

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

**Role of the RNA polymerase II C-terminal domain in
transcription termination and function of Spt5 in
3' RNA-processing factor recruitment**



Amelie Erna Schreieck
aus München, Deutschland

2013

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Patrick Cramer betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, 28.10.2013

Amelie Erna Schreieck

Dissertation eingereicht am	29.10.2013
1. Gutachter	Prof. Dr. Patrick Cramer
2. Gutachter	PD Dr. Dietmar Martin
Mündliche Prüfung am	19.12.2013

Acknowledgements

First of all, I would like to thank my supervisor Prof. Dr. Patrick Cramer. He really is an exceptional group leader and researcher. Thank you, Patrick, for your constant appreciation and support and for always motivating me by sharing your inspiring visions about the world of transcription. I am very grateful that I was able to work on my PhD thesis in your laboratory.

I also would like to thank my mentors and Thesis Advisory Committee members Dr. Dietmar Martin, Dr. Johannes Söding and Prof. Dr. Dirk Eick for their time, for fruitful discussions and a lot of scientific input on my projects during the last three years.

I am deeply grateful to Stefanie Etzold, my highly efficient co-worker and good friend, for sharing the Tyr1-P project with me and doing a lot of excellent work in very short time. Thank you, Steffi, for being my bench-neighbor for more than two years. I really enjoyed working with you and discussing a lot of (also non-scientific) things. And of course thank you for the little radio that made every day lab work so much more entertaining!

Special thanks also go to Michael Lidschreiber, for collaborations on many projects and for sharing all his ChIP experience with me. Thank you for lots of discussions on difficult subjects and also for tons of ChIP-chip profiles.

For additional technical help in the lab, I thank Kristin Leike for hundreds of ChIP experiments, Western Blots and everything else; that was really a great support during the last two years.

I am also very thankful to Andreas Mayer, who initially invited me to the lab to do my Master's thesis with him and then supervised me during the first year and introduced me to all the techniques and topics around Spt5 and the Pol II CTD. You made a great start in the lab possible for me.

Many thanks as well to all past and present members of the "Mediator group", especially Laurent Larivière, Martin Seizl, Rike Hög, Larissa Wenzek and Clemens Plaschka for endless help, many inspiring coffee breaks and "Mediator meetings" as well as the great working atmosphere with you.

I also thank my student, Katharina Hofmann, for a lot of work in the lab and for allowing me to practice my teaching skills.

A lot of people worked together with me on collaboration projects and I would like to thank them for that: Martin Heidemann and Roland Schüller from Prof. Dirk Eick's group shared CTD related projects with me. Thank you, Roland, for preparing all my yeast samples and for many interesting discussions. Ashley Easter and Lori Passmore collaborated with us on the

Glc7 project – thanks to you two for a lot of discussions, input and work on this topic, especially to Ashley for coming to Munich and having a week of great work and lots of fun. I think these are really examples for how collaborations should be like.

Of course, I also thank all other members of the Cramer lab for discussions, help and the nice atmosphere at work and especially Kerstin Maier not only for help in the lab but for giving me rides home from work and for watering our plants every summer!

Außerdem möchte ich mich bei meinen guten Freunden aus Studienzeiten bedanken: Francisco, Konstantin, Claudia, Lydia, Anna und Jens. Ihr habt mich fast vom ersten Tag an die letzten 8 Jahre begleitet. Vielen Dank für eure unzähligen Ratschläge während Kaffeepausen und unseren Mittagessen jeden Mittwoch und für eure Unterstützung in jeglicher Lebenslage.

Natürlich danke ich besonders euch von ganzem Herzen, liebe Mama, lieber Papa, lieber Maximilian, für eure bedingungslose Unterstützung in allem, was ich vorhatte. Ohne euch wäre vieles davon nicht möglich gewesen und ohne euch wäre ich auch nicht die, die ich heute bin.

Lieber Alex, vielen Dank, dass du mich die letzten Jahre in allem begleitet hast. Danke vor allem für deinen Optimismus, deinen Blick für das große Ganze und dafür, dass du mir in jeder Situation zeigen kannst, was das Wichtigste ist.

Summary

Protein-coding genes in eukaryotes are transcribed by RNA polymerase II (Pol II). This process is tightly regulated and coupled to RNA processing. Many transcription and RNA processing factors are recruited to Pol II via its conserved C-terminal domain (CTD) containing 27 heptapeptide repeats of the consensus sequence Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7 in *Saccharomyces cerevisiae*. These repeats can be differentially phosphorylated during the transcription cycle serving as a code for interacting factors. During transcription initiation, Ser5 is phosphorylated at the 5'-end of genes and this phosphorylation is required for RNA capping enzyme binding. During transcription elongation, the Pol II CTD becomes phosphorylated at Tyr1 and Ser2 and binds the elongation factor Spt5. Spt5 also contains a repetitive C-terminal region (CTR) required for cotranscriptional recruitment of proteins. At the 3'-end of genes, Ser2-phosphorylated Pol II associates with the cleavage and polyadenylation factor (CPF) and is dephosphorylated at Tyr1 residues.

This work shows that CPF is a Pol II CTD phosphatase and that its subunit Glc7 dephosphorylates Tyr1 *in vitro*. *In vivo*, Glc7 activity is required for normal Tyr1 dephosphorylation at the polyadenylation (pA) site, for recruitment of termination factors Pcf11 and Rtt103, and for normal Pol II termination. These results show that transcription termination involves Tyr1 dephosphorylation of the CTD and indicate that pre-mRNA processing and transcription termination are coupled via CPF-dependent Pol II Tyr1 dephosphorylation. Additionally, 19 kinases were tested for activity on Tyr1 in yeast by selective inhibition or knock-out *in vivo*. However, none of the candidates was identified as the Tyr1 kinase. Possibly this enzyme is an atypical kinase not known to be involved in transcription so far.

Furthermore, this work reports a new role of the Spt5 CTR in recruitment of RNA 3'-end processing factors. The results show that the Spt5 CTR as well as RNA is required for normal recruitment of the pre-mRNA cleavage factor (CF) I to the 3'-end of yeast genes. Genome-wide ChIP profiling detects occupancy peaks of CFI subunits around 100 base pairs downstream of the pA site of genes. CFI recruitment to this defined region may result from simultaneous binding to the Spt5 CTR, to nascent RNA containing the pA sequence, and to the elongating Pol II isoform that is phosphorylated at Ser2 of the CTD. Consistent with this model, the CTR interacts with CFI *in vitro*, but is not required for pA site recognition and transcription termination *in vivo*.

In summary, we characterized two new aspects of transcription and RNA processing regulation by two different C-terminal repetitive protein domains. CTD Tyr1 phosphorylation, which is removed by Glc7, regulates termination factor recruitment by masking their binding site, the Spt5 CTR is involved in recruitment of CFI. Both results greatly contribute to a more detailed understanding of the mechanisms involved in transcription termination and RNA 3'-end processing.

Publications

Parts of this work have been published or are in the process of being published:

Chapter 3:

Schreieck A*, Easter AD*, Etzold S, Wiederhold K, Lidschreiber M, Cramer P, Passmore LA. “RNA polymerase II termination involves CTD tyrosine dephosphorylation by CPF subunit Glc7”. Manuscript in revision at *Nature Structural and Molecular Biology*.

*These authors contributed equally.

Author contributions: AS performed dephosphorylation assays and *in vivo* experiments and wrote the manuscript, ADE purified the CPF complex, performed additional spot dilution assays and wrote the manuscript, SE performed *in vivo* and *in vitro* experiments, KW purified CPF, ML analyzed data, PC and LP designed and supervised the project and wrote the manuscript.

Chapters 1.4 and 2.2:

Mayer A*, Heidemann M*, Lidschreiber M, Schreieck A, Sun M, Hintermair C, Kremmer E, Eick D, Cramer P. (2012) “CTD tyrosine phosphorylation impairs termination factor recruitment to RNA polymerase II.” *Science* **336**:1723-1725

Author contributions: AM performed ChIP-chip and fluorescence anisotropy experiments, MH and CH validated antibodies, ML analyzed ChIP-chip data, MS carried out modeling, AS performed additional kinase ChIP assays. EK generated the 3D12 antibody. DE and PC designed and supervised research. AM and PC prepared the manuscript with help from all authors.

Chapters 1.3.3 and 4:

Mayer A*, Schreieck A*, Lidschreiber M, Leike K, Martin DE, Cramer P. (2012) “The Spt5 C-terminal region recruits yeast 3’ RNA cleavage factor I.” *Molecular and Cellular Biology* **32**: 1321-1331

*These authors contributed equally.

Author contributions: AM performed RNase-ChIP and ChIP-chip experiments and wrote the manuscript. AS performed all other experiments and wrote the manuscript, ML analyzed ChIP-chip data, KL performed ChIP-chip experiments, DEM designed research, PC designed and supervised research and wrote the manuscript.

Table of contents

Acknowledgements.....	I
Summary.....	III
Publications.....	IV
1. Introduction.....	1
1.1. Gene transcription in eukaryotes	1
1.2. The RNA polymerase II CTD	2
1.2.1. CTD sequence and structure	2
1.2.2. Posttranslational modifications of the CTD	4
1.3. The Pol II transcription cycle	7
1.3.1. Transcription initiation.....	7
1.3.2. Transcription elongation	8
1.3.3. The elongation factor Spt5	10
1.3.4. Transcription termination and RNA 3'-end processing.....	11
1.4. CTD Tyr1 phosphorylation impairs termination factor recruitment to Pol II	12
1.4.1. A monoclonal antibody against Tyr1-phosphorylated CTD	13
1.4.2. Tyr1 is phosphorylated in the coding region of genes	14
1.4.3. Tyr1-P negatively influences termination factor recruitment	14
1.4.4. The extended CTD code.....	15
1.5. Aims and scope of this work	16
2. Towards the identification of the CTD Tyr1 kinase in yeast.....	18
2.1. Protein kinases in yeast.....	18
2.2. Tyr1 of the Pol II CTD is not phosphorylated by canonical CTD kinases.....	19
2.3. Systematical approach to identify the Tyr1 kinase.....	21
2.3.1. Non-essential kinases	23
2.3.2. Essential kinases.....	24
2.4. Discussion.....	25
3. Pol II termination involves CTD Tyr1 dephosphorylation by CPF subunit Glc7	27
3.1. CPF subunit Glc7 dephosphorylates CTD Tyr1 <i>in vitro</i>	27
3.2. Glc7 is required for Tyr1 dephosphorylation <i>in vivo</i>	28

3.3.	Ssu72 does not dephosphorylate Tyr1	32
3.4.	Glc7 is involved in transcription termination	34
3.5.	Discussion.....	34
4.	The Spt5 C-terminal region recruits yeast 3' RNA cleavage factor I	37
4.1.	Investigation of elongation factor recruitment by Spt5 CTR	37
4.2.	Spt5 CTR is required for recruitment of CFI <i>in vivo</i>	39
4.3.	Spt5 CTR interacts with CFI <i>in vitro</i>	41
4.4.	RNA contributes to CFI recruitment	42
4.5.	CFI colocalizes with the Ser2-phosphorylated CTD downstream of the pA site.....	43
4.6.	CTR deletion does not impair termination	45
4.7.	CTR deletion does not alter pA site usage	47
4.8.	Discussion.....	47
5.	Future perspectives.....	50
5.1.	The CTD code extension Tyr1-P	50
5.1.1.	Identification of (the) Tyr1 kinase(s)	50
5.1.2.	A positive function of CTD Tyr1 phosphorylation.....	51
5.1.3.	Characterization of Thr4 and Ser7 phosphorylation in yeast.....	52
5.2.	Towards an overall CTD phosphorylation pattern	52
5.3.	New functions of the Spt5 CTR	53
6.	Materials and Methods.....	55
6.1.	Materials	55
6.1.1.	Bacterial and yeast strains	55
6.1.2.	Growth media and additives.....	57
6.1.3.	Plasmids and primers	58
6.1.4.	Antibodies	62
6.1.5.	Buffers and solutions.....	63
6.2.	Methods	65
6.2.1.	Molecular cloning methods	65
6.2.2.	Bacterial methods.....	65
6.2.3.	Yeast methods	66
6.2.4.	Protein methods.....	69

6.2.5.	RNA methods	71
6.2.6.	ChIP methods	72
6.2.7.	ChIP and tiling microarray hybridization (ChIP-chip)	74
References	78
List of figures	94
List of tables	96
Abbreviations	97

1. Introduction

1.1. Gene transcription in eukaryotes

The transfer of information encoded in the DNA sequence into proteins is essential and referred to as the central dogma of biochemistry (1). Transcription of genes from a DNA template into an RNA sequence is the first step of this pathway and therefore one of the most important processes in all living cells. The enzyme that carries out transcription is the DNA-dependent RNA polymerase (Pol) (2).

Eukaryotic cells have three nuclear RNA polymerases that perform transcription of DNA: Pol I, II and III. Pol I synthesizes ribosomal RNAs (rRNAs) and thereby makes up for 75% of transcription in the cell (3-5). Pol II produces protein-coding RNAs as well as small nuclear (snRNAs) and small nucleolar RNAs (snoRNAs) (6-8). Pol III is responsible for synthesis of transfer RNAs (tRNAs), the 5S rRNA and other small RNA molecules, which accounts for another 15% of transcription in cells (5, 9). Although Pol II performs only about 10% of transcription in cells, it transcribes about 85% of the whole yeast genome (10, 11).

All three enzymes are structurally related and consist of a conserved 10-subunit core and specific additional subunits (12). Pol II, however, is the only polymerase containing a flexible C-terminal repeat domain (CTD) on its largest subunit Rpb1 (13). It is commonly accepted that the CTD is necessary for Pol II regulation (14, 15).

Additional polymerases in eukaryotic cells are the mitochondrial polymerase, a single-subunit polymerase similar to the bacteriophage T7 polymerase (16), which transcribes the mitochondrial genome, and Pol IV and V in plant cells which regulate non-coding RNA mediated gene silencing (17).

RNA polymerase II transcribes through a chromatin environment in which the DNA is wrapped around nucleosomes like “beads on a string” (18, 19). Each nucleosome is wrapped in 147 bp of DNA (20) and consists of eight subunits: two histone H2A/H2B dimers and one histone H3/H4 tetramer (21). This packaging structure of DNA and proteins can prevent access of Pol II to the DNA and thus can inhibit transcription (19). As a consequence, chromatin remodelers with the help of sequence-specific regulators have to rearrange chromatin to control specific expression of genes (22).

1.2. The RNA polymerase II CTD

1.2.1. CTD sequence and structure

The CTD of Pol II, which is conserved from yeast to human, forms a flexible extension of the Pol II subunit Rpb1 consisting of multiple heptapeptide repeats of the consensus sequence Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7 (Y₁S₂P₃T₄S₅P₆S₇) (23). The number of repeats increases with the complexity of the organism ranging from 27 repeats in *Saccharomyces cerevisiae* (*S. cerevisiae*), S288C strain (24), to 42 repeats in *Drosophila melanogaster* (*D. melanogaster*) and 52 in *Homo sapiens* (*H. sapiens*) (23, 25) (Figure 1).

According to Figure 1, not only the CTD length but also the number of consensus repeats varies considerably between organisms. Whereas the yeast CTD contains 19 out of 27 consensus repeats (70%), the human CTD contains only 21 out of 52 (40%) and the *D. melanogaster* CTD only 2 out of 42 (5%). All consensus repeats of the human CTD are located in the proximal part of the CTD, closer to the body of Pol II and the linker. The distal part of the CTD consists exclusively of non-consensus repeats. A study in mice showed that the deletion of several non-consensus repeats of the CTD leads to defects in growth and development (26), whereas in yeast, a CTD with only consensus repeats does not show any defects (27). Therefore, it is possible that the longer CTD was acquired during evolution to take over more diverse functions. A polymerase is able to transcribe without a CTD *in vitro* (28), but the repeats are absolutely necessary for cell viability (24, 25, 29, 30). Yeast tolerates a truncation of the CTD down to eight repeats, but these cells grow slowly and show temperature sensitivity (27). In mice, a truncation to 29 of 52 repeats is possible (29). However, this truncation was done only in cell culture and the situation might be different in living organisms (31).

The CTD is thought to be mainly unstructured and flexible as it was not detected in the Pol II crystal structure (32), but it is able to form β -turns (33). CTD phosphorylation in turn leads to unfolding of secondary structure and extension of the repeat sequence (34).

As mentioned above, the consensus repeat $Y_1S_2P_3T_4S_5P_6S_7$ is conserved from yeast to human. Mutation of Ser2 or Ser5 of the CTD repeats in *S. cerevisiae* to alanine (Ala) or glutamate (Glu) as a phosphomimetic results in an inviable phenotype (27). An exchange of all Tyr1 residues to phenylalanine (Phe) is also lethal (27). In contrast to that, mutations of Thr4 or Ser7 to Ala are viable in yeast (35), whereas mutation of Ser7 to Glu is lethal (36). Thus, the amino acids of the consensus repeat are of different importance for CTD function with Tyr1, Ser2 and Ser5 being crucial for viability. Interestingly, the mutations of Tyr1 to Phe or Ser2 to Ala are not lethal in *Schizosaccharomyces pombe* (*S. pombe*) (37) but Thr4 to Ala is lethal in human cells (38). Taken together, these reports indicate that the CTD residues – although the sequence is conserved – may have diverging functions in other organisms.

Mutational studies by Stiller and Cook additionally showed that introducing Ala residues between every heptapeptide repeat is lethal in yeast, whereas an Ala residue after every second repeat is viable (39). Thus, the CTD repeats are functioning in pairs. By mutating the last residues of every diheptad to Ala, it was found in *S. pombe* that only a decapeptide YSPTSPSYSP is necessary for viability (40) and in *S. cerevisiae*, a mutant with YSPTSPSYSPTAAA repeats was still viable (41). The underlined sequence is termed the minimal functional unit of the CTD (41).

1.2.2. Posttranslational modifications of the CTD

The CTD can be reversibly posttranslationally modified as is schematically shown in Figure 2 for one repeat (42-46). These modifications are thought to serve as a CTD code that is recognized by interacting proteins, the “CTD readers”, which are specifically recruited to a distinctly modified CTD (47, 48).

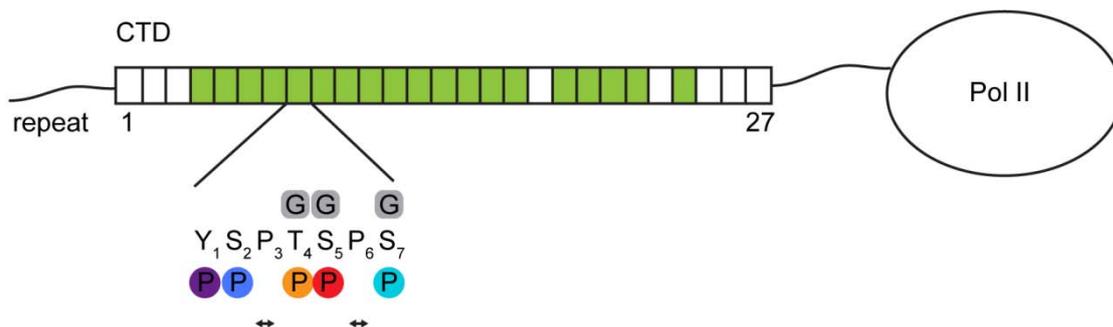


Figure 2: Model of the yeast Pol II CTD repeats and possible posttranslational modifications of a consensus repeat.

The 27 CTD repeats are depicted as boxes, green indicates consensus repeats. The consensus sequence with its possible modifications is shown below the repeats. G: Glycosylation (only shown in mammals, (45, 49)); P: phosphorylation (same color code is used during this work, (43, 44)); double arrow: cis/trans isomerization (46).

Five out of seven residues within one yeast CTD repeat can be phosphorylated, of which Ser2 and Ser5 are the most studied phosphorylation sites (38, 44, 50, 51). The CTD phosphoryla-

tions have been mapped along transcribed genes: Ser5 and Ser7 phosphorylation levels peak shortly after the transcription start site (TSS), with Ser7-P peaking slightly before Ser5-P (43). Tyr1 and Ser2 phosphorylations gradually increase along the transcribed region and reach peak levels around the polyadenylation (pA) site (44) (chapter 1.4). Thr4 phosphorylation has not been extensively studied in yeast.

The various CTD phosphorylations play an important role in the Pol II transcription cycle (chapter 1.3, (42)). Therefore, the kinases and phosphatases, termed “writers and erasers” of the CTD code (47), have been of great interest and are summarized in Table 1.

Table 1: Known *in vivo* CTD kinases and phosphatases in *S. cerevisiae* (adapted from (47)).

Specific CTD phosphorylation	Kinase	Phosphatase
Tyr1	--	--
Ser2	Bur1, Ctk1	Fcp1
Thr4	--	--
Ser5	Kin28, Srb10, Cdc28	Ssu72, (Rtr1)
Ser7	Kin28	Ssu72

Kin28, which is part of transcription factor TFIIH, introduces the first phosphorylations on Ser5 and Ser7 during the transcription cycle (50, 52). Other kinases were also proposed to phosphorylate Ser5 such as Srb10, a subunit of the Mediator, and Cdc28 (53-55). The main phosphatase for Ser5-P and Ser7-P, which signals decrease after transcription initiation, is Ssu72 (36, 56). However, as Ssu72 is a component of the cleavage and polyadenylation factor (CPF), it is recruited mainly at the 3'-end of genes and thus possibly not the enzyme that decreases Ser5-P levels directly after initiation. Another protein that was proposed to dephosphorylate Ser5-P in the 5' region of genes is the atypical phosphatase Rtr1 (57), although the mechanism of its recruitment remains unclear. Furthermore, it was shown recently that Rtr1 does not contain an active site and that it most likely only plays a regulatory role in Ser5-P dephosphorylation (58). Ser2 is phosphorylated during elongation by two kinases: Ctk1 and Bur1. Although Ctk1 is thought to be the major Ser2 kinase *in vivo* (59-61), its deletion does not completely abolish Ser2 phosphorylation. A second kinase, Bur1, additionally contributes to Ser2 phosphorylation (60). The main Ser2-P phosphatase is Fcp1 (59), which shows highest occupancy levels at the 3'-ends of genes in accordance with the decrease of Ser2-P levels (36). Interestingly, neither kinases nor phosphatases for Tyr1 or Thr4 residues have yet been identified in yeast. In mammals, Thr4 is phosphorylated by Plk3 (38) and Tyr1 by c-Abl (62). As Tyr1 is an essential residue of the CTD in yeast, it is of great interest to identify the enzymes which modulate phosphorylation of Tyr1. This issue is extensively addressed as part of this work (chapters 2 and 3).

Another posttranslational modification of the CTD, which has so far only been shown in mammals, is O-GlcNAcylation (49), the addition of O-linked N-acetylglucosamine (O-GlcNAc) to serine or threonine residues. This modification can be added to Thr4, Ser5 or Ser7 residues and is mutually exclusive with CTD phosphorylation (45). As the O-GlcNAc transferase has been shown to be a part of the preinitiation complex (PIC) in higher eukaryotes, it is proposed, that CTD glycosylation can regulate transcription initiation (45).

The proline residues of the CTD repeats can also be modified by cis-trans-isomerization (46). In *S. cerevisiae*, the essential isomerase Ess1 is responsible. The *trans* configuration is generally more stable, but it has been shown that the Ser5 phosphatase Ssu72 only binds the Ser5-phosphorylated CTD when the Ser5-Pro6 bond is in *cis* configuration (63). In contrast to that, the termination factor Pcf11 binds to Ser2-P when the Ser2-Pro3 bond is in *trans* configuration (64, 65). Mammalian non-consensus repeats have also been shown to be methylated on arginine at position 7 in repeat 31 (Arg1810) by coactivator-associated arginine methyltransferase 1 (CARM1) (66) and six lysines in the mouse CTD can be ubiquitinated by a ubiquitin ligase (67).

If all theoretically possible posttranslational modifications of one yeast CTD repeat are taken into account (phosphorylation and proline isomerization), 2 states for each residue would be possible, which results in $2^7 = 128$ combinations per repeat. For the complete yeast CTD, this would be $2^{173} = 1.2 \times 10^{52}$ possibilities. In mammals, with glycosylation and methylation and more repeats, this number would be even higher. This complexity raises the question why the CTD has evolved as it is now and if all of these possible states are present in living cells.

So far it is known that specific CTD marks serve as a code for recruitment of a myriad of interacting proteins, the CTD code readers (47). Table 2 gives an overview of important readers in yeast.

Table 2: Important Pol II CTD readers in *S. cerevisiae* (adapted from (42)).

Protein (complex)	Function	CTD phosphorylation state
TFIIE	transcription initiation	unphosphorylated CTD
TFIIF	transcription initiation	unphosphorylated CTD
TBP	transcription initiation	unphosphorylated CTD
Mediator	transcription initiation	unphosphorylated CTD
Ceg1 (capping enzyme)	guanylyltransferase, RNA 5' capping	Ser5-P
Abd1	methyltransferase, RNA 5' capping	<i>in vitro</i> phosphorylated CTD
Set1 (COMPASS)	histone methylation (H3K4me3)	Ser5-P
Spt6	nucleosome remodeling	Ser2-P

Protein (complex)	Function	CTD phosphorylation state
Nrd1	termination of non-polyadenylated transcripts	Ser5-P
Ess1	proline isomerization	<i>in vitro</i> phosphorylated CTD
Set2	histone methylation (H3K36me3)	Ser2-P/Ser5-P
Prp40	RNA splicing	<i>in vitro</i> phosphorylated CTD
Pcf11 (CFIA)	mRNA cleavage and polyadenylation	Ser2-P
Rna14 (CFIA)	mRNA cleavage and polyadenylation	<i>in vitro</i> phosphorylated CTD
Rna15 (CFIA)	mRNA cleavage and polyadenylation	<i>in vitro</i> phosphorylated CTD
Ydh1 (CPF)	mRNA cleavage and polyadenylation	<i>in vitro</i> phosphorylated CTD
Yhh1 (CPF)	mRNA cleavage and polyadenylation	<i>in vitro</i> phosphorylated CTD
Pta1 (CPF)	mRNA cleavage and polyadenylation	Ser5-P
Rtt103	interaction with exonucleases, transcription termination	Ser2-P
Yra1	mRNA export	Ser2-P/Ser5-P

The following chapter 1.3 includes all CTD readers mentioned above into a picture of the Pol II transcription cycle including RNA processing and important changes in the chromatin environment.

1.3. The Pol II transcription cycle

In the course of the Pol II transcription cycle, changes in CTD phosphorylation coordinate recruitment of transcription and RNA processing factors to the polymerase (14, 15, 68). Pol II CTD phosphorylation profiles along the transcribed region of yeast genes are shown in Figure 6 and (43). How these phosphorylations determine factor recruitment will be covered in the following chapter. Figure 3 depicts a simplified model of the Pol II transcription cycle and important CTD phosphorylations.

1.3.1. Transcription initiation

In the first step of the transcription cycle (Figure 3, Initiation), RNA Pol II with an hypophosphorylated CTD associates with general transcription factors (GTFs) and the mediator complex at the promoter DNA to form the PIC (15), whereby the GTFs TBP, TFIIE and TFIIIF as well as the Mediator complex directly interact with the CTD (69-72). Lastly, TFIIH is recruited to the PIC (69) and its kinase subunit Kin28 phosphorylates the CTD on Ser5 and Ser7 residues (52, 53). Ser5 phosphorylation leads to dissociation of Mediator and to promoter clearance (73).

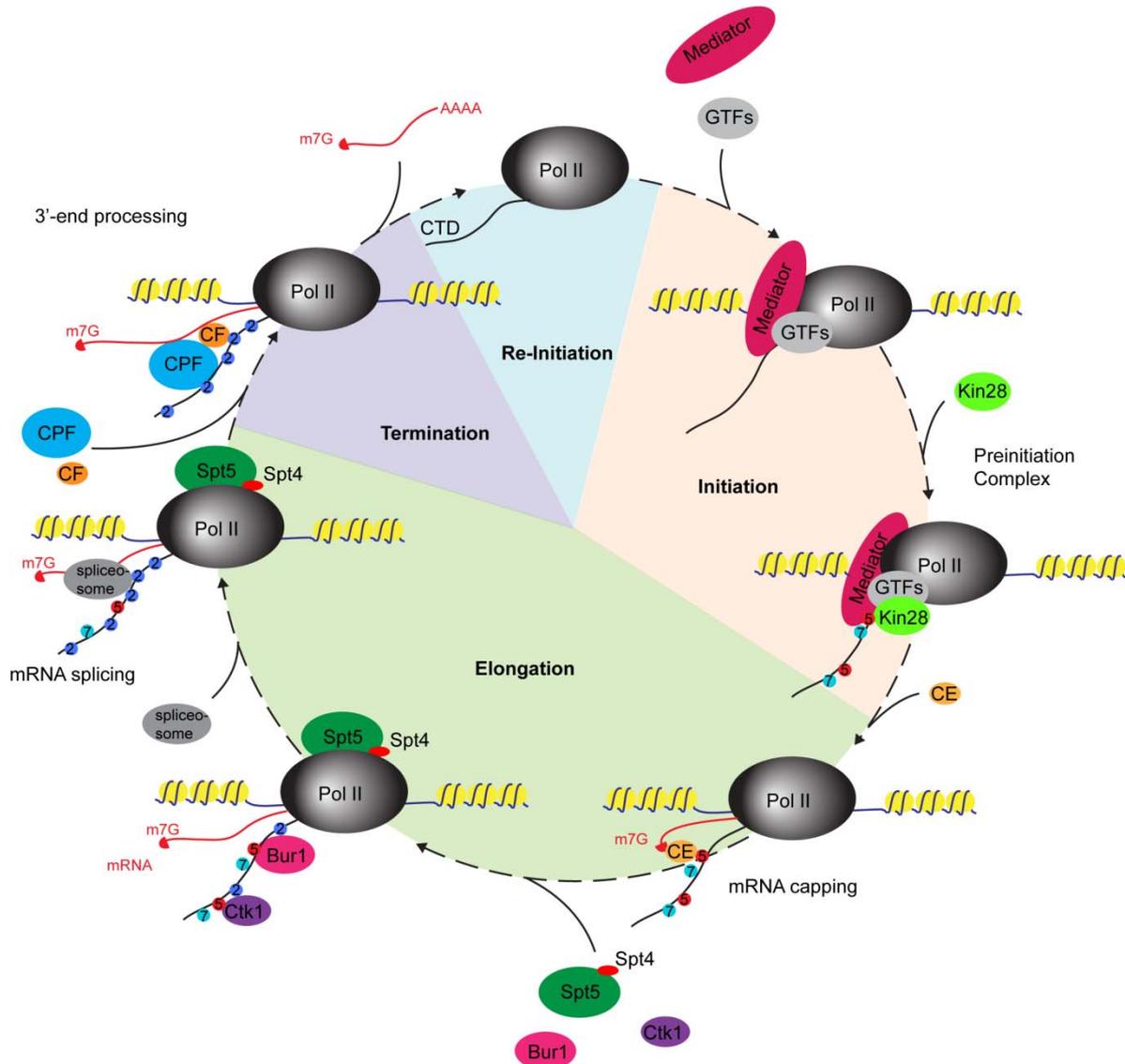


Figure 3: Schematic overview of the Pol II transcription cycle in *S. cerevisiae*.

DNA is depicted as a blue line with histones in yellow and RNA as a red line. Pol II and its CTD are in black. CTD phosphorylations Ser2-P, Ser5-P, Ser7-P are depicted as blue, red and light blue dots, respectively. (GTFs: general transcription factors; CE: capping enzyme; CF: cleavage factor; CPF: cleavage and polyadenylation factor). For details see within text.

1.3.2. Transcription elongation

After transcription initiation, the Ser5-phosphorylated CTD (Ser5-P) together with the C-terminal repeats of transcription elongation factor Spt5 (chapter 1.3.3) recruits the RNA 5'-capping machinery (50, 74, 75), which adds the 5' guanine-N7 methyl cap (m7G) to nascent RNAs to prevent them from degradation (76) and to enhance mRNA translation initiation (77). The capping reaction comprises three steps: In *S. cerevisiae* at first the RNA 5' triphosphatase Cet1 removes the γ -phosphate from the 5' triphosphate, then the guanylyltransferase Ceg1 adds an inverted guanylyl group by 5'-5' triphosphate linkage, and

finally the methyltransferase Abd1 adds a methyl group to the N7 atom of the terminal guanine group (78). Ceg1 and Abd1 thereby have been shown to interact directly with the Pol II CTD (74, 79). Ser5-P also recruits the early termination factor Nrd1 (80) and the histone methyltransferase Set1, which trimethylates histone 3 on lysine 4 (H3K4me3) (81), a chromatin mark for transcriptional activity. A function for the high level of Ser7 phosphorylation (Ser7-P) at the 5'-ends of genes has so far only been shown for mammals, where Ser7-P recruits the Integrator complex, which is involved in snRNA 3'-processing (82). A function of Ser7-P in yeast still has to be determined. However, as it is lethal in yeast to mutate Ser7 to Glu (chapter 1.2.1), this residue most likely has to be unphosphorylated at some stage and therefore, a negative function of Ser7-P is possible.

During transcription elongation, Ser5-P is gradually dephosphorylated either by Rtr1 or Ssu72 (57), whereas Ser2 is phosphorylated (Ser2-P) by the joint action of Bur1 and Ctk1 (43). In the course of this process, transcription elongation factors are recruited to Pol II to enable productive elongation. Spt4 and Spt5 are needed to ensure Pol II transcription processivity (chapter 1.3.3, (83, 84)). The Paf1 complex, the histone chaperone Spt6, and the FACT complex are important for ensuring proper transcription through chromatin: The Paf1 complex consists of the five proteins Paf1, Ctr9, Cdc73, Rtf1, and Leo1 in yeast and is involved in recruitment of histone methyltransferase Set1 (85). It is recruited via the phosphorylated C-terminal repeats of Spt5 (86, 87) (chapter 1.3.3). Spt6 directly interacts with Pol II via the CTD phosphorylated on Ser2 (88) and together with the FACT complex it is involved in nucleosome remodeling to enable transcription through chromatin and to prevent transcription from cryptic promoters (89, 90).

The Ser2-P/Ser5-P double phosphorylation mark of the CTD recruits the histone methyltransferase Set2, which trimethylates histone 3 on lysine 36 (H3K36me3) (91). The histone marks H3K4me3 and H3K36me3 marks add another level of transcriptional regulation to the CTD code, the "histone code" (92). Additionally, the general mRNA export factor Yra1 is recruited by the bivalent Ser2-P/Ser5-P mark on the yeast CTD (93). It binds the CTD as well as RNA and targets the mRNAs for nuclear export (94). Proline isomerase Ess1 and splicing factor Prp40 have been shown to interact with an *in vitro* hyperphosphorylated CTD (95, 96). The direct interaction between a splicing factor and the CTD suggests that the CTD can also regulate splicing. Pre-mRNA splicing removes non-coding parts of a gene (introns) and joins the two flanking DNA sequences (exons) (97). In *S. cerevisiae*, only 5% of the genes contain introns (98). The spliceosome complex, consisting of five small nuclear ribonucleic particles (snRNPs) and numerous other proteins, catalyzes the splicing reaction, which is performed in two transesterification reactions (97, 99).

With increasing Ser2-P levels during transcription elongation, RNA cleavage and polyadenylation factors are recruited (chapter 1.3.4).

complex (86, 87). Mammalian Spt5 recruits the activation-induced cytidine deaminase to DNA during antibody gene diversification (121). Yeast Spt4/5 recruits She2 to nascent RNA, coupling mRNA localization with Pol II transcription (122). Recruitment of factors can be mediated by the CTR of Spt5 (86, 123, 124). The yeast CTR recruits the Paf1 complex as well as the capping enzyme (75, 86, 87), and the fission yeast CTR binds the capping enzyme (119). Recently, the CTR was implicated to be involved in recruiting the histone deacetylase subunit Rco1 (125). The CTR forms a repeat structure similar to the Pol II CTD (103). The *S. cerevisiae* CTR consists of 15 hexapeptide repeats of the consensus sequence S[T/A]WGG[A/Q] (positions where alternative amino acids can occur between different repeats are indicated by brackets, and varying amino acids are indicated by slashes), whereas the human CTR consists of pentapeptide repeats with the consensus sequence GS[R/Q]TP (124) and the fission yeast CTR consists of nonapeptide repeats with the consensus sequence TPAWNSGSK (119). Deletion of the Spt5 CTR in yeast is not lethal (86, 87, 126) but leads to sensitivity to 6-AU and a slow-growth phenotype at 16°C (86, 87). The CTR deletion is synthetically lethal with the deletion of the gene coding for the Pol II CTD kinase Ctk1 (87). Deletion of the CTR in fission yeast leads to a slow-growth phenotype and abnormal cell morphology (123). The slow-growth phenotype is intensified if the Pol II CTD is truncated (123), suggesting that the CTR cooperates with the CTD. Deletion of the CTR impairs embryogenesis in zebrafish and leads to a derepression of gene transcription in zebrafish and human cells (127). Similar to the Pol II CTD, the CTR of Spt5 can be phosphorylated by the kinases Bur1 and P-TEFb in yeast and humans, respectively (86, 87, 124). CTR phosphorylation promotes transcription elongation in yeast and is important for the cotranscriptional recruitment of the Paf1 complex and for histone modification (86, 87). In human cells, CTR phosphorylation by P-TEFb converts Spt5 from a negative to a positive elongation factor (124). The Spt5 CTR may also play a role in the suppression of transcription-coupled nucleotide excision repair in yeast (126) and is involved in RNA 3'-end processing (chapter 4).

1.3.4. Transcription termination and RNA 3'-end processing

Protein-coding genes are transcribed until about 200 bp past the pA site in yeast, where Pol II is then dissociated from the DNA (128). Two models have been proposed for the termination process (129): In the “anti-terminator model”, Pol II dissociates from the DNA at a certain point when the pA site signal triggers the release of an anti-terminator factor within the transcription machinery (130, 131). In the “torpedo model”, the mRNA is cleaved at the pA site and as Pol II transcribes further downstream, a new uncapped 5'-end of this downstream product is exposed. This RNA is then degraded by the 5'-3' exonuclease Rat1/Rai1, which triggers Pol II termination (132, 133). The exonuclease complex also contains Rtt103, which directly interacts with Pol II phosphorylated on Ser2 of the CTD. This interaction enables tight coupling between transcription and termination (132, 134). The most important difference between the two models is that in the “torpedo” model, RNA cleavage is an

absolute prerequisite for termination, whereas it is not needed in the “anti-terminator model” (131). Most likely, termination takes place as a mixture of both models (135).

After Pol II has transcribed past the pA site, the nascent mRNA is processed in two steps: endonucleolytic cleavage and addition of a polyA tail (136). These processes depend on two multi-subunit complexes in yeast: cleavage factor I (CFI) and CPF (137). CFI can be separated into CFIA and CFIB (138, 139). CFIA consists of Clp1, Rna14, Rna15, and Pcf11 (140-142), CFIB consists of Hrp1 (138, 143, 144). Whereas all CFIA subunits have homologs in mammalian cells, no homologs of CFIB are currently known in higher eukaryotes (137). Both complexes can be recruited to the transcription machinery via the phosphorylated Pol II CTD (42). This recruitment mechanism is very important in addition to RNA binding of the 3'-end processing factors as deletion of the CTD leads to defects in RNA cleavage (145). This again indicates a tight coupling between the CTD and mRNA processing. The CFI subunit Pcf11 interacts with Ser2-phosphorylated CTD via its CTD-interacting (CID) domain (65). Ser2 phosphorylation peaks around the pA site and therefore serves as a termination signal (43, 146). Pcf11 bridges between Pol II and the RNA and by these two interactions it can dismantle elongation complexes *in vitro* (147). Two other subunits, Rna14 and Rna15 have also been shown to bind phosphorylated CTD *in vitro* (148). CPF subunits Ydh1, Yhh1 and Pta1 interact with the phosphorylated CTD (149-151). The necessary RNA signal sequences consist of an adenine-rich efficiency element, an adenine-rich positioning element about 30 bp upstream of the cleavage site and a uridine-rich element around cleavage and pA site (128). CPF and CFI bind to these RNA signals when Pol II transcribes over the pA site. The RNA is cleaved and polyadenylated by polyA polymerase Pap1 (141) and polyA binding proteins (PAB) then protect the 3'-end from degradation (152). When the mRNA is polyadenylated, it can be exported from the nucleus to enter translation. The Pol II CTD is dephosphorylated by Ssu72, a member of the CPF complex present at the 3'-ends of genes (153) (chapter 3) and Fcp1, which dephosphorylates Ser2-P (154, 155). The hypophosphorylated polymerase can then enter a new round of transcription (156).

The following chapter describes the discovery and first characterization of another important CTD modification in yeast, Tyr1-P, which was also found to play a role in transcription termination (44).

1.4. CTD Tyr1 phosphorylation impairs termination factor recruitment to Pol II

In the following chapter, parts of the recently published paper with the above title are presented (44). Author contributions are listed on page IV.

It was observed by Mayer et al. that the profiles of CTD Ser2 phosphorylation and Pcf11 occupancy do not correlate well (43). Pcf11 is recruited later than an interaction with Ser2-P might suggest. Therefore, the question was raised, whether Ser2-P could be masked along the coding region of genes and only become accessible for Pcf11 at the pA site (43). A possible

candidate for masking Ser2-P could be Tyr1 phosphorylation of the CTD (44). Although CTD phosphorylation on Tyr1 has long been known for human Pol II (62), it has never been shown in yeast. Since now specific antibodies against single phosphoresidues of the CTD are available, e.g. against Ser2-P (3E10) and Ser5-P (3E128) (generated by Dirk Eick and Elisabeth Kremmer, Helmholtz-Zentrum München), these modifications can be detected well. In this study, a similar monoclonal antibody for Tyr1-P (3D12) was generated (44).

1.4.1. A monoclonal antibody against Tyr1-phosphorylated CTD

Since the functional CTD unit is a pair of repeats (39), we determined antibody specificity using diheptapeptides bearing combinations of phosphorylations (Figure 5a, (44)). This revealed a high affinity for the Tyr1-phosphorylated CTD that was not impaired by adjacent Ser2-P, and no affinity to other CTD peptides (Figure 5a, (44)). The antibody immunoprecipitated Pol II from extracts of the yeast *S. cerevisiae* (Figure 5b), and the precipitated polymerases were also phosphorylated at Ser2, Ser5 and Ser7 (44). The antibody also recognized Pol II that was purified from human cells with antibody 1C7 (44) and phosphorylated *in vitro* by the Tyr1 kinase c-Abl (157) (Figure 5c). Thus, antibody 3D12 specifically recognizes the Tyr1-phosphorylated CTD, and Tyr1 phosphorylation occurs in yeast.



Figure 5: Pol II CTD is phosphorylated at Tyr1 (44).

(a) Part of the CTD sequence around phosphorylated Tyr1 (Y1). Residues Ser2, Pro3, Thr4, Ser5, Pro6, and Ser7 are denoted S2, P3, T4, S5, P6, and S7, respectively. CTD residues that interfere with 3D12 antibody binding upon phosphorylation are highlighted in black. (b) Western blot analysis of whole-cell extract from proliferating yeast (Input). Pol II was immunoprecipitated with antibodies 8WG16, 3D12, and 1C7 (IP Pol II) and probed with 8WG16 or 3D12. Isotype controls are shown. Ig, immunoglobulin. (c) Antibody 3D12 detects CTD Tyr1 phosphorylation in HeLa cells (Input). Unphosphorylated Pol II was immunoprecipitated with antibody 1C7 (IP 1C7, (38)) and incubated with c-Abl kinase, leading to a 3D12 signal (IP + c-Abl). The hyper- (II0) and hypophosphorylated forms (IIA) of Pol II are indicated.

1.4.2. Tyr1 is phosphorylated in the coding region of genes

To investigate whether genome-associated Pol II is phosphorylated at Tyr1, we used high-resolution ChIP profiling in proliferating yeast (43). Data from two biological replicates (Pearson correlation coefficient $R = 0.94$, (44)) were averaged and revealed strong signals over protein-coding and small nucleolar RNA genes. Gene-averaging of ChIP profiles (43) revealed Tyr1 phosphorylation in the coding region (Figure 6a, b). Whereas Tyr1-P signals were low at promoters, they increased downstream of the TSS. The gene-averaged profile resembled that for Ser2 phosphorylation, except that Ser2-P signals persist downstream of the pA site for ~200 bp, whereas Tyr1-P signals decrease already ~180 bp upstream of the pA site (Figure 6a, b). The point of Tyr1-P signal increase was dependent on the TSS, whereas the point of decrease was dependent on the pA site, but not on gene length or expression level (44). These results indicate that Tyr1-P marks are set and removed within the transcription cycle.

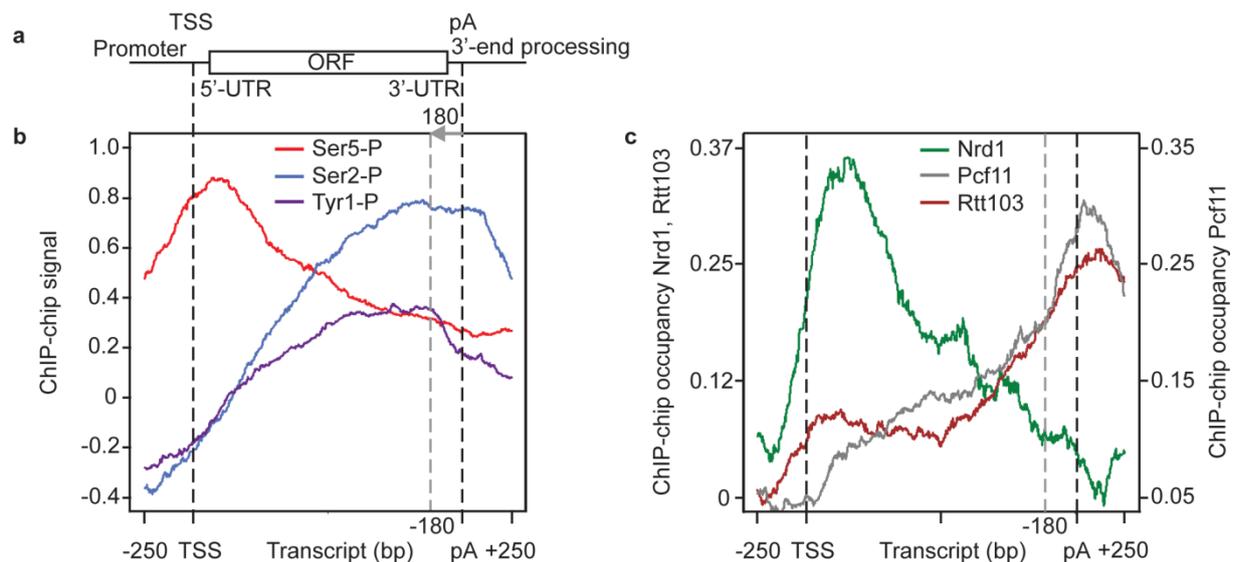


Figure 6: Gene-averaged ChIP profiles for CTD phosphorylations and termination factors (44).

(a) DNA frame with promoter, 5'-untranslated region (UTR), open reading frame (ORF), and 3'-UTR. Dashed black lines indicate the TSS and pA sites. The dashed gray line marks the position 180 bp upstream of the pA site. (b) Gene-averaged profiles for Ser5 (43), Ser2 (43), and Tyr1 phosphorylation for 339 genes of medium length (1238 ± 300 bp). (c) Gene-averaged profiles for Nrd1, Pcf11 (43) and Rtt103. ChIP-chip occupancy of Nrd1 and Rtt103 is on the left y-axis, Pcf11 occupancy on the right y-axis.

1.4.3. Tyr1-P negatively influences termination factor recruitment

To investigate whether Tyr1 phosphorylation influences factor recruitment to Pol II, we determined genomic occupancy profiles for termination factors Nrd1, Rtt103, and Pcf11, which contain a CTD-interacting domain. The gene-averaged Nrd1 occupancy peaked at the beginning of the transcribed region, 193 ± 44 bp downstream of the TSS (Figure 6c). This region also showed maximum signals in Ser5-P, and genomic Nrd1 and Ser5-P profiles

correlate ($R = 0.6$), consistent with Nrd1 binding to the Ser5-phosphorylated CTD (80). The general presence of Nrd1 at protein-coding genes extends previous results (158, 159) and befits a role of Nrd1 in early transcription termination (15, 160-162). Rtt103 showed peak occupancy at the end of genes, 112 ± 27 bp downstream of the pA site, where peak levels of Pcf11 were also observed (43) (Figure 6c, (44)). Since this region shows the maximum difference between Ser2 and Tyr1 phosphorylation signals, Tyr1-P may impair recruitment of Rtt103 and Pcf11 upstream of the pA site. Consistent with this, genome-wide occupancies of Rtt103 and Pcf11 do not correlate well with Ser2-P signals ($R = 0.4$, for both), although both proteins bind the Ser2-phosphorylated CTD (65, 163).

To test whether Tyr1 phosphorylation impairs CTD binding of termination factors, we determined the affinity of purified recombinant CIDs of yeast Nrd1, Pcf11, and Rtt103 for various CTD diheptad phosphopeptides using fluorescence anisotropy (44). None of the CIDs bound to an unphosphorylated CTD peptide. Consistent with previous results (80, 163), Pcf11-CID and Rtt103-CID bound to the Ser2-phosphorylated CTD peptide (Dissociation constant $K_D = 54 \pm 6 \mu\text{M}$ and $12 \pm 2 \mu\text{M}$, respectively; (44)), whereas the Nrd1-CID preferentially bound to a Ser5-phosphorylated CTD peptide ($K_D = 85 \pm 25 \mu\text{M}$). In contrast, none of the CIDs bound Tyr1-phosphorylated CTD peptides, regardless of whether additional phosphorylations were present or not. Thus, Tyr1 phosphorylation blocks CID binding to the CTD *in vitro*, consistent with the hypothesis that it impairs termination factor recruitment *in vivo*.

1.4.4. The extended CTD code

Our results extend the previously proposed CTD code (68, 164, 165), which was based on Ser2 and Ser5 phosphorylation, leading to an extended CTD code for the coordination of the transcription cycle with factor recruitment (Figure 7).

During initiation and early elongation, the CTD is phosphorylated on Ser5, which facilitates recruitment of the capping enzyme and Nrd1. 150-200 bp downstream of the TSS, peak occupancy levels are reached for Nrd1 and Pol II (43), likely marking a decision point where Pol II transiently pauses and either terminates or continues elongation (15). When Tyr1-P and Ser2-P levels rise, Pol II binds elongation factors stably and continues elongation. Tyr1-P releases Nrd1 and impairs recruitment of Rtt103 and Pcf11, suppressing termination during elongation. Before the pA site, Tyr1-P levels drop, whereas Ser2-P levels remain high. This enables recruitment of Rtt103 and Pcf11 that is enhanced by cooperative interactions between factors (163) and with nascent RNA (159), resulting in RNA 3'-end processing and transcription termination. Our results indicate that Tyr1 CTD phosphorylation is a target to activate transcription by suppressing Pol II termination, and explain why mutation of Tyr1 to phenylalanine, which lacks the oxygen atom required for phosphorylation, is lethal (27).

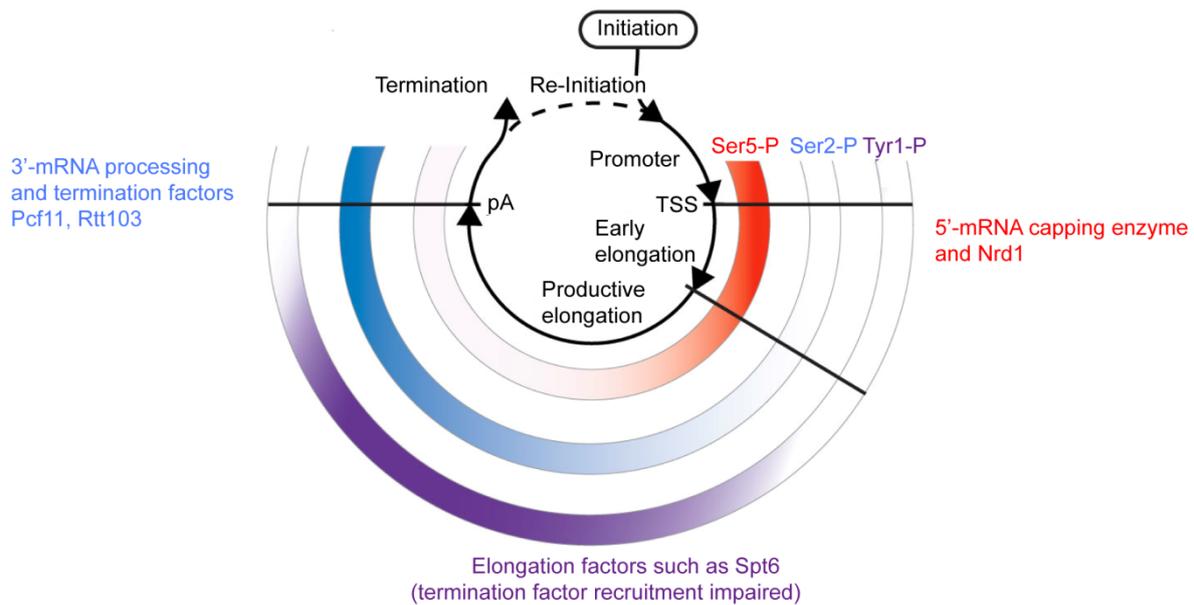


Figure 7: Extended CTD code for transcription cycle coordination (44).

During the cycle, levels of CTD phosphorylation at Tyr1, Ser2, and Ser5 residues change differently, as illustrated by gradients of red (Ser5-P), blue (Ser2-P) and violet (Tyr1-P) in three semicircles.

1.5. Aims and scope of this work

Gene transcription in eukaryotes is a highly regulated process (chapter 1.3). It has been known for a long time that the Pol II CTD plays a central role in regulating transcription (68). The CTD coordinates transcription with RNA processing steps in a chromatin environment. During the transcription cycle different posttranslational phosphorylations on the CTD facilitate the recruitment of many different factors by generating a CTD code (164). Phosphorylations on Ser5 and Ser2 residues are already established as main regulators of transcription and RNA processing: Ser5 phosphorylation is important for capping enzyme recruitment (50, 74), which in turn promotes the transition from transcription initiation to elongation (75). In contrast to that, Ser2 phosphorylation is involved in transcription termination as it recruits Pcf11 (65).

We found recently that the existing CTD code can be extended by an additional phosphorylation, Tyr1-P (chapter 1.4.4, (44)). Its genome-wide profile peaks towards the 3' region of genes before the pA site (Figure 6). A possible role of Tyr1-P in regulating termination was therefore assumed.

As we found the first evidence of Tyr1 phosphorylation in yeast, this modification, its kinase or phosphatase and its function have been completely uncharacterized until now. Three questions immediately arose, that were addressed in this work: (i) which kinase phosphorylates Tyr1 in yeast (chapter 2); (ii) which phosphatase dephosphorylates Tyr1-P (chapter 3); (iii) what function does Tyr1-P have in transcription termination (chapter 3).

To address these questions we mainly used ChIP, a method that measures the binding of proteins to specific DNA sequences (chapter 6.2.6, (166)). In short, by addition of formaldehyde DNA, proteins and RNA in close proximity are crosslinked, the DNA is then fragmented and the protein/modification of interest immunoprecipitated. DNA sequences that are bound to this protein/modification are enriched in the immunoprecipitated sample and after reversal of crosslinks can be quantified. Transcribing Pol II modifications or interacting factors along genes can then be quantitatively mapped to DNA regions. This mapping can either be performed for single genes by quantitative real-time PCR (qPCR) or genome-wide by ChIP-on-chip using microarray hybridization (chapter 6.2.7).

By selectively inhibiting the four canonical CTD kinases Bur1, Ctk1, Kin28 and Srb10, followed by single-gene ChIP analysis of Tyr1 phosphorylation levels before and after inhibition, we analyzed these enzymes for their phosphorylation activity on Tyr1 *in vivo* (chapter 2.2). A candidate list of 15 kinases was also screened for changes in Tyr1 phosphorylation levels by comparing either wild-type with kinase deletion strain (chapter 2.3.1) or untreated with conditionally kinase-depleted strain (chapter 2.3.2).

For the phosphatase search we took a slightly different approach. As the Tyr1 phosphorylation signal strongly decreases before the pA site in the 3' region, we found two likely candidate phosphatases within the CPF. One enzyme, Ssu72, has already been shown to function as a CTD phosphatase (56), whereas the second candidate, Glc7, has never been implicated in CTD dephosphorylation. By dephosphorylating phospho-Pol II with purified CPF *in vitro* and genome-wide ChIP of Tyr1-P in a candidate anchor-away strain *in vivo*, we aimed for identification of a Tyr1-P phosphatase (chapters 3.1, 3.2). By genome-wide ChIP of termination factors in this candidate anchor-away strain, we were able to characterize the role of Tyr1-P in transcription termination (chapter 3.4).

Another repetitive element within a protein that was shown to play a regulatory role in transcription is the Spt5 CTR (86, 123, 124). It shows a repeat structure similar to Pol II CTD, can also be phosphorylated and recruits factors to Spt5 and Pol II in a phosphorylation-dependent manner (87). Furthermore, the elongation factor Spt5 colocalizes with Pol II along the complete coding region of genes and past the pA site (43). Therefore, we asked whether more transcription-related proteins can be found that are recruited via the CTR and in which part of the transcription cycle they are important. We addressed these questions by single-gene ChIP of transcription and RNA processing factors in wild-type and Spt5 Δ CTR strains and analyzed the profiles for changes in recruitment of these factors (chapter 4.2). Additionally, we characterized CTR interactions with recruited factors *in vitro* (chapter 4.3) and analyzed functional implications of recruitment defects.

2. Towards the identification of the CTD Tyr1 kinase in yeast

Results presented in chapter 2.2 are published in (44). Detailed author contributions are listed on page IV.

Mayer et al. reported that the CTD of yeast can be phosphorylated at position Tyr1, which extends the CTD code of Ser2 and Ser5 phosphorylation (chapter 1.4.4, (44)). The responsible kinase remains unknown. In human, c-Abl is known to phosphorylate Tyr1 *in vitro* (Figure 5, (157)), but this kinase does not have a homolog in yeast.

Tyrosine residues are generally phosphorylated by typical tyrosine kinases or by dual-specificity kinases (167). The following chapters present an overview of protein kinases in yeast (chapter 2.1), show published ChIP data of CTD phosphorylation signals in strains with inhibited canonical CTD kinases (chapter 2.2) and unpublished ChIP data of CTD phosphorylation signals in strains with 15 other deleted or conditionally depleted kinases (chapter 2.3).

2.1. Protein kinases in yeast

The yeast genome contains 129 protein kinases classified in 7 groups (<http://www.yeastkinome.com/>, (168, 169)). Kinases are one of the largest groups of proteins and constitute around 2% of the eukaryotic genome (169). They can be subdivided into different kinase groups which are generally conserved among eukaryotic organisms but there are also subfamilies specific for yeast (169). These kinases mostly play a role in functions especially important for unicellular organisms like nutrient uptake, pseudohyphal growth (170), stress responses, cell wall signaling, cell cycle or small-molecule transport (169, 171). Table 3 gives an overview of the 7 kinase groups in yeast. Enzymes are grouped according to structural similarities (168, 169).

Table 3: 7 kinase groups in *S. cerevisiae* (adapted from (169) and <http://www.kinase.com>).

Kinase group	Members	Properties
AGC	17	Named after members protein kinase <u>A</u> , <u>G</u> and <u>C</u>
Atypical	13	No sequence similarity but kinase function
CAMK	22	Calcium-Calmodulin dependent
CK1	4	Casein kinase 1 group
CMGC	23	Named after members <u>CDK</u> , <u>MAPK</u> , <u>GSK3</u> and <u>CLK1</u>
STE	14	MAP kinase cascade kinases
Others	36	No strong similarities to the other groups

In comparison to multicellular organisms, *S. cerevisiae* does not have receptor-type protein kinases, possibly because cell-cell communication is only needed on a very basic level (170). Furthermore, no true protein-tyrosine kinase is present in yeast as these enzymes are mostly involved in signaling pathways (170). Therefore, the Pol II CTD Tyr1 kinase in yeast could be a Ser/Thr kinase with dual specificity, like for example members of the Wee1 or the MAP kinase kinase (MAPKK) family (170). Generally, Tyr1 specificity cannot be excluded for any Ser/Thr kinase before tested *in vitro* (170).

However, with 129 kinases already known and possibly more unconventional enzymes being discovered in the future, it is necessary to systematically search for the specific CTD Tyr1 kinase. As a consequence, we decided to first test the four canonical yeast CTD kinases.

2.2. Tyr1 of the Pol II CTD is not phosphorylated by canonical CTD kinases

It was tested whether CTD Tyr1 phosphorylation depends on one of the yeast CTD kinases Kin28, Srb10, Bur1, or Ctk1, which correspond to human Cdk7, Cdk8, Cdk9, and Cdk12, respectively (48). Therefore, ChIP of CTD phosphorylations on single genes in catalytically inactive kinase mutant strains was performed. The respective kinases in these analog-sensitive (as) strains are mutated in their ATP binding pocket, where a bulky amino acid side chain is changed to the smaller glycine. This creates a new binding pocket that can be used by inhibitors, in this case pyrazolo[3,4-d]pyrimidine compounds (172). These inhibitors or a solvent control are added to yeast cultures 60 min before formaldehyde crosslinking (Methods) to inhibit kinase activity. ChIP experiments were performed using the Tyr1-P antibody as well as the phospho-specific antibody against the known CTD substrate (Ser2-P or Ser5-P) or the Rpb3 antibody as a control. Finally, qPCR was carried out using primers for two housekeeping genes, *ADHI* and *PMAI* (Figure 8).

2. Towards the identification of the CTD Tyr1 kinase in yeast

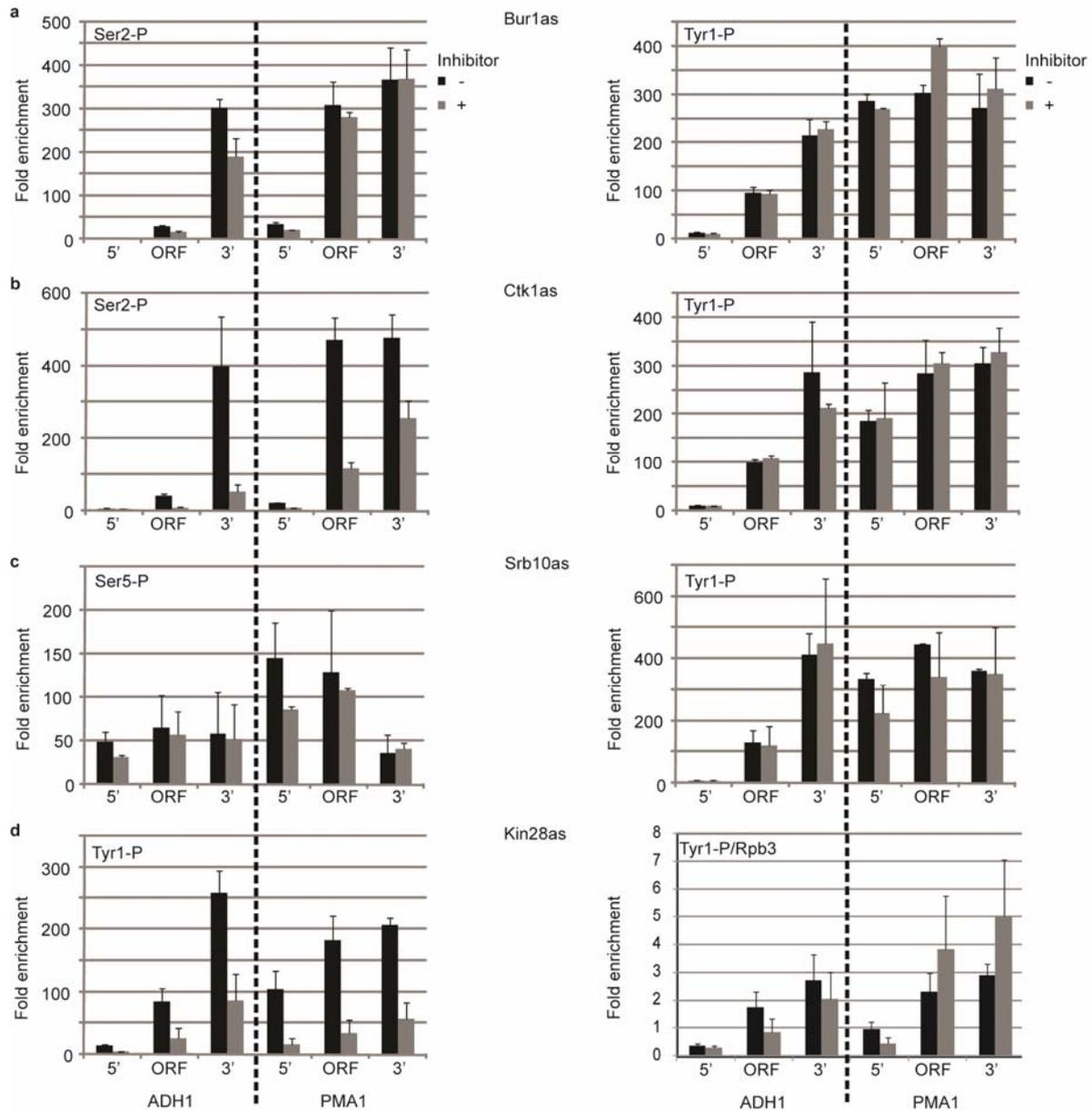


Figure 8: CTD Tyr1 phosphorylation levels are not changed upon inhibition of the canonical Pol II CTD kinases.

Kinase inhibition and ChIP experiments were performed with analog-sensitive yeast strains (as) (173). (a) Bur1as strain was inhibited with the small molecule inhibitor 3-MB-PP1, Ctk1as strain (b) and Srb10as strain (c) were inhibited with 1-NM-PP1, Kin28as strain (d) with 1-NA-PP1. ChIP occupancies of Pol II and of its different phosphorylated forms before and after kinase inhibition are indicated as black and gray bars, respectively. Results are shown for three different regions of the *ADH1* and *PMA1* gene. Fold enrichment values refer to a heterochromatic region on chromosome V that is not transcribed by Pol II. Standard deviations refer to at least two biological replicate measurements.

Inhibition of a potential Tyr1 kinase in as strains should result in a decrease of the Tyr1-P signal compared to the untreated sample. According to Figure 8, however, this is not the case for inhibition of Bur1, Ctk1, Kin28 or Srb10:

Bur1 inhibition slightly reduces Ser2 phosphorylation levels at the coding region (ORF) of *ADHI* (p-value: 0.02; T-test) and the 5' region of *PMAI* (p-value: 0.05) (Figure 8a, left panel). These results are in agreement with reported observations (60). Bur1 inhibition did not change Tyr1-P levels (Figure 8a, right panel). Ctk1 inhibition significantly reduces Ser2-P occupancy levels as expected (p-values ≤ 0.05) (Figure 8b, left panel), serving as a positive control, whereas Tyr1-P levels were not altered (Figure 8b, right panel). Srb10 inhibition neither changes Ser5-P occupancy levels nor the corresponding levels of Tyr1-P (Figure 8c). Kin28 inhibition leads to a reduction of Tyr1-P levels over the whole length of the transcribed region (Figure 8d, left panel). However, this was apparently due to an overall reduction of Pol II occupancy upon Kin28 inhibition as the ratios between Tyr1-P and Pol II subunit Rpb3 ChIP occupancies do not change within the range of the standard deviation (Figure 8d, right panel).

Taken together, inhibition of the four canonical CTD kinases *in vivo* did not significantly affect Tyr1 phosphorylation signals (Figure 8). This indicates that Tyr1 phosphorylation of the yeast CTD depends on a kinase other than the known CTD kinases. Consistent with this, Tyr1 phosphorylation in human cells is achieved by c-Abl (157), a kinase that lacks a yeast homolog.

2.3. Systematical approach to identify the Tyr1 kinase

The ChIP results above clearly indicate that Tyr1 is not phosphorylated by a known CTD kinase. To identify the kinase, a systematic search in a pool of kinase candidates *in vivo* seemed to be the most promising approach. The first candidates were chosen because of genetic or physical interaction with the transcription machinery, nuclear localization, or tyrosine kinase activity. Table 4 gives an overview of candidates, a short functional characterization and their relation to transcription or the Pol II CTD. This table does not include a complete list of possible Tyr1 kinases but rather a first selection:

Table 4: Tyr1 kinase candidates.

Kinase	Group	Function	Relation to CTD
Bck1	STE	MEK-kinase (174), cell wall integrity signaling (175)	genetic interaction with TFIID (176), Ctk1 (177)
Bdf1	atypical	localized uniformly along the length of chromosomes (178), two "bromodomains" (179)	human homolog Brd4 phosphorylates Ser2 (180)
Cdc5*	other	functions in mitosis and cytokinesis (181)	genetic interaction with Mex67 (182)
Cdk1/ Cdc28*	CMGC	catalytic subunit of the main cell cycle cyclin-dependent kinase (183)	involved in CTD Ser5 phosphorylation (55)

2. Towards the identification of the CTD Tyr1 kinase in yeast

Kinase	Group	Function	Relation to CTD
Fus3	CMGC	MAPK, control of cell proliferation (184)	physical interaction with Pol II (185)
Hrr25*	CK1	involved in chromosome segregation (186); homolog of mammalian casein kinase 1delta (187)	interaction with Ctk1 (177) and Pol II (188)
Kns1	CMGC	LAMMER family of protein kinases, dual specificity (189)	unknown
Kss1	CMGC	MAPK, control of cell proliferation (184)	physical interaction with Spt6 (185), genetic with TFIID (190)
Mck1	CMGC	dual-specificity protein kinase related to mammalian GSK-3 family (191, 192)	genetic interaction with Pol II, TFIID, Paf1, Mediator (193)
Mpk1	CMGC	MAPK, mediates PKC signaling (175)	physical interaction with Paf1 complex (160)
Pbs2	STE	MAPKK, osmoregulatory signal transduction cascade (194)	genetic interaction with Ctk1 (177), Pcf11 and Mediator (190)
Snf1	CAMK	phosphorylation of histone H3 (195)	genetic interaction with Ctk1 (177), physical interaction with Gcn5 (196)
Swe1	other	G2/M transition, inhibits Cdc28 through phosphorylation of Y19 (197)	physical interaction with TFIID (185), genetic interaction with Ccr4 (198)
Tpk1	AGC	PKA catalytic subunit, nutrient response via the Ras-cAMP signaling pathway (199, 200)	physical interaction with Spt5, Pol II (168)
Yak1	CMGC	serine-threonine protein kinase; component of a glucose-sensing system (201)	physical interaction with Ess1, splicing factor Cdc40 (202)

*essential genes

Yeast strains were cloned that contained either knock-outs of the non-essential kinases or anchor-away tags for the essential enzymes (Table 6). Conditional depletion of a nuclear protein by the anchor-away technique works as described (203): An FRB tag is fused to the protein of interest in a strain that carries an FKBP12-tag on the ribosomal subunit Rpl13A. Upon rapamycin addition, the anchor forms a ternary complex with the FRB tag and rapamycin, which leads to export of this complex from the nucleus within a few minutes as ribosomal subunits are rapidly exported to the cytoplasm. Most of these strains die in presence of rapamycin.

ChIP experiments were performed with all kinase strains using the Tyr1-P and the Pol II (Rpb3) antibody serving as a control. Knock-out strains were compared to wild-type, anchor-away strains were treated with rapamycin for 60 min before formaldehyde crosslinking and compared to DMSO treated cells (Methods). Chapters 2.3.1 and 2.3.2 provide an overview of the unpublished ChIP data for all kinase mutants.

2.3.1. Non-essential kinases

Kinase knock-out strains were analyzed by ChIP with the Tyr1-P and the Rpb3 antibody. Figure 9 depicts Tyr1-P occupancy levels in wild-type (wt) and 12 non-essential kinase knock-out strains normalized against total Pol II (Rpb3) levels.

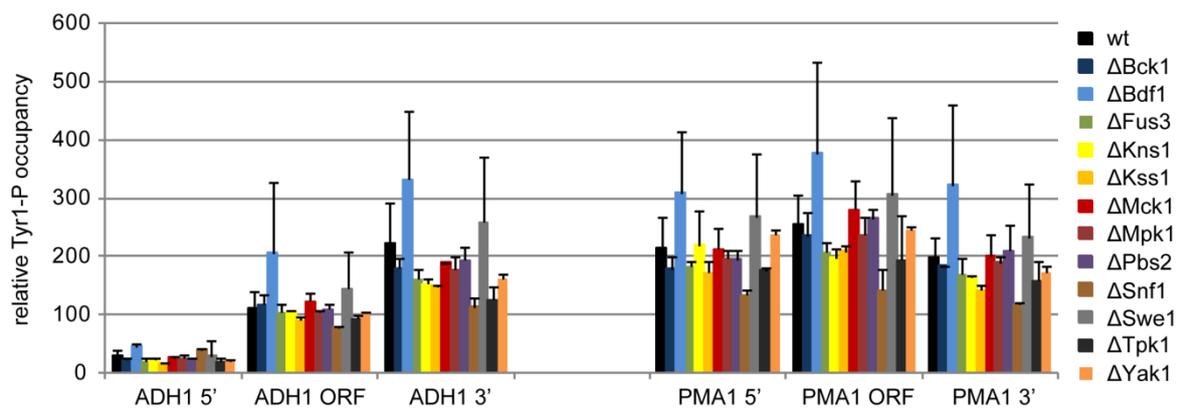


Figure 9: Tyr1-P occupancy levels are not decreased by deletion of 12 non-essential kinases.

Tyr1-P occupancy relative to Rpb3 (Methods) is shown for wild-type (black bars) and 12 mutants. Results are shown for three different regions of the *ADH1* and *PMA1* gene. Standard deviations refer to at least two independent ChIP experiments.

Deletion of a potential Tyr1 kinase should result in a decrease of the Tyr1-P signal in the mutant strain compared to the wild-type. According to Figure 9, however, this is not the case for 12 kinase deletions tested. Although the signals decrease slightly for some mutants, e.g. Δ Snf1 (brown bars), this decrease is within the range of the standard deviation for the wild-type (black bars). For Δ Bdf1 and Δ Swe1, a signal increase is detected, but the standard deviation is comparatively high. Taken together, the 12 kinases tested here are not identified as Tyr1 kinases.

2.3.2. Essential kinases

FRB tags on essential kinase genes were cloned and validated as described (Methods). Figure 10 shows spot dilutions of Cdc5-, Cdk1-, and Hrr25-FRB strains on YPD and YPD + rapamycin plates. These spot dilutions were performed in order to verify the successful export of the FRB-tagged protein from the nucleus upon rapamycin treatment resulting in cell inviability.

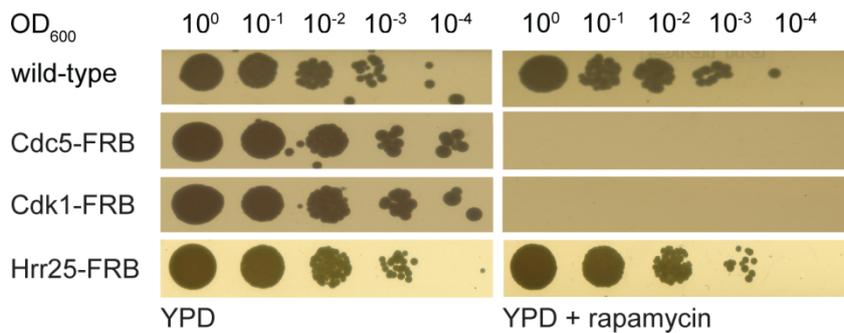


Figure 10: Growth analysis of anchor-away yeast strains.

Serial dilutions of wild-type yeast and strains with FRB-tagged kinases Cdc5, Cdk1 and Hrr25 plated on YPD (left panel) and YPD + rapamycin (right panel) show that rapamycin is lethal for Cdc5-FRB and Cdk1-FRB but it has no effect on Hrr25-FRB and wild-type growth.

To ensure that Hrr25 is exported from the nucleus, the protein was tagged with an additional GFP tag. Fluorescence microscopy pictures of the Hrr25-FRB-GFP strain without and with rapamycin are shown in Figure 11 (see Methods for experimental details).

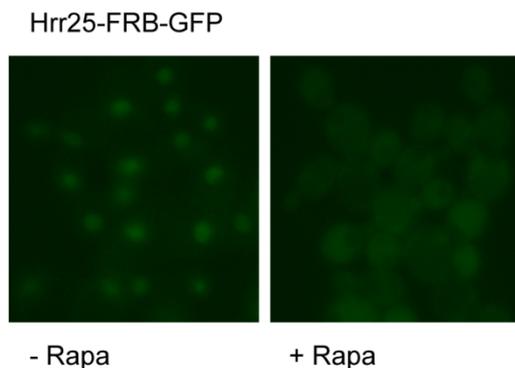


Figure 11: Fluorescence microscopy of the Hrr25-FRB-GFP strain.

without (left) and with rapamycin (Rapa) (right) shows that Hrr25 is exported from the nucleus upon rapamycin addition.

The fact that export of Hrr25 (Figure 11) from the nucleus is not lethal for the cells (Figure 10) possibly indicates that Hrr25 does not execute its essential function in the nucleus.

Figure 12 depicts Tyr1-P occupancy levels normalized against total Pol II (Rpb3) in the three essential kinase strains treated either with DMSO or rapamycin.

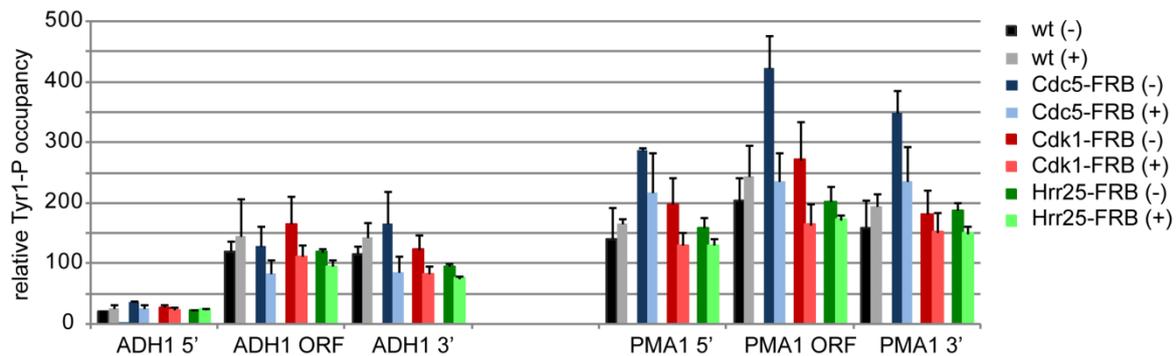


Figure 12: Tyr1-P levels in Cdc5, Cdk1 and Hrr25 anchor-away strains are not influenced by rapamycin treatment.

Tyr1-P occupancy relative to Rpb3 (Methods) is shown for untreated (-, dark bars) and treated samples (+, light bars). Results are shown for three different regions of the *ADH1* and *PMA1* gene. Standard deviations refer to two independent ChIP experiments for Cdc5-FRB, three for Hrr25-FRB and 11 for Cdk1-FRB.

Removal of a potential Tyr1 kinase from the nucleus by rapamycin treatment should result in a decrease of the Tyr1-P signal compared to the untreated sample. According to Figure 12, this is not the case for Cdk1 and Hrr25. For Cdc5, a slight decrease can be detected in the *PMA1* ORF region. A genome-wide Tyr1-P profile should clarify if this is a widespread effect. Taken together, the three enzymes tested here cannot be identified as Tyr1 kinases.

2.4. Discussion

20 years ago it was shown that the mammalian kinase c-Abl can phosphorylate the CTD on tyrosine both *in vivo* and *in vitro* (62). However, this enzyme does not have a homolog in *S. cerevisiae*. Jeronimo et al. tested the well-known CTD kinases and other potential candidates but could not observe kinase activity on Tyr1 in yeast (47).

The results shown here also demonstrate that inhibition of the canonical CTD kinases does not decrease Tyr1-P signals *in vivo*. Furthermore, deletion or conditional depletion of 15 other kinase candidates does not negatively influence Tyr1-P profiles. Although the results presented here generally confirm published data (47), they can be interpreted in different ways. Most probably all these enzymes tested are de facto not phosphorylating Tyr1 *in vivo*. However, there are also other possibilities: Another enzyme could compensate the deletion/depletion of the kinase, indicating more than one Tyr1 kinase. Alternatively, the deleted/depleted kinase could have only a slight effect on Tyr1 of the CTD, which cannot be detected by quantifying Tyr1-P levels on single-genes. Furthermore, indirect effects caused by deletion/depletion of a kinase which has other substrates than the CTD could influence the complete transcription process. This situation renders the data difficult to interpret.

To clarify these issues, Tyr1-P signals in the kinase mutants could be analyzed genome-wide. If none of the kinases influences Tyr1-P levels on a genome-wide scale, the candidate list would have to be extended. Another completely different approach could be to test purified *S. cerevisiae* kinases on a CTD substrate *in vitro*.

In summary, from our current point of view, although having tested 19 different kinases we cannot identify the Tyr1 kinase. Thus, it becomes increasingly likely, that the Tyr1 kinase is an enzyme that has not been mentioned in relation to transcription so far (47).

3. Pol II termination involves CTD Tyr1 dephosphorylation by CPF subunit Glc7

All results presented in chapter 3 were obtained in collaboration with Ashley Easter and Lori Passmore and are in the process of being published. For detailed author contributions see page IV.

As explained in chapter 1.4.3, Tyr1 phosphorylation prevents premature recruitment of termination factors within gene bodies (44). At the pA site, Tyr1 phosphorylation levels drop, whereas Ser2 phosphorylation remains, apparently enabling termination factor recruitment (44). Thus, the transition from transcription elongation to termination involves CTD Tyr1 dephosphorylation by an unknown phosphatase. We reasoned that the Tyr1 phosphatase might be part of the mRNA 3'-end processing and termination machinery. In particular, CPF contains two candidate phosphatases, Ssu72 (56, 204) and Glc7 (153). Ssu72 can dephosphorylate CTD residue Ser5 (56, 61, 205). Glc7 plays a role in transcription termination on snoRNA genes (206) but has no known CTD-related activity. Here, we show that Glc7 is a Tyr1 phosphatase both *in vitro* and *in vivo*. We further show that Glc7 is required for recruitment of termination factors and for normal Pol II termination *in vivo*.

3.1. CPF subunit Glc7 dephosphorylates CTD Tyr1 *in vitro*

To investigate whether Ssu72 or Glc7 has CTD Tyr1 phosphatase activity we purified endogenous CPF from yeast. Analysis of the entire CPF complex was necessary since isolated phosphatases generally have little specificity, and target-specific dephosphorylation often relies on cofactors (207). The resulting preparations contained all CPF subunits, including Ssu72 and Glc7 (Figure 13a). To test CPF preparations for CTD phosphatase activity, we used a substrate of purified yeast Pol II that had been pre-phosphorylated on Tyr1, Ser2, and Ser5 (Methods; Figure 13b, lane 1). Incubation with CPF led to almost complete dephosphorylation of Tyr1, Ser2 and Ser5, as revealed by Western blotting (Figure 13b, lanes 3-5).

To identify which of the two CPF phosphatases was responsible for Tyr1 dephosphorylation, we specifically inhibited Glc7 with EDTA or microcystin. EDTA chelates metal ions and abolishes activity of the metal ion-dependent phosphatase Glc7 (208, 209), but does not affect Ssu72 activity, which is metal ion-independent (204). Microcystin is a bacterial toxin that specifically inhibits PP1 and PP2A-type phosphatases (207) including Glc7, but not Ssu72. Addition of 10 mM EDTA or 200 nM microcystin to the reaction selectively inhibited Tyr1 and Ser2 dephosphorylation, whereas Ser5 dephosphorylation was still observed (Figure 13b, lanes 6-13). These results show that the Glc7 subunit of CPF functions as a CTD phosphatase for Tyr1 and Ser2 *in vitro*.

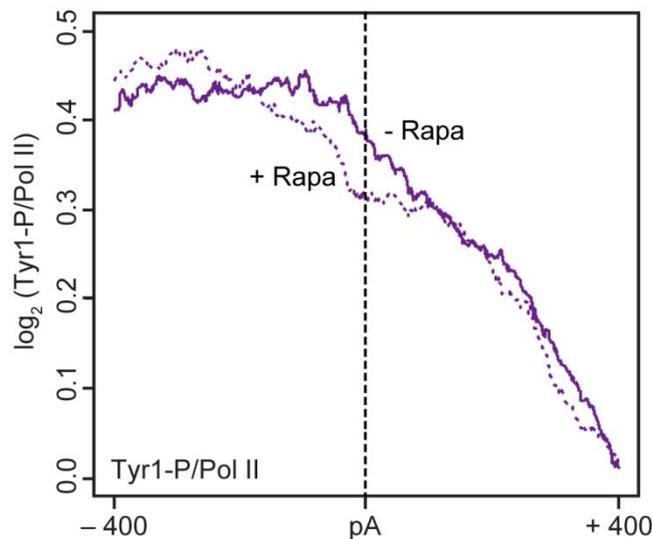


Figure 14: Genome-wide ChIP occupancy of Tyr1-phosphorylated Pol II around the pA site is not influenced by rapamycin in wild-type yeast.

ChIP-chip occupancy profiling of Tyr1-phosphorylated Pol II over 619 genes aligned at the pA site (dashed line) and normalized against the corresponding Rpb3 profile without and with rapamycin (-/+ Rapa, violet line and violet dotted line, respectively). The profile in a region from 400 bp upstream to 400 bp downstream of the polyA site is shown.

Nuclear depletion of Glc7 specifically affected Tyr1 dephosphorylation and not Ser2 dephosphorylation. ChIP occupancy of Ser2-phosphorylated Pol II normalized against total Pol II occupancy revealed no changes around the pA site (Figure 15d), whereas occupancy of Tyr1-phosphorylated Pol II increased (Figure 15c). Thus, the Glc7-dependent Ser2 dephosphorylation observed *in vitro* (Figure 13b) is likely non-specific and not used *in vivo*, consistent with the previously observed Ser2 dephosphorylation by Fcp1 *in vivo* (59, 61).

3. Pol II termination involves CTD Tyr1 dephosphorylation by CPF subunit Glc7

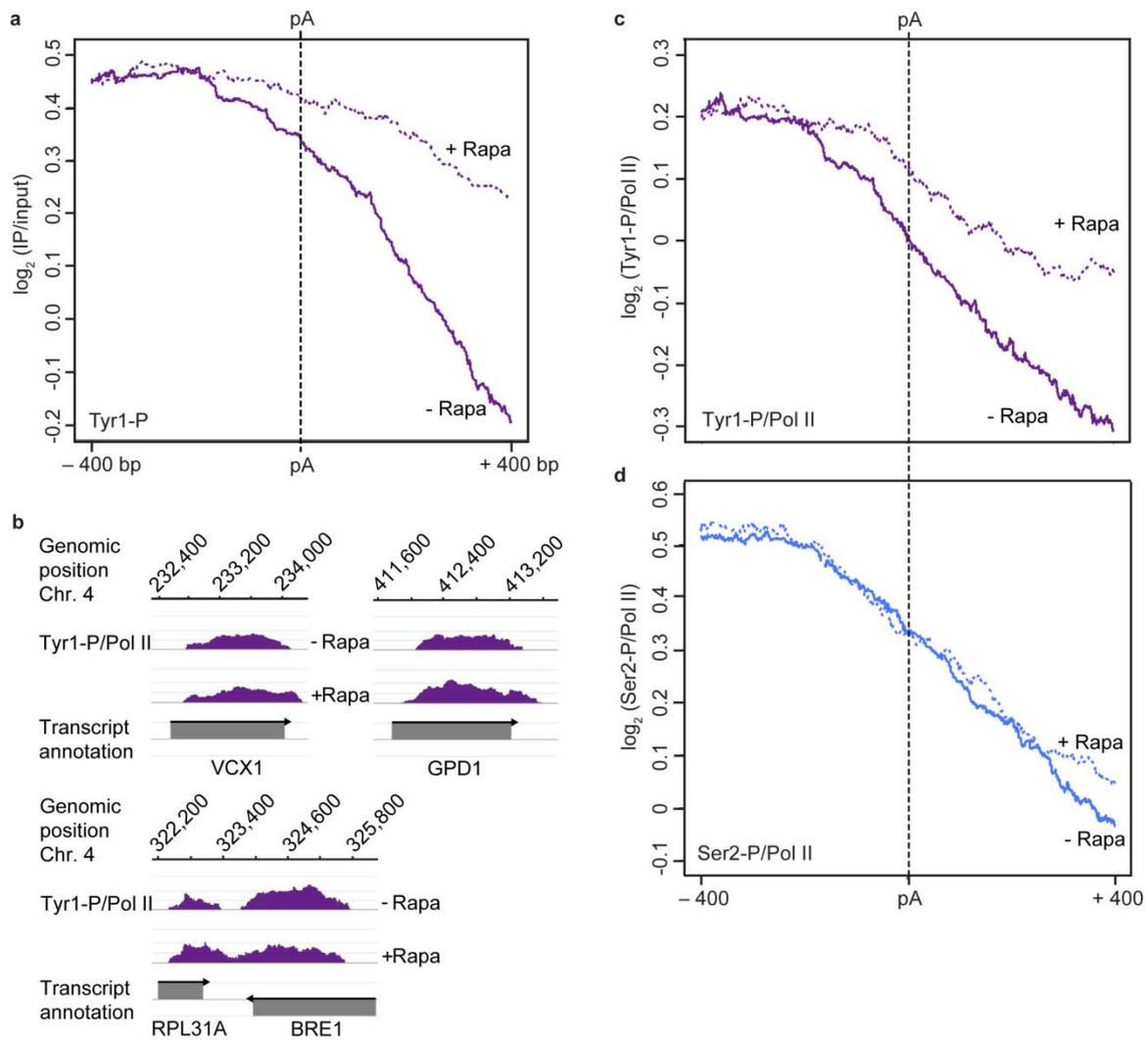


Figure 15: Glc7 is required for Tyr1 dephosphorylation *in vivo*.

(a) Metagenome analysis for genome-wide ChIP occupancy of Tyr1-phosphorylated Pol II around pA sites in the Glc7 anchor-away strain. Experiments were performed in biological duplicates and ChIP-chip signals averaged by taking the median signal ($\log_2(\text{IP}/\text{input})$) at each genomic position over a set of 619 representative genes (Methods). To better visualize changes in Tyr1 dephosphorylation around the pA site relative to the region upstream, the data were further normalized to have approximately equal occupancy levels upstream (around -400 bp) of the pA site for the ‘-’ and ‘+’ rapamycin treated profiles. Corresponding non-normalized data averaged over the entire gene length are shown in Figure 16. Occupancy of Tyr1-phosphorylated Pol II decreases sharply at the pA site (- Rapa, violet line), but Glc7 depletion largely abolishes this decrease (+ Rapa, violet dotted line). (b) Tyr1-P levels normalized to total Pol II (Figure 19a) on selected genes. Profiles represent normalized occupancies smoothed by a 150-nt window running median. Grey boxes indicate transcripts on the Watson (top) and Crick strands (bottom). (c, d) Depletion of Glc7 from the nucleus leads to a defect in Tyr1-P dephosphorylation but not in Ser2-P dephosphorylation. ChIP-chip occupancy profiling of Tyr1- (c) and Ser2- (d) phosphorylated Pol II over 619 genes aligned at the pA site (dashed line) and normalized against the corresponding Rpb3 profile without and with rapamycin (solid and dotted line, respectively). The profiles in a region from 400 bp upstream to 400 bp downstream of the polyA site are shown.

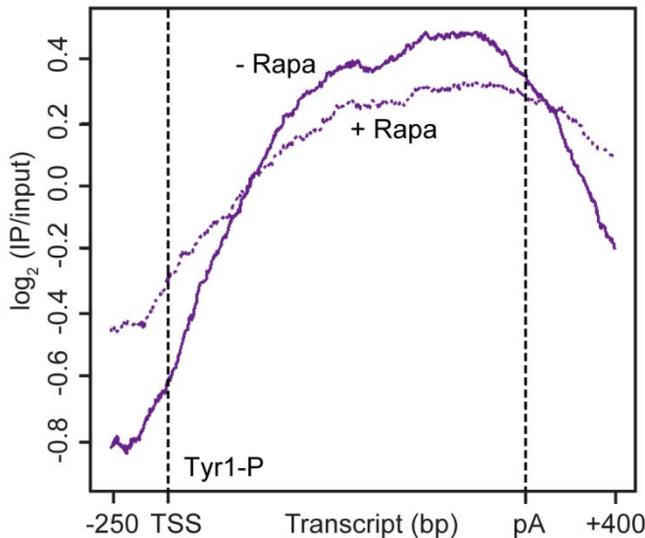


Figure 16: Depletion of Glc7 from the nucleus leads to a defect in Tyr1-P dephosphorylation.

ChIP-chip occupancy profiling of Tyr1-phosphorylated Pol II over 619 genes without and with rapamycin (-/+ Rapa, solid and dotted line, respectively) is shown (also see Figure 15a). Profiles in a region from 250 bp upstream of the TSS to 400 bp downstream of the pA site are shown. The higher level of Tyr1-P in the rapamycin treated sample around the TSS is probably caused by newly initiating Pol II that still contains remaining Tyr1-P from the previous round of transcription or from Pol II readthrough from upstream genes after failure to terminate.

To check whether the role of Glc7 in the nucleus is dependent on its catalytic activity, we transformed rescue plasmids expressing either wild-type or catalytically inactive mutants D94A or H247N of Glc7 into the Glc7 anchor-away strain. These mutations in the catalytic site are predicted not to disrupt the structure of Glc7 (209-211). Only wild-type Glc7 restored growth on rapamycin (Figure 18a), demonstrating that the catalytic activity of Glc7 is required for its nuclear function.

Although the mammalian homolog of Glc7, protein phosphatase 1 (PP1), acts primarily on phosphorylated serine and threonine residues (207, 211), the recombinant enzyme is active on phosphorylated tyrosine residues (212). Tyrosine dephosphorylation can also be performed by the related *Arabidopsis* PPP-Kelch phosphatase BSU1 (213, 214) and the serine/threonine protein phosphatase 2A (PP2A) if activated by the correct binding partner (215). Our results for Glc7 demonstrate that a PP1 homolog can act on a natural phosphorylated tyrosine substrate when present in the context of the correct multiprotein complex - in this case, CPF.

3. Pol II termination involves CTD Tyr1 dephosphorylation by CPF subunit Glc7

These results are consistent with a previous report that a point mutation in the CPF subunit Ref2 leads to Glc7 dissociation but has no impact on the integrity of the remaining complex (206). Taken together, these results indicate that Glc7 or Ssu72 can be depleted from CPF without destabilizing the complex, and that Glc7 catalyzes Tyr1 dephosphorylation, whereas Ssu72 does not.

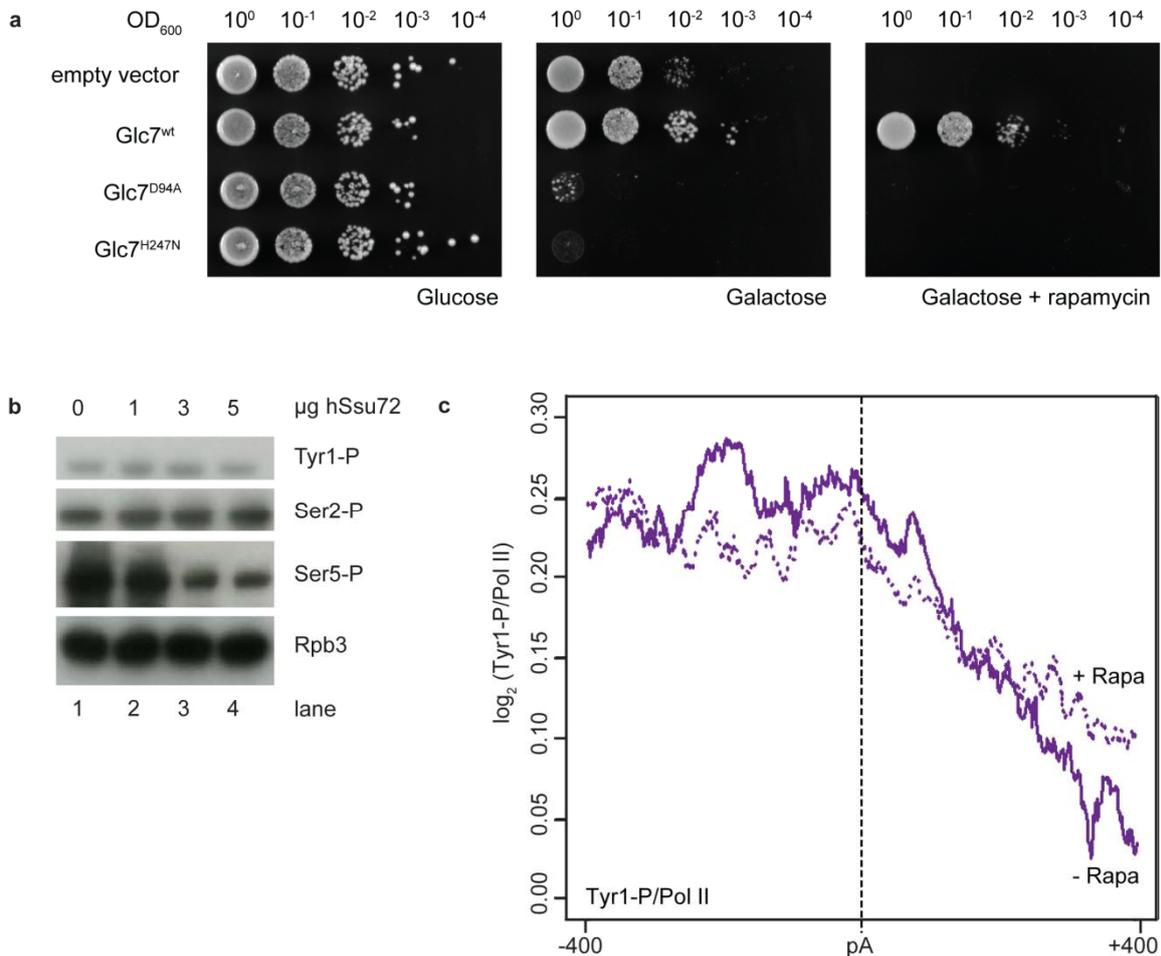


Figure 18: The catalytic activity of Glc7 in the nucleus is essential for cell viability and Ssu72 does not change Tyr1 phosphorylation *in vitro* and *in vivo*.

(a) Serial dilutions of the Glc7-FRB anchor-away strain transformed with either empty p415-GAL-S vector or the same vector with galactose-inducible wild-type (wt), D94A or H247N catalytic mutants of Glc7 were plated on -leucine plates containing glucose (left panel), galactose (middle panel) or galactose with additional rapamycin (right panel). Only wild-type Glc7 can recover wild-type growth upon depletion of Glc7-FRB from the nucleus. (b) Ssu72 does not dephosphorylate Tyr1-P *in vitro*: human Ssu72 (hSsu72) dephosphorylates Pol II at CTD residue Ser5 but not at residues Tyr1 and Ser2 (lanes 2-4) as shown by Western blotting with antibodies against Pol II subunit Rpb3 and Tyr1-, Ser2- and Ser5-phosphorylated CTD (3D12, 3E10 and 3E8, respectively). A 7- to 37-fold excess of hSsu72, *i.e.* a molar ratio of 0.27 to 1.42 of hSsu72 to each Pol II CTD heptad, was incubated with Pol II for 2 h. (c) Depletion of Ssu72 from the nucleus does not change Tyr1-P levels *in vivo*: ChIP-chip occupancy profiling of Tyr1-phosphorylated Pol II over 619 genes aligned at the pA site (dashed line) and normalized against the corresponding Rpb3 profile without and with rapamycin (-/+ Rapa, solid and dotted line, respectively). Profiles in a region from 400 bp upstream to 400 bp downstream of the polyA site are shown.

3.4. Glc7 is involved in transcription termination

Additional ChIP data indicated that Glc7 is required for transcription termination. The ChIP signal for total Pol II (Rpb3 subunit) normally decreases about 200 bp downstream of the pA site, due to transcription termination and Pol II release from DNA (Figure 19a, black line). Upon nuclear depletion of Glc7 however, strong Pol II ChIP signals remained further downstream (Figure 19a, dotted black line). These data suggest that Glc7 acts globally and is required for effective CTD Tyr1 dephosphorylation and normal transcription termination *in vivo*.

To investigate whether Tyr1 dephosphorylation is required for normal termination factor recruitment, we determined Pcf11 and Rtt103 ChIP-chip occupancy after Glc7 nuclear depletion (Figure 19b). These experiments revealed a strong decrease in termination factor occupancy. Both Pcf11 and Rtt103 showed much lower occupancies downstream of the pA site after rapamycin addition (Figure 19b). These data indicate that the defect in Pol II termination (Figure 19a) results from a defect in termination factor recruitment.

3.5. Discussion

Whereas previous work showed that the Pol II CTD is phosphorylated at Tyr1, and that Tyr1 phosphorylation impairs binding of transcription termination factors (44), it remained unknown whether Tyr1 dephosphorylation is required for termination *in vivo* and which enzyme dephosphorylates Tyr1. Here we establish the CPF subunit Glc7 as the CTD Tyr1 phosphatase, and show that Tyr1 dephosphorylation is required for normal recruitment of termination factors and transcription termination. These results support the previously proposed “extended CTD code” for the coordination of factor recruitment during the transcription cycle (44) and indicate a crucial role for Tyr1 dephosphorylation in the elongation-termination transition.

The data presented here and previously (44) thus lead to the following model for the elongation-termination transition of Pol II at the 3'-ends of protein-coding genes (Figure 19c). Elongating Pol II is phosphorylated mainly at Tyr1 (Y1P) and Ser2 (S2P), and this facilitates elongation factor binding. Tyr1 phosphorylation impairs premature recruitment of termination factors. When Pol II reaches the pA site, the Glc7 subunit of CPF dephosphorylates Tyr1, whereas Ser2 phosphorylation levels remain high. This allows for the binding of Pcf11 and Rtt103, termination factors that are not part of CPF and show peak occupancy ~100 bp downstream of the pA site (44). Further downstream, transcription terminates and Pol II is released from genes. When Glc7 is depleted from the nucleus, Tyr1 phosphorylation levels remain high downstream of the pA site, impairing termination and Pol II release.

3. Pol II termination involves CTD Tyr1 dephosphorylation by CPF subunit Glc7

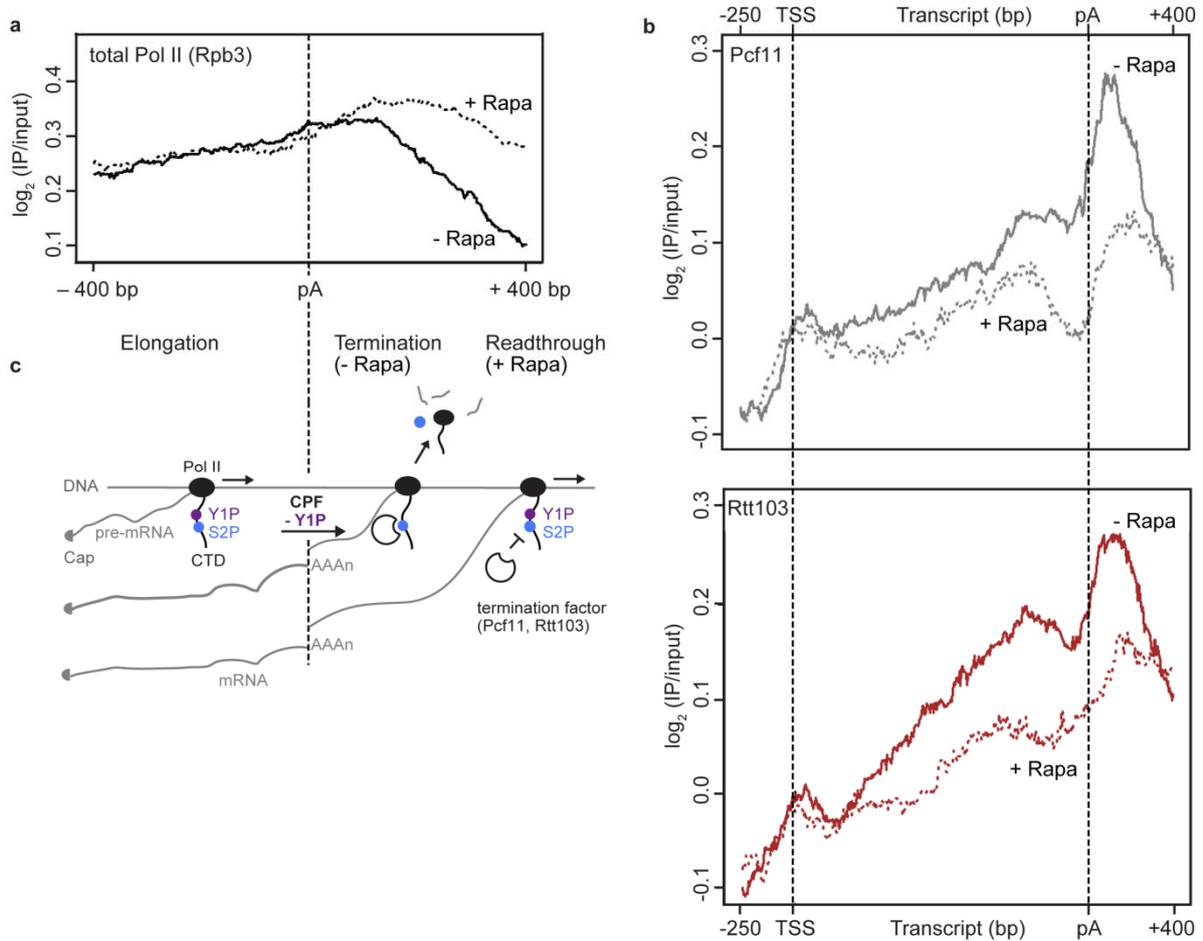


Figure 19: Tyr1 dephosphorylation by Glc7 is required for full termination factor recruitment and transcription termination *in vivo*.

(a) Metagenome analysis for genome-wide ChIP occupancy of total Pol II (subunit Rpb3) around pA sites in the Glc7 anchor-away strain. Experiments were performed in biological duplicates and ChIP-chip signals averaged by taking the median signal ($\log_2(\text{IP}/\text{input})$) at each genomic position over a set of 619 representative genes (Methods). We further normalized the data to have approximately equal occupancy levels upstream (around -400 bp) of the pA site for the '-' and '+' rapamycin treated profiles (see also legend Figure 15a). Occupancy of Pol II (Rpb3) decreases downstream of the pA site, indicating transcription termination (- Rapa, black line). Nuclear depletion of Glc7 impairs termination (+Rapa, black dotted line). (b) ChIP-chip occupancy profiling of TAP-tagged Pcf11 (top) and Rtt103 (bottom) in the Glc7 anchor-away strain. Experiments were performed in biological duplicates and ChIP-chip signals averaged by taking the median signal ($\log_2(\text{IP}/\text{input})$) at each genomic position over a set of representative medium length genes (1238 ± 300 bp, $n = 339$, see Methods). Occupancies of termination factors peak ~100 bp downstream of the pA site (- Rapa, grey and red lines), but Glc7 depletion largely decreases the peak signals (+ Rapa, dotted grey and red lines). (c) Model for the Pol II elongation-termination transition. DNA and RNA are depicted as grey lines, Pol II and its CTD are in black; Ser2-P is blue and Tyr1-P is purple. For details see within text.

In addition, we observed an influence of Glc7 depletion on Tyr1 phosphorylation levels near the TSS (Figure 16). This may be explained by initiation with polymerases that remained partially phosphorylated on Tyr1 residues, by termination defects at upstream genes (153), or by a possible role of Glc7 in transcription attenuation, which would be consistent with the role of Glc7 in Nrd1-dependent termination of snoRNA genes (206).

Our results also provide evidence that pA-dependent 3'-pre-mRNA processing is coupled to transcription termination via CPF-triggered Pol II dephosphorylation. A link between the pA site and transcription termination was established over 25 years ago (130). In the “anti-terminator model”, transcription of the pA site triggers a change in the Pol II machinery that allows for termination (14, 15, 128). In the “torpedo model”, pA-dependent RNA cleavage results in a new RNA 5'-end that is recognized by the Rat1/Rai1/Rtt103 exonuclease complex, which degrades nascent RNA and triggers termination (14, 15, 128). Our data are consistent with a combination of both models; Tyr1 phosphorylation would serve as an anti-terminator that is removed at the pA site by Glc7, allowing for recruitment of termination factors, including the torpedo nuclease complex that contains Rtt103.

4. The Spt5 C-terminal region recruits yeast 3' RNA cleavage factor I

All results presented in chapter 4 were obtained in collaboration with Andreas Mayer and are published in (216). For detailed author contributions see page IV.

Here we show that the Spt5 CTR is required for normal recruitment of CFI to the 3'-ends of yeast genes *in vivo* and interacts with CFI *in vitro*. High-resolution genome-wide occupancy profiling of CFI subunits reveals peak occupancy levels around 100 bp downstream of the pA site. Our results indicate that the Spt5 CTR cooperates with nascent RNA and the Ser2-phosphorylated form of the Pol II CTD to recruit CFI to the 3'-ends of genes.

4.1. Investigation of elongation factor recruitment by Spt5 CTR

To investigate whether the function of the yeast Spt5 CTR in recruiting Pol II-associated factors extends to elongation factors other than Paf1 (86, 87), we carried out ChIP analysis in strains lacking the Spt5 CTR. We generated a yeast strain with a CTR deletion (Methods). As reported previously, CTR deletion led to 6-AU sensitivity and a slow-growth phenotype at 16°C (Figure 20a) (86, 87). We also observed a slight growth defect at 30°C (Figure 20b), but not at 37°C (not shown). In contrast to observations in *S. pombe* (123), the morphology of *S. cerevisiae* was not altered upon CTR deletion (not shown). Quantitative Western blot analysis revealed approximately 1.7-fold higher Spt5 protein levels in cells lacking the CTR (Figure 20c).

For ChIP analysis, we fused a tandem affinity purification (TAP) tag to the C-terminus of elongation factors in the CTR deletion background. The occupancy levels of the elongation factors were determined by ChIP at different positions of genes *ADHI*, *ILV5*, *PDC1* and *TEF1*. We chose these genes for several reasons. First, these genes encode housekeeping proteins, are highly expressed (217), and are heavily occupied by Pol II in the mid-log phase of yeast growth (43). Second, their DNA elements, including the TSS and the pA site, are well characterized (10, 11, 105). Third, the transcription unit is long enough to distinguish between different binding levels at distinct positions of the gene.

4. The Spt5 C-terminal region recruits yeast 3' RNA cleavage factor I

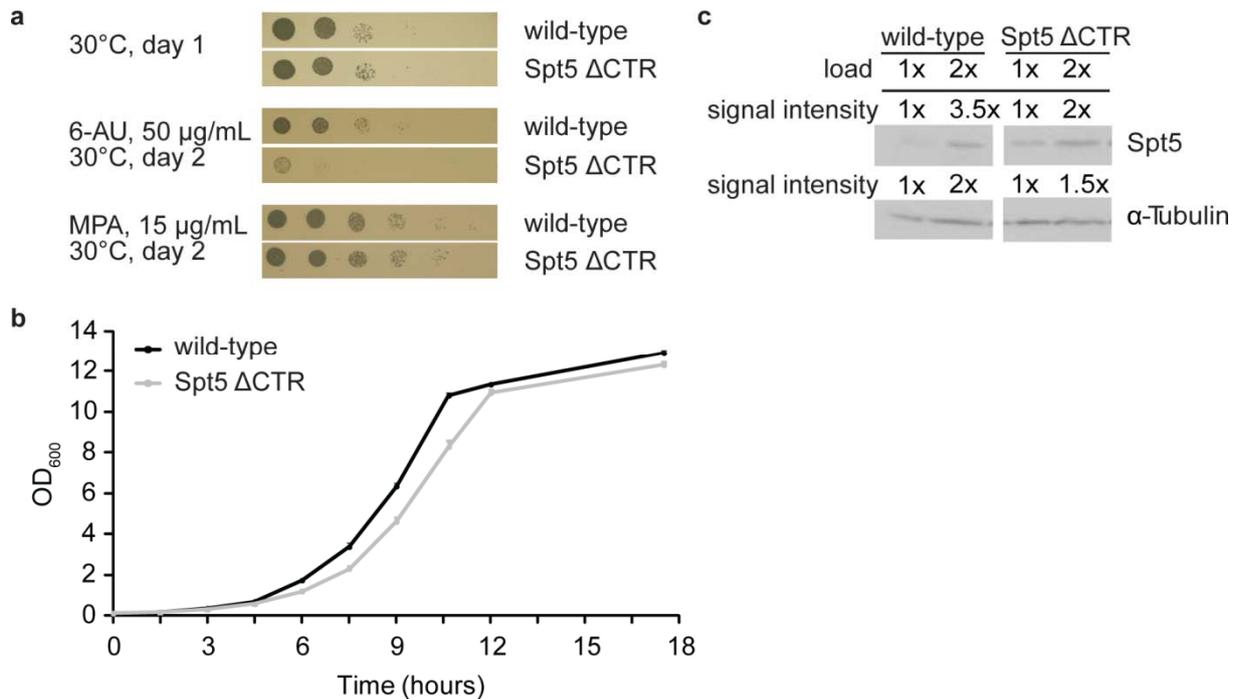


Figure 20: Spt5 CTR deletion leads to a growth defect in the presence of 6-AU and to a slight increase in Spt5 protein levels.

(a) The Spt5 ΔCTR mutant shows a strong growth defect in the presence of 50 μg/mL 6-AU as compared to wild-type. No effects are observed at 30°C or in the presence of 15 μg/mL mycophenolic acid (MPA) on solid medium. (b) The Spt5 ΔCTR mutant grown in liquid YPD medium (gray curve) shows a slight growth defect at 30°C as compared to the wild-type (black curve). Standard deviations of three independent measurements are indicated for each data point. (c) Spt5 protein levels in the ΔCTR mutant are upregulated 1.7-fold compared to wild-type. Quantitative Western blotting was performed with antibodies against Spt5 and α-Tubulin. Different amounts of protein were loaded and quantification of the intensities was performed. Normalization of the Spt5/α-Tubulin ratio obtained for Spt5 ΔCTR cells against the corresponding ratio of wild-type cells revealed that Spt5 protein levels were slightly elevated (1.7 fold) in cells lacking the CTR of Spt5.

We performed ChIP analyses for eight Pol II elongation factors that belong to the three different groups described recently (43): (i) Spt4 and Spt6 (group 1), (ii) Elf1 and Spn1 (group 2), (iii) Bur1, Ctk1, Paf1 and Spt16 (group 3). The results are shown in Figure 22c (Paf1) and Figure 21. These data revealed strong factor binding at the *ADHI* gene. A severe decrease in Paf1 occupancy to about 20% was detected at the *ADHI* gene (Figure 22c), consistent with previous reports (86, 87) and providing a positive control. The difference in Paf1 occupancy was not due to a difference in Pol II occupancy, which was unaffected by CTR deletion (Figure 22c). However, the other representative elongation factors tested did not show significant differences in their gene occupancies (Figure 21), showing that CTR deletions specifically reduce gene occupancy of Paf1.

4. The Spt5 C-terminal region recruits yeast 3' RNA cleavage factor I

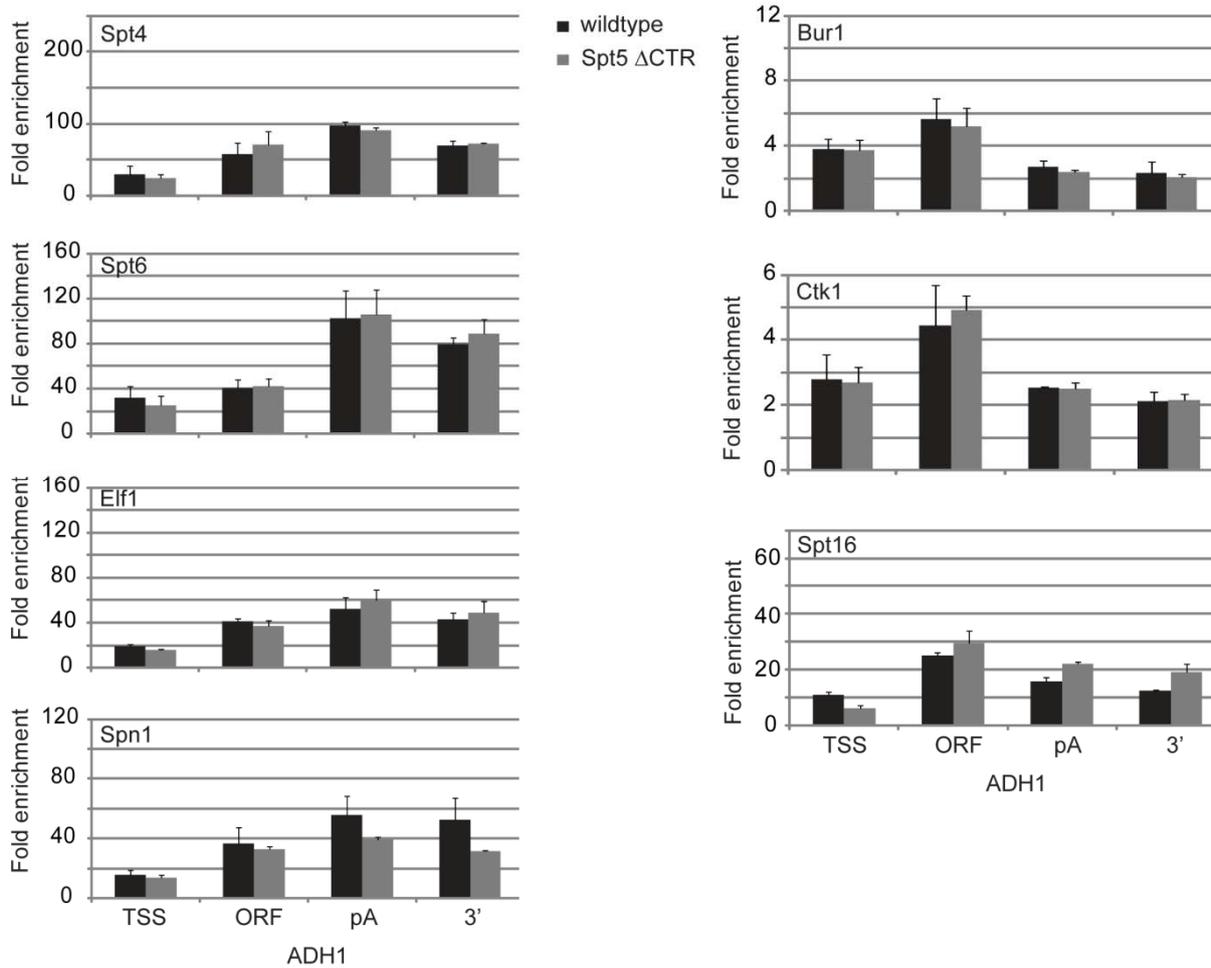


Figure 21: Apart from Paf1 (Figure 22c) no other Pol II transcription elongation factors are affected by the deletion of the *Spt5* CTR.

Although *Spn1* occupancy seems to be slightly reduced towards the 3'-end of the *ADH1* gene in *Spt5* Δ CTR cells as compared to wild-type cells, the differences are within the standard deviation and therefore are not significant. ChIP occupancies are indicated for wild-type and *Spt5* Δ CTR cells as black and gray bars, respectively. The standard deviations refer to at least two independent ChIP experiments.

4.2. *Spt5* CTR is required for recruitment of CFI *in vivo*

Since *Spt5* colocalizes with 3'-end processing factors (43, 105, 107) and copurifies with these factors (115), we tested whether it plays a role in the recruitment of 3'-end processing and transcription termination factors. ChIP analysis revealed that all CFIA subunits *Clp1*, *Pcf11*, *Rna14*, and *Rna15*, showed high occupancy at the 3'-end of protein-coding genes and the pA site (Figure 22a and Figure 23). In *Spt5* Δ CTR cells, occupancy of *Clp1*, *Rna14*, and *Rna15* was reduced more than 50% at the *ADH1* gene, and also markedly lower at other genes tested (Figure 23). However, we observed no difference in the occupancy of *Pcf11* between wild-type yeast and *Spt5* Δ CTR cells (Figure 22a).

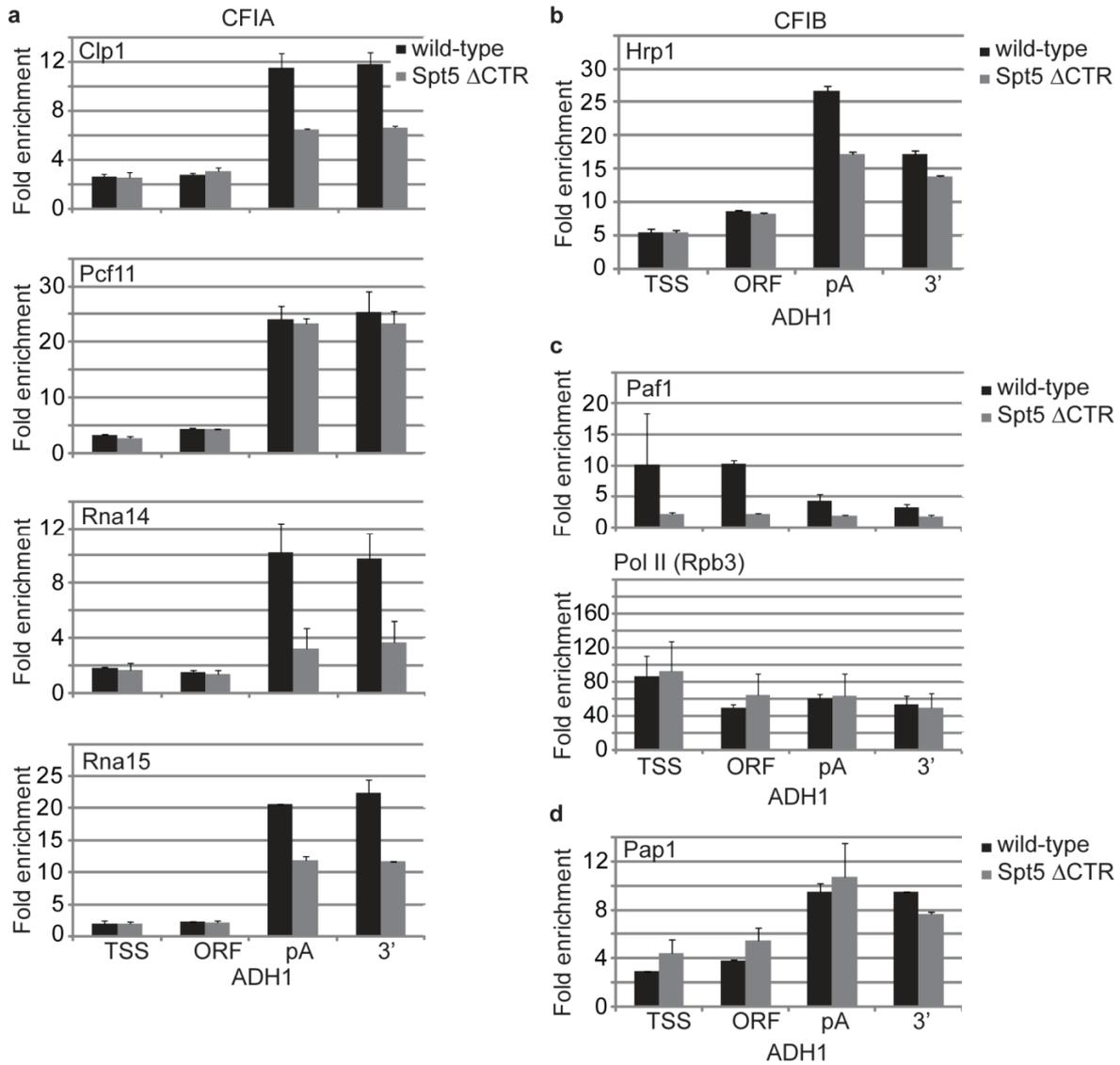


Figure 22: ChIP analysis reveals that CFI occupancy is reduced in Spt5 Δ CTR cells.

This is true for (a) CFIA subunits, except for Pcf11, as well as for (b) CFIB/Hrp1. (c) Whereas ChIP occupancy of Paf1 is strongly reduced in Spt5 Δ CTR cells, Pol II levels are not affected. (d) Pap1 occupancy is not changed in Spt5 Δ CTR cells. The fold enrichments at the *ADH1* gene over a nontranscribed region that is located near the centromere of chromosome V are given for the TSS, the ORF region, the pA site, and for the region 3' of the pA site. The color code is as in Figure 21. The standard deviations refer to at least two independent ChIP experiments.

Since CFIA is associated with CFIB/Hrp1, we investigated whether Hrp1 occupancy was affected by CTR deletion. Previous ChIP analyses have shown that Hrp1 crosslinks throughout the coding regions until the 3'-end of genes (50, 105). Our ChIP analysis revealed that Hrp1 shows strongest occupancy near the pA site, although it is recruited earlier than CFIA subunits (Figure 22b). Similar to CFIA subunits, Hrp1 binding was markedly reduced in Spt5 Δ CTR cells (Figure 22b). However, no occupancy difference could be observed for the poly(A) polymerase Pap1 (Figure 22d), which is also required for 3'-end processing (137,

218). Taken together, the Spt5 CTR plays a crucial role in recruitment of CFI, but not of Pap1, to the 3'-end of protein-coding genes.

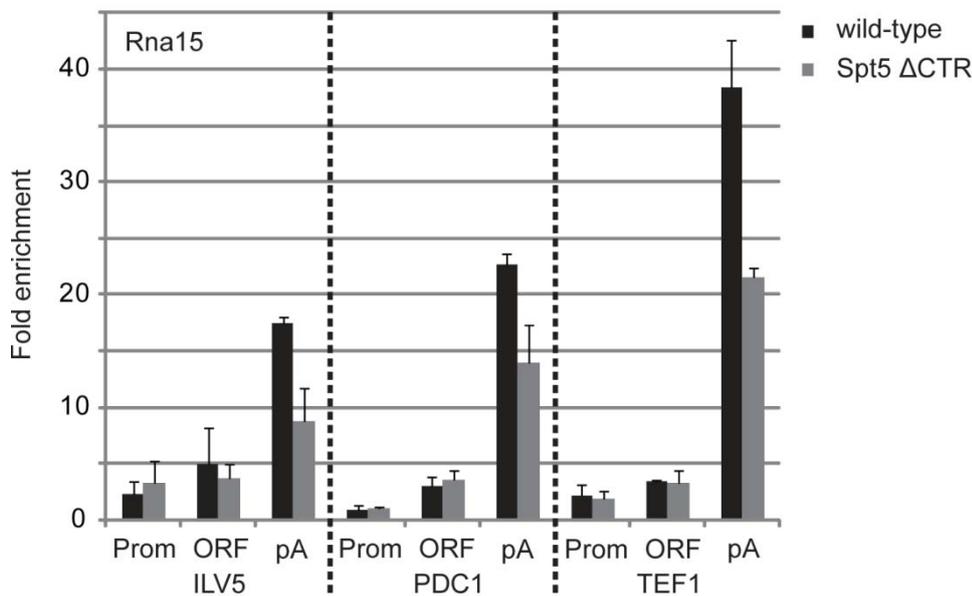


Figure 23: Rna15 ChIP occupancy levels at *ILV5*, *PDC1* and *TEF1* genes are also reduced in Spt5 Δ CTR cells.

The Rna15 occupancy levels are indicated for the promoter (prom) and ORF regions as well as for the pA site of all three genes in wild-type (black bars) and Spt5 Δ CTR cells (gray bars). The color code is as in Figure 21. The standard deviations refer to at least two independent ChIP experiments.

4.3. Spt5 CTR interacts with CFI *in vitro*

Since CFI recruitment to genes was impaired upon CTR deletion, we asked whether the Spt5 CTR interacts with CFI *in vitro*. We performed pulldown experiments with a recombinantly expressed glutathione S-transferase (GST)-tagged version of the Spt5 CTR and yeast cell lysates prepared from strains that expressed TAP-tagged versions of CFI subunits. Western blot analysis of the eluates with antibodies directed against the GST tag and the TAP tag of the respective CFI subunit revealed a coprecipitation of the GST-tagged Spt5 CTR and the CFI subunits Rna14, Rna15, and Hrp1, which was not detected with GST alone (Figure 24). A coprecipitation could not be observed for Pcf11 (not shown), in agreement with the ChIP data that showed that Pcf11 occupancy is not altered in Spt5 Δ CTR cells (Figure 22a). For Clp1, the input signal in the Western Blot analysis was much lower compared to other CFI subunits, such that a possible interaction between the Spt5 CTR and Clp1 cannot be ruled out. Taken together, these pulldown experiments revealed a previously unobserved interaction between the Spt5 CTR and CFI *in vitro*. Since the interaction was detected with the use of a lysate that naturally contains many non-specific competitor proteins, and since it was not observed with the GST tag alone, it must be regarded as being highly specific. It is however possible that the interaction is mediated by other proteins in the extract and thus indirect.

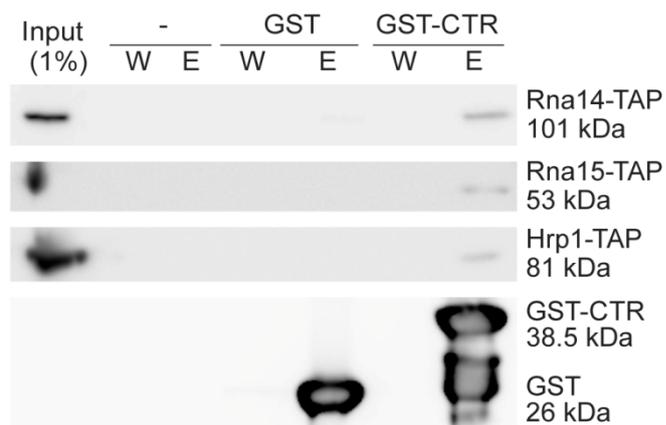


Figure 24: The Spt5 CTR interacts with CFI subunits *in vitro*.

GST pull-down experiments were performed with a GST-Spt5 CTR fusion protein (GST-CTR), with GST alone (GST) and without protein (-), serving as negative control. Western blotting was performed for the last washing fractions (W), the combined elution fractions (E) of the respective sample, and for 1% of the Rna14-TAP (first panel), Rna15-TAP (second panel), and Hrp1-TAP (third panel) yeast cell lysates (Input) with antibodies against the TAP tag of the corresponding CFI subunit and the GST tag (for details see Methods). The signals obtained for GST and GST-CTR were similar and are exemplarily shown for the pull-down experiments with Rna14-TAP (fourth panel).

4.4. RNA contributes to CFI recruitment

Since deletion of the Spt5 CTR led to a marked reduction in the occupancy of CFI subunits, but not to a complete loss, we asked which factors contribute to the residual binding of CFI. Since Rna15 and Hrp1 contain RNA recognition motifs (RRMs) that are known to bind RNA sequences *in vitro* (139, 219-221), we reasoned that nascent RNA may contribute to CFI recruitment. To address this, we performed an RNase-ChIP assay (222). In this assay, RNA is digested before the immunoprecipitation step, leading to a drop in factor occupancy if factor recruitment involves RNA.

We performed RNase-ChIP for Rna14 and Rna15 (Figure 25). First, we observed a decrease in Rna15 occupancy after RNase treatment, indicating an important role of RNA in Rna15 recruitment, both in wild-type and Spt5 Δ CTR cells. Second, Rna15 binding most strongly depended on RNA around the pA site of the *ADHI* gene. Third, the strongest reduction in Rna15 occupancy was observed when both the Spt5 CTR was deleted and when RNA was removed by RNase treatment. The additional decrease of the Rna15 occupancy level was highly reproducible and could be observed in all four independent biological replicates. A similar RNA dependence could be observed for Rna14. These results indicate that RNA contributes to CFI recruitment *in vivo* and that this can explain residual recruitment of CFI in cells lacking the Spt5 CTR.

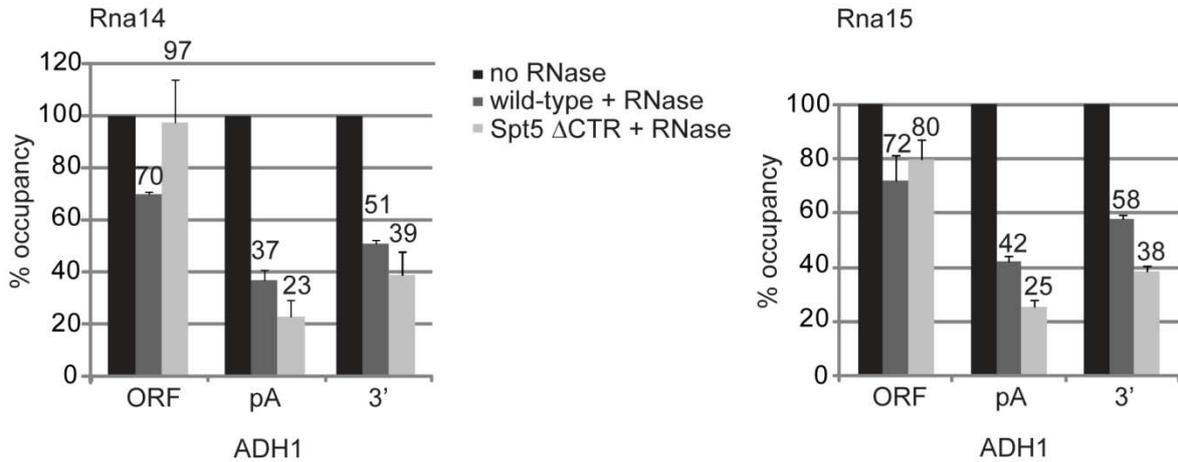


Figure 25: RNase-ChIP assays reveal that RNA contributes to CFI recruitment in Spt5 Δ CTR cells.

RNA-dependent binding is given for the ORF region, the pA site and a region 3' of the pA site of the *ADH1* gene. The ChIP occupancy signal without RNase treatment (black) was set to 100%. The relative ChIP signals of Rna14 and Rna15 for wild-type (dark gray) and Spt5 Δ CTR cells (light gray) after RNase treatment are indicated. The corresponding percentages are given above the bars. The standard deviations were calculated from four independent experiments.

4.5. CFI colocalizes with the Ser2-phosphorylated CTD downstream of the pA site

It was shown for mammalian cells that homologs of the yeast CFI complex are already recruited at the promoter region of genes (223, 224). One study in yeast showed that some RNA 3'-end processing factors, including Rna14 and Rna15, also crosslinked to the promoter and the early coding region of the *PMAI* gene, but not at other genes tested (105). However, our ChIP analysis (Figure 22 and Figure 23) and published data in yeast (105, 145) suggested that CFI subunits crosslinked near the pA site at the 3'-end of genes.

To investigate the preferred location of CFI subunits on a genome-wide level, we performed ChIP-chip analysis for Rna14 and Rna15. This revealed CFI recruitment at all protein-coding genes that are occupied by Pol II and its elongation factors (43). The ChIP-chip profiles showed sharp occupancy peaks for Rna14 and Rna15 105 bp and 108 bp downstream of the pA site, respectively (Figure 26a). We also detected weak Rna14 and Rna15 binding over the transcribed region, with an increase towards the 3'-end. These profiles were independent of gene length (Figure 26a and Figure 27). Comparison with previous profiles (Figure 26a) (43) revealed that peak occupancies of Rna14 and Rna15 coincided with Pcf11 occupancy, which peaked 52 bp downstream of the pA site (Figure 26a). CFI subunit peak occupancies further occurred in a region where the occupancies for Spt5 and the Ser2-phosphorylated Pol II were high (Figure 26b). The sharp occupancy drop of CFI subunits coincided with the drop for the ChIP signal of the Ser2-phosphorylated CTD, consistent with its role in CFI recruitment.

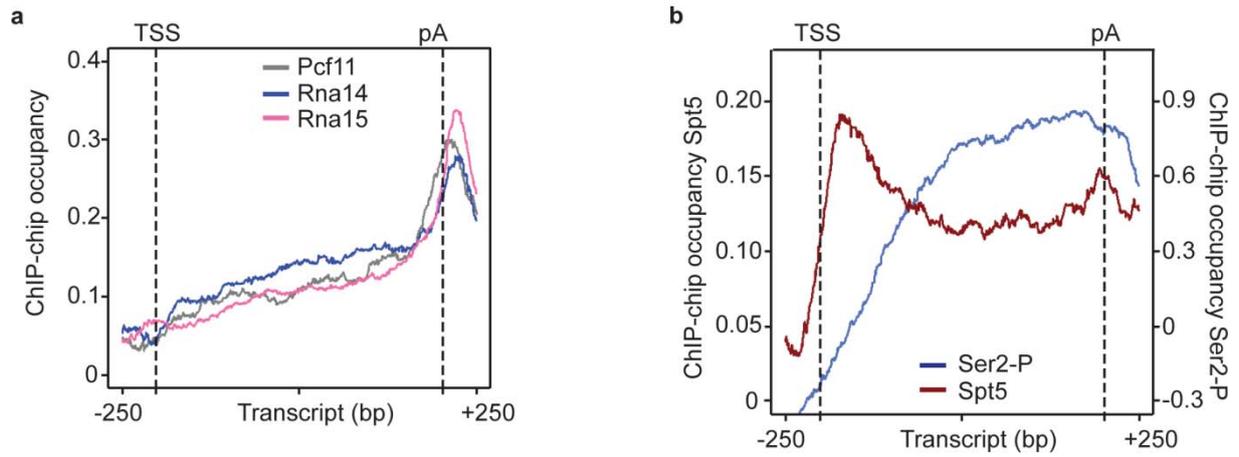


Figure 26: Genome-wide ChIP-chip occupancy profiling of CFI subunits, Ser2-P Pol II and Spt5 in yeast.

(a) Gene-averaged profiles for the long gene length class ($2,350 \pm 750$ bp, 299 genes, see Methods) for Pcf11 (43), Rna14, and Rna15. Profiles of other length classes are generally similar (Figure 27). Dashed black lines indicate the TSS and pA site. (b) Gene-averaged profiles as in (a) for the transcription elongation factor Spt5 (43) and the Ser2-phosphorylated CTD form of Pol II (43). Occupancies and signal intensities are given for Spt5 and the Ser2-phosphorylated form of Pol II on the left and right y-axes, respectively.

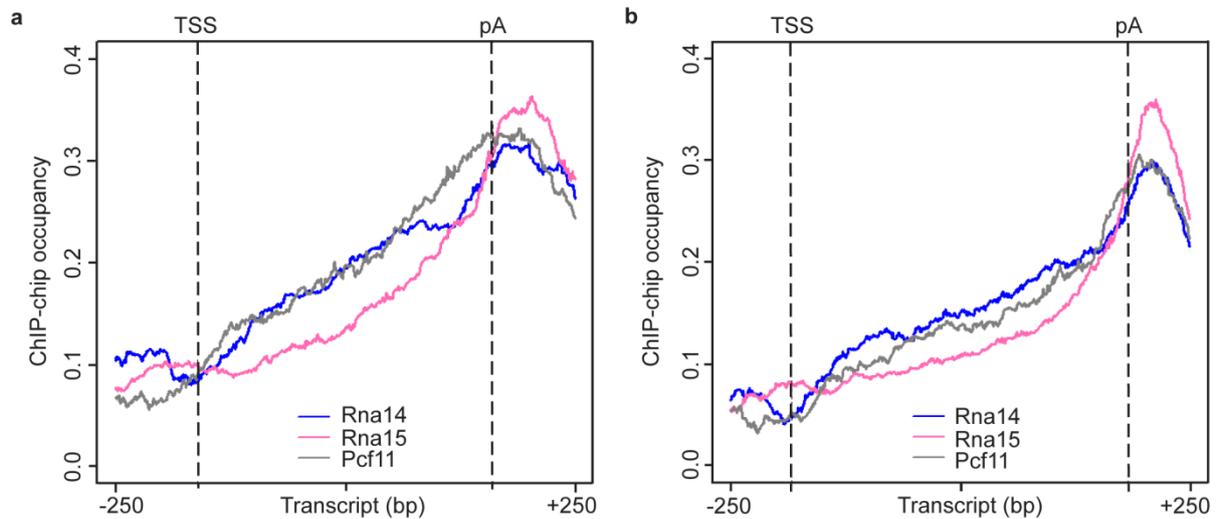


Figure 27: Genome-wide ChIP-chip occupancy profiling of CFI subunits in yeast for short and medium length genes.

Gene-averaged profiles (a) for the short gene length class (512-937 bp, 266 genes, see Methods), (b) for the medium length class (938-1537 bp, 339 genes) for Pcf11 (43), Rna14, and Rna15. Dashed black lines indicate the TSS and pA site.

4.6. CTR deletion does not impair termination

Previous studies showed that mutations in Pcf11, Rna14, and Rna15 can lead to defects in transcript cleavage and readthrough transcription beyond the termination site (225-227). We therefore investigated whether the reduced level of CFI recruitment observed in Spt5 Δ CTR cells leads to Pol II readthrough transcription at the *ACT1*, *PMA1* and *RNA14* genes. To detect transcriptional readthrough, we chose a PCR-based method with a gene-specific forward primer and different reverse primers positioned downstream of the normal transcript termination site, as described previously (146) (Figure 28a). In this assay, a prolonged transcript resulting from readthrough transcription would be detected by the generation of PCR products with a reverse primer that is located downstream of the termination site. As shown in Figure 28b, transcriptional readthrough could be detected when the function of Rna14 was impaired with the use of a temperature-sensitive yeast strain, serving as a positive control for this assay. However, no differences in the PCR products, and thus the length of the *PMA1* transcripts, were observed between wild-type and Δ CTR cells (Figure 28b). Similar results were obtained for the *ACT1* and *RNA14* genes (not shown). Thus, deletion of the Spt5 CTR does not result in a termination defect that would be detected by transcriptional readthrough.

Although we did not detect readthrough transcription at tested single genes, it may still occur at other genes. To investigate this, we measured high-resolution ChIP-chip occupancy profiles for the Pol II core subunit Rpb3 in wild-type and mutant yeast cells lacking the CTR. The high correlation between the Pol II profiles (Pearson correlation coefficient $R = 0.89$) and the high similarity of the gene-averaged profiles (Figure 28c) however indicated no difference in Pol II occupancy between wild-type and mutant cells. In addition, a difference profile calculated from Rpb3 occupancy in Δ CTR and wild-type cells did not reveal any clusters of altered occupancy. These results show that transcription readthrough does not occur in the Δ CTR strain.

4. The Spt5 C-terminal region recruits yeast 3' RNA cleavage factor I

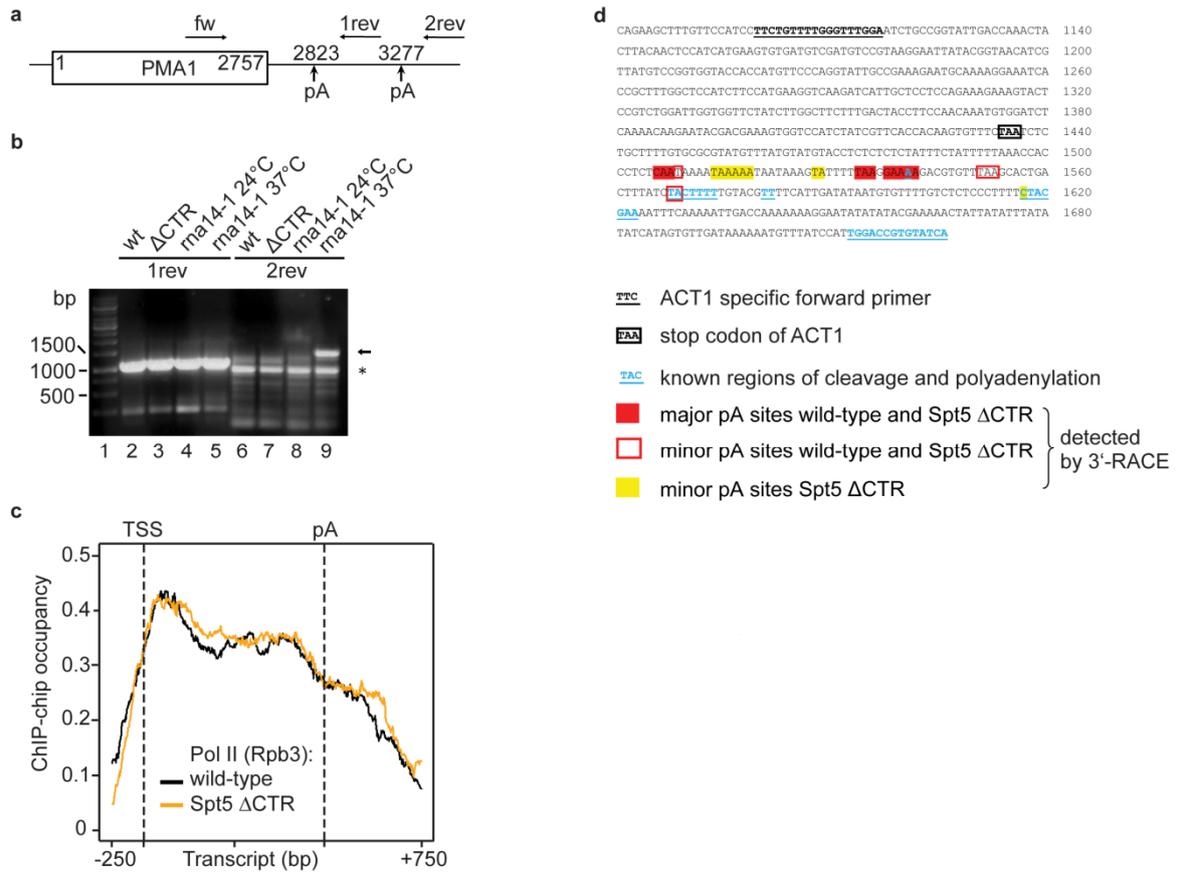


Figure 28: Spt5 CTR deletion provokes neither transcriptional readthrough of Pol II nor alternative pA site usage.

(a) Schematic representation of the yeast *PMA1* locus. The ORF region and the two pA sites according to (146) are indicated by a box and vertical arrows, respectively. The forward primer (fw) and two reverse primers (1rev and 2rev) that were used for Pol II readthrough detection are depicted as horizontal arrows. (b) Agarose gelelectrophoresis of the five PCR products as described in (a) for wild-type (wt), Spt5 Δ CTR cells (Δ CTR) and the *rna14-1* temperature sensitive strain grown at permissive (24°C) and restrictive temperature (37°C). The *rna14-1* mutant led to a readthrough transcript at 37°C (1.5 kb, black arrow), and served as a positive control. No differences in the length of the PCR products could be detected between wild-type and Spt5 Δ CTR cells. The unspecific PCR product also obtained with the second reverse primer (2rev) is marked by an asterisk. The height of the marker lanes in base pairs (bp) are shown in lane 1. (c) Gene-averaged occupancy profiles as in Figure 26 but for medium gene-length class ($1,238 \pm 300$ bp, 339 genes, see Methods) of Pol II (Rpb3) in wild-type and Spt5 Δ CTR cells. (d) The nucleotide sequence of the 3' region of yeast *ACT1* is shown. Key sequence elements are labeled. Blue sequences are sites predicted by (228). 3'-RACE revealed three major pA sites (red filled boxes) and three minor pA sites (red frames) that were equally used in wild-type and Spt5 Δ CTR cells. Additionally, 3'-RACE also led to the mapping of three rare pA sites that were exclusively used in Spt5 Δ CTR cells (yellow boxes). For details see Methods.

4.7. CTR deletion does not alter pA site usage

Since mutations of 3'-end processing factors were also shown to result in the usage of alternative pA sites (228), we investigated whether CTR deletion and the resulting reduction in CFI recruitment lead to alternative pA site usage. To detect a possible change in the usage of pA sites *in vivo*, we used rapid amplification of cDNA 3'-ends (3'-RACE), which allows the mapping of the 3'-ends of transcripts (229). We performed 3'-RACE for the *ACT1* and *PMAI* genes, which possess five and two pA sites, respectively (Figure 28d) (105, 228). The experiments revealed multiple pA sites for the *ACT1* gene, which map to a distinct region at the 3'-end of the gene, 1508-1617 bp from the TSS. We also quantified the usage of the different pA sites by sequencing 25 PCR products from both wild-type and Spt5 Δ CTR clones (Methods). For *ACT1*, these experiments revealed three major pA sites that were used in 75% of the cases, and six minor pA sites (Figure 28d). However, no significant difference in pA site usage was detected between wild-type and Spt5 Δ CTR cells (Figure 28d). Similar results were obtained for the *PMAI* gene (not shown). Thus, a reduced level of CFI recruitment in Spt5 Δ CTR cells does not lead to altered pA site usage *in vivo*.

4.8. Discussion

Two major transitions occur during the Pol II transcription cycle, the initiation-elongation transition at the 5'-ends of genes and the elongation-termination transition at the 3'-ends of genes, which is coupled to RNA processing. Whereas the first transition has been extensively studied (43, 50, 106, 163, 230-232), less is known about the second transition. Studies of the second transition revealed a role of the Ser2-phosphorylated Pol II CTD in the recruitment of 3'-end processing and termination factors (65, 145-147, 163). This transition also involves the Paf1 complex (85, 233, 234), elongation factor Spt6 (235), and the transcription regulator Sin1 (236).

Here we provide evidence for a role of the Spt5 CTR in the elongation-termination transition, in particular in the recruitment of the essential mRNA 3'-end processing factor CFI. ChIP analysis of wild-type and Spt5 Δ CTR in yeast cells detected a reduction in the occupancy of Paf1, as described previously (86, 87), but also of CFI subunits, indicating impaired CFI recruitment to the 3'-ends of genes *in vivo*. A pulldown assay additionally revealed an interaction between CFI subunits and the Spt5 CTR *in vitro*. This is consistent with a study showing that Rna14 can be copurified with Spt5 (115). RNase-ChIP experiments for CFI subunits Rna14 and Rna15 showed that RNA contributes to CFI recruitment *in vivo*. Genome-wide profiling by ChIP-chip revealed a sharp peak in Rna14 and Rna15 occupancy ~100 bp downstream of the pA site, which coincides with high occupancy of Spt5 and the Ser2-phosphorylated Pol II. These results show that the Spt5 CTR contributes to the recruitment of CFI to a defined region at the 3'-ends of yeast genes.

There is evidence that recruitment of RNA 3'-end processing factors also involves the Paf1 complex. Deletion of subunits of the Paf1 complex reduces the recruitment of Pcf11 (233) and interferes with Pol II binding of Cft1 (234), a component of the yeast CPF complex. Since Paf1 occupancy levels are markedly reduced in Spt5 Δ CTR cells, it may be argued that recruitment of CFI may occur via the Paf1 complex and that the observed reduction in CFI subunit occupancy may result from a loss of the Paf1 complex. However, several lines of evidence argue against this model and instead argue that CFI recruitment occurs via a direct interaction with the CTR. First, Pcf11 occupancy is not altered in Spt5 Δ CTR cells, despite the loss of the Paf1 complex (Figure 22a, c). Second, despite extensive interactomics studies, physical interactions between the Paf1 complex and CFI have never been observed. Third, the Paf1 complex clearly dissociates from the Pol II elongation complex upstream of the pA site (43, 105, 237), whereas CFI subunits are mainly recruited downstream of the pA site (Figure 26a) (43, 105, 237). These results argue for a Paf1 complex-independent mechanism of CFI recruitment in yeast.

We further showed that the Spt5 CTR is not required for normal pA site usage and transcription termination *in vivo*. This is consistent with the nonessential nature of the Spt5 CTR in yeast and may be due to residual CFI recruitment in Spt5 Δ CTR cells that likely results from binding of the CFI subunits to RNA and from the binding of Pcf11 to the Pol II CTD. Pcf11 contains an essential CID domain, which directly binds the Ser2-phosphorylated CTD (65, 131, 145) and may be responsible for normal recruitment of Pcf11 to genes upon CTR deletion. These observations indicate that CFI is recruited to a defined region downstream of the pA site by simultaneous interactions with the Spt5 CTR, nascent RNA, and the Pol II CTD (Figure 29). This model is consistent with the reported binding of Rna14 and Rna15 to the phosphorylated CTD (148) and with other published data. Rna15 can be crosslinked to RNA (139) and contains an RRM that binds GU-rich RNA (219, 220). Hrp1 has two RRMs that bind to AU-containing RNAs (221).

Finally, our results have implications for understanding the evolution of transcription-coupled events. Spt5 represents the only known RNA polymerase-associated factor that is conserved in all three domains of life (110); its bacterial homolog is called NusG (83). All Spt5 homologs contain two conserved domains, the NGN domain and a C-terminal KOW domain (113). Whereas the NGN domain binds to the polymerase clamp domain and closes the active center cleft, apparently to render transcription processive, the KOW domain extends from the polymerase surface toward exiting RNA (84, 101, 112) (Figure 29). In bacteria, the KOW domain interacts with the ribosome, thus coupling transcription to mRNA translation (238, 239). In eukaryotes, transcription and translation take place in different cellular compartments, and any coupling between these processes would likely occur via the mRNA that exits the nucleus (240, 241). Our data indicate that the CTR of Spt5 contributes to the coordination of transcription with RNA 3'-end processing, which in turn is coupled to mRNA export (242-244). Since the CTR occurs only in eukaryotic Spt5 homologs, it is likely that it

4. The Spt5 C-terminal region recruits yeast 3' RNA cleavage factor I

emerged during evolution to maintain coupling between transcription and translation after the spatial separation of these processes. Such coupling may be achieved by cotranscriptional Spt5-dependent loading of mRNA export factors onto the nascent RNA, before its maturation, nuclear export, and translation in the cytosol.

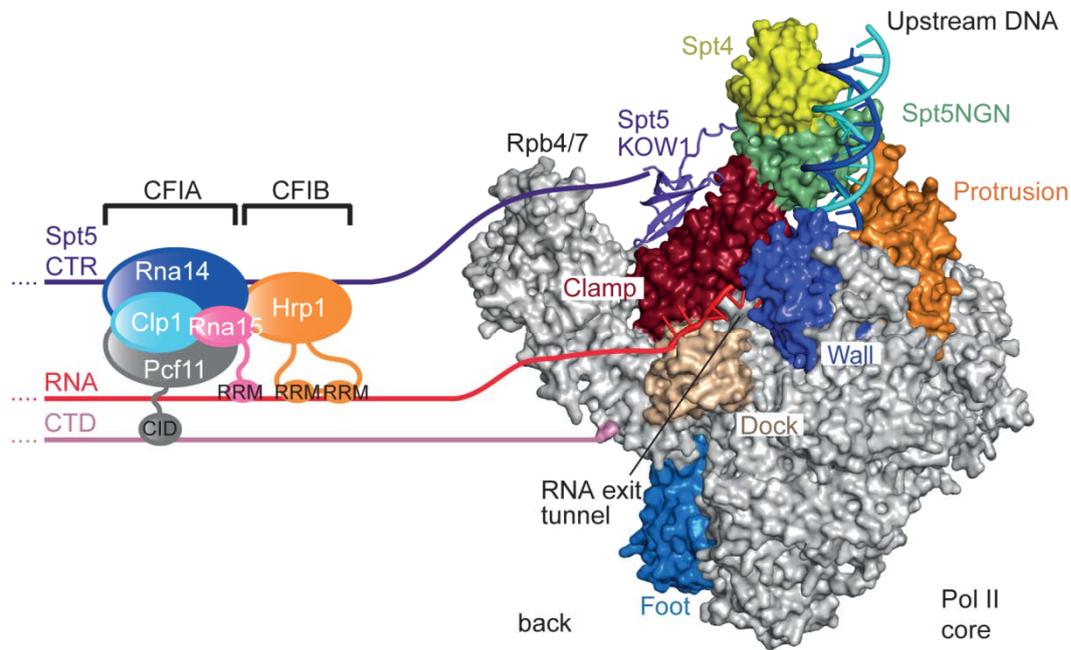


Figure 29: Model of CFI recruitment in yeast.

The complete yeast Pol II elongation complex with bound Spt4/5 is viewed from the back (84). Pol II and Spt4/5 are shown as molecular surfaces with key domains highlighted in color and labeled. Exiting RNA, the C-terminal KOW domains, the CTR of Spt5, and the Pol II CTD all extend from Pol II around the Rpb4/7 subcomplex, establishing a main interface for CFI recruitment. Rna14 may directly contact the Spt5 CTR, whereas the RNA is bound by the C-terminal RRM of Rna15 and by two internal RRMs of Hrp1. The Pol II CTD is bound by the N-terminal CID domain of Pcf11. CFI subunits are drawn to scale. Important protein domains are illustrated as extensions from the respective protein core.

5. Future perspectives

Gene transcription in eukaryotes is not only an essential but also a very highly regulated process. Transitions from transcription initiation to elongation and from elongation to termination as well as all processes involved in RNA maturation have to be coordinated. Reversible phosphorylations of C-terminal protein repeat domains of Pol II and elongation factor Spt5 have proven to be very important in this process (14, 15, 60, 68, 86). This work contributed to a more detailed understanding of the extended CTD code including Ser2, Ser5 and Tyr1 phosphorylation (44) and its role in transcription termination (chapter 3). Moreover, an additional function of the Spt5 CTR in CFIA recruitment was discovered (chapter 4).

Although these findings greatly enhance our knowledge about principles underlying transcriptional regulation, still a lot of important questions are unanswered and will be discussed in the following chapters.

5.1. The CTD code extension Tyr1-P

The generation of monoclonal residue-specific phospho-CTD antibodies was an important development in the CTD research field. Only with such a specific tool was it possible to characterize the Tyr1 phosphorylation pattern on transcribed genes in yeast (44). Based on this genome-wide profile, this work could discover a regulatory function of Tyr1-P in transcription termination as well as the Tyr1-P phosphatase Glc7, revealing more details about the elongation-termination transition.

The corresponding Tyr1 kinase remains to be identified (chapter 5.1.1). Moreover, the role of Tyr1-P in transcription termination is so far only a negative function as Tyr1-P prevents termination factor recruitment. Identification of the kinase could help to characterize a positive function of Tyr1-P (chapter 5.1.2).

5.1.1. Identification of (the) Tyr1 kinase(s)

As mentioned in chapter 2.4, the *in vivo* approach to identify kinases by deletion or conditional depletion of candidates could be continued by extending the candidate list (Table 4). Non-essential proteins could be deleted, essential proteins conditionally depleted by the anchor-away procedure (203) or inhibited chemically (172). Resulting yeast strains could then be analyzed by genome-wide ChIP of Tyr1-P and Rpb3 using ChIP-seq, where ChIP is coupled to high-throughput DNA sequencing (245).

Another possible approach is fractionation of wild-type *S. cerevisiae* nuclear extracts and incubation of these fractions with unphosphorylated Pol II or synthetic unphosphorylated CTD peptides. The samples can then be analyzed by Western blotting or Enzyme-linked Immunosorbent Assay (ELISA) to compare Tyr1-P phosphorylation levels using the 3D12

antibody. Extract fractions leading to a Tyr1-P signal could then be fractionated further and analyzed by mass spectrometry (MS).

Alternatively, kinase candidates could be purified *in vitro* and incubated either with unphosphorylated Pol II or CTD peptides. Tyr1-P signals would then be detected by Western blotting or ELISA as described above.

5.1.2. A positive function of CTD Tyr1 phosphorylation

The two phosphorylations of the initial CTD code, Ser5-P and Ser2-P, have two very special transcription related functions: Ser5-P is necessary for capping enzyme recruitment, therefore enabling 5' RNA capping (50, 74). Ser2-P binds the termination factor Pcf11, thus recruiting the RNA 3'-end processing machinery (65, 128). Tyr1-P was identified as the third CTD code phosphorylation. It is possible that Tyr1-P has a positive function in transcription of similar importance.

One approach to discover Tyr1-P functions could be to perform an interaction screen in which a Tyr1-P containing CTD peptide is used to pull down physical interactors from whole cell lysate. The unphosphorylated CTD peptide can be used as a control. MS could then analyze the pulldown fractions and identify Tyr1-P CTD binders. Whole cell lysates from different organisms, e.g. from yeast and human, can be tested with the same CTD peptide, opening up the possibility to characterize the function of Tyr1 phosphorylation in yeast and higher eukaryotes in parallel. Identified CTD interacting proteins could then be cocrystallized with a Tyr1-P peptide to characterize the binding interfaces.

A possible role of Tyr1-P could lie in regulating cotranscriptional splicing (246), as the Tyr1-P signal increases along the coding region of genes. Moreover, splicing has not been linked to one specific CTD phosphorylation before. Binding studies with CTD peptides and splicing factor proteins should elucidate a possible correlation.

A second approach relies on the identification of a Tyr1 kinase. Deletion or depletion of this enzyme *in vivo* enables the study of defects in this strain, which could be linked to the decrease in Tyr1 phosphorylation. If the enzyme is depleted conditionally e.g. in an anchor away strain, different time points after rapamycin treatment could be analyzed to determine the dynamics of involved processes.

A parallel characterization of Tyr1-P functions in the human system could provide insights into evolution of this CTD phosphorylation from yeast to human. Although the Tyr1 residue is essential both in yeast and human, the coding region of genes, where Tyr1-P levels are high, differs greatly. The human genome contains longer and less closely packed genes and many of the human genes also carry very long introns (48, 247).

5.1.3. Characterization of Thr4 and Ser7 phosphorylation in yeast

To complete the study of the CTD and its modifying phosphorylations in *S. cerevisiae*, Thr4-P and Ser7-P would have to be characterized as well. For Ser7-P the genome-wide profile was already published (43), but the function of this modification remains enigmatic. In the human system, Ser7-P contributes to recruitment of the Integrator complex involved in snRNA maturation (82). Mutation of Ser7 to Ala is not lethal (31, 35) in yeast, enabling cloning of a corresponding strain that could be examined for functional defects.

Thr4 phosphorylation has not been extensively studied in yeast until now. Neither kinase nor phosphatase is known (Table 1) and the function remains ambiguous although the genome-wide profile indicates an even distribution along the coding region of genes (44). The human Thr4 kinase Plk3 was recently identified and a function of Thr4-P in transcription elongation was implicated (38). Possible functions and modifying enzymes for Thr4-P in yeast could be found by similar approaches like for Tyr1-P (chapter 5.1.1 and 5.1.2) or by mutating all Thr4 residues to Ala and analyzing functional defects in this strain.

5.2. Towards an overall CTD phosphorylation pattern

Reversible posttranslational phosphorylations of the CTD are read by proteins that are specifically recruited to these phosphorylations, which therefore are termed the “CTD code” (47, 68). This code has served as a model for the varying CTD phosphorylation patterns within one CTD molecule. However, it is so far unknown how these phosphorylations are distributed throughout the complete CTD.

For characterization of the CTD phosphorylations, specific monoclonal antibodies are the best and widely used tools available at the moment. These antibodies are selected because of their specific recognition of epitopes on phosphorylated CTD peptides *in vitro*. The conditions *in vivo*, however, can be completely different. Only accessible epitopes are recognized by the antibodies but neighbouring phosphorylations or binding proteins can mask some of them. Thus, it is not clear how many phosphorylated epitopes really exist, only how many the antibody recognizes. Moreover, the antibodies cannot detect where on the CTD these modifications occur, how many times they can be found within one molecule or if the recognized phosphorylation depends on other modifications in the vicinity (48).

To address all these questions experimentally, the best approach would be MS of the complete CTD. As this peptide is too large in one piece, it has to be cut into smaller fragments, which in turn could be subjected to MS.

The experimental output could provide information about whether all possible residues within the CTD are phosphorylated (122 possible sites in *S. cerevisiae*), or how many residues within one repeat can be phosphorylated and in what combinations. Moreover, special

phosphorylated regions or phosphorylation patterns could be discovered. If such patterns can be found, synthetic peptides consisting of these phosphorylated sequences could be used for pulldown experiments to find specific interactors.

This method, however, can only provide information about the complete population of Pol II molecules within a cell. Special pools of Pol II molecules could be isolated, e.g. by arresting cells in the cell cycle. Data from this experiment could help to characterize the coupling of transcription to the cell cycle.

With the experiments mentioned above, the CTD code could be characterized in more detail. If the code for example changes along genes at defined positions, possible connections to the histone code in the same regions could be identified much easier.

5.3. New functions of the Spt5 CTR

It has been found in this and other studies that the C-terminal repeats of Spt5 can act as a recruitment platform for transcription and RNA processing factors similar to the Pol II CTD: The CTR is involved in recruitment of the 5' RNA capping enzyme (75, 123). Upon phosphorylation by Bur1 kinase, the CTR binds the Paf1 complex during transcription elongation (86, 87). At the 3' region of genes, the CTR contributes to CFIA recruitment (chapter 4). As a consequence, the question immediately arises if there are still additional transcription related factors to be identified that are recruited by the CTR of Spt5. New CTR interactors could be discovered by *in vitro* GST-CTR pulldown experiments using whole cell extracts. The binding fractions could be further analyzed by MS.

This study showed that CFIA is partly recruited by the Spt5 CTR and partly by RNA (chapter 4). It would be interesting to characterize these interactions further. Which subunit of CFIA physically interacts with the CTR could be discovered by fluorescence anisotropy measurements of CFIA subunits with a fluorescently labeled CTR peptide. The RNA binding of CFIA could be characterized by PAR-CLIP to identify specific RNA sequences or regions that are bound by this complex (248).

Additionally, the role of CTR phosphorylation in factor binding remains to be understood. It is known that the Paf1 complex is recruited in a phosphorylation dependent manner (86, 87) but this has not been tested for other proteins. If more factors are influenced by CTR phosphorylation, it is possible that a CTR phosphorylation cycle exists, leading to a “CTR code” similar to the known CTD code. This question could be addressed experimentally by generating monoclonal antibodies against a phosphorylated and an unphosphorylated CTR. ChIP using these antibodies could then provide insights into the CTR phosphorylation cycle. Comparison of the total Spt5, the CTR and phospho-CTR profiles with transcription factor occupancy profiles could allow identification of CTR- or phospho-CTR-dependent factors.

Most importantly, the discovery of a CTR phosphorylation cycle leading to a “CTR code” would add a new level of regulation complexity in addition to the CTD code and the histone code. In the future it will be most interesting to study the relationship between these three different codes and their influence on all transcription related processes.

6. Materials and Methods

6.1. Materials

6.1.1. Bacterial and yeast strains

Table 5: *E. coli* strains

Strain	Description	Source
XL-1 Blue	Rec1A; endA1; gyrA96; thi-1; hsdR17; supE44; relA1; lac[F'proAB lacIqZΔM15 Tn10(Tetr)]	Stratagene
BL21-CodonPlus (DE3)RIL	B; F-; ompT; hsdS(rB-, mB-); dcm+; Tetr; gal λ(DE3); endA; The [argU, ileY, leuW, Camr]	Stratagene

Table 6: *S. cerevisiae* strains

Strain	Description	Source
BY4741 (wild-type)	MATa; his3Δ1 leu2Δ0 met 15Δ0 ura3Δ0	Open Biosystems
wild-type-pRS316	BY4741; pRS316 [URA3]	This study
Bur1-TAP	BY4741; BUR1::TAP::HIS3MX6	Open Biosystems
Clp1-TAP	BY4741; CLP1::TAP::HIS3MX6	A. Mayer
Ctk1-TAP	BY4741; CTK1::TAP::HIS3MX6	Open Biosystems
Elf1-TAP	BY4741; ELF1::TAP::HIS3MX6	Open Biosystems
Hrp1-TAP	BY4741; HRP1::TAP::HIS3MX6	A. Mayer
Paf1-TAP	BY4741; PAF1::TAP::HIS3MX6	Open Biosystems
Pap1-TAP	BY4741; PAP1::TAP::HIS3MX6	Open Biosystems
Pcf11-TAP	BY4741; PCF11::TAP::HIS3MX6	Open Biosystems
Rna14-TAP	BY4741; RNA14::TAP::HIS3MX6	Open Biosystems
Rna15-TAP	BY4741; RNA15::TAP::HIS3MX6	Open Biosystems
Rpb3-TAP	BY4741; RPB3::TAP::HIS3MX6	Open Biosystems
Spn1-TAP	BY4741; SPN1::TAP::HIS3MX6	Open Biosystems
Spt4-TAP	BY4741; SPT4::TAP::HIS3MX6	Open Biosystems
Spt6-TAP	BY4741; SPT6::TAP::HIS3MX6	Open Biosystems
Spt16-TAP	BY4741; SPT16::TAP::HIS3MX6	Open Biosystems
Spt5 ΔCTR	BY4741; SPT5Δ931-1063::KANMX6	This study
Spt5 ΔCTR-pRS316	BY4741; SPT5Δ931-1063::KANMX6; pRS316 [URA3]	This study
Bur1-TAP Spt5 ΔCTR	BY4741; BUR1::TAP::HIS3MX6; SPT5Δ931-1063::KANMX6	This study
Clp1-TAP Spt5 ΔCTR	BY4741; CLP1::TAP::HIS3MX6; SPT5Δ931-1063::KANMX6	This study
Ctk1-TAP Spt5 ΔCTR	BY4741; CTK1::TAP::HIS3MX6; SPT5Δ931-1063::KANMX6	This study
Elf1-TAP Spt5 ΔCTR	BY4741; ELF1::TAP::HIS3MX6; SPT5Δ931-1063::KANMX6	This study

Strain	Description	Source
Hrp1-TAP Spt5 ΔCTR	BY4741; HRP1::TAP::HIS3MX6; SPT5Δ931-1063::KANMX6	This study
Paf1-TAP Spt5 ΔCTR	BY4741; PAF1::TAP::HIS3MX6; SPT5Δ931-1063::KANMX6	This study
Pap1-TAP Spt5 ΔCTR	BY4741; PAP1::TAP::HIS3MX6; SPT5Δ931-1063::KANMX6	This study
Pcf11-TAP Spt5 ΔCTR	BY4741; PCF11::TAP::HIS3MX6; SPT5Δ931-1063::KANMX6	This study
Rna14-TAP Spt5 ΔCTR	BY4741; RNA14::TAP::HIS3MX6; SPT5Δ931-1063::KANMX6	This study
Rna15-TAP Spt5 ΔCTR	BY4741; RNA15::TAP::HIS3MX6; SPT5Δ931-1063::KANMX6	This study
Rpb3-TAP Spt5 ΔCTR	BY4741; RPB3::TAP::HIS3MX6; SPT5Δ931-1063::KANMX6	This study
Spn1-TAP Spt5 ΔCTR	BY4741; SPN1::TAP::HIS3MX6; SPT5Δ931-1063::KANMX6	This study
Spt4-TAP Spt5 ΔCTR	BY4741; SPT4::TAP::HIS3MX6; SPT5Δ931-1063::KANMX6	This study
Spt6-TAP Spt5 ΔCTR	BY4741; SPT6::TAP::HIS3MX6; SPT5Δ931-1063::KANMX6	This study
Spt16-TAP Spt5 ΔCTR	BY4741; SPT16::TAP::HIS3MX6; SPT5Δ931-1063::KANMX6	This study
rna14-1	MATα, ura3-1, trp1-1, ade2-1, leu2-3,112, his3-11,15, rna14-1	(140)
Y40343	MATα ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3 GAL psi+ tor1-1 fpr1::NAT RPL13A-2×FKBP12::TRP1	Euroscarf
Glc7-FRB-KanMX	Y40343; GLC7::FRB::KANMX6	This study
Ssu72-FRB-KanMX	Y40343; SSU72::FRB::KANMX6	This study
Ssu72-FRB-Hygromycin	Y40343; SSU72::FRB::Hygromycin	S. Etzold
Ssu72-FRB-Hygromycin, Glc7-mCherryHis	Y40343; SSU72::FRB::Hygromycin, GLC7::mCherry::HIS3MX6	This study
Bur1 as	BY4741; BUR1 L149G – FLAG3::KANMX	S. Hahn
Ctk1 as	Mata, Δade2::hisG, his3Δ200, leu2Δ0, lys2Δ0, met15Δ0, trp1Δ63, ura3Δ0, CTK1 F260G – FLAG3::KANMX	S. Hahn
Kin28 as	Mata, Δade2::hisG, his3Δ200, leu2Δ0, lys2Δ0, met15Δ0, trp1Δ63, ura3Δ0, KIN28 L83G	S. Hahn
Srb10 as	Mata, Δade2::hisG, his3Δ200, leu2Δ0, lys2Δ0, met15Δ0, trp1Δ63, ura3Δ0, srb10Δ::KANMX, pSH599 (ars cen TRP1 srb10as-1 (Y236G))	S. Hahn
ΔBck1	BY4741; BCK1::KANMX	Euroscarf
ΔBdf1	BY4741; BDF1::KANMX	Euroscarf
ΔFus3	BY4741; FUS3::KANMX	Euroscarf
ΔKns1	BY4741; KNS1::KANMX	Euroscarf
ΔKss1	BY4741; KSS1::KANMX	Euroscarf

Strain	Description	Source
ΔMck1	BY4741; MCK1::KANMX	Euroscarf
ΔMpk1	BY4741; MPK1::KANMX	Euroscarf
ΔPbs2	BY4741; PBS2::KANMX	Euroscarf
ΔSnf1	BY4741; SNF1::KANMX	Euroscarf
ΔSwe1	BY4741; SWE1::KANMX	Euroscarf
ΔTpk1	BY4741; TPK1::KANMX	Euroscarf
ΔYak1	BY4741; YAK1::KANMX	Euroscarf
Cdc5-FRB-KanMX	Y40343; CDC5::FRB::KANMX6	This study
Cdk1-FRB-KanMX	Y40343; CDK1::FRB::KANMX6	This study
Hrr25-FRB-KanMX	Y40343; HRR25::FRB::KANMX6	This study
JWY104	MAT α pra1-1 prb1-1 prc1-1 cps1-3 ura3 Δ 5 leu2-3 his-	L. Passmore

6.1.2. Growth media and additives

Table 7: Growth media

Growth medium	Description	Species
LB	1% (w/v) tryptone; 0.5% (w/v) yeast extract; 0.5% (w/v) NaCl (+ 1.5% (w/v) agar for selective media plates)	<i>E. coli</i>
YPD	2% (w/v) peptone; 2% (w/v) glucose; 1.5% (w/v) yeast extract (+1.8% (w/v) agar for solid media plates)	<i>S. cerevisiae</i>
SC -ura	0.69% (w/v) yeast nitrogen base; 0.077% (w/v) drop-out -ura; 2% (w/v) glucose; (+ 2% (w/v) agar for solid media plates)	<i>S. cerevisiae</i>
SC -leu	0.69% (w/v) yeast nitrogen base; 0.069% (w/v) drop-out -leu; 2% (w/v) glucose; (+ 2% (w/v) agar for solid media plates)	<i>S. cerevisiae</i>
SC -leu + Gal plates	0.69% (w/v) yeast nitrogen base; 0.069% (w/v) drop-out -leu; 2% (w/v) galactose; + 2% (w/v) agar for solid media plates	<i>S. cerevisiae</i>

Table 8: Growth media additives

Supplement	Description	Working concentration
Ampicillin	Antibiotic	100 µg/mL
Chloramphenicol	Antibiotic	50 µg/mL
Hygromycin B	Antibiotic	200 µg/mL
Kanamycin	Antibiotic	30 µg/mL
Tetracycline	Antibiotic	12.5 µg/mL
clonNAT	Antibiotic	100 mg/mL
IPTG	Isopropyl-β-D-thiogalactopyranosid	0.5 mM
Geneticin (G418)	Antibiotic	400 µg/mL
6-Azauracil	Pyrimidine analog	50-200 µg/mL
Mycophenolic acid (MPA)	immunosuppressive	15-45 µg/mL
Rapamycin	Immunosuppressant, FKBP12 binding	1 µg/mL
1-NA-PP1	Kinase inhibitor	5 µM
1-NM-PP1	Kinase inhibitor	10 µM
3-MB-PP1	Kinase inhibitor	20 µM

6.1.3. Plasmids and primers

Table 9: Plasmids used in this study

Vector	Insert/restriction sites	Source
pFA6a KanMX6	--	(249)
pFA6a-FRB-KanMX6	FRB-tag	Euroscarf
pFA6a-TAP-KanMX6 (modified)	TAPS tag	L. Passmore, (250)
LL280/pYM-TAP	TAP-His3	L. Lariviere
LL592/pFA6a-mCherry	mCherry-His3	L. Lariviere
pRS316	--	(251)
p415-GAL-S	--	DualSystems Biotech
p415-GAL-S-Glc7	Glc7	L. Passmore/A. Easter
p415-GAL-S-Glc7-D94A	Glc7-D94A	L. Passmore/A. Easter
p415-GAL-S-Glc7-H247N	Glc7-H247N	L. Passmore/A. Easter
pET28b-GST-CTR	GST-CTR/NcoI, HindIII	This study
pGEX-4T-1	GST-tag	GE Healthcare

Table 10: Cloning primers used in this study

Primer name	Sequence (5' → 3')
KanMX_dCTR fw	AGCTGTAAATGCGCATGGAGGCTCAGGTGGTGGCGGTGTCTAAC GGATCCCCGGGTAAATTAAG
KanMX_dCTR fw_2	GCAAAACACCAGCTGTAAATGCGCATGGAGGCTCAGGTGGTGGC GGTGTCTAACGGATCCCCGGGTAAATTAAG
KanMX_dCTR rev	TTGATTTCTTCTTGGGTGATATTGGTTCTCCTTTTGGTGACGCATA GGCCACTAGTGGATC
KanMX_dCTR rev_2	GTCTTTTTTATTGATTTCTTCTTGGGTGATATTGGTTCTCCTTTTGG TGACGCATAGGCCACTAGTGGATC
KanMX_dCTR rev_3	CACATATGCATACATACATACATACGTATATGTAAAAGTTAGAAT CAAGGCGCATAGGCCACTAGTGGATC
dCTR_control_up fw	GTGAGGGAGGTGAAGGTA
dCTR_control_up rev	AATTC AACGCGTCTGTGA
dCTR_control_down fw	ATCATGCGTCAATCGTATGT
dCTR_control_down rev	GGAAAGTGGCAGAAAGAAAG
TAP control rev	CTGCAGCGAGGAGCCGTAAT
Cdc5 AA FRB fw	CTTTGATAAAGGAAGGTTTGAAGCAGAAGTCCACAATTGTTACCG TAGATCGGATCCCCGGGTAAATTAAC
Cdc5 AA FRB rev	CAATGGACTGGTAATTTTCGTATTCGTATTTCTTTCTACTTTAATAT TGGTGAATTCGAGCTCGTTTAAAC
Cdc5 AA Ctrl ORF fw	GCCGAGTATCTACGTTTGG
Cdc5 AA Ctrl down rev	CTCTAGAACCGCTCAGGC
Hrr25 AA FRB fw	CCACAACAGCCGCTCAAGATAAACCAGCTGGCCAGTCAATTTG GTTGCGGATCCCCGGGTAAATTAAC
Hrr25 AA FRB rev	CTATATATACATATGTTTATTTTTGTGCGTTTTGAGCAATATATGT TGCGAATTCGAGCTCGTTTAAAC
Hrr25 AA Ctrl ORF fw	CGCAGCCACAACCACAGC
Hrr25 AA Ctrl down rev	GCCAAGAGGTGCTCCGC
Cdk1 AA FRB fw	CCGGATTAGCGCCAGAAGAGCAGCCATCCACCCCTACTTCCAAG AATCACGGATCCCCGGGTAAATTAAC
Cdk1 AA FRB rev	GACAGTGCAGTAGCATTGTAATATAATAGCGAAATAGATTATA ATGCGAATTCGAGCTCGTTTAAAC
Cdk1 AA Ctrl ORF fw	CTTCAGTGGCGATAGTGAG
Cdk1 AA Ctrl down rev	CGACCTCGGTAAATACGTC
Glc7 AA-FRB fw new long	GAAAGTTTATTATGTTCTTTTCAAATTTTAAAGCCAGCCCAAAA AGTCTACCAAGGCAAGCTGGGGGTAGAAAGAAAAACGGATCCC CGGGTAAATTAAC
Glc7 AA FRB rev new	GCACTAAAGGGAATTATTTGTGTATATGACGAGTGATGATTGCAT CTCCGAATTCGAGCTCGTTTAAAC
Ssu72 AA FRB-fw	GTGGCAAAGCTCACATTCTCAACTACCGTCATTATACGCTCCTTC ATATTACCGATCCCCGGGTAAATTAAC
Ssu72 AA FRB-rev	CATGAGGGCCGCTTAATGCTTATGCTTTTCTACAGTAATTGACCG TTTTGTGAATTCGAGCTCGTTTAAAC
Glc7 AA Ctrl ORF fw	CCCGACGTTGGCTTATTATG
Ssu72 AA Ctrl ORF fw	CGACTTCGTTTTCACTTGTG

Primer name	Sequence (5' → 3')
KanMX Ctrl rev	CTGCAGCGAGGAGCCGTAAT
FRB Ctrl rev2	CCAGACTCTGAAGGAAACATC
GFP Hygro fw	GATCCGCTAGGGATAACAGGGTAATATCGTACGCTGCAGGTCGA CGGATCCC
GFP Hygro rev	GGGATCCGTCGACCTGCAGCGTACGATATTACCCTGTTATCCCTA GCGGATC
Ssu72 Hygro down rev	CATGAGGGCCGCTTAATGCTTATGCTTTTCTACAGTAATTGACCG TTTTGTATCGATGAATTCGAGCTCG
Hygro Ctrl fw	GTGGAAACCGACGCCCCAG
Ssu72 Ctrl down rev	CAGTCCTATATCTGGTGATG
S1Ref2-TAPS	GTCCCCATAGTAAAAAGAAATAAATATCCTCCAAGACCAGTACA C CTGGAAGTTCTGTTCCAGGGG
S2Ref2TAPS	GAGTATATATACTACATGTTTATGTATCAGCATGTCATAGC CATCGATGAATTCGAGCTCG

Table 11: Readthrough assay and 3'-RACE primers

Primer name	Sequence (5' → 3')
ACT1 (3'-RACE)	
Gene-specific 5' primer with <u>EcoRI</u> site	TATGAATTCCTTCTGTTTTGGGTTTGGA
OligodT anchor primer with <u>XhoI</u> site (229)	GACTCGAGTCGACATCGATTTTTTTTTTTTTTTTT
Anchor primer (229)	GACTCGAGTCGACATCGA
pET28b sequencing primers (GATC Biotech AG)	
T7 fw	TAATACGACTCACTATAGGG
pRSET-RP	GGGTTATGCTAGTTATTGC
PMA1 (readthrough transcription) (146)	
Upstream primer	CTATTATTGATGCTTTGAAGACCTCCAG
Downstream primer 1	CAAGAAAGAAAAAGTACCATCCAGAG
Downstream primer 2	GTAATTTGTATACGTTTCATGTAAGTG

Table 12: qPCR primers used in this study

Primer name	Sequence (5' → 3')
ADH1 5' fw	TTTCCTTCCTTCATTCACGCACA
ADH1 5' rev	TCAAGTAACTGGAAGGAAGGCCGTA
ADH1 TSS fw	TCCTTGTTTCTTTTTCTGCAC
ADH1 TSS rev	GAGATAGTTGATTGTATGCTTGG
ADH1 ORF fw	AGCCGCTCACATTCCTCAAG
ADH1 ORF rev	ACGGTGATACCAGCACACAAGA
ADH1 3' fw	CCTGTAGGTCAGGTTGCTTT
ADH1 3' rev	CGGTAGAGGTGTGGTCAA
ADH1 polyA fw	AAAACGAAAATTCTTATTCCTGA
ADH1 polyA rev	TACCTGAGAAAGCAACCTGA
ILV5 prom fw	ACCCAGTATTTTCCCTTCC
ILV5 prom rev	TTGTCTATATGTTTTTGCTTGC
ILV5 ORF fw	CTATCAAGCCATTGTTGACC
ILV5 ORF rev	CTTGAAGACTGGGGAGAAAC
ILV5 pA fw	CCGAAACGCGAATAAATAAT
ILV5 pA rev	GTCCCGATGAGGACTTATACA
PDC1 prom fw	TGCCCTTTTTCTGTTAGAC
PDC1 prom rev	AATAAGGTGGTGTGAACGA
PDC1 ORF fw	CAAGACCAAGAACATTGTCG
PDC1 ORF rev	AAAGTGGCGTTTCTGATCTT
PDC1 pA fw	TACCATGGAAAGACCAGACA
PDC1 pA rev	CCCAGACTTAAGCCTAACCA
PMA1 5' fw	TGACTGATACATCATCCTCTT
PMA1 5' rev	TTGGCTGATGAGCTGAAACAGAA
PMA1 ORF fw	AAATCTTGGGTGTTATGCCATGT
PMA1 ORF rev	CCAAGTGTCTAGCTTCGCTAACAG
PMA1 3' fw	GGTTTCTCTGGATGGTACTTT
PMA1 3' rev	TGACTTGTGTGCGTTTCATA
TEF1 prom fw	ACCACTTCAAACACCCAAG
TEF1 prom rev	ACGACACCCTAGAGGAAGAA
TEF1 ORF fw	TTGATTATTGCTGGTGGTGT
TEF1 ORF rev	TGTTCTCTGGTTTGACCATC
TEF1 pA fw	ATTTATCCCAGTCCGATTCA
TEF1 pA rev	CTGATGTGATTTTCGACCATT
YER fw	TGCGTACAAAAGTGTCAAGAGATT
YER rev	ATGCGCAAGAAGGTGCCTAT

All oligonucleotides were purchased from Thermo Fisher Scientific.

6.1.4. Antibodies

Table 13: Antibodies used in this study

Antibody	Stock solution	Host	Application	Source
Anti-Tubulin	1:1000	rat	Western Blot	Santa Cruz (sc-69971)
Anti-GST	1:5000	goat	Western Blot	GE Healthcare (RPN1236)
Anti-Spt5 (γ N-20)	1:400	goat	Western Blot	Santa Cruz (sc-26355)
Peroxidase Anti-Peroxidase (PAP)	1:2000	rabbit	Western Blot	SIGMA (P1291)
Anti-goat-HRP	1:3000	donkey	Western Blot	Santa Cruz (sc-2020)
Anti-mouse-HRP	1:3000	goat	Western Blot	Bio-Rad (170-6516)
Anti-rat-HRP	1:5000	goat	Western Blot	SIGMA (A9037)
Anti-Rpb3 (1Y26)	1:1000-2000, 5 μ L/IP	mouse	Western Blot, ChIP	NeoClone Biotechnology
Anti-Tyr1-P (3D12)	1:7, 100 μ L/IP	rat	Western Blot, ChIP	Elisabeth Kremmer/Dirk Eick
Anti-Ser2-P (3E10)	1:14, 25 μ L/IP	rat	Western Blot, ChIP	Elisabeth Kremmer/Dirk Eick
Anti-Ser5-P (3E8)	1:14, 20 μ L/IP	rat	Western Blot, ChIP	Elisabeth Kremmer/Dirk Eick
Anti-Ser7-P (4E12)	1:14, 50 μ L/IP	rat	Western Blot, ChIP	Elisabeth Kremmer/Dirk Eick

6.1.5. Buffers and solutions

Table 14: Buffers and solutions used in this study

Name	Description	Application
100x Protease Inhibitor	0.028 mg/ml Leupeptin, 0.137 mg/ml Pepstatin A, 0.017 mg/ml PMSF, 0.33 mg/ml Benzamidine in 100% ethanol	ChIP
10x PBS	1.37 M NaCl, 27 mM KCl, 10 mM Na ₂ HPO ₄ , 20 mM KH ₂ PO ₄ adjust pH to 7.4 with HCl, ad 1L H ₂ O, autoclave	e.g. Western Blot, magnetic bead ChIP
10x TBS	1.5 M NaCl, 0.5 M Tris/HCl pH 7.5	ChIP
1x TAE	4.84 g Tris, 1.14 mL acetic acid, 0.37 g EDTA, ad 1 L H ₂ O	Agarose gel electrophoresis
1x TBE	89 mM Tris, 89 mM boric acid, 2 mM EDTA	Agarose gel electrophoresis
20x phosphatase inhibitor mix	1 mM NaN ₃ , 1 mM NaF, 0.4 mM Na ₃ VO ₄	ChIP phosphoisoforms
5x SDS-PAGE loading dye	250 mM Tris-HCl pH 7.0, 50% (v/v) glycerol, 0.5% (w/v) bromophenol blue, 7.5% (w/v) SDS, 500 mM DTT	SDS-PAGE
ChIP elution buffer (Sepharose beads)	50 mM Tris/HCl pH 7.5, 10 mM EDTA, 1% SDS	Sepharose bead ChIP
ChIP wash buffer	10 mM Tris/HCl pH 8.0, 0.25 M LiCl, 0.5% NP40, 0.5% Sodium deoxycholate, 1 mM EDTA	Sepharose bead ChIP
Coomassie staining solution	50% (v/v) ethanol, 7% (v/v) acetic acid, 0.125% (w/v) Coomassie Brilliant Blue R-250	SDS-PAGE
CPF buffer A	20 mM HEPES pH 7.9, 150 mM KCl, 0.5 mM Mg(OAc) ₂ , 3 mM DTT, 0.2 mM PMSF, 0.2 mM benzamidine-HCl	CPF purification
CPF buffer B	20 mM HEPES pH 7.9, 400 mM KCl, 0.5 mM Mg(OAc) ₂ , 3 mM DTT, 0.2 mM PMSF, 0.2 mM benzamidine-HCl	CPF purification
CPF buffer C	20 mM HEPES pH 7.9, 150 mM KCl, 0.5 mM Mg(OAc) ₂ , 3 mM DTT	CPF purification, phosphatase assay
CPF lysis buffer	100 mM HEPES pH 7.9, 200 mM KCl, 0.5 mM MgCl ₂ , 0.5 mM Mg(OAc) ₂ , 10% glycerol	CPF purification
Destaining solution	5% (v/v) ethanol, 7.5% (v/v) acetic acid	SDS-PAGE
FA-lysis buffer (150)	150 mM NaCl, 50 mM HEPES, pH 7.5 with KOH, 1 mM EDTA, 1% Triton X-100, 0.1% Sodium deoxycholate, 0.1 % SDS	ChIP
FA-lysis buffer (500)	500 mM NaCl, 50 mM HEPES, pH 7.5 with KOH, 1 mM EDTA, 1% Triton X-100, 0.1% Sodium deoxycholate, 0.1 % SDS	Sepharose bead ChIP
HU buffer	5% (w/v) SDS, 0.2 M Tris/HCl pH 6.8, 10 mM EDTA, 215 mM β-mercaptoethanol, 8 M urea, 0.01% (w/v) bromophenolblue	Western Blot
IP elution buffer (magnetic beads)	1% SDS, 0.1 M NaHCO ₃	Magnetic bead ChIP

6. Materials and Methods

Name	Description	Application
LiCl wash buffer	100 mM Tris-HCl pH 7.5, 500 mM LiCl, 1% NP40, 1% sodiumdeoxycholat	Magnetic bead ChIP
LitPEG	40% (w/v) PEG 3350, in TELit pH 8.0, filter sterilized	Chemical transformation of yeast
LitSorb	18.2% (w/v) D-Sorbitol, in TELit, pH 8.0	Chemical transformation of yeast
Pulldown lysis buffer	50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM DTT	GST-pulldown
RNase storage buffer	10 mM HEPES pH 7.5, 20 mM NaCl, 0.1% Triton X-100, 1 mM EDTA and 50% (v/v) glycerol	RNase ChIP
SDS-PAGE running buffer	invitrogen	SDS-PAGE
Ssu72 buffer	50 mM Tris-HCl pH 8.0, 10 mM MgCl ₂ , 20 mM KCl, 2% glycerol, 5 mM DTT, 0.1 mM EDTA, 0.025% Tween 20	Dephosphorylation assay
TE buffer	10 mM Tris-HCl pH 7.5, 1 mM EDTA	ChIP
TELit, pH 8.0	155 mM LiOAc, 10 mM Tris/HCl, pH 8.0, 1 mM EDTA, pH 8.0	Chemical transformation of yeast
TFB-I	30 mM KOAc, 50 mM MnCl ₂ , 100 mM RbCl, 10 mM CaCl ₂ , 15% glycerol	Preparation of chemically competent <i>E. coli</i> cells
TFB-II	10 mM MOPS pH 7.0, 75 mM CaCl ₂ , 10 mM RbCl, 15% glycerol	Preparation of chemically competent <i>E. coli</i> cells
Western transfer buffer	invitrogen	Western Blot
Yeast lysis buffer	50 mM Tris-HCl pH 7.5, 10 mM EDTA, 5.15 μM β-mercaptoethanol	Yeast genomic DNA isolation

6.2. Methods

6.2.1. Molecular cloning methods

Polymerase chain reaction (PCR)

DNA fragments for homologous recombination were amplified using the Phusion High-Fidelity DNA Polymerase (Finnzymes), 25-100 ng of template, 0.4 μ M primers, 0.2 mM dNTPs and 35 cycles of a 3-step PCR reaction: 98°C for 7 sec, 50-60°C for 20 sec, 72°C for 1-3 min. Colony PCRs were performed using Taq polymerase (Fermentas), 3-5 μ L of crude cell lysate, 0.25 μ M primers, 2 mM MgCl₂, 0.1 mM dNTPs, and 35 cycles of a 3-step PCR reaction: 94°C for 1 min, 50-60°C for 30 sec, 72°C for 30 sec. Annealing temperatures and elongation times were optimized for specific primer pairs.

Restriction and ligation

Restriction digest was performed with appropriate restriction enzymes and the conditions defined by NEB double digest finder. Restricted plasmids and inserts were ligated in a volume ratio of 2:5 using T4 DNA ligase (Fermentas) at 16-20°C overnight.

Agarose gel electrophoresis

1-2% agarose, SybrSafe (Invitrogen, 1x) in 1x TAE or TBE buffer were covered with 1x TAE or TBE buffer. PCR samples (50 μ L) were mixed with 6x DNA sample buffer (10 μ L) and loaded on the gel. A voltage of 120-150 V was applied for approximately 45 min to separate DNA fragments according to their size. DNA bands were visualized by UV-illumination. If necessary, resulting bands of the PCR products were excised under UV light. Gel slices were processed with the QIAquick Gel Extraction Kit according to the manufacturer's instruction except for elution in 30 μ L of water.

6.2.2. Bacterial methods

Competent cells

For electroporation, *E. coli* cells from a preculture were grown at 37°C and 160 rpm to an OD₆₀₀ of 0.4-0.6. Cells were cooled on ice for 15 min, harvested, for electroporation washed twice with ice-cold sterile Millipore water, resuspended in 10% glycerol, for chemical transformation washed with TFB-I, resuspended in 2 mL TFB-II per 100 mL culture, in both cases flash-frozen and stored at -80°C.

Plasmid Miniprep

Single colonies were picked, transferred to 5 mL LB media containing respective antibiotics and incubated on the shaker at 37°C for 18 h. Next, cells were harvested and to extract the

plasmid DNA from the cells, the QiaPrep Spin MiniPrep Kit (QIAGEN) was applied according to the manufacturer's instructions but eluting in water.

Transformation

A 50 μ L aliquot of competent *E. coli* cells was transformed with 0.5 μ L of purified plasmid DNA or 5 μ L of ligation product. Chemically competent cells with DNA were kept on ice for 10 min, heat-shocked at 42 °C for 45 sec and after another 2 min on ice resuspended in 700 μ L of prewarmed LB media and incubated on the shaker at 37°C and 650 rpm for 30 min. Finally, cells were transferred to LB plates containing respective antibiotics and incubated at 37°C for 16 h. Electrocompetent cells were transformed with DNA by pulsing 2.5 kV for 5 msec with the MicroPulser™ Electroporator (BioRad) and incubation on ice for 5-10 min. Prewarmed LB was added and cells processed as above.

6.2.3. Yeast methods

Chemical transformation of yeast cells

A 50 mL yeast culture in YPD was grown from a preculture to an OD₆₀₀ of 0.5-0.7. Cells were harvested at 4,000 rpm for 10 min at room temperature, washed with 25 mL sterile water, then with 5 mL LitSorb, then resuspended in 360 μ L LitSorb. Salmon sperm DNA (Sigma) was heated for 10 min at 100°C and cooled on ice. 40 μ L of DNA were added to the cells. 10 μ L of PCR product or 1 μ L of purified plasmid were added to a 50 μ L aliquot of competent cells and diluted with 360 μ L of LitPEG. After incubation for 30 min at room temperature, 47 μ L of DMSO were added and cells incubated again for 15 min at 42°C. Finally, cells were centrifuged for 3 min at 2,000 rpm and the pellet resuspended in 1 mL of YPD and incubated for 3 h on a turning wheel at room temperature. Cells were centrifuged again, resuspended in 200 μ L of YPD, plated on G418-plates and incubated at 30°C for 3 days.

Transformation of yeast by electroporation

A 50 mL yeast culture in YPD was grown from a preculture to an OD₆₀₀ of 0.7-0.8. Cells were harvested by centrifugation at 4,000 rpm and 4°C, washed twice with 50 mL ice-cold sterile water and once with 50 mL ice-cold 1 M sorbitol. Next, the pellet was resuspended in 20 mL of 0.1 M LiOAc/30 mM DTT and incubated for 30 min at 160 rpm and 30°C. Cells were harvested and the pellet washed twice with 50 mL ice-cold 1 M sorbitol and finally resuspended in 100 μ L of 1 M sorbitol. For transformation, 5 μ L of DNA solution (at least 1 μ g in total) were added to 100 μ L of cell suspension and a pulse of 1.5 kV (5 msec) was applied to the cells in an electroporation cuvette in a MicroPulser™ Electroporator (Bio-Rad). 1 mL of prewarmed YPD media was added and the cell suspension incubated at 30°C and

700 rpm for 3-4 h. Cells were pelleted at 2,000 rpm for 3 min, resuspended in 200 μ L of media, and plated on corresponding marker plates and incubated at 30°C for 2-3 days.

Strain cloning and validation

S. cerevisiae kinase deletion strains were validated by colony PCR using a gene-specific upstream primer and a reverse primer within the KanMX cassette. *S. cerevisiae* strains Cdc5-, Cdk1-, Glc7-, Hrr25- and Ssu72-FRB-KanMX6 were cloned by amplifying the cassette from vector pFA6a-FRB-KanMX6 (Euroscarf) with suggested primer pairs F2, R1 (252) and transforming the parental strain Y40343 by homologous recombination. The functional insertion of the FRB tag was verified (i) by PCR and (ii) by spot diluting the strains on YPD + rapamycin plates (Figure 10 and Figure 17). The mCherry-tag on Glc7 was inserted by homologous recombination of a PCR product amplified from the plasmid pFA6a-mCherry-His3MX6. For recovery experiments, the Glc7-FRB strain was transformed with the plasmids p415-GAL-S as a control or the same plasmid expressing galactose-inducible wild-type, D94A or H247N catalytic mutants of Glc7.

S. cerevisiae strains containing C-terminally TAP tagged target proteins (TAP strains) were validated (i) by colony PCR using a gene-specific upstream primer and a reverse primer within the TAP tag and (ii) by Western blotting with an anti-TAP antibody (PAP, Sigma) to control the expression level of TAP-tagged proteins. The 15 C-terminal hexapeptide repeats, amino acids 931 to 1063, of Spt5 were deleted from TAP strains by homologous recombination with the KanMX6 cassette, amplified from the pFA6a-KanMX6 vector.

Yeast genomic DNA

Yeast cells from 1 mL overnight culture diluted to an OD₆₀₀ of 5.0 were harvested and resuspended in 180 μ L of yeast lysis buffer containing 1.8 units of lyticase (Sigma). Genomic DNA was isolated using the QIAcube (QIAGEN) and the standard protocol “DNeasy Blood & Tissue - Bacteria (Gram+) or yeast - Enzymatic lysis” according to the manufacturer’s instructions. The yield of DNA was 50-150 ng/ μ L.

Colony PCR

A single colony from the original transformation plate was transferred to a second plate. A pinhead-sized amount of cells was picked from this replated single colony and dissolved in 100 μ L of 20 mM NaOH. Approximately 50 μ L of 0.5 mm glass beads (Roth) were added to the solution which was boiled at 95°C and 1,400 rpm in a thermomixer for 5 min. Afterwards the sample was centrifuged at 13,000 rpm for 1 min. 5 μ L of the supernatant was used as template for a standard PCR reaction as described in chapter 6.2.1.

Phenotyping

For growth curve measurements, liquid overnight cultures of wild-type and mutant yeast strains were diluted with YPD to a starting OD₆₀₀ of 0.1. Yeast cells were grown for 18 h and the OD₆₀₀ was determined every 90 min. Biological triplicate measurements were performed.

For spot dilutions in general, yeast strains were grown in YPD or Synthetic Complete (SC) media at 30°C to stationary phase and diluted to an OD₆₀₀ of ~1.0 with water. 3 µL of cells were spotted on corresponding plates in 10-fold serial dilutions manually or using the epMotion 5075 (Eppendorf). Plates were incubated at 30°C and inspected daily.

For Figure 10 and Figure 17a, FRB strains were grown in YPD and spotted on YPD plates with or without 1 µg/mL f.c. rapamycin (LC Laboratories). For Figure 18a, the Glc7-FRB strain containing p415-GAL-S plasmids was grown in SC -leu media with 2% glucose, resuspended in SC -leu media without glucose and spotted on SC -leu plates with 2% glucose (repressed) or galactose (induced) with or without 1 µg/mL rapamycin. For Figure 20a, wild-type and ΔCTR cells transformed with the pRS316 plasmid were grown in SC -ura media and spotted on SC -ura plates containing 50 µg/ml of 6-AU or 15 µg/ml of mycophenolic acid (MPA).

Fluorescence microscopy

For fluorescence microscopy, cells were grown in 1 mL YPD + 40 µg/mL adenine hemisulfate (Sigma) to an OD₆₀₀ of ~0.8 and 0.5 mL of paraformaldehyde solution (10% paraformaldehyde (Sigma), 13 mM NaOH, 150 mM PBS) was added for 10 min. Cells were washed with 1x PBS once, resuspended in 50 µL 1x PBS and applied to a poly-L-lysine-coated glass slide (Sigma, Roth). Remaining liquid was removed and 1.0 ng/µL DAPI (Sigma) in 1x PBS was applied for 1 min. Fixed cells on glass slide were washed twice with 1x PBS and covered with a cover slip. DAPI and mCherry fluorescence was analyzed under the microscope (Leica DM2500, EL 6000). Camera DFC365FX and Software LAS AF 6000 Modular Systems Version 2.6.0.7266 (Leica) were used for image analysis.

Yeast cell lysates

Yeast cultures were grown in 200 ml YPD medium at 30°C to an OD₆₀₀ of ~ 0.8. Subsequent steps were performed at 4°C. Cells were collected by centrifugation, washed with 1x TBS and FA lysis buffer containing protease inhibitor (PI). Cell pellets were flash-frozen in liquid nitrogen and stored at -80°C. Pellets were thawed, resuspended in 1 ml FA lysis buffer containing PI and disrupted by bead beating (Retsch) in the presence of 1 ml silica-zirconia beads for 30 min at 4°C. Lysis efficiency was typically >80%. Cell debris was removed by centrifugation.

Denatured yeast cell lysates

A pinhead-sized amount of cells from a YPD plate was resuspended in 1 mL of cold water. 150 μ L of 7.5% β -Mercaptoethanol in 1.85 M NaOH were added and cells were incubated for 15 min on ice. Afterwards, 150 μ L of 55% TCA were added and the sample incubated for 10 min on ice. Next, cells were centrifuged for 10 min at 14,000 rpm and 4°C, the supernatant discarded and the pellet resuspended in 50 μ L HU buffer. The sample was neutralized using 5-15 μ L of 1 M Tris base and incubated for 10 min at 1,400 rpm and 65°C. Cells were centrifuged for 5 min at 14,000 rpm at room temperature and the supernatant was collected. Absorption of samples at 280 nm wavelength was determined at a BioPhotometer (Eppendorf) using UV cuvettes and a 1:20 dilution of the lysates in water.

6.2.4. Protein methods

GST protein expression and -pulldown

A synthetic DNA construct of the GST tag based on the pGEX3T vector fused to the DNA sequence coding for Spt5 CTR residues 931-1063 was synthesized (MrGene GmbH) and cloned into vector pET28b. Expression of the GST-Spt5CTR fusion construct (pET28b) and the GST tag alone (pGEX-4T-1) in *E. coli* was induced with 0.5 mM IPTG at 18 °C overnight in 1 L cultures. Cells were lysed by sonification in 50 mL pulldown lysis buffer with PI for 15 min (Branson Sonifier 250). The cell debris was removed by centrifugation. For pulldown experiments, *E. coli* cell lysates were incubated with 100 μ L Glutathione Sepharose beads (GE Healthcare) and pre-equilibrated in pulldown lysis buffer for 1 h at 4°C. Beads were washed eight times with pulldown lysis buffer and incubated with 300 μ L lysate from yeast strains expressing TAP-tagged versions of CFI subunits for 1 h at room temperature. Next, beads were washed eight times with pulldown lysis buffer and proteins were eluted from the beads eight times with pulldown lysis buffer containing 10 mM glutathione. All elution fractions were pooled and the protein precipitated with 10% TCA. The protein pellet was washed with ice-cold acetone and resuspended in 2x SDS-PAGE loading dye. Wash and elution fractions were analyzed by SDS-PAGE and Western blotting with antibodies against the TAP-tag of CFI subunits and the GST-tag.

CPF purification

Primers specific to 40-bp of the C-terminus and 3'-UTR of *REF2* were used to PCR amplify the TAP tag cassette from a modified pFA6a-TAP-kanMX6 vector (250), where a StrepII tag replaces the calmodulin binding peptide. The PCR product was transformed into *S. cerevisiae* strain JWY104 via homologous recombination. Cells were grown at 30°C in YPD to an OD₆₀₀ of 5.0-5.5, pelleted at 6,700 g, washed once with cold water, pelleted at 3,200 g, flash-frozen in liquid nitrogen and stored at -80 °C. Cells were mixed 1:1 (v:v) with CPF lysis buffer containing 1.6 μ g/mL DNase I and EDTA-free protease inhibitor cocktails tablets (Roche)

and lysed at 35 kPsi via cell disruption (Constant Systems). Lysate was clarified at 25,000 g for 25 min, bound to IgG sepharose (GE Healthcare) for 1 h and washed with CPF buffer A. CPF was eluted from the IgG resin with TEV protease at 16 °C for 1.5 h, and the supernatant applied to StrepTactin sepharose (GE Healthcare) for 1 h. StrepTactin resin was washed three times with CPF buffer B, three times with CPF buffer C, then eluted with 5 mM desthiobiotin in buffer C. Concentration was determined via absorbance at 280 nm using a calculated extinction coefficient. Bands were excised from Coomassie blue-stained SDS-PAGE gels (Figure 13a) and identified via tryptic digest mass spectrometry by the Passmore laboratory.

SDS-PAGE

Protein samples were separated by SDS-PAGE according to their size using NuPAGE 4-12% Bis-Tris gels (life technologies) in the novex Mini-cell chambers (Invitrogen) and MES and MOPS buffers according to the manufacturer's instructions. Protein samples were boiled in 5x SDS loading dye for 1.5-5 min before loading. PAGE was performed at 160-200 V. Gels were stained using Coomassie staining solution or Instant BlueTM (expedion) for 20 min and destaining overnight or used for Western Blot.

Western Blot

After SDS-PAGE using PageRulerTM Prestained Protein Ladder (Fermentas), the gel was blotted for 1-2 h at 100-150 V or overnight at 30 V in a Mini Trans-Blot System (Bio-Rad) filled with NuPAGE® transfer buffer containing 10% (v/v) ethanol at 4°C onto a polyvinylidene difluoride (PVDF) membrane (Roth), which had been pre-soaked in ethanol and transfer buffer. The membrane was blocked with 2% milk powder (Roth) in 1x PBS for 20 min or overnight, then probed with primary antibodies or HRP-coupled PAP antibody for TAP tag detection in 2% milk/PBS using the appropriate dilution (Table 13). Afterwards, membranes were washed three times for 5 min with 2% milk/PBS and incubated for 1 h with secondary antibodies in 2% milk/PBS using the appropriate dilution. Finally, membranes were washed three times with 1x PBS for 5 min. Antibody detection was performed using the Pierce® ECL Western Blotting Substrate (Thermo Scientific) and Amersham HyperfilmsTM ECL (GE Healthcare) and developing with X-omat M35 (Kodak) or using the LAS-3000 camera (FUJIFILM). For Figure 20c, protein signals were quantified relative to α -tubulin signals as loading control using the ImageQuant 5.0 software (GE Healthcare).

Dephosphorylation assay with RNA Pol II and CPF

After prephosphorylation with MAP Kinase (NEB) for 60 min at 30°C, with the kinase inactivated for 20 min at 65°C, 234 nM of purified Pol II (253) was incubated with 59 nM of CPF at 30°C for indicated periods in CPF buffer C. The CPF complex was purified and analyzed by Lori Passmore's group in Cambridge as described above. For Glc7 inhibition, 10 mM EDTA or 200 nM mycrocystin-LR (Sigma) was added. Samples were dissolved in

3x SDS-PAGE buffer and boiled at 95°C for 90 sec. After SDS-PAGE (0.5 µg of Pol II per lane), samples were blotted onto a PVDF membrane (Millipore) and the membrane probed with primary antibodies α -Tyr1P, α -Ser5P, α -Ser2P and α -Rpb3. Secondary antibodies α -rat IgG HRP and α -mouse IgG HRP were used. Antibody detection by Western blotting was performed as described above.

For Pol II dephosphorylation with human Ssu72 (hSsu72), 3 µg of Pol II was incubated with 1, 3 or 5 µg purified recombinant hSsu72 (204) in Ssu72 buffer for 2 h at 30°C. Samples were dissolved in 3x SDS-PAGE buffer and processed as described above.

6.2.5. RNA methods

RNA isolation

For RNA isolation, overnight yeast cultures were diluted in 20 mL fresh YPD medium to an OD₆₀₀ of ~ 0.1 and grown at 30°C to mid-log phase (OD₆₀₀ ~ 0.8). The RiboPure™-Yeast Kit (Applied Biosystems) was used to isolate yeast RNA according to the manufacturer's instructions.

3'-RACE

3'-RACE was performed for the *ACT1* gene as described (229). This gene was shown to possess several alternative pA sites and was used in 3'-end processing studies (146, 228). The *ACT1* cDNA was synthesized as described (146) using Superscript II reverse transcriptase (Invitrogen), 0.5 µM oligodT-anchor primer containing a XhoI restriction site, and 2 µg RNA template. Next, RNA was digested using RNase H (New England BioLabs Inc.) for 20 min at 37°C. The enzyme was inactivated at 65°C for 20 min. PCR reactions were conducted using the cDNA samples, 0.25 µM gene-specific upstream primer containing an EcoRI restriction site, an anchor primer and *Taq* DNA Polymerase (Fermentas). PCR product and the pET28b vector were digested with EcoRI-HF™ as well as XhoI (New England BioLabs Inc.), ligated and transformed into competent *E. coli* XL1-blue cells. Plasmids from 25 *E. coli* colonies from each strain were sequenced (GATC Biotech) using pET28b sequencing primers.

Readthrough assay

The readthrough was performed for the *PM1* gene according to (146), using a gene-specific forward primer and two reverse primers positioned at different regions downstream of the transcription termination site (Figure 28). RNA isolation, reverse transcription and PCR were conducted as described above and the PCR products for wild-type and Spt5 Δ CTR mutant were compared by standard agarose gelelectrophoresis. RNA isolation from the rna14-1 temperature sensitive mutant was performed as described before, but with the following modifications. Yeast cultures were either grown at 24°C to mid-log phase before RNA isolation (rna14-1 24°C, Figure 28b) or were grown to an OD₆₀₀ of ~ 0.8 at 24°C and then

transferred to 37°C (restrictive temperature) for 1 h (rna14-1 37°C, Figure 28b) before RNA was isolated.

6.2.6. ChIP methods

General ChIP with protein-specific antibodies

For ChIP experiments, yeast cultures were grown in 40 ml YPD medium at 30°C to mid-log phase ($OD_{600} \sim 0.8$) in biological duplicates, treated with 1% formaldehyde (Sigma) for 20 min at 20°C, and crosslinking was quenched with 5 ml 3 M glycine for 10 min. For Shokat strains or anchor-away strains, twice the volume of culture was grown and then splitted at an OD_{600} of 0.65. Both new cultures were incubated with equal volumes of kinase inhibitor and DMSO, respectively, for Shokat strains or rapamycin (1µg/mL f.c. in DMSO) and DMSO, respectively, for anchor-away strains at 30°C for another 60 min before formaldehyde crosslinking. Subsequent steps were performed at 4°C with pre-cooled buffers containing PI. Cells were collected by centrifugation, washed with 1x TBS, and twice with FA lysis buffer. Cell pellets were flash-frozen in liquid nitrogen and stored at -80°C. Pellets were thawed, resuspended in 1 ml FA lysis buffer, which contained 1x phosphatase inhibitor mix (Table 14) or a phosphatase inhibitor cocktail (PhosSTOP, Roche) when CTD phosphoisoform antibodies were used (Table 13), and disrupted by bead beating (Retsch) in the presence of 1 ml silica-zirconia beads for 30 min at 4°C. Lysis efficiency was typically >80%. Chromatin was solubilized and fragmented via sonication with a Bioruptor™ UCD-200 (Diagenode Inc.) connected to a water cooler system for 35 cycles 30 sec on/30 sec off at high intensity. After putting aside 30 µL chromatin solution as Input sample, immunoprecipitation was performed either using sepharose or magnetic beads. For the sepharose bead protocol, 670 µl of sample was immunoprecipitated with the appropriate amount of antibody per immunoprecipitation (IP) (Table 13) at 4°C overnight on a rotating wheel. 12.5 µL Protein G Sepharose 4 Fast Flow and 12.5 µL nProtein A Sepharose 4 Fast Flow (GE Healthcare) were then coupled to the antibodies in the chromatin solution for 90 min at 4°C on a rotating wheel. Immunoprecipitated chromatin on sepharose beads was washed in Ultrafree® Centrifugal Filters (Durapore PVDF 5.0 µm) three times with 380 µL FA lysis buffer, twice with FA lysis buffer containing 500 mM NaCl, twice with ChIP wash buffer and once with TE buffer. Proteins and DNA were eluted from the beads for 10 min at 65 °C with 120 µL ChIP elution buffer. Eluted chromatin was digested with Proteinase K (Sigma) at 37°C for 2 h and the reversal of crosslinks was performed at 65°C overnight. DNA was purified with the QIAquick PCR Purification Kit (Qiagen) according to the manufacturer's instructions. For the magnetic bead protocol, 670 µL of chromatin sample was incubated with 16.5 µL antibody-coated and prewashed magnetic Dynabeads® Protein G (life technologies) at 4°C overnight. Beads were prewashed four times with 5 mg/mL BSA in 1x PBS with PI, coated with the respective antibody for 30 min at 4°C and washed again three times. After immunoprecipitation, beads

were washed five times with LiCl wash buffer and once with TE buffer. Immunoprecipitated chromatin was eluted for 5 min at 95°C with IP elution buffer.

ChIP of TAP-tagged proteins

ChIP of TAP-tagged proteins was performed as described above with the following modifications: For the IP, 670 µl of TAP-tagged sample was immunoprecipitated with 25 µl IgG Sepharose 6 Fast Flow (GE Healthcare) at 4°C for 1 h.

RNase ChIP

The RNase-ChIP assay was performed as described above with the following adaptations: First, the crosslinking time was reduced to 5 min. Second, chromatin solution was divided into two samples before the IP step. One chromatin sample was treated with 7.5 U of RNase A and 300 U of RNase T1 (Ambion), the other sample was treated with the same volume of RNase storage buffer. After incubating for 30 min at room temperature, the IP was performed as above. The RNase ChIP data of Rna14 and Rna15 was further analyzed as follows: After averaging fold enrichments of the respective gene region between biological replicates, the values of the chromatin samples that were not treated with the RNase mix were set to 100% and all other values multiplied by the same factor.

Fragment size control

Chromatin fragment size was determined after DNA shearing by phenol-extracting a 100 µL sample twice and ethanol-precipitating over-night. The pellet was resuspended in 20 µl TE buffer and incubated with 10 µl RNase A/T1 Mix (2 mg/ml RNase A, 5000 U/ml RNase T1; Fermentas) at 37°C for 1 h. Alternatively, the chromatin sample was diluted with 92 µL TE buffer and 8 µL Proteinase K (Sigma) were added at 65°C overnight. RNase A/T1 Mix digest was performed at 37°C for 30 min. The resulting DNA sample was in both cases analyzed on a 1.5% agarose gel.

Quantitative real-time PCR (qPCR)

Input and immunoprecipitated samples were assayed by qPCR to assess the extent of protein occupancy at different genomic regions. Primer pairs directed against 60-70 nucleotide long regions of the transcriptional start site (TSS, 5'), coding (ORF), pA and 3' regions of the *ADH1*, *ILV5*, *PDC1*, *PMA1* and *TEF1* genes as well as against a heterochromatic control region of chromosome V were used to determine PCR efficiencies. All PCR efficiencies ranged between 95-100%. PCR reactions contained 1 µl DNA template, 2 µl of 10 µM primer pairs and 12.5 µl iTaq SYBR Green Supermix (Bio-Rad). qPCR was performed on a Bio-Rad CFX96 Real-Time System (Bio-Rad Laboratories, Inc.) using a 3 min denaturing step at 95°C, followed by 49 cycles of 30 sec at 95°C, 30 sec at 61°C and 15 sec at 72°C. Threshold cycle values were determined by application of the corresponding Bio-Rad CFX Manager

software versions 1.1 and 3.0 using the Ct determination mode “Regression”. Fold enrichment of any given region over an ORF-free heterochromatic region on chromosome V (control) was determined as described (254) using Microsoft Excel 2007 and the following formula:

$$\text{Fold enrichment} = 1.9^{\frac{Ct_{IP,control} - Ct_{Inp,control}}{Ct_{IP} - Ct_{Inp}}}$$

For chapter 2.3, Tyr1-P fold enrichments were normalized against Rpb3 fold enrichments using median-centered normalization: Tyr1-P values for every sample (e.g. replicate 1 and 2 of untreated or wild-type) and every primer pair were divided by the median of all corresponding Rpb3 values (median of replicate 1 and 2 of untreated or wild-type) and multiplied by the overall median of all Rpb3 values (median of untreated or wild-type 1 and 2 + treated or mutant 1 and 2). Resulting relative fold enrichments were averaged.

6.2.7. ChIP and tiling microarray hybridization (ChIP-chip)

ChIP for ChIP-chip

ChIP was performed in duplicates like described above with the following modifications: (i) yeast cultures were grown in 600 mL YPD to an OD₆₀₀ of ~0.8, or in 1.2 L when split afterwards; (ii) cell disruption by bead beating was performed either using the Retsch mill for 2 h with 5 min breaks on ice after 45 and 90 min or using the FastPrep®-24 Instrument (MP Biomedicals) with the QuickPrep adaptor 8 times for 40 sec at 4.0 m/s; (iii) the chromatin pellet was washed twice with FA lysis buffer by resuspending and centrifuging for 15 min at 13,000 rpm; (iv) when sepharose beads were used, the elution from the beads was done for 60 min at 65°C; (v) three parallel IPs were performed from each 600 mL culture and the DNA was combined before purification with the QIAquick PCR Purification Kit (Qiagen) according to the manufacturer’s instructions with an elution volume of 100 µL.

RNase digest and whole genome amplification

IP and input samples were digested with 5 µg RNase A for 20 min at 37°C, then a second PCR purification step was performed in which the three identical IPs were combined via purification in one column and elution in 50 µL. IP samples were concentrated using the Concentrator plus (Eppendorf) to 10 µL volume. IP and input samples were amplified with the GenomePlex© Complete Whole Genome Amplification 2 (WGA2) Kit (Sigma) and the WGA Protocol for ChIP-chip from the Farnham Lab (<http://www.genomecenter.ucdavis.edu/farnham/protocol.html>): Firstly, a library preparation step was performed to anneal uniform priming sites to the DNA fragments. 20 ng of Input DNA in 10 µL and the 10 µL concentrated IP DNA were incubated with 2 µL Library Preparation buffer and 1 µL Stabilization Solution in thin-walled 0.2 mL tubes for 2 min at 95°C, then cooled for 2 min. 1 µL Library Preparation Enzyme was added and the mixture incubated in a thermocycler as follows: 16°C for 20 min, 24°C for 20 min, 37°C for 20 min and 75°C for 5 min. Secondly,

the first amplification round was performed with 7.5 μ L 10x Amplification Master Mix and 5 μ L WGA DNA Polymerase in 75 μ L volume using the following program: 95°C for 3 min, then 14 cycles of 94°C for 15 sec and 65°C for 5 min. DNA was purified using the PCR Purification Kit (Qiagen) and concentration was determined with a ND-1000 Spectrophotometer (NanoDrop Technologies). Thirdly, a second amplification round was performed with 20 ng of the resulting DNA using the same reaction but adding 0.4 mM dUTPs (Promega) to be incorporated for enzymatic digestion. DNA was purified and concentration determined as above.

Enzymatic DNA fragmentation, labeling and microarray hybridization

These steps were performed according to the Affymetrix Chromatin Immunoprecipitation Assay Protocol P/N 702238 using the GeneChip[®] WT Double-Stranded DNA Terminal Labeling Kit (P/N 900812, Affymetrix). Enzymatic fragmentation and terminal labeling were performed by application of the GeneChip[®] WT Double-Stranded DNA Terminal Labeling Kit (P/N 900812, Affymetrix): 5.5 μ g of re-amplified DNA was fragmented in a 48 μ L reaction containing fragmentation buffer, 15 U UDG and 225 U APE1 at 37°C for 1 h 15 min and inactivation at 93°C for 2 min. Fragment size was determined on Invitrogen E-Gels (4%) and was usually around 50-70 bp. Fragments were labeled at the 3'-end in a 60 μ L reaction with 1 μ L of 5 mM GeneChip[®] DNA labeling reagent and 60 U terminal nucleotidyl transferase (TdT) at 37°C for 1 h and inactivation at 70°C for 10 min. For the following steps, the GeneChip[®] Hybridization, Wash, and Stain Kit (P/N 900720) was used: The complete labeled DNA sample was applied to a high-density custom-made GeneChip[®] *S. cerevisiae* tiling array (11) (Affymetrix, P/N 520055) in a 200 μ L volume containing 50 pM of control Oligonucleotide B2, 1x Hybridization mix and 7% DMSO in nuclease-free water and hybridized at 45°C for 16 h by rotation at 60 rpm in a GeneChip[®] Hybridization Oven 640 (Affymetrix). Arrays were washed and stained using the FS450_0001 protocol at the GeneChip[®] Fluidics Station 450 (Affymetrix) and scanned at the Affymetrix GeneChip[®] Scanner 3000 7G (Affymetrix) and the corresponding software. The resulting image files of the arrays were inspected manually for any damage and the CEL intensity files (.cel) were used for further data analysis.

Bioinformatics analysis

Tiling array output data consist of fluorescence intensities, which refer to the amount of DNA sample that is hybridized to a single target within the yeast genome, represented by probes on the array. The output of the microarray scanner is a .cel file containing the intensities. An additional array specific .bpmmap file then maps the intensities of the probes on the array to their position in the yeast genome.

Firstly, the data has to be normalized to be able to compare it to other array experiments and to account for saturation effects and sequence specific biases. This was done using R (255)

and Bioconductor (256). The R package Starr (257) was used for data import and conversion to the basic Bioconductor object class for microarray data, ExpressionSet. Data normalization was performed in three steps: (i) Quantile normalization between replicates – biological duplicates of each factor/experiment were analyzed, except for Figure 14 (one replicate); (ii) for each condition (including the reference measurements) we averaged the signal for each probe by calculating the geometric average over replicate intensities; (iii) data from TAP-tagged factors was normalized using a combined mock IP (ChIP with an untagged strain) plus input reference normalization; data from ChIP with specific antibodies was normalized by dividing through input reference intensities as no mock IP can be performed with the same antibody. Normalized signal was converted to occupancy values between 0-100% by setting the genome-wide 99.8% quantile to 100% occupancy and the 10% quantile to 0% occupancy. Further details of data normalization are described in (43).

Secondly, transcript-wise occupancy profiles were generated by calculating the normalized occupancy profiles as the median signal of all probes overlapping each nucleotide position (6.5 probes on average). Probe intensities were smoothed using sliding window smoothing (R package Ringo (258)) or running median smoothing (window half size of 75 bp).

Thirdly, occupancy profiles were averaged over sets of genes. Therefore, yeast genes were filtered first: Of all nuclear protein-coding genes of *S. cerevisiae* S288C that were classified by the Saccharomyces Genome Database as “verified” or “uncharacterized” (<http://www.yeastgenome.org/>, 5769 genes), only 4,366 genes have annotated TSS and pA sites (10), which are necessary to align genes. If the TSS was annotated downstream of the ATG codon or the pA site upstream of the Stop codon, the genes were excluded. This was done also for genes for which the distance between TSS and ATG or pA and Stop was longer than 200 bp to exclude wrongly annotated TSS/pA sites. 3448 genes remained. As the yeast genome is very compact and the resolution of ChIP-chip experiments limited, profiles can be influenced by neighbouring genes. This “spill-over” could be minimized by focusing on genes that show a minimum ORF (transcript) distance to neighbouring genes of 250 bp (200 bp), leaving 1786 genes for analysis. From these, we analyzed the 50% most highly expressed genes ((217), 1140 genes). Genes were divided into four length classes: Xtremely Short (XS, 256-511 bp, 93 genes), Short (S, 512-937 bp, 266 genes), Medium (M, 938-1537 bp, 339 genes) and Long (L, 1538-2895 bp, 299 genes). Profiles within these groups were scaled to median gene length and the gene-averaged profiles were calculated by taking the median over genes at each genomic position.

Profiles in chapter 3 were cut around the pA site from 400 bp upstream to 400 bp downstream. Since CPF acts around the pA site, we were especially interested in whether the profiles change around the pA site relative to the region upstream of the pA site. To better visualize these relative differences we further normalized the data to have approximately equal occupancy levels upstream (around -400 bp) of the pA site for the ‘-’ and ‘+’

rapamycin treated profiles. Corresponding non-normalized data averaged over the entire gene length (250 bp upstream of TSS to 400 bp downstream of pA site) are shown in Figure 16.

Only genes longer than 800 bp and with at least 400 bp distance to downstream genes were included (619 genes) and averaged at each genomic position.

References

1. Crick FH (1958) On protein synthesis. *Symposia of the Society for Experimental Biology* 12:138-163.
2. Weinzierl RO (1999) *Mechanisms of Gene Expression* (Imperial College Press, London).
3. Grummt I (2003) Life on a planet of its own: regulation of RNA polymerase I transcription in the nucleolus. *Genes & development* 17(14):1691-1702.
4. Russell J & Zomerdijk JC (2005) RNA-polymerase-I-directed rDNA transcription, life and works. *Trends in biochemical sciences* 30(2):87-96.
5. Werner M, Thuriaux P, & Soutourina J (2009) Structure-function analysis of RNA polymerases I and III. *Current opinion in structural biology* 19(6):740-745.
6. Wyers F, *et al.* (2005) Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* 121(5):725-737.
7. Davis CA & Ares M, Jr. (2006) Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America* 103(9):3262-3267.
8. Lykke-Andersen S & Jensen TH (2007) Overlapping pathways dictate termination of RNA polymerase II transcription. *Biochimie* 89(10):1177-1182.
9. Dieci G, Fiorino G, Castelnuovo M, Teichmann M, & Pagano A (2007) The expanding RNA polymerase III transcriptome. *Trends in genetics : TIG* 23(12):614-622.
10. Nagalakshmi U, *et al.* (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* 320(5881):1344-1349.
11. David L, *et al.* (2006) A high-resolution map of transcription in the yeast genome. *Proceedings of the National Academy of Sciences of the United States of America* 103(14):5320-5325.
12. Cramer P, *et al.* (2008) Structure of eukaryotic RNA polymerases. *Annual review of biophysics* 37:337-352.
13. Hsin JP & Manley JL (2012) The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes & development* 26(19):2119-2137.
14. Brannan K & Bentley DL (2012) Control of Transcriptional Elongation by RNA Polymerase II: A Retrospective. *Genetics research international* 2012:170173.
15. Buratowski S (2009) Progression through the RNA polymerase II CTD cycle. *Molecular cell* 36(4):541-546.
16. Ringel R, *et al.* (2011) Structure of human mitochondrial RNA polymerase. *Nature* 478(7368):269-273.
17. Haag JR & Pikaard CS (2011) Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. *Nature reviews. Molecular cell biology* 12(8):483-492.
18. Olins DE & Olins AL (2003) Chromatin history: our view from the bridge. *Nature reviews. Molecular cell biology* 4(10):809-814.

19. Rando OJ & Ahmad K (2007) Rules and regulation in the primary structure of chromatin. *Current opinion in cell biology* 19(3):250-256.
20. Li B, Carey M, & Workman JL (2007) The role of chromatin during transcription. *Cell* 128(4):707-719.
21. Luger K, Mader AW, Richmond RK, Sargent DF, & Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389(6648):251-260.
22. Venters BJ & Pugh BF (2009) How eukaryotic genes are transcribed. *Critical reviews in biochemistry and molecular biology* 44(2-3):117-141.
23. Chapman RD, Heidemann M, Hintermair C, & Eick D (2008) Molecular evolution of the RNA polymerase II CTD. *Trends in genetics : TIG* 24(6):289-296.
24. Nonet M, Sweetser D, & Young RA (1987) Functional redundancy and structural polymorphism in the large subunit of RNA polymerase II. *Cell* 50(6):909-915.
25. Zehring WA, Lee JM, Weeks JR, Jokerst RS, & Greenleaf AL (1988) The C-terminal repeat domain of RNA polymerase II largest subunit is essential in vivo but is not required for accurate transcription initiation in vitro. *Proceedings of the National Academy of Sciences of the United States of America* 85(11):3698-3702.
26. Litington Y, *et al.* (1999) Growth retardation and neonatal lethality in mice with a homozygous deletion in the C-terminal domain of RNA polymerase II. *Molecular & general genetics : MGG* 261(1):100-105.
27. West ML & Corden JL (1995) Construction and analysis of yeast RNA polymerase II CTD deletion and substitution mutations. *Genetics* 140(4):1223-1233.
28. Kim WY & Dahmus ME (1989) The major late promoter of adenovirus-2 is accurately transcribed by RNA polymerases IIO, IIA, and IIB. *The Journal of biological chemistry* 264(6):3169-3176.
29. Bartolomei MS, Halden NF, Cullen CR, & Corden JL (1988) Genetic analysis of the repetitive carboxyl-terminal domain of the largest subunit of mouse RNA polymerase II. *Molecular and cellular biology* 8(1):330-339.
30. Meininghaus M, Chapman RD, Horndasch M, & Eick D (2000) Conditional expression of RNA polymerase II in mammalian cells. Deletion of the carboxyl-terminal domain of the large subunit affects early steps in transcription. *The Journal of biological chemistry* 275(32):24375-24382.
31. Corden JL (2013) RNA Polymerase II C-Terminal Domain: Tethering Transcription to Transcript and Template. *Chemical reviews*.
32. Cramer P, Bushnell DA, & Kornberg RD (2001) Structural basis of transcription: RNA polymerase II at 2.8 Å resolution. *Science* 292(5523):1863-1876.
33. Meinhardt A, Kamenski T, Hoepfner S, Baumli S, & Cramer P (2005) A structural perspective of CTD function. *Genes & development* 19(12):1401-1415.
34. Bienkiewicz EA, Moon Woody A, & Woody RW (2000) Conformation of the RNA polymerase II C-terminal domain: circular dichroism of long and short fragments. *Journal of molecular biology* 297(1):119-133.

35. Stiller JW, McConaughy BL, & Hall BD (2000) Evolutionary complementation for polymerase II CTD function. *Yeast* 16(1):57-64.
36. Zhang DW, *et al.* (2012) Ssu72 phosphatase-dependent erasure of phospho-Ser7 marks on the RNA polymerase II C-terminal domain is essential for viability and transcription termination. *The Journal of biological chemistry* 287(11):8541-8551.
37. Schwer B & Shuman S (2011) Deciphering the RNA polymerase II CTD code in fission yeast. *Molecular cell* 43(2):311-318.
38. Hintermair C, *et al.* (2012) Threonine-4 of mammalian RNA polymerase II CTD is targeted by Polo-like kinase 3 and required for transcriptional elongation. *The EMBO journal* 31(12):2784-2797.
39. Stiller JW & Cook MS (2004) Functional unit of the RNA polymerase II C-terminal domain lies within heptapeptide pairs. *Eukaryotic cell* 3(3):735-740.
40. Schwer B, Sanchez AM, & Shuman S (2012) Punctuation and syntax of the RNA polymerase II CTD code in fission yeast. *Proceedings of the National Academy of Sciences of the United States of America* 109(44):18024-18029.
41. Liu P, Greenleaf AL, & Stiller JW (2008) The essential sequence elements required for RNAP II carboxyl-terminal domain function in yeast and their evolutionary conservation. *Molecular biology and evolution* 25(4):719-727.
42. Zhang DW, Rodriguez-Molina JB, Tietjen JR, Nemecek CM, & Ansari AZ (2012) Emerging Views on the CTD Code. *Genetics research international* 2012:347214.
43. Mayer A, *et al.* (2010) Uniform transitions of the general RNA polymerase II transcription complex. *Nature structural & molecular biology* 17(10):1272-1278.
44. Mayer A, *et al.* (2012) CTD tyrosine phosphorylation impairs termination factor recruitment to RNA polymerase II. *Science* 336(6089):1723-1725.
45. Ranuncolo SM, Ghosh S, Hanover JA, Hart GW, & Lewis BA (2012) Evidence of the involvement of O-GlcNAc-modified human RNA polymerase II CTD in transcription in vitro and in vivo. *The Journal of biological chemistry* 287(28):23549-23561.
46. Shaw PE (2007) Peptidyl-prolyl cis/trans isomerases and transcription: is there a twist in the tail? *EMBO reports* 8(1):40-45.
47. Jeronimo C, Bataille AR, & Robert F (2013) The Writers, Readers, and Functions of the RNA Polymerase II C-Terminal Domain Code. *Chemical reviews*.
48. Eick D & Geyer M (2013) The RNA Polymerase II Carboxy-Terminal Domain (CTD) Code. *Chemical reviews*.
49. Kelly WG, Dahmus ME, & Hart GW (1993) RNA polymerase II is a glycoprotein. Modification of the COOH-terminal domain by O-GlcNAc. *The Journal of biological chemistry* 268(14):10416-10424.
50. Komarnitsky P, Cho EJ, & Buratowski S (2000) Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes & development* 14(19):2452-2460.
51. Kim M, Suh H, Cho EJ, & Buratowski S (2009) Phosphorylation of the yeast Rpb1 C-terminal domain at serines 2, 5, and 7. *The Journal of biological chemistry* 284(39):26421-26426.

52. Akhtar MS, *et al.* (2009) TFIIH kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II. *Molecular cell* 34(3):387-393.
53. Hengartner CJ, *et al.* (1998) Temporal regulation of RNA polymerase II by Srb10 and Kin28 cyclin-dependent kinases. *Molecular cell* 2(1):43-53.
54. Liao SM, *et al.* (1995) A kinase-cyclin pair in the RNA polymerase II holoenzyme. *Nature* 374(6518):193-196.
55. Chymkowitch P, *et al.* (2012) Cdc28 kinase activity regulates the basal transcription machinery at a subset of genes. *Proceedings of the National Academy of Sciences of the United States of America* 109(26):10450-10455.
56. Krishnamurthy S, He X, Reyes-Reyes M, Moore C, & Hampsey M (2004) Ssu72 Is an RNA polymerase II CTD phosphatase. *Molecular cell* 14(3):387-394.
57. Mosley AL, *et al.* (2009) Rtr1 is a CTD phosphatase that regulates RNA polymerase II during the transition from serine 5 to serine 2 phosphorylation. *Molecular cell* 34(2):168-178.
58. Xiang K, Manley JL, & Tong L (2012) The yeast regulator of transcription protein Rtr1 lacks an active site and phosphatase activity. *Nature communications* 3:946.
59. Cho EJ, Kobor MS, Kim M, Greenblatt J, & Buratowski S (2001) Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. *Genes & development* 15(24):3319-3329.
60. Qiu H, Hu C, & Hinnebusch AG (2009) Phosphorylation of the Pol II CTD by KIN28 enhances BUR1/BUR2 recruitment and Ser2 CTD phosphorylation near promoters. *Molecular cell* 33(6):752-762.
61. Bataille AR, *et al.* (2012) A universal RNA polymerase II CTD cycle is orchestrated by complex interplays between kinase, phosphatase, and isomerase enzymes along genes. *Molecular cell* 45(2):158-170.
62. Baskaran R, Dahmus ME, & Wang JY (1993) Tyrosine phosphorylation of mammalian RNA polymerase II carboxyl-terminal domain. *Proceedings of the National Academy of Sciences of the United States of America* 90(23):11167-11171.
63. Werner-Allen JW, *et al.* (2011) cis-Proline-mediated Ser(P)5 dephosphorylation by the RNA polymerase II C-terminal domain phosphatase Ssu72. *The Journal of biological chemistry* 286(7):5717-5726.
64. Noble CG, *et al.* (2005) Key features of the interaction between Pcf11 CID and RNA polymerase II CTD. *Nature structural & molecular biology* 12(2):144-151.
65. Meinhart A & Cramer P (2004) Recognition of RNA polymerase II carboxy-terminal domain by 3'-RNA-processing factors. *Nature* 430(6996):223-226.
66. Sims RJ, 3rd, *et al.* (2011) The C-terminal domain of RNA polymerase II is modified by site-specific methylation. *Science* 332(6025):99-103.
67. Li H, *et al.* (2007) Wwp2-mediated ubiquitination of the RNA polymerase II large subunit in mouse embryonic pluripotent stem cells. *Molecular and cellular biology* 27(15):5296-5305.

68. Corden JL (2007) Transcription. Seven ups the code. *Science* 318(5857):1735-1736.
69. Maxon ME, Goodrich JA, & Tjian R (1994) Transcription factor IIE binds preferentially to RNA polymerase IIa and recruits TFIIH: a model for promoter clearance. *Genes & development* 8(5):515-524.
70. Kang ME & Dahmus ME (1995) The photoactivated cross-linking of recombinant C-terminal domain to proteins in a HeLa cell transcription extract that comigrate with transcription factors IIE and IIF. *The Journal of biological chemistry* 270(40):23390-23397.
71. Usheva A, *et al.* (1992) Specific interaction between the nonphosphorylated form of RNA polymerase II and the TATA-binding protein. *Cell* 69(5):871-881.
72. Myers LC, *et al.* (1998) The Med proteins of yeast and their function through the RNA polymerase II carboxy-terminal domain. *Genes & development* 12(1):45-54.
73. Sogaard TM & Svejstrup JQ (2007) Hyperphosphorylation of the C-terminal repeat domain of RNA polymerase II facilitates dissociation of its complex with mediator. *The Journal of biological chemistry* 282(19):14113-14120.
74. Schroeder SC, Schwer B, Shuman S, & Bentley D (2000) Dynamic association of capping enzymes with transcribing RNA polymerase II. *Genes & development* 14(19):2435-2440.
75. Lidschreiber M, Leike K, & Cramer P (2013) Cap completion and CTD kinase recruitment underlie the initiation-elongation transition of RNA polymerase II. *Molecular and cellular biology*.
76. Schwer B, Mao X, & Shuman S (1998) Accelerated mRNA decay in conditional mutants of yeast mRNA capping enzyme. *Nucleic acids research* 26(9):2050-2057.
77. Sonenberg N & Hinnebusch AG (2009) Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* 136(4):731-745.
78. Ghosh A & Lima CD (2010) Enzymology of RNA cap synthesis. *Wiley interdisciplinary reviews. RNA* 1(1):152-172.
79. McCracken S, *et al.* (1997) 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes & development* 11(24):3306-3318.
80. Vasiljeva L, Kim M, Mutschler H, Buratowski S, & Meinhart A (2008) The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. *Nature structural & molecular biology* 15(8):795-804.
81. Ng HH, Robert F, Young RA, & Struhl K (2003) Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Molecular cell* 11(3):709-719.
82. Egloff S, *et al.* (2007) Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. *Science* 318(5857):1777-1779.
83. Hartzog GA, Wada T, Handa H, & Winston F (1998) Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in *Saccharomyces cerevisiae*. *Genes & development* 12(3):357-369.

84. Martinez-Rucobo FW, Sainsbury S, Cheung AC, & Cramer P (2011) Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. *The EMBO journal* 30(7):1302-1310.
85. Jaehning JA (2010) The Paf1 complex: platform or player in RNA polymerase II transcription? *Biochimica et biophysica acta* 1799(5-6):379-388.
86. Zhou K, Kuo WH, Fillingham J, & Greenblatt JF (2009) Control of transcriptional elongation and cotranscriptional histone modification by the yeast BUR kinase substrate Spt5. *Proceedings of the National Academy of Sciences of the United States of America* 106(17):6956-6961.
87. Liu Y, *et al.* (2009) Phosphorylation of the transcription elongation factor Spt5 by yeast Bur1 kinase stimulates recruitment of the PAF complex. *Molecular and cellular biology* 29(17):4852-4863.
88. Yoh SM, Cho H, Pickle L, Evans RM, & Jones KA (2007) The Spt6 SH2 domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export. *Genes & development* 21(2):160-174.
89. Kaplan CD, Laprade L, & Winston F (2003) Transcription elongation factors repress transcription initiation from cryptic sites. *Science* 301(5636):1096-1099.
90. Belotserkovskaya R, *et al.* (2003) FACT facilitates transcription-dependent nucleosome alteration. *Science* 301(5636):1090-1093.
91. Kizer KO, *et al.* (2005) A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Molecular and cellular biology* 25(8):3305-3316.
92. Strahl BD & Allis CD (2000) The language of covalent histone modifications. *Nature* 403(6765):41-45.
93. MacKellar AL & Greenleaf AL (2011) Cotranscriptional association of mRNA export factor Yra1 with C-terminal domain of RNA polymerase II. *The Journal of biological chemistry* 286(42):36385-36395.
94. Strasser K & Hurt E (2000) Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. *The EMBO journal* 19(3):410-420.
95. Morris DP, Phatnani HP, & Greenleaf AL (1999) Phospho-carboxyl-terminal domain binding and the role of a prolyl isomerase in pre-mRNA 3'-End formation. *The Journal of biological chemistry* 274(44):31583-31587.
96. Morris DP & Greenleaf AL (2000) The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II. *The Journal of biological chemistry* 275(51):39935-39943.
97. Will CL & Luhrmann R (2011) Spliceosome structure and function. *Cold Spring Harbor perspectives in biology* 3(7).
98. Neugeglise C, Marck C, & Gaillardin C (2011) The intronome of budding yeasts. *Comptes rendus biologies* 334(8-9):662-670.
99. Jurica MS & Moore MJ (2003) Pre-mRNA splicing: awash in a sea of proteins. *Molecular cell* 12(1):5-14.
100. Kettenberger H, Armache KJ, & Cramer P (2003) Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell* 114(3):347-357.

101. Klein BJ, *et al.* (2011) RNA polymerase and transcription elongation factor Spt4/5 complex structure. *Proceedings of the National Academy of Sciences of the United States of America* 108(2):546-550.
102. Winston F, Chaleff DT, Valent B, & Fink GR (1984) Mutations affecting Ty-mediated expression of the HIS4 gene of *Saccharomyces cerevisiae*. *Genetics* 107(2):179-197.
103. Swanson MS, Malone EA, & Winston F (1991) SPT5, an essential gene important for normal transcription in *Saccharomyces cerevisiae*, encodes an acidic nuclear protein with a carboxy-terminal repeat. *Molecular and cellular biology* 11(8):4286.
104. Wada T, *et al.* (1998) DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes & development* 12(3):343-356.
105. Kim M, Ahn SH, Krogan NJ, Greenblatt JF, & Buratowski S (2004) Transitions in RNA polymerase II elongation complexes at the 3' ends of genes. *The EMBO journal* 23(2):354-364.
106. Pokholok DK, Hannett NM, & Young RA (2002) Exchange of RNA polymerase II initiation and elongation factors during gene expression in vivo. *Molecular cell* 9(4):799-809.
107. Venters BJ, *et al.* (2011) A comprehensive genomic binding map of gene and chromatin regulatory proteins in *Saccharomyces*. *Molecular cell* 41(4):480-492.
108. Viktorovskaya OV, Appling FD, & Schneider DA (2011) Yeast transcription elongation factor Spt5 associates with RNA polymerase I and RNA polymerase II directly. *The Journal of biological chemistry* 286(21):18825-18833.
109. Schneider DA, *et al.* (2006) RNA polymerase II elongation factors Spt4p and Spt5p play roles in transcription elongation by RNA polymerase I and rRNA processing. *Proceedings of the National Academy of Sciences of the United States of America* 103(34):12707-12712.
110. Werner F & Grohmann D (2011) Evolution of multisubunit RNA polymerases in the three domains of life. *Nature reviews. Microbiology* 9(2):85-98.
111. Knowlton JR, *et al.* (2003) A spring-loaded state of NusG in its functional cycle is suggested by X-ray crystallography and supported by site-directed mutants. *Biochemistry* 42(8):2275-2281.
112. Mooney RA, Schweimer K, Rosch P, Gottesman M, & Landick R (2009) Two structurally independent domains of *E. coli* NusG create regulatory plasticity via distinct interactions with RNA polymerase and regulators. *Journal of molecular biology* 391(2):341-358.
113. Hirtreiter A, *et al.* (2010) Spt4/5 stimulates transcription elongation through the RNA polymerase clamp coiled-coil motif. *Nucleic acids research* 38(12):4040-4051.
114. Ponting CP (2002) Novel domains and orthologues of eukaryotic transcription elongation factors. *Nucleic acids research* 30(17):3643-3652.
115. Lindstrom DL, *et al.* (2003) Dual roles for Spt5 in pre-mRNA processing and transcription elongation revealed by identification of Spt5-associated proteins. *Molecular and cellular biology* 23(4):1368-1378.
116. Swanson MS & Winston F (1992) SPT4, SPT5 and SPT6 interactions: effects on transcription and viability in *Saccharomyces cerevisiae*. *Genetics* 132(2):325-336.

117. Andrulis ED, Guzman E, Doring P, Werner J, & Lis JT (2000) High-resolution localization of *Drosophila* Spt5 and Spt6 at heat shock genes in vivo: roles in promoter proximal pausing and transcription elongation. *Genes & development* 14(20):2635-2649.
118. Kaplan CD, Morris JR, Wu C, & Winston F (2000) Spt5 and spt6 are associated with active transcription and have characteristics of general elongation factors in *D. melanogaster*. *Genes & development* 14(20):2623-2634.
119. Pei Y & Shuman S (2002) Interactions between fission yeast mRNA capping enzymes and elongation factor Spt5. *The Journal of biological chemistry* 277(22):19639-19648.
120. Wen Y & Shatkin AJ (1999) Transcription elongation factor hSPT5 stimulates mRNA capping. *Genes & development* 13(14):1774-1779.
121. Pavri R, *et al.* (2010) Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell* 143(1):122-133.
122. Shen Z, St-Denis A, & Chartrand P (2010) Cotranscriptional recruitment of She2p by RNA pol II elongation factor Spt4-Spt5/DSIF promotes mRNA localization to the yeast bud. *Genes & development* 24(17):1914-1926.
123. Schneider S, Pei Y, Shuman S, & Schwer B (2010) Separable functions of the fission yeast Spt5 carboxyl-terminal domain (CTD) in capping enzyme binding and transcription elongation overlap with those of the RNA polymerase II CTD. *Molecular and cellular biology* 30(10):2353-2364.
124. Yamada T, *et al.* (2006) P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. *Molecular cell* 21(2):227-237.
125. Drouin S, *et al.* (2010) DSIF and RNA polymerase II CTD phosphorylation coordinate the recruitment of Rpd3S to actively transcribed genes. *PLoS genetics* 6(10):e1001173.
126. Ding B, LeJeune D, & Li S (2010) The C-terminal repeat domain of Spt5 plays an important role in suppression of Rad26-independent transcription coupled repair. *The Journal of biological chemistry* 285(8):5317-5326.
127. Chen H, *et al.* (2009) Repression of RNA polymerase II elongation in vivo is critically dependent on the C-terminus of Spt5. *PLoS one* 4(9):e6918.
128. Mischo HE & Proudfoot NJ (2013) Disengaging polymerase: terminating RNA polymerase II transcription in budding yeast. *Biochimica et biophysica acta* 1829(1):174-185.
129. Buratowski S (2005) Connections between mRNA 3' end processing and transcription termination. *Current opinion in cell biology* 17(3):257-261.
130. Logan J, Falck-Pedersen E, Darnell JE, Jr., & Shenk T (1987) A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse beta maj-globin gene. *Proceedings of the National Academy of Sciences of the United States of America* 84(23):8306-8310.
131. Sadowski M, Dichtl B, Hubner W, & Keller W (2003) Independent functions of yeast Pcf11p in pre-mRNA 3' end processing and in transcription termination. *The EMBO journal* 22(9):2167-2177.
132. Kim M, *et al.* (2004) The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* 432(7016):517-522.

133. Connelly S & Manley JL (1988) A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes & Development* 2(4):440-452.
134. Proudfoot NJ (2011) Ending the message: poly(A) signals then and now. *Genes & development* 25(17):1770-1782.
135. West S, Proudfoot NJ, & Dye MJ (2008) Molecular dissection of mammalian RNA polymerase II transcriptional termination. *Molecular cell* 29(5):600-610.
136. Bentley DL (2005) Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Current opinion in cell biology* 17(3):251-256.
137. Mandel CR, Bai Y, & Tong L (2008) Protein factors in pre-mRNA 3'-end processing. *Cellular and molecular life sciences : CMLS* 65(7-8):1099-1122.
138. Gross S & Moore C (2001) Five subunits are required for reconstitution of the cleavage and polyadenylation activities of *Saccharomyces cerevisiae* cleavage factor I. *Proceedings of the National Academy of Sciences of the United States of America* 98(11):6080-6085.
139. Kessler MM, Zhao J, & Moore CL (1996) Purification of the *Saccharomyces cerevisiae* cleavage/polyadenylation factor I. Separation into two components that are required for both cleavage and polyadenylation of mRNA 3' ends. *The Journal of biological chemistry* 271(43):27167-27175.
140. Amrani N, *et al.* (1997) PCF11 encodes a third protein component of yeast cleavage and polyadenylation factor I. *Molecular and cellular biology* 17(3):1102-1109.
141. Minvielle-Sebastia L, Preker PJ, & Keller W (1994) RNA14 and RNA15 proteins as components of a yeast pre-mRNA 3'-end processing factor. *Science* 266(5191):1702-1705.
142. Minvielle-Sebastia L, Preker PJ, Wiederkehr T, Strahm Y, & Keller W (1997) The major yeast poly(A)-binding protein is associated with cleavage factor IA and functions in premessenger RNA 3'-end formation. *Proceedings of the National Academy of Sciences of the United States of America* 94(15):7897-7902.
143. Kessler MM, *et al.* (1997) Hrp1, a sequence-specific RNA-binding protein that shuttles between the nucleus and the cytoplasm, is required for mRNA 3'-end formation in yeast. *Genes & development* 11(19):2545-2556.
144. Minvielle-Sebastia L, *et al.* (1998) Control of cleavage site selection during mRNA 3' end formation by a yeast hnRNP. *The EMBO journal* 17(24):7454-7468.
145. Licatalosi DD, *et al.* (2002) Functional interaction of yeast pre-mRNA 3' end processing factors with RNA polymerase II. *Molecular cell* 9(5):1101-1111.
146. Ahn SH, Kim M, & Buratowski S (2004) Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. *Molecular cell* 13(1):67-76.
147. Zhang Z, Fu J, & Gilmour DS (2005) CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3'-end processing factor, Pcf11. *Genes & development* 19(13):1572-1580.
148. Barilla D, Lee BA, & Proudfoot NJ (2001) Cleavage/polyadenylation factor IA associates with the carboxyl-terminal domain of RNA polymerase II in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America* 98(2):445-450.

149. Kyburz A, Sadowski M, Dichtl B, & Keller W (2003) The role of the yeast cleavage and polyadenylation factor subunit Ydh1p/Cft2p in pre-mRNA 3'-end formation. *Nucleic acids research* 31(14):3936-3945.
150. Dichtl B, *et al.* (2002) Yhh1p/Cft1p directly links poly(A) site recognition and RNA polymerase II transcription termination. *The EMBO journal* 21(15):4125-4135.
151. Rodriguez CR, *et al.* (2000) Kin28, the TFIIH-associated carboxy-terminal domain kinase, facilitates the recruitment of mRNA processing machinery to RNA polymerase II. *Molecular and cellular biology* 20(1):104-112.
152. Moore MJ (2005) From birth to death: the complex lives of eukaryotic mRNAs. *Science* 309(5740):1514-1518.
153. Nedeá E, *et al.* (2003) Organization and function of APT, a subcomplex of the yeast cleavage and polyadenylation factor involved in the formation of mRNA and small nucleolar RNA 3'-ends. *The Journal of biological chemistry* 278(35):33000-33010.
154. Hausmann S, Erdjument-Bromage H, & Shuman S (2004) Schizosaccharomyces pombe carboxyl-terminal domain (CTD) phosphatase Fcp1: distributive mechanism, minimal CTD substrate, and active site mapping. *The Journal of biological chemistry* 279(12):10892-10900.
155. Ghosh A, Shuman S, & Lima CD (2008) The structure of Fcp1, an essential RNA polymerase II CTD phosphatase. *Molecular cell* 32(4):478-490.
156. Cho H, *et al.* (1999) A protein phosphatase functions to recycle RNA polymerase II. *Genes & development* 13(12):1540-1552.
157. Baskaran R, Escobar SR, & Wang JY (1999) Nuclear c-Abl is a COOH-terminal repeated domain (CTD)-tyrosine (CTD)-tyrosine kinase-specific for the mammalian RNA polymerase II: possible role in transcription elongation. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research* 10(6):387-396.
158. Kim H, *et al.* (2010) Gene-specific RNA polymerase II phosphorylation and the CTD code. *Nature structural & molecular biology* 17(10):1279-1286.
159. Creamer TJ, *et al.* (2011) Transcriptome-wide binding sites for components of the Saccharomyces cerevisiae non-poly(A) termination pathway: Nrd1, Nab3, and Sen1. *PLoS genetics* 7(10):e1002329.
160. Kim KY & Levin DE (2011) Mpk1 MAPK association with the Paf1 complex blocks Sen1-mediated premature transcription termination. *Cell* 144(5):745-756.
161. Steinmetz EJ, Conrad NK, Brow DA, & Corden JL (2001) RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. *Nature* 413(6853):327-331.
162. Rondon AG, Mischo HE, Kawauchi J, & Proudfoot NJ (2009) Fail-safe transcriptional termination for protein-coding genes in S. cerevisiae. *Molecular cell* 36(1):88-98.
163. Lunde BM, *et al.* (2010) Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. *Nature structural & molecular biology* 17(10):1195-1201.
164. Buratowski S (2003) The CTD code. *Nature structural biology* 10(9):679-680.

-
165. Egloff S & Murphy S (2008) Cracking the RNA polymerase II CTD code. *Trends in genetics : TIG* 24(6):280-288.
166. Aparicio O, *et al.* (2005) Chromatin immunoprecipitation for determining the association of proteins with specific genomic sequences in vivo. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.]* Chapter 21:Unit 21 23.
167. Lindberg RA, Quinn AM, & Hunter T (1992) Dual-specificity protein kinases: will any hydroxyl do? *Trends in biochemical sciences* 17(3):114-119.
168. Breitkreutz A, *et al.* (2010) A global protein kinase and phosphatase interaction network in yeast. *Science* 328(5981):1043-1046.
169. Manning G, Plowman GD, Hunter T, & Sudarsanam S (2002) Evolution of protein kinase signaling from yeast to man. *Trends in biochemical sciences* 27(10):514-520.
170. Hunter T & Plowman GD (1997) The protein kinases of budding yeast: six score and more. *Trends in biochemical sciences* 22(1):18-22.
171. Ball CA, *et al.* (2000) Integrating functional genomic information into the *Saccharomyces* genome database. *Nucleic acids research* 28(1):77-80.
172. Bishop AC, Buzko O, & Shokat KM (2001) Magic bullets for protein kinases. *Trends in cell biology* 11(4):167-172.
173. Knight ZA & Shokat KM (2005) Features of selective kinase inhibitors. *Chemistry & biology* 12(6):621-637.
174. Lee KS & Levin DE (1992) Dominant mutations in a gene encoding a putative protein kinase (BCK1) bypass the requirement for a *Saccharomyces cerevisiae* protein kinase C homolog. *Molecular and cellular biology* 12(1):172-182.
175. Lee KS, *et al.* (1993) A yeast mitogen-activated protein kinase homolog (Mpk1p) mediates signalling by protein kinase C. *Molecular and cellular biology* 13(5):3067-3075.
176. Milgrom E, West RW, Jr., Gao C, & Shen WC (2005) TFIID and Spt-Ada-Gcn5-acetyltransferase functions probed by genome-wide synthetic genetic array analysis using a *Saccharomyces cerevisiae* taf9-ts allele. *Genetics* 171(3):959-973.
177. Fiedler D, *et al.* (2009) Functional organization of the *S. cerevisiae* phosphorylation network. *Cell* 136(5):952-963.
178. Chua P & Roeder GS (1995) Bdf1, a yeast chromosomal protein required for sporulation. *Molecular and cellular biology* 15(7):3685-3696.
179. Lygerou Z, *et al.* (1994) The yeast BDF1 gene encodes a transcription factor involved in the expression of a broad class of genes including snRNAs. *Nucleic acids research* 22(24):5332-5340.
180. Devaiah BN, *et al.* (2012) BRD4 is an atypical kinase that phosphorylates serine2 of the RNA polymerase II carboxy-terminal domain. *Proceedings of the National Academy of Sciences of the United States of America* 109(18):6927-6932.

181. Alexandru G, Uhlmann F, Mechtler K, Poupart MA, & Nasmyth K (2001) Phosphorylation of the cohesin subunit Scc1 by Polo/Cdc5 kinase regulates sister chromatid separation in yeast. *Cell* 105(4):459-472.
182. Santos-Rosa H, *et al.* (1998) Nuclear mRNA export requires complex formation between Mex67p and Mtr2p at the nuclear pores. *Molecular and cellular biology* 18(11):6826-6838.
183. Mendenhall MD & Hodge AE (1998) Regulation of Cdc28 cyclin-dependent protein kinase activity during the cell cycle of the yeast *Saccharomyces cerevisiae*. *Microbiology and molecular biology reviews : MMBR* 62(4):1191-1243.
184. Boulton TG, *et al.* (1990) An insulin-stimulated protein kinase similar to yeast kinases involved in cell cycle control. *Science* 249(4964):64-67.
185. Ptacek J, *et al.* (2005) Global analysis of protein phosphorylation in yeast. *Nature* 438(7068):679-684.
186. Mehlgarten C & Schaffrath R (2003) Mutant casein kinase I (Hrr25p/Kti14p) abrogates the G1 cell cycle arrest induced by *Kluyveromyces lactis* zymocin in budding yeast. *Molecular genetics and genomics : MGG* 269(2):188-196.
187. Milne DM, Looby P, & Meek DW (2001) Catalytic activity of protein kinase CK1 delta (casein kinase I delta) is essential for its normal subcellular localization. *Experimental cell research* 263(1):43-54.
188. Phatnani HP, Jones JC, & Greenleaf AL (2004) Expanding the functional repertoire of CTD kinase I and RNA polymerase II: novel phosphoCTD-associating proteins in the yeast proteome. *Biochemistry* 43(50):15702-15719.
189. Lee K, Du C, Horn M, & Rabinow L (1996) Activity and autophosphorylation of LAMMER protein kinases. *The Journal of biological chemistry* 271(44):27299-27303.
190. Costanzo M, *et al.* (2010) The genetic landscape of a cell. *Science* 327(5964):425-431.
191. Lim MY, Dailey D, Martin GS, & Thorner J (1993) Yeast MCK1 protein kinase autophosphorylates at tyrosine and serine but phosphorylates exogenous substrates at serine and threonine. *The Journal of biological chemistry* 268(28):21155-21164.
192. Bianchi MW, Plyte SE, Kreis M, & Woodgett JR (1993) A *Saccharomyces cerevisiae* protein-serine kinase related to mammalian glycogen synthase kinase-3 and the *Drosophila melanogaster* gene shaggy product. *Gene* 134(1):51-56.
193. Collins SR, *et al.* (2007) Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature* 446(7137):806-810.
194. Maeda T, Takekawa M, & Saito H (1995) Activation of yeast PBS2 MAPKK by MAPKKKs or by binding of an SH3-containing osmosensor. *Science* 269(5223):554-558.
195. Lo WS, *et al.* (2001) Snf1--a histone kinase that works in concert with the histone acetyltransferase Gcn5 to regulate transcription. *Science* 293(5532):1142-1146.
196. Liu Y, Xu X, Singh-Rodriguez S, Zhao Y, & Kuo MH (2005) Histone H3 Ser10 phosphorylation-independent function of Snf1 and Reg1 proteins rescues a gcn5- mutant in HIS3 expression. *Molecular and cellular biology* 25(23):10566-10579.

197. Booher RN, Deshaies RJ, & Kirschner MW (1993) Properties of *Saccharomyces cerevisiae* wee1 and its differential regulation of p34CDC28 in response to G1 and G2 cyclins. *The EMBO journal* 12(9):3417-3426.
198. Traven A, *et al.* (2009) The Ccr4-Pop2-NOT mRNA deadenylase contributes to septin organization in *Saccharomyces cerevisiae*. *Genetics* 182(4):955-966.
199. Broach JR (1991) RAS genes in *Saccharomyces cerevisiae*: signal transduction in search of a pathway. *Trends in genetics : TIG* 7(1):28-33.
200. Santangelo GM (2006) Glucose signaling in *Saccharomyces cerevisiae*. *Microbiology and molecular biology reviews : MMBR* 70(1):253-282.
201. Moriya H, *et al.* (2001) Yak1p, a DYRK family kinase, translocates to the nucleus and phosphorylates yeast Pop2p in response to a glucose signal. *Genes & development* 15(10):1217-1228.
202. Fasolo J, *et al.* (2011) Diverse protein kinase interactions identified by protein microarrays reveal novel connections between cellular processes. *Genes & development* 25(7):767-778.
203. Haruki H, Nishikawa J, & Laemmli UK (2008) The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes. *Molecular cell* 31(6):925-932.
204. Meinhart A, Silberzahn T, & Cramer P (2003) The mRNA transcription/processing factor Ssu72 is a potential tyrosine phosphatase. *The Journal of biological chemistry* 278(18):15917-15921.
205. Xiang K, *et al.* (2010) Crystal structure of the human symplekin-Ssu72-CTD phosphopeptide complex. *Nature* 467(7316):729-733.
206. Nedeá E, *et al.* (2008) The Glc7 phosphatase subunit of the cleavage and polyadenylation factor is essential for transcription termination on snRNA genes. *Molecular cell* 29(5):577-587.
207. Shi Y (2009) Serine/threonine phosphatases: mechanism through structure. *Cell* 139(3):468-484.
208. Chu Y, Lee EY, & Schlender KK (1996) Activation of protein phosphatase 1. Formation of a metalloenzyme. *The Journal of biological chemistry* 271(5):2574-2577.
209. Egloff MP, Cohen PT, Reinemer P, & Barford D (1995) Crystal structure of the catalytic subunit of human protein phosphatase 1 and its complex with tungstate. *Journal of molecular biology* 254(5):942-959.
210. Huang HB, Horiuchi A, Goldberg J, Greengard P, & Nairn AC (1997) Site-directed mutagenesis of amino acid residues of protein phosphatase 1 involved in catalysis and inhibitor binding. *Proceedings of the National Academy of Sciences of the United States of America* 94(8):3530-3535.
211. Peti W, Nairn AC, & Page R (2013) Structural basis for protein phosphatase 1 regulation and specificity. *The FEBS journal* 280(2):596-611.
212. MacKintosh C, *et al.* (1996) Further evidence that inhibitor-2 acts like a chaperone to fold PP1 into its native conformation. *FEBS letters* 397(2-3):235-238.
213. Farkas I, Dombradi V, Miskei M, Szabados L, & Koncz C (2007) Arabidopsis PPP family of serine/threonine phosphatases. *Trends in plant science* 12(4):169-176.

-
214. Kim TW, *et al.* (2009) Brassinosteroid signal transduction from cell-surface receptor kinases to nuclear transcription factors. *Nature cell biology* 11(10):1254-1260.
215. Chao Y, *et al.* (2006) Structure and mechanism of the phosphotyrosyl phosphatase activator. *Molecular cell* 23(4):535-546.
216. Mayer A, *et al.* (2012) The spt5 C-terminal region recruits yeast 3' RNA cleavage factor I. *Molecular and cellular biology* 32(7):1321-1331.
217. Dengl S, Mayer A, Sun M, & Cramer P (2009) Structure and in vivo requirement of the yeast Spt6 SH2 domain. *Journal of molecular biology* 389(1):211-225.
218. Millevoi S & Vagner S (2010) Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. *Nucleic acids research* 38(9):2757-2774.
219. Leeper TC, Qu X, Lu C, Moore C, & Varani G (2010) Novel protein-protein contacts facilitate mRNA 3'-processing signal recognition by Rna15 and Hrp1. *Journal of molecular biology* 401(3):334-349.
220. Pancevac C, Goldstone DC, Ramos A, & Taylor IA (2010) Structure of the Rna15 RRM-RNA complex reveals the molecular basis of GU specificity in transcriptional 3'-end processing factors. *Nucleic acids research* 38(9):3119-3132.
221. Perez-Canadillas JM (2006) Grabbing the message: structural basis of mRNA 3'UTR recognition by Hrp1. *The EMBO journal* 25(13):3167-3178.
222. Abruzzi KC, Lacadie S, & Rosbash M (2004) Biochemical analysis of TREX complex recruitment to intronless and intron-containing yeast genes. *The EMBO journal* 23(13):2620-2631.
223. Glover-Cutter K, Kim S, Espinosa J, & Bentley DL (2008) RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. *Nature structural & molecular biology* 15(1):71-78.
224. Swinburne IA, Meyer CA, Liu XS, Silver PA, & Brodsky AS (2006) Genomic localization of RNA binding proteins reveals links between pre-mRNA processing and transcription. *Genome research* 16(7):912-921.
225. Aranda A & Proudfoot N (2001) Transcriptional termination factors for RNA polymerase II in yeast. *Molecular cell* 7(5):1003-1011.
226. Birse CE (1998) Coupling Termination of Transcription to Messenger RNA Maturation in Yeast. *Science* 280(5361):298-301.
227. Proudfoot N (2004) New perspectives on connecting messenger RNA 3' end formation to transcription. *Current opinion in cell biology* 16(3):272-278.
228. Mandart E & Parker R (1995) Effects of mutations in the *Saccharomyces cerevisiae* RNA14, RNA15, and PAP1 genes on polyadenylation in vivo. *Molecular and cellular biology* 15(12):6979-6986.
229. Frohman MA, Dush MK, & Martin GR (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proceedings of the National Academy of Sciences of the United States of America* 85(23):8998-9002.
230. Ahn SH, Keogh MC, & Buratowski S (2009) Ctk1 promotes dissociation of basal transcription factors from elongating RNA polymerase II. *The EMBO journal* 28(3):205-212.

-
231. Gao L & Gross DS (2008) Sir2 silences gene transcription by targeting the transition between RNA polymerase II initiation and elongation. *Molecular and cellular biology* 28(12):3979-3994.
232. Larson DR, Zenklusen D, Wu B, Chao JA, & Singer RH (2011) Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science* 332(6028):475-478.
233. Mueller CL, Porter SE, Hoffman MG, & Jaehning JA (2004) The Paf1 complex has functions independent of actively transcribing RNA polymerase II. *Molecular cell* 14(4):447-456.
234. Nordick K, Hoffman MG, Betz JL, & Jaehning JA (2008) Direct interactions between the Paf1 complex and a cleavage and polyadenylation factor are revealed by dissociation of Paf1 from RNA polymerase II. *Eukaryotic cell* 7(7):1158-1167.
235. Kaplan CD, Holland MJ, & Winston F (2005) Interaction between transcription elongation factors and mRNA 3'-end formation at the *Saccharomyces cerevisiae* GAL10-GAL7 locus. *The Journal of biological chemistry* 280(2):913-922.
236. Hershkovits G, Bangio H, Cohen R, & Katcoff DJ (2006) Recruitment of mRNA cleavage/polyadenylation machinery by the yeast chromatin protein Sin1p/Spt2p. *Proceedings of the National Academy of Sciences of the United States of America* 103(26):9808-9813.
237. Dermody JL & Buratowski S (2010) Leo1 subunit of the yeast paf1 complex binds RNA and contributes to complex recruitment. *The Journal of biological chemistry* 285(44):33671-33679.
238. Burmann BM, *et al.* (2010) A NusE:NusG complex links transcription and translation. *Science* 328(5977):501-504.
239. Proshkin S, Rahmouni AR, Mironov A, & Nudler E (2010) Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* 328(5977):504-508.
240. Hocine S, Singer RH, & Grunwald D (2010) RNA processing and export. *Cold Spring Harbor perspectives in biology* 2(12):a000752.
241. Martin W & Koonin EV (2006) Introns and the origin of nucleus-cytosol compartmentalization. *Nature* 440(7080):41-45.
242. Johnson SA, Cubberley G, & Bentley DL (2009) Cotranscriptional recruitment of the mRNA export factor Yral by direct interaction with the 3' end processing factor Pcf11. *Molecular cell* 33(2):215-226.
243. Luna R, Gaillard H, Gonzalez-Aguilera C, & Aguilera A (2008) Biogenesis of mRNPs: integrating different processes in the eukaryotic nucleus. *Chromosoma* 117(4):319-331.
244. Qu X, *et al.* (2009) Assembly of an export-competent mRNP is needed for efficient release of the 3'-end processing complex after polyadenylation. *Molecular and cellular biology* 29(19):5327-5338.
245. Park PJ (2009) ChIP-seq: advantages and challenges of a maturing technology. *Nature reviews. Genetics* 10(10):669-680.
246. Brugiolo M, Herzel L, & Neugebauer KM (2013) Counting on co-transcriptional splicing. *F1000prime reports* 5:9.
247. Sakharkar MK, Chow VT, & Kanguane P (2004) Distributions of exons and introns in the human genome. *In silico biology* 4(4):387-393.

-
248. Hafner M, *et al.* (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 141(1):129-141.
249. Bahler J, *et al.* (1998) Heterologous modules for efficient and versatile PCR-based gene targeting in *Schizosaccharomyces pombe*. *Yeast* 14(10):943-951.
250. Passmore LA, *et al.* (2003) Doc1 mediates the activity of the anaphase-promoting complex by contributing to substrate recognition. *The EMBO journal* 22(4):786-796.
251. 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(1):19-27.
252. Longtine MS, *et al.* (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* 14(10):953-961.
253. Sydow JF, *et al.* (2009) Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA. *Molecular cell* 34(6):710-721.
254. Fan X, Lamarre-Vincent N, Wang Q, & Struhl K (2008) Extensive chromatin fragmentation improves enrichment of protein binding sites in chromatin immunoprecipitation experiments. *Nucleic acids research* 36(19):e125.
255. RDevelopmentCoreTeam (2009) R: a language and environment for statistical computing. *Vienna, Austria*.
256. Gentleman RC, *et al.* (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome biology* 5(10):R80.
257. Zacher B, Kuan PF, & Tresch A (2010) Starr: Simple Tiling ARRAy analysis of Affymetrix ChIP-chip data. *BMC bioinformatics* 11:194.
258. Toedling J, *et al.* (2007) Ringo--an R/Bioconductor package for analyzing ChIP-chip readouts. *BMC bioinformatics* 8:221.

List of figures

Figure 1: Pol II CTD sequences in <i>S. cerevisiae</i> , <i>D. melanogaster</i> and <i>H. sapiens</i>	3
Figure 2: Model of the yeast Pol II CTD repeats and possible posttranslational modifications of a consensus repeat.	4
Figure 3: Schematic overview of the Pol II transcription cycle in <i>S. cerevisiae</i>	8
Figure 4: Schematic overview of Spt5 domain architecture (84, 86).	10
Figure 5: Pol II CTD is phosphorylated at Tyr1 (44).	13
Figure 6: Gene-averaged ChIP profiles for CTD phosphorylations and termination factors (44).	14
Figure 7: Extended CTD code for transcription cycle coordination (44).	16
Figure 8: CTD Tyr1 phosphorylation levels are not changed upon inhibition of the canonical Pol II CTD kinases.	20
Figure 9: Tyr1-P occupancy levels are not decreased by deletion of 12 non-essential kinases.	23
Figure 10: Growth analysis of anchor-away yeast strains.	24
Figure 11: Fluorescence microscopy of the Hrr25-FRB-GFP strain.	24
Figure 12: Tyr1-P levels in Cdc5, Cdk1 and Hrr25 anchor-away strains are not influenced by rapamycin treatment.	25
Figure 13: CPF subunit Glc7 is a Pol II CTD Tyr1 phosphatase <i>in vitro</i>	28
Figure 14: Genome-wide ChIP occupancy of Tyr1-phosphorylated Pol II around the pA site is not influenced by rapamycin in wild-type yeast.	29
Figure 15: Glc7 is required for Tyr1 dephosphorylation <i>in vivo</i>	30
Figure 16: Depletion of Glc7 from the nucleus leads to a defect in Tyr1-P dephosphorylation.	31
Figure 17: Growth analysis and fluorescence microscopy of anchor-away yeast strains.	32
Figure 18: The catalytic activity of Glc7 in the nucleus is essential for cell viability and Ssu72 does not change Tyr1 phosphorylation <i>in vitro</i> and <i>in vivo</i>	33
Figure 19: Tyr1 dephosphorylation by Glc7 is required for full termination factor recruitment and transcription termination <i>in vivo</i>	35
Figure 20: Spt5 CTR deletion leads to a growth defect in the presence of 6-AU and to a slight increase in Spt5 protein levels.	38

Figure 21: Apart from Paf1 (Figure 22c) no other Pol II transcription elongation factors are affected by the deletion of the Spt5 CTR.	39
Figure 22: ChIP analysis reveals that CFI occupancy is reduced in Spt5 Δ CTR cells.....	40
Figure 23: Rna15 ChIP occupancy levels at <i>ILV5</i> , <i>PDC1</i> and <i>TEF1</i> genes are also reduced in Spt5 Δ CTR cells.....	41
Figure 24: The Spt5 CTR interacts with CFI subunits <i>in vitro</i>	42
Figure 25: RNase-ChIP assays reveal that RNA contributes to CFI recruitment in Spt5 Δ CTR cells.....	43
Figure 26: Genome-wide ChIP-chip occupancy profiling of CFI subunits, Ser2-P Pol II and Spt5 in yeast.	44
Figure 27: Genome-wide ChIP-chip occupancy profiling of CFI subunits in yeast for short and medium length genes.	44
Figure 28: Spt5 CTR deletion provokes neither transcriptional readthrough of Pol II nor alternative pA site usage.	46
Figure 29: Model of CFI recruitment in yeast.....	49

List of tables

Table 1: Known <i>in vivo</i> CTD kinases and phosphatases in <i>S. cerevisiae</i> (adapted from (47)).	5
Table 2: Important Pol II CTD readers in <i>S. cerevisiae</i> (adapted from (42)).	6
Table 3: 7 kinase groups in <i>S. cerevisiae</i> (adapted from (169) and http://www.kinase.com).	18
Table 4: Tyr1 kinase candidates.	21
Table 5: <i>E. coli</i> strains.	55
Table 6: <i>S. cerevisiae</i> strains.	55
Table 7: Growth media.	57
Table 8: Growth media additives.	58
Table 9: Plasmids used in this study.	58
Table 10: Cloning primers used in this study.	59
Table 11: Readthrough assay and 3'-RACE primers.	60
Table 12: qPCR primers used in this study.	61
Table 13: Antibodies used in this study.	62
Table 14: Buffers and solutions used in this study.	63

Abbreviations

3'-RACE	rapid amplification of 3'-ends
6-AU	6-azauracil
A	alanine, adenine
Ala	alanine
APE1	apurinic endonuclease 1
as	analog-sensitive
ATP	adenosine 5'-triphosphate
Bp	base pairs
BSA	bovine serum albumin
C	cytosine
CARM1	coactivator-associated arginine methyltransferase 1
CF	cleavage factor
ChIP	chromatin immunoprecipitation
CID	CTD interacting domain
CPF	cleavage and polyadenylation factor
CTD	C-terminal domain of Rpb1 (Pol II subunit)
CTR	C-terminal repeats of Spt5
DAPI	4',6-Diamidin-2-phenylindol
<i>D. melanogaster</i>	<i>Drosophila melanogaster</i>
DMSO	dimethyl sulfoxide
dNTP	deoxynucleosin 5'-triphosphate
DTT	dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
Fw	forward
G	guanine, glycine
GFP	green fluorescent protein
Glu	glutamate
GST	Glutathion-S-Transferase
GTF	general transcription factor
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	horseradish peroxidase
<i>H. sapiens</i>	<i>Homo sapiens</i>
Ig(G)	immunoglobulin (G)
IP	immunoprecipitation
IPTG	Isopropyl- β -D-thiogalactopyranosid
K	lysine
kDa	kilodalton
KOW motif	Kyrpides/Ouzounis/Woese motif
LB	Luria-Bertani media
m7G	7-methylguanosine
MAPKK	MAP kinase kinase
MES	2-(N-morpholino)ethanesulfonic acid
MOPS	3-(N-morpholino)propanesulfonic acid
MPA	mycophenolic acid
mRNA	messenger RNA
MS	mass spectrometry
N	asparagine
NGN domain	NusG N-terminal homology domain
NP40	nonyl phenoxypropylpolyethoxyethanol
OD ₆₀₀	optical density at 600 nm wavelength
OAc	acetate
O-GlcNAc	O-linked N-acetylglucosamine
ORF	open reading frame

pA	polyadenylation
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
Phe	phenylalanine
PI	protease inhibitors
PIC	preinitiation complex
PMSF	phenylmethylsulfonylfluorid
Pol II	RNA polymerase II
PP1	protein phosphatase 1
Pro/P	proline
PVDF	polyvinylidene fluorid
Q	glutamine
qPCR	quantitative real-time PCR
R	arginine, Pearson correlation coefficient
Rapa	rapamycin
Rev	reverse
Rpm	rotations per minute
RRM	RNA recognition motif
rRNA	ribosomal RNA
Ser/S	serine
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
<i>S. pombe</i>	<i>Schizosaccharomyces pombe</i>
Ser	serine
Ser2-P/S2P	phosphorylated serine residue in position 2 of CTD repeats
Ser5-P/S5P	phosphorylated serine residue in position 5 of CTD repeats
Ser7-P/S7P	phosphorylated serine residue in position 7 of CTD repeats
SC	synthetic complete media
SDS	sodium dodecyl sulfate
snRNA	small nuclear RNA
snoRNA	small nucleolar RNA
T	thymine
TAE	Tris-acetate-EDTA
TAP	tandem affinity purification
TBS	Tris buffered saline
TBE	Tris-Borat-EDTA
TCA	trichloroacetic acid
TdT	terminal nucleotidyl transferase
Thr/T	threonine
Tris	Tris-(hydroxymethyl)-aminomethan
tRNA	transfer RNA
TSS	transcription start site
Tyr/Y	tyrosine
U	uridine
UDG	Uracil-DNA-Glycosylase
UTR	untranslated region
v/v	volume per volume
W	tryptophane
w/v	weight per volume
WGA	whole genome amplification
wt	wild-type
Tyr1	tyrosine
Tyr1-P/Y1P	phosphorylated tyrosine residue in position 1 of CTD repeats
YER	heterochromatic region on chromosome V
YPD	yeast extract – peptone – dextrose (glucose) media