Effect of CTCF and Cohesin on the dynamics of RNA polymerase II transcription and coupled pre-messenger RNA processing

Olga Liska

University of Oxford
Sir William Dunn School of Pathology
and
Lincoln College

Thesis submitted for the degree of Doctor of Philosophy at the University of Oxford

Hilary Term 2013

ABSTRACT

The CCCTC-binding factor (CTCF) is a versatile, multifunctional zinc-finger protein involved in a broad spectrum of cellular functions. In mammalian cells, CTCF functions together with the Cohesin complex, an essential regulator of sister chromatid cohesion. Together, CTCF and Cohesin have been shown to regulate gene expression at a genome-wide level in mammalian cells. In the yeast *Saccharomyces pombe*, Cohesin has been implicated in transcription termination of convergently transcribed genes, in a cell cycle dependent manner. The aim of this thesis was to investigate the possibility of direct transcriptional involvement of CTCF and Cohesin in human cells. The first model system applied for this experimental purpose was the β-globin gene with introduced canonical CTCF-binding sites replacing the endogenous Co-Transcriptional Cleavage (CoTC) element downstream of β-globin. The results obtained indicate that recruitment of CTCF to the β-globin 3` flanking region does not prevent read-through transcription. However, CTCF-binding does mediate RNA Polymerase II (Pol II) pausing at the site of recruited CTCF. This results in more efficient pre-mRNA 3` end processing and therefore rescues β-globin mRNA to wild type levels. Cohesin was not detected at the introduced CTCF-binding sites. These results are a contribution to our understanding of the spatio-temporal requirements for co-transcriptional events like 3` end pre-mRNA processing and Pol II kinetics.

The second part of my thesis presents an investigation on the involvement of CTCF and Cohesin in lipopolysaccharide (LPS)-induced *Tumor Necrosis Factor α* (*TNFα)* gene expression regulation in human monocytes and

differentiated M1- and M2-type macrophages. These studies provide first evidence of Cohesin recruitment to the *TNFα* gene body and its regulatory NFκB-binding sites. Differences in the recruitment profiles obtained indicate potential regulatory differences of *TNFα* among the three cell types. Preliminary data provide an insight into the effects on TNFα mRNA levels upon down-regulation of Cohesin subunits.

Acknowledgements

The finish line is reached - this thesis is the end of my journey in obtaining my DPhil at the University of Oxford, a place of highest academic standard, in the heart of Oxford, one of the world`s most unique places to be!
Throughout the entire time my work has been guided by the support and encouragement of numerous people including my supervisor, colleagues, friends and family.

First of all, I would like to express my gratitude to my supervisor Prof. Nicholas J. Proudfoot at the Sir William Dunn School of Pathology, who offered his continuous advice and encouragement throughout the course of this thesis.

I would like to thank my collaborators Prof. Irina Udalova and Thomas Krausgruber for their valuable input.

I am very proud to have worked alongside with supportive colleagues, in particular Priscilla, Eleanor, Steven, Shweta, Ashish, Joan and Alexandra, who have not only provided me with precious scientific advice and contributions but also opened their doors to me when I was in need of encouragement and understanding.

I take this opportunity to also thankfully acknowledge the Sir William Dunn School of Pathology for providing financial assistance by the EP Abraham Trust throughout my entire course of studies.

I feel lucky to be a member of Lincoln College College. I want to particularly thank Mrs. Carmella Elan-Gaston, who always had a cheerful smile for me and was of great help on numerous occasions.

Most importantly, I would have not overcome all the challenges and obstacles without the firm support from my family. I am deeply indebted to my parents Marie and Danny, my sister Nicole who always stood by my side through all the ups and downs. My close friends in Oxford, Priscilla & David, Hadeel, Justyna, Eleanor, Diu, Sarah and Laurel who have contributed a great deal to make this place particularly memorable.

This thesis is for my mentors, family & friends, with everlasting gratitude!

TABLE OF CONTENTS

GLOSSARY OF ABBREVIATIONS

A	adenine
ATP	adenosine triphosphate
ARE	AU rich element
bp	base pairs
C	cytosine
c-Myc	cellular homolog of the retroviral v-myconcogene
CBC	cap binding complex
CBE	CTCF-binding element
CD4	cluster of differentiation 4
Cdk	cyclin-dependent kinase
CdLS	Cornelia de Lange syndrome
cDNA	complementary deoxyribonucleic acid
CFIA	cleavage factor IA
CFIB	cleavage factor IB
CFIm	mammalian cleavage factor I
CFIIm	mammalian cleavage factor II
CHD8	Chromodomain helicase DNA binding protein 8
ChIA-PET	Chromatin Interaction Analysis by Paired-End Tag Sequencing
ChIP	chromatin immuno-precipitation
ChIP-Seq	ChIP in combination with high throughput sequencing

Ci	curie
CIITA	class II major histocompatibility complex transactivator
CID	CTD interacting domain
CLIP	crosslinking and immunoprecipitation
Clp1	cleavage and polyadenylation factor I subunit
CoTC	co-transcriptional cleavage
CPF	cleavage and polyadenylation factor
cps	counts per second
CPSF	cleavage and polyadenylation specificity factor
CstF	cleavage stimulatory factor
CTCF	CCCTC-binding factor
CTD	caroboxyl-terminal domain
CTP	cytosine triphosphate
CUTs	Cryptic unstable transcripts
DM1	dystrophia myotonica 1
DNA	deoxyribonucleic acid
DNA-PK	DNA- activated protein kinase
DNase	deoxyribonuclease
dNTP	deoxyribonucleotide triphosphate
DSE	downstream sequence element
DTT	dithiothreitol
DUE	downstream U-rich element
EC	elongation complex
EDTA	ethylenediaminetetraacetic acid (disodium salt)

EE	efficiency element
EMM	Edinburgh minimal medium
ER-α	estrogen receptor-alpha
ES	embryonic stem
Fip1	Factor interacting with PAP
G	guanine
g	g-force
GM-CSF	granulocyte-macrophage colony stimulating factor
GMP	guanine monophosphate
GTP	guanosine triphosphate
H19	hepatic fetal-specific mRNA
HC1	hydrochloric acid
HDE	histone downstream element
HEAT	Huntingtin, elongation factor 3 (protein phosphatase 2A or yeast kinase TOR1)
HEPES	N-2-hydroxyethylpiparizine-N'-2-ethanesulphonic acid
HIV	human immunodeficiency virus
HLA-DRB1	HLA-class II histocompatibility antigen DR beta1 chain
HLA-DQA1	HLA-class II histocompatibility complex DQ alpha1 chain
HLF	heat labile factor
hnRNP	heterogeneous nuclear ribonucleo protein
HS4	hyper sensitive site 4

HSF	heat shock factor
hscRACE	hybrid selection circular rapid amplification of complementary deoxyribonucleic acid ends
ICR	imprinted control region
IFN	interferon
Igf2	insulin-like growth factor 2
IgH	immunoglobulin heavy-chain
IgM	immunoglobulin M
IL3	interleukin-3
IP	immuno-precipitation
IVS	intervening sequence (intron)
kb	kilobase
KLF4	Krueppel-like factor 4
LCR	locus control region
LPS	lipopolysaccharide
LTR	long terminal repeat
MAZ	myc-associated zinc finger protein
M-CSF	macrophage colony stimulating factor
MCF-7	Michigan Cancer Foundation-7
MEM	minimal essential media
MHCII	major histocompatibility complex class II
miRNA	micro RNA
ml	millilitre
mRNA	messenger ribonucleic acid
MSA	mouse serum albumin

Nab	nuclear polyadenylated RNA binding protein
NELF	negative elongation factor
NIPBL	nipped-B like protein
Nf1	neurofibromatosis type 1 locus
NF-κB	nuclear factor kappa-light-chain-enhancer of activated B cells
nmt2	N-myristoyltransferase 2
Np13	Arginine methylation of yeast mRNA-binding protein 13
Nrd1	nuclear pre-mRNA down regulation
NRO	nuclear run-on
nt(s)	nucleotide(s)
NTP	ribonuceloside triphosphate
OCT4	octamer-binding transcription factor 4
OD	optical density
ORF	open reading frame
PABP	poly(A) binding protein
PAGE	polyacrylamide gel electrophoresis
Paf1C	RNA Polymerase-associated Factor 1 Complex
PAP	poly(A) polymerase
PARP1	poly(ADP-ribose)polymerase-1
PBS	phosphate buffered saline
PC4	positive cofactor 4
Pcf11	polyadenylation cleavage factor 11
PCR	polymerase chain reaction

Pds5	precocious dissociation of sisters
PE	positioning element
PEG	polyethylene glycol
PET	PNS/EDTA/trypsin
PIC	pre-initiation complex
PIPES	piperazine-N, N'-bis 2-ethanesulphonic acid
PMSF	phenylmethylsulfonyl fluoride
Pol	polymerase
Poly(A)	polyadenylation
pre-mRNA	pre-messenger ribonucleic acid
PSF	protein-associated splicing factor
PTB	polypyrimidine tract-binding protein
P-TEFb	positive transcription elongation factor b
Py(A)n	pyrimidine rich stretch at poly(A) site
Rad21	yeast Scc1; double-strand break repair protein
Rai1	rat interacting protein 1
Rat1	ribonucleic acid trafficking 1
RBS	Roberts-SC phocomelia syndrome
rDNA	ribosomal deoxyribonucleic acid
RFX	regulatory factor X
RPA	RNase protection assay
Rpb	RNA polymerase II subunit B
rpm	revolutions per minute
RNA	ribonucleic acid
RNAi	RNA interference

RNase	ribonuclease
rNTPs	ribonucleotide triphosphates
rRNA	ribosomal ribonucleic acid
RT	reverse transcription
rut	rho utilisation site
SA	stromalin (stromal antigen)
S. cerevisiae	*Saccharomyces cerevisiae*
Scc	sister chromatid cogesion
SDS	sodium dodecyl sulphate
Sen1	senataxin (yeast)
Ser2P	CTD Ser 2 phosphorylated
Ser5P	CTD Ser 5 phosphorylated
Ser7P	CTD Ser 7 phosphorylated
SETX	senataxin
siRNA	small interfering RNA
SMC	structural maintenance of chromosomes
snoRNA	small nucleolar RNA
snRNA	small nuclear ribonucleic acid
snRNP	small nuclear ribonucleoprotein
SOX2	SRY (sex determining region Y-) box 2
Spt5	suppressor of Ty5 homolog
S. pombe	*Schizosaccharomyces pombe*
SRp20	serine/arginine protein family protein 20 (= serine/arginine rich splicing factor 3)
ssDNA	single stranded deoxyribonucleic acid

Ssu72	suppressor of Sua7, gene 2
SUB1	suppressor of TFIIB mutations (yeast)
Suz12	suppressor of zeste 12
T	thymine
TAF3	TATA-box binding protein associated factor
Tat	trans-activator of transcription
TBP	TATA-box-binding protein
TCR	T-cell receptor
TF	transcription factor
TNF	tumour necrosis factor
TRAMP	Trf4/Air2/Mtr4 polyadenylation complex
tRNA	transfer ribonucleic acid
TSS	transcription start site
U	uracil
U2AF	U2 auxiliary factor
ura4	uracil requiring 4
USE	upstream sequence element
UTP	uridine triphosphate
UTR	untranslated region
UUE	upstream U rich element
UV	ultraviolet light
Wsb	WD repeat and SOCS box containing
WT	wild type
XL9	X-box-like region
Xrn2	5`-3` exoribonuclease 2

YB-1	Y-Box binding protein
YES	yeast extract with supplements
Yy1	Yin Yang 1

CHAPTER 1

INTRODUCTION

During my thesis I covered a wide spectrum of topics, which will be introduced in the following Chapter. I will start with describing our current understanding of gene transcription by RNA polymerase II, particularly addressing transcription termination and pre-mRNA 3` end processing. Further, I will describe the characteristics and functions of two essential and versatile proteins, CTCF and the Cohesin-complex, which are the key players throughout my work. In the last part, I will introduce the *TNFα* gene as my model for gene transcription regulation studies on endogenous level. I will also provide a brief insight into the field of macrophages, the main *TNFα* expressing cell type.

1.1 Transcription termination and pre-mRNA 3` end processing of RNA Pol II-transcribed protein-coding genes

Eukaryotic, nuclear gene transcription is performed by three structurally related, but functionally distinct, DNA dependent RNA polymerases (Pol), named RNA Pol I, Pol II and Pol III (Cramer et al, 2008). Pol I transcribes ribosomal RNA (rRNA) genes (with exception of 5S rRNA), while Pol III transcribes small non-coding RNAs including transfer RNAs (tRNAs) and 5S rRNA reviewed in (Paule et al, 2000). Pol II on the other hand, is responsible for the transcription of all protein-encoding genes, most of snRNAs as well as microRNAs. Studies on gene transcription regulation performed in this thesis are focused on Pol II-transcribed genes.

Genes are transcribed in a transcription cycle, which is a highly complex process composed of three tightly regulated steps: initiation, elongation and termination. During the initiation step, Pol II, accompanied by specific transcription initiation factors, is recruited to the gene promoter, where it synthesizes the first nucleotides of RNA. Pol II then leaves the promoter to enter the transcription elongation phase, producing a RNA copy of the transcribed gene. Termination occurs as the final step when RNA synthesis stops and the nascent transcript and Pol II are released from the DNA template.

Transcription of protein-encoding genes by Pol II, together with the processing of pre-messenger RNA (pre-mRNA), results in the formation of functional messenger RNA (mRNA), which is translated into protein. Pre-mRNAs are unstable and undergo three processing reactions in order to be protected from degradation and to be targeted to the cytoplasm for subsequent translation (Bernstein et al, 1989; Bousquet-Antonelli et al, 2000; Hilleren et al, 2001; Huang et al, 1996; Sachs et al, 1997; Saguez et al, 2008; Tarun et al, 1996; Vodala et al, 2008). The first reaction involves the addition of a trimethyl cap structure at the 5` end of the pre-mRNA, the so-called 5` cap (reviewed by (Shuman, 2001)). Pre-mRNAs usually consist of both coding exons and non-coding introns. Therefore, in the second processing reaction, introns are removed by splicing (reviewed by (Matlin et al, 2007)). The final reaction step is a cleavage event occurring at the 3` end of most per-mRNAs, which is coupled to the addition of a poly(A) tail to the 5` cleavage product. Transcription and pre-mRNA processing are interconnected and tightly

regulated and this way allow gene expression to occur correctly (Proudfoot et al, 2002).

The major focus of my thesis is to enhance understanding of how variations in a gene 3` end influence the final stage of the transcription cycle, the termination step. Transcription termination is of particular importance as it prevents interference with transcription of downstream gene`s (Greger et al, 1998) and is thought to allow recycling of Pol II for new rounds of transcription. Evidence has been adduced by Mapendano et al., who demonstrated that termination increases transcription re-initiation (Mapendano et al, 2010). Moreover, transcriptional termination elevates mRNA and protein levels by enhancing splicing and preventing degradation of the nascent transcript (West et al, 2009). It was further demonstrated that read-through transcription induces heterochromatin formation for some genes, which in turn can effect gene expression (Camblong et al, 2007; Gullerova et al, 2008). Finally, Pol II termination is tightly linked to 3` end processing. The following topics particularly focus on this linkage.

1.1.1 Mammalian pre-mRNA 3` end processing – cleavage and polyadenylation

The 3` end cleavage and polyadenylation reactions are directed by *cis* elements located in the 3` untranslated region of most pre-mRNAs. Three sequence elements are involved in mammalian 3` end processing together referred to as the poly(A) signal. This comprises the conserved hexanucleotide element AAUAAA or a variant thereof (Proudfoot et al, 1976),

the cleavage site, which generally contains a CA dinucleotide (Sheets et al, 1990) positioned approximately 10-30 nucleotides downstream of the AAUAAA element, and a more variable GU/U-rich downstream sequence element (DSE) found up to 30 nucleotides downstream of the cleavage site (reviewed in (Mandel et al, 2008) (Figure1.1). Some genes also possess a U-rich upstream sequence element (USE) (Hu et al, 2005; Moreira et al, 1998).

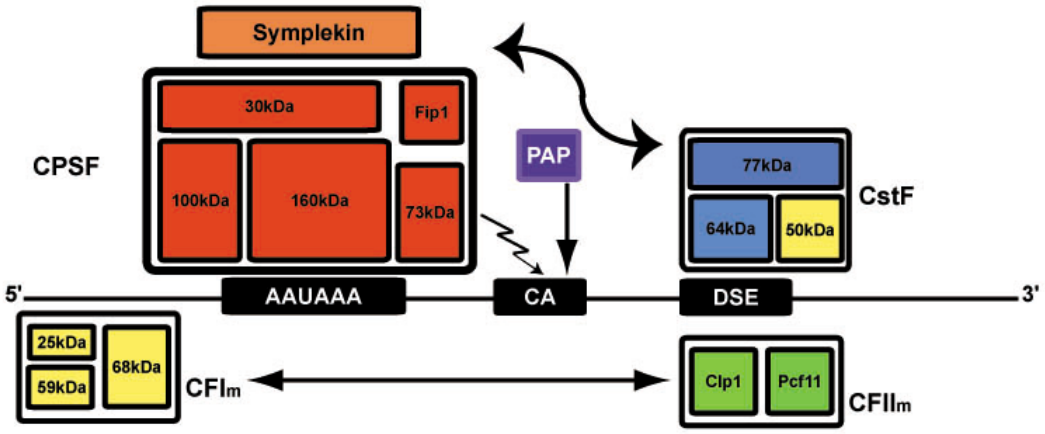


Figure 1.1: Mammalian 3' end processing complexes (adapted from Proudfoot., 2004). The pre-mRNA (black line) and *cis*-acting sequence elements (labelled black boxes, and referred to in the text) are indicated. *Trans*-acting factors are positioned above their points of action and contact with pre-mRNA. Contact between the different complexes is emphasised by the curved arrow. The cleavage site and poly(A) addition is indicated by a vertical arrow. This diagram is not drawn to scale.

In order to cleave the newly synthesized transcript and to add the poly(A) tail, the 3` end processing machinery employs a complex interplay of over 14 polypeptides in mammals and over 20 in yeast (reviewed in (Mandel et al, 2008)). The mammalian core machinery is further associated by

approximately 85 factors. Many of those mediate crosstalk with other processes (Shi et al, 2009). The core mammalian 3` end processing complex consists of the cleavage and polyadenylation specificity factor (CPSF), the cleavage stimulation factor (CstF), cleavage factor I (CFIm), cleavage factor II (CFIIm) poly(A) polymerase (PAP), poly(A) binding protein (PABP), and symplekin (Figure 1.1). CPSF contains 5 subunits, CPSF160, CPSF100, CPSF73, CPSF30, and Fip1 (Kaufmann et al, 2004; Murthy et al, 1992), all of which are required for both, cleavage and polyadenylation (Christofori et al, 1988). CPSF160 recognizes and binds the hexamer AAUAAA element, and it is believed that all the subunits of CPSF contribute to binding efficiency (Keller et al, 1991; Murthy et al, 1995). CstF is composed of three subunits, CstF77, CstF50 and CstF64. The latter makes direct contact with the UGUAN rich DSE (Beyer et al, 1997; Takagaki et al, 1997). CstF is responsible for cleavage but not for polyadenylation (Takagaki et al, 1989). The other member of the core 3` end processing complex, symplekin, is thought to bridge between CPSF160 and CstF64 acting as a scaffold (Hofmann et al, 2002; Takagaki et al, 2000). An additional factor, which binds the pre-mRNA at GU-rich sites upstream of the poly(A) site is CFIm, which is believed to regulate poly(A) site selection (Brown et al, 2003; Ruegsegger et al, 1996; Venkataraman et al, 2005). CFIIm on the other hand contains the highly conserved factors Pcf11 and Clp1. Clp1, which has a kinase activity, interacts with the CFIm and CPSF complexes and is proposed to maintain a 5`-phosphate on the downstream cleavage product aiding termination (de Vries et al, 2000; Weitzer et al, 2007). PAP is responsible for polyadenylation of the cleaved pre-mRNA, whilst PABP binds to the poly(A) tail protecting the pre-

mRNA from exonucleolytic degradation (Bernstein et al, 1989). Additionally, PABP enhances PAP activity and plays a role in controlling poly(A)tail length (Kerwitz et al, 2003; Kuhn et al, 2009). Both proteins are required for polyadenylation, but only PAP is required for cleavage (Colgan et al, 1997; Minvielle-Sebastia et al, 1997).

Assembly of the 3` end processing machinery on the pre-mRNA is a cooperative process. CPSF160, which directly recognizes AAUAAA, binds CstF77 and this way stabilizes the association of both CPSF and CstF with the poly(A) signal. The association between these two complexes is further stabilized by their interaction via symplekin (Takagaki et al, 2000). CPSF160 also interacts with CSTF100, Fip1, CFIm and PAP, reinforcing the interaction of CPSF with the pre-mRNA and recruits PAP to the complex (Colgan et al, 1997; Kaufmann et al, 2004; Murthy et al, 1995). Endonucleolytic RNA cleavage at the poly(A) addition site is most likely carried out by CPSF73 (Mandel et al, 2006; Ryan et al, 2004), followed by polyadenylation of the 5` cleavage product, while the 3` product is degraded.

It is well established that 3` end processing is required for transcription termination. First evidence was provided by studies on α-thalassemia, which was shown to be caused by mutation of the poly(A) signal of the α2-globin gene. This results in disruption of both 3` end processing and termination, and provided evidence that 3` end processing and termination are coupled events (Whitelaw et al, 1986). Later it was shown that the large subunit of Pol II, Rpb1, with its unique carboxy terminal domain (CTD), is required for pre-mRNA processing and transcription termination. Some studies suggested that

the CTD recruits 3` end processing factors to the elongating Pol II complex (McCracken et al, 1997; Zhao et al, 1999) while another proposed that CTD itself acts as a component of the 3` processing machinery (Hirose et al, 1998). Furthermore, the CTD has been shown to play a role in coupling transcription and pre-mRNA processing (Bird et al, 2004) by binding to proteins involved in mRNA biogenesis (reviewed in (Meinhart et al, 2005)).

1.1.2 Pol II termination

1.1.2.1 Termination models

To define the connection between 3` end processing and termination, two models were initially proposed. In both models it is the poly(A) signal that plays a central role. The “allosteric” model suggests that transcription of the poly(A) signal triggers conformational changes of the elongation complex resulting in transcription termination (Logan et al, 1987). These changes could be achieved by dissociation of elongation factors and/or association of termination factors. The second model, called the “torpedo” model, proposes that endonucleolytic cleavage of the pre-mRNA at the poly(A) site produces an unprotected transcript that is rapidly degraded by a 5`-3` exonuclease. In the moment when the exonuclease reaches the polymerase, the elongation complex is destabilised, which results in termination (Connelly et al, 1988). Each model has been supported by several studies, however, increasing evidence suggests that neither model fully explains the mechanism of termination. The present view combines features of both original models and

explains transcription termination via a so-called “hybrid” model (Kaneko et al, 2007; Luo et al, 2006; West et al, 2008).

1.1.2.2 Termination and associated factors

Although the mechanism of termination is only partially resolved, it is certain that a functional poly(A) site is important to render Pol II termination competent (Birse et al, 1997; Connelly et al, 1988; Dye et al, 1999; Logan et al, 1987; West et al, 2008; Whitelaw et al, 1986). Additionally, there is evidence for the particular importance of two Pol II subunits. Indications came form studies in yeast *S. cerevisiae*, which showed that mutations of Pol II subunits Rpb3 and Rpb11 resulted in transcriptional read-through. This lead to the idea that factor binding to the Rpb3/Rpb11 heterodimer causes an allosteric change to the elongating polymerase. Consequently, a “termination signal” might be transduced through the Rpb3/Rpb11 heterodimer (Steinmetz et al, 2006). Deletion and mutation analysis in *S. cerevisiae* revealed a number of factors, which were then identified as termination factors. Homologues of CstF64 (Rna15), CstF77 (Rna14), Pcf11, CPSF (Yhh1), CPSF73 (Ysh1) and Ssu72 all resulted in transcription read-through, when mutated (Birse et al, 1998; Dichtl et al, 2002; Garas et al, 2008; Steinmetz et al, 2003). ChIP experiments then confirmed association of 3` end processing factors with the Pol II elongation complex in yeast (Ahn et al, 2004; Kim et al, 2004a; Luo et al, 2006; Swinburne et al, 2006) and detected components of the 3` processing complex, including CPSF73 and CstF64, at the 3´ end of genes in mammals (Glover-Cutter et al, 2008; Swinburne et al, 2006).

In the course of several studies, the 3` end processing factor Pcf11 received particular attention for its role in transcription termination. Thus its deletion or mutation resulted in 3` end processing and termination defects (Licatalosi et al, 2002; Meinhart et al, 2004; Sadowski et al, 2003; Zhang et al, 2007). Pcf11 contains a conserved CTD-interacting domain (CID), is able to bind nascent RNA, and has been shown *in vitro* to act as a bridge between the CTD and pre-mRNA (Licatalosi et al, 2002; Meinhart et al, 2004; Sadowski et al, 2003; Zhang et al, 2007). Another mutation study in *S. cerevisiae* suggests that RNA cleavage activity is necessary for Pol II termination (Kim et al, 2006). In mammalian cells, Pcf11 depletion reduces termination efficiency and stabilizes the 3` pre-mRNA cleavage product, indicating a role in RNA degradation (West et al, 2008). It seems that Pcf11 acts in two ways: firstly, it supports the 5`-3` RNA exonuclease in degrading the 3` cleavage product and, secondly, it induces conformational changes in the elongating Pol II complex. A further possibility to render Pol II termination competent is the disassociation of factors from the elongation complex, which appear to act as anti-terminators while associated with Pol II. An example is the transcriptional co-activator PC4, which can interact with CstF64 and supposedly suppresses termination until it dissociates from the elongation complex at the poly(A) site (Calvo et al, 2001). Also the elongation factor Paf1C (Betz et al, 2002; Rondon et al, 2004) is found throughout the gene body but levels decrease after passing the through the poly(A) site, indicating dissociation from the elongation complex (Ahn et al, 2004; Kim et al, 2004a). Another example from *S. cerevisiae* is the mRNA binding protein Npl3, which has been shown to interact with the Pol II CTD to stimulate the elongation rate of Pol II (Dermody

et al, 2008). Npl3 also competes with CstF64 (Rna15 in yeast) for pre-mRNA binding and in this way prevents termination (Bucheli et al, 2005; Bucheli et al, 2007). Phosphorylation of Npl3 disrupts the interaction with the CTD and the pre-mRNA, exposing the pre-mRNA to CstF64 (Dermody et al, 2008). ChIP assays confirm this model by showing decreasing Npl3 levels downstream of the poly(A) site (Bucheli et al, 2005).

1.1.2.3 The torpedo model

Supporting evidence for the “torpedo” model of Pol II termination came with the identification of the 5`-3` exonuclease Rat1 in yeast and its human homolog Xrn2, who play an important role in transcription termination (Kim et al, 2004b; West et al, 2004). Mutated Rat1 has been shown to be unable to complement termination defects in a mutant rat1 strain demonstrating the significance of the exonuclease activity in termination (Kim et al, 2004b). ChIP analysis for Rat1 showed that this factor is recruited to the 3` end of genes, possibly through interaction with the 3` end processing factors (Kim et al, 2004b). In fact, Rat1 recruitment to the 3` end of genes is decreased by inactivation of the 3` end processing factor Pcf11 (Luo et al, 2006). Moreover, Rat1 is responsible for recruitment of Pcf11 and CstF64 to the 3` end of genes, suggesting that there is interdependent recruitment of Rat1 and 3` end processing factors. This indicates a dual role of Rat1, namely in 3` end processing as well as in termination (Luo et al, 2006). Although Rat1 has been shown to be part of the transcription termination machinery, Rat1 activity alone is not sufficient for termination *in vivo* (Kim et al, 2004b), indicating that

degradation of the 3` product of poly(A) site cleavage is not sufficient to terminate Pol II (Kim et al, 2006; Luo et al, 2006).

It has been shown that Sen1, an RNA/DNA dependent ATPase (Kim et al, 1999) cooperates with Rat1 to promote termination (Kawauchi et al, 2008). Sen1 interacts with the Pol II CTD and has been shown to resolve RNA:DNA hybrids, known as R-loops, which are formed downstream of the poly(A) signal during transcription, and accumulate in absence of Sen1 (Mischo et al, 2011). The current model suggests that Sen1 resolves R-loops, and this way exposes the 3` endonucleolytic cleavage product to the action of Rat1 (Kawauchi et al, 2008; Mischo et al, 2011).

In human cells, homologues of Rat1 and Sen1 have been shown to be similarly required for termination (Kaneko et al, 2007; Skourti-Stathaki et al, 2011; Suraweera et al, 2009; West et al, 2004). Deletion of the Rat1 homologue Xrn2 leads to a termination defect and accumulation of the 3` cleavage product of the human β-globin gene transcript (West et al, 2004). Xrn2 has also been detected at the 3` end of the human β-actin gene with the highest levels just after the poly(A) signal (Kaneko et al, 2007). This is consistent with the observation of Rat1 positioning in yeast (Kim et al, 2004b; Luo et al, 2006). Recruitment of Xrn2 to the 3` end processing machinery is required for degradation of the 3` cleavage product, but not for the cleavage reaction itself (Kaneko et al, 2007). Regarding the β-globin transcript, Xrn2 entry occurs following endonucleolytic cleavage at the co-transcriptional cleavage (CoTC) terminator element (discussed below) (West et al, 2004).

Alternatively for genes which do not possess CoTC terminators, cleavage at the poly(A) site creates an entry site for Xrn2 (Gromak et al, 2006).

Xrn2 has also been shown to interact with CstF64 and the multifunctional p54nrb/NonO(p54)-protein-associated splicing factor (PSF) (Kaneko et al, 2007). PSF and p54 interact with each other and have been implicated in transcription, splicing and polyadenylation (Liang et al, 2006; Rosonina et al, 2005). *In vitro* experiments suggest that, like Xrn2, p54/PSF is required for degradation of the 3` cleavage product. Moreover, p54/PSF appears to recruit Xrn2 to the 3` end processing machinery to ensure optimal Xrn2 activity (Kaneko et al, 2007). Recently, the human homologue of Sen1, Senataxin (SETX), has been shown to recruit Xrn2 to the 3` end of mammalian genes (Skourti-Stathaki et al). SETX has so far been linked to transcription initiation, termination, pre-mRNA processing and DNA damage repair (Skourti-Stathaki et al; Suraweera et al, 2007; Suraweera et al, 2009). Like Sen1, SETX interacts with Pol II and RNA processing factors. Mutation or depletion of SETX results in a changed Pol II profile, reduced mRNA levels and a termination defect (Skourti-Stathaki et al; Suraweera et al, 2009). As previously shown in yeast, R-loops are also formed in mammals in a transcription dependent manner and interestingly, their formation and resolution appears to reflect essential steps in transcription termination. R-loops further depend upon a functional poly(A) signal and a termination pause element (Skourti-Stathaki et al). In this regard, R-loop formation over the human β-actin gene termination pause element allows Pol II to pause downstream of the poly(A) signal. SETX then helps to recruit Xrn2 and

resolves the R-loops, which exposes the 3` poly(A) site cleavage product for degradation by Xrn2 (Skourti-Stathaki et al).

1.1.2.4 The hybrid termination model

So far evidence has been accumulated that a number of conditions are required for efficient termination to occur. These include a functional poly(A) signal, endonucleolytic cleavage of the nascent transcript, involvement of an exonuclease (Xrn2/Rat1), RNA:DNA hybrids and an RNA:DNA helicase (SETX/Sen1), the association of termination factors like Ssu72, Pcf11 and 3` end processing factors and, finally, dissociation of anti-termination factors like PC4 (Sub1) and Npl3. All these requirements point towards a combined Pol II termination model (Figure 1.2). More evidence for a hybrid model has been provided by the study from West et al. (2008) which shows that transcription of the poly(A) signal is necessary to render Pol II termination competent as proposed by the “allosteric” termination model (Logan et al, 1987). However, West et al. (2008) also showed that RNA cleavage at the poly(A) site or at a downstream CoTC mediated terminator element, creates an entry point for Xrn2 to degrade the 3` cleavage product, leading to termination as postulated in the torpedo model (Connelly et al, 1988; West et al, 2008). This study also shows that Pol II is only released from the DNA template following degradation of the 3` cleavage product when the poly(A) signal has been transcribed. Thus it becomes evident that the two original termination models are not only connected but appear to be non-exclusive.

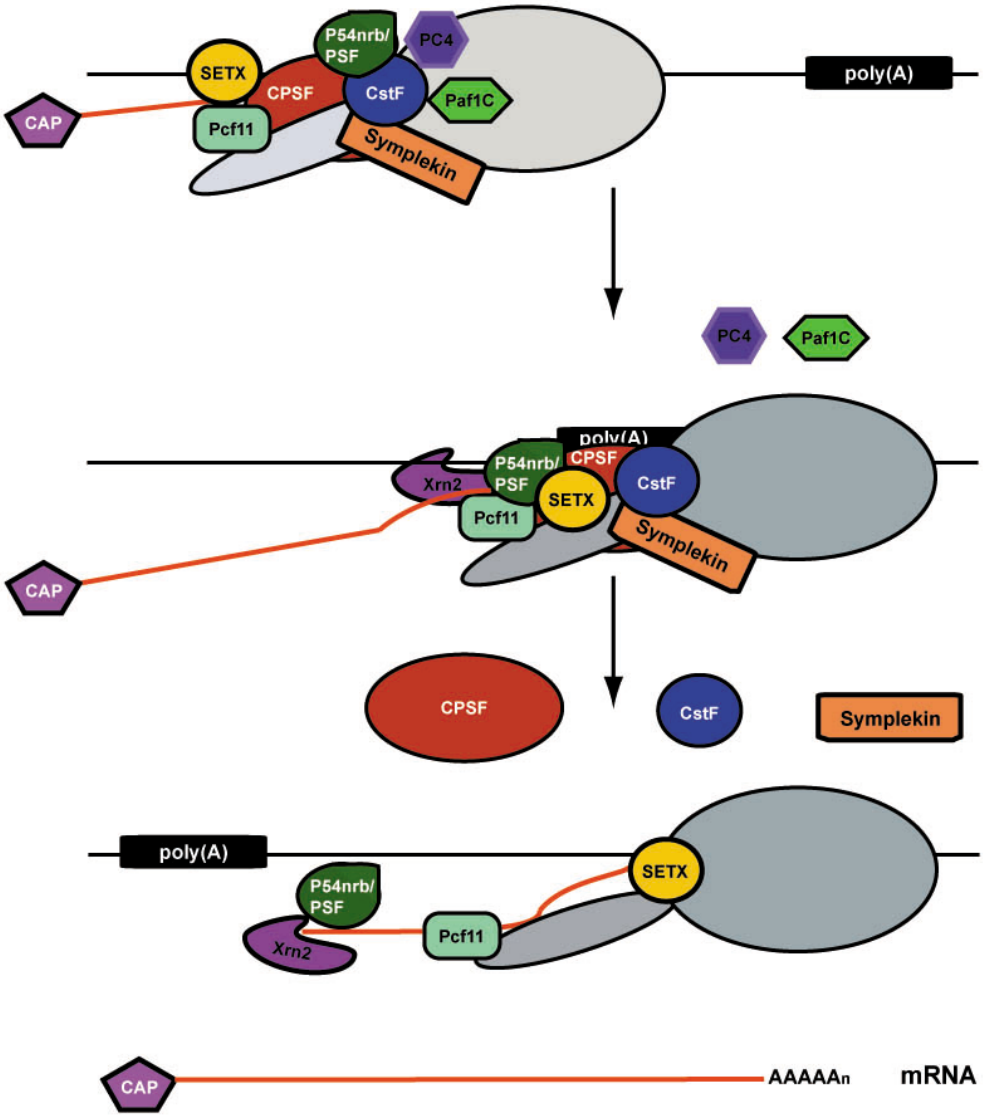


Figure 1.2: Mammalian Pol II termination. Elongating Pol II is represented by a light grey oval and is shown with several associated factors. Upon recognition of the poly(A) signal (black rectangle), Pol II becomes termination competent (dark grey oval). Factors that may act as anti-terminators, Paf1C and PC4, are released upon transcription of the poly(A) signal. Factors important for termination are recruited to Pol II. Xrn2 is probably recruited by p54nrb/PSF and Senataxin (SETX). Cleavage at the poly(A) signal leads to polyadenylation and release of the mRNA, and provides an entry site for Xrn2. Xrn2 degrades the 3′ RNA , with the help of SETX which resolves formation of R-loops. This leads to release of Pol II from the DNA template. Other factors, which are required for termination and not depicted here, include chromatin remodeling factors and individual Pol II subunits (Rpb3 and Rpb11). This diagram is not drawn to scale.

1.1.2.5 Terminator sequence elements

For a number of genes it has been shown that transcription termination requires sequence elements, named terminators, which are located downstream of the poly(A) site. So far, two major classes of terminator elements have been described, so called pause and CoTC elements.

1.1.2.5.1 Pause elements

Transcriptional pausing has been demonstrated by Pol II accumulation at promoter regions and downstream of poly(A) signal (Glover-Cutter et al, 2008). Several protein coding genes like human *α2-globin*, *γ-globin* and murine *IgM* contain sequences downstream of their poly(A) site which cause Pol II pausing and their deletion results in a decrease of termination efficiency (Ashfield et al, 1991; Enriquez-Harris et al, 1991; Gromak et al, 2006; Plant et al, 2005). An example for such a pause element is the MAZ element, which is a G-rich sequence found at the 3` end of the C2 complement gene. When repeated four times, MAZ mediates pausing and facilitates termination (Ashfield et al, 1994; Yonaha et al, 1999). Another G-rich region located downstream of the human *β-actin* poly(A) signal has been suggested to direct termination (Gromak et al, 2006). The efficiency of pause sites as transcriptional terminators, however, appears to correlate with the strength of the poly(A) signal (Gromak et al, 2006; Plant et al, 2005). Poly(A) signal strength, which determines the speed with which the 3` end processing apparatus assembles on the nascent transcript (Chao et al, 1999), has been suggested to pause Pol II. This is mediated by CPSF and appears independent of the CTD or any further downstream pause elements (Nag et

al, 2006; Nag et al, 2007; Orozco et al, 2002). In *S. pombe*, the *ura4* and *nmt2* genes both possess downstream sequence elements (DSE) located downstream of their poly(A) signal, which function as pause elements and are required for termination (Aranda et al, 1999; Birse et al, 1997).

Pause elements are likely to be a common feature of termination and are employed by Pol I, Pol II and Pol III complexes and even by bacterial polymerases. However, a consensus sequence has yet to be defined. Pause elements are so far thought to promote termination by slowing transcription elongation down and this way providing 5`-3`exonucleases with a bigger time window to degrade the nascent RNA and `catch up` with the transcribing Pol II (Gromak et al, 2006) (Figure 1.3a).

1.1.2.5.2 Co-transcriptional cleavage (CoTC) elements

The best characterized CoTC element is the 850bp AT-rich sequence located approximately 900bp downstream of the human β-globin poly(A) site, which has been shown to be required for Pol II termination (Dye et al, 2001; White et al, 2012). The term CoTC originates from the work of Dye and Proudfoot (2001), who demonstrated that the β-globin terminator element is cotranscriptionally cleaved at multiple positions within the human β-globin gene CoTC terminator transcript (Dye et al, 2001).

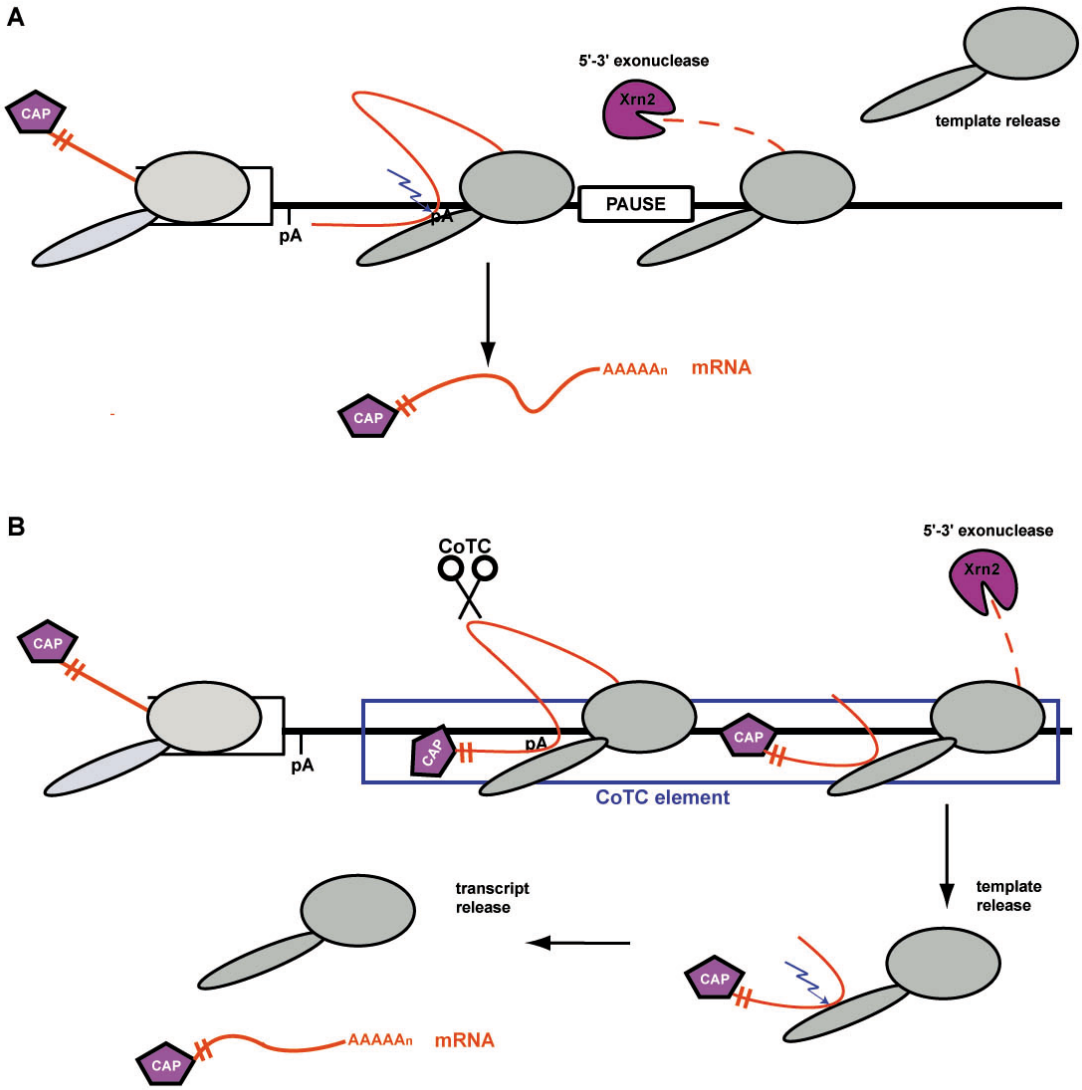


Figure 1.3: Mammalian Pol II termination at pause and CoTC terminator elements. **A)** Pause-mediated termination. Elongating Pol II (light grey oval), DNA (black line) and nascent RNA transcript (red line) are indicated. Upon recognition of the poly(A) signal (pA), Pol II becomes termination competent (dark grey oval). Cleavage at the poly(A) signal (lightning bolt), releases the mRNA and creates an entry site for the exonuclease Xrn2 (purple). At the pause element (white rectangle), the elongation rate decreases allowing more time for Xrn2 to degrade the nascent RNA and 'catch up' with Pol II leading to release of Pol II from the DNA template. **B)** CoTC-mediated termination. As above, recognition of the poly(A) signal causes Pol II to become termination competent. Cleavage of CoTC element (white rectangle) transcript, creates an entry site for the exonuclease Xrn2 (purple) to degrade the downstream transcript and 'catch up' with Pol II leading to release of Pol II from the DNA template with the pre-mRNA still attached to it. Only after cleavage at the poly(A) signal is the mRNA released from Pol II. This diagram is not drawn to scale.

Other examples of CoTC termination elements have been described for the human *ε-globin* gene (Dye et al, 2001) and the murine serum albumin (*MSA*) gene (West et al, 2006). The CoTC termination element of the *MSA* gene has multiple termination promoting elements within its sequence. The *MSA*-CoTC is located approximately 800bp downstream of the poly(A) signal. These elements are cleaved co-transcriptionally at multiple positions equivalent to the human β-globin gene (West et al, 2006).

The CoTC terminators described so far are located further downstream from the poly(A) site than the pause elements. Transcripts of the CoTC terminators are highly unstable and therefore provide an entry site for Xrn2 to degrade the 3` cleavage product (West et al, 2004). Transcript degradation then leads to release of Pol II with the pre-mRNA from the DNA template. The pre-mRNA is subsequently cleaved at the poly(A) site and mRNA is released from Pol II (West et al, 2008) (Figure 1.3b). This shows that 3` end processing of β-globin mRNA occurs post-transcriptionally. CoTC-mediated termination has also been shown to enhance mRNA levels. This is possibly a consequence of pre-mRNA release from the site of synthesis where nuclear RNA surveillance occurs (West et al, 2009). The mechanism of CoTC mediated cleavage event is not entirely resolved. One study suggests involvement of secondary structures in the human β-globin gene where autocatalytic cleavage by this structure resulted in transcriptional termination (Teixeira et al, 2004). However, very low cleavage efficiency occurs *in vitro* suggesting that an endonuclease is likely to be required *in vivo*.

1.1.3 Connections between capping, splicing, 3` end processing and termination

Pre-mRNA processing involves capping, splicing and 3` end cleavage/polyadenylation. All these processes are connected to each other and have to function in a coordinated manner. The first processing step is capping, which co-transcriptionally adds a trimethyl cap structure to the 5` end of the nascent transcript. The cap-binding complex (CBC), which is composed of the two subunits CBC20 and CBC80 binds the cap structure (McCracken et al, 1997; Visa et al, 1996). Evidence from *S. cerevisiae* suggests that a defect in capping leads to premature Rat1 dependent termination, indicating a link between capping and transcription termination (Jimeno-Gonzalez et al, 2010). In mammals, depletion of CBC80 leads to reduced 3` end cleavage. However, polyadenylation was not particularly affected (Flaherty et al, 1997). This could possibly be explained by a loss of physical contact between the CBC and the 3` end processing machinery and suggests that CBC affects Pol II termination indirectly (Flaherty et al, 1997).

Another interconnection occurs between pre-mRNA splicing and termination. An example for this is the case of terminal exon definition, which has been shown to co-transcriptionally enhance poly(A) site efficiency and in this way enhance termination (Dye et al, 1999; Niwa et al, 1991; Niwa et al, 1992) .

Interconnection between splicing and 3` end formation has been suggested by a study demonstrating enhancement of 3` end processing through

interactions of CPSF with components of the U2 snRNP, which associate with the intronic 3` splice site and the branch point (Kyburz et al, 2006). Moreover, U2AF65, a subunit of the splicing factor U2AF, co-purifies with CFIm and CFIIm complexes of the 3` end processing machinery. U2AF associates with the 3` splice site and contacts PAP, which enhances the poly(A) site function. Interaction with PAP then stimulates U2AF65 to bind to introns and enhance splicing (Vagner et al, 2000).

Another link between splicing and termination is the splicing factor SRp20, which has been identified in a genome wide RNAi screen for termination factors in *C. elegans* (Cui et al, 2008). SRp20 co-immunoprecipitates with CFIm (Dettwiler et al, 2004) and affects polyadenylation in Rous sarcoma virus (Maciolek et al, 2007). One possibility is that SRp20 is either involved in degradation of the 3` cleavage product or in release of Pol II from the DNA template (Cui et al, 2008).

The polypyrimidine tract-binding protein (PTB), a factor involved in splicing, was found at the 3` end of genes (Swinburne et al, 2006). Interestingly, PTB has been shown to inhibit 3` end processing by competing with CstF64 for binding to the GU/U rich DSE region of mammalian poly(A) sites (Castelo-Branco et al, 2004).

Finally, the human RNA:DNA helicase SETX, which has previously been described to play a role in termination, has also been shown to affect splicing efficiency and alternative splice site selection, possibly through inhibition of

the interaction of splicing factors with their cognate sequences, or due to a reduction in Pol II transcription elongation rate (Suraweera et al, 2009).
However, despite described evidences it is inevitable to consider that the effects of splicing on termination might also be of indirect nature by affecting cleavage efficiency.

An important example for how efficiency of one process can affect a concurrent or following process is provided by the study of West and Proudfoot (2009), which shows that efficiency of transcription termination can enhance gene expression by promoting pre-mRNA processing. This is suggested to be mediated by positive post-transcriptional effects on pre-mRNA splicing (West et al, 2009). Also, evidence has been provided for a link between termination and transcriptional initiation (Mapendano et al, 2010). Here, 3` end formation was shown to stimulate re-inititation by recycling of factors from the terminator to the promoter for continued transcription. This finding suggests that termination possibly releases transcription factors.

1.2. CTCF and Cohesin

The CCCTC-binding factor (CTCF) is a multivalent factor and the spectrum of its various roles in mammalian gene expression is steadily growing. Numerous studies over the recent years reported on interactions between CTCF and the Cohesin complex for gene expression regulation purposes (reviewed in (Phillips et al, 2009)). The key mechanism underlying such regulatory involvements includes CTCF-Cohesin mediated higher-order

chromatin structures, in short, chromatin loops. This thesis includes studies of the involvement of CTCF and Cohesin on steps of the transcription process itself (transcription termination studies) as well as studies on cell type specific influences of gene expression regulation.

1.2.1 CTCF

1.2.1.1 General introduction into the biology of CTCF

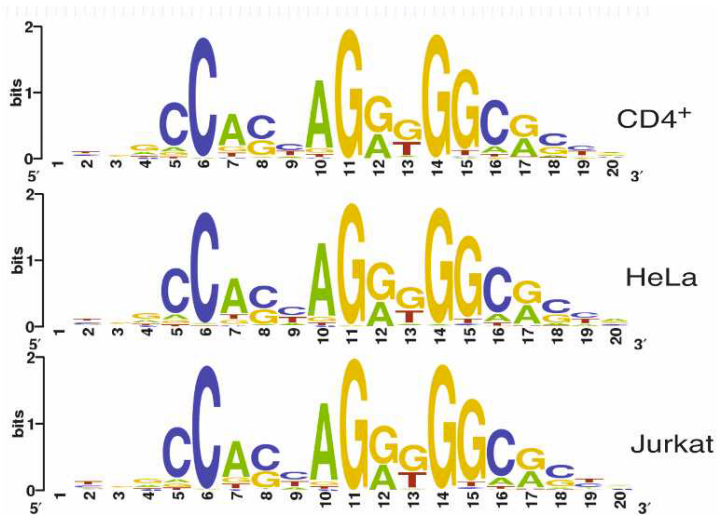
The CTCF factor is a highly conserved and ubiquitously expressed 11-zinc finger DNA-binding protein found in higher eukaryotes. The central DNA binding domain shares close to 100% homology between mouse, chicken and human, while the N- and C-termini around are slightly more divergent (Ohlsson et al, 2001). Through combinatorial use of different zinc fingers, CTCF is able to bind to a wide range of sequences as well as specific regulatory proteins (Filippova et al, 1996). This provides a basis for a variety of regulatory functions that CTCF is implicated in. Such regulatory involvements are: transcriptional activation and repression, genomic imprinting, hormone-responsive gene silencing, X-chromosome inactivation, enhancer blocking and/or barrier gene insulation and long-range chromatin interaction (Ohlsson et al, 2001; Phillips et al, 2009). CTCF expression levels and nuclear distribution patterns vary in a cell type-specific manner. CTCF expression levels appear to influence cellular functions, as ectopic overexpression and/or RNAi-based CTCF depletion in mammalian cells leads to transcription misregulation of many genes and linage-specific effects on

growth, differentiation, proliferation and apoptosis (Qi et al, 2003; Torrano et al, 2005).

The genome-wide distribution patterns of CTCF in different cell types has been investigated in several recent studies, employing chromatin immunoprecipitation (ChIP) in combination with tilling arrays (ChIP-chip) or high-throughput sequencing (ChIP-Seq) and Chromatin Interaction Analysis by Paired-End Tag Sequencing (ChIA-PET) (Fullwood et al, 2009). A ChIP-chip based report identified 13,804 CTCF binding sites in chromatin from human fibroblasts, globally distributed across chromosomes with 46% intergenic, 22% intronic, 12% exonic and 20% within 2.5 kb of promoters (Kim et al, 2007). Another ChIP-Seq based study identified 26,814 CTCF binding motifs in resting CD4+ T cells (Barski et al, 2007; Jothi et al, 2008) with a genome-wide distribution pattern of 45% intergenic, 7% 5`UTR, 3% exonic, 29% intronic, 2% 3`UTR and 13% within 5kb of the transcription start site (TSS). The genomes of mouse embryonic stem (ES) cells were shown to have 39,609 CTCF target sites (Chen et al, 2008), while for HeLa and Jurkat cells it was only 19,308 and 19,572, respectively (Cuddapah et al, 2009). Most recently, ChIA-PET analysis revealed 39,371 CTCF binding sites in mouse ES cells (Handoko et al, 2011), confirming the previously published count of 39,609 CTCF sites (Chen et al, 2008). It so far remains an open question, whether these cell type specific differences in CTCF occupancy are functionally significant or whether they merely reflect differences in computational and experimental procedures employed during analyses. Importantly, these genome-wide studies have enabled the identification of an

approximately 11-15 bp consensus sequence, which forms the core of CTCF-binding sites and appears highly consistent in all assessed cell types (Figure 1.4).

A



B

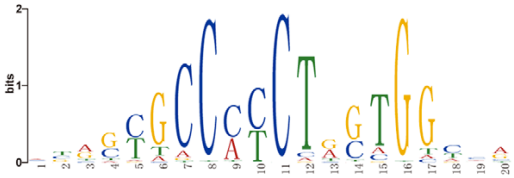


Figure 1.4 Significantly enriched CTCF consensus motifs within ubiquitous CTCF-binding sites taken from recent CTCF ChIP-sequencing (ChIPSeq) analyses. **A** Represents the enriched CTCF-binding motif from three cell lines: CD4+ cells, HeLa and Jurkat cells (taken from Cuddapah S et al, 2009). **B** Represents the CTCF-binding motif obtained from ChIPSeq analysis across 38 different cell lines (taken from Chen H et al, 2012)

This stands in contrast to earlier studies, which identified an extended sequence of 50-60 bp after stepwise deletion of single CTCF fingers (Ohlsson et al, 2001). It is suggested that binding may be partially stabilized by interactions of peripheral fingers and nucleotides surrounding the core

sequence. In fact, Ohlsson et al (2001) then proposed that during the formation of a CTCF-DNA complex, both DNA and CTCF allosterically customize their conformation to engage different zinc fingers (Ohlsson et al, 2001). This could serve the purpose of making base contacts or provide a target-specific surface that determines interactions between CTCF and other proteins. The current model accounting for the functional versatility of CTCF relies on these CTCF zinc finger conformations to explain divergent consensus sequences and different protein binding partners.

The first evidence for the transcriptional roles of CTCF came from a study that described the cloning of the CTCF gene (Lobanenkov et al, 1990). CTCF was found to bind sequences within the promoter-proximal regulatory region of the chicken *c-myc* gene (Klenova et al, 1993; Lobanenkov et al, 1990) and immediately downstream of the two conserved alternative transcriptional start sites in the human and mouse *c-myc* gene (Filippova et al, 1996). In these and other studies (Kohne et al, 1993), CTCF was shown to function as a transcriptional repressor. On the other hand, CTCF was also shown to serve as a transcriptional activator (Vostrov et al, 1997). However, these transcriptional activator/repressor roles were investigated *in vitro*. It still remains to be investigated, whether CTCF functions similarly in a native *in vivo* context of an endogenous genomic locus.

In line with the variety of CTCF`s functional involvements, several CTCF binding partners have been identified. These are involved in different processes and involve transcription factors Yy1, Kaiso and YB-1 (Chernukhin

et al, 2000; Defossez et al, 2005; Donohoe et al, 2007), chromatin modifying proteins (Sin3a, CHD8, Suz12) (Ishihara et al, 2006; Li et al, 2008; Lutz et al, 2000) the nuclear protein nucleophosmin (Yusufzai et al, 2004), RNA polymerase II (Chernukhin et al, 2007) and poly(ADP-ribose) polymerase-1 (PARP1) (Yu et al, 2004).

Nevertheless, genome-wide CTCF binding patterns suggest roles for CTCF, which are distinct from those of canonical transcription factors (Phillips et al, 2009). The genomic distribution of CTCF correlates with gene density but not with transcriptional start sites. Moreover, the binding pattern of CTCF does not seem to predict cell-specific gene expression (Chen et al, 2008). Domains with few or no CTCF binding sites often include clusters of transcriptionally co-regulated or developmentally related genes and these regions tend to be flanked by CTCF-binding sites (Kim et al, 2007). The same study also revealed that CTCF-rich regions contain genes which display extensive alternative promoter usage, naming examples like *Tcrβ, Tcrα/δ* and *Igλ* L chain loci (Kim et al, 2007).

1.2.1.2 CTCF and its insulator role

Insulators are classically defined by two characteristics. Firstly, by their ability to block communication between adjacent regulatory elements in a position-dependent manner (enhancer blocking) and secondly, by their capacity to buffer genes from position effects caused by the spread of repressive heterochromatin from adjacent sequences (barrier). The first link between CTCF and enhancer blocking insulation was provided by the discovery of CTCF-binding to the 5` HS4 insulator sequence upstream of the chicken *β*-

globin locus (Bell et al, 1999). This was followed by the identification of another CTCF-dependent insulator at the 3` end of this domain (Saitoh et al, 2000) and the discovery of four CTCF-binding sites within the Imprinted Control Region (ICR) of the mammalian *H19/Igf2* locus (Bell et al, 2000; Hark et al, 2000; Kanduri et al, 2000; Szabo et al, 2000). CTCF-dependency was also shown for the *DM1* locus (Filippova et al, 2001) and the boundaries of domains that escape the X chromosome inactivation (Filippova et al, 2005). Computational analyses have shown that CTCF binds to highly conserved, non-coding elements across 12 mammalian species, supporting the idea of a global role for CTCF as an insulator protein (Xie et al, 2007). Although CTCF is mainly considered to be involved in enhancer-blocking mechanisms, substantial genome-wide data also provide evidence for CTCF binding sites at the boundaries between active and repressive chromatin domains. This may reflect roles for CTCF in barrier insulation (Barski et al, 2007; Cuddapah et al, 2009).

1.2.1.2.1 CTCF-mediated gene insulation via chromatin loops

Linear models of gene transcription regulation provide only a partial picture and have recently been extended by discoveries of three-dimensional structures of chromatin within the nuclear space (Figure 1.5). The identification of CTCF`s involvement in mediating long-range chromatin interactions (Gerasimova et al, 2000; Yusufzai et al, 2004) was followed by the first direct evidence that loop formation can occur via contact between two CTCF-bound insulators *in vivo* (Hou et al, 2008).

1.2.1.2.2 CTCF-mediated loop formation at endogenous loci

Significant evidence for CTCF-mediated enhancer blocking insulation via long-range chromatin interactions in the native genomic context comes from studies on the imprinted *H19/Igf2* locus (Kurukuti et al, 2006; Murrell et al, 2004). Here, CTCF has been shown to play an important role in the establishment and maintenance of allele-specific DNA methylation imprinting and parent-of-origin *H19* and *Igf2* gene expression patterns during development. Four methylation-sensitive CTCF-binding sites are located in the ICR, immediately upstream of the *H19* gene. On the maternal allele, CTCF binds to the unmethylated ICR and forms a tightly coiled chromatin loop structure around the *Igf2* gene. This results in inaccessibility of the *Igf2* promoter to the enhancer downstream of *H19*. On the other hand, the ICR is methylated on the paternal allele, which abrogates CTCF binding and ICR-

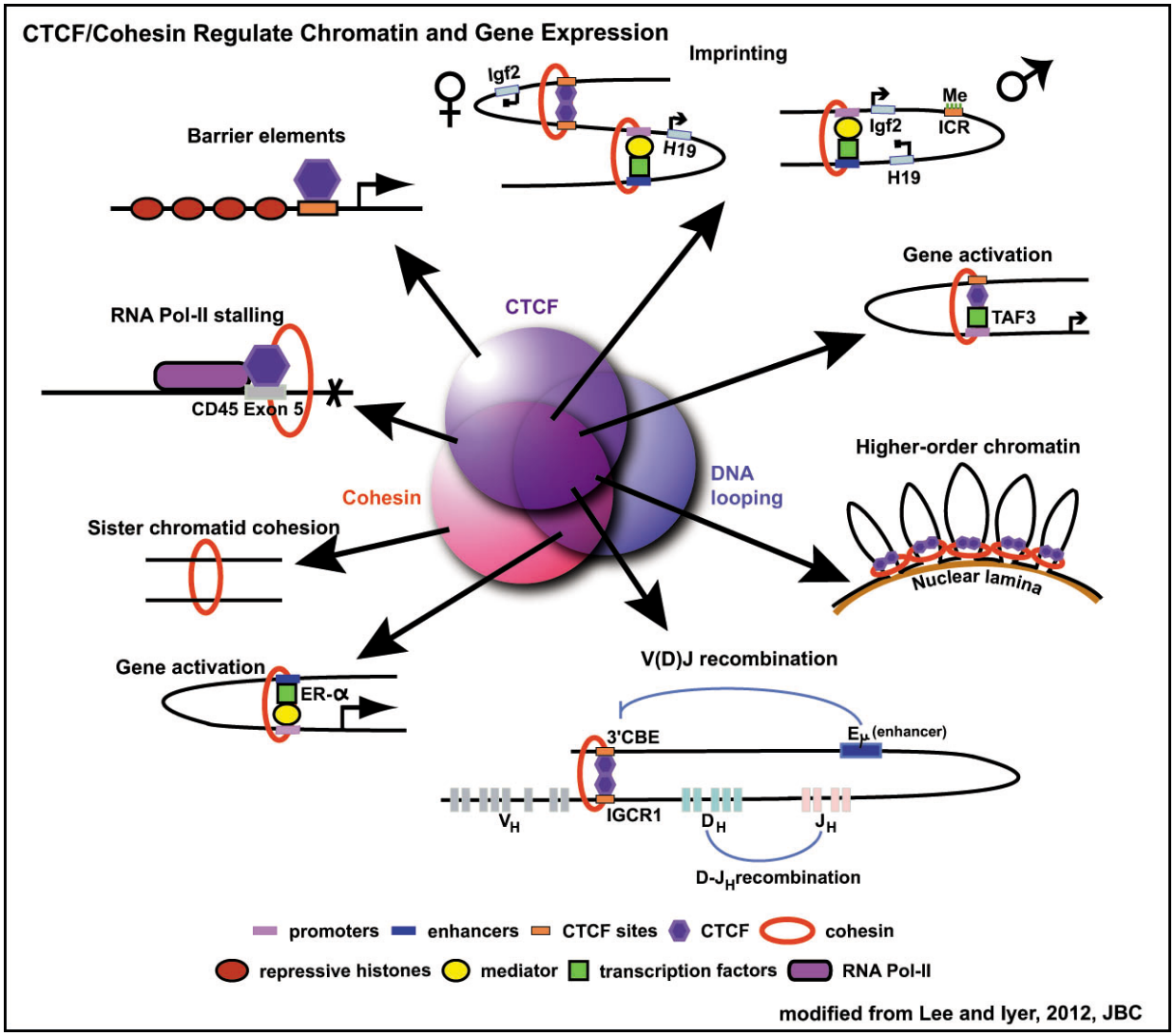


Figure 1.5. **Overview of CTCF and Cohesin functions.** The functional involvements of CTCF and Cohesin partially overlap, and some of these functions involve DNA looping. Distinct combinations of transcription factor binding and DNA looping are observed at different locations and under different cellular conditions. Maternally and paternally inherited chromosomes are indicated by appropriate symbols under *Imprinting*. Although this view is partly speculative, experimental evidence exists for many the presented aspects of CTCF and Cohesin functional contributions. 3' CBE, 3'-CTCF-binding element; Adapted from Lee and Iyer, 2012, JBC.

mediated insulation. This leads to functional communication between the *Igf2* promoter and the enhancer and results in active *Igf2* expression. *H19* is only expressed from the maternal allele, as loss of CTCF binding to the methylated ICR appears to silence *H19* on the paternal allele due to promoter methylation (Figure 1.5).

CTCF-based intra-chromosomal interactions between distal regulatory elements have also been reported at other developmentally regulated gene clusters, importantly, the *MHCII* and *β-globin* loci (Figure 1.5). CTCF binds to three sites upstream (5`HS85, 5`HS62/60 and 5`HS5 in the LCR) and one downstream (3`HS1) of the mouse *β-globin* locus in a cell type specific manner (Bulger et al, 2003; Farrell et al, 2002; Splinter et al, 2006). During erythroid differentiation, an "active chromatin hub" is formed where CTCF-bound regulatory sequences throughout the *β-globin* locus come into close spatial proximity. This occurs when stage-specific *globin* genes also interact with the LCR resulting in looping-out of transcriptionally silent globin genes (Palstra et al, 2003; Splinter et al, 2006; Tolhuis et al, 2002). Conditional CTCF deletion or mutation of the 3`HS1 CTCF-binding site has been shown to result in disruption of CTCF contacts in erythroid progenitors. Surprisingly, no effect could be measured on the kinetics of the *globin* gene expression levels during erythroid differentiation (Splinter et al, 2006). A possible explanation could be that latter may not be readily affected by mutations in one regulatory element. At this point it therefore remains unclear what is the exact functional significance of CTCF-mediated insulation at the 3`HS1 and 5`HS5 of the *β-globin* locus. Genetic disruption of CTCF-binding to HS1 has no effect on the transcription of neighboring olfactory receptor genes (Splinter et al, 2006) and deletion of 3`HS1 and 5`HS5 still shows seemingly unaffected *globin* gene expression (Bender et al, 2006; Bender et al, 1998).

Another example for CTCF-recruitment for enhancer-blocking insulator function is the *MHCII* locus. CTCF binds to the intergenic enhancer XL9,

which co-regulates the expression of the MHCII genes *HLA-DRB1* and *HLA-DQA1* and *in vitro* appears to have enhancer-blocking activity (Majumder et al, 2008). IFN-γ stimulation of non-MHCII-expressing epithelial cells results in expression of the transcriptional co-activator CIITA and the concurrent formation of chromatin interactions between XL9 and the two divergent promoters upstream of *HLA-DRB1* and *HLA-DQA1*. It appears that chromatin loop formation is required for gene expression and depends on XL9-bound CTCF as well as on the co-activator CIITA and the transcription factor RFX, constitutively bound to regulatory sequences within the proximal promoter of the *HLA-DRB1* and *HLA-DQA1* genes. CTCF is therefore involved in gene expression control of the human *MHCII* locus through its ability to form transcriptionally functional chromatin loops. However, it cannot be ruled out that XL9-bound CTCF acts to block inappropriate regulatory elements located outside the *HLA-DRB1* and *HLA-DQA1* region.

1.2.1.3 CTCF as a genome-wide organizer of chromatin architecture

Eukaryotic genomes are organized into functional architectures (Francastel et al, 2000). These higher order genome structures and their associated sub-nuclear compartments are key components, which contribute to several aspects of nuclear activities, including DNA transcription (Bolzer et al, 2005; Misteli, 2007). The spatial distribution and temporal regulation of long-range loops has been shown to be critical for cell identity and cell fate decision (Meshorer et al, 2006). Higher order chromatin organization includes interactions between DNA elements on the same (intra-chromosomal) or different chromosomes (inter-chromosomal) (Dekker et al, 2002; Dostie et al,

2006; Simonis et al, 2006). In line with the view that CTCF acts as a genome-wide organizer of chromatin is its involvement in inter-chromosomal interactions of the maternal (CTCF-bound) *H19/Igf2* ICR locus with the *Wsb/Nf1* locus (Ling et al, 2006). Furthermore, CTCF mediates inter-chromosomal interactions between homologous X chromosomes during the process of X chromosome inactivation (Tsai et al, 2008; Xu et al, 2007). Although the functional details of CTCF-mediated inter-chromosomal contacts remain to be fully elucidated, the *H19/Igf2* ICR at least provides an example, where such contacts appear to control epigenetic information *in trans* (Sandhu et al, 2009; Zhao et al, 2006) (Figure 1.5).

The significant importance of CTCF-mediated looping interactions at two levels in B-lymphocytes has recently been shown for Immunoglobulin heavy-chain (*IgH)* gene locus repositioning within the nucleus and locus contraction in preparation for gene recombination (Guo et al, 2011a). Concurrently, another report provided evidence for an *IgH* V(D)J recombination control region with a CTCF-binding element mediated control of V(D)J gene recombination in developing B-lymphocytes (Guo et al, 2011b). Another study, published within the same series of IgH studies in B-lymphocytes, also revealed a significant role for CTCF in the 3D structure of the *IgH* locus (Degner et al, 2011). Moreover, CTCF appears to be involved in the regulation of antisense germ line transcription and contributes to the compact nature of the *IgH* locus (Degner et al, 2011) (Figure 1.5). Another recent report has provided experimental evidence for CTCF-mediated configuration of the genome into distinct chromatin domains and sub-nuclear compartments, which exhibit unique epigenetic states and transcriptional

activities in mouse ES cells (Handoko et al, 2011). Contrary to the enhancer-blocking model, CTCF associated interactions appear to also promote communications between functional regulatory elements to regulate gene expression (Handoko et al, 2011).

Such major involvements of CTCF in genomic configurations are unlikely to be mediated without the involvement of additional factors. Genome-wide studies have recently provided mechanistic insights into CTCF-mediated chromatin loop formation by uncovering enrichment of the Cohesin protein complex at CTCF-binding sites (Parelho et al, 2008; Rubio et al, 2008; Stedman et al, 2008; Wendt et al, 2008). CTCF appeared to be required for Cohesin recruitment to these sites, as CTCF-depletion leads to loss of Cohesin binding. Cohesin subunits are believed to coalesce into a ring-like structure, which is known to mediate sister chromatid cohesion during mitosis. It is assumed that a similar mechanism could facilitate the stabilization of CTCF-based chromatin loops. Experimental evidence for the latter hypothesis has been provided by a study, which indicated requirement of Cohesin proteins for CTCF-mediated long-range intra-chromosomal interactions at the developmentally regulated *IFNg* locus (Hadjur et al, 2009).

Functional interaction between CTCF and Cohesin has also been observed at sites of CTCF-mediated enhancer-blocking insulator activity, such as the chicken *β-globin* locus HS4 and the *H19/Igf2* locus ICR (Parelho et al, 2008; Wendt et al, 2008). Furthermore Cohesin has also been detected at over 60 sites of CTCF-recruitment throughout the V_H region, all involved in the previously described three-dimensional structure and the compact nature of

the *IgH* locus and in the regulation of antisense germline transcription (Degner et al, 2011) (Figure 1.5). The discovery that a significant proportion of CTCF-binding sites are shared with Cohesin provides good evidence that Cohesin and CTCF act in concert at multiple genomic locations to mediate long-range intra- and inter-chromosomal interactions.

1.2.1.4 CTCF and RNA Pol II

CTCF has been shown to interact with the initiation (hypophosphorlyated) and elongation (hyperphosphorylated) forms of RNA Pol II *in vitro*, with preference for the initiation form *in vivo* (Chernukhin et al, 2007) (Figure 1.5). ChIP-analyses for CTCF-Pol II complex presence across different gene *loci* detected this complex at the β-globin insulator, but only in proliferating erythroblasts that do not express the globin genes. However, in differentiated cells that transcribe some of the globin genes, CTCF-Pol II association with the insulator is lost (Chernukhin et al, 2007). Another example within the same study is the *H19* ICR, where Pol II binding to the ICR requires functional CTCF target sites (Chernukhin et al, 2007). On the other hand, a single CTCF-binding site fused with a promoter-less reporter gene conferred transcriptional activity on the gene. This suggests that CTCF functions similarly to TATA-box-binding protein (TBP) by allowing accurate transcription initiation at some promoters (Zlatanova et al, 2009). Furthermore Pol II has been observed to stall at Cohesin/CTCF binding sites following gene activation possibly to modulate transcription speed and so the timing of RNA processing and the location of checkpoints for Pol II transcriptional regulation (Wada et al, 2009).

Recently the first evidence for a direct CTCF-involvement in a co-transcriptional process has been provided by demonstrating that CTCF can promote inclusion of weak upstream exons by mediating local Pol II pausing. This was shown for *CD45* at a mechanistic level, as well as genome-wide. CTCF-binding to *CD45* exon 5 DNA facilitates exon inclusion through local Pol II pausing, while DNA methylation inhibits CTCF binding and so leads to exon 5 exclusion (Shukla et al, 2011). A similar mechanism may occur for the *Igf2/H19* ICR where DNA methylation inhibits CTCF-binding, which in turn appears to affect splicing (Shukla et al, 2011) (Figure 1.5).

1.2.2 Cohesin

Elucidating the cellular involvement and functional interactions of CTCF, one protein complex stands out as providing a significant contribution to several of CTCF`s multifunctional roles. As described before, Cohesin has been identified in gene expression regulation in the context of CTCF-recruitment to various genomic *loci* (Parelho et al, 2008; Rubio et al, 2008; Stedman et al, 2008; Wendt et al, 2008) (Figure 1.5).

Cohesin is an evolutionarily conserved essential multi-protein complex that is important for higher-order chromatin organization. It plays pivotal roles in the maintenance of genome integrity through mitotic chromosome regulation, DNA repair and replication, as well as gene regulation associated with cellular differentiation. One of the first identified functions of the Cohesin-complex was its important role in mediating sister chromatid cohesion in mitosis (Klein et al, 1999). The complex is proposed to form a ring-like structure (Gruber et al,

2003; Haering et al, 2002) that holds sister chromatids together from the time of their synthesis in S-phase until their separation at the onset of anaphase (Uhlmann et al, 1998). This function not only ensures chromosome segregation at the right time point during mitosis but is also crucial for post-replicative DNA repair (Watrin et al, 2006), and the prevention of inappropriate recombination between repetitive regions (Hirano, 2006; Huang et al, 2006; Kobayashi et al, 2005; Lehmann, 2005; Nasmyth, 2005).

Mutation of human Cohesin or its regulators was found to cause severe developmental disorders known as Cornelia de Lange syndrome (CdLS) and Roberts-SC phocomelia syndrome (RBS) (Krantz et al, 2004; Liu et al, 2008; Musio et al, 2006).

1.2.2.1. Cohesin structure and loading onto DNA

The Cohesin complex is widely conserved and is composed of four core subunits, two long protein molecules known as SMC1 and SMC3 (*s*tructural *m*aintenance of *c*hromosomes) and two smaller subunits termed Scc1/Rad21 and Scc3/SA (Haering et al, 2002). Together, these proteins form a large circle-shaped structure of about 30–40 nm in diameter, which mediates cohesion by embracing sister chromatids as a ring (Gruber et al, 2003; Haering et al, 2002). Cleavage of the Scc1/Rad21 subunit at the onset of anaphase by separase releases Cohesin from chromosomes and allows chromatid segregation (Nasmyth et al, 2000; Uhlmann et al, 1999). The Cohesin core complex interacts with several other proteins (Dorsett, 2007), including the Scc2/Scc4 loading complex, which is required for loading of Cohesin onto the sister chromatids (Ciosk et al, 2000; Watrin et al, 2006).

Scc2/NIPBL is a HEAT-repeat protein that has been proposed to stimulate ATP hydrolysis by the SMC heads upon interaction with Cohesin (Arumugam et al, 2003). The mechanism of Cohesin loading by Scc2-Scc4 is not fully understood. However, ATP hydrolysis by the SMC proteins appears to be important (Arumugam et al, 2003). There are two models of how the DNA strands are incorporated inside the Cohesin ring. One model predicts that the opening of the ring occurs at the hinge (Gruber et al, 2006; Milutinovich et al, 2007). In contrast, the second model proposes that the DNA-hinge interactions stimulate ATP hydrolysis leading to disengagement of SMC heads to allow passage of the DNA between them (Hirano et al, 2006). In any case, both proposed mechanisms point to the importance of the hinge region, which appears to be a target of the regulatory factor Pds5, a modulator of the dynamic association of Cohesin with chromatin (Mc Intyre et al, 2007).

1.2.2.2 Genome-wide Cohesin distribution and functions

In *S. cerevisiae*, Cohesin is localized at centromers, pericentric domains, intergenic regions and at open reading frames (ORFs), without displaying a consensus DNA binding sequence (Glynn et al, 2004). ChIP-on-chip analysis reveals that there is no constant colocalization between Scc2 and Cohesin. Also, a significant proportion of detectable Cohesin is located between convergently transcribed genes in G2 phase of the cell cycle whereas in late G1 phase Cohesin is found at different locations. This leads to the idea that Cohesin and Scc2-Scc4 interact only at the initial loading site but then Cohesin slides towards the ends of genes pushed by the transcription machinery (Glynn et al, 2004; Lengronne et al, 2004). In comparison, the

Drosophila melanogaster Scc2-homolog Nipped-B and Cohesin colocalize throughout the detected sites, especially in actively transcribed regions and may even facilitate Cohesin loading (Misulovin et al, 2008). Mammalian cohesion is often detected at sites of CTCF-recruitment, which bear a conserved consensus sequence (Parelho et al, 2008; Rubio et al, 2008; Stedman et al, 2008; Wendt et al, 2008), as described above. The study revealing 13,894 CTCF-binding sites across the human genome also identified 8811 SCC1 sites, with an 89% overlap between both factors (Wendt et al, 2008). Like CTCF, the major proportion of Cohesin (49%) is located at intergenic regions. Another study mapping different Cohesin subunits such as SMC1 and SMC3 in mouse embryonic ES cells showed similar results (Kagey et al, 2010).

As described above, CTCF-Cohesin mediates long range chromatin loops (Gause et al, 2008; Hadjur et al, 2009; Mishiro et al, 2009; Nativio et al, 2009; Nitzsche et al, 2011; Schmidt et al, 2010). Although CTCF-binding sites are highly conserved across cell types (Cuddapah et al, 2009; Lee et al, 2012), loop formation at these CTCF sites occurs in a cell type-specific manner (Hou et al, 2010). How Cohesin-mediated cell type-specific looping is achieved, remains as yet unresolved. Also described before is the identification of Cohesin as an essential component of chromatin looping based regulation for the *β-globin* and *Igf2/H19* loci, (Parelho et al, 2008; Wendt et al, 2008), the *IFNg locus* (Hadjur et al, 2009) and the *IgH* region (Degner et al, 2011).

Although most Cohesin sites overlap with CTCF, a significant proportion of

sites are independent of each other, implying CTCF-independent functions of Cohesin as well as Cohesin-independent CTCF roles. A CTCF-independent role for Cohesin in gene regulation has been shown in multiple cell types (Schmidt et al, 2010). Here, Cohesin was functionally co-localized with ER-α in human MCF-7 breast cancer cells and with liver-specific transcription factors in hepatic cells (Figure 1.5). An integrative analysis of Cohesin occupation in MCF-7 cells with chromosomal interaction data from ChIA-PET demonstrated that Cohesin is enriched in ER-α bound regions, which were involved in intra-chromosomal loops (Fullwood et al, 2009). Another study shows that Cohesin-subunit Scc1 co-localizes with pluripotency-related factors, including KLF4, OCT4, SOX2, ESSRB, and NANOG, also in a CTCF-independent manner. This supported the idea that cohesin plays a role in ES cell identity (Nitzsche et al, 2011). Direct regulation of the oncogenic *c-Myc* gene by Cohesin has been reported in human cells and model organisms (Kawauchi et al, 2009; Liu et al, 2009; Rhodes et al, 2010; Rubio et al, 2008; Schaaf et al, 2009; Stedman et al, 2008). Expression profiling of mutant Cohesin cell lines identified several hundreds of deregulated genes. Many of those Cohesin sites were enriched in the promoters of the deregulated genes, suggesting direct regulation by Cohesin (Liu et al, 2009). Further, significant overlap has been shown between deregulated genes in Cohesin mutant cell lines and deregulated genes in mutants of Cohesin loading factor NIPBL (Scc2) (Kawauchi et al, 2009; Liu et al, 2009). Accordingly, depletion of Cohesin and NIPBL resulted in similar deregulation of Cohesin target genes (Kagey et al, 2010). Interestingly, these mutants did not reveal any significant defects in chromosome segregation. It appears likely that NIPBL affects gene

expression by regulating Cohesin loading at regulatory sites.

A recent study on functional Cohesin-mediated chromatin looping in gene regulation showed a direct interaction of Cohesin with the mediator complex, an important transcriptional co-activator. Here, Cohesin-loading factor NIPBL appears to interact with the Cohesin-mediator complex loading Cohesin onto promoters. This results in DNA looping between enhancers and promoters of key pluripotency genes such as OCT4 and NANOG (Kagey et al, 2010). In another example, TAF3, a core promoter factor, localized to a subset of CTCF- and Cohesin occupied sites in mouse ES cells to regulate genes by long-range chromatin looping. This mechanism appears indispensible for endoderm lineage differentiation and prevention of premature differentiation of neuroectoderm and mesoderm in mouse ES cells (Liu et al, 2011) (Figure 1.5). These examples suggest that chromatin looping often involves Cohesin. However, the location of the loop and the positioning of Cohesin depends on other factors, such as mediator or CTCF. As a final example of CTCF-independent involvement of Cohesin in gene regulation, the discovery that Cohesin selectively binds and regulates genes with paused Pol II should be emphasized (Fay et al, 2011). Cohesin often localises to genes, at positions of Pol II pausing just downstream of the TSS. Here, it appears that Cohesin does not inhibit binding of Pol II to promoters, neither does it physically block transcription elongation. However, at repressed genes, Cohesin hinders transition of paused Pol II into the elongation phase, at a step distinct from those controlled by Spt5 and the negative elongation factor (NELF) (Fay et al, 2011).

1.3. Human macrophages and TNFα expression regulation

The bacteriologist Ilya Mechnikov first described macrophages over a century ago. Macrophages are found in every organ in the body but their morphology and phenotypical appearance differs depending on the organ they reside in and their interactions with other cell types and matrix. Macrophages originate from haemopoietic progenitors (Geissmann et al, 2010). These differentiate either directly or via circulating monocytes into different subpopulations of tissue macrophages and myeloid dendritic cells (Steinman et al, 2010) (see Figure 1.6).

Monocytes originate from hematopoietic stem cells and differentiate sequentially (Auffray et al, 2009b; Geissmann et al, 2010) in a process termed monopoiesis. This is tightly regulated by microenvironmental signals, which adapt gene expression in the developing cells, often followed by irreversible changes to the cell phenotype and function. Initially monopoiesis passes through distinct proliferative progenitor stages. Hematopoietic stem cells produce common myeloid progenitors and the following stages are granulocyte/macrophage progenitors (Akashi et al, 2000) and then macrophage dendritic cell progenitors (Auffray et al, 2009a; Fogg et al, 2006), which are the common precursors of monocytes, macrophages and dendritic cells. Monocytes are under homeostatic conditions mainly located to bone marrow, where they arise from macrophage dendritic cell progenitors. They are also found in blood and spleen (Auffray et al, 2009a; Auffray et al, 2009b; Fogg et al, 2006; Swirski et al, 2009).

Bone-marrow derived monocytes no longer proliferate, but functionally specialize and can be distinguished into ‘classical’ and non-classical’ monocytes (Auffray et al, 2007; Geissmann et al, 2010; Ziegler-Heitbrock et al, 2010) (see Figure 1.6). Chemotactic signals mobilize monocytes from bone marrow and spleen to migrate to their target tissue and guide their further differentiation (reviewed recently in (Shi et al, 2011)). Due to the distinct expression of chemokine receptors, the different subsets of monocytes show different migration properties (Geissmann et al, 2003). Monocytes are recruited to sites of inflammation by chemokines that are expressed by many nucleated cell types due to pro-inflammatory or microbial stimuli (Martin et al, 2008; Serbina et al, 2006; Shi et al, 2011; Shi et al, 2011; Tacke et al, 2007). Both, macrophages and monocytes migrate through blood, bone marrow, lymphoid and nonhaematopoietic tissues. Once recruited to tissues monocytes differentiate under the control of tissue specific factors into resident macrophages (Gordon et al, 2005; Mosser et al, 2008) (see Figure 1.6). This differentiation process involves phenotypic, functional and morphological adaptation, which leads to a large spectrum of diverse macrophage populations within the body (Gordon et al, 2005; Mosser et al, 2008). In some tissues the major fraction of macrophages do not derive from monocytes but rather from yolk sac progenitors that constitute a different line of macrophages (Schulz et al, 2012).

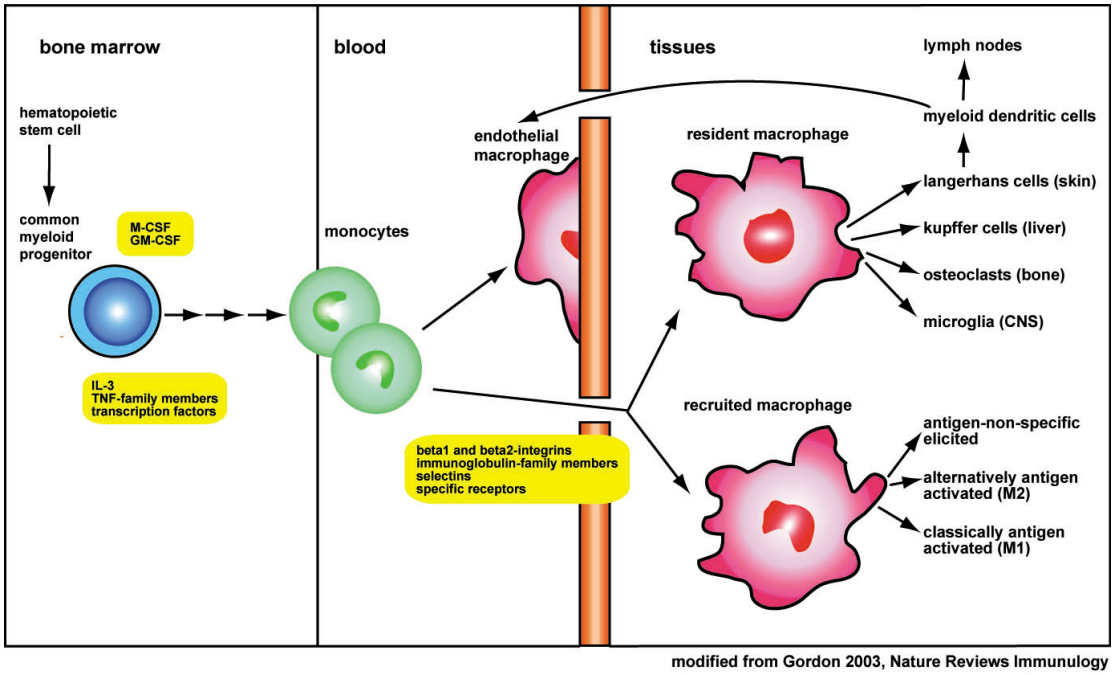


Figure 1.6. **Macrophage differentiantion, distribution and activation.** Lineage-determining cytokines such as macrophage colony-stimulating factor (M-CSF) and granulocyte-macrophage colony-stimulating factor (GM-CSF) and the interactions with stroma in haematopoietic organs induce growth and differentiation of macrophages from common myeloid progenitors, which descend from hematopoietic stem cells. Macrophage determination is known to involve interleukin-3 (IL3), tumour-necrosis factor (TNF)-family proteins, TNF-receptor-related moleculas and key transcription factors. Monocytes are mobilized from bone marrow to migrate to their target tissue while further differentiating. They are distributed via the blood stream and enter all different kinds of tissue where they adapt to their local microenvironment such as kupffer cells in the kidney and alveolar macrophages in the lung. The tissue specific phenotype of macrophages is influenced by surface and secretory products of neighbouring cells and the extracellular matrix. Cell migration from the blood is controlled by adhesion molecules such as integrins (beta1, beta2 and others), immunoglobulin-family members, selectins and specific receptors. They also control the transport through endothelia, the interstitium and epithelia as well as TRAP endothelium resistant macrophages. Monocyte-derived macrophages may re-enter the blood stream and differentiate into myeloid dendritic cells upon local stimuli such as phagocytosis. In response to inflammatory and immune stimuli tissue-resident macrophages undergo local activation. Also, monocytes and precursors from bone-marrow pools are recruited and lead to the accumulation of tissue macrophages with a distinct phenotype and increased turn-over rate. As a non-specific response to a foreign body or sterile inflammatory substance, they are classified as 'elicited'. In an antigen-specific immune response the recruited macrophages are either 'classically activated' or 'alternatively activated' and almost indistinguishable from resident macrophages as they also adapt to the particular microenvironment. Adapted from (Gordon, 2003), Nature Reviews Immunology

Macrophages that reside in organs permanently, such as kupffer cells in liver, alveolar macrophages in lung and microglia in the central nervous system, adapt to their microenvironment with enormous plasticity in phenotype and function (Krausgruber et al, 2011; Stout et al, 2005). They are involved in the removal of apoptotic cells, produce growth factors, induce angiogenesis and

serve as guardians of injury and infection (Pollard, 2009; Schafer et al, 2008). Macrophages are needed upon inflammation for pathogen clearance and during the following phase of tissue repair and wound healing in order to restore homeostasis (Schafer et al, 2008). In order to recognize foreign, normal and pathological cells as well as host-derived products, macrophages express a wide range of receptors. They contribute to tissue development and remodeling in developing or healing tissue and host defense in innate and adaptive immunity through endocytosis, phagocytosis and secretion of cytokines, growth factors and metabolites (Pollard, 2009; Schafer et al, 2008). Immunogenic signals such as inflammatory cytokines and toll-like receptor signals polarize the macrophages towards immunity and pathogen defense processes. Tissue macrophages are activated in response to inflammatory and immune stimuli, either as an antigen-non-specific response to a foreign object or sterile inflammatory substance or ‘classically activated’ or ‘alternatively activated’ via an antigen-specific immune response. Considering the polarization status of macrophages, a simplified categorization of tissue macrophages from mainly pro-inflammatory M1-macrophages (classically activated) up to mainly anti-inflammatory M2-macrophages (alternatively activated) was set up (Biswas et al, 2010; Gordon et al, 2010; Sica et al, 2008). Newly recruited, activated macrophages will be difficult to distinguish from previously resident macrophages as they adapt to the particular microenvironment. The M1/M2 classification oversimplifies the functional diversity of macrophages (Gordon et al, 2010; Mosser et al, 2008) and alternative classifications are based on macrophage function and their contribution to tissue remodeling and immune responses to different stimuli

(Mosser et al, 2008). Macrophages also play an important role in cancer. Tumor-associated macrophages can affect almost any stage of tumor development and progression via a wide range of anti-tumor functions such as production of cytotoxic factors, phagocytotic removal of cancer and metastatic cells and participation in cancer immuno-editing (O'Sullivan et al, 2012; Qian et al, 2010). They also have a wide range of pro-tumor functions via production of high levels of inflammatory mediators that may result in neoplastic transformation, and/or via production of components that support tumor growth, angiogenesis, immune suppression, tissue remodeling and metastatic spread of cancer cells (Mantovani et al, 2010; Sica et al, 2007; Sica et al, 2008; Tan et al, 2012). For a recent review that summarizes the current view on tumor-associated macrophages, especially their pro-tumor functions and how that knowledge can be applied in anti-tumor therapy see Richards et al. (Richards et al, 2012).

Tissue macrophages are terminally differentiated but can replicate locally and their turnover rate depends on the stimulus and the tissue environment. Macrophages are differentiated via lineage-determining cytokines (macrophage colony-stimulating factor (M-CSF), granulocyte-macrophage colony-stimulating factor GM-CSF)) and via contact with stroma in haematopoietic organs. Cytokines such as interleukin-3, tumor-necrosis factor (TNF)-family proteins as well as TNF-receptor related proteins and key transcription factors are necessary for macrophage determination. Other cytokines such as transforming growth factor-β, chemokines and growth factors influence gene expression in macrophages.

1.3.1 Tumor necrosis factor (TNF)

The cytokine TNF was originally known as tumor necrosis factor-α or cachectin. *TNF* plays important roles in the innate and adaptive immune and inflammatory response and also contributes to the normal function of lymphocytes, monocytes, macrophages, neutrophils, and dendritic cells (Aggarwal, 2000; Grivennikov et al, 2006). Multiple extracellular stimuli, including calcium signaling, mitogens, radiation, engagement of antigen receptor on T and B cells and infection by certain viruses and bacteria activate TNF expression (Aggarwal, 1995) and (Falvo et al, 2010). In T cells *TNF* is one of the first genes to be expressed upon cellular activation (Goldfeld et al, 1993). *TNF* expression is tightly regulated by specific stimuli to maintain cellular homeostasis and normal physiology in humans. Thus dysregulated TNF levels are linked to multiple diseases such as asthma, rheumatoid arthritis, cardiovascular diseases, Crohn's disease, type II diabetes, multiple sclerosis and several forms of cancer (Aggarwal et al, 2006; Sethi et al, 2008). In addition, dysregulation of *TNF* expression is linked to different levels of susceptibility to several major infectious diseases like tuberculosis and cerebral malaria (Jacobs et al, 2007; Lou et al, 2001). Therefore, studying *TNF* gene regulation involves studying cell type and stimulus specific gene regulation as well as helping to understand the direct translational implications on a variety of human diseases (Falvo et al, 2010).

Protein complexes termed 'enhanceosomes' are known to regulate *TNF* gene transcription (Barthel et al, 2003; Falvo et al, 2000b; Tsai et al, 2000; Tsytsykova et al, 2002) and consist of a set of transcription factors and co-activators that form higher-order structures with enhancer and promoter

regions of the gene and together drive transcription (Carey, 1998; Merika et al, 2001). The *TNF* enhanceosome assembly is cell type and stimulus specific with distinct transcription factors and coactivators (Barthel et al, 2003; Falvo et al, 2000a; Falvo et al, 2000b; Tsai et al, 2000; Tsytsykova et al, 2002).

The major family of transcription factors that is involved in the control of the immune response comprises the NF-κB proteins that activate transcription via a specific binding DNA motif (Hayden et al, 2008; Perkins, 2007; Vallabhapurapu et al, 2009). At first, *TNF* was described as one of the classical NF-κB-dependent pro-inflammatory cytokines. This was based on the murine TNF gene that was found to possess four DNA motifs in its promoter region that resemble the NF-κB-binding site (Collart et al, 1990; Shakhov et al, 1990). Also, the murine specific κB site has the highest affinity for a factor with NF-κB properties. It is located in the distal site of the murine *TNF* promoter and confers LPS inducibility upon a heterologous promoter but needs two or three copies (Collart et al, 1990; Drouet et al, 1991; Shakhov et al, 1990). Despite the discovery of three sites resembling NF-κB binding sites in the human *TNF* promoter region (Goldfeld et al, 1990) the main murine NF-κB-binding site (Collart et al, 1990; Drouet et al, 1991; Kuprash et al, 1999; Shakhov et al, 1990) is not conserved in the human TNF promoter (Leung et al, 2000). These NF-κB-binding sites were suggested to play a role in lipopolysaccharide (LPS)-induced murine TNF gene transcription (Collart et al, 1990; Drouet et al, 1991; Shakhov et al, 1990) but at least for human *TNF* gene transcription of the three NF-κB-like motifs (κ1-κ3) appear not to play a role as LPS- or virus-inducible enhancers (Falvo et al, 2000b; Goldfeld et al, 1990; Tsai et al, 2000; Yao et al, 1997). Although additional NF-κB-like sites

were found in the human TNF promoter, none could be convincingly linked to NF-κB-dependent LPS-mediated induction of TNF gene expression (Goldfeld et al, 1990; Kuprash et al, 1999; Udalova et al, 1998). Taken together, these data suggest that there is no conserved regulatory role for these NF-κB-binding sites in murine or human TNF gene transcription. Nonetheless, the proximal 200bp of the human TNF promoter, even when lacking two of the NF-κB-like motifs (κ1 and κ2) are sufficient for the *TNF* gene expression induction by LPS in several monocytic cell lines (Goldfeld et al, 1990; Kuprash et al, 1999; Trede et al, 1995; Tsai et al, 2000; Udalova et al, 1998; Yao et al, 1997). This minimal promoter region, located upstream of the mRNA cap site is highly conserved in mammals. It is also sufficient for activation of *TNF* expression and formation of specific enhancer complexes following other stimuli in various cell types (Barthel et al, 2003; Brinkman et al, 1999; Esensten et al, 2005; Falvo et al, 2000a; Goldfeld et al, 1990; Goldfeld et al, 1993; Leitman et al, 1992; Tsai et al, 2000; Tsai et al, 1996a; Tsai et al, 1996b; Tsytsykova et al, 2000; Tsytsykova et al, 2002).

TNF gene transcription requires specific spacing of the DNA motifs in its promoter region (Barthel et al, 2003; Tsytsykova et al, 2002) to fulfill the specific architectural requirements for formation of the multifactoral enhanceosome (Carey, 1998; Merika et al, 2001). Different factors can recognize the same *TNF* activator binding motifs, depending on the cell type, stimulus and the concentrations of the different factors in the nucleus (Barthel et al, 2003; Falvo et al, 2000a; Falvo et al, 2000b; Tsai et al, 2000; Tsai et al, 1996a; Tsai et al, 1996b), giving the *TNF* promoter a remarkable degree of

flexibility and allowing it to respond to a variety of stimuli in a specific manner using a short *cis*-regulatory region (Falvo et al, 2010).

A more recent study suggests that NF-κB does not regulate *TNF* gene expression through the 200bp promoter region, but rather plays a role in maintenance of TNF mRNA levels after the induction of the *TNF* gene transcription by viruses in T cells or by LPS in monocyte/macrophage type of cells (Tsytsykova et al, 2007). In addition, the interaction of NF-κB with the distal *TNF* promoter sites was suggested to play a role in LPS tolerance (Buckley et al, 2006).

A plethora of other nuclear factors are involved in activated T cells known to regulate *TNF* gene transcription. Equally, a large number of other transcription factors are involved in the activation of *TNF* gene transcription. The cell type and stimulus-specific control of the TNF gene at the transcriptional level and the known (key) components of the regulatory machineries are summarized in the review by Falvo et al. (Falvo et al, 2010).

A recent study provides indications for TNF locus regulation by CTCF in T-cells. The TNF locus comprises the genes Lymphotoxin A and B (*LTA* and *LTB*) and *TNFα* within the MHC class III region. In the course of that study, 3C analysis in Jurkat cells revealed that CTCF-mediated DNA looping occurs between the CTCF site and the *LTB* promoter, *LTA* promoter and most interestingly, the TNFα 3`UTR (Wicks et al, 2011).

1.4. Aims of thesis

A detailed study from this laboratory has shown that in *S. pombe*, Cohesin mediates transcription termination of convergently transcribed genes in a cell cycle dependent manner (Gullerova et al, 2008). In mammalian cells, Cohesin has been shown to be recruited by CTCF to canonical CTCF-binding sites for gene expression regulation functions (Parelho et al, 2008; Rubio et al, 2008; Stedman et al, 2008; Wendt et al, 2008). Thus, the purpose of my thesis has been to investigate whether Cohesin, recruited by CTCF, could be involved in mammalian gene transcription termination or other related processes that regulate gene transcription.

The first aim of my project was to investigate the involvement of CTCF and Cohesin in the transcription termination process in human cells, as presented in Chapter 3. In Chapter 4, I describe the impact of CTCF-recruitment on the 3` end of a *β-globin* gene construct lacking functional termination sequences. These results implicate CTCF-mediated effects on Pol II kinetics and pre-mRNA processing. I next focused my studies on an endogenous gene by studying the *TNFα* gene, expressed in human primary cells. *TNFα* provides a canonical CTCF-binding motif in closer proximity to the gene`s 3` end and has been shown to be part of a regulatory network mediated by chromatin loops (Wicks et al, 2011). To ensure best possible physiological conditions for studies of gene transcription regulation mechanisms, I addressed aspects of CTCF- and Cohesin-mediated transcription regulation in LPS-induced human monocytes and differentiated M1 and M2 macrophages. The results of these investigations are presented in Chapter 5.

CHAPTER 2

MATERIALS AND METHODS

2.1 Reagents

Most reagents used were obtained either from Sigma-Aldrich or from VWR for BDH and were of 'AnalaR' grade. Special reagents and those obtained from other sources are listed below:

[a^{32}P]rUTP 3000Ci/mmol (10mCi/ml)	GE healthcare
5x Transcription Buffer	Promega
Acrylamide (40%, acrylamide:bis 19:1)	Severn Biotech
Agar Bacteriological	Oxoid
Agarose	Melford
Blasticidin S HCl	Invitrogen
Bradford reagent	Biorad
Cell Line Nucleofector KitV for THP-1 cells	Lonza
Deoxyribonucleotide triphosphates (dNTPs)	Invitrogen
Dithiothreitol (DTT)	Invitrogen
Dulbecco's Modified Eagle Medium	PAA
Dulbecco's PBS	PAA
EDTA-free Proteinase Inhibitor tablets	Sigma-Aldrich
Ethylenediaminetetraacetic acid disodium salt	Sigma-Aldrich
ECL Western Blotting Detection Reagents	GE healthcare
Flp-In T-REX Core Kit	Invitrogen
Foetal Calf Serum	Gibco BRL
Glycogen	Roche
GM-CSF	Reprotech
HEPES	Melford

Hybond-C extra protein transfer membrane	GE healthcare
Hybond-NX nitrocellulose	GE healthcare
Hygromycin B	Invitrogen
Lipofectamine 2000	Invitrogen
Lipopolysaccharides from *Salmonella enteriditis*	Sigma-Aldrich
2-Mercaptoethanol	Gibco
MicroSpin G25 Columns	GE healthcare
Maxi-prep kit	Qiagen
M-CSF	Reprotech
N-Lauroylsarcosine sodium salt	Sigma-Aldrich
Non Essential Amino Acids	PAA
OptiMEM	Gibco BRL
Penicillin/Streptomycin	PAA
Phorbol 12-Myristate 13-Acetate (PMA)	Sigma
Ribonucleotide triphosphates (rNTPs)	Roche
RPMI1640	GIBCO
Streptavidin coated magnetic beads	Promega
Trizol	Invitrogen
RNase-free tRNA	Roche
TransIT -293 Transfection Reagent	Mirus
Trypsin	Gibco BRL
Zeocine	Invitrogen

2.2 Enzymes

Restriction enzymes were obtained from either Roche or New England Biolabs (NEB). Other enzymes were obtained from the following suppliers:

Shrimp Alkaline Phosphatase	Promega
T7 RNA Polymerase	Promega
Pfu Turbo DNA Polymerase	Stratagene
RNase out	Invitrogen
Proteinase K	Roche
RNase A	Roche
RNase H	NEB
T3 RNA Polymerase	Ambion
T4 DNA Ligase	Roche
T4 Polynucleotide Kinase	NEB
SuperScript III reverse transcriptase	Invitrogen
DNase I- RNase free	Roche
GoTaq DNA Polymerase	Promega
Phusion Taq DNA Polymerase	Finnzymes
Sensimix (Real Time PCR reaction mix)	Bioline

2.3 Antibodies

Rabbit-anti-CTCF (07-729)	Millipore
Rabbit-anti-Smc3 (ab9263)	Abcam
Rabbit-anti-Rad21 (ab992)	Abcam
Rabbit-anti-Cleavage stimulation factor2 (ab72297)	Abcam

Rabbit-anti-PolII (N-20) Santa Cruz Biotech

2.4 Nucleic Acid markers

DNA molecular weight markers (1Kb Plus DNA Ladder) for agarose gel electrophoresis were obtained from Invitrogen. RNA molecular weight markers (RNA Millenium size markers) were obtained from Ambion.

2.5 Tissue Culture solutions

DMEM or RPMI1640 were supplemented with 10% foetal calf serum, 1% Non Essential Amino Acids (NEAA), 100units/ml penicillin, and 100units/ml streptomycin.

Medium prepared with RPMI1640 was additionally supplemented with 2mM glutamine.

PBS/EDTA/Trypsin (PET) contained 0.5mM EDTA and 0.05% trypsin in PBS.

2.6 Solutions and Buffers

Organic solutions:

Phenol/Chloroform: Phenol (pH4 or 8 for RNA or DNA isolation respectively) and chloroform were mixed in equal volumes.

Electrophoresis buffers:

50x E buffer (for agarose gels):

Sodium Acetate 250mM

EDTA	50mM
Tris	2M
Glacial acetic acid	91ml/L

RNA loading buffer:

Formamide	80% (v/v)
EDTA	10mM
Bromophenol blue	0.25% (v/v)
Xylene cyanol	0.25% (v/v)

10x Orange G loading dye (for agarose gels):

Ficoll	25% (w/v)
Orange G	2.5mg/ml
EDTA	10mM

Solutions for DNA analysis:

TE

Tris HCl, pH8	10mM
EDTA	1mM

ChIP on HEK293 cells, THP-1cells and PBMC

Cell washing buffer

PBS + 1 tablet of EDTA-free Proteinase inhibitor inhibitors

Cell lysis buffer

PIPES pH8.0	5mM

KCl	85mM
Nonidet P-40 (NP40)	0.5%
PMSF	1mM
Pepstatin A	1µg/ml
Leupeptin	1µg/ml

Nuclei lysis buffer

SDS	1%
EDTA	10mM
Tris-HCl pH8.0	50mM
PMSF	0.5 mM
Pepstatin A	0.8µg/ml
Leupeptin	1µg/ml

Immunoprecipitation buffer

SDS	0.01%
Triton X-100	1.1%
EDTA	1.2mM
Tris-HCl pH8.1	16.7mM
NaCl	167mM
PMSF	0.5 mM
Pepstatin A	0.8µg/ml
Leupeptin	1µg/ml

Washing solutions

A: 0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl pH8.0, 150mM NaCl

B: 0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl pH8.0, 500mM NaCl

C: 0.25M LiCl, 1% NP40, 1% sodium deoxycholate, 1mM EDTA, 10mM Tris-HCl pH8.0

D: 1mM EDTA, 10mM Tris-HCl pH8.0

Elution buffer

SDS	1%
$NaHCO_3$	0.1M

Reverse cross-link

NaCl	0.3M
RNase A	0.3µg/ml
Proteinase K	20µg/ml

ChIP on human monocytes, M1 and M2 macrophages

Lysis buffer 1 (LB1)

Hepes-KOH pH7.5	50mM
NaCl	140mM
EDTA	1mM
Glycerol	10%
NP-40	0.5%

Triton X-100	0.25%

Lysis buffer 2 (LB2)

Tris-HCl pH8.0	10mM
NaCl	200mM
EDTA	1mM
EGTA	0.5mM

Lysis buffer 3 (LB3)

Tris-HCl pH8.0	10mM
NaCl	100mM
EDTA	1mM
EGTA	0.5mM
Na-Deoxycholate	0.1%
Na-Lauroylsarcosine	0.5%

Washing buffer

Hepes-KOH pH7.6	50mM
LiCl	500mM
EDTA	1mM
NP-40	1%
Na-deoxycholate	0.7%

Elution buffer

1xTE containing 2% SDS

Solutions for RNA analysis:

RNA solutions were either treated with 0.02% DEPC or made up with DEPC-treated water.

DNase buffer:

Tris HCl, pH7.5	10mM
Magnesium Chloride	10mM

20x SSC:

Sodium chloride	3M
Sodium citrate	300mM

This solution was adjusted to pH7 by addition of hydrochloric acid.

10x MOPS:

Sodium acetate	0.05M
EDTA	0.01M

This solution was adjusted to pH 7.0 by addition of hydrochloric acid.

Northern Hybridisation solution:

Formamide	50%
SDS	0.07%
Sodium chloride	5mM
Sodium phosphate	2.5mM

Northern wash solution 1:

SSC	2x
SDS	0.1%

Northern wash solution 2:

SSC	0.5x
SDS	0.1%

Solutions for NRO:

HLB/NP40:

Tris HCl, pH7.5	10mM
Sodium chloride	10mM
Magnesium chloride	2.5mM
NP40	0.5% (v/v)

HLB/NP40/sucrose:

10% sucrose in HLB/NP40

2x Transcription buffer:

Tris HCl, pH7.9	40mM
Potassium chloride	300mM
Magnesium chloride	10mM
Glycerol	40% (v/v)
Prior to use add:	
DTT	2mM

20x SSPE:

Sodium Chloride	3.6M
Sodium dihydrogen phosphate	200mM
EDTA	20mM

This solution was adjusted to pH7.4 by addition of sodium hydroxide.

Denaturing solution:

Sodium chloride	1.5M
Sodium hydroxide	0.5M

Neutralising solution:

Tris HCl, pH7.2	0.5M
Sodium chloride	1.5M
EDTA	1mM

100x Denhardts solution:

Ficoll	2%
Polyvinylpyrrolidone	2%
Bovine serum albumin	2%

Pre-hybridisation solution:

SSPE	6x
Formamide	50% (v/v)
SDS	0.1%
Denhardts	5x

tRNA	100mg/ml

Hybridisation solution:

SSPE	7x
Formamide	58% (v/v)
SDS	0.12%
Denhardts	5.8x
tRNA	100mg/ml

Solutions for protein analysis:

RIPA buffer:

Sodium chloride	15mM
Triton X-100	1%
Sodium deoxycholate	0.5%
SDS	0.1%
Glycerol	10%

Prior to use add:

PMSF	3mM
b-mercaptoethanol	5%

2x Laemmli:

Tris HCl, pH6.8	62mM
Glycerol	25%
SDS	2%

2-mercaptoethanol	10%
Bromophenol blue	0.25% (v/v)

PBS & Tween:

0.1% Tween in PBS

2.7 Plasmids and probes

DNA constructs were made by restriction digestion or long-range PCR using Pfu Turbo Polymerase. Details of each construct are given under the chapter in which they first appear. All plasmids were controlled for via DNA sequencing analysis.

Chapters 3 and 4

Plasmid construction

pβTERM, pβΔTERM and pTat were a kind gift from Dr. M Dye. The PCR-amplified core CTCF binding motif (sequence in Chapter 3, Figure 3.1) and the chicken β-globin FII site (Chapter 3, Figure 3.4) were inserted into pβΔTERM by blunt end ligation, giving rise to the plasmids pβΔTERM-CTCF2x/5x and pβΔTERM-FII1x/2x/4x, respectively.

pcDNA5/FRT/βTERM, pcDNA5/FRT/βΔTERM and pcDNA5/FRT/βΔTERM-FII4x were made by blunt end ligation of PCR products βTERM, βΔTERM and βΔTERM-FII4x ligated into the vector pcDNA5/FRT/T. Prior to ligation pcDNA5/FRT/T was linearized by PCR using primer pair pc5f/pc5r. The PCR

products βTERM, βΔTERM and βΔTERM-FII4x were generated from pβTERM, pβΔTERM and pβΔTERM-FII4x, respectively, using primer pair βallf/βallr. pcDNA5/FRT/βTERMmut was a kind gift from Dr. E White.

M13 probes

The β-globin gene specific ssDNA probes have been described in Ashe et al. (1997). Probes P and U3 have been described in Ashfield et al. (1994). Probe A has been described in Dye and Proudfoot (1999).

Northern Probes

Template for β-globin Exon 2 and β-actin Exon 4 probes was amplified from pβTERM plasmid or HeLa genomic DNA using the primers E2FT7 and E2RT3 or βACTEx4f and βACTEx4rT3, respectively. The probes were *in vitro* transcribed with T3 RNA Polymerase using the T7/T3 MAXI kit (Ambion) (Camblong et al., 2007).

RNA interference

For CTCF knock-down, the siGENOME SMARTpool siRNA Human CTCF was used. The siRNA pool contains the following sequences: GAAGAUGCCUGCCACUUAC, AAGAAUGAGAAGCGCUUUA, AAACAUACCGAGAACGAAA, GAAAGUGGUUGGUAAUAUG;

Chapter 5

RNA interference

For CTCF knock-down, the siGENOME SMARTpool siRNA Human CTCF was used. The CTCF siRNA pool contains the following sequences: GAAGAUGCCUGCCACUUAC, AAGAAUGAGAAGCGCUUUA, AAACAUACCGAGAACGAAA, GAAAGUGGUUGGUAAUAUG;

For Rad21 knock-down, the siGENOME SMARTpool siRNA Human Rad21 was used. The siRNA pool contains the following sequences: GGAAGAAGCAUUUGCAUUG, GAACAGAGCACCAGCAAUC, GAGCCCAACUUAGUGAUUA, GGGAGUAGUUCGAAUCUAU;

For SMC3 knock-down, the siGENOME SMARTpool siRNA Human SMC3 was used. The siRNA pool contains the following sequences: GAAAGCAUCUCCUUAAUGA, GAACGGAUCUUUAUGCAAA, GGAAUUAGGGUGUCAUUUA, GAAGAGAGAUUACAUACUC,

2.8 Table of oligonucleotides

The oligonucleotides are listed under the chapter in which they first appear.

Chapters 3 and 4

Af	5´GCATCCATCCGGAGTACTTCAAGAAC
Ar	5´CCTAGTTAGCCAGAGAGCTCCCAGG
Bf	5´CTGGCTAACTAGGGAACCCACTGC
Br	5´CAACTTCATCCACGTTCACCTTGC

Cf	5′CCAGAGGTTCTTTGAGTCCTTTGG
Cr	GGTTGTCCAGGTGAGCCAGG
Df	5′CCACAAGTATCACTAAGCTCGC
Dr	5′ATCCAGATGCTCAAGGCCC
Ef	5′CTGCAAACAGCTAATGCACA
Er	5′CTTGAATCCTTTTCTGAGGGATG
Ff	5′CAGGAAACTATTACTCAAAGGGTA
Fr	5′GTTTTCCCAGTCACGACGTT
5`UTRf	5′CCAGATCTGAGCCTGGGA
5`UTRr	5′GCACCATGAGCTTTATTGAG
rtF	5′CAGGAAACTATTACTCAAAGGGTA
rtR	5′AGAAAATACCGCATCAGGCGCCATT
pc5f	5′CTGGGGATGCGGTGGGCTCTA
pc5r	5′GCGTATATCTGGCCCGTACATCGC
βallf	5′CCATGATTACGAATTCGAGCTCGGTA
βallr	5′AGAAAATACCGCATCAGGCGCCATT
E2FT7	5′TAATACGACTCACTATAGGGAGACTTGGACCCAGAGGTTCTTTG
E2RT3	5′AATTAACCCTCACTAAAGGGAGAATCCACGTGCAGCTTGTCACA
Int1f	5′GGAGACCAATAGAAACTGGGC
Ex2r	5′GAGCCAGGCCATCACTAAAG
Ex2f	5′CCTGGCTCACCTGGACAACCTCAAG
Ex3r	5′GAAATTGGACAGCAAGAAAGCGAG
CTCFf	5′ATTCCTCAGGAGGGTGACTG
CTCFr	5′CAAGTGGACACCCAAATCAC

GAPDHf	5´CCACCCAGAAGACTGTGGAT
GAPDHr	5´TTCTAGACGGCAGGTCAGGT

Chapter 5

Qt	5´CCAGTGAGCAGAGTGACGAGGACTCGAGCTCAAGCTTTTTTTTT TTTTTTTTT
QO	5´CCAGTGAGCAGAGTGACG
QI	5´GAGGACTCGAGCTCAAGC
TNF1	5´CGGAGCTGAACAATAGGCTGTTCC
TNF2	5´CCCTGGCCTCTGTGCCTTCTTTTG
TNFex4f	5´ CAGAGGGCCTGTACCTCATC
TNFex4r	5´ GGAAGACCCCTCCCAGATAG
LTBf	5´ GATGTTGACGTACACCCTCTC
LTBr	5´ GCCTCTATTACCTCTACTGTCTCG
Af	5´CAGCAAGGACAGCAGAGGAC
Ar	5´TCCCGGATCATGCTTTCAGT
Bf	5´GAGCACTGAAAGCATGATCC
Br	5´CCACGATCAGAAGGAGAAG
Cf	5´GGTGTCTGGCACACAGAAGA
Cr	5´CATCGAGGGAGTCACCCTTA
Df	5´ CAGAGGGCCTGTACCTCATC
Dr	5´ GGAAGACCCCTCCCAGATAG
Ef	5´ATATTCCCCATCCCCCAGGAAACA

Er	5´CTGCAACAGCCGGAAATCTCACC
Ff	5´CTTCCAGAGATTCGGGTGTC
Fr	5´ AGGTCAGGCCTTCTCTCACA
Gf	5´CCTGGATGAAGGCTCTTCTG
Gr	5´GAGTGATCCTCCTGCCTCAG
Hf	5´CCTCCTCCCCAGAACTTCCT
Hr	5´CAGCAAGCACCACTCAGAAA
If	5´GACCATCATCCCTAGGAGCA
Ir	5´ GTGCTGCCGATAAAGATGCT
Jf	5´GATGTTGACGTACACCCTCTC
Jr	5´GCCTCTATTACCTCTACTGTCTCG
Kf	5´ GCAAGATACAACTCTCCACCAG
Kr	5´ CACTTCTCTGGTGACCTTGTTG
If	5´CCACAGCAATGGGTAGGAGAATGT
Ir	5´GAGGTCCTGGAGGCTCTTTCACT
IIf	5´GGAAGCCAAGACTGAAACCAGCA
IIr	5´CCGGGAATTCACAGACCCCACT
IIIf	5´CAGCAAGGACAGCAGAGGAC
IIIr	5´TCCCGGATCATGCTTTCAGT
IVf	5´ CAGAGGGCCTGTACCTCATC
IVr	5´ GGAAGACCCCTCCCAGATAG
Vf	5´ATATTCCCCATCCCCCAGGAAACA
Vr	5´CTGCAACAGCCGGAAATCTCACC
c-Mycf	5´CTCCTGCTCCTGCCCCCACCTG

c-Mycr	5´GCTGCAAAGCGTCTTTCCCTCCG
GAPDHpf	5´ AGCTCAGGCCTCAAGACCTTGGGCT
GAPDHpr	5´ GGCTGACTGTCGAACAGGAGGAGC
GAPDHexf	5´CCACCCAGAAGACTGTGGAT
GAPDHexr	5´TTCTAGACGGCAGGTCAGGT
GAPDH3`f	5´GTTGCCATGTAGACCCCTTG
GAPDH3`r	5´CCCAGCAAGAATGTCTCACC
Smc3f	5´ CTTTGAATGTGAACCAGCTTTCTA
Smc3r	5´ GGCTGTATCCCTGACATCTAACTT
Rad21f	5´GGGAAATTATGCACCTGGTC
Rad21r	5´GCAACGCTATTAAAGTTCCAAA
CTCFf	5´ATTCCTCAGGAGGGTGACTG
CTCFr	5´CAAGTGGACACCCAAATCAC
IGRf	5´TGACCCAGTCTGTGGTATTTTG
IGRr	5´GGAATGAGGCCAGCACCTAC

2.9 DNA manipulations

Cloning

Vectors were linearized by long range PCR (Barnes, 1994; see below) with dephosphorylated primers. The DNA fragments to be used as inserts were also amplified by PCR with phosphorylated primers. In the majority of cases, the DNA vector and insert were gel purified before being ligated using T4 DNA ligase as per the manufacturer's instructions. The ligation mixtures were transformed into *E.coli* strain NM554. The bacteria were made competent

using the calcium chloride method and stored at -80°C (kind gift from J. Monks). Plasmid DNA was isolated and analysed by restriction digest or colony PCR. The results were confirmed via DNA sequencing.

Plasmid isolation

For small scale and large scale preparations, the Qiagen Mini prep kit or Qiagen maxi prep kit were used, respectively, following manufacturer`s instructions.

Long range PCR

Long range PCR (Barnes, 1994) was used to open up a plasmid and to delete or insert sequences within a plasmid. This was achieved by using primers, which flanked the region to be deleted or, the region where the plasmid was going to be opened up at. For insertions, the primers contained the entire sequence to be inserted. The Pfu polymerase was used according to the manufacturer's instructions, with 20ng of plasmid DNA as template. The cycling conditions were chosen according to the manufacture`s guidelines.

Quantitative PCR

Quantitative PCR (qPCR) was performed using the SensiMix No ROX Kit in a Corbett Rotorgene 3000 or 6000 machine. PCR measurements were performed in triplicates and quantitated using the Rotorgene 3000 software. The cycling conditions were: 95°C 10 minutes, 45x [95°C 10 seconds, 55-68°C 15 seconds, 72°C 25 seconds]. A SYBR Green melt curve was

performed simultaneously to confirm the quality of the PCR product.

Generation of HEK293 expression cell lines

The Flp-In T-Rex Core Kit was used to create HEK293 Flp-In T-Rex Expression Cell Lines (hereafter referred to as the `Expression Cell Lines`) containing stable integrants of the pTat-inducible HIV-LTR/β-globin constructs pcDNA5/FRT/βTERM, pcDNA5/FRT/βΔTERM and pcDNA5/FRT/βΔTERM-FII4x, following the manufacturer's instructions. The HEK293 Flp-In T-Rex Host Cell Line (hereafter referred to as the `Host Cell Line`; Invitrogen; kind gift from Torben Jensen) was used as the parental cell line for integration of these constructs and contains stable genomic integrants of two plasmids:

- pFRT/*lacZeo*: the Flp-In target site vector creates one Flp Recombination Target (FRT) site upon stable integration into the genome and it contains a Zeocin resistance marker. Host cell lines contain only a single integration of pFRT/*lacZeo* (as confirmed by Southern Blot – Invitrogen).
- pcDNA6/TR: this vector constitutively expresses the Tet repressor and contains a Blasticidine resistance marker.

The *S. cerevisiae*-derived DNA recombination system using the Flp recombinase and site-specific recombination (Craig, 1988; Sauer, 1994) is the principle of creating the Expression Cell Line containing the previously described β-globin constructs. The Host Cell Line was grown in D-MEM media supplemented with 10% FCS, 1% NEAA, 100units/ml penicillin/streptomycin, 100μg/ml Zeocin, and 10μg/ml Blasticidine S HCl at 37^{o}C with 5% CO_2. On day 1, 10^6 cells for each β-globin construct were seeded in 6cm^2 plates, grown for 24 hours and were then co-transfected with 9μg of pOG44 (plasmid

transiently expressing Flp recombinase to mediate the rare and specific recombination event between the FRT sites) and 1μg pcDNA5/FRT/βTERM, pcDNA5/FRT/βΔTERM or pcDNA5/FRT/βΔTERM-FII4x in D-MEM media supplemented with 10% FCS and 1% NEAA (antibiotic-free). Simultaneously, an untransfected control was processed. On day 2 post transfection, cells were trypsinized and transferred into T175 flasks in D-MEM media supplemented with 10% FCS and 1% NEAA. On day 3, media was changed to D-MEM media supplemented with 10% FCS, 1% NEAA, 100units/ml penicillin/streptomycin, 10μg/ml Blasticidine S HCl and 100μg/ml Hygromycin B. Cells were then cultured at 37^{o}C with 5% $CO_{2,}$ for approximately two weeks (with media changed every 3-4 days) until Hygromycin and Blasticidine resistant colonies appeared. Individual colonies were selected and grown.

2.10 Tissue Culture

Growth and Maintenance

HeLa cells were grown in D-MEM media supplemented with 10% FCS, 1% non-essential amino acids (NEAA) and 100units/ml penicillin and streptomycin at 37^{o}C with 5% CO_2. HEK293 cells were also cultured in D-MEM media supplemented with 10% FCS, 1% NEAA, without penicillin and streptomycin but with Hygromycin and Blasticidine. Cells were passaged when they had reached confluence. They were washed in PBS and treated with PET until they detached from the dish (a few minutes, usually). The cells were subsequently re-suspended in D-MEM media and divided as required.

THP-1 cells are suspension cells and were grown in RPMI1640 supplemented with 10% FCS and 2mM glutamine. They were passaged upon reaching a cell density of $8x10^5$ cells/ml.

Human PBMC were provided by Dr. F Martinez, University of Oxford.

Human monocytes and GM-CSF and M-CSF differentiated macrophages were provided by Dr. T Krausgruber, Kennedy Institute of Rheumatology, Imperial College, London. Enriched populations of human monocytes were obtained from the blood of healthy donors by elutriation as described (Krausgruber et al., 2010). M1 and M2 macrophages were obtained after 5 days of culture of human monocytes in RPMI-1640 medium supplemented with GM-CSF (50 ng/ml) or M-CSF (100 ng/ml; Peprotech). For TNFα ChIP and mRNA analysis, cells were stimulated with LPS (100 ng/ml).

Transient transfections

HeLa cells:

$1X10^6$ cells were grown on $10cm^2$ plates and transfected with 5-10μg of reporter plasmid DNA and/or 1μg of Tat plasmid plus 12μl of Lipofectamine 2000 in OptiMEM as per manufacturer's instructions. The cells were harvested 24h post transfection.

HEK293 Expression Cell Lines:

For each cell line, $1X10^6$ cells were transfected with the reverse transfection approach using TransIT293 Transfection solution, as per manufacturer`s instructions. Briefly, 15μl TransIT293 Transfection solution were incubated

with 1μg of Tat plasmid in Opti-MEM cells. Cells were washed in PBS, trypsinized to detach from the plates, re-suspended in warm, fresh D-MEM medium containing 10%FCS and 1%NEAA and transferred into 10cm^2 plates. The transfection-mix was then added drop-wise to the cells. Cells were harvested after 24-36h post transfection.

THP-1 monocytes:

For RNAi treatment, 1x10^6 cells each sample were transfected with 300nM siRNA using electroporation. Cells were transfected applying the Amaxa Cell Line Nucleofector Kit V for THP-1 cells (Lonza) and electroporated using the Nucleofector® device with the program setting U-001 for high cell viability, following the manufacturer`s instructions.

RNA interference

HEK293 cells

HEK293 cells were transfected using Lipofectamine RNAiMAX following the reverse transfection approach. Cells were passaged in complete growth media the day before transfection to reach a cell confluence of 40% on the following day. The transfection protocol includes the following steps:

Per sample, 400pmol RNAi duplex were diluted in 2ml of Opti-MEM Medium in a tube and gently mixed. 15μl Lipofectamine RNAiMAX per sample were added to the diluted siRNA, gently mixed and incubated for 10-20min at room temperature. Cells were detached from the plate and resuspended in each 8ml of complete growth media without antibiotics. Each cell sample was then transferred into a new 10cm^2 plate and 2ml of transfection mix were added

drop-wise into each plate. Cells were incubated at 37°C for 56-72h and then harvested for further analysis.

Human M2 macrophages

The siGENOME SMARTpool siRNA (Dharmacon) designed to target human Smc3, Rad21 and CTCF were used for siRNA-mediated knockdown. DharmaFECT 1 was used as the siRNA transfection reagent according to the manufacturer's instructions (Dharmacon).

After siRNA treatment, cells were stimulated with LPS (100ng/ml), followed by subjection to further analysis.

Cell cycle synchronization

G2-phase cycle synchronization of HeLa cells was performed using the cdk-inhibitor Roscovitine, in combination with a Thymidine-induced cell cycle block. The two following protocols were tested:

Protocol I:

Cells at 70 % confluence in full medium

↓

add Thymidine to a final concentration of 2 mM

↓ 24 h, 37°C

wash with warm medium and release for 4h at 37°C

↓

add Roscovitine to a final concentration of 25 μM

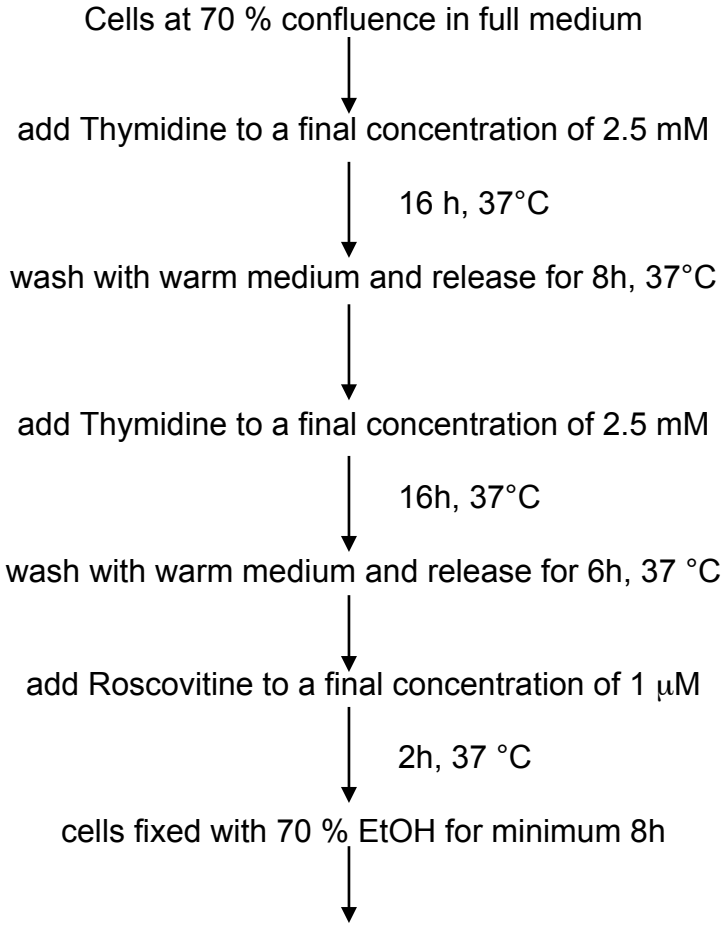
↓ 18h, 37°C

cells fixed with 70 % EtOH for minimum 8h

↓

Propidiumiodide-stain for FACS analysis

Protocol II:



Fluorecent Cell Sorting Analysis (FACS)

HeLa cells were grown in 10cm^2 dishes and transfected with the expression plasmids and pTat as described before. For FACS analysis, cells were washed with PBS, scraped of the plates and collected in tubes. After centrifugation at 500g for 5min at 4°C, the supernatant was removed and 2ml of 70% ethanol were added to each sample (pellet was not resuspended). Samples were incubated at 4°C for 30min, followed by centrifugation at 500g for 5min. The supernatant was discarded and the cell pellet washed with cold PBS, followed by another centrifugation step at 500g for 5min. The

supernatant was discarded again and the cells were resuspended in 1ml of a mixture of 0.1mg/ml Propidium Iodide in PBS+0.2mg/ml RNAseA. Cells were then incubated at 37°C for 30-40min in the dark, followed by a final centrifugation step at 500g for 5min and resuspended in 300µl of PBS. The DNA content of the fixed cell samples was then analysed using the FACScan system with the CellQuest software (Becton Dickinson, Franklin Lakes, NJ, USA).

2.11 Steady state RNA analysis

Mammalian total RNA isolation

Cells were washed with PBS and 1ml of Trizol reagent was added directly to the cells. Total RNA was then extracted following the manufacturer's instructions. The RNA pellet was resuspended in 50µl DNAse I buffer with 10 units of RNAse-free DNAse I and incubated at 37^{o}C for 30min and up to 60min, depending on the cell sample size used. DNase I was heat-inactivated at 75°C for 10min. DNase I was then removed by phenol/chloroform extraction and the RNA was ethanol precipitated. The RNA pellet was resuspended in DEPC-treated water, quantitated using a NanoDrop device, and stored at -20^{o}C.

Reverse-Transcription PCR

1µg of total RNA was reverse transcribed with either 50pmol of oligo dT or 5pmol of gene-specific primers using SuperScript III according to the manufacturer's instructions. For PCR amplification, 1/20th of the reverse

transcription reaction was used. The PCR conditions were chosen according to the manufacturer`s instructions for the SensiMix No ROX Kit.

Northern Blot Analysis

15-40µg of RNA were mixed with 4µl 10x MOPS, 7.2µl formaldehyde and 20µl formamide and heated at 65°C for 15 minutes. 4µg of RNA loading buffer was then added to the samples before being separated on 1% formaldehyde agarose gel. Separated RNA was then transferred from the gel to Hybond-N nitrocellulose membrane using the Vacuum transfer machine for 90min. Transferred RNA was cross-linked to the membrane in a UV-crosslinker (Stratalinker, Stratagene) with 700mJ x 100/cm^2 of energy. The membrane was then placed in hybridisation tubes containing 20ml of hybridisation solution and 50µg of denatured salmon sperm DNA was then added, followed by incubation from 2h to overnight at 62°C in a rotating hybridisation oven.

A radiolabelled RNA riboprobe was made by transcription of 0.1µg of PCR amplified template containing a T3 promoter. The probe was *in-vitro* transcribed in a reaction containing [α^{32}P]rUTP and using the T7/T3 maxiscript kit (Ambion) according to the manufacturer's instructions. The riboprobe was passed through a G25 column to separate labelled products from unincorporated nucleotides, denatured at 95°C for 30sec, and then added to the membrane containing fresh hybridisation buffer/salmon sperm DNA mixture. Hybridisation was conducted for 6h to overnight at 62°C in a rotating hybridisation oven. The blots were then washed for 5min at 62°C with washing solution 1, followed by 40min at 62°C in wash solution 2. Signals on

blots were visualized using autoradiography and quantitated using a PhosphoImager.

3`RACE

Rapid Amplification of cDNA 3` ends (3`RACE) is a PCR-based method for amplification of nucleic acid sequences from an mRNA template between a defined internal site and the 3´ end of the mRNA.

Total RNA was isolated from THP-1 cells and mRNAs were reverse transcribed into cDNA using an oligo dT-adapter primer (QT). TNFα cDNA was then specifically amplified by two rounds of PCR using gene-specific primers (TNF1 and TNF2), which anneal to regions in Exon 4 and adapter primers that target the poly(A) tail region (QO and QI, respectively). This captures the TNFα mRNA 3´ end sequence, which lies between the Exon and the poly(A) tail. The principle of 3`RACE is illustrated in Chapter 2, Figure 5.2B.

2.12 Nascent RNA analysis

Nuclear Run-On filter preparation

NRO filters were prepared using the slot blot apparatus from BIO-RAD. Three layers of Whatman paper were soaked in 5x SSPE buffer and placed in the slot blot underneath a pre-wet nitrocellulose hybridisation membrane. The blotting apparatus was assembled and each slot was washed out with 500µl of 5x SSPE. After this, 5µg of each of the M13 probes, diluted in 250µl of 5x SSPE was added to the appropriate slot. The slots were washed once more with 500µl of 5x SSPE. The filters were then placed in denaturing solution for

5min and then neutralising solution for 2min, before being air-dried. M13 probes were cross-linked to the filters by 2-minute exposure to UV light on a transilluminator. Filters were then placed in 10ml Nunc-tubes containing 2.5ml of prehybridisation solution and incubated overnight at 42°C in a rotating hybridisation oven.

Nuclear Run-On analysis of transfected HeLa nuclei

In the cold room, cells were rinsed in cold PBS, scraped into 10ml PBS, and pelleted by centrifugation at 500g for 5min. The cell pellet was resuspended in 4ml HLB/NP40, incubated on ice for 5min, and then underplayed with 1ml HLB/NP40/sucrose solution. The lysed cells were then centrifuged at 500g for 5min. The size of the nuclear pellet was estimated and resuspended in an equal volume of 2x transcription buffer before being transferred to a 1.5ml microcentifuge tube. 300mM rATP/rGTP/rCTP were added followed by 60mCi [α^{32}P]rUTP. The transcription was allowed to proceed for 15min at 30°C. The reaction was stopped by centrifugation at 15,800g for 30sec. The nuclear pellet was resuspended in 1ml Trizol reagent and RNA was isolated according to the manufacturer's instructions. The RNA was then resuspended in 60µl of DEPC-treated water, partially hydrolysed on ice for 5min by addition of 15µl of 1M sodium hydroxide, and finally, neutralised by addition of 30µl of 0.1M Tris/0.1M HCl. Prehybridisation solution was replaced with 650µl of hybridisation solution and the RNA mix and incubated at 42°C in a rotating hybridisation oven overnight. Filters were then washed in 1x SSPE/0.1% SDS on a shaking platform at room temperature for 20min. More stringent washes, at 65°C were carried out if necessary. The washing solution could also be

made more stringent by decreasing the SSPE down to 0.1x. Hybridisation signals were visualized by autoradiography and quantitated using a PhosphoImager.

2.13 DNA analysis

Chromatin Immuno-Precipitation (ChIP)

1. ChIP protocol applied for HEK293 cells, THP-1 cells and PBMC

Fixation of cells:

Cells at 70–80% confluence were cross-linked with 1% formaldehyde for 10 min at 37°C in slowly moving media to ensure equal distribution of the formaldehyde. Cross-linking is followed by quenching with 125mM glycine for 5min at 37°C.

Cells are then washed twice with cold washing solution (PBS+proteinase inhibitors), scraped off and collected by centrifugation at 500g for 5min at 4°C.

Cell- and nuclei lysis:

The cell pellet was resuspended in cell lysis buffer and incubated on ice for 10min, followed by collection of the nuclei by centrifugation at 550 x g for 5min at 4°C. The nuclei pellet was then lysed in nuclear lysis buffer for 10min on ice.

Sonication:

Sonication was performed on the Bioruptor UCD200.

Settings: 30 sec on/30 sec off, intensity level M, sonication time 10min for HeLa cells, and from 20 to 25min for HEK293 cells, respectively.

Sonication is followed by 10min centrifugation at 13,000 x g to remove cell debris. Immuno-precipitation buffer together with ProteinA beads were added

to the sonicated chromatin and incubated for 1h rotating at 4°C to pre-clear the chromatin before immune-preciptiation.

Chromatin immuno-precipitation:

After pre-clearing, beads were removed by centrifugation, 5min at 500 x g and the supernatant was divided into Eppendorf tubes.

Antibodies were added according to manufacturer's suggestions and incubated on the rotation wheel at 4°C over night.

Washes and elution:

On the next day, samples were washed with washing buffers 1 to 4, followed by elution in elution buffer for 2 x 25min rotating at room temperature. Before the fist wash step, the supernatant of the "no antibody control" samples was collected as the Input DNA samples and stored for further processing.

Reverse cross-linking:

After elution, NaCl, RNAse A and Proteinase K were added to all samples and left incubating at 65°C over night.

On the following day, reverse cross-linked samples were purified using the QIAquick PCR purification kit from Qiagen.

Purified immune-precipitated, non-precipitated controls and Input DNA were used as templates for quantitative real-time PCR, using the Corbett Research Rotor-Gene GG-3000 machine. The PCR mixture contained the SensiMix SYBR No-ROX master mix (Qiagen), 2 μl of the template DNA and the according primer pairs.

2. ChIP protocol applied for human monocytes, M1 and M2 macrophages

$1x10^7$ LPS-stimulated human monocytes and differentiated M1 and M2 macrophages, respectively, were collected for each immune-precipitation to be performed.

Fixation of cells:

Cells were fixed in warm cell culture medium directly supplemented with 1% formaldehyde and incubated for 10min. 125mM Tris pH 7.5 was added for neutralization, harvested and washed 4x with PBS. After the last washing step, the supernatant was removed.

Cell- and nuclei lysis:

First, proteinase inhibitors were added to each of the lysis buffers LB1, LB2 and LB3.

Cells were resuspended in LB1 (0,5ml per 10^7 cells), incubated on ice for 10min and harvested by centrifugation at 500 x g at 4°C for 5min. Pellet was resuspended in LB2 (0,5ml per 10^7 cells), incubated for 10min gently rocking at room temperature. Nuclei were harvested by centrifugation at 500 x g at 4°C for 5min. The pellet was then resuspended in LB3 (0,5ml per 10^7 cells).

Sonication:

Samples were distributed in tubes á 250µl and subjected to sonication performed on the Bioruptor UCD200.

Settings: 30 sec on/30 sec off, intensity level H, sonication time 4min;

Distributed samples were then pooled back in one tube, respectively and 50µl of 10% Triton X-100 was added to each. Centrifugation at 4°C at maximum speed removed remaining cell debris. 50µl of each cell lysate were stored at 20°C for further processing, to be used as the Input DNA samples.

Preparation of Protein G magnetic beads:

Dynabeads were blocked for 1h in 1xPBS/0.5%BSA, followed by two more washes in blocking solution. After the final wash, beads were resuspended in blocking solution, the according antibody was added and samples were incubated at room temperature for 1h rotating. After, beads were washed 3x in blocking solution and resuspended in 75µl blocking solution per 30µl beads.

Chromatin immune precipitation:

70µl of antibody/magnetic bead mix were added to each sonicated LB3 fraction and gently mixed at 4°C over night, rotating.

Washes and elution:

On the following day, supernatants were removed and samples were washed 6x in ice cold washing buffer. Before elution, samples were washed once more with TE buffer containing 50mM NaCl. For elution, beads were resuspended in 60µl elution buffer and incubated at 65°C for 15min. During this time samples were resuspended by brief vortexing every 2min. The eluates were then collected in fresh tubes, followed by another elution step performed as before.

Reverse cross-linking:

After elution, 2µl Proteinase K were added to each sample and left incubating at 65°C over night.

On the following day, reverse cross-linked samples were purified using the QIAquick PCR purification kit from Qiagen.

Purified immunoprecipitated, non-precipitated controls and Input DNA were used as templates for quantitative real-time PCR, using the Corbett Research

Rotor-Gene GG-3000 machine. The PCR mixture contained the SensiMix SYBR No-ROX master mix (Qiagen), 2μl of the template DNA and the according primer pairs.

2.14 Protein analysis

Western Blot analysis

The cells were washed and harvested in ice cold PBS. After centrifugation at 500g for 5min at 4°C, the cell pellet was resuspended in RIPA buffer, incubated on ice for 10min, and then centrifuged at 15,800g for 10min. Protein concentration was estimated by the Bradford assay. 10μg of protein extract was resuspended in 2x Laemmli buffer. The reactions were incubated at 100°C for 5min to denature the proteins fully. The proteins were resolved on a SDS-polyacrylamide gel and transferred to Hybond-C extra nitrocellulose membrane using a semi-dry Transblot apparatus. The membrane was blocked with 5% skimmed milk in PBS with 0.1% Tween (PBST) for 1h on a rocking platform, then washed three times for 10min with PBST on a rocking platform. Incubation with the primary antibody took place either for 1h at room temperature or at 4°C overnight in PBST with 5% skimmed milk. The membrane was washed in PBST (three times for 10min) and probed with the secondary antibody for one hour. The membrane was then washed three times for 10min with PBST and the blot was developed using the ECL Western Blotting Detection Kit according to the manufacturers instructions.

CHAPTER 3

STUDY ON THE ABILITY OF COHESIN AND CTCF TO TERMINATE TRANSCRIPTION OF THE β-GLOBIN GENE

3.1 Introduction

Transcription termination of mammalian Pol II requires a functional poly(A) signal and, in many cases, also a downstream terminator sequence.

Newly synthesized pre-mRNA molecules are usually cleaved at the poly(A) site, followed by pre-mRNA 3` end processing and Pol II termination. For genes which possess terminator elements, such as the *β-globin* gene, the first pre-mRNA cleavage event occurs at the terminator sequence, which is positioned in the gene`s 3` flanking region. This is then followed by a process leading to transcriptional termination of the polymerase. Pre-mRNA is then further cleaved at the poly(A) site for the 3` end processing machinery to generate stable mRNA.

Attempts to understand the mechanism of transcriptional termination have focused on the terminator of the human β-globin gene (Dye et al, 2001). This terminator region is an 800 base pair (bp) AT-rich sequence element located approximately 1kb downstream of the poly(A) signal and has been proven to be essential for efficient termination (Dye et al, 2001). This termination process occurs through a mechanism termed co-transcriptional cleavage (CoTC) (Dye et al, 2001), which, as the name implies, occurs co-transcriptionally. CoTC generates an unprotected 5`end, which becomes a substrate for degradation by the 5` to 3` RNA exonucleolase, Xrn2 (West et al, 2004). CoTC and subsequent Xrn2 activity leads to the release of Pol II from the DNA template (West et al, 2008).

Transcription termination of convergently transcribed genes in *Saccharomyces pombe (S. pombe)* is regulated by a mechanism that involves the highly conserved, ubiquitious and multi-functional Cohesin-complex (Gullerova et al, 2008). Here, dsRNA resulting from overlapping transcription of convergent genes in G1 phase of the cell cycle induces localized RNAi-dependent transient heterochromatin structure and Swi6 association, resulting in Cohesin recruitment (Gullerova et al, 2008). In G2, Cohesin is concentrated in the intergenic regions of the convergent genes, which results in a block to dsRNA formation by promoting gene-proximal transcription termination between the convergent genes (Gullerova et al, 2008). In mammalian cells, genome-wide ChIP-sequence analysis showed that Cohesin is often recruited to specific sites in a CTCF-dependent manner (Parelho et al, 2008; Rubio et al, 2008; Stedman et al, 2008; Wendt et al, 2008).

My project was intended to bridge our knowledge from *S. pombe* to mammalian cells by determining, whether Cohesin is also able to mediate transcriptional termination in a mammalian system. From a functional point of view I intended to identify, whether the known interaction of CTCF and Cohesin can potentially lead to transcription termination in mammals.

The arrangement of genes in the yeast and mammalian genome is fundamentally different. While the yeast genome is of very compressed nature, in the mammalian genome, functional genes can be up to several hundred kilobases apart. This striking difference in gene density makes it unlikely that similar gene arrangements occur, which would allow me to

analyse endogenous genes in mammals in the same context as in *S. pombe*. It was therefore necessary to set up of a model system that would allow me to address the role of Cohesin and CTCF in a simplified arrangement to facilitate mechanistic studies.

For this purpose, I made use of the established β-globin system, which comprises of a plasmid carrying the human β-globin gene driven by the Tat-inducible HIV-LTR promoter. The 3` end flanking region contains 240bp of post poly(A) site sequence (lacking terminator activity) followed by the wild-type terminator sequence (βTERM). βΔTERM, is the same as βTERM but lacks the terminator sequence and has been shown to be incapable of directing efficient termination (Dye et al, 2001). Both constructs were kind gifts of Dr. M. Dye.

Cohesin has not been shown to interact with any specific DNA motif. In order to achieve Cohesin recruitment to the 3` end of the β-globin gene, I made use of the previously described ability of CTCF to recruit Cohesin to CTCF-binding sites in human cells (Parelho et al, 2008; Rubio et al, 2008; Stedman et al, 2008; Wendt et al, 2008). For this purpose, I introduced CTCF-binding sites into the βΔTERM construct, in order to achieve CTCF and Cohesin recruitment. The CTCF-binding sites were cloned in tandem, two or five copies, respectively, resembling the minimum core sequence, which has been shown in *in vitro* gel mobility shift experiments to successfully bind CTCF (Renda et al, 2007). This resulted in the new constructs βΔTERM-CTCF2x and βΔTERM-CTCF5x. The plasmids and the sequence arrangement of the CTCF binding sites are shown in Figure 3.1 A and B. Figure 3.1C shows the

CTCF-binding motif acquired from a CTCF ChIP-Seq study in HeLa cells (Cuppadah S et al, 2009), published a year after I started my β-globin project. The outcome of this advanced study confirms the CTCF core-binding sequence I applied in my experimental set up.

3.2 Analysis of the efficiency of the core CTCF-binding sites to terminate transcription of the HIV-LTR β-globin gene plasmid constructs.

HeLa cells were transiently transfected with constructs βΔTERM-CTCF2x and βΔTERM-CTCF5x, as well as βTERM as a positive control and, βΔTERM as the negative control. In each case, a plasmid expressing the transactivator Tat (pTat) was co-transfected to activate efficient Pol II initiation/elongation from the HIV-LTR (for review see (Cullen, 1993)). A plasmid containing the VA I adenovirus gene (pVa) was used as a co-transfection control, confirming equal transfection efficiency. The transcription termination capacity of each sequence was first measured using quantitative RT-PCR (qRT-PCR) analysis. Total RNA was isolated from HeLa cells transfected with the different β-globin constructs and co-transfected with pTat. Isolated RNA was then reverse transcribed with primer rtR to select for RNA, which has not been cleaved in the terminator region (see Figure 3.2 A). This RNA-species results from read-through transcription due to a lack of transcription termination. RNA continuity in the 3` end was measured with the primer pair rtF/rtR and normalized to RNA amplified with primer pair Int1/Ex2 to detect nascent RNA levels.

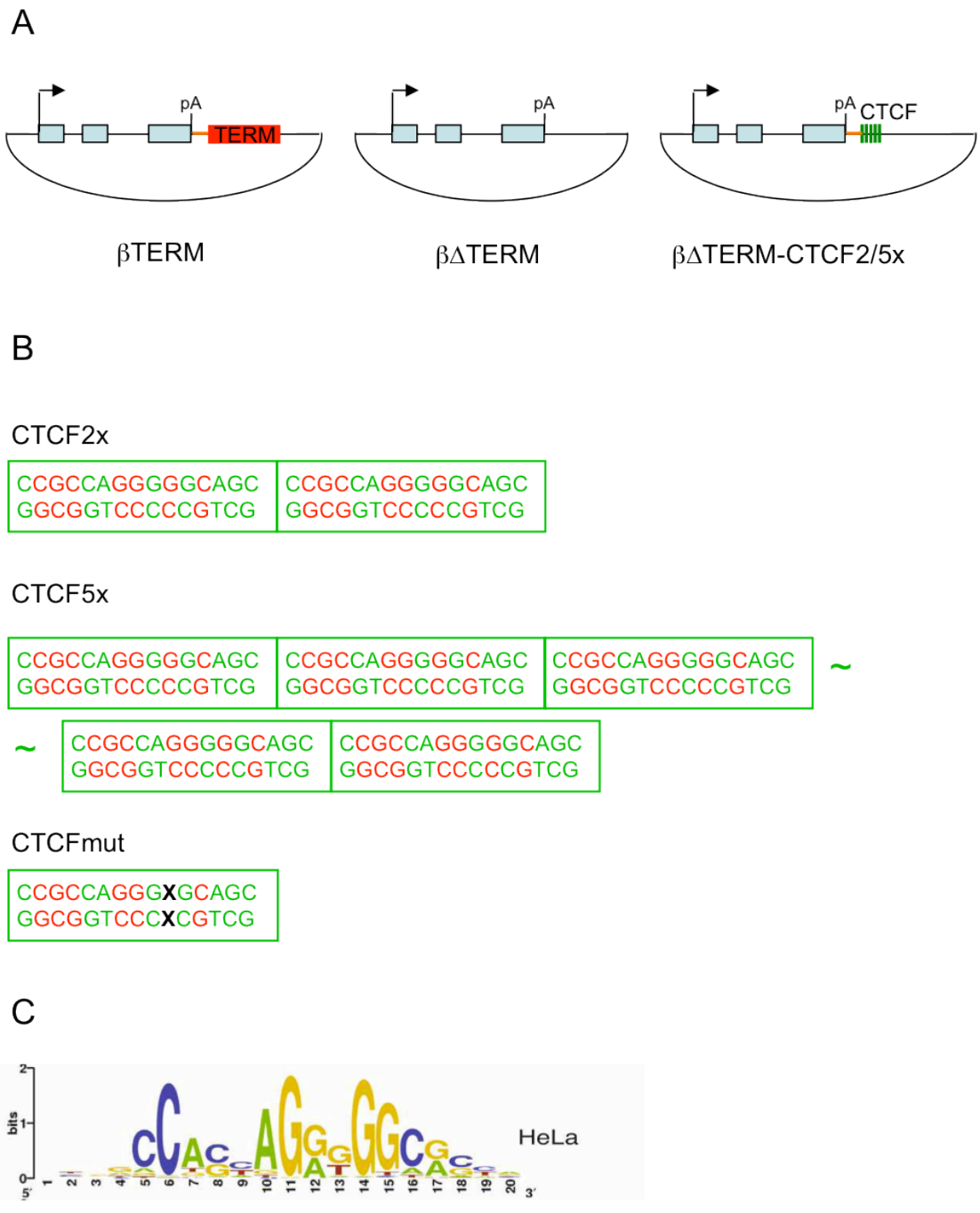


Figure 3.1: The human β-globin plasmid, Terminator sequence and CTCF binding sites **A.** The β-globin plasmids. DNA is indicated as a black line. The arrow indicates the start site of transcription, the light blue boxes indicate Exons, pA represents the poly(A) signal, the orange line indicates the 240bp post poly(A) sequence without terminator activity and the red and green boxes represent the Terminator and the CTCF-binding sites. βΔTERM is the Terminator-depleted β-globin construct **B.** Sequence of the core CTCF-binding site cloned in two and five copies in tandem into the βΔTERM plasmid, as well as a single copy mutant. Highlighted in red are the nucleotides shown to be essential for efficient CTCF binding. Indicated as a black **X** is the missing G-C pair in the CTCF mutant. **C.** Confirmation of the core CTCF-binding sequence used in the present β-globin project: representation of the CTCF-binding motif acquired from CTCF ChIP-sequencing (ChIPSeq) analysis in HeLa cells in a later study. Taken from Cuddapah S et al, 2009).

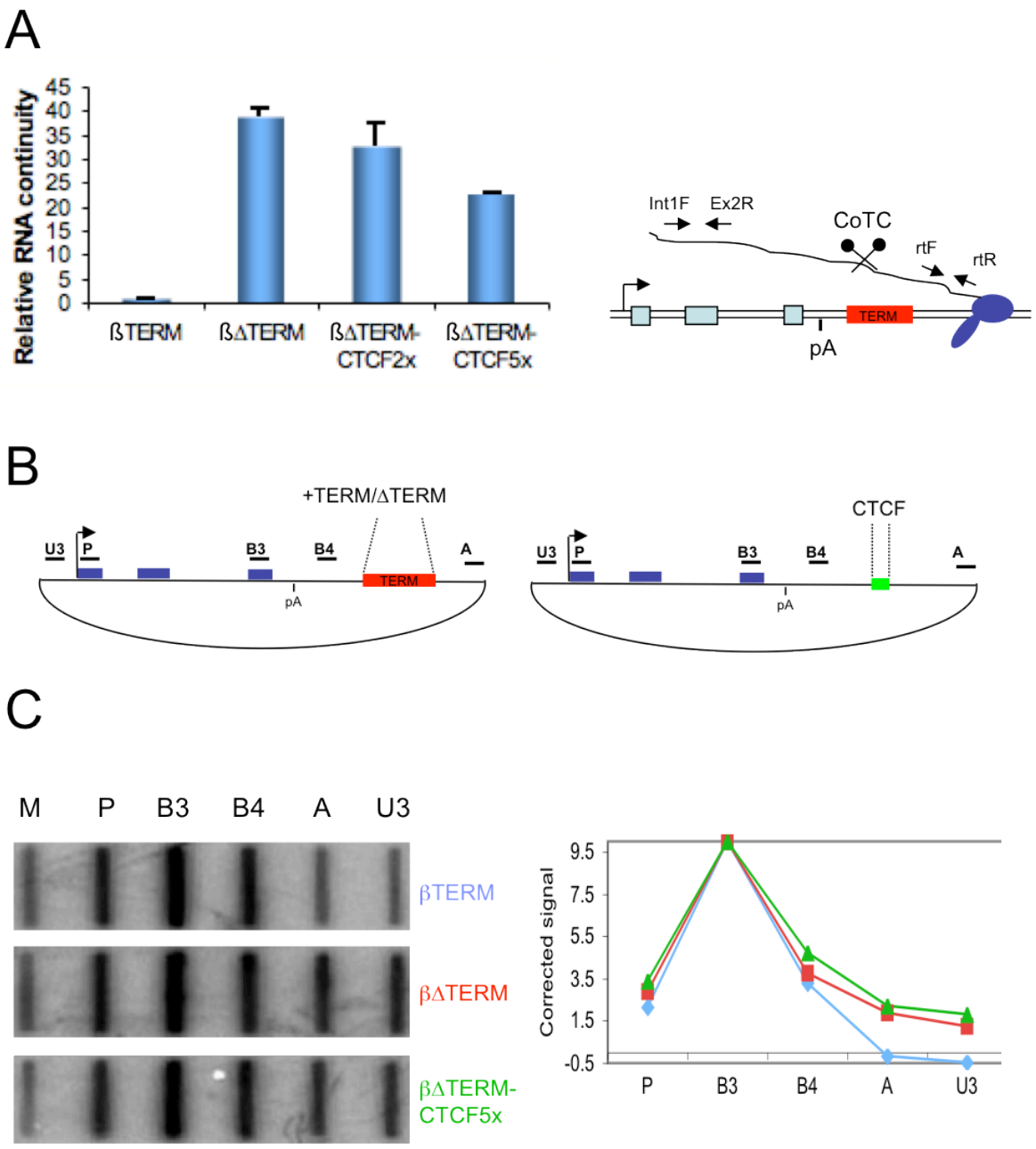


Figure 3.2: Transcription termination efficiency at CTCF-binding sites placed at the 3` end of human β-globin gene constructs. A. qRT-PCR analysis of total RNA isolated from HeLa cells transiently transfected with pTat and βTERM, βΔTERM or βΔTERM-CTCF2x/5x plasmids, respectively. The ratio of Terminator to Intron 1/Exon 2 RNA for βΔTERM-CTCF2x/5x and βΔTERM was calculated and corrected to the positive control, βTERM, which is set at 1. Error bars indicating standard deviation over multiple biological replicates are shown. Diagram showing the position of the PCR amplicons, and primers used for RT-PCR. Primer rtR was also used for reverse transcription of RNA into cDNA. PCR product detected following the Terminator region (CoTC) is indicative of the level of transcripts that are not cleaved over the Terminator region. **B.** Diagram of the b-globin plasmids. The transcription start site (arrow) and regions of complementarity of antisense ssDNA M13 probes are indicated (bold). **C.** NRO analyses of nascent transcripts of HeLa cells transiently transfected with pTat and βTERM, βΔTERM or βΔTERM-CTCF5x, respectively. NRO probes are indicated above their respective positions on the filters. M represents the M13 vector and is a background control. Graph on the right of the NRO data panels shows the strength of hybridisation signals after subtraction of the background signal M, and normalisation for the number of U residues present in each transcript.

As expected, βΔTERM resulted in high levels of read-through message. In comparison, βTERM showed a low abundance of 3` end-unprocessed RNA due to efficient termination. High levels of read-through RNA were detected in both constructs containing CTCF-binding sites. This indicates that under these conditions, CTCF-binding sites, with a hypothetical CTCF-recruitment, do not contribute to successful transcription termination. As RNA isolation followed by qRT-PCR measures steady state RNA levels, I extended these results with nascent RNA analysis by performing a Nuclear Run On (NRO) analysis. This technique measures nascent transcription mapping of the position of active polymerases. NRO was performed on nuclei isolated from HeLa cells transiently transfected with pTat and βTERM, βΔTERM or βΔTERM-CTCF5x. Single strand probes were employed covering the promoter (P), the final β-globin Exon (B3), an immediate downstream region of the poly(A) site (B4) as well as the vector backbone (A) and a region upstream of the HIV-LTR promoter (U3). These latter probes (A and U3) detect transcription read-through. Probe M covers a sequence in the M13 vector and detects the background signals. All probes were cloned into the single strand M13 phage vector. NRO analysis involves the radio-labeling of nascent transcripts by incubation with α^{32}P-UTP, followed by partial hydrolysation and hybridization to single stranded DNA probes fixed to the hybridization filters as shown in Figure. 3.2C. As expected, hybridization signals for the construct βTERM are high over probes P, B3 and B4 and lower over the read-through-indicating probes, A and U3 (Figure 3.2C, top panel). My results confirm that efficient termination occurs in the presence of the wt terminator element. In contrast, and as expected, hybridization signals were

detected throughout the plasmid for the βΔTERM construct, proving that deletion of the β-globin terminator inactivates termination (Figure 3.2C middle panel). Comparing the hybridization signals of βΔTERM-CTCF5x with the control constructs it is evident, that βΔTERM and βΔTERM-CTCF5x display similar behavior. The strong signals over region A and U3 indicate, that replacing the deleted terminator by CTCF-binding sites does not lead to efficient transcription termination (see Figure 3.2C, lowest panel). The strength of the NRO signals was quantified by PhosphoImager, normalized to the number of U resides in each probe, and displayed in the graph next to the autoradiograph (Figure 3.2C, diagram at the right).

3.3 Analysis of cell cycle dependency of Cohesin-mediated transcription termination on β-globin gene constructs

As demonstrated above, it appears that the CTCF-binding sites do not promote transcription termination in the β-globin gene assay. However, in *S. pombe,* Cohesin appears to act on transcription termination in a cell cycle dependent manner. While read-through transcription is detectable in the G1 phase, termination occurs through Cohesin recruitment in the G2 phase of the cell cycle (Gullerova et al, 2008).These data imply that such a cell cycle dependency for Cohesin recruitment could also occur in human cells mediated through recruitment to DNA-bound CTCF. I therefore elected to investigate the possibility that CTCF binding sites could elicit Pol II termination through Cohesin recruitment in a cell cycle specific manner. To do this, I tested a number of common cell cycle synchronization protocols.

The most efficient G2 phase synchronization method for HeLa cells was achieved by a combination of an initial thymidine block, followed by incubation of the cells with the cell cycle inhibitor Roscovitine. Roscovitine is a purine analogue that is a potent and selective inhibitor of cyclin-dependent kinases (cdk) (Meijer et al, 2003; Bain et al, 2003). In particular, it is a competitive inhibitor of cdc2/cyclin B, cdk2/cyclin A, cdk2/cyclin E, and cdk5/p35.

HeLa cells were first transiently transfected with pTat and the various β-globin constructs. 24h post transfection, Thymidine was added to a final concentration of 2mM and incubated for 24h at 37°C. Cells were then washed and released for 4h at 37°C. Afterwards, Roscovitine was added to the media to a final concentration of 25μM and incubated for 18h at 37°C. Finally, cells were fixed in 70% Ethanol, stained with Propidiumiodide and analyzed by flow cytometry, in particular by FACS (Fluorescence-activated cell sorting) analysis.

Figure 3.3A presents the graphical data from the CellQuest software (Becton Dickinson) which analyses the signals from the FACScan. The number of cell counts reflected in the height of the peaks over position 200 and 400 represent the distribution of cells residing in either the G1 or G2 phase, respectively. Cells in the G1 phase possess a single copy DNA content (n) (peak at 200), while G2-phase cells have replicated their DNA (2n) and therefore emit a higher fluorescent signal, represented by the peak at position 400 (x-axis). Next to the CellQuest cell cycle distribution graphs are the quantifications of the peaks, setting the total, counted cell number to 100 and

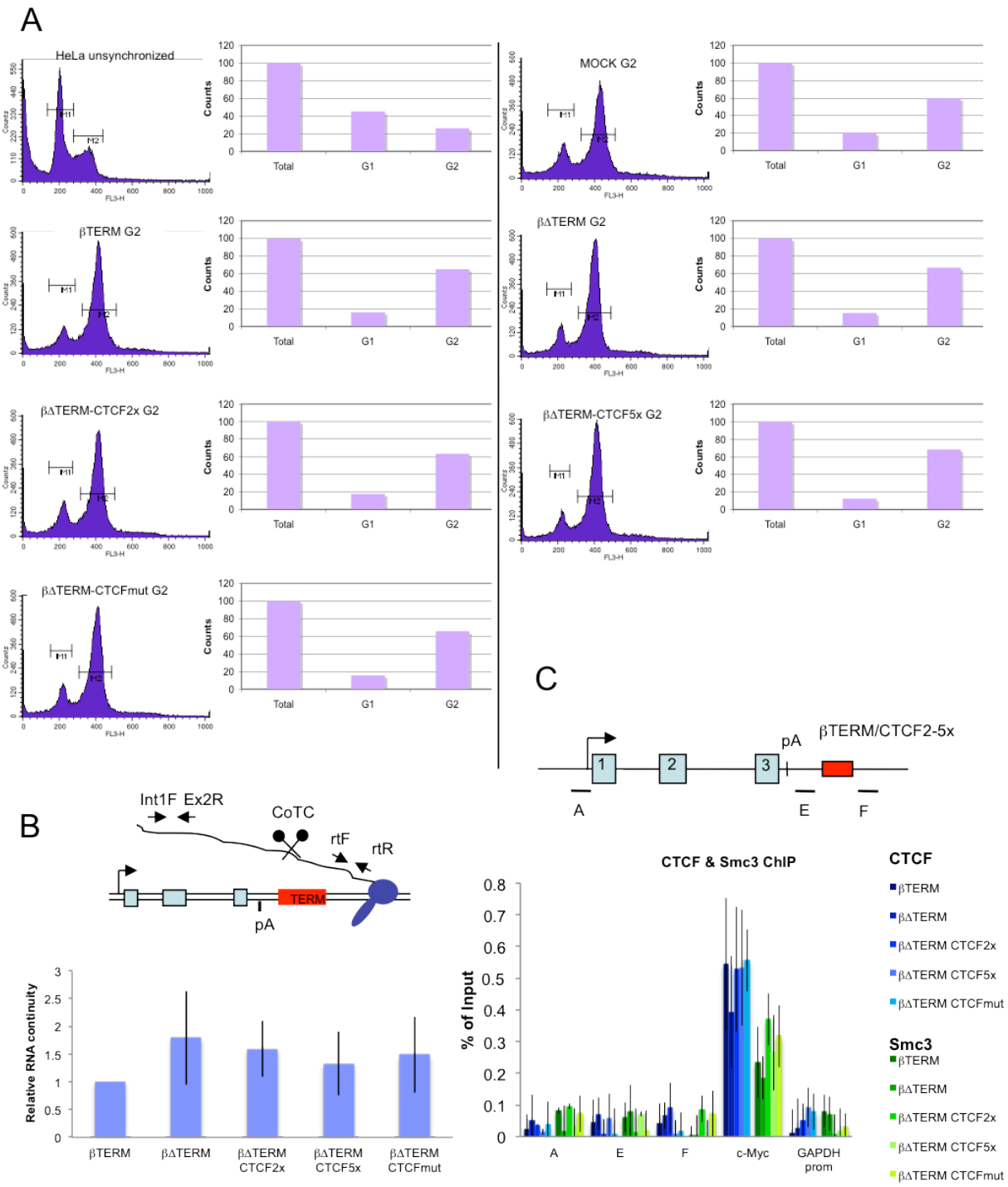


Figure 3.3: Transcription termination efficiency at CTCF-binding sites placed at the 3` end of human β-globin gene constructs in G2 synchronized HeLa cells. **A.** Cell cycle distribution of HeLa cells before and after cell cycle synchronization with Thymidine (2mM, 24h) and Roscovitine (25μM, 18h) was evaluated by flow cytometry. Cell DNA content was measured using the FACScan system with CellQuest software (Becton Dickinson, Franklin Lakes, NJ, USA). M1 and M2 indicate the DNA content in the cells and refer to G1 and G2 cell cycle phase. **B.** qRT-PCR analysis of total RNA isolated from G2-synchronized HeLa cells transiently transfected with pTat and βTERM, βΔTERM or βΔTERM-CTCF2x/5x/mut plasmids, respectively. The ratio of Terminator to Intron 1/Exon 2 RNA for βΔTERM-CTCF2x/5x/mut and βΔTERM was calculated and corrected to the positive control, βTERM, which is set at 1. Diagram above shows the position of the PCR amplicons, and primers used for RT-PCR. Primer rtR was used for reverse transcription of RNA into cDNA. PCR product detected after the Terminator region (CoTC) is indicative of the level of transcripts that are not cleaved over the Terminator region. **C.** qPCR quantitation of ChIP analysis for CTCF and Smc3 (Cohesin) recruitment to the CTCF binding sites. *c-Myc* serves as a positive control for CTCF- and Cohesin recruitment. Values are indicated as percent of total input DNA. Diagram of the β-globin gene shows the promoter (arrow), Exons (blue, numbered boxes), Terminator or CTCF binding sites (red boxe) and the position of the primers indicated below. All values are based on average values ± s.d. from three independent biological experiments.

showing the G1-G2 phase distribution within the total counts.

In unsynchronized samples, HeLa cells are predominantly found in the G1 phase, as shown in the first graph of Figure 3.3A. To ensure that the transfection procedure did not influence the cell synchronization, I performed the same analysis on untransfected, G2-synchronized cells (Figure 3.3A graph top right, Mock transfection). The results confirm that the transfection procedure itself did not influence the cell cycle synchronization ability of the cells.

The next steps, after synchronization of cells in the G2 phase, were to test for the transcription termination ability of the CTCF-site-containing β-globin constructs (Figure 3.3B) as well as to test for CTCF and Cohesin recruitment to the introduced CTCF-binding sites (Figure 3.3C). In these experiments I included one additional construct, with a mutated version of the core CTCF-binding site (βΔTERM-CTCFmut; see Figure 3.1B for sequence). Although βΔTERM-CTCFmut has only one copy of the mutated CTCF-binding site, I considered it to be a potentially valuable, additional control.

Transfected HeLa cells with pTat, pVa and each of the β-globin constructs were synchronized in the G2 cell cycle phase, using the protocol described above. A fraction of the cells was analysed by FACS (to confirm successful G2-synchronization) and the rest were used for total RNA isolation and ChIP analysis. For the transcription termination analysis, RNA was then, as described previously, reverse transcribed using primer rtR, followed by qRT-

PCR analysis with primers rtF/rtR and normalized to the values obtained from primer pair Int1/Ex2 (see diagram in Figure 3.3B). The result obtained turned out to be little meaningful as careful consideration highlights two issues, which led me to question the validity of this experiment. First, and most importantly, read-through transcription for the control construct βΔTERM was too low compared to the levels obtained in unsynchronized conditions (compare with Figure 3.2A). This indicates negative effects of the synchronization procedure on the transcription process. Second, technical difficulties appeared to bring high instability to repeat this experiment (reflected in the high error bars).

ChIP analysis was similarly affected by technical difficulties, which became obvious from the allover low signal strength and the high error bars as a sign of difficulties with experimental reproducibility. However, in terms of ratios, both, the positive control for CTCF and Cohesin recruitment, the CTCF binding site upstream of the TSS of the *c-Myc* gene and the negative control, the *GAPDH* promotor, behaved as expected (Figure 3.3C). Based on those ratios, the signals obtained across the β-globin constructs (probes A, E and F) fall into background levels, indicating unsuccessful recruitment of CTCF and Cohesin to the introduced CTCF-binding sites (probe E and F).

The major problem with the experiments performed on G2-synchronized HeLa cells was that they appeared very `sick` after cell cycle synchronization with Roscovitine. After subjecting these cells to the cdk-inhibitor I could optically recognize an increased level of cell stress, partially displaying signs of apoptosis. Furthermore, even though FACS analysis revealed a successful

synchronization of these cells (Figure 3.3A), only a small fraction of cells were analyzed (60.000) out of an entire sample (10^7). The samples used for total RNA isolation and ChIP were, despite washes before collection, necessarily a mixture of dead and living cells. Consequently, the G2 synchronized samples were questionable. It is most likely that the high stress status of these cells may have affected their transcriptional profile over the transfected plasmid.

In conclusion I predict that these experiment failed due to the unexpected side effects of the Rocovitine treatment on transcription. It is therefore impossible to draw any meaningful conclusions from these experiments.

3.4 Analysis of FII, the endogenous chicken β-globin 5`HS4 insulator sequence as a β-globin terminator in the established plasmid system.

CTCF is, as already mentioned, a highly conserved zinc finger protein implicated in diverse genomic regulatory functions. Combinatorial use of the eleven zinc fingers enables CTCF to bind to a wide range of variant sequences, as well as specific co-regulatory proteins. Gel mobility shift analysis by Renda et al. (Renda et al, 2007), have identified the CTCF DNA binding domain responsible for interaction of CTCF with various biologically relevant DNA-binding sites. Importantly, they have also been able to pin point the minimum, essential DNA-sequence requirement to bind CTCF, which resembles 12 base pairs forming the core. Within that sequence eight single, highly conserved nucleotides appear identical throughout a number of different CTCF-binding sites. Mutations within the core sequence abolish any

CTCF protein-DNA interaction. These experiments, having been performed *in vitro,* might not correspond to the requirements of an endogenous environment, implicating the complexity of the gene/locus regulatory elements, within which CTCF-binding sites are often located. The above described gel shift analysis identified CTCF zinc fingers (ZF) 4-8 to be the essential domain for CTCF DNA interaction. However, they suggest that, it is ZF4-7 that recognizes the core sequence, while ZF8 most likely interacts with sequences outside the core and might be responsible for additional interaction stability.

I was unsure, whether the minimal core CTCF-binding site alone, which I used in my first approach (Figure 3.1), was enough to ensure stable CTCF-DNA interaction in an *in vivo* environment. Therefore, I decided to apply an identical experimental procedure on a new β-globin construct, containing a larger CTCF-binding site sequence. I chose the CTCF-binding sequence located in the 5` HS4 site, a well-characterized insulator regulating the chicken *β-globin* locus (Chung et al, 1997; Reitman et al, 1988; Bell et al, 1999). The insulator activity in the chicken *β-globin* locus was pinpointed to a 250bp GC rich "core" fragment, which maps precisely with the HS4 site (Chung et al, 1997). DNase I footprinting of this core fragment divided the region into five sections (FI-FV), revealing the 42bp long FII sequence to posses an enhancer blocking activity in human cells (Bell et al, 1999). The same study identified the FII fragment to contain a conserved CTCF-binding site and to bind CTCF. I applied these findings to improve my strategy by cloning the entire 45 base pairs-long FII sequence into the 3` end of the Tat-inducible β-globin gene construct, as

before, replacing the β-globin terminator region. The newly designed plasmids are named βΔTERM-FII. To increase the chance of detecting an effect on termination, I cloned the FII-region in several copies in tandem, and confirmed their sequence accuracy and orientation by sequencing (Figure 3.4A depicting the β-globin plasmids and 3.4B showing the sequence details of the CTCF binding FII site).

As described before, the first approach to determine the transcription termination efficiency of the various β-globin constructs was to perform a quantitative RT-PCR analysis. HeLa cells were transiently co-transfected with pTat and the β-globin plasmids. Total RNA was isolated, reverse transcription performed with primer rtR, and qRT-PCR set up with primers rtF/rtR to measure read-through transcription. Primers Int1/Ex2 were again used for normalization purposes. The result clearly indicates high levels of read-through transcription detectable in all constructs containing the FII sites, and therefore displaying the same termination defect as βΔTERM (Figure 3.5A).
To confirm these data at the level of nascent RNA, I subjected isolated nuclei from transiently transfected HeLa cells to NRO analysis. Active Pol II density was, as described before, measured over regions that include the promoter (P), the final β-globin exon (B3), and the immediate region downstream of the poly(A) site (B4) as is shown in the diagram in Figure 3.5B. Probes A and U3, located in the vector backbone and over the HIV-LTR, indicate read-through transcription. I transfected the two control plasmids, βTERM and βΔTERM and only one of the constructs containing the CTCF-binding FII site (FII4x). As expected, hybridization signals for the construct βTERM were high over

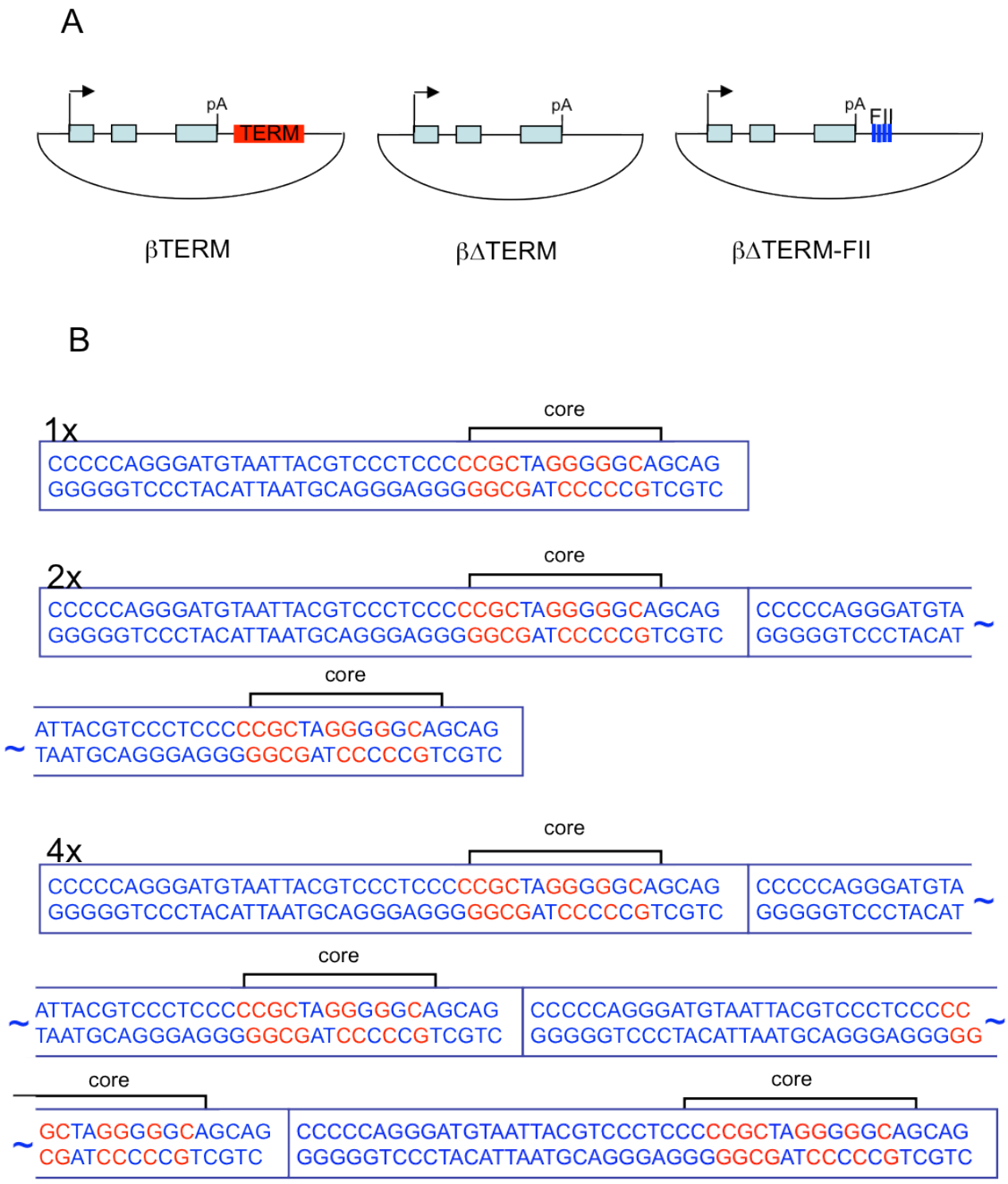


Figure 3.4: The β-globin plasmid, Terminator sequence and FII binding sites A. The β-globin plasmids. DNA is indicated as a black line. The arrow indicates the start site of transcription, the light blue boxes indicate the Exons, pA represents the poly(A) signal and the red and blue boxes represent the Terminator and the FII-binding sites in plasmids βTERM and βΔTERM-CTCF-FII, respectively. βΔTERM is the Terminator-depleted β-globin construct. **B.** Sequence of the endogenous CTCF binding chicken β-globin FII site cloned in one, two and four copies in tandem into the βΔTERM plasmid. Highlighted in red are the nucleotides shown to be essential for successful CTCF binding. The minimum, essential CTCF-binding sequence is indicated as the core.

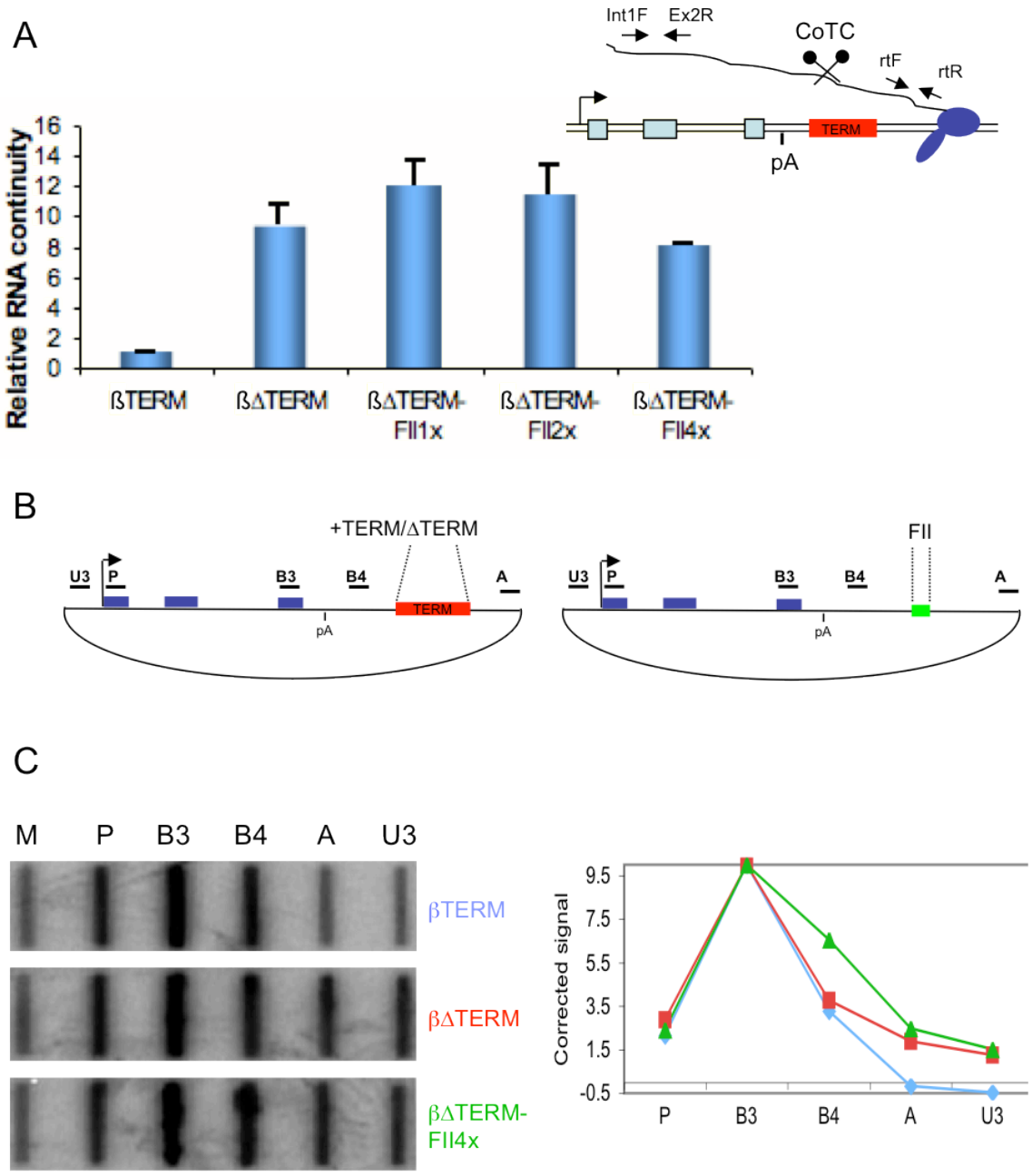


Figure 3.5: Transcription termination efficiency at CTCF-recruiting FII sites placed at the 3` end of human β-globin gene constructs. **A.** qRT-PCR analysis of total RNA isolated from HeLa cells transiently transfected with pTat and βTERM, βΔTERM or βΔTERM-FII4x plasmids, respectively. The ratio of Terminator to Intron 1/Exon 2 RNA for βΔTERM-FII4x and βΔTERM was calculated and corrected to the positive control, βTERM, which is set at 1. Error bars indicating standard deviation over three biological replicates are shown. Diagram showing the position of the PCR amplicons, and primers used for RT-PCR. Primer rtR was also used for reverse transcription of RNA into cDNA. PCR product detected after the Terminator region (CoTC) is indicative of the level of transcripts that are not cleaved over the Terminator region. **B.** Diagram of the b-globin plasmids. The start site of transcription (arrow) and regions of complementarity of antisense ssDNA M13 probes are indicated (bold). **C.** NRO analyses of nascent transcripts of HeLa cells transiently transfected with pTat and βTERM, βΔTERM or βΔTERM-FII4x, respectively. NRO probes are indicated above their respective positions on the filters. M represents the M13 vector and is a background control. Graph on the right of the NRO data panels shows the strength of hybridisation signals after subtraction of the background signal M, and normalisation for the number of U residues present in each transcript.

probes P, B3 and B4 and low over A and U3 (Figure 3.5C, top panel). This confirms the wt terminator element to be an efficient terminator. In contrast, hybridization signals for the βΔTERM construct were detected throughout the plasmid (Figure 3.5C middle panel). Comparing the hybridization signals of βΔTERM-FII4x with the control constructs it is evident that βΔTERM-FII4x is unable to terminate transcription efficiently (see Figure 3.5.C lowest panel). The strength of the NRO signals was quantified by PhosphoImager, normalized to the number of U resides in each probe, and displayed in the graph next to the autoradiograph (Figure 3.5C, diagram at the right). This nascent RNA analysis, therefore, confirms the result obtained from quantitative RT-PCR.

3.5 Analysis of the possibility of cell cycle dependency of Cohesin-mediated transcription termination – approach N°2

The cell synchronization approach, described in section 3.3, was technically challenging and resulted in an unclear termination efficiency read out, most probably due to the low viability of the HeLa cells after treatment with the cdk-inhibitor Roscovitine. Hoping that an improved synchronization protocol would increase cell survival to give reproducibility to the subsequent RNA analysis, I used a modified protocol, significantly minimizing the application of Roscovitine. Again, HeLa cells were first transiently transfected with pTat, pVa and the various β-globin constructs. 24h post transfection and with a cell confluence of approximately 70%, Thymidine was added to a final concentration of 2mM and incubated for 16h at 37°C. Cells were then washed

with warm media and released for 8h at 37°C. Afterwards, Thymidine was added again to a final concentration of 2mM and incubated with the cells for another 16h at 37°C. Followed by one more wash and a release phase of 6h at 37°C, Roscovitine was added, this time only to a final concentration of 1μM, and incubated with the cells for only 2h at 37°C. After that, cells were fixed with 70% Ethanol and stained with Propidiumiodide for FACS analysis. Figure 3.6A displays the graphs with the number of cell counts, reflected in the height of the peaks over position 200 (G1) and 400 (G2) of the x-axis. Next to the CellQuest cell cycle distribution graphs are the quantifications of the peaks, setting the total, counted cell number to 100 and showing the G1-G2 phase distribution within the total counts. The first control was again to ensure that transfection of the cells prior to synchronization does not influence the cell`s ability to synchronize (Figure 3.6A graph top right compared to all graphs of transfected G2-HeLa cells). Following the FACS results it appears that the modified protocol does successfully synchronize untransfected cells (Mock G2). However, β-globin plasmid transfected cells do appear to display a lower G2-synchronization efficiency compared to the first protocol applied (compare to Figure 3.3A). As the ratio of G1 to G2 cells was still shifted towards an increased number of G2-synchronized cells, I preceded with the analysis, expecting that if termination occurs at the CTCF-binding sites in G2, a shift in levels of read-though transcripts would be detectable. Cells were therefore subjected to total RNA isolation and quantitative RT-PCR. RNA was again reverse transcribed using primer rtR, followed by qRT-PCR analysis with primers rtF/rtR and normalized to the values obtained from primer pair Int1/Ex2 (see diagram in Figure 3.6B). Interestingly, this second approach

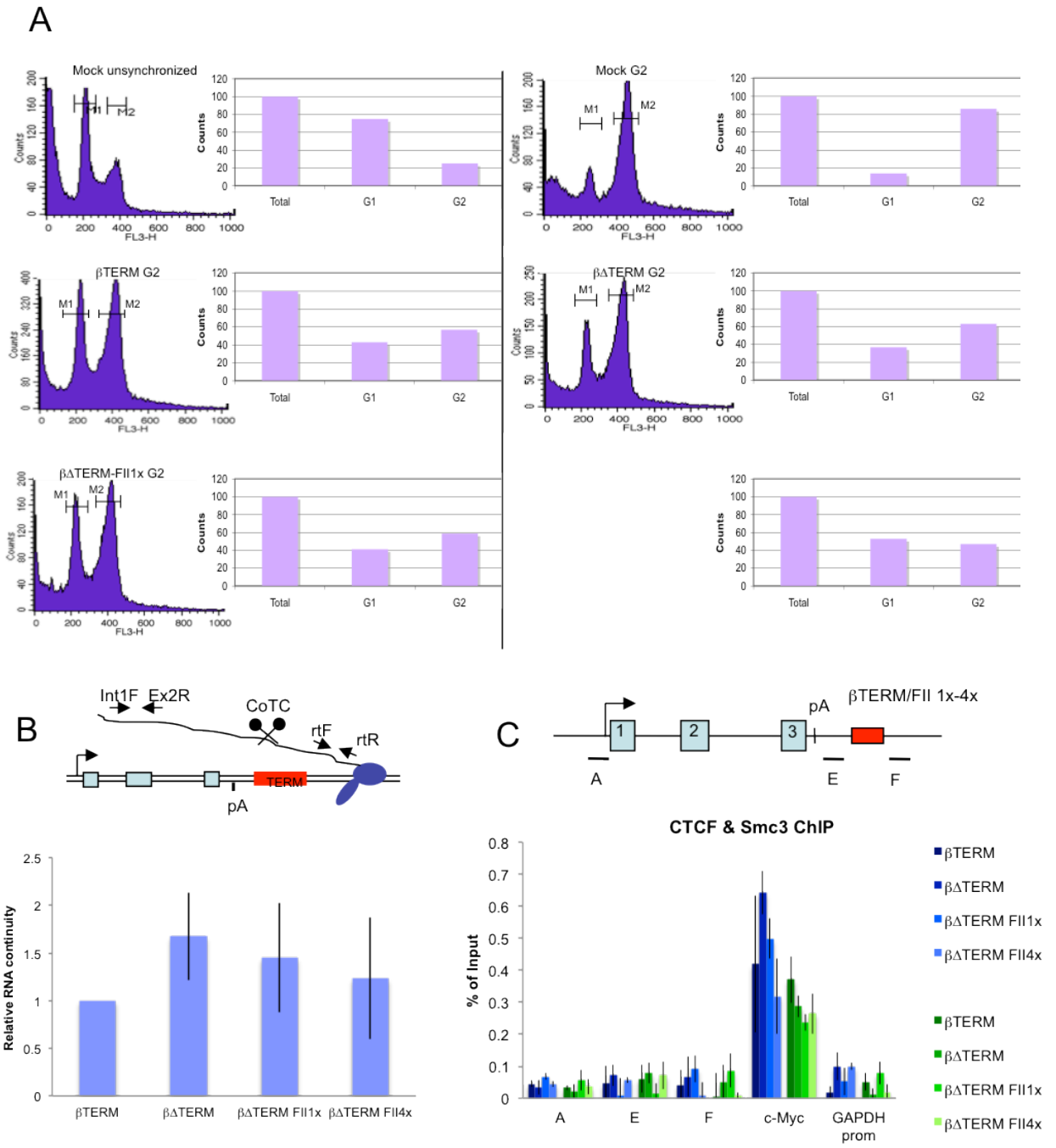


Figure 3.6: Transcription termination efficiency at CTCF-recruiting FII sites placed at the 3` end of human β-globin gene constructs in G2 synchronized HeLa cells. **A.** Cell cycle distribution of HeLa cells before and after cell cycle synchronization with Thymidine (2.5mM, 2x16h) and Roscovitine (1μM, 2h) was evaluated by flow cytometry. Cell DNA content was measured using the FACScan system with CellQuest software (Becton Dickinson, Franklin Lakes, NJ, USA). M1 and M2 indicate the DNA content in the cells and refer to G1 and G2 cell cycle phase. **B.** qRT-PCR analysis of total RNA isolated from G2-synchronized HeLa cells transiently transfected with either βTERM, βΔTERM or βΔTERM-FII1x/4x plasmids and pTat. The ratio of Terminator to Intron 1/Exon 2 RNA for βΔTERM-FII1x/4x and βΔTERM was calculated and corrected to the positive control, βTERM, which is set at 1. Diagram on the right showing the position of the PCR amplicons, and primers used for RT-PCR. Primer rtR was also used for reverse transcription of RNA into cDNA. PCR product detected after the Terminator region (CoTC) is indicative of the level of transcripts that are not cleaved over the Terminator region. **C.** qPCR quantitation of ChIP analysis for CTCF and Smc3 (Cohesin) recruitment to the CTCF binding sites. *c-Myc* serves as a positive control for CTCF- and Cohesin recruitment. Values are indicated as percent of total input DNA. Diagram of the β-globin gene shows the promoter (arrow), Exons (blue, numbered boxes), Terminator or FII sites (red boxe) and the position of the primers indicated below. All values are based on average values ± s.d. from three independent biological experiments.

with altered cell cycle synchronization conditions, and the introduction of an improved CTCF-binding site lead to a similar result as after the first synchronization approach (Figure 3.6B). Although the cell viability appeared improved, presumably due to lower Roscovitine concentrations in this synchronization protocol, the βΔTERM control again indicated negatively affected transcription. The high error bars again indicate problems in experimental reproducibility. Despite lower Roscovitine concentrations, affected cell viability or other indirect drug-mediated effects appeared to result in an obstacle to achieving reproducible data, which leads me to the conclusion that also these results cannot be considered reliable.

ChIP analysis for CTCF and Cohesin recruitment resulted in a similar scenario as before. Both, CTCF and Cohesin appeared to be recruited to the control site *c-Myc*, although the overall signal levels were again rather low (Figure 3.6C). Signals detected over the β-globin constructs (probes A, E, F) accounted for background as they did not differ from the signals detected at the GAPDH promotor. However, the high error across the experimental repeats indicates instability of the experimental system, despite improved experimental conditions.

I note that at this point I became concerned that Roscovitine, despite application in low concentration and improved cell viability, could have a major influence on transcription, and therefore lead to irreproducibility of the experiments.

In summary, despite two independent approaches to synchronize transfected HeLa cells in the G2 phase with subsequent RNA analysis, it appears that no meaningful conclusion can be drawn regarding termination events at the introduced CTCF-binding sites. Both experimental strategies suggest that negative effects on transcription, most likely caused by the synchronization procedure, distorted the experimental outcome. Attempting to find a possible explanation for this phenomenon, I considered the characteristics of Roscovitine and its biological impact on cells. Firstly, it has been observed that treatment of myeloma cells with Roscovitine induces rapid apoptosis (MacCallum et al, 2005). This observation is confirmed by the fact that the cell viability of the transfected HeLa cells used in my experiments was severely compromised. Comparing the overall levels of detectable RT-PCR products in unsynchronized and Roscovitine-treated cells, it becomes apparent that transcription may be significantly impaired by this drug treatment. This is consistent with the observation that Roscovitine negatively influences the mechanisms essential for Pol II CTD phosphorylation (MacCallum et al, 2005). Interestingly, work on drug induced, molecular alterations in cells, identified significant spectral changes in protein and nucleic acid structure upon treatment with different concentrations of Roscovitine (Akyuz et al, 2011).

3.6 Creation of a stable HEK293 β-globin expression system

Despite ease and utility of plasmid-based systems, particularly to study gene transcription at a functional level, I came across one major obstacle, which forced me to find an alternative experimental system. Having performed a

series of RNA analysis (as described above) to determine effects of CTCF and Cohesin on transcription termination, it remained an unresolved question, whether CTCF and Cohesin are actually recruited to the CTCF-binding sites that I engineered into the β-globin 3` end. The combination of high background signals and the restricted efficiency of the antibodies used affected the reliability of the acquired ChIP results in the plasmid system. Despite many attempts to increase stringency during washing steps, in order to reduce the high background signals, I could not find a way to entirely solve this technical problem in my experimental setup.

A major problem with transient transfection experiments is that only a small fraction of plasmid actually enters the nucleus and gets transcribed. The majority of plasmids remain adhered to the outside of the nuclear membrane. During ChIP experiments, nuclei are carefully isolated and the chromatin extracted. However, despite stringent washing steps, large amounts of untranscribed plasmid DNA `contaminates` the total input sample. This causes a high background in the ChIP analysis. This major obstacle led me to explore a different gene analysis system. As a solution to keep the gene constructs used in my transient transfection experiments, but to allow me to perform chromatin analysis in an endogenous environment, I created HEK293 cell lines stably expressing these gene constructs.

As depicted and described in Figure 3.7 and in Materials and Methods, HEK293 cells, containing an FRT-recombination site, were transfected with the pc5/FRT/TO plasmids carrying the β-globin constructs. The FRT site on

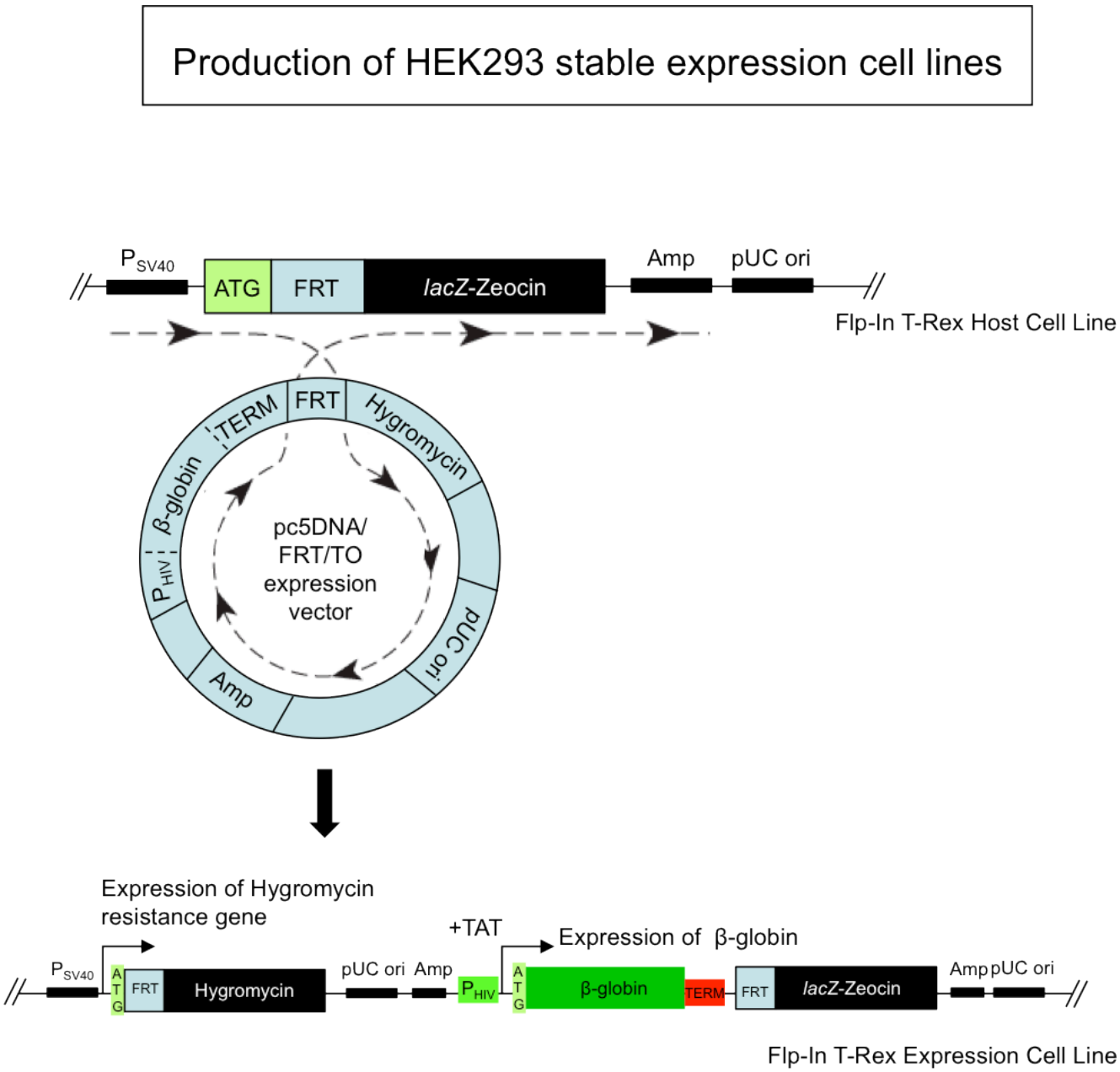


Figure 3.7: Generation of stable mammalian expression cell lines. Overview-chart depicting the integration of the expression vector into the host cell line, generating HEK293 cells stably expressing the Tat-inducible β-globin gene with the Terminator, mutated Terminator or the FII-sites, respectively (for simplicity, indicated by "TERM" as a representation for all). The β-globin gene is carried by the pc5DNA/FRT/TO expression vector, which integrates into the cell`s genomic DNA by homologous recombination between two unique FRT recombination sites, mediated by Flp-recombinase. This rare recombination event leads to disruption of the functional *lacZ*-Zeocin transcription unit caused by loss of the SV40 early promotor and the ATG initiation codon, which concurrently leads to acquisition of Hygromycin resistance. This is facilitated by the Hygromycin gene insertion downstream of the SV40 early promotor and the ATG initiation codon, the elements which before drove the *lacZ*-Zeocin gene expression. Individual colonies are therefore selected by their Hygromycin resistance. The host cell line contains a single FRT site only (verified by Southern Blot analysis by the providing company). Therefore, each Hygromycin-resistant colony, formed from a single cell, contains a single copy if the integrated, β-globin gene expressing pc5DNA/FRT/TO vector. Hygromycin resistance additionally prevents contamination of the colonies with cells, which experienced the highly unlikely event of random genomic integration of the expression vector. This is based on the fact that by FRT-site-independent, random integration the Hygromycin gene on the vector would remain without a promotor, which would leave such cells Hygromycin-sensitive.

the plasmid engages in homologous recombination with the FRT site engineered into the HEK293 genome, resulting in targeted stable integration of the β-globin gene into the cell`s genome. This approach resulted in the creation of an endogenous chromatin environment surrounding the gene of interest, while maintaining the model system requirements. These are, in this case, the Tat-inducible HIV-LTR driving expression of β-globin and the modified 3` ends of the gene. I created three cell lines in total, β-globin with its wt terminator sequence (βTERM; referred to as wild type), β-globin with the deleted terminator sequence (βΔTERM) and the CTCF-binding site containing β-globin gene, featuring the previously described FII site in a four-copy tandem, giving rise to βΔTERM-FII4x. Although the βΔTERM cell line showed expected behavior in regards to β-globin transcription termination (data not shown), valuable discussion with colleague Dr. Eleanor White in the NJP laboratory, lead me to the decision to apply her terminator mutant control stable HEK293 cell line in further experiments. This cell line displays a mutated form of the CoTC region, which has been shown to fail transcriptional termination (White et al, 2012). The reasoning behind applying this control cell line is the higher similarity in the sequence length between the new βTERMmut and the βTERM and βΔTERM-FII4x gene constructs. Results should then be more comparable and avoid potential indirect effects on transcription caused by higher sequence length variations in the terminator region.

To confirm β-globin expression from these new HEK293 cell lines, I performed Northern Blot analysis. This technique first allows quantitative comparison of

mRNA expression levels between different lines and, second it provides information about the size of the detected RNA. RNA isolated from wild type, mutant and FII-site containing stable cells before and after transcriptional activation with Tat, and also the parental HEK293 cell line, was separated on a 1% agarose formaldehyde gel, transferred to a nitrocellulose membrane and hydridized to a uniformly radiolabelled antisense RNA probe complementary to Exon 2 of the β-globin gene. The hybridized radiolabelled RNA probe was then visualized by PhosphoImager as shown in Figure 3.8A (lower band). In addition, a uniformely radiolabelled antisense RNA probe specific for the β-actin gene was used to detect β-actin mRNA as a loading control (Figure 3.8A upper band). The bands correspond to the full-length mature β-globin mRNA (626bp) confirming that the β-globin gene is expressed in all three stable cell lines (lanes 2, 4 & 6). The differences in the signal strengths show that the mutant cell line produces significantly lower levels of β-globin mRNA than the wild type. Interestingly, the FII-sites containing cell line expresses wild type levels of β-globin mRNA. In the absence of Tat, transcription at the HIV-LTR initiates at near background levels (lanes 3, 5 & 7). The lack of a signal in the parental HEK293 cell line confirms that the endogenous β-globin gene is not expressed in these cells.

I next wanted to confirm the recruitment of CTCF and Cohesin to the FII sites using ChIP analysis. The stable cell lines transfected with pTat were cross-linked with formaldehyde, chromatin was isolated as described in Materials & Methods and immuno-precipitated with the rabbit anti-CTCF and rabbit anti-Smc3 antibodies. Quantitative real time PCR, using primers along the gene

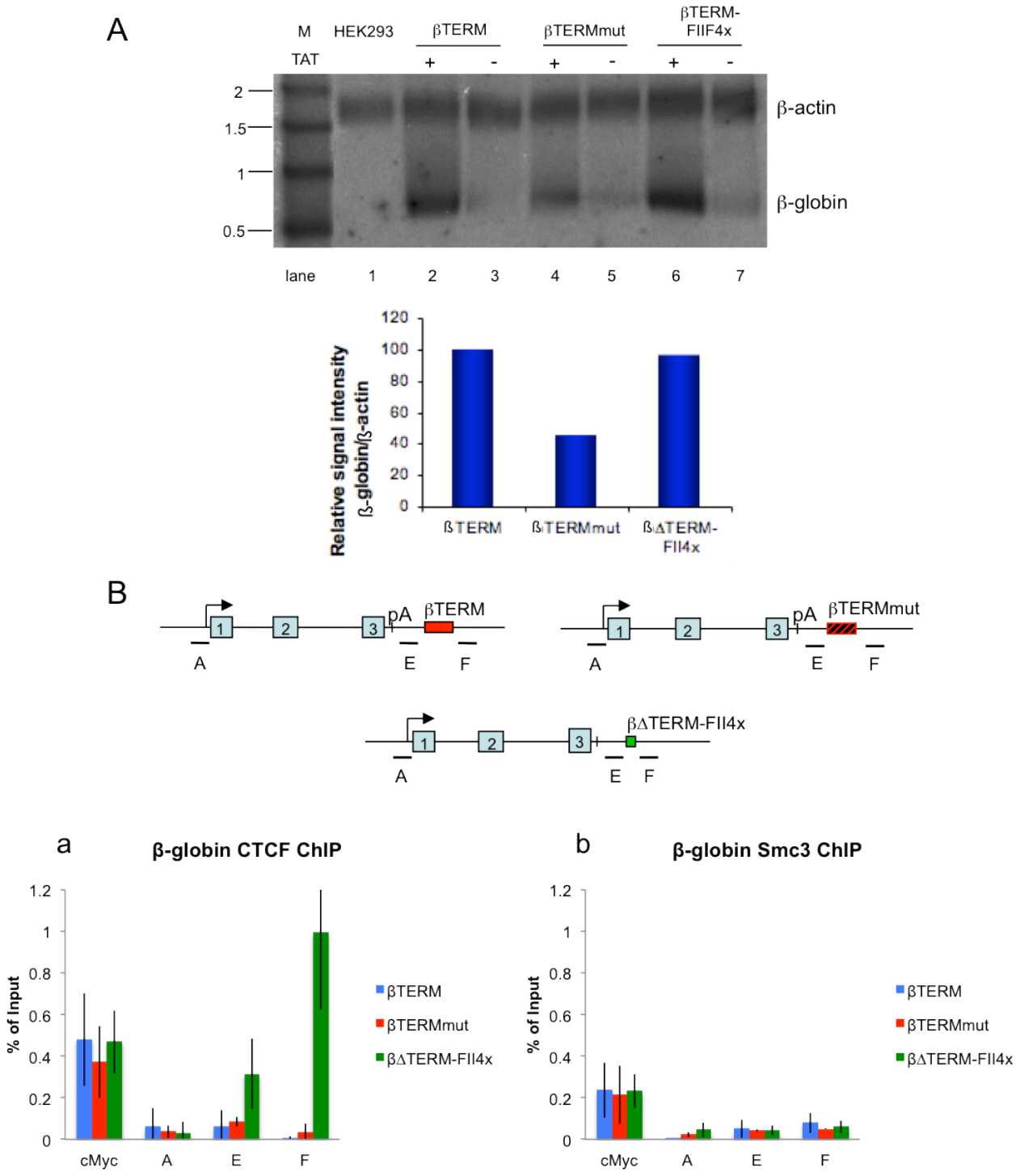


Figure 3.8: Proof for β-globin expression from stably integrated β-globin gene constructs in HEK293 cells and recruitment of CTCF to the inserted endogenous FII(CTCF) binding sites. A. Northern blot analysis of nuclear RNA isolated from HEK293 cells containing stably integrated constructs with the Terminator elements βTERM, βTERMmut and βΔTERM-FII4x and from the parental HEK293 line alone. An antisense RNA probe complementary to Exon 2 was used. β-actin was detected as a loading control. M represents the marker, lane numbers are indicated below. **B.** qPCR quantitation of ChIP analysis for CTCF and Smc3 (Cohesin) recruitment to the FII(CTCF) binding sites. *c-Myc* serves as a positive control for CTCF- and Cohesin recruitment. Values are indicated as percent of total input DNA. Diagram of the β-globin gene shows the promoter (arrow), Exons (blue, numbered boxes), Terminator, mutated Terminator and FII binding sites (red and green boxes) and the position of the primers indicated below. ChIP values are based on average values ± s.d. from three independent biological experiments.

body, indicates where significant CTCF- and Cohesin recruitment occurs. The graphs in Figure 3.8B display the position of the primer pairs along the β-globin gene. As a positive control for CTCF- and Cohesin recruitment, I used a primer pair for a CTCF-binding site in the 5` control region of the *c-Myc* gene, at position -1.97 kb upstream of the TSS. The diagrams in Figure 3.8B show positive CTCF recruitment (Figure 3.8Ba) and Cohesin (Figure 3.8Ba) recruitment to the *c-Myc* 5` regulatory region as expected, displaying similar signal levels across the three cell lines. This was used as a positive control, indicating the three cell lines were comparable. Since I expected CTCF- and Cohesin recruitment only at the 3` end of the β-globin gene, which contains the FII-sites, I used primer pair A, positioned in the HIV-LTR, as a negative control for recruitment of both proteins. As expected, in both cases only background signals were detectable at position A for all cell lines, as well as at positions E and F for the wild type and mutant cell lines. Since those cell lines don`t have any CTCF-binding sites in the β-globin 3` flank, the background signal levels confirmed the specificity of this analysis. Therefore, the significant signal increase for CTCF over probe E, and especially probe F, indicates specific CTCF positioning (Figure 3.8Ba). Interestingly, no Cohesin recruitment could be detected at any of these positions (Figure 3.8Bb). These data demonstrate that the FII sequence successfully recruits CTCF. Since E and F are in close proximity and immediately surround the actual FII site, both signals presumably reflect CTCF-recruitment to the same site. It was not possible to place a primer pair directly over the FII sequence due to low primer binding efficiency in this GC-rich region.

Having demonstrated that CTCF gets recruited to the cloned FII sites when placed in chromatin context, I next performed a quantitative RT-PCR analysis to determine, whether transcription of the β-globin gene would be successfully terminated. As before, total RNA from pTat-transfected stable HEK293 cell lines was reverse transcribed with primer rtR, followed by quantitative RT-PCR experiments with primer pair rtF/rtR to detect uncleaved pre-mRNA and primers Intr1/Ex2 for normalization purposes. The results in Figure 3.9 clearly indicate low levels of read-through transcripts in the wild type cell line and high levels of read-through transcription in the mutant cell line, confirming the expected outcome for both controls. However, the level of read-through RNA message in the FII containing cells is still significantly higher than detected in wild type cell line, though reduced compared to the mutant line. While variations in RNA stability could account for the differences in the levels of read-through RNA message between the mutant and FII-containing cells, it is obvious that transcription does not terminate as efficiently at the FII sites as compared to the wild type CoTC terminator. Therefore, I conclude that CTCF does not appear to support termination of β-globin transcription when recruited to the gene`s 3` flanking region.

3.7 Discussion

Evidence for direct involvement of the Cohesin complex in transcriptional termination has been so far provided only in yeast. As described before, in *S. pombe* Cohesin accumulates between convergently transcribed genes and leads to transcription termination in a cell cycle dependent manner (Gullerova et al, 2008). The mammalian and yeast genomes are fundamentally

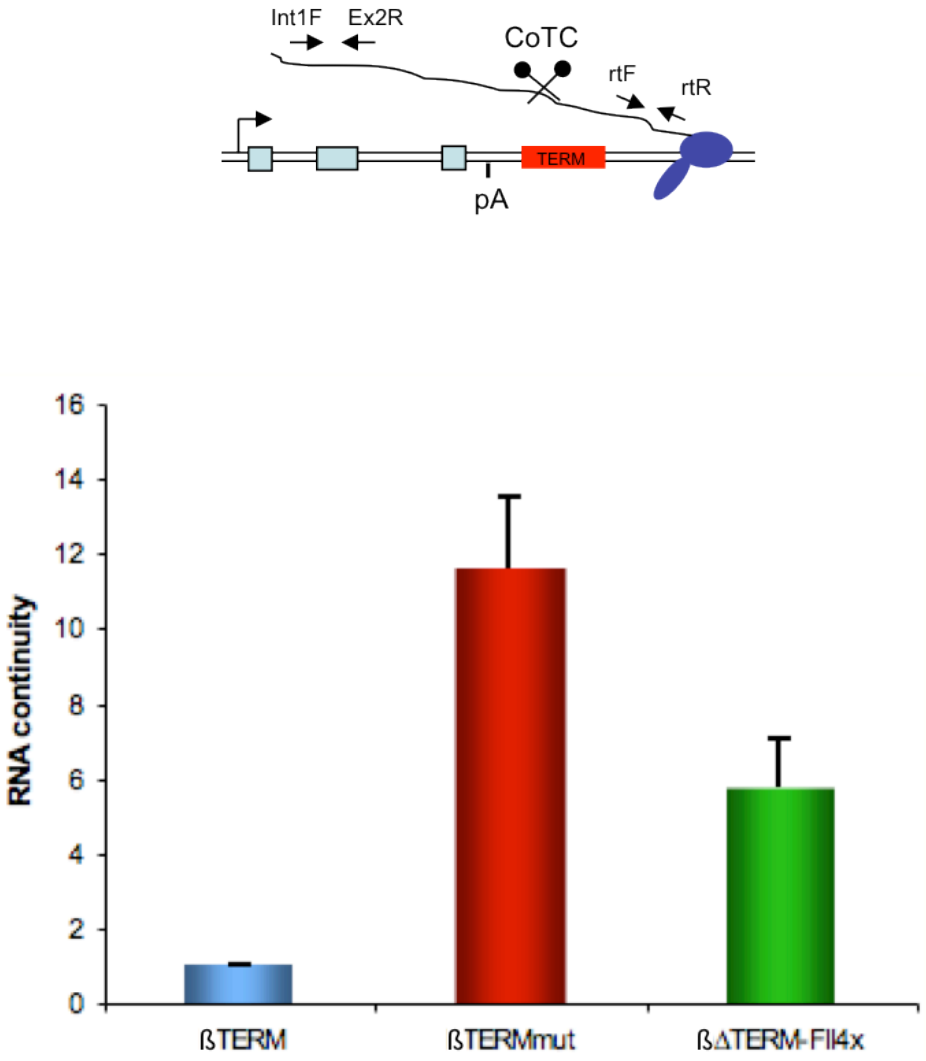


Figure 3.9: Effect of CTCF-recruitment on transcription termination of the human β-globin gene in the chromatin environment of stable HEK293 expression cell lines. qRT-PCR analysis of total RNA isolated from stable, β-globin expressing HEK293 cells transfected with pTat. The ratio of Terminator to Intron 1/Exon 2 RNA for βΔTERM-FII4x and βTERMmut was calculated and corrected to the positive control, βTERM, which is set at 1. Error bars indicating standard deviation of three biological replicates are shown. Diagram showing the position of the PCR amplicons, and primers used for RT-PCR. Primer rtR was also used for reverse transcription of RNA into cDNA. PCR product detected after the Terminator region (CoTC) is indicative of the level of transcripts that are not cleaved over the Terminator region.

different in many aspects of their structural and functional characteristics, which are mainly based on the compressed nature of the yeast genome. Furthermore, studies revealing the genome-wide distribution of Cohesin in mammalian cells detected a large proportion of Cohesin to be positioned at intergenic regions (Kagey et al, 2010; Wendt et al, 2008), but did not obtain any evidence for Cohesin accumulating at specific hot spots. Therefore, it was not feasible to study Cohesin-involvement in mammalian transcription termination in the same context as has been done in *S. pombe*. As described before in detail (see Chapter 1), mammalian Cohesin has been often reported to be recruited to CTCF-binding sites for functional purposes. To assess, whether Cohesin is able to terminate transcription in a mammalian gene context, I made use of the established β-globin gene construct system. I attempted several approaches to investigate CTCF- and Cohesin-recruitment to the cloned CTCF-binding sites in the 3`end of the plasmid encoded β-globin gene. In total, I used two approaches to find the right sequence requirements to ensure CTCF-binding to the β-globin 3` flanking region. First, the minimum CTCF core binding sequence (Figure 3.1) and later, the endogenous regulatory chicken β-globin FII sequence (Figure 3.4). Both approaches resulted in inefficient transcription termination of the β-globin gene in my model system (Figure 3.2A, C; Figure 3.5A, C). Applying nascent RNA analysis, as well as steady state RNA analysis, both confirmed that the transcribing Pol II moved beyond the point of the introduced CTCF-binding sites. In comparison to the co-transcriptionally cleaved wild type β-globin terminator, no CoTC occurred at the cloned CTCF-binding sites. There was no significant difference at the level of read-through message produced by

cells transfected with the deleted terminator β-globin construct (βΔTERM) or the CTCF-binding site containing constructs βΔTERM-CTCF or βΔTERM –FII.

Attempts to identify a potential cell cycle dependency for Cohesin-mediated transcription termination, as was described in *S. pombe*, were unsuccessful. The most efficient cell cycle synchronization protocol I tested for transfected HeLa cells made use of the cdk-inhibitor Roscovitine. However, application of this drug appeared to have a significant impact on cellular transcription and resulted in irreproducibility of my data (Figures 3.3 and 3.6).

To circumvent the technical difficulties I experienced using a transient transfection system and to test, whether a physiological chromatin context is required to achieve successful CTCF- and Cohesin-recruitment to the β-globin 3`end, I moved on to create a stable HEK293 β-globin expression system (Figure 3.7). Keeping the same β-globin promoter and 3` end modifications as in the plasmid system used before, I was first, able to confirm β-globin gene expression upon Tat-induction (Figure 3.8A) and second, provide evidence for successful recruitment of CTCF to the FII sites in the according cell line (Figure 3.8Ba). Following the results in Figure 3.8Bb, Smc3 (Cohesin) signals across the gene body including the FII sites remained mostly at background levels. Therefore, I must conclude that CTCF does not recruit Cohesin in this given β-globin gene context.

So far, the main conclusion derived from my analyses is that CTCF-recruitment to endogenous CTCF-binding FII sites alone does not prevent read-through transcription of the β-globin gene. This means CTCF does not

appear to act as a physical barrier to the transcribing polymerase. However, closer analysis of transcription over the FII site in stably β-globin expressing HEK293 cells revealed indications of possibly CTCF-mediated effects on β-globin mRNA levels and therefore moved this project on from basic questions regarding transcription termination, to questions about transcription kinetics, which may act to influence RNA stability. These results will be discussed in the following chapter.

CHAPTER 4

CTCF MEDIATED POL II PAUSING AT THE β-GLOBIN 3` FLANKING REGION RESCUES β-GLOBIN mRNA LEVELS IN THE ABSENCE OF THE WILD TYPE β-GLOBIN TERMINATOR

4.1 Introduction

As mentioned before, transcriptional termination of Pol II requires a functional poly(A) signal and, often, a downstream terminator sequence. Termination is triggered following recognition of the poly(A) signal by Pol II and subsequent pre-mRNA cleavage, occurring either at the poly(A) site or within the transcripts from the terminator element. It has been shown that the efficiency of transcriptional termination has significant influence on the levels of mRNA produced, which in turn may affect the levels of the resulting protein (West et al, 2009).

So far, I examined the hypothetical idea of Cohesin and/or CTCF involvement in transcription termination of the β-globin gene, when recruited to the gene`s 3` end. Based on my results in stable HEK293 cells, I came to the conclusion, that despite CTCF recruitment to the β-globin 3` end and despite the chromatin environment, transcription was not terminated efficiently. However, one result appeared striking. Comparing the levels of detected full length β-globin mRNA by Northern blot analysis across the HEK293 expression cell lines (Figure 3.8A), it became apparent that the strength of the bands in lane 2 and 6 was equal. Transcription termination influences the amounts of produced mRNA. Impaired termination may result in lower mRNA- and consequently lower protein-levels (West et al, 2009). Regardless of my result demonstrating that CTCF recruitment to the β-globin 3` end did not terminate transcription per se, the Northern Blot analysis proved the β-globin mRNA expression levels of βΔTERM-FII4x identical to the wild type (βTERM) and double compared to the levels detected for βTERMmut. This result moved the

project on from analysing transcription termination alone to investigating on how CTCF-recruitment to the β-globin 3` end can possibly enhance mRNA expression.

4.2 CTCF recruitment to the β-globin 3` flanking region causes Pol II pausing

Since the β-globin gene 3` flank is the only position where the three stable expression cell lines differ, I attempted to focus on transcription events occurring directly there. Therefore, I performed another ChIP experiment to obtain a Pol II profile across the β-globin gene to possibly provide evidence for a direct influence of CTCF-recruitment on the elongating polymerase levels. The ChIP result in Figure 4.1A shows the level of Pol II at each of the tested positions along the β-globin gene. Primer pair positions are indicated from A to F in the diagram above. At first, Pol II follows a rather typical distribution across the gene body, with higher presence at the promoter and within the first exon, decreasing downstream and, increasing slightly again around the poly(A) site. It is worth mentioning that although the obtained Pol II profile meets a generally acquired pattern for Pol II-transcribed genes, the differences in Pol II occupation between the beginning and the downstream body of the gene are more dramatic than usual. This may be explicable by the use of the HIV-LTR promoter to drive β-globin gene transcription. Originating from a virus, the HIV-LTR is a particularly strong promoter and it is reasonable to expect impact on Pol II kinetics.

The interesting result obtained from this ChIP profile is the notably increased

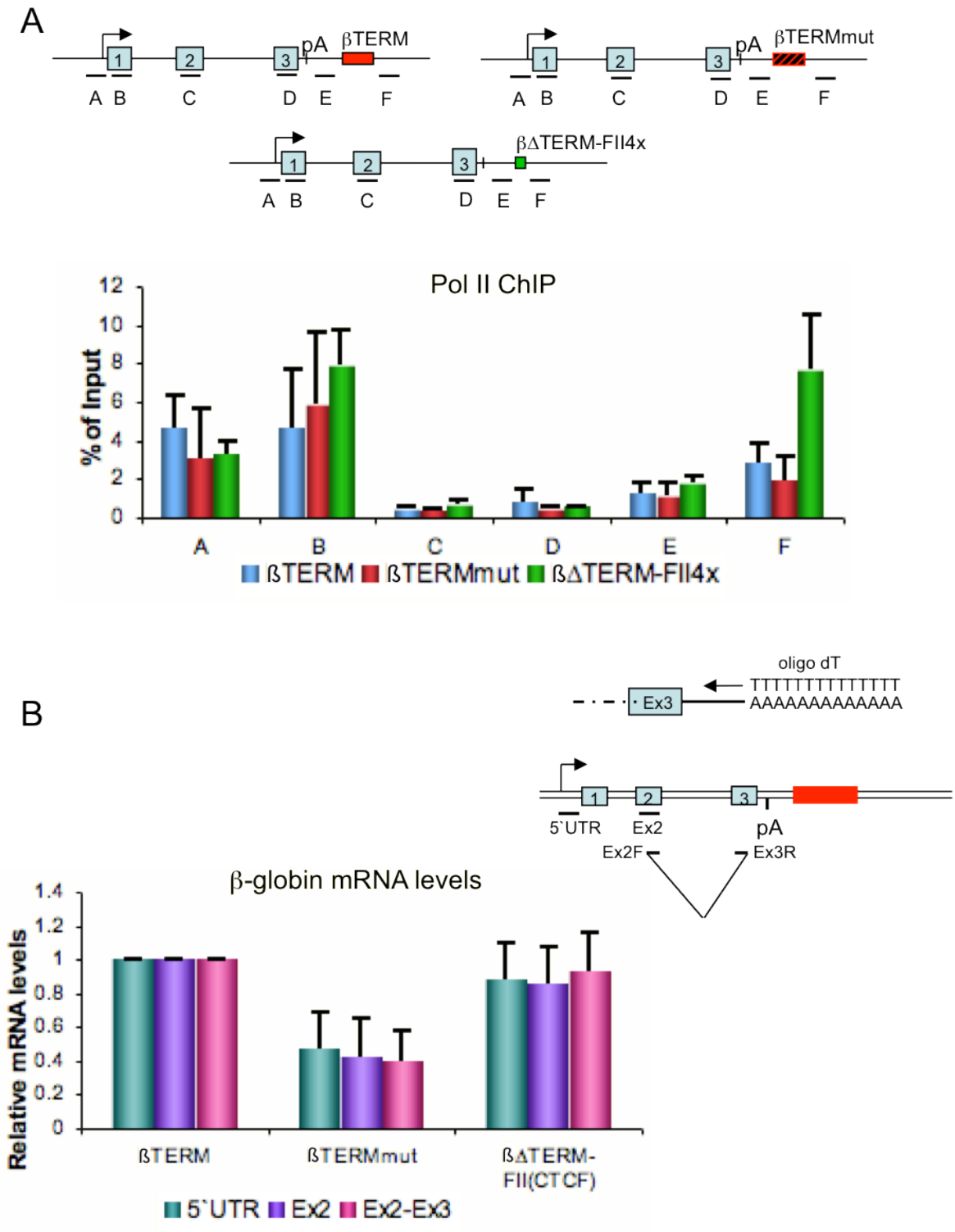


Figure 4.1: CTCF recruitment causes Pol II pausing at FII sites in stably β-globin expressing HEK293 cells. A. qPCR quantitation of ChIP analysis of Pol II along the β-globin gene. Values are indicated as percent of total input DNA. Diagram of the β-globin gene shows the promoter (arrow), Exons (blue boxes), Terminator, mutated Terminator and FII sites (red and green boxes) with the position of the primers indicated below. ChIP values are based on average values ± s.d. from three independent biological experiments. **B.** qRT-PCR analysis of total RNA isolated from stable HEK293 cells transiently transfected with pTat. Levels of mRNA have been measured by reverse transcription with oligo dT primers with subsequent qPCR for three different positions within the β-globin transcription unit. Values have been normalized to GAPDH mRNA levels in each sample. Error bars indicating standard deviation over multiple biological replicates are shown. Diagrams showing the position of the primers used for reverse transcription (upper) and RT-PCR (lower).

level of Pol II at, or, immediately after the CTCF-recruiting FII sites. Compared to the wild type and the mutant terminator, I measured an increase in Pol II accumulation of a minimum 3 fold (Figure 4.1A, primer position F, green bar). Increased Pol II levels at position F correspond exactly to the position of detected CTCF-recruitment (Figure 3.8B, green bar over probe F). However, for the βTERMmut control cell line, expectations would have suggested the detection of a higher Pol II ChIP signal over probe F, since the polymerase has been shown not to terminate at the mutated terminator. The low signal is surprising but might find an explanation in the dynamics of polymerase engagement in different processes at the 3` flanking region of the β-globin gene construct in each of the three cell lines. Probe F is, as indicated in the diagram, positioned closely after the terminator or FII-sites, respectively. In case of the wild type terminator, the signal increases a bit from probe E to probe F, suggesting the detection of Pol II, which is most probably engaged in the CoTC process. At the same positions, E to F, the signal also slightly increased at the mutated terminator, however, remained lower than in βTERM cells. While this did not meet the immediate expectation of an increased Pol II signal due to read-through transcription, the following, alternative suggestion might explain the happenings at position F in βTERMmut cells: As the polymerases reach the mutated terminator region, no CoTC occurs, which leads the polymerases on to extend the elongation process. The signal at probe F therefore appears similar to the signals detected over the other probes located in (probes C and D) and just after (probe E) the β-globin gene body. Therefore I suggest that if the Pol II complex does not engage in the termination process, elongation continues to a similar extend as it did before it

reached the terminator region. The involvement in the termination process might reduce Pol II dynamics and cause a slight retention, which therefore could lead to the increased Pol II ChIP-signal at the terminator in βTERM cells. In case of the CTCF-recruiting FII sites in βΔTERM-FII4x cells, Pol II appeared to accumulate as if a physical obstacle hindered elongation. Given the fact that I detected CTCF-recruitment at probe F, this could explain the sudden increase of Pol II in the βΔTERM-FII4x cell line. However, more clarity would be given by including more probes downstream of probe F. Extension of the Pol II profile downstream of the terminator or FII-sites would facilitate conclusions about possible Pol II engagements at the terminator or CTCF-binding sites.

To confirm the β-globin mRNA levels detected by Northern Blot, I performed qRT-PCR on isolated total RNA, reverse transcribed with oligo dT primers. Positions of the primer pairs applied in qRT-PCR are indicated in the diagram in Figure 4.1B. Consistent with the Northern Blot analysis, the levels of fully processed β-globin mRNA were almost identical in the βTERM and βΔTERM-FII4x cell lines, while the cell line βTERMmut expressed only half as much β-globin mRNA. So far, and under consideration of my previous explanation for the differences in Pol II ChIP signal at probe F across the three cell lines (Figure 4.1A), I conclude that CTCF-recruitment to the FII sites in the β-globin 3` flanking region possibly leads to Pol II accumulation. Nevertheless, read-through transcripts were detected in βΔTERM-FII4x cells (as shown in Figure 3.9), despite CTCF-recruitment to the FII sites (Figure 3.8Ba). Nevertheless, β-globin mRNA production in βΔTERM-FII4x cells appeared rescued to wild type levels (Figure 4.1B). The mechanism behind this phenomenon remains

to be elucidated.

4.3 CTCF-recruitment to the β-globin 3` flanking region itself is responsible for rescued β-globin mRNA levels in βΔTERM-FII4x cells.

So far, I detected CTCF-recruitment and Pol II accumulation at the CTCF-binding motifs and similar β-globin mRNA levels in βΔTERM-FII4x cells as compared to the wild type, βTERM. However, that alone did not prove Pol II pausing at the CTCF-binding sites was a direct consequence of CTCF-recruitment itself, which then rescued β-globin mRNA levels. In order to define, whether CTCF-recruitment or the nature of the CTCF-binding motif itself lead to the observed effects on Pol II occupation and mRNA production, I performed CTCF knock-down experiments, followed by quantitative RNA analysis. Stable HEK293 cell lines were RNAi-treated and subsequently transfected with pTAT. Total RNA was then isolated and reverse transcribed with oligo dT primers, followed by qRT-PCR to confirm CTCF knock-down efficiency and to measure β-globin mRNA levels. Figure 4.2A, upper panel, shows the achieved CTCF knock- down efficiency by comparing CTCF mRNA levels in CTCF depleted cells (siCTCF) and control knock-down cells (siLuc, siRNA against Luciferase). CTCF mRNA levels decreased to approximately 30-35%, resulting in knock-down efficiencies of about 65-70%, measured on RNA-level. Western blot analysis was performed to control for the effect of CTCF depletion at a protein level. The lower panel of Figure 4.2A displays the resulting protein bands of protein extracts from siLuc and siCTCF treated

cells. The upper lane shows the amount of CTCF protein in control and knock-down HEK293 cells, the lower lane displays GAPDH protein levels, which served as a loading control, to which the CTCF-signals were normalized to. Bands were quantified for their signal strength, showing that the knock-down efficiency on protein level was about the same as on mRNA level, namely about 70%. RNAi treated cells were then subjected to β-globin mRNA expression analysis. The result in Figure 4.2B shows that while β-globin mRNA levels remain unaffected by CTCF knock-down in the control cell lines, β-globin mRNA levels decrease in βΔTERM-FII4x cells. Although expected much stronger, the measured reduction is rather modest. However, since the CTCF-recruiting FII site is known to be very efficient in CTCF binding, it is likely that the remaining CTCF after the knock-down is, to a certain extend, still efficiently recruited to the FII sites. Nevertheless, the effect on β-globin mRNA levels appears significant as the controls remained entirely unaffected by the knock-down.

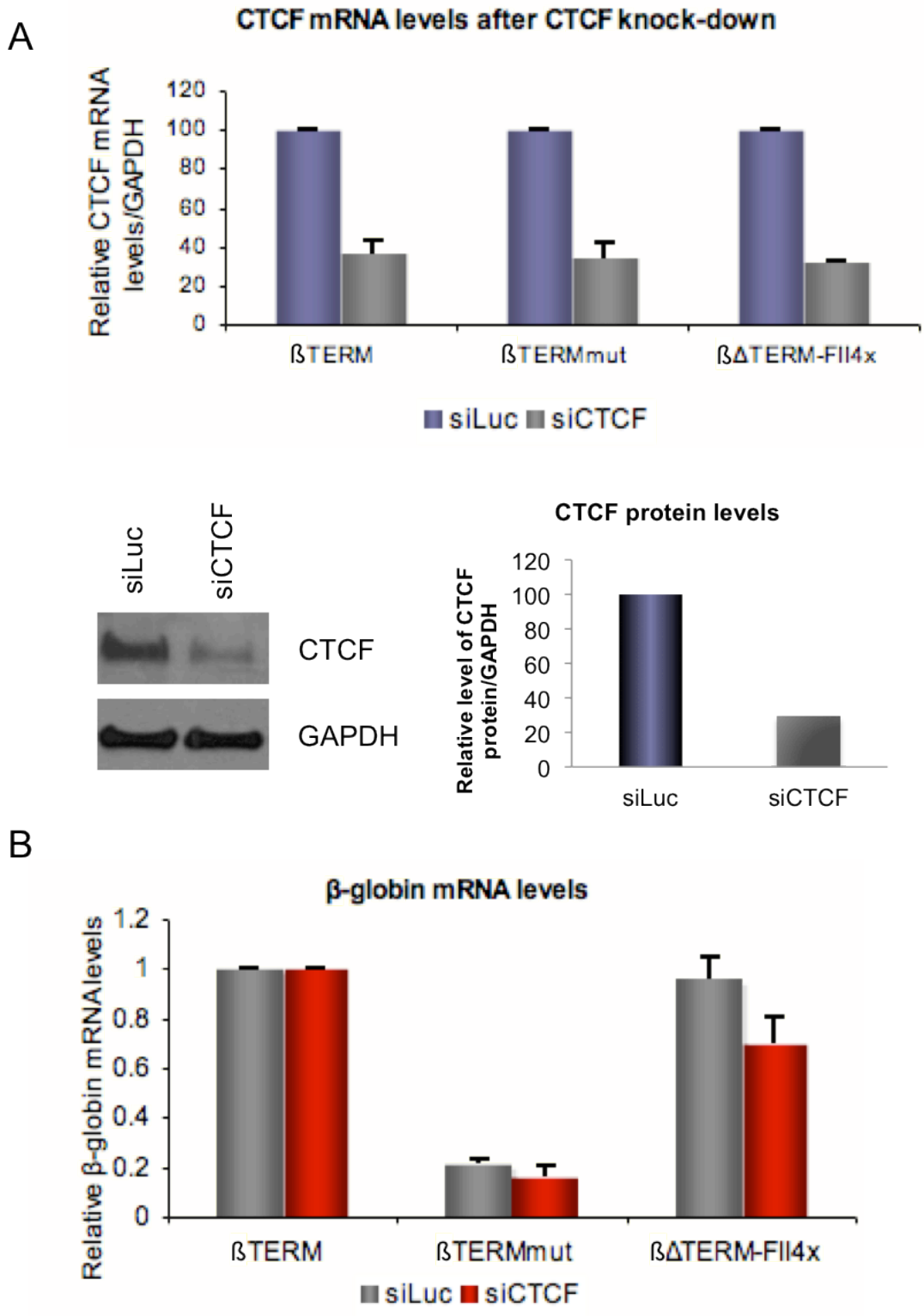


Figure 4.2 CTCF knock-down results in reduced β-globin mRNA levels in the βΔTERM-FII4x stable HEK293 cell line. A. RNAi-mediated depletion of CTCF. Upper panel: qRT-PCR analysis of siLuc (control) and siCTCF-treated stable HEK293 cells. Bars represent average values ± s.d. from three independent biological experiments. GAPDH mRNA was used as a control for normalization. The amount of CTCF mRNA in siLuc-treated cells was taken as 1. Below, left panel: representative western blot analysis of protein extracts from siLuc-treated and siCTCF-treated stable HEK293 cells using anti-CTCF antibody. GAPDH protein was used as a loading control. Diagram on the right shows ImageJ-based quantitation of signals obtained by western blot. CTCF signal values are normalized to those obtained from GAPDH. **B.** qRT-PCR analysis of β-globin mRNA levels in control cells (siLuc) and in siCTCF treated cells. Bars represent average values ± s.d. from three independent biological experiments. CTCF mRNA levels are normalized to GAPDH mRNA, respectively. Vales obtained from control siRNA- and siCTCF-treated wild type cells βTERM are set to 1.

4.4 Discussion

Summarizing the results obtained from the stable β-globin expression HEK293 cell lines, CTCF recruitment is detected at the introduced CTCF-binding sites in the β-globin 3` flanking region of βΔTERM-FII4x cells (Figure 3.8B). This potentially causes accumulation of elongating Pol II at the site of CTCF-binding and appears to be linked to rescued β-globin mRNA levels to wild type conditions, despite measured read-through transcription in βΔTERM-FII4x cells (Figures 4.1A and 3.9). The effect on mRNA levels appears to be linked directly to CTCF-recruitment to the FII sites as CTCF knock-down experiments resulted in reduced β-globin mRNA levels in βΔTERM-FII4x cells, while the control cell lines βTERM and βTERMmut, remained unaffected (Figure 4.2B). The observation of low Pol II levels at the terminator region in βTERMmut cells (Figure 4.1A, probe F) appeared, as discussed before, unexpected. However, I previously provided a possible explanation for this observation. The following discussion takes my interpretation into account and therefore, considers the increased Pol II level at the FII sites in βΔTERM-FII4x cells as a potential consequence of CTCF recruitment. Accumulation of Pol II in close proximity to the β-globin 3` end could then provide an explanation for the effects on β-globin mRNA levels, as discussed in the following section.

It has been shown that impaired transcriptional termination leads to reduced amounts of mRNA from the affected gene. This is caused by reduction of RNA stability due to impaired pre-mRNA 3` end processing and also, by failure to recycle polymerase for new transcription cycles (West et al, 2009).

Such an effect is expected to occur in the termination-deficient βTERMmut, as is indicated by the significantly reduced mRNA amounts (Figure 4.1). However, although CTCF-recruitment to the β-globin 3` flanking region did not prevent transcriptional read-through in βΔTERM-FII4x cells, pre-mRNA 3` end processing appeared to occur successfully, resulting in wild type β-globin mRNA levels. β-globin transcription termination requires a terminator sequence, named CoTC element, at which the transcript is cleaved prior to cleavage at the poly(A) site. Processing then occurs at the poly(A) site, most probably while the pre-mRNA is still in contact with the template-detached Pol II (Figure 4.3A). In the case of a deleted or mutated CoTC-element, cleavage at the poly(A) site does not seem to occur efficiently as the transcribing polymerase passed the point of termination. However, when CTCF-binding sites replace the CoTC-element, recruited CTCF appears to act as a barrier for the elongating Pol II. Even though this does not prevent transcriptional read-through, it does appear to keep the polymerase temporarily in proximity to the β-globin poly(A) signal region (Figure 4.3B). The reduced read-through transcription rate in βΔTERM-FII4x possibly results from the fact that as Pol II is paused, fewer polymerases continue elongation, which leads to lower detectable read-through RNA message. It cannot be excluded that a certain proportion of paused Pol II does eventually disengage from the DNA template. Although the majority of Pol II appears to escape the CTCF-barrier, Pol II pausing nevertheless may provide a time window to allow cleavage at the poly(A) site and further 3` end processing to occur. When and how the paused polymerase then escapes to continue elongation, remains to be elucidated. Further experiments would provide an answer to whether Pol II

does terminate at some later point, possibly involving the 5`-3`exonuclease Xrn2 in Pol II template release (Figure 4.3Ba), or whether Pol II remains associated with the DNA (Figure 4.3Bb).

Supportive evidence for Pol II accumulation at CTCF-bound sites is provided by Shukla et al., who showed that CTCF-recruitment to its binding site in Exon 5 of *CD4* leads to Pol II pausing, which influences pre-mRNA splicing (Shukla et al, 2011). It was shown that CTCF can autonomously promote Pol II transient spatio-temporal pausing but not complete arrest. CTCF appears to act as a direct impediment to transcription, independently of any particular nucleosome structure or chromatin context, as has been demonstrated on a naked DNA template. Paused Pol II at the CTCF-binding site in *CD4 has* moreover been shown to be Ser2-phosphorylated, a sign of elongating polymerase (Shukla et al, 2011). This provides further support that CTCF does not terminate transcription on an endogenous level.

In the β-globin model system used, the β-globin gene is driven by the HIV-LTR, which is a very strong viral promoter. However, it would now be interesting to investigate how CTCF-mediated Pol II pausing would influence a gene with its endogenous promoter. Therefore, it would be helpful to create new stable HEK293 expression cell lines with a terminator-dependent, inducible gene driven by its endogenous promoter, bearing the same modifications in the terminator region. This would confirm, whether the mRNA expression pattern across such gene constructions is similar to the pattern I obtained for the HIV-LTR driven β-globin gene.

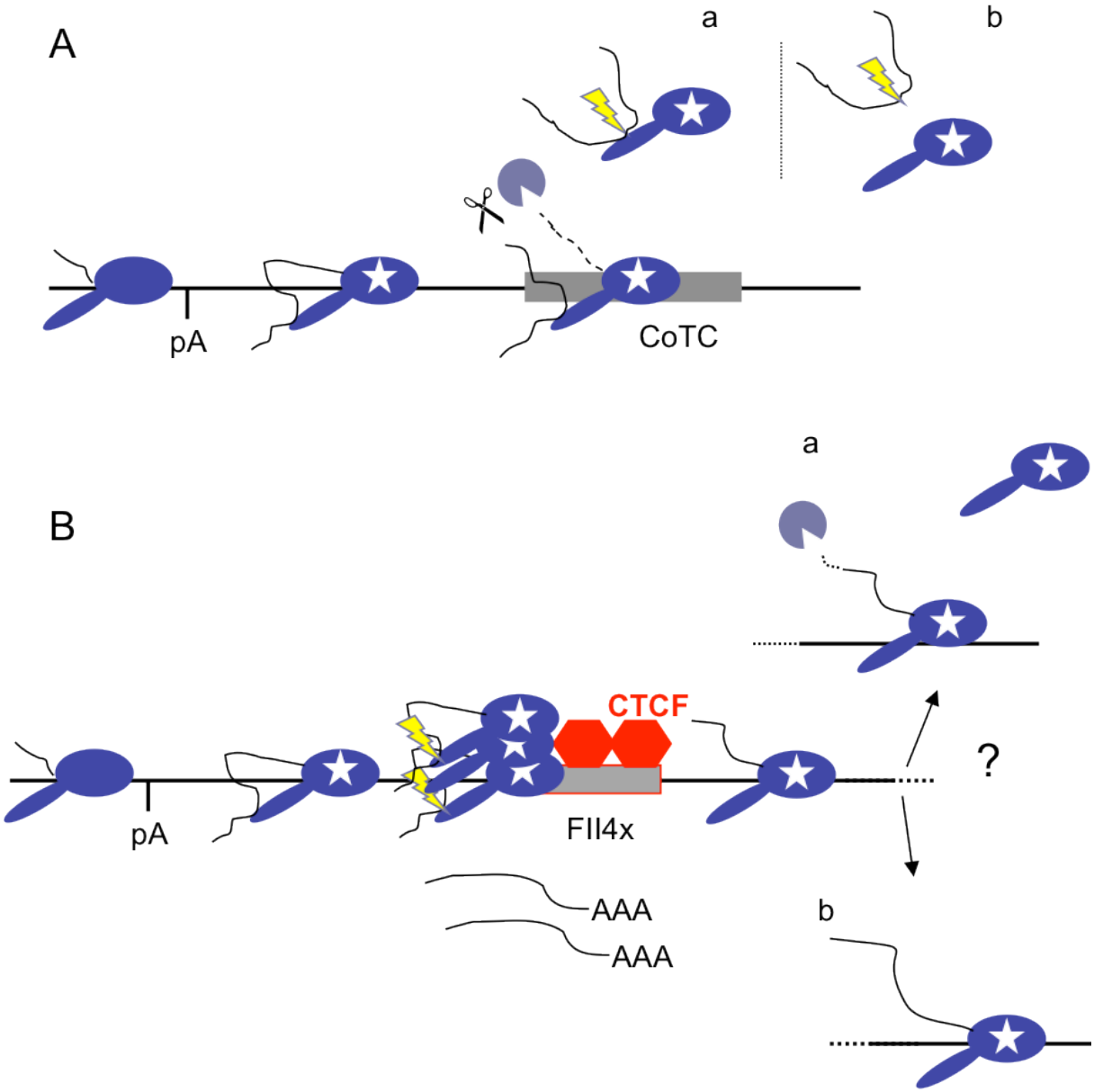


Figure 4.3: Sequence of events in Pol II Transcription Termination of β-globin featuring the CoTC element or CTCF-binding motifs, respectively.
(A) The human β-globin terminator: transcription of the poly(A) signal (pA) switches Pol II (blue icons) to a termination prone form (white star). CoTC of the terminator region RNA transcript, indicated by the scissors, is followed by degradation of the Pol II-associated pre-mRNA (dashed line), which leads to template release. Poly(A) site cleavage (lightning bolt) occurs on Pol II that is dissociated from the DNA template **(A-a)**. Eventually, cleavage of some released pre-mRNA occurs off Pol II **(A-b)**. **(B)** β-globin CTCF-binding FII sites replacing the terminator: transcription of the poly(A) signal switches Pol II to a termination prone form. Recruited CTCF (red hexagons) interacts with its binding motifs in the FII4x element (red framed grey box) and forms a barrier to the elongating Pol II. Pol II accumulates at the sites of CTCF-recruitment, which results in Poly(A) site cleavage and production of polyadenylated mRNA. All or part of the accumulated Pol II escape the CTCF barrier and continue elongation. Transcription then continues until degradation of the Pol II-associated pre-mRNA (dashed line) causes Pol II template release. **(B-a)**. Alternatively, Pol II remains attached with the template resulting in continued elongation **(B-b)**.

Another point to consider in future studies is the number of CTCF-binding sites cloned into the 3` flanking region of a gene. While in first place it appeared a logical thought to achieve more CTCF-recruitment by cloning several CTCF-binding motifs in tandem, multiple CTCF-binding events were eventually not supported by an increased number of binding sites. CTCF has been characterized by an unusually extensive DNase I footprint (51bp) when bound to its DNA site (Bell AC et al, 1999). Even though the endogenous CTCF-motif-containing β-globin FII site, which I applied in my studies, is 45pb in length and successfully recruits CTCF, it is quite likely that four FII sites in tandem do not provide enough space for multiple CTCF proteins to bind stably. It is therefore likely that the entire FII4x sequence bound only one, maximum two CTCF proteins, instead of the expected four. CTCF-binding motifs, which recruit CTCF or also Cohesin, occur mostly in single copies within endogenous regulatory regions, implying that one CTCF binding site is expected to be sufficient to obtain CTCF-mediated regulatory effects.

In summary, I have provided evidence that CTCF-recruitment to the 3` flanking region of a CoTC-depleted β-globin gene does not prevent read-through transcription. Nevertheless, Pol II accumulation at the sites of CTCF-binding resulted in rescued β-globin mRNA levels in βΔTERM-FII4x cells. This could be explained by CTCF-mediated alterations of Pol II kinetics, which provides a time window for pre-mRNA 3` end processing to occur, given the Pol II elongation complex passes a functional poly(A) site and pauses in close proximity of the poly(A) site. CTCF knock-down experiments provide an indication that the rescued β-globin mRNA levels in βΔTERM-FII4x cells are

most probably CTCF-dependent rather than being an indirect effect of the repetitive, GC-rich CTCF-binding motif itself. It further remains to be elucidated, whether, when and how Pol II eventually disengages from the DNA template following escape from CTCF-mediated pausing.

These studies provide some insight into how Pol II kinetics and the basic requirements for efficient pre-mRNA 3` processing and stable mRNA production may be interconnected. Based on the terminator-element-dependent β-globin gene, it appears that failure to produce mature β-globin mRNA efficiently due to termination deficiency, can be circumvented. This seems possible, when the Pol II complex is forced to reside in close proximity to the β-globin poly(A) site for long enough, possibly in order for cleavage at the poly(A) site to occur. This is then followed by 3` end processing and gives rise to stable mRNA. CTCF-mediated Pol II pausing may provide a time window for pre-mRNA processing to occur, despite the lack of actual transcription termination at the CoTC-element. This emphasizes the importance of the Pol II-poly(A) site proximity and the requirement of sufficient time for cleavage and processing. Whether poly(A) site cleavage occurs at the template, as is the case for genes, which do not require a CoTC-terminator element, or whether cleavage occurs off the template, appears secondary for mRNA production. In the HIV-LTR-β-globin system employed in my studies it is apparent that the expected failure to recycle Pol II efficiently upon dysfunctional termination, does not appear to influence β-globin mRNA levels in βΔTERM-FII4x cells. How Pol II recycling would affect the mRNA levels of a gene driven by its endogenous promoter needs to be addressed next.

CHAPTER 5

STUDY ON THE INVOLVEMENT OF CTCF AND COHESIN IN *TNFα* TRANSCRIPTIONAL REGULATION

5.1 Introduction

The genome-wide distribution of CTCF binding sites in human cells has been determined by ChIP-sequencing analysis. This shows that the majority of CTCF-recruitment occurs within introns (39%) and intergenic regions (43%; according to (Wendt et al, 2008). Only about 2% of CTCF-binding motifs were found to be located in proximity to the 3`end of genes. In the previous chapters I have described my investigations on how CTCF-recruitment to the β-globin gene 3`end influences transcription termination and mRNA stability. All these studies were, however, carried out in a model gene system, designed to allow mechanistic analysis within a well-defined environment. As a next step, I wanted to investigate on how CTCF-recruitment to the 3`end of an endogenous gene might influence its expression while it is located in its physiological environment. Finding an appropriate candidate gene proved challenging. Finally, I selected the end Tumor Necrosis Factor *α* (*TNFα*) gene as it appeared to meet many of the necessary criteria for my studies. First, it has a prominent CTCF-binding site within 2 kb distance from its 3`end and second, it is inducible and highly expressed upon stimulation. Analysis of the *TNF* locus in the UCSC genome browser shows that the CTCF-binding site is located in the final Exon (Exon 4) of Lymphotoxin B (*LTB*), a gene oriented convergently to *TNFα* (Figure 5.1). *LTB,* as *TNFα* and LTA, is member of the *TNF* superfamily forming the *TNF* locus within the MHC class III region. A recent study shows that CTCF recruitment to *LTB* Exon 4 contributes to *LTB* expression regulation in Jurkat cells (T-cell line) (Wicks et al, 2011). In the course of that study, 3C analysis revealed that CTCF-mediated DNA looping

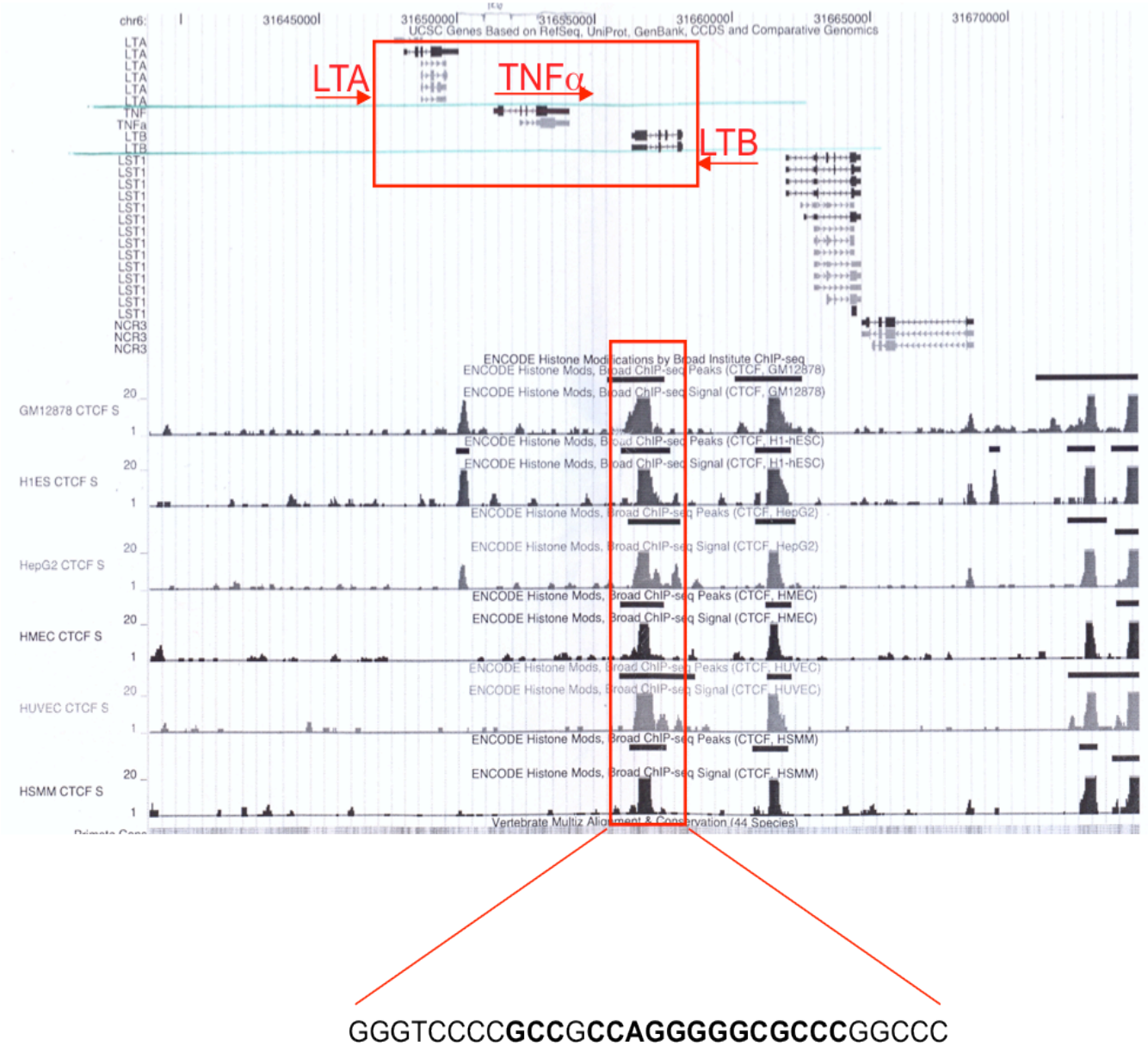


Figure 5.1: Snapshot image from the UCSC Genome Browser highlighting a cluster of genes encoding members of the TNF superfamily found at chromosome 6p21, comprising TNFα (Tumor Necrosis Factor α), LTA (lymphotoxin-α) and LTB (lymphotoxin-β). Arrows indicate the direction of transcription. Highlighted peaks indicate the only predicted CTCF-binding site found within this cluster among different cell lines. Below the diagram is the sequence of the CTCF-binding motif with the core of the binding sites highlighted in bold letters.

occurs between the CTCF site and the *LTB* promoter, LTA promoter and most interestingly, the *TNFα* 3`UTR.

I elected to investigate whether CTCF-binding to *LTB* Exon 4 also leads to Cohesin-recruitment to that site and whether that would have direct influence on *TNFα* transcription or regulate mRNA levels or fate. *TNFα* is a cytokine, which plays a critical role in systemic inflammation and host defense. However, persistent or inappropriately high *TNFα* expression leads to a number of autoimmune and inflammatory diseases reviewed in (Bradley, 2008). Essential genes like *TNFα* generally possess several levels of regulation to ensure a fine-tuned and fast response. *TNFα* is produced mainly by macrophages, although expression is also detectable in other cell types like CD4+ lymphocytes. In unstimulated cells, *TNFα* transcription is already active at basic levels. However, stimulation causes a significant increase in *TNFα* transcription and leads to protein secretion. An interesting way to fine-tune expression after mRNA production is the introduction of a limiting control step before the mRNA is translated. It was reported that *TNFα* mRNA from uninduced Macrophages failed to associate with polysomes, which resulted in undetectable *TNFα* protein production (Crawford et al, 1997). However, after stimulation of the cells with LPS, this study measured an increase in *TNFα* transcription itself but most strikingly, all the dormant mRNA located in the cytoplasm of uninduced cells became translationally active. The mechanism behind this phenomenon lies in a variation of the poly(A) tail length. While translationally silent mRNA proved to have only a limited poly(A) tail of about 25 nucleotides (hypoadenylated), LPS induction resulted in a poly(A) tail

increase by about 200 nucleotides (Crawford et al, 1997). Interestingly, the hypoadenylated *TNFα* mRNA species is stable in the cytoplasm, indicating that the inducible increase in poly(A) tail length could be part of a mechanism enabling a fast response to stimulation.

A further report showed that LPS-stimulation of Macrophages increases the expression levels of the 64-kDa Cleavage Stimulatory Factor (CstF64) and induces alternative poly(A)-site selection of several genes (Shell et al, 2005). Following the reports on CTCF-recruitment to *LTB* Exon 4 causing DNA looping to *TNFα* 3`UTR, *TNFα* poly(A)-tail variations and LPS-induced alterations in the levels of 3` end processing factors, I was intrigued to investigate, whether CTCF and/or Cohesin would play any regulatory role in such LPS-induced alterations.

5.2 Characterization of *TNFα* expression in human THP-1 cells

I began my investigations by determining the optimal system to address *TNFα* gene regulation. *TNFα* is expressed mainly in macrophages, however, mRNA is also detectable in induced monocytes, the precursor cells to differentiated macrophages. In the first instance, I decided to employ the Human acute monocytic leukemia cell line THP-1. The advantage of using a cell line over human primary cells is principally a constant supply of cells as well as a stable genetic background, which might allow increased reproducibility of experimental results.

The first step was to establish TNFα mRNA expression in induced THP-1 cells. Therefore, I isolated total RNA from LPS-induced (100ng/ml, different time points) and uninduced THP-1 cells and performed reverse transcription with oligo dT primers, followed by qPCR with primers within *TNFα* Exon 4. The results in Figure 5.2A (blue line) show, that TNFα mRNA expression is highest 1h post induction with LPS. The LPS-induction protocol applied corresponds to standard conditions and the measured TNFα mRNA levels are consistent with the generally obtained relative TNFα response. According to the literature *LTB*, is expressed in cells of the lymphoid linage, therefore it should not be detectable in THP-1 cells. This was confirmed by the background signals in Figure 5.2A (red line). Levels of TNFα and *LTB* mRNA are presented relative to the respective mRNA levels of the constitutively expressed housekeeping gene GAPDH.

The length of the *TNFα* mRNA (without poly(A)-tail) in LPS-stimulated and unstimulated mouse RAW cells was shown to be identical, indicating that only one poly(A)-site is used (Crawford et al, 1997). Interestingly, following the human *TNFα* gene sequence I predicted another potential poly(A) site about 0.5kb downstream from the first poly(A) signal. To identify, whether human monocytic THP-1 cells behave as mouse RAW cells regarding *TNFα* poly(A) site selection in LPS-induced and uninduced conditions, I performed a 3`RACE analysis to define the exact 3`end of all potential TNFα mRNA species. 3`RACE was carried out on total RNA isolated from LPS-induced and uninduced THP-1 cells and reverse transcribed with a modified oligo dT-primer, which is extended by an adaptor sequence to result in primer QT

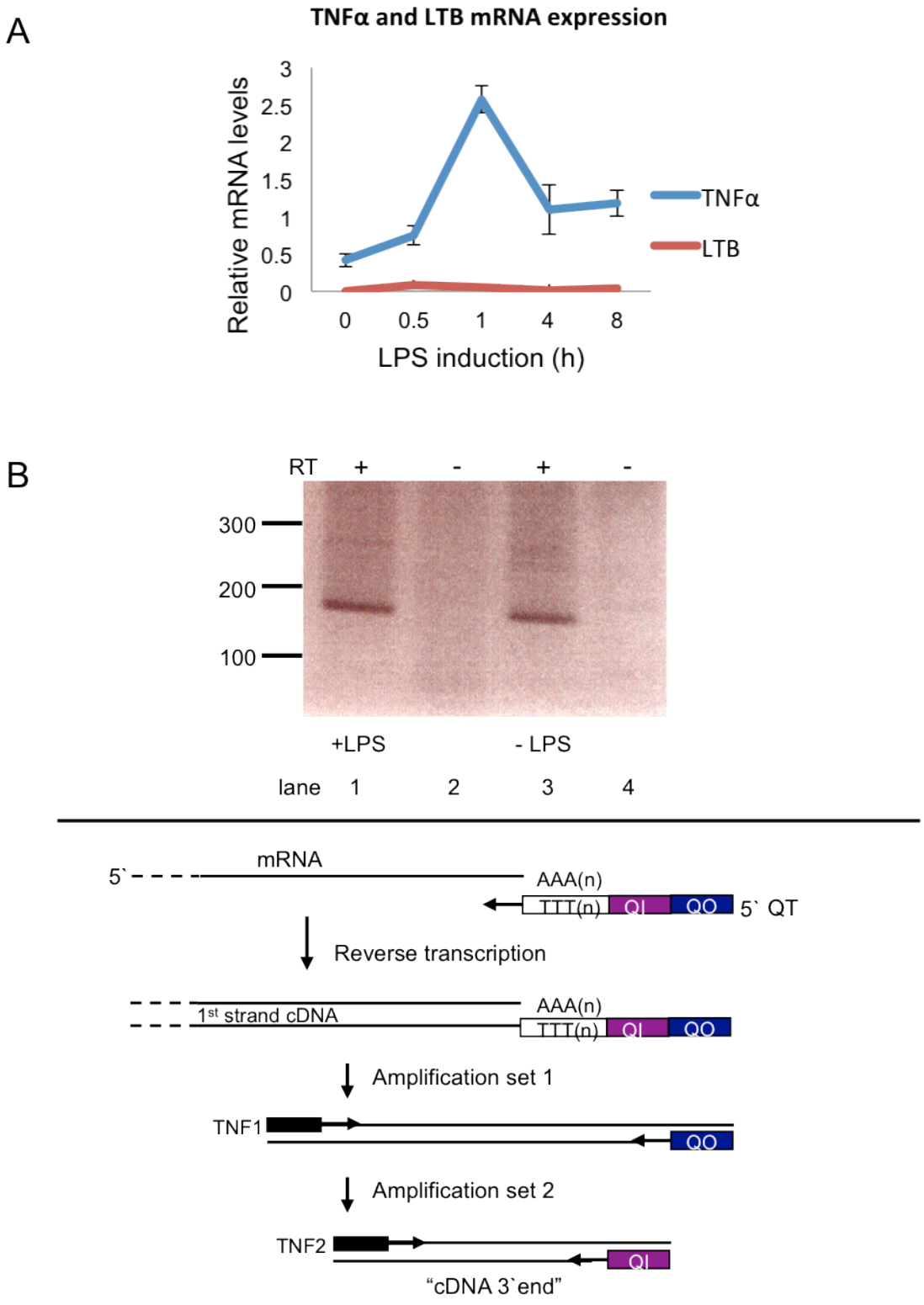


Figure 5.2 TNFα mRNA analysis in THP-1 cells. A qRT-PCR analysis of TNFα and LTB mRNA levels after induction of THP-1 cells with LPS. Lines indicate TNFα mRNA levels (blue line) and LTB mRNA levels (red line) obtained after induction with 100ng/ml LPS over a period of 8h. TNFα and LTB mRNA levels are normalized to the GAPDH mRNA levels in each sample. Error bars indicating standard deviation over three biological repeats are shown. **B** 3`RACE analysis of TNFα mRNA 3` ends in LPS induced and uninduced THP-1 cells. Depicted is a gel cut out showing the only bands obtained after electrophoretic separation of the PCR products on an agarose gel. LPS-induction and band size are indicated. Lanes 1 and 3 indicate the sample bands, lane 2 and 4 show the negative control without added reverse transcriptase in the first step of the 3`RACE. Diagram below displays the experimental steps of 3`RACE starting at the point of reverse transcription of isolated total RNA and depicting the single PCR steps involved throughout the procedure.

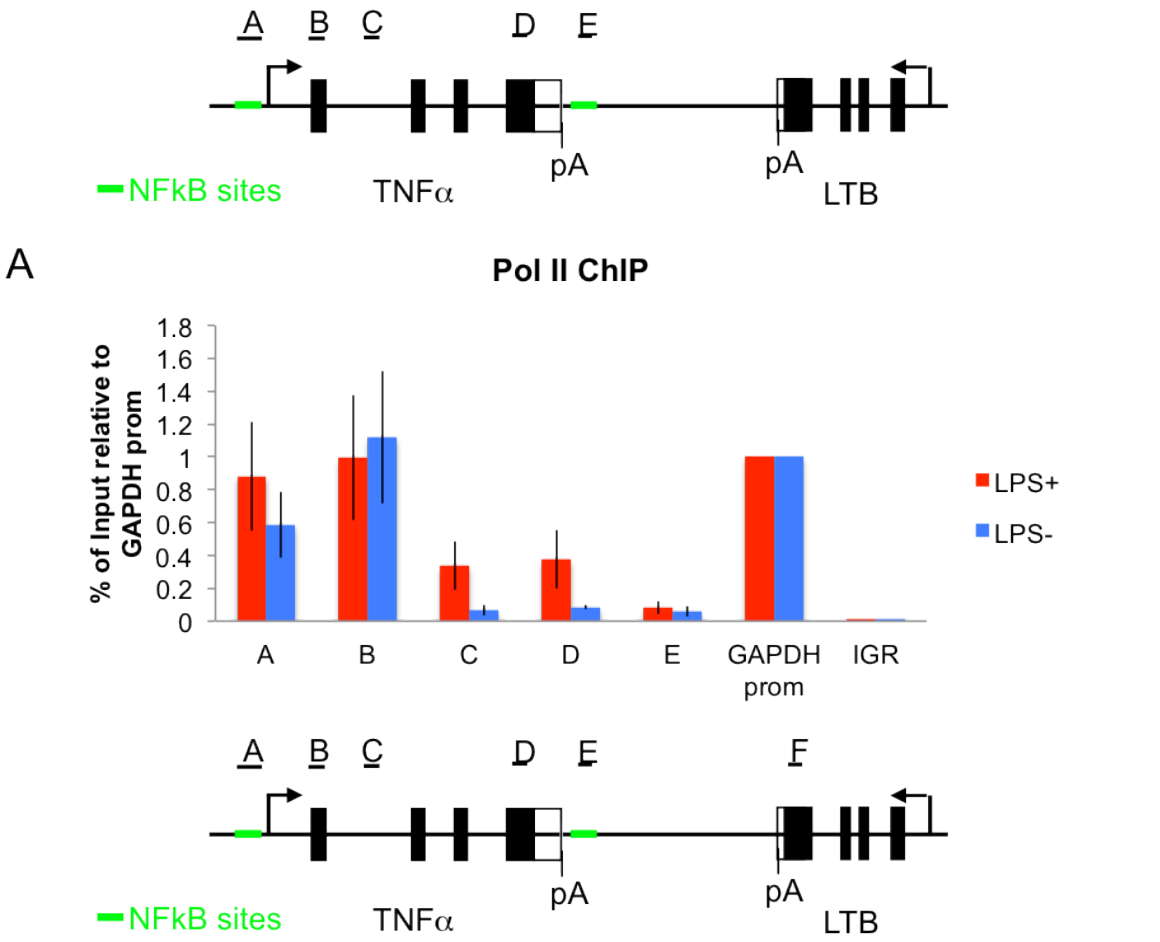
(primer names as well as primer adaptor/anchor sequences follow the published 3`RACE protocol (Scotto-Lavino et al, 2006)). As depicted in the diagram in Figure 5.2B, reverse transcription is followed by a sequence of PCR amplifications with primers within the sequence tail QT and *TNFα* specific primers. This results in cDNA 3`end PCR products, which are then analyzed by agarose gel electrophoresis and sequencing. PCR products shown in lane 1 and 3 (Figure 5.2B) demonstrate that the 3` end of the TNFα mRNA under both conditions is identical, and confirms that only the canonical poly(A) site is used in monocytic THP-1 cells. This result was also confirmed by sequencing of the 3` ends after 3`RACE (data not shown).

5.3 LPS-dependent variations in Pol II occupancy and CTCF, Cohesin and CstF64 recruitment across the *TNF* locus.

I next performed experiments to obtain chromatin-associated profiles of Pol II, CTCF, Cohesin and CstF64 across *TNFα* and partly *LTB* to investigate, whether LPS-induction leads to differences in occupancy and recruitment of these factors. Although I show above (Figure 5.2B) that there is no LPS-induction dependent alternative poly(A) site selection for *TNFα* in THP-1 cells, I did not exclude the possibility of variations in 3`end processing, which could be indicated by differences in cleavage factor recruitment. ChIP experiments were performed on LPS-induced (100ng/ml for 1h) and uninduced THP-1 cells after crosslinking with formaldehyde for 8min at 37°C. The optimized ChIP protocol and information regarding the antibodies used are given in Materials and Methods (Chapter 2). Figure 5.3A shows the observed Pol II profile starting at the *TNFα* promoter, continuing within the gene body and

downstream of the poly(A) signal. The first interesting observation regards the levels of polymerase recruited to the promoter before and after LPS stimulation. Although the amount of produced TNFα mRNA, 1h post induction, proved to be about 5 fold higher than in uninduced conditions (see Figure 5.1 A), the increase in Pol II levels at the LPS-induced promoter appeared to be less than 2 fold (Figure 5.3A, primer position A in diagram). Moving to Exon 1, the levels of Pol II were nearly equal each under both conditions (primer pair B), while the amount of Pol II then dropped significantly in uninduced cells on the way to Exon 4 (primer pairs C and D). After the poly(A) site, Polymerase levels were low in both cases indicative of transcription termination following the canonical poly(A) site. Pol II recruitment to the GAPDH promoter served as a positive control, which predictably remained unaltered by LPS-induction and was therefore used for normalization. Primers in a random region (primer position: Chr1p36.22) served as a negative control, providing the Pol II background levels of this ChIP experiment. These results indicate that LPS induction of TNFα acts at the transcriptional elongation phase.

I next performed a profile for CTCF recruitment to the *TNFα* gene body and the predicted CTCF binding site in the *LTB* Exon 4. Consistent with the lack of CTCF-binding motifs within the *TNFα* gene body, levels for CTCF recruitment to the tested regions A, B, C, D and E remained very low compared to position F, which covers the CTCF motif within *LTB* Exon 4 (see Figure 5.3B). Significantly, there was little to no LPS-induction dependent difference in CTCF recruitment to this region. Consistent with lack of *LTB* expression in monocytic cells (Figure 5.2A), I observed no induction dependent change in



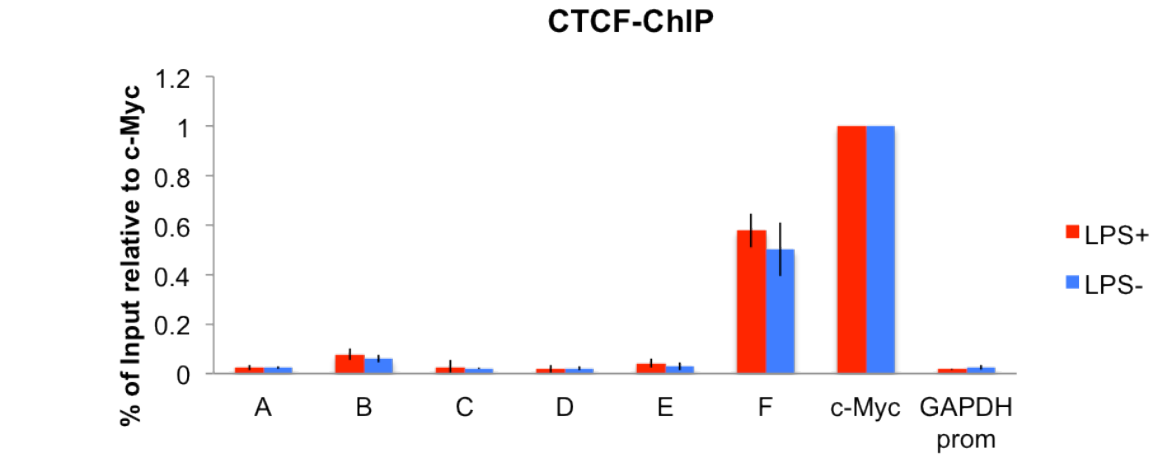


Figure 5.3: **Pol II occupancy and CTCF-recruitment to the *TNF* locus in human monocytic THP-1 cells upon LPS induction. A** qPCR quantitation of ChIP analysis for Pol II occupancy across *TNFα* and at the *GAPDH* promoter. **B** qPCR quantitation of ChIP analysis for CTCF recruitment to *TNFα* and *LTB* and the *c-Myc* regulatory region at position -1.97 kb to the *c-Myc* TSS. *GAPDH* and *c-Myc* serve as positive and negative controls for Pol II occupancy and CTCF-recruitment, respectively. Values are indicated as percent of total input DNA relative to the signal levels at the GAPDH promotor **(A)** and c-Myc **(B)**, respectively. Diagram of the *TNF* locus indicates the promoters (arrow), Exons (black boxes), NFκB binding sites in green and the positions of the primers pairs for qPCR given above. Red bars indicate +LPS and blue bars –LPS treated cells. ChIP values are based on average values ± s.d. from three independent biological experiments.

CTCF recruitment. It is possible that CTCF may be involved in repressing *LTB* expression in monocytes. However, such a premise remains only a hypothetical suggestion. Maybe not substantially significant compared to probe F, but still worth mentioning, was the slightly increased CTCF ChIP signal over probe B. Even if low in signal strength, probe B displays recruitment values of about 2 to 3 fold compared to A, C, D and E. *TNFα* Exon 1 does not contain any predicted CTCF-motif, therefore the slightly higher CTCF signal here can not be explained by CTCF-recruitment per se. However, comparing the CTCF ChIP profile to the Pol II profile described before, I detected high Pol II accumulation over Exon 1, which is independent of LPS induction. Examining the sequence of *TNFα* Exon 1 in more detail, I noticed an increased GC-content of this region compared to the surrounding sequences. An explanation for the increased Pol II levels over this region could relate to influences on Pol II kinetics by pausing or, at least, by slowing down Pol II elongating this region for regulatory purposes. A published report has also shown, that in certain cases, CTCF appears to interact with the large subunit of Pol II (Chernukhin et al, 2007). This has been shown to involve CTCF binding sites. It is difficult to evaluate, whether the slightly increased CTCF signal over *TNFα* Exon 1 reflects a true, Exon 1-specific CTCF accumulation or not. While the signal is significantly lower than at the known CTCF binding site (probe F), it still appears higher than background levels at other probes. Another, purely hypothetical suggestion could be that CTCF travels with the elongating Pol II through interaction with a Pol II subunit or, as part of the elongation complex. In such a scenario CTCF would probably appear as a detectable signal at the same position as Pol II pauses or is

slowed down. Depending on the nature and dynamics of such an interaction it might be predicted that higher signals would not be observed via ChIP analysis. A regulatory CTCF-binding site of the *c-Myc* gene serves as a positive control for CTCF recruitment. It is located 1.97kb upstream of the TSS. Also shown is the signal detected at the GAPDH promoter, showing background levels for CTCF (Figure 5.3B).

The next analysis in this ChIP series was to detect Cohesin recruitment to the TNFα gene, using an antibody against the Cohesin subunit Rad21. The results shown in Figure 5.4A present Rad21 ChIP signal levels across *TNFα*. The signals at the *TNFα* promoter at probe A may reflect LPS-induction dependent differences, indicating higher appearance of Rad21 upon increased transcription activity (Figure 5.4A, probe A). Although the difference between induced and uninduced cells was only about 2 fold, it interestingly reflects a similar fold increase in Pol II at the *TNFα* promoter, after induction. Lower in signal, but reflecting similar differences between LPS-induced and uninduced cells was the *TNFα* 3` end as indicated by probe E. Interestingly, both probes, A and E cover some of the NFκB binding sites in the 5` and the 3` end of *TNFα* and both seem to reflect similar LPS-dependent differences, even if at different signal levels. Signals over Intron1 and Exon 4 appear to fall into background levels. However, Rad21 detected over Exon 1 (probe B) remains puzzling. This is again the same region where I also detected increased Pol II and potentially CTCF levels. However, at this point it remains difficult to explain the presence of Rad21 at this location, especially, as it seems to be LPS-induction independent. The Rad21 ChIP analysis over

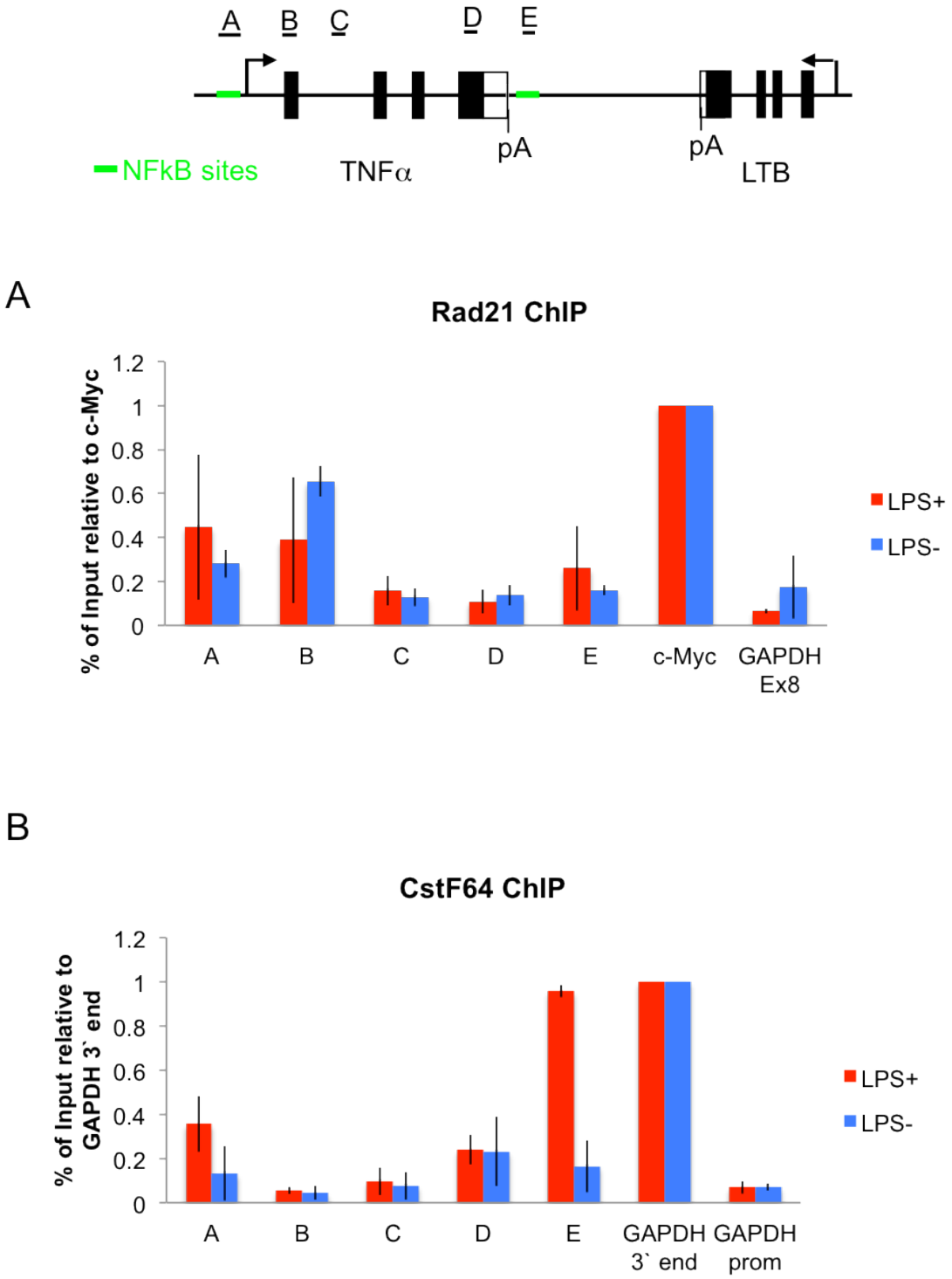


Figure 5.4: **Cohesin and CstF64-recruitment to *TNFα* in human monocytic THP-1 cells upon LPS induction. A** qPCR quantitation of ChIP analysis for Rad21 recruitment to *TNFα* and at the *c-Myc* -1.97kb regulatory region. **B** qPCR quantitation of ChIP analysis for CstF64 recruitment to *TNFα* and *GAPDH*. *GAPDH* and *c-Myc* serve as positive and negative controls, respectively. Values are indicated as percent of total input DNA relative to the signal levels at the *GAPDH* promotor **(A)** and *GAPDH* 3` end **(B)**, respectively. Diagram of the *TNF* locus indicates the promoters (arrow), Exons (black boxes), NFκB binding sites in green and the positions of the primers pairs for qPCR given above. Red bars indicate +LPS and blue bars – LPS treated cells. ChIP values are based on average values ± s.d. from three independent biological experiments.

TNFα went only as far as probe E and did not include the probe over the CTCF-binding site in *LTB* Exon 4. For unknown reasons the variation amongst the different experimental repeats was so high and inconclusive that I could not include these results in the dataset.

The final factor to be tested in this ChIP series was CstF64, the cleavage stimulation factor 64kDa subunit. Early work by the laboratory of JL Manley was the first to show, that upregulation or overexpression of CstF64 in B-lymphocytes plays a key role in regulating IgM heavy chain expression through alternative poly(A) site selection during B-cell differentiation (Takagaki et al, 1996). Later, it was published that monocytic mouse RAW cells increase their levels of CstF64 upon LPS-induction, which leads to alternative poly(A) site selection for several tested genes (Shell et al, 2005). As shown above by 3`RACE, LPS induction does not lead to alternative poly(A) site selection for *TNFα* in THP-1 cells, indicating that monocytic mouse and human *TNFα* are equally regulated in this respect. However, as *TNFα* is a LPS-responsive gene and appears to have stimulation-dependent variations in its 3`end, so far defined by differences in the poly(A) tail length (Crawford et al, 1997), I decided to include CstF64 in my ChIP analysis to investigate potential differences in its recruitment levels. CstF64 binds pre-mRNAs, however, being part of the mRNA 3` end processing complex, it resides in such close proximity to the DNA template so that it can be cross-linked to the DNA and therefore becomes detectable by ChIP analysis. Following the CstF64 profile in Figure 5.4B, three interesting observations became apparent. First, CstF64 recruitment to the *TNFα* 3`end (probe E) seems to be significantly LPS-

induction dependent. Second, I repeatedly obtained a small but significant CstF64 signal at the *TNFα* 5` end (probe A). The fact, that the 5` end peak also appeared to be LPS-dependent, argues that the promoter signal in LPS treated cells reflects a real signal. The only plausible explanation for how a 3` end processing factor could be detectable at the 5` end would be DNA looping. Interestingly, both probes, A and E, bind to regions of regulatory NFκB binding sites. It appears to be an attractive hypothetical idea to suggest that *TNFα* mRNA 3` end processing is differently organized in LPS-induced and uninduced cells, and that the underlying mechanism involves DNA looping between the regulatory regions of *TNFα*. It remains an unresolved question, as to how the induction-dependent length of the poly(A)-tail is regulated (Crawford et al, 1997). Polyadenlyation, as part of mRNA 3` end processing, is regulated at the transcription site. Therefore, it appears an intriguing possibility that the underlying mechanism to LPS-dependent poly(A) tail length is regulated by mechanisms which potentially involve DNA-looping. The adenylation of cytoplasmic, hypoadenylated TNFα mRNA might therefore reflect another regulatory process that helps to meet the requirement for a fast immune response. Comparing the CstF64 levels for chromatin from uninduced cells (blue bars) it is obvious that the signals were overall very low. Since uninduced monocytes still produce low levels of *TNFα* message, higher signals for CstF64 at the 3`end (probe E) might have been expected. A possible scenario, which could explain this result is that CstF64 is differently organized in the 3`end processing complex in uninduced cells and therefore ChIP analysis fails to detect it. Although this is only a hypothetical explanation, the CstF64 ChIP analysis presented remains interesting and

possibly hints at a regulatory mechanism behind active and inactive *TNFα* mRNA.

The next step in my investigation on the involvement of CTCF and Cohesin in *TNFα* regulation was to perform gene knock-down experiments to determine the level of regulatory importance of these factors for *TNFα* expression. Achieving reasonable transfection efficiency for THP-1 cells is generally rather difficult. Therefore, I applied a cell type specific electroporation protocol using a THP-1 specific transfection kit by Lonza to enable maximum siRNA transfection efficiency. Unfortunately, technical difficulties impeded my work on this cell line. Cells did not survive the combination of electroporation treatment and the knock-down of essential factors. I therefore elected to employ a different cellular system for these studies.

A more physiological context to study immune responses in human cells is to use primary cells. Happily, I was offered the possibility to employ human primary PBMC (Peripheral blood mononuclear cells) for my studies. PBMC constitute a very important part of the human peripheral immune system and consist mainly of monocytes, T-cells, B-cells as well as smaller numbers of NK (Natural Killer) cells and dentritic cells. The PBMC model can generally be used for cytokine and chemokine readouts, typically different Interleukins and *TNFα*. PBMC from donated human blood were kindly isolated by Dr. Martinez Estrada, using the standard Ficoll-Hypaque method. All experiments were performed as two biological replicates of each time two blood samples, provided by two donors and analyzed in parallel.

5.4 Study of LPS-induced expression and factor recruitment to *TNFα* in human PBMC

To determine *TNFα* expression of isolated PBMC after LPS-induction over time I generated an LPS response curve over an 8-hour period, after application of two concentrations of LPS. The results depicted in Figure 5.5 show that TNFα levels are significantly higher after induction with 100ng/ml of LPS compared to 10ng/ml LPS, applied in the second sample. The peak of response was in both cases 1h post induction. While THP-1 cells showed an approximately 5 fold increase in *TNFα* message after 1h of 100ng/ml LPS treatment, PBMC responded with an 85 fold TNFα increase following exposure to the same induction conditions.

I next optimized my ChIP protocol for application to human primary cells. Since cells were in suspension, the washing and crosslinking steps required a modified procedure of the protocol used before for adherent cells. PBMC were induced with 100ng/ml LPS, cross-linked with formaldehyde and processed according to the ChIP protocol for primary cells described in Chapter 2. Figure 5.6A shows the Pol II profile across the *TNFα* gene, extending into the 3` intergenic region. Upon induction, higher polymerase signal was detected at the promoter region (probe A). Interestingly, while the total TNFα mRNA levels show an up to 85 fold increase 1h post induction, this is reflected by only a 2 fold increase in polymerase at the promoter. At Exon 1, indicated by probe B, I similarly detected an LPS-dependent accumulation of Pol II. In contrast to the equal levels of Pol II detected over Exon 1 in induced and uninduced THP-1 cells, PBMC reflect a clear LPS-dependency in the signal

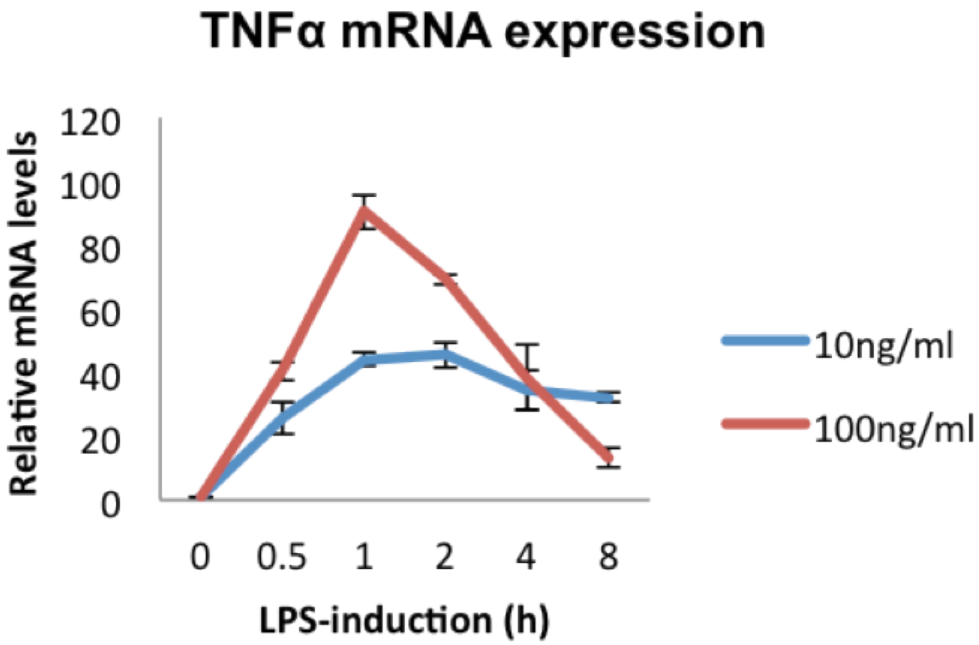


Figure 5.5 qRT-PCR analysis of TNFα mRNA levels after induction of PBMC with different concentrations of LPS. Lines indicate TNFα mRNA levels obtained after induction with 10ng/ml (blue line) and 100ng/ml LPS (red line) over a period of 8h. TNFα mRNA levels are normalized to the GAPDH mRNA levels in each sample. Error bars indicating standard deviation over four independent biological samples are shown.

levels. Pol II levels then dropped across the gene body (probes C and D), until a significant increase over the NFκB binding sites (probe E) appeared in LPS-induced cells only. This stands in contrast to the low Pol II level detected at this site in THP-1 cells. Pol II then dropped again and reached lowest signals after the 3` regulatory regions (probes F and G). As expected, Pol II levels reflect LPS-induction differences across all the analyzed sequence. The controls show the LPS-unaffected Pol II levels at the GAPDH promoter and the Pol II signal background over a transcriptionally inactive, intergenic region.

As well as Pol II, I performed ChIP analysis for CTCF, Smc3 and Rad21-recruitment on PBMC chromatin (plus and minus LPS-induction). To obtain an indication of the signal levels expected for positive recruitment of each factor, and to determine background levels, I analyzed recruitment of these factors to the known CTCF-binding site 1.97kb upstream of the *c-Myc* TSS (included in Figure 5.6B and Figure 5.7A and B, respectively). Regulation of *c-Myc* by Cohesin is evolutionary conserved (Rubio et al, 2008; Stedman et al, 2008), and Rad21, as well as the Cohesin loading factor NIPBL, have been detected at this site. Therefore, this same region served as a positive control for the recruitment of both, CTCF and Cohesin. Following the CTCF-recruitment profile across *TNFα* and *LTB* (Figure 5.6B) it appeared that only *LTB* is targeted by CTCF, evidently in a LPS-dependent manner. Compared to the signal level obtained at the known CTCF-binding site in Exon 4 of *LTB*, indicated by probe J, the CTCF signals detected across *TNFα* appeared to be at background levels. Signal levels detected at *LTB* Exon 1 (probe K) were also low, although displayed an increase after LPS induction, in contrast to

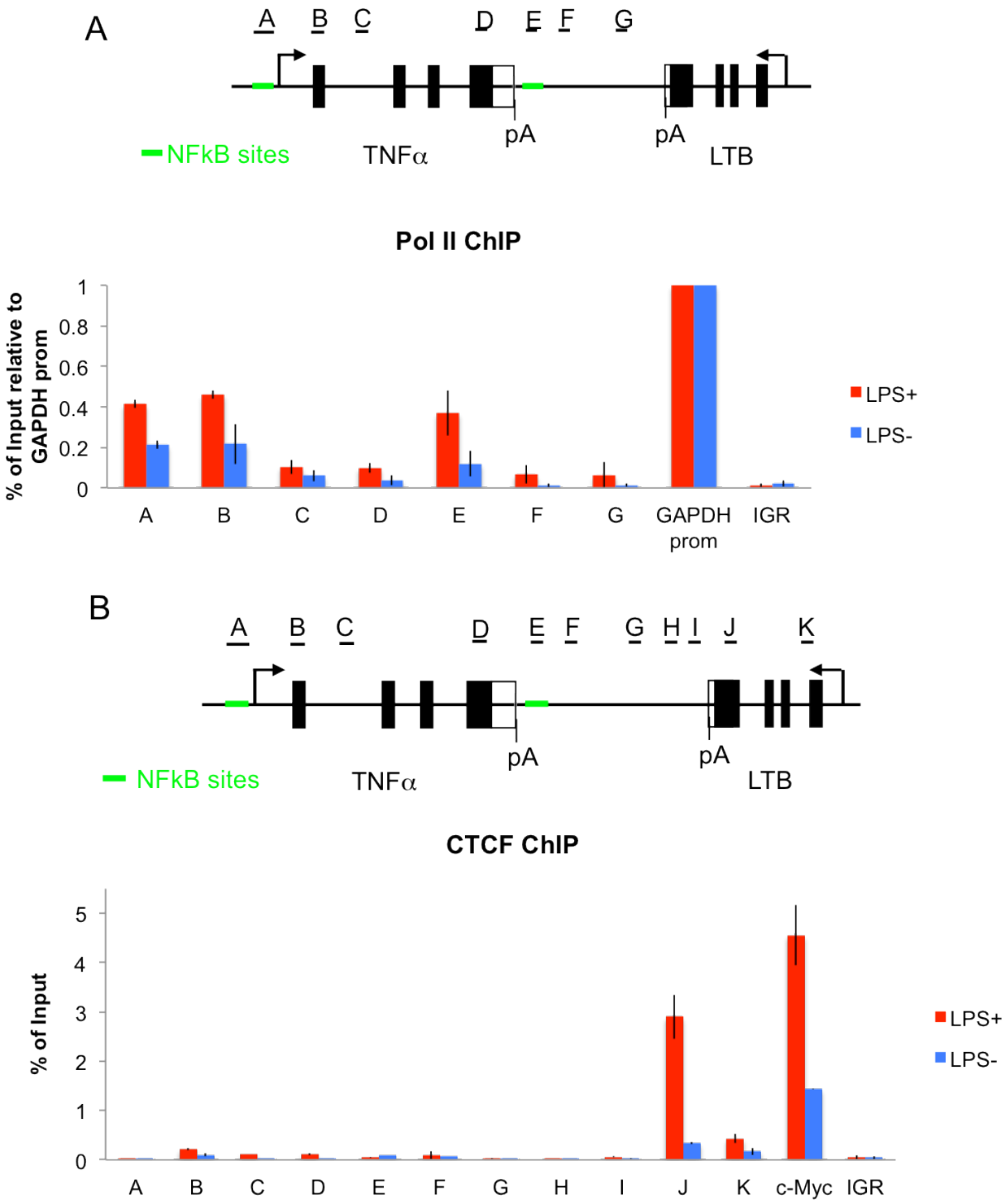
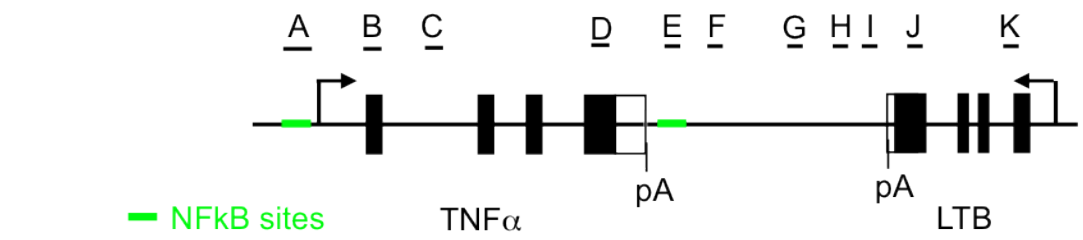


Figure 5.6: **Pol II and CTCF ChIP profiles across *TNFα* in human PBMC upon LPS induction. A** qPCR quantitation of ChIP analysis for Pol II occupancy across *TNFα* and at the *GAPDH* promoter. **B** qPCR quantitation of ChIP analysis for CTCF recruitment to *TNFα*, *LTB,* the *c-Myc* regulatory region at position -1.97 kb to the *c-Myc* TSS and a random intergenic region (IGR). *GAPDH*, *c-Myc* and IGR serve as positive and negative controls for Pol II occupancy and CTCF recruitment, respectively. Values are indicated as percent of total input DNA relative to the signal levels at the *GAPDH* promotor **(A)** or as percent of the total input DNA **(B)**. Diagram of the *TNF* locus indicates the promoter (arrow), Exons (black boxes), NFκB binding sites in green and the positions of the primers pairs for qPCR given above. Red bars indicate +LPS and blue bars –LPS treated cells. ChIP values are based on average values ± s.d. from four independent biological samples.

the allover low signal levels across *TNFα*. As only *LTB* Exon 4 has a canonical CTCF-binding site, it is solely expected to measure significant differences between that site and other regions, which possibly might only contact CTCF through regulatory, dynamic interactions. For this argument I refer to the results of Wicks and Knight, 2011, who were the first to detect CTCF in *LTB* Exon 4 and who obtained a 3C-interaction profile with other regions within the *TNF*-locus, amongst those, the 3`UTR of *TNFα*, which bears no canonical CTCF-motif. However, I argue that depending on the strength and dynamics of such DNA-loop interactions, CTCF could eventually become measurable at sites, which do not contain CTCF binding motifs per se. Another interesting aspect of the CTCF-profile across *TNFα* and *LTB* was the reproducible LPS-dependent difference in CTCF recruitment to *LTB* Exon 4 (probe J). So far, CTCF-binding to *LTB* Exon 4 has been described in the T-lymphocyte cell line Jurkat, where it most likely regulates gene expression by mediating DNA-loops. It appears a reasonable possibility that such DNA-loops are part of regulatory mechanisms with higher levels of complexity and dynamics, often cell type specific. The LPS-dependent difference in CTCF signals over probe J could therefore reflect direct or indirect effects on CTCF-recruitment among different cell types. As there is no evidence for direct inducibility of *LTB* by LPS, the obtained signals could therefore reflect a combination of signals coming from different cell types in the PBMCs, which eventually are part of different regulatory networks.

Interesting, but adding complexity to the PBMC system is the CTCF recruitment to the *c-Myc* control site, as it appeared to respond to LPS induction. This effect might be explained by the heterogeneity of cell types in



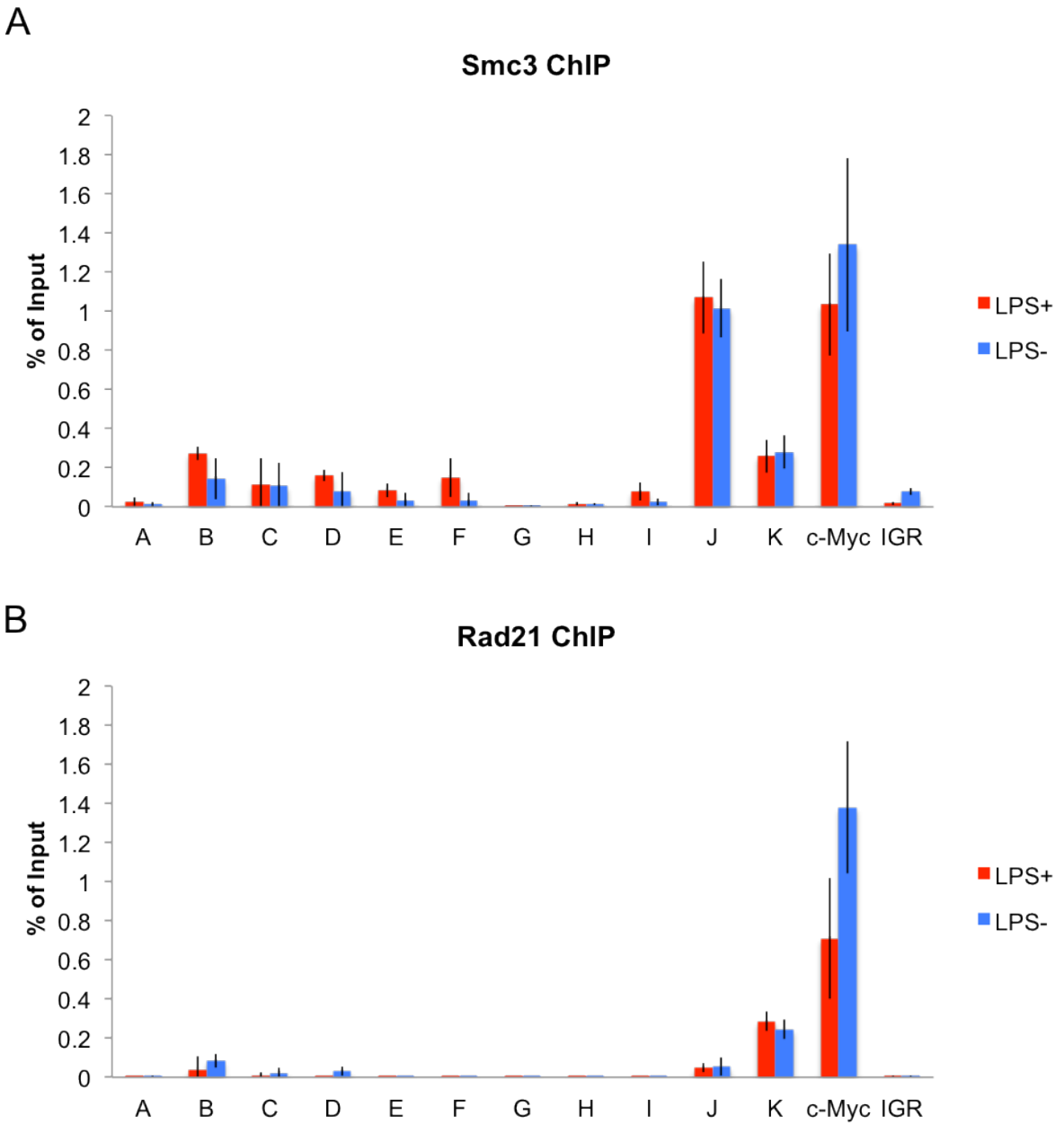


Figure 5.7: **Cohesin recruitment to the *TNF* locus in human PBMC upon LPS induction.** qPCR quantitation of ChIP analysis for Cohesin recruitment to the *TNF* locus and the *c-Myc* regulatory region at position -1.97 kb to the *c-Myc* TSS. Cohesin recruitment is indicated by signals obtained for Smc3 (**A)** and Rad21 **(B).** Values are indicated as percent of total input DNA. Diagram of the *TNF* locus indicates the promoters (arrow), Exons (black boxes), NFκB binding sites in green and the positions of the primers pairs for qPCR given above. Red bars indicate +LPS and blue bars –LPS treated cells. ChIP values are based on average values ± s.d. from four independent biological samples

PBMC and possible, cell specific differences in *c-Myc* regulation. LPS-induced alterations in *c-Myc* expression have, for instance, been shown in peritoneal macrophages (Introna et al, 1986). In regards to the present analysis, *c-Myc* served as a positive control for the ChIP itself but it could not be used for normalization purposes as the LPS-dependent CTCF variations at the *c-Myc* control site were most probably based on different reasons than the CTCF variations across *TNFα* or *LTB*. The random intergenic region (IGR, position Chr1p36.22) indicates the background levels of these ChIP experiments.

I next examined the distribution of two Cohesin subunits, Smc3 and Rad21, across *TNFα* and *LTB* (Figure 5.7) as well as their levels at the *c-Myc* control region, as mentioned before. Both Cohesin subunits appeared to be present at *c-Myc*, although the high error bars indicated high variation between biological repeats. Smc3 signals appeared at background levels across the *TNFα* gene. Only exception is possibly the signal above background at Exon1 (probe B), when cells were induced with LPS. A consistent and significant signal level was, however, detected in *LTB* Exon 4 (probe J), which is LPS-induction independent. This stands in contrast to the previously described CTCF signal over this same region, which is significantly LPS dependent. *LTB* Exon 1 (probe K), like probe B at *TNFα,* displays rather low but above background signals. The detected Rad21 signals from probe A until I across *TNFα* in Figure 5.7B reflect most likely background signals. Apart from the low signal strength, there is no correlation with the Smc3 pattern, and the error bars clearly indicate inconsistency amongst the biological repeats. Particularly unexpected are the significant differences in signal levels between Smc3 and

Rad21, detected at *LTB* Exon 4 (probe J). While Smc3 appeared to be present independently of LPS induction, Rad21 signals were at background levels. Since both, Smc3 and Rad21, were significantly detectable at the *c-Myc* control site, it is less likely that the difference is based on difficulties of the Rad21 ChIP analysis itself. However, so far there is no report indicating independent functions of the Cohesin subunits, which could help to explain this result.

5.5 Discussion: *TNFα* in THP-1 versus PBMC

I first obtained a TNFα mRNA expression profile from each cellular system: THP-1 cells and PBMC. Comparing the expression levels between the monocytic THP-1 cells and PBMC over different time points, it is apparent that primary cells (PBMC) respond with much higher *TNFα* expression activity upon LPS-induction than the cultured cell line. The induction range we are comparing at this point is a maximum 5 fold increase in mRNA expression in THP-1 cells (Figure 5.2A) with an up to 85 fold increase in PBMC upon exposure to 100ng/ml LPS for 1h (Figure 5.5). Taken from the perspective of the *TNFα* expression rate, primary cells represent a more complex but also more efficient system to measure regulatory effects on transcription.

Comparing the Pol II profiles across *TNFα* in both systems, I generally found an expected level of polymerase occupancy over each of the regions probed, which clearly indicated LPS-dependent differences. The promoters in both cell systems display increased Pol II levels upon induction, with a significantly higher difference among induced and uninduced PBMC. I next observed an

interesting accumulation of Pol II over Exon 1 (probe B) in both systems. However, while THP-1 cells displayed an equal signal level under induced and uninduced conditions, PBMC showed high induction dependency. As mentioned before, *TNFα* Exon 1 sequence has a high number of GC nucleotides, with extended G and C stretches. GC-rich sequences can influence polymerase elongation rates, which can significantly reduce the elongation speed, or potentially lead to Pol II pausing. It might be predicted that such effects form a part of regulatory mechanisms for co-transcriptional processes. The fact that I observe LPS-dependent variation in PBMC but equal Pol II levels in THP-1 cells could eventually point to regulatory differences between the different cell types. Another interesting contrast between the two cell systems is the significant LPS-dependent Pol II accumulation over the 3` NFκB sites in PBMC. This indicates potential regulatory differences between the cultured monocytic THP-1 cell line and primary cells. However, at this point it is not possible to determine, from which of the cell types present in the PBMC cell mix (monocytes, B-cells, T-cells, NK cells and dentritic cells) this increased Pol II signal originates.

Linking the Pol II profile to the CTCF-profile obtained in both cell systems, I first observed the slightly increased CTCF-signal measured at Exon 1 (probe B) in THP-1 as in PBMC. However, at this point it is important to draw attention to the generally low signal levels detected. Two possibilities could explain the obtained result. Generally speaking, either all CTCF-signals detected over the *TNFα* gene are only background levels, or the small peaks over probe B give hints for regulatory function. Potentially, except for probe B

in both systems, I would argue that all other CTCF-signals measured across *TNFα* fall into background levels. As the signals over probe B seem to be very slightly, but consistently, increased in both systems and appear to overlap with the region of increased Pol II signal, it could be possible that these CTCF signals are real. An explanation for such a scenario could be CTCF-dependent DNA-loops involving *TNFα* Exon 1, despite the lack of a CTCF-binding site within this sequence. This hypothetical idea is based on the observations of Wicks and Knight, who detected interactions between the canonical CTCF-binding site in *LTB* Exon 4 and the 3`UTR of *TNFα*, which does not contain any CTCF-binding motif either (Wicks et al, 2011). Such interactions are potentially cell type specific, dynamic and possibly do not require contacts between canonical CTCF-binding sites. That could also explain the low signals observed. Although the GC-rich sequence of Exon 1 alone could still be responsible for indirect ChIP signals, it is the dependency on LPS induction in PBMC that leaves the possibility for regulatory involvement still open. At this point it is appropriate to consider the consistency of Pol II and CTCF signals over Exon 1 in both systems to potentially be a signpost towards regulatory functions. However, without 3C-analysis it is not possible to draw further conclusions.

From all the factors analyzed in my ChIP experiments, Cohesin displays the highest variation. However, Rad21 detected at the regulatory NFκB binding sites in the 5` and the 3` end of *TNFα*, as well as above Exon 1 in THP-1 cells, could potentially indicate Cohesin recruitment for functional purposes. Support for this idea could additionally be given by the CstF64 profile,

displaying a peak at the expected *TNFα* 3` end but also a LPS-dependent signal on the *TNFα* 5` end. As discussed before, an interesting possibility is that chromatin-loops bring both ends together and eventually regulate LPS-dependent differences in *TNFα* 3` end processing. In contrast to THP-1 cells, Rad21 signals across *TNFα,* obtained from PBMC chromatin, are only at background levels. On the other hand, Smc3 in PBMC appears to reflect a possible profile indicating LPS-dependency, again at Exon 1. As Cohesin is known to mediate DNA-loop contacts (Parelho et al, 2008; Rubio et al, 2008; Stedman et al, 2008; Wendt et al, 2008), the detected signals might reflect functional involvement. However, depending on the way such contacts are mediated, it will be technically difficult to obtain meaningful information from ChIP experiments. 3C-analysis is the technique to deliver answers to these questions. However, it has to be considered that, depending on the dynamics and stability of DNA-loop interactions, even 3C-profiles could display significant variations.

Although there are several overlaps between the THP-1 and PBMC data sets, inconsistencies and increased variations among the PBMC ChIP profiles originate in the fact that PBMCs are a mix of different cell types. In order to obtain clearer results and provide a better basis for comparisons with the monocytic THP-1 cell line, primary monocytes can be sorted out from the PBMC cell mix. One possibility is the application of FACS sorting. Obtaining a pure monocytic cell sample would then provide a better basis for more consistency when comparing between the cell line and primary cells.

5.6 Effect of Cohesin and CTCF on *TNFα* expression mRNA expression in human blood derived monocytes and macrophages

I have so far described a range of mRNA and ChIP analyses in two independent cell systems in order to identify possible roles of Cohesin and/or CTCF in *TNFα* transcription regulation. Both, monocytic THP-1 cells and peripheral blood mononuclear cells (PBMC) delivered results, which potentially indicate roles for Cohesin and CTCF in the *TNF* locus regulatory machinery. While the *TNFα* response to LPS in THP-1 cells was comparably low, PBMC represented a more physiological system with strong *TNFα* expression upon induction. PBMC are widely used to study cytokine expression, which make a valuable system for immune response studies at the level of regulatory networks. However, studies at a single gene level appeared difficult as *TNFα,* as well as the other genes in the TNF locus, may be regulated differently in the different cell types present in PBMC. This could well be reflected in the increased variations in the recruitment of the tested factors, particularly Cohesin.

In order to continue my work at primary cell level but to avoid crosstalk of regulatory mechanisms for *TNFα* between different cells types, I started to focus on *TNFα* regulation in macrophages, which is the major *TNFα* secreting cell type. Therefore, I initiated a collaborative project with Dr. Udalova of Imperial College School of Medicine, London. This gave me the possibility to analyse *TNFα* expression in human blood derived macrophages. Additionally, I got the opportunity to apply optimized ChIP methodology, which is crucial

when working with human primary cells.

Macrophages are part of the mononuclear phagocyte system (MPS), which is defined on the basis of ontogeny and phagocytic activity and is particularly dynamic during inflammation or infection. Under these conditions, blood monocytes get recruited into the infected tissues, where they differentiate into macrophages (Gordon, 2007) or dentritic cells (DC) (Cheong et al, 2010). Macrophages then acquire distinct functional phenotypes, depending on the microenvironment they infiltrate. In this respect, macrophages demonstrate high plasticity, which allows them to respond efficiently to environmental signals and change their phenotype and physiology in response to cytokines and microbial signals (Mosser et al, 2008). Following their polarized phenotypes, macrophages are often divided into two categories referred to as classically activated macrophages (M1 macrophages) and alternatively activated macrophages (M2 macrophages). Each of these populations of cells displays distinct functions that are phenotypically characterized by the production of pro- and anti-inflammatory cytokines, respectively (Gordon et al, 2010). M1 macrophages are responsible for strong pro-inflammatory response mediating resistance to pathogens and contribute to tissue destruction. M2 macrophages carry an anti-inflammatory phenotype and promote tissue repair, remodeling as well as tumor progression (Gordon, 2003; Martinez et al, 2008). However, the literature and particularly recent scientific contributions indicate a grey shade area for the M1/M2 nomenclature. Depending on the trigger, more defined subpopulations emerge, which relate to one or other of the main phenotypes and exhibit

distinct characteristics in their responses to different stimuli reviewed in (Gordon et al, 2010). I was curious to determine, whether there might be differences in Cohesin or CTCF occupancy at the *TNFα* gene in M1 and M2-differentiated macrophages versus undifferentiated precursor monocytes. M1 macrophages are the main TNFα secreting cell type, however *TNFα* expression can be stimulated in all three cell types.

In the laboratory of Dr. Udalova, monocyte differentiation protocols are established with two macrophage-secreted cytokines, the Granulocyte macrophage colony-stimulating factor (GM-CSF) for M1 macrophages and the Macrophage stimulating factor (M-CSF) for M2-type macrophages. GM-CSF stimulates stem cells to produce granulocytes and monocytes. Monocytes then exit the circulation and migrate into tissue, where upon they mature into M1 macrophages and dentritic cells to follow the inflammatory cascade. M-CSF is involved in the proliferation, differentiation and survival of monocytes and macrophages and bone marrow progenitor cells (Stanley et al, 1997). M-CSF stimulates increased phagocytic and chemotactic activity of macrophages and increased tumor cell cytotoxicity (Nemunaitis, 1993). Upon stimulation of monocytes with GM-CSF and M-CSF (described in Chapter 2) cells were induced with LPS over a period of 8h followed by total RNA isolation. Reverse transcription was performed with oligo dT primers, followed by qRT-PCR analysis set up with a primer pair within TNFα Exon 4. TNFα mRNA levels were normalized to values obtained for GAPDH mRNA.

As shown in Figure 5.8, GM-CSF differentiated cells, which represent the pro-

inflammatory M1 macrophage phenotype, express more TNFα mRNA upon LPS stimulation compared to M-CSF (M2). The highest mRNA levels for both macrophage phenotypes were obtained after 1h of LPS induction, as already previously shown in monocytic THP-1 cells (Figure 5.2A) and PBMC (Figure 5.5). Although M2 macrophages represent the anti-inflammatory phenotype they are nevertheless inducible for *TNFα* expression upon LPS stimulation. Given the macrophage plasticity to respond to environmental stimuli, I suggest LPS treated M2 macrophages modify their characteristics according to the pro-inflammatory stimulus they were exposed to.

I next performed ChIP analysis to compare monocytes with M1 and M2 macrophages in order to identify possible differences in recruitment of Cohesin and CTCF. In order to perform this analysis on human blood isolated monocytes, GM-CSF and M-CSF differentiated macrophages I applied a ChIP protocol optimized by the laboratory of Dr. Udalova. The protocol details are described in Chapter 2. The results after quantitative real time PCR are presented in Figure 5.9. The graphs in 5.9 I, II and III show ChIP signals across *TNFα* and its regulatory regions in GM-CSF (M1), M-CSF (M2) macrophages and monocytes (Mono) over three LPS-stimulation times (0, 1 and 4h). Primer positions are given in the diagram above, covering the binding sites for NFκB and other regulatory factors and extending into the *TNFα* 5` UTR (probes I, II and III). Primer pair IV is located in *TNFα* Exon 4 and V covers the three *TNFα* poly(A)-proximal NFκB binding sites. The Pol II profile across *TNFα* in Figure 5.9A shows a clear difference between polymerase occupancy among M1 and M2 macrophages, and particularly

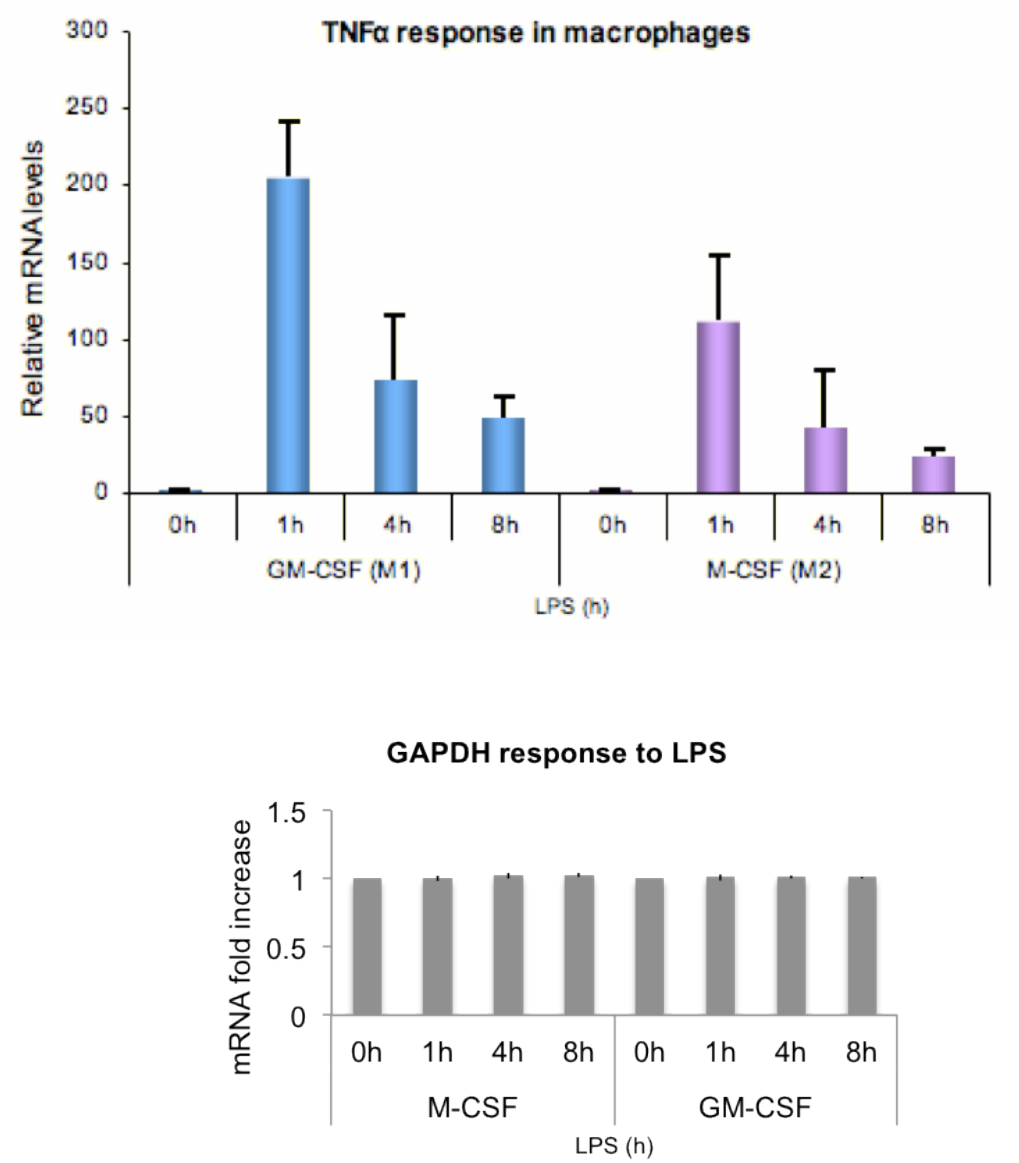


Figure 5.8: TNFα mRNA analysis in human M1 and M2 macrophages. qRT-PCR analysis of TNFα mRNA expression levels after 0-8h of 100ng/ml LPS treatment of GM-CSF (M1) and M-CSF (M2) differentiated macrophages. TNFα mRNA levels are normalized to GAPDH mRNA levels in each sample. Graph below indicates the fold increase of GAPDH mRNA levels after 0-8h of LPS treatment. Error bars indicating standard deviation over four independent biological samples are shown.

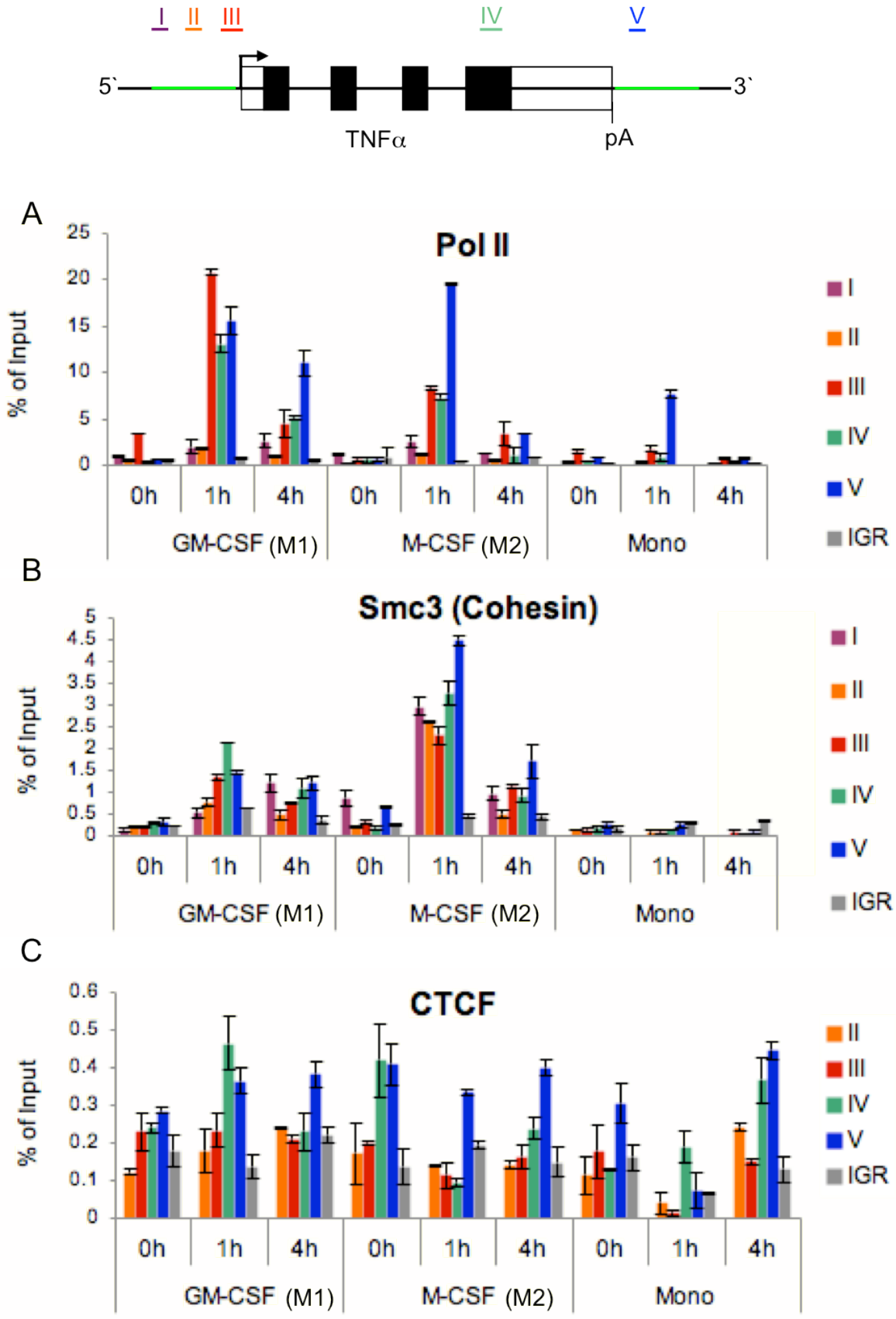


Figure 5.9: **Pol II, Cohesin and CTCF recruitment to *TNFα* upon LPS induction in human blood-derived monocytes and differentiated macrophages.** qPCR analysis of ChIP for Pol II **(A)** Smc3 (Cohesin) **(B)** and CTCF **(C)** recruitment to *TNFα* and its 5` and 3` regulatory regions. Values are indicated as percent of total input DNA. Diagram above indicates the positions of the primers pairs across *TNFα* and its regulatory regions. IGR stands for Intergenic Region and indicates a random, genomic region with no transcriptional activity. IGR provides the signal background levels of these experiments. ChIP values are based on average values ± s.d. from two independent biological samples.

monocytes. Even though all three cell types are inducible for *TNFα* expression, monocytes show lower levels of Pol II compared to both macrophage classes M1 and M2, which is consequently reflected in significantly lower TNFα mRNA levels. Moreover, the Pol II profile for monocytes does not show any differences upon LPS stimulation even after 1h and 4h, with exception of the significant peak over probe V. Interestingly, I observed a similar Pol II accumulation in PBMC at the same location, 1h post LPS-induction. More unexpectedly, Pol II profile differences appeared among M1 and M2 macrophages. Under uninduced conditions at time point 0h, both macrophage phenotypes display Pol II levels comparable to monocytes. 1h post LPS-induction, Pol II signals increase by up to 10 fold. However, the occupancy at different positions is significantly different among both macrophage types. M1 macrophages display background levels over the first two NFκB regulatory regions at probe I and II. The same applies to M2 macrophages. However, while Pol II shows high occupancy of the most TSS-proximal NFκB region/*TNFα* 5` UTR (probe III) in M1 cells, levels at this position are less than half in M2 cells. Approximately the same difference is reflected in the signals at *TNFα* Exon 4 (probe IV). Interestingly, both macrophage phenotypes display similarly high Pol II levels at the 3` NFκB regions covered by probe V. While at 4h post LPS-induction, *TNFα* still displays significant Pol II levels in M1 cells, levels in M2 macrophages dropped to nearly background. Interestingly, the high peak detected 1h post induction at probe III is strongly reduced after 4h of induced *TNFα* expression, while Pol II levels remain high at probe V. This again stands in contrast to Pol II levels detected after 4h LPS treatment in M2 cells. IGR again

indicates the control primer pair for background Pol II levels in a random, intergenic region (position Chr1p36.22).

While Cohesin ChIP analysis did not indicate very clear results in monocytic THP-1 cells or in PBMC, the Smc3 profile in macrophages shows significant recruitment levels, particularly 1h post induction, and highlights major differences across M1 and M2 cells (Figure 5.9B). Importantly, it is evident that monocytes do not display significant Smc3 recruitment to *TNFα*. Interestingly, while M1 macrophages display higher Pol II occupancy, Cohesin levels are lower compared to M2 cells. In contrast, M2 display less Pol II but higher Smc3 recruitment across the *TNFα* gene and its regulatory regions. Particularly striking is the fact that high Smc3 recruitment signals were detected over all the NFκB regulatory regions. The same regions display significantly lower Smc3 recruitment in M1 macrophages and additionally appear in a different pattern. These differences may also be linked to the regulation of *TNFα* at 1h post LPS-induction, as after 4h, the profiles balance out and display not only similar Smc3 levels but also identical distribution profiles.

CTCF-ChIP analysis across *TNFα* (Figure 5.9C) resulted in questionable profiles. While many of the detected signals are below or close to background levels (indicated by probe IGR), there is also no significant profile variation among the two macrophages phenotypes and monocytes at any of the measured LPS-induction time points. This leads me to the conclusion that CTCF is not recruited to the regions measured across the *TNFα* gene under

these conditions and in this experimental setup.

In view of the Cohesin (Smc3) recruitment profiles across *TNFα* for M2, it is appropriate to postulate, that Cohesin could play a regulatory role in *TNFα* expression. To determine, whether there is potentially direct influence on TNFα mRNA expression, I measured TNFα mRNA levels after knocking down Cohesin and CTCF (Figure 5.10). For the reason of limited, available cell samples I was only able to measure TNFα mRNA response to Cohesin- and CTCF knock-down in M2 macrophages. Briefly, M2 differentiated macrophages were transfected with siRNAs targeting CTCF, Smc3 and Rad21 (Cohesin), respectively. siRNA against Luciferase was applied as a control. Later, cells were induced with LPS for 0, 1, 4 or 8h and total RNA was isolated. Reverse transcription was performed with random hexamer primers, followed by qRT-PCR. Figure 5.10 (left graphs) shows the knock-down efficiency measured by qRT-PCR for Smc3, Rad21 and CTCF, respectively. The effects on TNFα mRNA levels are displayed in Figure 5.10 A-C, right graphs. Knock-down efficiencies vary between 65% and 80% for the two Cohesin subunits Smc3 and Rad21, and reach about 50% to 60% for CTCF. Comparing the resulting TNFα mRNA levels, the result in Figure 5.10A appears interesting, though unexpected. While Smc3 seems to be recruited to the *TNFα* gene and its regulatory regions (as shown in Figure 5.9B), knock-down of Smc3 does not influence TNFα mRNA levels at all after 1h of LPS-stimulation and causes a slight increase after 4h of induction. Rad21 knock-down on the other hand causes an about 50% reduction in TNFα mRNA levels after 1h LPS-induction, while after 4h and 8h, no difference was

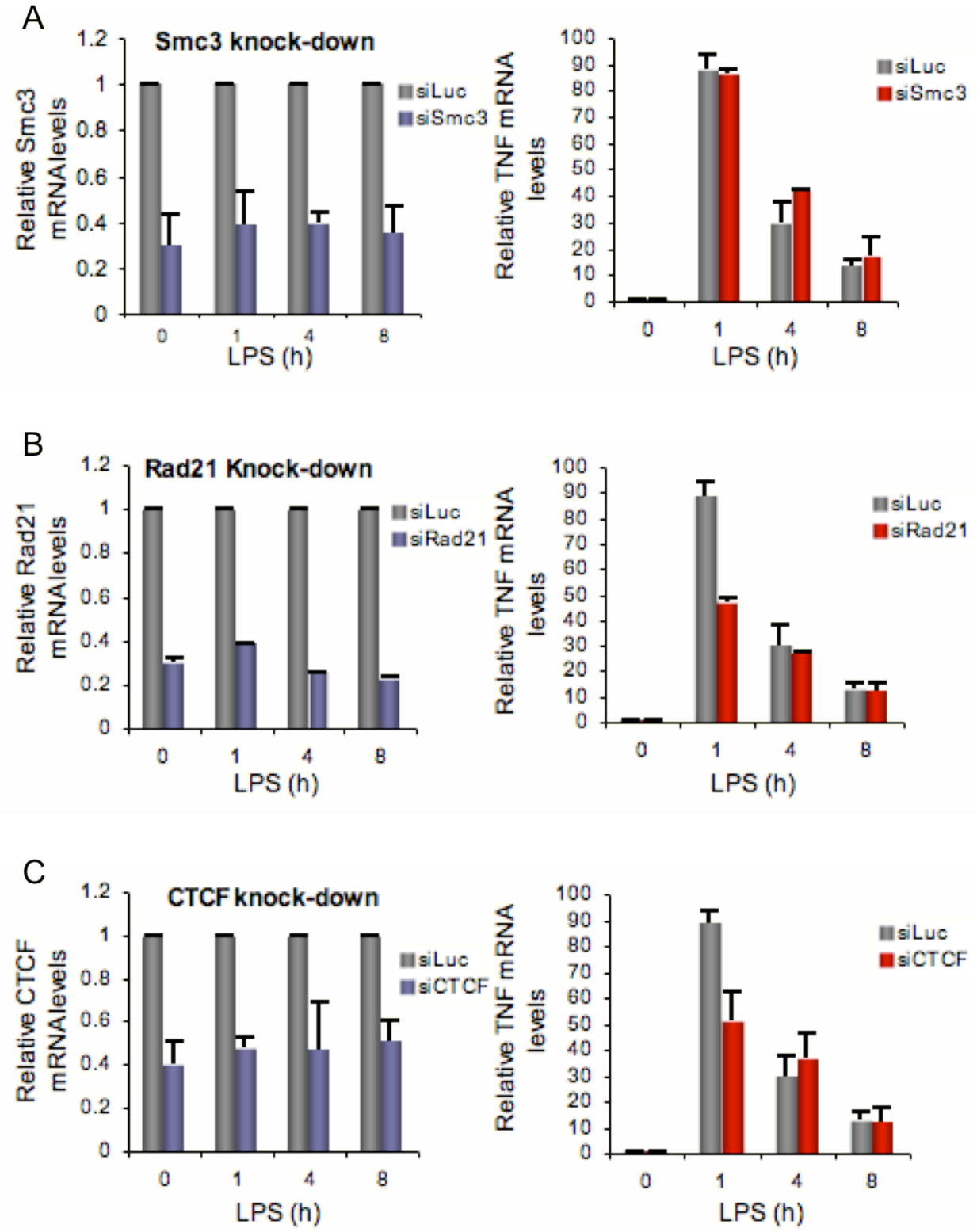


Figure 5.10 Effects of Cohesin and CTCF knock-down on TNFα mRNA expression levels in M2 macrophages. M2 macrophages were transfected with siRNA targeting Smc3 **(A)**, Rad21 **(B)** and CTCF **(C)** or Luciferase-targeting (control) siRNA (siLuc) and stimulated for 1h to 8h with LPS (100ng/ml). Smc3, Rad21 and CTCF levels as well TNFα mRNA were detected by qRT-PCR. Error bars indicate standard deviation over two independent biological experiments.

detectable (Figure 5.10B). It is surprising that while Smc3 and Rad21 both represent subunits of the Cohesin complex, only knock-down of Rad21 leads to altered *TNFα* expression levels. Smc3 knock-down does not cause any change. Furthermore, even though CTCF did not appear to be recruited to the *TNFα* gene, knock-down of CTCF mRNA to 50% causes a reduction of TNFα mRNA levels 1h post induction (Figure 5.10C).

5.7 Discussion: Effects of Cohesin and CTCF on *TNFα* expression in human blood derived monocytes and macrophages

Although monocytes and macrophages produce TNFα mRNA upon LPS-induction, differences in expression levels and mainly Pol II and Cohesin recruitment to the *TNFα* gene and its regulatory regions in the 5` and 3` end became apparent in human, blood-derived primary cells. Following the Pol II and Smc3 ChIP profiles in Figure 5.9A and B, it appears interesting that while TNFα is expressed in all three cell types, Pol II occupancy reflects significantly different patterns. As monocytes express little TNFα mRNA compared to macrophages, it is not surprising to obtain low levels of *TNFα* Pol II occupancy, compared to macrophages. However, the fact that the high Pol II levels in M1 and M2 macrophages display significant differences in distribution across *TNFα* leads to the view that despite efficient *TNFα* expression, regulation of the same gene is organized differently between M1 and M2 macrophages. This is supported by significant variations in the Cohesin-distribution profiles (represented by Cohesin subunit Smc3) across

TNFα between M1 and M2 cells. Smc3, in context with *TNFα* occupation, appears interesting in two aspects. First, the profile across *TNFα* evidently shows differences in the recruitment profiles between M1 and M2 cells. Secondly, the increased signals after 1h of LPS-induction match *TNFα* regulatory regions, indicating NFκB binding sites (Figure 5.9, diagram probes I and V). This led me to the idea that there could be a possible connection between NFκB recruitment and Cohesin-involvement in *TNFα* regulation. This is an attractive possibility particularly as the strongest Smc3 signal differences between M1 and M2 are at probe I and V, exclusively NFκB binding sites (Figure 5.9B, GM-CSF and M-CSF at 1h). Cohesin has been shown to have CTCF-independent gene regulatory functions by mediating long-range chromosomal *cis*-interactions with other factors in a cell type specific manner: the mediator complex in murine embryonic stem cells (Kagey et al, 2010), liver-specific transcription factors in HepG2 cells and estrogen receptor alpha in breast cancer MCF-7 cells (Schmidt et al, 2010). Moreover, Cohesin was recently found to play a role in T-cell receptor rearrangements and thymocyte differentiation (Seitan et al, 2011). It appears to be a reasonable assumption that macrophage-specific interactions with transcription factors potentially regulate *TNFα* expression. Particularly, as CTCF-levels across *TNFα* appear low (Figure 5.9C), this could point into the direction of another example for CTCF-independent Cohesin involvement in gene regulation in a cell type- or tissue specific manner. As M2 macrophages generally follow the anti-inflammatory pathway and do express *TNFα* upon LPS-induction I considered, whether the increased Cohesin recruitment to NFκB sites represents a certain M2-specific regulatory mechanism for *TNFα*. This could

result in significantly lower gene expression 1h post LPS treatment compared to M1 macrophages. Interestingly, the Smc3 profiles between M1 and M2 macrophages equalise after 4h of LPS induction, not only in levels but also in the distribution pattern. Also, the differences in TNFα mRNA levels among M1 and M2 at later LPS-induction times are less prominent (Figure 5.9, time points 4h and 8h). This opens the question, whether M2 macrophages, induced with a pro-inflammatory stimulus, respond with expression of pro-inflammatory factors, while they undergo a certain level of phenotypic modification. This idea is based on results suggesting macrophage phenotype plasticity between M1 and M2 cells, as for instance reported by (Krausgruber et al, 2011). At this point it would be interesting to investigate, whether the resulting TNFα mRNA from M2 macrophages is equal to TNFα from M1 macrophages, particularly in terms of processing and entry into translation.

Having obtained significant Cohesin-recruitment signals and variations across *TNFα*, it was rather unexpected to detect equal TNFα mRNA levels in M2 macrophages upon Smc3 knock-down, as shown in Figure 5.10A. In my knock-down experiments I was able to include another Cohesin subunit, Rad21, which in contrast to Smc3 did show a reduction in TNFα mRNA levels upon Rad21 knock-down. Given the fact that Smc3 and Rad21 are part of the same protein complex, this result is quite difficult to explain. As I was limited in the number of immuno-precipitations performed during each of the ChIP analyses, I could not obtain a Rad21 profile within the same ChIP series. Therefore, at this stage I cannot determine, whether Cohesin varies in the assembly of its subunits according to its functions in gene regulation or

whether an entirely different reason underlies this effect.

An additional remark concerns the significant Pol II peak in blood-derived monocytes over region E (Figure 5.9 V) and the detected Pol II signal at the same region in PBMC (Figure 5.6A, probe E). As with PBMC, it was not possible to conclude, in which cell type this Pol II accumulation in the *c* 3` end originated. The single peak in Figure 5.9A (monocytes 1h LPS induction) now resolves this question. Consequently, Pol II seems to be detectable as an accumulation at the 3` NFκB sites. The functional significance of this Pol II accumulation cannot be answered at this point.

5.8 A summary of *TNFα* and its potential regulation by Cohesin

I have investigated on the possibility of Cohesin- and CTCF-mediated regulation of *TNFα* gene expression. During the course of this project development, I applied three cellular systems. The aim was to find the most physiological conditions for *TNFα* expression, which also allowed experimental freedom in terms of possibilities to modify the system to a certain extend. I obtained several partly overlapping data sets across the cell lines, which I consider as potential indications towards regulatory mechanisms.

Summarizing, the ChIP profiles across *TNFα* provide insight into potential *TNFα* regulation via chromatin loops. In THP-1 monocytes, it is the detection of the cleavage stimulation factor CstF64 at the *TNFα* 3` end as well as at the

5` end in a LPS-induction dependent manner, which provided an indication that potentially chromatin loops bring both ends in proximity, when *TNFα* transcription is activated (Figure 5.4B). The Cohesin-profile for THP-1 cells, obtained by immune-precipitating the Cohesin subunit Rad21, eventually confirms this possibility by showing LPS-dependent signal increases at the sites of CstF64 detection (Figure 5.4A). Cohesin-ChIP analysis in PBMC cells showed less clear results (Figure 5.7). The high variations across the biological repeats could be explained by the fact that PBMC are a mixture of different cell types, of which several express *TNFα.* It is unclear, whether *TNFα* is regulated in the same way in each of these cell types. Potential differences, including factor involvement and the dynamics of regulatory structures, like chromatin loops, may account for variations among experimental repeats. CTCF was detected in low levels in THP-1 cells as in PBMC, but the signal observed in Exon 1 could eventually indicate regulatory involvements of CTCF (Figures 5.3B and 5.6). As discussed before, CTCF recruitment is readily detectable at canonical CTCF-binding sites. However, CTCF has also been detected at CTCF-motif-free locations, where it seemed to mediate regulatory gene looping (Wicks et al, 2011). Depending on the dynamics of such structures, it could be difficult to obtain significant signals for temporary CTCF-binding, compared to CTCF recruited to canonical binding sites.

The most reliable system for establishing *TNFα* gene expression signal strength and reproducibility so far, were the human blood-derived monocytes differentiated into M1 and M2 macrophages. Besides the high *TNFα*

expression levels obtained after LPS-stimulation (Figure 5.8), also the ChIP profiles exhibited significant signal strength, reproducibility and clear differences among the three cell types (monocytes, M1 and M2). Here, Cohesin is clearly recruited to the *TNFα* gene and it`s regulatory regions. Smc3 detected at the NFκB-binding sites, covered by probes I, II, III and V, exhibited major differences between M1 and M2 macrophages 1h post induction, while monocytes only show signal background levels (Figure 5.9B). CTCF signals, on the other hand, mostly appear very weak and close to background signal levels. This could potentially indicate another example for cell-type specific, regulatory chromatin loops, mediated by Cohesin recruited independently of CTCF but by cell type specific transcription factors. How Cohesin potentially regulates *TNFα* gene expression is not as clear yet. Knock-down experiments showed no change in TNFα mRNA expression levels after Smc3 down-regulation, while Rad21 resulted in a 2 fold reduction. Interestingly, CTCF knock-down seems to cause some level of TNFα mRNA reduction, although it did not appear to be recruited to the *TNFα* gene. My studies to date remain quite preliminary so that clear conclusions about the mechanism, by which Cohesin potentially regulates *TNFα* expression cannot yet be made. Further investigation is required to resolve these questions.

In spite of the preliminary nature of my studies on *TNFα* gene regulation, some clear conclusions can already be made. I provide the first evidence of Cohesin recruitment to the *TNFα* gene body and its regulatory NFκB-binding sites. In keeping with the known role of Cohesin in gene topology, my results may eventually lead to a mechanism of *TNFα* gene regulation. This may

occur via chromatin loops, mediated by cell type specific Cohesin-transcription factor interactions. Considering the complex expression dynamics *TNFα* follows in order to respond to inflammatory inducers, it can be expected that regulatory chromatin loops may play a key part in this process.

CHAPTER 6

REFERENCES

Aggarwal BB, Puri, R.K. (1995) Human cytokines: their role in disease and therapy. *Blackwell Science inc*

Aggarwal BB, Samanta, A., Feldmann, M. (2000) *TNF receptors*.

Aggarwal BB, Shishodia S, Takada Y, Jackson-Bernitsas D, Ahn KS, Sethi G, Ichikawa H (2006) TNF blockade: an inflammatory issue. *Ernst Schering Research Foundation workshop***:** 161-186

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

Akashi K, Traver D, Miyamoto T, Weissman IL (2000) A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* **404:** 193-197

Akyuz S, Ozel AE, Balci K, Akyuz T, Coker A, Arisan ED, Palavan-Unsal N, Ozalpan A (2011) Raman micro-spectroscopic analysis of cultured HCT116 colon cancer cells in the presence of roscovitine. *Spectrochimica acta Part A, Molecular and biomolecular spectroscopy* **78:** 1540-1547

Aranda A, Proudfoot NJ (1999) Definition of transcriptional pause elements in fission yeast. *Molecular and cellular biology* **19:** 1251-1261

Arumugam P, Gruber S, Tanaka K, Haering CH, Mechtler K, Nasmyth K (2003) ATP hydrolysis is required for cohesin's association with chromosomes. *Current biology : CB* **13:** 1941-1953

Ashfield R, Enriquez-Harris P, Proudfoot NJ (1991) Transcriptional termination between the closely linked human complement genes C2 and factor B: common termination factor for C2 and c-myc? *The EMBO journal* **10:** 4197-4207

Ashfield R, Patel AJ, Bossone SA, Brown H, Campbell RD, Marcu KB, Proudfoot NJ (1994) MAZ-dependent termination between closely spaced human complement genes. *The EMBO journal* **13:** 5656-5667

Auffray C, Fogg D, Garfa M, Elain G, Join-Lambert O, Kayal S, Sarnacki S, Cumano A, Lauvau G, Geissmann F (2007) Monitoring of blood vessels and tissues by a population of monocytes with patrolling behavior. *Science* **317:** 666-670

Auffray C, Fogg DK, Narni-Mancinelli E, Senechal B, Trouillet C, Saederup N, Leemput J, Bigot K, Campisi L, Abitbol M, Molina T, Charo I, Hume DA, Cumano A, Lauvau G, Geissmann F (2009a) CX3CR1+ CD115+ CD135+ common macrophage/DC precursors and the role of CX3CR1 in their response to inflammation. *The Journal of experimental medicine* **206:** 595-606

Auffray C, Sieweke MH, Geissmann F (2009b) Blood monocytes: development, heterogeneity, and relationship with dendritic cells. *Annual review of immunology* **27:** 669-692

Bain J, McLauchlan H, Elliott M, Cohen P (2003) The specificities of protein kinase inhibitors: an update. *The Biochemical journal* **371:** 199-204

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K (2007) High-resolution profiling of histone methylations in the human genome. *Cell* **129:** 823-837

Barthel R, Tsytsykova AV, Barczak AK, Tsai EY, Dascher CC, Brenner MB, Goldfeld AE (2003) Regulation of tumor necrosis factor alpha gene expression by mycobacteria involves the assembly of a unique enhanceosome dependent on the coactivator proteins CBP/p300. *Molecular and cellular biology* **23:** 526-533

Bell AC, Felsenfeld G (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* **405:** 482-485

Bell AC, West AG, Felsenfeld G (1999) The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98:** 387-396

Bender MA, Byron R, Ragoczy T, Telling A, Bulger M, Groudine M (2006) Flanking HS-62.5 and 3' HS1, and regions upstream of the LCR, are not required for beta-globin transcription. *Blood* **108:** 1395-1401

Bender MA, Reik A, Close J, Telling A, Epner E, Fiering S, Hardison R, Groudine M (1998) Description and targeted deletion of 5' hypersensitive site 5 and 6 of the mouse beta-globin locus control region. *Blood* **92:** 4394-4403

Bernstein P, Ross J (1989) Poly(A), poly(A) binding protein and the regulation of mRNA stability. *Trends in biochemical sciences* **14:** 373-377

Betz JL, Chang M, Washburn TM, Porter SE, Mueller CL, Jaehning JA (2002) Phenotypic analysis of Paf1/RNA polymerase II complex mutations reveals connections to cell cycle regulation, protein synthesis, and lipid and nucleic acid metabolism. *Mol Genet Genomics* **268:** 272-285

Beyer K, Dandekar T, Keller W (1997) RNA ligands selected by cleavage stimulation factor contain distinct sequence motifs that function as downstream elements in 3'-end processing of pre-mRNA. *The Journal of biological chemistry* **272:** 26769-26779

Bird G, Zorio DA, Bentley DL (2004) RNA polymerase II carboxy-terminal domain phosphorylation is required for cotranscriptional pre-mRNA splicing and 3'-end formation. *Molecular and cellular biology* **24:** 8963-8969

Birse CE, Lee BA, Hansen K, Proudfoot NJ (1997) Transcriptional termination signals for RNA polymerase II in fission yeast. *The EMBO journal* **16:** 3633-3643

Birse CE, Minvielle-Sebastia L, Lee BA, Keller W, Proudfoot NJ (1998) Coupling termination of transcription to messenger RNA maturation in yeast. *Science* **280:** 298-301

Biswas SK, Mantovani A (2010) Macrophage plasticity and interaction with lymphocyte subsets: cancer as a paradigm. *Nature immunology* **11:** 889-896

Bolzer A, Kreth G, Solovei I, Koehler D, Saracoglu K, Fauth C, Muller S, Eils R, Cremer C, Speicher MR, Cremer T (2005) Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS biology* **3:** e157

Bousquet-Antonelli C, Presutti C, Tollervey D (2000) Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell* **102:** 765-775

Bradley JR (2008) TNF-mediated inflammatory disease. *The Journal of pathology* **214:** 149-160

Brinkman BM, Telliez JB, Schievella AR, Lin LL, Goldfeld AE (1999) Engagement of tumor necrosis factor (TNF) receptor 1 leads to ATF-2- and p38 mitogen-activated protein kinase-dependent TNF-alpha gene expression. *The Journal of biological chemistry* **274:** 30882-30886

Brown KM, Gilmartin GM (2003) A mechanism for the regulation of pre-mRNA 3' processing by human cleavage factor Im. *Molecular cell* **12:** 1467-1476

Bucheli ME, Buratowski S (2005) Npl3 is an antagonist of mRNA 3' end formation by RNA polymerase II. *The EMBO journal* **24:** 2150-2160

Bucheli ME, He X, Kaplan CD, Moore CL, Buratowski S (2007) Polyadenylation site choice in yeast is affected by competition between Npl3 and polyadenylation factor CFI. *Rna* **13:** 1756-1764

Buckley JM, Wang JH, Redmond HP (2006) Cellular reprogramming by gram-positive bacterial components: a review. *Journal of leukocyte biology* **80:** 731-741

Bulger M, Schubeler D, Bender MA, Hamilton J, Farrell CM, Hardison RC, Groudine M (2003) A complex chromatin landscape revealed by patterns of nuclease sensitivity and histone modification within the mouse beta-globin locus. *Molecular and cellular biology* **23:** 5234-5244

Calvo O, Manley JL (2001) Evolutionarily conserved interaction between CstF-64 and PC4 links transcription, polyadenylation, and termination. *Molecular cell* **7:** 1013-1023

Camblong J, Iglesias N, Fickentscher C, Dieppois G, Stutz F (2007) Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in S. cerevisiae. *Cell* **131:** 706-717

Carey M (1998) The enhanceosome and transcriptional synergy. *Cell* **92:** 5-8

Castelo-Branco P, Furger A, Wollerton M, Smith C, Moreira A, Proudfoot N (2004) Polypyrimidine tract binding protein modulates efficiency of polyadenylation. *Molecular and cellular biology* **24:** 4174-4183

Chao LC, Jamil A, Kim SJ, Huang L, Martinson HG (1999) Assembly of the cleavage and polyadenylation apparatus requires about 10 seconds in vivo and is faster for strong than for weak poly(A) sites. *Molecular and cellular biology* **19:** 5588-5600

Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, Loh YH, Yeo HC, Yeo ZX, Narang V, Govindarajan KR, Leong B, Shahab A, Ruan Y, Bourque G, Sung WK, Clarke ND, Wei CL, Ng HH (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133:** 1106-1117

Cheong C, Matos I, Choi JH, Dandamudi DB, Shrestha E, Longhi MP, Jeffrey KL, Anthony RM, Kluger C, Nchinda G, Koh H, Rodriguez A, Idoyaga J, Pack M, Velinzon K, Park CG, Steinman RM (2010) Microbial stimulation fully differentiates monocytes to DC-SIGN/CD209(+) dendritic cells for immune T cell areas. *Cell* **143:** 416-429

Chernukhin I, Shamsuddin S, Kang SY, Bergstrom R, Kwon YW, Yu W, Whitehead J, Mukhopadhyay R, Docquier F, Farrar D, Morrison I, Vigneron M, Wu SY, Chiang CM, Loukinov D, Lobanenkov V, Ohlsson R, Klenova E (2007) CTCF interacts with and recruits the largest subunit of RNA polymerase II to CTCF target sites genome-wide. *Molecular and cellular biology* **27:** 1631-1648

Chernukhin IV, Shamsuddin S, Robinson AF, Carne AF, Paul A, El-Kady AI, Lobanenkov VV, Klenova EM (2000) Physical and functional interaction between two pluripotent proteins, the Y-box DNA/RNA-binding factor, YB-1, and the multivalent zinc finger factor, CTCF. *The Journal of biological chemistry* **275:** 29915-29921

Christofori G, Keller W (1988) 3' cleavage and polyadenylation of mRNA precursors in vitro requires a poly(A) polymerase, a cleavage factor, and a snRNP. *Cell* **54:** 875-889

Chung JH, Bell AC, Felsenfeld G (1997) Characterization of the chicken beta-globin insulator. *Proceedings of the National Academy of Sciences of the United States of America* **94:** 575-580

Ciosk R, Shirayama M, Shevchenko A, Tanaka T, Toth A, Shevchenko A, Nasmyth K (2000) Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. *Molecular cell* **5:** 243-254

Colgan DF, Manley JL (1997) Mechanism and regulation of mRNA polyadenylation. *Genes & development* **11:** 2755-2766

Collart MA, Baeuerle P, Vassalli P (1990) Regulation of tumor necrosis factor alpha transcription in macrophages: involvement of four kappa B-like motifs and of constitutive and inducible forms of NF-kappa B. *Molecular and cellular biology* **10:** 1498-1506

Connelly S, Manley JL (1988) A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes & development* **2:** 440-452

Cramer P, Armache KJ, Baumli S, Benkert S, Brueckner F, Buchen C, Damsma GE, Dengl S, Geiger SR, Jasiak AJ, Jawhari A, Jennebach S, Kamenski T, Kettenberger H, Kuhn CD, Lehmann E, Leike K, Sydow JF, Vannini A (2008) Structure of eukaryotic RNA polymerases. *Annual review of biophysics* **37:** 337-352

Crawford EK, Ensor JE, Kalvakolanu I, Hasday JD (1997) The role of 3' poly(A) tail metabolism in tumor necrosis factor-alpha regulation. *The Journal of biological chemistry* **272:** 21120-21127

Cuddapah S, Jothi R, Schones DE, Roh TY, Cui K, Zhao K (2009) Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome research* **19:** 24-32

Cui M, Allen MA, Larsen A, Macmorris M, Han M, Blumenthal T (2008) Genes involved in pre-mRNA 3'-end formation and transcription termination revealed by a lin-15 operon Muv suppressor screen. *Proceedings of the National Academy of Sciences of the United States of America* **105:** 16665-16670

Cullen BR (1993) Does HIV-1 Tat induce a change in viral initiation rights? *Cell* **73:** 417-420

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. *The EMBO journal* **19:** 5895-5904

Defossez PA, Kelly KF, Filion GJ, Perez-Torrado R, Magdinier F, Menoni H, Nordgaard CL, Daniel JM, Gilson E (2005) The human enhancer blocker CTC-binding factor interacts with the transcription factor Kaiso. *The Journal of biological chemistry* **280:** 43017-43023

Degner SC, Verma-Gaur J, Wong TP, Bossen C, Iverson GM, Torkamani A, Vettermann C, Lin YC, Ju Z, Schulz D, Murre CS, Birshtein BK, Schork NJ,

Schlissel MS, Riblet R, Murre C, Feeney AJ (2011) CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the Igh locus and antisense transcription in pro-B cells. *Proceedings of the National Academy of Sciences of the United States of America* **108:** 9566-9571

Dekker J, Rippe K, Dekker M, Kleckner N (2002) Capturing chromosome conformation. *Science* **295:** 1306-1311

Dermody JL, Dreyfuss JM, Villen J, Ogundipe B, Gygi SP, Park PJ, Ponticelli AS, Moore CL, Buratowski S, Bucheli ME (2008) Unphosphorylated SR-like protein Npl3 stimulates RNA polymerase II elongation. *PloS one* **3:** e3273

Dettwiler S, Aringhieri C, Cardinale S, Keller W, Barabino SM (2004) Distinct sequence motifs within the 68-kDa subunit of cleavage factor Im mediate RNA binding, protein-protein interactions, and subcellular localization. *The Journal of biological chemistry* **279:** 35788-35797

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

Donohoe ME, Zhang LF, Xu N, Shi Y, Lee JT (2007) Identification of a Ctcf cofactor, Yy1, for the X chromosome binary switch. *Molecular cell* **25:** 43-56

Dorsett D (2007) Roles of the sister chromatid cohesion apparatus in gene expression, development, and human syndromes. *Chromosoma* **116:** 1-13

Dostie J, Richmond TA, Arnaout RA, Selzer RR, Lee WL, Honan TA, Rubio ED, Krumm A, Lamb J, Nusbaum C, Green RD, Dekker J (2006) Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome research* **16:** 1299-1309

Drouet C, Shakhov AN, Jongeneel CV (1991) Enhancers and transcription factors controlling the inducibility of the tumor necrosis factor-alpha promoter in primary macrophages. *Journal of immunology* **147:** 1694-1700

Dye MJ, Proudfoot NJ (1999) Terminal exon definition occurs cotranscriptionally and promotes termination of RNA polymerase II. *Molecular cell* **3:** 371-378

Dye MJ, Proudfoot NJ (2001) Multiple transcript cleavage precedes polymerase release in termination by RNA polymerase II. *Cell* **105:** 669-681

Enriquez-Harris P, Levitt N, Briggs D, Proudfoot NJ (1991) A pause site for RNA polymerase II is associated with termination of transcription. *The EMBO journal* **10:** 1833-1842

Esensten JH, Tsytsykova AV, Lopez-Rodriguez C, Ligeiro FA, Rao A, Goldfeld AE (2005) NFAT5 binds to the TNF promoter distinctly from NFATp,

c, 3 and 4, and activates TNF transcription during hypertonic stress alone. *Nucleic acids research* **33:** 3845-3854

Falvo JV, Brinkman BM, Tsytsykova AV, Tsai EY, Yao TP, Kung AL, Goldfeld AE (2000a) A stimulus-specific role for CREB-binding protein (CBP) in T cell receptor-activated tumor necrosis factor alpha gene expression. *Proceedings of the National Academy of Sciences of the United States of America* **97:** 3925-3929

Falvo JV, Tsytsykova AV, Goldfeld AE (2010) Transcriptional control of the TNF gene. *Current directions in autoimmunity* **11:** 27-60

Falvo JV, Uglialoro AM, Brinkman BM, Merika M, Parekh BS, Tsai EY, King HC, Morielli AD, Peralta EG, Maniatis T, Thanos D, Goldfeld AE (2000b) Stimulus-specific assembly of enhancer complexes on the tumor necrosis factor alpha gene promoter. *Molecular and cellular biology* **20:** 2239-2247

Farrell CM, West AG, Felsenfeld G (2002) Conserved CTCF insulator elements flank the mouse and human beta-globin loci. *Molecular and cellular biology* **22:** 3820-3831

Fay A, Misulovin Z, Li J, Schaaf CA, Gause M, Gilmour DS, Dorsett D (2011) Cohesin selectively binds and regulates genes with paused RNA polymerase. *Current biology : CB* **21:** 1624-1634

Filippova GN, Cheng MK, Moore JM, Truong JP, Hu YJ, Nguyen DK, Tsuchiya KD, Disteche CM (2005) Boundaries between chromosomal domains of X inactivation and escape bind CTCF and lack CpG methylation during early development. *Developmental cell* **8:** 31-42

Filippova GN, Fagerlie S, Klenova EM, Myers C, Dehner Y, Goodwin G, Neiman PE, Collins SJ, Lobanenkov VV (1996) An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Molecular and cellular biology* **16:** 2802-2813

Filippova GN, Thienes CP, Penn BH, Cho DH, Hu YJ, Moore JM, Klesert TR, Lobanenkov VV, Tapscott SJ (2001) CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nature genetics* **28:** 335-343

Flaherty SM, Fortes P, Izaurralde E, Mattaj IW, Gilmartin GM (1997) Participation of the nuclear cap binding complex in pre-mRNA 3' processing. *Proceedings of the National Academy of Sciences of the United States of America* **94:** 11893-11898

Fogg DK, Sibon C, Miled C, Jung S, Aucouturier P, Littman DR, Cumano A, Geissmann F (2006) A clonogenic bone marrow progenitor specific for macrophages and dendritic cells. *Science* **311:** 83-87

Francastel C, Schubeler D, Martin DI, Groudine M (2000) Nuclear compartmentalization and gene activity. *Nature reviews Molecular cell biology* **1:** 137-143

Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, Orlov YL, Velkov S, Ho A, Mei PH, Chew EG, Huang PY, Welboren WJ, Han Y, Ooi HS, Ariyaratne PN, Vega VB, Luo Y, Tan PY, Choy PY, Wansa KD, Zhao B, Lim KS, Leow SC, Yow JS, Joseph R, Li H, Desai KV, Thomsen JS, Lee YK, Karuturi RK, Herve T, Bourque G, Stunnenberg HG, Ruan X, Cacheux-Rataboul V, Sung WK, Liu ET, Wei CL, Cheung E, Ruan Y (2009) An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature* **462:** 58-64

Garas M, Dichtl B, Keller W (2008) The role of the putative 3' end processing endonuclease Ysh1p in mRNA and snoRNA synthesis. *Rna* **14:** 2671-2684

Gause M, Schaaf CA, Dorsett D (2008) Cohesin and CTCF: cooperating to control chromosome conformation? *BioEssays : news and reviews in molecular, cellular and developmental biology* **30:** 715-718

Geissmann F, Jung S, Littman DR (2003) Blood monocytes consist of two principal subsets with distinct migratory properties. *Immunity* **19:** 71-82

Geissmann F, Manz MG, Jung S, Sieweke MH, Merad M, Ley K (2010) Development of monocytes, macrophages, and dendritic cells. *Science* **327:** 656-661

Gerasimova TI, Byrd K, Corces VG (2000) A chromatin insulator determines the nuclear localization of DNA. *Molecular cell* **6:** 1025-1035

Glover-Cutter K, Kim S, Espinosa J, Bentley DL (2008) RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. *Nature structural & molecular biology* **15:** 71-78

Glynn EF, Megee PC, Yu HG, Mistrot C, Unal E, Koshland DE, DeRisi JL, Gerton JL (2004) Genome-wide mapping of the cohesin complex in the yeast Saccharomyces cerevisiae. *PLoS biology* **2:** E259

Goldfeld AE, Doyle C, Maniatis T (1990) Human tumor necrosis factor alpha gene regulation by virus and lipopolysaccharide. *Proceedings of the National Academy of Sciences of the United States of America* **87:** 9769-9773

Goldfeld AE, McCaffrey PG, Strominger JL, Rao A (1993) Identification of a novel cyclosporin-sensitive element in the human tumor necrosis factor alpha gene promoter. *The Journal of experimental medicine* **178:** 1365-1379

Gordon S (2003) Alternative activation of macrophages. *Nature reviews Immunology* **3:** 23-35

Gordon S (2007) The macrophage: past, present and future. *European journal of immunology* **37 Suppl 1:** S9-17

Gordon S, Martinez FO (2010) Alternative activation of macrophages: mechanism and functions. *Immunity* **32:** 593-604

Gordon S, Taylor PR (2005) Monocyte and macrophage heterogeneity. *Nature reviews Immunology* **5:** 953-964

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

Grivennikov SI, Kuprash DV, Liu ZG, Nedospasov SA (2006) Intracellular signals and events activated by cytokines of the tumor necrosis factor superfamily: From simple paradigms to complex mechanisms. *International review of cytology* **252:** 129-161

Gromak N, West S, Proudfoot NJ (2006) Pause sites promote transcriptional termination of mammalian RNA polymerase II. *Molecular and cellular biology* **26:** 3986-3996

Gruber S, Arumugam P, Katou Y, Kuglitsch D, Helmhart W, Shirahige K, Nasmyth K (2006) Evidence that loading of cohesin onto chromosomes involves opening of its SMC hinge. *Cell* **127:** 523-537

Gruber S, Haering CH, Nasmyth K (2003) Chromosomal cohesin forms a ring. *Cell* **112:** 765-777

(Gullerova M, Proudfoot NJ (2008) Cohesin complex promotes transcriptional termination between convergent genes in S. pombe. *Cell* **132:** 983-995

Guo C, Gerasimova T, Hao H, Ivanova I, Chakraborty T, Selimyan R, Oltz EM, Sen R (2011a) Two forms of loops generate the chromatin conformation of the immunoglobulin heavy-chain gene locus. *Cell* **147:** 332-343

Guo C, Yoon HS, Franklin A, Jain S, Ebert A, Cheng HL, Hansen E, Despo O, Bossen C, Vettermann C, Bates JG, Richards N, Myers D, Patel H, Gallagher M, Schlissel MS, Murre C, Busslinger M, Giallourakis CC, Alt FW (2011b) CTCF-binding elements mediate control of V(D)J recombination. *Nature* **477:** 424-430

Hadjur S, Williams LM, Ryan NK, Cobb BS, Sexton T, Fraser P, Fisher AG, Merkenschlager M (2009) Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature* **460:** 410-413

Haering CH, Lowe J, Hochwagen A, Nasmyth K (2002) Molecular architecture of SMC proteins and the yeast cohesin complex. *Molecular cell* **9:** 773-788

Handoko L, Xu H, Li G, Ngan CY, Chew E, Schnapp M, Lee CW, Ye C, Ping JL, Mulawadi F, Wong E, Sheng J, Zhang Y, Poh T, Chan CS, Kunarso G, Shahab A, Bourque G, Cacheux-Rataboul V, Sung WK, Ruan Y, Wei CL (2011) CTCF-mediated functional chromatin interactome in pluripotent cells. *Nature genetics* **43:** 630-638

Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM (2000) CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* **405:** 486-489

Hayden MS, Ghosh S (2008) Shared principles in NF-kappaB signaling. *Cell* **132:** 344-362

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

Hirano M, Hirano T (2006) Opening closed arms: long-distance activation of SMC ATPase by hinge-DNA interactions. *Molecular cell* **21:** 175-186

Hirano T (2006) At the heart of the chromosome: SMC proteins in action. *Nat Rev Mol Cell Biol* **7:** 311-322

Hirose Y, Manley JL (1998) RNA polymerase II is an essential mRNA polyadenylation factor. *Nature* **395:** 93-96

Hofmann I, Schnolzer M, Kaufmann I, Franke WW (2002) Symplekin, a constitutive protein of karyo- and cytoplasmic particles involved in mRNA biogenesis in Xenopus laevis oocytes. *Molecular biology of the cell* **13:** 1665-1676

Hou C, Dale R, Dean A (2010) Cell type specificity of chromatin organization mediated by CTCF and cohesin. *Proceedings of the National Academy of Sciences of the United States of America* **107:** 3651-3656

Hou C, Zhao H, Tanimoto K, Dean A (2008) CTCF-dependent enhancer-blocking by alternative chromatin loop formation. *Proceedings of the National Academy of Sciences of the United States of America* **105:** 20398-20403

Hu J, Lutz CS, Wilusz J, Tian B (2005) Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation. *Rna* **11:** 1485-1493

Huang J, Brito IL, Villen J, Gygi SP, Amon A, Moazed D (2006) Inhibition of homologous recombination by a cohesin-associated clamp complex recruited to the rDNA recombination enhancer. *Genes Dev* **20:** 2887-2901

Huang Y, Carmichael GG (1996) Role of polyadenylation in nucleocytoplasmic transport of mRNA. *Molecular and cellular biology* **16:** 1534-1542

Introna M, Hamilton TA, Kaufman RE, Adams DO, Bast RC, Jr. (1986) Treatment of murine peritoneal macrophages with bacterial lipopolysaccharide alters expression of c-fos and c-myc oncogenes. *Journal of immunology* **137:** 2711-2715

Ishihara K, Oshimura M, Nakao M (2006) CTCF-dependent chromatin insulator is linked to epigenetic remodeling. *Molecular cell* **23:** 733-742

Jacobs M, Samarina A, Grivennikov S, Botha T, Allie N, Fremond C, Togbe D, Vasseur V, Rose S, Erard F, Monteiro A, Quesniaux V, Ryffel B (2007) Reactivation of tuberculosis by tumor necrosis factor neutralization. *European cytokine network* **18:** 5-13

Jimeno-Gonzalez S, Haaning LL, Malagon F, Jensen TH The yeast 5'-3' exonuclease Rat1p functions during transcription elongation by RNA polymerase II. *Molecular cell* **37:** 580-587

Jothi R, Cuddapah S, Barski A, Cui K, Zhao K (2008) Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data. *Nucleic acids research* **36:** 5221-5231

Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, Taatjes DJ, Dekker J, Young RA (2010) Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467:** 430-435

Kanduri C, Holmgren C, Pilartz M, Franklin G, Kanduri M, Liu L, Ginjala V, Ulleras E, Mattsson R, Ohlsson R (2000) The 5' flank of mouse H19 in an unusual chromatin conformation unidirectionally blocks enhancer-promoter communication. *Current biology : CB* **10:** 449-457

Kaneko S, Rozenblatt-Rosen O, Meyerson M, Manley JL (2007) The multifunctional protein p54nrb/PSF recruits the exonuclease XRN2 to facilitate pre-mRNA 3' processing and transcription termination. *Genes & development* **21:** 1779-1789

Kaufmann I, Martin G, Friedlein A, Langen H, Keller W (2004) Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *The EMBO journal* **23:** 616-626

Kawauchi J, Mischo H, Braglia P, Rondon A, Proudfoot NJ (2008) Budding yeast RNA polymerases I and II employ parallel mechanisms of transcriptional termination. *Genes & development* **22:** 1082-1092

Kawauchi S, Calof AL, Santos R, Lopez-Burks ME, Young CM, Hoang MP, Chua A, Lao T, Lechner MS, Daniel JA, Nussenzweig A, Kitzes L, Yokomori K, Hallgrimsson B, Lander AD (2009) Multiple organ system defects and transcriptional dysregulation in the Nipbl(+/-) mouse, a model of Cornelia de Lange Syndrome. *PLoS genetics* **5:** e1000650

Keller W, Bienroth S, Lang KM, Christofori G (1991) Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3' processing signal AAUAAA. *The EMBO journal* **10:** 4241-4249

Kerwitz Y, Kuhn U, Lilie H, Knoth A, Scheuermann T, Friedrich H, Schwarz E, Wahle E (2003) Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. *The EMBO journal* **22:** 3705-3714

Kim HD, Choe J, Seo YS (1999) The sen1(+) gene of Schizosaccharomyces pombe, a homologue of budding yeast SEN1, encodes an RNA and DNA helicase. *Biochemistry* **38:** 14697-14710

Kim M, Ahn SH, Krogan NJ, Greenblatt JF, Buratowski S (2004a) Transitions in RNA polymerase II elongation complexes at the 3' ends of genes. *The EMBO journal* **23:** 354-364

Kim M, Krogan NJ, Vasiljeva L, Rando OJ, Nedea E, Greenblatt JF, Buratowski S (2004b) The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* **432:** 517-522

Kim M, Vasiljeva L, Rando OJ, Zhelkovsky A, Moore C, Buratowski S (2006) Distinct pathways for snoRNA and mRNA termination. *Molecular cell* **24:** 723-734

Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green RD, Zhang MQ, Lobanenkov VV, Ren B (2007) Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* **128:** 1231-1245

Klein F, Mahr P, Galova M, Buonomo SB, Michaelis C, Nairz K, Nasmyth K (1999) A central role for cohesins in sister chromatid cohesion, formation of axial elements, and recombination during yeast meiosis. *Cell* **98:** 91-103

Klenova EM, Nicolas RH, Paterson HF, Carne AF, Heath CM, Goodwin GH, Neiman PE, Lobanenkov VV (1993) CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms. *Molecular and cellular biology* **13:** 7612-7624

Kobayashi T, Ganley AR (2005) Recombination regulation by transcription-induced cohesin dissociation in rDNA repeats. *Science* **309:** 1581-1584

Kohne AC, Baniahmad A, Renkawitz R (1993) NeP1. A ubiquitous transcription factor synergizes with v-ERBA in transcriptional silencing. *Journal of molecular biology* **232:** 747-755

Krantz ID, McCallum J, DeScipio C, Kaur M, Gillis LA, Yaeger D, Jukofsky L, Wasserman N, Bottani A, Morris CA, Nowaczyk MJ, Toriello H, Bamshad MJ, Carey JC, Rappaport E, Kawauchi S, Lander AD, Calof AL, Li HH, Devoto M, Jackson LG (2004) Cornelia de Lange syndrome is caused by mutations in

NIPBL, the human homolog of Drosophila melanogaster Nipped-B. *Nature genetics* **36:** 631-635

Krausgruber T, Blazek K, Smallie T, Alzabin S, Lockstone H, Sahgal N, Hussell T, Feldmann M, Udalova IA (2011) IRF5 promotes inflammatory macrophage polarization and TH1-TH17 responses. *Nature immunology* **12:** 231-238

Kuhn U, Gundel M, Knoth A, Kerwitz Y, Rudel S, Wahle E (2009) Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. *The Journal of biological chemistry* **284:** 22803-22814

Kuprash DV, Udalova IA, Turetskaya RL, Kwiatkowski D, Rice NR, Nedospasov SA (1999) Similarities and differences between human and murine TNF promoters in their response to lipopolysaccharide. *Journal of immunology* **162:** 4045-4052

Kurukuti S, Tiwari VK, Tavoosidana G, Pugacheva E, Murrell A, Zhao Z, Lobanenkov V, Reik W, Ohlsson R (2006) CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proceedings of the National Academy of Sciences of the United States of America* **103:** 10684-10689

Kyburz A, Friedlein A, Langen H, Keller W (2006) Direct interactions between subunits of CPSF and the U2 snRNP contribute to the coupling of pre-mRNA 3' end processing and splicing. *Molecular cell* **23:** 195-205

Lee BK, Bhinge AA, Battenhouse A, McDaniell RM, Liu Z, Song L, Ni Y, Birney E, Lieb JD, Furey TS, Crawford GE, Iyer VR (2012) Cell-type specific and combinatorial usage of diverse transcription factors revealed by genome-wide binding studies in multiple human cells. *Genome research* **22:** 9-24

Lehmann AR (2005) The role of SMC proteins in the responses to DNA damage. *DNA Repair (Amst)* **4:** 309-314

Leitman DC, Mackow ER, Williams T, Baxter JD, West BL (1992) The core promoter region of the tumor necrosis factor alpha gene confers phorbol ester responsiveness to upstream transcriptional activators. *Molecular and cellular biology* **12:** 1352-1356

Lengronne A, Katou Y, Mori S, Yokobayashi S, Kelly GP, Itoh T, Watanabe Y, Shirahige K, Uhlmann F (2004) Cohesin relocation from sites of chromosomal loading to places of convergent transcription. *Nature* **430:** 573-578

Leung JY, McKenzie FE, Uglialoro AM, Flores-Villanueva PO, Sorkin BC, Yunis EJ, Hartl DL, Goldfeld AE (2000) Identification of phylogenetic footprints in primate tumor necrosis factor-alpha promoters. *Proceedings of the National Academy of Sciences of the United States of America* **97:** 6614-6618

Li T, Hu JF, Qiu X, Ling J, Chen H, Wang S, Hou A, Vu TH, Hoffman AR (2008) CTCF regulates allelic expression of Igf2 by orchestrating a promoter-polycomb repressive complex 2 intrachromosomal loop. *Molecular and cellular biology* **28:** 6473-6482

Liang S, Lutz CS (2006) p54nrb is a component of the snRNP-free U1A (SF-A) complex that promotes pre-mRNA cleavage during polyadenylation. *Rna* **12:** 111-121

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. *Molecular cell* **9:** 1101-1111

Ling JQ, Li T, Hu JF, Vu TH, Chen HL, Qiu XW, Cherry AM, Hoffman AR (2006) CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. *Science* **312:** 269-272

Liu J, Krantz ID (2008) Cohesin and human disease. *Annual review of genomics and human genetics* **9:** 303-320

Liu J, Zhang Z, Bando M, Itoh T, Deardorff MA, Clark D, Kaur M, Tandy S, Kondoh T, Rappaport E, Spinner NB, Vega H, Jackson LG, Shirahige K, Krantz ID (2009) Transcriptional dysregulation in NIPBL and cohesin mutant human cells. *PLoS biology* **7:** e1000119

Liu Z, Scannell DR, Eisen MB, Tjian R (2011) Control of embryonic stem cell lineage commitment by core promoter factor, TAF3. *Cell* **146:** 720-731

Lobanenkov VV, Nicolas RH, Adler VV, Paterson H, Klenova EM, Polotskaja AV, Goodwin GH (1990) A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. *Oncogene* **5:** 1743-1753

Logan J, Falck-Pedersen E, Darnell JE, Jr., Shenk T (1987) A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse beta maj-globin gene. *Proceedings of the National Academy of Sciences of the United States of America* **84:** 8306-8310

Lou J, Lucas R, Grau GE (2001) Pathogenesis of cerebral malaria: recent experimental data and possible applications for humans. *Clinical microbiology reviews* **14:** 810-820, table of contents

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 & development* **20:** 954-965

Lutz M, Burke LJ, Barreto G, Goeman F, Greb H, Arnold R, Schultheiss H, Brehm A, Kouzarides T, Lobanenkov V, Renkawitz R (2000) Transcriptional repression by the insulator protein CTCF involves histone deacetylases. *Nucleic acids research* **28:** 1707-1713

MacCallum DE, Melville J, Frame S, Watt K, Anderson S, Gianella-Borradori A, Lane DP, Green SR (2005) Seliciclib (CYC202, R-Roscovitine) induces cell death in multiple myeloma cells by inhibition of RNA polymerase II-dependent transcription and down-regulation of Mcl-1. *Cancer research* **65:** 5399-5407

Maciolek NL, McNally MT (2007) Serine/arginine-rich proteins contribute to negative regulator of splicing element-stimulated polyadenylation in rous sarcoma virus. *Journal of virology* **81:** 11208-11217

Majumder P, Gomez JA, Chadwick BP, Boss JM (2008) The insulator factor CTCF controls MHC class II gene expression and is required for the formation of long-distance chromatin interactions. *The Journal of experimental medicine* **205:** 785-798

Mandel CR, Bai Y, Tong L (2008) Protein factors in pre-mRNA 3'-end processing. *Cellular and molecular life sciences : CMLS* **65:** 1099-1122

Mandel CR, Kaneko S, Zhang H, Gebauer D, Vethantham V, Manley JL, Tong L (2006) Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature* **444:** 953-956

Mantovani A, Sica A (2010) Macrophages, innate immunity and cancer: balance, tolerance, and diversity. *Current opinion in immunology* **22:** 231-237

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

Martin AP, Rankin S, Pitchford S, Charo IF, Furtado GC, Lira SA (2008) Increased expression of CCL2 in insulin-producing cells of transgenic mice promotes mobilization of myeloid cells from the bone marrow, marked insulitis, and diabetes. *Diabetes* **57:** 3025-3033

Martinez FO, Sica A, Mantovani A, Locati M (2008) Macrophage activation and polarization. *Frontiers in bioscience : a journal and virtual library* **13:** 453-461

Matlin AJ, Moore MJ (2007) Spliceosome assembly and composition. *Advances in experimental medicine and biology* **623:** 14-35

Mc Intyre J, Muller EG, Weitzer S, Snydsman BE, Davis TN, Uhlmann F (2007) In vivo analysis of cohesin architecture using FRET in the budding yeast Saccharomyces cerevisiae. *Embo J* **26:** 3783-3793

McCracken S, Fong N, Yankulov K, Ballantyne S, Pan G, Greenblatt J, Patterson SD, Wickens M, Bentley DL (1997) The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* **385:** 357-361

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

Meinhart A, Kamenski T, Hoeppner S, Baumli S, Cramer P (2005) A structural perspective of CTD function. *Genes & development* **19:** 1401-1415

Merika M, Thanos D (2001) Enhanceosomes. *Current opinion in genetics & development* **11:** 205-208

Meshorer E, Misteli T (2006) Chromatin in pluripotent embryonic stem cells and differentiation. *Nature reviews Molecular cell biology* **7:** 540-546

Milutinovich M, Unal E, Ward C, Skibbens RV, Koshland D (2007) A multi-step pathway for the establishment of sister chromatid cohesion. *PLoS Genet* **3:** e12

Minvielle-Sebastia L, Preker PJ, Wiederkehr T, Strahm Y, Keller W (1997) The major yeast poly(A)-binding protein is associated with cleavage factor IA and functions in premessenger RNA 3'-end formation. *Proceedings of the National Academy of Sciences of the United States of America* **94:** 7897-7902

Mischo HE, Gomez-Gonzalez B, Grzechnik P, Rondon AG, Wei W, Steinmetz L, Aguilera A, Proudfoot NJ Yeast Sen1 helicase protects the genome from transcription-associated instability. *Molecular cell* **41:** 21-32

Mishiro T, Ishihara K, Hino S, Tsutsumi S, Aburatani H, Shirahige K, Kinoshita Y, Nakao M (2009) Architectural roles of multiple chromatin insulators at the human apolipoprotein gene cluster. *The EMBO journal* **28:** 1234-1245

Misteli T (2007) Beyond the sequence: cellular organization of genome function. *Cell* **128:** 787-800

Misulovin Z, Schwartz YB, Li XY, Kahn TG, Gause M, MacArthur S, Fay JC, Eisen MB, Pirrotta V, Biggin MD, Dorsett D (2008) Association of cohesin and Nipped-B with transcriptionally active regions of the Drosophila melanogaster genome. *Chromosoma* **117:** 89-102

Moreira A, Takagaki Y, Brackenridge S, Wollerton M, Manley JL, Proudfoot NJ (1998) The upstream sequence element of the C2 complement poly(A) signal activates mRNA 3' end formation by two distinct mechanisms. *Genes & development* **12:** 2522-2534

Mosser DM, Edwards JP (2008) Exploring the full spectrum of macrophage activation. *Nature reviews Immunology* **8:** 958-969

Murrell A, Heeson S, Cooper WN, Douglas E, Apostolidou S, Moore GE, Maher ER, Reik W (2004) An association between variants in the IGF2 gene and Beckwith-Wiedemann syndrome: interaction between genotype and epigenotype. *Human molecular genetics* **13:** 247-255

Murthy KG, Manley JL (1992) Characterization of the multisubunit cleavage-polyadenylation specificity factor from calf thymus. *The Journal of biological chemistry* **267:** 14804-14811

Murthy KG, Manley JL (1995) The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. *Genes & development* **9:** 2672-2683

Musio A, Selicorni A, Focarelli ML, Gervasini C, Milani D, Russo S, Vezzoni P, Larizza L (2006) X-linked Cornelia de Lange syndrome owing to SMC1L1 mutations. *Nature genetics* **38:** 528-530

Nag A, Narsinh K, Kazerouninia A, Martinson HG (2006) The conserved AAUAAA hexamer of the poly(A) signal can act alone to trigger a stable decrease in RNA polymerase II transcription velocity. *Rna* **12:** 1534-1544

Nag A, Narsinh K, Martinson HG (2007) The poly(A)-dependent transcriptional pause is mediated by CPSF acting on the body of the polymerase. *Nature structural & molecular biology* **14:** 662-669

Nasmyth K (2005) How might cohesin hold sister chromatids together? *Philos Trans R Soc Lond B Biol Sci* **360:** 483-496

Nasmyth K, Peters JM, Uhlmann F (2000) Splitting the chromosome: cutting the ties that bind sister chromatids. *Science* **288:** 1379-1385

Nativio R, Wendt KS, Ito Y, Huddleston JE, Uribe-Lewis S, Woodfine K, Krueger C, Reik W, Peters JM, Murrell A (2009) Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. *PLoS genetics* **5:** e1000739

Nemunaitis J (1993) Macrophage function activating cytokines: potential clinical application. *Critical reviews in oncology/hematology* **14:** 153-171

Nitzsche A, Paszkowski-Rogacz M, Matarese F, Janssen-Megens EM, Hubner NC, Schulz H, de Vries I, Ding L, Huebner N, Mann M, Stunnenberg HG, Buchholz F (2011) RAD21 cooperates with pluripotency transcription factors in the maintenance of embryonic stem cell identity. *PloS one* **6:** e19470

Niwa M, Berget SM (1991) Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns. *Genes & development* **5:** 2086-2095

Niwa M, MacDonald CC, Berget SM (1992) Are vertebrate exons scanned during splice-site selection? *Nature* **360:** 277-280

O'Sullivan T, Saddawi-Konefka R, Vermi W, Koebel CM, Arthur C, White JM, Uppaluri R, Andrews DM, Ngiow SF, Teng MW, Smyth MJ, Schreiber RD, Bui JD (2012) Cancer immunoediting by the innate immune system in the absence of adaptive immunity. *The Journal of experimental medicine* **209:** 1869-1882

Ohlsson R, Renkawitz R, Lobanenkov V (2001) CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends Genet* **17:** 520-527

Orozco IJ, Kim SJ, Martinson HG (2002) The poly(A) signal, without the assistance of any downstream element, directs RNA polymerase II to pause in vivo and then to release stochastically from the template. *The Journal of biological chemistry* **277:** 42899-42911

Palstra RJ, Tolhuis B, Splinter E, Nijmeijer R, Grosveld F, de Laat W (2003) The beta-globin nuclear compartment in development and erythroid differentiation. *Nature genetics* **35:** 190-194

Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, Jarmuz A, Canzonetta C, Webster Z, Nesterova T, Cobb BS, Yokomori K, Dillon N, Aragon L, Fisher AG, Merkenschlager M (2008) Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132:** 422-433

Paule MR, White RJ (2000) Survey and summary: transcription by RNA polymerases I and III. *Nucleic acids research* **28:** 1283-1298

Perkins ND (2007) Integrating cell-signalling pathways with NF-kappaB and IKK function. *Nature reviews Molecular cell biology* **8:** 49-62

Phillips JE, Corces VG (2009) CTCF: master weaver of the genome. *Cell* **137:** 1194-1211

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. *Molecular and cellular biology* **25:** 3276-3285

Pollard JW (2009) Trophic macrophages in development and disease. *Nature reviews Immunology* **9:** 259-270

Proudfoot NJ, Brownlee GG (1976) 3' non-coding region sequences in eukaryotic messenger RNA. *Nature* **263:** 211-214

Proudfoot NJ, Furger A, Dye MJ (2002) Integrating mRNA processing with transcription. *Cell* **108:** 501-512

Qi CF, Martensson A, Mattioli M, Dalla-Favera R, Lobanenkov VV, Morse HC, 3rd (2003) CTCF functions as a critical regulator of cell-cycle arrest and death after ligation of the B cell receptor on immature B cells. *Proceedings of the National Academy of Sciences of the United States of America* **100:** 633-638

Qian BZ, Pollard JW (2010) Macrophage diversity enhances tumor progression and metastasis. *Cell* **141:** 39-51

Reitman M, Felsenfeld G (1988) Mutational analysis of the chicken beta-globin enhancer reveals two positive-acting domains. *Proceedings of the National Academy of Sciences of the United States of America* **85:** 6267-6271

Renda M, Baglivo I, Burgess-Beusse B, Esposito S, Fattorusso R, Felsenfeld G, Pedone PV (2007) Critical DNA binding interactions of the insulator protein CTCF: a small number of zinc fingers mediate strong binding, and a single finger-DNA interaction controls binding at imprinted loci. *The Journal of biological chemistry* **282:** 33336-33345

Rhodes JM, Bentley FK, Print CG, Dorsett D, Misulovin Z, Dickinson EJ, Crosier KE, Crosier PS, Horsfield JA (2010) Positive regulation of c-Myc by cohesin is direct, and evolutionarily conserved. *Developmental biology* **344:** 637-649

Richards DM, Hettinger J, Feuerer M (2012) Monocytes and Macrophages in Cancer: Development and Functions. *Cancer microenvironment : official journal of the International Cancer Microenvironment Society*

Rondon AG, Gallardo M, Garcia-Rubio M, Aguilera A (2004) Molecular evidence indicating that the yeast PAF complex is required for transcription elongation. *EMBO reports* **5:** 47-53

Rosonina E, Ip JY, Calarco JA, Bakowski MA, Emili A, McCracken S, Tucker P, Ingles CJ, Blencowe BJ (2005) Role for PSF in mediating transcriptional activator-dependent stimulation of pre-mRNA processing in vivo. *Molecular and cellular biology* **25:** 6734-6746

Rubio ED, Reiss DJ, Welcsh PL, Disteche CM, Filippova GN, Baliga NS, Aebersold R, Ranish JA, Krumm A (2008) CTCF physically links cohesin to chromatin. *Proceedings of the National Academy of Sciences of the United States of America* **105:** 8309-8314

Ruegsegger U, Beyer K, Keller W (1996) Purification and characterization of human cleavage factor Im involved in the 3' end processing of messenger RNA precursors. *The Journal of biological chemistry* **271:** 6107-6113

Ryan K, Calvo O, Manley JL (2004) Evidence that polyadenylation factor CPSF-73 is the mRNA 3' processing endonuclease. *Rna* **10:** 565-573

Sachs AB, Sarnow P, Hentze MW (1997) Starting at the beginning, middle, and end: translation initiation in eukaryotes. *Cell* **89:** 831-838

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

Saguez C, Schmid M, Olesen JR, Ghazy MA, Qu X, Poulsen MB, Nasser T, Moore C, Jensen TH (2008) Nuclear mRNA surveillance in THO/sub2 mutants is triggered by inefficient polyadenylation. *Molecular cell* **31:** 91-103

Saitoh N, Bell AC, Recillas-Targa F, West AG, Simpson M, Pikaart M, Felsenfeld G (2000) Structural and functional conservation at the boundaries of the chicken beta-globin domain. *The EMBO journal* **19:** 2315-2322

Sandhu KS, Shi C, Sjolinder M, Zhao Z, Gondor A, Liu L, Tiwari VK, Guibert S, Emilsson L, Imreh MP, Ohlsson R (2009) Nonallelic transvection of multiple imprinted loci is organized by the H19 imprinting control region during germline development. *Genes & development* **23:** 2598-2603

Schaaf CA, Misulovin Z, Sahota G, Siddiqui AM, Schwartz YB, Kahn TG, Pirrotta V, Gause M, Dorsett D (2009) Regulation of the Drosophila Enhancer of split and invected-engrailed gene complexes by sister chromatid cohesion proteins. *PloS one* **4:** e6202

Schafer M, Werner S (2008) Cancer as an overhealing wound: an old hypothesis revisited. *Nature reviews Molecular cell biology* **9:** 628-638

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 research* **20:** 578-588

Schulz C, Gomez Perdiguero E, Chorro L, Szabo-Rogers H, Cagnard N, Kierdorf K, Prinz M, Wu B, Jacobsen SE, Pollard JW, Frampton J, Liu KJ, Geissmann F (2012) A lineage of myeloid cells independent of Myb and hematopoietic stem cells. *Science* **336:** 86-90

Scotto-Lavino E, Du G, Frohman MA (2006) 3' end cDNA amplification using classic RACE. *Nature protocols* **1:** 2742-2745

Seitan VC, Hao B, Tachibana-Konwalski K, Lavagnolli T, Mira-Bontenbal H, Brown KE, Teng G, Carroll T, Terry A, Horan K, Marks H, Adams DJ, Schatz DG, Aragon L, Fisher AG, Krangel MS, Nasmyth K, Merkenschlager M (2011) A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. *Nature* **476:** 467-471

Serbina NV, Pamer EG (2006) Monocyte emigration from bone marrow during bacterial infection requires signals mediated by chemokine receptor CCR2. *Nature immunology* **7:** 311-317

Sethi G, Sung B, Aggarwal BB (2008) TNF: a master switch for inflammation to cancer. *Frontiers in bioscience : a journal and virtual library* **13:** 5094-5107

Shakhov AN, Collart MA, Vassalli P, Nedospasov SA, Jongeneel CV (1990) Kappa B-type enhancers are involved in lipopolysaccharide-mediated transcriptional activation of the tumor necrosis factor alpha gene in primary macrophages. *The Journal of experimental medicine* **171:** 35-47

Sheets MD, Ogg SC, Wickens MP (1990) Point mutations in AAUAAA and the poly (A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation in vitro. *Nucleic acids research* **18:** 5799-5805

Shell SA, Hesse C, Morris SM, Jr., Milcarek C (2005) Elevated levels of the 64-kDa cleavage stimulatory factor (CstF-64) in lipopolysaccharide-stimulated macrophages influence gene expression and induce alternative poly(A) site selection. *The Journal of biological chemistry* **280:** 39950-39961

Shi C, Jia T, Mendez-Ferrer S, Hohl TM, Serbina NV, Lipuma L, Leiner I, Li MO, Frenette PS, Pamer EG (2011) Bone marrow mesenchymal stem and progenitor cells induce monocyte emigration in response to circulating toll-like receptor ligands. *Immunity* **34:** 590-601

Shi C, Pamer EG (2011) Monocyte recruitment during infection and inflammation. *Nature reviews Immunology* **11:** 762-774

Shi Y, Di Giammartino DC, Taylor D, Sarkeshik A, Rice WJ, Yates JR, 3rd, Frank J, Manley JL (2009) Molecular architecture of the human pre-mRNA 3' processing complex. *Molecular cell* **33:** 365-376

Shukla S, Kavak E, Gregory M, Imashimizu M, Shutinoski B, Kashlev M, Oberdoerffer P, Sandberg R, Oberdoerffer S (2011) CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* **479:** 74-79

Shuman S (2001) Structure, mechanism, and evolution of the mRNA capping apparatus. *Progress in nucleic acid research and molecular biology* **66:** 1-40

Sica A, Bronte V (2007) Altered macrophage differentiation and immune dysfunction in tumor development. *The Journal of clinical investigation* **117:** 1155-1166

Sica A, Larghi P, Mancino A, Rubino L, Porta C, Totaro MG, Rimoldi M, Biswas SK, Allavena P, Mantovani A (2008) Macrophage polarization in tumour progression. *Seminars in cancer biology* **18:** 349-355

Simonis M, Klous P, Splinter E, Moshkin Y, Willemsen R, de Wit E, van Steensel B, de Laat W (2006) Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nature genetics* **38:** 1348-1354

Skourti-Stathaki K, Proudfoot NJ, Gromak N Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Molecular cell* **42:** 794-805

Splinter E, Heath H, Kooren J, Palstra RJ, Klous P, Grosveld F, Galjart N, de Laat W (2006) CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes & development* **20:** 2349-2354

Stanley ER, Berg KL, Einstein DB, Lee PS, Pixley FJ, Wang Y, Yeung YG (1997) Biology and action of colony--stimulating factor-1. *Molecular reproduction and development* **46:** 4-10

Stedman W, Kang H, Lin S, Kissil JL, Bartolomei MS, Lieberman PM (2008) Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. *The EMBO journal* **27:** 654-666

Steinman RM, Idoyaga J (2010) Features of the dendritic cell lineage. *Immunological reviews* **234:** 5-17

Steinmetz EJ, Brow DA (2003) Ssu72 protein mediates both poly(A)-coupled and poly(A)-independent termination of RNA polymerase II transcription. *Molecular and cellular biology* **23:** 6339-6349

Steinmetz EJ, Ng SB, Cloute JP, Brow DA (2006) cis- and trans-Acting determinants of transcription termination by yeast RNA polymerase II. *Molecular and cellular biology* **26:** 2688-2696

Stout RD, Jiang C, Matta B, Tietzel I, Watkins SK, Suttles J (2005) Macrophages sequentially change their functional phenotype in response to changes in microenvironmental influences. *Journal of immunology* **175:** 342-349

Suraweera A, Becherel OJ, Chen P, Rundle N, Woods R, Nakamura J, Gatei M, Criscuolo C, Filla A, Chessa L, Fusser M, Epe B, Gueven N, Lavin MF (2007) Senataxin, defective in ataxia oculomotor apraxia type 2, is involved in the defense against oxidative DNA damage. *The Journal of cell biology* **177:** 969-979

Suraweera A, Lim Y, Woods R, Birrell GW, Nasim T, Becherel OJ, Lavin MF (2009) Functional role for senataxin, defective in ataxia oculomotor apraxia type 2, in transcriptional regulation. *Human molecular genetics* **18:** 3384-3396

Swinburne IA, Meyer CA, Liu XS, Silver PA, Brodsky AS (2006) Genomic localization of RNA binding proteins reveals links between pre-mRNA processing and transcription. *Genome research* **16:** 912-921

Swirski FK, Nahrendorf M, Etzrodt M, Wildgruber M, Cortez-Retamozo V, Panizzi P, Figueiredo JL, Kohler RH, Chudnovskiy A, Waterman P, Aikawa E, Mempel TR, Libby P, Weissleder R, Pittet MJ (2009) Identification of splenic reservoir monocytes and their deployment to inflammatory sites. *Science* **325:** 612-616

Szabo P, Tang SH, Rentsendorj A, Pfeifer GP, Mann JR (2000) Maternal-specific footprints at putative CTCF sites in the H19 imprinting control region give evidence for insulator function. *Current biology : CB* **10:** 607-610

Tacke F, Alvarez D, Kaplan TJ, Jakubzick C, Spanbroek R, Llodra J, Garin A, Liu J, Mack M, van Rooijen N, Lira SA, Habenicht AJ, Randolph GJ (2007) Monocyte subsets differentially employ CCR2, CCR5, and CX3CR1 to accumulate within atherosclerotic plaques. *The Journal of clinical investigation* **117:** 185-194

Takagaki Y, Manley JL (1997) RNA recognition by the human polyadenylation factor CstF. *Molecular and cellular biology* **17:** 3907-3914

Takagaki Y, Manley JL (2000) Complex protein interactions within the human polyadenylation machinery identify a novel component. *Molecular and cellular biology* **20:** 1515-1525

Takagaki Y, Ryner LC, Manley JL (1989) Four factors are required for 3'-end cleavage of pre-mRNAs. *Genes & development* **3:** 1711-1724

Takagaki Y, Seipelt RL, Peterson ML, Manley JL (1996) The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. *Cell* **87:** 941-952

Tan KL, Scott DW, Hong F, Kahl BS, Fisher RI, Bartlett NL, Advani RH, Buckstein R, Rimsza LM, Connors JM, Steidl C, Gordon LI, Horning SJ, Gascoyne RD (2012) Tumor-associated macrophages predict inferior outcomes in classic Hodgkin lymphoma: a correlative study from the E2496 Intergroup trial. *Blood* **120:** 3280-3287

Tarun SZ, Jr., Sachs AB (1996) Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *The EMBO journal* **15:** 7168-7177

Teixeira A, Tahiri-Alaoui A, West S, Thomas B, Ramadass A, Martianov I, Dye M, James W, Proudfoot NJ, Akoulitchev A (2004) Autocatalytic RNA cleavage in the human beta-globin pre-mRNA promotes transcription termination. *Nature* **432:** 526-530

Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W (2002) Looping and interaction between hypersensitive sites in the active beta-globin locus. *Molecular cell* **10:** 1453-1465

Torrano V, Chernukhin I, Docquier F, D'Arcy V, Leon J, Klenova E, Delgado MD (2005) CTCF regulates growth and erythroid differentiation of human myeloid leukemia cells. *The Journal of biological chemistry* **280:** 28152-28161

Trede NS, Tsytsykova AV, Chatila T, Goldfeld AE, Geha RS (1995) Transcriptional activation of the human TNF-alpha promoter by superantigen in human monocytic cells: role of NF-kappa B. *Journal of immunology* **155:** 902-908

Tsai CL, Rowntree RK, Cohen DE, Lee JT (2008) Higher order chromatin structure at the X-inactivation center via looping DNA. *Developmental biology* **319:** 416-425

Tsai EY, Falvo JV, Tsytsykova AV, Barczak AK, Reimold AM, Glimcher LH, Fenton MJ, Gordon DC, Dunn IF, Goldfeld AE (2000) A lipopolysaccharide-specific enhancer complex involving Ets, Elk-1, Sp1, and CREB binding protein and p300 is recruited to the tumor necrosis factor alpha promoter in vivo. *Molecular and cellular biology* **20:** 6084-6094

Tsai EY, Jain J, Pesavento PA, Rao A, Goldfeld AE (1996a) Tumor necrosis factor alpha gene regulation in activated T cells involves ATF-2/Jun and NFATp. *Molecular and cellular biology* **16:** 459-467

Tsai EY, Yie J, Thanos D, Goldfeld AE (1996b) Cell-type-specific regulation of the human tumor necrosis factor alpha gene in B cells and T cells by NFATp and ATF-2/JUN. *Molecular and cellular biology* **16:** 5232-5244

Tsytsykova AV, Falvo JV, Schmidt-Supprian M, Courtois G, Thanos D, Goldfeld AE (2007) Post-induction, stimulus-specific regulation of tumor necrosis factor mRNA expression. *The Journal of biological chemistry* **282:** 11629-11638

Tsytsykova AV, Goldfeld AE (2000) Nuclear factor of activated T cells transcription factor NFATp controls superantigen-induced lethal shock. *The Journal of experimental medicine* **192:** 581-586

Tsytsykova AV, Goldfeld AE (2002) Inducer-specific enhanceosome formation controls tumor necrosis factor alpha gene expression in T lymphocytes. *Molecular and cellular biology* **22:** 2620-2631

Udalova IA, Knight JC, Vidal V, Nedospasov SA, Kwiatkowski D (1998) Complex NF-kappaB interactions at the distal tumor necrosis factor promoter region in human monocytes. *The Journal of biological chemistry* **273:** 21178-21186

Uhlmann F, Lottspeich F, Nasmyth K (1999) Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1. *Nature* **400:** 37-42

Uhlmann F, Nasmyth K (1998) Cohesion between sister chromatids must be established during DNA replication. *Current biology : CB* **8:** 1095-1101

Vagner S, Vagner C, Mattaj IW (2000) The carboxyl terminus of vertebrate poly(A) polymerase interacts with U2AF 65 to couple 3'-end processing and splicing. *Genes & development* **14:** 403-413

Vallabhapurapu S, Karin M (2009) Regulation and function of NF-kappaB transcription factors in the immune system. *Annual review of immunology* **27:** 693-733

Venkataraman K, Brown KM, Gilmartin GM (2005) Analysis of a noncanonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) site recognition. *Genes & development* **19:** 1315-1327

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

Vodala S, Abruzzi KC, Rosbash M (2008) The nuclear exosome and adenylation regulate posttranscriptional tethering of yeast GAL genes to the nuclear periphery. *Molecular cell* **31:** 104-113

Vostrov AA, Quitschke WW (1997) The zinc finger protein CTCF binds to the APBbeta domain of the amyloid beta-protein precursor promoter. Evidence for a role in transcriptional activation. *The Journal of biological chemistry* **272:** 33353-33359

Wada Y, Ohta Y, Xu M, Tsutsumi S, Minami T, Inoue K, Komura D, Kitakami J, Oshida N, Papantonis A, Izumi A, Kobayashi M, Meguro H, Kanki Y, Mimura I, Yamamoto K, Mataki C, Hamakubo T, Shirahige K, Aburatani H, Kimura H, Kodama T, Cook PR, Ihara S (2009) A wave of nascent transcription on activated human genes. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 18357-18361

Watrin E, Peters JM (2006) Cohesin and DNA damage repair. *Experimental cell research* **312:** 2687-2693

Watrin E, Schleiffer A, Tanaka K, Eisenhaber F, Nasmyth K, Peters JM (2006) Human Scc4 is required for cohesin binding to chromatin, sister-chromatid cohesion, and mitotic progression. *Current biology : CB* **16:** 863-874

Weitzer S, Martinez J (2007) hClp1: a novel kinase revitalizes RNA metabolism. *Cell cycle (Georgetown, Tex* **6:** 2133-2137

Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishiro T, Yahata K, Imamoto F, Aburatani H, Nakao M, Imamoto N, Maeshima K, Shirahige K, Peters JM (2008) Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451:** 796-801

West S, Gromak N, Norbury CJ, Proudfoot NJ (2006) Adenylation and exosome-mediated degradation of cotranscriptionally cleaved pre-messenger RNA in human cells. *Molecular cell* **21:** 437-443

West S, Gromak N, Proudfoot NJ (2004) Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature* **432:** 522-525

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

West S, Proudfoot NJ (2009) Transcriptional termination enhances protein expression in human cells. *Molecular cell* **33:** 354-364

West S, Proudfoot NJ, Dye MJ (2008) Molecular dissection of mammalian RNA polymerase II transcriptional termination. *Molecular cell* **29:** 600-610

White E, Kamieniarz-Gdula K, Dye MJ, Proudfoot NJ (2012) AT-rich sequence elements promote nascent transcript cleavage leading to RNA polymerase II termination. *Nucleic acids research*

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. *The EMBO journal* **5:** 2915-2922

Wicks K, Knight JC (2011) Transcriptional repression and DNA looping associated with a novel regulatory element in the final exon of the lymphotoxin-beta gene. *Genes and immunity* **12:** 126-135

Xie X, Mikkelsen TS, Gnirke A, Lindblad-Toh K, Kellis M, Lander ES (2007) Systematic discovery of regulatory motifs in conserved regions of the human genome, including thousands of CTCF insulator sites. *Proceedings of the National Academy of Sciences of the United States of America* **104:** 7145-7150

Xu N, Donohoe ME, Silva SS, Lee JT (2007) Evidence that homologous X-chromosome pairing requires transcription and Ctcf protein. *Nature genetics* **39:** 1390-1396

Yao J, Mackman N, Edgington TS, Fan ST (1997) Lipopolysaccharide induction of the tumor necrosis factor-alpha promoter in human monocytic cells. Regulation by Egr-1, c-Jun, and NF-kappaB transcription factors. *The Journal of biological chemistry* **272:** 17795-17801

Yonaha M, Proudfoot NJ (1999) Specific transcriptional pausing activates polyadenylation in a coupled in vitro system. *Molecular cell* **3:** 593-600

Yu W, Ginjala V, Pant V, Chernukhin I, Whitehead J, Docquier F, Farrar D, Tavoosidana G, Mukhopadhyay R, Kanduri C, Oshimura M, Feinberg AP,

Lobanenkov V, Klenova E, Ohlsson R (2004) Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation. *Nature genetics* **36:** 1105-1110

Yusufzai TM, Tagami H, Nakatani Y, Felsenfeld G (2004) CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Molecular cell* **13:** 291-298

Zhang Z, Klatt A, Henderson AJ, Gilmour DS (2007) Transcription termination factor Pcf11 limits the processivity of Pol II on an HIV provirus to repress gene expression. *Genes & development* **21:** 1609-1614

Zhao H, Kim A, Song SH, Dean A (2006) Enhancer blocking by chicken beta-globin 5'-HS4: role of enhancer strength and insulator nucleosome depletion. *The Journal of biological chemistry* **281:** 30573-30580

Zhao J, Hyman L, Moore C (1999) Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol Mol Biol Rev* **63:** 405-445

Ziegler-Heitbrock L, Ancuta P, Crowe S, Dalod M, Grau V, Hart DN, Leenen PJ, Liu YJ, MacPherson G, Randolph GJ, Scherberich J, Schmitz J, Shortman K, Sozzani S, Strobl H, Zembala M, Austyn JM, Lutz MB (2010) Nomenclature of monocytes and dendritic cells in blood. *Blood* **116:** e74-80

Zlatanova J, Caiafa P (2009) CTCF and its protein partners: divide and rule? *Journal of cell science* **122:** 1275-1284