Dissertation zur Erlangung des Doktorgrades der Fakultät für Chemie und Pharmazie der Ludwig-Maximilians-Universität München

The Prp19 complex is a novel transcription elongation factor required for TREX occupancy at transcribed genes

Sittinan Chanarat

aus Chiangmai, Thailand

2011

Erklärung

Diese Dissertation wurde im Sinne von § 13 Abs. 3 bzw. 4 der Promotionsordnung vom 29. Januar 1998 (in der Fassung der sechsten Änderungssatzung vom 16. August 2010) von Herrn Prof. Dr. Patrick Cramer betreut.

Ehrenwörtliche Versicherung

Diese Dissertation wurde selbstständig, ohne unerlaubte Hilfe erarbeitet.

München, den 07.06.2011

(Sittinan Chanarat)

Dissertation eingericht am 07.06.2011

1. Gutachter: Prof. Dr. Patrick Cramer

2. Gutachter: Dr. Dietmar Martin

Mündliche Prüfung am 20.07.2011

Summary

Gene expression is the process by which the information encoded in a gene is used to direct the synthesis of a protein. Several steps in the gene expression, including transcription, mRNA processing, nuclear mRNA export, translation and mRNA degradation are regulated and closely linked. In the budding yeast *S. cerevisiae* the conserved TREX (transcription/export) complex travels along the transcribed gene with RNA polymerase II (RNAPII) and couples transcription to nuclear mRNA export. Even though the TREX complex has been identified and studied extensively during the last decade, it is still unclear how the complex is recruited to the active transcription machinery.

Here I show that the Prp19 complex, a conserved non-snRNP spliceosome component essential for splicing, genetically and biochemically interacts with the TREX complex. The Prp19 complex is recruited to transcribed genes, interacts with RNAPII as well as the TREX complex and is absolutely required for the occupancy of the TREX complex at transcribed genes. Moreover, the C-terminus of Syf1, one of essential components of the Prp19 complex, is necessary for the occupancy of the Prp19 complex at transcribed genes as well as the interaction of the Prp19 complex with RNAPII. Importantly, the Prp19 complex is necessary for full transcriptional activity both *in vitro* and *in vivo*. Taken together, I identify the Prp19 splicing complex as a novel transcription elongation factor which is essential for TREX occupancy at transcribed genes and thus provide a novel link between transcription and messenger ribonucleoprotein (mRNP) formation.

1. Introduction

In the past decades, many studies have revealed that several steps in gene expression are intimately connected and form a complex network. The close coupling processes within this sophisticated system include transcription, capping, splicing, polyadenylation, cleavage, mRNA maturation, mRNA export, translation and ultimately mRNA degradation. To ensure efficient and accurate gene expression of the genetic information, the interplay between those procedures has to be highly regulated in living cells. (for review see Maniatis and Reed 2002; Proudfoot et al. 2002; Reed and Hurt 2002; Cole and Scarcelli 2006).

1.1 Transcription

Transcription is the first step leading to gene expression. It is the process of creating a complementary RNA copy of a DNA sequence. During transcription, a sequence of DNA is transcribed by RNA polymerases. In eukaryotes, when the gene transcribed encodes a protein, the result of transcription is messenger RNA (mRNA) which is transcribed by a specific RNA polymerase known as RNA polymerase II (RNAPII).

Transcription by RNAPII is facilitated by a large number of transcription factors. Transcription initiation factors control transcription and recruit RNAPII to a promoter. When RNAPII begins to synthesize the mRNA, transcription initiation factors are substituted with transcription elongation factors (Pokholok et al. 2002). Transcription elongation factors enhance the rate of transcription and/or increase the processivity, the ability of elongating RNAPII to travel the entire length of the gene, by facilitating chromatin passage and mRNA processing (Hirose and Manley 2000; Orphanides and Reinberg 2000;

Orphanides and Reinberg 2002; Ahn et al. 2004; Mason and Struhl 2005; Perales and Bentley 2009).

1.2 Transcription, nuclear mRNA export and TREX

Several genetic and biochemical approaches showed the existence of an extensive network between the factors which couple transcription with other processes, *e.g.* nuclear mRNA export. A key player in coupling transcription and nuclear mRNA export is the highly conserved TREX complex (Strasser et al. 2002). In budding yeast, TREX is composed of the heterotetrameric THO subcomplex, the two mRNA export proteins Sub2 and Yra1, the two SR-like proteins Gbp2 and Hrb1 and a protein of unknown function Tex1. THO, consisting of Tho2, Thp2, Mft1 and Hpr1, is associated with actively transcribed genes and is necessary for efficient transcription elongation *in vivo*. (Chavez and Aguilera 1997; Piruat and Aguilera 1998; Chavez et al. 2000; Strasser et al. 2002). Even though THO components are not essential in yeast, the lack of them leads to a DNA/RNA hybrid formation, so-called R-loop, which causes an elongation block for the next elongating RNAPII (Huertas et al. 2006). The THO complex has also been implicated in a crucial network between transcription, RNA quality control and genome stability (Jimeno et al. 2002) as THO mutants show transcription-dependent hyperrecombination. This hyperrecombination phenotype might be a secondary effect of a defect in transcription elongation which causes exposure of the transcribed DNA template to the recombination machineries (Reed 2003). Importantly, THO is also necessary to load the mRNA export proteins Sub2 and Yra1 co-transcriptionally onto the nascent mRNA for its subsequent nuclear export (Strasser et al. 2002; Abruzzi et al. 2004). Sub2 is a DECD-box RNA helicase which is involved in both splicing and mRNA export. Sub2

interacts with and recruits the mRNA export factor Yra1, which is able to bind directly to the mRNA (Portman et al. 1997), to the nascent transcript (Strasser and Hurt 2001). The SR-like proteins Gbp2 and Hrb1 associate with TREX and with actively transcribed genes. Even though their involvement in mRNA export has not been elucidated, they interact with and are exported together with nascent mRNA to the cytoplasm in a THO-dependent manner. (Hacker and Krebber 2004; Hurt et al. 2004).

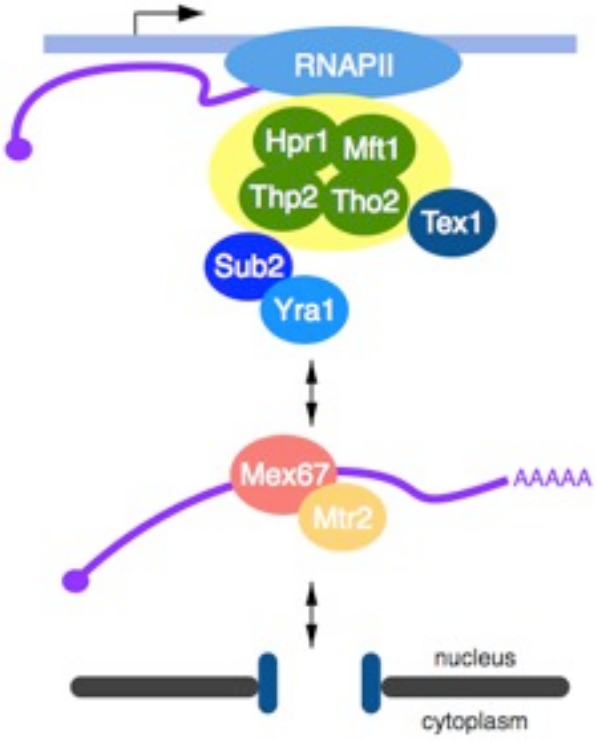

Figure 1.1: TREX couples transcription and mRNA export
TREX consists of the heterotetrameric THO complex, comprising Tho2, Hpr1, Mft1, and Thp2, the nuclear mRNA export factors Sub2 and Yra1, the SR-like proteins Gbp2 and Hrb1, and Tex1. TREX travels along the transcribed gene with RNAPII, interacts with the nascent mRNA and directs it into the mRNA export pathway via the mRNA export factors Sub2 and Yra1. Yra1 then directly binds the mRNA exporter Mex67-Mtr2, which in turn binds directly to the mRNA as well as to nuclear pore proteins and transports the mRNP through the nuclear pore complex to the cytoplasm

1.3 Splicing and recruitment of the TREX complex

Pre-mRNA splicing is an important process and plays a crucial role in the regulation of gene expression. The splicing reaction is catalyzed by spliceosome, which is composed of five small nuclear RNAs (snRNAs), U1, U2, U4, U5, and U6 in the forms of small nuclear ribonucleoprotein (snRNP) complexes (Will and Lührmann 1997, 2006; Staley and Guthrie 1998; Burge et al. 1999; Wahl et al. 2009). Spliceosome assembly is a stepwise process which involves sequential interaction of snRNPs and other protein components.

The spliceosome assembly begins with the recognition of the 5'-splice site by U1 snRNP and the polypyrimidine tract by the U2 snRNP auxiliary factor (U2AF), which initiates the formation of an ATP-independent early complex (E complex). The interaction of U2AF to the polypyrimidine tract recruits the U2 snRNP to the branch point sequence, forming the A complex. Next, the B complex is formed by the association of the U4/U6.U5 tri-snRNP. Subsequently, a series of protein/RNA rearrangements alter the B complex to the C complex, an active and transient spliceosome intermediate. These short-lived spliceosomal complexes can be distinguished by their snRNP compositions as shown in Figure 1.2.

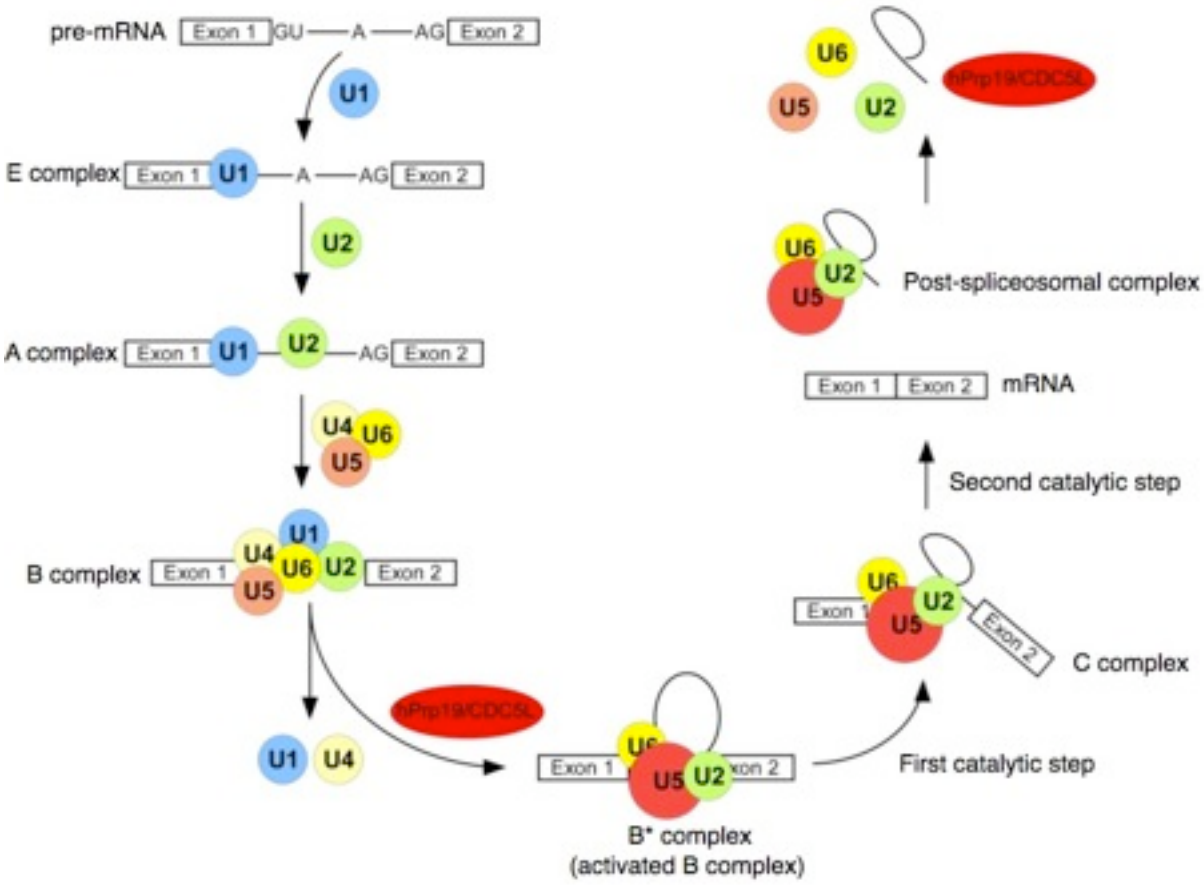

Figure 1.2: The stepwise spliceosome assembly

During spliceosome assembly, firstly the U1 snRNP associates with the pre-mRNA forming the E complex. Then the U2 snRNP binds and forms the A complex. Recruitment of the pre-assembled U4/U6.U5 tri-snRNP leads to formation of the pre-catalytic B complex. By structural rearrangements, the U1 and U4 snRNPs dissociate and incorporation of the Prp19 complex leads to remodeling of U5 and formation of the activated B complex. The C complex is formed after the first step of splicing occurs. Subsequently, the second step of splicing takes place. Finally, the mature mRNA and the post-spliceosomal complex are released as well as the splicing factors are recycled.

Even though the gene expression machineries are highly conserved from yeast to metazoans, a fundamental difference between these organisms is that most yeast genes lack introns whereas the converse is true in higher

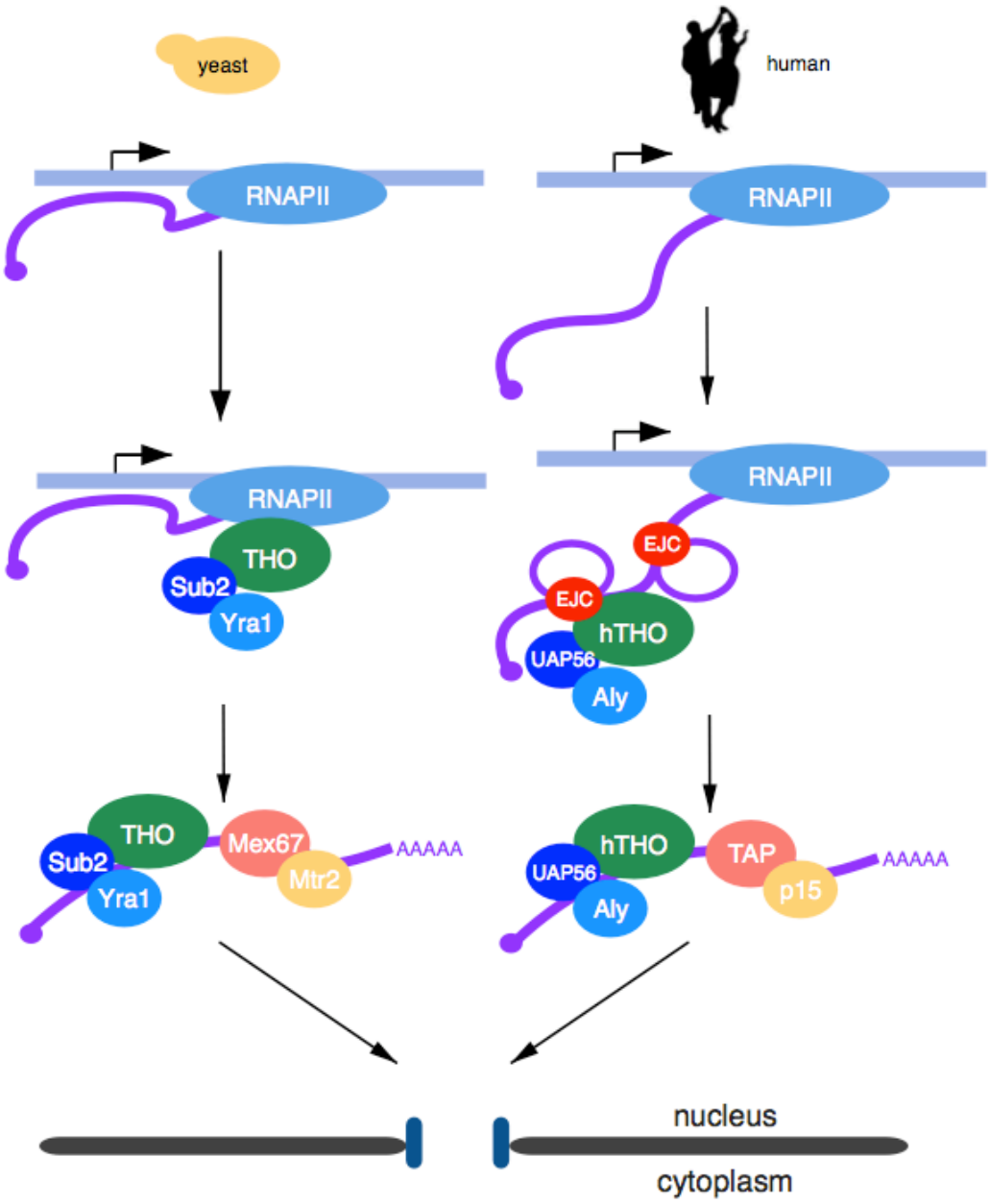

Figure 1.3: Model of mRNA export in yeast and human

In the yeast *S. cerevisiae*, mRNA export is coupled to transcription (left panel), whereas in higher eukaryotes it is coupled to splicing in a cap-dependent manner (right panel). In yeast, the THO subcomplex is recruited to the nascent mRNA during ongoing transcription and recruits the mRNA export factors Sub2 and Yra1. Then Mex67-Mtr2 binds to the mRNP through the interaction with Yra1 and transports the mRNP through the nuclear pore complex (left panel). In higher eukaryotes, the TREX complex, together with the exon-exon junction complex (EJC), is recruited to the mRNA during splicing (right panel). As in yeast, the mRNA is exported to the cytoplasm by the interaction of Aly, the human homolog of Yra1, with Tap-p15, the human homolog of yeast Mex67-Mtr2.

organisms. In human cells, where most of protein-encoding genes contain introns and have to be spliced, it has been shown that TREX complex is recruited to transcribed genes during splicing instead of transcription. *In vitro* studies of spliceosome assembly and subsequent proteomic analyses revealed that human TREX associates with the C complex (Wahl et al. 2009). However, in yeast, the TREX complex appears to be recruited by the transcription machinery in a splicing-independent manner (Figure 1.3). Even though the function of the TREX complex is conserved, its recruitment is not. It has been a longstanding question how the TREX complex is recruited to transcribed genes and which factors are required for this recruitment.

1.4 Splicing and the Prp19 complex

The spliceosome assembly is a strictly ordered process which requires several snRNPs and other splicing factors. A Prp19 complex, or so-called the nineteen complex (NTC), is a non-snRNP splicing complex, which interacts with the spliceosome immediately after the U4 dissociation, and is required for the activation process (Tarn et al. 1993, 1994). The complex is composed of at least eight proteins, *i.e.* Cef1, Clf1, Isy1, Ntc20, Snt309, Syf1, Syf2 and Prp19. Deficiency in Prp19 complex function also results in failure in spliceosome recycling due to impaired U4/U6 biogenesis (Chen et al. 2006).

Among Prp19 complex components, *SYF1* was identified as "synthetic lethal with cdc forty" in a yeast genetic screen for genes that interact with *CDC40/PRP17*, which is a splicing factor important for the catalytic step and plays a role in cell cycle progression (Ben-Yehuda et al. 2000). The Syf1 protein is an essential gene and highly conserved throughout the eukaryotic organisms. It contains fifteen copies of the tetratricopeptide repeat (TPR) motif, which is thought to be a scaffold domain for protein-protein interactions (for review, see D'Andrea and Regan 2003). Two-hybrid analysis revealed that Syf1 interacts with most identified Prp19 complex components, indicating that Syf1 may serve as a platform in maintaining the integrity of Prp19 complex (Chen et al. 2002).

1.5 XAB2 in human

Interestingly, Syf1 has a homolog in human known as XPA-binding protein 2 (XAB2) which also contains TPR motifs. It has been shown that XAB2 is a multifunctional factor involved in pre-mRNA splicing, transcription and transcription-coupled repair (TCR) (Kuraoka et al. 2008). XAB2 has been

identified firstly by virtue of its ability to interact with Xeroderma Pigmentosum group A protein (XPA), a protein product of the gene that causes a disease with DNA-damage-processing defects (Nakatsu et al. 2000). Moreover, XAB2 interacts with TCR-specific factors Cockayne Syndrome group A and B proteins (CSA and CSB) and RNAPII and is involved in transcription and TCR. In human cells, knockdown of XAB2 gene results in an ultraviolet hypersensitivity and a decrease in mRNA synthesis. XAB2 forms a hexameric protein complex consisting of XAB2(Syf1), hPRP19(Prp19), hISY1(Isy1), CCDC16, hAquarius, and PPIE. Importantly, though, it is still unclear whether these multiple functions of XAB2 are conserved among eukaryotes and whether these similar functions as TREX complex could implicate their interplay in gene expression.

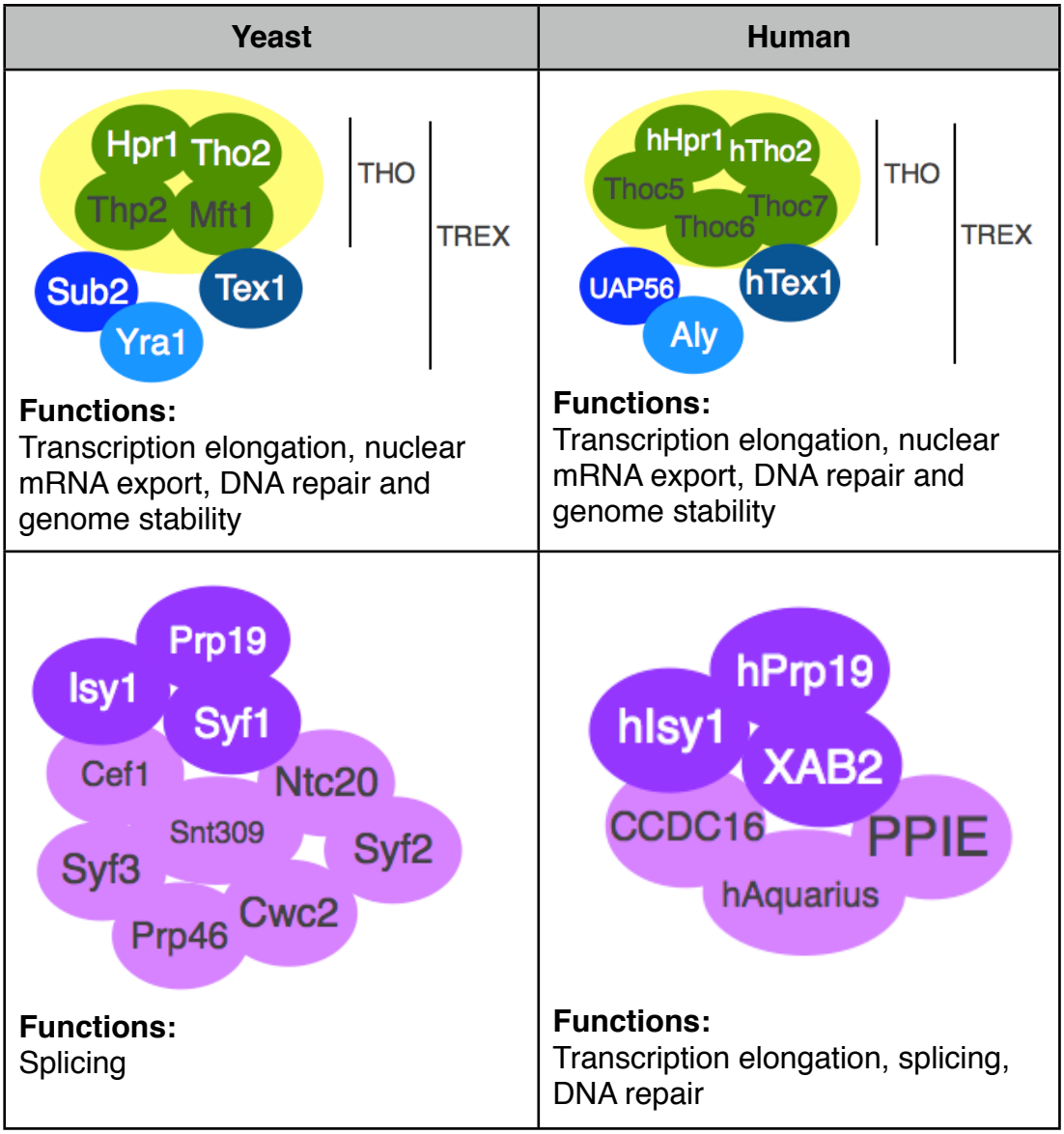

Figure 1.4 : Comparison of yeast and human TREX and Prp19 complex

Components of each complex and known functions are shown. Components with homologous proteins found in both species are written in white, otherwise in black.

Aim of this work

In budding yeast, the conserved TREX complex, which couples transcription to nuclear mRNA export, is recruited to actively transcribed genes and travels along the genes together with RNA polymerase II (RNAPII). However, it is still unclear how TREX is recruited to the transcription machinery and which factors mediate this interaction. Interestingly, Syf1, a component of the Prp19 splicing complex, is homologous to human XAB2 (XPA binding protein 2), a multifunctional protein involved not only in splicing, but also in transcription and transcription-coupled DNA repair (TCR). Since XAB2 and TREX have several similar functions, I hypothesized that Syf1 in yeast also functions in transcription together with TREX.

The aim of this work was to investigate the cellular function of Syf1 particularly in transcription in conjunction with TREX. Firstly, the genetic and biochemical interaction between *SYF1* and the TREX complex should be investigated. Then a possible function of Syf1 in transcription should be assessed. Finally, the functional relationship between Syf1 and TREX in gene expression should be characterized.

The results of this study should provide an insight into a potential novel function of Syf1 and the Prp19 complex in gene expression as well as their functional connection with the TREX complex. This work might eventually provide a basis for further studies in the field of gene expression.

2. Results

2.1 Generation of temperature sensitive alleles of *SYF1*

Since XAB2, the human homolog of Syf1, has several functions similar to the TREX complex in human, it is interesting to investigate whether Syf1 also plays the same roles as XAB2 and has a functional connection with TREX in yeast. As *SYF1* is an essential gene in yeast, one cannot simply knockout the gene to generate a mutant for a study. Thus, I decided to isolate conditional temperature sensitive (ts) alleles of *SYF1* by a random mutagenesis. In Figure 2.1, 10-fold serial dilutions of each ts strain were spotted on YPD, a complete medium for yeast growth, and incubated at 30ºC, 33ºC and 37ºC. Cells carrying one of the five ts alleles grew normally as wild-type cells at the permissive temperature, 30ºC, but were lethal at the restrictive temperature, 37ºC. Interestingly, all alleles grew differently at the intermediate temperature, 33ºC.

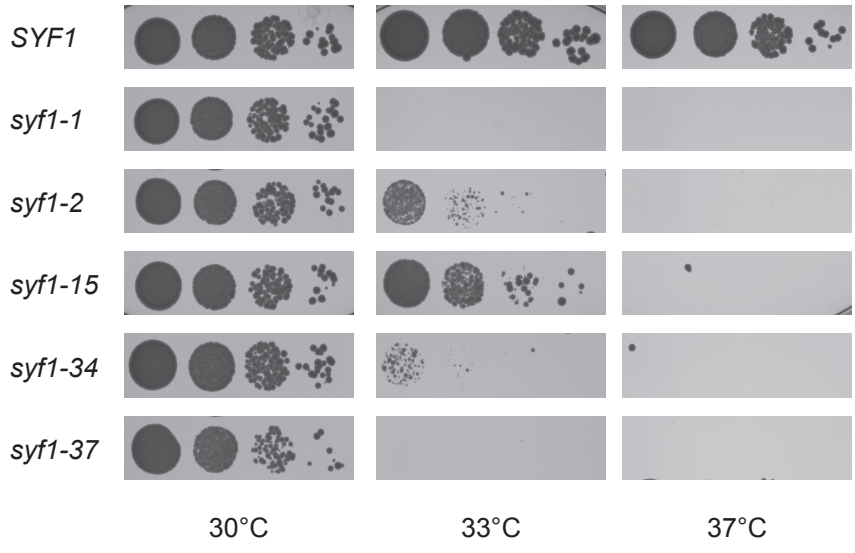

Figure 2.1: Temperature sensitive alleles of *SYF1*
Ten-fold serial dilutions of yeast cells expressing wild-type *SYF*1 or one of five ts alleles of *SYF1* were spotted onto YPD plates and incubated for 2 days at 30°C (permissive temperature), 33°C (intermediate temperature), or 37°C (restrictive temperature), respectively.

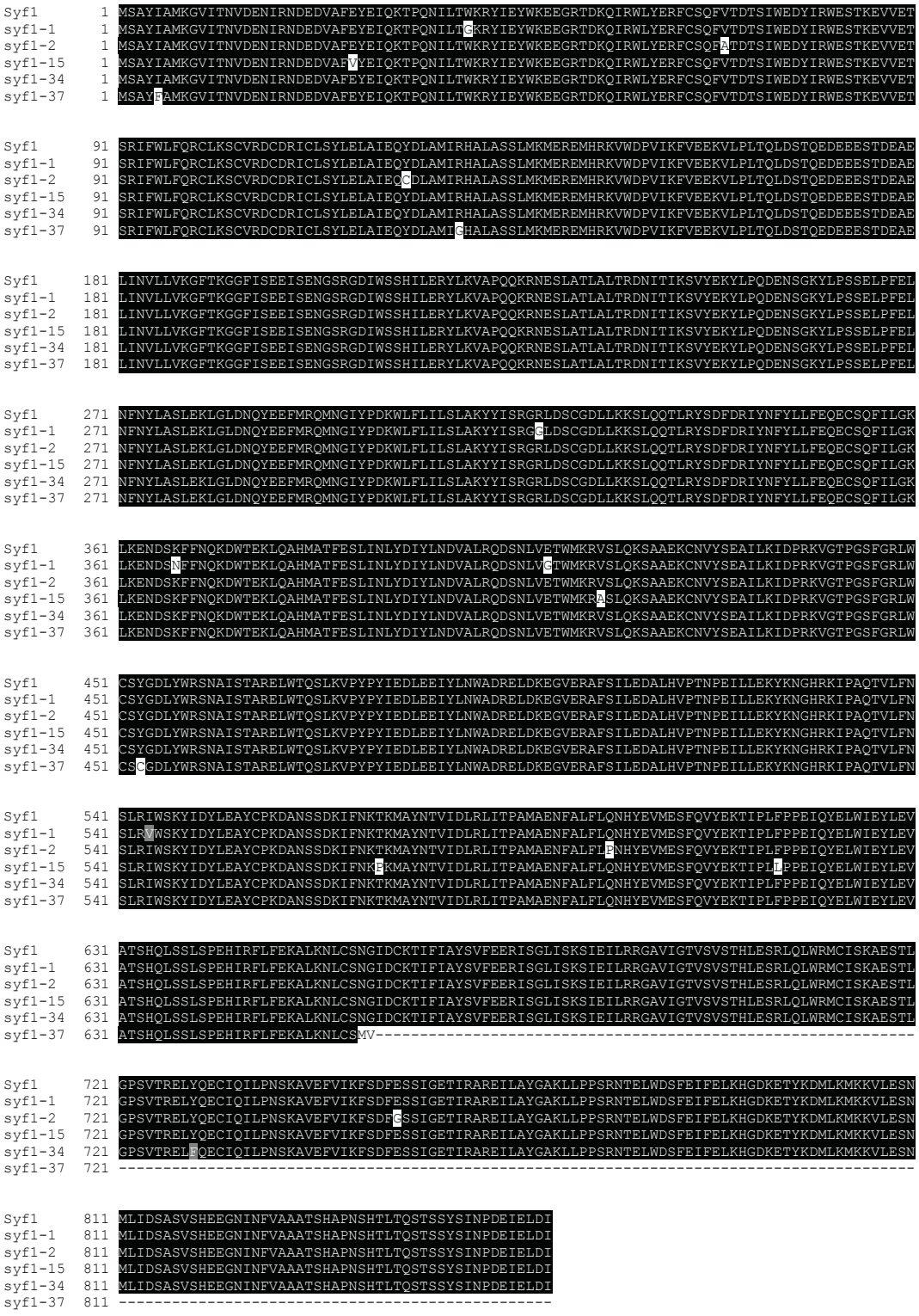

Figure 2.2: An amino acid sequence alignment of Syf1 and its ts mutants
Darkened areas show the residue identity and shaded areas represent mutations.

Then, the plasmids carrying the putative ts alleles were rescued and sequenced in order to identify the mutations. Amino acid alignment of all *syf1* ts alleles is shown in Figure 2.2.

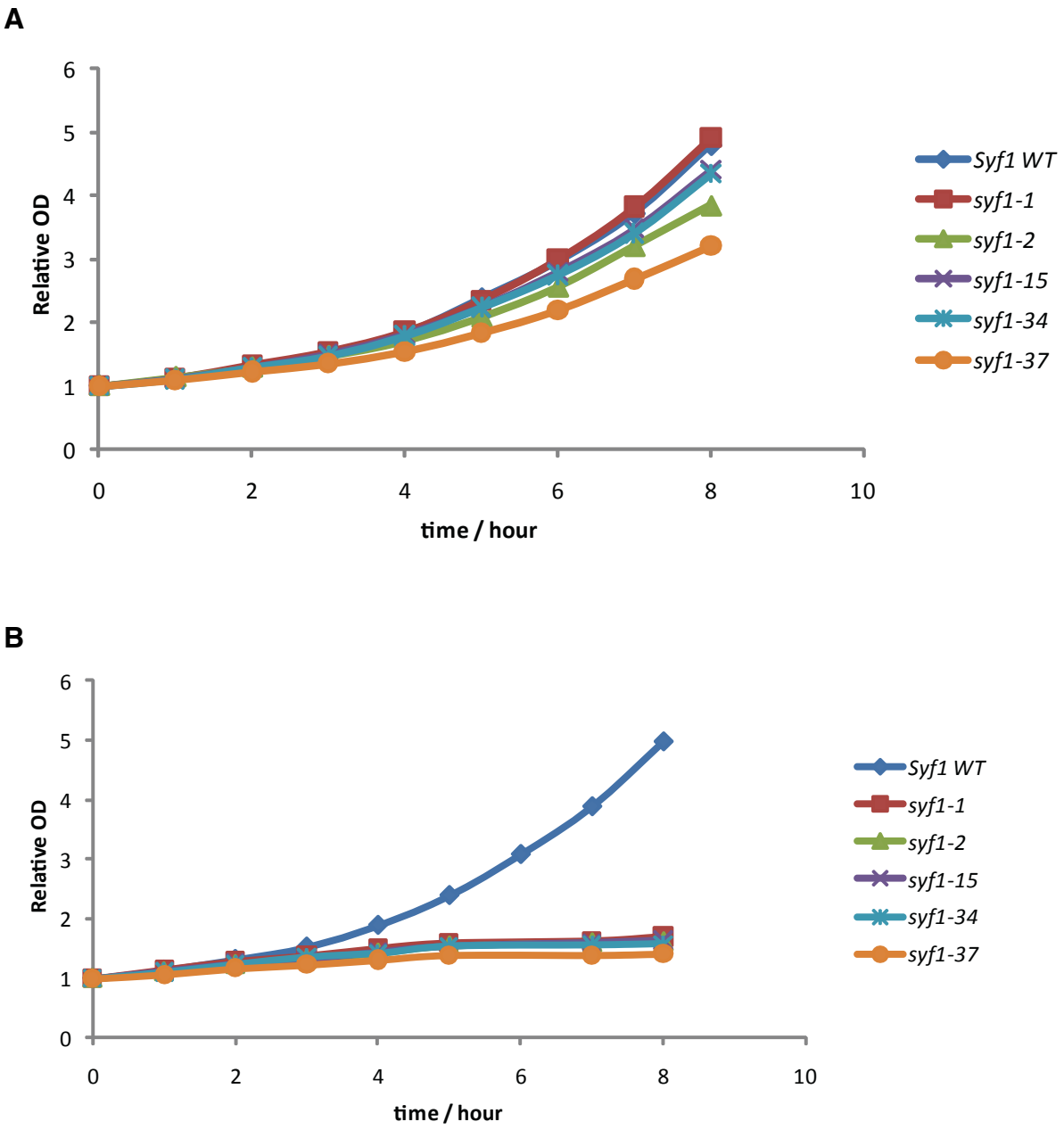

Figure 2.3: Growth curves for *SYF1* and its ts alleles
Cells were grown under permissive, 30ºC (A), and restrictive conditions, 37ºC (B).

In order to acquire growth information of those strains, growth curves were determined at 30ºC and 37ºC as shown in Figure 2.3. Though some mutants

grow slightly slower than wild-type cells, all strains could duplicate and grow logarithmically at the permissive temperature, 30ºC. However, after growth at the non-permissive temperature, 37ºC, for about three hours the ts mutants stop growing.

2.2 Genetic interaction between *SYF1* and the THO complex

2.2.1 The principle of synthetic lethality

For the first indication of a functional overlap of Syf1 and the TREX complex, the ts alleles of *SYF1* and the deletion mutants of the THO complex were used to test for a genetic interaction. Alleles having a synthetic lethal relationship, one of several approaches to assess genetic interaction, can often, but not always, be explained based on the functions of their proteins. For instance, they might be unambiguously redundant with regard to an essential function, be two subunits of an essential multi-protein complex, be two interconnected factors in an important linear pathway, or involve in parallel pathways that are together essential for live. By exploring a synthetic lethal relationship, it may help address questions about how cellular processes work when the protein products expressed by two different genes have an effect together but not separately.

Generally, two alleles are synthetically lethal when the combination of two viable mutations in these genes leads to cell death. As one approach to address the synthetic lethality between two genes, *X* and *Y*, a double knockout strain of the two genes of interest that carries the wild-type copies of both proteins on *URA3* plasmids (double shuffle strain) is generated. These strains are transformed with combinations of plasmids carrying either wild-type copy of gene *X* or *Y*, or mutated *x* or *y*. The wild-type copies of both genes

encoded by the *URA3* plasmids are then removed by streaking the cells onto plates containing 5-fluoroorotic acid (5-FOA). If cells carrying both mutated genes are lethal on 5-FOA plate, the mutant alleles are synthetically lethal (Figure 2.4).

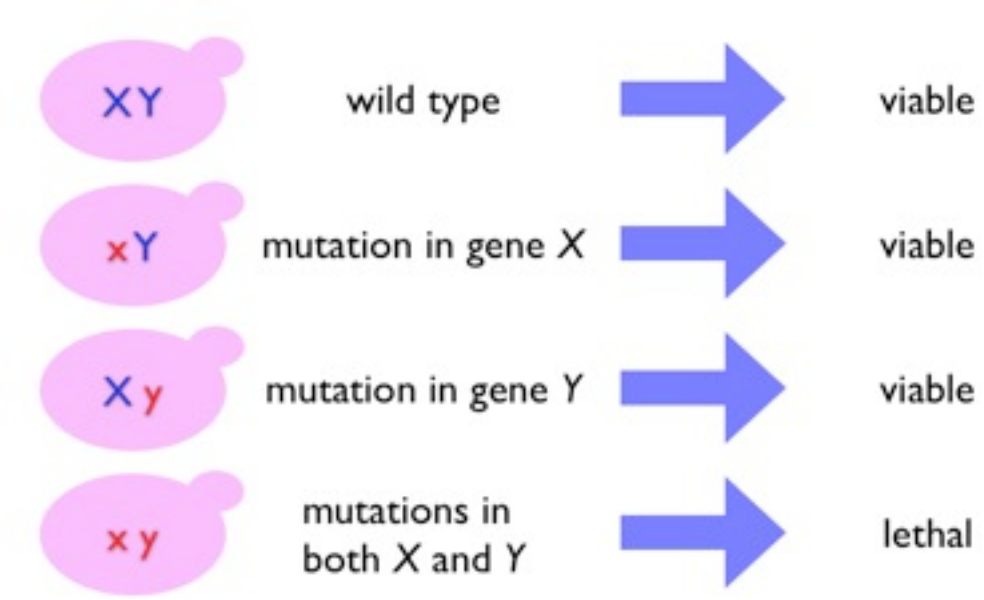

Figure 2.4: Principle of synthetic lethality
Two genes are synthetic lethal only when their simultaneous mutation results in cellular or organismal death, *i.e.*, mutation of either gene *X* or gene *Y* does not affect viability whereas mutation of both genes at the same time is lethal.

2.2.2 *SYF1* is synthetic lethal with *HPR1* and *MFT1*

As XAB2, a human homolog of the yeast *SYF1*, has several functions similar to the TREX complex in higher eukaryotes, I hypothesized that *SYF1* in yeast plays an overlapping role with the TREX complex. To test this hypothesis, I assessed the genetic relationship between *SYF1* and the THO complex, a subcomplex of the TREX complex, by synthetic lethality. Double shuffle stains of *SYF1* and *HPR1* or *MFT1* were generated. They were transformed with *SYF1* or one of the *syf1* ts alleles and *HPR1*, *MFT1* or an empty plasmid.

Transformants were streaked onto plates containing 5-FOA and their growth was analyzed after incubation at 30ºC for 5 days.

As shown in Figure 2.5, combination of *syf1-37* and *Δhpr1* or *Δmft1* leads cells to lethality, indicating the synthetic lethal relationship between *SYF1* with *HPR1* and *MFT1*. Since, *SYF1* is not in the same complex with THO, *SYF1* might be functionally redundant with respect to the THO complex, or be in the same linear pathway with THO, or involve in parallel pathways with THO that are together crucial for cells.

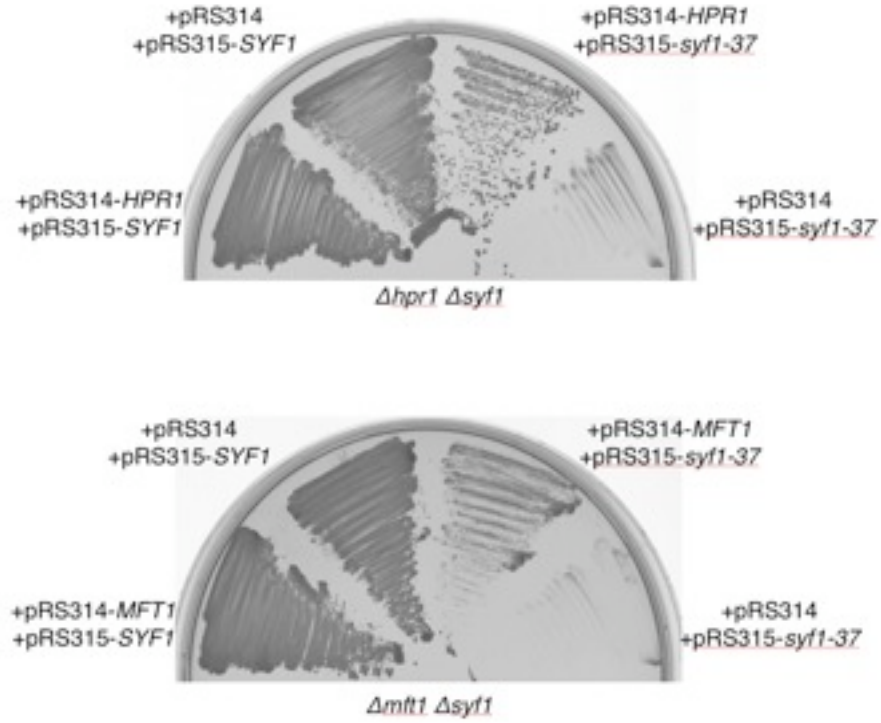

Figure 2.5 : *SYF1* genetically interacts with the THO complex components, *HPR1* and *MFT1*

SYF1 HPR1 and *SYF1 MFT1* double shuffle strains were transformed with an empty plasmid or plasmids encoding wild-type *HPR1* or *MFT1*, respectively, and a plasmid encoding wild-type *SYF1* or *syf1-37*. Synthetic lethality was assessed by restreaking the transformants on plates containing 5-FOA.

Remarkably, among five *syf1*-ts alleles, *syf1-37* is the only allele that genetically interacts with the THO complex while the other ts alleles do not as shown in Table 1.

	Δhpr1	*Δmft1*
SYF1	-	-
syf1-1	-	-
syf1-2	ss	ss
syf1-15	-	-
syf1-34	-	-
syf1-37	sl	sl

Table 1: *SYF1* genetically interacts with *HPR1* and *MFT1* in an allele-specific manner

SYF1 HPR1 and *SYF1 MFT1* shuffle strains were transformed with an empty plasmid or plasmids encoding wild-type *HPR1* or *MFT1*, respectively, and a plasmid encoding wild-type *SYF1* or one of the five *syf1*-ts alleles. An sl interaction was assessed by restreaking the respective transformants on 5-FOA containing plates. sl: synthetic lethal; ss: synthetic sick; -: no genetic interaction.

2.3 Physical interaction between Syf1 and the THO complex

Since *SYF1* genetically interacts with the THO complex, it is possible that they also interact biochemically with each other and function together in the cells. To test this, I performed a coimmunoprecipitation assay using protein extracts from Syf1-hemagglutinin(HA)- and Hpr1-tandem-affinity-purification(TAP)-tagged cells. Hpr1-TAP was immunoprecipitated by IgG sepharose, and the immunoprecipitates were probed by α-HA. Syf1-HA could be coimmuno-precipitated with Hpr1, whereas no signal was detected in the control (Figure 2.6).

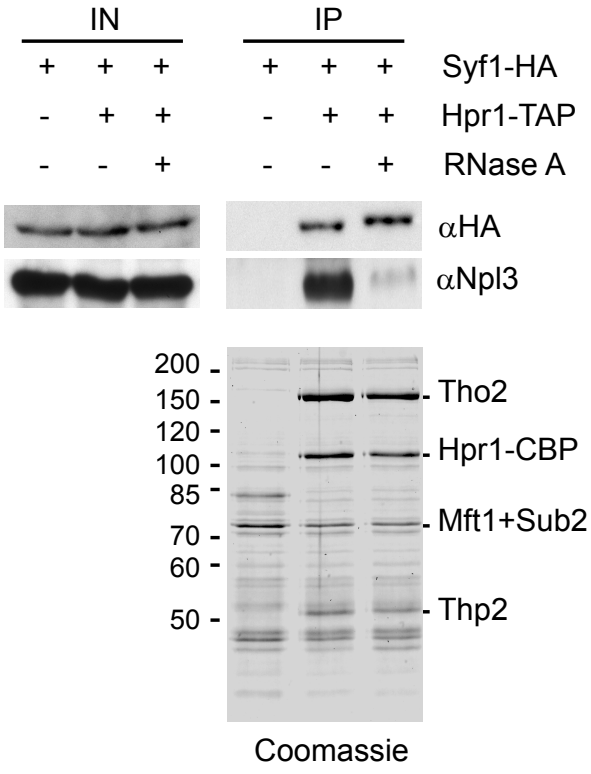

Figure 2.6: Syf1 interacts with TREX in an RNA-independent manner
The TREX complex was purified using TAP-tagged Hpr1 and copurification of HA-tagged Syf1 assessed by western blotting with α-HA antibody (α-HA). To assess whether the interaction between TREX and Syf1 is mediated by RNA, cell extract was treated with RNase A prior to purification. Npl3, an SR-protein involved in mRNA export that interacts with TREX in an RNA-dependent manner, served as positive control for the RNase treatment (α-Npl3). IN: input, IP: immunoprecipitation, Coomassie: Tobacco-etch-virus protease (TEV) eluate of the TAP purification.

Since Syf1 is a component of the splicing machinery and THO is able to interact with RNA, I assessed whether this interaction is mediated by RNA. After RNase treatment, Syf1-HA was still coimmunoprecipitated with Hpr1 indicating that Syf1 binds the THO complex in an RNA-independent manner (Figure 2.6).

Taken together Syf1 interacts with TREX both genetically and physically.

2.4 Deletion of the C-terminal domain of Syf1 causes 6-azauracil sensitivity and synthetic lethal with *Δhpr1* or *Δmft1*

The genetic interaction between *SYF1* and THO indicates that Syf1 might have overlapping functions with the THO/TREX complex, namely transcription and/or mRNA export. To assess whether *SYF1* also plays a role in transcription, I tested *syf1*-ts cells for sensitivity to 6-azauracil (6-AU), a drug that impairs transcription elongation by reducing GTP/UTP from intracellular nucleotide pools. Interestingly, among the five *syf1*-ts alleles, *syf1-37*, the only allele that genetically interacts with THO, causes 6-AU sensitivity (Figure 2.7). This result suggests for the first time that Syf1 might play a role in transcription elongation and that the genetic interaction between *SYF1* and THO might be caused by their overlapping function in transcription elongation.

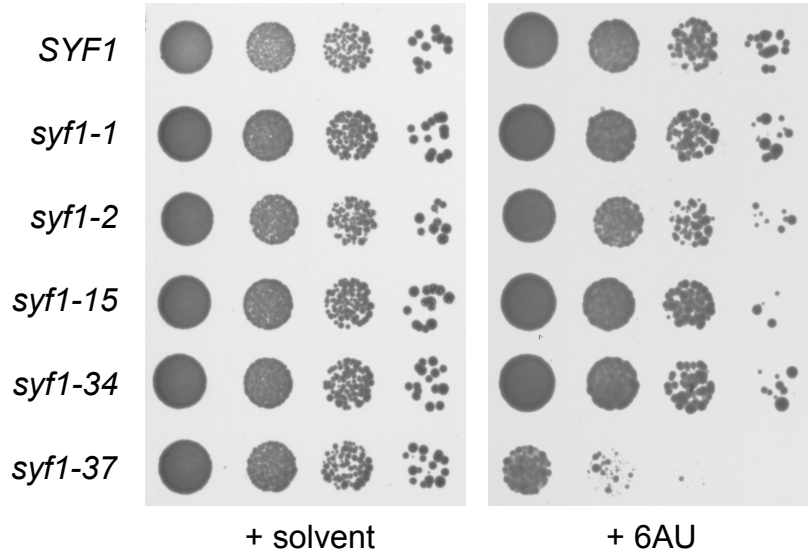

Figure 2.7: The *syf1-37* ts mutant is sensitive to 6-AU
10-fold serial dilutions of yeast cells expressing wild-type *SYF1* or one of five different ts alleles of *SYF1* and containing pRS316 were spotted onto SDC(-ura) plates containing solvent or 50 µg/ml 6-AU and incubated for 2 days at 30°C.

Interestingly, the *syf1-37* ts allele is the only mutant that encodes a C-terminal-truncated protein (Figure 2.2). To confirm whether the mutant does not generate a full-length protein due to a read-through of the nonsense codon during translation as previously shown for some proteins, different dilutions of the whole cell extracts of TAP-*syf1-37* (1-100x) and of TAP-*SYF1* cells (100x) were separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and the amount of Syf1 was assessed by western blotting using an antibody recognizing the protein A tag. As shown in Figure 2.8, no full-length Syf1 could be detected as a result of frame-shifting during the translation of the *syf1-37* allele, *i.e.* if any, the product of the read-through of the nonsense codon is less than 1% of the truncated protein.

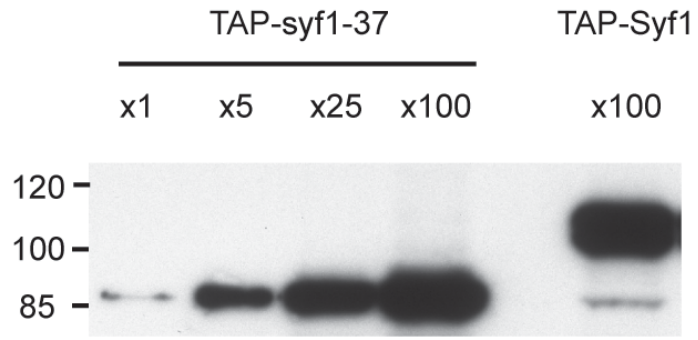

Figure 2.8: The *syf1-37* ts mutant is a C-terminal truncated protein
No full-length Syf1 can be detected as a consequence of frame-shifting during translation of the *syf1-37* allele. Different dilutions of whole cell extracts of TAP-*syf1-37* (1 to 100x) and of TAP-*SYF1* cells (100x) were separated by SDS-PAGE and the amount of TAP-syf1-37 and TAP-Syf1 assessed by western blotting using an antibody recognizing protein A. Frame-shifting during translation of the *syf1-37* allele to produce full-length Syf1 is, if any, less than 1%.

Syf1 is mainly composed of 15 tetratricopeptide repeat (TPR) motifs. TPR motif exhibits a high degree of sequence diversity, but structural comparison reveals a highly conserved three dimensional structure, which consists of two anti-parallel alpha helices separated by a turn. The motif has been found in many organisms ranging from bacteria to humans and is thought to be a scaffold domain for protein complex assembly. Since syf1-37 lacks the four most C-terminal TPR motifs, I asked whether this truncation causes *SYF1* to interact genetically with *HPR1* and *MFT1*. To answer this question, I gradually truncated the C-terminal part of Syf1 proteins and tested for the synthetic lethality of those different mutants with *Δhpr1* or *Δmft1* (Figure 2.9). Deletion of the C-terminal unstructured region of Syf1 (*syf1-ΔC*), the last (15th) and the last two (14th and 15th) TPR motifs (*syf1-Δ15* and *syf1-Δ14*, respectively) were not synthetic lethal with both *Δhpr1* and *Δmft1*. When I deleted one more TPR motif (*syf1-Δ13*), it became synthetic lethal with *Δhpr1* and *Δmft1*. Additional deletion of the next (12th) TPR motif (*syf1-Δ12*) also caused the synthetic lethality. Note that this syf1-Δ12 is 20 amino acids longer than the syf1-37.

Thus, the deletion of the C-terminal part of Syf1 is indeed responsible for the genetic interaction with the THO complex.

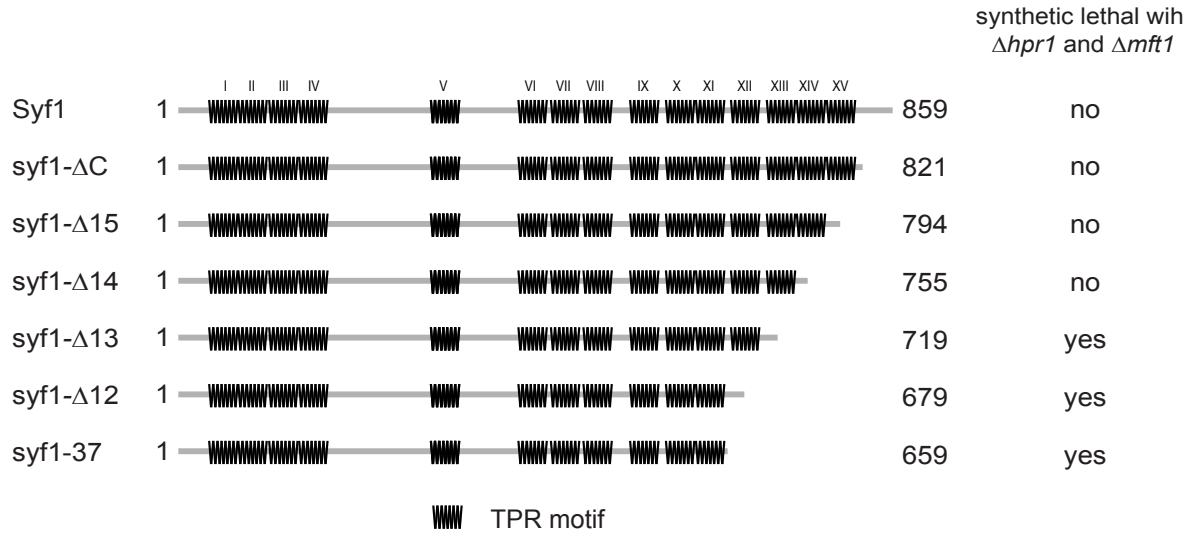

Figure 2.9: The C-terminus of Syf1 is important for the genetic interaction of *SYF1* with *HPR1* and *MFT1*

Schematic diagram showing the domain structure of Syf1 with each TPR motif indicated by a zigzag. Deletion of the three most C-terminal TPR motifs (*syf1-Δ13*) leads to synthetic lethality with *Δhpr1* and *Δmft1*.

2.5 Syf1 functions in transcription

2.5.1 Syf1 and Prp19 complex are recruited to transcription machinery *in vivo*

Genetic interaction with THO components, 6-AU sensitivity of a *syf1* mutant and transcription activity of the human homolog XAB2 suggest that Syf1 might also play a role in transcription in yeast. Thus, I hypothesized that if Syf1 is a general transcription factor, it should be present at the transcription machinery not only on intron-containing genes where splicing occurs, but also on intronless genes. To assess this, I performed chromatin immunoprecipitation

(ChIP) assays using a *SYF1*-TAP strain and assessed the recruitment of the protein to four intronless genes (*ADH1*, *PGK1*, *PMA1*, and *PDR5;* Figure 2.10) and two intron-containing genes (*ACT1* and *DBP2;* Figure 2.10). Occupancy on these six genes was calculated as the enrichment of Syf1 at the respective gene relative to a non-transcribed region (NTR) serving as a negative control. Strikingly, Syf1 is specifically recruited to all of these six genes, but not to the NTR (Figure 2.11). Since Syf1 is recruited to intronless genes, it might have a general function in transcription.

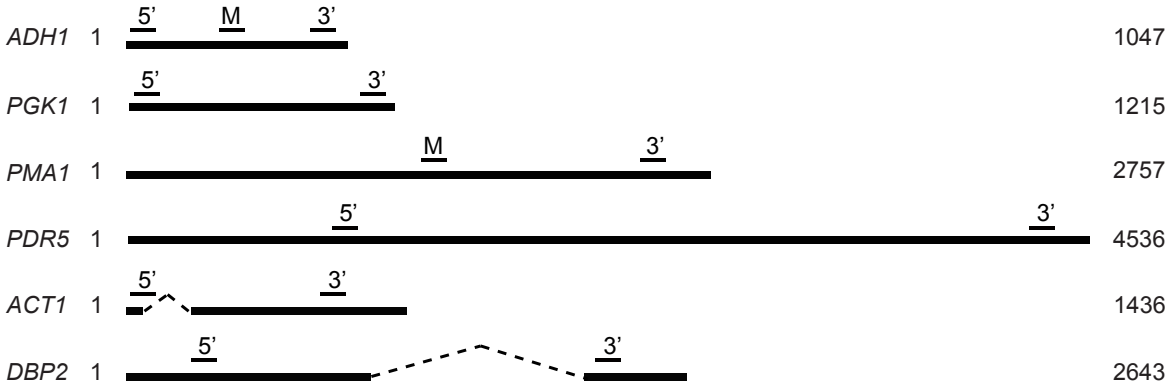

Figure 2.10: Schematic diagram of the *ADH1*, *PGK1*, *PMA1*, *PDR5*, *ACT1*, and *DBP2* genes

The open reading frames (ORFs) are represented by a solid and the introns by a hatched line. Bars above the genes show the positions of the primer pairs used for the ChIP analysis.

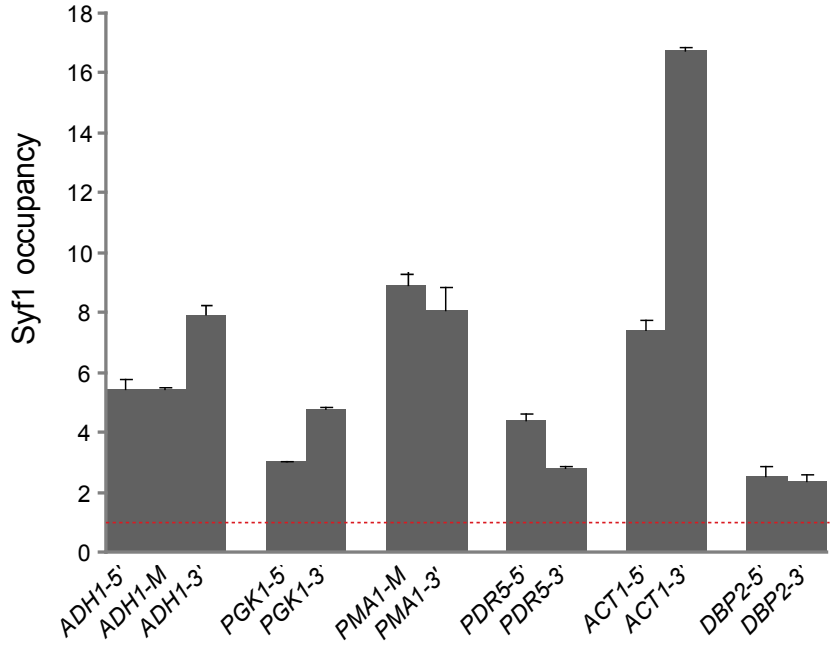

Figure 2.11: Syf1 is recruited to transcribed genes

The occupancy of Syf1 at the intronless *ADH1*, *PGK1*, *PDR5*, and *PMA1* and the intron-containing *ACT1* and *DBP2* genes was calculated as the enrichment of Syf1 at the respective gene relative to its presence at the NTR. The occupancy at the NTR was set to 1 as indicated by a dashed line.

Next, I asked whether Syf1 alone or the whole Prp19 complex is recruited to the genes. To do this, I assessed the occupancy of two other Prp19 complex components, namely Prp19 and Ntc20, to the six genes shown above. Prp19 (Figure 2.12) and Ntc20 (Figure 2.13) were also present at these genes suggesting that the whole Prp19 complex is recruited to transcribed genes.

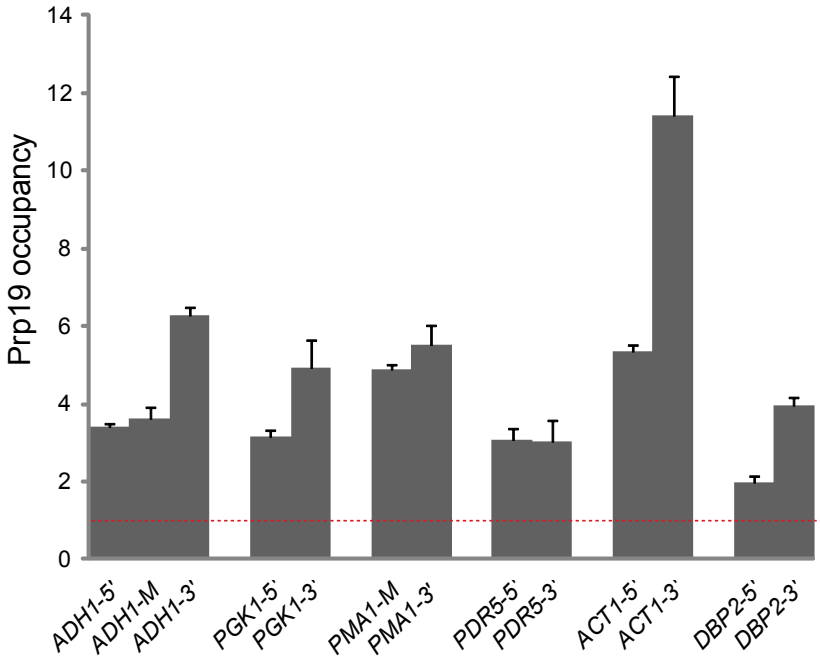

Figure 2.12: Prp19 is recruited to transcribed genes

Experiment as in Figure 2.11, except the occupancy of Prp19 was assessed.

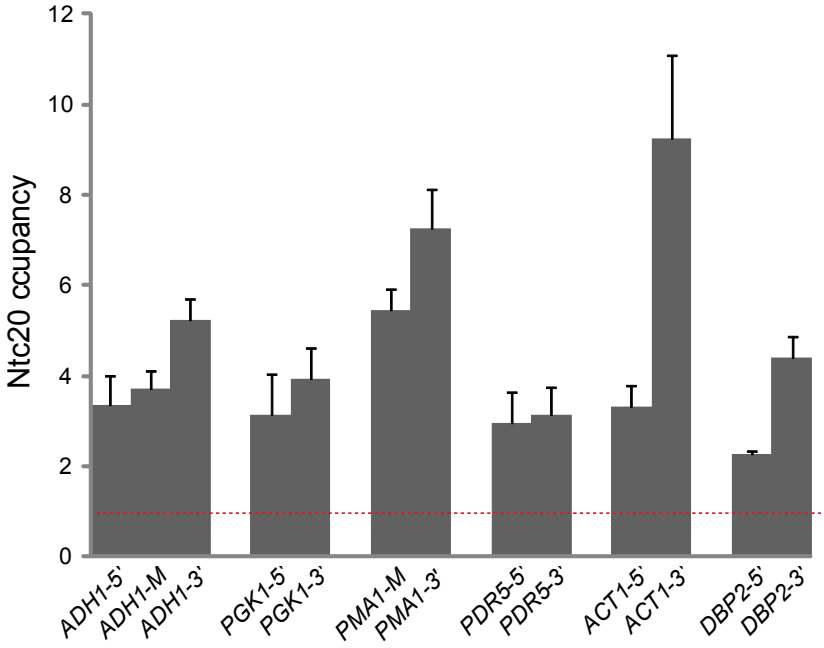

Figure 2.13: Ntc20 is recruited to transcribed genes

Experiment as in Figure 2.11, except the occupancy of Ntc20 was assessed.

2.5.2. Recruitment of Syf1 to transcribed genes is transcription dependent

Next, I asked whether the recruitment of Syf1 to transcribed genes is dependent on an active transcription. To test this, I assessed the occupancy of Syf1 at the inducible *GAL1* gene. The *GAL1* gene is not expressed when cells are grown in raffinose-containing medium and is induced when cells are grown in galactose-containing medium. The recruitment of Syf1 to the *GAL1* gene was observed only when the gene is induced, but was not in the non-induced condition (Figure 2.14). This indicates that Syf1 recruitment to the transcription machinery is transcription dependent.

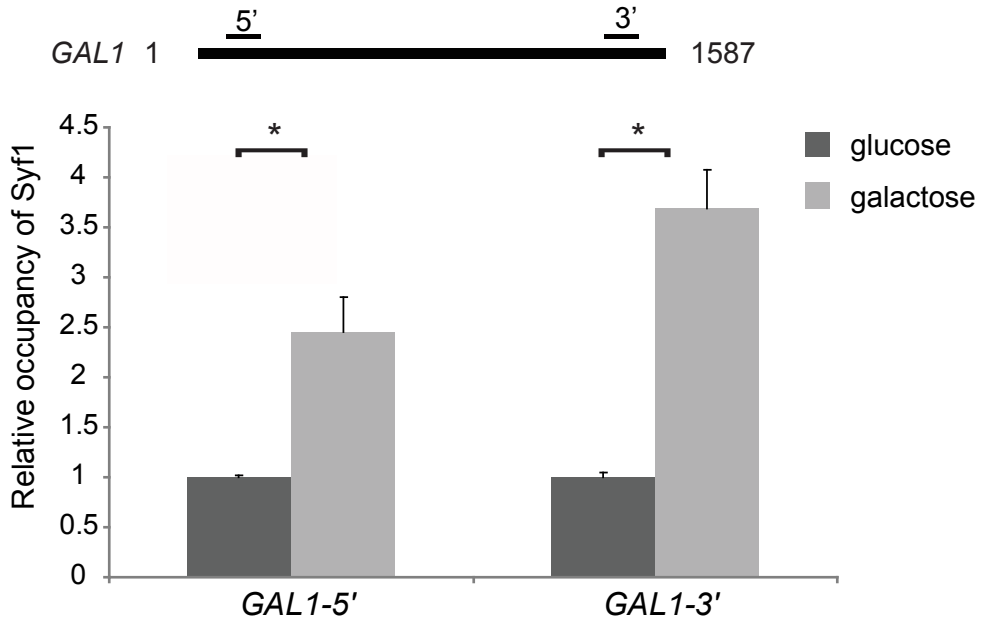

Figure 2.14: Recruitment of Syf1 to a gene is transcription dependent
Occupancy of Syf1 to the *GAL1* gene under non-induced (cells grown in glucose-containing medium, dark grey bars), which was set to 1, and induced conditions (cells grown in galactose-containing medium, light grey bars). Primer pairs amplify the 5' and 3' region of *GAL1* as indicated in the upper panel.

2.5.3. Functional Syf1 is necessary for effective transcription *in vivo*

As Syf1 is physically present at transcribing genes in a transcription-dependent manner, it is very likely that Syf1 plays a role in transcription and affects transcription activity *in vivo*. To assess this, I performed an *in vivo* transcription assay. Among the *SYF1* truncation mutants, *syf1-Δ13* is the longest allele and synthetic lethal with *Δhpr1* and *Δmft1*. However, because *syf1-Δ13* cells grow more slowly than the *syf1-37* ts mutant, I chose the latter mutant for performing further experiments in order to avoid unspecific effects due to slow growth (see Figure 2.15).

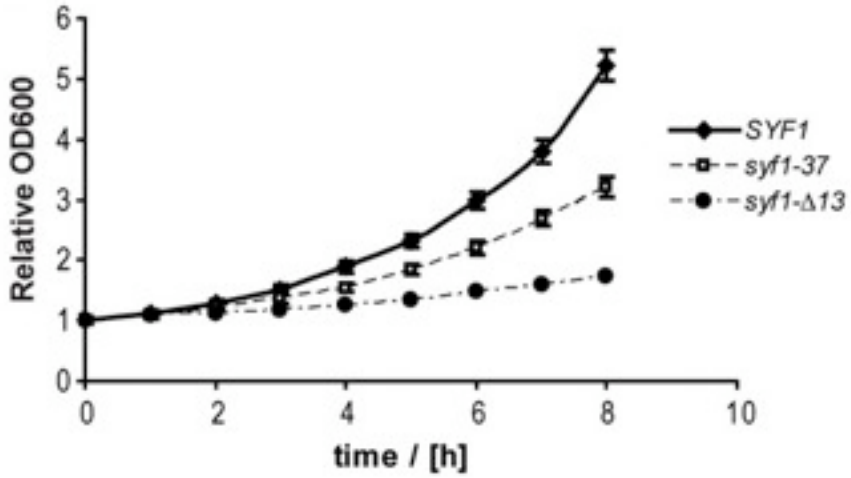

Figure 2.15: Growth curves of *SYF1*, *syf1-37*, and *syf1-Δ13* cells at 30°C
Growth curves of *SYF1*, *syf1-37*, and *syf1-Δ13* cells in YPD at 30°C by measuring optical density at 600 nm.

In the transcription assay, I used two reporter constructs: *ACT1* as an intron-containing gene and *GAL10* as an intronless gene, both of which were controlled by the regulatable *GAL10* promoter (*GAL10::ACT1* and *GAL10::GAL10*, respectively.) Wild-type and mutant cells were grown in raffinose-containing medium at permissive temperature. Then, galactose was added and cells harvested at 0, 15, 30 and 60 minutes (min) after induction.

Next, RNA was extracted and the amount of the *ACT1* and *GAL10* mRNAs quantified. As shown in Figure 2.16, the amount of both *ACT1* and *GAL10* mRNAs were reduced in the *syf1-37* mutant compared to the wild-type cells, indicating that in this mutant mRNA synthesis of both intron-containing and intronless genes are decreased due to the mutation of the *SYF1* gene.

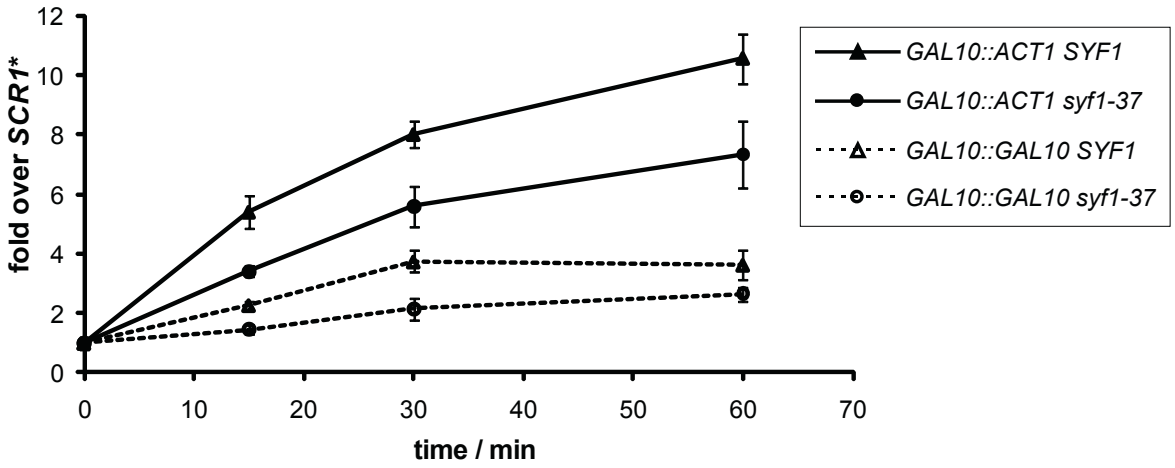

Figure 2.16: Functional Syf1 is necessary for efficient transcription *in vivo*
GAL10::ACT1 and *GAL10::GAL10* were used as reporters for an intron-containing and an intronless gene, respectively. The level of *ACT1* (solid) and *GAL10* (dotted) mRNAs of cells grown for 0, 15, 30, and 60 min after galactose induction at the permissive temperature (30°C) was assessed by primer extension and normalized to the amount of *SCR1* RNA, an RNAPIII transcript.

Since Syf1 is known as a splicing factor, I asked whether this decreased transcription activity is related to splicing activity and/or a secondary effect of defective splicing. By using the same *GAL10::ACT1* reporter construct, I quantified the ratio of the amount of spliced and unspliced mRNAs and subsequently calculated the splicing efficiency. Expression of the *GAL10::ACT1* reporter was induced for 1 hour and the amount of spliced and unspliced *ACT1* mRNAs assessed (Figure 2.17). As expected for a ts mutant of a splicing factor, unspliced transcripts accumulated in the *syf1-37* strain at the restrictive temperature. More importantly, the *syf1-37* mutant did not show

a splicing defect at the permissive temperature, indicating that the decrease in *ACT1* mRNA level in the *syf1-37* strain is most likely not a secondary effect of defective splicing.

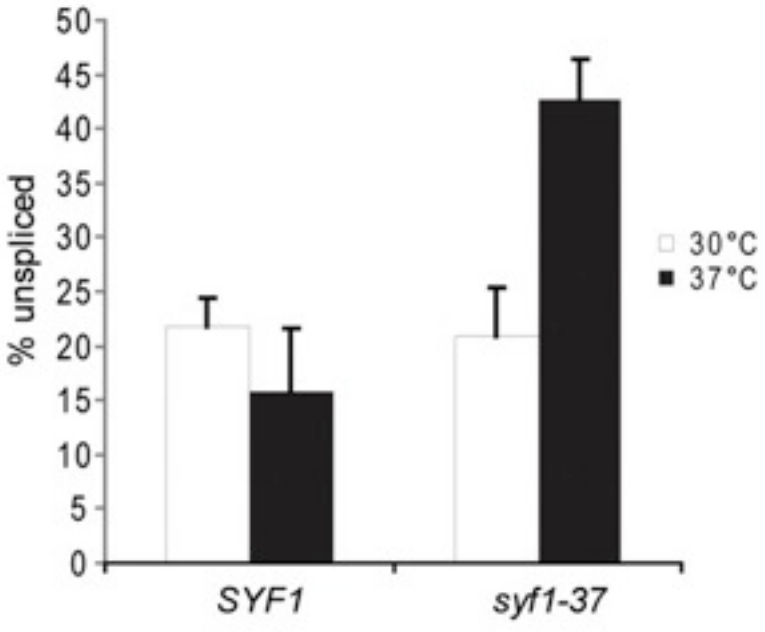

Figure 2.17: Splicing is not affected in *syf1-37* cells
The *GAL10::ACT1* reporter construct was induced for 1 hour with galactose in *SYF1* and *syf1-37* cells, which were grown at the permissive temperature (30°C) or shifted for 1 h to the restrictive temperature (37°C). The amount of spliced and unspliced *ACT1* RNA was assessed by PCR. The splicing efficiency was calculated as the ratio of unspliced versus total *ACT1* RNA.

Taken together, *SYF1* is necessary for effective transcription *in vivo* and the function in transcription is independent of its splicing activity.

2.5.4 Functional Syf1 is necessary for effective transcription *in vitro*

As most cellular processes *in vivo* are highly complex and intimately linked, the decreased mRNA synthesis in *syf1-37* ts strain could also be a secondary effect of any step in gene expression that is impaired in the cells. In order to clarify whether Syf1 is directly involved in transcription, I decided to make use of an *in vitro* transcription assay.

I performed a plasmid-based *in vitro* transcription assay with either a *HIS4* or a *SER3* promoter (Koschubs et al. 2009) and prepared nuclear extracts from the *SYF1* and the *syf1-37* cells. At permissive temperature (18°C), the *syf1-37* nuclear extract had slightly lower transcriptional activity than the wild-type extract (Figure 2.18, lane 1-2), which could be due to the mutation or to variations in the nuclear extract preparation. Importantly, though, when the temperature during the transcription reaction was raised to the restrictive temperature (30°C), the activity of the *syf1-37* nuclear extract decreased significantly compared to the wild-type extract (Figure 2.18).

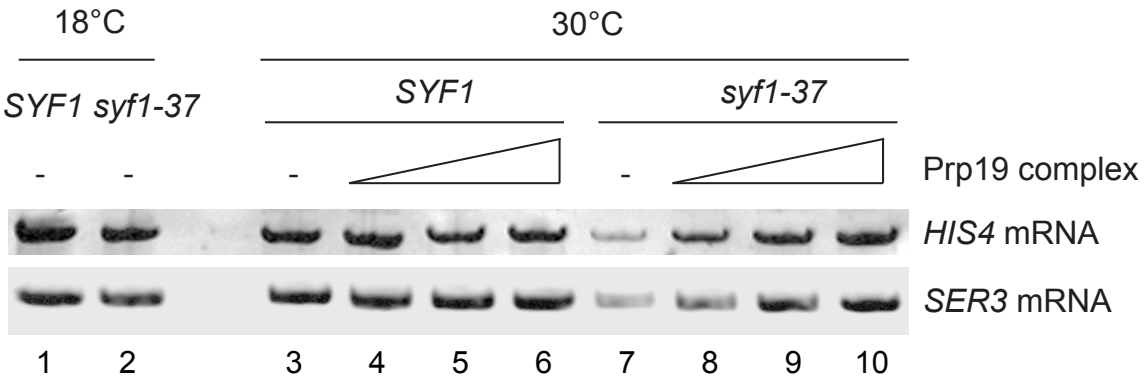

Figure 2.18: Functional Syf1 is necessary for efficient transcription *in vitro*
The transcription activity of nuclear extracts from *SYF1* and *syf1-37* cells was assessed at 18°C (permissive temperature, lanes 1 and 2) and 30°C (restrictive temperature, lanes 3 and 7) by an *in vitro* transcription assay using a template containing either a *HIS4* or a *SER3* promoter. Add-back of 50, 125 or 250 ng of Prp19 complex purified from *S. cerevisiae* rescued the impaired transcription activity of the *syf1-37* extract at the restrictive temperature (lanes 7-10).

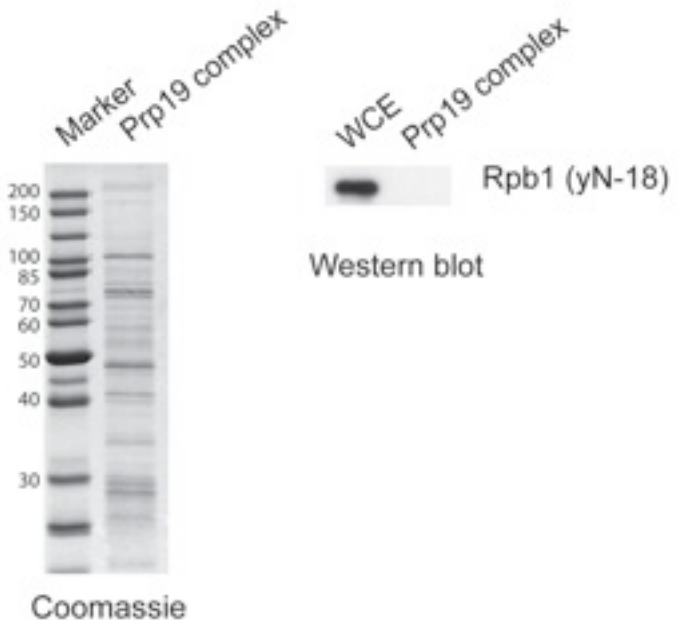

Figure 2.19: The purified Prp19 complex used for the add-back in the *in vitro* transcription experiment does not contain RNAPII

250 ng Prp19 complex were separated by SDS-PAGE and stained with Coomassie (left panel) or assessed for the presence of Rpb1, the largest subunit of RNAPII, by western blotting with an antibody recognizing the N-terminus of Rpb1 (yN-18, right panel). WCE: whole cell extract.

To assess further whether this reduction is indeed due to an inactive Syf1 protein in *syf1-37* nuclear extract but not the other proteins inactivated during the restrictive temperature, I performed an add-back experiment with the Prp19 complex purified from yeast. Strikingly, the addition of the Prp19 complex rescued the transcriptional activity of the *syf1-37* extract in a dose-dependent manner (Figure 2.18, lane 7-10). Because RNAPII was not found in the purified Prp19 complex added back to the extract assessed by western blotting (Figure 2.19) and the addition of the purified protein could not increase the transcriptional activity of the wild-type extract (Figure 2.18, lane 3-6), the activity rescued in the *syf1-37* (Figure 2.18, lane 7-10) was not simply caused by adding more active transcription machinery or other stimulating

transcription factor copurifying with the Prp19 complex. Taken together, this result indicates that the decreased activity in the *syf1-37* extract is due to impaired function of the syf1-37 ts protein and that the Syf1/Prp19 complex is indeed necessary for transcriptional activity *in vitro*. In summary, functional Syf1 is necessary for effective transcription *in vitro*.

2.5.5 Syf1 is necessary for full RNAPII processivity

As *syf1-37* cells are sensitive to 6-AU, a transcription elongation inhibitor, and the fact that the THO/TREX complex plays a role in transcription elongation, Syf1 is most likely also involved in the elongation phase of transcription. In principle, a decreased transcriptional activity during elongation can be caused by a reduction in the rate of transcription, namely the speed of elongating RNAPII, or by a decrease in the processivity of RNAPII, namely the capability of transcribing RNAPII to travel the entire length of the gene. Interestingly, it has been shown that deletion of the THO complex components reduces RNAPII processivity (Mason and Struhl 2005). It is possible that the Syf1/Prp19 complex functions in the same way as the THO complex and plays a role in RNAPII processivity. To clarify this, I measured RNAPII processivity by quantifying the level of RNAPII associated at five positions along the 8-kb-long *YLR454* gene under the control of *GAL1* promoter in cells grown in galactose-containing medium.

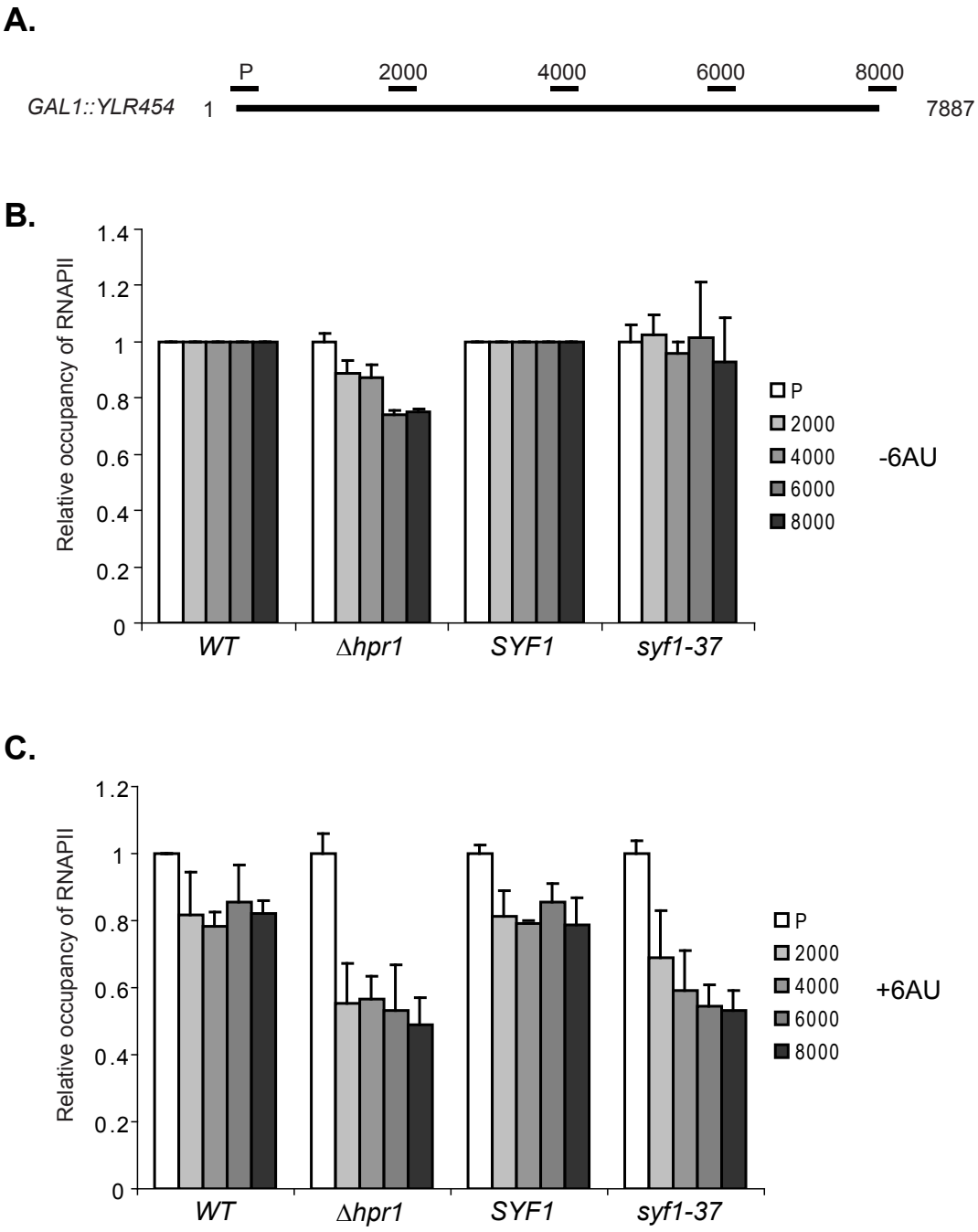

Figure 2.20: RNAPII processivity is decreased in *Δhpr1* and *syf1-37* cells

(A) Schematic diagram of the *GAL1::YLR454* reporter gene used for analyzing RNAPII processivity in wild-type, *Δhpr1*, *SYF1*, and *syf1*-37 cells grown at permissive temperature, 30°C. Positions of primer pairs used for the ChIP analysis are indicated above the gene.

(B) Relative RNAPII occupancy within the *YLR454* gene in wile-type, *Δhpr1*, *SYF1*, and *syf1-37* cells grown in the medium containing galactose. The occupancy in wild-type cells at each position was set to 1.

(C) Relative RNAPII occupancy as in (B) in the presence of 50 μg/ml 6-AU.

As shown previously, deletion of *HPR1* caused a reduction in RNAPII processivity and addition of 6-AU increased the defect. Processivity of RNAPII was not affected by the *syf1-37* mutation in the absence of 6-AU. Importantly, in the presence of 6-AU, RNAPII processivity was reduced in *syf1-37* cells indicating a transcription elongation defect in this mutant (Figure 2.20). Thus, Syf1 plays a role in transcription elongation and is required for efficient RNAPII processivity when transcription elongation is impaired by 6-AU.

2.6 The C-terminus of Syf1 is required for recruitment of the Prp19 complex to transcribed genes and for the interaction of the Prp19 complex with RNAPII

So far, I have identified the *syf1-37* allele which has a decreased transcription activity and lacks the four most C-terminal TPR motifs, known as scaffold domains for protein-protein interactions. Thus, it might be possible that by the lack of its C-terminal TPR motifs, Syf1 does not interact with some other proteins necessary for efficient transcription activity or is not recruited to the RNAPII transcription machinery and, as a result, transcription efficiency is reduced. To test this, I assessed recruitment of syf1-37 to three representative genes, namely *ADH1*, *PMA1* and *DBP2*, by the ChIP assay. While RNAPII recruitment to genes in *syf1-37* cells is comparable to its recruitment in *SYF1* cells, the occupancy of syf1-37 relative to RNAPII is significantly reduced to approximately 50% compared to the wild-type cells, indicating that syf1-37 could not be recruited efficiently to transcription machinery *in vivo* (Figure 2.21).

Since Syf1 is a component of the Prp19 complex, I asked whether not only Syf1 recruitment but also recruitment of the whole Prp19 complex is reduced

in *syf1-37* cells. Therefore, I assessed the occupancy of two other Prp19 complex components, namely Prp19 and Ntc20, in *SYF1* and *syf1-37* cells. In fact, the occupancy of both Prp19 and Ntc20 in *syf1-37* cells was also decreased to nearly 50% compared to the wild-type cells (Figure 2.22). Thus, I concluded that Syf1 functions in transcription elongation as part of the Prp19 complex.

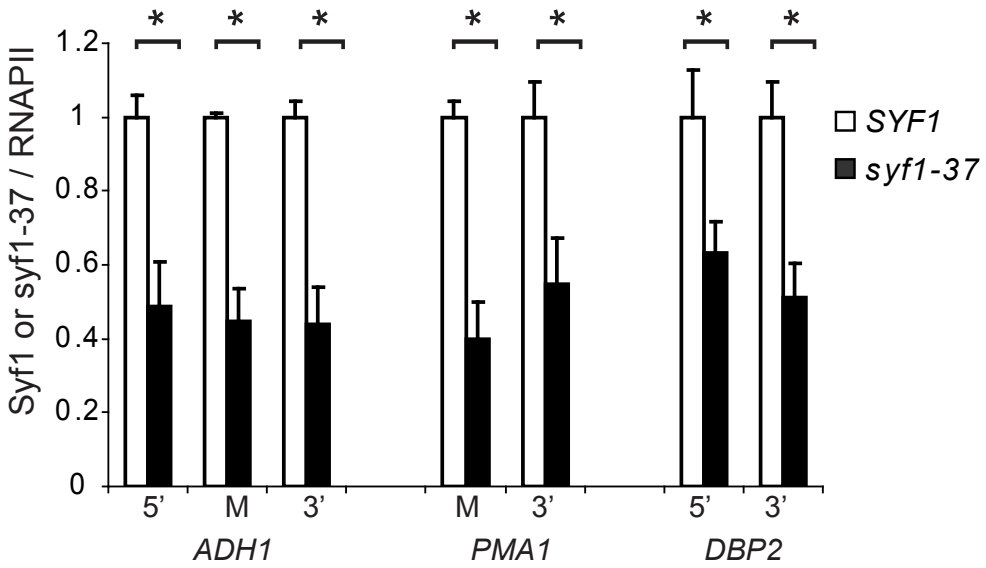

Figure 2.21 : The *syf1-37* mutation leads to loss of Syf1 from the transcribed genes

ChIPs of TAP-Syf1, TAP-syf1-37 and Rpb3-TAP (RNAPII) in *SYF1* and *syf1-37* cells grown at permissive temperature (30°C) were quantified by real-time (RT)-PCR using the primer pairs shown in Figure 2.10. The occupancy ratio of Syf1 over RNAPII over the NTR was calculated and the values for *SYF1* were set to 1.

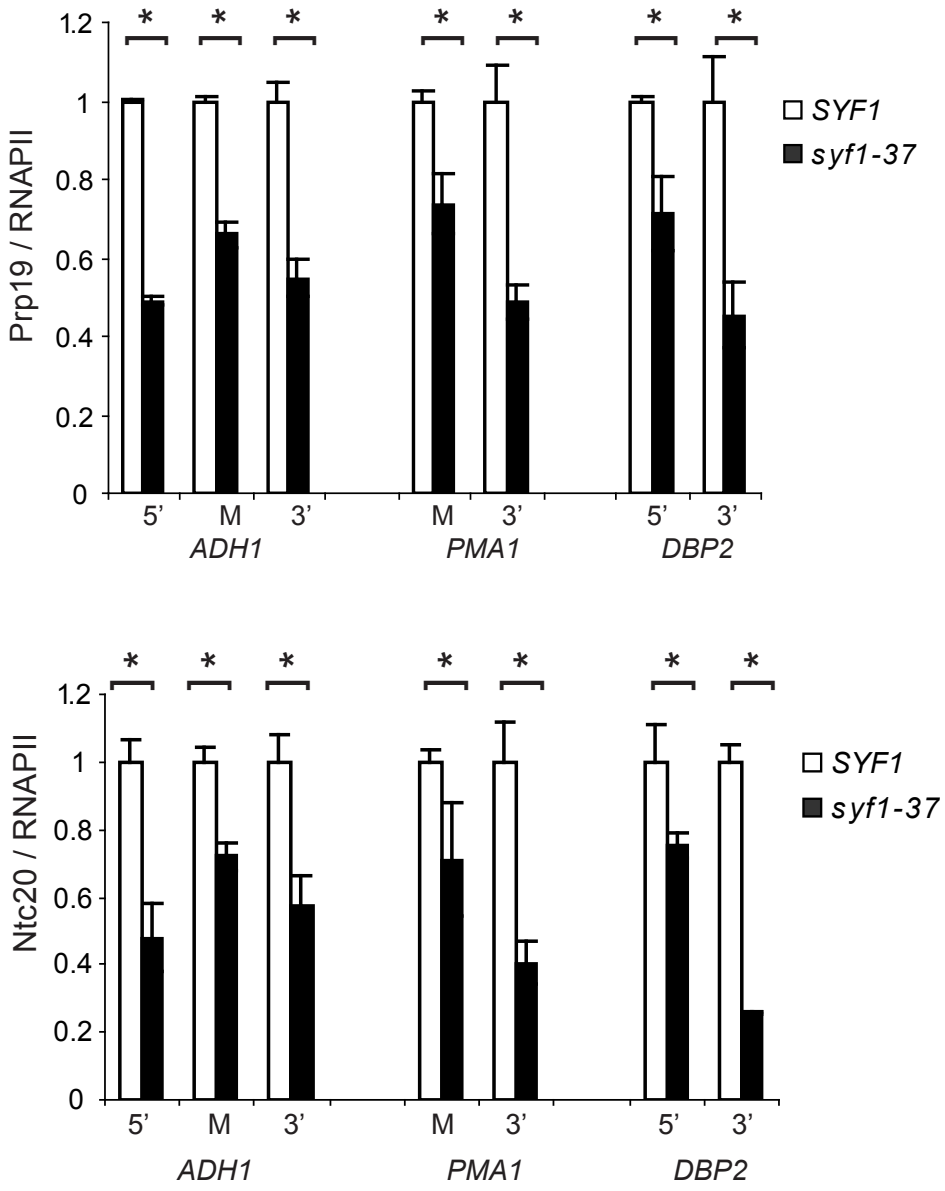

Figure 2.22: The *syf1-37* mutation leads to loss of Prp19 and Ntc20 from the transcribed genes

Experiment as in Figure 2.21, except the occupancy of Prp19 and Ntc20 were assessed.

The *syf1-37* mutation causes a lower level of the whole Prp19 complex at the transcribed genes most likely due to the lack of its C-terminal region. I wondered whether in this mutant the ability of syf1-37 to interact with other proteins would also be affected, *e.g.* Syf1/syf1-37 with RNAPII or Syf1/syf1-37 with THO/TREX. Thus, I assessed this by purifying TAP-tagged Syf1/syf1-37

and analyzing copurification of RNAPII and THO/TREX by western blotting. Both Syf1 and syf1-37 copurified the same amounts of Prp19 complex component and Hpr1 indicating that without its C-terminal region *syf1-37* could still interact with the whole Prp19 complex and THO/TREX. However, copurification of Rpb1, the largest subunit of RNAPII, with *syf1-37* was largely decreased suggesting that the interaction between RNAPII and the Prp19 complex is impaired by the lack of C-terminal region of Syf1 (Figure 2.23). This result therefore indicates that the C-terminus of Syf1 is necessary for the interaction between the Prp19 complex and RNAPII and consequently for the recruitment of the whole Prp19 complex to the transcription machinery.

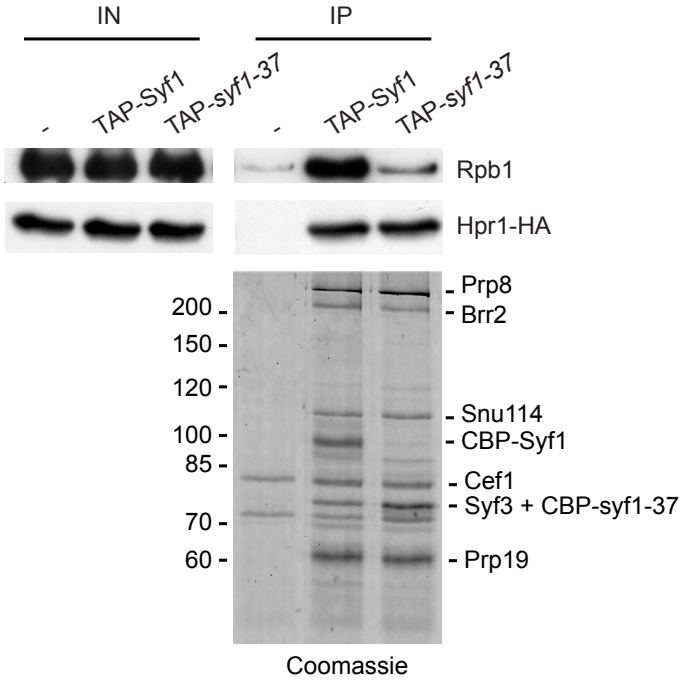

Figure 2.23 : The C-terminus of Syf1 mediates the interaction between the Prp19 complex and RNAPII

The interaction of syf1-37 with RNAPII was strongly decreased while the interaction with the Prp19 and the TREX complex was not affected. TAP-Syf1 and TAP-syf1-37 were purified from cells expressing Hpr1-HA and copurification of RNAPII and Hpr1 assessed by western blotting using the 8WG16 antibody, recognizing the largest subunit of RNAPII Rpb1, or an α-HA antibody, respectively. Copurifying proteins were identified by mass spectrometry and indicated on the right. A non-tagged strain served as negative control. IN: input, IP: immunoprecipitation, Coomassie: EGTA eluates of the TAP purifications.

2.7 The Prp19 complex is required for TREX occupancy at transcribed genes

The genetic interaction between *SYF1* and the THO complex and the function of Syf1 in transcription indicate that there might be an interplay between the Prp19 and the TREX complex. Therefore, I assessed whether the occupancy of Hpr1, a component of the THO complex, is influenced by Syf1 and *vice versa*. While the occupancy of Syf1 relative to RNAPII was not affected by deletion of *HPR1* (Figure 2.24), Hpr1 occupancy relative to RNAPII was reduced in *syf1-37* cells approximately to 50% compared to the wild-type cells which is coincident with nearly 50% reduction in the Prp19 complex recruitment in *syf1-37* cells (Figure 2.25). Interestingly, the decrease of THO recruitment in *syf1-37* cells was not due to a general or nonspecific loss of transcription factors at the site of transcription because the recruitment of RNAPII (Figure 2.26) and two other transcription elongation factors Paf1 (Figure 2.27), a component of the PAF complex, and Spt5 (Figure 2.28) was not affected by the *syf1-37* mutation. Thus, the Prp19 complex is necessary for THO occupancy at transcribed genes.

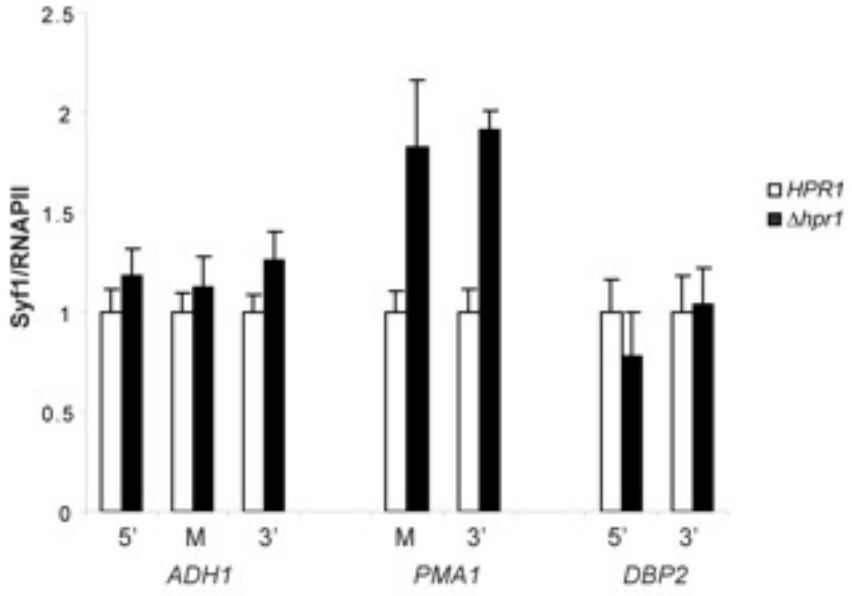

Figure 2.24: Deletion of *HPR1* does not affect occupancy of Syf1

ChIP experiments of Syf1-TAP in wild-type and *Δhpr1* cells were quantified by RT-PCR using the primer pairs shown in Figure 2.10. The ratio of Syf1 occupancy relative to RNAPII over the NTR was calculated and the values for wild-type were set to 1.

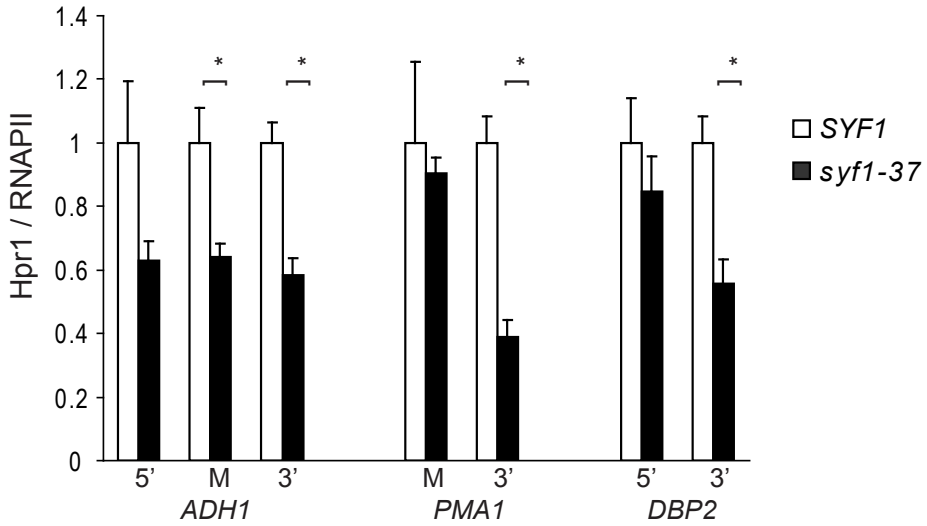

Figure 2.25: Mutation of *SYF1* leads to loss of the THO complex from the transcribed gene

ChIPs of Hpr1-TAP in *SYF1* and *syf1-37* cells grown at permissive temperature, 30°C, were quantified by RT-PCR using the primer pairs shown in Figure 2.10. The ratio of the occupancy of Hpr1 relative to the occupancy of RNAPII over the NTR was calculated and the values for *SYF1* were set to 1.

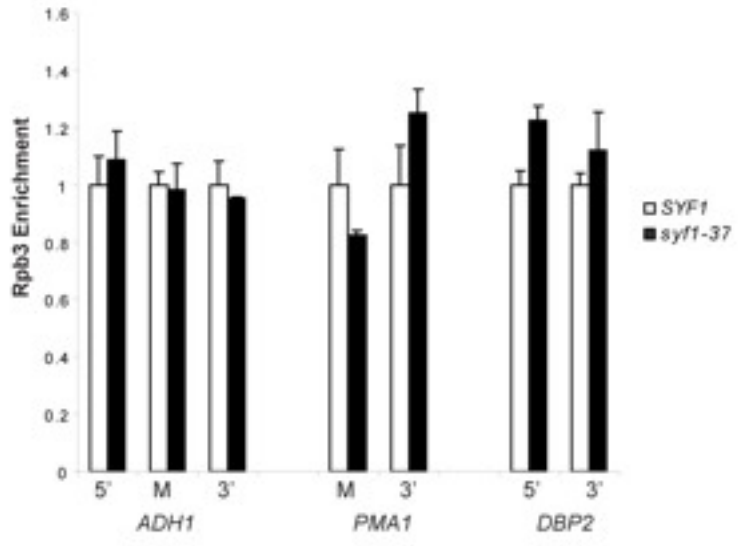

Figure 2.26: Mutation of *SYF1* does not lead to loss of the RNAPII from the transcribed gene

ChIPs of Rpb3-TAP in *SYF1* and *syf1-37* cells grown at the permissive temperature, 30°C, were quantified by RT-PCR using the primer pairs shown in Figure 2.10. The ratio of the occupancy of Rpb3 over the NTR was calculated and the values for *SYF1* were set to 1.

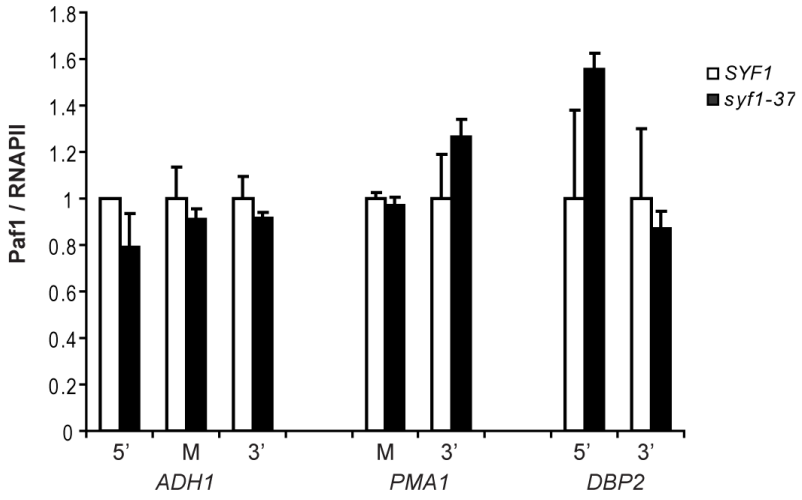

Figure 2.27: Mutation of *SYF1* does not lead to loss of the Paf1 from the transcribed gene

ChIPs of Paf1-TAP in *SYF1* and *syf1-37* cells grown at the permissive temperature, 30°C, were quantified by RT-PCR using the primer pairs shown in Figure 2.10. The ratio of the occupancy of Paf1 relative to the occupancy of RNAPII over the NTR was calculated and the values for *SYF1* were set to 1.

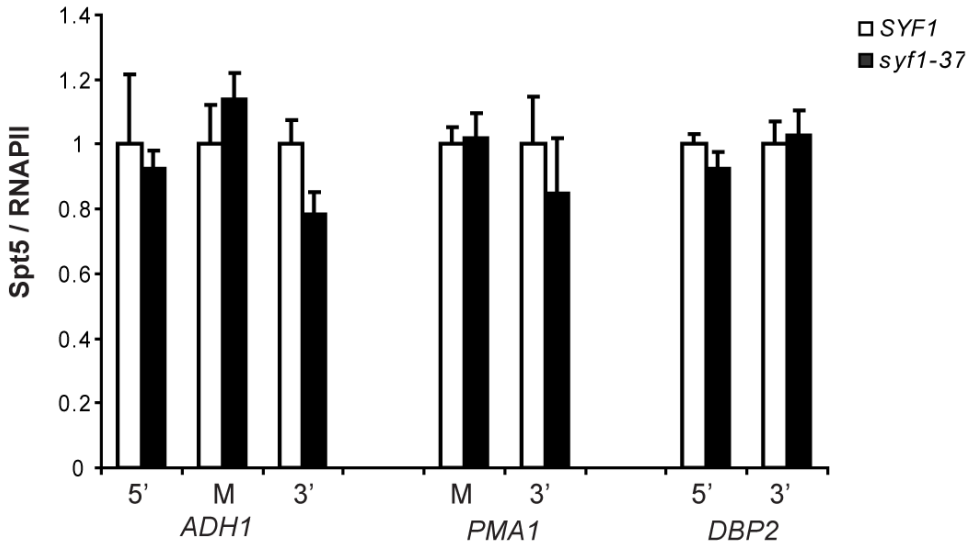

Figure 2.28: Mutation of *SYF1* does not lead to loss of the Spt5 from the transcribed gene

ChIPs of Spt5-TAP in *SYF1* and *syf1-37* cells grown at the permissive temperature, 30°C, were quantified by RT-PCR using the primer pairs shown in Figure 2.10. The ratio of the occupancy of Spt5 relative to the occupancy of RNAPII over the NTR was calculated and the values for *SYF1* were set to 1.

It has been shown that the THO subcomplex is required for the efficient recruitment of other TREX components such as mRNA export factors Sub2 and Yra1 to the site of transcription. Since the recruitment of THO is reduced in *syf1-37* cells, I wondered whether Sub2 and Yra1 recruitment is also affected by the mutation. I assessed the recruitment of both proteins in *syf1-37* cells and found that their occupancies were also decreased (Figure 2.29). However, the reduction in occupancy of both Sub2 and Yra1 at transcribed genes is lower than the decrease of Hpr1. This is most likely due to the fact that Sub2 and Yra1 are still partially recruited to genes even in the complete absence of Hpr1 (Zenklusen et al. 2002). Taken together, the Prp19 complex is needed for effective recruitment of the TREX complex to transcribed genes.

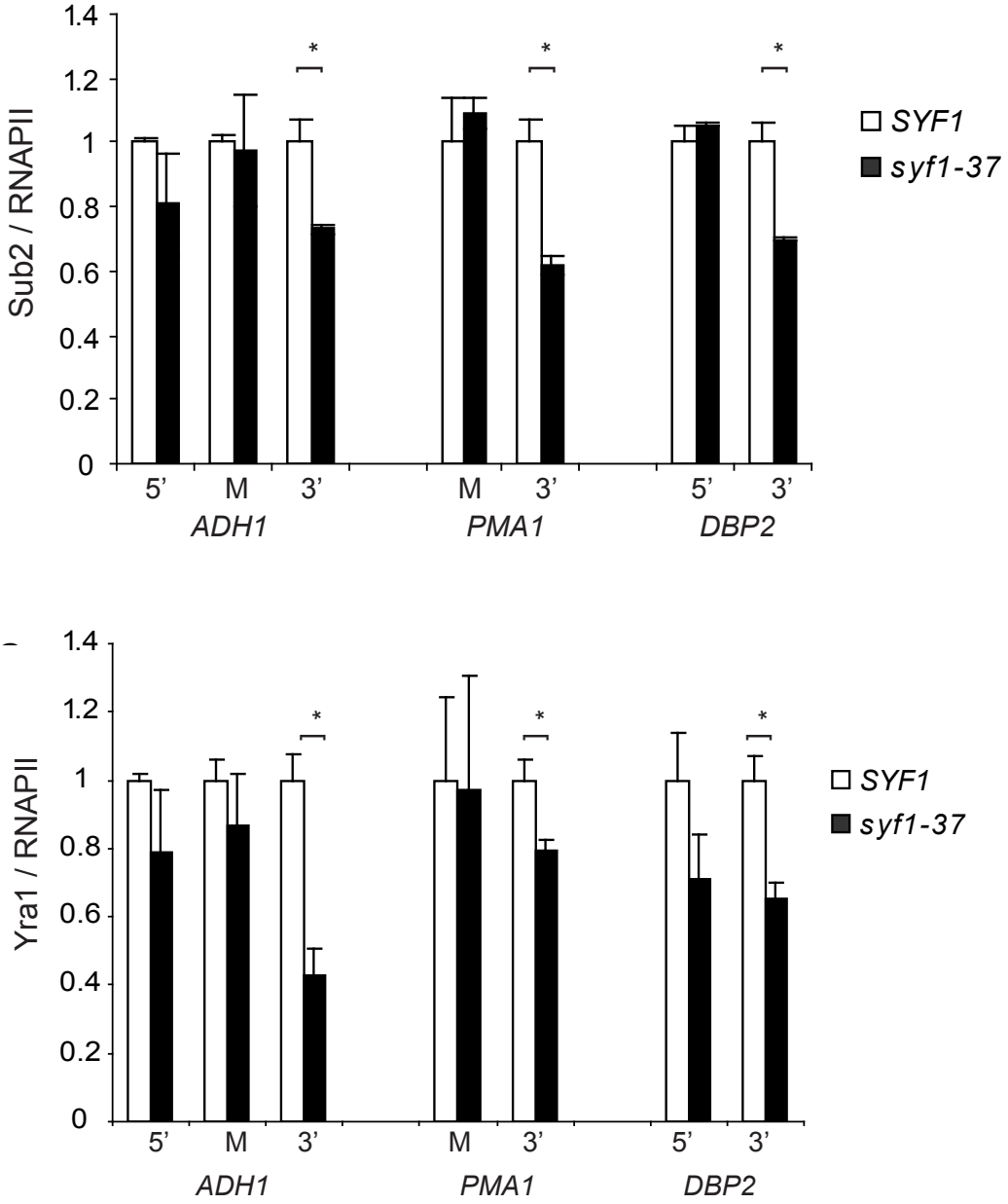

Figure 2.29: Syf1 is required for an efficient recruitment of TREX complex to the transcribed genes

Experiment as in Figure 2.25, except the occupancy of Sub2 and Yra1 were assessed.

2.8 The C-terminal TPR motifs of Syf1, which are important for the interaction of the Prp19 complex with RNAPII, are not interchangeable

Deletion of the last three C-terminal TPR motifs of Syf1 significantly impairs an interaction between the Prp19 complex and RNAPII and causes a reduction in Prp19 occupancy at transcribed genes *in vivo*. Thus, the C-terminal domain of Syf1 is an important binding interface between the Prp19 complex and RNAPII.

Since there are 15 TPR motifs in Syf1, I then asked whether these motifs, especially at the C-terminal region of Syf1, are interchangeable. Within the C-terminus of Syf1, the 13th of the 15 TPR motifs seems to be specifically crucial for the interaction since additional deletion of only 13th TPR motif (*syf1-Δ13*) leads to synthetic lethal with *Δhpr1*, while deletion of the two most C-terminal TPR motifs (XIV and XV, *syf1-Δ14*) does not cause any genetic interaction with *Δhpr1*. To answer this question, I constructed fusion mutants of six different combinations of two other TPR motifs (I+II, III+IV, VI+VII, VII+VIII, IX +X, or XI+XII) of Syf1 to the *syf1-Δ13* deletion mutant (Figure 2.30). Surprisingly, all the mutants are still synthetic lethal with *Δhpr1*. Therefore, the mere addition of TPR motifs to the 3'-end of the truncated Syf1 does not rescue functionality of Syf1 in transcription indicating that its C-terminal TPR motifs are not interchangeable.

Similarly, deletion of the N-terminal five TPR motifs of Syf1 (I-V) does not cause a synthetic lethality with *Δhpr1* (Figure 2.31). This results indicates that these TPR motifs are most likely not crucial for the physical interaction between the Prp19 complex and RNAPII as well as the function of the Prp19 complex in transcription. Interestingly, however, when the 6th TPR motif of

Syf1 is deleted, the cells become lethal indicating its essential role for the cells.

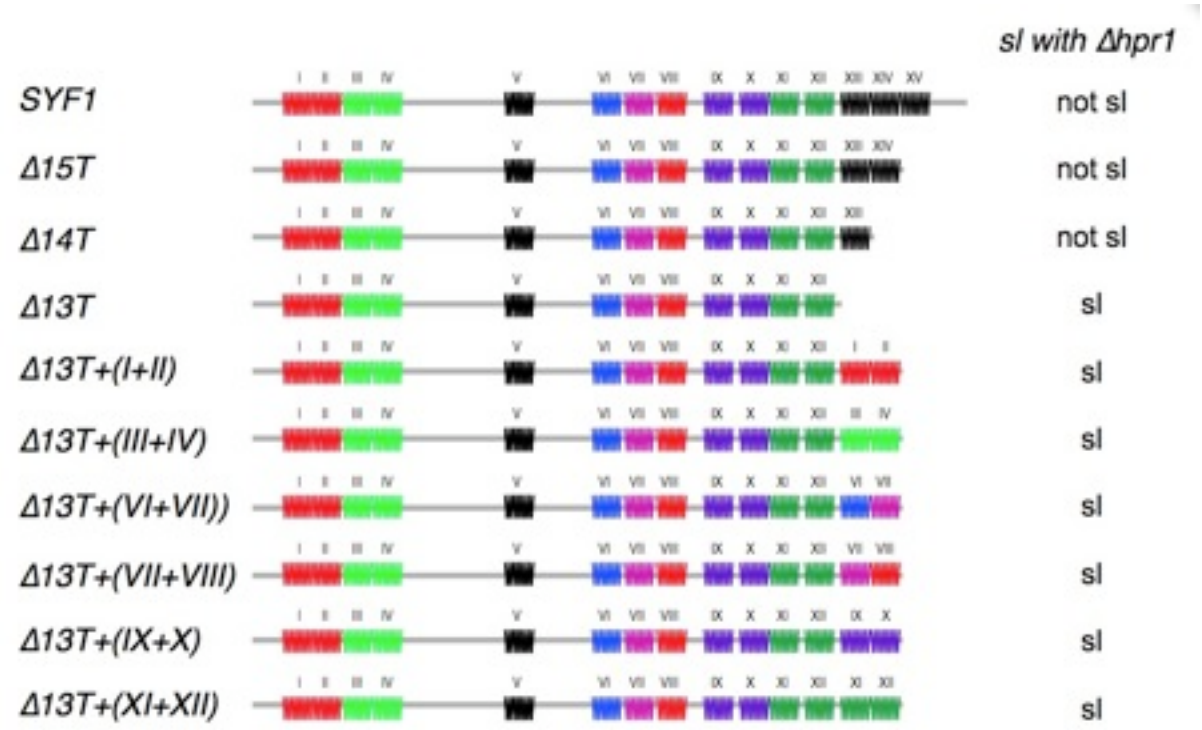

Figure 2.30: The C-terminal TPR motifs of Syf1 are important for the interaction of the Prp19 complex with RNAPII and are not interchangeable

Addition of six different combinations of two other TPR motifs of Syf1 to the syf1-*Δ13* deletion mutation cannot rescue synthetic lethality with *Δhpr1.*

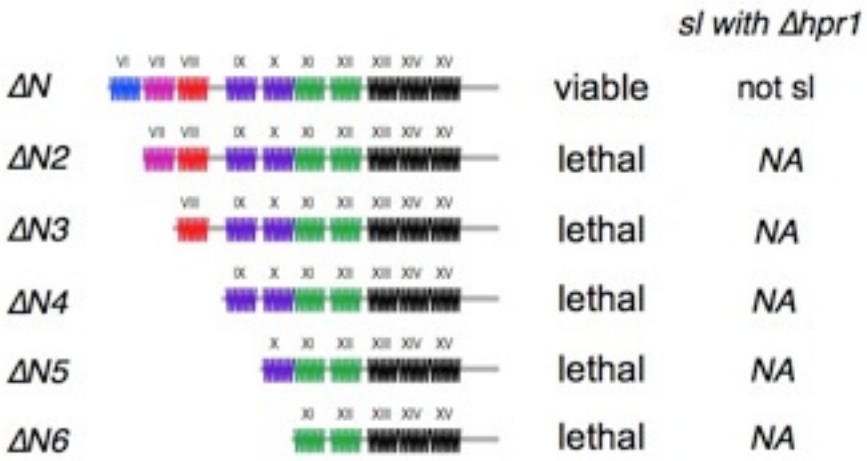

Figure 2.31: The deletion of N-terminal TPR motifs of Syf1 does not lead to synthetic lethality with *Δhpr1*

2.9 The C-terminus of Syf1 and the interaction with Yju2

The C-terminus of Syf1 seems to be important for an interaction between the Prp19 complex and RNAPII. Interestingly, it has been reported recently that the C-terminus of Syf1 (484-754 amino acids) interacts with Yju2, a novel essential splicing factor which promotes the first catalytic reaction of splicing (Chang et al. 2009). Therefore, the interaction of Yju2 with Prp19 complex could be impaired in *syf1-37* cells, expressing Syf1 that lacks the four most C-terminal TPR motifs, and as a result transcription efficiency is decreased.

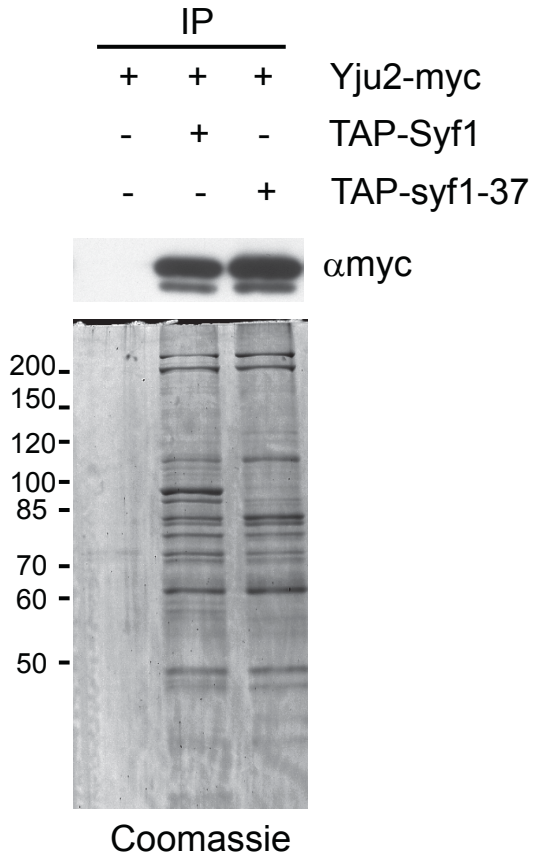

Figure 2.32: Yju2 associates with Prp19 complex in *syf1-37* cells
Prp19 complex was purified using either TAP-Syf1 or TAP-syf1-37 and copurification of myc-tagged Yju2 assessed by western blotting with α-myc antibody (α-myc). A non-tagged strain served as negative control. IP: immunoprecipitation, Coomassie: TEV eluate of the TAP purification.

To answer these questions, first Prp19 complex was purified using a strain expressing either TAP-Syf1 or TAP-syf1-37 and copurification of myc-tagged Yju2 was assessed by western blotting. As shown in Figure 2.32, Yju2 copurified with the Prp19 complex regardless of presence or absence of the C-terminus of Syf1.

Second, to test whether *YJU2* is implicated in transcription, I tested *yju2*-ts cell for a sensitivity to 6-AU, a drug that impairs transcription elongation. As shown in Figure 2.33, *yju2* ts cell did not confer 6-AU sensitivity. These results suggest that *YJU2* might not involve in transcription and is most likely not the cause of the impaired transcription in the *syf1-37* cells.

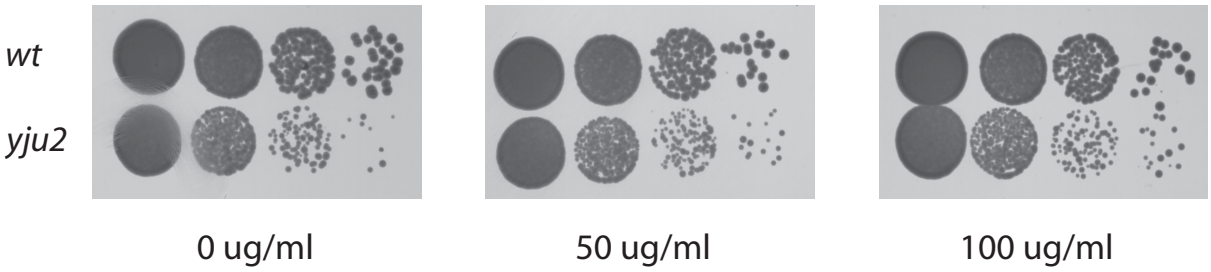

Figure 2.33: *yju2* ts cell did not confer 6-AU sensitivity
10-fold serial dilutions of yeast cells expressing wild-type *YJU2* or *yju2* ts allele were spotted onto SDC-(ura) plates containing solvent (DMSO) or 50 or 100 µg/ml 6-AU and incubated for 2 days at 30°C.

2.10 The *syf1-37* mutation does not lead to an mRNA export defect

The genetic interaction between *SYF1* and the THO complex and the reduced recruitment of the TREX components in *syf1-37* cells suggest a role of Syf1 in mRNA export. Therefore I assessed an mRNA export defect in *syf1-37* cells. As a positive control for nuclear accumulation of mRNA *Δhpr1* cells were used. While deletion of the *HPR1* gene led to a block of mRNA export at 37°C,

syf1-37 cells showed no defect on mRNA export (Figure 2.34). This result suggests that although the recruitment of TREX to transcribed genes in the *syf1-37* cells is reduced, it is still sufficient for effective mRNA export to the cytoplasm.

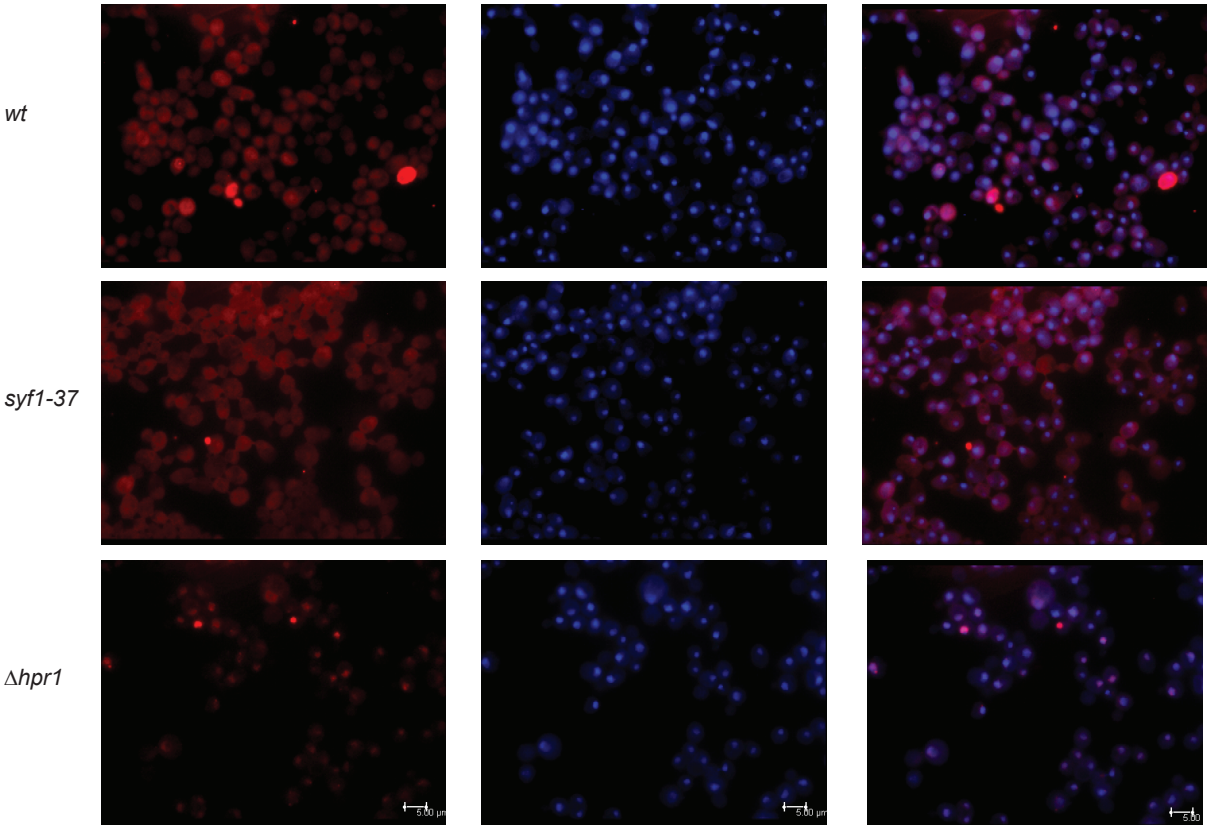

Figure 2.34: The *syf1-37* cells do not display an mRNA export defect
Cells were grown at 37°C for three hours and oligo-dT *in situ* hybridization was performed. The poly(A)-tail of bulk mRNA was visualized by using a Cy3-labelled oligo(dT) probe and nuclei of the cells were stained with 4',6-diamidino-2-phenylindole (DAPI). While *Δhpr1* cells display mRNA export defect, wild-type (wt) and *syf1-37* cells do not show any mRNA accumulation in the nucleus.

2.11 Transcription impairment in the *syf1-37* cells does not lead to a hyperrecombination phenotype

It has been shown previously that the mutation of the THO complex impairs transcription *in vivo*. In these mutants, the transcribed mRNA can form RNA-DNA hybrid, so-called R-loop, which links an impairment of both transcription and transcription-associated recombination. Moreover, this hyper-recombination is only observed when transcription goes through either the DNA repeats or the intervening regions flanked by the repeats, suggesting that transcription in crucial for the hyperrecombination phenotype.

Since transcription is impaired in *syf1-37* cells, I asked whether this transcription impairment would cause transcription-associated recombination. To answer this question, a hyperrecombination assay was performed by using the L-*lacZ* and L-*PHO5* system, containing either *lacZ* or *PHO5* gene flanked by *leu2* direct repeats (Figure 2.35, upper panel) under the control of constitutive *LEU2* promoter. In the absence of the THO complex, R-loops are formed due to impaired transcription elongation and as a result the mutant shows a hyperrecombination phenotype (Figure 2.35, lower panel, *Δhpr1*). Importantly, though the *syf1-37* mutant displays a transcription defect, its recombination rates are comparable to wild-type (Figure 2.35, lower panel, *syf1-37*). This result suggests that although the recruitment of TREX to transcribed genes and the transcription efficiency in the *syf1-37* cells are reduced, the amount of TREX is probably sufficient for maintaining the genomic stability.

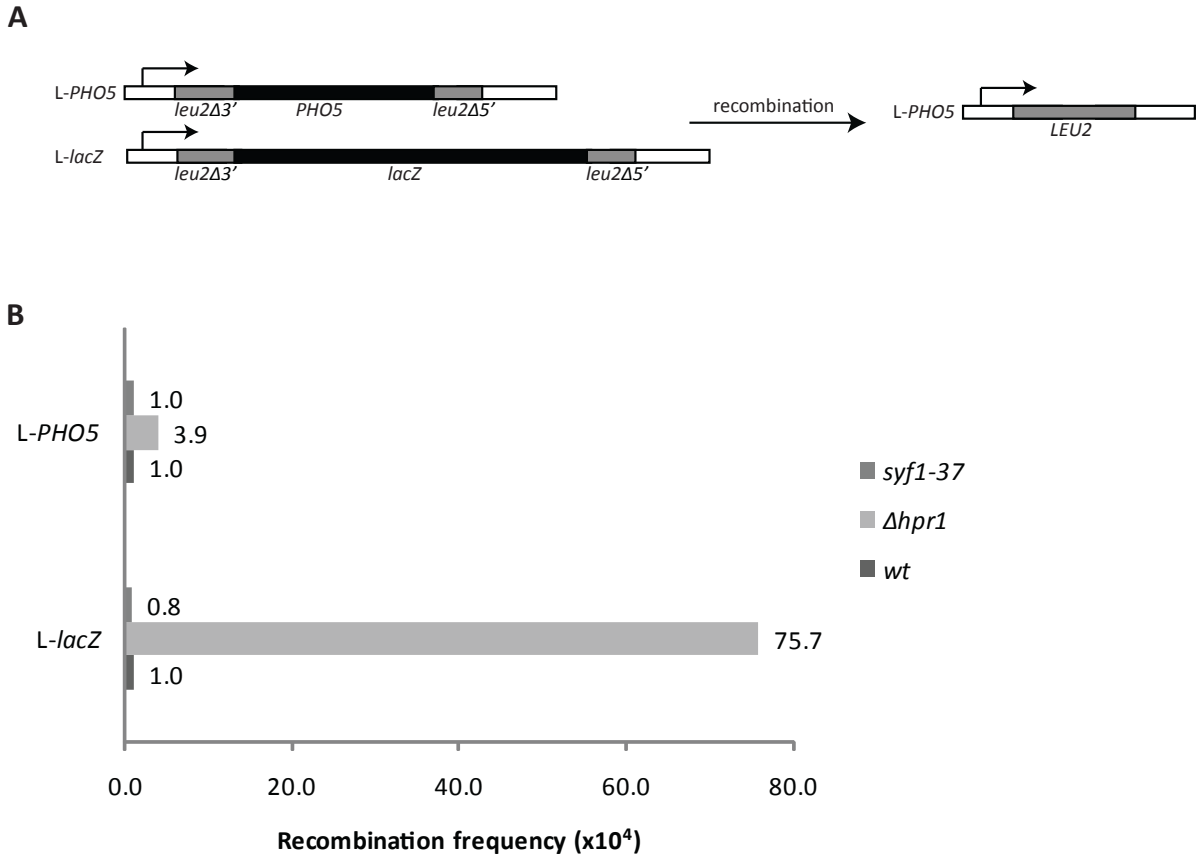

Figure 2.35: Recombination frequencies of wild-type, *syf1-37* and *Δhpr1* cells using direct repeat systems

(A) Schematic diagram of the deletion resulting from recombination between the direct repeats used. (B) Recombination frequencies in wild-type, *syf1-37* and *Δhpr1* cells transformed with the plasmid containing either L-*PHO5* or L-*lacZ* construct.

3. Discussion

3.1 The Prp19 complex functions in transcription elongation and is necessary for TREX occupancy at transcribed genes

In the last decade, the Prp19 complex has been studied extensively in its conserved function, *i.e.* splicing. It forms a heteromeric complex which is also known as the nineteen complex (NTC) in budding yeast and the hPrp19/CDC5L complex in higher cells. The Prp19 complex is one of several non-snRNPs which play a crucial role in splicing and is necessary for the activation of the spliceosome.

Here I show that in budding yeast the Prp19 complex has an additional role in gene expression, *i.e.* transcription elongation. The complex is recruited to both intron-containing and intronless genes in a transcription-dependent manner. Furthermore, mutation of *SYF1*, an essential and conserved component of the Prp19 complex, causes a decreased occupancy of the whole Prp19 complex at transcribed genes and reduces transcriptional activity. Moreover, the *syf1-37* mutation reduces RNAPII processivity indicating a role in transcription elongation.

In addition, I identify *SYF1* as the first factor required for occupancy of the TREX complex since TREX recruitment to genes is specifically decreased in *syf1-37* cells. However, for some genes, *e.g. PMA1* and *DBP2*, the Prp19 complex is not necessary for TREX recruitment at the 5'-end. Therefore, there might be other factors responsible for TREX recruitment to genes at these sites. Notably, the occupancy of two other transcription elongation factors, *i.e.* Paf1 and Spt5, is not affected by the mutation of *SYF1*, indicating that the reduction of TREX occupancy is specific to *SYF1* mutation and most likely to

the C-terminal region of Syf1. It is interesting to remark that the PAF complex leaves the transcribing RNAPII downstream of the polyadenylation site as TREX does (Kim et al. 2004), but apparently depends on other factors than *SYF1* or TREX for recruitment. It has been shown that THO is important for efficient transcription elongation by ensuring the stability of the transcription machinery on the chromatin template (Mason and Struhl 2005). Consistent with a requirement of the THO complex for RNAPII processivity, the *syf1-37* cells show impaired processivity of RNAPII in the presence of 6-AU. Thus, the decreased THO/TREX occupancy in the *syf1* mutant cells seems to be sufficient for efficient RNAPII processivity under normal conditions, but causes impaired RNAPII processivity when transcription is blocked in the presence of 6-AU.

Interestingly, even though the occupancy of the components of the TREX complex Hpr1, Sub2 and Yra1 is decreased to approximately 50% in the *syf1-37* cells compared to the wild-type, the defect on nuclear mRNA export and the hyperrecombination phenotype are not observed suggesting that the amount of TREX is still sufficient for its functions.

3.2 Prp19 E3 ligase activity and transcription

Interestingly, Prp19 contains a U-box domain and has E3 ubiquitin ligase activity *in vitro* (Hatakeyama et al. 2001; Ohi et al. 2003). It has been shown recently that this catalytic activity of Prp19 is crucial for structural rearrangements during the splicing reaction (Song et al. 2010). Prp19 ubiquitylates Prp3, a component of U4 snRNP, where non-proteolytic K63-linked chains are formed. The modification of Prp3 enhances the interaction between Prp3 and Prp8, the component of U5 snRNP, and consequently stabilizes the U4/U6.U5 tri-snRNP. In the *syf1-37* mutant, the interaction

between Prp19 complex and the components of U5 snRNP, *e.g.* Prp8 and Brr2, does not seem to be affected as shown in Figure 2.23. Nevertheless, it is still questionable that, in addition to its structural function to stabilize the recruitment of TREX to genes, the E3 ligase activity of Prp19 might be essential for its role in transcription.

Remarkably, one of the component of the TREX complex Hpr1 is polyubuiquitylated and degraded. Interestingly, though, the ubiquitylated form of Hpr1 interacts with the ubiquitin-associated (UBA) domain of the mRNA export receptor Mex67. This interaction recruits Mex67 to transcribed genes and transiently protects Hpr1 from degradation by the proteasome (Gwizdek et al. 2006). However, when I purified Hpr1 from *SYF1* and *syf1-37* cells and quantified Hpr1 ubiquitylation, the modification was not affected. This result is consistent with the idea that the E3 ligase activity of Prp19 is not affected in *syf1-37* cells. Therefore, it is still a challenge whether the E3 ligase activity of Prp19 is crucial for its function in transcription as well as what its substrate(s) might be.

3.3 The TPR motifs of Syf1

The TPR motif displays a large degree of sequence diversity but structural comparison shows a highly conserved three dimensional structure consisting of two anti-parallel alpha helices separated by a turn. Syf1 contains 15 TPR motifs and interestingly the motifs are not interchangeable. This suggests that the different TPR motifs of Syf1, especially at the C-terminus, might contain a specific residue important for a post-translational modification and/or for protein-protein interaction and consequently for its function. Thus, it is still an open question in terms of what the difference between each TPR motif is and how it regulates the function of Syf1.

Interestingly, it has been shown recently that the C-terminus of Syf1 interacts with Yju2, a novel splicing factor which promotes the first catalytic reaction of splicing (Chang et al. 2009). However, Yju2 binds to syf1-37, which lacks the four most C-terminal TPR motifs, and does not confer 6-AU sensitivity. Other regions of the C-terminus of Syf1 might play a role in this interaction. These results are consistent with the fact that in *syf1-37* cells the Prp19 complex is still intact and splicing not defective.

3.4 The function of the Prp19 complex in recruiting TREX to genes might be conserved among eukaryotes

In higher eukaryotic cells, where splicing occurs at most sites of mRNA genes, TREX is recruited to the mRNA during splicing (Masuda et al. 2005; Wahl et al. 2009). In yeast, however, TREX is recruited in a transcription-dependent manner. Here I show that the Prp19 complex, so far known as a component of the splicing machinery, recruits TREX to genes in yeast.

During splicing, a highly dynamic molecular machine, the spliceosome, is assembled in a stepwise manner onto the mRNA and forms several intermediates called complex E, A, B, B* and C ((Wahl et al. 2009) and references therein). Briefly, the U1 snRNP binds 5'-splice site and forms the E complex. Then the U2 snRNP recognizes the branch point sequence of the pre-mRNA and forms the A complex. Subsequently, binding of the pre-assembled U4/U6.U5 tri-snRNP leads to the B complex. Additional factors as well as extensive structural rearrangements destabilizing the association of U1 and U4 snRNPs with the catalytic core of spliceosome forms the catalytically activated B* complex. After the first transesterification reaction, the C complex is formed and catalyzes the second step of splicing.

Interestingly, the hPrp19 and its associated complexes containing XAB2/hSyf1 are present in the B complex while TREX components, such as THOC1/hHpr1 and UAP56/hSub2, are found only later when the C complex is formed. As the Prp19 complex is part of the spliceosome before the presence of TREX, it might recruit or help recruit the TREX complex to the spliceosome during the splicing reaction. Therefore, the role of the Prp19 complex in recruiting TREX to the mRNA might be conserved in higher eukaryotes.

In summary, I have identified the Prp19 complex as a novel transcription elongation factor which is the first factor essential for TREX occupancy at transcribed genes.

4. Materials

4.1 Consumables and chemicals

Consumables and chemicals were purchased from the following companies: Acros Organics (Geel, Belgium), Applichem (Darmstadt), Applied Biosciences (Darmstadt), Apollo Scientific Limited (Bredbury, UK), Axon (Kaiserslautern), Becton Dickinson, (Heidelberg), Beckman Coulter (Krefeld), Biaffin (Kassel), Biomol (Hamburg), Biorad (Munich), Biozym (Hess. Oldendorf), Chemicon (Temecula, Canada), Fermentas (St. Leon- Rot), Formedium (Norwich, UK), GE Healthcare (München), Gilson (Bad Camberg), Invitrogen (Karlsruhe), Macherey & Nagel (Düren), Medac (Hamburg), Medigenomix (München), Membra Pure (Bodenheim), Merck Biosciences (Darmstadt), Millipore (Molsheim, France), Mobitec (Göttingen), MP Biomedical (Illkirch, France), NEB (Frankfurt), Neolab (Heidelberg), Nunc (Wiesbaden), Peske (Aindling-Arnhofen), Promega (Mannheim), Qiagen (Hilden), Roche (Mannheim), Roth (Karlsruhe), Santa Cruz (Santa Cruz, USA), Sarstedt (Nümbrecht), Semadeni (Düsseldorf), Serva (Heidelberg), Sigma (Taufkirchen), Stratagene (Amsterdam, The Netherlands) and VWR (Ismaning).

4.2 Commercially available kits

The following commercially available kits were purchased from these companies: DC Protein Assay (Biorad, München), ECL kit (Applichem, Darmstadt), Nucleobond AX PC100, Nucleospin Mini, Nucleospin extract (all Macherey & Nagel, Düren), RNeasy MinElute Cleanup Kit (Quiagen, Hilden), and Phusion Hot Start II High-Fidelity DNA Polymerase (Finzymes, Finland)

4.3 Equipments

The following commercially available equipments were purchased from these companies: Beckman DU650 spectrophotometer, L80 Ultracentrifuge, SW32 rotor, SW40 rotor, (Beckman Coulter, Krefeld), T3 Thermocycler (Biometra, Goettingen), Gel dryer model 583, Mini-Protean II system (Biorad, München), DNA engine (Biozym, Hess Oldendorf), Lumat LB9507 (EG&G Berthold), bead mill (Fritsch, Idar-Oberstein), Thermomixer compact, Eppendorf centrifuge 5415D, 5415R (Eppendorf, Wesseling-Berzdorf), peristaltic pump (GE Healthcare, München), Rotanda, 46R (Hettich, Tuttlingen), Ika Vibrax VXR basic (Ika, Staufen), Heraeus Hera Freeze (Kendro, Langenselbold), Kodak X omat M35 (Kodak), Kühner ISF-1-V (Kühner AG, Switzerland), Elektrophoresis Power Supply Consort E835, Heidolph shaker duomax 1030, Rotator, Thermostatcabinet Aqualytic, Vortex Genie (Neolab, Heidelberg), Innova 44 shaking incubator (New Brunswick Scientific, Nürtingen), Olympus BX60 fluorescence microscope (Olympus, Hamburg), Semi-dry blotting device (Peqlab, Erlangen), Tri Carb Liquid Scintillation Analyzer (Perkin Elmer, Rodgau), Dissection microscope manual MSM (Singer, Somerset, UK), RP80AT-364, Sorvall Evolution, Sorvall RCM120Ex, SLC6000, SS34 rotor (Thermo Fisher Scientific), Image reader LAS-3000 (Fujifilm), STORM 860 and Typhoon 9400 (GE Healthcare)

4.4 Radioactivity

Radioactivity, [γ-^{32}P]-ATP (370 MBq/ml, 10 mCi/ml), was obtained from GE Healthcare.

4.5 Enzymes

Calf Intestine Alkaline Phosphatase (1 U/µl; Fermentas, St. Leon-Rot), restriction endonucleases (BamHI, EcoRI, HindIII, NheI, NotI, SphI, PstI, SacI, SalI, SpeI, SphI, XbaI, XhoI; Fermentas, St. Leon-Rot), RevertAid H Minus M-MuLV Reverse Transcriptase (Fermentas, St. Leon-Rot), T4-DNA Ligase (Fermentas, St. Leon-Rot), RiboLock Ribonuclease Inhibitor (Fermentas, St. Leon-Rot), Micrococcal Nuclease (Roche, Mannheim) and Zymolyase 20000T/100000T (Medac, Hamburg)

4.6 Antibodies

	Source	Dilution	Company
Primary antibodies			
α-Pgk1	mouse	1:1000	Molecular Probes
α-HA	mouse	1:1000	Roche Applied Science
α-Npl3	rabbit	1:5000	Siebel and Guthrie, 1996
α-Rpb1 (N-term; yN-18)	goat	1:1000	Santa Cruz
α-Rpb1 (CTD; 8WG16)	mouse	1:1000	Covance
Peroxidase α-Peroxidase complex (PAP)	rabbit	1:5000	Sigma
Secondary antibodies			
α-mouse IgG HRPO	goat	1:3000	Biorad
α-rabbit IgG HRPO	goat	1:3000	Biorad
α-goat IgG HRPO	rabbit	1:5000	Sigma

4.7 Oligonucleotides

Oligo	Sequence	Description
Cloning primers		
TAP-PstI-f	ATTCTGCAGATGGCAGGCCTTGCGCAAC	for cloning pNOPTAP
TAP-BamHI-r	ATTGGATCCAAGTGCCCCGGAGGATGAGA	for cloning pNOPTAP
pNOP-XhoI-f	ATTCTCGAGTCCTGTACTTGCTGATATTGTGG	for cloning pNOPTAP
pNOP-PstI-r	ATTCTGCAGTACTGTTTTAGTTGATTTGAGTT	for cloning pNOPTAP
SpeI-*SYF1*-f	TAGACTAGTTGTGCATAGAGCCCTTGTCT	for cloning *SYF1* into pRS315
SYF1-XhoI-r	ATGCTCGAGAACCTTCGGAAAGGGAAAAA	for cloning *SYF1* into pRS315
BamHI-*SYF1*-f	TCGGAAAGTTTGGGATCCTG	for cloning C-terminal truncation of *SYF1*
NheI-*SYF1*-821-r	ATTGCTAGCTACTCATGGCTAACACTAGCTGAATCAATGAGCATG	for cloning C-terminal truncation of *SYF1*
NheI-*SYF1*-794-r	AATGCTAGCTTACTTATCCCCATGTTTTAACTC	for cloning C-terminal truncation of *SYF1*
NheI-*SYF1*-755-r	AATGCTAGCTATATAGAGCTCTCGAAATCTG	for cloning C-terminal truncation of *SYF1*
NheI-*SYF1*-719-r	AATGCTAGCTAAGTAGACTCGGCCTTGGAAA	for cloning C-terminal truncation of *SYF1*
NheI-*SYF1*-679-r	AATGCTAGCTACAGACCACTTATCCTTTCTT	for cloning C-terminal truncation of *SYF1*
SYF1-SalI-f	ATTGTCGACATGTCAGCATACATCGCAATG	for subcloning *SYF1* and *syf1-37* into pNOPTAP
SYF1-SpeI-rev	ATACTAGTAAAGCGGCTAGAGGGGACT	for subcloning *SYF1* and *syf1-37* into pNOPTAP

Oligo	Sequence	Description
GAL10-prom-SalI-f	ATCGTCGACTGGTTATGCAGCTTTTCCATT	for cloning pRS316-*GAL10-ACT1*
GAL10-prom-SpeI-r	ACTACTAGTCTTTCACTTTGTAACTGAGCTGTC	for cloning pRS316-*GAL10-ACT1*
ACT1-cds-SpeI-f	TATACTAGTATGGATTCTGGTATGTTCTAGCG	for cloning pRS316-*GAL10-ACT1*
ACT1-cds-NotI-r	ATTGCGGCCGCTCTTCTGGGGCAACTCTCAA	for cloning pRS316-*GAL10-ACT1*
TAP-tagging primers		
NTC20-f	GAATAATAAGGAAAAGGGTATTACAGGAACCCGACCGTGACAATGATGATAGCGGCTCCATGGAAAAGAGAAG	for TAP tagging *NTC20*
NTC20-r	CTTTCCCAGCCAACCCAATGAAGTCTTGCGCACGCAATTGGTTGTAATTATCCGATACGACTCACTATAGGG	for TAP tagging *NTC20*
PRP19-f	CCGCTATTCTGAAGACAAATGATAGTTTCAATATTGTTGCATTGACACCCTCCATGGAAAAGAGAAG	for TAP tagging *PRP19*
PRP19-r	GTTACTACTATTACACAGGTTTATTTAGAAAGTACAAACGTGTCAGCGTATTACGACTCACTATAGGG	for TAP tagging *PRP19*
PAF1-f	GTTCATACTGAACAAAAACCAGAGGAAGAAAAGGAAACTTTACAAGAAGAATCCATGGAAAAGAGAAG	for TAP tagging *PAF1*
PAF1-r	GTAAAAAGAACTACAGGTTTAAAATCAATCTCCCTTCACTTCTCAATATTTACGACTCACTATAGGG	for TAP tagging *PAF1*

Oligo	Sequence	Description
SPT5-f	CAACCAAGGAAATAAGTCAAACTATGGTGGTAACAGTACATGGGGAGGTCATTCCATGGAAAAGAGAAG	for TAP tagging *SPT5*
SPT5-r	GTCTTTTTTATTGATTTCTTCTTGGGTGATATTGGTTCTCCTTTTGGTGATACGACTCACTATAGGG	for TAP tagging *SPT5*
Primer extension oligos		
U1	CAAGGGAATGGAAACGTCAG	*in vivo* transcription assay
SCR1	TTTACGACGGAGGAAAGACG	*in vivo* transcription assay
GAL10	TTCCAGACCTTTTCGGTCAC	*in vivo* transcription assay
ACT1	GTTTTCCCAGTCACGACGTT	*in vivo* transcription assay
RT-PCR primers		
NTR-f	TGCGTACAAAAAGTGTCAAGAGATT	ChIP experiment
NTR-r	ATGCGCAAGAAGGTGCCTAT	ChIP experiment
ADH1-5'-f	GTTGTCGGCATGGGTGAAA	ChIP experiment
ADH1-5'-r	GGCGTAGTCACCGATCTTCC	ChIP experiment
ADH1-M-f	AGCCGCTCACATTCCTCAAG	ChIP experiment
ADH1-M-r	ACGGTGATACCAGCACACAAGA	ChIP experiment
ADH1-3'-f	TTGGACTTCTTCGCCAGAGG	ChIP experiment
ADH1-3'-f	GCCGACAACCTTGATTGGAG	ChIP experiment
PGK1-5'-r	TTGCCAACCATCAAGTACGTTT	ChIP experiment
PGK1-5'-r	CCCAAGTGAGAAGCCAAGACA	ChIP experiment
PGK1-3'-f	GAAGACGGCACCAGCCAAT	ChIP experiment

Oligo	Sequence	Description
PGK1-3'-r	TGACAAGATCTCCCATGTCTCTACTG	ChIP experiment
PMA1-M-f	GTTTTTCGTCGGTCCAATTCA	ChIP experiment
PMA1-M-r	AACCGGCAGCCAAAATAGC	ChIP experiment
PMA1-3'-f	CAGAGCTGCTGGTCCATTCTG	ChIP experiment
PMA1-3'-r	GAAGACGGCACCAGCCAAT	ChIP experiment
PDR5-5'-f	TCTCCATCTGTGGATCCAAATTT	ChIP experiment
PDR5-5'-r	CGGAATCCAAACCCCTTGT	ChIP experiment
PDR5-3'-f	GGTGTTGCTAACGTAGACGTCAA	ChIP experiment
PDR5-3'-r	CGGATGGTGGTGTGAATTCTAG	ChIP experiment
ACT1-5'-f	TTTTCACGCTTACTGCTTTTTTCTT	ChIP experiment
ACT1-5'-r	GCTAGAACATACCAGAATCCATTGTTAA	ChIP experiment
ACT1-3'-f	TCAGAGCCCCAGAAGCTTTG	ChIP experiment
ACT1-3'-r	TTGGTCAATACCGGCAGATTC	ChIP experiment
DBP2-5'-f	CCAAAGCCAATCACCACTTTC	ChIP experiment
DBP2-5'-r	AGCCTTCACTTCATTCAAAACGT	ChIP experiment
DBP2-3'-f	CTTCACCGAACAAAACAAAGGTT	ChIP experiment
DBP2-3'-r	TCGGGAGGAATATTTTGATTAGCT	ChIP experiment
YLR454-p-f	TGGTACCGTCAGGCTAAAATCC	ChIP experiment
YLR454-p-r	TCGTAGACCCTTGATGAAAAGCT	ChIP experiment
YLR454-1000-f	GATCTCGTGTATGCAATGGAAAAA	ChIP experiment

Oligo	Sequence	Description
YLR454-1000-r	ACGGTGACGTTCATTTCTTTCA	ChIP experiment
YLR454-2000-f	CGGTTTTTATGCCAAGAGATGTT	ChIP experiment
YLR454-2000-r	GCCATCAATTAGGTCCTTTGAACT	ChIP experiment
YLR454-4000-f	GCCTTCATGTGGCATTGAAA	ChIP experiment
YLR454-4000-r	TGTATTCTTCGGAACGCCTTTT	ChIP experiment
YLR454-8000-f	GGAGAATAAAATTGCTGACGACACT	ChIP experiment
YLR454-8000-r	GGCTCCGTGTAGGAATTATAAGTTG	ChIP experiment
GAL1-5'-f	GTGACTTCTCGGTTTTACCTTTAGC	ChIP experiment
GAL1-5'-r	AATGGATGGATTTTTCTCGTTCA	ChIP experiment
GAL1-3'-f	GGTGGTTGTACTGTTCACTTGGTT	ChIP experiment
GAL1-3'-r	TCATTGGCAAGGGCTTCTTT	ChIP experiment

4.8 Plasmids

Plasmid	Description	Reference
pBS1479	plasmid used for C-terminal TAP-tagging of a protein by genomic integration	(Puig, Caspary et al. 2001)
pNOPPATA1L	plasmid containing the *NOP1* promoter, sequences coding for protein A and the TEV cleavage site, and the *ADH1* terminator	Hellmuth, K. (2002)
pRS314	an *ARS-CEN* vector with a *TRP1* marker	(Sikorski and Hieter 1989)
pRS315	an *ARS-CEN* vector with a *LEU2* marker	(Sikorski and Hieter 1989)
pRS314-*MFT1*	the coding sequence of *MFT1* plus about 500 bp downstream and upstream of the ORF was amplified by PCR creating a SalI site and a BamHI site and cloned into the same sites of pRS314	(Hurt, Luo et al. 2004)
pRS314-*HPR1*	the coding sequence of *HPR1* plus about 500 bp downstream and upstream of the ORF was amplified by PCR creating a SacI site and a HindIII site and cloned into the same sites of pRS314	(Strasser, Masuda et al. 2002)

Plasmid	Description	Reference
pNOPTAP	the promoter sequence of *NOP1* was PCR amplified from pNOPPATA1L adding XhoI and PstI sites and the coding sequence of protA-TEV-CBP was PCR amplified from pBS1761 adding PstI and BamHI sites and cloned into XhoI and BamHI sites of pRS315	this study
pRS315-*SYF1*	a HincII-SpeI fragment of a genomic clone containing *SYF1* was cloned into the SmaI and SpeI sites of pRS315	this study
pRS315-*syf1-Δ12*	the sequence encoding amino acids 1-679 of *SYF1* was PCR amplified from pRS315-*SYF1* adding BamHI and NheI and cloned into the same sites of pRS315-*SYF1*	this study
pRS315-*syf1-Δ13*	the sequence encoding amino acids 1-719 of *SYF1* was PCR amplified from pRS315-*SYF1* adding BamHI and NheI and cloned into the same sites of pRS315-*SYF1*	this study
pRS315-*syf1-Δ14*	the sequence encoding amino acids 1-755 of *SYF1* was PCR amplified from pRS315-*SYF1* adding BamHI and NheI and cloned into the same sites of pRS315-*SYF1*	this study

Plasmid	Description	Reference
pRS315-*syf1-Δ15*	the sequence encoding amino acids 1-794 of *SYF1* was PCR amplified from pRS315-*SYF1* adding BamHI and NheI and cloned into the same sites of pRS315-*SYF1*	this study
pRS315-*syf1-ΔC*	the sequence encoding amino acids 1-821 of *SYF1* was PCR amplified from pRS315-*SYF1* adding BamHI and NheI and cloned into the same sites of pRS315-*SYF1*	this study
pRS315-*syf1-1*	as pRS315-*SYF1* but carrying *syf1-1*	this study
pRS315-*syf1-2*	as pRS315-*SYF1* but carrying *syf1-2*	this study
pRS315-*syf1-15*	as pRS315-*SYF1* but carrying *syf1-15*	this study
pRS315-*syf1-34*	as pRS315-*SYF1* but carrying *syf1-34*	this study
pRS315-*syf1-37*	as pRS315-*SYF1* but carrying *syf1-37*	this study
pNOPTAP-*SYF1*	the coding sequence of *SYF1* was PCR amplified from pRS315-*SYF1* adding SpeI site was cloned into the same site of pNOPTAP	this study

Plasmid	Description	Reference
pNOPTAP-*syf1-37*	the coding sequence of *syf1-37* was PCR amplified from pRS315-*syf1-37* adding SpeI site and cloned into the same site of pNOPTAP	this study
pRS316-*GAL10-ACT1*	the promoter sequence of *GAL10* was PCR amplified from genomic DNA adding SalI and SpeI sites and the sequence of the *ACT1* ORF PCR amplified from genomic DNA adding SpeI and NotI and cloned into the SalI and NotI sites of pRS316	this study

4.9 Strains

E. coli strains

DH5α F$^-$, *ϕ*80d*lacZ*ΔM15, Δ(*lacZYA*-*argF*)U169, *deoR*, *recA1*, *endA1*, *hsdR*17(rk$^-$, mk$^+$), *phoA*, *supE44*, *λ*$^-$, *thi*-1, *gyrA96*, *rel*A1

S. cerevisiae strains

Strain	Genotype	Reference
W303	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535*	(Thomas and Rothstein 1989)
BY4741	*MAT a; his3Δ1 leu2Δ met15Δ ura3Δ*	EUROSCARF
Δhpr1 SYF1 shuffle	*MAT a; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; hpr1::HIS3; syf1::kanMX4;* pRS316-*SYF1*	this study
Δmft1 SYF1 shuffle	*MAT a; ade2-1; his3-11,15; ura3-52; leu2-3,112; trp1-1; mft1::kanMX4; syf1::kanMX4;* pRS316-*SYF1*	this study
GAL1::YLR454	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; TRP1-KL::GAL1-YLR454*	(Mason and Struhl 2005)
GAL1::YLR454 RPB3-TAP	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; RPB3-TAP::TRP1-KL; TRP1-KL::GAL1-YLR454*	this study
GAL1::YLR454 RPB3-TAP Δhpr1	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; RPB3-TAP::TRP1-KL; TRP1-KL::GAL1-YLR454; hpr1::HIS3*	this study

Strain	Genotype	Reference
GAL1::YLR454 RPB3-TAP SYF1	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; RPB3-TAP::TRP1-KL; TRP1-KL::GAL1-YLR454; syf1::kanMX4;* pRS315-*SYF1*	this study
GAL1::YLR454 RPB3-TAP syf1-37	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; RPB3-TAP::TRP1-KL; TRP1-KL::GAL1-YLR454; syf1::kanMX4;* pRS315-*syf1-37*	this study
HPR1-HA TAP-syf1-37	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; syf1::kanMX4; HPR1-3HA::HIS3;* pNOPTAP-*syf1-37*	this study
HPR1-TAP SYF1	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; HPR1-AP::TRP1-KL; syf1::kanMX4;* pRS315-*SYF1*	this study
HPR1-TAP syf1-37	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; HPR1-TAP::TRP1-KL; syf1::kanMX4;* pRS315-*syf1-37*	this study
HPR1-TAP Syf1-HA	*MAT alpha; ade2-1; his3-11, 15; ura3-52; leu2-3, 112; trp1-1; HPR1-TAP::TRP1-KL; SYF1-HA::kanMX4*	this study
NTC20-TAP SYF1	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; NTC20-TAP::TRP1-KL; syf1::kanMX4;* pRS315-*SYF1*	this study
NTC20-TAP syf1-37	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; NTC20-TAP::TRP1-KL; syf1::kanMX4;* pRS315-*syf1-37*	this study

Strain	Genotype	Reference
PAF1-TAP SYF1	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; PAF1-TAP::TRP1-KL; syf1::kanMX4;* pRS315-*SYF1*	this study
PAF1-TAP syf1-37	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; PAF1-TAP::TRP1-KL; syf1::kanMX4;* pRS315-*syf1-37*	this study
PRP19-TAP SYF1	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; PRP19-TAP::TRP1-KL; syf1::kanMX4;* pRS315-*SYF1*	this study
PRP19-TAP syf1-37	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; PRP19-TAP::TRP1-KL; syf1::kanMX4;* pRS315-*syf1-37*	this study
RPB3-TAP SYF1	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; RPB3-TAP::TRP1-KL; syf1::kanMX4;* pRS315-*SYF1*	this study
RPB3-TAP syf1-37	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; RPB3-TAP::TRP1-KL; syf1::kanMX4;* pRS315-*syf1-37*	this study
SPT5-TAP SYF1	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; SPT5-TAP::TRP1-KL; syf1::kanMX4;* pRS315-*SYF1*	this study
SPT5-TAP syf1-37	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; SPT5-TAP::TRP1-KL; syf1::kanMX4;* pRS315-*syf1-37*	this study

Strain	Genotype	Reference
SUB2-TAP SYF1	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; SUB2-TAP::TRP1-KL; syf1::kanMX4;* pRS315-*SYF1*	this study
SUB2-TAP syf1-37	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; SUB2-TAP::TRP1-KL; syf1::kanMX4;* pRS315-*syf1-37*	this study
SYF1	*MAT a; his3Δ1; leu2Δ0; ura3Δ0; syf1::kanMX4;* pRS315-*SYF1*	this study
syf1-37	*MAT a; his3Δ1; leu2Δ0; ura3Δ0; syf1::kanMX4;* pRS315-*syf1-37*	this study
SYF1 GAL10::ACT1	*MAT a; his3Δ1; leu2Δ0; ura3Δ0; syf1::kanMX4; pRS315-SYF1;* pRS316-*GAL10-ACT1*	this study
SYF1 shuffle	*MAT a; his3Δ1; leu2Δ0; ura3Δ0; syf1::kanMX4;* pRS316-*SYF1*	this study
syf1-37 GAL10::ACT1	*MAT a; his3Δ1; leu2Δ0; ura3Δ0; syf1::kanMX4;* pRS315-*syf1-37;* pRS316-*GAL10-ACT1*	this study
SYF1-HA	*MAT alpha; ade2-1; his3-11, 15; ura3-52; leu2-3, 112; trp1-1; SYF1-HA::kanMX4*	this study
TAP-SYF1	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; syf1::kanMX4;* pNOPTAP-*SYF1*	this study
TAP-SYF1 Δhpr1	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; HPR1::HIS3; syf1::kanMX4;* pNOPTAP-*SYF1*	this study

Strain	Genotype	Reference
TAP-syf1-37	*MAT a; ade2-1; his3-11, 15; ura3-1; leu2-3, 112; trp1-1; can1-100; rad5-535; syf1::kanMX4;* pNOPTAP-*syf1-37*	this study
yju2 ts	*MAT a ura3Δ0 leu2Δ0 his3Δ1 can1Δ::LEU2-MFA1pr::His3 yju2-ts::URA3*	Ben-Aroya et al. 2008

5. Methods

5.1 Standard methods

Cloning procedures such as DNA digest with restriction enzymes, CIAP dephosphorylation, DNA ligations, and transformation of newly generated vectors in *Escherichia coli* and separation of DNA in agarose gels were done according to Sambrook and Russell, CSHL Press, 2001. Commercial available kits were used according to manufacturer's instructions.

Amplification of yeast genes or TAP-tags was usually done in a 20 µl Phusion PCR reaction using 0.5 µM forward-primer, 0.5 µM reverse-primer, 1 µg yeast genomic DNA or 50 pg plasmid DNA mixed with 10 µl 2xPhusion Hot Start II High-Fidelity DNA Polymerase (Finzyme). The following amplification protocol was used: 30 s 98°C, [10 s 98°C, 30 s at the respective melting temperature, 30 s/1000 bp 72°C], 35 cycles, 5 min 72°C.

To check for the correct integration of the disruption cassette, colony PCR was performed. A 25 µl reaction was assembled, containing 1 µM forward-primer, 1 µM reverse-primer, 62.5 µM of each dNTP, 750 µM $MgCl_2$ and 1x Taq buffer (Fermentas, St. Leon-Rot). Freshly growing cells were picked with a toothpick and added to the reaction. The PCR reaction was boiled for 15 min (optional), before 1.5 U Taq polymerase (Fermentas, St. Leon-Rot) was added. Amplification was performed using the following protocol: [30 s 95°C, 30 s 45°C, 60 s 72°C] 25 cycles, 2 min 72°C.

5.2 Yeast specific techniques

5.2.1 Culture of *S. cerevisiae*

Yeast cells were grown in either full or synthetic complete medium. Full medium contained 1% of yeast extract, 2% of Bacto-peptone, and 2% of glucose. Synthetic complete medium contained 0.67% of yeast nitrogen base, 0.06% of complete synthetic mix (containing all essential amino acids except three amino acids and two nucleobases used as auxotrophy markers, *i.e.* leucine, tryptophan, histidine, uracil and adenine) and 2% of glucose. 5-FOA was added to a final concentration of 0.1%. 6-AU was used at the concentration of 50 μg/ml.

5.2.2 Transformation of yeast cells

Yeasts were grown in a 50-ml culture to an optical density of 0.5 to 1.0, harvested and washed once with distilled water. Cells were then washed with solution I (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 100 mM CH_3COOLi) and resuspended in 250 μl of the same buffer. Fifty-microliter of cell suspension was mixed with 1-5 μg of plasmid DNA, 10 μg of single strand carrier DNA and 300 μl of solution II (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 100 mM CH_3COOLi, 40% PEG-4000) and incubated at room temperature (RT) for 30 min on a rotator. The suspension was then incubated at 42°C for 10 min, following by 3-min incubation on ice. Cells were washed once with water and plated on selective medium.

5.2.3 Genomic integration of a tandem-affinity-purification (TAP) tag

The genomic integration of a TAP-tag is achieved by homologous recombination of the C-terminal region of the respective gene with a transformed PCR-product containing the TAP-tag and an auxotrophy marker as described by Seraphin B. For C-terminal tagging, pRS1479 (*TRP1*) was

used as a plasmid template of a PCR. The 5'-primer consisted of ~50 nucleotides of yeast genomic sequence flanking the integration site prior to the stop codon of the gene of interest and 5'-TCCATGGAAAAGAGAAG-3', which anneals to the upstream sequence of calmodulin binding protein (CBP) on the plasmid template. The 3'-primer consisted of ~50 nucleotides of yeast genomic sequence flanking the integration site directly downstream of the stop codon of the gene of interest and 5'-TACGACTCACTATAGGG-3', which anneals downstream of the selection marker on the plasmid template. The PCR product was then purified by NucleoSpin Extract II (Macherey-Nagel, Dueren, Germany). The purified DNA was then transformed into the yeast cells to achieve integration into the genome by homologous recombination. Transformants were then tested for the presence of the TAP-tag by western blotting.

5.2.4 Crossings of yeast strains to test for synthetic lethality

Synthetic lethality can be assessed by either combining the deletion of two nonessential genes or by analyzing combinations between different allelic mutations in case of essential genes. Thus, double shuffle strain has to be generated. Haploid parental strains carrying opposite mating types were mated, diploids picked and sporulated on a YPA plate, on which the diploid cells underwent meiosis, divided into four haploid spores, enclosed in a tetrad. The tetrads were treated with Zymolyase 100T to lyse yeast cell wall. Spores were then dissected and subsequently auxotrophy marker analyzed.

5.2.5 Dot spots

One loopful (1 μl) of freshly-growing cells was resuspended in 1 ml of distilled water. After five ten-fold serial dilutions were made, 6-8 μl of cell suspensions were spotted onto a media plate.

5.2.6 Isolation of temperature sensitive alleles

Random mutagenesis of the entire *SYF1* gene was carried out by using error-prone PCR in limiting dATP conditions, where 0.2 mM of dATP and 1 mM of the other three dNTPs were used. The mutant PCR product was transformed along with the XhoI/NheI-digested pRS315-*SYF1* into *SYF1* shuffle strain for cloning by gap repair. Transformants on the SDC-Leu master plate were replica plated onto two 5-FOA plates. One 5-FOA plate was incubated at 37°C and the other at 30°C. Putative ts mutant colonies were then picked from the master plate and restreaked onto a new 5-FOA plate, and the ts phenotype was confirmed by repeating the above procedure. Subsequently, colony plasmid rescue was performed. Plasmids were sequenced to identify the mutations.

5.3 Tandem affinity purification (TAP)

5.3.1 Cell harvest and lysis

For purification of protein complex from yeast *S. cerevisiae*, a 2-litre overnight culture of an optical density of 3-4 was harvested by centrifugation. Cells were washed with water, equilibrated with 25 ml TAP buffer (100 mM NaCl, 50 mM Tris-HCl pH 7.5, 1.5 mM $MgCl_2$, 0.15% NP-40) containing protease inhibitor mix (1.3 μg/ml pepstatin A, 0.28 μg/ml leupeptin, 170 μg/ml phenylmethylsulfonyl fluoride (PMSF), 330 μg/ml benzamidine), and flash-frozen in liquid nitrogen. After thawed in 37ºC water bath, 10 ml of cells were mixed with an equal volume of TAP buffer containing 1 mM dithiothreitol (DTT) and protease inhibitor mix and 20 ml of glass beads in a bead mill (Fritsch, Idar-Oberstein). Cells were lysed according to the following protocol: 3 x (4 min, 500 rpm, 2 min break). After the lysis, whole cell lysate was collected,

glass beads washed with 15 ml of TAP buffer, wash fraction collected and mixed with original whole cell lysate. Then, the lysate was cleared from precipitates by centrifuging at 4 krpm for 10 min at 4°C. The supernatant was subjected to an ultracentrifugation at 27 krpm for one hour at 4°C using an SW32 rotor. Top fatty phase was removed and clear lysate was collected, mixed with glycerin to obtain a final concentration of 5%, flash-frozen in liquid nitrogen.

5.3.2 Purification and trichloro-acetic acid (TCA) precipitation

Four-hundred-microliter bed-volume of pre-washed IgG sepharose was incubated with the whole cell lysate at 4˚C overnight, collected into a Mobicol column and washed with 10 ml of TAP buffer containing 0.5 mM DTT. Next 5 μl of purified recombinant TEV protease was added together with 120 μl TAP buffer into the column. After 1-hour incubation at 16°C, the eluate was collected by centrifugation, added to 250 μl bed-volume of calmodulin sepharose beads, which is pre-washed with calmodulin binding buffer (TAP buffer containing 1 mM DTT and 2 mM $CaCl_2$), incubated on the rotator for 1 hour at 4°C and washed with 5 ml of calmodulin binding buffer. Bound proteins were eluted by incubation with 600 μl of elution buffer (10 mM Tris-HCl pH 8.0, 5 mM EGTA) at 37°C for 15 min. The eluate was collected, precipitated by 10% TCA and washed once with acetone. Protein pellet was dried and resuspended in 35 μl SDS-sample buffer (62.5 mM Tris-HCl pH 6.8, 2.3% SDS, 10% glycerol, 0.05% bromophenolblue and 25 mM DTT). The proteins were then separated by SDS-PAGE.

5.4 *in vivo* transcription assay

Total cellular RNA was isolated by acidic phenol extraction from cells grown in raffinose or shifted to 2% galactose for the indicated time points. The primer extension analysis was performed according to a standard method. Briefly, 1 pmol of ^{32}P-labeled primer (4x10^6 cpm/pmol) specific for *ACT1* mRNA, *GAL10* mRNA, or *SCR1* RNA, an RNAPIII transcript that served as a loading control, was mixed with 5 µg of RNA, hybridized at 37°C for 5 min, extended by reverse transcriptase at 42°C for 1 h, and analyzed on a 6% urea/acrylamide sequencing gel. Quantification was performed with a STORM 860 and the ImageQuant Software (GE Healthcare).

5.5 *in vitro* transcription assay

Yeast nuclear extracts were prepared from 3-liter culture. Transcription reactions were carried out in a 25 *µ*l volume. The reaction mixture contained 200 *µ*g nuclear extract, 150 ng of plasmid template containing either a *HIS4* or a *SER3* promoter (Koschubs et al., 2009), 1x transcription buffer (10 mM HEPES pH 7.6, 50 mM potassium acetate, 0.5 mM EDTA, 2.5 mM magnesium acetate), 2.5 mM DTT, 192 *µ*g of phosphocreatine, 0.2 *µ*g of creatine phosphokinase, 10 U of Ribolock RNase inhibitor (Fermentas), 100 *µ*M NTPs and 200 ng of recombinant Gcn4. The reaction was incubated at 18°C or 30°C for 20 min and RNA was isolated by using RNeasy MinElute Cleanup kit (Qiagen). Transcripts were analyzed by primer extension using 0.125 pmol of a fluorescently labeled 5'-Cy5-oligo. Quantification was performed with a Typhoon 9400 and the ImageQuant Software (GE Healthcare). For the add-back experiment, the Prp19 complex was purified until the TEV eluate using a *SYF1*-TAP strain by TAP method as described above except that the TAP buffer containing 500 mM NaCl and the HIS-tagged TEV protease was depleted using Ni-NTA agarose.

5.6 Splicing assay

Analysis of splicing efficiency was performed by reverse transcription (RT)-PCR. Products were separated on 2% agarose gels, images recorded digitally and ratios between spliced and unspliced products quantified using the ImageJ software (NIH, Bethesda, MD, USA).

5.7 SDS-PAGE and western blotting

SDS-PAGE was performed according to Laemmli using the Mini-Protean II cell (Biorad, Munich). Separated proteins were transferred onto nitrocellulose membrane using a semi-dry blotting apparatus (Peqlab, Erlangen). After transfer, the membrane was incubated with blocking buffer (5% nonfat dry milk in PBS) for 30 min and incubated with blocking buffer containing the primary antibody. After that, the membrane was washed three times with blocking buffer and incubated at RT for two hours with blocking buffer containing the secondary antibody conjugated with HRP. Then the membrane was washed three times with PBS and visualized with an ECL-kit according to manufacture's protocol (Applichem).

5.8 Chromatin immunoprecipitation

Cross-linked cells were lysed with an equal volume of glass beads by vortexing for 6x 3 min with 3-min breaks on ice. Chromatin lysate was used for immunoprecipitation with 15 μl of IgG- or α-Yra1-coupled (Yra1 ChIP) Dynabeads for 3 hours at 20°C. After washing and elution, the IP eluates as

well as input samples were treated with proteinase K overnight at 65°C. DNA was purified by phenol/chloroform extraction. Quantitative PCR with input and IP samples was done on an Applied Biosystems StepOnePlus cycler, using Applied Biosystems' Power SYBRGreen PCR Master Mix. As negative control, primers for the non-transcribed region (NTR) amplifying 174131-174200 on chromosome V (Rother et al. 2010) was used. PCR efficiencies (E) were determined with standard curves. The occupancy of each factor at the respective gene was calculated as its enrichment at the respective gene relative to NTR according to $[E\text{^}(C_T\text{IP-}C_T\text{Input})]_{NTR}$ / $[E\text{^}(C_T\text{IP-}C_T\text{Input})]_{gene}$. For evaluation of processivity, data is expressed as the ratio (distribution) of RNAPII occupancy in a given strain (or 6-AU condition) relative to the promoter position, followed by normalization of each value to the corresponding position in the wild-type or untreated strain, which was defined as 1.

For statistical analyses, data are means ± SEM from n=3. Asterisks indicate statistical significance (Student's t-test; *p<0.05).

5.9 oligo(dT)-*in situ* hybridization

10 ml of cells were grown to mid-log phase. 1.25 ml 37% HCOH were added and the cell suspension was incubated for 90 min on a turning wheel. For *mex67-5* cells, a 10 ml culture was shifted for 15 min to 37°C prior to addition of formaldehyde. Cells were collected (5 min, 3 krpm) and washed twice with 5 ml 0.1 M K-phosphate, pH 6.4. After washing with 1 ml of spheroblasting buffer (0.1 M K-phosphate, 1.2 M sorbitol), the cells were resuspended in 200 µl spheroblasting buffer with 100 µg of Zymolyase 100T and incubated for 30 min at 30°C. Zymolyase treatment was stopped by centrifugation for 4 min at 2000 rpm and the spheroblasts were attached to pre-coated polylysine slide

glass. After 5 min, non-adherent cells were removed by aspiration. Fixation of the adherent cells was achieved by putting the slide for 6 min into -80°C-cold methanol and for 30 s into -80°C-cold acetone. After air-drying, the cells were first rehydrated for 10 min in 100 µl 2x SSC (0.3 M NaCl, 30 mM Na-citrate, pH 7.0), before 12 µl prehybridisation buffer (50% formamide, 10% dextran sulphate, 125 µg/ml of E.coli tRNA, 500 µg/ml herring sperm DNA, 4x SSC, 0.02% polyvinyl pyrrolidone, 0.02% BSA, 0.02% Ficoll-400) were applied to the cells. After an incubation for 1 h at 37°C, 0.75 µl of 1 pmol/µl oligo-d$(T)_{50}$ probe was added, and the slide was incubated overnight in a humid chamber at 37°C. Slides were then washed with 100 ml 0.5x SSC for 30 min at RT in a staining jar, 5 µg 4',6-diamidino-2-phenylindole (DAPI) were added, followed by an incubation for 3 min and further washed with 0.5x SSC for 5 min at RT. The slide was dried at RT in the dark, a solution of 80% glycerol in PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4) was added and the slide was covered with a cover slip. Finally, slides were analyzed using an Olympus BX60 fluorescence microscope.

5.10 Frequency of recombination

For assessing the recombination frequencies, six to twelve independent colonies per transformants were obtained from SC-trp and recombinants were counted as Leu+ colonies on SC-trp-leu plates. Recombination frequencies were calculated as the median frequency of those cultures.

References

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

Ahn, S.H., Kim, M., and Buratowski, S. 2004. Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. *Mol Cell* **13:** 67-76.

Ajuh, P., Kuster, B., Panov, K., Zomerdijk, J.C., Mann, M., and Lamond, A.I. 2000. Functional analysis of the human CDC5L complex and identification of its components by mass spectrometry. *EMBO J* **19:** 6569-6581.

Ben-Yehuda, S., Dix, I., Russell, C.S., McGarvey, M., Beggs, J.D., and Kupiec, M. 2000. Genetic and physical interactions between factors involved in both cell cycle progression and pre-mRNA splicing in Saccharomyces cerevisiae. *Genetics* **156:** 1503-1517.

Blatch, G.L. and Lassle, M. 1999. The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays* **21:** 932-939.

Carmody, S.R. and Wente, S.R. 2009. nuclear mRNA export at a glance. *J Cell Sci* **122:** 1933-1937.

Chan, S.P., Kao, D.I., Tsai, W.Y., and Cheng, S.C. 2003. The Prp19p-associated complex in spliceosome activation. *Science* **302:** 279-282.

Chang, K.J.,Chen, H.C., and Cheng S.C. 2009. Ntc90 is required for recruiting first step factor Yju2 but not for spliceosome activation *RNA* **15:** 1729-1739

Chavez, S., Aguilera, A. 1997. The yeast *HPR1* gene has a functional role in transcriptional elongation that uncovers a novel source of genome instability. *Genes & Dev* **11:** 3459-3470

Das, A.K., Cohen, P.W., and Barford, D. 1998. The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions. *EMBO J* **17:** 1192-1199.

David, C.J., Boyne, A.R., Millhouse, S.R., Manley, J.L. 2011. The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65–Prp19 complex. *Genes & Dev* **25:** 972-983

Gaillard, H., Wellinger, R.E., and Aguilera, A. 2007. A new connection of mRNP biogenesis and export with transcription-coupled repair. *Nucleic Acids Res* **35:** 3893-3906.

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

Hatakeyama, S., Yada, M., Matsumoto, M., Ishida, N., and Nakayama, K.I. 2001. U box proteins as a new family of ubiquitin-protein ligases. *J Biol Chem* **276:** 33111-33120.

Hirose, Y. and Manley, J.L. 2000. RNA polymerase II and the integration of nuclear events. *Genes & Dev* **14:** 1415-1429.

Huertas, P. and Aguilera, A. 2003. Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Mol Cell* **12:** 711-721.

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

Iglesias, N. and Stutz, F. 2008. Regulation of mRNP dynamics along the export pathway. *FEBS Lett* **582:** 1987-1996.

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

Johnson, S.A., Cubberley, G., and 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.

Katahira, J. and Yoneda, Y. 2009. Roles of the TREX complex in nuclear export of mRNA. *RNA Biol* **6:** 149-152.

Kim, M., Ahn, S.H., Krogan, N.J., Greenblatt, J.F., and Buratowski, S. 2004. Transitions in RNA polymerase II elongation complexes at the 3' ends of genes. *EMBO J* **23:** 354-364.

Kohler, A. and Hurt, E. 2007. Exporting RNA from the nucleus to the cytoplasm. *Nat Rev Mol Cell Biol* **8:** 761-773.

Komili, S. and Silver, P.A. 2008. Coupling and coordination in gene expression processes: a systems biology view. *Nat Rev Genet* **9:** 38-48.

Koschubs, T., Seizl, M., Lariviere, L., Kurth, F., Baumli, S., Martin, D.E., and Cramer, P. 2009. Identification, structure, and functional requirement of the Mediator submodule Med7N/31. *EMBO J* **28:** 69-80.

Kuraoka, I., Ito, S., Wada, T., Hayashida, M., Lee, L., Saijo, M., Nakatsu, Y., Matsumoto, M., Matsunaga, T., Handa, H. et al. 2008. Isolation of XAB2 complex involved in pre-mRNA splicing, transcription, and transcription-coupled repair. *J Biol Chem* **283:** 940-950.

Lamb, J.R., Tugendreich, S., and Hieter, P. 1995. Tetratrico peptide repeat interactions: to TPR or not to TPR? *Trends Biochem Sci* **20:** 257-259.

Lei, E.P. and Silver, P.A. 2002. Intron status and 3'-end formation control cotranscriptional export of mRNA. *Genes & Dev* **16:** 2761-2766.

Li X.L., Manley J.L. 2006. Cotranscriptional processes and their influence on genome stability. *Genes & Dev* **20:** 1838-1847

Libri, D., Graziani, N., Saguez, C., and Boulay, J. 2001. Multiple roles for the yeast SUB2/yUAP56 gene in splicing. *Genes & Dev* **15:** 36-41.

Luo, M.L., Zhou, Z., Magni, K., Christoforides, C., Rappsilber, J., Mann, M., and Reed, R. 2001. Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. *Nature* **413:** 644-647.

Makarova, O.V., Makarov, E.M., Urlaub, H., Will, C.L., Gentzel, M., Wilm, M., and Luhrmann, R. 2004. A subset of human 35S U5 proteins, including Prp19, function prior to catalytic step 1 of splicing. *EMBO J* **23:** 2381-2391.

Maniatis, T. and Reed, R. 2002. An extensive network of coupling among gene expression machines. *Nature* **416:** 499-506.

Mason, P.B. and Struhl, K. 2005. Distinction and relationship between elongation rate and processivity of RNA polymerase II *in vivo*. *Mol Cell* **17:** 831-840.

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

Nakatsu, Y., Asahina, H., Citterio, E., Rademakers, S., Vermeulen, W., Kamiuchi, S., Yeo, J.P., Khaw, M.C., Saijo, M., Kodo, N. et al. 2000. XAB2, a novel tetratricopeptide repeat protein involved in transcription-coupled DNA repair and transcription. *J Biol Chem* **275:** 34931-34937.

Nottrott, S., Urlaub, H., and Luhrmann, R. 2002. Hierarchical, clustered protein interactions with U4/U6 snRNA: a biochemical role for U4/U6 proteins. *EMBO J* **21:** 5527-5538.

Ohi, M.D. and Gould, K.L. 2002. Characterization of interactions among the Cef1p-Prp19p-associated splicing complex. *RNA* **8:** 798-815.

Ohi, M.D., Vander Kooi, C.W., Rosenberg, J.A., Chazin, W.J., and Gould, K.L. 2003. Structural insights into the U-box, a domain associated with multi-ubiquitination. *Nat Struct Biol* **10:** 250-255.

Orphanides, G. and Reinberg, D. 2000. RNA polymerase II elongation through chromatin. *Nature* **407:** 471-475.

Orphanides, G. and Reinberg, D. 2002. A unified theory of gene expression. *Cell* **108:** 439-451.

Perales, R. and Bentley, D. 2009. "Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions. *Mol Cell* **36:** 178-191.

Pokholok, D.K., Hannett, N.M., and Young, R.A. 2002. Exchange of RNA polymerase II initiation and elongation factors during gene expression *in vivo*. *Mol Cell* **9**: 799-809.

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

Reed, R. 2003. Coupling transcription, splicing and mRNA export. *Curr Opin Cell Biol* **15:** 326-331.

Reed, R. and Cheng, H. 2005. TREX, SR proteins and export of mRNA. *Curr Opin Cell Biol* **17:** 269-273.

Rondon, A.G., Jimeno, S., and Aguilera, A. 2010. The interface between transcription and mRNP export: From THO to THSC/TREX-2. *Biochim Biophys Acta* **1799:** 533-538.

Rondon, A.G., Jimeno, S., Garcia-Rubio, M., and Aguilera, A. 2003. Molecular evidence that the eukaryotic THO/TREX complex is required for efficient transcription elongation. *J Biol Chem* **278:** 39037-39043.

Rother, S., Burkert, C., Brunger, K.M., Mayer, A., Kieser, A., and Strasser, K. 2010. Nucleocytoplasmic shuttling of the La motif-containing protein Sro9 might link its nuclear and cytoplasmic functions. *RNA* **16:** 1393-1401.

Rougemaille, M., Dieppois, G., Kisseleva-Romanova, E., Gudipati, R.K., Lemoine, S., Blugeon, C., Boulay, J., Jensen, T.H., Stutz, F., Devaux, F. et al. 2008. THO/Sub2p functions to coordinate 3'-end processing with gene-nuclear pore association. *Cell* **135:** 308-321.

Sikorski RS, 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.

Song, E.J., Werner, S.L., Neubauer, J., Stegmeier, F., Aspden, J., Rio, D., Harper, J.W., Elledge, S.J., Kirschner, M.W., and Rape, M. 2010. The Prp19 complex and the Usp4Sart3 deubiquitinating enzyme control reversible ubiquitination at the spliceosome. *Genes & Dev* **24:** 1434-1447.

Strasser, K. and Hurt, E. 2000. Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. *EMBO J* **19:** 410-420.

Strasser, K. and Hurt, E. 2001. Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p. *Nature* **413:** 648-652.

Strasser, K., Masuda, S., Mason, P., Pfannstiel, J., Oppizzi, M., Rodriguez-Navarro, S., Rondon, A.G., Aguilera, A., Struhl, K., Reed, R. et al. 2002. TREX is a conserved complex coupling transcription with messenger RNA export. *Nature* **417**: 304-308.

Stutz, F. and Izaurralde, E. 2003. The interplay of nuclear mRNP assembly, mRNA surveillance and export. *Trends Cell Biol* **13:** 319-327.

Tarn, W.Y., Hsu, C.H., Huang, K.T., Chen, H.R., Kao, H.Y., Lee, K.R., and Cheng, S.C. 1994. Functional association of essential splicing factor(s) with PRP19 in a protein complex. *EMBO J* **13:** 2421-2431.

Thomas BJ, Rothstein R. 1989. Elevated recombination rates in transcriptionally active DNA. *Cell* **56:** 619-630.

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

Will, C.L. and Lührmann, R. 2006. Spliceosome structure and function. *in RF Gesteland, TR Chech, and JF Atkins (ed), The RNA world, 3rd ed Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY*: 369-400.

Zenklusen, D., Vinciguerra, P., Strahm, Y., and Stutz, F. 2001. The yeast hnRNP-Like proteins Yra1p and Yra2p participate in mRNA export through interaction with Mex67p. *Mol Cell Biol* **21:** 4219-4232.

Zenklusen, D., Vinciguerra, P., Wyss, J.C., and Stutz, F. 2002. Stable mRNP formation and export require cotranscriptional recruitment of the mRNA export factors Yra1p and Sub2p by Hpr1p. *Mol Cell Biol* **22:** 8241-8253.

Abbreviations

°C	degree celcius
5-FOA	5-Fluoroorotic acid
6-AU	6-azauracil
ATP	adenosine triphosphate
bp	basepair
CBP	calmodulin binding protein
Da	dalton
DAPI	4',6-diamidino-2-phenylindole
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleosid triphosphate
E. coli	*Escherichia coli*
ECL	enhanced chemoluminiscence
g	gram
h	hour
HA	hemagglutinin
l	litre
m	milli
M	molar
min	minute
mRNA	messenger ribonucleic acid
mRNP	messenger ribonucleoprotein
nt	nucleotide
NTC	nineteen complex

OD	optical density
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
pH	potential of hydrogen
PMSF	phenylmethylsulfonyl fluoride
RNA	ribonucleic acid
RNAPII	RNA polymerase II
rpm	revolutions per minute
RT	room temperature
SDS	sodium dodecyle sulphate
TAP	tandem affinity purification
TCA	trichloro-acetic acid
TEV	Tobacco etch virus
TREX	transciption and export
ts	ts
YPD	yeast extract, peptone, glucose; glucose containing medium
YPG	yeast extract, peptone, galactose; galactose containing medium
wt	wild-type
μ	micro

Acknowledgements

I would like to thank all the people who have contributed to this work in one way or another.

First of all, I want to thank Dr. Katja Sträßer for giving me an opportunity to work with her, for allowing me the freedom to work with my own ideas and for helping me see things more positively and clearly than I usually do.

I also would like to thank Prof. Dr. Patrick Cramer, my official "Doktorvater", for his advices and a chance to collaborate.

I want to thank all the members of the Sträßer lab, past and present, for everything I learned from them, for all little things they helped me out with, for making the lab a good place to work in. Especially I would like to thank: Dr. Susanne Röther and Dr. Emanuel Clausing for teaching me all the methods in yeast genetics; Dr. Lina Karakasili, Britta Coordes, Katharina Brünger, Cornelia Burkert and Dominik Meinel for all help and many fruitful discussions; Anja Kieser and Juliane Horenk for their excellent technical supports; and my practical and master students, Amelie Schreieck, Simon Geiger and Erik Tilch for their great work.

I would like to thank Dr. Martin Seizl and the members of the Cramer lab for showing me the *in vitro* transcription assay and their kind assistance.

I also would like to thank Prof. Dr. Olaf Rieß for all of his supports and advices during my study in Tübingen, where I learned not only great science but also the positive attitude about my scientific life.

I cannot finish my acknowledgements without expressing my gratitude to Prof. Dr. Hiroshi Handa for his supervision during my best time in Tokyo. His

unforgettable message, "Science is the fruit of the uniqueness and self-respect.", is always in my heart. I also want to thank Dr. Yuki Yamaguchi, Dr. Takashi Narita and Dr. Masaki Endou for their kind and careful supervisions. During my three years with them, they taught me how to develop my own creativity, improve critical thinking and raise the quality of diligent work.

Lastly, I would like to thank all the parts of my family, who are always supporting me and being the reason I want to do my best constantly.

Publications

Chanarat S, Seizl M and Sträßer K. 2011 The Prp19 Complex is a novel transcription elongation factor required for TREX occupancy at transcribed genes. ***Genes and Development*** 25(11): 1147-1158

Clausing E, Mayer A*, **Chanarat S***, Müller B, Germann SM, Cramer P, Lisby M, and Sträßer K. 2010 The Transcription Elongation Factor Bur1/Bur2 Interacts with Replication Protein A to Maintain Genome Stability. ***J Biol Chem*** 285(53): 41665-41674 **these authors contributed equally to this paper*

Nuber S, Franck T, Wolburg H, Schumann U, **Chanarat S**, Teismann P, Schulz JB, Luft AR, Jürgen Tomiuk J, Wilbertz J, Bornemann A, Krüger R and Riess O. 2010 Overexpression of the alpha-synuclein interacting protein synphilin-1 leads to behavioural and neuropathological alterations in transgenic mice. ***Neurogenetics*** 11(1):107-20

Grundmann K, Reischmann B, Vanhoutte G, Hübener J, Teismann P, Hauser TK, Bonin M, Wilbertz J, Horn S , Nguyen HP, Kuhn M, **Chanarat S**, Wolburg H, Van der Linden A, Riess O. 2007. Overexpression of human wild-type torsinA causes a neurological phenotype in a transgenic mouse model. ***Neurobiol Dis*** 27(2):190-206

Yamaguchi Y, Mura T, **Chanarat S**, Okamoto S, Chen Y, Handa H. 2007. Hepatitis delta antigen binds to the RNA polymerase II clamp and modulates transcriptional fidelity. ***Genes Cells*** 12(7): 863-75

Narita T, Yamaguchi Y, Yano K, Sugimoto S, **Chanarat S**, Wada T, Kim DK, Hasegawa J, Omori M, Inukai N, Endoh M, Yamada T, Handa H. 2003. Human transcription elongation factor NELF: identification of novel subunits and reconstitution of the functionally active complex. ***Mol Cell Biol*** 23(6):1863-73

Curriculum vitae

Personal data:

Name:	Sittinan Chanarat
Date of Birth:	January 28, 1980
Place of Birth:	Chiangmai, Thailand
Nationality:	Thai

Education

since June 2007	Gene Center, University of Munich; Munich, Germany Ph.D. Candidate in Biochemistry
2005-2006	University of Tübingen; Tübingen, Germany
2003-2004	Tokyo Institute of Technology; Tokyo, Japan M.Eng. in Biological Informations
1999-2002	Tokyo Institute of Technology; Tokyo, Japan B.Eng. in Bioengineering
1998	Tokyo University of Foreign Studies; Tokyo, Japan Certificate of Japanese Language Proficiency

Fellowship / Scholarship

2005-2006	Marie Curie Early Stage Training Fellowship
1998-2004	Japanese Government Scholarship (Monbusho)