University of Oxford

Medical sciences division

Sir William Dunn School of Pathology

Merton College

‘Role of Pcf11 post-translational modifications in gene expression’

Adam Volanakis

DPhil. Thesis

Supervisor: Prof. Nicholas J. Proudfoot

Oxford

September 2016

Acknowledgments 4
Abstract 6
Abbreviations 9
Chapter 1 11
Introduction 11
1. Introduction 12
1.1. Gene expression 12
1.2. Transcription initiation and elongation 15
1.3. Transcription termination 20
1.4. mRNA export 25
1.5. Pol II CTD modifications 28
1.6. Pcf11 30
1.7. Wnk1 kinase 34
1.8 Synopsis 37
Chapter 2 39
Materials and methods 39
2.1 Methods 40
2.1.1 Recombinant Pcf11 generation 40
2.1.2 *In-vitro* kinase assay with whole cell HeLa extracts 41
2.1.3 Recombinant Wnk1 *in-vitro* kinase assay 42
2.1.4 Western Blot 42
2.1.5 PhosTAG gel 43
2.1.6 Wnk1 CRISPR knock-out generation 44
2.1.7 cDNA synthesis and qPCR 45
2.1.8 PCR directed DNA mutagenesis 46
2.1.9 Chromatin Immunoprecipitation 47
2.1.10 Immunofluorescence staining of HeLa cells 49
2.1.11 mRNA Fluorescence *in-situ* hybridization (FISH) 50
2.1.12 Co-Immunoprecipitation 51
2.1.13 Cytoplasmic and Nuclear fractionation 52
2.1.14 Nuclear Envelope fractionation 53
2.1.15 'Heavy' Chromatin detection 54
2.2 Reagents 55
2.2.1 Oligos 55
2.2.2 Antibodies 55
Chapter 3 57
Phosphorylation of Pcf11 CID by Wnk1 kinase 57
3.1. Introduction 58
3.2. Results 60
3.2.1. Wnk1 interacts with Pcf11 60
3.2.2. Wnk1 nuclear localisation 61
3.2.3. Wnk1 phosphorylates Pcf11 CID 63
3.3 Conclusions 71
Chapter 4 73
Pcf11 CID phosphorylation 73

4.1 Introduction ... **74**
4.2 Results ... **75**
4.2.1 Role of CID phosphorylation in Pcf11 localisation ... 75
4.2.2 CID phosphorylation and Pol II CTD interaction ... 80
4.2.3 CID phosphorylation and Pcf11 recruitment ... 82
4.3 Conclusions ... **84**

Chapter 5 ... **86**

Role of Pcf11 CID phosphorylation ... **86**
5.1 Introduction ... **87**
5.2 Results ... **89**
5.2.1 Role of Wnk1 in mRNA export ... 89
5.2.2 Wnk1 and export factor recruitment ... 94
5.2.3 Mechanism of mRNA export regulation by Wnk1 ... 97
5.3 Conclusions ... **99**

Chapter 6 ... **101**

Discussion ... **101**
6.1 Pcf11 post-translational modifications ... **102**
6.2 Wnk1 kinase ... **103**
6.3 Role of Pcf11 CID phosphorylation ... **105**
6.4 mRNA export and Pcf11 CID phosphorylation ... **107**
6.5 Summary ... **108**
6.6 Future plans ... **109**

References ... **112**

Permissions ... **125**

Acknowledgments

My DPhil. has been an extremely challenging but at the same time rewarding experience. My work in the laboratory of Prof. N. J. Proudfoot for the past 4 years has been a great opportunity for me to meet and interact with bright and talented scientists, tackle challenging scientific questions but also enjoy a nice working environment.

First and foremost, I would like to thank Prof. N. J. Proudfoot for giving me the opportunity to work in his lab. Nick has been a big inspiration for me personally and scientifically. He has always been helpful both in work related matters but also on any personal problems that I encountered. Finally, I would like to thank Nick for his support and contribution now that I am planning the next steps of my career.

I would also like to thank my collaborator on my DPhil. project Dr. Kinga Kamieniarz-Gdula. This project started from her original observations and I would like to thank her for sharing this project with me. Although we didn't always see eye to eye, Kinga has always been really helpful and supportive.

Next, I would like to thank certain people from the NJP lab. Dr. Claudia Ribeiro de Almeida has been a great friend with whom we always discussed experiments and science and she was always the first person to hear about my "exciting" new results. Furthermore, I would like to thank Dr. Pawel Grzechnik and Dr. Hannah Mischo for our discussions and their support. They are both people that I look up to and appreciate. Finally I would also like to

thank the rest of the NJP lab for creating a nice and inspiring work environment.

These 4 years in the lab, some work relationships turned into great friendships. At this point I would like to thank Esther Griesbach for making everyday life happier. Finally, I would like to thank Dr. Margarita Schlakow for being a great friend and an amazing support in extremely difficult and challenging times for me. Without her, I do not think I would be able to see those 4 years through.

Finally, I would like to thank my father Emmanouel Volanakis and my mother Evgenia Saklampanaki for their endless support not only during my DPhil. but throughout all the long years of my studies.

All experiments presented in this thesis were performed by Adam Volanakis. Experiments presented in Chapter 4, Figures 4.8 and 4.9 were performed by Dr. Kinga Kamieniarz-Gdula

Abstract

Role of Pcf11 post translational modifications in gene expression.
Adam Volanakis, Kinga Kamieniarz-Gdula and Nicholas J. Proudfoot
Sir William Dunn school of Pathology, University of Oxford.

mRNA export is one of the major steps in the regulation of gene expression as it provides the link between gene transcription in the nucleus and mRNA translation in the cytoplasm. Efficient export of mRNAs requires the formation of mRNPs at the sites of transcription consisting of the mRNA in complex with export and transcription factors. Formation of the mRNPs is tightly coupled with the co-transcriptional processes of capping, splicing and polyadenylation of the transcript.

Our work focuses on the core termination factor Pcf11, a component of the Cleavage Factor II complex in mammals. Pcf11 is an essential protein in yeast and it has been shown to participate in 3' end formation of the transcript (cleavage and polyadenylation) as well as in the release of RNA polymerase II from the DNA.

We discovered the phosphorylation of 2 residues in the CID domain of human Pcf11. The CID domain is responsible for the interaction of the factor with the CTD domain of RNA polymerase II. Furthermore, we identified Wnk1 as the kinase responsible for this phosphorylation. Our experiments show that phosphorylation of Pcf11 CID by Wnk1 disrupts its interaction with RNA polymerase II. Further analysis of Wnk1 and its role in the nucleus revealed that this kinase participates in mRNA export.

We propose a model, where phosphorylation of Pcf11 CID by Wnk1 is required for proper mRNP assembly and release from the site of transcription. Our data identify a new role for Wnk1 in gene expression regulation through mRNA export and new insights into the cross-talk between transcription termination and mRNA export.

Abbreviations

APA	Alternative poly adenylation
CDK	Cyclin Dependent kinase
CFIA	Cleavage Factor I
CFII	Cleavage Factor II
ChIP	Chromatin Immunoprecipitation
CoTC	Co-Transcriptional Cleavage
CPSF	Cleavage and Polyadenylation Specificity Factor
CSTF	Cleavage stimulation factor
CTD	Carboxy Terminal domain
DNA	Deoxyribonucleic Acid
DOGs	Downstream of Gene
eRNA	Enhancer RNA
FISH	Fluorescent *in-situ* hybridization
IP	Immunoprecipitation
miRNA	micro RNA
mRNA	messenger RNA
mRNPs	messenger ribonucleic protein
NELF	Negative elongation factor
NLS	Nuclear localisation signal
NMD	Nonsense mediated decay
NPC	Nuclear pore complex
ORF	Open Reading Frame

pA	Poly adenylation
PIC	Pre-Initiation Complex
RNA	Ribonucleic acid
RNA Pol II	RNA polymerase II
snRNA	Small nuclear RNA
TAD	Topologically associated domains
TBP	TATA binding protein
TEFb	Transcription Elongation Factor b
TFII	Transcription Factor II
UTR	Untranslated region
WNK1	With no Lysine (K)

Chapter 1

Introduction

1. Introduction

1.1. Gene expression

Gene expression is the process through which the genetic information stored in DNA is transcribed into RNA, which in turn is either translated to protein or performs other functions as non-coding RNA. It is a tightly regulated process that responds to temporal and environmental cues. This regulation is essential for the development of multicellular organisms or for the adaptation to environmental conditions and survival of single cellular organisms. In particular, the development of multicellular organisms from single cells is a great example of this regulation. The spatial and time dependent expression of genes is responsible for the differentiation of cells that share a common set of genes into tissues with distinct features that perform different roles in the organism.

Expression of a protein-coding gene starts with transcription of the DNA template into RNA by the RNA polymerase II (RNA Pol II). Transcription initiation, elongation and termination are highly regulated processes that are accomplished through a complex network of transcription factors, chromatin-modifying enzymes and other proteins interacting with the gene. The RNA transcript must also be modified or processed before it can be exported from the nucleus. The RNA processing reactions involve the 5' capping of the transcript, removal of introns by splicing, 3' end formation and poly-adenylation of the transcript. A mature mRNA is then exported through the nuclear pore complex into the cytoplasm where the process of translation can

start. Each of these processes, in combination with quality control and turnover mechanisms can provide points for the regulation of gene expression (Figure 1.1). Some occur co-transcriptionally but others post-transcriptionally even though it is usually difficult to make this distinction as they often happen simultaneously and in a highly coordinated manner.

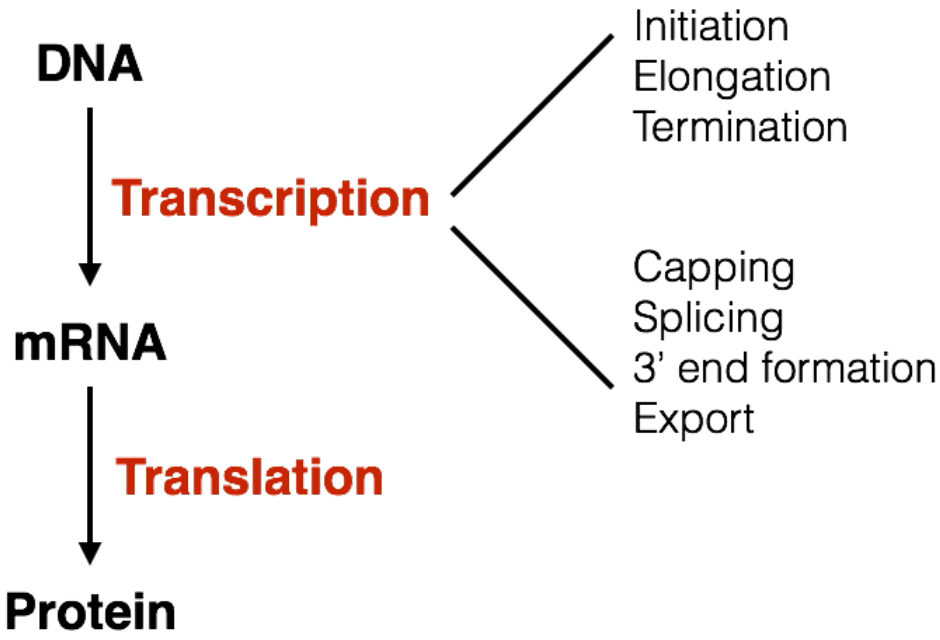


Figure 1.1

Gene expression starts with the transcription of protein coding genes to mRNA. Transcription initiation, elongation and termination are highly regulated processes that require the cooperative action of a plethora of factors. At the same time the transcript needs to undergo processing events to produce a functional, export competent mRNA

1.2. Transcription initiation and elongation

Transcription initiation is one of the most well studied mechanisms of regulation of gene expression. The binding of transcription factors on gene promoters in combination with the effects of distant chromosomal regions called enhancers result in changes in the chromatin structure and the epigenetic modifications of genes resulting in their activation and expression.

The basic model for transcription initiation of protein coding genes involve the assembly of PolII with the general transcription factors TFIIB, TFIID, TFIIE, TFIIF and TFIIH on the gene promoter to form the pre-initiation complex (Conaway and Conaway 1993; Roeder 1996; Kornberg 1998; Reinberg et al. 1998). Binding of this complex to the promoter region is required for the escape of Pol II from the promoter and the initiation of RNA synthesis into the productive elongation phase. TFIID contains the TATA binding protein (TBP) along with several TBP-associating factors (TAFs). Although TBP binding to the TATA box is a common component of transcription initiation at TATA containing promoters, TBP-associating factors are diverse and show gene specificity. Initiation of transcription starts with the recognition of the TATA box by TBP. The TATA box is usually located 30 nucleotides upstream of the transcription start site and has the consensus TATAWAWR. The formation of the TFIIB-TBP-promoter complex results in the recruitment and binding of Pol II-TFIIF. Subsequently, the TFIIE and TFIIH are recruited to form the pre-initiation complex (PIC). When the PIC is formed, melting of the double strand DNA over the promoter happens that leads to the formation of the ‘transcription bubble’ and the transition to the productive

elongation phase of transcription (Grunberg et al. 2012; He et al. 2013). The above account outlines the general core mechanism of transcription initiation (Figure 1.2). Gene specific factors and associated DNA elements provide specific expression profiles and regulation of individual genes.

Early experiments identified DNA sequences separate from gene promoters that are also required for the correct regulation of gene expression (Banerji et al. 1981). These regions are called enhancers and were initially shown to induce transcription of genes at distances ranging from hundreds of bases to megabases (Lettice et al. 2003). Regions classified as enhancers usually have several binding sites for sequence-specific transcription factors. Activation of an enhancer begins with the binding of transcription factors and chromatin remodeling enzymes, events which induce local nucleosome remodeling and recruitment of RNA Pol II (Consortium 2012; West et al. 2014). While it is common for a repressed enhancer to be 'buried' in heterochromatin regions, certain 'pioneer' factors are recruited that are able to bind their recognition motifs in these condensed chromatin states and allow the activation of the enhancers (Zaret et al. 2008). This kind of enhancer regulation is common in the regulation of transcription in developmental processes and the activation of lineage-specific genes and pathways. Activation of an enhancer leads to its transcription by RNA Pol II and in most cases the production of enhancer RNAs (eRNAs) (De Santa et al. 2010; Kim et al. 2010). eRNAs are usually short, non poly-adenylated transcripts that are rapidly degraded by the nuclear RNA surveillance mechanisms. It has been shown that eRNA expression has a functional link with gene expression but

the specific interactions that lead to this kind of regulation remains elusive. In general, it is believed that the interaction of promoters and enhancers involves looping factors like cohesin and the mediator complex that promote an association between the two regions. This looping is thought to facilitate transcription and productive elongation of the gene (Schmidt et al. 2010; Koch et al. 2011). The role of enhancers in the gene expression highlights the fact that the three dimensional structure of chromatin in the nucleus has an important role in gene expression regulation. The formation of topologically associating domains (TADs) is an important aspect of gene expression regulation either through promoter-enhancer interactions or through 'insulation' processes. This latter process can be used for the negative regulation of transcription (Dixon et al. 2012; Sanyal et al. 2012).

The release of RNA Pol II from its paused state into productive elongation is the next highly regulated step in gene expression. Accumulation of RNA Pol II at promoters is thought to be a rate-limiting step for transcription that acts to provide a quality checkpoint before the transition to productive elongation (Adelman and Lis 2012; Kwak et al. 2013). A key player in this process is the C-terminal domain of the Rpb1 subunit of RNA Pol II called the CTD. This domain has a heptad repeat structure of 52 units with the amino acid sequence YSPTSPS and is subject to different modifications during the transcription cycle. It is thought that the CTD can provide a platform for the recruitment of transcription and processing factors to the transcribed locus. Promoter-associated transcription factors in combination with the negative elongation factor (NELF) and DRB-sensitivity factor (DSIF) are responsible for

the pausing of RNA Pol II (Li and Gilmour 2013; Weber et al. 2014). Release of paused polymerase is achieved by the recruitment of the positive transcription elongation factor-b (P-TEFb) complex formed by cyclin T1 and cyclin-dependent kinase 9 (CDK9). P-TEFb is recruited to promoters through its interactions with specific transcription factors and other co-factors in a gene specific manner. It is believed that it promotes the transition of RNA Pol II to the productive elongation phase through the phosphorylation of Ser2 on the RNA Pol II CTD and also on NELF, which leads to its dissociation from RNA Pol II (Lis et al. 2000; Peterlin and Price 2006). When productive elongation starts the speed of transcription varies across the gene and it is believed to provide an opportunity for the co-transcriptional processing of the transcript (splicing, cleavage and poly-adenylation). Transcription rate can be regulated through histone marks and specific DNA sequences (Alexander et al. 2010; Danko et al. 2013; Jonkers et al. 2014; Nojima et al. 2016).

To sum up, the first steps of gene expression comprise highly complex mechanisms that are regulated at every step. From initiation to elongation the action of a huge set of proteins and DNA elements provides the necessary regulatory steps for regulation of gene expression in response to the environment and time in the life of the cell.

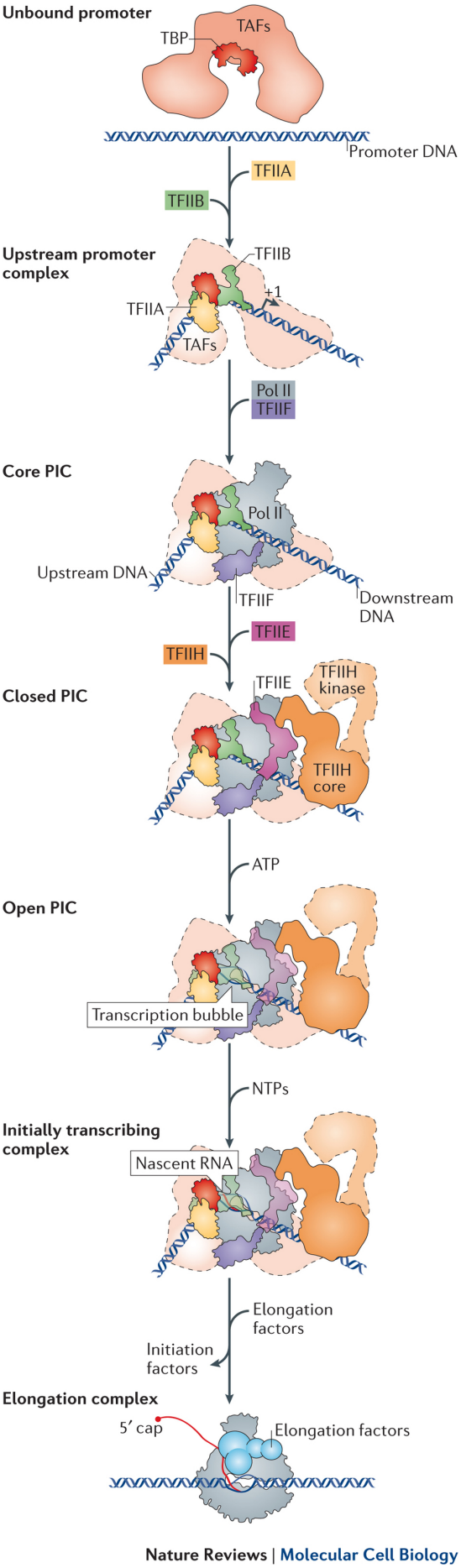


Figure 1.2

Schematic overview of transcription initiation. TBP binds to the TATA box at the promoter region causing a bend of the template. Subsequently more transcription factors are recruited with the eventual recruitment of RNA Pol II along with TFIIF. The final step in the formation of the pre-initiation complex (PIC) is the recruitment of TFIIH and TFIIF. After PIC formation melting of the DNA occurs that leads to the formation of the transcription bubble and the transition to the productive elongation phase of transcription.

Source: Nature Reviews Molecular cell biology. DOI: 10.1038/nrm3952

1.3. Transcription termination

Termination is the final step of transcription. In mammalian protein coding genes, it is a highly complex process that involves the coordinated action of several termination factors. Transcription termination is not just a constitutive process occurring at the 3' ends of genes to define the boundaries of transcription units. It is also a regulatory process that can affect the cellular fate of the transcript as well as the form of the protein produced. The interaction of the termination machinery with modifying and degradation enzymes provides a regulatory role in the process. Transcription termination and the 3' end formation of mRNA coding genes leads to the production of stable RNAs that are exported to the cytoplasm for translation. On the other hand, in the case of many non-coding RNA genes, transcription termination is linked to their nuclear retention and subsequent degradation. Furthermore, transcription termination is a mechanism that can keep pervasive transcription in check. Large parts of the genome are transcribed although they may not correspond to specific annotated coding regions. The function of those transcripts is largely unknown. Early termination of these transcripts is crucial to stop these transcription events from interfering with transcription of coding regions and to prevent the accumulation of non-functional RNAs (Jacquier 2009; Jensen et al. 2013).

Transcription termination occurs when RNA Pol II ceases RNA synthesis and is released from the DNA template after the cleavage and release of the co-transcriptionally processed newly synthesized RNA. Pol II

termination can occur through different pathways depending on the termination signals on the gene or nascent RNA as well as on the termination factors that are recruited to the end of the gene. In the majority of mRNA coding genes, the pA dependent transcription termination pathway is used (Whitelaw and Proudfoot 1986; Logan et al. 1987). The core termination factors involved in this pathway are the cleavage and poly adenylation specificity factor (CPSF), cleavage stimulatory factor (CstF) and cleavage factors I and II (CFI, CFII). RNAs terminated through this pathway are poly-adenylated and in general rapidly exported out of the nucleus for cytoplasmic translation.

As RNA Pol II transcribes a gene, its CTD domain is progressively phosphorylated on Ser2, which is recognized by CFI and CFII. CPSF binds directly to the body of the polymerase, whereas specific signals on the transcript (including the poly adenylation signal) are recognized by CstF and CPSF. When the RNA Pol II transcribes past the pA site it pauses and the transcript is cleaved by the CPSF component 73 (CPSF73). The transcript is then poly adenylated and engaged in the process of nuclear export. Transcript cleavage generates a 5' end on the nascent RNA, which provides an entry point for the exonuclease Xrn2 (Rat1 in yeast) that starts degrading the transcript. Xrn2 potentially acts like a 'torpedo', which in combination with conformational changes in the elongating polymerase leads to its dissociation from the DNA template and the completion of transcription termination (Plant et al. 2005; Kuehner et al. 2011; Porrua and Libri 2015; Proudfoot 2016) (Figure 1.3).

The place and time that transcription occurs can have regulatory effects on the fate of the transcript. Delayed termination can result in interference of the transcribing polymerase with downstream transcription units as well as with the replication process (Greger et al. 1998; Greger and Proudfoot 1998; Shearwin et al. 2005; Gaillard et al. 2015). Furthermore transcription termination upstream or inside ORFs can lead to down regulation of gene expression (Kaida et al. 2010) whereas impaired termination at the 3' ends of genes can affect transcription initiation possibly through a feedback mechanism (West and Proudfoot 2008; Mapendano et al. 2010). Finally, the choice of termination sites of a gene can lead to alternative poly adenylation of the transcript. Alternative poly adenylation (APA), depending on where it occurs can produce transcripts with different properties. Cryptic pA sites inside introns can generate different isoforms of a protein whereas usage of alternative pA sites in the 3' ends of genes can produce transcript with longer or shorter 3' UTRs. Alternative 3' UTRs may affect the stability, localisation and function of the mRNA by generating different protein binding sites and miRNA recognition sites that can affect the fate of the transcript (Proudfoot 2011; Elkon et al. 2013; Tian and Manley 2013) (Figure 1.4). Recent studies have shown that under stress conditions some genes fail to terminate at their regular termination sites, this results in the production of long extended transcripts called DOGs (Downstream of gene). These may play a role in protecting the integrity of the nucleus under cellular stress conditions (Vilborg et al. 2015).

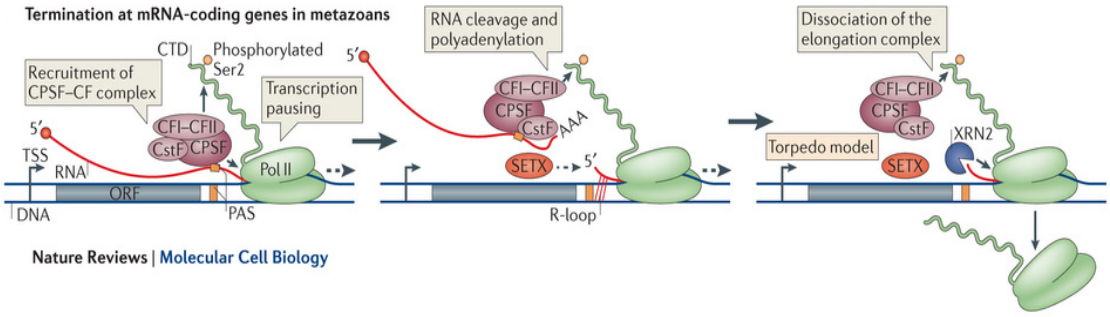


Figure 1.3

Schematic representation of the transcription termination process of mRNA coding genes. During transcription elongation Pol II CTD domain gets phosphorylated on Ser2. This phosphorylation, along with signals on the transcript leads to the recruitment of the core termination factors. When Pol II transcribes past the pA site it pauses and the transcript is cleaved by CPSF73 and poly adenylated. Senataxin (SETX) resolves the RNA-DNA hybrid formed between the nascent RNA and the DNA template followed by the cleavage of the transcript. This generates an available 5' end entry point for the exonuclease Xrn2 which starts degrading the nascent RNA. This in combination with conformational changes to RNA Pol II leads to its dissociation from the DNA template and the termination of transcription.

Source: Nature Reviews Molecular Cell Biology, DOI: 10.1038/nrm3943

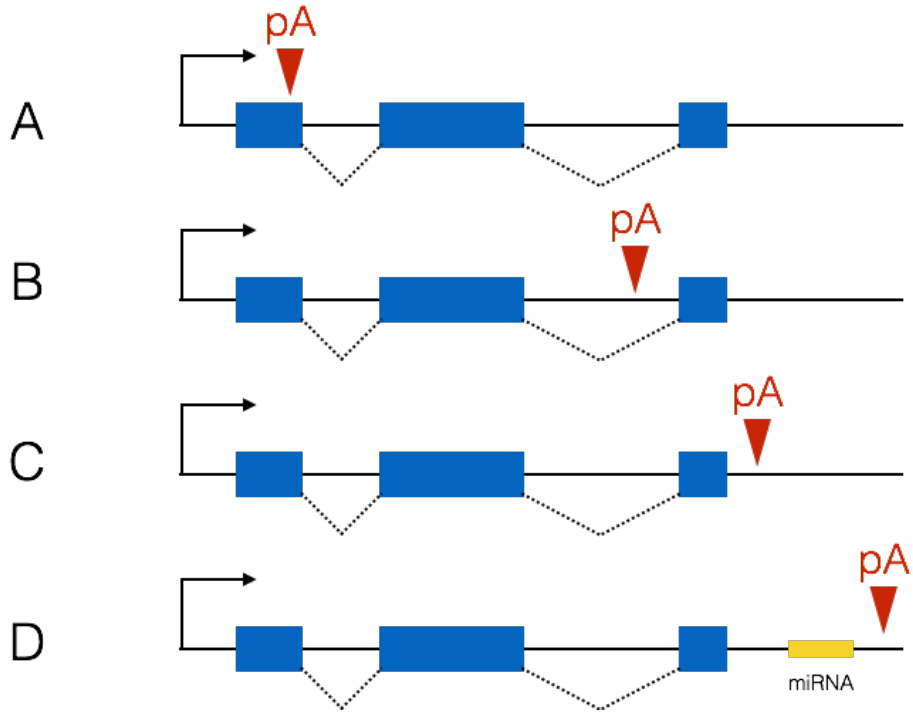


Figure 1.4

Examples of alternative poly adenylation. A. Usage of cryptic pA sites inside the ORF of a gene can result in premature termination and the subsequent degradation of the transcript. B. Usage of alternative pA sites can lead to exon exclusion thus producing different isoforms of a protein C, D. Usage of a proximal or distal pA site can change the length of the 3' UTR of a transcript. Often lengthening of a UTR results in the inclusion of miRNA binding sites that can affect the fate of the transcript.

1.4. mRNA export

The next step of gene expression for mRNA genes following termination of transcription is nuclear mRNA export. During transcription and mRNA processing (Capping, splicing, 3' end formation) the transcript is bound by several RNA binding proteins and sometimes ncRNAs to form mRNPs. Nuclear export of mRNPs into the cytoplasm occurs through the Nuclear Pore Complex (NPC), a multiprotein barrel-like structure stretching across the nuclear membrane. The action of several export factors is required for mRNA export some of which are recruited co-transcriptionally to the gene locus through interactions with transcription factors. Coupling of transcript processing with recruitment of export factors is required for efficient export (Overview of mRNA export Figure 1.5) . Release of mRNPs from transcription loci is a regulated process such that any defect in 3' end formation or mRNP assembly will trigger quality control mechanisms that will degrade the aberrant molecules (Hilleren et al. 2001; Hilleren and Parker 2001; Jensen et al. 2003; Jensen and Rosbash 2003; Houseley et al. 2006).

In metazoans and yeast, the export factors UAP56/Sub2 and Aly/Yra1 are co-transcriptionally recruited to the elongation complex through interactions with splicing and transcription factors, an event that leads to the assembly of the TREX complex (Chi et al. 2013). Some components of the TREX complex are then transferred onto the mRNA (Aly, UAP56), marking it competent for export. The mRNA is then bound by the export receptor TAP/Mex67, which through interactions with nucleoporins and export factors

acts to guide the mRNA into the NPC and to subsequent export (reviewed in (Bjork and Wieslander 2014)).

Interestingly, it has been shown that mutation of certain termination factors (Yeast CFIA subunits, RNA14, RNA15, Pcf11) as well as poly(A) polymerase can have inhibitory effects on mRNA export which causes the retention of the transcripts in the nucleus. Experiments in yeast have shown that this defect in mRNA export is caused by the interaction of the termination factor Pcf11 and the export factor Yra1 (Aly in mammals) (Johnson et al. 2009). Yra1 is recruited to transcribed loci through its association with Pcf11 which is required for the correct assembly of an export competent mRNP. Furthermore it was shown *in vitro* that mammalian Aly and Pcf11 can also interact. Defects in mRNA export lead to the accumulation of termination factors at the 3' ends of genes, suggesting that mRNA export and the release of mRNPs from the transcription locus require the coordinated action of export factors like Aly with 3' processing factors (Rougemaille et al. 2008). Finally it has been suggested in *S.cerevisiae* that recruitment of the export factor Yra1 to genes through its interaction with Pcf11 can modulate transcription termination possibly by competition with the Pcf11 co-factor Clp1 (Johnson et al. 2011). All the above suggests that the co-ordination between the processes of transcription termination and mRNA export is critical to productive gene expression. However the mechanisms by which this is achieved are largely unknown especially in metazoans.

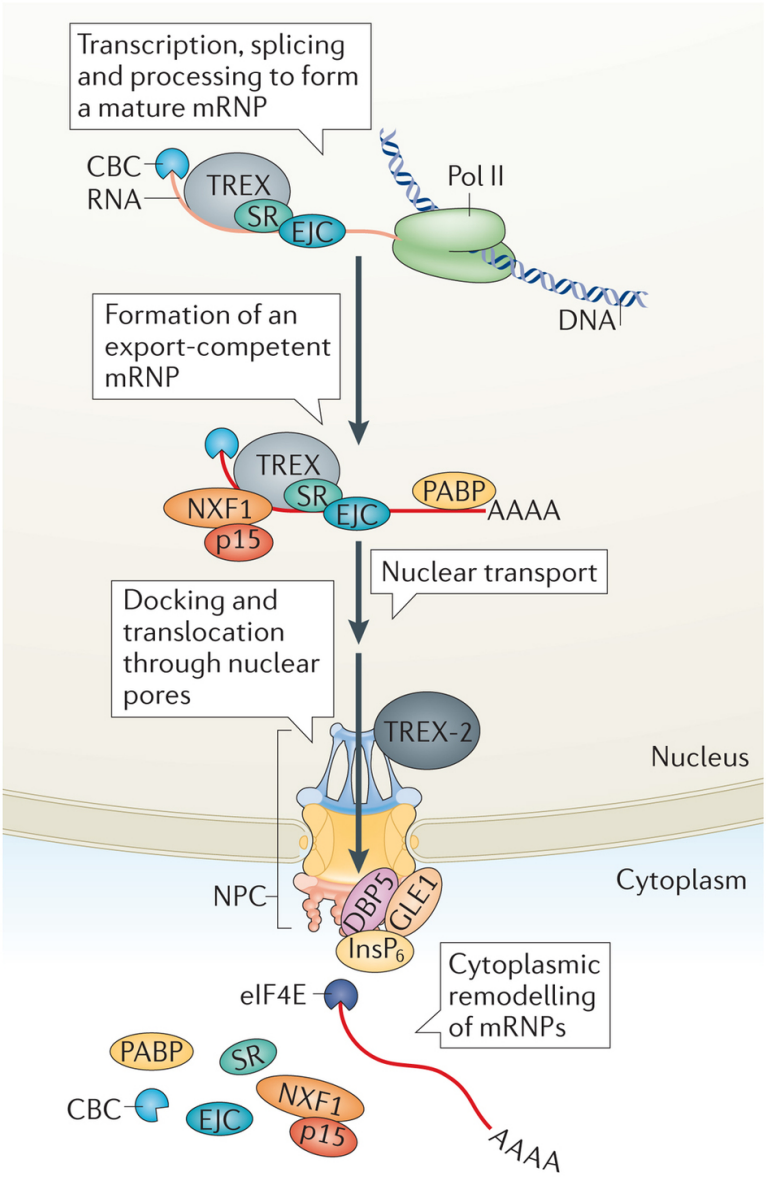


Figure 1.5

Representation of the mRNA export pathway. Assembly of export factors starts co-transcriptionally with the recruitment of the TREX complex by splicing factors and the exon junction complex (EJC), further export proteins are recruited during 3' end formation (NXF1, p15) leading to the formation of a mature mRNP. After release from the transcribed locus the mRNP is docked at the Nuclear Pore Complex and exported to the cytoplasm.

Source: Nature Reviews Molecular Cell Biology. DOI: 10.1038/nrm4010

1.5. Pol II CTD modifications

In eukaryotes, all protein-coding genes are transcribed by RNA Pol II. The largest subunit of RNA Pol II, Rpb1, contains highly conserved C-terminal domain (CTD) that consists of multiple tandem repeats of seven amino acids (Tyr-Ser-Pro-Thr-Ser-Pro-Ser). The number of repeats varies between organisms. *Saccharomyces cerevisiae* has 26 CTD repeats whereas in humans there are 52 repeats in total. The CTD of the polymerase plays a critical role in transcription and its regulation. The residues of the CTD repeats can be modified by phosphorylation (Ser, Tyr, Thr) whereas the Prolines can assume *trans- or cis-* isomerization states (Chapman et al. 2008; Liu et al. 2010). The modification of the CTD shows specific profiles during each stage of transcription and can turn the CTD into a scaffold for the precise and timely recruitment of transcription factors based on the type of modifications (reviewed in (Egloff and Murphy 2008; Egloff et al. 2012)). In general, the CTD plays a major role in coupling the processes of transcription with RNA processing and mRNA export. Consequently, its deletion is lethal in several model organisms.

Phosphorylation of Ser5 and Ser2 of the CTD are the best-characterized modifications. Both modifications are required for the regulation of transcription and RNA processing. In eukaryotes, CDK7, a subunit of the TFIIH initiation factor is thought to be the kinase responsible for the majority of Ser5 phosphorylation. In general, Ser5 phosphorylation is a characteristic of RNA PolII during transcription initiation as well as on the borders of exons. During transcription initiation the Ser5 mark is required for the recruitment of

capping enzymes which act to add the 5' cap structure to the nascent RNA (Ghosh et al. 2011). Ser5 phosphorylation is also involved in the recruitment of other factors including the termination factor Nrd1, as well as the histone modifying enzymes Set1 and Rpd3S. Finally, Ser5 phosphorylation is present in the 5' ends of exons at the sites where pausing of the polymerase is detected. In yeast it is thought that RNA Pol II is newly phosphorylated on Ser5 and this modification plays a role in its pausing/release (Alexander et al. 2010). Ser2 phosphorylation shows a somewhat opposite profile to Ser5. Ser2 phosphorylation levels increase during the elongation phase of transcription and are characteristic of the 3' ends of genes. CDK9 seems to be the major kinase responsible for this modification even though other enzymes have also been identified as potential Ser2 phosphorylation kinases (Bres et al. 2008; Bartkowiak and Greenleaf 2011). The Ser2 phosphorylation mark is the dominant CTD modification at the 3' ends of genes and is recognized by several termination factors such as Pcf11 and Rtt103 in yeast (Lunde et al. 2010). Some proteins recognize CTD repeats that are simultaneously phosphorylated on Ser2 and Ser5, like the splicing factor U2AF and Prp40. The efficient recruitment of splicing factors is a requirement not only for splicing but also for the initial recruitment of certain export factors (Bono and Gehring 2011; David et al. 2011).

Modifications of the other residues of the CTD repeats are also present during transcription. Their role is less defined and they may have a more gene specific action. Ser7 phosphorylation seems to play a role in the transcription of specific snRNA genes whereas Thr4 phosphorylation appears to be

present mostly over the termination regions of genes. The involvement of these modifications in transcription has not yet been illuminated (Egloff et al. 2007; Nojima et al. 2016).

1.6. Pcf11

Pcf11 is one of the core termination factors in eukaryotes. It is an essential protein in yeast (Amrani et al. 1997) and has been shown to participate in all steps of pA dependent transcription termination (Cleavage, poly adenylation, Pol II release) as well as CoTC dependent termination (West and Proudfoot 2008). Pcf11 is recruited to actively transcribed genes through its CTD interacting domain (CID) that contacts the CTD of RNA Pol II. It is believed that Pcf11 preferentially interacts with CTD repeats phosphorylated on Ser2 (Barilla et al. 2001; Licatalosi et al. 2002; Sadowski et al. 2003). Mutations of Pcf11 can lead to defects of transcript cleavage, in the poly adenylation of the mRNA and in the release of RNA Pol II from the DNA template (de Vries et al. 2000; Zhang and Gilmour 2006; West and Proudfoot 2008; Ghazy et al. 2012). Furthermore, Pcf11 has a significant role in alternative poly adenylation (APA) and its recruitment may be the rate-limiting step for the cleavage and poly adenylation reaction (Li et al. 2015).

In yeast, Pcf11 is a 72 kDa protein associated with the transcribed gene through its interaction with the polymerase CTD. Pcf11 has a Clp1 binding domain flanked by Zinc finger domains as well as a polyglutamine domain. At the N-terminus, Pcf11 has a CID domain that recognizes the CTD

heptad repeats (Tyr-Ser-Pro-Thr-Ser-Pro-Ser) and more specifically shows a preference for repeats with phosphorylated Ser2 and trans-isomerized Pro3 (Meinhart and Cramer 2004).

In detail, Pcf11 CID forms a right-handed superhelix consisting of eight α-helices. This superhelix interacts with the CTD through hydrogen bonds in the groove formed by α-helices 2, 4 and 7. More specifically seven hydrogen bonds are generated between the CTD and five residues of the CID. Phosphorylated Ser2 does not interact directly with the CID. This observation explains why Pcf11 can interact both with Ser2- phosphorylated and unphosphorylated CTD. Ser2 phosphorylation may act to enhance the CID-CTD interaction by stabilizing the β-turn structure of the CTD through the generation of a hydrogen bond with Thr4. Stabilisation of this β-turn structure is thought to greatly enhance and facilitate the interaction with the CID domain (Meinhart and Cramer 2004; Noble et al. 2005).

In *Saccharomyces cerevisiae*, Pcf11 is an essential protein. It is part of the Cleavage Factor Ia complex (CFIA), which is composed of the proteins Rna14, Rna15, Pcf11 and Clp1 (Minvielle-Sebastia et al. 1994; Minvielle-Sebastia et al. 1997). The preferential interaction of Pcf11 with Ser2 phosphorylated CTD is involved in the recruitment of CFIA at the transcribed locus. Although CTD is not absolutely required for transcription termination in yeast, its absence greatly reduces the efficiency of termination (Licatalosi et al. 2002). While Pcf11 participates both in transcript cleavage and termination of transcription, these functions can be uncoupled from each other (Sadowski et al. 2003). It has been shown that CTD binding of Pcf11 is required for

correct and efficient transcription termination in a pA dependent manner. On the other hand, CTD binding is not required for the cleavage of the transcript and 3' end formation.

Recent studies in yeast have shown that Pcf11, apart from its role in the transcription termination and 3' end formation of protein coding genes, it is also required for Nrd1-dependent termination, which is specific for short non-coding RNAs (CUTs, SUTs, XUTs) (Grzechnik et al. 2015). In these non-coding RNA genes, Nrd1 is required for the recruitment of Pcf11 and Pcf11 in turn acts to position Nrd1 on the gene. Furthermore, Pcf11 recruitment on these genes promotes CTD Ser2 phosphorylation, possibly through the restriction of Nrd1-CTD interaction. This facilitates the recruitment of Sen1, which is required for Nrd1-dependent transcription termination *in-vivo* (Finkel et al. 2010; Chinchilla et al. 2012; Porrua and Libri 2013).

Moreover, it has been shown that through its zinc finger/Clp1-binding domain Pcf11 is further responsible for the co-transcriptional recruitment of the export factor Yra1 in *Saccharomyces cerevisiae*. Yra1 (ALY/REF in humans) is a member of the highly conserved REF family of export factors that are recruited to the transcript and convey the mRNA to the export proteins Mex67/Nxf (Johnson et al. 2009). The involvement of Pcf11 in mRNA export provides a clear link between the processes of export and 3' end formation provides insight into the coordination of the gene expression process.

In humans, Pcf11 is a 170 kDa protein, component of the Cleavage Factor A II complex (CFIIAm) which is essential for transcription termination

(de Vries et al. 2000). Pcf11 has a role in the efficient cleavage of the transcript but is also required for the degradation of the 3' product of the pA site cleavage (West and Proudfoot 2008), suggesting that it is somehow involved in the recruitment of Xrn2 exonuclease as it has been suggested for its yeast homologue Rat1 (Luo et al. 2006). These results, in combination with *in-vitro* studies in yeast and *Drosophila melanogaster* where Pcf11 has been shown to dismantle the RNA Pol II elongation complex (Zhang et al. 2005; Zhang and Gilmour 2006), suggest that Pcf11 participates in transcription termination. Indeed, these data imply a termination mechanism that incorporates characteristics of both the torpedo and the allosteric models of termination.

Finally genome wide profiling of poly adenylated transcripts revealed that Pcf11 has a significant effect in pA site selection (Li et al. 2015). Knock-down of Pcf11 in human cells caused the lengthening of the 3' UTR generated by the usage of a downstream pA site in transcription termination. In contrast depletion of other core termination and poly adenylation factors either had no effect on 3' UTR length or caused its shortening. These results provide only limited mechanistic insight in the participation of Pcf11 in pA site selection and alternative poly adenylation. However it is clear that Pcf11 recruitment may be the rate-limiting step for the cleavage and poly-adenylation reaction.

1.7. Wnk1 kinase

WNKs (With no lysine [K]) are a family of Serine/Threonine protein kinases with a unique placement of a lysine residue in their catalytic domain that is involved in ATP binding and the catalysis of the phosphate transfer reaction (Verissimo and Jordan 2001). More specifically, WNK kinases lack a catalytic lysine normally present in β strand 3 of typical kinase domains and instead this catalytic lysine is found in β strand 2. In general, WNK kinase domains do not share more than 29% sequence identity with other protein kinases. They are restricted to multicellular organisms, with four homologues expressed in human cells.

Wnk1 kinase is a member of the WNK family and is globally and ubiquitously expressed in human cells (Xu et al. 2000). It has several splice variants and isoforms that are expressed in different tissues suggesting that certain isoforms may have tissue specific functions (Verissimo and Jordan 2001; O'Reilly et al. 2003). Most notably the dominant isoform expressed in kidneys lacks 384 N-terminal amino acids, resulting in a protein without a catalytic domain. Wnk1 is thought to have a variety of functions including participation in the ion export pathways, chromosome segregation and cardiovascular development (O'Reilly et al. 2003; Tu et al. 2011; Serysheva et al. 2013; Xie et al. 2013; Liu et al. 2015). In mice, disruption of the *WNK1* gene is lethal early in development. Finally mutation of Wnk1 and certain of its isoforms can be used for the treatment of chronic neuropathic pain caused by peripheral nerve injury (Kahle et al. 2016).

Wnk1 is a 230 kDa protein with an N-terminal kinase domain (amino acids 218-483) (Min et al. 2004; Xu et al. 2005a). Mutagenesis analysis has identified Ser-382 as an auto-phosphorylated residue whose phosphorylation is required for Wnk1 activity. In human cells the four WNK kinases share an 80% sequence identity of the catalytic domain. Biochemical experiments have identified an auto-inhibitory domain next to the catalytic domain of the protein between residues 490 and 550. The rest of the protein lacks any identifiable structured domains apart from two coiled-coil regions at the C-terminus. The high conservation of these C-terminal domains suggests that they may have functional roles.

It has proved difficult so far to express and purify full length Wnk1 in bacterial or insect cells (Min et al. 2004). However, the expression and crystallization of the catalytic domain has been successful. The structure of the catalytic domain differs from typical kinase domains by having a six-stranded β-sheet that forms a nearly complete β-barrel in contrast with the typical five-stranded β-barrel. The defining feature of Wnk1 kinase is the lack of the catalytic Lysine residue in β3 strand. In typical kinase domains this Lysine is responsible for interacting with ATP phosphates and is a common mutagenesis target for the generation of catalytic mutants. In Wnk1, this Lysine is substituted by a Cysteine and its distance from the active site makes it improbable that it can participate in catalysis. But a Lysine (Lys-233) is present in the β2 strand of Wnk1, which according to the structure can reach the catalytic centre. Mutation of Lys-233 abolishes the catalytic function of Wnk1 suggesting that the catalytic Lys of WNKs is located on strand β2

instead of β3. This unique catalytic residue suggests that the ATP binding site of Wnk1 has a different shape to typical kinases. Analysis of the structure of the active site and the activation loop structure of Wnk1 provides insight in the activation of the protein through auto-phosphorylation. The exterior positioning of the activation loop of Wnk1 enables it to be inserted in the active site of a second Wnk1 kinase and be activated through its auto phosphorylation.

The physiological role of WNK kinases was largely unknown until it was discovered that large intronic deletions of Wnk1 and Wnk4 that result in the over-expression of the proteins can cause pseudohypoaldosteronism type II, a genetic disease featuring hypertension (Wilson et al. 2001). Further studies have shown that Wnk1 is responsible for the regulation of ion co-transporters in human cells. Specifically, Wnk1 is required for the regulation of the BK channel, a large calcium conductance and voltage activated potassium channel, through the inhibition of the ERK1/2 signaling pathway. A catalytically active Wnk1 is required for this type of regulation (Liu et al. 2015). Moreover, Wnk1 participates in the regulation of sodium transporters through the activation of Sgk1 kinase, which in turn is required for the activation of epithelial sodium channels (Xu et al. 2005b).

Disruption of the Wnk1 gene in mice displays a lethal phenotype early in development, suggesting that Wnk1 has an important role in the early mouse developmental stages (Zambrowicz et al. 2003). Studies have shown that Wnk1 is required for the activation of Osr1, an oxidative stress-response kinase. Failure to activate Osr1 leads to defective angiogenesis in the brain, reduced myocardial trabeculation in hearts and defective maturation of large

vessels in yolk sacs (Xie et al. 2013). In *Drosophila melanogaster,* WNK kinases have been identified as positive regulators to the Wnt/β-catenin signaling pathway, which plays a pivotal role in development, more specifically in the polarization of cells in the epithelial plane. Furthermore its regulation is important for cancer prevention (Serysheva et al. 2013).

Localisation studies of Wnk1 in human cells have shown a strong localisation of the protein on mitotic spindles during mitosis. Knock-down experiments showed that loss of Wnk1 causes aberrant mitotic spindles and defective chromosome segregation. Although the mechanism through which Wnk1 regulates the mitotic spindle formation it is elusive, there is strong evidence that Wnk1 can act as a mitotic kinase and participate in cell cycle regulation (Tu et al. 2011).

1.8 Synopsis

In this study I investigate the termination factor Pcf11. Our work identified two residues (S120 and T121) in the CID domain of Pcf11 that are phosphorylated. Furthermore we identified the kinase responsible for this modification.

In Chapter 3, I describe the identification and verification of the phosphorylation of Pcf11 CID. I use *in-vitro* assays with recombinant Pcf11 CID to verify the modifications. Furthermore I investigate the role of Wnk1 kinase. My experiments provide evidence for the nuclear localisation of the

kinase as well as show that it can phosphorylate Pcf11 CID on residues S120 and/or T121.

In Chapter 4, I study the effect of Pcf11 CID phosphorylation. I used microscopy experiments and cell fractionations to identify the localisation of Pcf11 and how this is affected by CID phosphorylation. Furthermore using Co-Immunoprecipitation experiments and biochemical approaches we show that CID phosphorylation can affect the interaction of Pcf11 with RNA Pol II and can potentially have an effect on its recruitment at transcribed genes.

Finally, in the 5th Chapter I study the role of Pcf11 CID modifications in gene expression. Using mRNA FISH experiments I show that phosphorylation of Pcf11 is required for efficient mRNA export. Furthermore, I investigate the mechanism through which this regulation is happening. My experiments show that Pcf11 CID phosphorylation is not affecting recruitment of mRNA export factors but acts through a different downstream mechanism. Our data provide hints for a potential mechanism where Pcf11 CID phosphorylation is required for the disassembly of the transcription termination/ mRNP complexes and for efficient mRNA export.

Chapter 2

Materials and methods

2.1 Methods

2.1.1 Recombinant Pcf11 generation

Pcf11 CID was amplified from Pcf11 cDNA using primers (#11, #12, see section 2.2.1.). With the amplification, restriction enzyme sites BamH1 and Xho1 were added to the 5' and 3' end of the product respectively. The CID DNA was then ligated in pGEX-4T1 N-terminal GST tag vector. The expression vector was transformed in BL21 *E.coli* cells for expression. 5ml cultures were inoculated and grown overnight from single colonies. From these cultures, 2 lt cultures were started at OD 600= 0.1 and grown to OD 600= 0.5. 10mM IPTG was added to the cultures for the induction of expression. The cultures were incubated at 25^{o}C for 3 hours. Cells were collected with centrifugation and cell pellets were lysed in 10ml Lysis buffer (50mM Tris pH 7.5, 100mM NaCl, 1% Triton, 5% Glycerol, 1mM DTT, DNAse, 300ng/ul Lysozyme, Protease inhibitor cocktail) for 30' at 4^{o}C. Lysates were sonicated 10 x 1 minute bursts. The lysate was then centrifuged at 13,000 rpm for 10' and the supernatant was collected.

Lysates were mixed 1:1 with Binding buffer (50mM Tris, 150mM NaCl; pH 8.0) and applied to a Glutathione Column (Pierce Glutathione Cartridge 16110). After washes with Binding buffer, the protein was eluted with Binding buffer with 10mM Glutathione. The eluate was applied to a Pierce Protein Concetrator column (PES, 10k MWCO, CN88517) and centrifuged 15,000g for 15 minutes. The concentrated sample was collected and run on an acrylamide gel. The gel was stained with colloidal blue stain (Invitrogen,

LC6025) and the protein concentration was determined using Bradford (BioRad, 500002).

2.1.2 *In-vitro* kinase assay with whole cell HeLa extracts

HeLa cells were lysed with Assay Lysis buffer (1% NP40, 150mM NaCl, 50mM Tris pH 7.2, Protease Inhibitor cocktail, Phosphatase inhibitor PhoSTOP) for 10 minutes on ice. After centrifugation the supernatant was collected.

The kinase reaction consisted of 200ng recombinant CID, 10ul of whole cell extract, 0.5mM ATP, 0.5ul ^{32}P γ-ATP, Protease and phosphatase inhibitors in Kinase assay buffer (100mM HEPES pH 7.9, 100mM $MgCl_2$, 100mM $MnCl_2$, 1.5M NaCl). The samples were incubated at 37^{o}C shaking for 1 hour.

Glutathione sepharose beads were prepared for purification of CID from the kinase assay mix. Beads were washed with 1x Kinase assay buffer and added to the kinase reaction. The samples were incubated with the beads at 4^{o}C for 1 hour. After 3 x 5 minute washes with 1x Kinase assay buffer the beads were boiled in SDS gel loading buffer and loaded on a 12% Acrylamide gel (Invitrogen Novex). The gel was dyed with InstaBlue and scanned.

For the drying of the gel, it was first equilibrated for 5' in gel drying solution (30% Methanol, 5% Glycerol) and then dried with a hot air vacuum pump at 80^{o}C for 3 hours. After drying the gel was exposed on Kodak biomax film.

2.1.3 Recombinant Wnk1 *in-vitro* kinase assay

For this assay commercially available recombinant Wnk1 catalytic domain was used (ProQinase, CN 1111-0000-11) with the recombinant Pcf11 CID that I generated and purified (Chapter 2.1.1).

1µg of recombinant Wnk1 and 5µg of recombinant CID were incubated with kinase reaction buffer (100mM HEPES pH 7.9, 100mM $MgCl_2$, 100mM $MnCl_2$, 1.5M NaCl, Protease and Phosphatase inhibitor cocktail) and the addition of 0.2ul ^{32}P γ-ATP. For the control reaction, recombinant Wnk1 was incubated with 0.5 µg of MBP positive control substrate in the same buffer and conditions. The samples were incubated at 37^{o}C for 10 minutes. SDS loading buffer was added to the reaction followed by boiling at 95^{o}C for 3 minutes. The samples were then run on a 12% Bis-Tris gel (Novex, NP0341). After completion of electrophoresis the gel was prepared for drying by incubation for 15 minutes in drying solution (30% Methanol, 5% Glycerol). The gel was then dried with a hot air vacuum pump apparatus for 3 hours at 80^{o}C.

2.1.4 Western Blot

Cell pellets were resuspended in 50ul RIPA buffer (150mM NaCl, 5mM EDTA, 50mM Tris pH 8, 1% NP40, 0.5% Na-deoxycholate, 0.1% SDS) supplemented with protease and phosphatase inhibitor cocktail. The samples were incubated for 15 minutes on ice for efficient lysis and subsequently centrifuge for removal of the debris. An equal volume of 2x SDS loading buffer

was added to the lysates (200mM β-mercaptoethanol, 100mM Tris pH 6.8, 4% SDS w/v, 20% glycerol v/v, 0.2% bromophenol blue w/v) and the samples were boiled at 95^{o}C for 3 minutes.

The samples were loaded to the appropriate gel for electrophoresis. For the transfer, PVDF membrane was used and a wet-transfer method. Membranes were stained with Ponceau red and then blocked in 5% milk in PBS with 0.1% Tween (PBST) for 1 hour. The primary antibody was diluted in 5% milk in PBST and was incubated with the membrane overnight at 4^{o}C. The blot was subsequently washed with PBST for 15 minutes total with changing of the wash solution every 5 minutes. The secondary antibody was diluted 5% milk and PBST and incubated with the membrane for 45 minutes in room temperature. The membrane was then washed 3 times for 15 minutes each in PBST. After the washes the blot was air-dried and ECL (Pierce, 32106) was added on the membrane, which was then exposed on a film.

2.1.5 PhosTAG gel

PhosTAG is a proprietary compound (Wako chemical industries, AAL-107) that recognizes and binds phosphorylated amino acids. The compound is added to protein acrylamide gels and causes a delay in the migration of phosphorylated versus unphosphorylated proteins. 100μM PhosTAG was added to the mix for the resolving part of the protein gel.

HeLa cells were lysed with RIPA buffer (150mM NaCl, 5mM EDTA, 50mM Tris pH 8, 1% NP40, 0.5% Na-deoxycholate, 0.1% SDS) with the

addition of phosphatase and protease inhibitors. Part of the samples was treated with Lambda protein phosphatase (NEB, P-0753) for 30 minutes at 30^{o}C. The samples and the de-phosphorylated samples were run on the PhosTAG gel (30mA for 4 hours). The gel was washed with 10mM EDTA in transfer buffer 3 times for 15 minutes to remove the Manganese ions present in the gel, which is a requirement for the efficient transfer of the PhosTAG gel. After the transfer I proceeded as described in the Western blot section (Chapter 2.1.4).

2.1.6 Wnk1 CRISPR knock-out generation

Two guide RNAs (gRNAs) were designed to target the first exon of WNK1 gene (gRNA 1: AGTCACAGATGACGCTCCGGCGG and gRNA 2: CCGCCGCCACACTATGGACAAGG) and cloned into pX459 vector which also contained the gene expressing Cas9 D10A mutation which causes the enzyme to generate nicks on the DNA instead of double strand breaks.

The two plasmids were transfected to HeLa cells. 24 hours after transfection Pyromycin selection was applied to the cells (1μg/ml) for 3 days. After selection cells were collected and diluted to a concentration of 1 cell / 10ul and then added to a 96 well plate so that each well would contain 1 cell. Cells were left to grow for one week and checked every two days. Wells that contained single colonies were selected and the cells transferred to 6 well plates.

The cells were expanded further and genomic DNA was extracted. Using the genomic DNA as template a region of the WNK1 gene was amplified using primers upstream and downstream of the gRNA targeting sites (oligos #13, #14). This region was then cloned in a PCR TOPO II vector (Invitrogen, CN 450159). The plasmids were transformed in bacteria and 10 individual bacterial clones were selected for each Wnk1 knock-out clone. The plasmids were sequenced for verification.

For further verification of the clones, the positive clones that appeared from the sequencing were selected. Protein extracts were generated from these cells and tested by Western blot for expression of Wnk1 kinase (Chapter 2.1.4). Furthermore, RNA was extracted from these cells. After cDNA synthesis using oligo-dT primers (Chapter 2.1.7) I used quantitative real time PCR to detect Wnk1 (oligos #13, #14) mRNA levels.

2.1.7 cDNA synthesis and qPCR

For cDNA synthesis the SuperScript III enzyme was used (Invitrogen, CN 18080). 4 μg of RNA were incubated with DNAse I for 30 minutes at 37^{o}C. The enzyme was then heat inactivated at 75^{o}C for 10 minutes. 1μg of DNase treated RNA was used per reaction. For cDNA synthesis with a gene specific primer, 2μM of the oligo were used, for oligo-dT (Anchored oligo dT, ThermoFisher, AB1247) , 50μM and for random hexamers (ThermoFisher, SO142) 50 ng/ul. The RNA and the oligo were incubated with 1ul 10mM dNTPs at 65^{o}C for 5 minutes. Then, cDNA synthesis mix (5mM $MgCl_2$,

Reverse Transcription buffer, 10mM DTT, RNasin) and 1ul of SuperScript enzyme where added and the reaction was incubated at 50°C for 50 minutes. The reaction was terminated with a 5-minute incubation at 85 °C.

For the qPCR, 2ul from each cDNA synthesis reaction was used. The Real Time PCR reaction contained the required oligos (0.2µM) and SYBR green qPCR mix (Qiagen, CN 208052). Genomic DNA of known concentration was used for generation of standard curve and the calculation of sample concentration.

2.1.8 PCR directed DNA mutagenesis

For PCR directed mutagenesis of Pcf11 CID, I designed DNA oligos bearing the desired mutations: AGC to GCT and ACT to GCC for the S120A T121A mutant, AGC to GAC for the S120E mutant and ACT to GAC for the T121E mutant. These mutated sequences were flanked with 20 base pairs identical to the WT sequence. A reverse and forward version of each primer was used as pairs for the reaction. 50 ng of pcDNA3.1 plasmid bearing the Pcf11 cDNA was used as template for the PCR reaction with high fidelity PrimeStar DNA polymerase (TAKARA). The PCR cycle was 18 cycles long. After the PCR, 10 Units of the restriction enzyme Dpn1 was added directly in the 50ul reaction and incubated at 37°C for 3 hours for the cleavage of the original template.

High efficiency DH5a bacteria were transformed with the mutated plasmid and single colonies were selected the following day. Clones were screened by plasmid extraction and sequencing.

2.1.9 Chromatin Immunoprecipitation

Chromatin preparation

10cm plates containing 5-6 x 10^6 cells where cross-linked with the addition of 1% formaldehyde in the medium. Cells where incubated at 37^{o}C for 15 minutes with mild shaking. 125mM of Glycine was added to stop the cross-linking reaction after which cells where washed 3 times with ice cold PBS and collected by scraping.

Cells where lysed with L1 Lysis buffer (50mM Tris pH 8, 2mM EDTA, 0.1% NP40, 10% Glycerol) with the addition of protease inhibitors. The samples were centrifuged for 5 minutes at 800g and nuclei where collected. The nuclei were resuspended in L2 Lysis buffer (1% SDS, 10mM EDTA, 50mM Tris pH 8) with added protease inhibitors. The samples were sonicated using a Bioruptor (High setting, 30 seconds sonication with 30 seconds intervals). After centrifugation the debris were removed and the supernatant containing the chromatin collected in new tubes and their concentration was measured with nanodrop (an aliquot was kept at this point to verify the sonication efficiency).

Immunoprecipitation

100µg of chromatin was used per IP. The required amount was diluted 10x with Dilution buffer (0.5% NP40, 200mM NaCl, 50mM Tris pH 8, PIC) and pre-cleared with incubation with Dynabeads (1:1 protein A and protein G) for 1 hour at 4^{o}C. After pre-clearing, the beads were removed from the chromatin samples and an aliquot (10%) was kept for use as Input. 5µg of the required antibody were added to each chromatin IP sample and incubated at 4^{o}C on a rotator overnight. Subsequently 30ul of blocked beads (see Bead preparation section) were added to each IP sample and incubated at 4^{o}C for 1 hour.

After the IP the beads were collected using a magnetic rack and washed 3 times with cold wash buffer 1 (150mM NaCl, 0.5% NP40, 2mM EDTA, 20mM Tris pH 8, 0.1% SDS) for a total wash time of 45 minutes, followed by a single wash with wash buffer 2 (250mM NaCl, 0.5% NP40, 2mM EDTA, 20mM Tris pH 8, 0.1% SDS). After the washes, the beads were incubated for 15 minutes at 30^{o}C with elution buffer (1% SDS, 100mM $NaHCO_3$). After elution, beads were removed from the sample.

De-cross linking and DNA extraction

For the de-crosslinking, 1ul of RNase A was added to the IP samples and Inputs and incubated for 4 hours – overnight at 65^{o}C. This was followed with proteinase K treatment (20µg / sample) at 42^{o}C for 1 hour. After de-crosslinking an equal volume of phenol/chloroform was added to the samples. After a 15-minute centrifugation at 13,000 rpm, the aqueous phase was collected and the DNA was precipitated with Ethanol. The samples were then used for Real time PCR.

Bead preparation.

For the preparation of beads, a mix (1:1) of protein A and protein G was prepared and washed with Dilution buffer. The beads were then resuspended in dilution buffer and blocked with 0.5% BSA overnight. Before addition to the IPs, beads were washed with Dilution buffer.

2.1.10 Immunofluorescence staining of HeLa cells

Cells were grown on coverslips coated with poly-L-lysine. 48 hours after transfection the cells were washed 2 times with PBS. Subsequently, they were treated with 0.2% Triton in PBS for 5 minutes for the removal of soluble proteins (Melan and Sluder 1992), followed by fixation for 15 minutes with 4% Paraformaldehyde in PBS. The Triton treatment step was skipped in the experiments were the removal of the soluble proteins was not required. In those cases after the initial washes, cells were directly treated with fixation solution (4% Paraformaldehyde in PBS). The coverslips were then washed with PBS and the cells further permeabilised with 0.5% Triton in PBS.

The cells were blocked in 5% Normal Goat serum, 0.1% Tween in PBS for 1 hour at room temperature. The primary antibodies were diluted in blocking solution and incubated with the cells for 1 hour at room temperature. The coverslips were washed 3 times for 5 minutes in cold PBS and incubated with the secondary antibodies diluted in blocking buffer for 30 minutes at room temperature in a dark chamber. The cells were subsequently washed 2 times

with 0.2% Triton in PBS for 5 minutes and once with 0.1% Triton in PBS for 5 minutes. This was followed with a short wash in PBS and mounting on a slide using ProLong Gold mounting reagent, containing DAPI. The slides were then left to dry for an hour in a dark chamber and were sealed with nail polish.

2.1.11 mRNA Fluorescence *in-situ* hybridization (FISH)

Cells were grown on 6-well plates on poly-L-Lysine covered coverslips. 48 hours after transfection, the coverslips were washed with ice cold PBS. Cells were fixed with 4% Paraformaldehyde (PFA) in PBS for 10 minutes at room temperature and then incubated in 100% Methanol for another 10 minutes. After fixation, 70% Ethanol was added to the cells for 15 minutes to rehydrate the cells followed by a 5 minute wash in 100mM Tris pH 8.

For the preparation of the probe, 1ng/ul oligo dT (T_{23}) conjugated with Alexa488 fluorophore was added to hybridization buffer (yeast tRNA 1mg/ml BSA 0.005% w/v, Dextran sulfate 10% w/v, 25% formamide in 2x SSC buffer). The probe in the hybridization buffer was then pipetted on the coverslip and the samples were sealed in a dark, humidified chamber and incubated overnight at 37^{o}C.

The samples were washed once with 4x SSC followed with two washes for 5 minutes with 2x SSC. The primary antibody was diluted in 2x SSC and 0.1% Triton and dropped on the coverslips for a 1 hour incubation at room temperature. The coverslips were subsequently washed three times with 2x SSC and 0.1% Triton. Then the appropriate dilution of the secondary antibody

was diluted in 2x SSC and 0.1% Triton and pipetted on the coverslips and incubated for 30 minutes at room temperature. The slides were washed for 3 times in 2x SSC. The coverslips were then dried and mounted on the slides using ProLong Gold DAPI containing mounting solution. The slides where left to dry in a dark chamber for 1 hour and then were sealed using nail polish.

2.1.12 Co-Immunoprecipitation

Cells were grown in 15cm plates and harvested at 80% confluency (~ 2 x 10^7 cells) and harvested in PBS. After centrifugation cell pellet was resuspended in 500ul lysis buffer (150mM NaCl, 1% NP40, 50mM Tris pH 8, Benzonase and Protease inhibitors) and incubated at 4^o C for 30 minutes. The lysate was centrifuged for 5 minutes at 13,000rpm to remove lysis debris.

2-5 µg of antibody was added to 200ul of cell extract with the addition of 200ul of 150mM NaCl and incubated overnight on a rotating wheel at 4^oC overnight. 10% of the cell extract was kept to be used as input. Protein A and Protein G DynaBeads were washed in lysis buffer and blocked in 0.1% BSA overnight.

The beads were added to the IP samples and incubated for 1 hour at 4^oC. The beads were collected with centrifugation at 1,500 rpm and washed 3 times with IPH buffer (300mM NaCl, 1% NP40, 50mM Tris pH 8). After the washes the beads were re-suspended in SDS loading buffer and loaded in an acrylamide gel along with different amounts of inputs.

2.1.13 Cytoplasmic and Nuclear fractionation

Cells were collected and washed in ice-cold PBS and pelleted at 500g, 4^{o}C. The cell pellet was re-suspended in 100ul / 2 x 10^{6} cells ice-cold lysis buffer (50mM Tris pH 8, 300mM Sucrose, 4mM MgAc, 12.5mM KCl, 1mM DTT, 10mM β-mercaptoethanol, 1mM PMSF) and incubated on ice for 5 minutes after which 0.3% NP40 was added and incubated on ice for another 5 minutes. The lysates were centrifuged at 500g for 5 minutes at 4^{o}C. The supernatant (Cytoplasmic fraction) was collected in a tube. The nuclear pellet was subsequently re-suspended in 100ul Buffer B (50mM Tris pH 7.5, 25mM KCl, 5mM $MgCl_2$, 0.3% NP40) and incubated on ice for 5 minutes. In a new tube 200ul of a sucrose cushion was added (Buffer B with 0.2M Sucrose) and the nuclei were overlayed slowly on it. The tubes were centrifuged for 5 minutes at 4,500rpm to wash the nuclei and minimize the levels of cytoplasmic contamination. The nuclear pellet was then resuspended in Buffer C (20mM HEPES pH 7.9, 400mM NaCl, 1mM EDTA, 1mM EGTA, 1% NP40) and incubated on ice for 30 minutes for lysis. The samples were centrifuged at 13,000 rpm for 5 minutes to remove the nuclear lysis debris and collect the nuclear lysate in a new tube.

For RNA extraction an equal volume of phenol-chloroform was added to the nuclear and cytoplasmic extracts. Samples were centrifuged at 13,000rpm for 15 minutes and the aqueous phase was moved to a new tube (200ul). 1ml of Trizol and 200 ul of chloroform were then added to the lysate

and the samples were centrifuged for 15 minutes at 13,000 rpm. After the phenol and trizol extraction the RNA was precipitated with 100% Ethanol. RNA pellet was re-suspended in RNase free water and its concentration measured on nanodrop.

2.1.14 Nuclear Envelope fractionation

For the nuclear envelope fractionation experiment, the commercially available Minute Nuclear envelope protein extraction kit was used (Invent biotechnologies, NE-013).

2 x 10^7 cells were harvested and washed with cold PBS and resuspended in 500ul Buffer A (Proprietary composition) with protease and phosphatase inhibitors. Cells were lysed for 10 minutes on ice. The lysate was subsequently applied to a filter column and centrifuged for 30 seconds at 13,000rpm. The supernatant was discarded and the pelleted nuclei were resuspended in PBS followed by a wash with 1 ml PBS. The nuclei were then resuspended in 300ul Buffer B (Proprietary composition) and incubated on ice for 15 minutes with vortexing intervals. The samples were then centrifuged for 5 minutes at 8,000rpm and the supernatant transferred to a new tube. 800ul of cold PBS is then added to the sample and mixed by inverting the tube. This leads to the precipitation of the nuclear envelope. The sample was centrifuged for 15 minutes at 13,000 rpm and the nuclear envelope containing pellet is collected. The pellet was then resuspended in SDS loading buffer and used for Western blotting.

2.1.15 'Heavy' Chromatin detection

6-7 x 10^6 cells were collected 48 hours after transfection. Cells were cross-linked with 1% Formaldehyde for 15 minutes at 37^{o}C and the cross linking reaction stopped with the addition of 125mM Glycine. After collection of the cells with centrifugation (500g for 5 minutes) they were resuspended in Lysis buffer 1 (50mM Tris pH 8, 2mM EDTA, 0.1% NP40, 10% Glycerol) and nuclei collected with centrifugation at 800g for 5 minutes. Nuclei were subsequently resuspended in Lysis buffer 2 (1% SDS, 10mM EDTA, 50mM Tris pH 8) and the samples were sonicated for 15 minutes (Max setting, 30'' ON/OFF). After sonication each sample was split in two tubes one tube was centrifuged at 18,000g for 10 minutes and the other at 2,000g for 10 minutes. From the 18,000g centrifugation both the pellet and the supernatant were kept whereas from the 2,000g centrifugation, the supernatant only was kept. The pellet was resuspended in Lysis buffer 2 and DNA was extracted from both samples using phenol:chloroform. An equal volume of phenol:chloroform was added to the samples followed by centrifugation for 15 minutes at 13,000rpm. The aqueous phase was collected to new tubes and the DNA was precipitated with Ethanol.

The samples were then analysed with Real time PCR using gene specific oligos as described in section 2.1.7.

2.2 Reagents

2.2.1 Oligos

1	GAAAAGTTTATTTAAGTTACGTGCTGCATGGGATGAAATATTCCCTTG	S120A T121A mutation
2	CAAAGGGAATATTTCATCCCATGCAGCACGTAACTTAAATAAACTTTTC	
3	GAAAAGTTTATTTAAGTTACGTGAAACATGGGATGAAATATTCCC	S120E mutation
4	GGGAATATTTCATCCCATGTTTCACGTAACTAAATAAACTTTTC	
5	GTTTATTTAAGTTACGTTCTGAGTGGGATGAAATATTCCCTTTG	T121E mutation
6	CAAAGGGAATATTTCATCCCACTCAGAACGTAACTTAAATAAAC	
7	CACCTGCAGCGGCAAACAACAACAGCATCGATTAAG	shRNA resistance Pcf11
8	CTTAATCGATGCTGTTGTTGTTTTGGGGCTCGAGGTG	
9	AATTTTATCAGTCATCGCTCGAAGACCTGAC	Pcf11 CID amp.
10	TCGAAAGTTTCTTCAAAGGGAATATTTCATCCC	
11	AGCCATCTGTTGTTTGCGGCCGCGCCACCATGTCAGAGCAGACGCCGGC	Pcf11 cDNA amp.
12	CACTTGATGTACTGCGCTAGCGGTGGCTGCAGAAACTGACTCGACTGTGTC	
13	ATGTCTAACGATGGCCGCTTTCTC	Wnk1 Exon1
14	CTGCCTCTCAGACTTTGTTAATTTTCGATC	
15	TCAGACGACGGTAGGAATCAG	Myc 3’
16	CCTGACCAGGCTTAGATGTTG	
17	GATCGCGCTGAGTATAAAAGC	Myc TSS
18	TGCCTCTCGCTGGAATTACTA	
19	CAGAACAACTGCTATGCACCA	Act 3’
20	ACACCAGCAGAGAATGAAGGA	
21	AATCACCTTGCATCTGTTTGG	Gene desert
22	AAAAGGAGAAACCCAGTGGAA	
23	AATCCCTATCTTTAGTCCAATGATGCC	TBP
24	GCTGCGGTACAATCCCAGAACTCT	

2.2.2 Antibodies

a-Pcf11	Ab134391	Abcam
a-Wnk1	A301-515A	Bethyl
a-Pol II (N20)	Sc-899	Santa Cruz
a-Aly	Ab202894	Abcam
a-Aly [11G5]	Ab6141	Abcam
a-GST	Ab19256	Abcam
a-HA	Ab9110	Abcam
a-Lamin B1 Alexa-647	Ab194108	Abcam
a-Lamin B1	Ab8982	Abcam
a-Myc tag	Ab9132	Abcam
a-NPC	Ab24609	Abcam
a-Nxf1	Ab50609	Abcam
a-Pol II [8wg16]	Ab817	Abcam
a-Thoc5	Ab86070	Abcam
a-Thoc7	Ab155218	Abcam
Alexa Fluor 488 goat anti-rabbit	Ab150077	Abcam
Alexa Fluor 594 goat anti-mouse	Ab150116	Abcam

Chapter 3

Phosphorylation of Pcf11 CID by Wnk1 kinase

3.1. Introduction

Pcf11 is part of the CFIIA termination complex in mammals. It is one of the core termination factors that has a role in every step of transcription termination (Cleavage, poly adenylation, Pol II release) (Amrani et al. 1997; Sadowski et al. 2003; West and Proudfoot 2008; Grzechnik et al. 2015). It is recruited to transcribing genes through its interaction with the CTD domain of RNA Pol II. Pcf11 shows a higher affinity for Ser2 phosphorylated CTD with trans isomerized Pro3 (Noble et al. 2005). Once recruited to the transcribing polymerase, it is involved in cleavage of the transcript as well as in the recruitment of the exonuclease Xrn2 and the dissociation of the transcribing polymerase from the DNA template. Apart from it's role in 3' end formation and transcription termination, it has been shown in yeast that Pcf11 is involved in the recruitment of export factor Yra1 (Aly in mammals). It is believed that this interaction provides the necessary cross-talk between the processes of 3' end formation and mRNA export (Johnson et al. 2009; Johnson et al. 2011). The CID domain of Pcf11 is responsible for its interaction with RNA Pol II CTD and is a highly conserved domain located in the N-terminus of the protein. The CID consists of eight α-helices that form a right hand super-helix (Meinhart and Cramer 2004). This structure forms a well-defined hydrophobic pocket in which the residues responsible for the CTD interaction are located.

Wnk1 is a member of a unique family of kinases that lack a characteristic Lysine residue in their catalytic domain. It is a Ser/Thr kinase ubiquitously expressed in metazoans. Alternative splicing and an alternative

TSS are responsible for the expression of several different isoforms some of which are tissue specific (O'Reilly et al. 2003). Wnk1 has been shown to localize primarily in the cytoplasm (Tu et al. 2011). It has been described as an essential protein for the regulation of the ion channels and its mutations cause hypertension (Zambrowicz et al. 2003; Min et al. 2004). Furthermore, Wnk1 localises at the mitotic spindles during mitosis and its depletion causes defects in their formation and chromosome segregation though it is not known how Wnk1 participates in these processes. In mice, knock out of Wnk1 kinase is lethal during development and leads to defective angiogenesis (Xie et al. 2013).

Mass-spectrometry analysis of Pcf11 performed by Dr. Kinga Kamieniarz-Gdula in our lab, identified Wnk1 kinase as an interacting protein. This was an unexpected result as Wnk1 has been shown to localize in the cytoplasm and there have been no results to indicate a nuclear role for the protein. I therefore verified the nuclear localisation of Wnk1 and its interaction with Pcf11 in nuclear extracts.

This interaction of Pcf11 with Wnk1 could imply that Pcf11 is a substrate for phosphorylation. *In-silico* analysis showed that Pcf11 could be a substrate for Wnk1. More specifically S120 and T121 of its CID could be targets for phosphorylation by Wnk1. Further Mass-spectrometry analysis of Pcf11 identified that S120 and T121 of its CID domain are indeed phosphorylated (Data not shown).

Interestingly, S120 and T121 are located in the hydrophobic pocket responsible for the CID-CTD interaction and T121 is one of the residues in

direct contact with the CTD. A potential phosphorylation of these residues could have an effect on this interaction and potentially play a regulatory role in Pcf11 recruitment and transcription termination.

I performed Co-IP experiments to verify the interaction between Wnk1 and Pcf11. Furthermore, Immuno-fluorescence microscopy experiments where used to verify the nuclear localisation of Wnk1. Finally I performed biochemical kinase assays to test whether Wnk1 phosphorylates Pcf11 on S120 and T121.

3.2. Results

3.2.1. Wnk1 interacts with Pcf11

Mass-spectrometry analysis of Pcf11 identified Wnk1 as an interacting partner (Kinga Kamieniarz-Gdula, personal communication, data not shown). This is a surprising result as Wnk1 had only been described in literature to localise in the cytoplasm. *In-silico* analysis of the Wnk1 protein sequence identified two Nuclear Localisation Signals (NLS) suggesting that the protein could potentially be imported to the nucleus (Kinga Kamieniarz-Gdula personal communication). This result was verified with Co-IP experiments of Wnk1 with Pcf11 using nuclear extracts. Wnk1 is co-immunoprecipitated with Pcf11, which verifies the initial observation that the two proteins interact (Kinga Kamieniarz-Gdula, personal communication).

3.2.2. Wnk1 nuclear localisation

The fact that Pcf11-Wnk1 Co-IPs where performed in nuclear extracts and Pcf11 localizes primarily in the nucleus suggests that Wnk1 is present in the nucleus where this interaction takes place. Previous studies using immunofluorescence microscopy have shown that Wnk1 localises primarily in the cytoplasm (Tu et al. 2011). To test these observations I used an α-Wnk1 antibody to study Wnk1 localisation (Bethyl A301-504). A high-resolution microscopy approach was employed to acquire a more sensitive and precise result. It is clear from this experiment that while Wnk1 localises in the cytoplasm, similar levels are detected in the nucleus (Figure 3.1). Surprisingly it shows a strong perinuclear localisation. As a control I used an siRNA against Wnk1 to deplete it from the cells (Figure 3.2) and as expected Wnk1 fluorescent signal is lost. Furthermore as an additional control, I used an antibody against RNA Pol II, which clearly localises exclusively in the nucleus. Although my results contrast the published observation, I believe that this is a result of the plethora of isoforms of Wnk1 in the cell. It is tempting to speculate that certain isoforms of the protein show specific cellular localisation patterns.

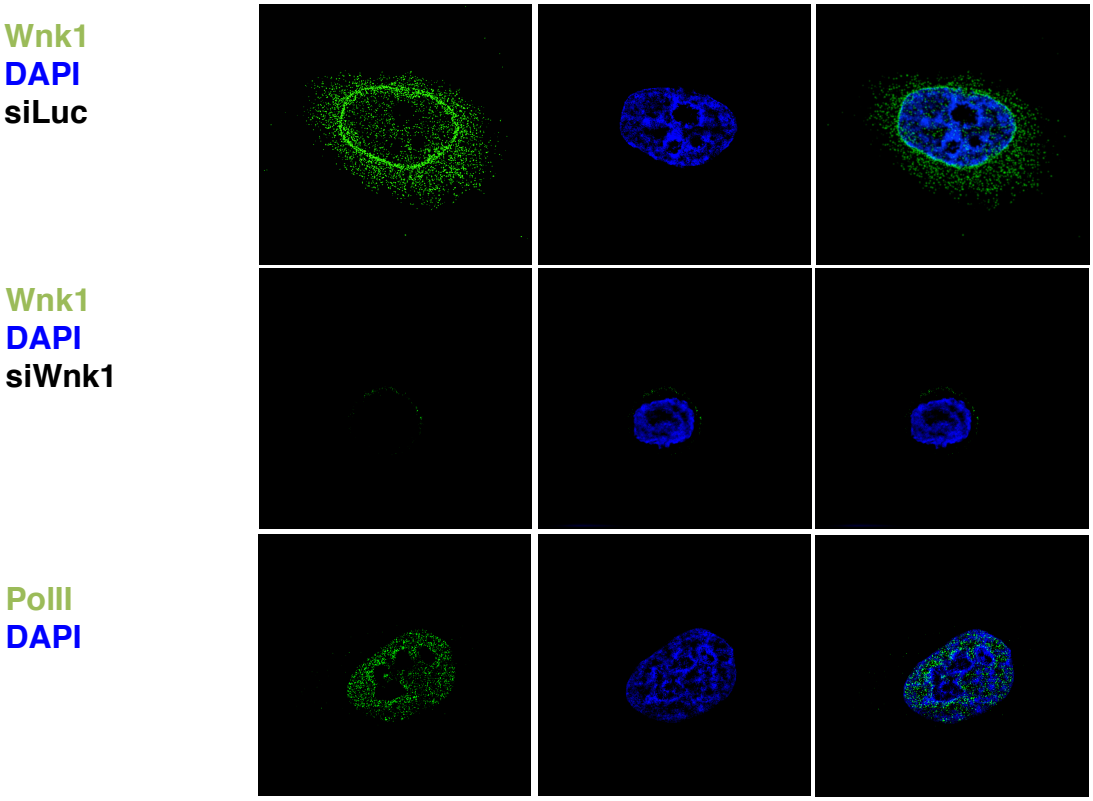


Figure 3.1
Super resolution microscopy and detection of Wnk1 kinase. Upon knock-down of Wnk1 the signal disappears. RNA PolII staining was used as a control for nuclear localisation. The same localisation pattern was observed in all cells imaged (n=20). The imaging was done using OMX-V3 microscope (GE Healthcare,Micron Oxford facility) with a 60x/1.42 NA oil objective (Olympus). The same laser intensity was used for all images. SIM reconstruction was performed with SoftWorx software (Applied precision). Different image channels were aligned using 200nm TetraSpeck beads (Thermo scientific) as reference for the transformation.

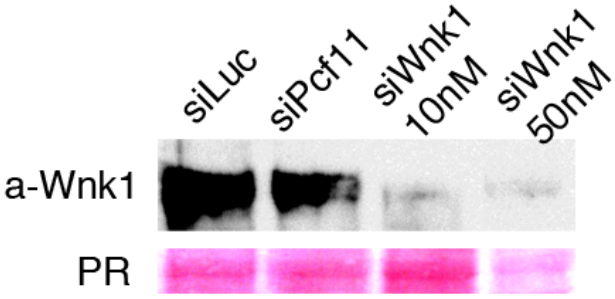


Figure 3.2
Wnk1 knock-down efficiency using an siRNA against Wnk1. siRNA against Pcf11 and against Luciferase were used as a negative controls. Ponceau Red staining of the membrane shown as loading control.

3.2.3. Wnk1 phosphorylates Pcf11 CID

Since Wnk1 interacts with Pcf11 it is tempting to speculate that Pcf11 is a substrate for phosphorylation. Mass-spectrometry analysis of Pcf11 upon pull-down identified that two residues on its CID domain can be found phosphorylated (Kinga Kamieniarz-Gdula, personal communication, data not shown). S120 and T121 of the CID domain are two residues that are part of the hydrophobic pocket where the interaction with the CTD takes place. Furthermore T121 is one of the residues in direct contact with the CTD. *In-silico* analysis of Pcf11 sequence identified that these two residues could be Wnk1 substrate (Kinga Kamieniarz-Gdula, personal communication).

To verify the modification of S120 and/or T121 I generated constructs of the CID domain of Pcf11 for bacterial expression. I expressed and purified in bacteria recombinant GST tagged Pcf11 CID with the Wild Type (WT) sequence, with the mutations S120A, T121A and the double mutant S120A T121A (Figures 3.3, 3.4). I then used the recombinant CIDs to perform an *in-vitro* kinase assay using whole cell extracts of HeLa cells. The recombinant proteins where incubated with the cell extracts and ^{32}p-γATP for one hour at 37^{o}C. After the end of the reaction the GST-CID proteins where pulled down using GST columns followed by electrophoresis in 12% Acrylamide gels. As expected WT CID is indeed phosphorylated (Figure 3.5). Upon mutation of either S120A or T121A a slight decrease in the levels of phosphorylation occurs, but in the double mutant S120A T121A an almost complete loss of the phosphorylation is observed. My results therefore suggest that Pcf11 CID can be phosphorylated *in-vitro* specifically on S120 and/or T121.

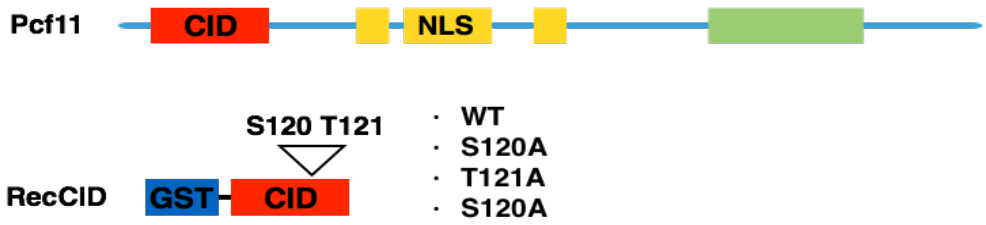


Figure 3.3
Generation of constructs for expression of recombinant Pcf11 CID (RecCID). Pcf11 CID was amplified and cloned in pGEX3.1 bacterial expression vector. Site directed mutagenesis by PCR was used to produce S120A, T121A and S120A T121A mutants.

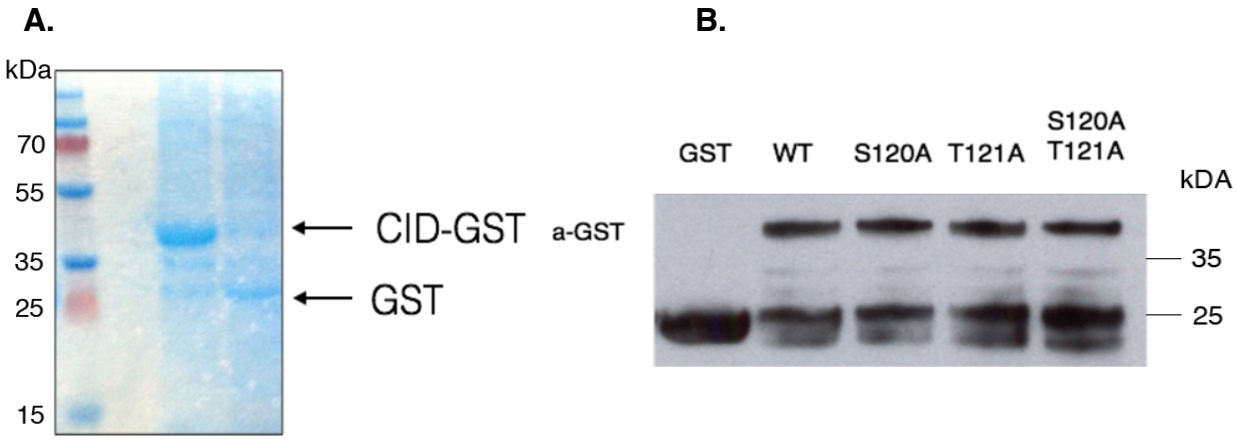


Figure 3.4
A.Expression in bacteria and purification of recombinant GST tagged Pcf11 CID upon IPTG induction. Lower bands correspond to cleavage products (verified by mass-spectrometry). B. Western blot with the purified recombinant proteins using an α-GST antibody.

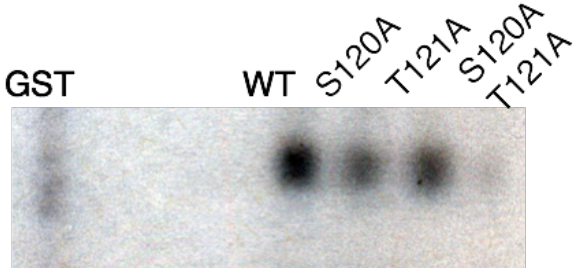


Figure 3.5
In-vitro phosphorylation assay of Pcf11 CID. Recombinant CID was incubated with HeLa whole cell extracts and ^{32}p-γATP for 1 hour. After purification with a GST column the samples were run on a 12% Acrylamide gel. WT CID is clearly phosphorylated whereas the single mutants show a decrease. In the double S120A T121A mutant the phosphorylation is almost completely lost.

The next step was to test whether Wnk1 specifically can phosphorylate S120 and/or T121 of Pcf11 CID. I therefore performed *in-vitro* kinase assays using recombinant Wnk1 kinase domain and recombinant GST tagged Pcf11 CID. The two recombinant proteins were incubated in kinase assay buffer (See Chapter 2, section 2.1.3) with ^{32}p-γATP for 15 minutes followed by electrophoresis on a 12% acrylamide gel. It is clear from the assay that (Figure 3.6) Wnk1 can phosphorylate Pcf11 CID in a dosage dependent manner. Autophosphorylation of Wnk1 which is part of its activation procedure is also visible in this experiment. As a control I repeated the kinase assay and used mass spectrometry to detect the modification of S120 and T121 (Data not shown). As expected phosphorylation of these residues upon incubation of Pcf11 CID with Wnk1 kinase was detectable. To verify that Wnk1 phosphorylates Pcf11 CID specifically on S120 and T121 I performed the same assay using a S120A T121A recombinant CID. It is clear that upon mutation of these residues there is a significant decrease in the phosphorylation levels of the protein (Figures 3.7, 3.8). My results indicate that Wnk1 can phosphorylate Pcf11 CID specifically in S120 and/or T121 *in-vitro.*

The next question I addressed was whether Wnk1 kinase can phosphorylate Pcf11 *in vivo.* To answer this question I generated knock-out cell lines of Wnk1 using the CRISPR/Cas9 system (Mali et al. 2013). I designed and generated guide RNAs (gRNAs) targeting the first exon of the Wnk1 gene (Figure 3.9). We used the D10A version of Cas9 that generates nicks on the DNA. The gRNAs and the Cas9 were transfected in HeLa cells

and single cell clones were selected. To screen for mutants, I isolated genomic DNA from the clones and PCR amplified the 5' end of WNK1 using a forward primer upstream of the first gRNA binding site and a reverse primer downstream of the second gRNA binding site (Oligos #13, #14). The PCR product was subsequently cloned in a PCR TOPOII Vector and sequenced. Several clones with small deletions in this region were identified. These clones were tested by Western blot for Wnk1 expression (Figure 3.10). Three clones showed no detectable Wnk1 protein. I verified this result using quantitative RT PCR to detect Wnk1 mRNA upon oligo-dT cDNA synthesis. The knock-out clones show greatly reduced Wnk1 mRNA (Figure 3.11) which is probably due to the effect of the Nonsense Mediated Decay (NMD) quality control mechanism. Sequencing of clone 2 verified that the knockout is caused by a 61 base pair deletion that generates a STOP codon in the first exon of Wnk1 (Figure 3.12).

To see whether Wnk1 affects Pcf11 phosphorylation *in-vivo* I decided to run an acrylamide gel using the PhosTAG reagent. PhosTAG is a compound that binds phosphorylated residues of proteins resulting in a slower migration on a protein gel. Protein extracts were obtained from HeLa cells and Wnk1 knockout cells. Western blotting using an α-Pcf11 antibody (Abcam ab134391) showed that in the Wnk1 knockout line Pcf11 migrates faster than the WT, suggesting that Pcf11 is less phosphorylated in the knockout line (Figure 3.13). As a control, the same extracts were treated with lambda phosphatase. Upon phosphatase treatment the two samples migrate at the same speed suggesting that the difference that we see in their migration is

caused by phosphorylation of Pcf11. This result indicates that Wnk1 phosphorylates Pcf11 *in-vivo.* Although the PhosTAG approach does not distinguish the specific sites of phosphorylation, this result in combination with the *in-vitro* kinase assays strongly suggests that Wnk1 phosphorylates S120 and/or T121 of the Pcf11 CID.

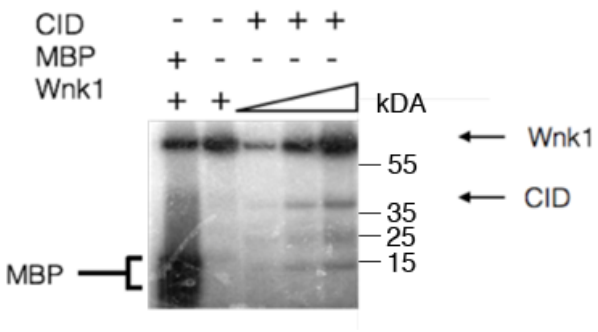


Figure 3.6
In-vitro kinase assay with recombinant Wnk1 and Pcf11 CID. MBP was used as a positive control. Lower bands correspond to cleavage products of the CID-GST protein (mass spec verified).

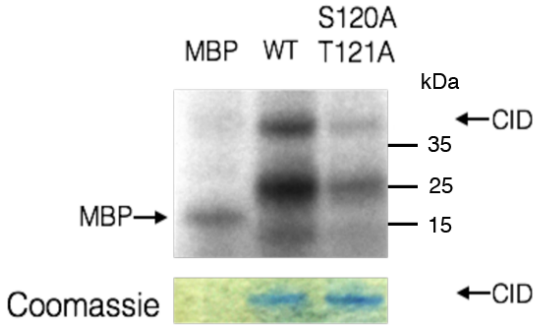


Figure 3.7
In-vitro kinase assay with recombinant Wnk1 and Pcf11 CID. Phosphorylation of WT and S120A T121A Pcf11 by recombinant Wnk1. Lower bands correspond to cleavage products of the CID-GST protein (mass spec verified).

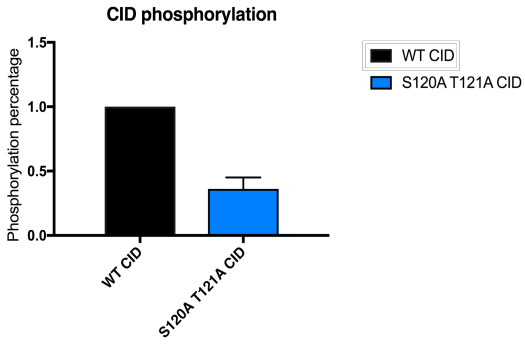


Figure 3.8
Quantitation of phosphorylation levels in the *in-vitro* kinase assay (n=3). There is a 60% drop in phosphorylation levels of S120A T121A CID versus WT CID.

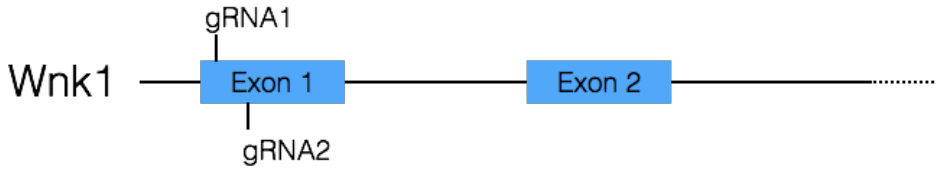


Figure 3.9
gRNAs targeted to the first exon of Wnk1 gene to generate knock-out cell lines.

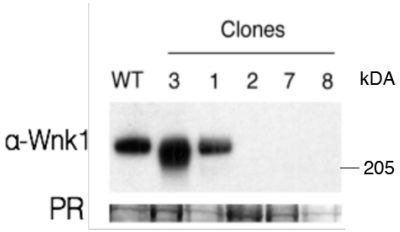


Figure 3.10
Western blot from whole cell extract of the selected clones. Clones 2, 7 and 8 show no Wnk1 expression. Ponceau Red stain was used as loading control.

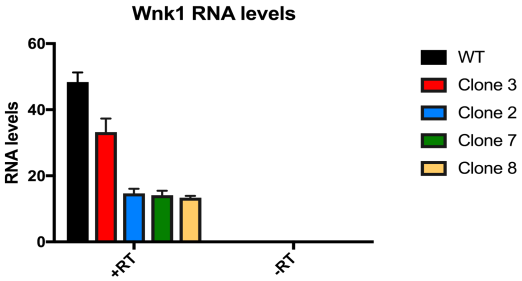


Figure 3.11
mRNA levels of the Wnk1 knock-out clones with no detectable protein levels. Clones 2, 7 and 8 show greatly reduced mRNA levels

```
          *         *         *         *         *         *         *         *         *         *
1>GCGTCGTCCGGGTCGGCGCGAACCCGCCCGGCCGCGGTTCCCTGCAGACCTCTGCGCGGGCGGCTCGGCCCTTCACGCCCTTTTCGTTCACGAATCCGAG>100
1>GCGTCGTCCGGGTCGGCGCGAACCCGCCCGGCCGCGGTGCCCTGCAGACCTCTGCGCGGGCGGCTCGGCCCTTCACGCCCTTTTCGTTCACGAATCCGAG>100

          *         *         *         *         *         *         *         *         *         *
101>CCCGCTCGCCTCTCTCCAGCGAACCGACCATGTCTGGCGGCGCCGCAGAGAAGCAGAGCAGCACTCCCGGTTCCCTGTTCCTCTCGCCGCCGGCTCCTGC>200
101>CCCGCTCGCCTCTCTCCAGCGAACCGACCATGTCTGGCGGCGCCGCAGAGAAGCAGAGCAGCACTCCCGGTTCCCTGTTCCTCTCGCCGCCGGCTCCTGC>200

          *         *         *         *         *         *         *         *         *         *
201>CCCCAAGAATGGCTCCAGCTCCGATTCCTCCGTGGGGGAGAAACTGGGAGCCGCGGCCGCCGACGCTGTGACCGGCAGGACCGAGGAGTACAGGCGCCGC>300
201>CCCCAAGAATGGCTCCAGCTCCGATTCCTCCGTGGGGGAGAAACTGGGAGCCGCGGCCGCCGACGCTGTGACCGGCAGGACCGAGGAGTACAGGCGCCGC>300

          *         *         *         *         *         *         *         *         *         *
301>CGCCACACTATGGACAAGGACAGCCGTGGGGCGGCCGCGACCACTACCACCACTGAGCACCGCTTCTTCCGCCGGAGCGTCATCTGTGACTCCAATGCCA>400
301>CGCCACACTATGG----------------------------------------------------------------GAGCGTCATCTGCGACTCCAATGCCA>339

          *         *         *
401>CTGCACTGGAGCTTCCCGGCCTTCCTCTTT~~~>430
340>CTGCACTGGAGCTTCCCGGCCTTCCTCTTTAAG>372
```

Figure 3.12
Sequencing of clone 2 Wnk1 knock-out. Our CRISPR strategy generated a 61bp deletion on the first exon of the gene that generates a stop codon downstream in the sequence. Differences in sequence from the WT highlighted in red

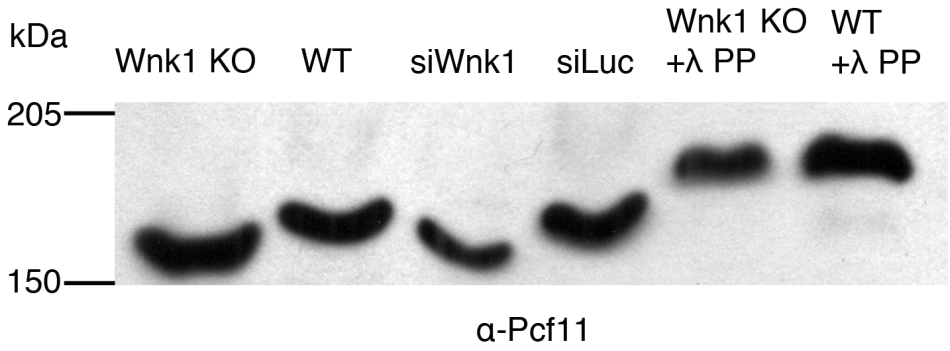


Figure 3.13
Whole cell extracts from Wnk1 knockout cells and WT cells along with Wnk1 knock down cells were run on a PhosTAG acrylamide gel. Absence of Wnk1 results in a faster migration of Pcf11 compared to the WT. As a control, the Wnk1 knock out and WT samples were treated with λ-phosphatase which caused them to migrate at the same speed. This suggests that the difference in migration speed is causes by phosphorylation of Pcf11, which is reduced upon Wnk1 depletion.

3.3 Conclusions

I have identified a new role for Wnk1 kinase in the phosphorylation of the core termination factor Pcf11. Furthermore, I have shown that Wnk1 kinase localises in the nucleus in addition to its previously reported cytoplasmic localisation. Although my results contradict the previously published localisation studies they also provide new insight into Wnk1 and its function. It is known that Wnk1 kinase has a plethora of isoforms generated by alternative splicing or alternative TSS. Many of these isoforms show tissue or cell type specificity. It is tempting to speculate that a particular Wnk1 isoform shows specific localisation to the cytoplasm whereas others may also localise to the nucleus and the nuclear envelope.

I show that Pcf11 is phosphorylated on S120 and T121 of its CID domain. Interestingly, mass-spectrometry experiments performed in our lab by Dr. Kinga Kamieniarz-Gdula, identified Wnk1 as the unique kinase interacting with Pcf11. Unfortunately my experimental approach does not distinguish whether S120 or/and T121 are phosphorylated. According to the structure of Pcf11 CID it is probable that phosphorylation of both residues simultaneously is sterically impossible. Interestingly, these two residues of the CID domain are part of the structure that interacts with the CTD. Specifically T121 is one of the residues in direct contact with CTD. It is possible that modification of these amino acids by phosphorylation could somehow interfere with the CTD-CID interaction. If this is the case it is possible that Wnk1 can be used to regulate the recruitment or eviction of Pcf11 to the transcription locus by modulating its

interaction with RNA Pol II. This hypothesis forms the main focus of the experiments described in Chapter 4.

Chapter 4

Pcf11 CID phosphorylation

4.1 Introduction

Pcf11 CID is the domain responsible for the interaction with RNA Pol II and has a critical role in the recruitment of Pcf11 at transcribing genes. Phosphorylation of S120 and or T121 can potentially have an effect on this interaction and thus play a role in the recruitment of Pcf11.

The results presented in Chapter 3 show that Wnk1 kinase localises primarily in the nuclear envelope but also inside the nucleus. Several studies have shown that actively transcribed genes often associate with the nuclear envelope, localizing close to the Nuclear Pore Complexes (NPCs). This localisation is believed to contribute to a more fast and efficient mRNA export but has been also linked to gene expression memory through chromatin modifications. The association of active genes with the nuclear membrane and the NPC is often required for their optimal expression as it has been shown, predominantly in yeast (Brickner and Walter 2004; Casolari et al. 2005; Taddei et al. 2006). Perinuclear localisation of Wnk1 kinase could provide an insight for Pcf11 phosphorylation and its potential role in gene expression.

In this chapter we investigate how phosphorylation of Pcf11 CID affects the function and localisation of this termination factor. I employed microscopy experiments to test the cellular localisation of phosphorylated and unphosphorylated Pcf11. Furthermore, I used co-immunoprecipitation experiments to determine how this phosphorylation affects the interaction of Pcf11 with RNA Pol II and how this affects its recruitment to transcribed genes.

4.2 Results

4.2.1 Role of CID phosphorylation in Pcf11 localisation

In order to investigate the role of Pcf11 S120 T121 phosphorylation, I decided to use the CRISPR/Cas9 system to mutate these residues to amino acids that cannot be phosphorylated. I designed guide RNAs targeting the Cas9 enzyme to the CID domain of Pcf11 to generate a double strand break. Then I used a single strand DNA oligo with 50bp homologous arms to the genomic DNA with S120A and T121A mutations. After transfections single clone colonies where selected upon serial dilutions and the CID region was amplified with PCR. Unfortunately after screening of more than 100 clones none were found positive for the mutation. This result suggests that this mutation is not viable or it has a severe phenotype for the cells. Similarly, in *S. cerevisiae* Pcf11 is an essential protein, it is probable that in human cells the mutation of Pcf11 is also non viable or shows an extremely severe phenotype. To test this, I used the DCas9/CRISPR system to try and generate insertion/deletion mutations of PCF11 gene (Figure 4.1). After screening of single clone colonies, I identified several mutants with short deletions. These deletions were always 6 or 9 base pair long effectively maintaining the ORF of the gene (Figure 4.2). This result suggests that Pcf11 is an essential gene in humans and that it may be happlo-insufficient. The next approach that I tried was to generate constructs for the ectopic expression of Pcf11. Pcf11 cDNA was amplified and inserted into pEF3-HA vector. The plasmid was transfected to HeLa cells and its expression was tested with western blot 48 hours after

transfection (Figure 4.3). I proceeded to the generation of Pcf11 mutants by PCR directed mutagenesis. With this approach, I generated constructs with mutation of S120 and T121 to Alanine (A) representing forms of Pcf11 that cannot be phosphorylated, as well as mutations to the phosphomimetic amino acid Aspartate (E) that should represent the phosphorylated form of Pcf11 (Figure 4.4). Furthermore, I generated constructs bearing a mutation that gives resistance to an shRNA targeting a specific region of endogenous Pcf11. My strategy was to use these constructs in Chromatin Immunoprecipitation (ChIP) experiments to test whether phosphorylation of Pcf11 affects its recruitment to transcribed genes. I transfected HeLa cells with an shRNA against endogenous Pcf11 along with the plasmids expressing mutant Pcf11. Unfortunately, I was not able to ChIP ectopically expressed tagged versions of Pcf11 even though I employed a variety of tags both on the amino (N-) and carboxyl (C-) terminus of the protein. It is possible that the conformation of the protein with these tags does not allow immunoprecipitation using tag specific antibodies.

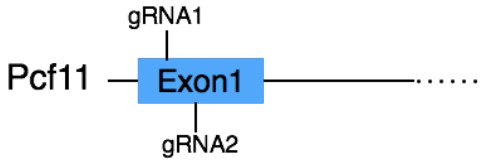


Figure 4.1
Guide RNAs targeting the first exon of the Pcf11 gene.

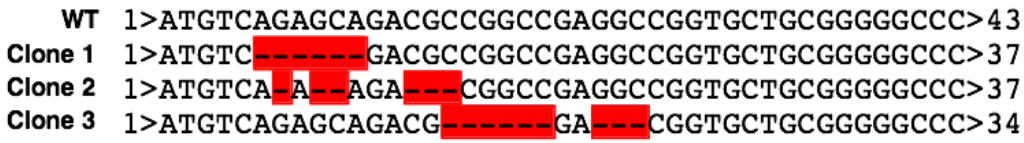


Figure 4.2
Examples of the deletions of endogenous Pcf11 upon CRISPR/Cas9 cleavage and Non-Homologous end joining repair. All positive clones had deletions of 6 or 9 nucleotides within the ORF, resulting in no changes in Pcf11 expression.

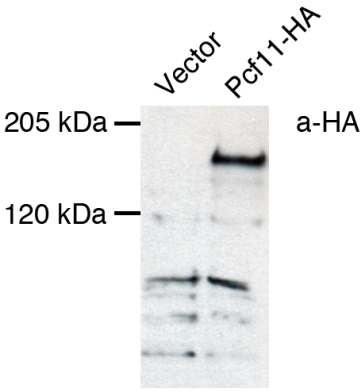


Figure 4.3
Ectopic expression of Pcf11-HA in HeLa cells. Cells were transfected with pEF3-Pcf11-HA plasmid and cell extracts were acquired after 48 hours. Western blot with a-HA antibody (ab9101).

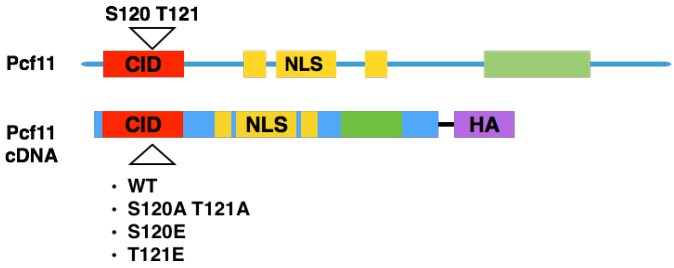


Figure 4.4
Generation of Pcf11-HA mutants. PCR directed mutagenesis was used for the mutation of S120 and T121 to Alanine (S120A T121A) and to Aspartate (S120E, T121E).

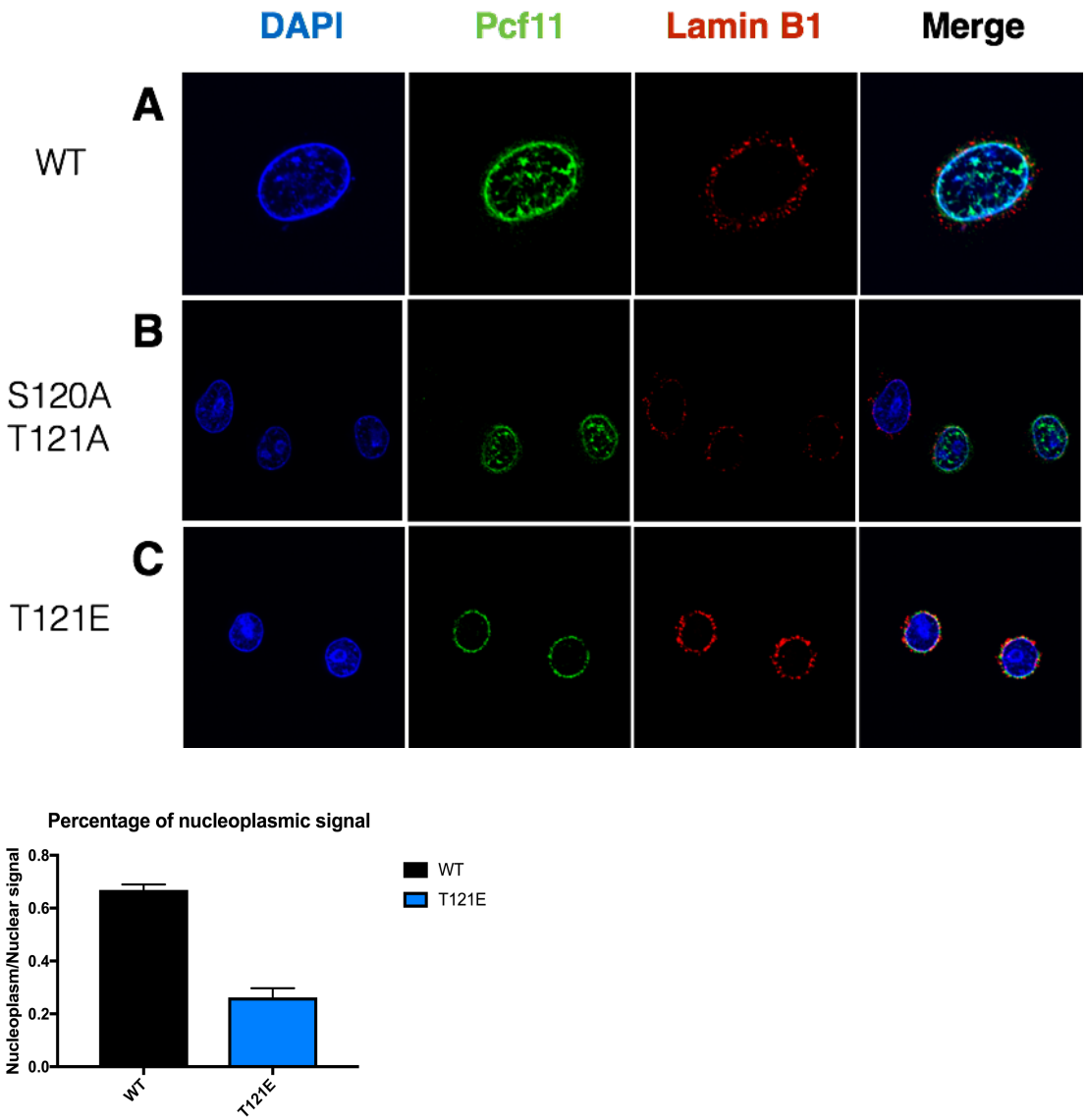


Figure 4.5

Ectopic expression of Pcf11-HA and immunostaining with a-HA and Alexa488 secondary antibody. LaminB1 was used as a nuclear envelope marker stained with a-LaminB1 and Alexa594 secondary antibody. Images were obtained on a confocal microscope system (Fluoview FV1000; Olympus) using a 60x/1.4 NA oil objective and Fluoview software.

A. WT Pcf11 localises in the nucleus in certain foci that could correspond to chromosomal locations. It also shows a perinuclear localisation. **B.** Unphosphorylated Pcf11 localises similarly to WT. **C.** The phosphomimetic mutant T121E shows near exclusive perinuclear localisation.

Quantitation of Pcf11 localisation in WT and T121E cells (N=10). Using FIJI software the integrated intensity of fluorescence was measured in the nucleoplasm and in the nucleus (nucleoplasm and nuclear envelope, LaminB1 was used to mark nuclear envelope). The percentage of Nucleoplasmic/ Nuclear signal is shown on the graph.

To test the localisation of phosphorylated and unphosphorylated Pcf11, I transfected HeLa cells with constructs expressing WT, S120A T121A and S120E, T121E HA-tagged versions of Pcf11. 48 hours after transfection the cells where fixed and stained with an a-HA antibody and an Alexa-488 secondary antibody. Confocal microscopy was used for imaging the localisation of the different Pcf11 mutants. The WT Pcf11 as expected localises exclusively in the nucleus in specific regions that possibly represent chromosomal loci (Figure 4.5a). Interestingly it also shows localisation to the nuclear membrane. Pcf11 with S120A and T121A mutations, representing the unphosphorylated version of the protein shows a similar localisation to the WT (Figure 4.5b). This result suggests that Pcf11 CID phosphorylation is not required for the correct localisation of the protein. Interestingly, the phosphomimetic Pcf11 mutants localise almost exclusively on the nuclear envelope and do not show any levels of localisation on chromosomal loci (Figure 4.5c). This result suggests that the phosphorylation of S120 and/or T121 can potentially affect the recruitment or localisation of Pcf11 possibly through regulating the interaction with RNA Pol II. To verify the nuclear envelope localisation of Pcf11, I obtained separate fractions of the nucleus, cytoplasm and nuclear envelope in HeLa cells. After cytoplasmic lysis the nuclei of the cells were collected, washed and lysed. A commercially available fractionation kit (Minute Nuclear envelope extraction kit, Invent Biotechnologies) was used to precipitate the nuclear envelope fraction. Western blot was then performed with these separate cellular fractions. LaminB1 and Tubulin were used as controls for the fractionations (Figure 4.6).

It is clear from this experiment that endogenous Pcf11 localises in the nucleus and in the nuclear envelope similarly to the ectopically expressed WT Pcf11-HA as shown by the microscopy images (Figure 4.5).

4.2.2 CID phosphorylation and Pol II CTD interaction

Our original hypothesis for the phosphorylation of Pcf11 CID was that the modification of S120 and/or T121 due to their critical position in the domain, could affect the interaction with the polymerase. To test this idea, I performed co-immunoprecipitation experiments using ectopically expressed HA-tagged Pcf11. HeLa cells were transfected with the plasmids expressing WT-Pcf11-HA and S120E-Pcf11-HA and an a-HA antibody was used for immunoprecipitation from whole cell extracts. Followed by Western blot using a-HA and a-Pol II antibodies. In the case of WT Pcf11 we can detect a clear co-immunoprecipitation of Pcf11 with Pol II (Figure 4.7) However, in the phosphomimetic mutant S120E this interaction is abolished. To further verify this result pull-down experiments were performed using phosphorylated and unmodified synthetic peptides of Pcf11 CID (experiment performed by Dr. Kinga Kamieniarz-Gdula). A Western blot was performed using a-Pol II antibodies. Again, the unmodified CID peptide interacts with the polymerase whereas in the modified peptides this interaction is lost (Figure 4.8). These results were further verified using mass-spectrometry analysis after the peptide pull-down. Mass-spec results verify our observations, as it is clear that Pol II preferentially interacts with the unmodified Pcf11 CID peptides (Figure 4.9).

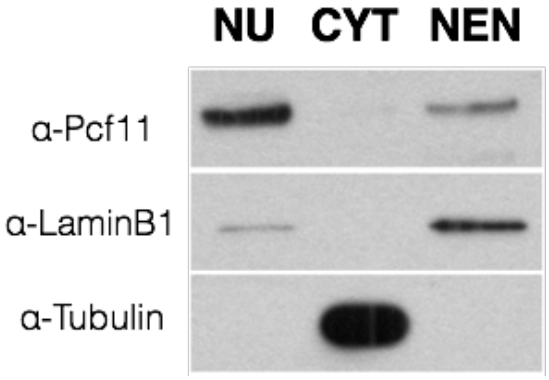


Figure 4.6
Western blot with Nuclear (NU), Cytoplasmic (CYT) and Nuclear envelope (NEN) protein fractions. Pcf11 localises is present in the nuclear fraction as well as in the nuclear envelope fraction. LaminB1 and Tubulin were used as controls for the fractionations.

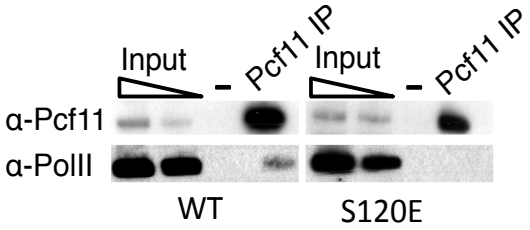


Figure 4.7
Co-IP experiment of Pcf11 with RNA Pol II. Ectopically expressed WT and S120E Pcf11 was immunoprecipitated using a-HA antibody. This was followed by Western blot with a-HA and a-Pol II antibody (N20).

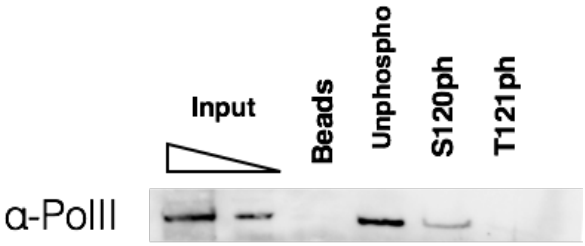


Figure 4.8
Pull-down using Pcf11 CID peptides and Western blot of eluates. Unmodified Pcf11 peptide interacts with RNA Pol II (N20). This interaction is greatly reduced in S120 phosphorylated peptide and abolished in T121 phosphorylated peptide

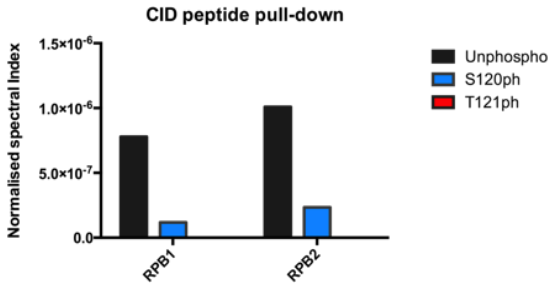


Figure 4.9
Mass spectrometry of eluates upon Pcf11 CID peptide pull-down. Spectral indexes of Pol II subunit peptides are shown. The unmodified Pcf11 CID peptide strongly interacts with Pol II. In the phosphorylated peptides this interaction is abolished.

4.2.3 CID phosphorylation and Pcf11 recruitment

The next step was to see how Pcf11 CID phosphorylation affects the recruitment of Pcf11 to the chromatin. To address this, I used an siRNA against Wnk1 kinase and tested the localisation of Pcf11 on chromatin by ChIP analysis. I probed for the MYC gene to detect recruitment of Pcf11 since experiments in our lab (Kinga Kamieniarz-Gdula, personal communication) have identified the localisation of Pcf11 in the 3' end of the gene in high levels (Data not shown). ChIP analysis was performed 48 hours following transfection of the cells with siWnk1. It is clear from these results that upon depletion of Wnk1 kinase, Pcf11 recruitment is increased in the 3' end of MYC (Figure 4.10) and also on the TSS region. This result is in agreement with the Co-IP and immunostaining experiments that suggest that phosphorylation of Pcf11 restricts its recruitment to the DNA possibly by affecting its ability to interact with RNA Pol II. The observed increase in the recruitment of Pcf11 to MYC upon Wnk1 depletion may be caused by increased Pol II occupancy or possibly by altered transcriptional activity. To address this possibility, I performed Pol II ChIP experiments upon depletion of Wnk1 with siRNA. My results suggest that Pol II occupancy at the MYC gene does not change along the gene. Thus the increased Pcf11 levels that we detect are not due to increased Pol II levels (Figure 4.11).

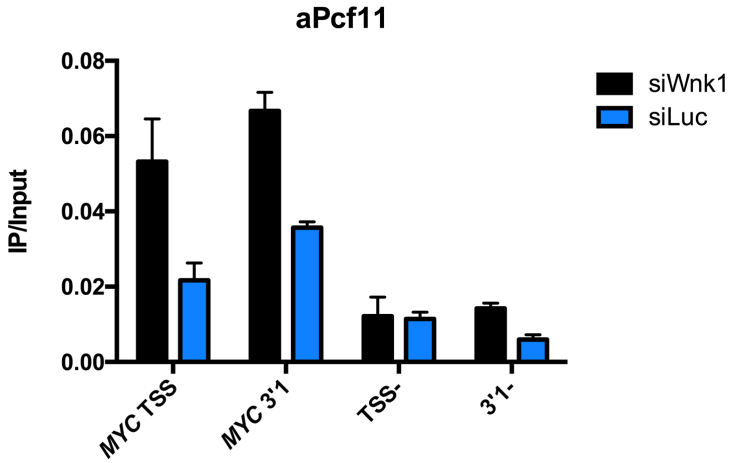


Figure 4.10
Chromatin immunoprecipitation of Pcf11 upon Wnk1 knock down. Pcf11 occupancy is increased at the 3' end and at the TSS of the MYC gene. No antibody (TSS-, 3.1-.), beads only samples for each region used as negative controls. Error bars represent standard deviation.

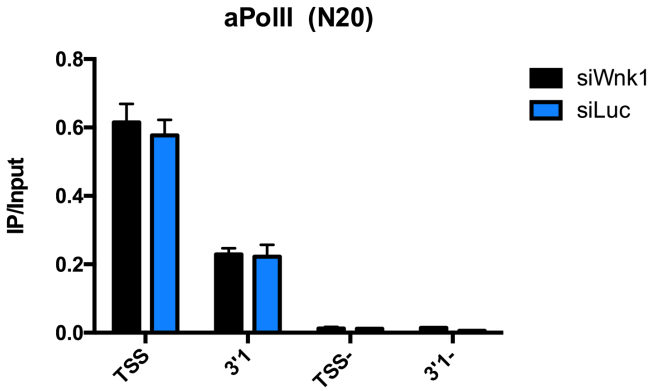


Figure 4.11
Chromatin immunoprecipitation of RNA Pol II upon Wnk1 knock down. TSS and 3' end regions of MYC gene were tested. We have no detectable differences in RNA Pol II recruitment at the MYC locus upon Wnk1 depletion. No antibody (TSS-, 3.1-.), beads only samples for each region used as negative controls. Error bars represent standard deviation.

4.3 Conclusions

My results show that Pcf11 localises exclusively in the nucleus of the cell but it also shows a perinuclear localisation. This is in accordance with several studies showing the association of actively transcribed genes with the nuclear periphery. Perinuclear localisation is believed to increase the efficiency and speed of mRNA export as well as playing a role in the transcriptional memory of certain genes. DNA elements are required for the association of genes with the NPCs. RNA processing factors and export factors are often required for the stabilization of this interaction in yeast (Abruzzi et al. 2006; Dieppois et al. 2006; Taddei et al. 2006). Interestingly the phosphomimetic mutant of Pcf11 (S120E, T121E) shows an exclusive localisation to the nuclear periphery. Phosphorylation of certain proteins can sometimes affect their import to the nucleus, although in the case of Pcf11 this does not seem to be the case, since the unphosphorylated mutant (S120A T121A) behaves like the WT and does not show any signs of cytoplasmic retention. I suggest that the phosphorylated Pcf11 localises to the nuclear envelope due to its inability to be recruited at genomic loci. I have shown that phosphorylated Pcf11 is unable to interact with RNA Pol II suggesting that this modification may be used for the regulation of its recruitment or release from the transcription locus. I have further verified this result by showing that Pcf11 levels are increased in the 3' ends of the MYC gene upon depletion of Wnk1. This increase is not due to increased Pol II association. It is tempting to speculate that Wnk1 phosphorylation is somehow

required for the release of Pcf11 from the 3' end of genes, possibly through the disruption of the Pcf11 CID – Pol II CTD interaction.

Chapter 5

Role of Pcf11 CID phosphorylation

5.1 Introduction

I have demonstrated, that phosphorylated Pcf11 displays reduced interaction with RNA Pol II. In accordance with my original hypothesis, phosphorylation of S120 and/or T121 abolishes the interaction with the CTD domain of Pol II (Figure 4.7). The perinuclear localisation of phosphorylated Pcf11 together with Wnk1 hints that their function may be related to mRNA export. Indeed, Pcf11 has been implicated in mRNA export in *S. cerevisiae* through the recruitment of the export factor Yra1 (Johnson et al. 2009). It is also believed that recruitment of Yra1 can regulate 3' end formation in yeast through its competition with Pcf11 co-factor Clp1 (Johnson et al. 2011). It is tempting to speculate that Pcf11 may play a similar role in human cells, providing the cross-talk between the processes of transcription termination and mRNA export. For efficient mRNA export, export factors are recruited during transcription through interactions with transcription factors. Export factors are then loaded on the newly synthesized mRNA, which is then guided to the NPC and exported to the cytoplasm (reviewed in (Bjork and Wieslander 2014). The fact that phosphorylated Pcf11 does not interact with RNA Pol II suggests that Wnk1 acts downstream of the process of transcription.

In this chapter, I investigate the role of Pcf11 CID phosphorylation by Wnk1. I use qPCR and RNA fish approaches to show that this modification has an effect on mRNA export. Next, I try to uncover the mechanism through which Wnk1 regulates mRNA export. I use ChIP and Co-IP experiments to analyse the recruitment of export factors and their association with Pcf11.

Finally, I propose a model for the role of Wnk1 kinase in transcription termination and mRNA export.

5.2 Results

5.2.1 Role of Wnk1 in mRNA export

The localisation experiments on Wnk1 kinase and Pcf11 gave a hint to the potential role of Pcf11 CID phosphorylation. The fact that Wnk1 shows a strong localisation on the nuclear envelope suggests that its function may be related to mRNA export. To test this hypothesis I performed nuclear and cytoplasmic fractionations of cells following Wnk1 depletion. I transfected cells with an siRNA against Wnk1 kinase and after 48 hours extracted total RNA from the nuclear and cytoplasmic fractions and performed cDNA synthesis using an oligo-dT primer. Gene specific oligos were employed to measure the mRNA levels of the MYC and TBP genes (Figure 5.1). It is clear that upon depletion of Wnk1 there is a decrease in the cytoplasmic levels of both mRNAs whereas in the nuclear fraction a minor increase of the mRNA levels is apparent. These results point to a role for Wnk1 in mRNA export although I cannot be sure whether it is a direct effect of the kinase through the phosphorylation of Pcf11. To further verify this result I performed an mRNA Fluorescence In Situ Hybridization (FISH) experiment. HeLa cells were transfected with an siRNA for Wnk1 depletion and harvested after 72 hours. A fluorescent oligo-dT probe was used for the detection of total mRNAs in the cell. Cells were incubated with the probe for 24 hours followed by confocal microscopy imaging. Control cells, transfected with an siRNA against Luciferase show localisation of mRNAs in the cytoplasm and in specific nuclear foci (Figure 5.2). Upon Wnk1 depletion we see a clear decrease in the accumulation of mRNAs in the cytoplasm and a corresponding increase of

mRNA levels in the nucleus. The observed changes in nuclear/cytoplasmic mRNA levels upon depletion of Wnk1 strongly suggest that Wnk1 has a general role in mRNA export.

My results show a clear connection of Wnk1 kinase with mRNA export. Depletion of Wnk1 results in a strong and general defect in nuclear export. The next question that I tried to address was whether this defect is an independent function of Wnk1 or if it is related to the phosphorylation of Pcf11 CID. To address this question, I performed mRNA FISH in cells ectopically expressing WT Pcf11 or S120A T121A Pcf11 tagged with HA. Cells were transfected with the shRNA resistant Pcf11 expression plasmids and simultaneously with an shRNA against endogenous. This strategy was intended to cause depletion of endogenous Pcf11 and the simultaneous expression of the mutant Pcf11 in Wnk1 depletion conditions. The same fluorescent oligo-dT probe was used in combination with a-HA staining to select for cells expressing the mutant Pcf11. Expression of WT Pcf11 does not change the profile of mRNA levels and localisation (Figure 5.3) similarly to Figure 5.2 we can see mRNA in the cytoplasm and in specific foci in the nucleus. When unphosphorylated Pcf11 (S120A T121A) is expressed an mRNA export defect is apparent as the levels of mRNA in the cytoplasm are reduced similarly to the Wnk1 depleted cells (Figure 5.2). The above results suggest that Wnk1 participates in the process of mRNA export, through the phosphorylation of Pcf11 CID. Furthermore I show that the S120 and T121 residues are important for mRNA export, possibly by modulating the association of Pcf11 with RNA Pol II.

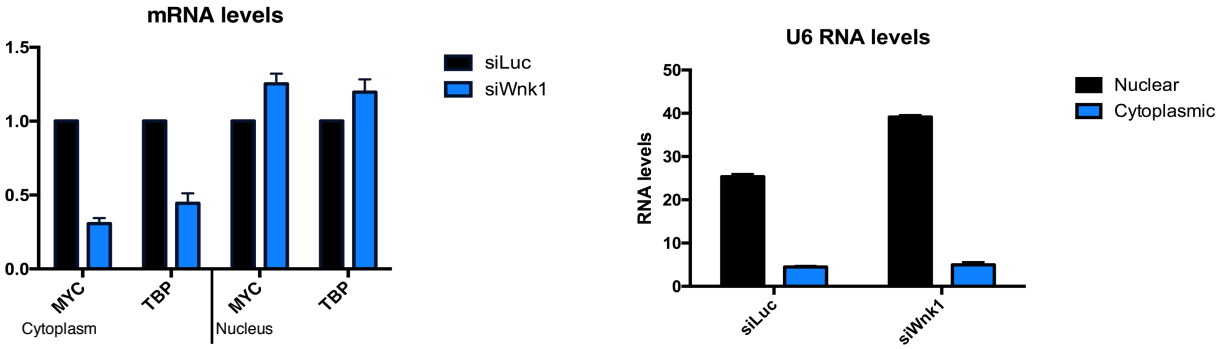
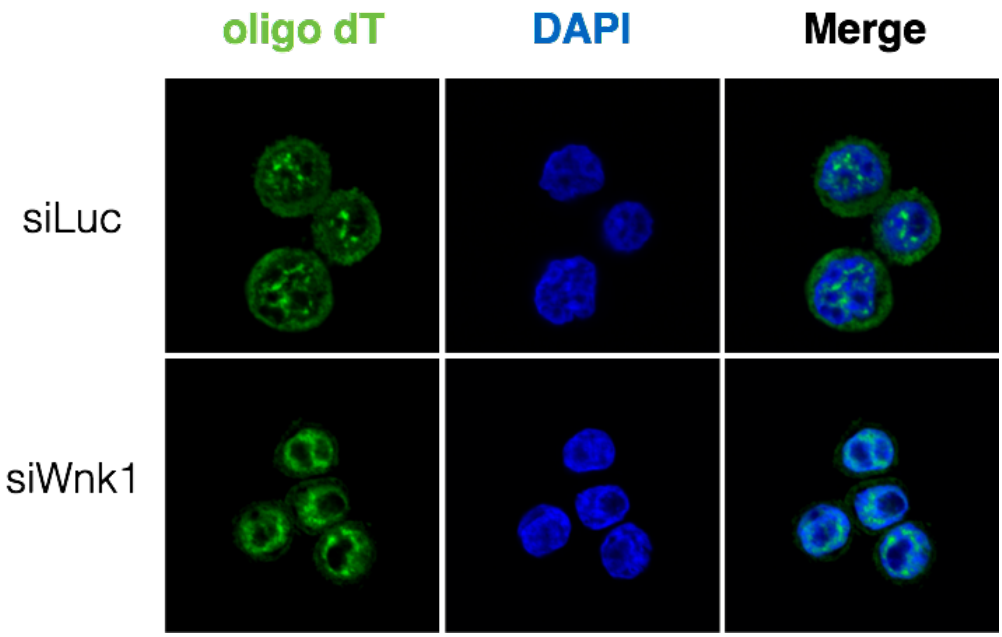


Figure 5.1
mRNA levels of MYC and TBP genes in nuclear and cytoplasmic extracts. Upon depletion of Wnk1 the cytoplasmic levels of mRNA are reduced whereas there is an increase in the nuclear levels.
U6 RNA levels were measured in the same samples for the verification of the fractionation quality.



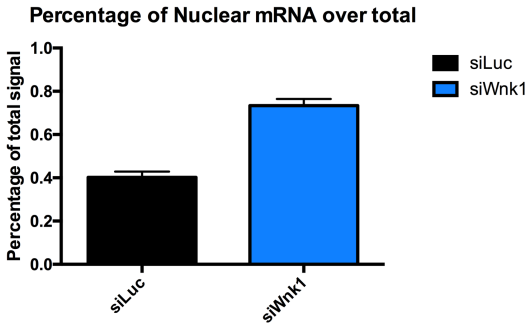
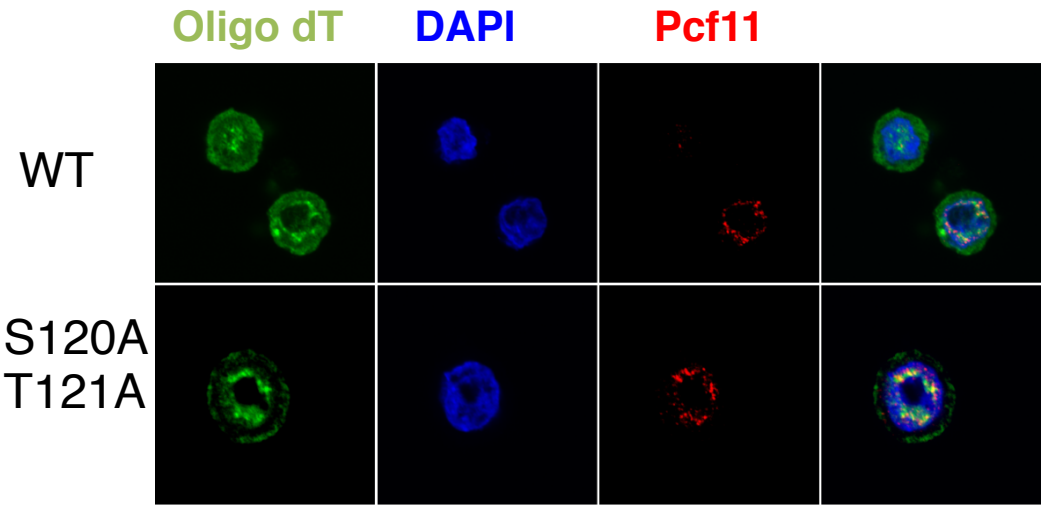


Figure 5.2
Fluorescent oligo-dT FISH for the detection of mRNA levels. In the control cells (upper panels) we can detect mRNAs primarily in the cytoplasm and in specific foci in the nucleus. In Wnk1 depleted cells (lower panels) the cytoplasmic mRNA levels drop and there is an increase in the nuclear mRNA levels.
Quantitation of percentage of nuclear localised mRNA over total. The integrated intensity of the nuclear fluorescent signal and the total fluorescent signal was calculated for every cell (n=30) and the ratio between the two is presented in the graph.



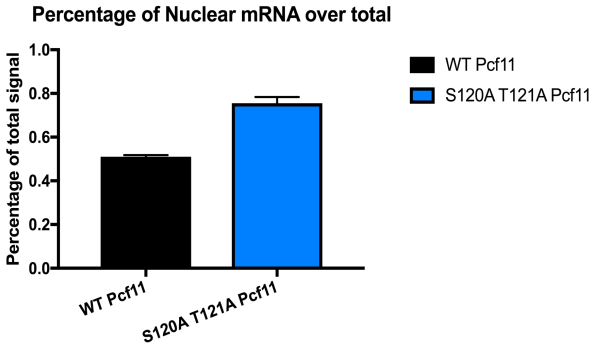


Figure 5.3
Fluorescent oligo-dT FISH for the detection of mRNA levels in HeLa cells. Ectopically expressed Pcf11-HA was stained with a-HA antibody and Alexa594 secondary antibody. Cells expressing WT Pcf11 show normal mRNA distribution with high levels in the cytoplasm and in the nucleus in specific foci. Cells expressing the unphosphorylated mutant S120A T121A of Pcf11 show an mRNA export defect with a decrease in cytoplasmic mRNA levels.
Quantitation of percentage of nuclear localised mRNA over total. The integrated intensity of the nuclear fluorescent signal and the total fluorescent signal was calculated for every cell (n=10) and the ratio between the two is presented in the graph.

5.2.2 Wnk1 and export factor recruitment

Although it is clear from these experiments that phosphorylation of S120 and/or T121 by Wnk1 is required for mRNA export, the mechanism by which this happens remains elusive. In yeast, it has been shown that Pcf11 is required for the recruitment of the export factor Yra1 (Aly in mammals) (Johnson et al. 2009). Yra1 interacts with Pcf11 through a domain downstream of the CID, the same domain that is responsible for the interaction of Pcf11 with its co-factor Clp1. Furthermore, the recruitment of Yra1 can regulate 3' end formation by competing with the Clp1 for Pcf11 interaction (Johnson et al. 2011).

I initially investigated whether Pcf11 phosphorylation can affect the recruitment of export factors to the DNA. Therefore, I performed ChIP experiments against the export factor Aly, following depletion of Wnk1. I tested the TSS and 3' ends of the β-Actin and cMYC genes, a gene desert region was used as a negative control. In the case of Actin, I detect an increase in the levels of Aly towards the 3' end of the gene but no effect upon Wnk1 depletion (Figure 5.4). Similarly, recruitment of Aly at the MYC gene is not affected by the absence of the kinase. An interesting observation in this experiment is that the export factor Aly appears to be recruited early on the genes. This observation is contradictory to yeast where recruitment of Yra1 can compete with Clp1 for their interaction with Pcf11 and thus have a negative effect on transcription termination (Johnson et al. 2011). In human cells, this mechanism may differ and Aly may be recruited earlier in the transcription process.

Another approach to determine whether phosphorylation by Wnk1 can affect recruitment of export factors was to use mutant Pcf11 and test whether unphosphorylated Pcf11 interacts with Aly. I therefore ectopically expressed WT and the phosphomimetic Pcf11-HA followed by co-immunoprecipitation experiments. It is apparent that WT Pcf11 interacts with the export factor Aly as it has been previously reported using an *in-vitro* approach (Figure 5.5) (Johnson et al. 2009). This interaction appears unaffected in the phosphomimetic Pcf11 mutant. This result is in accordance with previous studies that identify the interacting domain of Pcf11 with export factor Aly (Johnson et al. 2009). According to this study Pcf11 and Aly interact through the Clp1 interacting domain with no participation of the CID domain. My result suggests that phosphorylation of the CID domain of Pcf11 does not affect recruitment of export factors at the transcribed genes. Furthermore, I tested by the same approach whether phosphorylation of Pcf11 affects its interaction with NPC. As expected the WT Pcf11 interacts with NPC but this interaction is not affected by phosphorylation. The above results are in accordance with my localisation experiments showing that Pcf11 is associated with the nuclear envelope (Figures 4.5, 4.6). Surprisingly phosphorylation of Pcf11 CID does not affect localisation and interaction with the NPC. This result suggests that Pcf11 phosphorylation affects mRNA export through a different mechanism and not by affecting its interaction with export factors and the NPC but possibly through the release of the mRNP from the polymerase.

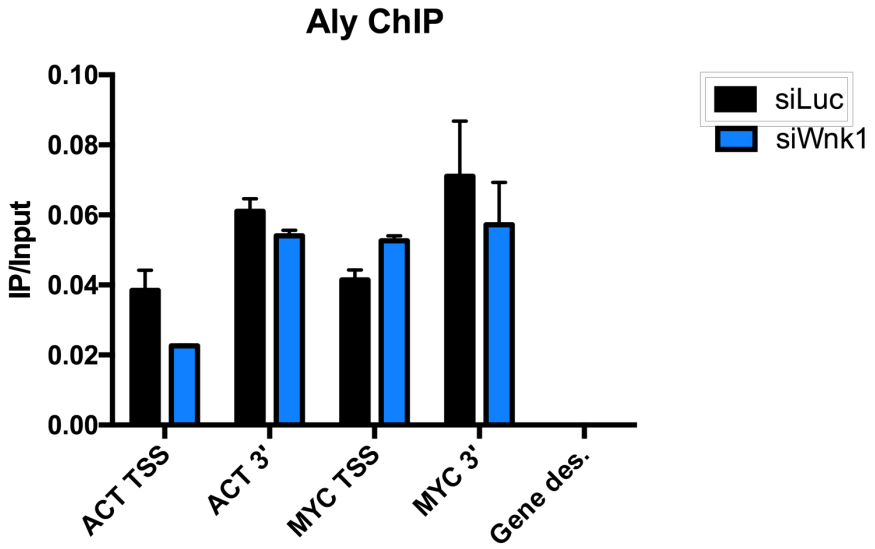


Figure 5.4
Chromatin immunoprecipitation experiment upon Wnk1 depletion in HeLa cells. Aly is recruited at the 5' ends of Actin and MYC genes but it's occupancy increases towards the 3' ends of the genes. Wnk1 depletion does not seem to affect recruitment of Aly. Error bars represent standard deviation.

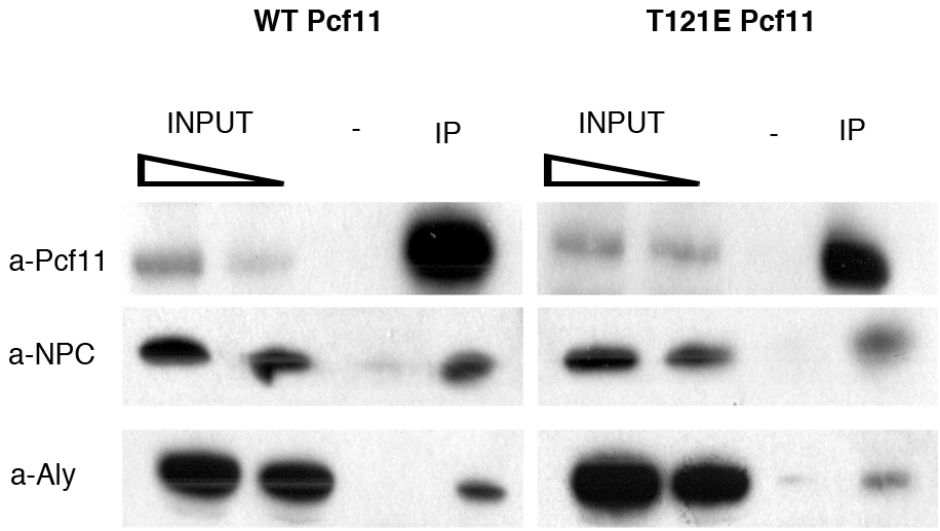


Figure 5.5
Co-IP experiment with WT and S120E phosphomimetic Pcf11. Pcf11 interacts with the NPC but this interaction does not change between the phosphomimetic and WT Pcf11. Similarly Pcf11 interacts with the export factor Aly but the interaction is not affected by phosphorylation of Pcf11 CID.

5.2.3 Mechanism of mRNA export regulation by Wnk1

To further investigate the mechanisms through which Pcf11 CID phosphorylation might affect mRNA export, I decided to use an approach previously described in yeast for the identification of factors involved in mRNA export (Rougemaille et al. 2008). Depletion of factors that are involved in mRNA 3' end formation but also have an effect on mRNA export can lead to the formation of 'heavy' chromatin. Essentially, if somehow the cross-talk between transcription termination and mRNA export is disrupted then the mRNP cannot be released from the DNA and this leads to the formation of 'heavier' fractions of chromatin in the 3' ends of genes. I depleted Wnk1 from HeLa cells and extracted DNA from the supernatant and pellet after an 18,000g centrifugation as described in Figure 5.6. Using qPCR, I observe that upon Wnk1 knock down there is an accumulation of DNA in the pellet fraction, as the ratio between Supernatant and pellet dramatically drops, suggesting that Wnk1 depletion leads to an export defect possibly due to failure to release the mRNP from the chromatin (Figure 5.7). Surprisingly, in this "heavy" chromatin fraction, I detected not only the 3' end of the MYC gene but also the TSS although in lower levels. A possible explanation for this result could be the association of the 5' and 3' end of the gene through a looping mechanism (O'Sullivan et al. 2004; Ansari and Hampsey 2005; Singh and Hampsey 2007).

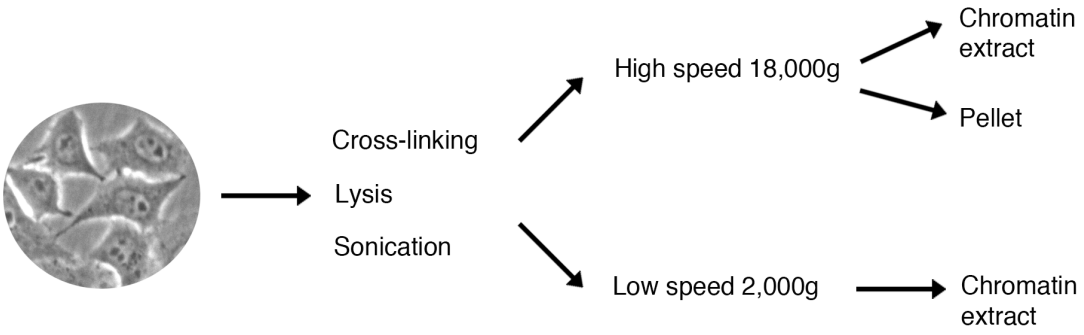


Figure 5.6
Schematic representation of the "heavy" chromatin detection experiment. 48 hours after transfection with an siRNA against Wnk1, cells were cross-linked, lysed and the extracts were sonicated. Samples were divided between two tubes and centrifuged at 18,000g and 2,000g respectively. Chromatin was extracted from the pellet and supernatant of the 18,000g sample and chromatin from the 2,000g sample. Genomic regions present in the 18,000g pellet represent 'heavy' chromatin which presumably is heavier due to it's packing with proteins.

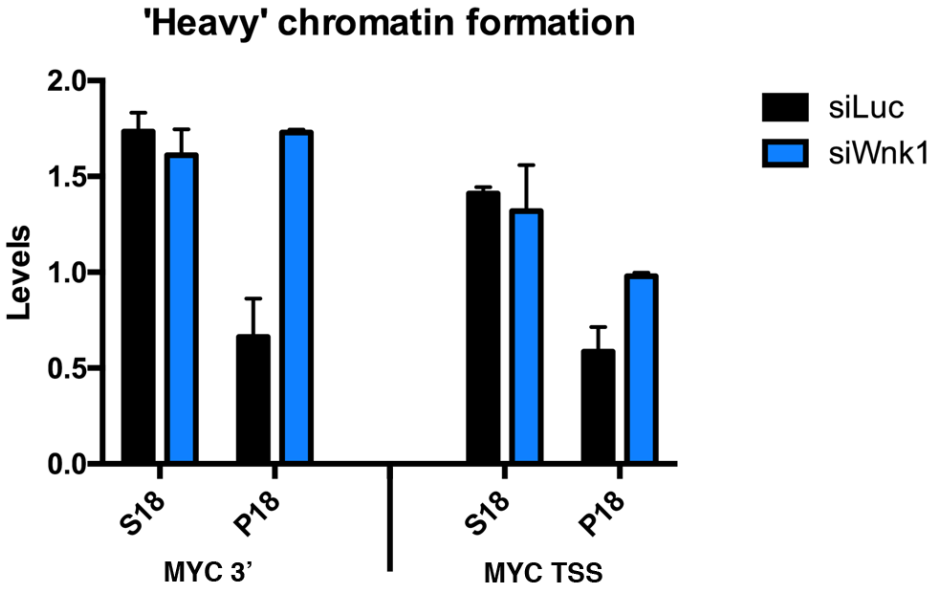


Figure 5.7
Levels of MYC regions in the supernatant (S18) and pellet fractions (P18) after 18,000g centrifugation normalized to levels in supernatant of 2,000g centrifugation. The MYC 3' end region is present primarily in the S18 fraction in the control sample (siLuc). Upon Wnk1 depletion MYC 3' end levels increase in the P18 fraction. Wnk1 knock-down causes the formation of 'heavy chromatin' with which MYC 3' end region is associated. The MYC TSS region shows a similar pattern as the 3' end. In the control sample the majority of the DNA is present in the S18 fraction. Upon Wnk1 depletion there is an increase of the levels in the P18 fraction reminiscent of 'heavy' chromatin formation. Error bars represent standard deviation.

5.3 Conclusions

In this chapter I have investigated the role of Pcf11 CID phosphorylation by Wnk1. I show that phosphorylation of these two residues is required for efficient mRNA export. This result strengthens the idea that Pcf11 provides the 'cross-talk' between mRNA 3' end formation and export. In yeast, Pcf11 is one of the transcription factors involved in the initial recruitment of export factors and is thought to provide a platform for the co-ordination of the processes of termination and export through its competitive interaction with export factor Yra1 (Aly in mammals) and its co-factor Clp1. My results suggest that in human cells there is an extra layer of regulation for transcription termination and mRNA export through the phosphorylation of Pcf11 CID. Interestingly S120 and T121 of Pcf11 CID are only present in higher eukaryotes suggesting that this mechanism of regulation has emerged later in evolution.

I have shown that recruitment of export factor Aly is not affected by Pcf11 CID phosphorylation. Furthermore, this modification does not affect the association of Pcf11 with the NPC. These results suggest that the regulation of mRNA export through Pcf11 CID phosphorylation is independent of the recruitment of export factors indicating a different mechanism. Failure to phosphorylate Pcf11 CID leads to the formation of 'heavy' chromatin. This type of chromatin is thought to consist of DNA in a more condensed form resulting either through chromatin modifications or from the tight association of proteins with the DNA. This result, in combination with the fact that phosphorylated Pcf11 is not capable of interacting with RNA Pol II, provide a

hint for the potential mechanism of this regulation. My data can be explained by a simple model in which following the termination of transcription, phosphorylation of Pcf11 CID by the kinase Wnk1 is required to weaken the interaction of the termination factor with the polymerase. This allows the disassembly of the transcription complex and the consequent release of the mRNP which is then exported to the cytoplasm through the NPC.

Chapter 6

Discussion

6.1 Pcf11 post-translational modifications

Pcf11 is one of the core transcription termination factors in eukaryotes. In humans, it is part of the cleavage factor II complex and it has been shown to have a role in every step of transcription termination and 3' end formation (Amrani et al. 1997; Noble et al. 2005; West and Proudfoot 2008; Grzechnik et al. 2015). Recent studies in human cells show that Pcf11 also plays a role in pA site selection and suggest that its recruitment may be the rate limiting step in pA-dependent transcription termination (Li et al. 2015). Pcf11 is recruited to transcribed genes through its CID domain that directly interacts with the CTD of RNA Pol II preferentially when the Ser2 of the CTD repeats is phosphorylated. Mass-spectrometry analysis of Pcf11 (Kinga Kamieniarz-Gdula, unpublished results, personal communication) identified that two residues of the CID domain of Pcf11 are phosphorylated. Interestingly these two residues S120 and T121 are part of the hydrophobic pocket responsible for the CID – CTD interaction and more specifically S120 is one of the residues in direct contact with the CTD (Noble et al. 2005). The critical position of these two residues suggests that their potential phosphorylation has a role in Pcf11 function since it can potentially affect its interaction with the polymerase. Interestingly, in yeast Pcf11 CID residues that correspond to these positions (120 and 121) are Leucine and Lysine respectively. This suggests that S120 and/or T121 phosphorylation and its potential involvement in the function of Pcf11 is a characteristic of higher eukaryotes.

6.2 Wnk1 kinase

Pcf11 mass-spectrometry analysis has also identified interacting proteins (Kinga Kamieniarz-Gdula, personal communication) including a single kinase, Wnk1.

Wnk1 belongs to a unique family of kinases (WNKs) with a characteristic catalytic domain where a Lysine (K) normally responsible for ATP binding is missing and located at a further downstream position (Xu et al. 2000; Min et al. 2004; Xu et al. 2005a). Wnk1 is a protein ubiquitously expressed in metazoans and has multiple isoforms generated by alternative splicing and alternative TSSs. Expression of the isoforms shows tissue specificity with a characteristic example being the non-catalytic isoform expressed in kidneys (O'Reilly et al. 2003). Wnk1 participates in the regulation of ion channels and in mice is an essential protein for the development of the cardiovascular system (O'Reilly et al. 2003; Zambrowicz et al. 2003; Xu et al. 2005b; Xie et al. 2013; Liu et al. 2015; Kahle et al. 2016). Furthermore, Wnk1 has been also shown to play a role in the cell cycle and correct mitotic spindle formation (Tu et al. 2011). Although, it has been previously shown that Wnk1 localises primarily in the cytoplasm, my results suggest that Wnk1 also localises in the nucleus but primarily is found in the nuclear envelope. There may be specific isoforms of the kinase that show this distinct perinuclear localisation. Possibly this had not been previously reported due to the specificity of antibodies employed in past studies which could be specific only for certain cytoplasmic isoforms.

The fact that Wnk1 was the only kinase found to interact with Pcf11 made it a potential candidate for the phosphorylation of Pcf11 CID. I generated and purified recombinant Pcf11 CID domain from bacteria. I used this recombinant Pcf11 CID domain with a commercially available Wnk1 kinase to set up an *in-vitro* kinase assay. In these assays it is clear that Wnk1 kinase can phosphorylate Pcf11 CID. Further experiments showed that this phosphorylation is specific to S120 and/ or T121. Unfortunately neither the Mass-spectrometry approach nor the *in-vitro* kinase assay allows us to determine whether both residues are phosphorylated. According to structural predictions, the simultaneous phosphorylation of S120 and T121 would be sterically restrictive. My experiments using a PhosTag gel showed that Pcf11 is also phosphorylated by Wnk1 *in-vivo* but unfortunately this approach does not allow the identification of the exact residues of Pcf11 that are phosphorylated. Taken together these experiments demonstrate that Wnk1 phosphorylates Pcf11 CID in the nucleus.

These results hint at a novel role of Wnk1 kinase in the nucleus and potentially in gene expression that was not previously anticipated. This adds yet another function to the repertoire of roles that Wnk1 plays in organisms. Finally I have managed to identify new post-translational modifications for the termination factor Pcf11. The critical position of the phosphorylated residues in the CID of the protein could potentially add another layer of complexity in the regulation of transcription termination through post-translational modifications of termination factors.

6.3 Role of Pcf11 CID phosphorylation

The first experiment I did to investigate the role of Pcf11 CID phosphorylation was to test its localisation using a microscopy approach. I generated mutant versions of Pcf11 to mimic the phosphorylated and unphosphorylated state of the protein (S120A T121A, S120E, T121E). WT Pcf11 localises exclusively in the nucleus where it shows specific foci of localisation that possibly correspond to actively transcribed genomic regions. Interestingly it also shows a perinuclear localisation similar to Wnk1 kinase. This result is not surprising since the association of actively transcribed genes with the nuclear periphery has been well documented (Brickner and Walter 2004; Taddei et al. 2006). The association of actively transcribed genes with the NPC is thought to enhance mRNA export but also contribute to transcriptional memory by promoting specific chromatin modifications.

Phosphorylation of Pcf11 does not seem to affect its import to the nucleus since the unphosphorylated mimic localises similarly to the WT. Surprisingly the phosphomimetic Pcf11 (S120E and T121E) shows an exclusive perinuclear localisation and does not seem to be associated with chromatin. This result suggests that phosphorylation of the CID domain abolishes the recruitment of Pcf11 with actively transcribed genes possibly by blocking its interaction with RNA Pol II CID. The fact that these residues (S120 and T121) are located in critical positions for this interaction suggests that their phosphorylation could potentially be employed to regulate Pcf11 recruitment to actively transcribed genes.

Localisation of phosphorylated Pcf11 and Wnk1 at the nuclear periphery provided a hint for the potential role of CID phosphorylation. Using cell fractionation and qPCR, I was able to detect an mRNA export defect in cells depleted for Wnk1. I further verified these results by an mRNA FISH approach that also showed that Wnk1 depletion causes widespread mRNA export defects. This effect of Wnk1 on mRNA export appears to be through the phosphorylation of Pcf11 CID since expression of the unphosphorylated Pcf11 also shows the same phenotype.

A surprising observation from these experiments is the widespread nature and severity of this phenotype. Wnk1 knock-out mice show lethal phenotypes early in development but this phenomenon is attributed to the participation of Wnk1 to the cardiovascular development process (Xie et al. 2013). I was able to generate homozygous Wnk1 knock-out HeLa lines using a CRISPR/Cas9 approach which only show a mild, slow-growth phenotype. It would be expected that a widespread mRNA export defect caused by Wnk1 would make it an essential protein for the cells. Possibly redundancy is built into these important and fundamental cellular processes. It is probable that upon Wnk1 knock-out the cells adjust and employ alternative mechanisms and pathways for the regulation of mRNA export. A second explanation for this phenomenon would be that the short deletion on the first exon of Wnk1 that I generated (Figures 3.11-3.14) may abolish the expression of a dominant isoform of the protein but it could potentially lead to the expression of a different isoform of the protein that could compensate for the loss of the canonical Wnk1.

6.4 mRNA export and Pcf11 CID phosphorylation

In yeast Pcf11 is responsible for the recruitment of export factor Yra1. Furthermore it is thought that Pcf11 provides the necessary cross-talk for the coordination of mRNA 3' end processing and export since its co-factor Clp1 and the export factor Yra1 compete for the interaction with Pcf11 (Johnson et al. 2009; Johnson et al. 2011). It is tempting to speculate that phosphorylation of Pcf11 CID could affect mRNA export through the modulation of export factor recruitment. However, when I tested the recruitment of the export factor Aly, CID phosphorylation does not seem to have a detectable effect on its recruitment. Furthermore phosphorylation of Pcf11 CID does not seem to affect its interaction with the NPC since both the phosphomimetic and unphosphorylated mutants show a similar localisation to the nuclear periphery. I suspect that Pcf11 CID phosphorylation provides an extra layer of regulation of mRNA export not through the recruitment of export factors.

I have shown that phosphorylation of Pcf11 CID abolishes its interaction with RNA Pol II. Furthermore, the phosphomimetic mutant of Pcf11 localises exclusively on the nuclear envelope. These observations led me to test the hypothesis that CID phosphorylation could be used for dismantling the transcription termination complex and the mRNP. I explored this possibility with an indirect approach and the detection of 'heavy' chromatin. This method has been shown in previous studies (Rougemaille et al. 2008) to correspond to mRNPs that have failed to release from their respective genes. Upon depletion of Wnk1 kinase, I can clearly detect formation of 'heavy' chromatin in the genes that I tested. This result suggests that Pcf11 phosphorylation is

required for the release of the mRNP and its subsequent export through the NPC. I propose a model where phosphorylation of the CID disrupts the interaction of Pcf11 with RNA Pol II. This disruption is required for the disassembly of the termination complex and mRNP from the transcription site (Figure 6.1). The mRNP can be then exported through the NPC to the cytoplasm. Previous studies in *Drosophila melanogaster* it have shown that Pcf11 and its interaction with the CTD is required for the efficient disassembly of the transcription elongation complex. However, the mechanisms through which this might be happening are still unclear. Two studies suggest that Pcf11 and its CID create a bridge between the nascent transcript and the CTD and this interaction is required for the efficient disassembly of the elongation complex (Zhang et al. 2005; Zhang and Gilmour 2006). In mammalian cells, possibly phosphorylation of Pcf11 CID by Wnk1 is required for the disruption of this bridge and the final disassembly and release of the mRNP from the gene template.

6.5 Summary

My work has identified a new role for the kinase Wnk1 in the nucleus. I have shown that Wnk1 can phosphorylate Pcf11 CID on S120 and/or T121. Furthermore, I have linked this phosphorylation to the process of mRNA export. Wnk1 phosphorylates Pcf11 CID and this phosphorylation is required for the proper release and nuclear export of mRNAs. My data suggests that Wnk1 does not affect recruitment of export factors to transcribed genes but can abolish the interaction of Pcf11 with RNA Pol II. Possibly, the loss of this

interaction is required for the disassembly of the transcription termination complex and the release of the mRNA for nuclear export.

My data provide new insight for the regulation and coordination of the processes of transcription termination and mRNA export. Furthermore, I provide new evidence for the regulation of gene expression processes through the post-translational modifications of a termination factor. This adds an extra layer of regulation to the coordination of gene expression mechanisms that lead to the production and nuclear export of an mRNA.

6.6 Future plans

My work on this project provides the framework for a deeper investigation on the cross talk between transcription termination and mRNA export. The next step on this project will be to investigate the generality of this regulation. I plan to perform genome wide mRNA sequencing in nuclear and cytoplasmic fractions upon depletion of Wnk1 kinase to see whether there is any gene specificity to this regulatory process. I will also measure how widespread the effect of this mechanism is. Several studies have shown that mRNA export mechanisms show gene specificity (reviewed in (Wickramasinghe and Laskey 2015)). I would like to test whether Wnk1 can provide this kind of regulation in the export of specific mRNAs to the cytoplasm.

Furthermore, it is crucial to further verify and study exactly how CID phosphorylation affects the disassembly of the termination complexes and release of the mRNA. In order to do this, I would like to set up an *in-vitro* system to determine the precise effect of Wnk1 kinase in releasing transcription complexes from the DNA template. Furthermore, I would like to further analyze the formation of 'heavy' chromatin. I am planning to use mass spectrometry to identify the differences between phosphorylated and unphosphorylated Pcf11 in their association with certain proteins. I would also like to test how this phosphorylation affects the interaction of Pcf11 with chromatin DNA and mRNA to further verify the disassembly hypothesis.

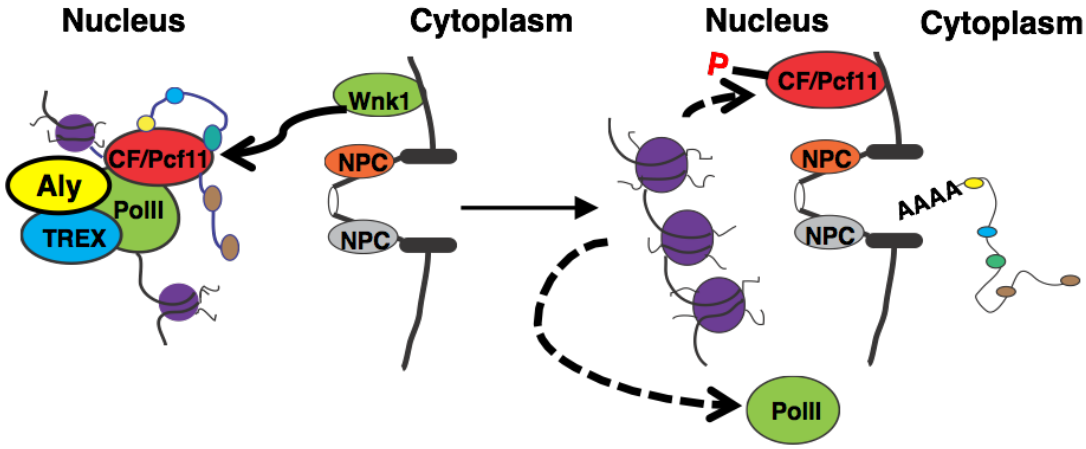


Figure 6.1
Proposed model for the role of Pcf11 CID phosphorylation in mRNA export. Wnk1 phosphorylates Pcf11 CID. This phosphorylation disrupts the interaction of Pcf11 with RNA Pol II and leads to the disassembly of the transcription complex. The mRNP is then released from the transcribed locus for export to the cytoplasm through the NPC.

References

Abruzzi KC, Belostotsky DA, Chekanova JA, Dower K, Rosbash M. 2006. 3'-end formation signals modulate the association of genes with the nuclear periphery as well as mRNP dot formation. *EMBO J* **25**: 4253-4262.

Adelman K, Lis JT. 2012. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet* **13**: 720-731.

Alexander RD, Innocente SA, Barrass JD, Beggs JD. 2010. Splicing-dependent RNA polymerase pausing in yeast. *Mol Cell* **40**: 582-593.

Amrani N, Minet M, Wyers F, Dufour ME, Aggerbeck LP, Lacroute F. 1997. PCF11 encodes a third protein component of yeast cleavage and polyadenylation factor I. *Mol Cell Biol* **17**: 1102-1109.

Ansari A, Hampsey M. 2005. A role for the CPF 3'-end processing machinery in RNAP II-dependent gene looping. *Genes Dev* **19**: 2969-2978.

Banerji J, Rusconi S, Schaffner W. 1981. Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**: 299-308.

Barilla D, Lee BA, Proudfoot NJ. 2001. Cleavage/polyadenylation factor IA associates with the carboxyl-terminal domain of RNA polymerase II in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **98**: 445-450.

Bartkowiak B, Greenleaf AL. 2011. Phosphorylation of RNAPII: To P-TEFb or not to P-TEFb? *Transcription* **2**: 115-119.

Bjork P, Wieslander L. 2014. Mechanisms of mRNA export. *Semin Cell Dev Biol* **32**: 47-54.

Bono F, Gehring NH. 2011. Assembly, disassembly and recycling: the dynamics of exon junction complexes. *RNA Biol* **8**: 24-29.

Bres V, Yoh SM, Jones KA. 2008. The multi-tasking P-TEFb complex. *Curr Opin Cell Biol* **20**: 334-340.

Brickner JH, Walter P. 2004. Gene recruitment of the activated INO1 locus to the nuclear membrane. *PLoS Biol* **2**: e342.

Casolari JM, Brown CR, Drubin DA, Rando OJ, Silver PA. 2005. Developmentally induced changes in transcriptional program alter spatial organization across chromosomes. *Genes Dev* **19**: 1188-1198.

Chapman RD, Heidemann M, Hintermair C, Eick D. 2008. Molecular evolution of the RNA polymerase II CTD. *Trends Genet* **24**: 289-296.

Chi B, Wang Q, Wu G, Tan M, Wang L, Shi M, Chang X, Cheng H. 2013. Aly and THO are required for assembly of the human TREX complex and association of TREX components with the spliced mRNA. *Nucleic Acids Res* **41**: 1294-1306.

Chinchilla K, Rodriguez-Molina JB, Ursic D, Finkel JS, Ansari AZ, Culbertson MR. 2012. Interactions of Sen1, Nrd1, and Nab3 with multiple phosphorylated forms of the Rpb1 C-terminal domain in Saccharomyces cerevisiae. *Eukaryot Cell* **11**: 417-429.

Conaway RC, Conaway JW. 1993. General initiation factors for RNA polymerase II. *Annu Rev Biochem* **62**: 161-190.

Consortium EP. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**: 57-74.

Danko CG, Hah N, Luo X, Martins AL, Core L, Lis JT, Siepel A, Kraus WL. 2013. Signaling pathways differentially affect RNA polymerase II initiation, pausing, and elongation rate in cells. *Mol Cell* **50**: 212-222.

David CJ, Boyne AR, Millhouse SR, Manley JL. 2011. The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65-Prp19 complex. *Genes Dev* **25**: 972-983.

De Santa F, Barozzi I, Mietton F, Ghisletti S, Polletti S, Tusi BK, Muller H, Ragoussis J, Wei CL, Natoli G. 2010. A large fraction of extragenic RNA pol II transcription sites overlap enhancers. *PLoS Biol* **8**: e1000384.

de Vries H, Ruegsegger U, Hubner W, Friedlein A, Langen H, Keller W. 2000. Human pre-mRNA cleavage factor II(m) contains homologs of yeast proteins and bridges two other cleavage factors. *EMBO J* **19**: 5895-5904.

Dieppois G, Iglesias N, Stutz F. 2006. Cotranscriptional recruitment to the mRNA export receptor Mex67p contributes to nuclear pore anchoring of activated genes. *Mol Cell Biol* **26**: 7858-7870.

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**: 376-380.

Egloff S, Dienstbier M, Murphy S. 2012. Updating the RNA polymerase CTD code: adding gene-specific layers. *Trends Genet* **28**: 333-341.

Egloff S, Murphy S. 2008. Cracking the RNA polymerase II CTD code. *Trends Genet* **24**: 280-288.

Egloff S, O'Reilly D, Chapman RD, Taylor A, Tanzhaus K, Pitts L, Eick D, Murphy S. 2007. Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. *Science* **318**: 1777-1779.

Elkon R, Ugalde AP, Agami R. 2013. Alternative cleavage and polyadenylation: extent, regulation and function. *Nat Rev Genet* **14**: 496-506.

Finkel JS, Chinchilla K, Ursic D, Culbertson MR. 2010. Sen1p performs two genetically separable functions in transcription and processing of U5 small nuclear RNA in Saccharomyces cerevisiae. *Genetics* **184**: 107-118.

Gaillard H, Garcia-Muse T, Aguilera A. 2015. Replication stress and cancer. *Nat Rev Cancer* **15**: 276-289.

Ghazy MA, Gordon JM, Lee SD, Singh BN, Bohm A, Hampsey M, Moore C. 2012. The interaction of Pcf11 and Clp1 is needed for mRNA 3'-end formation and is modulated by amino acids in the ATP-binding site. *Nucleic Acids Res* **40**: 1214-1225.

Ghosh A, Shuman S, Lima CD. 2011. Structural insights to how mammalian capping enzyme reads the CTD code. *Mol Cell* **43**: 299-310.

Greger IH, Demarchi F, Giacca M, Proudfoot NJ. 1998. Transcriptional interference perturbs the binding of Sp1 to the HIV-1 promoter. *Nucleic Acids Res* **26**: 1294-1301.

Greger IH, Proudfoot NJ. 1998. Poly(A) signals control both transcriptional termination and initiation between the tandem GAL10 and GAL7 genes of Saccharomyces cerevisiae. *EMBO J* **17**: 4771-4779.

Grunberg S, Warfield L, Hahn S. 2012. Architecture of the RNA polymerase II preinitiation complex and mechanism of ATP-dependent promoter opening. *Nat Struct Mol Biol* **19**: 788-796.

Grzechnik P, Gdula MR, Proudfoot NJ. 2015. Pcf11 orchestrates transcription termination pathways in yeast. *Genes Dev* **29**: 849-861.

He Y, Fang J, Taatjes DJ, Nogales E. 2013. Structural visualization of key steps in human transcription initiation. *Nature* **495**: 481-486.

Hilleren P, McCarthy T, Rosbash M, Parker R, Jensen TH. 2001. Quality control of mRNA 3'-end processing is linked to the nuclear exosome. *Nature* **413**: 538-542.

Hilleren P, Parker R. 2001. Defects in the mRNA export factors Rat7p, Gle1p, Mex67p, and Rat8p cause hyperadenylation during 3'-end formation of nascent transcripts. *RNA* **7**: 753-764.

Houseley J, LaCava J, Tollervey D. 2006. RNA-quality control by the exosome. *Nat Rev Mol Cell Biol* **7**: 529-539.

Jacquier A. 2009. The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs. *Nat Rev Genet* **10**: 833-844.

Jensen TH, Dower K, Libri D, Rosbash M. 2003. Early formation of mRNP: license for export or quality control? *Mol Cell* **11**: 1129-1138.

Jensen TH, Jacquier A, Libri D. 2013. Dealing with pervasive transcription. *Mol Cell* **52**: 473-484.

Jensen TH, Rosbash M. 2003. Co-transcriptional monitoring of mRNP formation. *Nat Struct Biol* **10**: 10-12.

Johnson SA, Cubberley G, Bentley DL. 2009. Cotranscriptional recruitment of the mRNA export factor Yra1 by direct interaction with the 3' end processing factor Pcf11. *Mol Cell* **33**: 215-226.

Johnson SA, Kim H, Erickson B, Bentley DL. 2011. The export factor Yra1 modulates mRNA 3' end processing. *Nat Struct Mol Biol* **18**: 1164-1171.

Jonkers I, Kwak H, Lis JT. 2014. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *Elife* **3**: e02407.

Kahle KT, Schmouth JF, Lavastre V, Latremoliere A, Zhang J, Andrews N, Omura T, Laganiere J, Rochefort D, Hince P et al. 2016. Inhibition of the kinase WNK1/HSN2 ameliorates neuropathic pain by restoring GABA inhibition. *Sci Signal* **9**: ra32.

Kaida D, Berg MG, Younis I, Kasim M, Singh LN, Wan L, Dreyfuss G. 2010. U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* **468**: 664-668.

Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S et al. 2010. Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465**: 182-187.

Koch F, Fenouil R, Gut M, Cauchy P, Albert TK, Zacarias-Cabeza J, Spicuglia S, de la Chapelle AL, Heidemann M, Hintermair C et al. 2011. Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. *Nat Struct Mol Biol* **18**: 956-963.

Kornberg RD. 1998. Mechanism and regulation of yeast RNA polymerase II transcription. *Cold Spring Harb Symp Quant Biol* **63**: 229-232.

Kuehner JN, Pearson EL, Moore C. 2011. Unravelling the means to an end: RNA polymerase II transcription termination. *Nat Rev Mol Cell Biol* **12**: 283-294.

Kwak H, Fuda NJ, Core LJ, Lis JT. 2013. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science* **339**: 950-953.

Lettice LA, Heaney SJ, Purdie LA, Li L, de Beer P, Oostra BA, Goode D, Elgar G, Hill RE, de Graaff E. 2003. A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum Mol Genet* **12**: 1725-1735.

Li J, Gilmour DS. 2013. Distinct mechanisms of transcriptional pausing orchestrated by GAGA factor and M1BP, a novel transcription factor. *EMBO J* **32**: 1829-1841.

Li W, You B, Hoque M, Zheng D, Luo W, Ji Z, Park JY, Gunderson SI, Kalsotra A, Manley JL et al. 2015. Systematic profiling of poly(A)+ transcripts modulated by core 3' end processing and splicing factors reveals regulatory rules of alternative cleavage and polyadenylation. *PLoS Genet* **11**: e1005166.

Licatalosi DD, Geiger G, Minet M, Schroeder S, Cilli K, McNeil JB, Bentley DL. 2002. Functional interaction of yeast pre-mRNA 3' end processing factors with RNA polymerase II. *Mol Cell* **9**: 1101-1111.

Lis JT, Mason P, Peng J, Price DH, Werner J. 2000. P-TEFb kinase recruitment and function at heat shock loci. *Genes Dev* **14**: 792-803.

Liu P, Kenney JM, Stiller JW, Greenleaf AL. 2010. Genetic organization, length conservation, and evolution of RNA polymerase II carboxyl-terminal domain. *Mol Biol Evol* **27**: 2628-2641.

Liu Y, Song X, Shi Y, Shi Z, Niu W, Feng X, Gu D, Bao HF, Ma HP, Eaton DC et al. 2015. WNK1 activates large-conductance Ca2+-activated K+ channels through modulation of ERK1/2 signaling. *J Am Soc Nephrol* **26**: 844-854.

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. *Proc Natl Acad Sci U S A* **84**: 8306-8310.

Lunde BM, Reichow SL, Kim M, Suh H, Leeper TC, Yang F, Mutschler H, Buratowski S, Meinhart A, Varani G. 2010. Cooperative interaction of

transcription termination factors with the RNA polymerase II C-terminal domain. *Nat Struct Mol Biol* **17**: 1195-1201.

Luo W, Johnson AW, Bentley DL. 2006. The role of Rat1 in coupling mRNA 3'-end processing to transcription termination: implications for a unified allosteric-torpedo model. *Genes Dev* **20**: 954-965.

Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. 2013. RNA-guided human genome engineering via Cas9. *Science* **339**: 823-826.

Mapendano CK, Lykke-Andersen S, Kjems J, Bertrand E, Jensen TH. 2010. Crosstalk between mRNA 3' end processing and transcription initiation. *Mol Cell* **40**: 410-422.

Meinhart A, Cramer P. 2004. Recognition of RNA polymerase II carboxy-terminal domain by 3'-RNA-processing factors. *Nature* **430**: 223-226.

Melan MA, Sluder G. 1992. Redistribution and differential extraction of soluble proteins in permeabilized cultured cells. Implications for immunofluorescence microscopy. *J Cell Sci* **101 (Pt 4)**: 731-743.

Min X, Lee BH, Cobb MH, Goldsmith EJ. 2004. Crystal structure of the kinase domain of WNK1, a kinase that causes a hereditary form of hypertension. *Structure* **12**: 1303-1311.

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**: 1702-1705.

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. *Proc Natl Acad Sci U S A* **94**: 7897-7902.

Noble CG, Hollingworth D, Martin SR, Ennis-Adeniran V, Smerdon SJ, Kelly G, Taylor IA, Ramos A. 2005. Key features of the interaction between Pcf11 CID and RNA polymerase II CTD. *Nat Struct Mol Biol* **12**: 144-151.

Nojima T, Gomes T, Carmo-Fonseca M, Proudfoot NJ. 2016. Mammalian NET-seq analysis defines nascent RNA profiles and associated RNA processing genome-wide. *Nat Protoc* **11**: 413-428.

O'Reilly M, Marshall E, Speirs HJ, Brown RW. 2003. WNK1, a gene within a novel blood pressure control pathway, tissue-specifically generates radically different isoforms with and without a kinase domain. *J Am Soc Nephrol* **14**: 2447-2456.

O'Sullivan JM, Tan-Wong SM, Morillon A, Lee B, Coles J, Mellor J, Proudfoot NJ. 2004. Gene loops juxtapose promoters and terminators in yeast. *Nat Genet* **36**: 1014-1018.

Peterlin BM, Price DH. 2006. Controlling the elongation phase of transcription with P-TEFb. *Mol Cell* **23**: 297-305.

Plant KE, Dye MJ, Lafaille C, Proudfoot NJ. 2005. Strong polyadenylation and weak pausing combine to cause efficient termination of transcription in the human Ggamma-globin gene. *Mol Cell Biol* **25**: 3276-3285.

Porrua O, Libri D. 2013. A bacterial-like mechanism for transcription termination by the Sen1p helicase in budding yeast. *Nat Struct Mol Biol* **20**: 884-891.

-. 2015. Transcription termination and the control of the transcriptome: why, where and how to stop. *Nat Rev Mol Cell Biol* **16**: 190-202.

Proudfoot NJ. 2011. Ending the message: poly(A) signals then and now. *Genes Dev* **25**: 1770-1782.

-. 2016. Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. *Science* **352**: aad9926.

Reinberg D, Orphanides G, Ebright R, Akoulitchev S, Carcamo J, Cho H, Cortes P, Drapkin R, Flores O, Ha I et al. 1998. The RNA polymerase II general transcription factors: past, present, and future. *Cold Spring Harb Symp Quant Biol* **63**: 83-103.

Roeder RG. 1996. The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem Sci* **21**: 327-335.

Rougemaille M, Dieppois G, Kisseleva-Romanova E, Gudipati RK, Lemoine S, Blugeon C, Boulay J, Jensen TH, Stutz F, Devaux F et al. 2008. THO/Sub2p functions to coordinate 3'-end processing with gene-nuclear pore association. *Cell* **135**: 308-321.

Sadowski M, Dichtl B, Hubner W, Keller W. 2003. Independent functions of yeast Pcf11p in pre-mRNA 3' end processing and in transcription termination. *EMBO J* **22**: 2167-2177.

Sanyal A, Lajoie BR, Jain G, Dekker J. 2012. The long-range interaction landscape of gene promoters. *Nature* **489**: 109-113.

Schmidt D, Schwalie PC, Ross-Innes CS, Hurtado A, Brown GD, Carroll JS, Flicek P, Odom DT. 2010. A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Res* **20**: 578-588.

Serysheva E, Berhane H, Grumolato L, Demir K, Balmer S, Bodak M, Boutros M, Aaronson S, Mlodzik M, Jenny A. 2013. Wnk kinases are positive

regulators of canonical Wnt/beta-catenin signalling. *EMBO Rep* **14**: 718-725.

Shearwin KE, Callen BP, Egan JB. 2005. Transcriptional interference--a crash course. *Trends Genet* **21**: 339-345.

Singh BN, Hampsey M. 2007. A transcription-independent role for TFIIB in gene looping. *Mol Cell* **27**: 806-816.

Taddei A, Van Houwe G, Hediger F, Kalck V, Cubizolles F, Schober H, Gasser SM. 2006. Nuclear pore association confers optimal expression levels for an inducible yeast gene. *Nature* **441**: 774-778.

Tian B, Manley JL. 2013. Alternative cleavage and polyadenylation: the long and short of it. *Trends Biochem Sci* **38**: 312-320.

Tu SW, Bugde A, Luby-Phelps K, Cobb MH. 2011. WNK1 is required for mitosis and abscission. *Proc Natl Acad Sci U S A* **108**: 1385-1390.

Verissimo F, Jordan P. 2001. WNK kinases, a novel protein kinase subfamily in multi-cellular organisms. *Oncogene* **20**: 5562-5569.

Vilborg A, Passarelli MC, Yario TA, Tycowski KT, Steitz JA. 2015. Widespread Inducible Transcription Downstream of Human Genes. *Mol Cell* **59**: 449-461.

Weber CM, Ramachandran S, Henikoff S. 2014. Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. *Mol Cell* **53**: 819-830.

West JA, Cook A, Alver BH, Stadtfeld M, Deaton AM, Hochedlinger K, Park PJ, Tolstorukov MY, Kingston RE. 2014. Nucleosomal occupancy changes locally over key regulatory regions during cell differentiation and reprogramming. *Nat Commun* **5**: 4719.

West S, Proudfoot NJ. 2008. Human Pcf11 enhances degradation of RNA polymerase II-associated nascent RNA and transcriptional termination. *Nucleic Acids Res* **36**: 905-914.

Whitelaw E, Proudfoot N. 1986. Alpha-thalassaemia caused by a poly(A) site mutation reveals that transcriptional termination is linked to 3' end processing in the human alpha 2 globin gene. *EMBO J* **5**: 2915-2922.

Wickramasinghe VO, Laskey RA. 2015. Control of mammalian gene expression by selective mRNA export. *Nat Rev Mol Cell Biol* **16**: 431-442.

Wilson FH, Disse-Nicodeme S, Choate KA, Ishikawa K, Nelson-Williams C, Desitter I, Gunel M, Milford DV, Lipkin GW, Achard JM et al. 2001. Human hypertension caused by mutations in WNK kinases. *Science* **293**: 1107-1112.

Xie J, Yoon J, Yang SS, Lin SH, Huang CL. 2013. WNK1 protein kinase regulates embryonic cardiovascular development through the OSR1 signaling cascade. *J Biol Chem* **288**: 8566-8574.

Xu B, English JM, Wilsbacher JL, Stippec S, Goldsmith EJ, Cobb MH. 2000. WNK1, a novel mammalian serine/threonine protein kinase lacking the catalytic lysine in subdomain II. *J Biol Chem* **275**: 16795-16801.

Xu BE, Lee BH, Min X, Lenertz L, Heise CJ, Stippec S, Goldsmith EJ, Cobb MH. 2005a. WNK1: analysis of protein kinase structure, downstream targets, and potential roles in hypertension. *Cell Res* **15**: 6-10.

Xu BE, Stippec S, Chu PY, Lazrak A, Li XJ, Lee BH, English JM, Ortega B, Huang CL, Cobb MH. 2005b. WNK1 activates SGK1 to regulate the epithelial sodium channel. *Proc Natl Acad Sci U S A* **102**: 10315-10320.

Zambrowicz BP, Abuin A, Ramirez-Solis R, Richter LJ, Piggott J, BeltrandelRio H, Buxton EC, Edwards J, Finch RA, Friddle CJ et al. 2003. Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention. *Proc Natl Acad Sci U S A* **100**: 14109-14114.

Zaret KS, Watts J, Xu J, Wandzioch E, Smale ST, Sekiya T. 2008. Pioneer factors, genetic competence, and inductive signaling: programming liver and pancreas progenitors from the endoderm. *Cold Spring Harb Symp Quant Biol* **73**: 119-126.

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 Dev* **19**: 1572-1580.

Zhang Z, Gilmour DS. 2006. Pcf11 is a termination factor in Drosophila that dismantles the elongation complex by bridging the CTD of RNA polymerase II to the nascent transcript. *Mol Cell* **21**: 65-74.

Permissions

NATURE PUBLISHING GROUP LICENSE TERMS AND CONDITIONS

Sep 13, 2016

This Agreement between Adam Volanakis ("You") and Nature Publishing Group ("Nature Publishing Group") consists of your license details and the terms and conditions provided by Nature Publishing Group and Copyright Clearance Center.

License Number	3947120768761
License date	Sep 13, 2016
Licensed Content Publisher	Nature Publishing Group
Licensed Content Publication	Nature Reviews Molecular Cell Biology
Licensed Content Title	Structural basis of transcription initiation by RNA polymerase II
Licensed Content Author	Sarah Sainsbury, Carrie Bernecky, Patrick Cramer
Licensed Content Date	Feb 18, 2015
Licensed Content Volume Number	16
Licensed Content Issue Number	3
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Figures	Figure 1
Author of this NPG article	no
Your reference number	
Title of your thesis / dissertation	Role of Pcf11 Post translational modifications in gene expression
Expected completion date	Oct 2016
Estimated size (number of pages)	120
Requestor Location	Adam Volanakis Sir William Dunn School of Pathology South Parks Road Oxford, OX1 3RE United Kingdom Attn: Adam Volanakis
Billing Type	Invoice
Billing Address	Adam Volanakis

Sir William Dunn School of Pathology
South Parks Road

Oxford, United Kingdom OX1 3RE
Attn: Adam Volanakis

Total 0.00 GBP

Terms and Conditions

Terms and Conditions for Permissions

Nature Publishing Group hereby grants you a non-exclusive license to reproduce this material for this purpose, and for no other use,subject to the conditions below:

1. NPG warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to Nature Publishing Group and does not carry the copyright of another entity (as credited in the published version). If the credit line on any part of the material you have requested indicates that it was reprinted or adapted by NPG with permission from another source, then you should also seek permission from that source to reuse the material.

2. Permission granted free of charge for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to the work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.Where print permission has been granted for a fee, separate permission must be obtained for any additional, electronic re-use (unless, as in the case of a full paper, this has already been accounted for during your initial request in the calculation of a print run).NB: In all cases, web-based use of full-text articles must be authorized separately through the 'Use on a Web Site' option when requesting permission.

3. Permission granted for a first edition does not apply to second and subsequent editions and for editions in other languages (except for signatories to the STM Permissions Guidelines, or where the first edition permission was granted for free).

4. Nature Publishing Group's permission must be acknowledged next to the figure, table or abstract in print. In electronic form, this acknowledgement must be visible at the same time as the figure/table/abstract, and must be hyperlinked to the journal's homepage.

5. The credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)
   For AOP papers, the credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

   Note: For republication from the *British Journal of Cancer,* the following credit lines apply.
   Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)For AOP papers, the credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

6. Adaptations of single figures do not require NPG approval. However, the adaptation should be credited as follows:

   Adapted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)

Note: For adaptation from the *British Journal of Cancer,* the following credit line applies.
Adapted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

7. Translations of 401 words up to a whole article require NPG approval. Please visit http://www.macmillanmedicalcommunications.com for more information.Translations of up to a 400 words do not require NPG approval. The translation should be credited as follows:

   Translated by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication).

   Note: For translation from the *British Journal of Cancer,* the following credit line applies.
   Translated by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

We are certain that all parties will benefit from this agreement and wish you the best in the use of this material. Thank you.
Special Terms:
v1.1

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

NATURE PUBLISHING GROUP LICENSE TERMS AND CONDITIONS

Sep 13, 2016

This Agreement between Adam Volanakis ("You") and Nature Publishing Group ("Nature Publishing Group") consists of your license details and the terms and conditions provided by Nature Publishing Group and Copyright Clearance Center.

License Number	3947120894914
License date	Sep 13, 2016
Licensed Content Publisher	Nature Publishing Group
Licensed Content Publication	Nature Reviews Molecular Cell Biology
Licensed Content Title	Transcription termination and the control of the transcriptome: why, where and how to stop
Licensed Content Author	Odil Porrua, Domenico Libri
Licensed Content Date	Feb 4, 2015
Licensed Content Volume Number	16
Licensed Content Issue Number	3
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Figures	Figure 2
Author of this NPG article	no
Your reference number	
Title of your thesis / dissertation	Role of Pcf11 Post translational modifications in gene expression
Expected completion date	Oct 2016
Estimated size (number of pages)	120
Requestor Location	Adam Volanakis Sir William Dunn School of Pathology South Parks Road Oxford, OX1 3RE United Kingdom Attn: Adam Volanakis
Billing Type	Invoice

Billing Address

Adam Volanakis
Sir William Dunn School of Pathology
South Parks Road

Oxford, United Kingdom OX1 3RE
Attn: Adam Volanakis

Total 0.00 USD

Terms and Conditions

Terms and Conditions for Permissions

Nature Publishing Group hereby grants you a non-exclusive license to reproduce this material for this purpose, and for no other use,subject to the conditions below:

1. NPG warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to Nature Publishing Group and does not carry the copyright of another entity (as credited in the published version). If the credit line on any part of the material you have requested indicates that it was reprinted or adapted by NPG with permission from another source, then you should also seek permission from that source to reuse the material.

2. Permission granted free of charge for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to the work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.Where print permission has been granted for a fee, separate permission must be obtained for any additional, electronic re-use (unless, as in the case of a full paper, this has already been accounted for during your initial request in the calculation of a print run).NB: In all cases, web-based use of full-text articles must be authorized separately through the 'Use on a Web Site' option when requesting permission.

3. Permission granted for a first edition does not apply to second and subsequent editions and for editions in other languages (except for signatories to the STM Permissions Guidelines, or where the first edition permission was granted for free).

4. Nature Publishing Group's permission must be acknowledged next to the figure, table or abstract in print. In electronic form, this acknowledgement must be visible at the same time as the figure/table/abstract, and must be hyperlinked to the journal's homepage.

5. The credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)
   For AOP papers, the credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

   Note: For republication from the *British Journal of Cancer,* the following credit lines apply.
   Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)For AOP papers, the credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

6. Adaptations of single figures do not require NPG approval. However, the adaptation should be credited as follows:

   Adapted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference

citation), copyright (year of publication)

Note: For adaptation from the *British Journal of Cancer,* the following credit line applies.
Adapted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

7. Translations of 401 words up to a whole article require NPG approval. Please visit http://www.macmillanmedicalcommunications.com for more information.Translations of up to a 400 words do not require NPG approval. The translation should be credited as follows:

   Translated by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication).

   Note: For translation from the *British Journal of Cancer,* the following credit line applies.
   Translated by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

We are certain that all parties will benefit from this agreement and wish you the best in the use of this material. Thank you.
Special Terms:
v1.1

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

NATURE PUBLISHING GROUP LICENSE TERMS AND CONDITIONS

Sep 13, 2016

This Agreement between Adam Volanakis ("You") and Nature Publishing Group ("Nature Publishing Group") consists of your license details and the terms and conditions provided by Nature Publishing Group and Copyright Clearance Center.

License Number	3947120984273
License date	Sep 13, 2016
Licensed Content Publisher	Nature Publishing Group
Licensed Content Publication	Nature Reviews Molecular Cell Biology
Licensed Content Title	Control of mammalian gene expression by selective mRNA export
Licensed Content Author	Vihandha O. Wickramasinghe, Ronald A. Laskey
Licensed Content Date	Jun 17, 2015
Licensed Content Volume Number	16
Licensed Content Issue Number	7
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Figures	Figure 1
Author of this NPG article	no
Your reference number	
Title of your thesis / dissertation	Role of Pcf11 Post translational modifications in gene expression
Expected completion date	Oct 2016
Estimated size (number of pages)	120
Requestor Location	Adam Volanakis Sir William Dunn School of Pathology South Parks Road Oxford, OX1 3RE United Kingdom Attn: Adam Volanakis
Billing Type	Invoice
Billing Address	Adam Volanakis

Sir William Dunn School of Pathology
South Parks Road

Oxford, United Kingdom OX1 3RE
Attn: Adam Volanakis

Total 0.00 USD

Terms and Conditions

Terms and Conditions for Permissions

Nature Publishing Group hereby grants you a non-exclusive license to reproduce this material for this purpose, and for no other use,subject to the conditions below:

1. NPG warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to Nature Publishing Group and does not carry the copyright of another entity (as credited in the published version). If the credit line on any part of the material you have requested indicates that it was reprinted or adapted by NPG with permission from another source, then you should also seek permission from that source to reuse the material.

2. Permission granted free of charge for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to the work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.Where print permission has been granted for a fee, separate permission must be obtained for any additional, electronic re-use (unless, as in the case of a full paper, this has already been accounted for during your initial request in the calculation of a print run).NB: In all cases, web-based use of full-text articles must be authorized separately through the 'Use on a Web Site' option when requesting permission.

3. Permission granted for a first edition does not apply to second and subsequent editions and for editions in other languages (except for signatories to the STM Permissions Guidelines, or where the first edition permission was granted for free).

4. Nature Publishing Group's permission must be acknowledged next to the figure, table or abstract in print. In electronic form, this acknowledgement must be visible at the same time as the figure/table/abstract, and must be hyperlinked to the journal's homepage.

5. The credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)
   For AOP papers, the credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

   Note: For republication from the *British Journal of Cancer,* the following credit lines apply.
   Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)For AOP papers, the credit line should read:
   Reprinted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME], advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].XXXXX)

6. Adaptations of single figures do not require NPG approval. However, the adaptation should be credited as follows:

   Adapted by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication)

Note: For adaptation from the *British Journal of Cancer,* the following credit line applies.
Adapted by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

7. Translations of 401 words up to a whole article require NPG approval. Please visit http://www.macmillanmedicalcommunications.com for more information.Translations of up to a 400 words do not require NPG approval. The translation should be credited as follows:

   Translated by permission from Macmillan Publishers Ltd: [JOURNAL NAME] (reference citation), copyright (year of publication).

   Note: For translation from the *British Journal of Cancer,* the following credit line applies.
   Translated by permission from Macmillan Publishers Ltd on behalf of Cancer Research UK: [JOURNAL NAME] (reference citation), copyright (year of publication)

We are certain that all parties will benefit from this agreement and wish you the best in the use of this material. Thank you.
Special Terms:
v1.1

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.