Universidade de Lisboa

Faculdade de Ciências

Departamento de Biologia Vegetal


# The role of chromatin on RNA Polymerase II transcription termination

**Ana Filipa Batalha Martins**

Dissertação

Mestrado em Biologia Molecular e Genética

2012

Universidade de Lisboa

Faculdade de Ciências

Departamento de Biologia Vegetal


# The role of chromatin on RNA Polymerase II transcription termination

**Ana Filipa Batalha Martins**

Dissertação orientada pelo Doutor Sérgio Fernandes de Almeida e pelo Prof. Doutor Júlio Duarte

Mestrado em Biologia Molecular e Genética

2012

**Table of Contents**

**Sumário**

Tem-se tornado cada vez mais evidente que a transcrição no ambiente da cromatina é um processo extremamente complexo e finamente regulado. É agora claro que a maquinaria transcricional eucariótica está adaptada para explorar a presença de nucleossomas de variadíssimas e sofisticadas maneiras. Após activação, os genes sofrem drásticas mudanças na estrutura da cromatina. Isto é possível devido a um sistema de montagem da cromatina dependente da transcrição que integra a acção coordenada de numerosos factores proteicos capazes de modificar as propriedades da cromatina. Estes factores vão cooperar ou competir de forma a alterar o estado da cromatina entre permissivo e não permissivo, levando à activação ou repressão da transcrição. O trabalho apresentado nesta tese focou-se na terminação da transcrição. O objectivo foi investigar de que forma está a cromatina implicada na fase final da transcrição. Assim, levantou-se a hipótese que características específicas da cromatina são observadas nas regiões do DNA onde a RNA polimerase II se dissocia do DNA molde. Para testar esta hipótese, foram utilizadas técnicas bioquímicas, como por exemplo, imunoprecipitação de cromatina e digestão com nucleases. Os resultados revelaram um comportamento dinâmico da RNA polimerase II dependente da transcrição em regiões *downstream* do limite 3´ dos genes. Além disto, os resultados sugerem que os nucleossomas localizados nesta região são desmontados ineficientemente pelos complexos transcricionais da RNA polimerase II. De facto, experiências adicionais demonstraram que regiões a jusante do local de poliadenilação, onde a terminação ocorre, exibem reduzido recrutamento de chaperones de histonas. Finalmente, foi possível confirmar um enriquecimento de modificações específicas de histonas nas regiões 3' que flanqueiam os genes. Em conjunto, os resultados aqui apresentados permitem sugerir um modelo pelo qual o estado da cromatina a jusante do local de poliadenilação facilita os eventos moleculares que levam à terminação da transcrição.

Palavras-chave: Cromatina, Nucleossomas, Chaperones de histonas, Terminação da transcrição pela RNA polimerase II

**Resumo**

A célula eucariótica armazena a sua informação genética em moléculas de DNA que podem ter mais de um 1 metro de comprimento. De forma a caber no volume limitado do núcleo, o DNA é compactado cerca de 20 000 vezes. A forma mais compactada na qual o DNA pode ser encontrado nas células é produzida durante a mitose: o cromossoma metafásico. Cada cromossoma consiste numa longa molécula de DNA organizada em níveis crescentes de compactação por proteínas com as quais se encontra associado. O primeiro nível de organização é a compactação do DNA numa fibra de 10nm formando uma estrutura designada nucleossoma. Os nucleossomas são a unidades estrutural da cromatina e são compostos pela associação do DNA com uma família de pequenas proteínas básicas designadas histonas. Estas proteínas são as mais abundantemente associadas ao DNA eucariótico e são denominadas H1, H2A, H2B, H3 e H4. Cada nucleossoma é composto por duas cópias de cada histona (excepto a histona H1) agregadas numa estrutura octamérica que é envolvida por cerca de 146-147 pares de bases de DNA. Os nucleossomas estão separados por uma sequência de DNA com tamanho variável à qual está associada a histona H1. As histonas são caracterizadas pela presença de um domínio globular formado pela região C-terminal e por uma extensão N-terminal, designada de cauda. A característica mais relevante desta cauda é o elevado número de modificações pós-traducionais que os seus resíduos podem sofrer tais como, acetilações, metilações ou fosforilações. Estas modificações têm sido extensivamente implicadas na reparação do DNA, na activação transcricional e no silenciamento génico. Além disto, estas caudas assumem também um papel importante na estabilidade e montagem dos nucleossomas, contribuindo para a compactação da cromatina. O nível seguinte de compactação da estrutura da cromatina é a fibra de 30nm formada por arranjos de nucleossomas. Esta estrutura representa uma compactação de cerca de 50 vezes, sendo necessário o compactamento adicional do DNA. A arquitectura destes níveis organizacionais superiores não é clara e permanece incompreendida.

Apesar de crucial para as células, o empacotamento do DNA impõe limitações na sua acessibilidade influenciando assim todos os processos dependentes deste tais como a reparação, replicação recombinação e transcrição. Consequentemente a conversão entre estados de cromatina com diferentes graus de compactação é um processo altamente regulado. É frequentemente necessário rearranjar ou mobilizar os nucleossomas de forma a facilitar tais processos. Como tal, as células desenvolverem múltiplas estratégias para o controlo da estrutura da cromatina e da acessibilidade do DNA tais como: enzimas modificadoras de histonas, variantes de histonas, complexos remodeladores da cromatina e chaperones de histonas. O modelo actualmente aceite visa que estes mecanismos agem em

conjunto e coordenadamente de forma a restringir o acesso da maquinaria celular ao DNA durante determinados períodos de tempo e em locais precisos. Os nucleossomas são muito mais dinâmicos e ajustáveis do que anteriormente previsto. Têm sido feitos progressos consideráveis no esclarecimento dos mecanismos que permitem a passagem da RNA Polimerase II (RNAPII) através dos nucleossomas aquando da transcrição. Presentemente, os nucleossomas são vistos como componentes essenciais deste processo em vez de uma simples barreira passiva. Vários chaperones de histonas têm sido implicados em facilitar a transcrição, particularmente o *Facilitator of Chromatin Transcription* (FACT). Vários estudos sugerem a existência de mecanismos que expulsam as histonas nucleossomais em frente da RNAPII durante o elongamento e mecanismos que coordenam a posterior recuperação dos nucleossomas imediatamente atrás desta. A remontagem dos nucleossomas após a transcrição é fundamental de forma a manter a integridade da estrutura da cromatina, impedindo assim a iniciação da transcrição intragénica.

Só recentemente se começou a compreender esta regulação da transcrição mediada pela cromatina. Os aspectos mecanísticos deste processo estão longe de serem completamente compreendidos. Inúmeros exemplos de um acoplamento funcional entre a cromatina e a transcrição pela RNAPII têm sido reportados durante as fases de iniciação e elongação. No entanto, o papel da cromatina na terminação da transcrição não tem sido devidamente estudado. Assim, o principal objectivo desta tese foi averiguar como é que a cromatina influencia a última fase da transcrição.

De forma a investigar-se a dinâmica da cromatina após activação da transcrição, tirou-se partido do programa de expressão génica activado durante a *unfolded protein response* (UPR), uma reacção celular accionada pela presença abundante de proteínas *misfolded* no retículo endoplasmático. Durante a UPR, determinados genes são activados de forma a restaurar a homeostase celular. Entre estes genes encontram-se *CHOP*, *ERP70* e *HERPUD*. A UPR pode ser induzida através da adição do agente redutor ditiotreitol (DTT) ao meio de cultura celular. Assim, células HeLa expostas a este composto mostraram um aumento significativo nos níveis de mRNA o que sugere um aumento da taxa de transcrição dos mesmos. A ocupação da RNAPII ao longo dos genes foi avaliada antes e depois da activação transcricional pelo DTT. Estas experiências revelaram um comportamento dinâmico da RNAPII dependente da transcrição mesmo em regiões a jusante do local de poliadenilação (poli(A)). Visto que os nucleosomas podem agir como uma barreira à passagem da RNAPII e causar o seu abrandamento, estes podem desempenhar um papel na terminação da transcrição impedindo a transcrição espúria de genes adjacentes e induzindo os acontecimentos mecanísticos que levam à dissociação da RNAPII do DNA molde. De forma a investigar esta hipótese, os níveis das histonas H2B e H3 foram determinados. Após a activação transcricional dos genes *CHOP*, *ERP70* e *HERPUD*, estes

demonstraram uma redução dos níveis da histona H3 ao longo de todas as regiões génicas estudadas. No entanto, esta redução não foi consistente e é possível observar uma diminuição da mesma à medida que avançamos em direcção da região terminal dos genes. Para a histona H2B os resultados apresentados foram similares. De forma a estudar mais aprofundadamente a localização dos nucleossomas, foi realizado um ensaio com nuclease microcócica (Mnase), que permite mapear a localização dos nucleossomas. Estas experiências revelaram que após activação da transcrição, os nucleossomas localizados *downstream* do local poli(A) não são eficientemente ejectados pela RNAPII. Uma vez que os chaperones de histonas são cruciais para este processo, foi analisado o recrutamento do FACT. O resultado destas experiências permitiu conlcuir que o mecanismo pelo qual o FACT é recrutado para as regiões codificantes é interrompido após o local poli(A). Por último, a distribuição no genoma humano de três modificações da histona H3 (H3K4me, H3K4me2 e H3K9me) obtida através de uma análise bioinformática efectuada no laboratório de acolhimento revelou um enriquecimento específico na região de terminação da transcrição. Estes resultados foram validados experimentalmente através da quantificação destas marcas em genes individuais.

Em conjunto, os dados obtidos nesta tese permitem propor um novo modelo para a terminação da transcrição. Segundo este modelo, após o local poli(A), os nucleossomas actuam de modo a impedir o avanço da RNAPII porque os complexos de transcrição não são capazes de recrutar eficientemente chaperones de histonas nesta região. A pausa que deste modo será induzida na RNAPII facilitará os eventos moleculares que culminam na terminação de transcrição. Além disto a descoberta de modificações na histona H3 presentes na região terminal dos genes alimenta o interesse pela pesquisa adicional do papel destas marcas na terminação da transcrição. Estes dados proporcionam uma nova visão para a terminação da transcrição, um processo fundamental, que continua a ser uma das fases menos compreendidas do ciclo de transcrição.

**Abstract**

It has become increasingly evident that transcription in a chromatin environment is an extremely complex and finely tuned process. It is now clear that the eukaryotic transcriptional machinery is adapted to exploit the presence of nucleosomes in a variety of sophisticated ways. Nucleosomes are now seen as a crucial component of this process, rather than a simple passive barrier. Upon activation, genes undergo severe changes in chromatin structure. This is achieved through a transcription-dependent chromatin assembly system that involves the orchestrated action of numerous protein factors capable of modifying chromatin properties. These proteins will cooperate or compete to change the chromatin state between permissive and non-permissive, leading to activation or repression of transcription. Herein we focused on transcription termination. We aimed at investigating whether chromatin is a determinant of the final stage of transcription. In this sense, we hypothesized that specific chromatin features are observed at the DNA regions where RNA Polymerase II dissociates from the transcribed template. To test this hypothesis, we used a biochemical approach, which includes techniques such as chromatin immunoprecipitation and nuclease digestion assays. Our results reveal a transcription-dependent dynamic behavior of RNA Polymerase II molecules in regions downstream the 3' boundary of genes. Our data further suggest that the nucleosomes occupying this region are inefficiently disassembled by the RNA polymerase II transcription complexes. Notably, the regions downstream the poly(A) site, where termination takes place, exhibited reduced recruitment of histone chaperones. Finally, we were able to confirm the presence of histone modifications enriched at the 3' flanking region of genes. Altogether, these data allow us to envisage a model by which the chromatin landscape downstream the poly(A) site facilitates the molecular events that drive transcription termination.

Keywords: Chromatin, Nucleosomes, Histone chaperones, RNA polymerase II transcription termination

## Introduction

### 1. Chromatin basics

The eukaryotic cell stores its genetic information in DNA molecules that can be over 1 meter in length (Mariño-Ramírez et al., 2005). In order to fit into the limited volume of the nucleus, the DNA is compacted about 20,000-fold (Németh and Langst, 2004). The most highly condensed form of DNA that can be encountered in cells is produced during mitosis: the metaphase chromosomes (Belmont et al., 1987). Each chromosome consists of a single, long molecule of DNA, organized into increasing levels of condensation by proteins with which it is intricately complexed. The first level of organization is the compaction of the DNA into a 10nm fiber which confers a 5- to 10- fold compaction (Bell et al., 2011). This is achieved through the association of DNA with specific proteins named histones, forming a nucleoprotein complex that resembles “beads-on-a-string” (Li and Reinberg, 2011). The beadlike structures are termed nucleosomes and represent the primary structural unit of the chromatin (see 1.2 Nucleosome structure). The next level of chromatin organization is the 30-nm fiber, which is composed of packed nucleosome arrays mediated by core histone internucleosomal interactions (Mariño-Ramírez et al., 2005). This produces an approximately 50-fold net compaction (Németh and Langst, 2004). Additional folding is required to compact DNA further. The architecture of these high order organizational levels is unclear but several evidences support the loop model (Németh and Langst, 2004). This model postulates that the 30nm chromatin fibers form large loops of DNA that are tethered at their bases to a proteinacious structure referred to as nuclear scaffold (Mardsen, 1979). Although histones are the predominant proteins in chromatin, nonhistone proteins are also involved in organizing chromatin structure (Mardsen, 1979). Nonhistone proteins provide the structural scaffold for these loops (Mardsen, 1979). Additional folding of the scaffold has been proposed to compact the structure into the highly condensed form of chromatin (Belmont and Bruce, 1994; Belmont et al., 1987). The precise structure of chromatin beyond the 30nm fiber remains uncharacterized and is poorly understood (Horn and Peterson, 2002; Luger, 2003).

Compaction also protects the DNA from damage, confers an overall organization to each molecule of DNA and only DNA packaged into a chromosome can be transmitted efficiently to daughter cells during cell division (Hamilton et al., 2011). Furthermore, it adds an additional level to the regulation of DNA-dependent processes as it influences the access to the DNA packaged in chromatin (Bai and Morozov, 2010; Bell et al., 2011). Consequently, the interconversion between chromatin states is a tightly regulated process and the cell developed multiple strategies for the most favorable use of chromatin (Lusser and Kadonaga, 2003) (see 1.3 Nucleosome dynamics and genome accessibility).

In summary, compaction of eukaryotic genomes into condensed fibers is required due to spatial restrictions imposed by the nucleus of cells (Németh and Langst, 2004). This compaction is achieved through the association of DNA with proteins (mainly histone proteins) (Horn and Peterson, 2002) resulting in a hierarchy of folding levels that range from the 10nm diameter chromatin fiber to the highly compacted metaphase chromosomes. (Belmont, et al., 1987) Although crucial to the cell, this packaging imposes limitations to the DNA accessibility from which the cell takes advantage off as an additional level to the regulation of DNA-dependent processes such as DNA repair, replication, recombination and transcription (Bell et al., 2011).

## 2. Nucleosome structure

Nucleosomes represent the first level of organization of chromatin and consist in the association between DNA and histone proteins (Chakravarthy et al., 2005) . Histones are a family of small, basic proteins (Mariño-Ramírez et al., 2005). They are the most abundant proteins associated with eukaryotic DNA and cells commonly contain five abundant histones: H1, H2A, H2B, H3 and H4 (Hondele and Ladurner, 2011). These proteins are rich in positively charged basic aminoacids which interact with the negatively charged phosphodiester backbone of DNA resulting in a tight bond (Richmond et al., 1997). Each nucleosome is composed of two copies of histones H2A, H2B, H3 and H4, assembled in an octameric core (Németh and Langst, 2004). Tightly wrapped around this octamer in a left-handed superhelix are about 146-147 bp of DNA (Mariño-Ramírez et al., 2005). This complex forms the nucleosome core particle, which represents the basic repeating unit of chromatin (Chakravarthy et al., 2005).

Each core histone contains two separate functional domains: a histone-fold motif and a N-terminal extension called “tail” (Horn and Peterson, 2002). The histone fold motif is a globular domain that mediates the heterodimeric interactions between core histones (Mariño-Ramírez et al., 2005). H3 and H4 histones first form heterodimers that then come together to form a tetramer which binds to DNA and directs the subsequent association of histones H2A-H2B dimers (Németh and Langst, 2004). This domain is sufficient for both histone-histone and histone-DNA contacts within the nucleosome (Horn and Peterson, 2002). The N-terminal tails are exposed on the outside of the DNA appearing as unstructured random coils and are also involved in internucleosomal interactions (Mariño-Ramírez et al., 2005). A striking feature of the N-terminal tails is that they are subjected to extensive post-translational modifications (Bannister and Kouzarides, 2011; Kouzarides, 2007; Zhou et al., 2011). Their residues possess a large number of distinct modifications, such as acetylation, methylation, phosporylation and ubiquitylation (Kouzarides, 2007). These modifications are the keystone

of epigenetics (Jenuwein and Allis, 2001). This term refers to heritable changes in an organism that alter gene expression without altering the DNA sequence (Bird, 2007). Implications in DNA replication, transcriptional activation and silencing have also been widely reported (Kouzarides, 2007). Furthermore these tails also play important roles in nucleosome stability and chromatin assembly and contribute to define the condensed state of the chromatin fiber and higher order structures (see 1.3 Nucleosome dynamics and genome accessibility).

Nucleosomes are connected by a segment of DNA of variable length called linker DNA (Luger, 2003). Associated with the linker DNA is histone H1, also termed linker histone, which further tightens the association of the DNA with the nucleosome (Mariño-Ramírez et al., 2005). Linker histones are not related in sequence to the core histones, but they also contain a globular domain and a N-terminal tail (Horn and Peterson, 2002). Although only the linker histone globular domain is essential for binding to nucleosomes, the tail is believed to be important for chromatin folding (Horn and Peterson, 2002). Binding of histone H1 is therefore implicated in facilitating the formation of high-order chromatin structures (Németh and Langst, 2004).

### 3. Nucleosome dynamics and genome accessibility

Every nuclear process that requires access to DNA functions in the context of chromatin (Bell et al., 2011). The accessibility of DNA that is sequestered in chromatin differs dramatically from that of linear protein-free DNA (Chakravarthy et al., 2005). This has fundamental implications for all biological processes that use DNA as a template, such as transcription, replication, DNA repair, and recombination (Bell et al., 2011; Chakravarthy et al., 2005). Nucleosomes, as the main packaging element of DNA within the nucleus, are the primary determinant of DNA accessibility (Richmond et al., 1997). To facilitate DNA-directed processes in chromatin, it is often necessary to rearrange or to mobilize the nucleosomes (Bell et al., 2011). This process is termed chromatin remodeling and it encompasses a range of structural transitions between different chromatin conformational states (Flaus and Owen-Hughes, 2001). The mechanisms that control chromatin structure and thus DNA accessibility involve the targeted action of four broad classes of players: histone modifiers, histone variants, chromatin remodelers and histone chaperones (Avvakumo et al., 2011). These mechanisms operate in concert restricting the access of the cellular machinery to their target genomic regions (Mariño-Ramírez et al., 2005).

**Histone modifiers:** As mentioned, a striking feature of the histone N-terminal tail is the set of post-translational modifications it can host. Such modifications are carried out by a group of enzymes commonly termed histone modifiers (Kouzarides, 2007). Some examples of these modifications are acetylation, methylation, phosporylation and ubiquitylation (Bannister and Kouzarides, 2011; Bell et al., 2011; Kouzarides, 2007; Mariño-ramírez et al., 2005). These modifications may translate either directly, by altering the nucleosome properties, or indirectly, by functioning as signaling and docking platforms for the recruitment of numerous proteins, to an altered level of chromatin compaction (Bannister and Kouzarides, 2011; Kouzarides, 2007). The best-studied example of how histone modifications can directly affect chromatin structure is the acetylation of lysine residues. Acetylation neutralizes the positive charge of lysine and thus weakens the interaction with DNA providing a more permissive chromatin state (Petesch and Lis, 2012). Another extensively studied modification is the trimethylation of lysine 36 of histone H3 (H3K36me3). This modification is highly correlated with transcription elongation, splicing and is necessary for repression of cryptic transcription initiation (Almeida et al., 2011; Berger, 2007; Luco et al., 2011; Moore and Proudfoot, 2009). Ubiquitylation of H2BK123 is also associated with transcriptionally active regions (Petesch and Lis, 2012). The indirect role provided by these modifications is given by the ability of some proteins to recognize them in a very precise manner (Bottomley, 2004). These interactions are mediated, for example, by specific protein domains called bromodomains and chromodomains (Bottomley, 2004). Bromodomain-containing proteins interact with acetylated histone tails whereas chromodomain-containing proteins interact with methylated histone tails (Bottomley, 2004). Accumulating evidence suggests a link between different patterns and specific biological events. This has led to the “histone code” hypothesis (Strahl and Allis, 2000). This hypothesis predicts that distinct histone modifications act in combination to form alternative patterns that are “read” by other proteins to bring about unique and distinct outcomes in chromatin-templated processes (Strahl and Allis, 2000). The discovery of additional effectors and new modifications will provide new insight into how particular histone modifications affect chromatin structure and composition.

**Histone variants:** Histone variants are specialized histones that can replace core histones and exhibit specific spatial and temporal patterns (Chakravarthy et al., 2005). They serve to demarcate particular regions or confer specialized functions to the nucleosome into which they were incorporated (Bell et al., 2011). Due to their unique structural properties, replacement of canonical histones with variant forms may alter the interaction surfaces and the overall stability of nucleosomes and contribute to the formation of altered chromatin structures (Bell et al., 2011). Almost all eukaryotic organisms have variants of both histone

H2A and histone H3 and some contain histone H2B variants, but no variants of histone H4 appear to exist (Petesch and Lis, 2012). There are numerous examples of different histones variants, for instance, H2A.Z and H3.3, whose presence in the nucleosomes correlates positively with transcription (Kamakaka and Biggins, 2005). The deposition of H2A.Z into chromatin is important for the proper transcription of many genes and essential for many organisms (Petesch and Lis, 2012). H3.3 is deposited onto transcriptionally active genes in a replication-independent manner and, remarkably, only differs from the canonical H3 by four aminoacids (Deal, 2010; Petesch and Lis, 2012). Nucleosomes containing these variant histones occupy regions surrounding the promoter and 5' ends of transcribed genes and are thought to contribute to 5' nucleosome depleted regions and facilitate RNA Polymerase II access to the underlying DNA (Petesch and Lis, 2012).

**Chromatin remodelers:** Nucleosome stability is also influenced by large protein complexes referred to as chromatin remodeling complexes (Flaus and Owen-hughes, 2001; Lusser and Kadonaga, 2003). These multi-protein complexes facilitate changes in nucleosome occupancy or interaction with the DNA using the energy of ATP hydrolysis (Hargreaves and Crabtree, 2011). There are multiple types of remodeling complexes in a cell and they can have as few as two subunits or more than ten subunits (Hargreaves and Crabtree, 2011). Although the ATP hydrolyzing subunit is relatively well-conserved, the remaining subunits differ among these complexes and can, for example, target them to particular locations in the genome (Flaus and Owen-hughes, 2001). Currently there are four known main families of chromatin remodelers that carry out these functions: switch/sucrose nonfermentable (SWI/SNF), imitation switch (ISWI), chromodomain-helicase- DNA-binding protein (CHD) and inositol-requiring 80 (INO80) (Petesch and Lis, 2012). The SWI/SNF complex has been better studied for its role in nucleosome depletion near promoters and the ISWI family has been shown to help RNA Polymerase II overcome the nucleosomal barrier during transcription (Petesch and Lis, 2012; Tolkunov et al., 2011). The major reported function of CHD is to slide nucleosomes into ordered arrays throughout gene bodies and promote assembly of chromatin whereas INO80 facilitates transcription elongation by its ability to exchange histones (Lusser, 2005; Petesch and Lis, 2012). In either situation, these complexes alter the chromatin structure and, as a result, the DNA accessibility (Hargreaves and Crabtree, 2011).

**Histone chaperones:** Histone chaperones prevent nonspecific interactions between DNA and histones promoting proper nucleosome assembly (Avvakumov et al., 2011). They further associate with histones upon their synthesis, escort them into the nucleus, and assist their association with DNA during different processes such as DNA replication, repair, or

transcription (Avvakumov et al., 2011). Contrary to chromatin remodelers, these factors do not use the energy of ATP but are instead endowed with a strong affinity for specific surfaces of histones (Petesch and Lis, 2012). Histone chaperones also participate in nucleosome disassembly and reassembly (Park and Luger, 2008; Akey and Luger, 2003). Numerous histone chaperones have been found to associate with specific histones, histone post-translational modifications and elongation factors (Petesch and Lis, 2012). Some of the most extensively studied histone chaperones are the Facilitator of Chromatin Transcription (FACT), anti-silencing factor 1 (ASF1) and suppressor of ty homolog 6 (SPT6) (Petesch and Lis, 2012). FACT consists of two subunits, SPT16 and SSRP1 and was first identified by its ability to facilitate RNA Polymerase II transcription through a nucleosome template in vitro (Orphanides et al., 1998, 1999). ASF1 and SPT6 possess H3–H4 tetramer chaperone activity (Petesch and Lis, 2012). Lack of all three histone chaperones results in aberrant transcription initiation from cryptic start sites within transcribed coding regions (Bell et al., 2011; Workman, 2009). Several studies suggest roles for FACT and ASF1 in core histone displacement, while SPT6 seems to be required for the assembly of evicted histones enabling chromatin reconstitution after RNA Polymerase II passage (Belotserkovskaya et al., 2003; Schwabish and Struhl, 2006).

## 4. Transcriptional control of gene expression

Control of gene expression, the process of selectively using genetic information, enables different cells, at different times to differentially activate specific sets of genes. This is usually regulated at the level of transcription. In eukaryotes transcription is carried out by three RNA polymerases (Richard and Manley, 2009). RNA polymerase I (RNAPI) transcribes ribosomal RNAs. RNA polymerase III (RNAPIII) transcribes noncoding RNAs such as transfer RNAs, 5S rRNA, and U6 spliceosomal snRNA (Richard and Manley, 2009). Transcription of protein-coding genes is carried out by RNA polymerase II (RNAPII). RNAPII is also responsible for transcribing many noncoding RNAs, including spliceosomal small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), microRNA (miRNA) precursors, and cryptic unstable transcripts (CUTs) (Richard and Manley, 2009). These polymerases are structurally related and share several subunits (Woychik et al., 1990). Human RNAPII is a 550 kDa complex of twelve subunits, Rpb1 to Rpb12 (Cramer et al,. 2001). RNAPII uniquely possesses an unstructured carboxy-terminal domain (CTD) of its larger subunit, Rpb1 (Buratowski, 2009). This domain consists of tandem repeats of the heptameric sequence Tyr1-Ser2-Pro3-Thr4-Ser5-Pyro6-Ser7 (Buratowski, 2009). The CTD is targeted by a wide range of post-translational modifications, of which the best-studied is phosporylation (Buratowski, 2009). The CTD is a key platform for the recruitment of factors involved in transcription, mRNA

processing, and histone modifications (Buratowski, 2009; Kuehner et al., 2011). Different phosphorylation patterns predominate at each stage of transcription providing a mean for RNAPII regulation and coordination (Buratowski, 2005, 2009).

Transcription of protein-coding genes by RNAPII is a repetitive, cyclic process that enables the synthesis of multiple RNA molecules from the same template (Svejstrup, 2004). It consists of three main stages: initiation, elongation and termination. The transcription cycle starts with RNAPII gaining access to the promoter, unwinding DNA and initiating RNA synthesis (Shandilya and Roberts, 2012; Svejstrup, 2004). RNAPII must then get a stable grip on both the template DNA and the growing RNA chain and elongate through the entire length of the gene (Shandilya and Roberts, 2012; Svejstrup, 2004). Finally, RNA synthesis ceases and both the polymerase and the nascent RNA are released from the DNA template (Kuehner et al., 2011; Shandilya and Roberts, 2012). There are specific sequences that signal transcription termination. These sequences constitute the poly(A) signal and result in a characteristic AAUAAA element that emerges in the nascent transcript (Kuehner et al., 2011; Shandilya and Roberts, 2012). The poly(A) signal is recognized by RNA-binding factors prompting changes in the polymerase-associated machinery (Shandilya and Roberts, 2012; Svejstrup, 2004). Ultimately this triggers transcription termination (Kuehner et al., 2011; Richard and Manley, 2009). After being released, RNAPII can then reinitiate a new round of transcription (Shandilya and Roberts, 2012). All stages of the transcription cycle are functionally coupled to a specific event in pre-mRNA processing (5′-capping, splicing and 3′-end formation) and each of them is highly controlled and involves a large number of specific factors (Buratowski, 2005; Moore and Proudfoot, 2009; Proudfoot et al., 2002).

## 5. Interplay between chromatin and transcription

Considerable progress has been made towards the clarification of the mechanistic aspects of transcription through nucleosomal templates. The current view posits that histone chaperones work in concert with chromatin-remodeling and histone-modifying enzymes to allow the progression of RNAPII (Avvakumov et al., 2011). Several mechanisms evict nucleosomal histones in front of elongating RNAPII, while others coordinate the subsequent recovery of the nucleosomes immediately behind transcribing RNAPII (Workman, 2009). Recent studies predict that, in a transcription context, the nucleosome has two distinct components, a interchangeable H2A–H2B dimer and a more stable H3–H4 tetramer (Mellor, 2006). This is because histones H2A-H2B and H3-H4 behave differently during transcription. H3-H4 tetramers are considerably less mobile than the H2A-H2B dimers during transcription, meaning that the interactions of H2A-H2B histones with DNA are disrupted preferentially relative to H3-H4 causing a biased displacement of the H2A-H2B dimers (Mellor, 2006).

Several histone chaperones have been implicated in facilitating transcription. Particularly, FACT and SPT6 have been extensively studied for the assembly and disassembly of core histones within nucleosomes (Buratowski, 2009; Petesch and Lis, 2012; Workman, 2009). FACT tracks with elongating RNAPII and enhances transcription elongation through its interaction with H2A-H2B dimers, destabilizing the nucleosome during polymerase passage (Mason and Struhl, 2003; Orphanides et la., 1999; Saunders et al., 2003). It is also associated with various factors that aid RNAPII elongation (Petesch and Lis, 2012). SPT6 also closely associates with elongating RNAPII (Andrulis et al., 2000). Loss of this chaperone results in a genome-wide reduction in H3 density across many transcriptionally active genes suggesting that SPT6 is required for the assembly of evicted histones (Ivanovska et al., 2011). Nucleosome reassembly after transcription-induced disassembly is critical to maintain the integrity of chromatin structure prohibiting cryptic transcription initiation (Workman, 2009).

This chromatin-mediated regulation of transcription is just beginning to be understood. The mechanistic aspects of this complex process are far from being fully understood and further studies are needed to unravel all its features. For instance, how does RNAPII make its way through the hostile chromatin milieu during transcription and how do selected histone modifications influence each stage of the transcription process (initiation, elongation and termination) are outstanding questions in the field far from being clarified. Numerous examples of the functional coupling between chromatin and transcription have previously been provided for transcription initiation or RNAPII elongation. However, the role of chromatin on transcription termination has not been adequately addressed. Therefore the major goal of this thesis is to understand how does chromatin influence the final stage of RNAPII transcription. We expect the results to reveal fundamental aspects of chromatin dynamics implicated in transcription termination and hopefully help to address fundamental questions in the field.

## Experimental Procedures

### 1. Cell culture and drug treatment

HeLa cells grown as monolayers in Dulbecco's modified Eagle medium - DMEM (Invitrogen), supplemented with 10% (v/v) fetal bovine serum (FBS) and 100U/mL Penicillin-Streptomycin. Drugs were added to HeLa cells in media at final concentrations of 2mM dithiothreitol (DTT, Sigma-Aldrich) and 50µM 5,6-dichloro-1-*β*-D-ribobenzimidazole (DRB, Sigma-Aldrich). Cells were exposed to DTT for 60 minutes and to DRB for 30 minutes. Data from cell cultures without any treatment is shown in all cases.

### 2. RNA isolation and cDNA synthesis

Total cellular RNA was extracted using PureZOL ™. cDNA was made using Transcriptor High Fidelity cDNA Synthesis Kit (Roche) according to the manufacturer's instructions. For mRNA levels quantification, the synthesized cDNAs were analyzed by real-time PCR (see below 2.5 Quantitative real time PCR).

### 3. Chromatin Immunoprecipitation (ChIP)

ChIP was performed as described in de Almeida, SF et al (de Almeida et al., 2011). Briefly, cell extracts were sonicated with a Sanyo Soniprep 150 at an amplitude of 10 microns with six 20 seconds bursts, shearing chromatin into 200–400 bp fragments. The DNA fragments crosslinked to proteins were then immunoprecipitated with specific antibodies and protein A sepharose beads (Sigma). DNA from immunoprecipitated samples was extracted with Chelex 100 (BioRad) as described previously (Nelson et al., 2006). DNA from input samples was extracted with UltraPure Phenol:Chloroform:Isoamyl Alcohol 25:24:1 (Invitrogen). Samples were analyzed by quantitative real-time PCR (see below 2.5 Quantitative real time PCR). The following antibodies were used: anti-Histone H3 (Abcam, ab1791); anti-Histone H2B (Abcam, ab1790), total; anti-Spt16 (Santa Cruz, sc-28734); anti-Spt6 (Abcam, ab49066); anti-RNA Polymerase II, N20 (Santa Cruz, sc-899).

### 4. Micrococcal Nuclease Assay (MNase)

The assay was performed as described in Nature Methods (Nature Publishing Group 2005. Shortly, after cell harvesting by trypsinization, the nuclei were extracted and digested with one unit of Micrococcal Nuclease for 0; 5; 10 or 20 minutes at 28ºC. The reaction was stopped by adding MNase *stop buffer* (100 mM EDTA, 10 mM EGTA pH 7.5). DNA was extracted with UltraPure Phenol:Chloroform:Isoamyl Alcohol 25:24:1 (Invitrogen) and treated

with RNase A (10 mg/ml) for total RNA digestion. DNA quantification was performed using Nanodrop and samples were resolved by gel electrophoresis (1.8% TAE agarose gel). DNA fragments were stained with Gel Red and photographed under a UV light transilluminator (AlphaImager). Samples were further analyzed by quantitative real-time PCR (see below 2.5 Quantitative real time PCR).

**5. Quantitative real time PCR (qRT-PCR)**

qRT-PCR was performed using Real-Time PCR System 7000 (Applied Biosystems) using iTaq$^{TM}$ SYBR Green Supermix with Rox (Bio-Rad). The results were estimated as follows: $2^{Ct\ (Input) - Ct\ (IP)}$. For ChIP samples the Ct (Input) and the Ct (IP) are the threshold cycles of qRT-PCR on DNA samples from input and specific immunoprecipitations, respectively. Input samples were diluted 1:40 for the real time PCR reaction. For MNase samples the Ct (Input) and the Ct (IP) are the threshold cycles of qRT-PCR on DNA samples undigested and digested for 20 minutes, respectively. All real time PCR reactions were performed using 10ng of DNA. For mRNA levels quantification the Ct (Input) and the Ct (IP) are the threshold cycles of qRT-PCR on DNA samples amplified whit an U6 snRNA primer and with a specific primer for each UPR-responsive gene. The cDNA reaction product was diluted 1:15 for the real time PCR reaction. Gene-specific primer pairs used are presented in Annex I.

## Results

### 1. Transcriptional modulation of UPR-responsive genes reveals a dynamic behavior of RNA Polymerase II downstream of poly(A) sites

In order to monitor chromatin dynamics during transcription we took advantage of the gene expression program activated by the unfolded protein response (UPR), a cellular reaction to the flooding of the endoplasmic reticulum (ER) with misfolded proteins (Schröder and Kaufman, 2005). During the URP the transcription of a number of genes is enhanced as a means to restore cellular homeostasis following an insult that disrupts protein folding (Schröder and Kaufman, 2005). Amongst UPR-responsive genes are *CHOP*, *ERP70* and *HERPUD* (Kokame et al., 2000; Oyadomari and Mori, 2004; Schröder and Kaufman, 2005). The UPR can be induced by treatment of cells with the reducing agent dithiothreitol (DTT), which interferes with disulphide bond formation causing an accumulation of improperly folded proteins in the ER and by doing so, triggers an UPR (Costa, 2009). Accordingly, HeLa cells cultured in the presence of DTT during 60 minutes showed a striking increase in the mRNA levels of these genes, which is indicative of augmented transcription rate (Fig. 1).

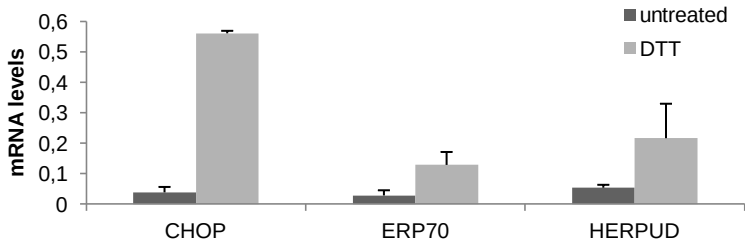



**Figure 1. Transcriptional modulation of UPR-responsive genes**

mRNA levels for *CHOP*, *ERP70* and *HERPUD* in HeLa cells before (untreated, dark grey) and after DTT treatment (light grey). Exposure to DTT causes an upregulation of these genes. The graph depicts the mean for at least three independent experiments. Error bars denote standard deviation.

Since the focus of this thesis is on transcription termination we designed three sets of primers located downstream the poly(A) site of each UPR-responsive gene. Each set differs in the distance from the poly(A) site. The primers were labeled Poly(A) 1, 2 and 3 and are located between 0.1 - 0.6 kilobases, 1.3 – 1.7 kilobases and 2.7 – 3.1 kilobases downstream the poly(A) site, respectively. For a more general outlook of our results we also designed a primer in the promoter and coding regions of each gene. This allowed us to inspect particular features of the regions downstream the poly(A) site.

To measure RNAPII occupancy along the UPR-responsive genes, we performed ChIP before and after transcriptional activation with DTT. Untreated cells displayed low levels of RNAPII on the promoter region, throughout the gene and downstream of the poly(A) site

(Fig. 2). The reduced transcription rate is consistent with the low mRNA levels of these genes (Fig. 1). Upon DTT treatment, RNAPII occupancy increased throughout the genes (Fig. 2). This increase was more striking on the promoter region (Fig. 2). These results are in accordance with a transcriptional upregulation following DTT treatment (Fig. 1). Notably, this increase in RNAPII occupancy was also observed on the 3' flanking regions of the analyzed genes, spanning up to ~3 kilobases downstream of poly(A) sites (Fig. 2).

To further investigate the chromatin alterations that shepherd transcription inhibition, DTT-treated cells were further incubated in the presence of DRB. This compound renders RNAPII incompetent to escape the promoter-proximal pause and resume elongation (Yamaguchi et al., 1998). Accordingly, addition of DRB to HeLa cells that had been cultured for one hour in the presence of DTT led to a dramatic accumulation of RNAPII on the promoter region, where it became arrested (Fig. 2). In contrast, downstream from the promoter region, RNAPII occupancy decreased significantly and it could only be detected at low levels (Fig. 2). Notably, the accumulation of RNAPII past the poly(A) site observed on transcriptionally active genes was reduced to marginal levels following treatment with DRB (Fig. 2). These findings reveal a dynamic behavior of RNAPII molecules at the region downstream the 3' boundary of the genes that is related to the transcription rate.

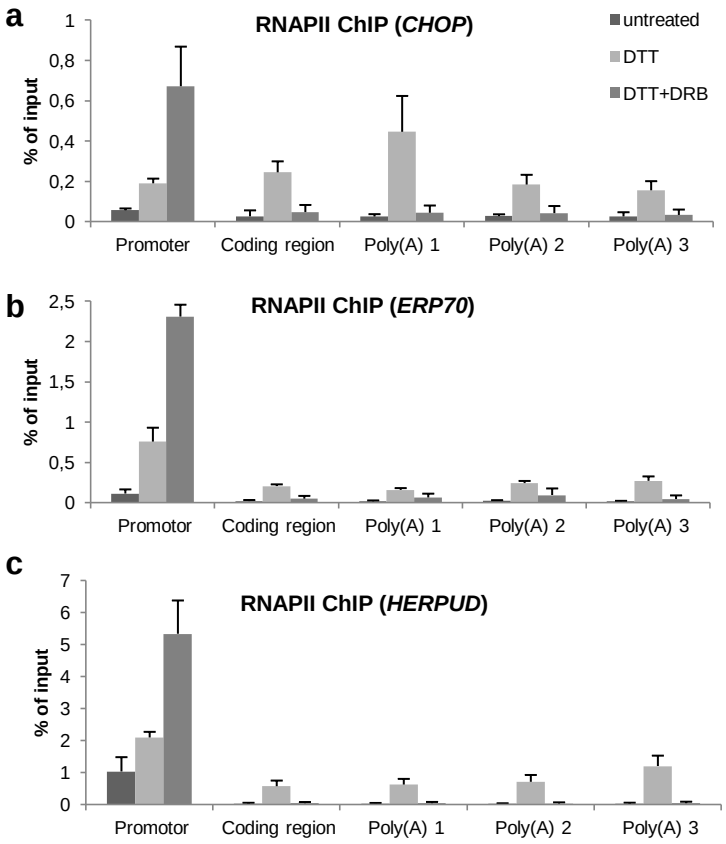



**Figure 2. RNAPII levels downstream poly(A) sites are transcriptionally related**

ChIP of RNAPII in HeLa cells along the genes *CHOP* **(a)**, *ERP70* **(b)** and *HERPUD* **(c)** before (untreated, dark grey) and after transcription upregulation (DTT, light grey) and after transcription inhibition (DTT+DRB, grey). Note the DTT-induced increase in RNAPII levels downstream the poly(A) sites. The y axis represents the percent of input DNA that was recovered by ChIP. All graphs depict the mean for at least three independent experiments. Error bars denote standard deviation.

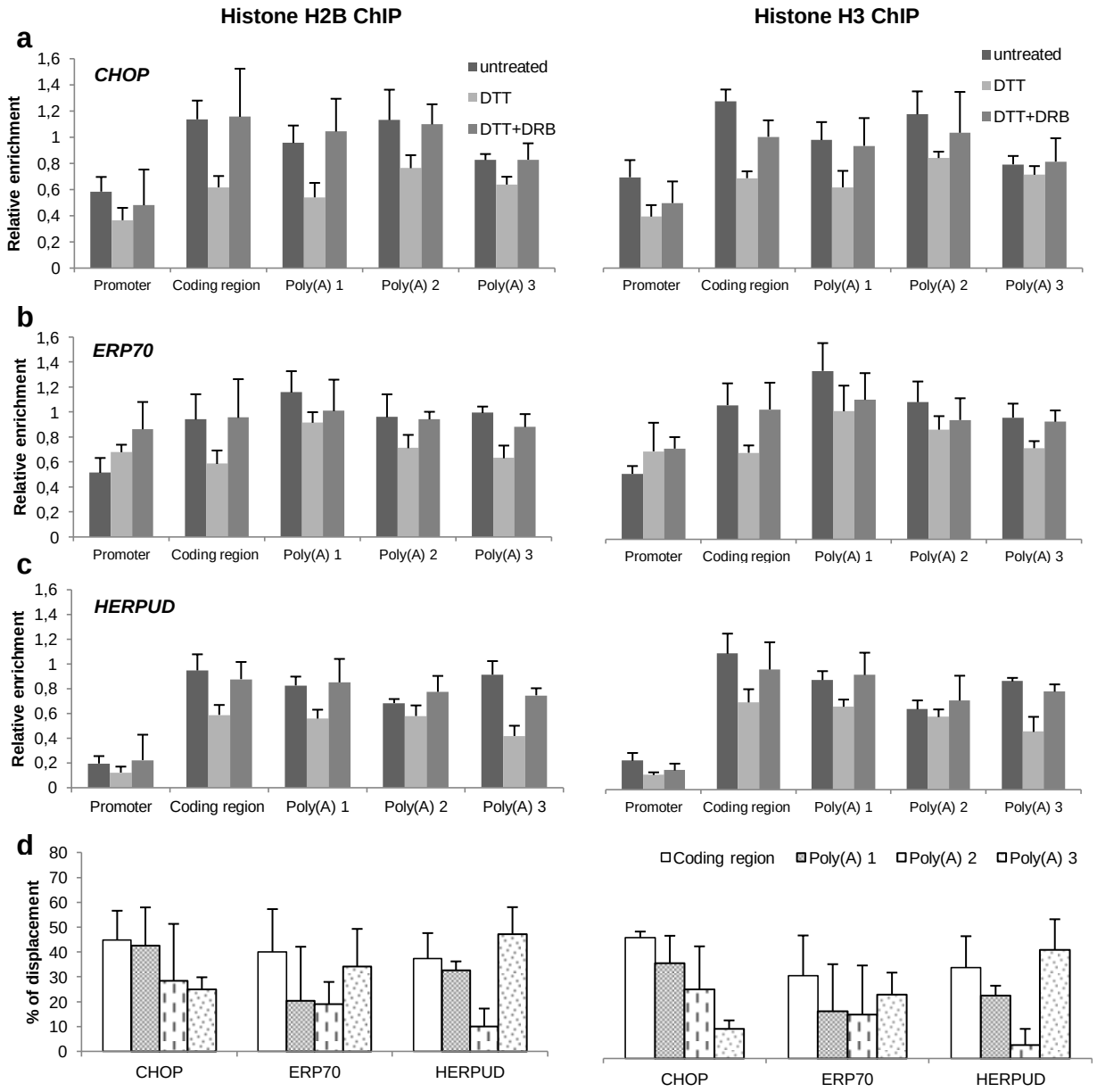



**Figure 3. Nucleosomal histones exhibit reduced displacement downstream of poly(A) sites**

**(a-c)** Histone density across UPR-responsive genes detected by ChIP in HeLa cells before (untreated, dark grey) and after transcription upregulation (DTT, light grey) and after transcription inhibition (DTT+DRB, grey). Histone H2B **(left side)** and Histone H3 **(right side)** ChIP signals along the genes *CHOP* **(a)**, *ERP70* **(b)** and *HERPUD* **(c)** are genes are shown as percentage of immunoprecipitated DNA relatively to a non-transcribed intergenic region. **(d)** The percentage of histone displacement along *CHOP*, *ERP70* and *HERPUD* genes after DTT treatment was obtained using the formula *%displacement=100-{(histone ChIP*$_{after\ DTT}$ */ histone ChIP*$_{non\text{-}treated\ cells}$*)x100}*. Histone displacement downstream the poly(A) site exhibits lower levels when compared with the coding region of genes. All graphs depict the mean for at least three independent experiments. Error bars denote standard deviation.

## 2. Genomic regions downstream of poly(A) sites display reduced histone displacement and impaired nucleosome eviction

Nucleosomes pose a barrier for RNAPII elongation during transcription. To overcome such barrier, RNAPII relies on a number of tools that promote the displacement of nucleosomes and facilitate the progression of the transcription complexes along the gene templates (Petesch and Lis, 2012). On the other hand, nucleosome reassembly on the wake of RNAPII elongation is mandatory in order to maintain chromatin structure and prevent spurious transcription initiation (Németh and Langst, 2004). Importantly, nucleosomes may also play a role in transcription termination by serving as speed bumps that slow down RNAPII preventing transcription read-through and instigating the mechanistic events that lead to RNAPII dissociation from the DNA template. To investigate these hypotheses the levels of histone H3 and histone H2B on UPR-responsive gene templates were measured by ChIP. Following transcriptional activation with DTT, the levels of histone H3 decreased when compared to the amount recovered prior to the drug addition (Fig. 3a, b, c, right side). Notably, this reduction was not consistent throughout the entire gene length, since histone displacement decreased towards the 3' flanking region of genes (Fig. 3, right side). Interestingly, nucleosomes located after the poly(A) site displayed reduced disassembly of their core histones when compared to the coding region of genes as revealed by the levels of histone H3 in DTT treated and untreated cells (Fig. 3d). A similar result was observed for histone H2B (Fig. 3 left side). After DRB treatment the levels of both histones increased throughout the entire gene length to the levels observed in non-treated cells (Fig. 3a, b and c) in agreement with the view that nucleosomes are evicted during RNAPII elongation.

To further investigate the nucleosome occupancy in response to transcription rate modulation we performed a micrococcal nuclease (MNase) assay. This enzyme digests all DNA except the sequences protected by the nucleosomes. MNase has the ability to digest the linker DNA between nucleosomes yielding DNA fragments of approximately 147 base pairs (that correspond to the DNA comprised in a single nucleosome) or a multiple number (which accounts for oligonucleosomes). The separation of the fragments partition by gel electrophoresis after MNase digestion generates a characteristic "ladder" pattern visible in Figure 4a. Quantification of the DNA sequences that resist MNase digestion by real-time quantitative PCR reveals the nucleosome occupancy of a given gene segment. Following a 20 minutes MNase digestion, the amount of DNA recovered in the coding regions of *CHOP*, *ERP70* and *HERPUD* upon DTT treatment was significantly lower than that recovered from the untreated cells (Fig. 4b, c and d). This result is consistent with increased nucleosome displacement by the transcription machinery and progressively vanished as we walked into the genomic regions located after the poly(A) sites (Fig. 4b, c and d). DRB treatment caused

nucleosome occupancy to rise back to levels comparable to those off untreated cells cultured throughout the entire gene length. Surprisingly, we observed an increased nucleosome occupancy in the past-poly(A) regions of the analyzed genes after DRB addition, which was particularly noticeable on the poly(A) 3 amplicon of *HERPUD* (Fig. 4b, c and d). This result most likely reflects the formation of high order chromatin structures – which strongly protect the DNA from MNase due to the massive genome-wide inhibitory effect of DRB.

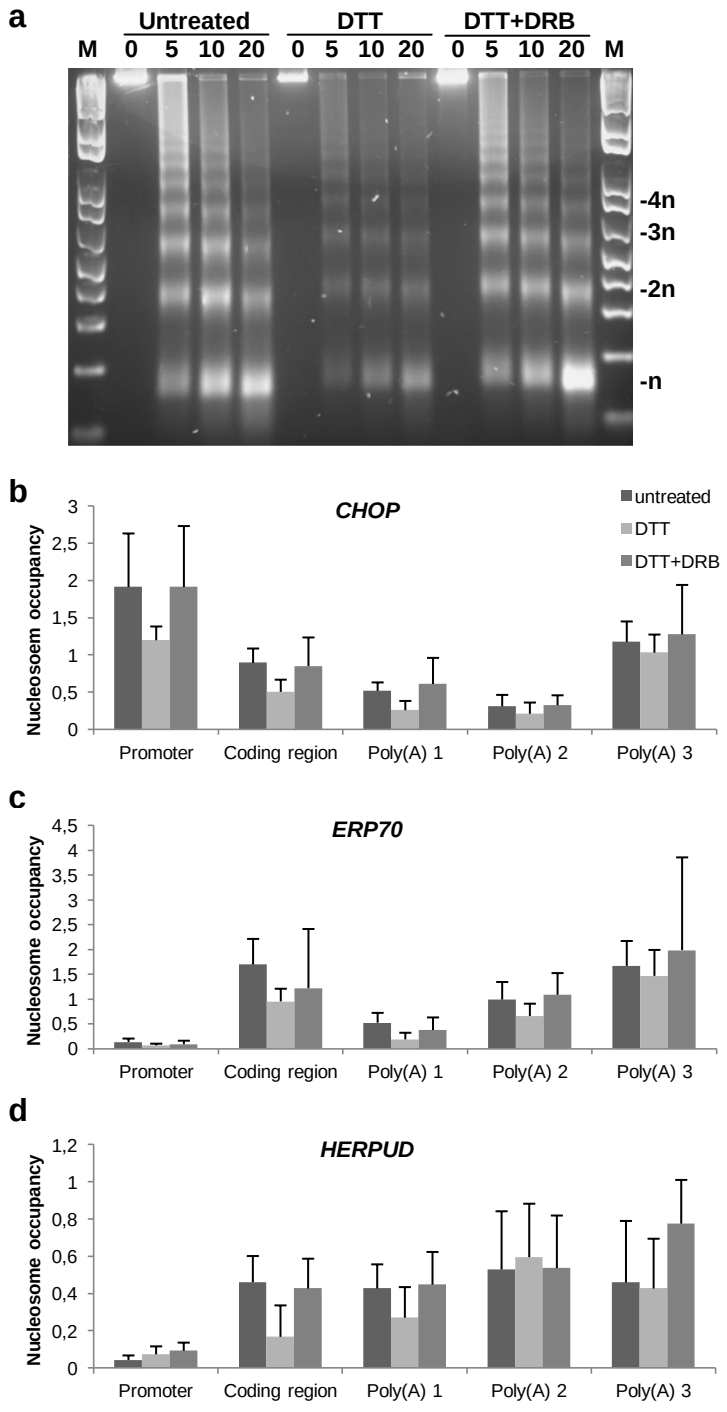



**Figure 4. Nucleosomes resist to disassembly downstream of poly(A) sites**

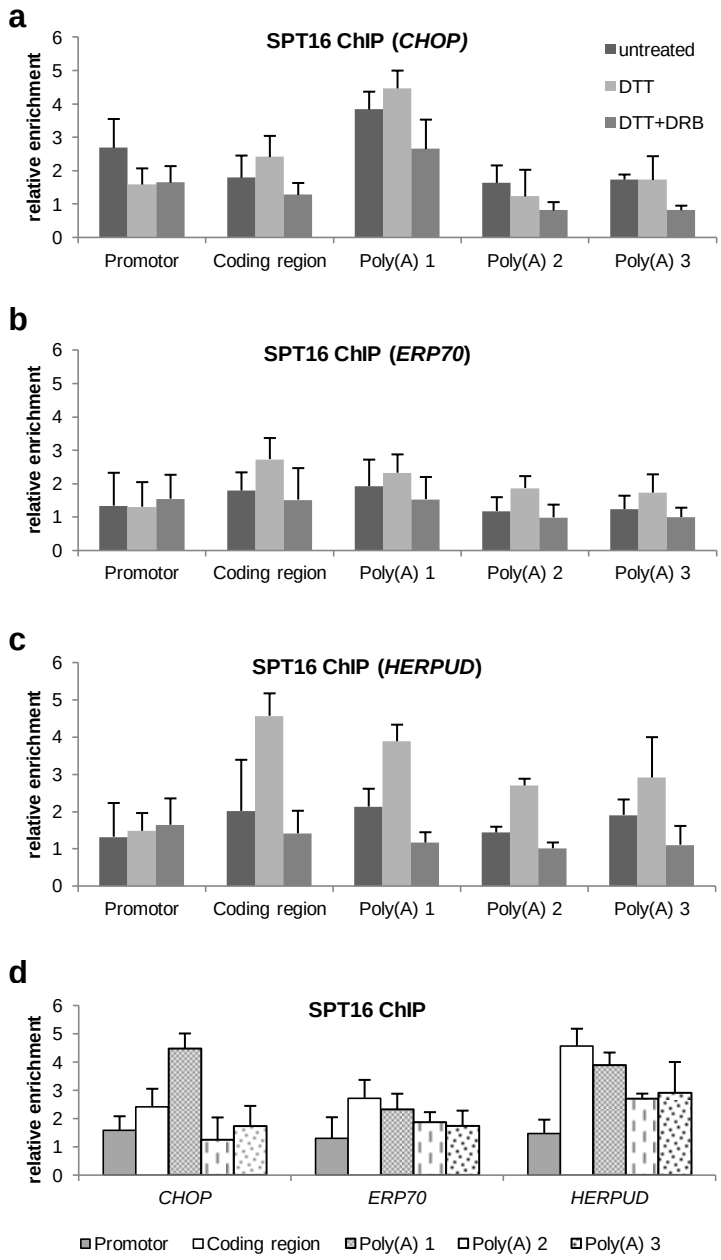
**(a)** Gel electrophoresis of a MNase assay performed in HeLa cells exposed to DTT or DTT+DRB. Nuclei were digested with one unit of MNase for 0; 5; 10 or 20 minutes. At the right is indicated the position of the DNA fragments assembled on one or more nucleosomes (1=n; 2=2n; X=Xn, where n indicates approximately 147 bp of DNA). "M" denotes the lanes containing the DNA ladder. **(b-d)** Nucleosome occupancy profile along *CHOP* **(b),** *ERP70* **(c)** and *HERPUD* **(d)** in HeLa cells before (untreated, dark grey) and after transcription activation (DTT, light grey) and after transcription inhibition (DTT+DRB, grey). The nucleosome occupancy was determined by normalizing the amount of the MNase digested DNA to the amount of undigested DNA (y axis). All graphs depict the mean for at least three independent experiments. Error bars denote standard deviation.

### 3. FACT recruitment decreases downstream of poly(A) sites

Comparison of the RNAPII profile (Fig. 2) and nucleosome occupancy at the 3' flanking region (Fig. 3 and Fig. 4) suggests that after the poly(A) site nucleosomes are not efficiently negotiated by the elongating RNAPII. Since histone chaperones play a major role in releasing RNAPII from nucleosomal constraint, the distribution of SPT16 (a subunit of the FACT complex) and SPT6 was assessed by ChIP (Fig. 5). FACT complex and SPT6 are two histone chaperones whose activity promotes histone displacement from nucleosomes and subsequent histone deposition and nucleosome reassembly in the wake of RNAPII elongation (Németh and Langst, 2004; Petesch and Lis, 2012). Upon transcriptional activation of the UPR-responsive genes the levels of SPT16 bound to chromatin templates increased significantly on the coding regions (Fig. 5a, b and c). Notably, this recruitment was considerably lower on regions downstream of the poly(A) site (Fig. 5d). After transcription inhibition with DRB, SPT16 levels decreased throughout the entire gene length to values comparable to the untreated condition (Fig. 5a, b and c). Intriguingly, SPT6 did not show a significant enrichment in none of the experimental conditions tested when compared to the intergenic control region (Fig. 5e). In contrast, we could observe the recruitment of SPT6 to another (non UPR-related) gene (Fig. 5e). This result provides evidence that this histone chaperone does not participate in the co-transcriptional remodeling of the chromatin templates of the studied genes. Altogether, these data suggest that the mechanism of FACT recruitment that operates along the coding region is interrupted at the poly(A) site.

### 4. H3K4me, H3K4me2 and H3K9me histone marks are enriched at the 3' flanking region of genes

Previous studies revealed that the recruitment of histone chaperones is in part driven by post translational modification of histones (Kouzarides, 2007). Recently, our lab performed a bioinformatics analysis of the genome-wide distribution of different histone marks and identified a specific enrichment of H3K4me, H3K4me2 and H3K9me at the 3' flanking region of active protein-coding genes (data not shown). This enrichment in the terminal region of genes could implicate an involvement of these modifications on transcription termination. In order to validate this observation the distribution of these modifications along the UPR-responsive genes was assessed by ChIP after transcriptional activation (Fig. 6). In agreement with the genome-wide data, *CHOP, HERPUD* and *ERP70* displayed higher levels of the three histone modifications in the 3' flanking region when compared with the coding region. This result paves the way for the investigation of a novel role of these histone marks in the mechanism of recruitment of histone chaperones after the poly(A) site, which could be directly implicated on transcription termination.



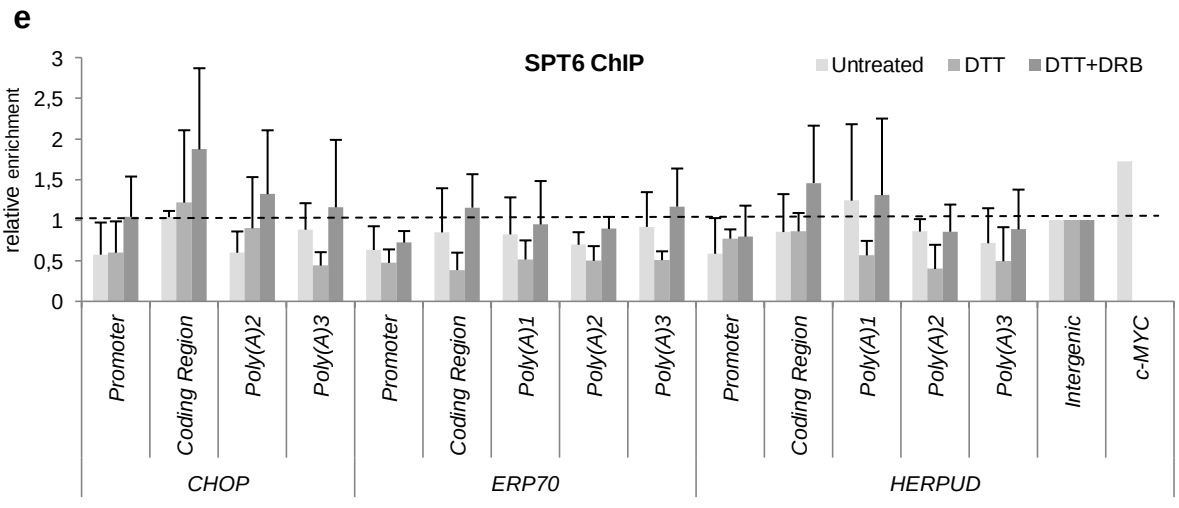




**Figure 5. FACT is inefficiently recruited to genomic regions downstream of poly(A) site**

**(a-c)** ChIP of SPT16, a FACT subunit, along *CHOP* **(a)**, *ERP70* **(b)** and *HERPUD* **(c)** in HeLa cells before (untreated, dark grey) and after transcription stimulation (DTT, light grey) and after transcription inhibition (DTT+DRB, grey). **(d)** ChIP of SPT16 along *CHOP*, *ERP70* and *HERPUD* genes in HeLa cells treated with DTT. The y axis represents the results expressed as "relative enrichment", calculated as the enrichment of SPT16 relatively to the levels detected in a non-transcribed intergenic region. **(e)** ChIP signal of SPT6 in HeLa cells before (untreated, dark grey) and after transcription upregulation (DTT, light grey) and after transcription inhibition (DTT+DRB, grey). The y axis represents the percent of input material immunoprecipitated. All graphs depict the mean for at least three independent experiments. Error bars denote standard deviation.

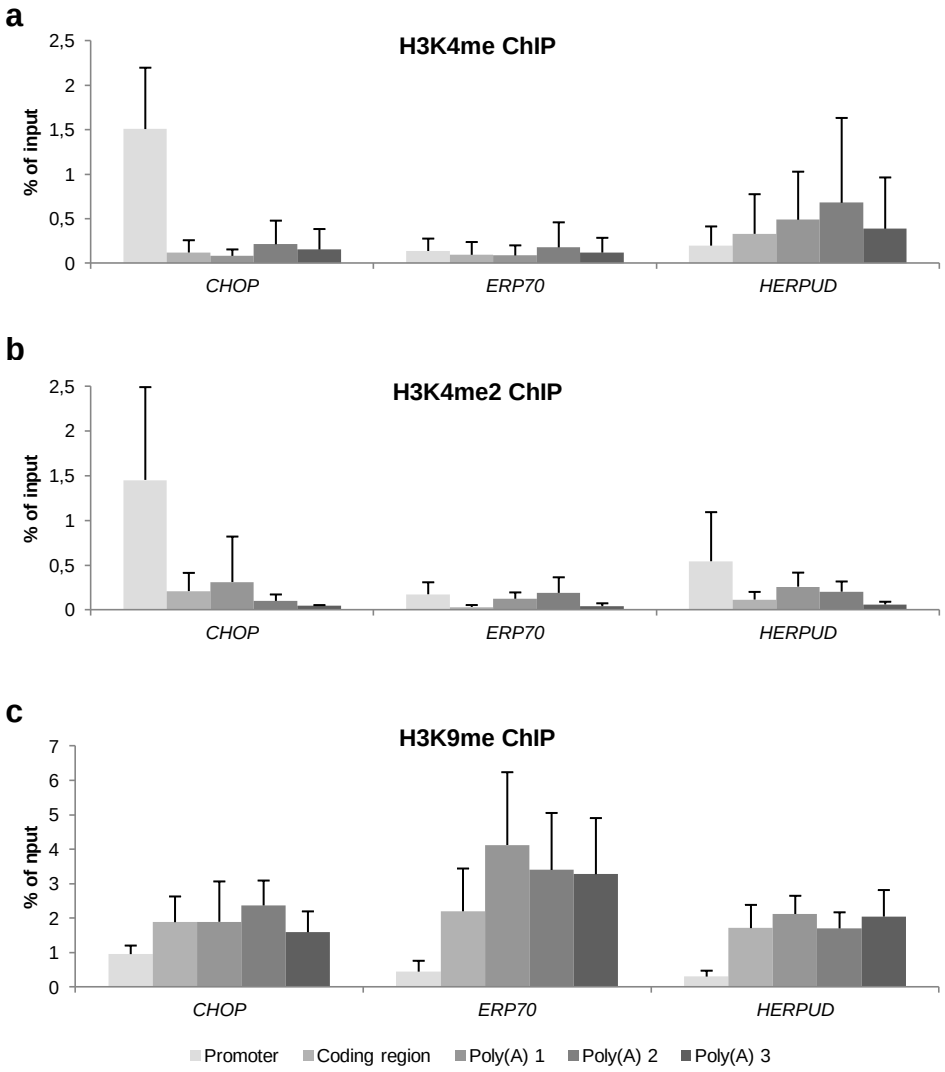



**Figure 6. Histone marks enrichment on regions downstream the poly(A) site**

ChIP of histone marks H3K4me **(a)**, H3K4me2 **(b)** and H3K9me **(c)** in HeLa cells treated with DTT along *CHOP*, *ERP70* and *HERPUD*. These histone marks are enriched preferentially on regions downstream the poly(A) site. The y axis represents the percent of input material immunoprecipitated. All graphs depict the mean for at least three independent experiments. Error bars denote standard deviation.

## Discussion

By taking advantage of UPR-responsive genes we were able to set an experimental model system to track RNAPII density and map changes in the chromatin structure upon transcriptional activation and inhibition. Our ChIP analysis revealed a transcription-induced accumulation of RNAPII at regions up to three kilobases downstream the poly(A) site (Fig. 2). This accumulation suggests pausing of RNAPII, a widespread phenomenon that can be observed in different regions of the transcribed DNA template (Oesterreich et al., 2011; Glover-cutter et al., 2008; Levine, 2011).For instance, pausing in the promoter-proximal region is important for coupling transcription and mRNA processing and is also involved in regulating the transition from initiation to elongation (Core and Lis, 2008). Promoter-proximal pausing is mediated by negative elongation factor (NELF) and DRB sensitivity-inducing factor (DSIF) (Core and Lis, 2008). Pausing downstream the poly(A) site has also been extensively reported and several studies have shown its role in transcription termination (Glover-cutter et al., 2008; Gromak, West, and Proudfoot, 2006; Kuehner et al., 2011). A recent study suggested that the presence of nucleosomes after the poly(A) site induces pausing of polymerase (Grosso et al., 2012). In support of this, our results reveal that nucleosome occupancy on *CHOP*, *ERP70* and *HERPUD* correlates with the RNAPII profile. Upon transcriptional upregulation nucleosome occupancy is higher at the distal 3' flanking region of genes (Fig. 4). This region also displays impaired histone eviction (Fig. 4) as confirmed by the diminished levels of histone displacement observed in regions downstream of the poly(A) site (Fig 3). Given that the activity of histone chaperones is critical for the eviction of nucleosomal histones (Avvakumov et al., 2011; Park and Luger, 2008; Ãkey and Luger, 2003) we hypothesized that altered nucleosome dynamics observed downstream the poly(A) site relates to unproductive recruitment of histone chaperones. Concordantly, our results revealed reduced recruitment of FACT, a major player in transcription regulation (Belotserkovskaya et al., 2003; Petesch and Lis, 2012; Winkler and Luger, 2011), to the 3' flanking region of genes (Fig. 5). Notably, in agreement with unpublished analysis of genome-wide data from our lab we verified an enrichment of H3K4me, H3K4me2 and H3K9me on the genomic regions downstream the poly(A) site (Fig. 6), which may interfere with the recruitment of histone chaperones such as FACT.

Altogether, the findings reported on this thesis disclose a novel mechanism of transcription arrest acting at the 3' flanking region, which is where termination takes place. At this region, nucleosomes stall RNAPII because transcription complexes are not able to efficiently recruit histone chaperones after the poly(A) site. It is possible to envisage a model by which pausing of RNAPII facilitates the molecular events that culminate in transcription termination, preventing unproductive transcription read-through. The finding of histone

modifications specifically enriched at the 3’ flanking region discloses a role for these marks on nucleosome eviction and fuels the interest for additional research aimed at investigating the contribution of these epigenetic traits to transcription termination. Taken together, our results provide novel insights to transcription termination, a fundamental process that remains one of the least understood stages of the transcription cycle.

**Acknowledgments**

This work is the result of a very enjoyable period at Instituto de Medicina Molecular in the group of Prof. Maria do Carmo-Fonseca, to whom I would like to express my gratitude for providing me the opportunity to carry out this work in her lab.

I am most grateful to Dr. Sérgio Fernandes de Almeida, for giving me a challenging and most exciting project to work on. His constant patience, encouragement, support, and invaluable advices and insight made this work successful and an extraordinary experience. He was an outstanding supervisor and has set an example I hope to match some day. I would also like to thank his time and effort in reviewing this work. His corrections and input on scientific matters and on the English improved this thesis very much. Thank you for everything.

I am most indebted to Ana Raposo and Sílvia Carvalho who made the lab an excellent and friendly environment for working. They were the ones who endure me every day and made the many hours spent in the lab noteworthy. Thank you for your great support in all my struggles and for being great friends throughout my stay at IMM. I wish you both the best luck with your future careers.

I would also like to thank Sílvia Lourenço who was an incredible friend through this year.

Special thanks to Sérgio Marinho for getting me started with biochemical labwork and for his friendship.

Many thanks to Ana Rita Grosso for performing the bioinformatics data analysis.

Moreover, I am thankful for both scientific and social support from all the members of the UBCe. Special thanks to Ram.

I would also like to thank Prof. Júlio Duarte who willingly accepted to be my supervisor and addressed all thesis related issues most expeditiously.

I must also acknowledge Fundação para a Ciência e Tecnologia (FCT) which made this research possible through its financial support.

To my friends,
For all the craziness, for all the stupidities, for all the nights we didn't sleep =)

To my parents,
They are the ones who deserve my deepest gratefulness. They supported me mentally, materially and cake-wise and thus contributed to everything I have reached so far. They always believed in me. I doubt that I will ever be able to convey my appreciation fully.

**References**


Akey, C. W., & Luger, K. (2003). Histone chaperones and nucleosome assembly. Current Opinion in Structural Biology, 13, 6–14.

Andrulis, E. D., Guzma, E., Do, P., Werner, J., & Lis, J. T. (2000). High-resolution localization of Drosophila Spt5 and Spt6 at heat shock genes in vivo : roles in promoter proximal pausing and transcription elongation, 14, 2635–2649.

Avvakumov, N., Nourani, A., & Co, J. (2011). Histone Chaperones : Modulators of Chromatin Marks. Molecular Cell, 41, 502-513.

Bai, L., & Morozov, A. V. (2010). Gene regulation by nucleosome positioning. Trends in Genetics, 26, 476–483.

Bannister, A. J., & Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell research, 21, 381–95.

Bell, O., Tiwari, V. K., Thomä, N. H., & Schübeler, D. (2011). Determinants and dynamics of genome accessibility. Nature Reviews Genetics, 12, 554–564.

Belmont, A. S., and Bruce, K., (1994). Visualization of G1 Chromosomes : A Folded, Twisted, Supercoiled Chromonema Model of Interphase Chromatid Structure. The Journal of Cell Biology, 127, 287–302.

Belmont, A. S., Sedat, J. W. and Agard, D. A. (1987). A three-dimensional approach to mitotic chromosome structure: evidence for a complex hierarchical organization. The Journal of Cell Biology, 127, 287–302, Vol. 105, 77–92.

Belotserkovskaya, R., Oh, S., Bondarenko, V. A., Orphanides, G., Studitsky, V. M., & Reinberg, D. (2003). FACT Facilitates Transcription-Dependent Nucleosome Alteration. Science, 301, 1090–1093.

Berger, S. L. (2007). The complex language of chromatin regulation during transcription. Nature, 447, 407–412.

Bird, A. (2007). Perceptions of epigenetics. Nature, 447, 396–398.

Bottomley, M. J. (2004). Structures of protein domains that create or recognize histone modifications. EMBO reports, 5, 464–9.

Buratowski, S. (2005). Connections between mRNA 3' end processing and transcription termination. Current opinion in cell biology, 17, 257–61.

Buratowski, S. (2009). Progression through the RNA polymerase II CTD cycle. Molecular cell, 36, 541–6.

Carrillo Oesterreich, F., Bieberstein, N., & Neugebauer, K. M. (2011). Pause locally, splice globally. Trends in cell biology, 21, 328–35.

Chakravarthy, S., Park, Y., Chodaparambil, J., Edayathumangalam, R. S., & Luger, K. (2005). Structure and dynamic properties of nucleosome core particles. FEBS Letters, 579, 895–898.

Core, L. J., & Lis, J. T. (2008). Transcription Regulation Through Promoter-Proximal Pausing of RNA Polymerase II. Science, 319, 1791–1792.

Costa, V. M. (2009). ER Stress-Inducible Factor CHOP Affects the Expression of Hepcidin by Modulating C / EBPalpha Activity. PLos One, 4, e6618.

Cramer, P., Bushnell, D. a, & Kornberg, R. D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science, 292, 1863–76.

de Almeida, S. F., Grosso, A. R., Koch, F., Fenouil, R., Carvalho, S., Andrade, J., Levezinho, H., et al. (2011). Splicing enhances recruitment of methyltransferase HYPB / Setd2 and methylation of histone H3 Lys36. Nature Structural & Molecular Biology, 18, 977–83.

Deal, R. B. et al. (2010). Genome-Wide Kinetics of Nucleosome Turnover Determined by Metabolic Labeling of Histones, Science, 328, 1161–1164.

Flaus, A., & Owen-hughes, T. (2001). Mechanisms for ATP-dependent chromatin remodelling. Current Opinion in Genetics & Development, 11, 148–154.

Glover-cutter, K., Kim, S., Espinosa, J., & Bentley, D. L. (2008). RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. Nature Structural & Molecular Biology, 15, 71–78.

Gromak, N., West, S., & Proudfoot, N. J. (2006). Pause Sites Promote Transcriptional Termination of Mammalian RNA Polymerase II. Molecular and Cellular Biology, 26, 3986–3996.

Grosso, A.R., et al,. (2012). Dynamic transitions in RNA polymerase II density profiles during transcription termination, Genom. Res., 22, 1477-1456

Hamilton, C., Hayward, R. L., & Gilbert, N. (2011). Global chromatin fibre compaction in response to DNA damage. Biochemical and biophysical research communications, 414, 820–5.

Hargreaves, D. C., & Crabtree, G. R. (2011). ATP-dependent chromatin remodeling : genetics , genomics and mechanisms. Cell Research, 21, 396–420.

Hondele, M., & Ladurner, A. G. (2011). The chaperone – histone partnership : for the greater good of histone traffic and chromatin plasticity. Current Opinion in Structural Biology, 21, 698–708.

Horn, P. J., & Peterson, C. L. (2002). Chromatin Higher Order Folding : Wrapping up Transcription. Science, 297, 1824–1827.

Ivanovska, I., Jacques, P.-É., Rando, O. J., Robert, F., & Winston, F. (2011). Control of chromatin structure by spt6: different consequences in coding and regulatory regions. Molecular and cellular biology, 31, 531–41.

Jenuwein, T., & Allis, C. D. (2001). Translating the Histone Code. Science, 293, 1074–1080.

Kamakaka, R. T., & Biggins, S. (2005). Histone variants: deviants? Genes & development, 19, 295–310.

Kokame, K., Agarwala, K. L., Kato, H., & Miyata, T. (2000). Herp, a new ubiquitin-like membrane protein induced by endoplasmic reticulum stress. The Journal of biological chemistry, 275, 32846–53.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell, 128, 693–705.

Kuehner, J. N., Pearson, E. L., & Moore, C. (2011). Unravelling the means to an end: RNA polymerase II transcription termination. Nature reviews. Molecular cell biology, 12, 283–94.

Levine, M. (2011). Paused RNA Polymerase II as a Developmental Checkpoint. Cell, 145, 502–511.

Li, G., & Reinberg, D. (2011). Chromatin higher-order structures and gene regulation. Current Opinion in Genetics & Development, 21, 175–186.

Luco, R. F., Allo, M., Schor, I. E., Kornblihtt, A. R., & Misteli, T. (2011). Epigenetics in alternative pre-mRNA splicing. Cell, 144, 16–26.

Luger, K. (2003). Structure and dynamic behavior of nucleosomes Karolin Luger. Current Opinion in Genetics & Development, 127–135.

Lusser, A., & Kadonaga, J. T. (2003). Chromatin remodeling by ATP-dependent molecular machines. BioEssays, 25.12, 1192–1200.

Mardsen, M., P., F., and Laemmli, U., K., (1979). Metaphase chromosome structure: Evidence for a radial loop model. Cell, 17, 849-858

Mariño-ramírez, L., Kann, M. G., & Shoemaker, B. A. (2005). Histone structure and nucleosome stability. Exp. Rev. Proteomics, 5, 719–729.

Mason, P. B., & Struhl, K. (2003). The FACT Complex Travels with Elongating RNA Polymerase II and Is Important for the Fidelity of Transcriptional Initiation In Vivo, Molecular and Cellular Biology 23, 8323–8333.

Mellor, J. (2006). Dynamic nucleosomes and gene transcription. Trends in genetics, 22, 320-329.

Moore, M. J., & Proudfoot, N. J. (2009). Review Pre-mRNA Processing Reaches Back to Transcription and Ahead to Translation. Cell, 136, 688-700.

Nelson, J. D., Denisenko, O., Sova, P., & Bomsztyk, K. (2006). Fast chromatin immunoprecipitation assay. Nucleic acids research, 34, e2.

Németh, A., & Langst, G. (2004). Chromatin higher order structure : Opening up chromatin for transcription. Briefings in Functional Genomics and Proteomics, 2, 334–343.

Orphanides, G, Wu, W. H., Lane, W. S., Hampsey, M., & Reinberg, D. (1999). The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. Nature, 400, 284–8.

Orphanides, George, Leroy, G., Chang, C., Luse, D. S., & Reinberg, D. (1998). FACT , a Factor that Facilitates Transcript Elongation through Nucleosomes. Cell, 92, 105–116.

Oyadomari, S., & Mori, M. (2004). Roles of CHOP/GADD153 in endoplasmic reticulum stress. Cell death and differentiation, 11, 381–9.

Park, Y., & Luger, K. (2008). Histone chaperones in nucleosome eviction and histone exchange. Current Opinion in Structural Biology, 18, 282–289.

Petesch, S. J., & Lis, J. T. (2012). Overcoming the nucleosome barrier during transcript elongation. Trends in Genetics, 28, 285–294.

Proudfoot, N. J., Furger, A., Dye, M. J., & William, S. (2002). Integrating mRNA Processing with Transcription. Cell, 108, 501–512.

Ransom, M., Williams, S. K., Dechassa, M. L., Das, C., Linger, J., Adkins, M., Liu, C., et al. (2009). FACT and the Proteasome Promote Promoter Chromatin Disassembly and Transcriptional Initiation *. Journal of Biological Chemistry, 284, 23461–23471.

Richard, P., & Manley, J. L. (2009). Transcription termination by nuclear RNA polymerases Transcription termination by nuclear RNA polymerases. Genes & Development, 1247–1269.

Richmond, R. K., Sargent, D. F., Richmond, T. J., Luger, K., & Ma, A. W. (1997). Crystal structure of the nucleosome core particle resolution at 2 .8 A. Nature, 389, 251-260.

Saunders, A., Werner, J., Andrulis, E. D., Nakayama, T., Hirose, S., Reinberg, D., & Lis, J. T. (2003). Tracking FACT and the RNA polymerase II elongation complex through chromatin in vivo. Science, 301, 1094–6.

Schröder, M., & Kaufman, R. J. (2005). The mammalian unfolded protein response. Annual review of biochemistry, 74, 739–89.

Schwabish, M. A., & Struhl, K. (2006). Asf1 Mediates Histone Eviction and Deposition during Elongation by RNA Polymerase II. Molecular Cell, 22, 415–422.

Shandilya, J., & Roberts, S. G. E. (2012). The transcription cycle in eukaryotes: From productive initiation to RNA polymerase II recycling. Biochimica et biophysica acta, 1819, 391–400.

Strahl, B. D., & Allis, C. D. (2000). The language of covalent histone modifications, 403, 41–45.

Svejstrup, J. Q. (2004). The RNA polymerase II transcription cycle : cycling through chromatin. Biochimica and Biophysica acta, 1677, 64 – 73.

Tolkunov, D., Zawadzki, K. A., Singer, C., Elfving, N., Morozov, A. V., & Chang, F. (2009). Chromatin remodelers clear nucleosomes from intrinsically unfavorable sites to establish nucleosome-depleted regions at promoters. Molecular Biology of the Cell, 22, 2106-2118.

Winkler, D. D., & Luger, K. (2011). The Histone Chaperone FACT : Structural Insights and Mechanisms for Nucleosome. Journal of Biological Chemistry, 286, 369–374.

Workman, J. L. (2009). Nucleosome displacement in transcription. Genes & Development, 20, 2009–2017.

Woychik, N. a, Liao, S. M., Kolodziej, P. a, & Young, R. a. (1990). Subunits shared by eukaryotic nuclear RNA polymerases. Genes & development, 4, 313–23.

Yamaguchi, Y., Wada, T., & Handa, H. (1998). Interplay between positive and negative elongation factors: drawing a new view of DRB. Genes to cells, 3, 9–15.

Zhou, V. W., Goren, A., & Bernstein, B. E. (2011). Charting histone modifications and the functional organization of mammalian genomes. Nature reviews. Genetics, 12, 7–18.

## Annex I

| Gene | Primer designation | Primer sequence (5'-3') |
|---|---|---|
| ***CHOP*** | Promotor Forward | GTGGGGGTAAAACGGCGGGT |
| | Promotor Reverse | GGTCGCCCCTAGTCGGTCGT |
| | Coding Region Forward * | TCGAGCGCCTGACCAGGGAA |
| | Coding Region Reverse * | TCCAAGCCTTCCCCCTGCGT |
| | Poly(A) 1 Forward | CTGCTGGCTTCGGGGACGTT |
| | Poly(A) 1 Reverse | TTTGGCCCTGCCGCTTCCTC |
| | Poly(A) 2 Forward | GTTCCTGCCCGTTGCCTGAGG |
| | Poly(A) 2 Reverse | TCCGGGATGCCTTGCGCAGT |
| | Poly(A) 3 Forward | GGGTCCAGGGCTCAGAGAGTGT |
| | Poly(A) 3 Reverse | TGGTGCCTATCTCAGCTCTTCTGC |
| ***ERP70*** | Promotor Forward | CCCAGGCTCCGCCTCTCCTGC |
| | Promotor Reverse | CTCCGAGCCCTAACGTGAGGTGCC |
| | Exon1 Forward * | TGCCGGCGTCAGTCTGGGAT |
| | Exon1 Reverse * | CGGGGGAGCCGGAAAAACCC |
| | Coding Region Forward | TGGGCCCCCTCACCTGTTCC |
| | Coding Region Reverse | ACCAGGGGCAGGGCGTACTT |
| | Poly(A) 1 Forward | CATCACCGTCCTCACCCCGC |
| | Poly(A) 1 Reverse | AATGCCTCCTGCCCCCACCA |
| | Poly(A) 2 Forward | CCTGCAAAAGGCGGGACCACT |
| | Poly(A) 2 Reverse | CACCGATCACGCTGGGCTGT |
| | Poly(A) 3 Forward | CCAGGAGCCCCATGGACAAATCT |
| | Poly(A) 3 Reverse | GCTGCACCGGGTAGGGTCTGA |
| ***HERPUD*** | Promotor Forward | GGTTGCATCAGCCCGTGCCC |
| | Promotor Reverse | CTGCAACGACAGTTCACGTCTCT |
| | Exon1 Forward * | TCCGAGACCGAACCCGAGCC |
| | Exon1 Reverse * | CAGGTGGGCCTTGAGGTGGC |
| | Coding Region Forward | GCGGATGAATGCACAAGGTGGC |
| | Coding Region Reverse | ACAACGGTGGCCCCCATGAC |
| | Poly(A) 1 Forward | CAGAGACTCCTGTCATCCTAGCAG |
| | Poly(A) 1 Reverse | ACCGGCACAAACAGTCCTCTTCT |
| | Poly(A) 2 Forward | CCAGCTCAAGGTAAGAAGGGTGGC |
| | Poly(A) 2 Reverse | GCTGTTTGGCAGGAAGAGCACG |
| | Poly(A) 3 Forward | CCTGTCCAAGGTCTCAAACCCCT |
| | Poly(A) 3 Reverse | CTCCTGGCCCCCACGCCATA |
| ***INTERGENIC*** | Intergenic Forward | GGCTAATCCTCTATGGGAGTCTGTC |
| | Intergenic Reverse | CCAGGTGCTCAAGGTCAACATC |
| ***U6 snRNA*** | U6 Forward | GCTTCGGCAGCACATATACTA |
| | U6 Reverse | AAATATGGAACGCTTCACGA |
| ***c-MYC*** | Exon3 Forward | CCTGAGCAATCACCTATGAACTTG |
| | Exon 3 Reverse | CAAGGTTGTGAGGTTGCATTTG |

* These primers were used in mRNA levels quantification