Gene type-specific control of pol II transcription elongation

Hadeel Al-Rawaf

Linacre College

University of Oxford

A thesis submitted for the degree of Doctor of Philosophy at the University of Oxford, Trinity Term, 2012

Abstract

Gene type-specific control of pol II transcription elongation

Hadeel Al-Rawaf, Linacre College, University of Oxford

Eukaryotic RNA polymerase II (RNA pol II) transcribes both protein-coding and snRNA genes. A hallmark of RNA pol II that distinguishes it from the other RNA polymerases is the unusual carboxy-terminal domain (CTD) of the large subunit. In mammals, the CTD comprises 52 repeats of the heptapeptide YSPTSPS. In the expression of protein-coding genes, phosphorylation of Ser2 of the CTD heptapeptide repeats by the positive transcription elongation factor b (P-TEFb) is associated with productive transcriptional elongation. However, a number of studies show that P-TEFb is not required for elongation of transcription on snRNA genes, but functions only to activate 3' processing. In addition, we find that Negative Elongation Factor (NELF), which is an important factor in the formation of the elongation checkpoint of protein-coding genes, functions as a termination factor on the U2 snRNA genes. Thus, elongation and transcription termination appear to be controlled differently in the transcription of snRNA and protein-coding genes. In order to understand the molecular mechanism underlying these differences, I have analyzed the association of different RNA pol II elongation factors with the β-actin protein-coding gene and the U2 snRNA genes. The elongation factors TFIIS and FACT associate with both gene types, suggesting that these factors are unlikely to be directly involved in the differential control of elongation. However, I have found that a number of other well-characterized elongation factors are differentially associated with the two gene types. Thus, this analysis highlights key players in the gene-type specific switch to productive elongation of transcription.

Table of contents

Abstract i
Table of contents ii
List of Tables: v
Abbreviations vi
1 Chapter One: Introduction 1
1.1: Transcription of human genes 1
1.1.1: RNA polymerase II 2
1.1.2: Transcriptional pre-initiation 5
1.1.3: Transcription initiation in protein coding genes 7
1.1.4: Transcriptional elongation 8
1.1.5: mRNA processing 16
1.1.6: Structure of the β-actin gene 18
1.1.7: Transcription of human snRNA genes 21
2 Chapter Two: Results 29
2.1: The entire transcribed region of the U2 snRNA gene is nucleosome depleted. 29
2.1.1: Preface 29
2.1.2: Distribution of pol II across the β-actin and U2 snRNA genes 30
2.1.3: The chromatin landscape of the RNU2 locus and the β-actin gene. 34
3 Chapter Three: Results 36
3.1: Some elongation factors are associated with both the β-actin and U2 snRNA genes. 36
3.1.1: Preface 36
3.1.2: Cloning of TFIIS and SSRP1 (FACT). 37
3.1.3: TFIIS and FACT are associated with both snRNA genes and protein-coding genes 40
4 Chapter Four: Results 43
4.1: Some transcription elongation factors are differentially associated with the U2 snRNA and the β-actin genes. 43
4.1.1: Spt6 is differentially associated with the β-actin and U2 snRNA genes. .. 43
4.1.2: PAF is differentially associated with the β-actin and U2 snRNA genes. .. 46
4.1.3: DSIF is present at different levels on the β-actin and the U2 snRNA genes. 50
4.2.1: Set1 is differentially associated with the β-actin and snRNA genes. 55
4.2.2: H3K36me2 is differentially associated with the β-actin and U2 snRNA genes. 58
4.2.4: Association of RAD6A and RAD6B with the β-actin and U2 snRNA genes. 64
4.2.5: Association of BRE1 with the β-actin and U2 snRNA genes. 69
5 Chapter Five: Results 72
5.1: The PAF complex controls levels of Pol II, Spt5, and H2B specifically on the β-actin gene. 72
6 Chapter Six: Discussion 77
6.1: TFIIS levels correlate with pol II levels on both the β-actin and U2 snRNA genes. 77
6.2: The histone chaperone, FACT is recruited to both gene types but shows differential association. 78

6.3: DSIF and phospho-DSIF are present at different levels on the β-actin and U2 snRNA genes. 79
6.4: The elongation factors Spt6 and PAF and the histone modification enzyme Set1 are differentially associated with the β-actin and U2 snRNA genes. 79
6.5: H2BK120ub1 is found only on the β-actin gene. 81
6.6: Functions of the early elongation checkpoint. 82
6.7: Why does transcription elongation terminate with the first nucleosome encountered on the U2 snRNA genes? 85
7 Chapter Seven: Materials and Methods 89
7.1Materials 89
7.1.1: Antibodies 89
7.1.2: Antibiotics 91
7.1.3: Buffers and solutions 91
7.1.4: Chemicals and other reagents 91
7.1.5: Chromatin Immunoprecipitation (ChIP) materials and buffers 92
7.1.6: Kits and enzymes 93
7.1.8: Bacteria and cell line 96
7.1.9: Bacterial and tissue culture media 96
7.1.10: Cloning vector 96
7.2: Methods 97
7.2.1: DNA related methods 97
7.2.2: RNA related methods 101
7.2.3: Protein related methods 104

List of Figures:

Figure (1.1): Transcription elongation events on protein-coding genes. ……....11
Figure (1.2): Schematic diagram of the β-actin gene……..……....20
Figure (1.3): Structure of snRNA and protein-coding genes…….....22
Figure (1.4): Function of P-TEFb phosphorylation in snRNA gene transcription...…...26
Figure (2.1): Localization of pol II on the β-actin and U2 snRNA genes……...33
Figure (2.2): Localization of H3 and H2B on the β-actin and U2 snRNA genes.....35
Figure (3.1). Western blot analysis of TFIIS and SSRP1. ……..39
Figure (3.2). TFIIS and FACT are associated with both snRNA and protein-coding genes. ……..42
Figure (4.1) Spt6 is differentially associated with the β-actin and U2 snRNA genes….54
Figure (4.2): PAF is differentially associated with the β-actin and U2 snRNA genes…49
Figure (4.3): DSIF/phospho DSIF is differentially associated with the β-actin and U2 snRNA genes..……...……..53
Figure (4.4): Levels of DSIF/phospho DSIF relative to pol II on both β-actin and U2 snRNA genes. …….54
Figure (4.5): Set1 is differentially associated with the β-actin and U2 snRNA genes…57
Figure (4.6): H3K36me2 is differentially associated with the β-actin and U2 snRNA genes.....60
Figure (4.7): H2BK120ub1 is differentially associated with the β-actin and U2 snRNA genes…….....63
Figure (4.8): RAD6A and RAD6B are associated with both snRNA and protein-coding genes……67
Figure (4.9): BRE1 is associated with both snRNA and protein-coding genes…….71
Figure (5.1): Loss of PAF alters levels of Pol II, Spt5, and H2B on the β-actin gene…75
Figure (6.1): conclusion: differenential factor recruitment on the β-actin and the U2 genes……88

List of Tables:

Table (1.1): Elongation factors of RNA polymerase II involved in this study……15
Table (7.1): Primers……..95

Abbreviations

bp	base pair
BSA	Bovine Serum Albumin
CDK9	Cyclin-dependent kinase 9
ChIP	Chromatin Immunoprecipitation
CTD	C-terminal domain
DMSO	dimethylsulphoxide
DNA	Deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DRB	5,6-dicloro-1-β-D-ribofuranosylbenzimidazole
DSE	Distal Sequence Element
DTT	Dithiothreitol
EDTA	ethylenediaminetetraacetic acid
FACT	Facilitates Chromatin Transcription Elongation Factor
HATs	Histone Acetyltransferases
IP	Immunoprecipitation
kDa	kilo Dalton
LTR	Long-terminal repeat
mRNA	messenger RNA
NaCl	Sodium Chloride
NELF	Negative Elongation Factor
NP-40	Nonidet P40
NRO	Nuclear Run-On
PBS	Phosphate Buffered Saline

PCR	Polymerase Chain Reaction
PMSF	Phenylmethylsulfonlyfluoride
Pol II	RNA Polymerase II
PSE	Proximal Sequence Element
P-TEFb	Positive Transcription Elongation factor b
PTM	Post Translational Modification
qPCR	quantitative PCR
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
RT	Room Temperature
RT-PCR	Real time PCR
SDS	Sodium Dodecyl Sulphate
snRNA	small nuclear RNA
Spt16	Suppressor of Ty protein 16
SSRP1	Structure-specific recognition protein 1
SYBR	SybrGreen
TBS	Tris Buffered Saline
TFIIS	transcription elongation factor S-II
Tris-HCl	tris (hydroxymethyl) aminomethane hydrochloride
TSS	transcription start si

1 Chapter One: Introduction

At each stage during the life of a cell, the correct set of ribonucleic acid (RNA) transcripts and proteins must be present. Accordingly, gene expression is controlled at all levels, from transcription through to translation and post-translational modification. Regulation of the first step in this process, initiation of transcription, is fundamental to ensuring the appropriate expression patterns within each cell and has been the subject of intensive study for many years. Historically, the next step of producing a full-length transcript, elongation of transcription, has been less studied. However, it is becoming increasingly clear that elongation of transcription of genes transcribed by RNA polymerase II (pol II) is subject to complex regulatory processes. This thesis describes the results of my study of the control of pol II-transcribed genes at the level of elongation. In this Introduction I will give an overview of transcription of the pol II-dependant genes in general in addition to an in-depth review of what is known of transcription elongation.

1.1: Transcription of human genes

The eukaryotic genome contains genes that encode messenger RNA (mRNA), 18 and 28S ribosomal (r)RNA, and various small non-coding RNAs, including the small nuclear (sn)RNAs and the cytoplasmic 5S ribosomal and transfer (t)RNAs. The group with the largest number of individual members is that of the mRNA-encoding genes.

Transcription is the first step of gene expression. During that process segment of DNA is copied into RNA by various multisubunit RNA polymerase enzymes. In eukaryotic cells, there are three distinct RNA polymerases: RNA polymerase I (pol I) that transcribes the rRNA genes; RNA polymerase II (pol II) that transcribes the protein-coding genes, small nucleolar (sno)RNA genes and some of the snRNA genes; RNA polymerase III (pol III) that transcribes the tRNA genes, the 5S rRNA gene and some of the snRNA genes.

Transcription of protein-encoding genes consists of pre-initiation, inititation, promoter-escape, elongation and termination. During pre-inititaion the RNA polymerase and its cofactors assemble on the DNA template. During initiation, the first phosphodiester bond is synthesized. Promoter-escape marks dissociation of RNA polymerase from the promoter. RNA strand is synthesized during transcription elongation and polymerase disengage from the DNA during termination.

1.1.1: RNA polymerase II

Pol II is a multisubunit enzyme, containing 12 subunits (Choder, 2004). The largest, RPB1 has a unique carboxy-terminal domain (CTD) structure (Corden, 1990; Egloff & Murphy, 2008a; Meinhart et al, 2005). The CTD comprises multiple copies of the heptapeptide sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser (YSPTSPS) (Egloff & Murphy, 2008b), and the number of these repeats differs among species: 26 in yeast, 32 in C.elegans, 45 in Drosophila, and 52 in mammals (Chapman et al, 2008; Egloff & Murphy, 2008a; Sims et al, 2004a; Stiller & Hall, 2002)

The CTD functions as a platform to faciliate molecular interactions during the transcription cycle, through extensive post-translational modification. Collectively, different combinations of different modifications create a "CTD code" that is read by

a wide range of proteins involved in transcription and processing of RNA (Egloff & Murphy, 2008a). The most-studied and so far, best-understood CTD modification is phosphorylation, which can occur on the tyrosine, threonine, and all three serines that make up the CTD repeat (Egloff & Murphy, 2008a). According to the phosphorylation state of the CTD, Rpb1 has two forms, the hypophosphorylated form which is designated IIa, and a hyperphosphorylated form known as IIo. The hypophosphorylated Pol II (IIa) is selectively recruited to the preinitiation complex (PIC). In a way, newly-recruited pol II resembles an empty vessel ready to "load" the phosphorylation marks (Phatnani & Greenleaf, 2006). In order to obtain the right combinations of phosporylation marks at different stages of the transcription cycle, the CTD undergoes dynamic modification by a range of kinases and phosphatases, which are responsible for the addition or removal of phosphates, respectively (Egloff & Murphy, 2008a). The cyclin-dependant kinase 7 (CDK7), which is a subunit of the basal transcription factor TFIIH (Feaver et al, 1997) phosphorylates serine Ser5 in the YSPTSPS heptapeptide early in transcription (Komarnitsky et al, 2000; Phatnani & Greenleaf, 2006). The cyclin-dependant kinase 9 (CDK9) subunit of the elongation factor P-TEFb is responsible for Ser2 phosphorylation later in the transcription cycle (Komarnitsky et al, 2000). More recent studies have demonstrated that Ser7 of the CTD heptapeptide is also phosphorylated (Chapman et al, 2007; Egloff et al, 2007).

The notion that the CTD and its phosphorylation acts to coordinate transcription with co-transcriptional pre-mRNA processing, by the sequential recruitment of different RNA processing factors, has flourished during the last decade (Corden & Patturajan, 1997). For example, multiple studies have shown that the CTD phosphorylated at Ser5 is vital for recruiting capping enzymes to the 5' end of a newly emerging RNA transcript and capping is not activated by a CTD-less Pol II (Kanin et al, 2007;

McCracken et al, 1997a). The phospho-Ser5 CTD directly recruits guanlylytransferases (Schroeder et al, 2000).

3' end formation of the transcript is linked to phosphorylation of Ser2 of the CTD heptapeptide (Ahn et al, 2004). Simultaneous phosphorylation of Ser5 and Ser2 helps to recruit the histone methyltransferase (HMT) Set2 that directs di- and trimethylation of H3K36 during transcription of the body of the gene (Li et al, 2005; Strahl et al, 2002).

Ser7 phosphorylation is detected on snRNA genes (Egloff et al, 2007) and is also high at the 3'-end of protein-coding genes (Chapman et al, 2007). Interestingly, loss of Ser7 phosphorylation, by replacement of Ser7 to alanine, causes a defect in snRNA 3'-end processing, with little effect on the production of mRNA. Previous chromatin immunoprecipitation (ChIP) studies have shown that a large macromolecular complex known as the Integrator complex is detected on both the U1 and U2 snRNA genes and mediates snRNA 3'-end processing (Baillat et al, 2005; Egloff et al, 2008). This complex interacts with the CTD phosphorylated at Ser7 as mutations of Ser7 abrogate Integrator complex recruitment to snRNA genes (Egloff & Murphy, 2008b).

X-ray crystallographic studies of pol II have shown that the CTD is situated close to the RNA exit channel, a strategic position that helps to recruit different RNA binding factors to the transcript (Cramer et al, 2001).

There is a general requirement of CTD phosphorylation throughout the transcription cycle. However, there may be differences in the phosphorylation requirements for specific genes.

1.1.2: Transcriptional pre-initiation

Pre-initiation complex (PIC) is composed of pol II and general transcription factors (GTFs) including TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH. PIC is assembling on the DNA at the promoter, defined as the minimal sequence required for accurate transcription initiation *in vitro* (Roeder, 1996). The general transcription factor TFIID contains TATA-binding protein TBP and up to 13 TBP-associated factors TAFs (Dynlacht BD et al, 1991; Matangkasombut et al, 2004; Papai et al, 2011; Zaborowska et al, 2012). TFIID structure is synamic and different TFIID composition affect its conformation (Liu et al 2008). Biochemical experiments demonstarted that TAFs contain the histone fold domain (HFD) that facilitates heterodimer formation (Gangloff YG et al. 2001). However, pol II and the GTFs copurify together with other protein complexes, including the Mediator complex (Borggrefe & Yue, 2011) and Swi/Snf (Wilson et al, 1996). Pol II and the proteins that are co-purified with it have been termed a holoenzyme and this may be the form of the enzyme associated with transcription factories (Cook, 2010; Myers & Kornberg, 2000). Formation of the PIC usually begins with TFIID binding to the core promoter element TATA box located 25 bp upstream of the transcription start site. Other described promoter elements, the initiator (Inr) and core promoter element (DPE) can be present alone or in combination (Smale & Kadonaga, 2003). Formation of this promoter-bound complexes is sufficient for a basal level of transcription. For activator-dependent transcription, general cofactors transmit regulatory signals (Thomas and Chiang 2006). Interestingly, activator can function in the absence of ATP by stimulating transcription by increasing the number of functional PICs (Lin and Green, 1991).

Chromating that is compacted must be “loosened” for transcription to take place by chromatin remodelling or modyfing complexes (Gerber & Shilatifard, 2003; Li et al, 2007a). Collectively, enhancer-binding factors and specific DNA sequences direct the genes to active centers through interaction with promoter selectivity factors (Lemon & Tjian, 2000).

There are different sets of factors at different promoters. This reflects the spatiotemporal profile of gene expression in organisms (Muller & Tora, 2004). Components of TFIIB, TFIIE, TFIIF, TFIIH and pol II and encoded by single-copy genes (Aoyagi & Wassarman, 2000). Pol II transcription can be initiated at certain promoters with PICs containing only a partial set GTFs (Eichner et al, 2010; Muller & Tora, 2004). In contrats to GTFs, there are two additional families in metazoans encoding fators related to TBP (Veenstra and Wolffe, 2001). Different TBP-like fators could facilitate differential gene expression.

The spectrum of TBP- and TAF-type factors allows to form multiple promoter recognition complexes. *In vitro* studies demonstrated that single human cell containing TFIID complexes with or without TAF10 exibits diferent functional properties (Brou et al. 1993). There are TFIID complexes containing cell type-specific TAFs, for example, TAF4b is enriched in differentiatied human B lymphocytes (Dikstein et al. 1996). Various TFIID complexes have following function:

- Regonizing different promoters
- Interacting with distinct sets of GTFs
- Mediating differential responses from activators to the PIC (Muller & Tora, 2004).

In yeast, promoters can be grouped into two classes based on their requirement for TAFs (Kuras et al, 2000):

- TAF-dependent promoters require TAFs for transcription
- TAF-independent promoters do not require TAFs for activity.

Various TAF contiaing complexes that lact TBP were isolated from yeast and also human cells (Muratoglu et al, 2003; Wieczorek et al, 1998). Certain human TAFs have been identified as sequence-specific DNA binding proteins (Smale, 2001). Therefore, TAF combination is at least partially core promoter-dependent (Green, 2000) and TBP-independent transcription adds another level to the complexity of regulation of gene expression.

To sum up, a variety of TAF-containing complexes are recruited to specific promoters. This has with implications for the regulation of transcription (Muller & Tora, 2004; Zaborowska et al, 2012).

1.1.3: Transcription initiation in protein coding genes

Pol II exists in a hypophosphorylated form, Pol IIa, and a hyperphosphorylated form Pol IIo (Palancade & Bensaude, 2003). Pol IIa assembles with its transcription machinery at the promoter. TFIIH interacts specifically with the pol IIa form. Transcription-associated CDKs regulate transcription and related events through phosphorylation of the CTD, as well as many other potential substrates [2]. A simplified model of early transcriptional events involve CDK7. CDK7 is a subunit of the general transcription factor TFIIH and is responsible for Ser5 phosphorylation of promoter associated Pol II, which is required for pre-mRNA capping (see Section 1.5). This shifts PolII from an early, abortive initiation, into processive initiation (Sims et al, 2004b; Marshall & Price, 1992). During abortive initiation, Pol II appears to synthesize only short, about 50 nt-long, transcripts (Aida et al, 2006)

1.1.4: Transcriptional elongation

Elongation stage of transcription is highly regulated (Sims et al, 2004a). The main rate-limiting step of transcription elongation is Pol II pausing. Negative transcription elongation factors (N-TEFs): DRB sensitivity-inducing factor (DSIF) and negative elongation factor (NELF), are recruited to the elongation complex in the promoter proximal region to induce pol II pausing (Yamaguchi et al., 1999; Andrulis et al, 2000; Egloff et al, 2009b; Wu et al, 2003). DSIF is composed of Spt4 (p14) and Spt5 (p160) subunit (Yamaguchi et al, 2001). NELF consists of four subunits: NelfA, NelfB, NelfC-D, NelfE (Yamaguchi et al, 1999a).

As mentioned before, the modified CTD functions as a platform for the association of factors that regulate Pol II elongation. Phosphorylation of Ser5 residues has been extensively characterised. Genome wide mapping of these modifications shows that Ser5 is enriched at promoter regions (Dahmus, 1996; Komarnitsky et al, 2000). Meanwhile, Ser2 is associated with productively elongating Pol II and accumulates along the gene body, peaking at the 3' end of genes (Buratowski 2009). This phosphorylation marks the processive phase of transcription elongation (Komarnitsky et al, 2000).

The CDK9 subunit of P-TEFb phosphorylates Ser2 of the CTD, as well as the elongation factors NELF and DSIF, which releases paused polymerase into productive elongation (Buratowski 2009; Peterlin and Price 2006; Saunders et al, 2006; Bartkowiak & Greenleaf, 2011; Fujinaga et al, 2004; Garber & Jones, 1999; Lavoie et al, 2001; Yamada et al, 2006). Phospho NELF dissociates from pol II (Gomes et al, 2006) and phospho DSIF turn into a transcriptional activator (Yamada

et al, 2006),. Phospho DSIF remains associated with pol II during active transcription (Wu et al, 2003).

P-TEFb is composed of Cdk9 kinase and a cyclin. It interacts with T-type cyclins (Fu et al, 1999). Cdk9 is a Ser/Thr proline-directed kinase (Gerber & Jones, 1999) and it exists in two isoforms Cdk9-42 and Cdk9-55 (Marshall and Price, 1995). The ratio of two isoforms is cell type-dependent (Liu and Herrmann, 2005). A major function of P-TEFb is to eliminate the negative effect of DSIF and NELF (Figure1.1), and this reversal requires the kinase activity of the factor (Marshall et al, 1996). In addition, P-TEFb activity is dispensable for transcription in the absence of DSIF *in vitro* (Wada et al, 1998). P-TEFb can be recruited to paused Pol II by different ways. The recruitment can be directed activators (Eberhardy and Farnham 2002), the co-activator Brd4 (Jang et al 2005) and through Mediator complex (Yamaguchi personal communication).

Much of the mechanism of P-TEFb was elucidated using 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) (Marciniak & Sharp, 1991), an ATP-competitor that inhibits the activity of kinases including CDK9 (Marshall & Price, 1995). Treatment of cells with DRB results in the production of shortened transcripts (Tamm et al, 1980) and general decrease of mRNA synthesis (Sehgal et al, 1976).

As mentioned before, the negative transcription elongation factors DSIF and NELF also represent major CDK9 substrates. Their phosphorylation is required to overcome the proximal pause of the polymerase.

DSIF associates with Pol II at an early stage of elongation and it capable of promoating and inhibiting Pol II elongatiom (Wada et al, 1998; Yamada et al 2006). It associates at the start of transcription (Gomes et al, 2006) and can also be detected throughout the transcribed region (Egloff et al, 2009b). DSIF was shown to physically interact with other elongation factors such as FACT (for facilitate chromatin

transcription). FACT facilitates passage of Pol II through nucleosomes (Reinberg and Sims 2006).

In addition, the Spt5 subunit and its phosphorylation play an important role in recruitment of the PAF complex, which in turn is known to recruit a number of histone modifications enzymes such as histone H3K4, H3K36 methyltransferases and histone H2B ubiquitylases (Kim et al, 2009; Krogan et al, 2003a; Liu et al, 2009; Ng et al, 2003; Squazzo et al, 2002) (see Table 1.1).

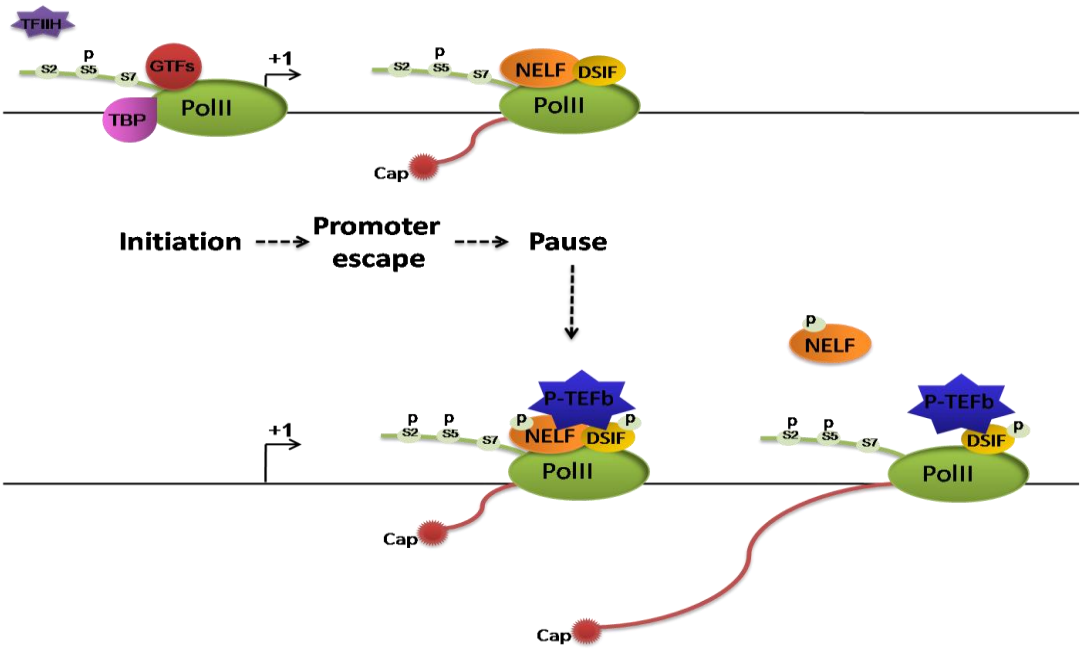


Figure (1.1): Transcription elongation events on protein-coding genes.

Transcription is initiated by the recruitment of the pre-initiation complex (PIC) which contains the general transcription factors (GTFs) and pol II to the promoter through the TATA-binding protein (TBP). After transcription is initiated, the CDK7 subunit of the TFIIH complex phosphorylates Ser5 of the pol II CTD to facilitate promoter escape. Thereafter, pol II engages in the transcription of a nascent RNA molecule of approximately 20-40 nt. Pol II then enters a stage known as promoter proximal pausing, in which pol II is under the effect of two negative elongation factors (NTEFs); DSIF and NELF. At this stage, the mRNA is capped by the action of the capping enzymes. Promoter proximal pausing is overcome by P-TEFb phosphorylation, through its kinase subunit (CDK9) and with the help of positive elongation factors.

1.1.4.1: Elongation factors

After this transition from abortive to processive elongation, pol II is acted upon by general elongation factors (Reines et al, 1999). Polymerization by pol II *in vivo* takes place at approximately 1000-2000 nts per minute. *In vitro* transcription experiments on naked DNA, using a purified preparation of pol II, have indicated a RNA synthesis rate of approximately a few hundred nucleotides per minute. Therefore, the involvement of elongation factors and/or chromatin remodelling factors that speed up transcription was predicted (Ucker & Yamamoto, 1984). Several elongation factors have been isolated that can increase the processivity of pol II. All of these interact directly with elongating polymerases and boost elongation rate *in vitro*. Mechanistic studies demonstrate that these factors often function by suppressing transient pausing by pol II. Several lines of evidence suggest that elongation factors may function similarly by helping to maintain the nascent transcript in proper alignment with the catalytic site. These factors include TFIIF, DSIF and CSB (Sims et al, 2004b). Some of them also act on chromatin and therefore influence elongation rate indirectly (Hartzog et al, 2002).

Pol II transcribing different genes can differ in the ability to processively elongate (Perales & Bentley, 2009). This can be determined by the promoter (reviewed in Spencer and Groudine, 1990). For example, with regards to the c-myc gene, the appearance of a block to elongation at the site called T2 is dependent on which promoter is being used (Miller et al., 1989). When transcription is initiated at an alternate promoter, the specific block to elongation is not functional (Spencer & Groudine, 1990). In addition, the TATA box appears to be involved in decreasing the elongation properties of pol II (Yankulov et al, 1994). Therefore, the species of TFIID

may be important in determining the elongation processivity of pol II, much like the scenario in yeast (Spencer & Groudine, 1990). (Yankulov et al, 1994) also found that the competence of pol II to elongate is controlled by activators. Hence, they conclude that there are certain TFIID compositions that cannot respond to regulatory signals and others that can, which might result in the difference of elongation capacity. It has been hypothesized that these differences may involve the differential association of ancillary elongation factors with the CTD that are dependent on the specific phosphorylation state of the CTD. Therefore, elongation complexes are not uniform in their composition. The rate of elongation may be important to provide sufficient time for subsequent processing (Reinberg & Roeder, 1987).

Polymerase associated factor (PAF) first identified 15 years ago, was shown to accompany yeast pol II. The phosphorylation state of the pol II CTD does not dictate its association with the PAF complex as it is detected on the promoter and throughout the coding region of genes regardless of whether the CTD is unphosphorylated or Ser2/5 phosphorylated (Jaehning, 2010). The PAF complex functions as a platform that coordinates recruitment of factors ranging from histone modification enzymes to RNA 3'-end processing factors and elongation factors. For example, two essential events that lead to the production of different histone modifications are H2B ubiquitylation and CTD Ser2 phosphorylation (Jaehning, 2010). H2B ubiquitylation is required for the ultimate trimethylation of H3K4 (Shilatifard, 2008). PAF also has a role in achieving high levels of Ser2 phosphorylation of the CTD (Mueller et al, 2004), which are in turn necessary for trimethylation of H3K36-M3 through Set2 recruitment (Kizer et al, 2005). Additionally, PAF was reported to interact with the Spt5 subunit of the DSIF complex and together stimulate transcription (Chen et al,

2009). Finally, a connection between PAF and the CF1A cleavage and polyadenylation complex has also been shown (Jaehning, 2010).

Elongation Factor	Subunits	Properties
DSIF	Spt4 Spt5	Stimulates elongation, suppresses early transcript termination, stimulates capping
TFIIS		Stimulates pol II-mediated cleavage of nascent transcript to alleviate arrest
Spt6		Stimulates elongation, modulates chromatin structure, histone chaperone activity
FACT	Spt16 SSRP1	Facilitates elongation through chromatin, modulates chromatin structure, histone chaperone activity
Set1		Methylates histone H3-K4, localized to promoter and coding regions
PAF	Paf1 Cdc73 Ctr9 Leo1 hSki8	Modulates H2B monoubiquitination, recruiting H3K4 and H3K36 methyltransferases

(Sims et al, 2004a)

Table (1.1): Elongation factors of RNA polymerase II involved in this study.

Listed in this table are the elongation factors studied throughout this thesis along with their subunits and functions.

1.1.5: mRNA processing

The pre-mRNA has to be processed into a messenger RNA (mRNA). This includes:

- adding a cap structure at the 5′ end

- splicing out of introns from the pre-mRNA

- generation of a 3′ end. The CTD of the large subunit of pol II is coupling pol II transcription to mRNA processing (Egloff & Murphy, 2008a; Hirose & Manley, 2000). In particular, post-tranlsational modifications of the CTD are important for the coordination of the processing steps (Egloff & Murphy, 2008a).

1.1.5.1: Splicing

Before mRNA is exported to cytoplasm the RNA splicing takes place. Splicing is removal of introns and ligation of exons performed by a complex called spliceosome. Spliceosome is composed of U1, U2, U4, U5 and U6 and additional proteins (Chen and Manley 2009). It recognizes conserved sequences: the 5′ splice site (5'ss), the 3′ splice site (3'ss), the branch point (BP) and the pyrimidine tract (PPT). Several steps are required for splicing:

1) Binding of U1 to the 5'ss
2) Binding of splicing factor 1 (SF1) to the branch site
3) Binding of U2 auxilary factor (U2AF) to the PPT and the 3'ss (Berglund et al 1997; Berglund et al 1998)
4) ATP-dependent step of SF1 diplacement by U2
5) Recruitment of the U4, U5 and U6 (Wu and Manley 1991).

Introns are removed in two chemical steps:

1) 5'ss cleavange and lariat formation. Lariat is formed as the cut 5′ end of the intron becomes covalently linked to the adenine nucleotide 2′ -OH

2) Formation of a phosphodiester bond between the last base of the upstream exon and the first base of the downstream exon (Proudfoot et al 2002). The two exons are joined together and intron released (Moore & Sharp, 1993).

1.1.5.2: mRNA 3′ end processing

The formation of the 3′ end of mRNA requires co-transcriptional cleavage followed by polyadenylation. Cleavage is directed by a conserved AAUAAA sequence (polyA site) and a downstream GU rich sequence (DSE) (Gil & Proudfoot, 1984). The polyA site is recognized by the 160 kDa subunit of the cleavage and polyadenylation specificity factor (CPSF160) (Proudfoot et al, 2002). CPSF160 interacts with TFIID and pol II CTD, linking polyadenylation to transcription (Dantonel et al., 1997). The 64kDa subunit of CstF recognizes the DSE (Proudfoot et al, 2002). Endonucleolytic cleavage targets a CA dinucleotide sequence that is situated between the DSE and AAUAAA sequence. After cleavage, an enzyme called poly-A polymerase (PAP) adds up to 200 ATP nts to the 3′ end.

Eggermont & Proudfoot (1993) investigated if a functional poly(A) signal and a downstream transcriptional pause site can prevent transcriptional interference of RNA polymerase II-activated promoters. It turn out, that the expression plasmid containing two tandem retroviral long terminal repeats (LTRs), transfected into HeLa cells, resulted in transcriptional interference of the LTR promoters (Eggermont & Proudfoot, 1993).

1.1.5.3: Role of the CTD in polyadenylation

McCracken et al. (1997b) demonstrated that the CTD of pol II is required for efficient RNA processing. The authors demonstrated that the CTD associates with the 3'-processing factors (McCracken et al, 1997b). In particular, loss of the phosphorylation of pol II CTD at Ser2 causes major alterations in the 3'RNA processing (Ahn et al, 2004; Barilla et al, 2001). Ser2 phosphorylation occurs later in the transcription cycle and is generally highest towards the 3' end of the genes. In addition, P-TEFb inhibitors affect elongation of transcription of protein-coding genes, splicing and polyadenylation (Bird et al., 2004, Ni et al., 2004). Moreover, Pcf11 factor, that is a component of the cleavage factor IA (CFIA) complex, binds directly to CTD heptapeptide repeats that are phosphorylated on Ser2 (Meinhart & Cramer, 2004).

1.1.6: Structure of the β-actin gene

Since I have used the β-actin gene as a model protein-coding gene, I shall give a brief introduction to the structure of this gene here.

The β-actin gene

β-actin is one of the most abundant proteins in all mammalian non-muscle cells. There is only a single functional β-actin gene located on chromosome 7 (Ng et al, 1985) with a transcription unit of approximately 5kb (Gromak et al, 2006). The gene contains five introns and encodes transcripts that are spliced and polyadenylated (Nakajima-Iijima et al, 1985). Pseudogenes have also been reported and reach at least

20 in number (Ponte et al, 1983). The β-actin gene is a member of the immediate-early genes, where its expression is induced upon the exposure of serum response factor (SRF) (Danilition et al, 1991; Schratt et al, 2001). The high levels of the expressed protein reflect a particularly strong promoter of the β-actin gene (Ng et al, 1985). The promoter can be characterized by the presence of three cis-acting regulatory elements: the TATA, CCAAT, and CCArGG box consensus sequences. The TATA box, found upstream of the transcription start site at -29 to -23 (Danilition et al, 1991; Ng et al, 1985), is crucial for achieving maximum activity of the promoter**.** Furthermore, the promoter selectivity transcription factor, Sp1, binds to the TATA associated factor TAF110 (TAF4 according to new TAFs nomenclature) (Hoey et al, 1993) together with another factor known as the β-actin CCAAT binding factor (Danilition et al, 1991). The CCArGG box (-62 to -53) is recognised by the serum response factor (SRF) (Danilition et al, 1991) (Figure 1.2 A and B).

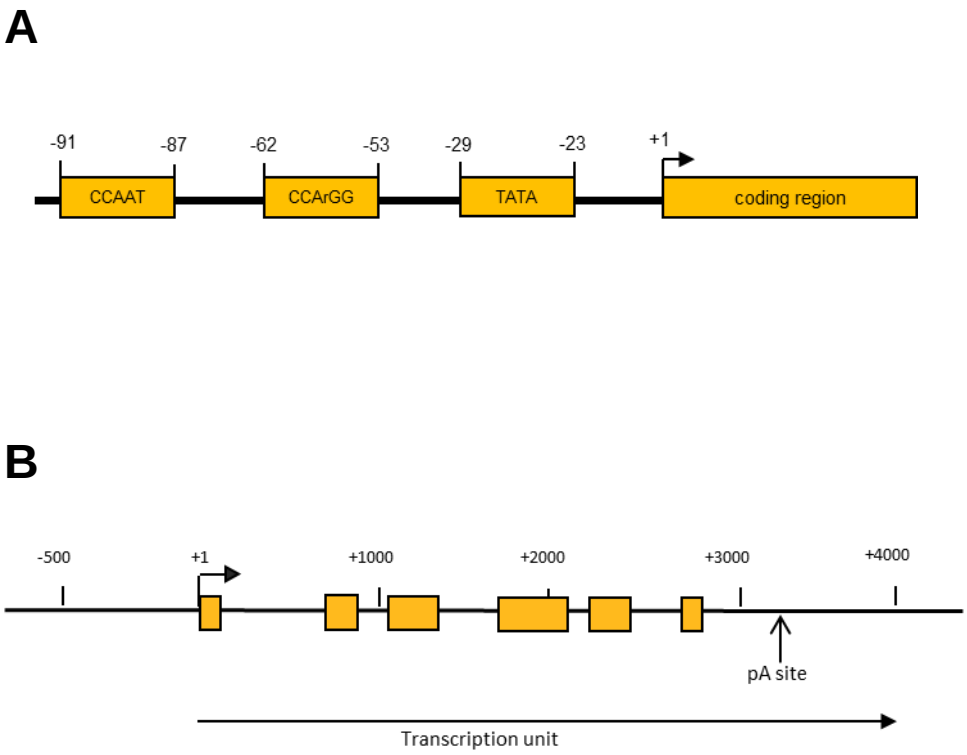


Figure (1.2): Schematic diagram of the β-actin gene.

A) The relative locations of various individual repeat elements of the β-actin promoter are indicated. +1 is the first base of the β-actin coding region.

B) Diagram of the human β-actin gene with the exons depicted as open boxes and the polyadenylation (pA) signals indicated with an arrow.

1.1.7: Transcription of human snRNA genes

Small nuclear RNA (snRNA) genes are short, intron-less and encode short structural RNAs (Will & Luhrmann, 2001). The majority of snRNA are required for splicing, for example, U1, U2, U4, U5 and U6 snRNAs (Hernandez, 2001). The 7SK snRNA has been implicated in the regulation of transcriptional elongation (Yang et al, 2001). U snRNA genes are typically present in multiple copies, for instance, around 20 U2 snRNA genes are clustered in 6.2 kb tandem repeats (Westin et al., 1984). Most snRNA genes are transcribed by pol II and some by pol III (Hernandez, 2001; Schramm et al, 2000).

The snRNA genes have a characteristic promoter region and a 3′ end processing signal specific to pol II-transcribed snRNA genes (Figure 1.3). The pol II snRNA core promoters contain essential element, called the proximal sequence element (PSE). The distal sequence element (DSE) functions to enhance transcription from the core promoter (Hernandez, 2001; Mattaj et al. 1985).

The snRNA promoters differ from those pol II promoters that direct mRNA synthesis. Therefore, snRNA genes are an interesting model for the study of pol II transcription. In my research, I have used the U2 gene which is one of the best-characterized, pol II-transcribed snRNA genes.

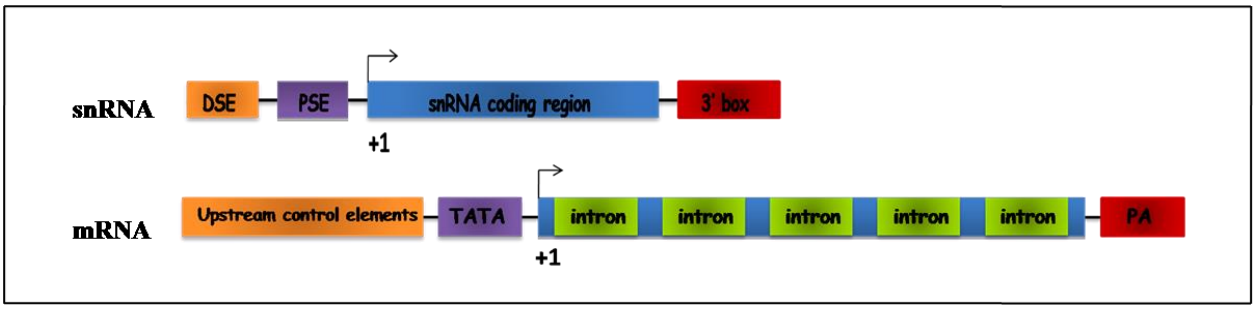


Figure (1.3): Structure of snRNA and protein-coding genes.

The diagram at the top shows a model of an snRNA gene with its distinct PSE (the proximal sequence element) and DSE (the distal sequence element), containing promoter and a 3′box processing element. The diagram underneath shows a model of a protein coding gene with its TATA box containing promoter and various upstream control elements. Protein coding genes also have introns and a polyadenylation (PA) 3′processing signal.

1.1.7.1: snRNA promoter

The PSE element, located 50nts upstream of the TSS of the U2snRNA gene, defines the start of transcription and also serves as a binding site for the snRNA-specific PSE bidning factor PTF. PTF consists of five subunits PTFα (SNAP190), PTFβ (SNAP50), PTFδ (SNAP45), PTFγ (SNAP43) and SNAP19 (Hernandez, 2001; Yoon et al., 1995).

The DSE element, is located 220 bp upstream of the start site of transcription. It functions to activate transcription and contains at least one octamer motif (Murphy et al 1992; Murphy et al. 1997). It contains binding sites for transcription factor Oct-1 (Murphy et al. 1997) and Sp1 binding sites (Ares et al, 1987).

Activators of transcription of mRNA genes, such as VP16, cannot activate transcription of the pol II transcribed snRNA genes (Das et al, 1995). It indicates that there are differences in the components of the PIC between snRNA and mRNA genes. The following GTFs: TFIIB, TFIIA, TFIIF, TFIIE and TFIIH are required for the transcription of snRNAs (Kuhlman et al, 1999). Bernues et al. (1993) analyzed transcription factor requirements for the human U1 and U6 genes. PSE-binding protein (PBP) was shown to be essential for transcription of both genes. The authors also demonstrated that TBP is required for in vitro snRNA tarnscrition, but not TFIID (Bernues et al. 1993). This suggests that the U1 snRNA gene promoter may be TAF-independent or require a specific combination of TAFs, as the case with the U2 snRNA, where a specific TAF complex was found to bind the U2 snRNA exclusively and was denoted the snTAFC (Zaborowska et al, 2012). Interestingly, TAF1, TAF10 and TAF4, are not detected on snRNA genes (Zaborowska et al, 2012).

The CDK7 subunit of TFIIH was recently shown by chromatin immunoprecipitation to be associated with the promoter of the U2 snRNA gene (Glover-Cutter et al, 2009).

Although transcription of snRNA genes is highly efficient *in vivo*, there is no efficient transcriptional initiation of snRNA genes *in vitro* (Gunderson et al., 1990). This highlights that additional factors need to be identified.

1.1.7.2: snRNA gene transcription

The 3′ end formation of pol II-transcribed snRNAs depends on the 3'box RNA processing element (Medlit et al. 2003) that is a 13 nucelotide sequence downstream of the coding region. 3' end formation of snRNAs it is a two step process:

1) Generation of precursors of snRNA with extra nucelotides at the 3' end
2) Shortening of the precurosors to the mature snRNA (Hernandez, 1985).

Cuello et al. 1997, demonstarted using nuclear run-on analysis of endogenous genes, that transcription continues beyond the 3' box of the U2 genes in vivo. This further suggest that the 3' box is not an efficient transcriptional terminator but processing element.

Sequences upstream from the 3' box of the U2 snRNA gene influence 3' end formation. In addition, the phosphorylated CTD of pol II activates the snRNA processing event in vitro (Uguen & Murphy, 2003). In particular, phosphorylation of serine-7 facilitates interaction with the snRNA gene-specific Integrator complex (Egloff et al. 2007). Two subunits of the Integrator complex, Int9 and Int11, are responsible for the first cleavage during the formation of the 3' end of pre-snRNAs (Baillat et al., 2005).

3′ box processing of snRNA is promoter-dependent

There has been evidence inidating that the 3'box processing of snRNA genes is only supported by snRNA promoter. Pilch & Marzluff (1991) demonstrated that placing both a U1 snRNA promoter and a U1 snRNA 3' end around a histone coding region

results appropriate U1 ends. In case for mRNA genes, polyadenylation factors are recruited to the mRNA promoter (Dantonel et al. 1997). Similarly, the snRNA promoter might recruit 3' end processing factors (Zaborowska et al. 2012).

The CTD links snRNA promoter and 3′ box processing

Medlin and collegues demonstrated that phosphorylation of pol II CTD on the Ser2 is required for correct 3' end snRNA formation. Truncation of the CTD of RNA polymerase II and treatment of cells with CTD kinase inhibitors resulted in a reduction in proper 3' end formation of the snRNA transcripts (Medlin et al. 2003). CTD kinase inhibitors didn't affect initiation and elongation of snRNA genes (Medlin et al. 2003). Contrary, treatment of CTD inhibtors affects transcriptional elongation of mRNA (Price, 2000). Furthermore, the double phosphorylation mark of pol II CTD on Ser2P/Ser7P is necessary for efficient binding of Integrator to the CTD (Egloff et al., 2010; Figure 1.4). This further highlights the role of the CTD in snRNA processing.

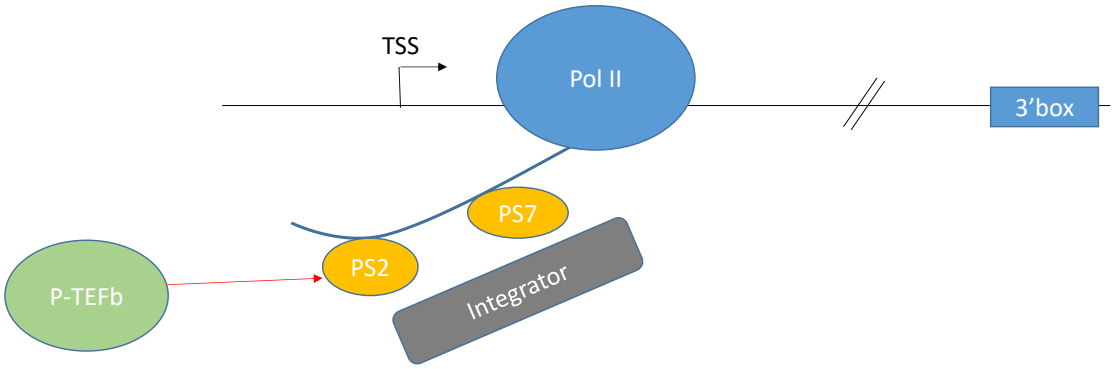


Figure (1.4): Function of P-TEFb phosphorylation in snRNA gene transcription. The P-TEFb phosphorylates Ser2 (PS2) of pol II CTD. The double phosphorylation on Ser2 and Ser7 recruits Integrator for efficient RNA processing

1.1.7.3: Transcriptional elongation of snRNA genes

Transcription complexes at the snRNA promoter are not able to elongate efficiently over long DNA distances (Sheldon et al, 1993). As the elongation capacity is defined by the TAF composition (Yankulov et al, 1994), it is not surprising that the snRNA gene-specific TAF subcomplex (snTAFc) is loaded onto snRNA genes (Zaborowska et al. 2012). This complex contains TBP, TAF5, TAF6, TAF8, TAF9, TAF11, and TAF13. Interestingly, the proposed snTAFc does not contain the classical core TAFs (TAF4 and TAF10) and lacks the highest-molecular weight TAFs, TAF1 and TAF2 (Zaborowska et al. 2012).

As mentioned before, two negative transcription elongation factors DSIF and NELF, are recruited to the elongation complex on mRNA genes to induce pol II pausing (Yamada et al., 2006, Andrulis et al., 2000, Price, 2000). Release from the pause site is mediated by the action of the Cdk9 subunit of P-TEFb. CDK9 phosphorylates DSIF, NELF and Ser2 of the pol II CTD. P-TEFb is detected by ChIP on snRNAs (Medlin et al, 2005) but the snRNA gene transcription is not affected by P-TEFb inhibition. However, at the start of these studies, the presence of NELF and DSIF on the U2 gene had not been reported. It is tempting to speculate that the differeces between mRNA and snRNA genes in NELF and DSIF recruitment might cause the difference in transcriptional regulation between mRNA and snRNA genes. Alternatively, chromatin structure or the absence of elongation factors critical for long-range transcription of protein-coding genes could undelie the differences in elongation properties of pol II initiated at snRNA promoter.

The experiments described in this thesis were designed to test the last two possibilities. I undertook a systematic analysis of the chromatin structure of the β-actin and snRNA genes and the known elongation factors associated with these genes.

2 Chapter Two: Results

2.1: The entire transcribed region of the U2 snRNA gene is nucleosome depleted.

2.1.1: Preface

Members of the snRNA gene group are transcribed by either pol II (U1, U2, U3, U4, U5) or pol III (U6, 7SK). The U1 and U2 snRNA genes have been extensively used as models to study transcription of snRNA genes in general (Egloff et al, 2008). The work presented in this thesis has been carried out on the U2 snRNA genes. The human U2 snRNA gene cluster is found on chromosome 17 and consists of 10 to 20 genes per haploid genome, organized as a nearly perfect tandem array of an approximately 6-kb region (RNU2 locus) (Lindgren et al, 1985; Westin et al, 1984).

In comparison to protein-coding genes, snRNA genes are intronless, contain a 3′ box processing element instead of a polyadenylation signal, and are also much shorter (less than 1 kb) than mRNA genes, which can span megabases in length. Additionally, snRNA gene promoters are TATA-less and include a ‘proximal sequence element’ (PSE) and a ‘distal sequence element’ (DSE) (Egloff et al, 2008) (Figure 1.3). P-TEFb is recruited to these genes during transcription and CDK9 inhibitors drastically inhibit snRNA processing at the specialized 3′ box processing element (Medlin et al, 2005).

Unlike mRNA genes, P-TEFb activity is not required for elongation of transcription of snRNA genes, highlighting significant differences in the control of elongation of transcription of snRNA and protein-coding genes.

It has been shown previously that the chromatin state of genes can influence transcription (Li et al, 2007a). The chromatin structure and histone modification patterns within the transcribed regions of these different gene types may therefore give insights into any mechanisms associated with control of elongation which are operating in a gene-type specific manner.

To determine the chromatin structure, I have carried out a comparative analysis of the U2 snRNA and the β-actin protein-coding gene in HeLa cells with respect to association of the H3 and H2B core histones by performing high-resolution ChIP, coupled with quantitative real-time PCR (qRT-PCR) analysis. This work has been recently published in MCB (Egloff et al, 2009b).

2.1.2: Distribution of pol II across the β-actin and U2 snRNA genes

High-resolution pol II ChIP was performed to map the transcription units of the β-actin and the U2 snRNA in HeLa cells. In a ChIP experiment, cells are initially treated with formaldehyde to crosslink DNA-binding proteins to DNA. HeLa cells are lysed and the DNA is sheared into small fragments by sonication. DNA fragments are immunoprecipitated by treatment with antibodies against the protein of interest.the crosslinked precipitates of protein/DNA complexes are then heated to reverse the crosslinking interaction. This is then followed by the protease treatment. Real time PCR is used to quantify the purified DNA of interest enriched with

immunoprecipitation of the specific protein under study (see Materials and Methods (Chapter 6)).

Figure (2.1A and B) shows a model of the U2 snRNA (RNU2) locus and the β-actin gene with the relative locations of the PCR primers indicated. With regard to the β-actin gene, primer pair 1 is located within a region of the gene considered not to be transcribed and as expected, pol II does not associate with this part of the gene (Figure 2C). Primer pair 2, which amplifies the promoter region, shows a strong signal for pol II association. However, in comparison to primer 2, primer pair 3 shows a fourfold decrease in pol II levels with levels continuing to drop towards the end of the transcription unit with primers 4-6. The pol II antibody (Santa Cruz, N-20) used in this study is known to immunoprecipitate both the pre-initiation (hypophosphorylated) and transcriptionally engaged (hyperphosphorylated) forms of pol II. Previous studies have shown that high concentrations of pol II are typically found at the promoter regions of protein-coding genes and considerably lower levels of pol II are found through the coding region (Cheng & Sharp, 2003). It is therefore likely that a significant proportion of the pol II detected at the promoter includes both the hyper- and hypophosphorylated forms while the levels associated within the body of the gene are mainly the active hypophosphorylated form. Using primer pair 6, only very little pol II can be detected. However, it has been shown that transcription proceeds beyond primer 6 and that the β-actin mRNA contains exon 6 (wherein primer pair 6 is located) (Gromak et al, 2006). Thus there may be some active RNA pol II in this region but at levels that are not detected by this assay. In conclusion, on the β-actin gene, pol II is recruited to all regions tested with a stronger association seen at the promoter region.

For the U2 snRNA gene, primer pairs 1, 2, and 3 give a very low signal. This is expected, as these target non-transcribed regions. Primer pair 4, which amplifies the PSE region of the U2 snRNA gene promoter, shows the strongest signal of all tested regions. In common with the pol II found at the β-actin promoter, at least some of the polymerase detected here may be inactive. Primer pair 5 which amplifies just downstream of the 3′ box, shows a slightly lower signal than that detected on primer 4. A further twofold decrease is detected on the region amplified by primer pair 6. The U2 snRNA gene ChIP results are in accordance with previous studies where the whole of the U2 snRNA gene transcription unit has been mapped by nuclear run-on analysis to extend to approximately 1 kb downstream from the start of transcription (Medlin et al, 2003).

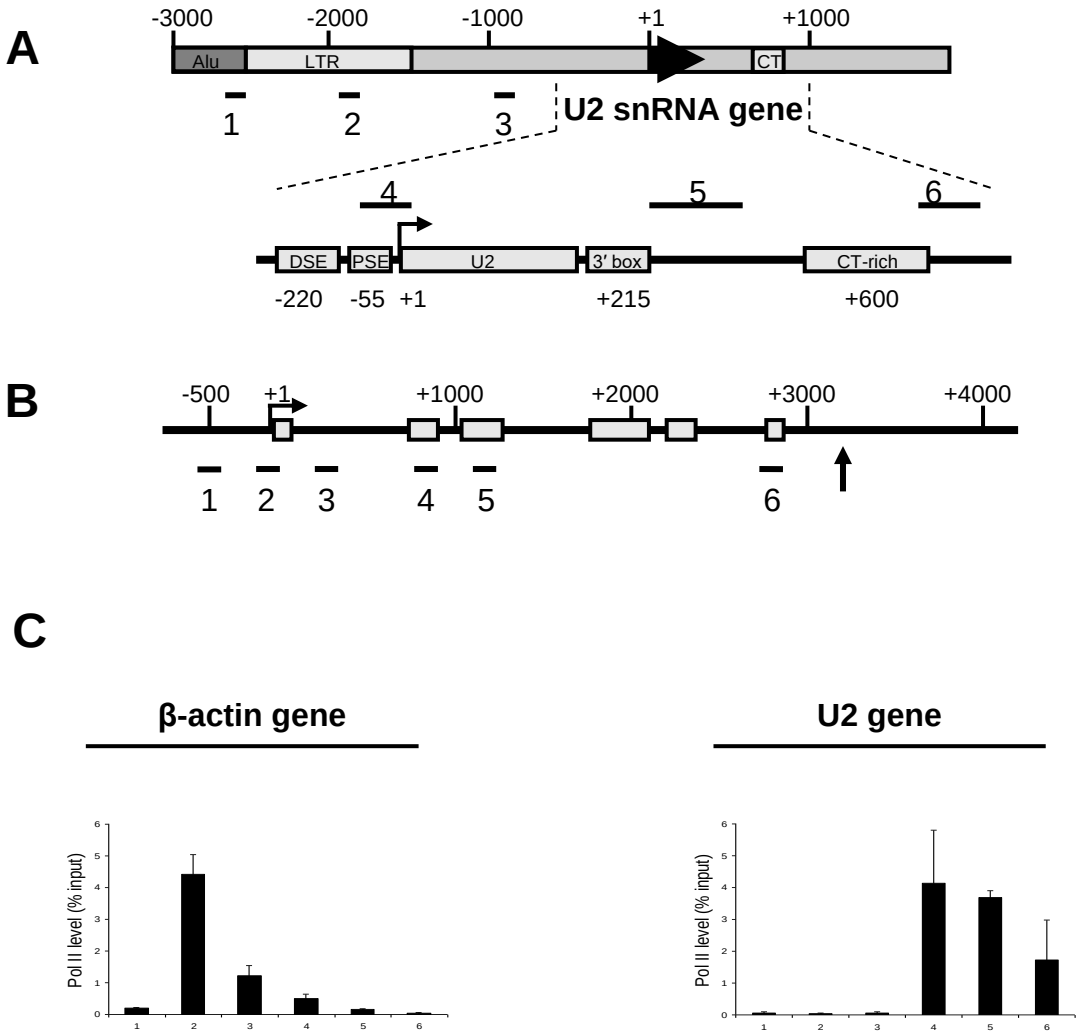


Figure (2.1): Localization of pol II with the β-actin and U2 snRNA genes.

A) Model of the human U2 snRNA gene and its surrounding regions (RNU2 locus) within the human genome. The 6-kb repeat of the human RNU2 locus, with U2 gene promoter and RNA processing elements noted in the expanded diagram below. The repeat also contains a CT microsatellite, two tandem Alu repeats, and a long-terminal-repeat element (LTR). The positions of the pairs of primers used for ChIP are shown below the 6-kb repeat and above the expanded gene diagram. Arrows indicate the start of transcription.

B) Diagram of the human β-actin gene with the exons depicted as open boxes and the polyadenylation signals indicated with an arrow. The positions of primer pairs used for ChIP are shown below the diagram.

C) qRT-PCR of pol II ChIP analysis on the β-actin and RNU2 locus. The y axes represent the percentages of input DNA immunoprecipitated. Error bars represent standard deviation of 3 independent experiments.

2.1.3: The chromatin landscape of the RNU2 locus and the β-actin gene.

In vivo, transcription takes place within the context of chromatin, where pol II traverses nucleosomes that constitute the histone proteins H2A, H2B, H3, and H4. A previous study in *Drosophila* showed that a large proportion of active genes have a nucleosome-free region where the +1 nucleosome is typically located at position +62 (relative to TSS) (Mavrich et al, 2008). Since the transcription unit of the U2 snRNA gene differs from the β-actin gene in terms of size (5 kb vs 1 kb for the β-actin and U2, respectively) (Gromak et al, 2006), structure (U2 lack exons) and yields transcripts that are not polyadenylated, ~200 bp from the TSS (Yuo et al, 1985) I wondered whether the chromatin landscape also changes. Towards this aim, I analysed the association of the H3 and H2B core histones with the human β-actin protein-coding genes and the U2 snRNA genes by high resolution ChIP, coupled with quantitative real-time PCR (qRT-PCR) (Figure 2.2).

The β-actin and U2 snRNA genes both have an extended area of nucleosome depletion of about 800 bp downstream from the start of transcription (Figure 2.2). Notably, this covers the entire transcription unit of the U2 gene whereas the density of nucleosomes increase within the rest of the coding region of β-actin gene (Figure 2.2) see primers 4-6) and in the non-transcribed region of the U2 repeat (Figure 2.2) see primers 1-3). In addition, the nucleosome profile across the β-actin gene also differs from previously published analyses of other protein-coding genes, for example, the GAPDH and γ-actin genes have a much shorter promoter-proximal region of nucleosome depletion and nucleosomes are at a high level approximately 200 bp downstream from the start of transcription (Egloff et al, 2009b).

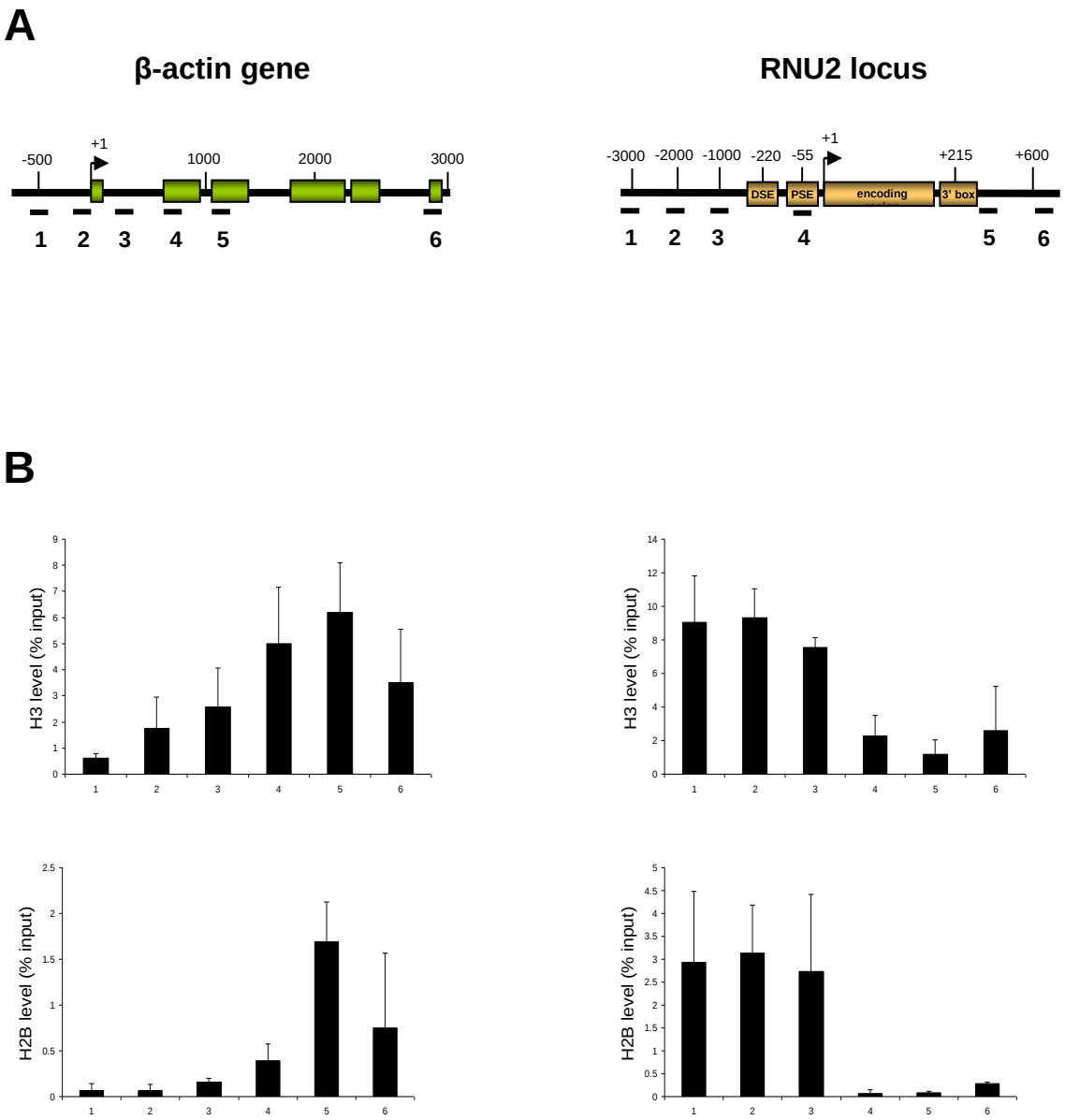


Figure (2.2): Localization of H3 and H2B on the β-actin and U2 snRNA genes.

RT-PCR of ChIP analysis of the β-actin and U2 snRNA genes using antibodies against histones H3 and H2B. The error bars represent the standard error of three repeats.

3 Chapter Three: Results

3.1: Some elongation factors are associated with both the β-actin and U2 snRNA genes.

3.1.1: Preface

Several studies provide evidence that nucleosomes can efficiently pause transcribing pol II *in vitro* and *in vivo* and that histone acetyltransferases (HATs), chromatin remodelers (Swi/snf, FACT), TFIIS, and activation domains of transcription factors are all implicated in relief of the pause (Bondarenko et al, 2006; Brown et al, 1996; Brown & Kingston, 1997; Carey et al, 2006; Kireeva et al, 2005; Li et al, 2007a; Petesch & Lis, 2008; Wada et al, 2000). On both the β-actin and U2 snRNA genes, NELF recruitment correlates with the increase in nucleosome density, suggesting that NELF and chromatin together cause pol II to pause or, in the case of the snRNA genes, terminate (Egloff et al, 2009b). However, it is not clear why pol II cannot negotiate the NELF/nucleosome block on the U2 snRNA genes, as P-TEFb activity is present. One testable possibility is that differential association of factors required for transcription of nucleosomal templates may underlie differences in NELF function on protein-coding and snRNA genes. Accordingly, I have first analysed the association of TFIIS and FACT (Facilitates Chromatin Transcription) with the U2 snRNA and β-

actin genes (Figure 3.1), as these factors have been found to play roles in elongation of transcription on chromatin templates (Guermah et al, 2006; Kireeva et al, 2005; Orphanides et al, 1998).

TFIIS is considered the best biochemically studied transcription elongation factor (Wind & Reines, 2000). TFIIS is known to increase the efficiency of elongating pol II by reactivating the arrested pol II and interacting directly with it to induce its endonucleolytic activity (Palangat et al, 2005; Reines & Mote, 1993) which results in cleavage of the expelled RNA and therefore alleviates the pol II arrest mode. Pol II arrest is the result of a prolonged pause and extensive backtracking of the enzyme (Kettenberger et al, 2003).

FACT was first characterized in 1998 and is identified as a factor that aids pol II in transcription through nucleosomes (Orphanides et al, 1998). FACT has been reported to destabilize nucleosomal structure during passage of pol II by removing the histone H2A/H2B dimers (Belotserkovskaya et al, 2003; Reinberg & Sims, 2006). In addition, it is also able to re-assemble the H2A/H2B dimers back onto the DNA and thus can function as a histone chaperone (Belotserkovskaya et al, 2003). Subsequent studies showed that FACT is composed of two subunits, SSRP1 and Spt16 with the molecular weights of 80 and 140 kDA respectively (Orphanides et al, 1999).

3.1.2: Cloning of TFIIS and SSRP1 (FACT).

I have cloned TFIIS and the SSRP1 subunit of FACT, since no commercial antibodies were available. TFIIS and SSRP1 were obtained by PCR amplification from HeLa cell cDNA, using appropriate 5' and 3' end primers (Materials and Methods Chapter

6), which was then cloned into a Flag-tagged vector (pFLAG-CMV-2) as described in Materials and Methods (Chapter 6) (Figure 3.1A). To determine whether these clones were expressed, HeLa cells were transfected with plasmids encoding either Flag-tagged TFIIS or SSRP1 and expressed TFIIS and SSRP1 were detected by western blot analysis with an anti-Flag antibody. Non-transfected HeLa cells were used as a negative control. Western blot analysis detected expression of a 35 and 80 kDa size protein product of the transiently transfected TFIIS and SSRP1 respectively, as expected (Figure 3.1B).

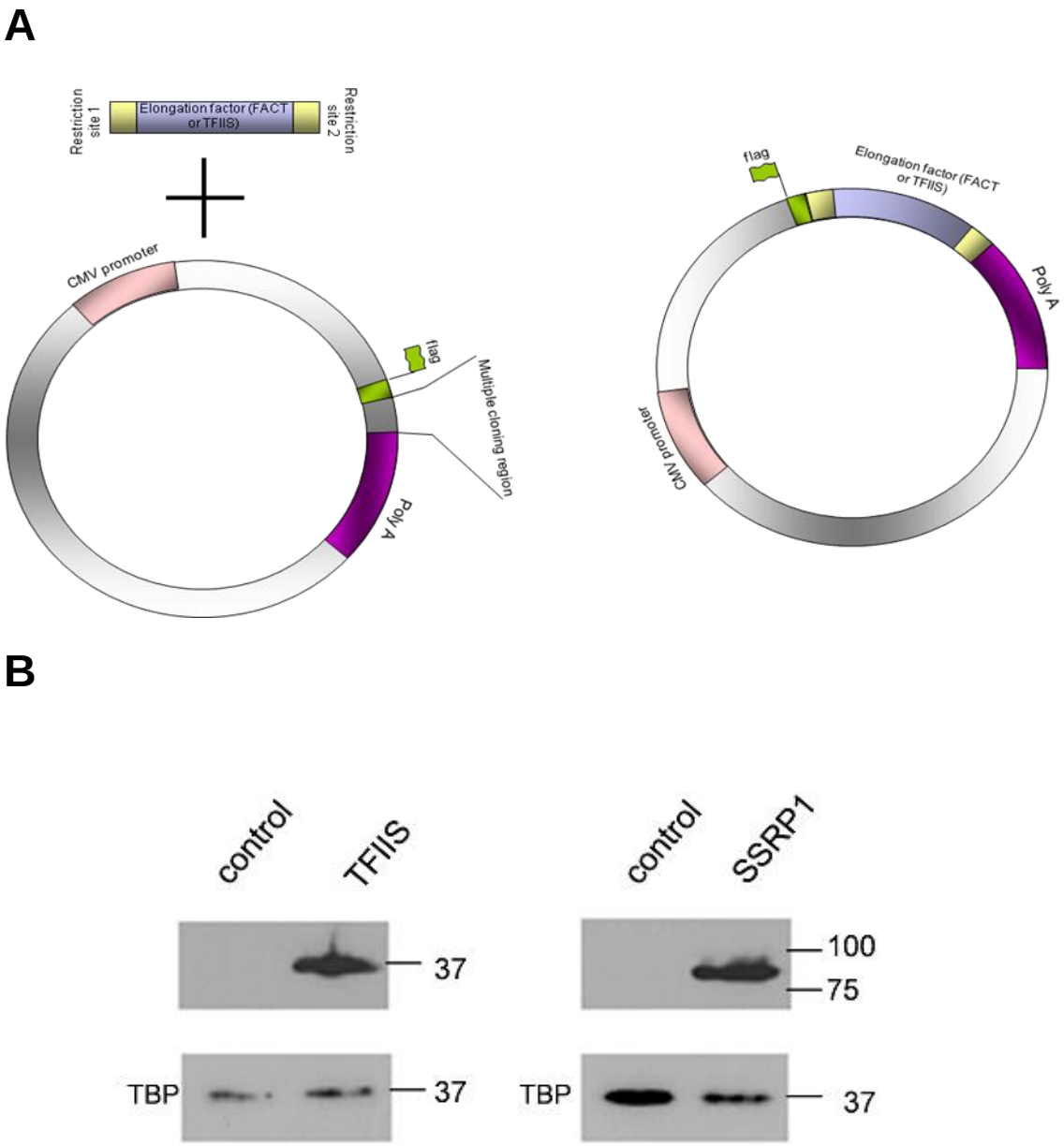


Figure (3.1). Western blot analysis of TFIIS and SSRP1.

A) Vector Flag-tagging. A schematic diagram of the cloning technique used for the production of a Flag- tagged elongation factor TFIIS and SSRP1 (FACT subunit).

B) Western blot analysis of TFIIS (left) and SSRP1 (right) protein in HeLa cells after transfection of Flag-tagged TFIIS and SSRP1 plasmids. TATA Binding Protein (TBP) was used as loading control. Molecular weight markers (kDa) are indicated on the right of each figure.

3.1.3: TFIIS and FACT are associated with both snRNA genes and protein-coding genes.

As described above, TFIIS and FACT have been shown to play roles in elongation of transcription on chromatin templates (Guermah et al, 2006; Kireeva et al, 2005; Orphanides et al, 1998). Association of TFIIS and FACT elongation factors with the β-actin and snRNA genes was assessed by ChIP using antibodies against Flag-tagged TFIIS and SSRP1, and a commercial antibody against the endogenous Spt16 subunit of FACT.

TFIIS is associated with both genes and correlates with pol II occupancy (compare Figure 3.2B-*top graph* and Figure 2.1C). In addition, the ratio of TFIIS to pol II is approximately the same for both genes (Egloff et al, 2009b), suggesting that TFIIS travels with the polymerase *in vivo* regardless of chromatin structure. The two subunits of FACT are also found to be associated with both genes (Figure 3.2B). However, on snRNA genes, FACT association largely correlates with pol II occupancy (compare Figure 3.2B-*middle and bottom graphs* and Figure 2.1C*)*. In contrast, FACT appears to be recruited to the β-actin gene at the same position where nucleosome density increases, which also coincides with NELF recruitment (compare Figure 3.2B-*top graph* and Figure 2.2B*)* (Egloff et al., 2009). As previously described (Egloff et al, 2009b), and in accordance with Gomes et al. (Gomes et al, 2006), FACT association is drastically reduced after treatment of cells with the CDK9 inhibitor, DRB, indicating that P-TEFb activity is required for FACT association, most likely through Ser2 phosphorylation of the CTD (Gomes & Espinosa, 2010). This association is further supported by the close correlation between FACT localization and Ser2 phosphorylation on both types of genes (Egloff et al, 2009b). Thus, although

the transcribed region of U2 gene is largely nucleosome-depleted, TFIIS and FACT are still recruited to the genes.

These results therefore indicate that the failure of pol II to negotiate the NELF-dependent block at the end of the U2 snRNA genes is not due to the failure to recruit TFIIS and FACT. The role of FACT in elongation of transcription of snRNA genes is unclear, since DRB treatment abolishes FACT association (Gomes et al, 2006) without affecting transcription of snRNA genes (Medlin et al, 2003) (Egloff et al., 2009).

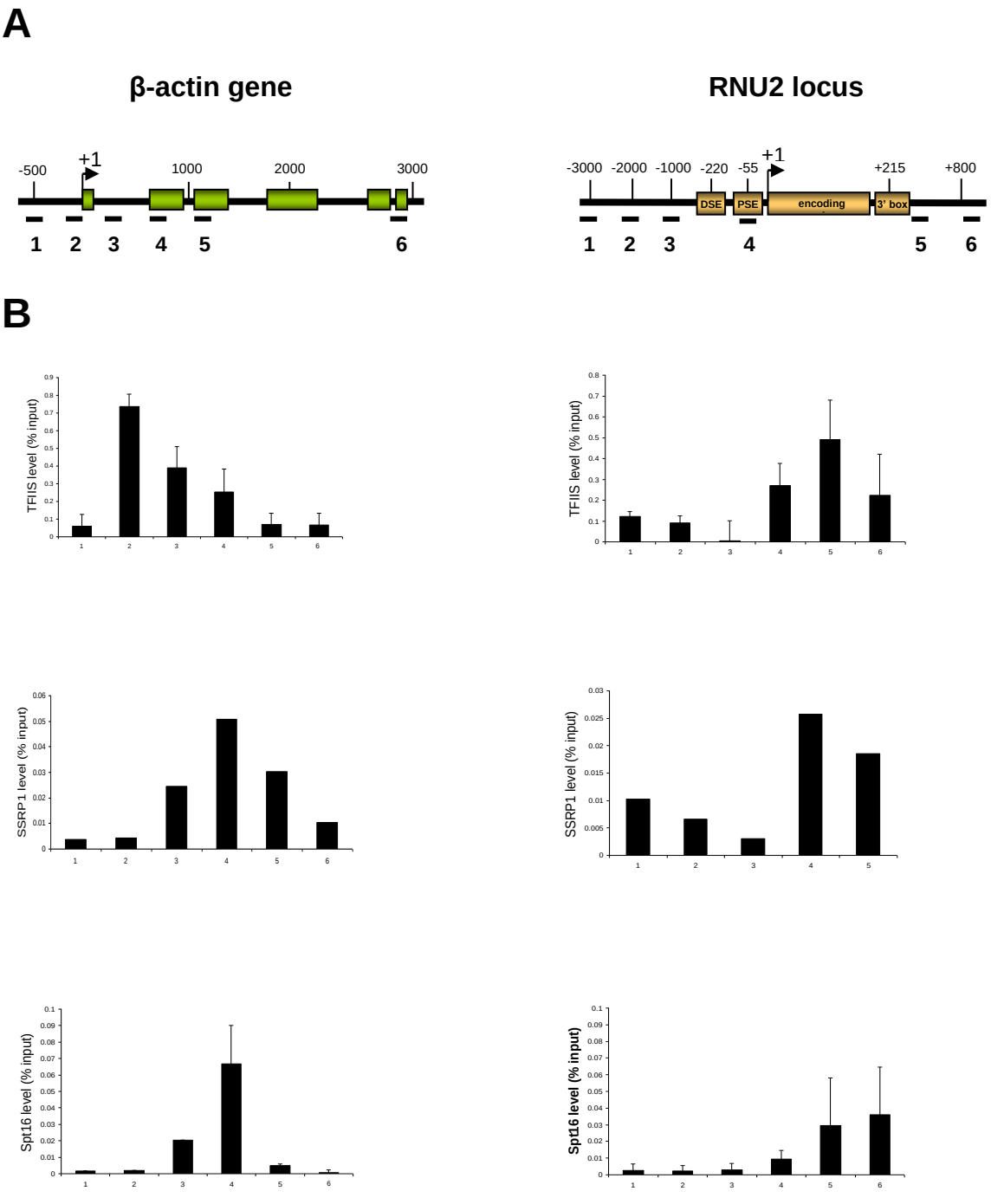


Figure (3.2). TFIIS and FACT are associated with both snRNA and protein-coding genes.

A) schematic diagram of the β-actin gene and the RNU2 locus.

B) RT-PCR of ChIP analysis of the RNU2 locus and β-actin genes, using antibodies against the Flag-tagged TFIIS (top graphs), Flag-tagged SSRP1 (FACT) (middle graphs) and the endogenous Spt16 (bottom graphs). The y axes represent the percentages of input DNA immunoprecipitated. Error bars represent standard deviation of 3 independent experiments.

4 Chapter Four: Results

4.1: Some transcription elongation factors are differentially associated with the U2 snRNA and the β-actin genes.

4.1.1: Spt6 is differentially associated with the β-actin and U2 snRNA genes.

Spt6 (suppressor of Ty) was first identified in *Saccharomyces cerevisiae* as a member of a class of suppressors of δ insertion mutations in the 5′ noncoding regions of HIS4 and LYS2 genes together with Spt4, Spt5, and Spt16 (Clark-Adams & Winston, 1987; Sims et al, 2004a; Winston et al, 1984). It is an essential protein involved in transcriptional elongation and was initially characterized as a histone chaperone (Andrulis et al, 2000; Hartzog et al, 1998; Kaplan et al, 2000). Spt6 is a transcription elongation factor known to associate with pol II and move along with the RNA elongation complex towards the 3′ region, reassembling the chromatin on recently transcribed sequences after the passage of the polymerase (Bres et al, 2008; Endoh et al, 2004; Ivanovska et al, 2011). The chaperone activity of Spt6 acts on H3/H4 and is important for the maintenance of chromatin in yeast (Bortvin & Winston, 1996; Selth et al, 2010). Accordingly, the absence of Spt6 causes a reduction in the level of histones from yeast genes, leading eventually to transcription initiation at cryptic sites

(Kaplan et al, 2003). Further to the histone chaperone activity of Spt6, earlier studies in yeast highlighted genetic interactions between Spt6 and Spt4 and Spt5 (DSIF) (Swanson & Winston, 1992). Similarly, interactions of Spt6 with the Paf1 complex, TFIIS, FACT, and pol II have been demonstrated in a number of studies (Costa & Arndt, 2000; Hartzog et al, 2002; Krogan et al, 2002b; Squazzo et al, 2002; Swanson & Winston, 1992).

The distribution of Spt6 on the β-actin and U2 snRNA genes was assessed by ChIP using a commercial anti-Spt6 antibody. For the β-actin gene, Spt6 levels correlate with pol II occupancy (compare Figure 4.1B and Figure 2.1C). In contrast, no signal above background is found on the U2 gene (Figure 4.1B). The low signal found on the transcription unit of the U2 snRNA gene is similar to the levels detected on the non-transcribed regions (primer pairs 1, 2 and 3), suggesting that it likely shows nonspecific background levels. The observation that Spt6 is highest at the promoter of the β-actin gene contrasts with previous data suggesting that human Spt6 is only recruited to pol II phosphorylated on Ser2 (Yoh et al, 2007), but is in agreement with the finding that in yeast, Spt6 is recruited to promoters together with Paf1 and FACT (Adelman et al, 2006).

In conclusion, Spt6 seems to associate exclusively with the β-actin gene, in contrast to the histone chaperone FACT (described above), which is found on both the β-actin and U2 snRNA genes. This result therefore suggests that Spt6 may play a gene-type specific role in elongation of transcription.

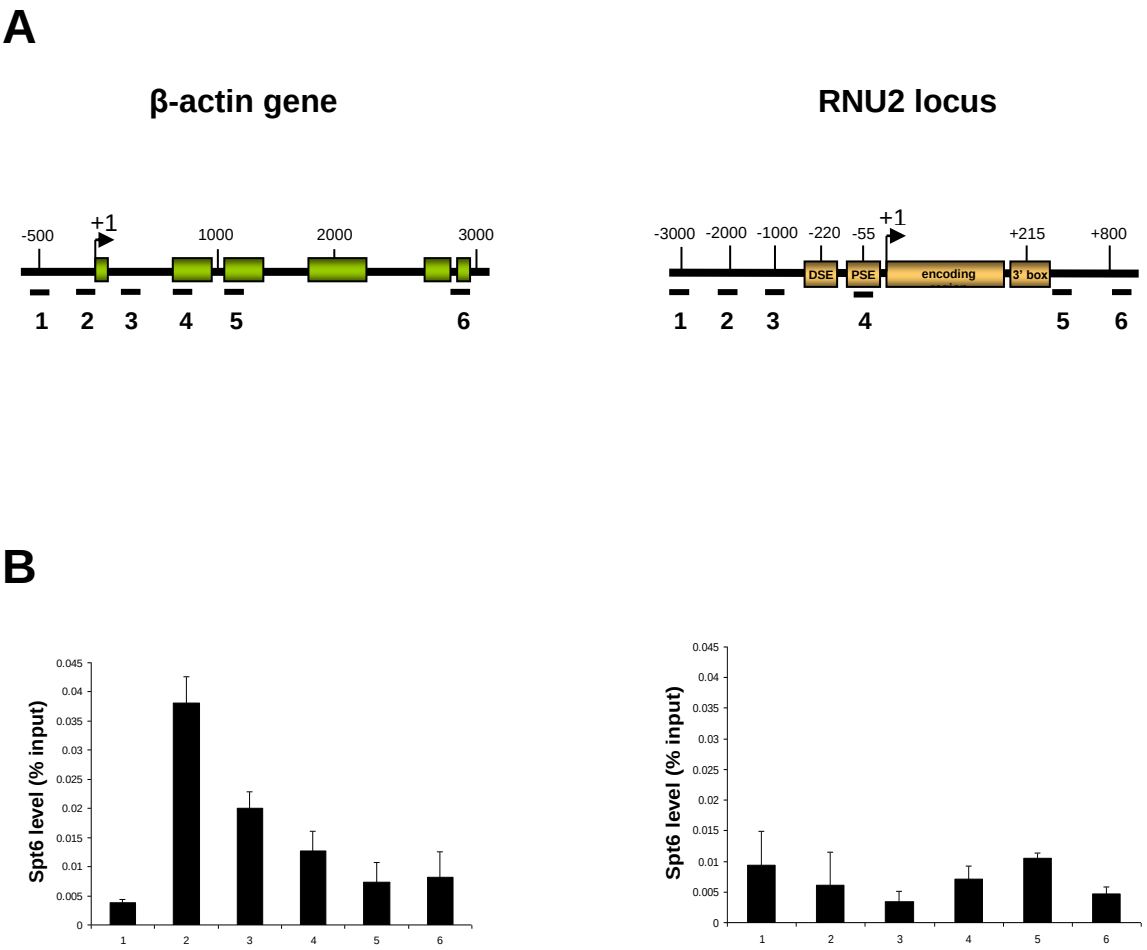


Figure 4.1 Spt6 is differentially associated with the β-actin and U2 snRNA genes.

A) Schematic diagrams of the β-actin gene and the RNU2 locus. In the β-actin schematic, the arrow represents the transcription start, and boxes and lines denote exons and introns, respectively. In the RNU2 locus schematic, boxes represent the distal sequence element (DSE) and proximal sequence element (PSE). Horizontal lines, indicated in both schematics, mark the position of the ChIP primers.

B) RT-PCR of ChIP analysis of the RNU2 locus and β-actin genes, using antibodies against the endogenous Spt6. The y axes represent the percentages of input DNA immunoprecipitated. Error bars represent standard deviation of 3 independent experiments.

4.1.2: PAF is differentially associated with the β-actin and U2 snRNA genes.

PAF (polymerase associated factor) is a complex that was initially identified in yeast and contains Paf1, Cdc73, Ctr9, Leo1 and Rtf1 subunits (Mueller & Jaehning, 2002; Squazzo et al, 2002). Paf1 and Cdc73 were the first two subunits to be identified in a screening study for transcription factors that associate with yeast pol II (Wade et al, 1996). The role of PAF in elongation has been suggested by ChIP experiments, which showed that PAF subunits localize to the open reading frame of transcribed genes (Krogan et al, 2002b). The human form of the PAF complex contains an additional subunit named hSki8, a part of the SKI complex involved in 3′-5′ mRNA degradation (Zhu et al, 2005a). In addition, the human PAF complex was shown to associate with pol II and colocalize with the enzyme on actively transcribed genes (Chen et al, 2009; Kim et al, 2010; Rozenblatt-Rosen et al, 2005). Cdc73 was the first characterized subunit of the human PAF complex, and was discovered through mutations of the parafibromin gene in hyperparathyroidism-jaw tumor syndrome (HPT-JT) (Carpten et al, 2002). Mutations in Cdc73 have also been linked with different types of cancer; such as breast, renal and gastric (Newey et al, 2009). In contrast to a number of transcription factors that interact with the phosphorylated form of pol II, PAF was shown to associate with the unphosphorylated form of pol II (Rozenblatt-Rosen et al, 2005) and can be detected at or near promoters and within the coding regions of the genes (Kim et al, 2010; Pokholok et al, 2002; Rozenblatt-Rosen et al, 2009). Furthermore, it has been proposed that PAF acts as a platform to recruit factors such as RAD6/Bre1 and Set1, which are responsible for histone H2B monoubiquitination

and histone H3 lysine 4 trimethylation (H3K4-Me3) respectively, to elongation-engaged pol II (Kim et al, 2010; Qiu et al, 2006).

The yeast Bur1 kinase, which is the yeast equivalent of human CDK9 (P-TEFb), has previously been shown to phosphorylate the C-terminal region (CTR) of the Spt5 subunit of DSIF (Wood & Shilatifard, 2006). Similarly, the Thr-4 residues in the pentapeptide repeats of the CTR of human Spt5, undergoes phosphorylation by P-TEFb (Ivanov et al, 2000; Kim & Sharp, 2001; Yamada et al, 2006). A relatively recent study described that phosphorylation of Spt5 by Bur1 kinase is important for recruiting the PAF complex (Liu et al, 2009).

In summary and in relation to my study, one of the major functions of PAF is to help coordinate a subset of histone modifications; PAF is required for the ultimate recruitment of the methyltransferases Set1 and Set2, resulting in H3K4-Me3 and H3K36-M3 respectively (Chu et al, 2007; Krogan et al, 2003a; Ng et al, 2003).

Association of the PAF elongation factor was assessed with ChIP using an antibody against the endogenous Paf1 subunit of the complex. The immunoprecipitated DNA fragments associated with PAF were then quantified by real-time PCR.

On the β-actin gene, PAF is detected mainly on primer pairs 3 and 4 while it is very low or absent from the U2 snRNA gene (Figure 4.2B). In agreement with previous studies (Adelman et al, 2006; Kim et al, 2010; Kim et al, 2004; Pokholok et al, 2002), PAF is detected throughout the coding region of the β-actin gene with a signal starting downsteam of the promoter at +350 bp and decreasing towards the 3′ end of the transcribed region (Figure 4.2B). As the U2 snRNA gene has very low histone occupancy across the whole of the transcribed region of about 800 bp (Egloff et al, 2009b), the absence of PAF on the U2 snRNA gene correlates with its function in elongation through chromatin (Kim et al, 2010). Thus, taken together, PAF is another

factor that associates exclusively with the β-actin gene, suggesting that it may act in a gene-type specific manner.

Interestingly, PAF shows a different profile to Spt6, which is highest at the promoter similarily to Pol II. The “early” elongation checkpoint on the β-actin gene occurs approximately +800 bp from the transcription start, which is in contrast to most mRNA genes occurring within the first 100 bp (Li & Gilmour, 2011). Therefore it is possible that the PAF profile on the β-actin gene is linked to the position of the checkpoint rather than the PIC.

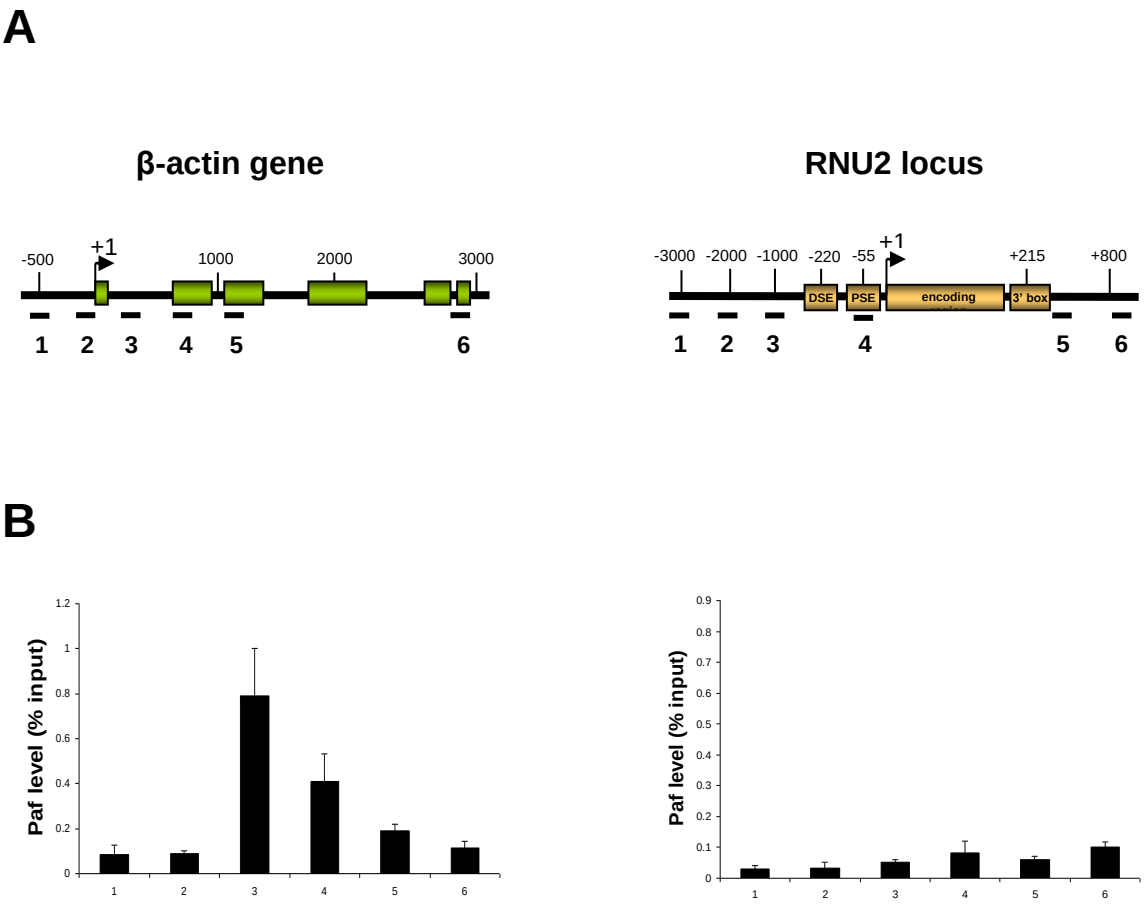


Figure (4.2): PAF is differentially associated with the β-actin and U2 snRNA genes.

A) schematic diagrams of the β-actin gene and the RNU2 locus.

B) RT-PCR of ChIP analysis of the RNU2 locus and β-actin genes, using antibodies against the endogenous Paf1 subunit of the PAF complex. The y axes represent the percentages of input DNA immunoprecipitated. Error bars represent standard deviation of 3 independent experiments.

4.1.3: DSIF is present at different levels on the β-actin and the U2 snRNA genes.

DSIF (DRB-sensitivity-inducing factor) was first discovered as a factor that causes pol II pausing in the presence of the P-TEFb inhibitor 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) (Wada et al, 1998). DSIF is a heterodimeric complex composed of Spt4 and Spt5 subunits. SPT genes were first discovered in *Saccharomyces cerevisiae* in an analysis that identified suppressors of δ insertion mutations in the 5′ noncoding regions of HIS4 and LYS2 genes (Winston et al, 1984). DSIF is a factor of dual activities with both positive and negative effects. Initially, DSIF was only known as an inhibitor of transcription elongation through its engagement in a rate limiting step of the elongation process known as promoter proximal pausing (Saunders et al, 2006; Wu et al, 2003). In this step, and soon after transcriptional initiation, DSIF helps to recruit NELF, which arrests transcription, causing a checkpoint block (Yamaguchi et al, 2002). Subsequent phosphorylation of the Spt5 subunit of the DSIF complex, the NELF-E subunit of NELF, and serine 2 of the CTD by the P-TEFb complex (Fujinaga et al, 2004; Yamada et al, 2006), which contains the cyclin-dependent kinase (CDK9) and cyclin T (Sims et al, 2004b), results in the release of the negative block and allows pol II to resume elongation (Bres et al, 2008).

Several studies have shown that DSIF interacts with a number of factors such as FACT, TFIIS, Spt6, and PAF (Costa & Arndt, 2000; Hartzog et al, 1998; Krogan et al, 2002b; Lindstrom et al, 2003; Mueller & Jaehning, 2002; Orphanides et al, 1999; Squazzo et al, 2002).

DSIF is also known to associate with pol II and they are thought to track along the DNA together (Wada et al, 1998; Yamaguchi et al, 1999b). In support of this notion, a recent genome wide ChIP-seq study on murine embryonic stem (ES) cells revealed that DSIF corelates with pol II throughout the genome (Rahl et al, 2010).

Association of the DSIF elongation factor was assessed with ChIP using an antibody against the endogenous Spt5. The immunoprecipitated DNA fragments associated with both Spt5 was then quantified by real-time PCR. For both the β-actin and U2 snRNA genes, ChIP results indicate that Spt5 is found where pol II is detected (compare Figure 4.3 and Figure 2.1C). However, on the U2 snRNA gene, the level of DSIF is approximately four times less than that observed on the β-actin gene. When the level of Spt5 is divided on the level of pol II, the difference is even more drastic, with at least 10 fold less Spt5 on the U2 snRNA gene than found on the β-actin gene (Figure 4.4).

ChIP analysis, using an antibody to phospho-Spt5 (kind gift of Yuki Yamaguchi, Tokyo Institute of Technology), indicated that the phosphorylated form of Spt5 is also present on both the β-actin and snRNA genes but again the level on the U2 sRNA genes is approximately 4 times lower. Moreover, when the level of phospho-Spt5 is ratioed to the level of pol II, the difference between the β-actin and U2 snRNA genes is even more extreme with also 10 fold less phospho-Spt5 on the U2 snRNA gene than that found on the β-actin gene

In summary, phosphorylated Spt5 was shown in a number of studies to act as a prerequisite for PAF recruitment, which in turn, as described above, recruits a wide spectrum of factors such as RAD6/BRE1 and Set1/Set2. On the basis of this, I speculate that the relatively low level of phospho-Spt5 on the U2 snRNA gene is

critical to the differential recruitment of several important elongation factors and may help explain the failure of pol II to transcribe beyond the nucleosomal block.

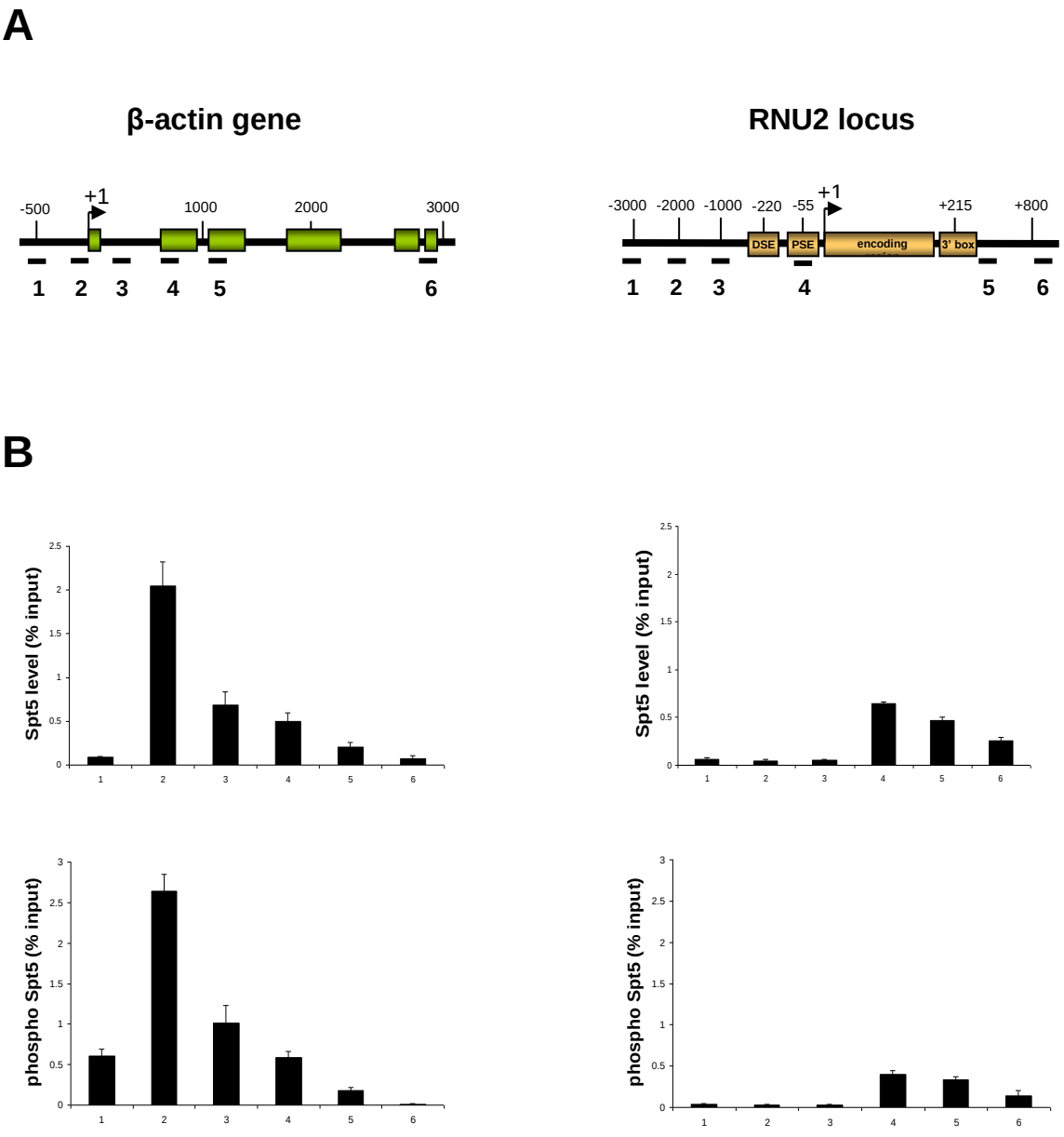


Figure (4.3): Spt5 (DSIF)/phospho Spt5 (DSIF) is differentially associated with the β-actin and U2 snRNA genes.

C) schematic diagrams of the β-actin gene and the RNU2 locus.

D) RT-PCR of ChIP analysis of the RNU2 locus and β-actin genes, using antibodies against the endogenous Spt5 (Top graphs), phosphor Spt5 (Bottom graphs) The y axes represent the percentages of input DNA immunoprecipitated. Error bars represent standard deviation of 3 independent experiments.

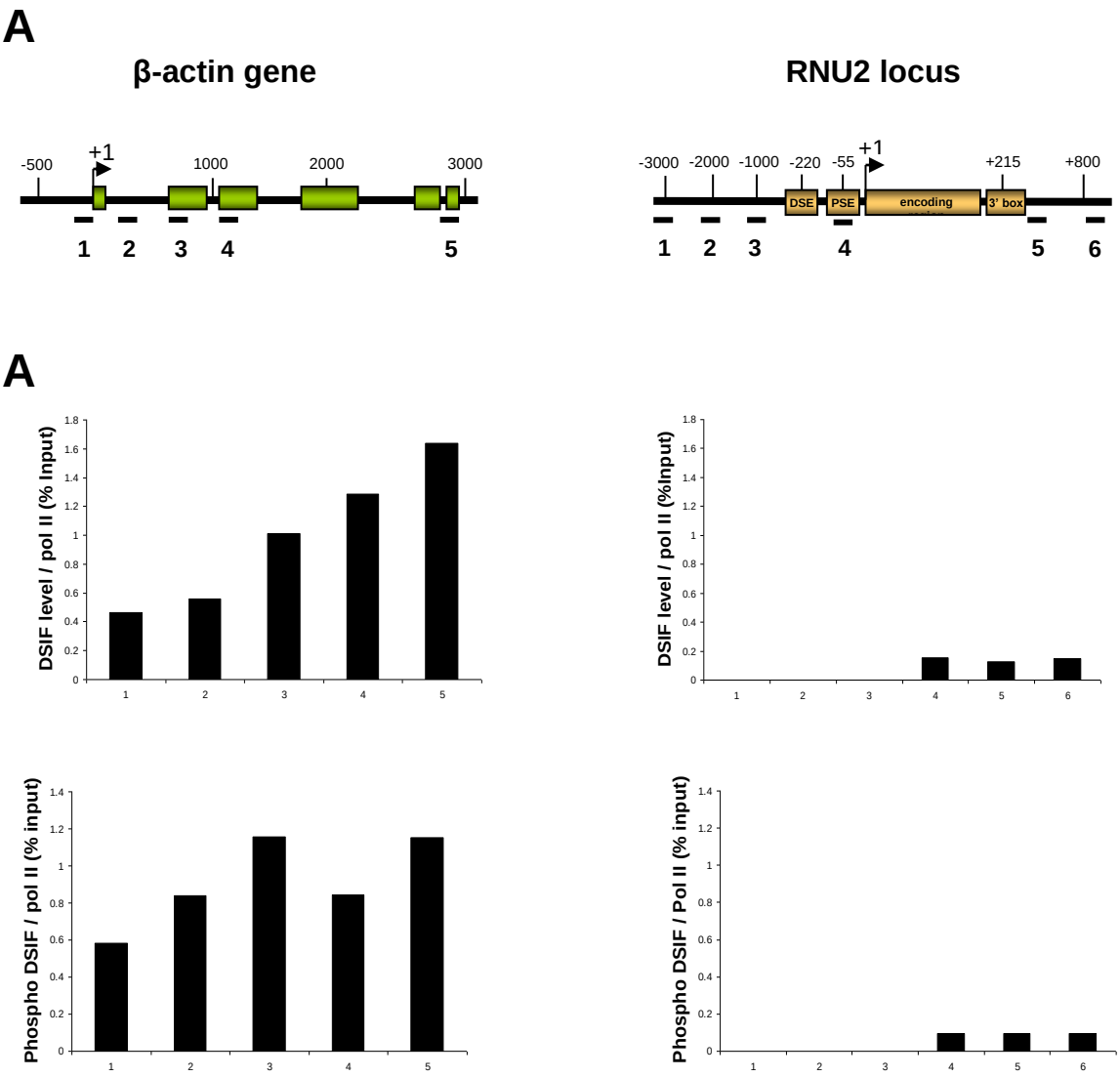


Figure (4.4): Levels of Spt5 (DSIF)/phosphor Spt5 (DSIF) relative to pol II on both β-actin and U2 snRNA genes.

A) schematic diagrams of the β-actin gene and the RNU2 locus.

B) Levels of DSIF (Top) and phospho DSIF (Bottom) normalized to pol II, on the β-actin and U2 genes.

4.2: Histone modifications and histone modification enzymes associated with the RNU2 locus and the β-actin gene.

4.2.1: Set1 is differentially associated with the β-actin and snRNA genes.

The N-terminal tails of histones undergo a number of posttranslational modifications which have an impact on the fate of transcription at all stages. Some histone modifications are associated with active transcription such as acetylation of histones H3 and H4 by histone acetyltransferases (HATs) or trimethylation of H3K4 and H3K36 by methyltransferases (HMTs) (Li et al, 2007a), whereas other histone modifications such as H3K9 trimethylation mirror gene repression.

In yeast, Set1 is the methyltransferase for histone H3 lysine 4 (H3K4) (Briggs et al, 2001; Krogan et al, 2002a; Miller et al, 2001; Noma & Grewal, 2002). Generally speaking, histone methylation is linked to both active and inhibitory events of the chromatin (Briggs et al, 2001). It is known that both histone H3 and H4 undergo extensive methylation, where lysine residues 4, 9, 27, and 36 in H3 and lysine 20 in the H4 histone are substrates for a number of methyltransferases (Briggs et al, 2001). Set1 is one of the members of the SET-domain containing proteins (Nislow et al, 1997; Schultz et al, 2000). H3K4 occurs as either mono-, di-, or trimethylation, and the resulting methylation ranges between promoters and within coding regions of active genes (Sims et al, 2004a). For example, H3K4 trimethylation is known to associate with active transcription and is generally detected surrounding the transcription start sites (Egloff et al, 2009b). This is in accordance with data from genome-wide analysis showing the enrichment of H3K4 trimethylation at the 5′ end

of actively transcribing genes (Barski et al, 2007; Bernstein et al, 2005). Set1 binds pol II CTD phosphorylated on Ser5 (Ng et al, 2003) and requires PAF as a prerequisite (Krogan et al, 2003a; Ng et al, 2003).

As an understanding of the chromatin lansdscape can lead to a better understanding of the process of transcription, I have determined the levels of Set1 on the RNU2 locus and the β-actin gene by ChIP using a commercial anti-Set1 antibody.

For the β-actin gene, Set1 occupies the promoter proximal region of the gene. In contrast, no signal above background is found on the U2 snRNA gene (Figure 4.5B).

Since it is known that methylation of H3K4 by Set1 requires the presence of PAF (Krogan et al, 2003a; Laribee et al, 2005), the fact that Set1 is absent from the U2 snRNA gene is in accordance with a previous result I have shown, where on the U2 snRNA gene, PAF is also not detected. Therefore the absence of Set1 on the U2 gene could be a direct consequence of the absence of PAF. However, a previous study conducted by the Murphy lab showed that the H3K4 trimethylation histone mark is present on both the β-actin and U2 snRNA genes (Egloff et al, 2009b). In other words, on the U2 snRNA gene, H3K4me3 is present while Set1 is missing. In yeast, Set1 is specific for H3K4me3. However, in human cells, more than 10 H3K4 methyltransferase candidates exist (Ruthenburg et al, 2007). Thus, it is likely that the H3K4me3 mark on the U2 snRNA and β-actin genes is added by a different methyltransferase.

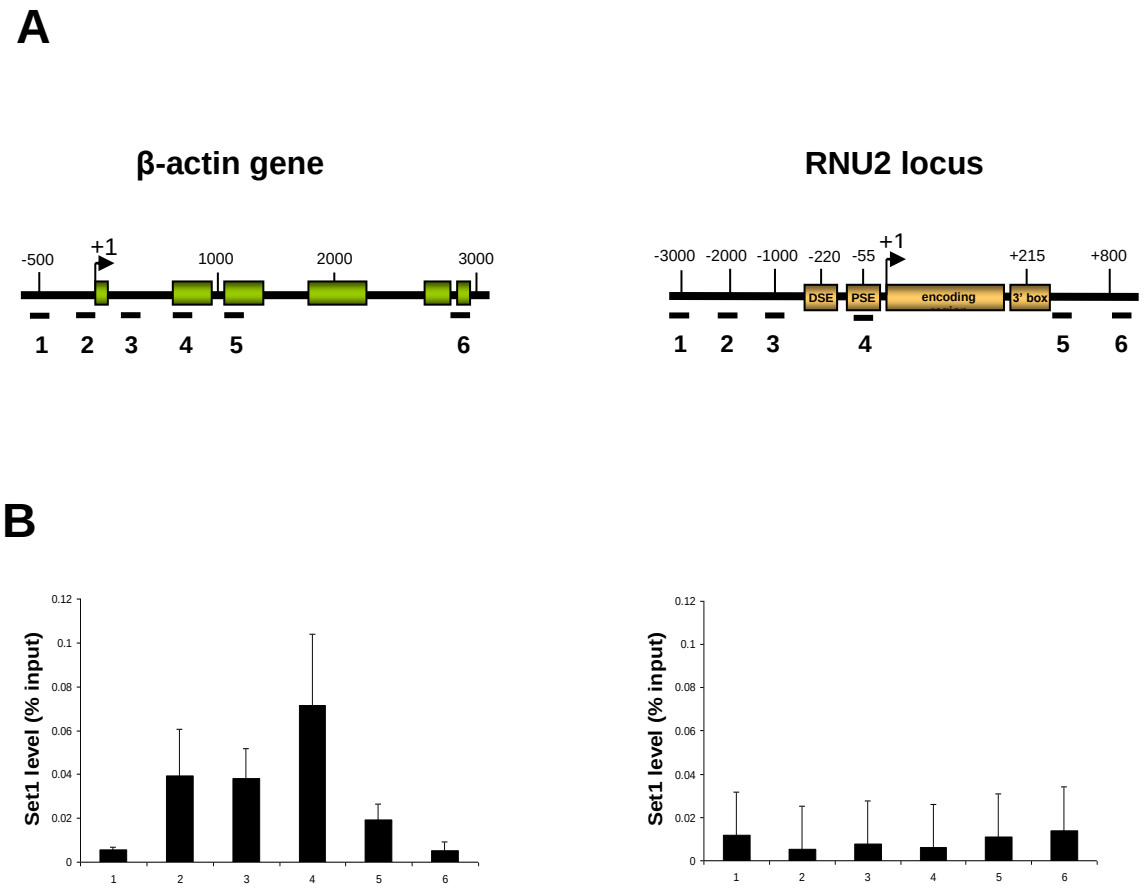


Figure (4.5): Set1 is differentially associated with the β-actin and U2 snRNA genes.

A) schematic diagrams of the β-actin gene and the RNU2 locus.

B) RT-PCR of ChIP analysis of the RNU2 locus and β-actin genes, using antibodies against the endogenous Set1. The y axes represent the percentages of input DNA immunoprecipitated. Error bars represent standard deviation of 3 independent experiments.

4.2.2: H3K36me2 is differentially associated with the β-actin and U2 snRNA genes.

H3K36 dimethylation (me2) carried out by Set2 is a vital histone modification (Strahl et al, 2002). H3K36me2 is a posttranslational modification that generally mirrors a state of active gene transcription (Li et al, 2007a). Generally speaking, histone methylation is a modification that is identified by its function in signaling a number of different physiological processes. For example, it was previously proposed that histone methylation mediates the switch between acetylation and deacetylation states of the chromatin in yeast, where the histone deacetylase Rpd3s was found to recognize only the methylated form of histones and minimally needs H3K36me2 for successful recruitment (Li et al, 2009).

Human Set2 directly binds the pol II CTD phosphorylated on Ser5 and Ser2 (Li et al, 2005). As a result, H3K36me2 is strictly associated with the elongating form of pol II and is scarcely present on upstream regions of genes (Rao et al, 2005). Furthermore, Set2 recruitment in yeast requires PAF as deletion of PAF results in less recruitment of Set2 (Krogan et al, 2003b). Thus, Set2 recruitment is associated with the elongation phase of transcription (Kizer et al, 2005; Li et al, 2002; Schaft et al, 2003).

The distribution of H3K36me2 on the β-actin and U2 snRNA genes was assessed by ChIP using a commercial anti-H3K36me2 antibody. The result of real-time PCR analysis is shown in (Figure 4.6). H3K36me2 associates with the β-actin gene at approximately 800 bp downstream from the start of transcription (primer pair 4), a site that correlates with Ser2 phosphorylation and the “late” transcription checkpoint (Egloff et al, 2009b). No detectable levels upstream or downstream from the

checkpoint were found. In contrast, no signal above background is found over the entire U2 snRNA gene (Figure 4.6B).

Although Ser2 phosphorylation is found on pol II transcribing both the β-actin and U2 snRNA genes, it appears to differentially influence the chromatin landscape. In yeast, H3K36 methylation helps to recruit deacetylases to inhibit cryptic-gene transcription initiation on long or infrequently transcribed genes (Li et al, 2007b). Set2 activity may therefore be unnecessary for the correct expression of the short U2 genes.

Furthermore, H3K36 trimethylation associates with the β-actin gene and not to the U2 snRNA gene (Egloff et al, 2009b), where the mark was shown to peak at the checkpoint and through out the body of the gene denoting the elongating form of Pol II (Egloff et al, 2009b; Shilatifard, 2006).

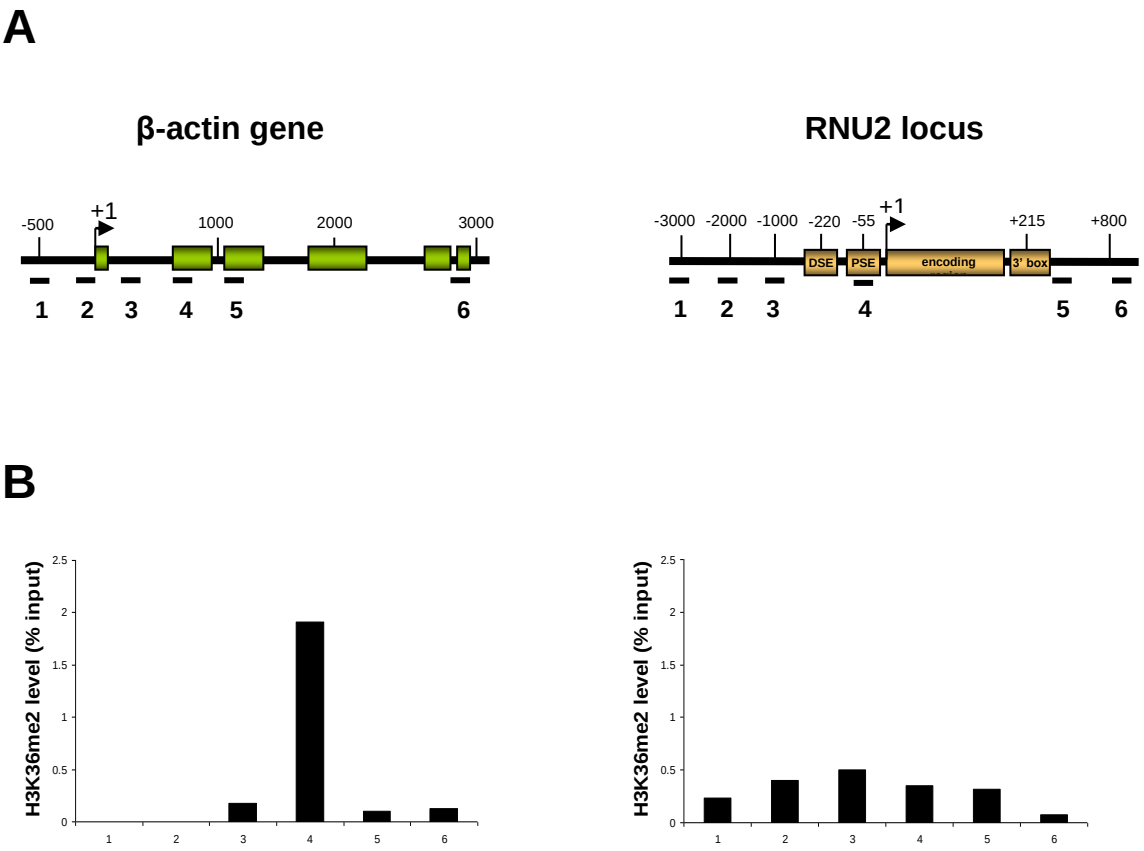


Figure (4.6): H3K36me2 is differentially associated with the β-actin and U2 snRNA genes.

A) schematic diagrams of the β-actin gene and the RNU2 locus.

B) RT-PCR of ChIP analysis of the RNU2 locus and β-actin genes, using antibodies against the endogenous H3K36me2. The y-axes represent the percentages of input DNA immunoprecipitated. The HK36me2 level is shown relative to H3.

4.2.3: H2Bub1 is differentially associated with the β-actin and U2 snRNA genes.

H2B ubiquitination is carried out by the action of three enzymes; E1, E2, and E3 which are the ubiquitin activation enzyme, ubiquitin conjugating enzyme, and ubiquitin ligase enzymes respectively (Hershko & Ciechanover, 1998). Initially, studies identified Rad6 and Bre1 as the E2 and E3 enzymes in yeast (Hwang et al, 2003; Robzyk et al, 2000; Wood et al, 2003a). Likewise, and after the development of a monoclonal antibody that exclusively recognizes ubiquitinated lysine 120 of H2B in human (Minsky et al, 2008), a number of studies emerged in the field that also identified Rad6 and Bre1 as the human E2 and E3 ubiquitination enzymes respectively (Kim et al, 2009). Additionally, a vital factor for H2B ubiquitination is the PAF complex, where a tight linkage exist between the Ser5 phosphorylated form of the CTD of pol II, PAF and the ubiquitination machinery (Kim et al, 2009). Furthermore, being part of an intricate cycle, H2B ubiquitination was also shown to be an essential factor implicated in H3K4 and H3K79 methylation events by Set1 and Dot1, respectively (Dover et al, 2002; Kim et al, 2009). In addition, an in vitro study which was conducted in the context of chromatin, showed H2B ubiquitination to be indispensible for FACT-mediated transcription (Pavri et al, 2006).

The distribution of H2B lysine 120 monoubiquitination was assessed by ChIP using a monoclonal anti-H2BK120ub1 antibody. H2BK120ub1 is detected at the promoter proximal region of the the β-actin gene on primer pairs 2, 3 and 4. In contrast, no signal above background was found on U2 snRNA gene (Figure 4.7B). The observation that H2BK120ub1 is found only on the β-actin gene fits with the previous result, mentioned above, that the PAF complex is also missing from the U2 snRNA

genes. As described above, PAF is a factor involved in H2B ubiquitination through interactions between Bre1 and PAF (Kim et al, 2009). Thus, in the absence of PAF, no H2BK120ub1 occurs.

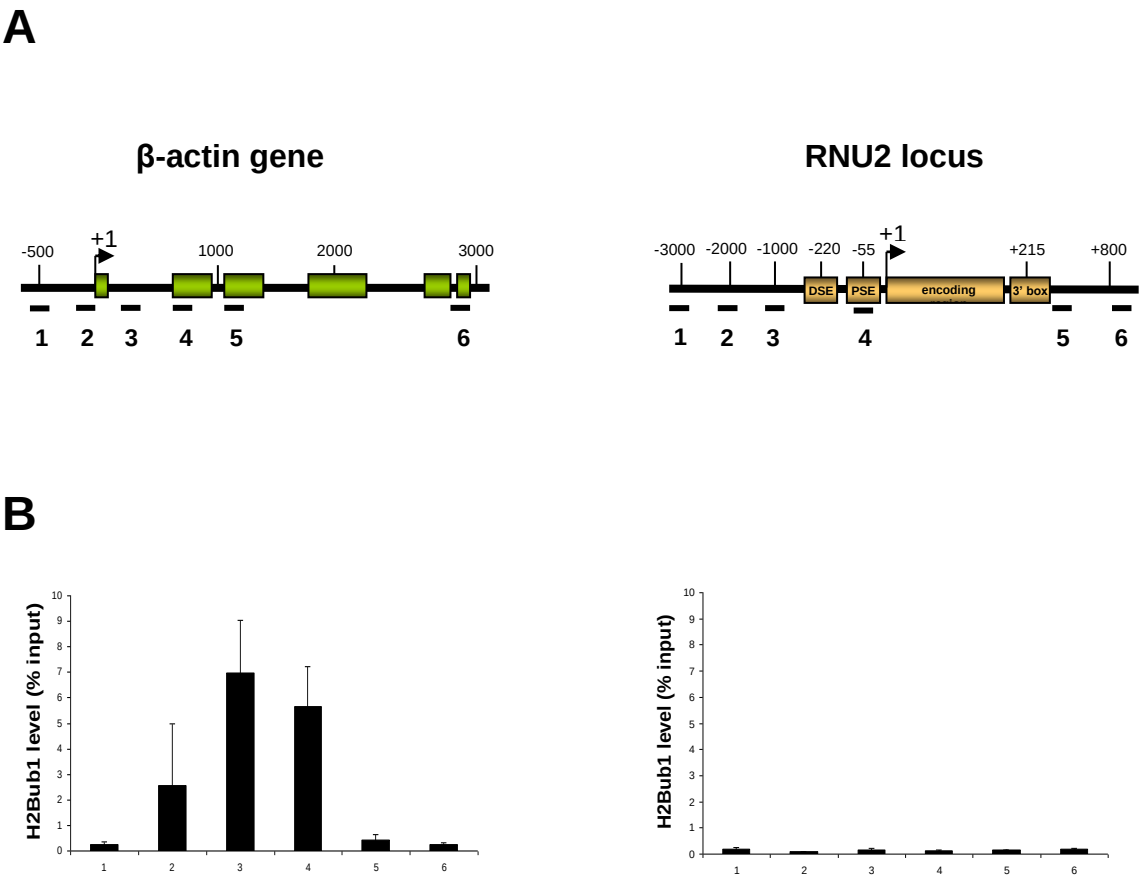


Figure (4.7): H2BK120ub1 is differentially associated with the β-actin and U2 snRNA genes.

A) schematic diagrams of the β-actin gene and the RNU2 locus.

B) RT-PCR of ChIP analysis of the RNU2 locus and β-actin genes, using antibodies against the endogenous H2BK120ub1. The y axes represent the percentages of input DNA immunoprecipitated. Error bars represent standard deviation of 3 independent experiments. The H2BK120ub1 level is shown relative to H2B.

4.2.4: Association of RAD6A and RAD6B with the β-actin and U2 snRNA genes.

Since I have shown that there is no H2B monoubiquitination associated with the U2 snRNA genes, and in order to understand the molecular mechanism underlying this difference between snRNA and protein-coding genes, I have analysed the association of proteins involved in H2B monoubiquitination with the U2 snRNA and β-actin genes.

H2B monoubiquitination requires the action of three classes of enzymes – E1 ubiquitin enzyme, E2 ubiquitin conjugating enzyme, and E3 ubiquitin ligase (Kim & Roeder, 2011). Firstly, E1 activates ubiquitin in the presence of ATP resulting in the formation of a thioester bond to the E2 enzyme. The E3 enzyme then transfers ubiquitin from E2 to the H2B lysine 120 (Hochstrasser, 1996; Scheffner et al, 1995; Weake & Workman, 2008). H2B ubiquitylation has been implicated in active transcription as it is known to regulate histone H3 methylation at lysine 4, which is associated with the promoter regions of active genes (Weake and Workmann, Mol Cell, 2008).

The E2 (ubiquitin conjugating) enzyme, known as Rad6, was first studied in *Saccharomyces cerevisiae* (Robzyk et al, 2000; Weake & Workman, 2008). Although there are two homologs of Rad6 in humans: Rad6A and Rad6B (Koken et al, 1991), earlier studies indicated that a different E2 ubiquitin conjugating enzyme, known as UbcH6, monoubiquitinated H2B in conjunction with the E3 ubiquitin ligase complex, BRE1, in human cells (Zhu et al, 2005b). However, a recent study showed that the human Rad6, and not UbcH6, could complement the function of yeast RAD6 mutants for H2B ubiquitination and therefore it is likely that either Rad6A or Rad6B is the

primary E2 ubiquitin conjugating enzyme in human cells and works in conjunction with the BRE1 complex to monoubiquitinate H2B at lysine 120 in human cells (Kim et al, 2009). Recently, it was shown that depletion of RAD6A, but not RAD6B, results in overall reduction of the Rad6 function, indicating that Rad6A accounts for the majority of E2 ligase activity (Kim et al, 2009).

However, it has been shown that recruitment of RAD6-BRE1 per se is not sufficient for H2B monoubiquitination. The PAF complex was shown in several studies to play an essential role in RAD6-mediated ubiquitination and subsequent H3-K4 methylation (Wood et al, 2003b; Zhu et al, 2005b). More specifically, several studies have shown that the PAF complex functions as an activator for the RAD6-BRE1 complex rather than a prerequisite for its recruitment (Wood et al, 2003b).

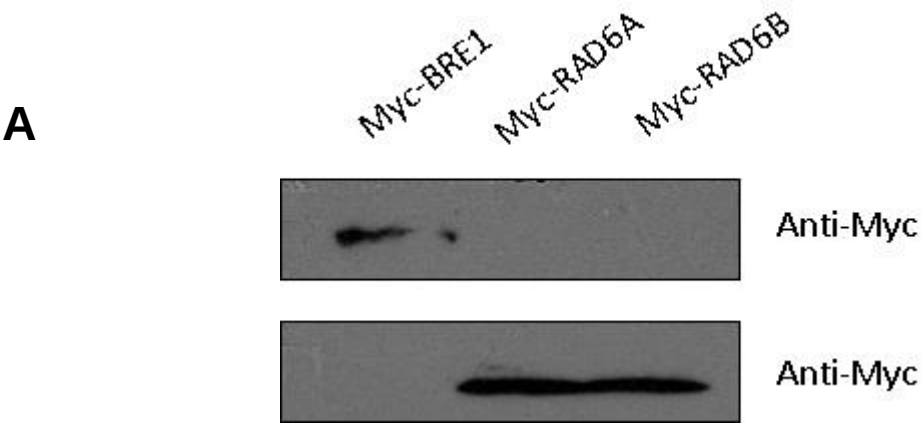
Mechanistically, the RAD6 complex requires the BRE1 for its recruitment to actively transcribed genes (Wood et al, 2003a) and RAD6A and RAD6B are both known to associate with the PAF complex only in the presence of BRE1 (Kim et al, 2009). This was confirmed by the demonstration that the paf1 subunit binds to BRE1 (Kim et al, 2009). Additionally, in yeast, the Rtf1 subunit of the PAF complex was shown to aid the interaction between pol II, the Set1-containing comlex COMPASS, and RAD6 (Wood et al, 2003b). Depletion of the CTR9 component of the PAF complex results in a global decrease of the H2B monoubiquitination and H3K4 trimethylation levels (Pirngruber et al, 2009), emphasizing that PAF is vital for H2B monoubiquitination (Kim et al, 2009; Pirngruber et al, 2009; Zhu et al, 2005a).

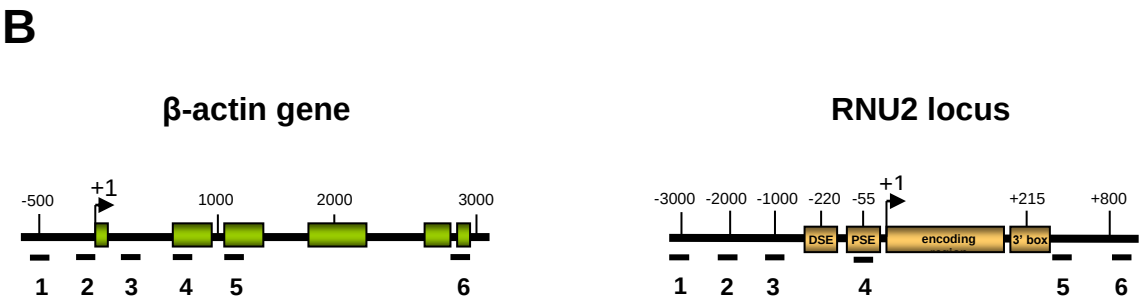
In order to study the localization of RAD6A and RAD6B on the β-actin and U2 snRNA genes, I utilized plasmids expressing tagged RAD6A and RAD6B genes, with three tandem Myc epitopes at their C terminus. To determine whether the Myc-RAD6A and Myc-RAD6B proteins are expressed, HeLa cells were transfected with

these plasmids, total protein was isolated 48 hours later and subjected to Western blot analysis, using an anti-Myc antibody. The amino acid sequence of both RAD6A and RAD6B is 96% identical (Lyakhovich & Shekhar, 2004), hence both proteins exhibit nearly the same molecular weight. The analysis detected expression of a 17 kDa sized protein product of the transiently transfected RAD6A and RAD6B, as expected (Figure 4.8).

Having confirmed the expression of RAD6A and RAD6B in HeLa cells, the association of both RAD6A and RAD6B with β-actin and U2 snRNA genes was then analysed by ChIP, coupled with quantitative real-time PCR analysis. On the β-actin gene, RAD6A and RAD6B peak at primer pair 3, which designates the promoter proximal region, and the signal drops dramatically after this point. In the case of RAD6A, the signal starts coming up again at primer pair 5, where the checkpoint is positioned. Interestingly, another peak is detected at primer pair 6 for both RAD6A and RAD6B. This may be indicative of a downstream promoter. In agreement with this, a recent study showed that the TATA binding protein (TBP) was also found to associate at this position (Zaborowska et al, 2012), Alternatively, these results may indicate the presence of a loop between the promoter and this downstream region (Hampsey et al, 2011).

On the U2 snRNA gene, RAD6A levels peak at the 3′ box and the CTCF binding site (figure 4.8). RAD6B is detected on the promoter and the 3′ box region of the U2 gene. However, the level of RAD6B is approximately four times lower than that observed on the β-actin gene.


A
Myc-BRE1
Myc-RAD6A
Myc-RAD6B
Anti-Myc
Anti-Myc


B
β-actin gene
-500
+1
1000
2000
3000
1 2 3 4 5 6
RNU2 locus
-3000 -2000 -1000 -220 -55 +1 +215 +800
DSE
PSE
encoding
3' box
1 2 3 4 5 6

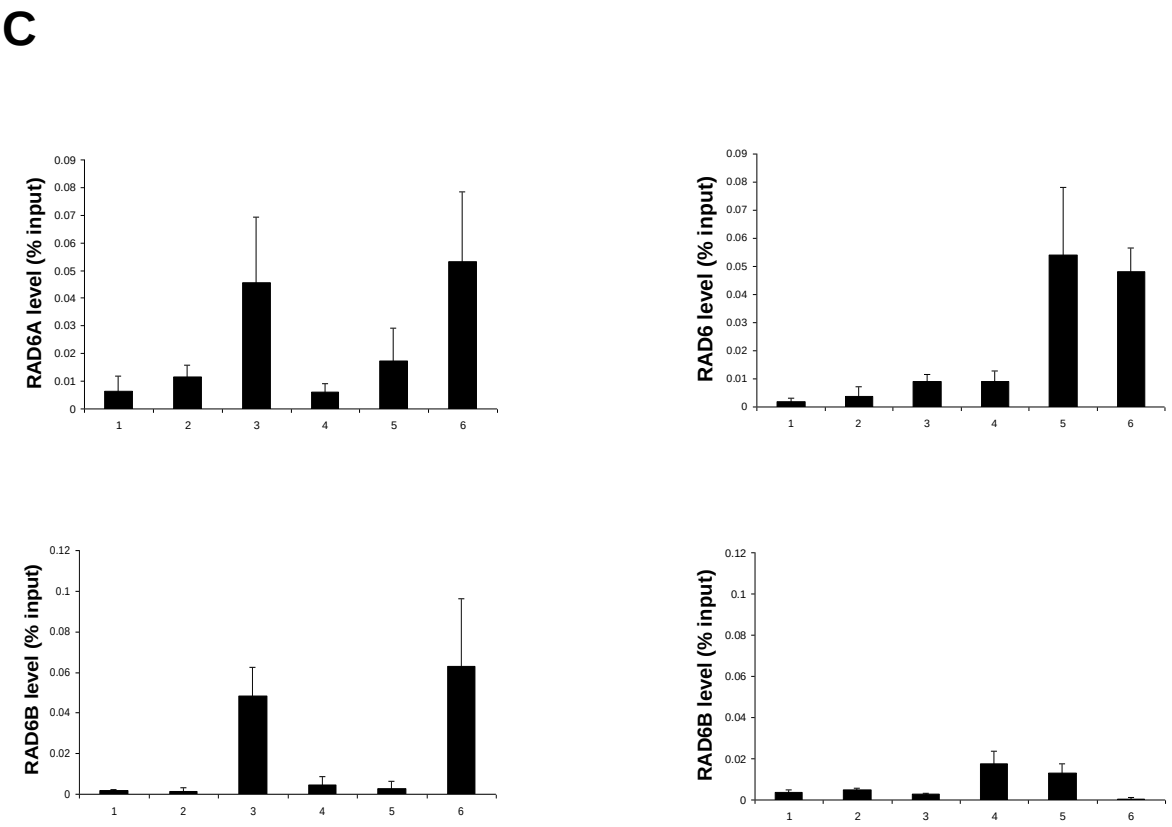

C
RAD6A level (% input)
RAD6 level (% input)
RAD6B level (% input)
RAD6B level (% input)

Figure (4.8) RAD6A and RAD6B are associated with both snRNA and protein-coding genes.

A) Western blot analysis of BRE1 (top left), RAD6A (bottom middle), and RAD6B (bottom right) protein in HeLa cells after transfection of Myc-tagged BRE1, RAD6A, and RAD6B.

B) schematic diagram of the β-actin gene and the RNU2 locus.

C) RT-PCR of ChIP analysis of the RNU2 locus and β-actin genes, using antibodies against the Myc-tagged RAD6A (top graph), Myc-tagged RAD6B (bottom graph). The y axes represent the percentages of input DNA immunoprecipitated. Error bars represent standard deviation of 3 independent experiments.

4.2.5: Association of BRE1 with the β-actin and U2 snRNA genes.

RDA6 was previously shown to bind three different RING finger containing ubiquitin ligases: Ubr1, Rad18, or Rad5 (Bailly et al, 1994; Dohmen et al, 1991; Johnson et al, 1992) However, none of these ubiquitin ligases was found to function in the H2B monoubiquitination pathway.

The first reported ubiquitin ligase required for H2B monoubiquitination was the RING finger containing protein BRE1 in *Saccharomyces cerevisiae* (Hwang et al, 2003). BRE1 was shown to act as the E3 ubiquitin ligase of the H2B monoubiquitination enzymatic cascade (Kim et al, 2005; Wood et al, 2003a).

The human homologs of the yeast BRE1 are BRE1A (RNF20) and BRE1B (RNF40) (Hwang et al, 2003). BRE1A and BRE1B share 47% and 43% sequence resemblance, respectively, with the yeast BRE1 (Hwang et al, 2003). Additionally, BRE1A binds to BRE1B (Zhu et al, 2005b) and are both required to monoubiquitinate H2B in vivo and in vitro (Kim et al, 2009).

In yeast, BRE1 was shown to be essential to recruit RAD6 to actively-transcribed genes (Wood et al, 2003b). Although the BRE1 complex has two components in humans, BRE1A and BRE1B, RAD6 only binds to the complex and not to the individual forms of BRE1 (Kim et al, 2009). Furthermore, RAD6 binds BRE1 in a RING finger independent interaction (Kim et al, 2009).

Similar to RAD6, studies where BRE1 was either depleted or overexpressed demonstrated that this protein plays a vital role in H2B monoubiquitination and subsequent H3K4 and H3K79 methylation (Kim et al, 2005). BRE1 alone is not capable of carrying out H2B monoubiquitination (Kim et al, 2009). BRE1 is known to

connect between the PAF and RAD6 complexes. More specifically, the Paf1 subunit of the PAF complex binds directly to the BRE1 complex (Kim et al, 2009).

In this study, I analysed the association of BRE1A (RNF20) with the β-actin and U2 snRNA genes. Similar to RAD6 results described above, the BRE1A used here is Myc-tagged. Expression of the Myc-BRE1A was conducted by transfecting the expression plasmid in HeLa cells. The total protein was then isolated 48 hours later and expression analysed by Western blot. The analysis detected a 124 kDa sized protein, as expected (Figure 4.8 (A)).

The localization of BRE1A on the β-actin and U2 snRNA genes was then tested using ChIP and further analysed by real-time PCR. On both β-actin and U2 snRNA genes, BRE1 shows a similar profile to RAD6A, i.e., the protein level peaks at primer pair 3, 5, and 6. On the U2, BRE1A is found at both the 3' box (primer pair 5) and the CTCF binding region (primer pair 6).

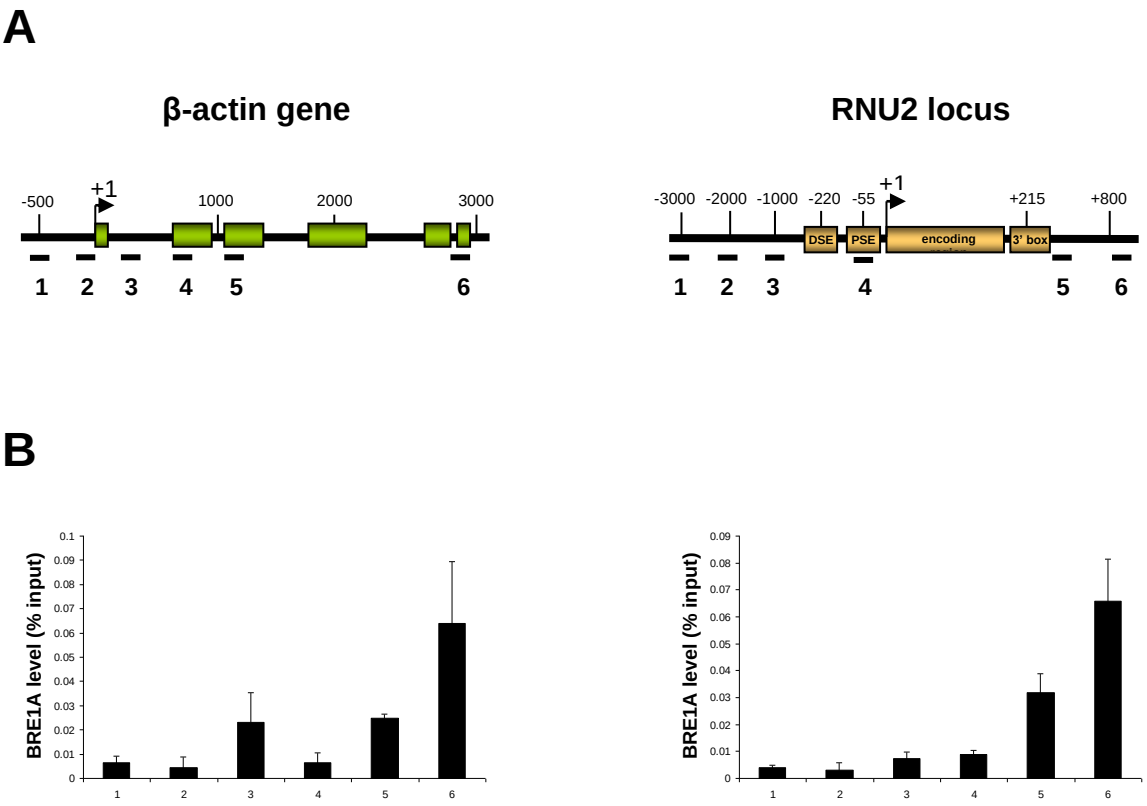


Figure (4.9) BRE1A is associated with both snRNA and protein-coding genes.

A) schematic diagram of the β-actin gene and the RNU2 locus.

B) RT-PCR of ChIP analysis of the RNU2 locus and β-actin genes, using antibodies against the Myc-tagged BRE1A. The y axes represent the percentages of input DNA immunoprecipitated. Error bars represent standard deviation of 3 independent experiments.

5 Chapter Five: Results

5.1: The PAF complex controls levels of Pol II, Spt5, and H2B specifically on the β-actin gene.

The PAF complex was shown to act as a platform for the recruitment of a wide variety of factors ranging from different transcription factors, histone modifications, and termination factors.

The Rtf1 subunit of PAF was first identified in *Saccharomyces cerevisiae* and was specifically characterized by its ability to direct binding of the TATA box binding protein (TBP) to transcription start sites (Stolinski et al, 1997). Moreover, the Rtf1 subunit is an important component of the transcription elongation machinery, as demonstrated from studies using loss-of-function and temperature sensitive (TS) mutants or siRNAs to reduce Rtf1 levels. In each study, disruption of Rtf1 function or de-regulation of Rtf1 mRNA levels resulted in defects in the elongation process (Costa & Arndt, 2000)

Earlier studies suggested that the human PAF complex lacks the Rtf1 counterpart (Zhu et al, 2005a). However, only recently, Rtf1 was confirmed as part of this PAF complex (Kim et al, 2010). Moreover, it has been shown that Rtf1 is required for PAF complex association with Pol II as loss of Rtf1 leads to dissociation of the remaining subunits of the PAF complex (Paf1, Ctr9, Leo1, Cdc73, and Ski8) from the transcription machinery (Mueller et al, 2004).

Since a number of elongation factors and histone modifications are missing from the U2 snRNA gene (Chapter 3), an interesting question is whether one key factor differentially recruited to the β-actin gene acts as the main regulator and central player of this intricate system. PAF is a good candidate for such a role, given the many different interactions co-ordinated by this complex.

In order to determine whether PAF association with the β-actin gene underpins the recruitment and activation of any of the elongation factors differentially recruited to this gene, I have knocked down the level of PAF using siRNA against the Rtf1 subunit of the complex.

HeLa cells were transfected twice, 24 hours apart with Rtf1 siRNA and harvested 48 hours post the second transfection. In order to confirm that the knockdown was efficient, total protein was isolated and subjected to Western blot analysis, using an anti-Rtf1 antibody. The analysis detected a protein size of 77 kDa, as expected and confirmed that the level of Rtf1 was considerably reduced following knockdown (Figure 5.1 (A)).

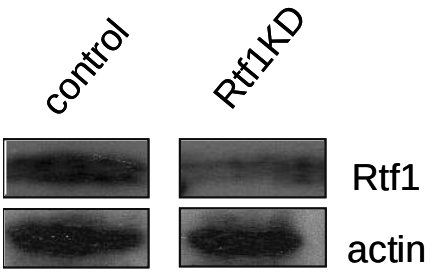
To fully assess the role of PAF, ChIP was performed on the Rtf1-knockdown cells, followed by quantitative real-time PCR analysis. On the β-actin gene, Rtf1 knockdown causes a decrease in the levels of pol II and Spt5 over primer set 1, where these factors are at their highest, and an increase over the rest of the gene. Previous studies conducted on *Saccharomyces cerevisiae* have shown that PAF, through its Rtf1 subunit, interacts directly with Spt5 (Costa & Arndt, 2000; Squazzo et al, 2002), which would explain its loss when Rtf1 is knocked down. However, on the U2 snRNA gene, the levels of both pol II and Spt5 do not change appreciably. This is in agreement with the analysis presented in Chapter 4, Figure 4.2, which indicates that PAF is not recruited to the U2 snRNA gene and therefore, it would be expected that

reduction in Rtf1 levels would have little effect on the association of pol II and Spt5 on this gene. Taken together, this confirms that PAF is dispensable for the transcription of the U2 snRNA gene.

The reduction of pol II and Spt5 is also associated with an increase in H2B levels across the β-actin gene when Rtf1 levels are reduced (Figure 5.1 (C)), which possibly reflects increased nucleosome re-assembly as a result of the inhibition of transcription. Changes in the level of H2B on primer sets 4, 5 and 6 of the U2 snRNA gene are more difficult to explain. Unfortunately, it was not possible to assess the effect of Rtf1 knockdown on H2B monoubiquitination as ChIP with the anti-H2b monoubiquitination antibody after Rtf1 knockdown did not give results above background with the control sample.

In conclusion, these results emphasize that the U2 snRNA elongation factor complement is different from that existing on the β-actin gene and suggest that the PAF complex plays a major role in the differential elongation properties of the β-actin and U2 snRNA genes.

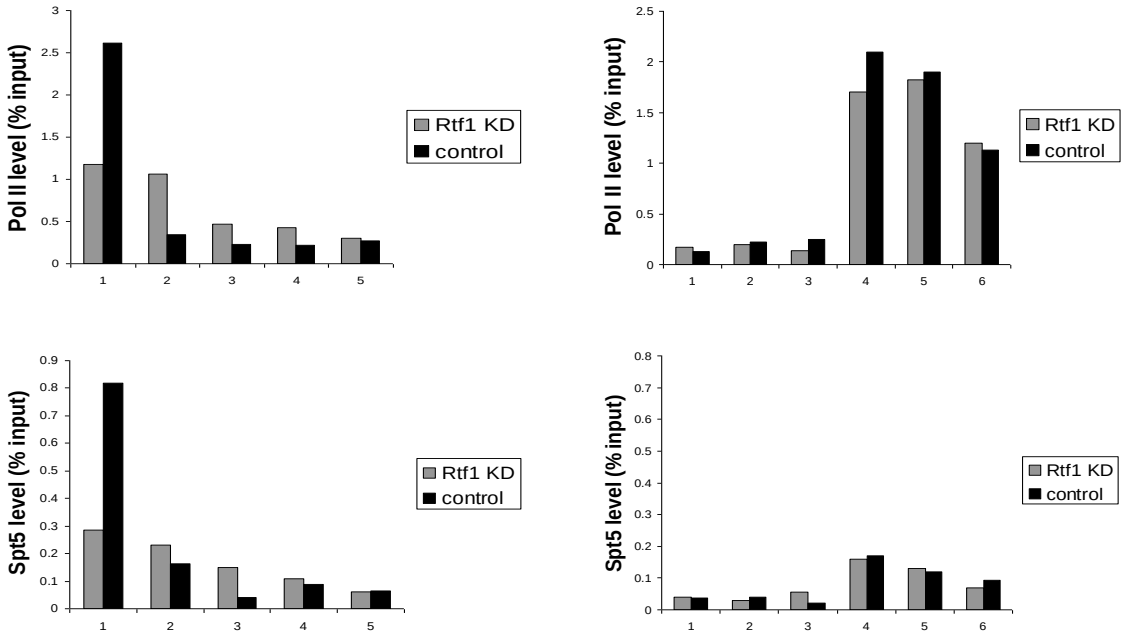
A


control
Rtf1KD
Rtf1
actin

B


-500
+1
1000
2000
3000
1
2
3
4
5
6
-3000
-2000
-1000
-220
-55
+1
+215
+800
DSE
PSE
encoding region
3' box
1
2
3
4
5
6

C


Pol II level (% input)
Rtf1 KD
control
Pol II level (% input)
Rtf1 KD
control
Spt5 level (% input)
Rtf1 KD
control
Spt5 level (% input)
Rtf1 KD
control

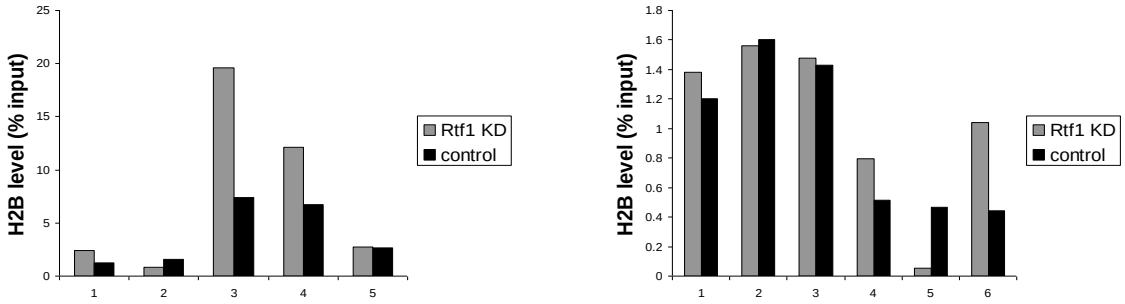


Figure (5.1): Loss of PAF alters levels of Pol II, Spt5, and H2B on the β-actin gene.

A) Western blot analysis of Rtf1 knockdown (top right), and control (top left) from HeLa cell extract. Actin was used as a loading control.

B) schematic diagram of the β-actin gene and the RNU2 locus.

C) RT-PCR of ChIP analysis of the RNU2 locus and β-actin genes, using antibodies against Pol II (top graph), Spt5 (middle graph) and H2B (bottom graph) after and before Rtf1 knockdown. The y axes represent the percentages of input DNA immunoprecipitated.

6 Chapter Six: Discussion

6.1: TFIIS levels correlate with pol II levels on both the β-actin and U2 snRNA genes.

TFIIS correlates largely with pol II on both the β-actin and U2 snRNA genes (Chapter 3, section 3.1.3). On both genes, the levels of both RNA pol II and TFIIS peak at the promoter regions. This result may be expected considering the role of TFIIS as a general ancillary transcription factor for pol II (Wind & Reines, 2000). Thus, TFIIS is likely to be an indispensible factor for transcription by pol II. However, the ratio of TFIIS to RNA pol II is higher over the coding region than at the promoter on both genes. This result suggests that TFIIS associates better with the actively transcribing pol II.

6.2: The histone chaperone, FACT is recruited to both gene types but shows differential association.

FACT is recruited to both genes but has a different pattern of association on the different gene types. On the U2 snRNA genes, FACT follows the pol II profile, whereas its association with the β-actin gene is detected only after about 800 bp (Chapter 3, Figure 3.2).

As mentioned previously, recent findings in the Murphy lab (Egloff et al, 2009a) have shown that a NELF-dependent checkpoint is located approximately 800 bp from the beginning of the β-actin gene transcription unit and that P-TEFb recruitment and phosphorylation of Ser2 of the pol II CTD occur at this point. Interestingly, NELF is involved in termination of transcription of the U2 snRNA gene and P-TEFb recruitment and CTD Ser2 phosphorylation occurs early in transcription of the U2 snRNA gene, where FACT is first detected (Egloff et al, 2009a). These results are consistent with a role for CTD Ser2 phosphorylation in FACT recruitment (see below). Interestingly, treatment of cells with CDK9 inhibitors results in loss of FACT recruitment (Gomes et al, 2006) but has no effect on transcription of the U2 snRNA genes (Medlin et al., 2005). Thus, FACT does not appear to be an essential transcription factor for snRNA genes and may simply be recruited through CTD Ser2 phosphorylation.

6.3: DSIF and phospho-DSIF are present at different levels on the β-actin and U2 snRNA genes.

The Spt5 subunit of DSIF and its phosphorylated form are detected on both the β-actin and U2 snRNA genes (Chapter 4, Section 4.1.3). However, there are considerably higher levels of both factors on the β-actin gene, particularly when DSIF and phospho DSIF levels are divided by the level of pol II (Chapter 4, Section 4.1.3). This difference in the recruitment of Spt5/phospho-Spt5, between the β-actin and U2 snRNA genes, may contribute to the differential requirement for P-TEFb for transcription. The molecular mechanism of differential Spt5 recruitment is however, unclear.

6.4: The elongation factors Spt6 and PAF and the histone modification enzyme Set1 are differentially associated with the β-actin and U2 snRNA genes.

The data presented in Chapter 4, demonstrate that Spt6, PAF and Set1 are all exclusively loaded on the β-actin gene, which taken together, indicates that the elongation complement on the β-actin gene differs significantly from that on U2 snRNA genes, suggesting that these factors play a gene-type-specific role in elongation of transcription.

Given the absence of nucleosomes from the U2 snRNA gene transcription unit, it is perhaps not surprising that Spt6 and PAF are absent. Both factors have been implicated in transcription of chromatin templates in a number of studies. On the β-actin template instead, pol II encounters nucleosomes downstream from the NELF-dependent elongation checkpoint and both the histone chaperone Spt6 and the elongation factor PAF would be critical for transcription for the rest of the gene to make a functional transcript.

As Spt5 has been shown to interact with PAF (Chen et al, 2009) and in yeast Spt5 phosphorylation have been shown to facilitate PAF recruitment (Liu et al, 2009), the absence of PAF may be directly related to the low level of Spt5 recruited to the U2 snRNA gene.

Set1 is also absent from the U2 snRNA gene, although H3K4me3, a mark generally placed by Set1 (Noma & Grewal, 2002) is present. Other methyltransferases of the same family are likely involved in placing this mark and highlights a further difference in control of expression of protein-coding and snRNA genes.

6.5: H2BK120ub1 is found only on the β-actin gene.

The histone modification H2BK120ub1 is generally found within the body of protein-coding genes, from just downstream of the transcription start site (Pirngruber et al, 2009). Accordingly, this mark is readily detected on the β-actin gene (Chapter4, section 4.2.3), from around the TSS, peaking at approximately +350 upstream of the position of the checkpoint. In contrast, this mark is undetectable on the U2 snRNA gene. As PAF has been shown to be required for recruitment/activation of the ubiquitin ligase Bre1 (Kim & Roeder, 2009), the lack of the H2BK120ub1 mark on the U2 snRNA gene is not unexpected. In addition, this result functionally reinforces the finding that PAF is undetectable by ChIP on the U2 snRNA genes.

Surprisingly, RAD6A and BRE1, which are both involved in H2B monoubiquitination, are found associated with the U2 snRNA genes. However, RAD6B is present at much higher levels on the β-actin genes than on the U2 snRNA genes. These results suggest that RAD6A and BRE1 are not sufficient to give H2B monoubiquitination and that Paf and/or RAD6B are also required.

H3K4 trimethylation is a histone mark that signifies active transcription (Santos-Rosa et al, 2002), and this mark relies at least in part, on H2B monoubiquitnation which helps attracts the Set1 methyltransferase (Weake & Workman, 2008). PAF, and consequently H2B monoubiquitination and Set1 are missing from the U2 snRNA gene (Chapter 4). It is therefore interesting and unexpected that H3K4 is detected at the promoter of the U2 snRNA gene. Taken together, these findings prompt the question of whether H3K4 trimethylation observed on the U2 snRNA gene is a result of a different methyltransferase.

In yeast, Set1 is recruited to the transcription machinery through the PAF complex and catalyses H3K4 trimethylation (Ng et al, 2003). Additionally Set1 is known to act as the sole H3K4 methyltransferase in yeast (Kim et al, 2005). However, in humans, a number of methyltransferases are candidates for H3K4 trimethylation, including Set1, Set7/9, MLL-1, Mll-2, MLL-3, and ALR (Sims et al, 2003).

My analysis of the association of RAD6A and B and BRE1 indicated that on the U2 snRNA gene, RAD6 and BRE1 are recruited to these genes but are apparently not activated, and therefore fail to carry out monoubiquitination of H2B.

Earlier reports (Weake & Workman, 2008) suggest that H2B monoubiquitination may serve as a key player in elongation checkpoint formation. In yeast, H2B deubiquitination, which occurs after H2B monoubiquitination, is carried out by Ubp8 within the SAGA complex, which results in the recruitment of Ctk1 kinase (a homologue of the human CDK9), which leads to Ser2 phosphorylation of the Pol II CTD. As a consequence of this phosphorylation, Set2 is recruited to the elongating polymerase and results in HK36me3, a mark signifying productive elongation (Weake & Workman, 2008). These findings have prompted the suggestion that H2B ubiquitylation plays a key role in checkpoint formation. Lack of this mark on the U2 snRNA gene may therefore be connected to the failure of pol II to make the transition to productive elongation and the failure of P-TEFb to impact on elongation.

6.6: Functions of the early elongation checkpoint.

It has been proposed that a potential function of the early elongation checkpoint is to help ensure efficient 5′ capping. Capping constitutes an essential step in the

processing and maturation of mRNA and is known to occur co-transcriptionally (Bentley, 2005; Cho et al, 1997). Capping can be defined as the addition of an m7GpppG cap by the sequential action of three enzymes, namely an RNA 5′ triphosphatase (RTP), a guanylyltransferase (GT), and a methyltransferase (MT) that together constitute the capping enzyme (CE). The early addition of the cap is required to prevent degradation of the nascent mRNA (Mandal et al, 2004). The 5′ cap is then bound by the cap binding complex (CBC), which contains two proteins, CBP20 and CBP80 (Proudfoot et al, 2002), and is necessary to regulate nuclear export of the mRNA. Once in the cytoplasm the CBC is replaced by the translation factor eIF-4E (Shatkin & Manley, 2000). Replacement of the CBC by the translation factors has a role in enhancing translation (Sachs et al, 1997). Association of the CE with the transcribing pol II is facilitated through interaction of the GT and MT with the Ser5 phosphorylated CTD of the pol II (Fabrega et al, 2003; McCracken et al, 1997a; Wen & Shatkin, 1999). The yeast MT Abd1 interacts with the elongation checkpoint factors (Schroeder et al, 2004). Recruitment of the human guanylyltransferase CE is enhanced by binding to the largest subunit of DSIF, Spt5 (Wen & Shatkin, 1999). It has also been shown that the CBC interacts with NELF (Narita et al, 2007). Additionally, recruitment of the CE disables NELF-mediated repression of transcription *in vitro* (Mandal et al, 2004). Thus, productive elongation would require the prior recruitment of the enzymes involved in capping, thereby helping to ensure that capping is efficient.

Another potential function of the early elongation checkpoint is to regulate transcription. For example, HIV transcription is blocked by an early NELF-dependent checkpoint, and Tat and P-TEFb are required for the release of the block and productive elongation (Chiu et al, 2002). In addition, Ser2 phosphorylation of pol II

CTD by P-TEFb occurs at the checkpoint block on the β-actin gene (Egloff et al, 2009a), where FACT is recruited, raising the possibility that FACT recruitment is dependent on P-TEFb activity. This is supported by the finding that FACT association is lost after P-TEFb inhibition (Gomes et al, 2006). In addition, the NELF-dependant checkpoint correlates with the end of a nucleosome-depleted region (Egloff et al, 2009b), implicating chromatin in the block. Thus, it is possible that pol II pauses at the checkpoint block to allow the recruitment of factors such as FACT to increase pol II processivity through chromatin. In contrast, Spt6 occurs at the promoter before the checkpoint block, where in yeast its recruitment requires the PAF complex (Adelman et al, 2006; Pavri et al, 2006), raising the possibility that failure to recruit Spt6 on the U2 genes may be due to the absence of PAF. Additionally, Spt6 and PAF may be required in addition to FACT to go through the checkpoint block. Thus, differential recruitment of factors required to transcribe through chromatin may underlie the failure of P-TEFb to promote transcription through the checkpoint block at the U2 gene.

6.7: Why does transcription elongation terminate with the first nucleosome encountered on the U2 snRNA genes?

On the U2 snRNA gene, it has been shown previously that transcription terminates when pol II encounters the nucleosome/NELF checkpoint (Egloff et al, 2009b). To further elucidate, it was previously suggested that when a nucleosome is encountered, negative elongation factor (NELF) recruitment blocks transcriptional elongation, P-TEFb releases the block by phosphorylation of NELF, DSIF, and Ser2 of the CTD and productive elongation is resumed (Egloff et al, 2009b). Results shown here indicate that on the β-actin gene, TFIIS, Spt6 and DSIF are recruited to the pol II at/near the promoter and associate with it along the whole transcription unit. FACT is also recruited at the same position as NELF, where Ser2 of the pol II CTD is phosphorylated. Additionally, PAF, Set1 and the histone modification H2BK120ub1 are found on the promoter proximal region of the β-actin gene whereas H3K36me2 is detected at the block.

On the U2 snRNA genes, P-TEFb is recruited at the promoter where it phosphorylates Ser2 of the CTD (Egloff et al, 2009b). The timing of P-TEFb recruitment and the resulting Ser2 phosphorylation is different between U2 snRNA and β-actin genes. On the U2 snRNA genes, pol II will already harbour phosphorylation on its Ser2 when the nucleosome/NELF roadblock is encountered, thus possibly triggering termination rather than traversing the block. Additionally, TFIIS, DSIF and FACT are also recruited at the promoter and associate with pol II along the nucleosome-depleted transcription unit of the U2 gene (Egloff et al, 2009b). However, DSIF, which is indispensable for the recruitment of a wide range of elongation factors and histone modification enzymes, is recruited to the promoter of the U2 gene in much lower

levels than found on the β-actin gene. Therefore lack of DSIF could be the key to the differential recruitment of elongation factors. Also, it is possible that a combination of ‘’early’’ phosphorylation of Ser2 on the pol II CTD along with the lack of sufficient amounts of DSIF affects interactions between the pol II CTD and elongation/termination factors at the nucleosomal checkpoint, resulting in the failure of pol II to transcribe through nucleosomes.

Thus, this comparative study of elongation factors and histone modification enzymes recruited to the structurally and functionally different β-actin and U2 snRNA genes highlights key differences in elongation complex composition, and chromatin structure, which could readily explain differences in the requirement for P-TEFb for gene expression. This approach allows the formulation of a testable model where key elongation factors are responsible for the differential behaviour of pol II when it meets the NELF/nucleosome block.

Figure 5.1 shows a model where a high level of phospho-Spt5 downstream of the checkpoint allows recruitment of PAF and in turn Spt6 to the β-actin gene, which allows transcription through chromatin. On the U2 snRNA gene instead, the level of Spt5 is too low to allow efficient recruitment of PAF and Spt6. This results in termination of transcription at the NELF/nucleosome block. In this model, the complex of elongation factors found on the U2 snRNA genes may represent the minimum set of factors needed to establish a default elongation. Since mRNA genes generally contain nucleosome downstream of the transcription start, chromatin modelling factors are recruited specifically, allowing processive, long-range transcription of the generally much longer transcription units of these protein-coding genes. Hence, the outcome of pol II is dictated by the set of elongation factors.

This model can be tested by knockdown of either PAF or Spt6, followed by analysis by, for example, nuclear run-on and pol II ChIP of transcription of the β-actin gene.

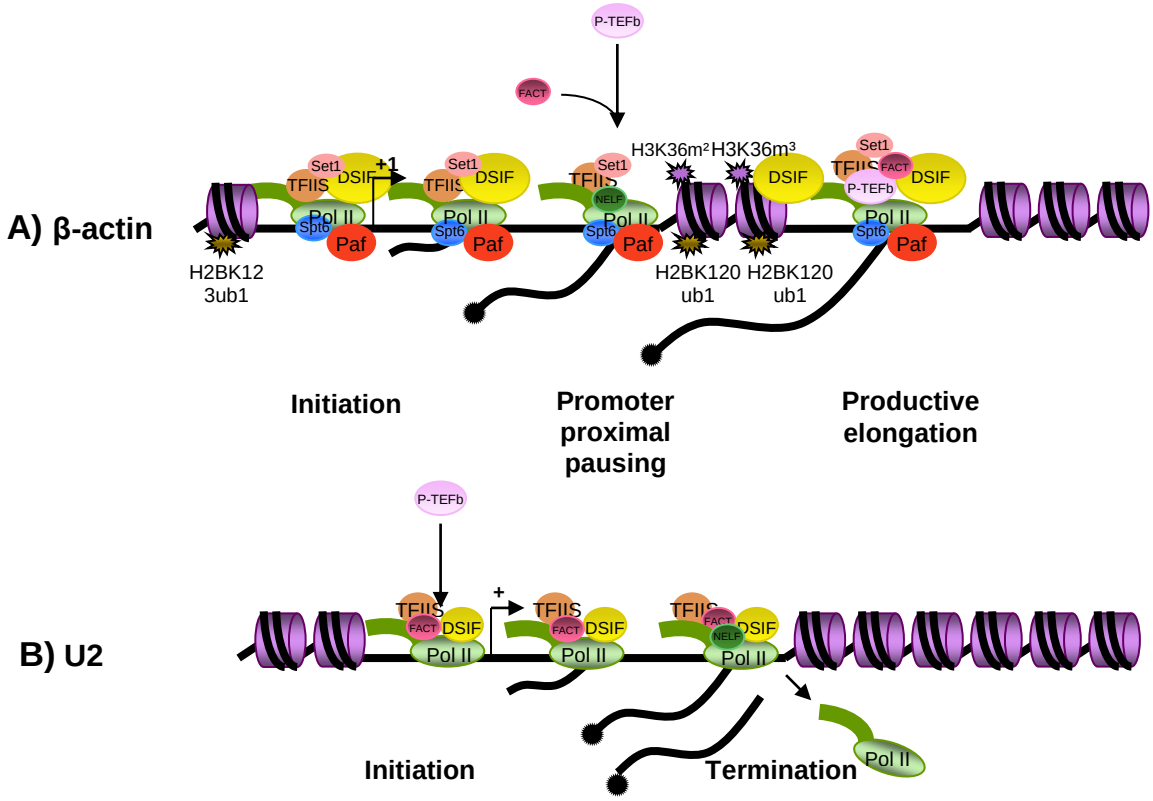


Figure (6.1): conclusion: differenential factor recruitment on the β-actin and the U2 genes.

A) Previous results suggest that when a nucleosome is encountered, negative elongation factor (NELF) recruitment blocks transcriptional elongation, P-TEFb releases the block by phosphorylation of NELF, DSIF, and Ser2 of the CTD and productive elongation is resumed. Results shown here indicate that on the β-actin gene, TFIIS, (SPT5) DSIF, Spt6, PAF and Set1 are recruited to the RNA pol II at the promoter and associates with it along the whole transcription unit. FACT is also recruited at the same position as NELF. Additionally, the histone modification H2BK120ub1 is found on the promoter proximal region of the β-actin gene whereas H3K36m² is detected at the block

B) On the U2 snRNA genes, P-TEFb is recruited at the promoter where it phosphorylates Ser2 of the CTD. TFIIS, DSIF, and FACT are also recruited at the promoter and associate with RNA pol II along the nucleosome-depleted transcription unit of the U2 gene. On the basis of these results we speculate that when RNA pol II encounters a nucleosome, NELF is recruited and transcription terminates in the absence of H3K36m², H2BK120ub1, Spt6, PAF and reduced levels of phospho Spt5.

7 Chapter Seven: Materials and Methods

7.1Materials

7.1.1: Antibodies

Antibodies were stored at 4 ^{o}C or at -20 ^{o}C

Antibodies	**Company**
- Anti-Flag M2 (mouse monoclonal)	Sigma
- Anti-Histone H3 (rabbit polyclonal)	Abcam
- Anti-RNA Polymerase II (N-20-rabbit polyclonal)	Santa Cruz
-Anti- Histone H2B (rabbit polyclonal)	Abcam
- Anti-Histone H2B ubiquitination (mouse monoclonal)	Medimabs
- Anti-Histone H3 (dimethyl K36-rabbit polyclonal)	Millipore
- Anti-Paf1 (rabbit polyclonal)	Abcam

- Anti-Spt16 (H-300-rabbit polyclonal)	Santa Cruz
-Anti-Spt5 (H-300-rabbit polyclonal)	Santa Cruz
- Anti-SUPT6H (rabbit polyclonal)	Bethyl
- Anti-Set1 (rabbit polyclonal)	Bethyl
-Goat anti –mouse HRP	Upstate
-Rabbit anti-mouse	Jackson Laboratories

7.1.2: Antibiotics

Antibiotics for cloning experiments were stored at -20 ^{o}C.

Stocks were made at 100× concentration.

Antibiotic	**Company**
- Ampicillin (100 μg/ml)	Sigma Aldrich

7.1.3: Buffers and solutions

- Phosphate Buffered Saline (PBS): 140 mM NaCl, 3.4 mM KCl, 10 mM Na_2HPO_4, 1.8 mMKH_2PO_4 – pH 7.4.
- Tris Buffered Saline: 150 mM NaCl, 40 mM Tris-HCl, pH 7.4.
- Proteinase K buffer (5×): 50 mM Tris-HCl, 25 mM EDTA, 1.25% SDS, pH 7.5.
- TBE Buffer: 10 mM Tris, 250 mM orthoborate, 2 mM EDTA, pH 8.3.
- Ligase Buffer (5×): 5 mM ATP, 5 mM DTT, 25 mM $MgCl_2$ and 330 mM Tris-HCl, pH 7.6.

7.1.4: Chemicals and other reagents

Chemical reagents were all of analytical grade, and were purchased from VWR International or Sigma, except for the reagents listed below:

Material	Company
- DNA molecular weight marker: 1 kb ladder	Invitrogen
- Bromophenol blue-containing DNA loading dye	Biolabs
- Ethidium Bromide (10 mg/ml stock)	BDH
- Trizol	Invitrogen
- Nuclease-free water	Qiagen
-Lipofectamine	Invitrogen

7.1.5: Chromatin Immunoprecipitation (ChIP) materials and buffers

Material	Company
- BSA(10 mg/ml)	New England Biolabs
- Herring Sperm DNA(10 mg/ml)	Sigma
- Formaldehyde	Sigma
- Protease inhibitor cocktail- one pellet per ml PBS	Roche
- Qiaquick PCR purification kit	Qiagen
- Staphlococcus aureus cell membranes (Staph A)	Calbiochem

All the buffers used in ChIP assays were stored at RT:

- Cell lysis buffer: 5 mM PIPES pH8.0, 85 mM KCl, 0.5% NP40.

- Nuclear lysis buffer: 50 mM Tris-HCl pH 8.1, 10 mM EDTA, 1% SDS.

- Dialysis Buffer: 2 mM EDTA, 50 mM Tris-HCl pH 8.0, 0.2 % Sarkosyl.

- Dilution buffer: 0.01% SDS, 1.1% Triton x-100, 1.2 mM EDTA, 16.7 mM Tris-HCl pH 8.1, 167 mM NaCl.

- Wash buffer 1 (Rabbit polyclonal antibodies): 100 mM Tris-HCl pH 9.0, 500 mM LiCl, 1% NP40, 1% Deoxycholic acid.

- Wash buffer 2 (Rat or mouse monoclonal antibodies): 100 mM Tris-HCl pH 8.0, 250 mM LiCl, 0.5% NP40, 0.5% Deoxycholic acid.

- Elution buffer: 50 mM $NHCO_3$, 1% SDS.

- Proteinase K Buffer (5×): 50 mM Tris-HCl pH 7.5, 25 mM EDTA, 1.2% SDS.

- EB buffer: 10 mM Tris-HCl, pH8.1.

7.1.6: Kits and enzymes

Material	**Company**
- Plasmid mini and maxi kits	Qiagen
- Qiaquick PCR purification kit	Qiagen
- QIAEX II gel extraction kit	Qiagen
- QuantiTect SYBR Green PCR master mix	Qiagen
- Superscript III for cDNA synthesis	Invitrogen
- HotStar Master Mix for PCR	Qiagen
- Pfu Ultra	Stratagene
- Proteinase K	Promega

- T4 DNA Ligase	Roche
- Restriction endonucleases	Roche

7.1.7: PCR Oligonucleotides

All oligonucleotides used in PCR assay were synthesized by Eurofins MWG. A standard annealing temperature of 57 ^{o}C was used in all PCR reactions and a single amplicon, of the appropriate size, was determined by agarose gel electrophoresis or as a single melt peak when analysed by real-time PCR. ChIP primers were designed to give a PCR product of approximately 100 bp in length.

Sequences of primers used in all assays are listed 5′ to 3′ in the following page.

Primers used in this study

Table (7.1): Primers

β-actin 1 (R)	5'CACTTAGAAGTCGCAGGACC
β-actin 1 (F)	5'GCTGCGGCTGGGTAGGTTTG
β-actin 2 (R)	5'GGTGTGGACGGGCGGCGGATC
β-actin 2 (F)	5'CCAATCAGCGTGCGCCGTTCCGA
β-actin 3(R)	5'ACGCAGTTAGCGCCCAAAGG
β-actin 3 (F)	5'GGGCAACCGGCGGGGTCTTT
β-actin 4 (R)	5'AGCCAGCTCCCCTACCTGGT
β-actin 4 (F)	5'CACAGCGCGCCCGGCTATTC
β-actin 5 (R)	5'CACCTGGGTCATCTTCTCGCGGTT
β-actin 5 (F)	5'CCCCATCGAGCACGGCATCGTC
β-actin 6 (R)	5'TTGCGGTGGACGATGGAGGGGCC
β-actin 6 (F)	5'CATTGCTCCTCCTGAGCGCAAGTA
U2-1 (R)	5'CAGGGTTCCTGCACATTCACAGTTA
U2-1 (F)	5'CAGCTGTGGCTGGACAGGTTGGACC
U2-2 (R)	5'AATCAGGATATTCACGTGGACTACT
U2-2 (F)	5'GAACTGACCACATTTCCAGGGCTC
U2-3 (R)	5'AGAGTGTGAGCCCTCATTCACGCCC
U2-3 (F)	5'GGAGCGGAGCGTTCTCTGTCTCCC
U2-4 (R)	5'CACTTGATCTTAGCCAAAAGG
U2-4 (F)	5'ATGAGAGTGGGACGGTGAC
U2-5 (R)	5'CTCCGGGTGGGTCCCATTCCTTTAA
U2-5 (F)	5'ACGAGTCCTGTGACGCGCCGGCTTG
U2-6 (R)	5'GGACAAATAGCCAACGCATGCGG
U2-6 (F)	5'CCTCCCCGCCTCTCCCTCGCTC
TFIIS (R)	5'ATTATCTAGATCAACAGAACTTCCATCGAT
TFIIS (F)	5'ATATGCGGCCGCGGAGGACGAAGTGGTCCGCTT
SSRP1 (R)	5'ATTATCTAGACTACTCATCGGATCCTGACG
SSRP1 (F)	5'ATATGCGGCCGCGGCAGAGACACTGGAGTTCAA

7.1.8: Bacteria and cell line

- Escherichia coli NM554 competent cells were used to amplify plasmid DNA.

- HeLa adherent cells were used for ChIP, transient transfection and knockdown experiments.

7.1.9: Bacterial and tissue culture media

- NM554 cells were cultured in 2-YT media at 37 ^{o}C, shaking overnight.

- HeLa cells were grown in ‘complete’ media consisting of Dulbecco’s Modified Eagle’s Media (DMEM), supplemented with 2 mM L-glutamine (Invitrogen) and 10% heat inactivated foetal bovine serum (Invitrogen).

7.1.10: Cloning vector

- pFLAG-CMV-2 (Sigma-Aldrich). This vector was used for TFIIS and SSRP1 (FACT) cloning.

7.2: Methods

7.2.1: DNA related methods

7.2.1.1: Polymerase Chain Reaction (PCR)

PCR was used for cloning experiments and analysis of reverse transcribed total RNA products.

For all PCR assays a Taq-containing master mix (2×) was used (HotStar). Typical PCR reactions contained either 100 ng of genomic DNA, 10 ng of plasmid DNA, or $1/50^{th}$ of a cDNA reaction, as template DNA, plus 1 μl of a 10 pmoles/μl primer pair stock in a total volume of 15 μl.

For all cloning experiments 2 U of a proofreading enzyme, typically pfu Ultra was added to each reaction. Thermal cycling was carried out as follows - 95 ^{o}C / 15 min, 25-30 cycles at 96 ^{o}C / 1 min, 56 ^{o}C / 1 min and 72 ^{o}C / 1 min, followed by 72 ^{o}C / 7 min. The products were analysed by electrophoresis on 1-2% agarose gels. PCR products were purified from agarose gels using the QIAEX II gel extraction kit, according to manufacturer's instructions.

7.2.1.2: Real-time quantitative Polymerase Chain Reaction (qPCR)

qPCR was carried out on a Rotorgene 6000 thermal cycler (Corbett) using 2× master mix with SYBR green. Typical reaction volumes were 10μl, consisting of 5 μl of 2× master mix, 1 μl of a 10 pmoles/μl primer pair stock and usually $1/50^{th}$ of ChIP DNA. Thermal cycling consisted of: 95 ^{o}C for 15 min, followed by 40 cycles at 94 ^{o}C for 15

sec, 57 ºC for 15 sec and 72 ºC for 30 sec. All amplification were followed by a 10 min melt cycle to confirm that a single PCR product was generated per reaction. This involves ramping the temperature from 50 ºC to 99 ºC by 1 ºC every 5 sec. A standard curve was constructed for each primer pair using 1/5 step-wise dilutions of the 'Input DNA' sample, and all other samples were quantitated relative to this, giving a "percent input DNA" value for each sample. This value was adjusted for background signal by subtracting the "percent input DNA" given by the no antibody sample, giving the final value plotted on the graphs shown in this thesis.

At the end of the run a threshold value is set for all samples and standards over the linear range of the PCR reaction. The Rotorgene software converts the cycle number (Ct) for each standard to copies/µl and generates a standard curve. The Ct values for each sample are quantitated relative to this standard. To allow normalisation of β-actin and U2 snRNA transcript levels in HeLa, the concentration (in copies/µl) of the β-actin and snRNA-specific amplicon were divided by the total input DNA for each immunoprecipitation, giving the final value plotted on the graphs shown in this thesis.

7.2.1.2: Agarose gel electrophoresis

First the DNA samples were mixed with bromophenol blue-containing DNA loading dye. Seocndly they were loaded on 1 – 2% agarose gels with 0.002% Ethidium Bromide (10 mg/ml stock). Electrophoresis was carried out at 100 mA in 1× TBE buffer, also containing 0.002% Ethidium Bromide. DNA bands were visualized under UV light. The 100 bp ladder was used a sizing marker.

7.2.1.3: DNA sequencing

DNA sequencing was carried out to confirm that all clones generated were as expected. The sequencing reaction was carried out using mini-prep or maxi-prep (Qiagen) purified plasmid along with the appropriate primer for the clone in question (see Table 6.1). All sequencing reactions were sourced out to Source BioScience UK limited, Department of Biochemistry, University of Oxford, UK.

7.2.1.4: Restriction enzyme digestion

500 ng – 1 µg of plasmid DNA or PCR product was used in all enzyme digestion reactions with digestion buffer (1×) and 5 – 10 U of restriction enzyme(s). Digestions were incubated at 37 ^{o}C up to 1 h for plasmid DNA and up to 3 h for PCR products.

7.2.1.5: DNA ligation

DNA fragments with blunt or overlapping ends were incubated with 1 U of T4 DNA ligase in 1× ligation buffer for 5 min at RT. Typically, 10 ng of vector DNA to 50 ng of insert were incubated for in a total reaction volume of 10 µl.

7.2.1.6: Plasmid DNA purification

Qiagen miniprep and maxiprep kits were used to purify plasmid DNA from 200 µl to 100 ml transformed bacterial cultures, respectively.

Mini-prep

For miniprep analysis, typically 200 µl of 2YT media with appropriate antibiotic, was inoculated with a single colony and incubated for 6 h at 37 ºC with shaking. Miniprep DNA preparations were generally used in sequencing reactions to confirm the expected sequence of clones generated in this thesis.

Maxi-prep

For maxi-prep analysis, a 100 ml of 2YT media was inoculated with 100 µl of a 5 ml overnight culture. This DNA was used in downstream application, including ChIP and transfection assays.

7.2.2: *RNA related methods*

7.2.2.1: Purification of total RNA and cDNA purification

To clone TFIIS and SSRP1 (FACT), it was necessary to first purify the total RNA from cells and generate cDNA templates. Total RNA was isolated from HeLa cells or HUES-2 using Trizol reagent according to manufacturer's instructions. Briefly 1 ml of Trizol reagent was added directly to cells, adhered to a 9.6 cm^2 well, lysed thoroughly and incubated for 5 min at RT. The 1 ml lysates were transferred to eppendorfs and 0.2 ml of chloroform was added to each sample. Following a vigorous shake for 15 sec at RT samples were centrifuged at 4 ^oC, 11,290 RCF for 15 min. The aqueous phase was carefully removed and the total RNA precipitated in an equal volume of isopropanol for 10 min at RT. Following centrifugation at 4 ^oC, 11,290 RCF for 10 min, the supernatant was removed and the pellet washed in 75% Ethanol, re-centrifuged at 4410 RCF for 5 min at 4 ^oC and air-dried at RT for 5 min. The pellet was dissolved in 20 µl of RNase-DNase-free water. Following this, the RNA sample was DNase treated using the DNase I as per manufacturer's instructions. RNA was quantitated using a Nanodrop spectrophotometer (Thermo scientific). cDNA was synthesized using 500 ng of total RNA, 200 ng random hexamers (Roche) and superscript III (Invitrogen) according to manufacturers' instructions.

7.2.2.2: Molecular cloning of TFIIS and SSRP1 (FACT)

To clone TFIIS and SSRP1 (FACT), cDNA synthesized from HeLa cell total RNA was used as a template for the amplification of both TFIIS and SSRP1. Primers used

for the amplification are listed in Table 7.1. PCR using Taq polymerase (Go Taq) and Pfu was carried out on the cDNA. The amplified template fragments (TFIIS and SSRP1) were cloned (by ligation) into NotI and XbaI sites within the multiple cloning region of pFLAG-CMV-2 vector (Sigma-Aldrich). Ligated products were transformed into *E.coli* strain NM554. Sequencing of single clones were determined and analysed.

7.2.2.3: siRNA mediated knockdown

To fully assess the role of the PAF complex, small interfering RNA (siRNA) mediated knockdown was approached using siRNA against the Rtf1 subunit. siRNA oligonucleotides were purchased from Santa Cruz Biotechnology. Cells were transfected when 40%-60% confluent using Lipofectamine 2000 (Invitrogen) as the transfection reagent and according to the manufacturer's instructions. As a control, Lipofectamine 2000 alone and without siRNA was used. Cells were subjected to double transfection points, each separated by 24 hours. Finally, cells were collected after 48 hours from the first transfection.

7.2.3: *Protein related methods*

7.2.3.1: Cell Culture

Cell culture was performed in a category 2 laminar flow hood, using sterile technique. HeLa cells were incubated at 37 ^{0}C in the presence of 5% CO_2. Typically, adherent HeLa cells, in 75 cm^2 flasks, were lifted off by incubation with 1 ml of PBS plus 1 mM EDTA for 5 min at RT. This was followed by gentle tapping of the side of the flask. Cells were resuspended in 10 ml of complete DMEM media, by pipetting several times, before passage into new flasks at 1:10 dilution.

Freezing

Cells stored at -70 ^{o}C were resuspended in 1 ml of complete DMEM media, supplemented with 10% foetal bovine serum (20% final) and 10% DMSO, per 75 cm^2 flask. Cells were put into a cryotube. In order to defreeze, cells were removed from storage and grown in DMEM complete media supplemented with 10% foetal bovine serum (20% final) for the first passage. For all subsequent passages cells were grown in complete DMEM media.

7.2.3.2: Transfection experiments in HeLa cells

HeLa cells were plated in complete DMEM media in 6-well Costar dishes at a density of 2×10^5 cells per well. Each DNA sample to be transfected was diluted in 50 μl of Opti-MEM 1 (Invitrogen) reduced serum medium and incubated for 5 min at RT. Lipofectamine 2000 was also diluted in Opti-MEM 1 reduced serum medium and

incubated for 5 min at RT - typically 2 µl of LipofectAMINE 2000 was used per µg of DNA in 50 µl of Opti-MEM 1 reduced serum medium. The DNA and Lipofectamine 2000 were combined and incubated for a further 30 min at RT. 100 µl of complexes was added drop wise to the adherent HeLa cells. Media was changed after 6 h and cells were harvested within 24 to 48 h.

7.2.3.3: ChIP analysis

Chromatin immunoprecipitation (ChIP) analysis was carried out to measure the density of RNA pol II, TFIIS, DSIF, FACT, Spt6, PAF, H3, H2B, H3K36me2, H2BK120ub1, and Set1 elongation factors and Histone modifications associated with the gene of interest. It was performed as described by Farnhamn lab (O'Geen et al. 2000). On average, 1 x 10^6 HeLa cells were used per immunoprecipitation. HeLa cells were transfected with TFIIS and SSRP1, and ChIP analysis was carried out 48 hr after transfection. For ChIP analysis of the endogenous proteins (pol II, Spt16, Spt6, DSIF, PAF, H3, H2B, H3K36me2, H2BK120ub1, Set1), 80-90% confluent HeLa cells were carried out throughout the analysis. To crosslink the proteins, formaldehyde was added to HeLa cell cultures at a final concentration of 1% for 10 minutes at room temperature (RT). To stop the crosslinking reaction, glycine was added at a final concentration of 0.125 M for 5 minutes at RT. To harvest HeLa cells, plates were rinsed twice with cold PBS then scraped. Cells were collected by centrifugation and then washed with cold PBS of pH 7.4 plus 100 mM phenylmethylsulfonyl fluoride (PMSF)**.** Cells were lysed using cell lysis buffer (5 mM PIPES, 85 mM KCL, 0.5 % NP40) plus protease inhibitors and samples were incubated on ice for 20 minutes. Nuclei were collected by centrifugation at 2300 × g for 5 minutes at 4°C. After that,

nuclei lysis buffer (50 mM Tris-HCl pH 8.1, 10 mM EDTA, 1% SDS) plus protease inhibitors (Roche) was added and incubated on ice for 10 minutes. This was then proceeded by sonication on ice to an average length of about 300 bp by using Bioruptor sonicator (Diagenode) for 10 min (10-s pulse), and samples were then centrifuged at 18,300 × g for 10 minutes at 4°C. The chromatin was precleared by adding blocked staph A cells and incubated on a rotating platform for 15 minutes at 4°C. 100 µl of Staph A were blocked by adding 10 µl of 10 mg/ml herring sperm DNA and 10 µl of 10 mg/ml BSA and were incubated for 4 hours at 4°C. Precleared chromatin was then incubated with Anti-pol II (N-20) (Santa Cruz Biotechnology), 1 µg Anti-Flag antibody (Sigma), 1 µg Anti-Spt16 (Santa Cruz Biotechnology), 3.5 µg Anti-Spt6 (Bethyl), 1 µg Anti-DSIF (Santa Cruz), 3 µg Anti-PAF (Abcam), 1 µg Anti-H3 (Abcam), 1 µg Anti-H2B (Abcam), 2 µg Anti-H3K36me2 (Millipore), 3 µg Anti-H2BK120ub1 (Medimabs), 5 µg Anti-Set1 (Bethyl) and rotated for 12 hours at 4°C (a no antibody sample was included as negative control). The next day, blocked staph A cells were added and incubated on a rotating platform for 15 minutes at RT, then samples were centrifuged and supernatant from the sample lacking the antibody was used as a total input chromatin. Pellets were washed twice with 1.4 ml of 1X dialysis buffer (2 mM EDTA, 50 mM Tris-HCl pH 8.0) plus 0.2 % sarkosyl for polyclonal antibodies, and then washed four times with 1.4 ml IP wash buffer (100 mM Tris-HCl with either pH 9.0 for polyclonal antibodies or pH 8.0 for monoclonal antibodies, 500 mM LiCl, 1% NP40, 1 % deoxycholic acid). To elute antibody/protein/DNA complexes, two rounds of 150 µl of IP elution buffer (50 nM NaHCO3, 1% SDS) was added and vortexed at each round for at least 15 minutes then centrifuged at 18,300 × g for 3 minutes at 4°C to collect the supernatant. Crosslinks were reversed by addition of NaCl to a final concentration of 0.3 M and 10

ug of RNase A was also added to remove RNA, then this was followed by incubation at 67°C for 4 – 5 hrs. Samples were then precipitated at -20°C overnight by the addition of 2.5 volumes of ethanol, the following day, samples were centrifuged at 18,300 × g for 20 minutes at 4°C. After that, pellets were resuspended in 100 ul of TE (10mM Tris-HCl pH 7.5, 25 mM EDTA, 1.25% SDS), and 1.5 ul of protein K then incubated at 45°C waterbath for 2 hrs. DNA purification was carried out using QIAquick PCR purification kit (QIAGEN) according to the manufacturer's instructions. Quantitative Real time PCR (qRT-PCR) was performed using Qiagen QuantiTect SYBR Green PCR master mix with a Corbett Rotogene system. PCR reactions were typically 40 cycles. 1 μl of 10 pmol/μl of each primer pair was used in the PCR reaction (Table 1). % Input was calculated by (IP average – no antibody average)*100/Input average).

Adelman K, Wei W, Ardehali MB, Werner J, Zhu B, Reinberg D, Lis JT (2006) Drosophila Paf1 modulates chromatin structure at actively transcribed genes. *Mol Cell Biol* **26:** 250-260

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. *Mol Cell* **13:** 67-76

Aida M, Chen Y, Nakajima K, Yamaguchi Y, Wada T, Handa H (2006) Transcriptional pausing caused by NELF plays a dual role in regulating immediate-early expression of the junB gene. *Mol Cell Biol* **26:** 6094-6104

Andrulis ED, Guzman E, Doring P, Werner J, Lis JT (2000) High-resolution localization of Drosophila Spt5 and Spt6 at heat shock genes in vivo: roles in promoter proximal pausing and transcription elongation. *Genes Dev* **14:** 2635-2649

Aoyagi N, Wassarman DA (2000) Genes encoding Drosophila melanogaster RNA polymerase II general transcription factors: diversity in TFIIA and TFIID components contributes to gene-specific transcriptional regulation. *J Cell Biol.* **150:**F45-50.

Ares M, Jr., Chung JS, Giglio L, Weiner AM (1987) Distinct factors with Sp1 and NF-A specificities bind to adjacent functional elements of the human U2 snRNA gene enhancer. *Genes Dev* **1:** 808-817

Baillat D, Hakimi MA, Naar AM, Shilatifard A, Cooch N, Shiekhattar R (2005) Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* **123:** 265-276

Bailly V, Lamb J, Sung P, Prakash S, Prakash L (1994) Specific complex formation between yeast RAD6 and RAD18 proteins: a potential mechanism for targeting RAD6 ubiquitin-conjugating activity to DNA damage sites. *Genes Dev* **8:** 811-820

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

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

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

Belotserkovskaya R, Oh S, Bondarenko VA, Orphanides G, Studitsky VM, Reinberg D (2003) FACT facilitates transcription-dependent nucleosome alteration. *Science* **301:** 1090-1093

Bentley DL (2005) Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Curr Opin Cell Biol* **17:** 251-256

Berglund JA, Chua K, Abovich N, Reed R, Rosbash M (1997) The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. *Cell* **89:** 781-787

Berglund, J. A., Abovich, N. & Rosbash, M. (1998) A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes Dev*, **12**, 858-67.

Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ, 3rd, Gingeras TR, Schreiber SL, Lander ES (2005) Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120:** 169-181

Bernues J, Simmen KA, Lewis JD, Gunderson SI, Polycarpou-Schwarz M, Moncollin V, Egly JM, Mattaj IW (1993) Common and unique transcription factor requirements of human U1 and U6 snRNA genes. *EMBO J* **12:** 3573-3585

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. *Mol Cell Biol*, **24**, 8963-9.

Bondarenko VA, Steele LM, Ujvari A, Gaykalova DA, Kulaeva OI, Polikanov YS, Luse DS, Studitsky VM (2006) Nucleosomes can form a polar barrier to transcript elongation by RNA polymerase II. *Mol Cell* **24:** 469-479

Borggrefe T, Yue X (2011) Interactions between subunits of the Mediator complex with gene-specific transcription factors. *Semin Cell Dev Biol* **22:** 759-768

Bortvin A, Winston F (1996) Evidence that Spt6p controls chromatin structure by a direct interaction with histones. *Science* **272:** 1473-1476

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

Briggs SD, Bryk M, Strahl BD, Cheung WL, Davie JK, Dent SY, Winston F, Allis CD (2001) Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in Saccharomyces cerevisiae. *Genes Dev* **15:** 3286-3295

Brou C, Chaudhary S, Davidson I, Lutz Y, Wu J, Egly JM, Tora L, Chambon P (1993) Distinct TFIID complexes mediate the effect of different transcriptional activators. *EMBO J* **12:** 489-499

Brown SA, Imbalzano AN, Kingston RE (1996) Activator-dependent regulation of transcriptional pausing on nucleosomal templates. *Genes Dev* **10:** 1479-1490

Brown SA, Kingston RE (1997) Disruption of downstream chromatin directed by a transcriptional activator. *Genes Dev* **11:** 3116-3121

Buratowski, S. (2009) Progression through the RNA Polymerase II CTD Cycle. *Mol. Cell.* **36** 541-546

Carey M, Li B, Workman JL (2006) RSC exploits histone acetylation to abrogate the nucleosomal block to RNA polymerase II elongation. *Mol Cell* **24:** 481-487

Carpten JD, Robbins CM, Villablanca A, Forsberg L, Presciuttini S, Bailey-Wilson J, Simonds WF, Gillanders EM, Kennedy AM, Chen JD, Agarwal SK, Sood R, Jones MP, Moses TY, Haven C, Petillo D, Leotlela PD, Harding B, Cameron D, Pannett AA, Hoog A, Heath H, 3rd, James-Newton LA, Robinson B, Zarbo RJ, Cavaco BM, Wassif W, Perrier ND, Rosen IB, Kristoffersson U, Turnpenny PD, Farnebo LO, Besser GM, Jackson CE, Morreau H, Trent JM, Thakker RV, Marx SJ, Teh BT, Larsson C, Hobbs MR (2002) HRPT2, encoding parafibromin, is mutated in hyperparathyroidism-jaw tumor syndrome. *Nat Genet* **32:** 676-680

Chapman RD, Heidemann M, Albert TK, Mailhammer R, Flatley A, Meisterernst M, Kremmer E, Eick D (2007) Transcribing RNA polymerase II is phosphorylated at CTD residue serine-7. *Science* **318:** 1780-1782

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

Chen, J. L., Attardi, L. D., Verrijzer, C. P., Yokomori, K. & Tjian, R. (1994). Assembly of recombinant TFIID reveals differential coactivator requirements for distinct transcriptional activators. *Cell,* **79:** 93-105.

Chen Y, Yamaguchi Y, Tsugeno Y, Yamamoto J, Yamada T, Nakamura M, Hisatake K, Handa H (2009) DSIF, the Paf1 complex, and Tat-SF1 have nonredundant, cooperative roles in RNA polymerase II elongation. *Genes Dev* **23:** 2765-2777

Cheng C, Sharp PA (2003) RNA polymerase II accumulation in the promoter-proximal region of the dihydrofolate reductase and gamma-actin genes. *Mol Cell Biol* **23:** 1961-1967

Chiu YL, Ho CK, Saha N, Schwer B, Shuman S, Rana TM (2002) Tat stimulates cotranscriptional capping of HIV mRNA. *Mol Cell* **10:** 585-597

Cho EJ, Takagi T, Moore CR, Buratowski S (1997) mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev* **11:** 3319-3326

Choder M (2004) Rpb4 and Rpb7: subunits of RNA polymerase II and beyond. *Trends Biochem Sci* **29:** 674-681

Chu Y, Simic R, Warner MH, Arndt KM, Prelich G (2007) Regulation of histone modification and cryptic transcription by the Bur1 and Paf1 complexes. *EMBO J* **26:** 4646-4656

Clark-Adams CD, Winston F (1987) The SPT6 gene is essential for growth and is required for delta-mediated transcription in Saccharomyces cerevisiae. *Mol Cell Biol* **7:** 679-686

Cook PR (2010) A model for all genomes: the role of transcription factories. *J Mol Biol* **395:** 1-10

Corden JL (1990) Tails of RNA polymerase II. *Trends Biochem Sci* **15:** 383-387

Corden JL, Patturajan M (1997) A CTD function linking transcription to splicing. *Trends Biochem Sci* **22:** 413-416

Costa PJ, Arndt KM (2000) Synthetic lethal interactions suggest a role for the Saccharomyces cerevisiae Rtf1 protein in transcription elongation. *Genetics* **156:** 535-547

Cramer P, Bushnell DA, Kornberg RD (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292:** 1863-1876

Cuello P, Boyd DC, Dye MJ, Proudfoot NJ, Murphy S (1999) Transcription of the human U2 snRNA genes continues beyond the 3' box in vivo. *Embo J* **18:** 2867-2877

Dahmus ME (1996) Reversible phosphorylation of the C-terminal domain of RNA polymerase II. *J Biol Chem* **271:** 19009-19012

Danilition SL, Frederickson RM, Taylor CY, Miyamoto NG (1991) Transcription factor binding and spacing constraints in the human beta-actin proximal promoter. *Nucleic Acids Res* **19:** 6913-6922

Dantonel JC, Murthy KG, Manley JL, Tora L (1997) Transcription factor TFIID recruits factor CPSF for formation of 3' end of mRNA. *Nature* **389:** 399-402

Das G, Hinkley CS, Herr W (1995) Basal promoter elements as a selective determinant of transcriptional activator function. *Nature* **374:** 657-660

Dikstein R, Zhou S, Tijan R (1996) Human TAFII 105 is a cell type-specific TFIID subunit related to hTAFII130. *Cell* **4:**137-46.

Dohmen RJ, Madura K, Bartel B, Varshavsky A (1991) The N-end rule is mediated by the UBC2(RAD6) ubiquitin-conjugating enzyme. *Proc Natl Acad Sci U S A* **88:** 7351-7355

Dover J, Schneider J, Tawiah-Boateng MA, Wood A, Dean K, Johnston M, Shilatifard A (2002) Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. *J Biol Chem* **277:** 28368-28371

Dynlacht BD, Hoey T, Tijan R (1991) Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell* **66:**563-576.

Eberhardy SR, Farnham PJ (2002) Myc recruits P-TEFb to mediate the final step in the transcriptional activation of the cad promoter. *J Biol Chem.* **277**(42):40156–40162.

Eggermont J, Proudfoot NJ (1993) Poly(A) signals and transcriptional pause sites combine to prevent interference between RNA polymerase II promoters. *EMBO J* **12:** 2539-2548

Egloff S, Al-Rawaf H, O'Reilly D, and Murphy S (2009a) Chromatin and CTCF are implicated in "late" NELF-dependent checkpoints on the U2 snRNA and β-actin genes (submitted for publication).

Egloff S, Al-Rawaf H, O'Reilly D, Murphy S (2009b) Chromatin structure is implicated in "late" elongation checkpoints on the U2 snRNA and beta-actin genes. *Mol Cell Biol* **29:** 4002-4013

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

Egloff S, Murphy S (2008b) Role of the C-terminal domain of RNA polymerase II in expression of small nuclear RNA genes. *Biochem Soc Trans* **36:** 537-539

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

Egloff S, O'Reilly D, Murphy S (2008) Expression of human snRNA genes from beginning to end. *Biochem Soc Trans* **36:** 590-594

Egloff S, Szczepaniak SA, Dienstbier M, Taylor A, Knight S, Murphy S (2010) The integrator complex recognizes a new double mark on the RNA polymerase II carboxyl-terminal domain. *J Biol Chem* **285:** 20564-20569

Eichner J, Chen HT, Warfield L, Hahn S (2010) Position of the general transcription factor TFIIF within the RNA polymerase II transcription preinitiation complex. *EMBO J* **29:** 706-716

Endoh M, Zhu W, Hasegawa J, Watanabe H, Kim DK, Aida M, Inukai N, Narita T, Yamada T, Furuya A, Sato H, Yamaguchi Y, Mandal SS, Reinberg D, Wada T, Handa H (2004) Human Spt6 stimulates transcription elongation by RNA polymerase II in vitro. *Mol Cell Biol* **24:** 3324-3336

Fabrega C, Shen V, Shuman S, Lima CD (2003) Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Mol Cell* **11:** 1549-1561

Feaver WJ, Henry NL, Wang Z, Wu X, Svejstrup JQ, Bushnell DA, Friedberg EC, Kornberg RD (1997) Genes for Tfb2, Tfb3, and Tfb4 subunits of yeast transcription/repair factor IIH. Homology to human cyclin-dependent kinase activating kinase and IIH subunits. *J Biol Chem* **272:** 19319-19327

Fu TJ, Peng J, Lee G, Price DH & Flores O (1999) Cyclin K functions as a CDK9 regulatory subunit and participates in RNA polymerase II transcription. *J Biol Chem*, **274:** 34527-30.

Fujinaga K, Irwin D, Huang Y, Taube R, Kurosu T, Peterlin BM (2004) Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. *Mol Cell Biol* **24:** 787-795

Gangloff YG, Romier C, Thuault S, Werten S, Davidson I (2001) The histone fold is a key structural motif of transcription factor TFIID. *Trends Biochem Sci* **26:**250-7.

Garber ME, Jones KA (1999) HIV-1 Tat: coping with negative elongation factors. *Curr Opin Immunol* **11:** 460-465

Gil A, Proudfoot NJ (1984) A sequence downstream of AAUAAA is required for rabbit beta-globin mRNA 3'-end formation. *Nature* **312:** 473-474

Glover-Cutter K, Larochelle S, Erickson B, Zhang C, Shokat K, Fisher RP, Bentley DL (2009) TFIIH-associated Cdk7 kinase functions in phosphorylation of C-terminal domain Ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II. *Mol Cell Biol* **29:** 5455-5464

Gomes NP, Bjerke G, Llorente B, Szostek SA, Emerson BM, Espinosa JM (2006) Gene-specific requirement for P-TEFb activity and RNA polymerase II phosphorylation within the p53 transcriptional program. *Genes Dev* **20:** 601-612

Gomes NP, Espinosa JM (2010) Gene-specific repression of the p53 target gene PUMA via intragenic CTCF-Cohesin binding. *Genes Dev* **24:** 1022-1034

Green MR (2000) TBP-associated factors (TAFIIs): multiple, selective transcriptional mediators in common complexes. *Trends Biochem Sci* **25:** 59-63

Gromak N, West S, Proudfoot NJ (2006) Pause sites promote transcriptional termination of mammalian RNA polymerase II. *Mol Cell Biol* **26:** 3986-3996

Guermah M, Palhan VB, Tackett AJ, Chait BT, Roeder RG (2006) Synergistic functions of SII and p300 in productive activator-dependent transcription of chromatin templates. *Cell* **125:** 275-286

Gunderson SI, Knuth MW & Burgess RR (1990) The human U1 snRNA promoter correctly initiates transcription in vitro and is activated by PSE1. Genes Dev, 4, 2048-60.

Hampsey M, Singh BN, Ansari A, Laine JP, Krishnamurthy S (2011) Control of eukaryotic gene expression: gene loops and transcriptional memory. *Adv Enzyme Regul* **51:** 118-125

Hartzog GA, Speer JL, Lindstrom DL (2002) Transcript elongation on a nucleoprotein template. *Biochim Biophys Acta* **1577:** 276-286

Hartzog GA, Wada T, Handa H, Winston F (1998) Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in Saccharomyces cerevisiae. *Genes Dev* **12:** 357-369

Hernandez N (1985) Formation of the 3' end of U1 snRNA is directed by a conserved sequence located downstream of the coding region. *EMBO J* **4:** 1827-1837

Hernandez N (2001) Small nuclear RNA genes: a model system to study fundamental mechanisms of transcription. *J Biol Chem* **276:** 26733-26736

Hershko A, Ciechanover A (1998) The ubiquitin system. *Annu Rev Biochem* **67:** 425-479

Hirose Y, Manley JL (2000) RNA polymerase II and the integration of nuclear events. *Genes Dev* **14:** 1415-1429

Hochstrasser M (1996) Ubiquitin-dependent protein degradation. *Annu Rev Genet* **30:** 405-439

Hoey T, Weinzierl RO, Gill G, Chen JL, Dynlacht BD, Tjian R (1993) Molecular cloning and functional analysis of Drosophila TAF110 reveal properties expected of coactivators. *Cell* **72:** 247-260

Hwang WW, Venkatasubrahmanyam S, Ianculescu AG, Tong A, Boone C, Madhani HD (2003) A conserved RING finger protein required for histone H2B monoubiquitination and cell size control. *Mol Cell* **11:** 261-266

Ivanov D, Kwak YT, Guo J, Gaynor RB (2000) Domains in the SPT5 protein that modulate its transcriptional regulatory properties. *Mol Cell Biol* **20:** 2970-2983

Ivanovska I, Jacques PE, Rando OJ, Robert F, Winston F (2011) Control of chromatin structure by spt6: different consequences in coding and regulatory regions. *Mol Cell Biol* **31:** 531-541

Jaehning JA (2010) The Paf1 complex: platform or player in RNA polymerase II transcription? *Biochim Biophys Acta* **1799:** 379-388

Jang MK, Mochizuki K, Zhou M, Jeong H-S, Brady JN, Ozato K (2005) The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol Cell* 19(4):523–534.

Johnson RE, Henderson ST, Petes TD, Prakash S, Bankmann M, Prakash L (1992) Saccharomyces cerevisiae RAD5-encoded DNA repair protein contains DNA helicase and zinc-binding sequence motifs and affects the stability of simple repetitive sequences in the genome. *Mol Cell Biol* **12:** 3807-3818

Kanin EI, Kipp RT, Kung C, Slattery M, Viale A, Hahn S, Shokat KM, Ansari AZ (2007) Chemical inhibition of the TFIIH-associated kinase Cdk7/Kin28 does not impair global mRNA synthesis. *Proc Natl Acad Sci U S A* **104:** 5812-5817

Kaplan CD, Laprade L, Winston F (2003) Transcription elongation factors repress transcription initiation from cryptic sites. *Science* **301:** 1096-1099

Kaplan CD, Morris JR, Wu C, Winston F (2000) Spt5 and spt6 are associated with active transcription and have characteristics of general elongation factors in D. melanogaster. *Genes Dev* **14:** 2623-2634

Kettenberger H, Armache KJ, Cramer P (2003) Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell* **114:** 347-357

Kim J, Guermah M, McGinty RK, Lee JS, Tang Z, Milne TA, Shilatifard A, Muir TW, Roeder RG (2009) RAD6-Mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. *Cell* **137:** 459-471

Kim J, Guermah M, Roeder RG (2010) The human PAF1 complex acts in chromatin transcription elongation both independently and cooperatively with SII/TFIIS. *Cell* **140:** 491-503

Kim J, Hake SB, Roeder RG (2005) The human homolog of yeast BRE1 functions as a transcriptional coactivator through direct activator interactions. *Mol Cell* **20:** 759-770

Kim J, Roeder RG (2009) Direct Bre1-Paf1 complex interactions and RING finger-independent Bre1-Rad6 interactions mediate histone H2B ubiquitylation in yeast. *J Biol Chem* **284:** 20582-20592

Kim J, Roeder RG (2011) Nucleosomal H2B ubiquitylation with purified factors. *Methods* **54:** 331-338

Kim JB, Sharp PA (2001) Positive transcription elongation factor B phosphorylates hSPT5 and RNA polymerase II carboxyl-terminal domain independently of cyclin-dependent kinase-activating kinase. *J Biol Chem* **276:** 12317-12323

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

Kireeva ML, Hancock B, Cremona GH, Walter W, Studitsky VM, Kashlev M (2005) Nature of the nucleosomal barrier to RNA polymerase II. *Mol Cell* **18:** 97-108

Kizer KO, Phatnani HP, Shibata Y, Hall H, Greenleaf AL, Strahl BD (2005) A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Mol Cell Biol* **25:** 3305-3316

Koken MH, Reynolds P, Jaspers-Dekker I, Prakash L, Prakash S, Bootsma D, Hoeijmakers JH (1991) Structural and functional conservation of two human

homologs of the yeast DNA repair gene RAD6. *Proc Natl Acad Sci U S A* **88:** 8865-8869

Komarnitsky P, Cho EJ, Buratowski S (2000) Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev* **14:** 2452-2460

Krogan NJ, Dover J, Khorrami S, Greenblatt JF, Schneider J, Johnston M, Shilatifard A (2002a) COMPASS, a histone H3 (Lysine 4) methyltransferase required for telomeric silencing of gene expression. *J Biol Chem* **277:** 10753-10755

Krogan NJ, Dover J, Wood A, Schneider J, Heidt J, Boateng MA, Dean K, Ryan OW, Golshani A, Johnston M, Greenblatt JF, Shilatifard A (2003a) The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Mol Cell* **11:** 721-729

Krogan NJ, Kim M, Ahn SH, Zhong G, Kobor MS, Cagney G, Emili A, Shilatifard A, Buratowski S, Greenblatt JF (2002b) RNA polymerase II elongation factors of Saccharomyces cerevisiae: a targeted proteomics approach. *Mol Cell Biol* **22:** 6979-6992

Krogan NJ, Kim M, Tong A, Golshani A, Cagney G, Canadien V, Richards DP, Beattie BK, Emili A, Boone C, Shilatifard A, Buratowski S, Greenblatt J (2003b) Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. *Mol Cell Biol* **23:** 4207-4218

Kuhlman TC, Cho H, Reinberg D, Hernandez N (1999) The general transcription factors IIA, IIB, IIF, and IIE are required for RNA polymerase II transcription from the human U1 small nuclear RNA promoter. *Mol Cell Biol* **19:** 2130-2141

Kuras L, Kosa P, Mencia M, Struhl K (2000) TAF-Containing and TAF-independent forms of transcriptionally active TBP in vivo. *Science* **288:** 1244-1248

Laribee RN, Krogan NJ, Xiao T, Shibata Y, Hughes TR, Greenblatt JF, Strahl BD (2005) BUR kinase selectively regulates H3 K4 trimethylation and H2B ubiquitylation through recruitment of the PAF elongation complex. *Curr Biol* **15:** 1487-1493

Lemon B, Tjian R (2000) Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev* **14:** 2551-2569

Li B, Carey M, Workman JL (2007a) The role of chromatin during transcription. *Cell* **128:** 707-719

Li B, Gogol M, Carey M, Pattenden SG, Seidel C, Workman JL (2007b) Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription. *Genes Dev* **21:** 1422-1430

Li B, Jackson J, Simon MD, Fleharty B, Gogol M, Seidel C, Workman JL, Shilatifard A (2009) Histone H3 lysine 36 dimethylation (H3K36me2) is sufficient to recruit the

Rpd3s histone deacetylase complex and to repress spurious transcription. *J Biol Chem* **284:** 7970-7976

Li J, Gilmour DS (2011) Promoter proximal pausing and the control of gene expression. *Curr Opin Genet Dev* **21:** 231-235

Liu H and Herrmann CH (2005). Differential localization and expression of the Cdk9 42k and 55k isoforms. *J Cell Physiol*, **203:** 251-60

Li J, Moazed D, Gygi SP (2002) Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation. *J Biol Chem* **277:** 49383-49388

Li M, Phatnani HP, Guan Z, Sage H, Greenleaf AL, Zhou P (2005) Solution structure of the Set2-Rpb1 interacting domain of human Set2 and its interaction with the hyperphosphorylated C-terminal domain of Rpb1. *Proc Natl Acad Sci U S A* **102:** 17636-17641

Lin YS, Green MR (1991) Mechanism of action of an acidic transcriptional activator in vitro. Cell **64:** 971-981

Lindgren V, Ares M, Jr., Weiner AM, Francke U (1985) Human genes for U2 small nuclear RNA map to a major adenovirus 12 modification site on chromosome 17. *Nature* **314:** 115-116

Lindstrom DL, Squazzo SL, Muster N, Burckin TA, Wachter KC, Emigh CA, McCleery JA, Yates JR, 3rd, Hartzog GA (2003) Dual roles for Spt5 in pre-mRNA processing and transcription elongation revealed by identification of Spt5-associated proteins. *Mol Cell Biol* **23:** 1368-1378

Liu WL, Coleman RA, Grob P, King DS, Florens L, Washburn MP, Geles KG, Yang JL, Ramey V, Nogales E, Tijan R (2008) Structural changes in TAF4b-TFIID correlate with promoter selectivity. *Mol Cell.* **29:**81-91.

Liu Y, Warfield L, Zhang C, Luo J, Allen J, Lang WH, Ranish J, Shokat KM, Hahn S (2009) Phosphorylation of the transcription elongation factor Spt5 by yeast Bur1 kinase stimulates recruitment of the PAF complex. *Mol Cell Biol* **29:** 4852-4863

Lyakhovich A, Shekhar MP (2004) RAD6B overexpression confers chemoresistance: RAD6 expression during cell cycle and its redistribution to chromatin during DNA damage-induced response. *Oncogene* **23:** 3097-3106

Mandal SS, Chu C, Wada T, Handa H, Shatkin AJ, Reinberg D (2004) Functional interactions of RNA-capping enzyme with factors that positively and negatively regulate promoter escape by RNA polymerase II. *Proc Natl Acad Sci U S A* **101:** 7572-7577

Marciniak RA, Sharp PA (1991) HIV-1 Tat protein promotes formation of more-processive elongation complexes. *EMBO J* **10:** 4189-4196

Marshall NF, Peng J, Xie Z, Price DH (1996) Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. *J Biol Chem* **271:** 27176-27183

Marshall NF, Price DH (1992) Control of formation of two distinct classes of RNA polymerase II elongation complexes. *Mol Cell Biol* **12:** 2078-2090

Marshall NF, Price DH (1995) Purification of P-TEFb, a transcription factor required for the transition into productive elongation. *J Biol Chem* **270:** 12335-12338

Matangkasombut O, Auty R, Buratowski S (2004) Structure and function of the TFIID complex. *Adv Protein Chem* **67:** 67-92

Mattaj IW, Lienhard S, Jiricny J, De Robertis EM (1985) An enhancer-like sequence within the Xenopus U2 gene promoter facilitates the formation of stable transcription complexes. *Nature* **316:** 163-167

Mavrich TN, Jiang C, Ioshikhes IP, Li X, Venters BJ, Zanton SJ, Tomsho LP, Qi J, Glaser RL, Schuster SC, Gilmour DS, Albert I, Pugh BF (2008) Nucleosome organization in the Drosophila genome. *Nature* **453:** 358-362

McCracken S, Fong N, Rosonina E, Yankulov K, Brothers G, Siderovski D, Hessel A, Foster S, Shuman S, Bentley DL (1997a) 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev* **11:** 3306-3318

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

Medlin J, Scurry A, Taylor A, Zhang F, Peterlin BM, Murphy S (2005) P-TEFb is not an essential elongation factor for the intronless human U2 snRNA and histone H2b genes. *EMBO J* **24:** 4154-4165

Medlin JE, Uguen P, Taylor A, Bentley DL, Murphy S (2003) The C-terminal domain of pol II and a DRB-sensitive kinase are required for 3' processing of U2 snRNA. *EMBO J* **22:** 925-934

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 Dev* **19:** 1401-1415

Miller T, Krogan NJ, Dover J, Erdjument-Bromage H, Tempst P, Johnston M, Greenblatt JF, Shilatifard A (2001) COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc Natl Acad Sci U S A* **98:** 12902-12907

Minsky N, Shema E, Field Y, Schuster M, Segal E, Oren M (2008) Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells. *Nat Cell Biol* **10:** 483-488

Moore MJ, Sharp PA (1993) Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature* **365:** 364-368

Mueller CL, Jaehning JA (2002) Ctr9, Rtf1, and Leo1 are components of the Paf1/RNA polymerase II complex. *Mol Cell Biol* **22:** 1971-1980

Mueller CL, Porter SE, Hoffman MG, Jaehning JA (2004) The Paf1 complex has functions independent of actively transcribing RNA polymerase II. *Mol Cell* **14:** 447-456

Muller F, Tora L (2004) The multicoloured world of promoter recognition complexes. *EMBO J* **23:** 2-8

Murphy S, Yoon JB, Gerster T. & Roeder RG (1992) Oct-1 and Oct-2 potentiate functional interactions of a transcription factor with the proximal sequence element of small nuclear RNA genes. Mol Cell Biol, 12, 3247-61.

Murphy S (1997) Differential in vivo activation of the class II and class III snRNA genes by the POU-specific domain of Oct-1. *Nucleic Acids Res* **25:** 2068-2076

Muratoglu S, Georgieva S, Papai G, Scheer E, Enunlu I, Komonyi O, Cserpan I, Lebedeva L, Nabirochkina E, Udvardy A, Tora L, Boros I. *Mol Cell Biol.***23:**306-21

Myers LC, Kornberg RD (2000) Mediator of transcriptional regulation. *Annu Rev Biochem* **69:** 729-749

Nakajima-Iijima S, Hamada H, Reddy P, Kakunaga T (1985) Molecular structure of the human cytoplasmic beta-actin gene: interspecies homology of sequences in the introns. *Proc Natl Acad Sci U S A* **82:** 6133-6137

Narita T, Yung TM, Yamamoto J, Tsuboi Y, Tanabe H, Tanaka K, Yamaguchi Y, Handa H (2007) NELF interacts with CBC and participates in 3' end processing of replication-dependent histone mRNAs. *Mol Cell* **26:** 349-365

Newey PJ, Bowl MR, Thakker RV (2009) Parafibromin--functional insights. *J Intern Med* **266:** 84-98

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

Ng SY, Gunning P, Eddy R, Ponte P, Leavitt J, Shows T, Kedes L (1985) Evolution of the functional human beta-actin gene and its multi-pseudogene family: conservation of noncoding regions and chromosomal dispersion of pseudogenes. *Mol Cell Biol* **5:** 2720-2732

Ni Z, Schwartz BE, Werner J, Suarez JR & Lis JT (2004). Coordination of transcription, RNA processing, and surveillance by P-TEFb kinase on heat shock genes. *Mol Cell*, 13, 55-65.

Nislow C, Ray E, Pillus L (1997) SET1, a yeast member of the trithorax family, functions in transcriptional silencing and diverse cellular processes. *Mol Biol Cell* **8:** 2421-2436

Noma K, Grewal SI (2002) Histone H3 lysine 4 methylation is mediated by Set1 and promotes maintenance of active chromatin states in fission yeast. *Proc Natl Acad Sci U S A* **99 Suppl 4:** 16438-16445

O'Geen H, Nicolet CM, Blahnik K, Green R, Farnham PJ (2006) Comparison of sample preparation methods for ChIP-chip assays. *Biotechniques* 41:577-80

Orphanides G, LeRoy G, Chang CH, Luse DS, Reinberg D (1998) FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell* **92:** 105-116

Orphanides G, Wu WH, Lane WS, Hampsey M, Reinberg D (1999) The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. *Nature* **400:** 284-288

Palancade B, Bensaude O (2003) Investigating RNA polymerase II carboxyl-terminal domain (CTD) phosphorylation. *Eur J Biochem* **270:** 3859-3870

Palangat M, Renner DB, Price DH, Landick R (2005) A negative elongation factor for human RNA polymerase II inhibits the anti-arrest transcript-cleavage factor TFIIS. *Proc Natl Acad Sci U S A* **102:** 15036-15041

Papai G, Weil PA, Schultz P (2011) New insights into the function of transcription factor TFIID from recent structural studies. *Curr Opin Genet Dev* **21:** 219-224

Park JM, Werner J, Kim JM, Lis JT, Kim YJ (2001) Mediator, not holoenzyme, is directly recruited to the heat shock promoter by HSF upon heat shock. *Mol Cell* **8:** 9-19

Pavri R, Zhu B, Li G, Trojer P, Mandal S, Shilatifard A, Reinberg D (2006) Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell* **125:** 703-717

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

Petesch SJ, Lis JT (2008) Rapid, transcription-independent loss of nucleosomes over a large chromatin domain at Hsp70 loci. *Cell* **134:** 74-84

Peterlin, B. & Price, D. (2006) Controlling the Elongation Phase of Transcription with P-TEFb. Mol. Cell. 23, 297-305

Phatnani HP, Greenleaf AL (2006) Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev* **20:** 2922-2936

Pilch DR, Marzluff WF (1991) Expression of histone-U1 snRNA chimeric genes: U1 promoters are compatible with histone 3' end formation. *Gene Expr* **1:** 41-53

Pirngruber J, Shchebet A, Schreiber L, Shema E, Minsky N, Chapman RD, Eick D, Aylon Y, Oren M, Johnsen SA (2009) CDK9 directs H2B monoubiquitination and controls replication-dependent histone mRNA 3'-end processing. *EMBO Rep* **10:** 894-900

Pokholok DK, Hannett NM, Young RA (2002) Exchange of RNA polymerase II initiation and elongation factors during gene expression in vivo. *Mol Cell* **9:** 799-809

Ponte P, Gunning P, Blau H, Kedes L (1983) Human actin genes are single copy for alpha-skeletal and alpha-cardiac actin but multicopy for beta- and gamma-cytoskeletal genes: 3' untranslated regions are isotype specific but are conserved in evolution. *Mol Cell Biol* **3:** 1783-1791

Price DH (2000) P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. *Mol Cell Biol* **20:** 2629-2634

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

Qiu H, Hu C, Wong CM, Hinnebusch AG (2006) The Spt4p subunit of yeast DSIF stimulates association of the Paf1 complex with elongating RNA polymerase II. *Mol Cell Biol* **26:** 3135-3148

Rahl PB, Lin CY, Seila AC, Flynn RA, McCuine S, Burge CB, Sharp PA, Young RA (2010) c-Myc regulates transcriptional pause release. *Cell* **141:** 432-445

Rao B, Shibata Y, Strahl BD, Lieb JD (2005) Dimethylation of histone H3 at lysine 36 demarcates regulatory and nonregulatory chromatin genome-wide. *Mol Cell Biol* **25:** 9447-9459

Reinberg D, Roeder RG (1987) Factors involved in specific transcription by mammalian RNA polymerase II. Purification and functional analysis of initiation factors IIB and IIE. *J Biol Chem* **262:** 3310-3321

Reinberg D, Sims RJ, 3rd (2006) de FACTo nucleosome dynamics. *J Biol Chem* **281:** 23297-23301

Reines D, Conaway RC, Conaway JW (1999) Mechanism and regulation of transcriptional elongation by RNA polymerase II. *Curr Opin Cell Biol* **11:** 342-346

Reines D, Mote J, Jr. (1993) Elongation factor SII-dependent transcription by RNA polymerase II through a sequence-specific DNA-binding protein. *Proc Natl Acad Sci U S A* **90:** 1917-1921

Renner DB, Yamaguchi Y, Wada T, Handa H, Price DH (2001) A highly purified RNA polymerase II elongation control system. *J Biol Chem* **276:** 42601-42609

Robzyk K, Recht J, Osley MA (2000) Rad6-dependent ubiquitination of histone H2B in yeast. *Science* **287:** 501-504

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

Rozenblatt-Rosen O, Hughes CM, Nannepaga SJ, Shanmugam KS, Copeland TD, Guszczynski T, Resau JH, Meyerson M (2005) The parafibromin tumor suppressor protein is part of a human Paf1 complex. *Mol Cell Biol* **25:** 612-620

Rozenblatt-Rosen O, Nagaike T, Francis JM, Kaneko S, Glatt KA, Hughes CM, LaFramboise T, Manley JL, Meyerson M (2009) The tumor suppressor Cdc73 functionally associates with CPSF and CstF 3' mRNA processing factors. *Proc Natl Acad Sci U S A* **106:** 755-760

Ruthenburg AJ, Allis CD, Wysocka J (2007) Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol Cell* **25:** 15-30

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

Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL, Mellor J, Kouzarides T (2002) Active genes are tri-methylated at K4 of histone H3. *Nature* **419:** 407-411

Saunders A, Core LJ, Lis JT (2006) Breaking barriers to transcription elongation. *Nat Rev Mol Cell Biol* **7:** 557-567

Schaft D, Roguev A, Kotovic KM, Shevchenko A, Sarov M, Neugebauer KM, Stewart AF (2003) The histone 3 lysine 36 methyltransferase, SET2, is involved in transcriptional elongation. *Nucleic Acids Res* **31:** 2475-2482

Scheffner M, Nuber U, Huibregtse JM (1995) Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade. *Nature* **373:** 81-83

Schramm L, Pendergrast PS, Sun Y, Hernandez N (2000) Different human TFIIIB activities direct RNA polymerase III transcription from TATA-containing and TATA-less promoters. *Genes Dev* **14:** 2650-2663

Schratt G, Weinhold B, Lundberg AS, Schuck S, Berger J, Schwarz H, Weinberg RA, Ruther U, Nordheim A (2001) Serum response factor is required for immediate-early gene activation yet is dispensable for proliferation of embryonic stem cells. *Mol Cell Biol* **21:** 2933-2943

Schroeder SC, Schwer B, Shuman S, Bentley D (2000) Dynamic association of capping enzymes with transcribing RNA polymerase II. *Genes Dev* **14:** 2435-2440

Schroeder SC, Zorio DA, Schwer B, Shuman S, Bentley D (2004) A function of yeast mRNA cap methyltransferase, Abd1, in transcription by RNA polymerase II. *Mol Cell* **13:** 377-387

Schultz J, Copley RR, Doerks T, Ponting CP, Bork P (2000) SMART: a web-based tool for the study of genetically mobile domains. *Nucleic Acids Res* **28:** 231-234

Sehgal PB, Darnell JE, Jr., Tamm I (1976) The inhibition by DRB (5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole) of hnRNA and mRNA production in HeLa cells. *Cell* **9:** 473-480

Selth LA, Sigurdsson S, Svejstrup JQ (2010) Transcript Elongation by RNA Polymerase II. *Annu Rev Biochem* **79:** 271-293

Shatkin AJ, Manley JL (2000) The ends of the affair: capping and polyadenylation. *Nat Struct Biol* **7:** 838-842

Sheldon M, Ratnasabapathy R, Hernandez N (1993) Characterization of the inducer of short transcripts, a human immunodeficiency virus type 1 transcriptional element that activates the synthesis of short RNAs. *Mol Cell Biol* **13:** 1251-1263

Shilatifard A (2006) Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu Rev Biochem* **75:** 243-269

Shilatifard A (2008) Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Curr Opin Cell Biol* **20:** 341-348

Sims RJ, 3rd, Belotserkovskaya R, Reinberg D (2004a) Elongation by RNA polymerase II: the short and long of it. *Genes Dev* **18:** 2437-2468

Sims RJ, 3rd, Mandal SS, Reinberg D (2004b) Recent highlights of RNA-polymerase-II-mediated transcription. *Curr Opin Cell Biol* **16:** 263-271

Sims RJ, 3rd, Nishioka K, Reinberg D (2003) Histone lysine methylation: a signature for chromatin function. *Trends Genet* **19:** 629-639

Smale ST (2001) Core promoters: active contributors to combinatorial gene regulation. *Genes Dev* **15:** 2503-2508

Smale ST, Kadonaga JT (2003) The RNA polymerase II core promoter. *Annu Rev Biochem* **72:** 449-479

Spencer CA, Groudine M (1990) Transcription elongation and eukaryotic gene regulation. *Oncogene* **5:** 777-785

Squazzo SL, Costa PJ, Lindstrom DL, Kumer KE, Simic R, Jennings JL, Link AJ, Arndt KM, Hartzog GA (2002) The Paf1 complex physically and functionally associates with transcription elongation factors in vivo. *EMBO J* **21:** 1764-1774

Stiller JW, Hall BD (2002) Evolution of the RNA polymerase II C-terminal domain. *Proc Natl Acad Sci U S A* **99:** 6091-6096

Stolinski LA, Eisenmann DM, Arndt KM (1997) Identification of RTF1, a novel gene important for TATA site selection by TATA box-binding protein in Saccharomyces cerevisiae. *Mol Cell Biol* **17:** 4490-4500

Strahl BD, Grant PA, Briggs SD, Sun ZW, Bone JR, Caldwell JA, Mollah S, Cook RG, Shabanowitz J, Hunt DF, Allis CD (2002) Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Mol Cell Biol* **22:** 1298-1306

Swanson MS, Winston F (1992) SPT4, SPT5 and SPT6 interactions: effects on transcription and viability in Saccharomyces cerevisiae. *Genetics* **132:** 325-336

Tamm I, Kikuchi T, Darnell JE, Jr., Salditt-Georgieff M (1980) Short capped hnRNA precursor chains in HeLa cells: continued synthesis in the presence of 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole. *Biochemistry* **19:** 2743-2748

Thomas MC, Chiang CM (2006) The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol* **41:**105-78

Ucker DS, Yamamoto KR (1984) Early events in the stimulation of mammary tumor virus RNA synthesis by glucocorticoids. Novel assays of transcription rates. *J Biol Chem* **259:** 7416-7420

Uguen P, Murphy S (2003) The 3' ends of human pre-snRNAs are produced by RNA polymerase II CTD-dependent RNA processing. *EMBO J* **22:** 4544-4554

Wada T, Orphanides G, Hasegawa J, Kim DK, Shima D, Yamaguchi Y, Fukuda A, Hisatake K, Oh S, Reinberg D, Handa H (2000) FACT relieves DSIF/NELF-mediated inhibition of transcriptional elongation and reveals functional differences between P-TEFb and TFIIH. *Mol Cell* **5:** 1067-1072

Wada T, Takagi T, Yamaguchi Y, Ferdous A, Imai T, Hirose S, Sugimoto S, Yano K, Hartzog GA, Winston F, Buratowski S, Handa H (1998) DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes Dev* **12:** 343-356

Wade PA, Werel W, Fentzke RC, Thompson NE, Leykam JF, Burgess RR, Jaehning JA, Burton ZF (1996) A novel collection of accessory factors associated with yeast RNA polymerase II. *Protein Expr Purif* **8:** 85-90

Weake VM, Workman JL (2008) Histone ubiquitination: triggering gene activity. *Mol Cell* **29:** 653-663

Wen Y, Shatkin AJ (1999) Transcription elongation factor hSPT5 stimulates mRNA capping. *Genes Dev* **13:** 1774-1779

Westin G, Zabielski J, Hammarstrom K, Monstein HJ, Bark C, Pettersson U (1984) Clustered genes for human U2 RNA. *Proc Natl Acad Sci U S A* **81:** 3811-3815

Wieczorek E, Brand M, Jacq X, Tora L (1998) Function of TAF(II)-containing complex without TBP in transcription by RNA polymerase II. *Nature* **393:** 187-191

Will CL, Luhrmann R (2001) Spliceosomal UsnRNP biogenesis, structure and function. *Curr Opin Cell Biol* **13:** 290-301

Wilson CJ, Chao DM, Imbalzano AN, Schnitzler GR, Kingston RE, Young RA (1996) RNA polymerase II holoenzyme contains SWI/SNF regulators involved in chromatin remodeling. *Cell* **84:** 235-244

Wind M, Reines D (2000) Transcription elongation factor SII. *Bioessays* **22:** 327-336

Winston F, Chaleff DT, Valent B, Fink GR (1984) Mutations affecting Ty-mediated expression of the HIS4 gene of Saccharomyces cerevisiae. *Genetics* **107:** 179-197

Wong MW, Henry RW, Ma B, Kobayashi R, Klages N, Matthias P, Strubin M, Hernandez N (1998) The large subunit of basal transcription factor SNAPc is a Myb domain protein that interacts with Oct-1. *Mol Cell Biol* **18:** 368-377

Wood A, Krogan NJ, Dover J, Schneider J, Heidt J, Boateng MA, Dean K, Golshani A, Zhang Y, Greenblatt JF, Johnston M, Shilatifard A (2003a) Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. *Mol Cell* **11:** 267-274

Wood A, Schneider J, Dover J, Johnston M, Shilatifard A (2003b) The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. *J Biol Chem* **278:** 34739-34742

Wood A, Shilatifard A (2006) Bur1/Bur2 and the Ctk complex in yeast: the split personality of mammalian P-TEFb. *Cell Cycle* **5:** 1066-1068

Wu CH, Yamaguchi Y, Benjamin LR, Horvat-Gordon M, Washinsky J, Enerly E, Larsson J, Lambertsson A, Handa H, Gilmour D (2003) NELF and DSIF cause promoter proximal pausing on the hsp70 promoter in Drosophila. *Genes Dev* **17:** 1402-1414

Veenstra GJ, Wolffe AP (2001) Gene-selective developmental roles of general transcription factors. *Trends Biochem Sci.***26:**665-71.

Yamada T, Yamaguchi Y, Inukai N, Okamoto S, Mura T, Handa H (2006) P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. *Mol Cell* **21:** 227-237

Yamaguchi Y, Inukai N, Narita T, Wada T, Handa H (2002) Evidence that negative elongation factor represses transcription elongation through binding to a DRB sensitivity-inducing factor/RNA polymerase II complex and RNA. *Mol Cell Biol* **22:** 2918-2927

Yamaguchi Y, Narita T, Inukai N, Wada T, Handa H (2001) SPT genes: key players in the regulation of transcription, chromatin structure and other cellular processes. *J Biochem (Tokyo)* **129:** 185-191

Yamaguchi Y, Takagi T, Wada T, Yano K, Furuya A, Sugimoto S, Hasegawa J, Handa H (1999a) NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell* **97:** 41-51

Yamaguchi Y, Wada T, Watanabe D, Takagi T, Hasegawa J, Handa H (1999b) Structure and function of the human transcription elongation factor DSIF. *J Biol Chem* **274:** 8085-8092

Yang Z, Zhu Q, Luo K, Zhou Q (2001) The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* **414:** 317-322

Yankulov K, Blau J, Purton T, Roberts S, Bentley DL (1994) Transcriptional elongation by RNA polymerase II is stimulated by transactivators. *Cell* **77:** 749-759

Yoh SM, Cho H, Pickle L, Evans RM, Jones KA (2007) The Spt6 SH2 domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export. *Genes Dev* **21:** 160-174

Yoon JB, Murphy S, Bai L, Wang Z, Roeder RG (1995) Proximal sequence element-binding transcription factor (PTF) is a multisubunit complex required for transcription of both RNA polymerase II- and RNA polymerase III-dependent small nuclear RNA genes. *Mol Cell Biol* **15:** 2019-2027

Yuo CY, Ares M, Jr., Weiner AM (1985) Sequences required for 3' end formation of human U2 small nuclear RNA. *Cell* **42:** 193-202

Zaborowska J, Taylor A, Murphy S (2012) A novel TBP-TAF complex on RNA polymerase II-transcribed snRNA genes. *Transcription* **3:** 92-104

Zhu B, Mandal SS, Pham AD, Zheng Y, Erdjument-Bromage H, Batra SK, Tempst P, Reinberg D (2005a) The human PAF complex coordinates transcription with events downstream of RNA synthesis. *Genes Dev* **19:** 1668-1673

Zhu B, Zheng Y, Pham AD, Mandal SS, Erdjument-Bromage H, Tempst P, Reinberg D (2005b) Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. *Mol Cell* **20:** 601-611

Zhu H, Joliot V, Prywes R (1994) Role of transcription factor TFIIF in serum response factor-activated transcription. *J Biol Chem* **269:** 3489-3497