drops significantly for the wild type and all three *npl3* mutants, with Yra1 and Sto1 almost completely disappearing after RNase A treatment (Figure 21). Sub2, which is already reduced in the *npl3-RRM1* mutant without RNase A treatment, drops after treatment even more than the wild type (Figure 21A). The *npl3-Linker* mutant shows a very similar pattern in the Npl3-pulldown as the *npl3-RRM1* mutant, execpt that Sub2 is not as strongly reduced as it is in the *npl3-RRM1* mutant (Figure 21B). In the *npl3-RRM2* pulldown the levels of Npl3 itself as well as Nab2 are significantly increased compared to wild type. The level of Mex67 tends to be increased as well (Figure 21C). This agrees with the result from the nuclear mRNP purification (Figure 20) and can probably be explained by dimerisation of the *npl3-RRM2* mutant. Interestingly, other co-purified proteins show the same level as in the *npl3-RRM1* mutant.

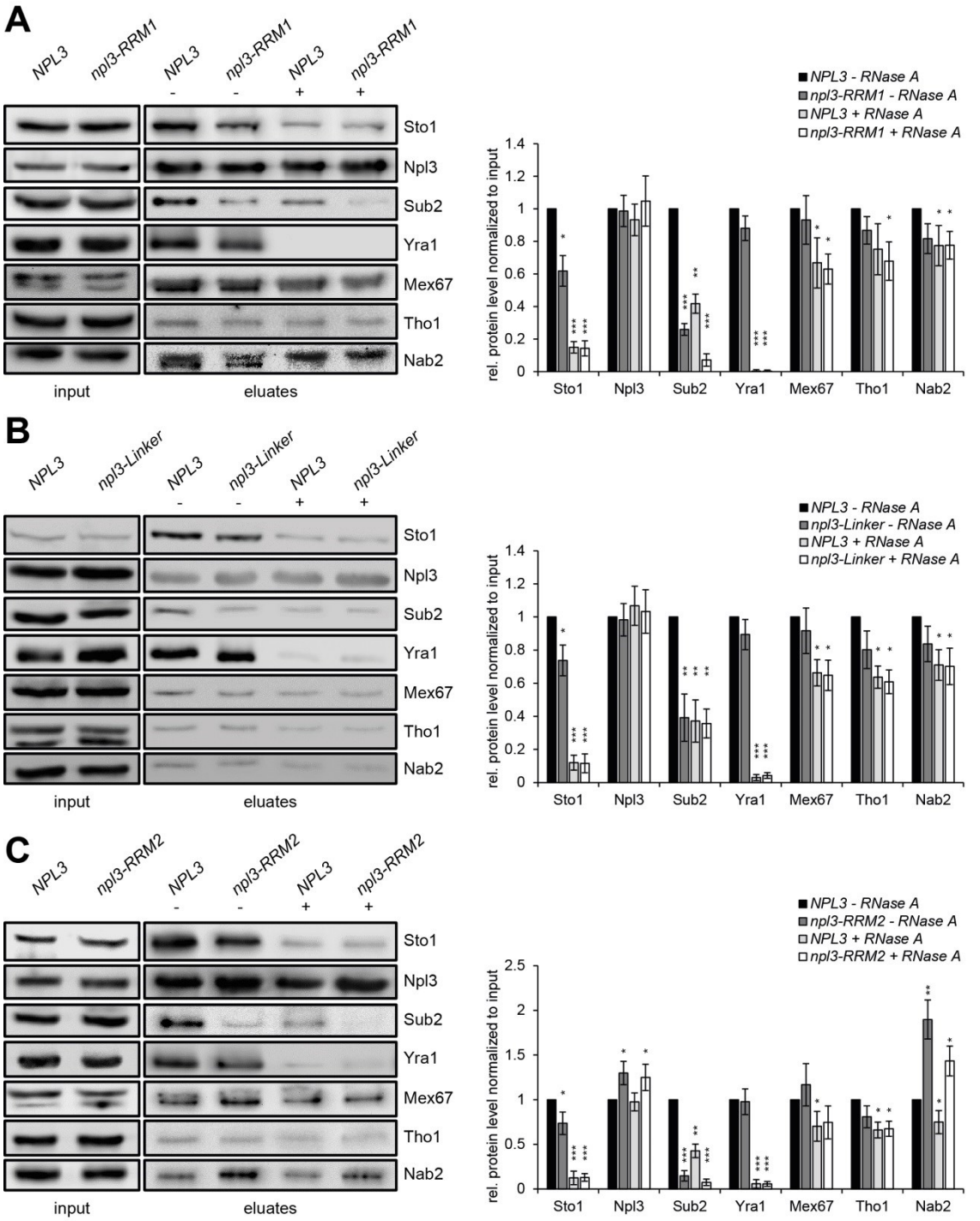



**Figure 21: Quantification of TAP-Npl3 co-purified proteins.** TAP-tagged Npl3 was purified until TEV-elution from wild-type, *npl3-RRM1* **(A)**, *npl3-Linker* **(B)** and *npl3-RRM2* **(C)** cells and selected co-purifiers were determined by Western blot (left) and quantified (right) of three biological independent replicates. Samples were treated with or without RNase A before IgG-sepharose was added to the lysates. The values for wild type were set to 1. * P value < 0.05, ** P value ≤ 0.01, *** P value ≤ 0.001, students t-test.

In addition, the TAP-purified samples of nuclear mRNPs (Cbc2) and Npl3 were analyzed by mass spectrometry with label-free quantification to get insights which other co-purifying proteins change in abundance besides the selected ones, which had been investigated in Western blot. The quantification of the co-purified proteins showed a clear decrease in the *npl3-RRM1* and *npl3-Linker* mutant, whereas the *npl3-RRM2* mutant

showed an enrichment in several proteins (Figure 22). In *npl3-Linker* cells most components of the TREX complex (Tho2, Hpr1, Mft1, Thp2, Yra1, Gb2p, Hrb1) were significantly depleted from nuclear mRNPs (Figure 22B). Tex1 shows almost a significant reduction, whereas Sub2 shows only a tendency to be reduced. Besides the TREX components, several ribosome biogenesis factors, rRNA-related proteins and translation factors like Slf1, Sro9, Noc2 and Tif4632 were also depleted. Sro9 is a shuttling protein like Npl3 and is also involved in transcription. Moreover, some DNA repair and chromatin-related proteins were also reduced in the *npl3-Linker* cells. Besides all the RNA-binding and biogenesis factors, some unexpected co-purifiers appeared as well (e. g. Imd-proteins and mitochondrial related proteins). As had been observed for nuclear mRNP purifications from *npl3-Linker* cells, a broad range of proteins co-purified significantly worse with nuclear mRNPs from *npl3-RRM1* cells than from wild type (Figure 22A). Although more proteins are lost from *npl3-RRM1* mRNPs than from *npl3-Linker* mRNPs, TREX components are unaffected in *npl3-RRM1* mRNPs, except Hrb1, which is one of the TREX components. Surprisingly, Mtr2 seems to be significantly reduced as well, while Mex67 is not affected and no mRNA export is observable in *npl3-RRM1* cells (Figure 18 and Figure 22).

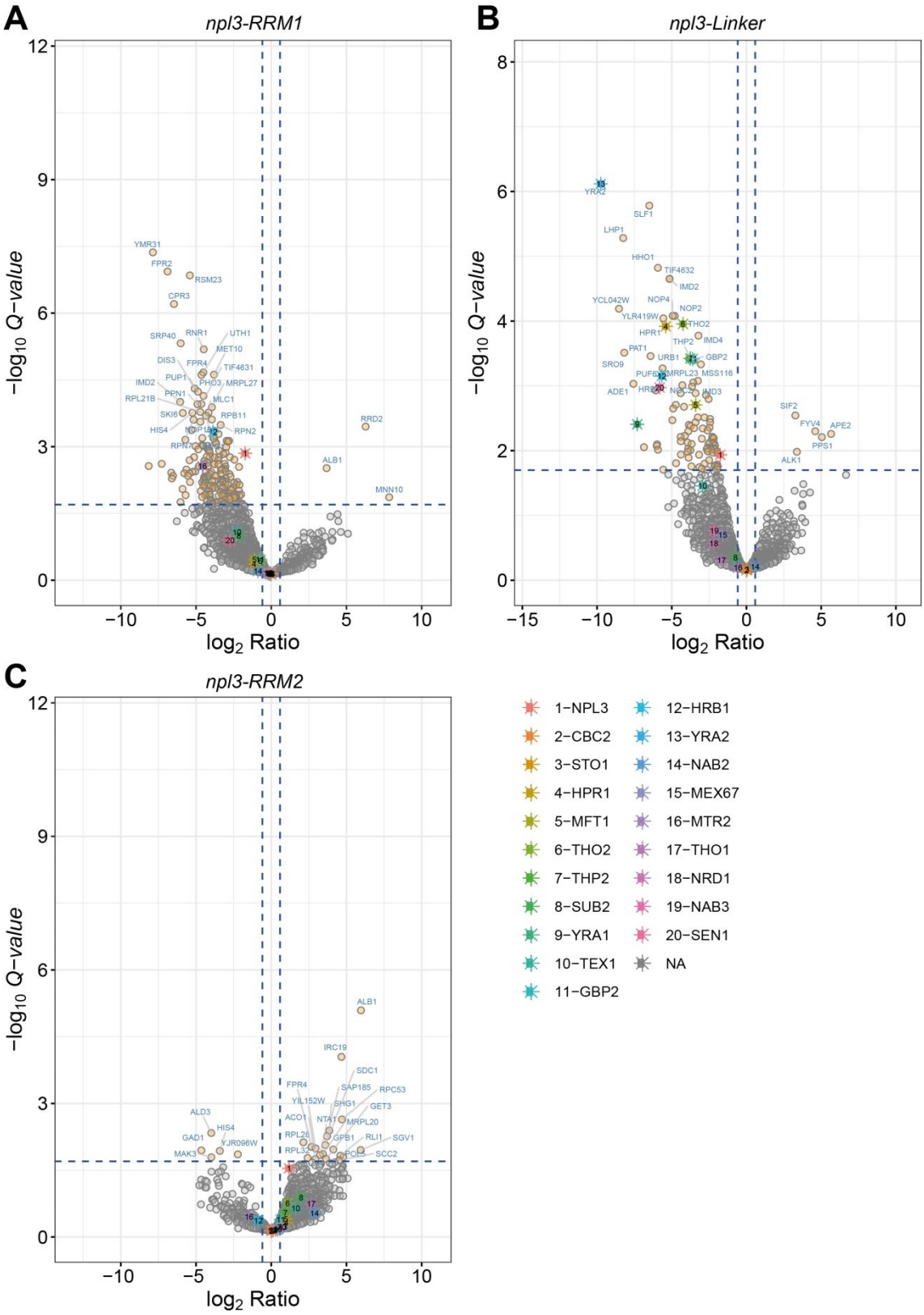



**Figure 22: MS quantification of co-purified proteins of Cbc2-TAP in Npl3 wild type and *npl3* mutants.** TAP-tagged Cbc2 was purified until TEV-elution from *NPL3* wild type, *npl3-RRM1* **(A)**, *npl3-Linker* **(B)** and *npl3-RRM2* **(C)** cells and co-purifiers were quantified by MS of three biological independent replicates. MS quantification and data analysis performed by Monika Raabe and Ivan Silbern.

Besides the nuclear mRNP purifications, TAP-purified Npl3 samples were quantified by MS as well. First, levels of co-purified proteins were compared for TAP-Npl3 (wild-type) and the TAP-*npl3* mutants (Figure 23). According to MS analysis, the levels of co-purified proteins in each mutant compared to wild type have only minor differences. Although all three *npl3* mutants bind less RNA *in vivo* (Figure 12), there is no major difference in the amount of co-purified protein detected. This indicates that the protein-protein interactions of Npl3 to other proteins are not disturbed in the *npl3* mutants. However, note that the changed levels of Sto1 and Sub2 that had been observed by Western blot could not be verified in the MS analysis (Figure 21 and Figure 23).

Next, the levels of co-purified proteins were compared between untreated and RNase A-treated samples of Npl3 wild type and each *npl3* mutant (Figure 24). This comparison revealed which of the Npl3 interactions are RNA-dependent and which are not. As expected, a large number of co-purifiers are significantly reduced upon RNase A treatment (Figure 24). However, some proteins seem to be enriched after RNase A treatment as well. TAP-purification of Npl3 wild type ± RNase A treatment demonstrated that most components of the TREX complex are strongly reduced after RNase A treatment, indicating that the association of Npl3 with the TREX complex is not by direct protein-protein interaction (Figure 24A). There are two reasons why co-purified proteins would not be affected by RNase treatment: i) They interact with Npl3 independently of RNA and therefore no change is detectable after RNase A treatment, ii) They show a low intensity in MS quantification before and after RNase A treatment, and are unlikely to be true interactors. Cbc2 and Sto1, the two subunits of the CBC as well as Yra1 are strongly reduced in RNase A-treated samples in wild type and *npl3* mutants, as shown in quantitative Western blot (Figure 21). Moreover, Gbp2 and Hrb1, the only two other SR-like proteins in *S. cerevisiae* besides Npl3, are significantly reduced as well. Interestingly, Nab3 and Nrd1 are significantly reduced in Npl3 wild type ± RNase A, whereas Sen1 shows no change. In contrast, in the quantification of Cbc2-TAP, only Sen1 showed a significant reduction in the *npl3-Linker* mutant (Figure 24 and Figure 22). This indicates that Npl3 might associate with the NNS-complex *via* direct interaction with Sen1. Nrd1-Nab3-Sen1 form together the NNS-complex that is implicated in the termination of cryptic unstable transcripts (CUTs), snoRNAs and some mRNAs (Arigo *et al.*, 2006; Creamer *et al.*, 2011; Steinmetz *et al.*, 2006; Webb *et al.*, 2014).

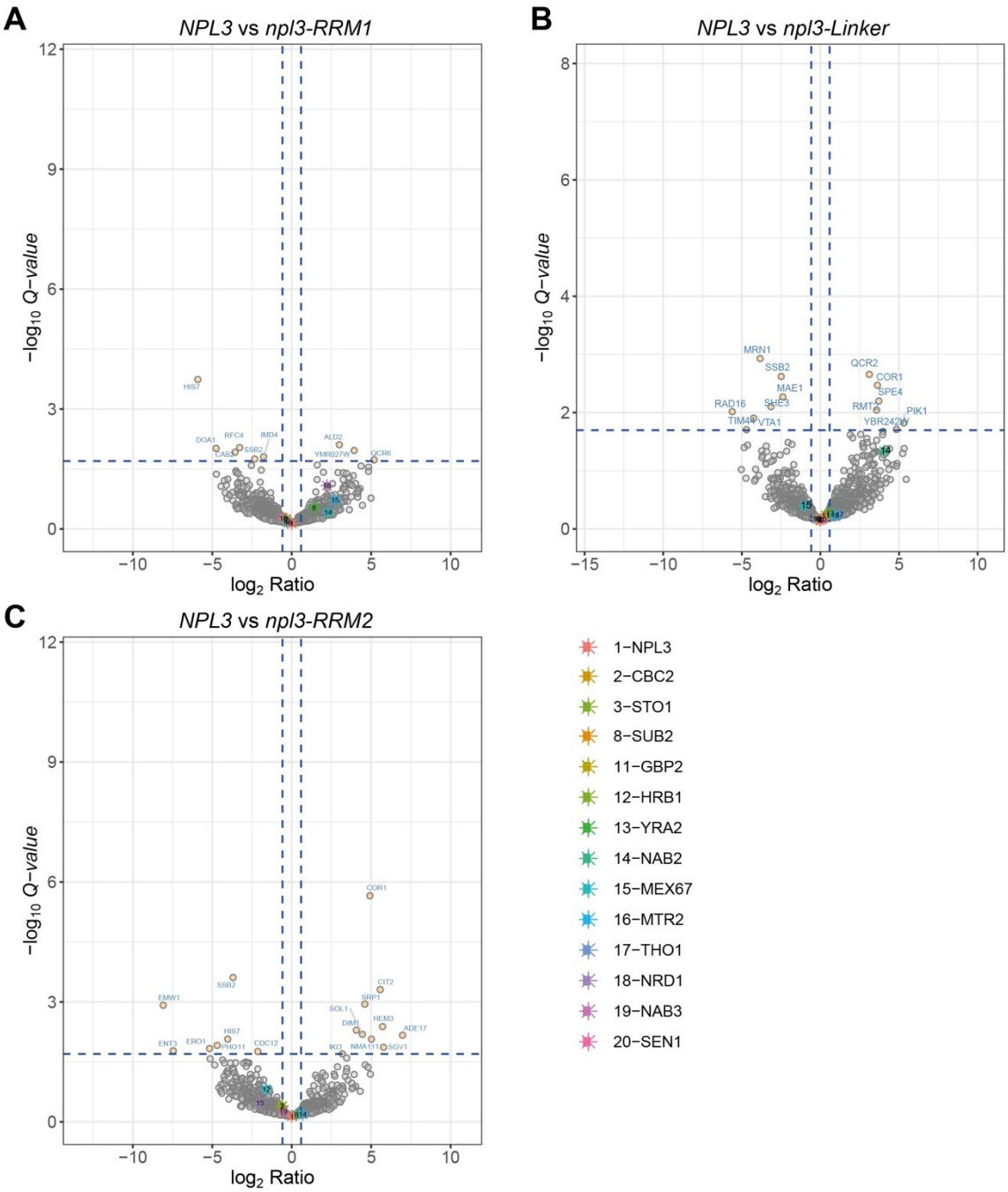



**Figure 23: MS quantification of TAP purified Npl3 wild type vs *npl3* mutants.** TAP-tagged Npl3 was purified until TEV-elution from *NPL3* wild-type, *npl3-RRM1* **(A)**, *npl3-Linker* **(B)** and *npl3-RRM2* **(C)** cells and co-purifiers were quantified by MS of three biological independent replicates. After quantification the levels of co-purified proteins were compared between *NPL3* wild type and the three *npl3* mutants. Each comparison shows only minor differences in the level of co-purified proteins. MS quantification and data analysis performed by Monika Raabe and Ivan Silbern

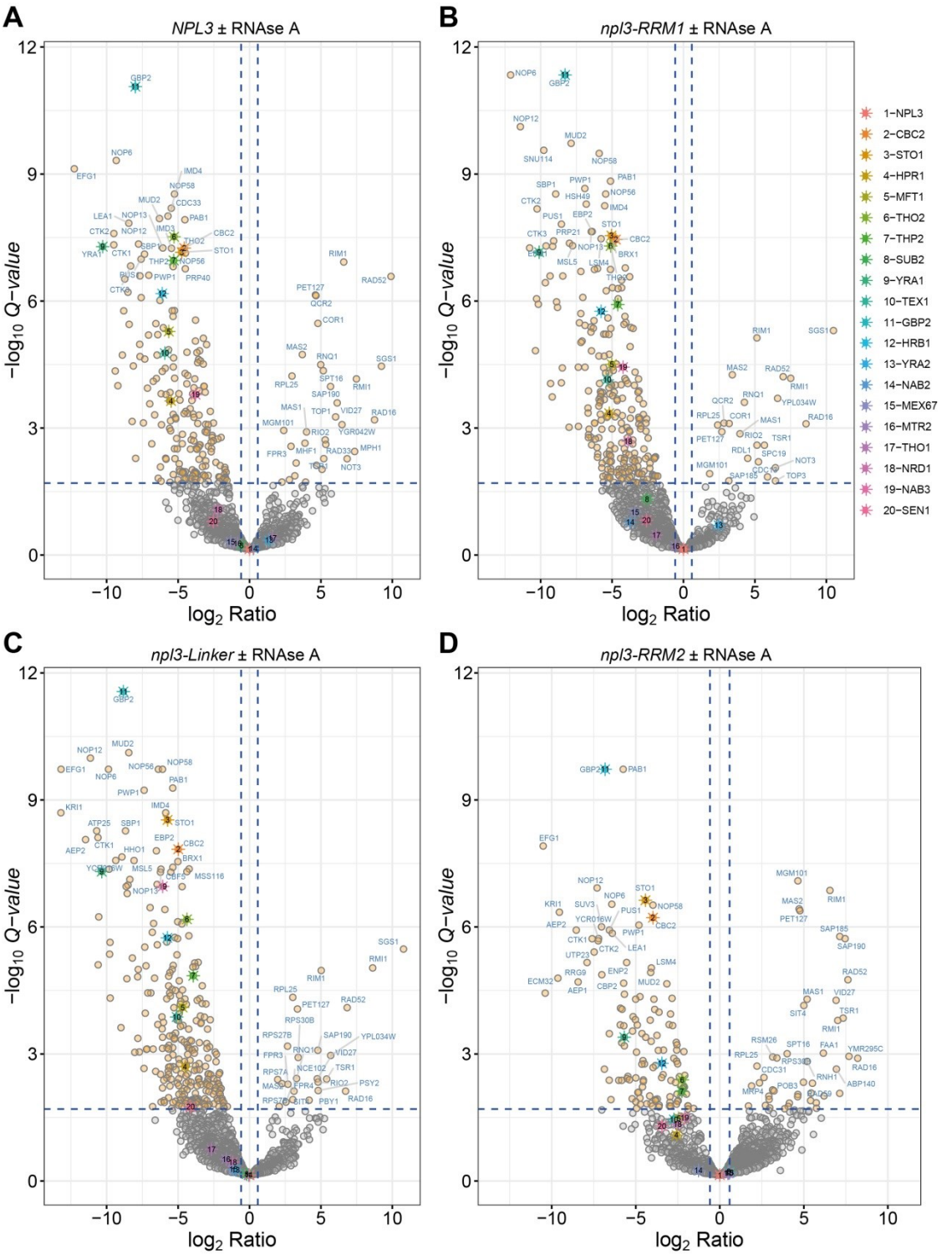



**Figure 24: MS quantification of TAP purified Npl3 wild type and *npl3* mutants, comparison of RNase A treated vs untreated samples.** TAP-tagged Npl3 was purified until TEV-elution from *NPL3* wild-type **(A)**, *npl3-RRM1* **(B)**, *npl3-Linker* **(C)** and *npl3-RRM2* **(D)** cells and co-purifiers were quantified by MS of three biological independent replicates. Samples were treated with or without RNase A before IgG-sepharose was added to the lysates. MS quantification and data analysis performed by Monika Raabe and Ivan Silbern

## *6.12* RNA binding of several mRNP components is reduced in *npl3* mutants

Because we observed changed levels of RBPs within nuclear mRNPs, in particular reduced levels of the TREX complex, Tho1 and Mex67 in the *npl3-Linker* mutant and an increased level of Nab2 in *npl3-RRM2* mutant cells (Figure 16 and 18), the association of these proteins with RNA *in vivo* was analyzed in the respective mutant background by RIP. The TREX components Hpr1, Sub2 and Yra1, Tho1 as well as the mRNA exporter subunit Mex67 associate with three selected transcripts to a lesser degree (*PMA1*, *CCW12* and *YEF3*) in the *npl3-Linker* mutant than in the wild type (Figure 20).

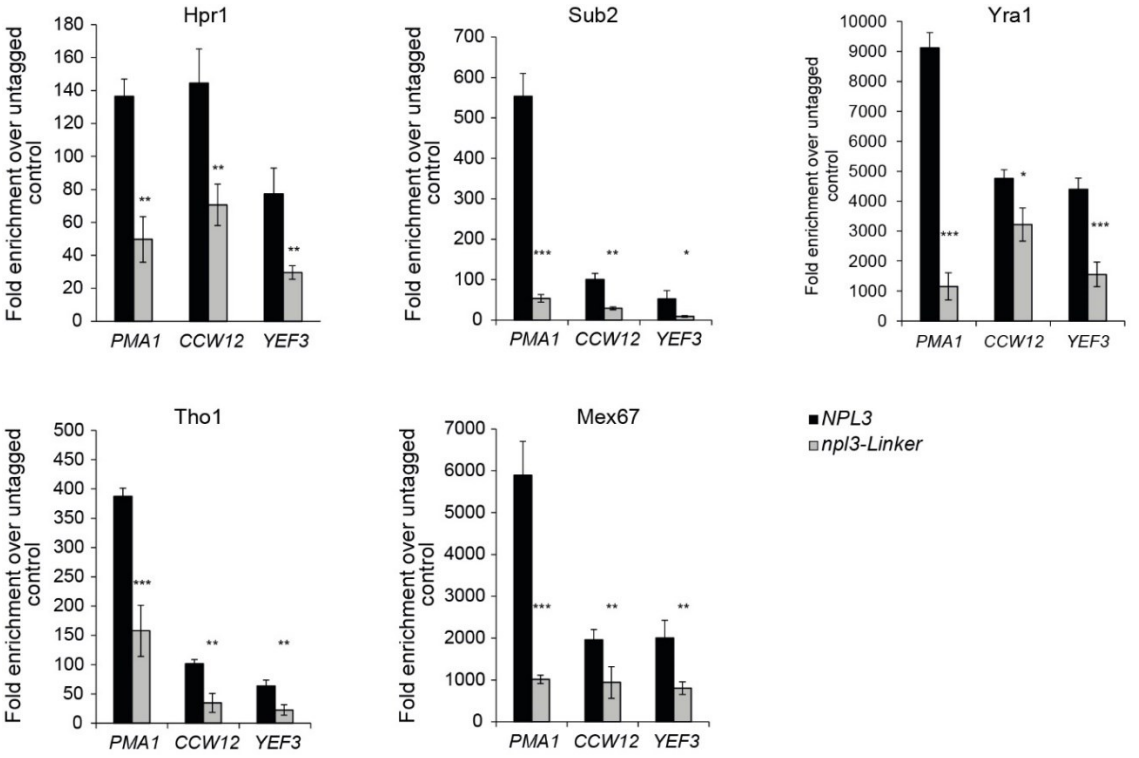



**Figure 25: RNA immunoprecipitation of RBPs that showed a reduction in nuclear mRNPs in *npl3* Linker cells.** All tested mRNP components (Hpr1, Sub2, Yra1, Tho1 and Mex67) bind significantly weaker to mRNA *in vivo* for three analyzed transcripts (*PMA1*, *CCW12*, *YEF3*) shown by RT-qPCR. The level of mRNA was determined as enrichment over the level of a non-tagged control. Bars represent the mean ± standard deviation of three independent biological replicates, p-values of students t-test: * ≤ 0.05; ** ≤ 0.01; *** ≤ 0.001. Schematic of the transcripts with appropriate primer position used for the RT-qPCR is depicted in Figure 9B. Experiment performed by Kristin Hühn.

Nab2, which shows an increased level in the Cbc2-TAP purification in the *npl3-RRM2* mutant, binds more mRNA shown by a RT-qPCR-analyzed RIP (Figure 21). Thus, Npl3 is required to recruit these nuclear mRNP components to the mRNP or is needed to retain them.

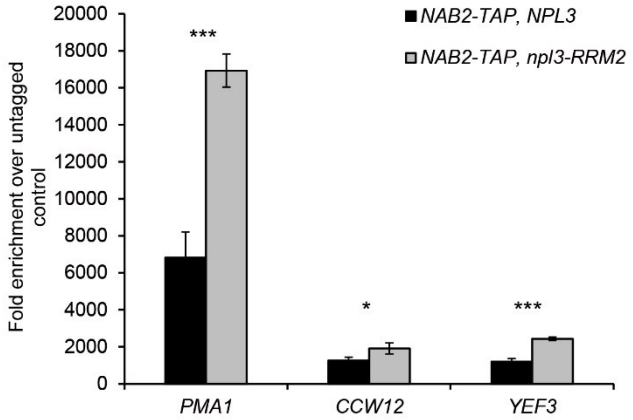



**Figure 26: RNA immunoprecipitation of Nab2-TAP in npl3-RRM2.** Nab2 association with mRNA *in vivo* is increased for three analyzed transcripts (*PMA1*, *CCW12*, *YEF3*) shown by RT-qPCR. The level of mRNA was determined as enrichment over the level of a non-tagged control. Bars represent the mean ± standard deviation of three independent biological replicates, p-values of students t-test: * ≤ 0.05; *** ≤ 0.001. Schematic of the transcripts with appropriate primer position used for the RT-qPCR is depicted in Figure 9B. Experiment performed by Kristin Hühn.

## 6.13 The occupancy of some mRNP components at transcribed genes is decreased in *npl3* mutants

Already during transcription mRNP components are recruited to the nascent mRNA, and, in consequence, to transcribed genes. The changed levels of nuclear mRNP components in the *npl3* mutants might be due to failed recruitment of those components to transcribed genes. To test this hypothesis, the occupancy of mRNP components was analyzed by chromatin immunoprecipitation (ChIP) at three tested genes (*PMA1*, *CCW12* and *YEF3*). First, the occupancy of RNA polymerase II (RNAPII) was determined since the recruitment of RBPs depends on active transcription. A reduced level of RNAPII in the mutant cells compared to wild type would reduce recruitment of mRNP component even if mRNP assembly was otherwise unaltered. The occupancy of RNAPII is increased in *npl3-Linker* cells (Figure 22B), whereas no change in the occupancy of RNAPII was detected in the *npl3-RRM1* and *npl3-RRM2* cells (Figure 23A and 24A). The occupancy of Npl3 is reduced in all three *npl3* mutant cells (Figure 22C and D, 23B and C, 24B and C). In agreement with the deceased levels of Hpr1, Yra1 and Mex67 in the nuclear mRNP, the occupancy of those mRNP components at transcribed genes in *npl3-Linker* cells is reduced (Figure 22E, G, I). Although the levels of Sub2 and Tho1 are reduced in nuclear mRNPs in *npl3-Linker* cells (Figure 16), their occupancies at transcribed genes are unaffected (Figure 22F and H). Moreover, the occupancy of Nab2 at genes in *npl3-RRM2* mutant cells is also unchanged (Figure 24D).

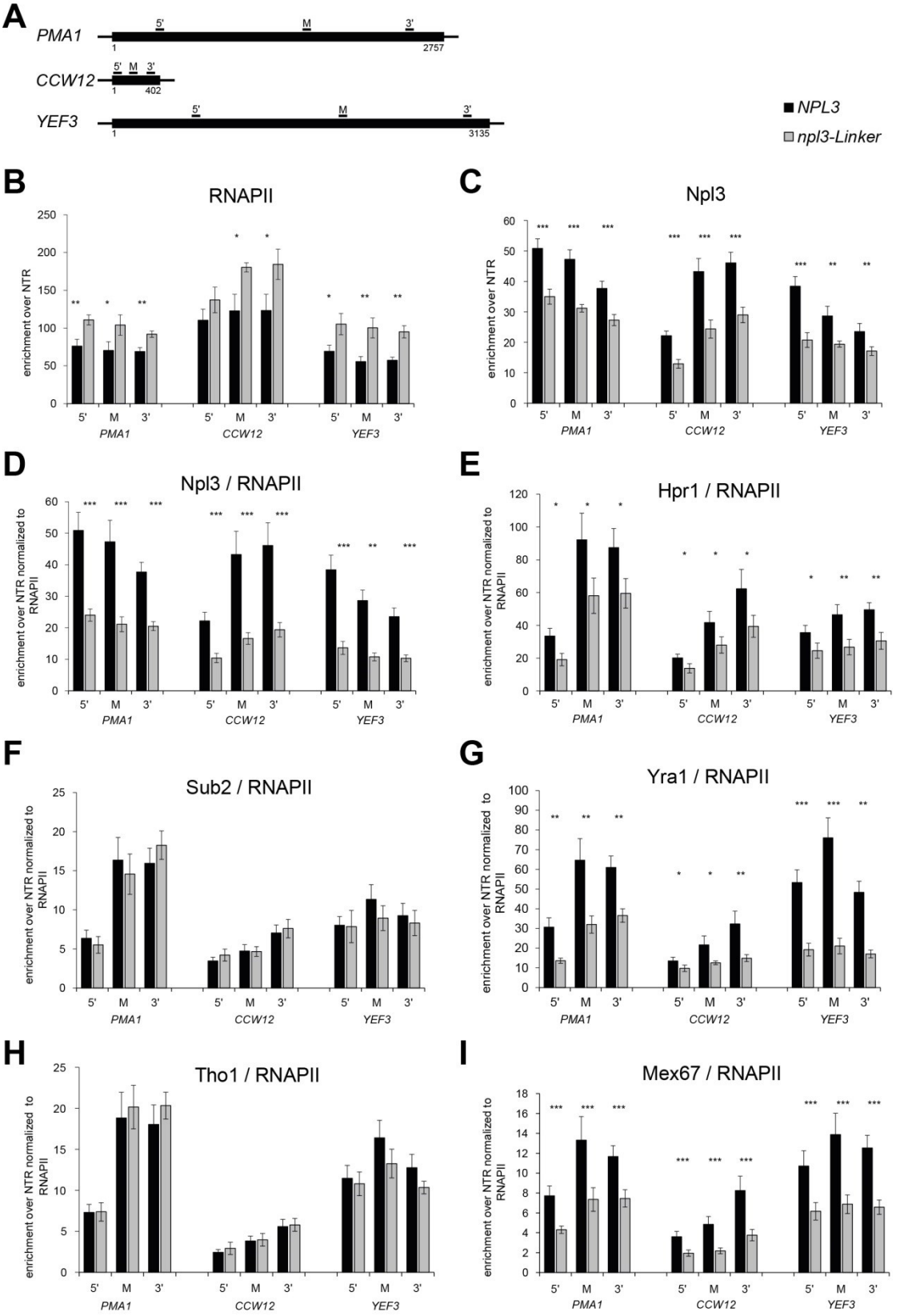



**Figure 27: The occupancy of Npl3, Hpr1, Yra1 and Mex67 at transcribed genes is reduced in *npl3-Linker* cells.** Occupancy of target proteins at transcribed genes was determined by chromatin immunoprecipitation (ChIP). **(A)** Schematic of the three tested genes (*PMA1*, *CCW12* and *YEF3*). Open reading frames (ORFs) are represented by a solid line. Positions of primer pairs used for the qPCR are indicated by bars above the solid lines. 5': 5' end of the ORF, M: middle, 3': 3' end. **(B)** The occupancy of RNAPII is increased in *npl3-Linker* cells. **(C and D)** The occupancy of Npl3 (C) and the occupancy of Npl3 normalized to the occupancy of RNAPII (D) is decreased in *npl3-Linker* cells. **(E)** The occupancy of Hpr1

normalized to RNAPII is decreased in *npl3-Linker* cells. **(F)** The occupancy of Sub2 normalized to RNAPII is not affected in *npl3-Linker* cells. **(G)** The occupancy of Yra1 normalized to RNAPII is decreased in *npl3-Linker* cells. **(H)** The occupancy of Tho1 normalized to RNAPII is unaffected in *npl3-Linker* cells. **(I)** The occupancy of Mex67 normalized to RNAPII is decreased in *npl3-Linker* cells. Bars represent the mean ± standard deviation of three independent biological replicates, p values of students t-test: * ≤ 0.05; ** ≤ 0.01; *** ≤ 0.001.

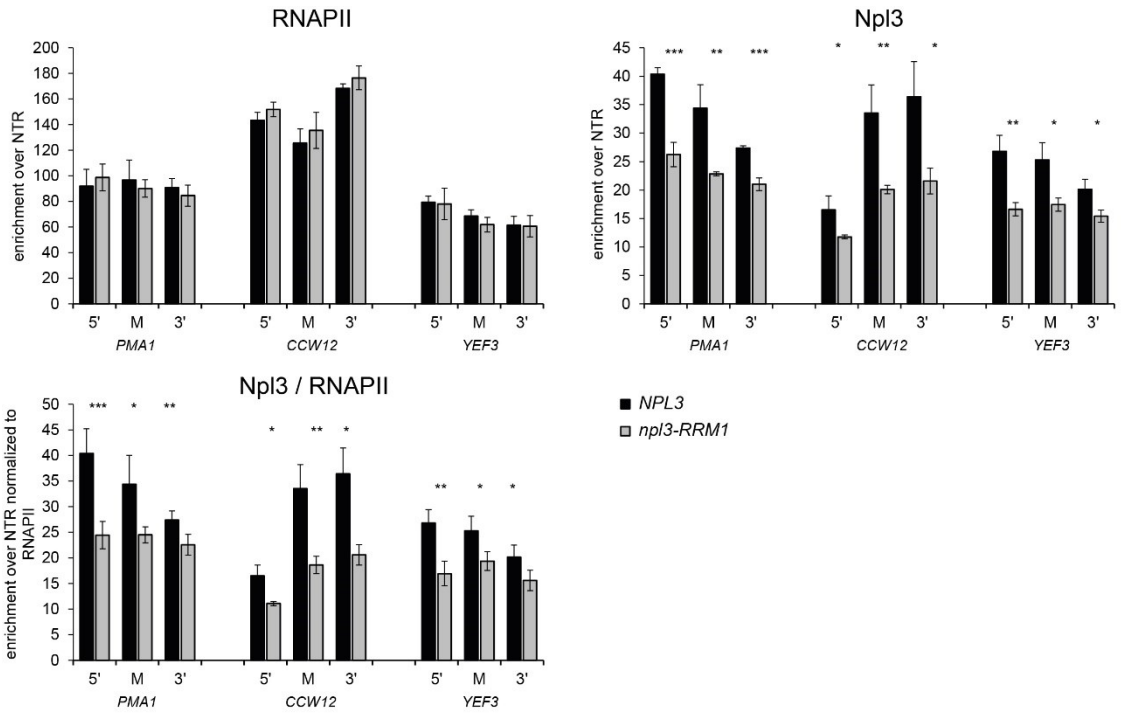



**Figure 28: The occupancy of Npl3 at transcribed genes is reduced in *npl3-RRM1* cells.** The occupancy of target proteins at transcribed genes was determined by chromatin immunoprecipitation (ChIP). The same primer pairs for analyzing the ChIP were used as in 22A. **(A)** The occupancy of RNAPII is unaffected in *npl3-RRM1* cells. **(B and C)** The occupancy of Npl3 (B) and the occupancy of Npl3 normalized to the occupancy of RNAPII (C) is decreased in *npl3-RRM1* cells. Bars represent the mean ± standard deviation of three independent biological replicates, p values of students t-test: * ≤ 0.05; ** ≤ 0.01; *** ≤ 0.001.

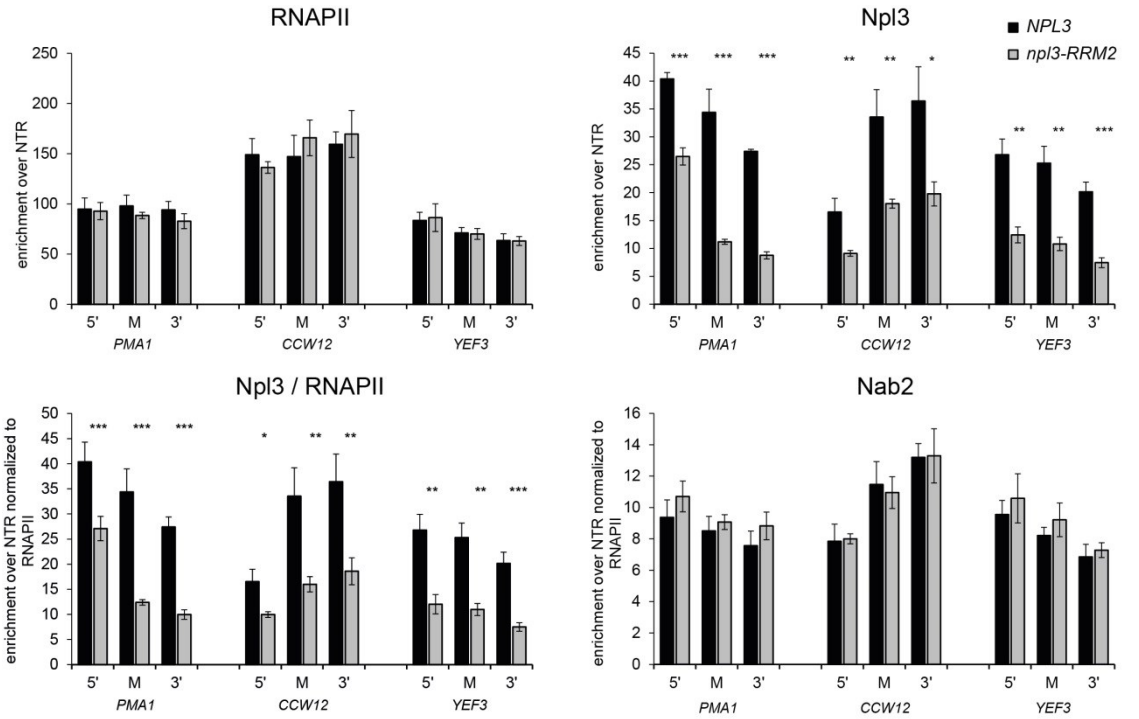



**Figure 29: The occupancy of Npl3 at transcribed genes is reduced in *npl3-RRM2* cells.** The occupancy of target proteins at transcribed genes was determined by chromatin immunoprecipitation (ChIP). The same primer pairs for analyzing the Chip were used as in Figure 27A. **(A)** The occupancy of RNAPII is unaffected in *npl3-RRM2* cells. **(B and C)** The occupancy of Npl3 (B) and the occupancy of Npl3 normalized to the occupancy of RNAPII (C) is decreased in *npl3-RRM2* cells. **(D)** The occupancy of Nab2 normalized to RNAPII is unaffected in *npl3-RRM2* cells. Bars represent the mean ± standard deviation of three independent biological replicates, p values of student's t-test: * ≤ 0.05; ** ≤ 0.01; *** ≤ 0.001.

# 7 Discussion

The aim of the present study was to understand how the RNA-binding activity of various nuclear mRNP-binding components contributes to their function. To this end, we first identified the amino acids involved in RNA contacts for several nuclear mRNP-binding proteins *via* the RNP$^{XL}$ to be able to mutate those putative binding sites. The SR-like protein Npl3 was chosen to analyze the functional relevance of the newly identified RNA binding sites. Npl3 is a protein with pleiotropic functions and, to date, its mode of action is not fully understood at a molecular level. Three binding sites in Npl3 were selected and mutated to investigate how these binding sites contribute to the functions of Npl3 in mRNA assembly, mRNA export and splicing. All three *npl3* mutants showed reduced RNA binding *in vivo* and *in vitro* and displayed overlapping but overall distinct phenotypes. This showed that the different RNA binding sites of Npl3 are important for independent tasks of each domain. Moreover, we were able to show that Npl3 is involved in either the recruitment of mRNP components to the mRNA or their retention.

## 7.1 Identification of RNA-binding sites

RBPs bind nascent mRNAs already co-transcriptionally and accompany the mRNA to regulate RNA biogenesis and function until the RNA is degraded. Therefore, the identification of RBPs as well as their interaction sites is crucial to understanding their dynamic and sophisticated roles along the processes in which they are involved. RBPs are indispensable for regulating post-transcriptional RNA events such as 5' capping, splicing, 3' end processing, mRNA export, localization, stability and translation. Classical RBPs harbor one or multiple canonical RNA-binding domains that bind RNA, like KH domain (Valverde *et al.*, 2008), zinc-finger (Lu *et al.*, 2003), DEAD box helicase domains (Linder and Jankowsky, 2011), RGG domain (Birney *et al.*, 1993) or RRM domain (Cléry *et al.*, 2008). However, in recent years novel RBPs that do not contain any annotated classical RNA-binding domain have been identified *in vivo* by proteome-wide studies (Baltz *et al.*, 2012; Castello *et al.*, 2012; Mitchell *et al.*, 2013). To unravel how these RBPs interact with RNA without harboring any known RNA binding domain (RBD) high-throughput methods were developed to identify the RBDs of hundreds of novel RBPs (Castello *et al.*, 2016; Kramer *et al.*, 2014; Panhale *et al.*, 2019). One of those methods is the RNP$^{XL}$ method (Kramer *et al.*, 2014) that has been used in our study to identify the RNA-binding sites of RBPs involved in mRNP formation.

In general, the identification of RBPs and their RNA-binding sites *in vivo* by RNA interactome capture methods is based on UV-crosslinking to covalently connect proteins

to RNA followed by an enrichment step of the RNA-protein crosslinks followed by identification of the RNA-binding sites by mass spectrometry. To enrich or purify the RNA-protein crosslinks different variants were developed. Utilizing oligo(dT)-coupled beads to pull down poly(A)-tail containing RNAs is one of the most common methods. To improve the purification of RNA-protein complexes and since RBPs do not only bind mRNAs, a variety of methods like phenol-toluol extraction (PTex), orthogonal organic phase separation (OOPS), identification of RNA-associated peptides (iRAP), complex capture (2C) and total RNA-associated protein purification (TRAPP) were developed in recent years to capture all possible RNA-protein crosslinks (Asencio *et al.*, 2018; Peil *et al.*, 2018; Shchepachev *et al.*, 2019; Smith *et al.*, 2020; Urdaneta *et al.*, 2019). In this study, a more targeted approach of $RNP^{XL}$ was used. The RNA-binding proteins known to be involved in mRNP formation were genomically TAP-tagged and purified until TEV-protease digestion after *in vivo* crosslinking of 4-thiouracil-labeled RNA. Because of the purification step, the UV-induced crosslinks of the target proteins and co-purifiers were much more enriched compared to proteome-wide approaches. Even though the former studies were successful and jointly identified more than 1,000 RNA-binding sites in a huge number of RBPs, our targeted approach revealed many additional RNA-interacting residues in proteins involved in mRNP formation. For example, 24 unique crosslinks were identified in Sub2, compared to only very few that were identified in previous studies. The vast majority of these contacts had not been observed before (Kramer *et al.*, 2014; Peil *et al.*, 2018), underlining the increased sensitivity of our targeted approach. Most crosslinks occur within known RBDs, validating our approach. Taken together, various new binding sites in 16 established mRNP components were identified. Analysis of these binding sites on molecular level will afford new insights into how mRNP assembly takes place and how the different mRNP components regulate different processes.

## 7.2 Growth defect of *npl3* mutants

The identified crosslinks for Npl3 mostly occur within the RRM domains and the RGG domain; this was expected since both are known to have RNA-binding ability (Birney *et al.*, 1993; Cléry *et al.*, 2008; Deka *et al.*, 2008). The RRM1 of Npl3 is a classical RNA recognition motif with conserved RNP1 and RNP2 motifs whereas the RRM2 is a cryptic RRM without conserved RNP motifs (Deka *et al.*, 2008). Unexpectedly, a crosslink was also found in the linker region that connects both RRMs.

In the dot spot assay, a Δ*npl3* strain was spotted as control. Each of the new *npl3* mutants (*npl3-RRM1*, *npl3-Linker* and *npl3-RRM2*) showed a less severe growth defect than the

Δ*npl3* strain (Figure 8). This indicates that the new *npl3* mutants are partial loss-of-function mutants. Furthermore, combination of the mutations either exacerbated the growth defect or resulted in cell death (Figure 8). Thus, the combination of the different *npl3* mutations has an additive effect. The effect might be due to a stronger RNA-binding defect in the double mutants compared to the single mutants or that several functions get abrogated at once. Interestingly, combination of the *npl3-RRM2* mutation with any other mutation resulted in cell death. In contrast, the Δ*npl3* strain is viable at 25°C and 30°C, although growth is poor (Figure 8). That the *npl3* double mutants induce a stronger growth defect than a Δ*npl3* strain, suggests that the *npl3* mutants inhibit by their presence the function of other RBPs or might block binding of other proteins. Although any combination with *npl3-RRM2* results in cell death, an overexpression of *npl3-RRM1* and *npl3-Linker* are lethal at 37°C showing a dominant negative effect, while an overexpression of *npl3-RRM2* has no effect on growth at all (Figure 10). That the *npl3-RRM1* and the *npl3-Linker* overexpression induces cell death at 37°C indicates once more that both mutated proteins act in a disruptive mode in some RNA-related process induced by their presence under heat stress condition. The observed growth defects of the *npl3* mutants are not caused by mutation-induced lower Npl3 expression, as the three mutant *npl3* proteins are expressed at wild-type level (Figure 11).

## 7.3 RNA-binding ability

To prove that the putative RNA-binding sites identified by crosslinking MS are truly involved in RNA binding, RNA immunoprecipitation (RIP) and NMR analysis were performed to assess RNA binding of the *npl3* mutants *in vivo* and *in vitro*. The results of the *in vitro* RNA binding assays coincide very well with the *in vivo* results. Association of all three *npl3* mutants with mRNAs *in vivo* was significantly decreased for three representative transcripts (*PMA1*, *CCW12*, *YEF3*) if compared to Npl3 wild type (Figure 12). Moreover, *in vitro* assays with the purified RNA-binding domain construct ($Npl3^{120-280}$) showed that the mutations strongly reduced binding to both tested RNA oligos (Figure 12, Figure 14 and Figure 15).

Surprisingly, the strongest reduction in RNA binding was observed for the *npl3-Linker* mutant, both *in vivo* and *in vitro*. The linker region that connects both RRM domains strongly contributes to RNA binding, presumably by mediating important electrostatic contacts. In the *npl3-Linker* mutant the P196 and the A197 are both substituted by an aspartic acid. Thus, two negatively charged side chains are introduced into the linker region without affecting the structural integrity of the protein (Figure 13C and D). Taking the structural model with electrostatic surface potential into account (Figure 13B), we

suggest that these negative charges are responsible for the reduction of the overall RNA binding capacity of Npl3. This is supported by dot spot assays where growth defects only resulted when the P196 of Npl3 is mutated into an amino acid that has a negatively charged side chain, but not a neutrally charged one (Supplementary figure 1). However, the growth defect of the P196D substitution is not rescuable by a substitution of A197 with an amino acid with a positively charged side chain. Moreover, the growth defect of the single P196D is enhanced when the A197 is substituted by an aspartic acid as well, even though the A197D alone shows no growth phenotype (Figure 7, Supplementary figure 1).

Although the *npl3-Linker* mutation has the strongest effect, the RNA-binding ability is strongly reduced in the *npl3-RRM1* and the *npl3-RRM2* mutant as well. In the *npl3-RRM1* mutant, F162 is substituted by a tyrosine. The RRM1 domain is a canonical RRM with conserved RNP1 and RNP2 motifs (Figure 30). Since the F162 is one of the highly conserved aromatic amino acids involved in RNA binding, a reduction in the RNA-binding ability was expected in this mutant. The RRM2 domain of Npl3 is a cryptic RRM without conserved RNP motifs. Only the F245 which was substituted by an isoleucine is conserved (Figure 6). However, the RRM2 domain is unfolded in the *npl3-RRM2* mutant as shown by NMR (Figure 13).

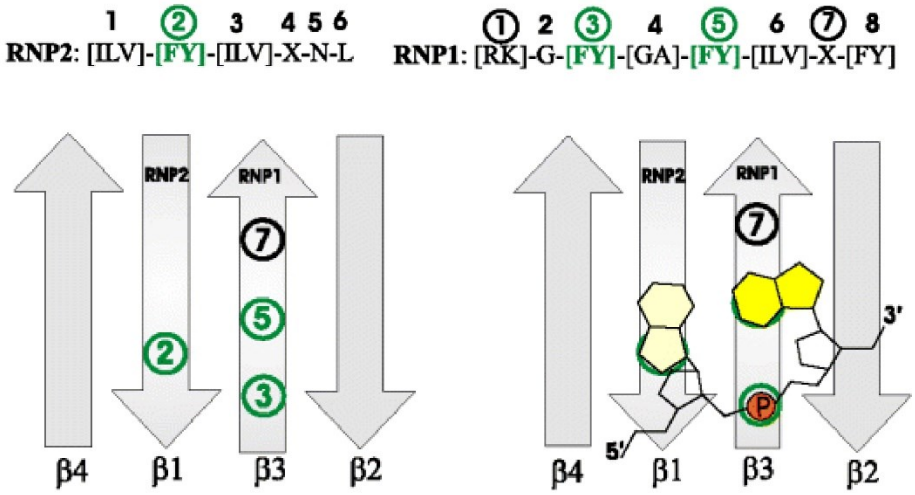



**Figure 30: Conserved RNP1 and RNP2 motif within RRM domains.** The aromatic residues indicated by circled numbers according to each conserved RNP sequence motif. The conserved aromatic residues are depicted by green circles. Modified after (Maris *et al.*, 2005).

RRM1 is the main interaction site for the UAA-containing RNA, as shown for wild-type Npl3; in contrast, the UGG-containing RNA oligo induces additional chemical shift perturbations in the RRM2 domain and is bound with much higher affinity overall, suggesting that the RRM2 domain contributes to its binding (Figure 14). However, binding to the UGG-containing RNA was more strongly affected in the *npl3-RRM1* mutant

than in the *npl3-RRM2* mutant, although the RRM2 domain is fully unfolded in the *npl3-RRM2* mutant under the assay conditions (Figure 13 and Figure 15), suggesting that the majority of Npl3's RNA-binding activity is provided by RRM1.

The higher preference of the Npl3 RNA-binding domains for UGG-containing RNAs was observed before (Bucheli *et al.*, 2007; Deka *et al.*, 2008). However, full-length Npl3 binds RNA sequence-independently with even higher affinity than the isolated RRM domains (Deka *et al.*, 2008; Holmes *et al.*, 2015). Nevertheless, a preference of Npl3 to bind UGG- and UAA-rich sequence motifs supports the model that it functions in 3′-end formation by competing with Rna15 (Bucheli *et al.*, 2007). Apart from confirming that our targeted mutations indeed disrupt the ability of Npl3 to bind RNA, these results show that RRM1 most substantially contributes to Npl3 RNA-binding activity and confirms a preference for UGG containing RNA.

## 7.4 Global expression level of transcripts is changed in *npl3* mutants

The expression levels of all mRNAs were analyzed by RNA-seq in the *npl3* mutants. Each of the *npl3* mutants showed a change in expression for more than 3,500 transcripts at 30°C compared to Npl3 wild-type cells (Figure 16A and C). At 37°C less mRNAs change in their expression level in each mutant but there is more variability between the mutants (Figure 16A and C). Expression levels might be affected both by differences in transcription and changes in transcript stability. However, when half-lives of six exemplary transcripts were analyzed by thiolutin treatment to block transcription and subsequent RT-qPCR analysis (Figure 17), the stability of all six transcripts was not affected by *npl3* mutation. The expression level of these transcripts was changed in the *npl3* mutants shown by the RNA-seq data, except for *YEF3*. Thus, we think it is likely that Npl3 affects RNA expression through changes in transcription rather than altered stability of mRNAs. For Δ*npl3* cells, gene expression was shown to be altered due to an increased occurrence of transcriptional read-through (Holmes *et al.*, 2015).

## 7.5 Functional analysis of the *npl3* mutants

In *S. cerevisiae*, Npl3 is the only SR-like protein that is involved in splicing. It has a role in early spliceosome recruitment to the pre-mRNA by interacting with U1 and U2 snRNP components (Kress *et al.*, 2008). Therefore, it is expected that splicing is affected in *npl3* mutants that bind mRNA only weakly. Indeed, gene expression analysis revealed that specifically the levels of intron-containing transcripts are down-regulated in the *npl3* mutants, which already suggests that splicing is affected (Figure 16B); a splicing defect in *npl3* mutants could be confirmed by splice junction analysis, and is particularly

pronounced at 37°C (Figure 16D). Strikingly, the different *npl3* mutants show distinct defects in splicing. This suggests that the different RNA-binding regions contribute differently to the recruitment of Npl3 to specific RNA targets; alternatively, the mutations may affect protein-protein interactions in addition to RNA binding. Why the *npl3-Linker* mutant shows a slight increase in splicing efficiency although RNA binding is strongly reduced remains enigmatic.

Since Npl3 is a multi-functional RBP involved in multiple aspects of post-transcriptional gene expression (see 3.2), the functional impact on RNA metabolism of the three *npl3* mutants was investigated. Nuclear mRNA export was determined by fluorescence *in situ* hybridization using an oligo$(dT)_{50}$ probe coupled to Cy3. Compared to the other two *npl3* mutants, the *npl3-Linker* mutant shows by far the strongest mRNA export defect at 30°C, which is even enhanced at 37°C. Surprisingly, the *npl3-RRM2* mutant shows a mild nuclear export defect at 37°C while the *npl3-RRM1* mutant shows almost no export defect (Figure 18) although all three *npl3* mutants show decreased association with mRNA *in vivo* (Figure 12). Interestingly, the intensity of the mRNA export defect of the *npl3-mutants* does not correlate well with the growth defects (Figure 8) indicating that processes other than mRNA export are disturbed in these mutants. Npl3 is known to serve as an adaptor for the mRNA export receptor Mex67-Mtr2 to mediate nuclear mRNA export (Gilbert and Guthrie, 2004; Iglesias and Stutz, 2008; Santos-Rosa *et al.*, 1998). Besides Npl3, other adaptor proteins like Yra1 or Nab2 can recruit Mex67-Mtr2 to the mRNA in a redundant fashion (Iglesias *et al.*, 2010; Strässer and Hurt, 2000; Zenklusen *et al.*, 2001). We analyzed mRNP composition in the *npl3* mutants by pulling on the cap-binding complex and were able to show by quantitative Western blot and MS that several components were significantly depleted form nuclear mRNPs in the *npl3-Linker* cells. The change in the composition of nuclear mRNPs in *npl3* mutant cells is consistent with the observed mRNA export defects. Among others, Yra1 and Mex67 showed a significant reduction in the *npl3-Linker* mutant (Figure 20). Even though Nab2 was unaffected in *npl3-Linker* cells, the strong export defect can easily be explained by the reduction of Yra1 and Npl3; since both are known to be important Mex67-Mtr2 adaptors, their absence likely explains the reduction in the level of Mex67. The *npl3-RRM1* cells do not show an mRNA export defect and besides Npl3, only Hpr1 was depleted from nuclear mRNPs in these cells (Figure 18 and Figure 20). Hpr1 seems to be involved in the export of only a very limited pool of mRNAs under certain stress conditions (Bretes *et al.*, 2014); hence, its depletion from mRNPs in *npl3-RRM1* cells is not expected to induce a general export defect. A reduced level of Npl3 seems to be compatible with proper mRNA export as long as Yra1 and Nab2 are present, confirming that the adaptors

work in a redundant manner. The *npl3-RRM2* mutant shows a mild mRNA export defect. We were able to show that the mutation of *npl3-RRM2* results in destabilization of the RRM2 domain. At the same time, the levels of Npl3 and Nab2 are increased at the nuclear mRNPs. The increased level of misfolded Npl3 is probably induced by dimerization or oligomerization of mutated Npl3. However, since the mRNP components are recruited normally in *npl3-RRM2*, except for Nab2, the export defect is limited.

## 7.6 mRNP assembly in *npl3* mutants

As mentioned above, nuclear mRNP composition was examined by TAP-purification of Cbc2 (small subunit of cap binding complex, CBC) and subsequent analysis by quantitative Western Blot and MS (Figure 20). This revealed that several RBPs are significantly changed in their level at the nuclear mRNPs in the *npl3* mutant cells. Interestingly, Hpr1, Sub2 and Yra1 were significantly reduced in the *npl3-Linker* cells shown by Western blot (Figure 20). This suggested that recruitment of the entire TREX complex to mRNPs is reduced. Indeed, quantitative MS analysis revealed that all TREX components (Tho2, Hpr1, Thp2, Hrb1, Mft1, Tex1, Gbp2, Yra1) except Sub2 are significantly depleted from nuclear mRNPs in the *npl3-Linker* mutant. This result indicated that Npl3 is involved in recruitment of the TREX complex or its retention at the mRNA. This was very surprising and is a novel function of Npl3. Even though both the TREX complex and Npl3 are recruited to the mRNA early during transcription, no association between the TREX complex and Npl3 has been reported. The distribution of Npl3 along transcripts is highest at the 5' end of mRNAs shown by PAR-CLIP and CRAC analysis (Baejen *et al.*, 2014; Holmes *et al.*, 2015). This would support the model that Npl3 binds very early to nascent mRNAs, enabling it to recruit other components. However, purification of TAP-tagged Npl3 indicated that Npl3 is not associated with TREX (Hurt *et al.*, 2004). This suggests that the effect of Npl3 on the recruitment or retention of TREX at the mRNA is not mediated by a direct interaction. This is also supported by our purification of TAP-tagged Npl3 with subsequent MS quantification, which confirmed that the interaction with most TREX components was mediated by RNA (Figure 24A).

Besides TREX, several other proteins are depleted from nuclear mRNPs in *npl3-Linker* cells. Tho1 is significantly reduced in the Western blot and also reduced, but not significantly, in MS quantification. Tho1 is a protein identified as a suppressor of the phenotypes of Δ*hpr1* (THO complex; (Jimeno *et al.*, 2002; Piruat and Aguilera, 1998)) and thereby assumed to be functionally connected to the THO/TREX complex. Later on, overexpression of Tho1 was shown to suppress phenotypes of Δ*npl3* as well (Santos-

Pereira *et al.*, 2013). Moreover, Sro9 and its paralog Slf1 are strongly reduced in the Cbc2 purification quantified by MS in *npl3-Linker* cells. Like Npl3, Sro9 is recruited to actively transcribed genes, is exported in an mRNA-dependent manner and is involved in the regulation of translation in the cytoplasm. Moreover, it has been co-purified in a Npl3 pulldown (Röther *et al.*, 2010). Furthermore, all three components of the NNS-complex (Nrd1, Nab3 and Sen1) are depleted from the mRNP purification quantified by MS, even though the effect was statistically significant only for Sen1. The NNS-complex is involved in 3' end processing of ncRNAs such as snRNAs, snoRNAs and cryptic unstable transcripts (CUTs) and some mRNAs (Conrad *et al.*, 2000; Steinmetz *et al.*, 2001). Npl3 was shown to be involved in surveillance and/or transcription termination of both mRNAs and ncRNAs like the NNS-complex (Holmes *et al.*, 2015). In the absence of Npl3, about 30 % of all protein-coding genes and diverse ncRNAs show transcription termination defects and increased read-through. These termination defects cause broad changes in gene expression through inappropriate termination (Holmes *et al.*, 2015). In the MS-quantified Npl3 purification, only Sen1 was unaffected in presence of RNase A (Figure 24), indicating that Npl3 directly associates with Sen1.

In *npl3-RRM2* cells, Nab2 and Npl3 are both enriched at the nuclear mRNP shown by Western blot (Figure 20). Despite strong differences in structure and domain architecture (Deka *et al.*, 2008; Grant *et al.*, 2008; Marfatia *et al.*, 2003), both proteins have partially redundant functions. Both bind to poly(A)-containing RNAs and function as Mex67-Mtr2 adaptor proteins. However, Nab2 preferentially binds directly to poly(A) tails, while Npl3 binds rather with no sequence specificity, although the distribution of Npl3 is highest at the 5' and 3' ends of transcripts. Moreover, Nab2 and Npl3 are both shuttling proteins that are exported to the cytoplasm together with the mRNA and are imported back after releasing the mRNA (Baejen *et al.*, 2014; Gilbert and Guthrie, 2004; Green *et al.*, 2002; Kelly *et al.*, 2010; Lee *et al.*, 1996). *npl3-RRM2* showed overall only minor significant changes in MS analysis but overall a strong tendency to an upregulation of several proteins (Figure 22C). In *npl3-RRM1* cells only Npl3 itself and Hpr1 were depleted from nuclear mRNPs as shown by Western Blot (Figure 20). Overall, considerably more proteins showed a significant reduction in the *npl3-RRM1* mutant compared to the *npl3-Linker* mutant in the mRNP purification. However, proteins that are known to function in mRNP assembly that showed a reduction in the *npl3-Linker* mutant were not changed in *npl3-RRM1*. This indicates once more, that the different domains of Npl3 have different functions. Surprisingly, Mtr2 showed a significant reduction in *npl3-RRM1* although Mex67, with which it forms a heterodimer, is not affected (Figure 22A). Because *npl3-RRM1* cells also do not show any pronounced mRNA export defect, we regard the

reduction of Mtr2 to be very unlikely and consider it an artifact of value imputation (see below). Previous studies revealed that some phenotypes of Δ*npl3* cells can be rescued by overexpression of several RBPs, including the DEAD box helicase Sub2, the poly(A)-binding protein Nab2 and Tho1. The hyperrecombination phenotype of Δ*npl3* cells is completely rescued by overexpression of Sub2. Moreover, overexpression of Npl3 is able to suppress the hyperrecombination phenotype of Δ*hpr1,* indicating a connection between TREX and Npl3 at early steps of mRNP biogenesis (Santos-Pereira *et al.*, 2014) and references within). Whether the changed levels of certain proteins in mRNPs shown by quantitative MS and Western blot are directly induced by the lack of Npl3 in mRNPs or secondary effects has to be investigated.

Some proteins show a significant change in Western Blot analysis but not in the MS quantification. However, both assays display changes in the same direction. The discrepancies between Western Blot and MS quantification are probably due to the fact that for MS quantification three independent biological replicates were used, while more than three were analyzed by Western Blot, which improves the statistics for the latter. Furthermore, a low level of unique peptides in those proteins might be a reason as well. Wherever MS signals were below the detection limit, background values were imputed during data analysis. Moreover, only unique peptides are used for quantification. All identified proteins in MS were plotted against their abundance. Tho1 and Nab2 are two of the candidates that show significant changes in Western blot but not in MS, and are among the co-purified proteins with the lowest abundance (Supplementary figure 2). Although, some candidates were not significant in MS, the effects of Western blot and MS always showed the same trend.

Taking the TAP-tagged Npl3 purifications into account, which were analyzed by MS as well (Figure 23 and Figure 24), we conclude that the changed levels of co-purified proteins in the CBC-purifications are not due to disturbed protein-protein interactions of the mutant *npl3* protein. The purified *npl3* mutants revealed only very few changes compared to Npl3 wild type (Figure 23). In particular, this is also true for all candidates that showed a changed level in the nuclear mRNP purification. Furthermore, each Npl3 purification was compared with and without RNase A treatment. RNase treatment eliminated the same co-purifying proteins from Npl3 wild type, *npl3-RRM1* and *npl3-Linker* purifications. The DEAD box helicase Sub2, part of the TREX complex, showed a significant reduction in all Western Blots of TAP-purified *npl3* mutants and a significant reduction in the Cbc2-TAP purification in *npl3-Linker* cells. However, Sub2 did not show a significant change in any of the MS-quantified samples. Due to the reduced level of Sub2 in these purifications, the interaction of Npl3 and Sub2 was investigated *in*

*vitro* by GST-pulldown and NMR-titration (data not shown). With this approach, no interaction could be detected. Since all other TREX components are depleted from nuclear mRNPs in *npl3-Linker* cells, we still consider it very likely that Sub2 as part of the TREX complex is at least partially reduced as well. To mention it, Sub2 is much more abundant than every other TREX component and acts as well in a TREX independent way.

The quantitative MS analysis of the Cbc2-TAP purification revealed several unexpected proteins to be enriched or depleted from nuclear mRNPs in the *npl3* mutants. For example, several mitochondrial related proteins (Ymr31, Cpr3, …), numerous ribosomal proteins (Rsm23, Srp40, …), and the inosine monophosphate dehydrogenase components (Imd2, Imd3, Imd4), which catalyse the rate-limiting step in the synthesis of GTP, were reduced. Npl3 (named Mts1 at that time) was shown to be required for correct mitochondrial protein targeting, which might explain the reduction in mitochondrial proteins detected in MS analysis (Ellis and Reid, 1993; Gratzer *et al.*, 2000). More likely, the changed level of mitochondrial proteins is an indirect effect caused by disturbed expression of a mitochondrial targeting factor. The reason that several ribosomal proteins were significantly reduced in the *npl3* mutants could be explained due to the fact that Npl3 stays partially at the mRNA after nuclear export and associates with ribosomes to influence translation (Baierlein *et al.*, 2013; Estrella *et al.*, 2009; Rajyaguru and Parker, 2012; Windgassen *et al.*, 2004). The reduced RNA binding in the *npl3* mutants might influence the association of the ribosomes to the mRNA. Even though the CBC leaves the nucleus together with the mRNA (Shen *et al.*, 2000; Visa *et al.*, 1996), it is thought that the CBC is immediately replaced in the cytoplasm by the eIF4F translation initiation complex and imported back to the nucleus (Dias *et al.*, 2009). Therefore, it is expected to find mainly nuclear RBPs in a Cbc2-TAP purification and cytoplasmic proteins should be rather rare. Nevertheless, it is possible that at least a minor pool of CBC stays in the cytoplasm long enough for cytoplasmic proteins to bind the exported mRNP or CBC directly. The identification and quantification of co-purified proteins by MS is completely unbiased. Although proteins like the Imd-proteins and others are unexpected in nuclear mRNP purification, it was previously shown that these proteins bind to mRNA identified in similar MS-quantified pulldowns and that abundant proteins are common contaminants in similar MS experiments (Ho *et al.*, 2002; Mitchell *et al.*, 2013).

## 7.7 Recruitment of RBPs to mRNPs and transcribed genes is disturbed in *npl3* mutants

Diverse mRNP components like TREX (Hpr1, Sub2, Yra1), Tho1 and Mex67 are depleted from nuclear mRNPs in the *npl3-Linker* cells as shown by Western blot and MS (Figure 20 and Figure 22). The reduced levels of these components in nuclear mRNPs in the *npl3-Linker* mutant indicate that these proteins bind less mRNA *in vivo*. The performed RNA immunoprecipitation experiments confirmed that these mRNP components associate with RNA less in the *npl3-Linker* mutant (Figure 25), whereas Nab2 showed an increased level in mRNPs in the *npl3-RRM2* mutant and binds more mRNA in the *npl3-RRM2* mutant (Figure 26). Consequently, we suggest that the correct function of Npl3 is required for the recruitment of these nuclear mRNP components to the mRNP or for their retention within the mRNP.

Since RBPs that are involved in mRNP assembly are recruited to the mRNA co-transcriptionally, consequently to actively transcribed genes, the reduced level of some nuclear mRNP components might be due to an impaired recruitment to the transcribed genes. This recruitment depends on the amount of RNA, thus on the level of transcription. Therefore, the occupancy of RNAPII was determined. Unexpectedly, the level of RNAPII is increased at actively transcribed genes in the *npl3-Linker* mutant (Figure 27B), whereas the level of Npl3 is decreased (Figure 27C and D). The reason for this could be that the cell tries to compensate for the global reduction in transcript levels that is observed in the *npl3-Linker* mutant. However, the two other *npl3* mutants (*npl3-RRM1* and *npl3-RRM2*) reduce the expression levels of transcripts to a similar level. Especially at 30°C, the majority of changed transcript levels is overlapping in all three mutants (Figure 16A), RNAPII is at wild-type level in the *npl3-RRM1* and *npl3-RRM2* mutants (Figure 28A and Figure 29A). Therefore, this explanation is rather unlikely. However, Npl3 was shown to form a complex with RNAPII in an RNA-independent manner (Lei *et al.*, 2001). Since the recruitment of Npl3 to transcribed genes is at least partially RNA dependent (Meinel *et al.*, 2013), the reduced level of Npl3 at transcribed genes in all *npl3* mutants is probably a direct consequence of their reduced RNA-binding ability (Figure 12). A previous study suggested that the CTD of RNAPII is recruiting Npl3 during transcription (Santos-Pereira *et al.*, 2014). Therefore, the opposite explanation may be equally likely, that by increasing the level of RNAPII at transcribed genes the cell is trying to compensate the reduced level of functional Npl3 in the *npl3-Linker* cells.

The occupancies of Hpr1, Yra1 and Mex67 at transcribed genes are significantly reduced in *npl3-Linker* cells, mirroring their reduction in nuclear mRNPs (Figure 27 and Figure 20). The reduced occupancy of Hpr1, but not Sub2, indicates that co-transcriptional recruitment of the THO complex depends on Npl3. The occupancy of Yra1 and Mex67 seems to be dependent on Npl3 as well. Yra1, like Npl3, acts as an adaptor protein for the mRNA export adaptor Mex67-Mtr2 and thereby promotes nuclear export of mRNAs (Strässer and Hurt, 2000). Moreover, Yra1 is thought to act in 3' end processing, thereby aiding the generation of an export-competent mRNP; this is mediated by direct interaction with Pcf11, a CPF complex subunit (Johnson *et al.*, 2009). A reduced level of Yra1 in nuclear mRNPs reflects the strong nuclear mRNA export defect in *npl3-Linker* cells. Whether the *npl3-Linker* cells show 3' end processing defects as well has to be investigated. That Npl3 is involved in 3' processing by competing with 3' end processing factors like Rna15 was shown before (Bucheli and Buratowski, 2005; Bucheli *et al.*, 2007). The occupancies of Sub2 and Tho1 at transcribed genes are not reduced in *npl3-Linker* cells (Figure 27). This suggests that the decrease of nuclear mRNP components has different reasons. Hpr1, Yra1 and Mex67 are not recruited in a proper way to actively transcribed genes, whereas Sub2 and Tho1 are not correctly relocated from site of transcription to the forming mRNP. In both cases, fully functional Npl3 is required.

## 7.8 Conclusions

The functions of the proteins involved in mRNP assembly, mRNA export and RNA processing have been largely analyzed by deletion or depletion of the whole protein or at least entire protein domains, which is likely to abrogate several functions of each protein at once. In this study we were able to identify more than 100 RNA-binding sites within 16 RNA-binding proteins involved in mRNP formation by UV-light crosslinking and subsequent mass spectrometric analyses (including Npl3, Nab2, Tho1, Mex67-Mtr2 and in most components of the TREX complex; Tho2, Hpr1, Mft1, Hrb1, Sub2 and Yra1). For 70 of these binding sites, a single amino acid that crosslinked to RNA *in vivo* could be unambiguously identified. We mutated the newly identified binding sites of the SR-like protein Npl3 and investigated the functional consequences. The detailed analysis of three RNA binding-deficient mutants of Npl3 revealed that mutation of different domains of Npl3 has clearly individual outcomes *in vivo*. Npl3 was shown in previous studies to be involved in various aspects of mRNA metabolism, including transcription elongation and termination, 3' end processing, splicing, mRNP assembly and nuclear mRNA export. However, we here elucidated a novel function for Npl3: Using the *npl3-Linker* RNA binding-deficient mutant we were able to show that Npl3 is involved in the recruitment

and transfer of nuclear mRNP components to the mRNA. Surprisingly, the TREX complex is one of the components that are not properly transferred or fail to be retained at the mRNA without functional Npl3.

In the future, it would be interesting to investigate how Npl3 associates with other RBPs that showed a changed level in nuclear mRNPs and to understand how Npl3 contributes to their regulation. One interesting candidate is Sen1, part of the NNS complex. Moreover, our identification of novel RNA-binding sites in other mRNPs components makes the targeted creation of RNA binding-deficient mutants possible, analogous to what we have done for Npl3, offering the possibility to unravel new functions of these RBPs or gain mechanistic insight into their mode of action.

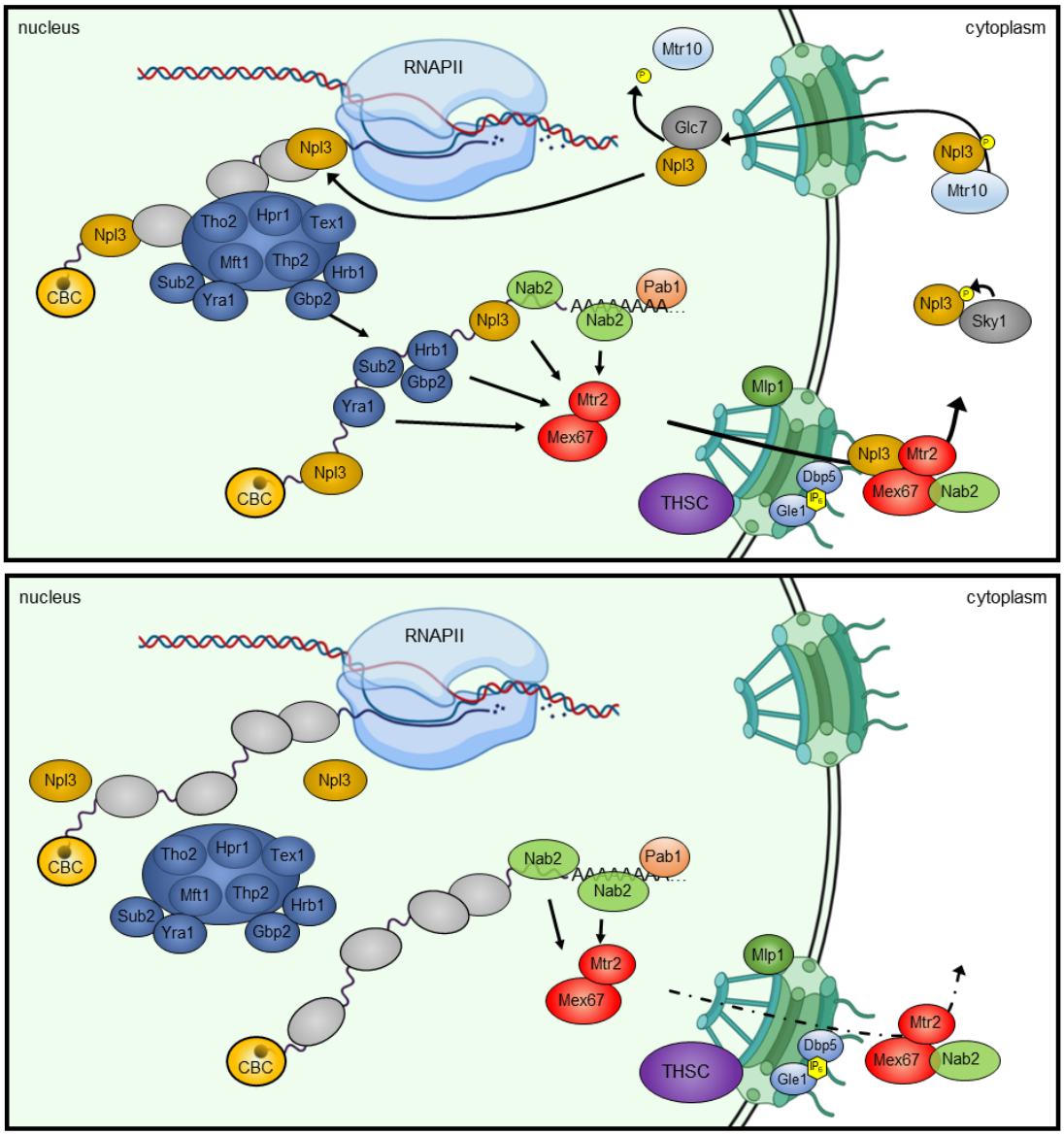



**Figure 31: Model for *npl3-Linker* mutant.** Upper panel: Wild-type Npl3 is recruited to nascent mRNA and is able to recruit other RBPs like the TREX complex to the mRNA or retain them. The mRNA is properly processed and exported to the cytoplasm. Lower panel: *npl3-Linker* is not able to bind mRNA anymore and thereby various RBPs cannot be recruited or retained at the mRNA. The lack of adaptors, except Nab2,

leads to reduced recruitment of Mex67-Mtr2, resulting in an mRNA export defect. Grey ovals represent other RBPs.

# 8 References

**Adams, R. L., Wente, S. R.** (2020) Dbp5 associates with RNA-bound Mex67 and Nab2 and its localization at the nuclear pore complex is sufficient for mRNP export and cell viability. *PLoS genetics* 16, e1009033.

**Adkins, M. W., Tyler, J. K.** (2006) Transcriptional activators are dispensable for transcription in the absence of Spt6-mediated chromatin reassembly of promoter regions. *Mol Cell* 21, 405-416.

**Aitchison, J. D., Blobel, G., Rout, M. P.** (1996) Kap104p: a karyopherin involved in the nuclear transport of messenger RNA binding proteins. *Science (New York, N.Y.)* 274, 624-627.

**Alber, F. *et al.*** (2007) The molecular architecture of the nuclear pore complex. *Nature* 450, 695-701.

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

**Amberg, D. C., Goldstein, A. L., Cole, C. N.** (1992) Isolation and characterization of RAT1: an essential gene of Saccharomyces cerevisiae required for the efficient nucleocytoplasmic trafficking of mRNA. *Genes & development* 6, 1173-1189.

**Andersen, P. R. *et al.*** (2013) The human cap-binding complex is functionally connected to the nuclear RNA exosome. *Nature structural & molecular biology* 20, 1367-1376.

**Aparicio, O., Geisberg, J. V., Sekinger, E., Yang, A., Moqtaderi, Z., Struhl, K.** (2005) Chromatin immunoprecipitation for determining the association of proteins with specific genomic sequences in vivo. *Current protocols in molecular biology* Chapter 21, Unit 21 23.

**Aparicio, O., Geisberg, J. V., Struhl, K.** (2004) Chromatin immunoprecipitation for determining the association of proteins with specific genomic sequences in vivo. *Current protocols in cell biology* Chapter 17, Unit 17.17.

**Arigo, J. T., Eyler, D. E., Carroll, K. L., Corden, J. L.** (2006) Termination of Cryptic Unstable Transcripts Is Directed by Yeast RNA-Binding Proteins Nrd1 and Nab3. *Molecular Cell* 23, 841-851.

**Asencio, C., Chatterjee, A., Hentze, M. W.** (2018) Silica-based solid-phase extraction of cross-linked nucleic acid–bound proteins. *Life Science Alliance* 1, e201800088.

**Baejen, C., Torkler, P., Gressel, S., Essig, K., Söding, J., Cramer, P.** (2014) Transcriptome maps of mRNP biogenesis factors define pre-mRNA recognition. *Mol Cell* 55, 745-757.

**Baierlein, C., Hackmann, A., Gross, T., Henker, L., Hinz, F., Krebber, H.** (2013) Monosome formation during translation initiation requires the serine/arginine-rich protein Npl3. *Molecular and cellular biology* 33, 4811-4823.

**Baltz, A. G. *et al.*** (2012) The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol Cell* 46, 674-690.

**Barillà, D., Lee, B. A., Proudfoot, N. J.** (2001) Cleavage/polyadenylation factor IA associates with the carboxyl-terminal domain of RNA polymerase II in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* 98, 445-450.

**Batisse, J., Batisse, C., Budd, A., Böttcher, B., Hurt, E.** (2009) Purification of nuclear poly(A)-binding protein Nab2 reveals association with the yeast transcriptome and a messenger ribonucleoprotein core structure. *The Journal of biological chemistry* 284, 34911-34917.

**Becalska, A. N., Gavis, E. R.** (2009) Lighting up mRNA localization in Drosophila oogenesis. *Development (Cambridge, England)* 136, 2493-2503.

**Berkovits, B. D., Mayr, C.** (2015) Alternative 3' UTRs act as scaffolds to regulate membrane protein localization. *Nature* 522, 363-367.

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

**Blencowe, B. J., Bowman, J. A., McCracken, S., Rosonina, E.** (1999) SR-related proteins and the processing of messenger RNA precursors. *Biochemistry and cell biology = Biochimie et biologie cellulaire* 77, 277-291.
**Blobel, G.** (1985) Gene gating: a hypothesis. *Proc Natl Acad Sci U S A* 82, 8527-8529.
**Bobola, N., Jansen, R. P., Shin, T. H., Nasmyth, K.** (1996) Asymmetric accumulation of Ash1p in postanaphase nuclei depends on a myosin and restricts yeast mating-type switching to mother cells. *Cell* 84, 699-709.
**Boisvert, F. M., Chénard, C. A., Richard, S.** (2005) Protein interfaces in signaling regulated by arginine methylation. *Science's STKE : signal transduction knowledge environment* 2005, re2.
**Bossie, M. A., DeHoratius, C., Barcelo, G., Silver, P.** (1992) A mutant nuclear protein with similarity to RNA binding proteins interferes with nuclear import in yeast. *Molecular biology of the cell* 3, 875-893.
**Bousquet-Antonelli, C., Presutti, C., Tollervey, D.** (2000) Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell* 102, 765-775.
**Bretes, H., Rouviere, J. O., Leger, T., Oeffinger, M., Devaux, F., Doye, V., Palancade, B.** (2014) Sumoylation of the THO complex regulates the biogenesis of a subset of mRNPs. *Nucleic acids research* 42, 5043-5058.
**Brockmann, C. *et al.*** (2012) Structural basis for polyadenosine-RNA binding by Nab2 Zn fingers and its function in mRNA nuclear export. *Structure (London, England : 1993)* 20, 1007-1018.
**Bucheli, M. E., Buratowski, S.** (2005) Npl3 is an antagonist of mRNA 3' end formation by RNA polymerase II. *The EMBO journal* 24, 2150-2160.
**Bucheli, M. E., He, X., Kaplan, C. D., Moore, C. L., Buratowski, S.** (2007) Polyadenylation site choice in yeast is affected by competition between Npl3 and polyadenylation factor CFI. *RNA (New York, N.Y.)* 13, 1756-1764.
**Bui, K. H. *et al.*** (2013) Integrated structural analysis of the human nuclear pore complex scaffold. *Cell* 155, 1233-1243.
**Busch, A., Hertel, K. J.** (2012) Evolution of SR protein and hnRNP splicing regulatory factors. *Wiley interdisciplinary reviews. RNA* 3, 1-12.
**Casañal, A. *et al.*** (2017) Architecture of eukaryotic mRNA 3'-end processing machinery. *Science (New York, N.Y.)* 358, 1056-1059.
**Castello, A. *et al.*** (2012) Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* 149, 1393-1406.
**Castello, A., Fischer, B., Frese, C. K., Horos, R., Alleaume, A. M., Foehr, S., Curk, T., Krijgsveld, J., Hentze, M. W.** (2016) Comprehensive Identification of RNA-Binding Domains in Human Cells. *Mol Cell* 63, 696-710.
**Caudron-Herger, M., Rusin, S. F., Adamo, M. E., Seiler, J., Schmid, V. K., Barreau, E., Kettenbach, A. N., Diederichs, S.** (2019) R-DeeP: Proteome-wide and Quantitative Identification of RNA-Dependent Proteins by Density Gradient Ultracentrifugation. *Mol Cell* 75, 184-199.e110.
**Chan, S., Choi, E. A., Shi, Y.** (2011) Pre-mRNA 3'-end processing complex assembly and function. *Wiley interdisciplinary reviews. RNA* 2, 321-335.
**Chávez, S., Beilharz, T., Rondón, A. G., Erdjument-Bromage, H., Tempst, P., Svejstrup, J. Q., Lithgow, T., Aguilera, A.** (2000) A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in Saccharomyces cerevisiae. *The EMBO journal* 19, 5824-5834.
**Chen, M., Manley, J. L.** (2009) Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nature reviews. Molecular cell biology* 10, 741-754.
**Cho, E. J., Kobor, M. S., Kim, M., Greenblatt, J., Buratowski, S.** (2001) Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. *Genes & development* 15, 3319-3329.
**Cho, E. J., Rodriguez, C. R., Takagi, T., Buratowski, S.** (1998) Allosteric interactions between capping enzyme subunits and the RNA polymerase II carboxy-terminal domain. *Genes & development* 12, 3482-3487.

**Cho, E. J., Takagi, T., Moore, C. R., Buratowski, S.** (1997) mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes & development* 11, 3319-3326.

**Cho, H., Kim, T. K., Mancebo, H., Lane, W. S., Flores, O., Reinberg, D.** (1999) A protein phosphatase functions to recycle RNA polymerase II. *Genes & development* 13, 1540-1552.

**Cléry, A., Blatter, M., Allain, F. H.** (2008) RNA recognition motifs: boring? Not quite. *Current opinion in structural biology* 18, 290-298.

**Cléry, A., Sinha, R., Anczuków, O., Corrionero, A., Moursy, A., Daubner, G. M., Valcárcel, J., Krainer, A. R., Allain, F. H. T.** (2013) Isolated pseudo–RNA-recognition motifs of SR proteins can regulate splicing using a noncanonical mode of RNA recognition. *Proceedings of the National Academy of Sciences* 110, E2802.

**Colombo, C. V., Trovesi, C., Menin, L., Longhese, M. P., Clerici, M.** (2017) The RNA binding protein Npl3 promotes resection of DNA double-strand breaks by regulating the levels of Exo1. *Nucleic acids research* 45, 6530-6545.

**Conrad, N. K., Wilson, S. M., Steinmetz, E. J., Patturajan, M., Brow, D. A., Swanson, M. S., Corden, J. L.** (2000) A yeast heterogeneous nuclear ribonucleoprotein complex associated with RNA polymerase II. *Genetics* 154, 557-571.

**Corden, J. L., Cadena, D. L., Ahearn, J. M., Jr., Dahmus, M. E.** (1985) A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. *Proceedings of the National Academy of Sciences of the United States of America* 82, 7934-7938.

**Cosma, M. P.** (2004) Daughter-specific repression of Saccharomyces cerevisiae HO: Ash1 is the commander. *EMBO Rep* 5, 953-957.

**Cox, J., Mann, M.** (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology* 26, 1367-1372.

**Creamer, T. J., Darby, M. M., Jamonnak, N., Schaughency, P., Hao, H., Wheelan, S. J., Corden, J. L.** (2011) Transcriptome-wide binding sites for components of the Saccharomyces cerevisiae non-poly(A) termination pathway: Nrd1, Nab3, and Sen1. *PLoS genetics* 7, e1002329.

**Cronshaw, J. M., Krutchinsky, A. N., Zhang, W., Chait, B. T., Matunis, M. J.** (2002) Proteomic analysis of the mammalian nuclear pore complex. *The Journal of cell biology* 158, 915-927.

**Czudnochowski, N., Bösken, C. A., Geyer, M.** (2012) Serine-7 but not serine-5 phosphorylation primes RNA polymerase II CTD for P-TEFb recognition. *Nat Commun* 3, 842.

**Deka, P., Bucheli, M. E., Moore, C., Buratowski, S., Varani, G.** (2008) Structure of the yeast SR protein Npl3 and Interaction with mRNA 3'-end processing signals. *Journal of molecular biology* 375, 136-150.

**Dermody, J. L. *et al.*** (2008) Unphosphorylated SR-like protein Npl3 stimulates RNA polymerase II elongation. *PloS one* 3, e3273.

**Derti, A., Garrett-Engele, P., Macisaac, K. D., Stevens, R. C., Sriram, S., Chen, R., Rohl, C. A., Johnson, J. M., Babak, T.** (2012) A quantitative atlas of polyadenylation in five mammals. *Genome research* 22, 1173-1183.

**Devos, D. P., Gräf, R., Field, M. C.** (2014) Evolution of the nucleus. *Current opinion in cell biology* 28, 8-15.

**Dias, S. M., Wilson, K. F., Rojas, K. S., Ambrosio, A. L., Cerione, R. A.** (2009) The molecular basis for the regulation of the cap-binding complex by the importins. *Nature structural & molecular biology* 16, 930-937.

**Dichtl, B., Blank, D., Sadowski, M., Hübner, W., Weiser, S., Keller, W.** (2002) Yhh1p/Cft1p directly links poly(A) site recognition and RNA polymerase II transcription termination. *The EMBO journal* 21, 4125-4135.

**Douglass, S. M., Leung, C. S., Johnson, T. L.** (2019) Extensive splicing across the Saccharomyces cerevisiae genome. *bioRxiv* 515163.

**Dreyfuss, G., Kim, V. N., Kataoka, N.** (2002) Messenger-RNA-binding proteins and the messages they carry. *Nature reviews. Molecular cell biology* 3, 195-205.

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

**Ellis, E. M., Reid, G. A.** (1993) The Saccharomyces cerevisiae MTS1 gene encodes a putative RNA-binding protein involved in mitochondrial protein targeting. *Gene* 132, 175-183.

**Estrella, L. A., Wilkinson, M. F., Gonzalez, C. I.** (2009) The shuttling protein Npl3 promotes translation termination accuracy in Saccharomyces cerevisiae. *Journal of molecular biology* 394, 410-422.

**Fair, B. J., Pleiss, J. A.** (2017) The power of fission: yeast as a tool for understanding complex splicing. *Current genetics* 63, 375-380.

**Fairbanks, G., Steck, T., Wallach, D.** (1971) Electrophoretic Analysis of the Major Polypeptides of the Human Erythrocyte Membrane. *Biochemistry* 10, 2606-2617.

**Fasken, M. B., Stewart, M., Corbett, A. H.** (2008) Functional significance of the interaction between the mRNA-binding protein, Nab2, and the nuclear pore-associated protein, Mlp1, in mRNA export. *The Journal of biological chemistry* 283, 27130-27143.

**Fischer, T., Rodríguez-Navarro, S., Pereira, G., Rácz, A., Schiebel, E., Hurt, E.** (2004) Yeast centrin Cdc31 is linked to the nuclear mRNA export machinery. *Nature cell biology* 6, 840-848.

**Fischer, T., Strässer, K., Rácz, A., Rodriguez-Navarro, S., Oppizzi, M., Ihrig, P., Lechner, J., Hurt, E.** (2002) The mRNA export machinery requires the novel Sac3p-Thp1p complex to dock at the nucleoplasmic entrance of the nuclear pores. *The EMBO journal* 21, 5843-5852.

**Flach, J., Bossie, M., Vogel, J., Corbett, A., Jinks, T., Willins, D. A., Silver, P. A.** (1994) A yeast RNA-binding protein shuttles between the nucleus and the cytoplasm. *Molecular and cellular biology* 14, 8399-8407.

**Flaherty, S. M., Fortes, P., Izaurralde, E., Mattaj, I. W., Gilmartin, G. M.** (1997) Participation of the nuclear cap binding complex in pre-mRNA 3' processing. *Proc Natl Acad Sci U S A* 94, 11893-11898.

**Gallardo, M., Aguilera, A.** (2001) A new hyperrecombination mutation identifies a novel yeast gene, THP1, connecting transcription elongation with mitotic recombination. *Genetics* 157, 79-89.

**Gallardo, M., Luna, R., Erdjument-Bromage, H., Tempst, P., Aguilera, A.** (2003) Nab2p and the Thp1p-Sac3p complex functionally interact at the interface between transcription and mRNA metabolism. *The Journal of biological chemistry* 278, 24225-24232.

**Galy, V., Gadal, O., Fromont-Racine, M., Romano, A., Jacquier, A., Nehrbass, U.** (2004) Nuclear retention of unspliced mRNAs in yeast is mediated by perinuclear Mlp1. *Cell* 116, 63-73.

**Geisberg, J. V., Moqtaderi, Z., Fan, X., Ozsolak, F., Struhl, K.** (2014) Global analysis of mRNA isoform half-lives reveals stabilizing and destabilizing elements in yeast. *Cell* 156, 812-824.

**Ghaemmaghami, S., Huh, W. K., Bower, K., Howson, R. W., Belle, A., Dephoure, N., O'Shea, E. K., Weissman, J. S.** (2003) Global analysis of protein expression in yeast. *Nature* 425, 737-741.

**Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3rd, Smith, H. O.** (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods* 6, 343-345.

**Gilbert, W., Guthrie, C.** (2004) The Glc7p nuclear phosphatase promotes mRNA export by facilitating association of Mex67p with mRNA. *Mol Cell* 13, 201-212.

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

**Grant, R. P., Hurt, E., Neuhaus, D., Stewart, M.** (2002) Structure of the C-terminal FG-nucleoporin binding domain of Tap/NXF1. *Nature structural biology* 9, 247-251.

**Grant, R. P., Marshall, N. J., Yang, J. C., Fasken, M. B., Kelly, S. M., Harreman, M. T., Neuhaus, D., Corbett, A. H., Stewart, M.** (2008) Structure of the N-terminal Mlp1-binding domain of the Saccharomyces cerevisiae mRNA-binding protein, Nab2. *Journal of molecular biology* 376, 1048-1059.

**Gratzer, S., Beilharz, T., Beddoe, T., Henry, M. F., Lithgow, T.** (2000) The mitochondrial protein targeting suppressor (mts1) mutation maps to the mRNA-binding domain of Npl3p and affects translation on cytoplasmic polysomes. *Molecular microbiology* 35, 1277-1285.
**Green, D. M., Marfatia, K. A., Crafton, E. B., Zhang, X., Cheng, X., Corbett, A. H.** (2002) Nab2p is required for poly(A) RNA export in Saccharomyces cerevisiae and is regulated by arginine methylation via Hmt1p. *The Journal of biological chemistry* 277, 7752-7760.
**Grohmann, D., Werner, F.** (2011) Cycling through transcription with the RNA polymerase F/E (RPB4/7) complex: structure, function and evolution of archaeal RNA polymerase. *Res Microbiol* 162, 10-18.
**Häcker, S., Krebber, H.** (2004) Differential export requirements for shuttling serine/arginine-type mRNA-binding proteins. *The Journal of biological chemistry* 279, 5049-5052.
**Hackmann, A., Gross, T., Baierlein, C., Krebber, H.** (2011) The mRNA export factor Npl3 mediates the nuclear export of large ribosomal subunits. *EMBO Rep* 12, 1024-1031.
**Hackmann, A., Wu, H., Schneider, U.-M., Meyer, K., Jung, K., Krebber, H.** (2014) Quality control of spliced mRNAs requires the shuttling SR proteins Gbp2 and Hrb1. *Nature Communications* 5, 3123.
**Hafner, M. *et al.*** (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 141, 129-141.
**Hall, T. M.** (2005) Multiple modes of RNA recognition by zinc finger proteins. *Current opinion in structural biology* 15, 367-373.
**Hartzog, G. A., Wada, T., Handa, H., Winston, F.** (1998) Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in Saccharomyces cerevisiae. *Genes & development* 12, 357-369.
**He, F., Jacobson, A.** (2001) Upf1p, Nmd2p, and Upf3p regulate the decapping and exonucleolytic degradation of both nonsense-containing mRNAs and wild-type mRNAs. *Molecular and cellular biology* 21, 1515-1530.
**Henry, M. F., Silver, P. A.** (1996) A novel methyltransferase (Hmt1p) modifies poly(A)+-RNA-binding proteins. *Molecular and cellular biology* 16, 3668-3678.
**Hentze, M. W., Castello, A., Schwarzl, T., Preiss, T.** (2018) A brave new world of RNA-binding proteins. *Nature reviews. Molecular cell biology* 19, 327-341.
**Hertel, K. J., Graveley, B. R.** (2005) RS domains contact the pre-mRNA throughout spliceosome assembly. *Trends in biochemical sciences* 30, 115-118.
**Herzel, L., Ottoz, D. S. M., Alpert, T., Neugebauer, K. M.** (2017) Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. *Nature reviews. Molecular cell biology* 18, 637-650.
**Hieronymus, H., Silver, P. A.** (2003) Genome-wide analysis of RNA-protein interactions illustrates specificity of the mRNA export machinery. *Nature genetics* 33, 155-161.
**Ho, C. K., Shuman, S.** (1999) Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. *Mol Cell* 3, 405-411.
**Ho, Y. *et al.*** (2002) Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature* 415, 180-183.
**Holmes, R. K. *et al.*** (2015) Loss of the Yeast SR Protein Npl3 Alters Gene Expression Due to Transcription Readthrough. *PLoS genetics* 11, e1005735.
**Hooks, K. B., Delneri, D., Griffiths-Jones, S.** (2014) Intron Evolution in Saccharomycetaceae. *Genome Biology and Evolution* 6, 2543-2556.
**Huang, T., Böhlenius, H., Eriksson, S., Parcy, F., Nilsson, O.** (2005) The mRNA of the Arabidopsis gene FT moves from leaf to shoot apex and induces flowering. *Science (New York, N.Y.)* 309, 1694-1696.
**Hurt, E., Luo, M.-j., Röther, S., Reed, R., Sträßer, K.** (2004) Cotranscriptional recruitment of the serine-arginine-rich (SR)-like proteins Gbp2 and Hrb1 to nascent mRNA via the TREX complex. *Proceedings of the National Academy of Sciences* 101, 1858.

**Iglesias, N., Stutz, F.** (2008) Regulation of mRNP dynamics along the export pathway. *FEBS letters* 582, 1987-1996.
**Iglesias, N., Tutucci, E., Gwizdek, C., Vinciguerra, P., Von Dach, E., Corbett, A. H., Dargemont, C., Stutz, F.** (2010) Ubiquitin-mediated mRNP dynamics and surveillance prior to budding yeast mRNA export. *Genes & development* 24, 1927-1938.
**Jaehning, J. A.** (2010) The Paf1 complex: platform or player in RNA polymerase II transcription? *Biochimica et biophysica acta* 1799, 379-388.
**Jan, C. H., Friedman, R. C., Ruby, J. G., Bartel, D. P.** (2011) Formation, regulation and evolution of Caenorhabditis elegans 3'UTRs. *Nature* 469, 97-101.
**Jimeno, S., Rondón, A. G., Luna, R., Aguilera, A.** (2002) The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability. *The EMBO journal* 21, 3526-3535.
**Johnson, S. A., Cubberley, G., Bentley, D. L.** (2009) Cotranscriptional recruitment of the mRNA export factor Yra1 by direct interaction with the 3' end processing factor Pcf11. *Mol Cell* 33, 215-226.
**Jones, J. C., Phatnani, H. P., Haystead, T. A., MacDonald, J. A., Alam, S. M., Greenleaf, A. L.** (2004) C-terminal repeat domain kinase I phosphorylates Ser2 and Ser5 of RNA polymerase II C-terminal domain repeats. *The Journal of biological chemistry* 279, 24957-24964.
**Kadowaki, T. *et al.*** (1994) Isolation and characterization of Saccharomyces cerevisiae mRNA transport-defective (mtr) mutants. *The Journal of cell biology* 126, 649-659.
**Katahira, J., Strässer, K., Podtelejnikov, A., Mann, M., Jung, J. U., Hurt, E.** (1999) The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human. *The EMBO journal* 18, 2593-2609.
**Kelly, S. M., Corbett, A. H.** (2009) Messenger RNA export from the nucleus: a series of molecular wardrobe changes. *Traffic (Copenhagen, Denmark)* 10, 1199-1208.
**Kelly, S. M., Leung, S. W., Apponi, L. H., Bramley, A. M., Tran, E. J., Chekanova, J. A., Wente, S. R., Corbett, A. H.** (2010) Recognition of polyadenosine RNA by the zinc finger domain of nuclear poly(A) RNA-binding protein 2 (Nab2) is required for correct mRNA 3'-end formation. *The Journal of biological chemistry* 285, 26022-26032.
**Keren, H., Lev-Maor, G., Ast, G.** (2010) Alternative splicing and evolution: diversification, exon definition and function. *Nature reviews. Genetics* 11, 345-355.
**Kim, J., Roeder, R. G.** (2009) Direct Bre1-Paf1 complex interactions and RING finger-independent Bre1-Rad6 interactions mediate histone H2B ubiquitylation in yeast. *The Journal of biological chemistry* 284, 20582-20592.
**Klass, D. M., Scheibe, M., Butter, F., Hogan, G. J., Mann, M., Brown, P. O.** (2013) Quantitative proteomic analysis reveals concurrent RNA-protein interactions and identifies new RNA-binding proteins in Saccharomyces cerevisiae. *Genome research* 23, 1028-1038.
**Klein, B. J., Bose, D., Baker, K. J., Yusoff, Z. M., Zhang, X., Murakami, K. S.** (2011) RNA polymerase and transcription elongation factor Spt4/5 complex structure. *Proc Natl Acad Sci U S A* 108, 546-550.
**Knockenhauer, K. E., Schwartz, T. U.** (2016) The Nuclear Pore Complex as a Flexible and Dynamic Gate. *Cell* 164, 1162-1171.
**Köhler, A., Hurt, E.** (2007) Exporting RNA from the nucleus to the cytoplasm. *Nature reviews. Molecular cell biology* 8, 761-773.
**Kramer, K., Sachsenberg, T., Beckmann, B. M., Qamar, S., Boon, K. L., Hentze, M. W., Kohlbacher, O., Urlaub, H.** (2014) Photo-cross-linking and high-resolution mass spectrometry for assignment of RNA-binding sites in RNA-binding proteins. *Nature methods* 11, 1064-1070.
**Kress, T. L., Krogan, N. J., Guthrie, C.** (2008) A single SR-like protein, Npl3, promotes pre-mRNA splicing in budding yeast. *Mol Cell* 32, 727-734.
**Kressler, D. *et al.*** (2012) Synchronizing nuclear import of ribosomal proteins with ribosome assembly. *Science (New York, N.Y.)* 338, 666-671.

**Krishnamurthy, S., He, X., Reyes-Reyes, M., Moore, C., Hampsey, M.** (2004) Ssu72 Is an RNA polymerase II CTD phosphatase. *Mol Cell* 14, 387-394.
**Kyburz, A., Sadowski, M., Dichtl, B., Keller, W.** (2003) The role of the yeast cleavage and polyadenylation factor subunit Ydh1p/Cft2p in pre-mRNA 3'-end formation. *Nucleic acids research* 31, 3936-3945.
**Laemmli, U. K.** (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* 227, 680-685.
**Langie, S. A. S. *et al.*** (2015) Causes of genome instability: the effect of low dose chemical exposures in modern society. *Carcinogenesis* 36 Suppl 1, S61-S88.
**Lee, D. C., Aitchison, J. D.** (1999) Kap104p-mediated nuclear import. Nuclear localization signals in mRNA-binding proteins and the role of Ran and Rna. *The Journal of biological chemistry* 274, 29031-29037.
**Lee, M. S., Henry, M., Silver, P. A.** (1996) A protein that shuttles between the nucleus and the cytoplasm is an important mediator of RNA export. *Genes & development* 10, 1233-1246.
**Lei, E. P., Krebber, H., Silver, P. A.** (2001) Messenger RNAs are recruited for nuclear export during transcription. *Genes & development* 15, 1771-1782.
**Liao, S. M., Zhang, J., Jeffery, D. A., Koleske, A. J., Thompson, C. M., Chao, D. M., Viljoen, M., van Vuuren, H. J., Young, R. A.** (1995) A kinase-cyclin pair in the RNA polymerase II holoenzyme. *Nature* 374, 193-196.
**Lidschreiber, M., Leike, K., Cramer, P.** (2013) Cap completion and C-terminal repeat domain kinase recruitment underlie the initiation-elongation transition of RNA polymerase II. *Molecular and cellular biology* 33, 3805-3816.
**Linder, P., Jankowsky, E.** (2011) From unwinding to clamping - the DEAD box RNA helicase family. *Nature reviews. Molecular cell biology* 12, 505-516.
**Lindstrom, D. L., Squazzo, S. L., Muster, N., Burckin, T. A., Wachter, K. C., Emigh, C. A., McCleery, J. A., Yates, J. R., 3rd, Hartzog, G. A.** (2003) Dual roles for Spt5 in pre-mRNA processing and transcription elongation revealed by identification of Spt5-associated proteins. *Molecular and cellular biology* 23, 1368-1378.
**Liu, Q., Dreyfuss, G.** (1995) In vivo and in vitro arginine methylation of RNA-binding proteins. *Molecular and cellular biology* 15, 2800-2808.
**Liu, X., Bushnell, D. A., Wang, D., Calero, G., Kornberg, R. D.** (2010) Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism. *Science (New York, N.Y.)* 327, 206-209.
**Liu, Y., Warfield, L., Zhang, C., Luo, J., Allen, J., Lang, W. H., Ranish, J., Shokat, K. M., Hahn, S.** (2009) Phosphorylation of the transcription elongation factor Spt5 by yeast Bur1 kinase stimulates recruitment of the PAF complex. *Molecular and cellular biology* 29, 4852-4863.
**Lu, D., Searles, M. A., Klug, A.** (2003) Crystal structure of a zinc-finger-RNA complex reveals two modes of molecular recognition. *Nature* 426, 96-100.
**Lund, M. K., Guthrie, C.** (2005) The DEAD-box protein Dbp5p is required to dissociate Mex67p from exported mRNPs at the nuclear rim. *Mol Cell* 20, 645-651.
**Lunde, B. M. *et al.*** (2010) Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. *Nature structural & molecular biology* 17, 1195-1201.
**Mandel, C. R., Bai, Y., Tong, L.** (2008) Protein factors in pre-mRNA 3'-end processing. *Cellular and molecular life sciences : CMLS* 65, 1099-1122.
**Mao, X., Schwer, B., Shuman, S.** (1995) Yeast mRNA cap methyltransferase is a 50-kilodalton protein encoded by an essential gene. *Molecular and cellular biology* 15, 4167-4174.
**Marfatia, K. A., Crafton, E. B., Green, D. M., Corbett, A. H.** (2003) Domain analysis of the Saccharomyces cerevisiae heterogeneous nuclear ribonucleoprotein, Nab2p. Dissecting the requirements for Nab2p-facilitated poly(A) RNA export. *The Journal of biological chemistry* 278, 6731-6740.
**Maris, C., Dominguez, C., Allain, F. H.** (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *The FEBS journal* 272, 2118-2131.

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

**Mayer, A., Lidschreiber, M., Siebert, M., Leike, K., Söding, J., Cramer, P.** (2010) Uniform transitions of the general RNA polymerase II transcription complex. *Nature structural & molecular biology* 17, 1272-1278.

**Mayer, A., Schreieck, A., Lidschreiber, M., Leike, K., Martin, D. E., Cramer, P.** (2012) The spt5 C-terminal region recruits yeast 3' RNA cleavage factor I. *Molecular and cellular biology* 32, 1321-1331.

**McKinney, J. S., Sethi, S., Tripp, J. D., Nguyen, T. N., Sanderson, B. A., Westmoreland, J. W., Resnick, M. A., Lewis, L. K.** (2013) A multistep genomic screen identifies new genes required for repair of DNA double-strand breaks in Saccharomyces cerevisiae. *BMC genomics* 14, 251.

**Meinel, D. M. *et al.*** (2013) Recruitment of TREX to the transcription machinery by its direct binding to the phospho-CTD of RNA polymerase II. *PLoS genetics* 9, e1003914.

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

**Mitchell, S. F., Jain, S., She, M., Parker, R.** (2013) Global analysis of yeast mRNPs. *Nature structural & molecular biology* 20, 127-133.

**Moehle, E. A., Ryan, C. J., Krogan, N. J., Kress, T. L., Guthrie, C.** (2012) The yeast SR-like protein Npl3 links chromatin modification to mRNA processing. *PLoS genetics* 8, e1003101.

**Mohr, D., Frey, S., Fischer, T., Güttler, T., Görlich, D.** (2009) Characterisation of the passive permeability barrier of nuclear pore complexes. *The EMBO journal* 28, 2541-2553.

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

**Morris, D. P., Greenleaf, A. L.** (2000) The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II. *The Journal of biological chemistry* 275, 39935-39943.

**Mosley, A. L., Pattenden, S. G., Carey, M., Venkatesh, S., Gilmore, J. M., Florens, L., Workman, J. L., Washburn, M. P.** (2009) Rtr1 is a CTD phosphatase that regulates RNA polymerase II during the transition from serine 5 to serine 2 phosphorylation. *Mol Cell* 34, 168-178.

**Müller-McNicoll, M., Neugebauer, K. M.** (2013) How cells get the message: dynamic assembly and function of mRNA-protein complexes. *Nature reviews. Genetics* 14, 275-287.

**Neve, J., Burger, K., Li, W., Hoque, M., Patel, R., Tian, B., Gullerova, M., Furger, A.** (2016) Subcellular RNA profiling links splicing and nuclear DICER1 to alternative cleavage and polyadenylation. *Genome research* 26, 24-35.

**Ng, H. H., Robert, F., Young, R. A., Struhl, K.** (2003) Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell* 11, 709-719.

**Nilsen, T. W., Graveley, B. R.** (2010) Expansion of the eukaryotic proteome by alternative splicing. *Nature* 463, 457-463.

**Noble, C. G., Hollingworth, D., Martin, S. R., Ennis-Adeniran, V., Smerdon, S. J., Kelly, G., Taylor, I. A., Ramos, A.** (2005) Key features of the interaction between Pcf11 CID and RNA polymerase II CTD. *Nature structural & molecular biology* 12, 144-151.

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

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

**Ozsolak, F., Kapranov, P., Foissac, S., Kim, S. W., Fishilevich, E., Monaghan, A. P., John, B., Milos, P. M.** (2010) Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation. *Cell* 143, 1018-1029.

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

**Paine, P. L., Feldherr, C. M.** (1972) Nucleocytoplasmic exchange of macromolecules. *Experimental cell research* 74, 81-98.

**Paine, P. L., Moore, L. C., Horowitz, S. B.** (1975) Nuclear envelope permeability. *Nature* 254, 109-114.

**Palancade, B., Zuccolo, M., Loeillet, S., Nicolas, A., Doye, V.** (2005) Pml39, a novel protein of the nuclear periphery required for nuclear retention of improper messenger ribonucleoparticles. *Molecular biology of the cell* 16, 5258-5268.

**Panhale, A., Richter, F. M., Ramírez, F., Shvedunova, M., Manke, T., Mittler, G., Akhtar, A.** (2019) CAPRI enables comparison of evolutionarily conserved RNA interacting regions. *Nature Communications* 10, 2682.

**Panté, N., Kann, M.** (2002) Nuclear pore complex is able to transport macromolecules with diameters of about 39 nm. *Molecular biology of the cell* 13, 425-434.

**Pascual-García, P., Govind, C. K., Queralt, E., Cuenca-Bono, B., Llopis, A., Chavez, S., Hinnebusch, A. G., Rodríguez-Navarro, S.** (2008) Sus1 is recruited to coding regions and functions during transcription elongation in association with SAGA and TREX2. *Genes & development* 22, 2811-2822.

**Pei, Y., Shuman, S.** (2002) Interactions between fission yeast mRNA capping enzymes and elongation factor Spt5. *The Journal of biological chemistry* 277, 19639-19648.

**Peil, L., Waghmare, S., Fischer, L., Spitzer, M., Tollervey, D., Rappsilber, J.** (2018) Identification of RNA-associated peptides, iRAP, defines precise sites of protein-RNA interaction. *bioRxiv* 456111.

**Pemberton, L. F., Rosenblum, J. S., Blobel, G.** (1997) A distinct and parallel pathway for the nuclear import of an mRNA-binding protein. *The Journal of cell biology* 139, 1645-1653.

**Pérez-Martínez, L., Öztürk, M., Butter, F., Luke, B.** (2020) Npl3 stabilizes R-loops at telomeres to prevent accelerated replicative senescence. *EMBO Rep* 21, e49087-e49087.

**Piruat, J. I., Aguilera, A.** (1998) A novel yeast gene, THO2, is involved in RNA pol II transcription and provides new evidence for transcriptional elongation-associated recombination. *The EMBO journal* 17, 4859-4872.

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

**Qiu, H., Hu, C., Hinnebusch, A. G.** (2009) Phosphorylation of the Pol II CTD by KIN28 enhances BUR1/BUR2 recruitment and Ser2 CTD phosphorylation near promoters. *Mol Cell* 33, 752-762.

**Qu, X., Lykke-Andersen, S., Nasser, T., Saguez, C., Bertrand, E., Jensen, T. H., Moore, C.** (2009) Assembly of an export-competent mRNP is needed for efficient release of the 3'-end processing complex after polyadenylation. *Molecular and cellular biology* 29, 5327-5338.

**Radu, A., Moore, M. S., Blobel, G.** (1995) The peptide repeat domain of nucleoporin Nup98 functions as a docking site in transport across the nuclear pore complex. *Cell* 81, 215-222.

**Rajyaguru, P., Parker, R.** (2012) RGG motif proteins: modulators of mRNA functional states. *Cell cycle (Georgetown, Tex.)* 11, 2594-2599.

**Rajyaguru, P., She, M., Parker, R.** (2012) Scd6 targets eIF4G to repress translation: RGG motif proteins as a class of eIF4G-binding proteins. *Mol Cell* 45, 244-254.

**Reichelt, R., Holzenburg, A., Buhle, E. L., Jr., Jarnik, M., Engel, A., Aebi, U.** (1990) Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components. *The Journal of cell biology* 110, 883-894.

**Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., Séraphin, B.** (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nature biotechnology* 17, 1030-1032.

**Rodríguez-Navarro, S., Fischer, T., Luo, M. J., Antúnez, O., Brettschneider, S., Lechner, J., Pérez-Ortín, J. E., Reed, R., Hurt, E.** (2004) Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. *Cell* 116, 75-86.

**Rodriguez, C. R., Cho, E. J., Keogh, M. C., Moore, C. L., Greenleaf, A. L., Buratowski, S.** (2000) Kin28, the TFIIH-associated carboxy-terminal domain kinase, facilitates the recruitment of mRNA processing machinery to RNA polymerase II. *Molecular and cellular biology* 20, 104-112.

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

**Röther, S.** (2007) Functional Analysis of the RNA Polymerase II C-terminal Domain Kinase Ctk1 in the Yeast Saccharomyces cerevisiae.

**Röther, S., Burkert, C., Brünger, K. M., Mayer, A., Kieser, A., Strässer, K.** (2010) Nucleocytoplasmic shuttling of the La motif-containing protein Sro9 might link its nuclear and cytoplasmic functions. *RNA (New York, N.Y.)* 16, 1393-1401.

**Rout, M. P., Aitchison, J. D., Suprapto, A., Hjertaas, K., Zhao, Y., Chait, B. T.** (2000) The yeast nuclear pore complex: composition, architecture, and transport mechanism. *The Journal of cell biology* 148, 635-651.

**Russell, I. D., Tollervey, D.** (1992) NOP3 is an essential yeast protein which is required for pre-rRNA processing. *The Journal of cell biology* 119, 737-747.

**Sakharkar, M. K., Chow, V. T., Chaturvedi, I., Mathura, V. S., Shapshak, P., Kangueane, P.** (2004) A report on single exon genes (SEG) in eukaryotes. *Frontiers in bioscience : a journal and virtual library* 9, 3262-3267.

**Sambrook, J. F., Russell, D.**, *Molecular Cloning: A Laboratory Manual (3-Volume Set)*. (2001), vol. 1.

**Santos-Pereira, J. M., Herrero, A. B., Garcia-Rubio, M. L., Marin, A., Moreno, S., Aguilera, A.** (2013) The Npl3 hnRNP prevents R-loop-mediated transcription-replication conflicts and genome instability. *Genes & development* 27, 2445-2458.

**Santos-Pereira, J. M., Herrero, A. B., Moreno, S., Aguilera, A.** (2014) Npl3, a new link between RNA-binding proteins and the maintenance of genome integrity. *Cell cycle (Georgetown, Tex.)* 13, 1524-1529.

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

**Schneider, M., Hellerschmied, D., Schubert, T., Amlacher, S., Vinayachandran, V., Reja, R., Pugh, B. F., Clausen, T., Köhler, A.** (2015) The Nuclear Pore-Associated TREX-2 Complex Employs Mediator to Regulate Gene Expression. *Cell* 162, 1016-1028.

**Sen, R., Barman, P., Kaja, A., Ferdoush, J., Lahudkar, S., Roy, A., Bhaumik, S. R.** (2019) Distinct Functions of the Cap-Binding Complex in Stimulation of Nuclear mRNA Export. *Molecular and cellular biology* 39,

**Senger, B., Simos, G., Bischoff, F. R., Podtelejnikov, A., Mann, M., Hurt, E.** (1998) Mtr10p functions as a nuclear import receptor for the mRNA-binding protein Npl3p. *The EMBO journal* 17, 2196-2207.

**Shchepachev, V., Bresson, S., Spanos, C., Petfalski, E., Fischer, L., Rappsilber, J., Tollervey, D.** (2019) Defining the RNA interactome by total RNA-associated protein purification. *Mol Syst Biol* 15, e8689-e8689.

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

**Shen, E. C., Stage-Zimmermann, T., Chui, P., Silver, P. A.** (2000) 7The yeast mRNA-binding protein Npl3p interacts with the cap-binding complex. *The Journal of biological chemistry* 275, 23718-23724.

**Shepard, P. J., Hertel, K. J.** (2009) The SR protein family. *Genome biology* 10, 242.
**Shevchenko, A., Wilm, M., Vorm, O., Mann, M.** (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Analytical chemistry* 68, 850-858.
**Shibagaki, Y., Itoh, N., Yamada, H., Nagata, S., Mizumoto, K.** (1992) mRNA capping enzyme. Isolation and characterization of the gene encoding mRNA guanylytransferase subunit from Saccharomyces cerevisiae. *Journal of Biological Chemistry* 267, 9521-9528.
**Shieh, G. S. *et al.*** (2011) H2B ubiquitylation is part of chromatin architecture that marks exon-intron structure in budding yeast. *BMC genomics* 12, 627.
**Shuman, S.** (2001) Structure, mechanism, and evolution of the mRNA capping apparatus. *Progress in nucleic acid research and molecular biology* 66, 1-40.
**Siebel, C. W., Feng, L., Guthrie, C., Fu, X. D.** (1999) Conservation in budding yeast of a kinase specific for SR splicing factors. *Proc Natl Acad Sci U S A* 96, 5440-5445.
**Siebel, C. W., Guthrie, C.** (1996) The essential yeast RNA binding protein Np13p is methylated. *Proc Natl Acad Sci U S A* 93, 13641-13646.
**Sikorski, R. S., Hieter, P.** (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics* 122, 19-27.
**Skrisovska, L., Allain, F. H.** (2008) Improved segmental isotope labeling methods for the NMR study of multidomain or large proteins: application to the RRMs of Npl3p and hnRNP L. *Journal of molecular biology* 375, 151-164.
**Smibert, P. *et al.*** (2012) Global patterns of tissue-specific alternative polyadenylation in Drosophila. *Cell reports* 1, 277-289.
**Smith, T. *et al.*** (2020) Organic phase separation opens up new opportunities to interrogate the RNA-binding proteome. *Current Opinion in Chemical Biology* 54, 70-75.
**Smolka, M. B., Albuquerque, C. P., Chen, S. H., Zhou, H.** (2007) Proteome-wide identification of in vivo targets of DNA damage checkpoint kinases. *Proc Natl Acad Sci U S A* 104, 10364-10369.
**Sonnleitner, E., Prindl, K., Bläsi, U.** (2017) The Pseudomonas aeruginosa CrcZ RNA interferes with Hfq-mediated riboregulation. *PloS one* 12, e0180887.
**Soucek, S., Corbett, A. H., Fasken, M. B.** (2012) The long and the short of it: the role of the zinc finger polyadenosine RNA binding protein, Nab2, in control of poly(A) tail length. *Biochimica et biophysica acta* 1819, 546-554.
**Spies, N., Burge, C. B., Bartel, D. P.** (2013) 3' UTR-isoform choice has limited influence on the stability and translational efficiency of most mRNAs in mouse fibroblasts. *Genome research* 23, 2078-2090.
**Steinmetz, E. J., Brow, D. A.** (2003) Ssu72 protein mediates both poly(A)-coupled and poly(A)-independent termination of RNA polymerase II transcription. *Molecular and cellular biology* 23, 6339-6349.
**Steinmetz, E. J., Conrad, N. K., Brow, D. A., Corden, J. L.** (2001) RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. *Nature* 413, 327-331.
**Steinmetz, E. J., Warren, C. L., Kuehner, J. N., Panbehi, B., Ansari, A. Z., Brow, D. A.** (2006) Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Mol Cell* 24, 735-746.
**Strässer, K., Hurt, E.** (2000) Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. *The EMBO journal* 19, 410-420.
**Strässer, K., Hurt, E.** (2001) Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p. *Nature* 413, 648-652.
**Strässer, K. *et al.*** (2002) TREX is a conserved complex coupling transcription with messenger RNA export. *Nature* 417, 304-308.
**Stutz, F., Bachi, A., Doerks, T., Braun, I. C., Séraphin, B., Wilm, M., Bork, P., Izaurralde, E.** (2000) REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. *RNA (New York, N.Y.)* 6, 638-650.

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

**Tamura, K., Fukao, Y., Iwamoto, M., Haraguchi, T., Hara-Nishimura, I.** (2010) Identification and Characterization of Nuclear Pore Complex Components in <em>Arabidopsis thaliana</em&gt. *The Plant Cell* 22, 4084.

**Taylor, R. G., Walker, D. C., McInnes, R. R.** (1993) E. coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic acids research* 21, 1677-1678.

**Terry, L. J., Wente, S. R.** (2007) Nuclear mRNA export requires specific FG nucleoporins for translocation through the nuclear pore complex. *The Journal of cell biology* 178, 1121-1132.

**Thakurta, A. G., Gopal, G., Yoon, J. H., Saha, T., Dhar, R.** (2004) Conserved nuclear export sequences in Schizosaccharomyces pombe Mex67 and human TAP function in mRNA export by direct nuclear pore interactions. *The Journal of biological chemistry* 279, 17434-17442.

**Thomas, B. J., Rothstein, R.** (1989) Elevated recombination rates in transcriptionally active DNA. *Cell* 56, 619-630.

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

**Towbin, H., Staehelin, T., Gordon, J.** (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences* 76, 4350.

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

**Tsukamoto, T., Shibagaki, Y., Imajoh-Ohmi, S., Murakoshi, T., Suzuki, M., Nakamura, A., Gotoh, H., Mizumoto, K.** (1997) Isolation and characterization of the yeast mRNA capping enzyme beta subunit gene encoding RNA 5'-triphosphatase, which is essential for cell viability. *Biochemical and biophysical research communications* 239, 116-122.

**Tuck, A. C., Tollervey, D.** (2013) A transcriptome-wide atlas of RNP composition reveals diverse classes of mRNAs and lncRNAs. *Cell* 154, 996-1009.

**Urdaneta, E. C. *et al.*** (2019) Purification of cross-linked RNA-protein complexes by phenol-toluol extraction. *Nature Communications* 10, 990.

**Valverde, R., Edwards, L., Regan, L.** (2008) Structure and function of KH domains. *The FEBS journal* 275, 2712-2726.

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

**Visa, N., Izaurralde, E., Ferreira, J., Daneholt, B., Mattaj, I. W.** (1996) A nuclear cap-binding complex binds Balbiani ring pre-mRNA cotranscriptionally and accompanies the ribonucleoprotein particle during nuclear export. *The Journal of cell biology* 133, 5-14.

**Vitaliano-Prunier, A., Babour, A., Hérissant, L., Apponi, L., Margaritis, T., Holstege, F. C., Corbett, A. H., Gwizdek, C., Dargemont, C.** (2012) H2B ubiquitylation controls the formation of export-competent mRNP. *Mol Cell* 45, 132-139.

**Wahl, M. C., Will, C. L., Lührmann, R.** (2009) The spliceosome: design principles of a dynamic RNP machine. *Cell* 136, 701-718.

**Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S. F., Schroth, G. P., Burge, C. B.** (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470-476.

**Webb, S., Hector, R. D., Kudla, G., Granneman, S.** (2014) PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast. *Genome biology* 15, R8.

**Wen, Y., Shatkin, A. J.** (1999) Transcription elongation factor hSPT5 stimulates mRNA capping. *Genes & development* 13, 1774-1779.

**Wilkening, S., Pelechano, V., Järvelin, A. I., Tekkedil, M. M., Anders, S., Benes, V., Steinmetz, L. M.** (2013) An efficient method for genome-wide polyadenylation site mapping and RNA quantification. *Nucleic acids research* 41, e65.

**Will, C. L., Lührmann, R.** (2011) Spliceosome structure and function. *Cold Spring Harbor perspectives in biology* 3,

**Wilson, S. M., Datar, K. V., Paddy, M. R., Swedlow, J. R., Swanson, M. S.** (1994) Characterization of nuclear polyadenylated RNA-binding proteins in Saccharomyces cerevisiae. *The Journal of cell biology* 127, 1173-1184.

**Windgassen, M., Sturm, D., Cajigas, I. J., González, C. I., Seedorf, M., Bastians, H., Krebber, H.** (2004) Yeast shuttling SR proteins Npl3p, Gbp2p, and Hrb1p are part of the translating mRNPs, and Npl3p can function as a translational repressor. *Molecular and cellular biology* 24, 10479-10491.

**Wong, C., Sridhara, S., Bardwell, J. C., Jakob, U.** (2000) Heating greatly speeds Coomassie blue staining and destaining. *BioTechniques* 28, 426-428, 430, 432.

**Wood, W. B.** (1966) Host specificity of DNA produced by Escherichia coli: bacterial mutations affecting the restriction and modification of DNA. *Journal of molecular biology* 16, 118-133.

**Xia, C. *et al.*** (2018) Elucidation of the Mechanisms of Long-Distance mRNA Movement in a Nicotiana benthamiana/Tomato Heterograft System. *Plant Physiol* 177, 745-758.

**Yamada, T., Yamaguchi, Y., Inukai, N., Okamoto, S., Mura, T., Handa, H.** (2006) P-TEFb-Mediated Phosphorylation of hSpt5 C-Terminal Repeats Is Critical for Processive Transcription Elongation. *Molecular Cell* 21, 227-237.

**Yao, W., Roser, D., Köhler, A., Bradatsch, B., Bassler, J., Hurt, E.** (2007) Nuclear export of ribosomal 60S subunits by the general mRNA export receptor Mex67-Mtr2. *Mol Cell* 26, 51-62.

**Yun, C. Y., Fu, X. D.** (2000) Conserved SR protein kinase functions in nuclear import and its action is counteracted by arginine methylation in Saccharomyces cerevisiae. *The Journal of cell biology* 150, 707-718.

**Zenklusen, D., Vinciguerra, P., Strahm, Y., Stutz, F.** (2001) The yeast hnRNP-Like proteins Yra1p and Yra2p participate in mRNA export through interaction with Mex67p. *Molecular and cellular biology* 21, 4219-4232.

**Zhang, D. W., Rodríguez-Molina, J. B., Tietjen, J. R., Nemec, C. M., Ansari, A. Z.** (2012) Emerging Views on the CTD Code. *Genet Res Int* 2012, 347214.

**Zhang, Z., Fu, J., Gilmour, D. S.** (2005) CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3'-end processing factor, Pcf11. *Genes & development* 19, 1572-1580.

**Zhang, Z. *et al.*** (2013) USP49 deubiquitinates histone H2B and regulates cotranscriptional pre-mRNA splicing. *Genes & development* 27, 1581-1595.

**Zhou, K., Kuo, W. H., Fillingham, J., Greenblatt, J. F.** (2009) Control of transcriptional elongation and cotranscriptional histone modification by the yeast BUR kinase substrate Spt5. *Proc Natl Acad Sci U S A* 106, 6956-6961.

# 9 List of figures

# 10 List of tables

# 11 Abbreviations

**Table 14: Abbreviations**

| Abbreviation | Description |
|---|---|
| 4-tU | 4-Thiouracil |
| 5-FOA | 5-Fluoroorotic acid |
| A | Alanine |
| AA | Amino acid |
| APA | Alternative polyadenylation |
| CBC | cap-binding complex |
| ChIP | Chromatin immunoprecipitation |
| CPF | cleavage and polyadenylation factor |
| CRAC | cross-linking and analysis of cDNA |
| CTD | carboxy-terminal domain of RNAPII |
| CTR | C-terminal region of Spt5 |
| D | Aspartic acid |
| E | Glutamic acid |
| *E. coli* | *Escherichia coli* |
| EtOH | ethanol |
| F | Phenylalanine |
| FG- | phenylalanine glycine- |
| FISH | Fluorescence *in situ* hybridization |
| FTpA | FLAG-TEV-prot A tag |
| G | Glycine |
| hnRNP | heterogeneous nuclear ribonucleoprotein |
| I | Isoleucine |
| ITC | Isothermal titration calorimetry |
| K | Lysine |
| KH | K homology |
| L | Leucine |
| MDa | Megadalton |
| mRNA | messenger ribonucleic acid |
| mRNP | messenger ribonucleoprotein particle |
| NPC | nuclear pore complex |
| Nup | nucleoporin |
| P | Proline |
| PAB | poly(A)-binding protein |
| PAR-CLIP | photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation |
| PCR | polymerase chain reaction |
| poly(A) | Poly-adenaylation |
| R | Arginine |
| RBP | RNA binding protein |
| RGG | arginine/glycine/glycine |
| RIP | RNA immunoprecipitation |
| RNAPII | RNA polymerase II |
| RNP | ribonucleoprotein |
| RRM | RNA recognition motif |
| rRNA | Ribosomal ribonucleic acid |
| (RT)-qPCR | (Reverse transcription)- quantitative polymerase chain reaction |
| S | Serine |
| *S. cerevisiae* | *Saccharomyces cerevisiae* |
| snRNP | small nuclear RNP |

| | |
|---|---|
| SR | serine-arginine |
| T | Threonine |
| TAP | tandem affinity purification |
| TEV | Tobacco etch virus |
| TREX | transcription and export |
| V | Valine |
| WB | Western blot |
| Y | Tyrosine |
| YPD | Yeast extract, peptone, dextrose |
| ZnF | zinc finger domain |

# 12 Publications

**The SR-like protein Npl3 transfers mRNP components from the transcription site 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

2021 (manuscript in process)

# 13 Danksagung

Zunächst geht mein Dank an Katja Sträßer, die mir die Möglichkeit und das Vertrauen entgegengebracht hat in ihrem Labor zu arbeiten, das spannende und herausfordernde Thema bearbeiten zu dürfen sowie die Unterstützung und produktiven Diskussionen.

Mein Dank geht selbstverständlich auch an meine Zweitgutachterin Prof. Dr. Sandra Hake sowie den anderen Mitgliedern des Prüfungskomitees Apl. Prof. Dr. Elena Evguenieva-Hackenberg und Prof. Dr. Lienhard Schmitz für die durchaus zeitraubende Tätigkeit.

Des Weiteren möchte ich mich bei meinen Kollaborationspartnern für deren wertvollen Beitrag zu dieser Arbeit bedanken: Henning Urlaub, Alexander Wulf, Monika Raabe und Ivan Silbern für die unfassbar zeitraubenden Messungen und Auswertung der MS-Daten, Michael Sattler und Nitin Kachariya für die diversen NMR-Analysen, Janine Altmüller, Kathi Zarnack und Samira Reuscher für die Sequenzierung und Auswertung der Transkriptom-Daten.

Auch wenn nicht jeder hier aufgeführt wurde, geht auch ein großer Dank an all die ehemaligen und derzeitigen Mitglieder der gesamten Biochemie. Sei es für die nette Arbeitsatmosphäre, die die vielen Stunden im Labor durchaus angenehm machen konnten, aber auch die konstruktiven Diskussionen, die das Projekt vorangebracht haben. Ihr habt dazu beigetragen, dass ich diese Promotion in den Guten wie auch in den frustrierenden Zeiten schlussendlich vollbracht habe. Ein extra Dankeschön geht hierbei an Kristin, die all die RIPs für dieses Projekt gemacht hat. Bei Heike möchte ich mich besonders für all ihre Unterstützung im Labor bedanken. Ohne dich hätte ich wohl jedes Wochenende durcharbeiten müssen. Ganz vielen Dank geht auch an Peter und Wolle, dass ich jederzeit zu euch ins Büro kommen konnte, wenn es um jedwede Unterstützung bedurfte. Sei es die Reparatur von Geräten jeglicher Art, der gemeinsamen Einführung der Gibson-Klonierung im Labor (yeay!) oder Hilfe bei Computerproblemen aller Art. Bei Cornelia bedanke ich mich für die vielen konstruktiven Beiträge - insbesondere hinsichtlich der schriftlichen Arbeit!

Meiner Familie möchte ich an dieser Stelle ebenfalls meinen Dank aussprechen. Ohne meine Eltern hätte ich wohl nie mit dem Studium angefangen und somit auch nicht bis zu diesem Punkt geschafft. Danke, dass ihr mich über die Jahre unterstützt und stets an mich geglaubt habt.

Zum Schluss möchte ich noch meinen ganz besonderen und herzlichen Dank an Birte richten, die das gesamte Projekt über die Jahre mit getragen (und ertragen) hat. Sie war jederzeit für mich da und hat mich sowohl in privater als auch wissenschaftlicher Hinsicht unterstützt wo sie konnte. Gerade die letzten beiden Jahre waren zuweilen sehr kräftezehrend und ich habe vieles von ihr unbewusst vorausgesetzt, das nicht immer selbstverständlich war. Nichtsdestotrotz warst du stets für mich da. DANKE!

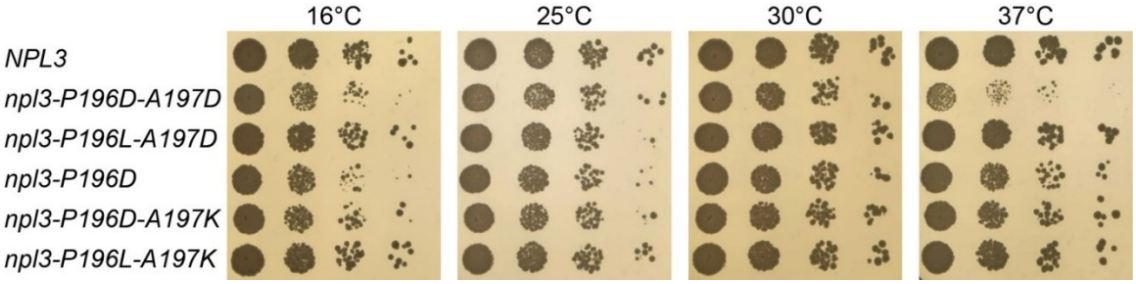
# 14 Eidesstattliche Erklärung

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.

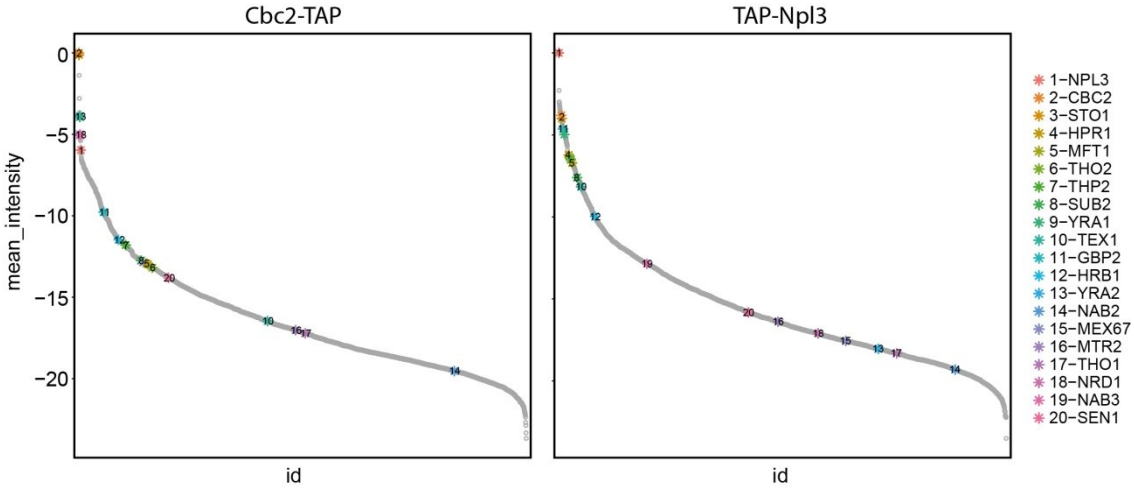
Gießen, den: ..... Unterschrift: .....

(Vorname, Nachname)

# 15 Appendix




**Supplementary figure 1:** Mutation in the *npl3-Linker* mutant has to carry a negative charged side chain to induce growth defects at position P197 and is not rescued by positive charged side chain at position A197. A197 substituted to aspartic acid, thus a negative charged side chain, enhances effect of P196D.




**Supplementary figure 2: Proteins ranked according to their intensity in MS quantification.** All proteins identified and quantified by MS in TAP-purifications of Cbc2 and Npl3. MS quantification and data analysis performed by Monika Raabe and Ivan Silbern.

**Supplementary table 1:** Identified peptides and amino acids of different mRNP components that crosslink *in vivo* to RNA.

| Protein | Cross-linked peptide | Cross-linked amino acid |
|---|---|---|
| Npl3 | 127-LFVRPFPLDVQESELNEIFGPFGPMK-152 | 127-152 (RRM1) |
| Npl3 | 182-SFANQPLEVVYSKLP(*)AK-198 | P196 |
| Npl3 | 182-SFANQPLEVVYSKLPAK(*)-198 | K198 |
| Npl3 | 273-DDNPPP(*)IR-280 | P278 |
| Npl3 | 273-DDNPPPIR(*)R-281 | R280 |
| Npl3 | 385-GGYDGPRGDYGPPRDAYR(*)TR-404 | R402 |
| Npl3 | 371-NDYGP(*)PR-377 | P375 |
| Npl3 | 392-GDYG(*)P(*)PR-398 | G395 and P396 |
| Sto1/Cbp80 | 7-R(*)GDFDEDENYR-17 | R7 |
| Sto1/Cbp80 | 18-DF(*)RPR-22 | F19 |

| | | |
|---|---|---|
| Sto1/Cbp80 | 18-DFRP(*)R-22 | P21 |
| Sto1/Cbp80 | 18-DFRPRMP(*)K-25 | P24 |
| Sto1/Cbp80 | 107-NNVAGKSIINYFFEELQK-124 | K112 or S113 |
| Sto1/Cbp80 | 578-K(*)NDLYFR-584 | K578 |
| Cbc2/Cbp20 | 17-RL(*)DTPSR-23 | L18 |
| Cbc2/Cbp20 | 31-R(*)NPNGLQELR-40 | R31 |
| Cbc2/Cbp20 | 134-GK(*)SGGQVSDELR-145 | K135 |
| Cbc2/Cbp20 | 146-FDFDASRGGFAIPFAER-162 | 146-162 |
| Cbc2/Cbp20 | 163-VGVP(*)HSR-169 | P166 |
| Tho2 | 1183-LPSSALIGHLK-1193 | G1190 or H1191 |
| Tho2 | 1357-TLIQNPQNP(*)DFAEK-1370 | P1365 |
| Hpr1 | 534-IP(*)TGLDK-540 | P535 |
| Mft1 | 189-YRIYDDFSK-197 | I191 or Y192 |
| Hrb1 | 51-FADTY(*)RGSR-59 | Y55 |
| Hrb1 | 137-GDYGP(*)LLAR-145 | P141 |
| Hrb1 | 190-ADIITSRGHHR-200 | S195 or R196 |
| Sub2 | 53-K(*)GSYVGIHSTGFK-65 | K53, G54, Y56 |
| Sub2 | 53-KG(*)SY(*)VGIHSTGFK-65 | K53, G54, Y56 |
| Sub2 | 54-GSYVGIHST(*)GFKDFLLKPELSR-75 | T62 |
| Sub2 | 66-DFLLKP(*)ELSR-75 | K70 |
| Sub2 | 141-ELAY(*)QIR-147 | Y144 |
| Sub2 | 162-TAVFYGGTP(*)ISK-173 | P170 |
| Sub2 | 180-NKDTAPHIVVATPGR-194 | N180 or K181 |
| Sub2 | 195-LK(*)ALVR-200 | K196 |
| Sub2 | 197-ALVR(*)EK-202 | R200 |
| Sub2 | 201-EKYIDLSHVK-210 | E201 or K202 |
| Sub2 | 203-Y(*)IDLSHVK-210 | Y203 |
| Sub2 | 211-NFVIDEC(*)DK-219 | C217 |
| Sub2 | 228-RDVQEIFR-235 | R228 or D229 |
| Sub2 | 228-RDVQEIFR(*)ATPR-239 | R235 |
| Sub2 | 260-RF(*)LQNPLEIFVDDEAK-275 | F261 |
| Sub2 | 315-STTRANELTK-324 | T316 or T317 |
| Sub2 | 325-LLNASNFPAITVHGHMK-341 | G338 or H339 |

| | | |
|---|---|---|
| Sub2 | 351-AFKDF(*)EK-357 | F355 |
| Sub2 | 359-IC(*)VSTDVFGR-368 | C360 |
| Sub2 | 369-G(*)IDIER-374 | G369 |
| Sub2 | 375-INLAINYDLTNEADQYLH(*)R-393 | H392 |
| Sub2 | 404-G(*)LAISFVSSK-413 | G404 |
| Sub2 | 414-EDEEVLAK(*)IQER-425 | K421 |
| Yra1 | 8-SLDEIIG(*)SNKAGSNR-22 | G14 and K17 or A18 |
| Yra1 | 8-SLDEIIGSNKAG(*)SNR-22 | G19 |
| Yra1 | 25-V(*)G(*)G(*)T(*)R(*)GNGPR-34 | 25-29 |
| Yra1 | 36-VGK(*)QVGSQR-44 | K38 |
| Yra1 | 45-R(*)SLPNR-50 | R45 |
| Yra1 | 51-RGP(*)IR-55 | P53 |
| Yra1 | 57-NTR(*)APPNAVAR-67 | R59 |
| Yra1 | 57-NTRAPPNAVAR-67 | P62 or N63 |
| Yra1 | 60-APPNAVAR(*)VAK-70 | R67 |
| Yra1 | 68-VAK(*)LLDTTR-76 | K70 |
| Yra1 | 80-VNVEGLPRDIK-90 | P86 or R87 |
| Yra1 | 108-VLLSYNER(*)GQSTGMANITFK-127 | R115 |
| Yra1 | 139-FNGSPIDGGRSR-150 | 146-GGRSR-150 |
| Yra1 | 153-LNLIVDP(*)NQR(*)PVK-165 | P159 and R162 |
| Yra1 | 153-LNLIVDPNQRP(*)VK-165 | P163 |
| Yra1 | 178-GGNAP(*)RPVK-186 | P182 |
| Yra1 | 210-KSLEDLDK-217 | K210 or S211 |
| Yra2 | 38-LGFAPSDAASR(*)SK-50 | R48 |
| Tho1 | 75-KEVSSEPK-82 | 75-82 |
| Tho1 | 129-ALDLLNK(*)K-136 | K135 |
| Tho1 | 140-ANKF(*)GQDQADIDSLQR-155 | F143 |
| Tho1 | 181-KNEPESGNNGKFK-193 | G190 or K191 or F192 |
| Tho1 | 214-SGYRR-218 | G215 or Y216 or R217 |
| Nab2 | 210-GGGAVGK(*)NR-218 | K216 |
| Nab2 | 230-NNNSTRFNPLAK(*)-241 | K241 |

| | | |
|---|---|---|
| Nab2 | 314-TREEFQK-320 | 314-320 |
| Nab2 | 322-KADL(*)LAAK-329 | L325 |
| Nab2 | 322-KADLLAAK(*)R-330 | K329 |
| Nab2 | 394-EVKPISQK(*)K-402 | K401 |
| Nab2 | 403-AAPPP(*)VEK-410 | P407 |
| Nab2 | 411-SLEQC(*)K-416 | C415 |
| Nab2 | 417-FGTHC(*)TNK-424 | C421 |
| Nab2 | 433-SHIMC(*)R-438 | C437 |
| Sac3 | 460-ALSH(*)TLNK-467 | H463 |
| Sac3 | 527-TYLTC(*)LER-534 | C531 |
| Sac3 | 862-NLIFSPVNDEFNK(*)FATHLTK-881 | K874 |
| Sac3 | 1084-YDK(*)TLR-1089 | K1086 |
| Sac3 | 1116-KMLEKEK(*)-1122 | K1122 |
| Cdc31 | 51-ALGFELP(*)KR-59 | P57 |
| Cdc31 | 113-ISIK(*)NLR-119 | K116 |
| Cdc31 | 121-VAK(*)ELGETLTDEELR-135 | K123 |
| Mex67 | 428-YNH(*)GYNSTSNNK-439 | H430 |
| Pab1 | 307-QYEAYRLEK-315 | Y308 or Y311 |

**Supplementary table 2:** Generated *npl3* point mutants; including pre-screen (6.3).

| Mutation | Growth | Expression level of Npl3 |
|---|---|---|
| F128I | Like wild type | n.a. |
| F128Y | Like wild type | n.a. |
| F128E | Like wild type | n.a. |
| F128A | Like wild type | n.a. |
| F160Y,F162Y, F165Y | Lethal 37°C, slow growth at 16, 25 and 30°C | Like wild type |
| F160Y | Like wild type | Like wild type |
| F162Y | Lethal 37°C, slow growth at 16, 25 and 30°C | Like wild type |
| F165Y | Like wild type | Like wild type |
| F160Y,F162Y | Lethal 37°C, slow growth at 16, 25 and 30°C | Like wild type |
| F160Y,F165Y | Like wild type | Like wild type |
| F162Y,F165Y | Lethal 37°C, slow growth at 16, 25 and 30°C | Like wild type |

| | | |
|---|---|---|
| F160A,F162A, F165A | Lethal at 16 and 37°C, slow growth at 25 and 30°C | Strongly reduced |
| F160V,F162I | Lethal 37°C, slow growth at 16, 25 and 30°C | n.a. |
| P196D | Slight growth defect at 16 and 37°C | Like wild type |
| A197D | Like wild type | Like wild type |
| P196D,A197D | Slow growth at 16 and 37°C, slightly slower at 25 and 30°C | Like wild type |
| P196L,A197D | Like wild type | Like wild type |
| P196D,A197K | Slight growth defect at 16 and 37°C | Like wild type |
| P196L,A197K | Like wild type | Like wild type |
| F229E | Like wild type | n.a. |
| F229A,R235A | Like wild type | Like wild type |
| F245Y | Lethal 37°C, slow growth at 16°C, slight defect at 25 and 30°C | Like wild type |
| F245I | Lethal 37°C, slow growth at 16°C, slight defect at 25 and 30°C | Like wild type |
| P196D,A197D,F162Y | Lethal 37°C, slow growth at 16, 25 and 30°C | Like wild type |
| P196D,A197D,F245Y | Lethal | n.a. |
| F162Y,F245Y | Lethal | n.a. |
| P196D,A197D, F162Y,F245Y | Lethal | n.a. |

**Supplementary table 3:** Data that were generated in collaboration

| page | generated data (method) | performed by / in collaboration with |
|---|---|---|
| 46 | Identification of *in vivo* RNA-protein crosslinks ($RNP^{XL}$) | Henning Urlaub[1], Alexander Wulf[1] |
| 54 | RNA-binding to *npl3* mutants *in vivo* (RIP) | Kristin Hühn[2] |
| 55 | Structural integrity of truncated *npl3* mutants (NMR) | Michael Sattler[3,4], Nitin Kachariya[3,4] |
| 57 | RNA-binding to *npl3* mutants *in vitro* (NMR) | Michael Sattler[3,4], Nitin Kachariya[3,4] |
| 58 | RNA-binding to *npl3* mutants *in vitro* (ITC) | Michael Sattler[3,4], Nitin Kachariya[3,4] |
| 60 | Analysis of RNA-seq data (RNA-Seq / bioinformatic analysis) | Samira Reuscher[5], Kathi Zarnack[5], Janine Altmüller[6] |
| 68 | MS quantification of Cbc2-TAP purification (MS) | Henning Urlaub[1], Monika Raabe[1], Ivan Silbern[1] |
| 70 & 71 | MS quantification of TAP purified Npl3 (MS) | Henning Urlaub[1], Monika Raabe[1], Ivan Silbern[1] |
| 72 | RNA binding of RBPs that showed a reduction in nuclear mRNPs in *npl3* Linker cells (RIP) | Kristin Hühn[2] |
| 73 | RNA binding of Nab2 in *npl3-RRM2* (RIP) | Kristin Hühn[2] |
| 112 (supp. figure) | Protein ranking according to their intensity in MS quantification (MS) | Henning Urlaub[1], Monika Raabe[1], Ivan Silbern[1] |
| 112-116 (supp. table) | Identified *in vivo* RNA-protein cross (MS) | Henning Urlaub[1], Monika Raabe[1], Ivan Silbern[1] |

[1] Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Goettingen; University Medicine, Institute for Clinical Chemistry, Robert-Koch-Strasse 40, 37075 Göttingen, Germany

[2] Institute of Biochemistry, FB08, Justus Liebig University Giessen, Heinrich-Buff-Ring 17, 35392 Giessen, Germany

[3] Bavarian NMR Center (BNMRZ), Department of Chemistry, Technical University of Munich, Lichtenbergstrasse 4, 85748 Garching

[4] Institute of Structural Biology, Helmholtz Center Munich, 85764 Neuherberg, Germany

[5] Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, Max-von-Laue-Straße 15, 60438 Frankfurt a.M., Germany

[6] Cologne Center for Genomics (CCG), University of Cologne, 50931 Cologne, Germany