

UNIVERSITÀ DEGLI STUDI DI NAPOLI
“FEDERICO II”



Scuola di Dottorato in Medicina Molecolare
DOTTORATO DI RICERCA IN PATOLOGIA E FISIOPATOLOGIA
MOLECOLARE
XXVI CICLO

Mechanism of retinoic acid-induced transcription: histone
code, DNA oxidation and formation of chromatin loops

COORDINATORE

Chiar. mo Prof. Vittorio Enrico Avvedimento

CANDIDATA

Dott.ssa Candida Zuchegna

Tutor

Chiar. mo Prof. Antonio Porcellini

Anno Accademico 2013-2014

Table of contents

Abstract.....	4
Introduction.....	5
The chromatin structure.....	6
Epigenetics.....	6
<i>Histone post-translational modifications</i>	7
<u>Histone methylation</u>	9
<i>DNA methylation</i>	15
Transcription.....	17
<i>Transcription factors</i>	18
<i>Transcriptional regulation</i>	22
Nuclear receptors.....	26
<i>Retinoic acid and receptors</i>	29
<u>Caspase 9</u>	31
<u>CYP26A1</u>	33
DNA Base Excision Repair and Nucleotide Excision Repair.....	34
Aim of the study.....	39
Methods.....	42
Cells and transfections.....	43
RNA extraction and qRT-PCR and qPCR.....	43
Chromatin Immuno-Precipitation (ChIP).....	44
8-Oxo-7, 8-dihydro-2'-deoxyguanosine (8-oxo-dG) DNA Assay...	45
Chromosome conformation capture (3C).....	46

Antibodies used for the experiments.....	47
LSD1 Activity/Inhibition Assay.....	47
Statistical analysis.....	48
Results.....	50
Recruitment of RA receptor and activation of RNA polymerase II at RA-target promoters.....	51
Histone H3 K4 and K9 methylation marks induced by RA.....	57
Recruitment of base (BER) or nucleotide (NER) excision repair enzymes to the RARE-promoter chromatin following RA induction.....	69
Formation of dynamic chromatin loops governing the selection of 5' and 3' borders of RA-induced transcription units.....	76
Discussion and conclusions.....	82
References.....	87
List of publications.....	105

Abstract

Histone methylation changes and formation of chromatin loops involving enhancers, promoters and 3' end regions of genes have been variously associated with active transcription in eukaryotes. It is not known if these events are mechanistically linked and their specific role in transcription initiation. We have studied the effect of activation of the Retinoic A receptor, at the RARE-promoter chromatin of *CASP9* and *CYP26A1* genes, at 15 and 45 min following RA exposure, and we found that histone H3 lysine 4 and 9 are demethylated by the lysinespecific demethylase, LSD1 and by the JMJ-domain containing demethylase, D2A. The action of the oxidase (LSD1) and a dioxygenase (JMJD2A) in the presence of Fe⁺⁺ elicits an oxidation wave that locally modifies the DNA locally and recruits the enzymes involved in base and nucleotide excision repair (BER and NER). These events are essential for the formation of chromatin loop(s) that juxtapose the RARE element with the 5' transcription start site and the 3' end of the genes. The RARE bound-receptor governs the 5' and 3' end selection and directs the productive transcription cycle of RNA polymerase.

This is the first demonstration that chromatin loops, histone methylation changes and localized DNA repair are mechanistically linked.

Introduction

The chromatin structure

DNA consists of four nucleotide bases [adenine (A), guanine (G), cytosine (C), and thymine (T)] that are paired together (A-T and C-G) to give DNA its double helical shape. Nucleotide base sequences are the genetic code or instructions for protein synthesis. The associated DNA and histone proteins are collectively called chromatin (Fig. 1); the complex is tightly bonded by attraction of the negatively charged DNA to the positively charged histones. Genetic information encoded in DNA is largely identical in every cell of a eukaryote. However, cells in different tissues and organs can have widely different gene expression patterns and can

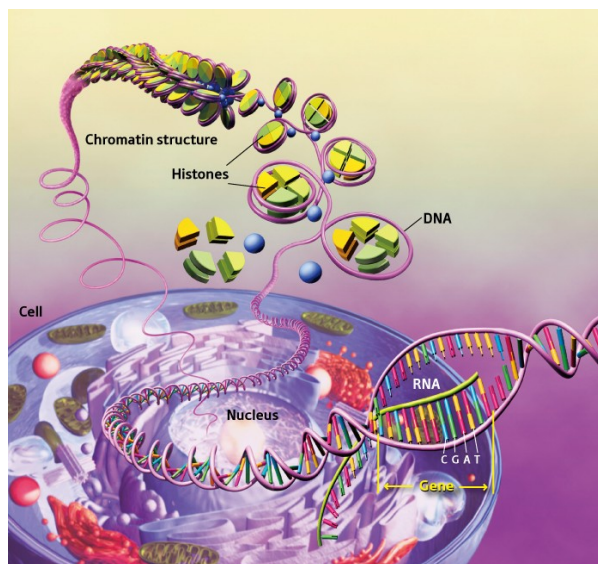


Figure 1 Schematic representation of chromatin structure. The chromatin fiber that makes up chromosomes is composed of nucleosome units, each consisting of DNA wrapped around histone proteins.

exhibit specialized functions. Gene expression in different cell types needs to be appropriately induced and maintained and also has to respond to developmental and environmental changes; inappropriate expression patterns lead to disease. Chromatin is not simply a packaging tool; it is also a dynamically entity that contains the regulatory signals necessary to program appropriate cellular pathways and is believed to contribute to the control of gene expression.

Epigenetics

Epigenetics is defined today as the study of changes in gene function that are transmitted through both mitosis and meiosis without involving any change in the DNA sequence (Wu C. & Morris J. R., 2001). The term is made of two parts: Greek prefix “epi”, which means upon or over and “genetics”, which is the science of genes, heredity, and variation in living organisms. This word was first defined by Conrad Waddington as the branch of biology which studies the causal

interactions between genes and their environment that create the phenotype (Waddington C. H., 2012). It is driven by specialized mechanisms that include DNA methylation, small non-coding regulatory RNAs, histone variants and histone post-translational modifications (Margueron R. & Reinberg D., 2010). The interaction between different epigenetic mechanisms controls the accessibility of genes by the transcriptional machinery.

Histone post-translational modifications

In eukaryotes, 147 bp of DNA is wrapped around an octamer of histones consisting of two copies of H2A, H2B, H3 and H4 with one molecule of histone H1 bound to the DNA as it enters the nucleosome core particle (Fig. 2). The resulting nucleosomes are further compacted to form higher-order chromatin structures, which remain poorly understood. The core histones (H2A, H2B, H3 and H4) have two domains: a histone fold domain, which is involved in interactions with other histones and in wrapping DNA around the nucleosome core particle, and an amino-terminal tail domain, which protrudes from the nucleosome (Luger K. et al., 1997) and can be subject to post-translational modifications (Fig. 3), such as acetylation, methylation, phosphorylation and monoubiquitylation, as well as other modifications that are less well studied (Kouzarides T., 2007; Vaquero A. et al., 2003; Strahl B. D. and Allis C. D., 2000). These modifications are thought to contribute to the control of gene expression through influencing chromatin compaction or signaling to other protein complexes. Therefore, an appropriate balance of stability and dynamics in histone post-translational modifications is necessary for accurate gene expression. Chromatin structure or landscape is a composite of various domains characterized by the local enrichment of a specific combination of histone post-translational modifications, histone variants, nucleosome occupancy, DNA methylation patterns and nuclear localization.

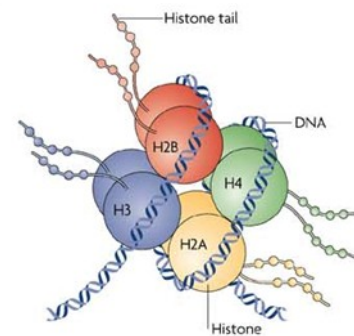


Figure 2 Structure of a nucleosome, with the DNA wrapped around an octameric histone core.

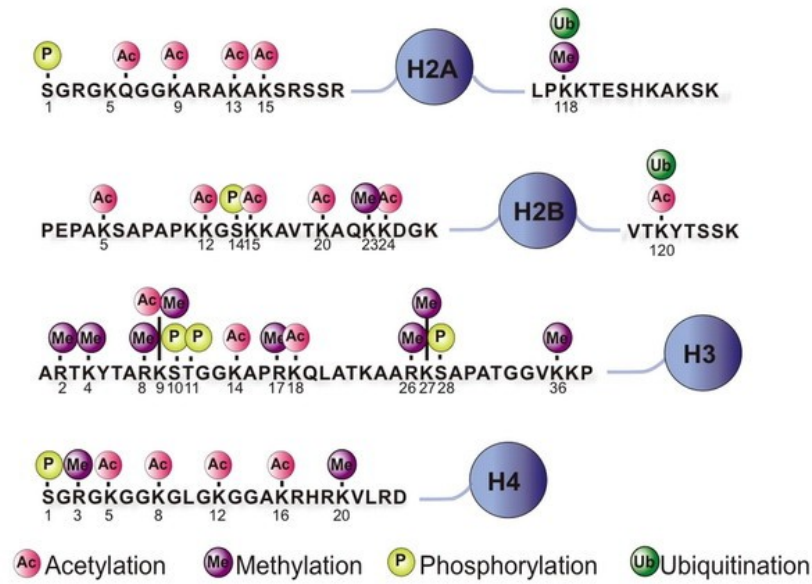


Figure 3 Histone modifications involved in chromatin reorganization. Histone N-tails are post-translationally modified, and certain combinations of histone modifications appear to generate a “histone code”.

Although some proteins that regulate chromatin structure are well defined, exactly how the histone-modifying enzymes, histone modifications and modification-recognizing proteins are localized and restricted to specific loci is currently unclear. Genome-wide profiling (using chromatin immunoprecipitation followed by microarray (ChIP–chip) or sequencing (ChIP–seq)) has provided a partial picture of the chromatin landscape, including the localization of histone post-translational modifications and histone variants, DNA methylation patterns and nucleosome occupancy. Moreover, the discovery of protein domains — including chromodomains, bromodomains, plant homeodomains (PHDs), tudor domains and malignant brain tumour (MBT) domains — that specifically recognize a defined histone modification has advanced the understanding of the role of histone post-translational modifications (Ruthenburg A. J. et al. 2007; Campos E. I. and Reinberg D., 2009). Some of them contribute to the transmission of epigenetic information or participate in the process of transcription (the so-called “active marks”), and others are probably restricted to “structural functions” (Berger S.L. et al., 2009; Trojer P. and Reinberg D., 2006). Although specific histone post-translational modifications have been correlated with defined functions, such as gene regulation, it is clear that a single type of histone post-translational

modification does not originate a single outcome. For example, histone 3 lysine 9 trimethylation (H3K9me3) is found both in silent heterochromatin and at some active genes (Campos E. I. & Reinberg D., 2009).

Histone methylation

Histone methylation occurs on all basic residues: arginines (Byvoet P. et al., 1972), lysines (Murray K., 1964) and histidines (Fischle W. et al., 2008). Lysines can be monomethylated (me1) (Murray K., 1964), dimethylated (me2) (Paik W. K. & Kim S., 1967) or trimethylated (me3) (Haempel K. et al., 1968) on their ϵ -amine group (Fig. 4); arginines can be monomethylated, symmetrically dimethylated or asymmetrically dimethylated on their guanidinyll group (Borun T. W. et al., 1972), and histidines have been reported to be monomethylated (Borun T. W. et al., 1972; Gershey E. L. et al., 1969), although this methylation seems to be rare and has not been further characterized. The most extensively studied histone methylation sites include histone H3 lysine 4 (H3K4), 9, 27, 36, 79 (Fig. 5) and histone H4 lysine 20. Sites of arginine (R) methylation include H3R2, H3R8, H3R17, H3R26 and H4R3. However, many other basic residues throughout the histone proteins have also recently been identified as methylated by mass spectrometry and quantitative proteomic analyses (reviewed in Young N. L. et al., 2010).

The functional effects and

the regulation of the newly identified methylation events remain to be determined. In general, methyl groups are believed to turn over more slowly than many other post-translational modifications, and histone methylation was originally thought to be irreversible (Byvoet P. et al., 1972). The discovery of an H3K4 demethylase, lysine-specific demethylase 1A (KDM1A; also known as LSD1), revealed that histone methylation is, in fact, reversible and dynamic (Shi Y. et al., 2004). Three

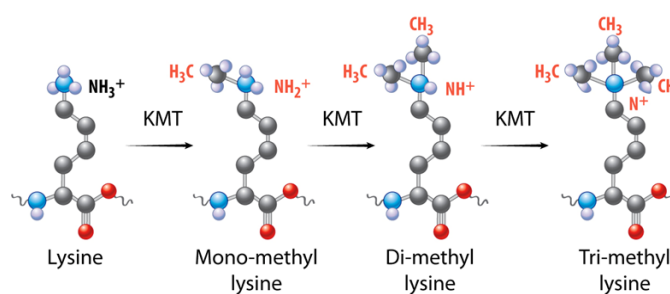


Figure 4 Lysine mono-, di- and trimethylation. The diversity of chemical states obtained by selective and sequential methylation of lysine residues.

families of enzymes have been identified thus far that catalyse the addition of methyl groups donated from S-adenosylmethionine to histones. The SET-domain-containing proteins (Rea S. et al., 2000) and DOT1-like proteins (Feng Q. et al.,

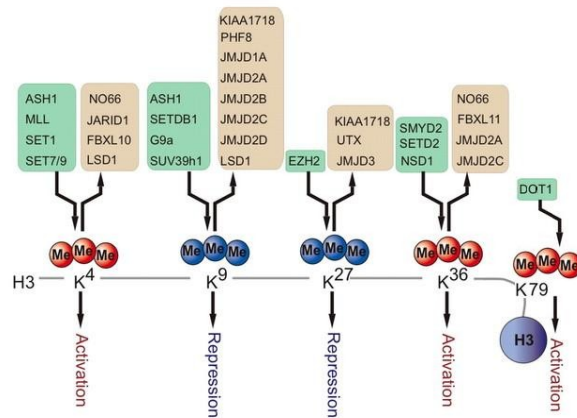


Figure 5 Methylation of different lysine residues in the "tail" of histone H3 has different effects on gene transcription.

2002) have been shown to methylate lysines, and members of the protein arginine N-methyltransferase (PRMT) family have been shown to methylate arginines (Bannister A. J. & Kouzarides T., 2011). These histone methyltransferases have been shown to methylate histones that are incorporated into chromatin and also free histones and non-histone proteins (Huang J. & Berger S. L., 2008).

Two families of demethylases have been identified thus far that mediate the removal of methyl groups from different lysine residues on histones. These are the amine oxidases (Shi et al., 2004) and jumonji C (JmjC)-domain-containing, iron-dependent dioxygenases (Tsukada Y. et al., 2006; Whetstine J. R. et al., 2006) (Fig. 6). These enzymes are highly conserved from yeast to humans and demethylate histone and non-histone substrates. Arginine demethylases remain more elusive. Although an initial report suggested that one of the JmjC domain proteins, JMJD6, demethylates arginines (Chang B. et al., 2007), a more recent study indicates that the main function of JMJD6 is to hydroxylate an RNA-splicing factor (Webby C. J. et al., 2009). Monomethyl arginines have also been shown to be converted by protein arginine deiminase type 4 (PADI4) to citrulline. However, PADI4 is not an arginine demethylase, as it works on both methylated and unmethylated arginine (Cuthbert G. L. et al., 2004).

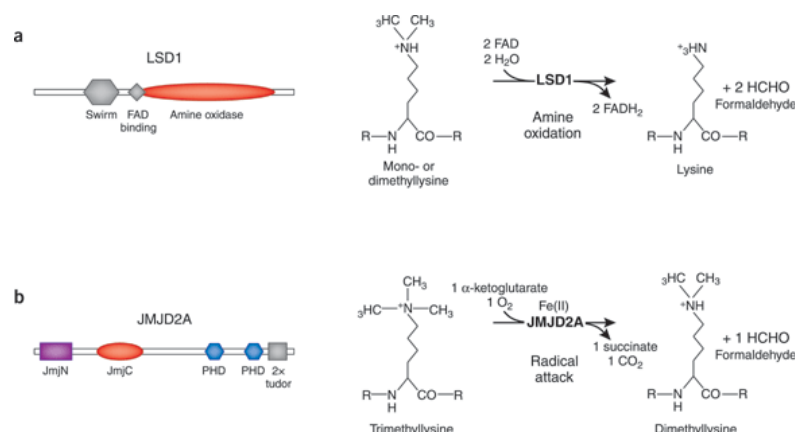


Figure 6 Structure and mechanism of action of the demethylases LSD1 (a) and JMJD2A (b). (a) Removal of methyl group(s) from mono- and dimethylated lysine residues is an oxidative process catalyzed by flavin-dependent amine oxidases from the LSD1 family. The substrate is oxidized by FAD to generate an imin intermediate, which is then hydrolyzed. This mechanism requires a protonated nitrogen and therefore precludes the use of trimethylated lysines as a substrate. (b) Histone demethylation catalyzed by JmjC domain-containing proteins. The oxidative demethylation mechanism used by these metalloenzymes requires Fe(II) and alpha-

Variation in patterns of methylations of histone tails are essential components of epigenetic regulation, as they reflect and modulate chromatin structure and function. In contrast to DNA CpG methylation, which is only prominent in some higher eukaryotes, histone methylation is present in organisms such as *C. elegans* and *D. melanogaster*, in which DNA methylation is largely absent. Depending on the biological context, some methylation events may have to be stably maintained (for example, in terminal differentiation) whereas others may need to be available to change (for example, in response to stimuli). Indeed, methylation at different lysine residues on histones has been shown to display differential turnover rates (Zee B. M. et al., 2010). An appropriate balance between stable and dynamic histone methylation is thus necessary to maintain normal biological function (Fig. 7). Methyl-modifying enzymes have a crucial role in almost every aspect of biology, and disruption of their function leads to developmental defects, diseases or ageing.

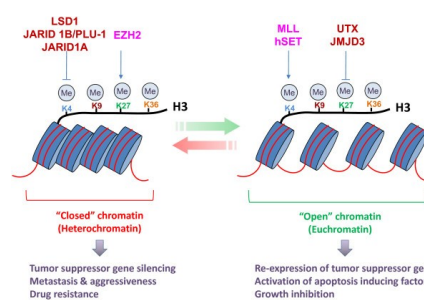


Figure 7 Model of dynamic interplay of enzymes mediating methylation of histone lysines. Methylases are shown in pink and demethylases are shown in red.

A current model suggests that methylated histones are recognized by chromatin effector molecules (“readers”), causing the recruitment of other molecules to alter the chromatin and/or transcription states (Taverna S. D. et al., 2007). The location of the methyl-lysine residue on a histone tail and the degree of methylation (whether me1, me2 or me3) have been associated with differential gene expression status. For example, H3K4me3 is generally associated with active transcription (Bernstein B. E. et al., 2002; Santos-Rosa H. et al., 2002), or with genes that are poised for activation, whereas H3K27me3 is associated with repressed chromatin. H3K4me1 is often associated with enhancer function (Heintzman N. D. et al., 2007), whereas H3K4me3 is linked to promoter activity. H3K79me2 is important for cell cycle regulation, whereas H3K79me3 is linked to the WNT-signalling pathway (Mohan M. et al., 2010). However, there are instances in which the same modifications can be associated with opposing activities, such as transcriptional activation and repression. This is the case for example of H3K4me2 and H3K4me3. Probably, the change in activity is due to different effector proteins. For instance, when H3K4me2 or H3K4me3 marks are bound by the PHD-domain containing co-repressor protein inhibitor of growth family member 2 (ING2), they are associated with transcriptional repression (Shi X. et al., 2006) through the stabilization of a histone deacetylase complex. Combined marks can also have different roles to the same marks appearing in isolation. Although H3K4me3 and H3K27me3 are marks associated with active and repressive transcription, respectively, when they are present together, they appear to have a role in poising genes for transcription (Bernstein B. E. et al., 2006). Combinatorial histone modifications are efficiently recognized by proteins with multiple domains to effect specific outcomes. For instance, the chromatin regulator TRIM24 has a PHD domain and a bromodomain, which recognize unmethylated H3K4 and acetylated H3K23 on the same histone tail (Tsai W. W. et al., 2010): this binding leads to estrogen-dependent gene activation. Combinatorial action of methyl-modifying enzymes is also context-specific. Histone methylation dynamics are known to have important roles in many biological processes, including cell-cycle regulation, DNA damage and stress response, development and differentiation (Eissenberg J. C. & Shilatifard A.,

2010). The importance of the tight regulation of histone methylation is demonstrated by emerging links of histone methylation to disease and ageing. Several studies have begun to address the role of histone modifications at specific stages of transcription. Most of the associations between histone methylation status and transcription are based on correlations between gene expression level and genome-wide or locus-specific chromatin immunoprecipitation (ChIP) studies. It appears that histone methylation has a role in many levels of transcriptional regulation from chromatin architecture to specific loci regulation through the recruitment of cell-specific transcription factors (Fig. 8) and interaction with initiation and elongation factors. In addition, histone methylation influences RNA processing.

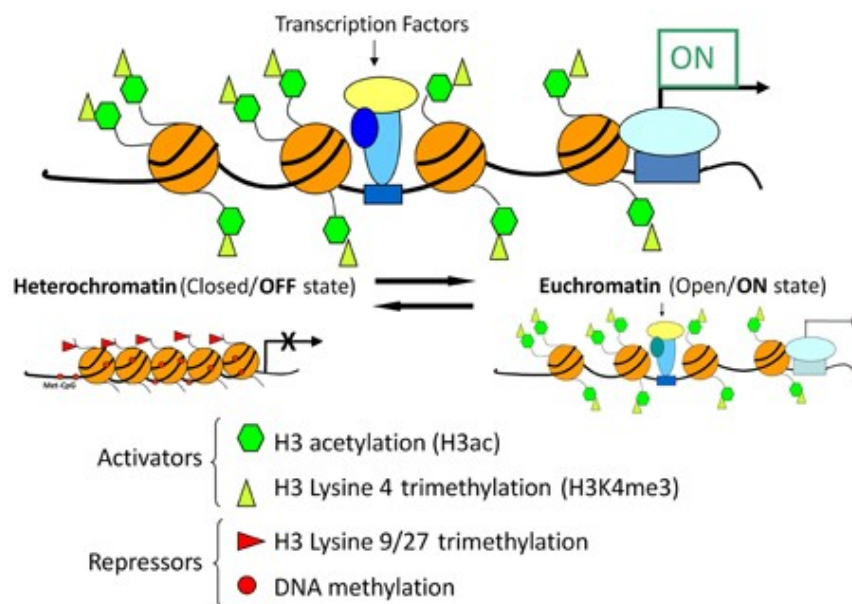


Figure 8 Chromatin state and gene transcription. Model for chromatin factors interacting with transcription factors to regulate transcription of a gene. Histone modifications and DNA methylation are important factors in regulating the chromatin from active to repressed and vice versa. Histone H3 acetylation and histone H3 methylation and lysine 4, are both associated with an active chromatin state. In contrast, histone H3 methylation at lysine 9 or lysine 27 as well as DNA methylation are associated with repressive chromatin state. Chromatin is in a dynamic equilibrium between the two states.

An interesting and untested hypothesis is that histone methylation could influence transcription by bringing physically separate regions of chromatin close together through chromosomal looping. This could include enhancer and promoter regions

or, in the case of repressive interactions, it could include insulator elements (Deng W. & Blobel G. A., 2010). However, whether chromosomal looping is a cause or a consequence of transcriptional regulation remains to be determined. Histone modifications can affect the higher-order chromatin structure directly (Shogren-Knaak M. et al., 2006) or indirectly by recruiting chromatin-remodelling complexes (Suganuma T. & Workman J. L., (2011; Bell O. et al., 2011). Inaccessible chromatin domains can be ‘opened’ by so-called pioneering factors (Cirillo L. A. et al., 2002), which are sequence-specific DNA-binding transcription factors (such as forkhead box protein A1, FOXA1, and GATA4). After binding of the pioneering factors, DNA methylation and histone modifications could participate in making the chromatin more accessible for other transcription factors, the pre-initiation complex (PIC) and RNA polymerase II (RNAPII) (Serandour A. A. et al., 2011). Certain histone methylation patterns (such as stretches of chromatin that are marked by a high density of H3K4 and H3K79 methylation) also appear to be necessary for binding of transcription factors, presumably by providing a euchromatic environment, which facilitates sequence-specific binding. It is still unclear whether the recruitment of chromatin-remodelling machinery to sites of transcription (Fuda N. J. et al., 2009) enables more efficient transcription and/or is necessary for elongation to begin. Despite advancements in understanding the role of histone methylation in transcriptional control, there is still a lot of uncertainty regarding the order of events.

Histone methylation has been implicated also in the control of RNA splicing. Interestingly, the average exon length of many eukaryotic species is similar to the length of DNA wrapped around one nucleosome (Zhu L. et al., 2009), whereas intron length varies greatly. The association of the splicing factor U2 small nuclear ribonucleoprotein (snRNP) with chromatin is enhanced by histone H3 lysine 4 trimethylation (H3K4me3) (Vermeulen M. et al., 2010; Sims R. J. et al., 2007). Moreover, recent global chromatin immunoprecipitation followed by sequencing (ChIP-seq) analyses in *C. elegans*, mice and humans show that exons are enriched for H3K36me3 compared to introns and that alternatively spliced exons have lower levels of H3K36me3 than constitutively spliced exons (Kolasinska-Zwierz P. et al., 2009). *In vitro* assays have shown that the rate of

transcriptional elongation can affect splicing. As H3K36me3 can recruit a histone deacetylase complex (Guccione E. et al., 2006), which represses transcription, a kinetic model for splicing has been proposed in which the histone methylation can affect the rate of transcription and thus can influence splicing (Luco R. F. et al., 2011).

DNA methylation

DNA methylation is the covalent modification of the cytosine residues by the addition of a methyl group at the 5'-carbon position. In mammals, DNA is methylated specifically at the C's that precede G's in the DNA chain (CpG dinucleotides). In normal somatic cells, most (over 50%) CpG islands are unmethylated. DNA methylation is important for the regulation of non-CpG islands, CpG islands promoters, and repetitive sequences to maintain genome stability. This covalent modification is correlated with reduced transcriptional activity of genes that contain high frequencies of CpG dinucleotides in the vicinity of their promoters (Bird A. P. & Wolffe A. P., 1999) and it has been implicated in development and differentiation (Li E. et al., 1992), imprinting (Li E. et al., 1993), X chromosome inactivation (Panning B. & Jaenisch R., 1998), and cancer (Laird P. W. et al., 1995; Baylin S. B. & Ohm J. E., 2006). Aberrant methylation has been found in cancer cells (Cho Y. H. et al., 2010) and it was shown that DNA methylation is associated with DNA damage and repair (Cuozzo C. et al., 2007) and that methylation is reduced by transcription of the repaired regions as a mechanism of adaptation to environmental challenges (Morano A. et al., 2014).

The enzymes that catalyze this modification are called DNA methyltransferases (DNMTs), and are well characterized and conserved in mammals and plants (Law J. A. et al., 2010). There are two categories of DNMTs: *de novo* and maintenance (Goll M. G. & Bestor T. H., 2005). Patterns of DNA methylation are initially established by the *de novo* DNA methyltransferases DNMT3A and DNMT3B during the blastocyst stage of embryonic development (Okano M. et al., 1999). These methyl marks are then faithfully maintained during cell divisions through the action of the maintenance methyltransferase, DNMT1, which has a preference for hemi-methylated DNA (Hermann A. et al., 2004). Both the establishment and

maintenance of DNA methylation patterns are crucial for development. The methylation of DNA is a general mechanism by which control of transcription in vertebrates is linked to chromatin structure (Fig. 9).

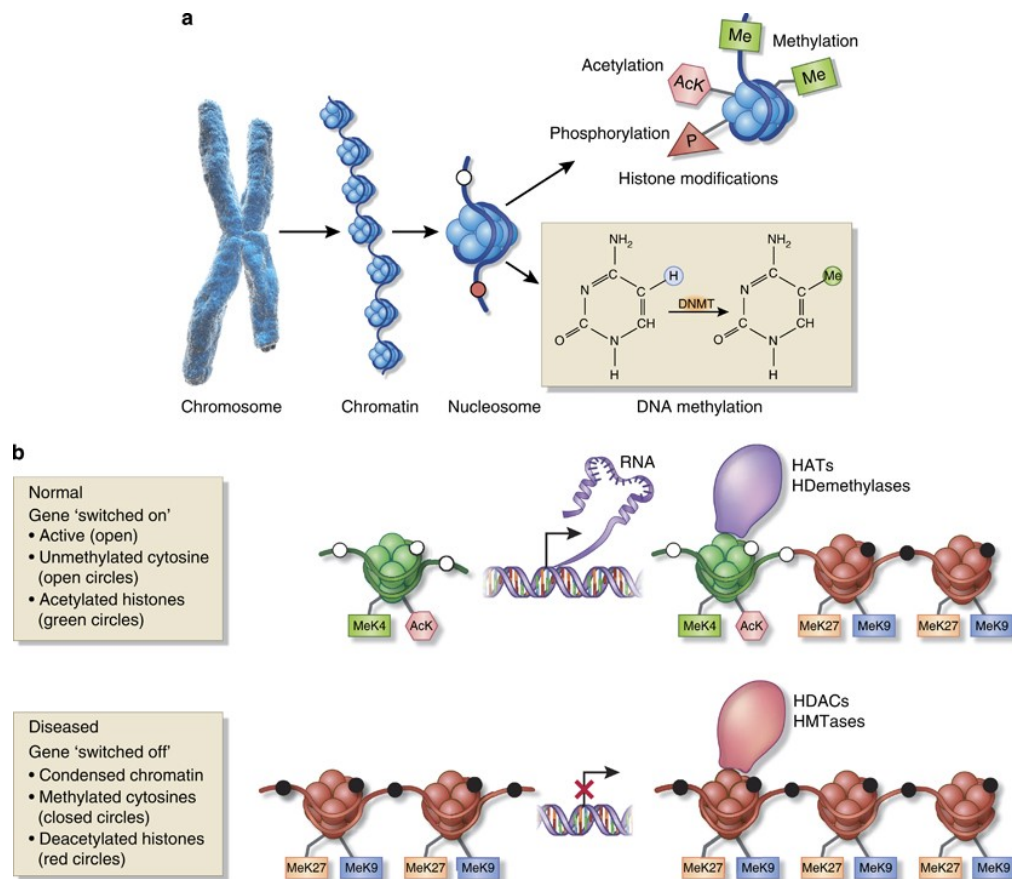


Figure 9 Mechanism of epigenetic modifications. (a) Epigenetic modifications of chromatin structure. Modifications of histones and DNA methylation provide a unique epigenetic signature that regulates chromatin organization and gene expression. (b) Epigenetic changes associated with disease states. In normal healthy tissues, promoter regions of actively transcribed genes are without DNA methylation at CpG dinucleotides (open circles) within CpG islands, and histones are modified with predominantly active marks (lysine 4 methylation and acetylation of lysine 9). The transcriptional start site is open and free of nucleosomes. This state is maintained by enzymes that modify the histone tails (histone acetyltransferases and histone demethylases). Intra- and intergenic regions have predominately methylated CpGs and inactive histone modifications (lysine 9 and lysine 27 methylation). In diseased states, methylation of CpG sites within CpG islands is associated with repressive chromatin marks, with these changes resulting from the presence of histone deacetylases

Methylation inhibits transcription through the action of a protein, MeCP2, that specifically binds to methylated DNA and represses transcription (Lewis J. D. et al., 1992). Interestingly, MeCP2 functions as a complex with histone deacetylase,

linking DNA methylation to alterations in histone acetylation and nucleosome structure.

DNA methylation also seems to have a role in directing histone methylation (Bartke T. et al., 2010). It was shown that H3K4me3 and DNA methylation are inversely correlated (Meissner A. et al., 2008) and, at least in plants, DNA methylation and H2A.Z are mutually exclusive (Zilberman D. et al., 2008) and H3K9me2, H3K9me3 and DNA methylation are highly coincident (Bernatavichute Y. V. et al., 2008). The example of H3K9me has been examined in several organisms. Knockdown of DNMT1 in mammals resulted in decreased levels of H3K9me2 and H3K9me3 (Espada J. et al., 2004). The exact mechanisms connecting replication to DNA methylation and H3K9me are not completely elucidated and might depend on the genomic locus and time of replication. However they seem to be tightly associated to replication through PCNA. Histone post-translational modifications can play an important role in the recruiting for methyl-modifying enzymes to specific genomic locations and, in some cases, in the determination of their substrate specificity (Kouzarides T., 2007).

Transcription

Transcription is the process by which the information in DNA is copied into RNA. It is performed by RNA polymerase. In the nucleus of eukaryotes, transcription is carried out by three different RNA polymerases, RNA polymerase I, II and III (Pol I, II and III) that transcribe distinct classes of genes: Pol I is responsible for the transcription of the large ribosomal RNA genes (28S, 18S and 5.8S), Pol II for the transcription of the protein-coding genes and some small nuclear RNAs and Pol III transcribes some structural and catalytic RNAs, including most small nuclear RNAs, tRNAs and 5S rRNA (Sentenac A., 1985). All three of the nuclear RNA polymerases are complex enzymes, consisting of 8 to 14 different subunits each. Transcription has three main steps: initiation, elongation and termination. Initiation consists in the binding of RNA polymerase to double-stranded DNA; this step involves a transition to single-strandedness in the region of binding; RNA polymerase binds to the DNA at a specific area called the promoter region. This region contains binding sites for RNA polymerase and

the transcription factors (TFs) necessary for normal transcription: RNA polymerase II cannot bind to promoters in eukaryotic DNA without the help of transcription factors (Struhl K., 1999). In many eukaryotic organisms, the promoter contains a conserved gene sequence called the TATA box. Various other consensus sequences also exist and are recognized by the different TF families. Transcription is initiated when one TF binds to one of these promoter sequences, initiating a series of interactions between multiple proteins (activators, mediators

and repressors) at the same site, or other promoter, regulator, and enhancer sequences (Fig. 10).

Ultimately, a transcription complex is formed at the promoter that facilitates

binding and transcription by RNA polymerase. During the elongation the covalent addition of nucleotides to the 3' end of the growing polynucleotide chain occurs; this involves the development of a short

stretch of DNA that is transiently single-stranded. The termination is the final step: RNA polymerase recognizes the terminator sequence and detaches from the DNA.

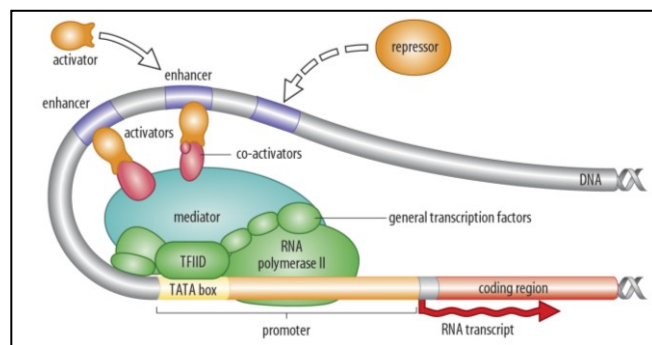


Figure 10 Assembly of the transcription initiation complex on a promoter. The basal machinery (RNA pol II and the general initiation factors (GIFs) assembles on the core promoter elements. The function of the general initiation factors is modulated by regulatory factors which recognize gene specific promoter proximal and distal enhancer elements. Proteins bound at the enhancer element interact with proteins bound at the promoter region to form a transcription initiation complex and to initiate transcription at a high rate.

Transcription factors

RNA Pol II needs to interact with specific proteins (called transcription factors, TFs) to initiate transcription. Two general types of transcription factors have been defined. General transcription factors are involved in transcription from all polymerase II promoters and therefore constitute part of the basic transcription machinery. Additional transcription factors bind to DNA sequences that control

the expression of individual genes and are thus responsible for regulating gene expression. Transcription factors are regulatory proteins whose function is to activate (or more rarely, to inhibit) transcription of DNA by binding to specific DNA sequences. TFs have defined DNA-binding domains with up to 106-fold higher affinity for their target sequences than for the remainder of the DNA strand. These highly conserved sequences have been used to categorize the known TFs into various "families," such as the MADS box-containing proteins, SOX proteins, and POU factors (Reményi A. et al., 2004). Transcription factors can also be classified by their three-dimensional protein structure, including basic

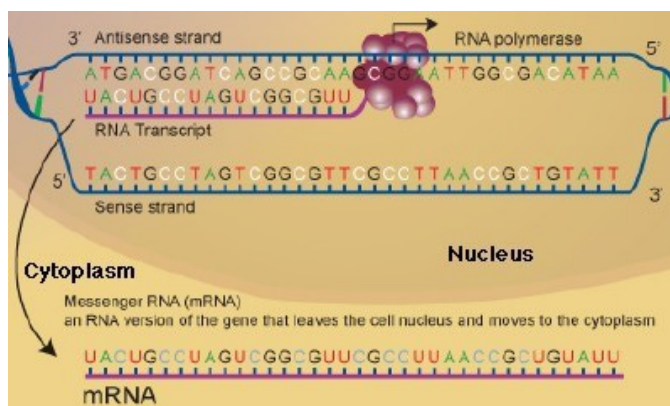


Figure 11 The transcription process. RNA synthesis involves separation of the DNA strands and synthesis of an RNA molecule in the 5' to 3' direction by RNA

polymerase to transcribe helix-turn-helix, helix-loop-helix, and zinc finger proteins. These different structural motifs result in transcription factor specificity for the consensus sequences to which they bind. These proteins unwind the DNA strand and allow RNA polymerase to transcribe

only a single strand of DNA into a single stranded RNA polymer called messenger RNA (mRNA). The strand that serves as the template is called the antisense strand. The strand that is not transcribed is called the sense strand (Fig. 11). RNA polymerase and the group of protein that directly interact with it (general factors) are called the basal transcription apparatus (Fig 12). This is the apparatus that is directly responsible for transcription. Five general transcription factors are required for initiation of transcription by RNA polymerase II in reconstituted *in vitro* systems. The first step in formation of a transcription complex is the binding of a general transcription factor called TFIID to the TATA box. TFIID is itself composed of multiple subunits, including the TATA-binding protein (TBP), which binds specifically to the TATAA consensus sequence, and 10-12 other polypeptides,

called TBP-associated factors (TAFs). TBP then binds a second general transcription factor (TFIIB) forming a TBP-TFIIB complex at the promoter. TFIIB in turn serves as a bridge to RNA polymerase, which binds to the TBP-TFIIB complex in association with a third factor, TFIIIF. Following recruitment of RNA polymerase II to the promoter, the binding of two additional factors (TFIIIE and TFIIH) is required for initiation of transcription. TFIIH is a multisubunit factor that appears to play at least two important roles. First, two subunits of TFIIH are helicases, which may unwind DNA around the initiation site. Another subunit of TFIIH is a protein kinase that phosphorylates repeated sequences present in the C-terminal domain (CTD) of the largest subunit of RNA polymerase II. Phosphorylation of these sequences is thought to release the

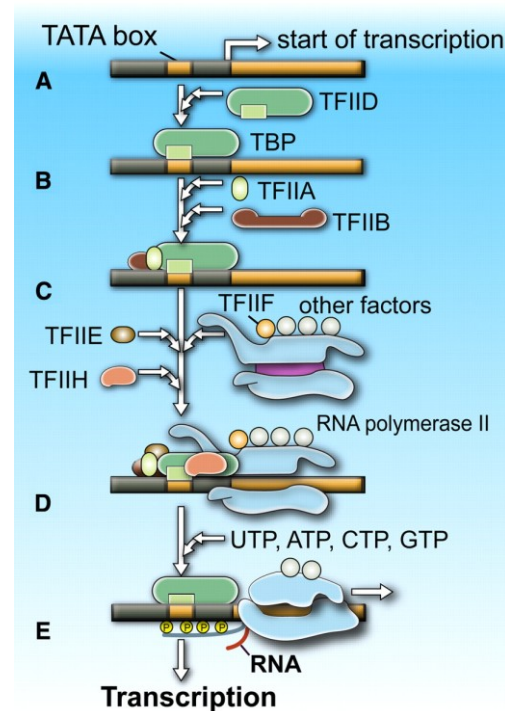


Figure 12 Schematic model of the assembly of the general transcription factors and RNA polymerase II on the promoter and beginning of the transcription.

polymerase from its association with the initiation complex, allowing it to proceed along the template as it elongates the growing RNA chain. After Pol II leaves the promoter, TFIIB and TFIIF are released, whereas other factors such as activators, TBP, Mediator, TFIIH and TFIIIE remain largely promoter-associated and form what is termed a reinitiation intermediate or scaffold, to facilitate subsequent rounds of transcription (Yudkovsky N. et al., 2000).

In addition to a TATA box, the promoters of many genes transcribed by RNA polymerase II contain a second important sequence element (an initiator, or Inr, sequence) that spans the transcription start site. Moreover, some RNA polymerase II promoters contain only an Inr element, with no TATA box. Initiation at these promoters still requires TFIID (and TBP), even though TBP obviously does not recognize these promoters by binding directly to the TATA sequence. Instead, other subunits of TFIID (TAFs) appear to bind to the Inr sequences. This binding recruits TBP to the promoter, and TFIIB, polymerase II, and additional transcription factors then assemble as already described. TBP thus plays a central role in initiating polymerase II transcription, even on promoters that lack a TATA box.

Other factors, those that interact directly or through a coactivator with the proteins of the basal transcription apparatus, are also important for transcription. These generally have a positive effect on transcription, but occasionally they can repress gene expression through transcription. These factors are called upstream factors and they are unique to each promoter. Finally, some factors are turned in a temporal or spatial manner, or directly in response to the environment. These factors provide the final link in controlling gene expression. These are termed inducible factors. Histone post-translational modifications, DNA methylation and nucleosome occupancy are pivotal to determining which response elements will be bound by a particular transcription factor either directly, by regulating the affinity of a transcription factor for its binding site, or indirectly, through factors that recognize a defined chromatin environment. However, the relationship between transcription factors and chromatin can also work another way: transcription factors can form regulatory loops (including positive-feedback loops) that impose epigenetic regulation (Ptashne M., 2007). The importance of

transcription factors in the chromatin landscape has been shown by reprogramming experiments. Transforming cells from their fully differentiated state into pluripotent ES cells requires the transient expression of only a few transcription factors, which initiate a dramatic restructuring of the chromatin landscape (Wernig M. et al., 2007). Despite the development of *in vitro* systems and the characterization of several general transcription factors, much remains to be learned concerning the mechanism of polymerase II transcription in eukaryotic cells.

Transcriptional regulation

Transcription is regulated at all steps by a variety of mechanisms. In eukaryotic cells it is controlled by proteins that bind to specific regulatory sequences and modulate the activity of RNA polymerase. In eukaryotes, regulation of gene expression requires the coordinated interactions of multiple proteins. The so-called housekeeping genes, are needed by almost every type of cell and appear to be unregulated or constitutive. But the regulation of gene expression in a tissue-specific manner is essential for cellular differentiation. Genes that regulate cell identity are turned on under very specific temporal, spatial, and environmental conditions to ensure that a cell is able to perform its designated function.

Gene expression is controlled on two levels. First, transcription is controlled by limiting the amount of mRNA that is produced from a particular gene. The second level of control is through post-transcriptional events that regulate the translation of mRNA into proteins. Even after a protein is made, post-translational modifications can affect its activity.

The state of chromatin structure at a specific region in eukaryotic DNA, along with the presence of specific transcription factors, works to regulate gene expression in eukaryotes. Sequence-specific TFs are considered an important mechanism of gene regulation in both prokaryotic and eukaryotic cells. Many activating TFs are generally bound to DNA until removed by a signal molecule, while others might only bind to DNA once influenced by a signal molecule. The binding of one type of TF can influence the binding of others, as well. Thus, gene expression in eukaryotes is highly variable, depending on the type of activators

involved and what signals are present to control binding. Even when transcription factors are present in a cell, transcription does not always occur, because often the TFs cannot reach their target sequences. The association of the DNA molecule with proteins is the first step in its silencing. The state of chromatin can limit access of transcription factors and RNA polymerase to DNA promoters, contributing to the restrictive ground state of gene expression. In order for gene transcription to occur, the chromatin structure must be unwound (Fig. 13). It allows simultaneous regulation of functionally or structurally related genes that tend to be present in widely spaced clusters or domains on eukaryotic DNA (Sproul D. et al., 2005).

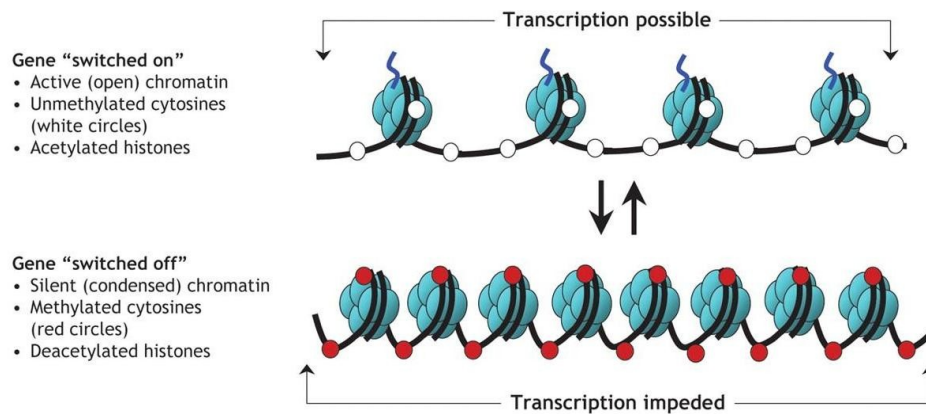


Figure 13 Representation of histone acetylation and transcription activation. Acetylated core histone proteins or unmethylated cytosines lead to a more open chromatin conformation resulting in a transcriptionally active state. On the other hand, removal of acetyl group by HDAC or DNA cytosines methylation repress the transcription and chromatin becomes more condensate.

Interactions of chromatin with activators and repressors can result in domains of chromatin that are open, closed, or poised for activation. Chromatin domains have various sizes and different extents of stability. These variations allow for phenomena found solely in eukaryotes, such as transcription at various stages of development and epigenetic memory throughout cell division cycles. They also allow for the maintenance of differentiated cellular states, which is crucial to the survival of multicellular organisms (Struhl K., 1999). Besides the chromatin structure and the presence of specific transcription factors there are other control activities in the cell, such as epigenetic mechanisms, including DNA methylation

and imprinting, noncoding RNA and histone post-translational modifications (Phillips T., 2008).

Many genes in mammalian cells are controlled by regulatory sequences located farther away (sometimes more than 10 kilobases) from the transcription start site. These sequences, called enhancers, were first identified by Walter Schaffner in 1981 during studies of the promoter of another virus, SV40 (Banerji J. et al., 1981). In addition to a TATA box and a set of six GC boxes, two 72-base-pair repeats located farther upstream are required for efficient transcription from this promoter. These sequences were found to stimulate transcription from other promoters as well as from that of SV40, and, surprisingly, their activity depended on neither their distance nor their orientation with respect to the transcription initiation site. They could stimulate transcription when placed either upstream or downstream of the promoter, in either a forward or backward orientation.

Enhancers, like promoters, function by binding transcription factors that then regulate RNA polymerase, even when separated by long distances from transcription initiation sites. This is possible because of DNA looping, which allows a transcription factor bound to a distant enhancer to interact with RNA polymerase or general transcription factors at the promoter. Transcription factors bound to distant enhancers can thus work by the same mechanisms as those bound adjacent to promoters, so there is no fundamental difference between the actions of enhancers and those of *cis*-acting regulatory sequences adjacent to transcription start sites. Interestingly, although enhancers were first identified in mammalian cells, they have subsequently been found in bacteria—an unusual instance in which studies of eukaryotes served as a model for the simpler prokaryotic systems (Cooper G. M., 2000).

Transcription by RNA polymerase II is coupled to RNA processing, including capping, splicing and cleavage/ polyadenylation. The C-terminal repeat (CTD) of RNA pol II orchestrates both processes by recruiting RNA processing factors. Indeed, CTD directly binds polyadenylation factors and its truncation inhibits transcript cleavage *in vivo*. A protein phosphatase that catalyzes the dephosphorylation of the C-terminal domain of RNA polymerase II is Ssu72. Genetic and physical interactions between Ssu72 and RNAP II have been

demonstrated and it has been hypothesized a role for Ssu72 in basal (noninduced) transcription by RNAP II (Pappas D. L. Jr & Hampsey M., 2000) (Fig. 14). Ssu72 was initially identified in a screen for suppressors of *sua7-1*, a cold-sensitive mutation in yeast TFIIB, hence Ssu72 (Suppressor of *sua7-1* clone 2) (Sun Z.W. & Hampsey M., 1996). Recent analysis revealed that Ssu72 dephosphorylates Ser-5 in the CTD of RNA pol II and regenerates initiation competent hypophosphorylated RNA pol II (Krishnamurthy S. et al., 2004). Large-scale analysis of protein complexes in yeast identified Ssu72 as a component of a cleavage and polyadenylation factor (CPF) complex, it interacts directly with the Pta1 subunit of CPF and is implicated in transcript cleavage and termination (Dichtl B. et al., 2002). Alternatively, Ssu72 may exert two independent functions: transcription/RNA processing in the nucleus and as yet to be defined activity in the cytoplasm (St-Pierre B. et al., 2005). Interestingly, there are no apparent Ssu72 homologs in bacterial or archaeal genomes, implying that Ssu72 function is specific to eukaryotes.

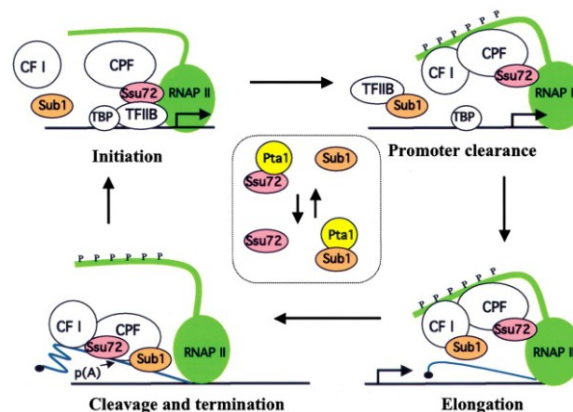


Figure 14 A model illustrating how Ssu72 might function at different points in the transcription cycle through interactions with TFIIB, RNAP II, cleavage/polyadenylation factor (CPF), and CF I. In the initiation stage, Ssu72 helps to correctly position RNAP II at the promoter through direct interactions with TFIIB and RNAP II. Ssu72 also recruits CPF to the promoter through its Pta1 partner and/or the weaker interactions with other CPF subunits. Ssu72 and Sub1 act as positive elongation factors. Recognition of processing signals by CPF and CF I triggers transcription termination. The inset depicts the mutually exclusive interaction of Ssu72 and Sub1 with Pta1 during the transcription cycle. The blue line represent pre-mRNA, capped at the 5' end and the 3' processing site shown by p(A).

Nuclear receptors

Nuclear hormone receptors are ligand-activated transcription factors that bind lipophilic molecules and regulate gene expression by interacting with specific DNA sequences of their target genes. Nuclear receptor ligands are chemically diverse, including hydrophobic molecules such as steroid hormones (e.g. estrogens, glucocorticoids, progesterone, mineralocorticoids, androgens, vitamin D3, ecdysone, oxysterols and bile acids), retinoic acids (all-trans and 9-cis isoforms), thyroid hormones, fatty acids, leukotrienes and prostaglandins (Escriva et al., 2000; Laudet and Gronemeyer, 2002). Because ligands are nonpolar, they can just diffuse across the plasma membrane. There are a total of 48 nuclear receptor family members in the human genome (Robinson-Rechavi et al., 2001). For some of these receptors, the physiological function and endogenous natural ligand are not known: these are termed orphan receptors. Some of the original orphan receptors have now had their endogenous ligands identified ("adopted orphans"). As early as 1968 a two-step mechanism of action was proposed for these receptors based upon the observation of an inactive and an active state of the receptors. The first step involves activation through binding of the hormone; the second step consists of receptor binding to DNA and regulation of transcription. Genes that are regulated by nuclear receptors contain particular DNA sequences (response elements) in their promoters, where the nuclear receptor binds. A hormone response element (HRE) is a specific DNA sequence that a receptor recognizes with markedly increased affinity and typically contains two consensus hexameric half-sites. Thus each receptor protein dimer that binds the DNA has to recognize the sequence, spacing and orientation of the half-sites within their response element. For dimeric HREs, the half-sites can be configured as palindromes, inverted palindromes, or direct repeats. There are two main classes of nuclear receptors. Steroids like testosterone, estrogens, cortisols are type I ligands and bind to inactive cytosolic receptors bound to heat shock proteins. Binding of the hormones activate them by dissociating heat shock proteins from the receptors. These activated receptors move into the nucleus and bind as homodimers to their specific hormone response elements (HRE), which are mostly located in the enhancer region of the gene promoter/regulatory regions.

Vitamin D, thyroid hormones and retinoids are type II ligands: receptors are retained in the nucleus bound to DNA as heterodimers regardless of the ligand binding status. In the absence of ligand, the receptors are complexed with corepressor proteins. Hormone binding causes the dissociation of corepressor and the recruitment of coactivator proteins, which recruit additional proteins to activate transcription.

Nuclear receptors share a common structural organization (Fig 15).



Figure 15 Structure of nuclear receptors. Schematic diagram for a common domain structure of NRs which include N-terminal activation function 1 (AF-1), DNA binding domain (DBD) consisting of two zinc fingers, hinge region (Hinge), ligand binding domain (LBD), and C-terminal AF-2.

The N-terminal region (A/B domain) is highly variable, and contains at least one constitutionally active transactivation region (AF-1) and several autonomous transactivation domains (AD); A/B domains are variable in length, from less than 50 to more than 500 amino acids, and their 3D structure is not known. The most conserved region is the DNA-binding domain (DBD, C domain), which notably contains the P-box, a short motif responsible for DNA-binding specificity on sequences typically containing the AGGTCA motif, and is involved in dimerization of nuclear receptors. This dimerization includes homodimers as well as heterodimers. The 3D structure of the DBD has been resolved for a number of nuclear receptors and contains two highly conserved zinc fingers, the four cysteines of each finger chelating one Zn^{2+} ion. Between the DNA-binding and ligand-binding domains is a less conserved region (D domain) that behaves as a flexible hinge between the C and E domains, and contains the nuclear localization signal (NLS), which may overlap on the C domain. The largest domain is the moderately conserved ligand-binding domain (LBD, E domain), whose secondary structure of 12 α -helices is better conserved than the primary sequence. The central DBD is responsible for targeting the receptors to their hormone response elements (HRE). The DBD binds as a dimer with each monomer recognizing a six

base pair sequence of DNA. The reading helix of each monomer makes sequence specific contacts in the major groove of the DNA at each half-site. These contacts allow the dimer to read the sequence, spacing and orientation of the half-sites within its response element, and thus discriminate between sequences. These proteins exhibit, however, a flexibility in recognizing DNA sequences and also accept a variety of amino-acid substitutions in their reading helix without abolishing binding. The LBD participates in several activities including hormone binding, homo- and/or heterodimerization, formation of the heat-shock protein complex and transcriptional activation and repression. The binding of the hormone induces conformational changes that seem to control these properties and influence gene expression. The conformational changes that accompany the transition between the liganded and unliganded forms of the nuclear hormone receptors affect dramatically their affinity for other proteins. The proteins that associate with the receptor may be activators or repressors of transcription. The general term for this type of protein is coregulator. A particular receptor may associate with different groups of coregulators in different cell types. In the absence of ligand, an inhibitory complex associates with the ligand-binding domain. Ligand binding causes a conformational change so that the inhibitory complex dissociates. This allows the receptor to travel to the nucleus, bind to DNA, and associate with the coactivator protein complex (Fig.16).

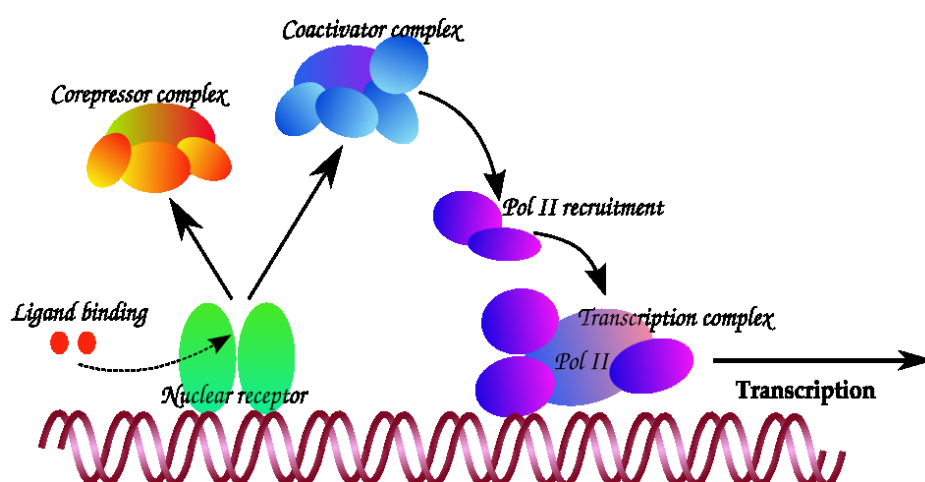


Figure 16 Regulation of transcription induced by nuclear receptors through the dimerization and the interaction with co-regulators proteins.

Retinoic acid and receptors

Retinoic acid (RA) is a lipophilic molecule and an active metabolite of vitamin-A (all-trans-Retinol), which belongs to the retinoids, a class of chemical compounds each composed of three basic parts: a trimethylated cyclohexene ring that is a bulky hydrophobic group, a conjugated tetraene side chain that functions as a linker unit, and a polar carbon-oxygen functional group. Biochemical conversion of carotenoid or other retinoids to retinoic acid (RA) is essential for normal regulation of a wide variety of biological processes like cell proliferation, development, differentiation and apoptosis. RA exerts its action by the transcriptional regulation of specific genes via a family of nuclear receptors called retinoic acid receptors, RARs, and retinoid X receptors, RXRs. The RAR family is activated both by all-trans-RA and by 9-*cis*-RA, whereas the RXR family is activated exclusively by 9-*cis*-RA. The RXRs play a central role in dimerization of nuclear receptors and in nuclear receptor signaling, as they are partners for different receptors that bind as heterodimers to DNA (Zhang X. K. et al., 1992). A two-step model for heterodimeric binding to DNA has been proposed. First, RXR would form heterodimers in solution with its partner through their dimerization interfaces contained in the LBDs, and in a second step, the DBDs would be able to bind with affinity to the DNA (Mangelsdorf D. J. & Evans R. M., 1995). According to a current model of transcriptional activation, in the absence of ligand, RAR/RXR heterodimers are bound to DNA, and they recruit co-repressors with HDAC (histone deacetylase) activity, resulting in chromatin condensation and gene silencing (Dilworth F. J. & Chambon P., 2001). Upon ligand binding, RAR and RXR undergo conformational changes that favor the dissociation of co-repressors and the recruitment of other proteins with histone acetylase activity, which opens up the chromatin, making it accessible to transcriptional machinery to initiate transcription (Fig. 17). Experiments with knock-out mice have clearly shown that the RXR/RAR heterodimer is responsible for different biological effects of retinoids on development (Kastner P. et al., 1997).

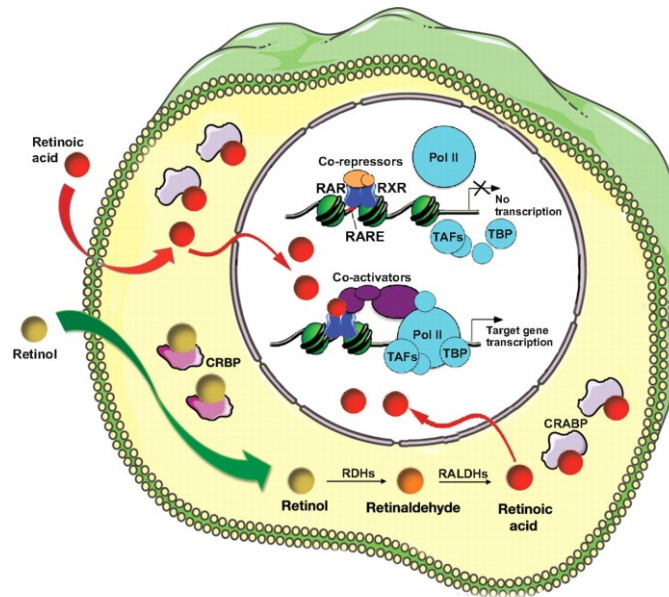


Figure 17 Summary of the RA signalling pathway. RA, synthesized intracellularly from circulating retinol or diffusing from an adjacent cells, eventually reaches the nucleus. Cellular retinoic acid-binding proteins (CRABPs) may be involved in this transfer. Cellular retinol-binding proteins (CRBPs) may help present retinol to retinol dehydrogenases (RDHs). Dimers of RA receptors (RARs) and retinoid X receptors (RXRs) are able to bind to RA-response elements (RAREs) in their target genes in the absence of ligand, interacting with protein complexes (co-repressors) that stabilise the chromatin nucleosomal structure and prevent access to the promoter. Upon RA binding, a conformational change in the helicoidal structure of the RAR ligand-binding domain changes its protein-protein interaction properties, releasing the co-repressors and recruiting co-activator complexes that destabilise the nucleosomes and/or facilitate assembly of the transcription pre-initiation complex, which contains RNA polymerase II (Pol II), TATA-binding protein (TBP) and TBP-associated factors (TAFs).

Three distinct but highly homologous RAR isotypes have been described termed RAR α , RAR β and RAR γ , encoded by three separate genes. In addition, several isoforms of each RAR isotype, which vary in both the length and amino acid sequence of the N-terminal A domain, have been identified, generated by alternative promoters and differential splicing. The targets of RA include a multitude of structural genes, oncogenes, transcription factors and cytokines (Balmer J. E. & Blomhoff R., 2002). Like all nuclear receptors RARs also have a conserved modular structure consisting of an AF-1 or A/B (Amino-Terminal Activating Factor-1 Transcriptional Activation) Domain; a zinc-finger DBD or C (DNA-Binding Domain); a CoR or D (Hinge/Corepressor Binding) Domain; a LBD or AF-2 or E (Ligand-Binding/Transcriptional Activation) Domain; and a

variable F (Carboxyl-Terminal) Domain. In general, the RARs contain six regions from A-F. The binding site for RAR/RXR heterodimers DBD is a specific DNA sequence known as a RARE (RA response element). RAREs consist of a direct repeat of a core hexameric sequence, PuG(G/T)-TCA, separated by 1, 2 or 5 base-pairs (DR1, DR2 and DR5) (Chambon P., 1996). In addition to RAR, two other proteins, termed cellular retinoic acid-binding proteins (CRABP-I and CRABP-II), bind RA with high affinity and specificity (Dong D. et al., 1999). CRABPs are small (~ 14 kDa) soluble proteins that are members of the family of intracellular lipid binding proteins. It is generally believed that CRABPs function to solubilize and protect RA in the aqueous space of the cytosol, but accumulating evidence suggests that they also play more specific roles in modulating signaling by RA. In regard to the biological functions of CRABP-II, it has been showed that this protein transports RA from the cytosol to the nucleus where it directly associates with RAR and that the resulting complex mediates “channeling” of RA to the receptor, thereby facilitating its ligation and enhancing its transcriptional activity (Budhu A. et al., 2001). More than 500 genes with diverse functions are regulated by RA, and RAREs have been localized in many of these genes including RAR β , CRABPI, CRABPII and members of the Hox and HNF gene families (Balmer J. E. & Blomhoff R., 2002). It was showed that after hormone binding, an active receptor complex induces covalent modifications at the N-terminal tails of nucleosomal histones and assembles an active transcription complex on chromatin (Shahhoseini M., 2013).

Caspase 9

Human *CASP9* (apoptosis-related cysteine peptidase) is located on chromosome 1 (1p36.1-1p36.3). It is approximately 35 Kb long and has 9 exons and 8 introns (Hadano S. et al., 1999). Alternative splicing results in multiple transcript variants. This gene encodes a member of the cysteine-aspartic acid protease (caspase) family, which is thought to play a central role in apoptosis and to be a tumor suppressor. The mammalian caspase family consists of 14 members (Earnshaw W.C. et al., 1999). Caspases are involved in the signal transduction pathways of apoptosis, necrosis and inflammation. They have been implicated in

the pathogenesis of many disorders including stroke, Alzheimer's disease, myocardial infarction, cancer, and inflammatory disease. These enzymes can be divided into two major classes - initiators and effectors (Fig. 18). The initiator

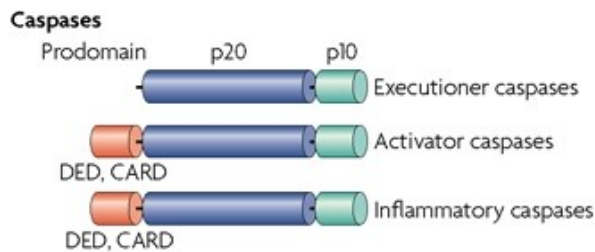


Figure 18 Schematic representation of caspases main domains: a prodomain and large (p20) and small (p10) catalytic subunits. The large domain contains the active site Cys residue. Activation of caspases involves removal of the prodomain and separation of the p20 and p10 subunits. The prodomains of activator and inflammatory caspases contain protein-protein-interaction domains, such as the caspase-recruitment domain (CARD) and the death-effector domain (DED).

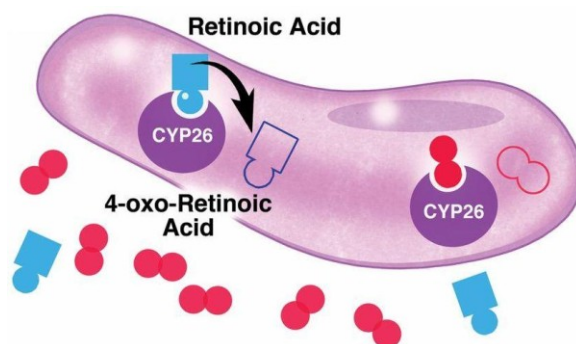
isoforms (caspases-1,-4,-5,-8,-9,-10,-11,-12) are activated by, and interact with, upstream adaptor molecules through protein-protein interaction domains known as CARD and DED. Effector caspases (-3,-6,-7) are responsible for cleaving downstream substrates and are sometimes referred to as the executioner caspases. More than 400 caspase substrates have so far been identified.

Caspases exist as inactive proenzymes which undergo proteolytic processing at conserved aspartic residues to produce two subunits, large (20 kDa) and small (10 kDa), that dimerize to form the active enzyme (Kuida K., 2000). They can also be found intracellularly as part of large multiprotein complexes. Caspase 9 can undergo autoproteolytic processing and activation by the apoptosome, a protein complex of cytochrome c and the apoptotic peptidase activating factor 1 (Apaf-1); this step is thought to be one of the earliest in the caspase activation cascade (Li P. et al., 1997). Binding of caspase-9 to Apaf-1 leads to activation of the protease which then cleaves and activates caspase-3. Caspases are regulated by inhibitors of apoptosis and by dominant negative isoforms. *CASP9* contains within a stretch of 8 kb upstream of the start site (1p36.3), a potential RARE composed of the noncanonical DR-2 sequence AGGTCAGcAGTTTCG at position -1690, but this element does not function as a RARE. An additional potential RARE, composed of the consensus DR-2 sequence AGGTCAGgAGTTTCA, was found in the second intron of the gene, 9.461 bp downstream of the start site. It was demonstrated that caspase 9 is a direct target for RAR signaling, and that the RARE responsible for

this response is likely to be this DR-2 element located in the second intron of the gene (Donato L. J. & Noy N., 2005).

CYP26A1

CYP26A1 encodes a member of the cytochrome P450 superfamily of enzymes, which include *CYP26A1*, *CYP26B1*, and *CYP26C1*. The cytochrome P450 proteins are monooxygenases which catalyze many reactions involved in drug metabolism and synthesis of cholesterol, steroids and other lipids. This endoplasmic reticulum protein acts on retinoids, including all-trans-retinoic acid (RA), and converts it to more polar metabolites through 4-oxidation, 4-hydroxylation, 18-hydroxylation and 5,6-epoxydation activities (White J. A. et al., 1996; Fujii H. et al., 1997) (Fig. 19). *CYP26A1* was first isolated from zebrafish as a gene product induced by RA during regeneration of adult caudal fin (White J. A. et al., 1996). Subsequently, homologs have been isolated from human, mouse, chick, and *Xenopus* with all the genes exhibiting a high degree of sequence conservation. *CYP26A1* metabolizes all-trans RA



but not the 9-cis or 13-cis RA isomers and regulates the cellular level of retinoic acid, which is

involved in regulation of gene expression in both embryonic and adult tissues. Two alternatively spliced transcript variants of *CYP26A1* gene, which encode the distinct isoforms, have been reported. Analysis of *CYP26A1* expression in cultured human cells shows that exogenous RA can strongly induce the expression of this gene indicating that regulation of RA catabolism may include a positive feedback loop (Sonneveld E. et al., 1998). Some of the RA inducibility of *CYP26A1* is due to regulation at the transcriptional level, mediated by a highly conserved RARE. Analysis of human, mouse, and zebrafish proximal regions of the *CYP26A1* promoter allowed to determine the presence of a canonical RARE (R1) within a conserved 32-bp sequence (in the first 200 bp of the *CYP26A1*

Figure 19 Control of RA distribution regulated by the CYP26A1 enzyme, which converts it to 4-oxo-retinoic acid.

promoter), which was shown to be recognized by the RAR γ /RXR α heterodimer (Loudig O. et al., 2000). Subsequently, it was uncovered a conserved second RARE (R2) occurring 2 kb upstream of the transcription start site, which appears to function synergistically with the R1 element to provide maximal induction of *CYP26A1* in response to RA, but it was unable to support transcription in the absence of R1 and its surrounding sequences, so it was proposed that the *CYP26A1* gene contains a single promoter that includes R1 and that R2 is an upstream enhancer element that is necessary for complete RA inducibility, but is not by itself sufficient for transcription (Loudig O. et al., 2005).

DNA Base Excision Repair and Nucleotide Excision Repair

Normal metabolic processes generate reactive oxygen species (ROS), which modify bases by oxidation. Both purine and pyrimidine bases are subject to oxidation. The most common mutation is guanine oxidized to 8-oxo-7,8-dihydroguanine, resulting in the nucleotide 8-oxo-deoxy-guanosine (8-oxo-dG). The 8-oxo-dG is capable of base pairing with deoxyadenosine, instead of pairing with deoxycytidine as expected. If this error is not detected and corrected by mismatch repair enzymes, the DNA subsequently replicated will contain a C→A point mutation. ROS may also cause depurination, depyrimidination, and single-strand or double strand breaks in the DNA.

Oxygen radicals generate mostly non-bulky DNA lesions, most of them are substrates for Base Excision Repair (BER). This repair system involves multiple enzymes to excise and replace a single damaged nucleotide base (Fig. 20). Key enzymes of the BER pathway are DNA-glycosylases: a DNA glycosylase cleaves the bond between the nucleotide base and ribose, leaving the ribose phosphate chain of the DNA intact but resulting in an apurinic or apyrimidinic (AP) site. 8-Oxoguanine DNA glycosylase I (OGG1) removes 7,8-dihydro-8-oxoguanine (8-oxoG), one of the base mutations generated by reactive oxygen species. Polymorphism in the human OGG1 gene is associated with the risk of various cancers such as lung and prostate cancer. Uracil DNA glycosylase, another BER enzyme, excises the uracil that is the product of cytosine deamination, thereby

preventing the subsequent C→T point mutation. N-Methylpurine DNA glycosylase (MPG) is able to remove a variety of modified purine bases. The AP sites in the DNA that result from the action of BER enzymes, as well as those that result from depyrimidination and depurination actions, are repaired by the action of AP-endonuclease 1 (APE1). APE1 cleaves the phosphodiester chain 5' to the AP site. The DNA strand then contains a 3'-hydroxyl group and a 5'-abasic deoxyribose phosphate. DNA polymerase β (Pol β) inserts the correct nucleotide and removes the deoxyribose phosphate through its associated AP-lyase activity. The presence of X-ray repair cross-complementing group 1 (XRCC1) is necessary to form a heterodimer with DNA ligase III (LIG3). XRCC1 acts as a scaffold protein to present a non-reactive binding site for Pol β , and bring the Pol β and LIG3 enzymes together at the site of repair (Lindahl T. & Wood R. D., 1999). Poly(ADP-ribose) polymerase (PARP-1) interacts with XRCC1 and Pol β and is a necessary component of the BER pathway (Caldecott K.W. et al., 1996; Dantzer F. et al., 2000). The final step in the repair is performed by LIG3, which connects the deoxyribose of the replacement nucleotide to the deoxyribosylphosphate backbone. This pathway has been named "short-patch BER" (Srivastava D. K. et al., 1998). An alternative pathway called "long-patch BER" replaces a strand of nucleotides with a minimum length of 2 nucleotides. Repair lengths of 10 to 12 nucleotides have been reported (Ranalli T. A. et al., 2002; Sattler U. et al., 2003). Longpatch BER requires the presence of proliferation cell nuclear antigen (PCNA), which acts as a scaffold protein for the restructuring enzymes (Fortini P. et al., 1998). Other DNA polymerases, possibly Pol δ and Pol ϵ (Klungland A. & Lindahl T., 1997), are used to generate an oligonucleotide flap. The existing nucleotide sequence is removed by flap endonuclease-1 (FEN1). The oligonucleotide is then ligated to the DNA by DNA ligase I (LIG1), sealing the break and completing the repair. The process used to determine the selection of short-patch versus long patch BER pathways is still under investigation (Sung J. S. & Demple B., 2006).

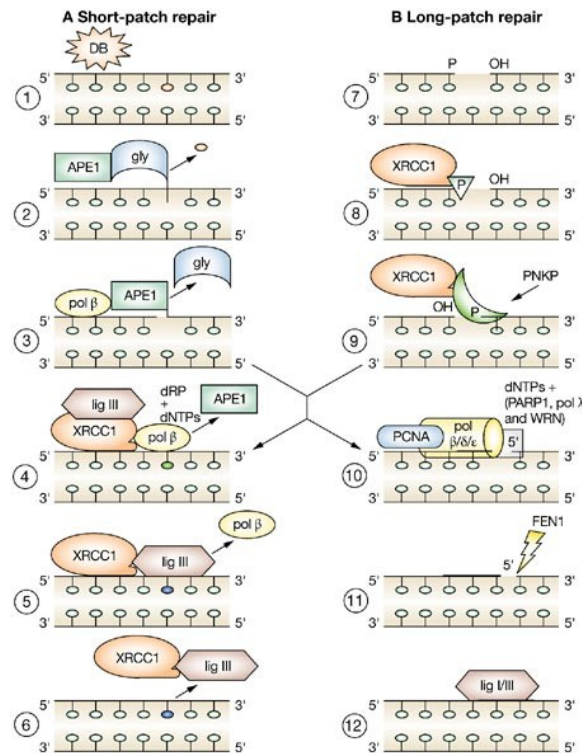


Figure 20 Scheme of base excision repair showing the two subpathways: (A) the 'short-patch' or single-nucleotide pathway, and (B) the 'long-patch' pathway. Crossing over of the pathways can occur at points (3) and (9). There are essentially four steps in the base excision repair pathway. First, when an altered base is detected (1) the surveillance glycosylases remove that base (2). Next, the endonuclease that is specific for an apurinic or apyrimidinic site cleaves the strand on the 5' side of the abasic site (3). This is followed by filling in of the gap with a correct nucleotide by DNA pol β , and at the same time releasing the dRP (4). Finally, DNA ligase III ligates the newly introduced nucleotide with the downstream sequence (5), thereby restoring the repaired DNA (6). Sometimes, other DNA polymerases such as DNA polymerase δ or ϵ , along with PCNA, are involved in filling larger sized gaps, also in a strand-displacement manner (long-patch repair [B]; steps 7–12). (Rao K. S., 2007).

While BER may replace multiple nucleotides via the long-patch pathway, the initiating event for both short-patch and long-patch BER is damage to a single nucleotide, resulting in minimal impact on the structure of the DNA double helix. Nucleotide Excision Repair (NER) repairs damage to a nucleotide strand containing at least 2 bases and creating a structural distortion of the DNA. NER acts to repair single strand breaks in addition to serial damage from exogenous sources such as bulky DNA adducts and UV radiation (Balajee A. S. & Bohr V. A., 2000). The same pathway may be used to repair damage from oxidative stress (Gros L. et al., 2002).

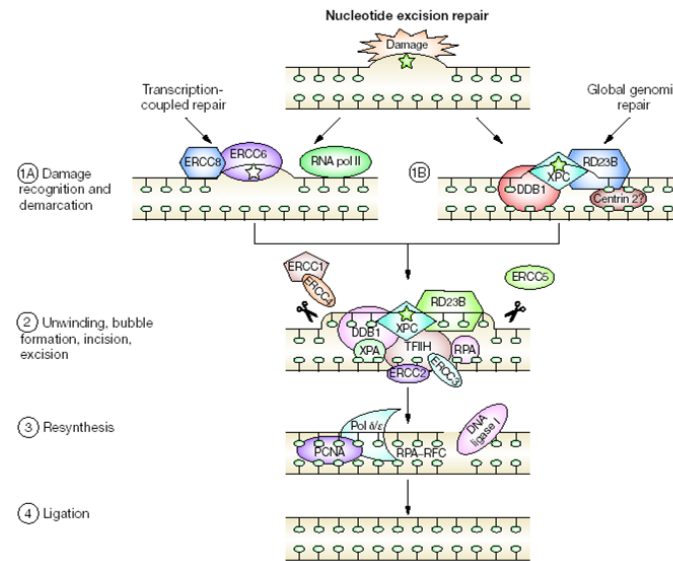


Figure 21 A model of the nucleotide excision repair pathway, which includes global genomic repair (1B) and transcription-coupled repair (1A). The damaged base in the DNA is indicated by a green star. In global genomic repair, the damage is recognized by the heterotrimeric complex of XPC, RD23B and centrin 2, whereas when the damage is in a gene that is being actively transcribed by RNA pol II, ERCC8 and ERCC6 have a crucial role in stalling the transcription process so that repair of the transcribed gene can be initiated. From this point onwards, the repair pathway is common to both mechanisms, and it proceeds by recruiting several other factors, as shown in (2), to effect unwinding, bubble formation of the strand harboring the damage, incision of the strand at discrete points on the 5' and 3' sides, and excision of the fragment containing the damage. In step (3), the gap created by the excision of the damaged strand is resynthesized by DNA pol δ/ϵ , with the help of auxiliary factors such as PCNA and RPA–RFC. Finally, DNA ligase I ligates the newly synthesized fragment to the downstream strand to complete the repair process and yield the repaired product (4). (Rao K. S., 2007).

Over 20 proteins are involved in the NER pathway in mammalian cells (Fig. 21): the XPA protein (and possibly also XPC) initiates repair by recognizing damaged DNA and forming complexes with other proteins involved in the repair process. These include the XPB and XPD proteins, which act as helicases that unwind the damaged DNA. In addition, the binding of XPA to damaged DNA leads to the recruitment of XPF (as a heterodimer with ERCC1) and XPG to the repair complex. The XPA protein binds to replication protein A (RPA) which enhances the affinity of XPA for damaged DNA and is essential for NER. XPF/ERCC1 and XPG are endonucleases, which cleave DNA on the 5' and 3' sides of the damaged site, respectively. This cleavage excises an oligonucleotide consisting of approximately 30 bases. The resulting gap then appears to be filled in by DNA

polymerase δ or ϵ (in association with replication factor C and PCNA) and sealed by ligase (You J. S. et al., 2003). Global genomic NER (GGR) repairs damage throughout the genome, while a specific NER pathway called Transcription Coupled Repair (TCR) repairs genes during active RNA polymerase transcription (Hanawalt P. C., 2002).

A connection between transcription and repair was first suggested by experiments showing that transcribed strands of DNA are repaired more rapidly than nontranscribed strands in both *E. coli* and mammalian cells (Mellon I. & Hanawalt P. C., 1989; Mellon I. et al., 1987). Since DNA damage blocks transcription, this transcription-repair coupling is thought to be advantageous by allowing the cell to preferentially repair damage to actively expressed genes. Although the molecular mechanism of transcription-repair coupling in mammalian cells is not yet known, it is noteworthy that the XPB and XPD helicases are components of a multisubunit transcription factor (called TFIIH) that is required to initiate the transcription of eukaryotic genes. Thus, these helicases appear to be required for the unwinding of DNA during both transcription and nucleotide-excision repair, providing a direct biochemical link between these two processes.

Aim of the study

Retinoic acid (RA), an active derivative of vitamin A, plays a role in regulation of embryonic development, homeostasis and differentiation of adult tissues. RA metabolites, collectively known as retinoids, are well-characterized inhibitors of cancer cell proliferation or inducers of stem cell differentiation. It was demonstrated that retinoids potently inhibit the growth of breast cancer cell lines and can inhibit mammary carcinogenesis in animals models. Among the mechanisms proposed to explain the inhibition of breast cancer cell growth by retinoids there are cyclin D degradation, RAR β induction, inhibition of AP-1 activity and alterations in specific downstream IGF signaling elements. In fact, RA regulates the PI 3-kinase/Akt pathway reducing IRS-1 protein levels and tyrosine phosphorylation (del Rincón S. V. et al., 2003). In breast cancer cell line MCF7 it has been found that RA signaling regulates the expression of many genes that have been implicated in breast carcinogenesis and/or whose expression is indicative for the clinical outcome of breast cancer. Interestingly, RAR α /RAR γ exhibit extensive colocalization of their genomic binding regions with ER α in the vicinity of genes that are antagonistically regulated by estrogen and RA and in breast tumor samples, the expression of RAR targets identified in MCF-7 cells predicts a positive clinical outcome (Hua S. et al., 2009).

Despite extensive studies on RA-induced transcription, it is not known if there is a common set of histone modifications or how the initiation transcription complex is assembled on regulatory regions. The H3 methylation changes reported so far that are associated with activation of the receptor(s) by RA, may be secondary to repression of transcription (Angrisano T. et al., 2011) or induced by the establishment of complex phenotypes, such as stem cell differentiation (Shahhoseini M. et al., 2012; Compe E. & Egly J. M., 2012). Although large DNA domains have been studied and histone modifications have been recorded during development, the mechanism used by RA to activate transcription still remains elusive. To address this issue, we studied two prototypic genes induced by RA: caspase 9 (*CASP9*) and Cyp26A1 (*CYP26A1*). *CASP9* contains a functional RARE located 9.5 Kb downstream of the transcription start site (Donato L. J. & Noy N., 2005), whereas the RA-induced *CYP26A1* expression is driven by a compact RARE-promoter (Balmer J. E. & Blomhoff R., 2002). We

first analyzed recruitment of RA receptor and RNA polymerase II to the promoter and RARE sites. Second, we studied the changes of methylation of lysine 4 (K4) and lysine 9 (K9), following the recruitment on the chromatin sites of 2 demethylating enzymes, LSD1 (KDM1A) and JMJD2A (KDM4A). It is proposed that LSD1 demethylates H3K4me2 or K9me2 and JMJD2A demethylates H3K4-K9 me3 (Kooistra S. M. & Helin K., 2012). Third, we analyzed the formation after RA exposure of specific chromatin-DNA domains that connect the 5' end-promoter-RARE and the 3' end site of the RA-target gene.

Methods

Cells and transfections

Human breast cancer MCF-7 cells were grown at 37°C in 5% CO₂ in Dulbecco's modified Eagle's medium (DMEM) supplemented with phenol red, L-glutamine (2 mM), insulin (10 µg/ml), hydrocortisone (3.75 ng/ml), and 10% fetal bovine serum (FBS, South America origin, Brazil, Invitrogen, Rockville, MD, USA). Cells were provided with fresh medium every 3 days. To evaluate the effect of retinoic acid challenge, cells were grown in phenol red-free DMEM containing 10% dextran–charcoal-stripped FBS for 1 to 3 days, before being challenged with 300 nM retinoic acid for different times according to the experimental needs.

To obtain LSD1 knock down with siRNA, cells were transiently transfected, using a Neon® Transfection System, with siRNA GS23028 (Qiagen Inc., USA) in medium without serum to a final concentration of 10 nM and incubation was continued for 48 h. Scrambled RNA, at the same concentration, was used as negative control. The same procedure was used to get JMJD2A, OGG1 and APE1 knock down with the specific siRNAs (JMJD2A, SR306452C; OGG1, SR303282; APE1, SR300230; OriGene Technologies, Inc., USA). To determine rescue of LSD1 activity in knock down experiments with siRNAs, LSD1 full-length cDNA was inserted into the CMV 3xFLAG expression vector (Sigma-Aldrich, St. Louis, MO, USA). To obtain rescue of JMJD2A activity, cells were transfected with pCMV6-AC-GFP plasmid containing JMJD2A full-length (RG200574, OriGene Technologies, Inc., USA). To assess the transfection efficiency at single cell level, all transfections were traced with pEGFP Vector (Clontech) or with BLOCK-iT Alexa Fluor® Red Fluorescent Control and analysed by FACS.

RNA extraction and qRT-PCR and qPCR

Total RNA was extracted using Triazol (Gibco/Invitrogen). cDNA was synthesized in a 20 µl reaction volume containing 1 µg of total RNA, 100 units of Superscript III Reverse Transcriptase (Invitrogen), and 2 µl random hexamer (20 ng/µl) (Invitrogen). mRNA was reverse-transcribed for 1 h at 50 °C, and the reaction was heat inactivated for 15 min at 70 °C. The products were stored at -20 °C until use. Quantitative reverse Transcription Polymerase Chain Reaction (qRT-

PCR) and Quantitative Polymerase Chain Reaction (qPCR) were performed three times in six replicates on a 7500 Real Time PCR System (Applied Biosystems) using the SYBR Green-detection system (FS Universal SYBR Green MasterRox/Roche Applied Science). The complete list of oligonucleotides used is reported in Table 1.

Chromatin Immuno-Precipitation (ChIP)

Cells were transfected and/or treated as indicated in the legends of the figures. The cells ($\sim 2.5 \times 10^6$ for each antibody) were fixed for 10 minutes at room temperature by adding 1 volume of 2% formaldehyde to a final concentration of 1%, the reaction was quenched by the addition of glycine to a final concentration of 125 mM. Fixed cells were harvested and the pellet was resuspended in 1 ml of Lysis Buffer (10 mM Tris-HCl pH 8.0, 10 mM NaCl, 0.2 % NP40) containing 1X protease inhibitor cocktail (Roche Applied Science). The lysates were sonicated in order to have DNA fragments from 300 to 600 bp. Sonicated samples were centrifuged and supernatants diluted 2 fold in the ChIP Buffer (1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris-HCl pH 8.0). An aliquot (1/10) of sheared chromatin was further treated with proteinase K (4U every 1×10^6 nuclei), extracted with 1 volume of phenol/chloroform/isoamyl alcohol (25:24:1) and precipitated in LiCl 0,4 M/ ethanol 75% to determine DNA concentration and shearing efficiency (input DNA). The ChIP reaction was set up according to the manufacturer's instructions. Briefly, the sheared chromatin was precleared for 2 h with 1 μ g of non-immune IgG (Santa Cruz Biotechnology, Santa Cruz, CA, USA) and 20 μ l of Protein A/G PLUS-Agarose (Santa Cruz Biotechnology) saturated with salmon sperm (1 mg/ml). Precleared chromatin was divided in aliquots and incubated at 4 °C for 16 h with 1 μ g of the specific antibody (for the codes, see below) and non-immune IgG respectively. The immuno-complexes were recovered by incubation for 3 h at 4 °C with 20 μ l of protein-A/G PLUS agarose, beads were washed with wash buffers according to the manufacturer's instructions and immunoprecipitated DNA was recovered through phenol/chloroform/isoamyl alcohol extraction and ethanol precipitation and redissolved in TE buffer (10 mM Tris-HCl, 1mM EDTA, pH 8,0). Samples were subjected to qPCR using the

primers indicated in the legend of the specific figures, primers sequences are reported in Table 1. Real Time-qPCRs were performed using FastStart Universal SYBR Green Master (Rox) (Roche Applied Science) with cycle conditions as follows:

CASP9 Promoter: 95 °C 10 min; 5x (95 °C 45 sec, 68 °C 30 sec, 72 °C 30 sec); 40x (95 °C 45 sec, 65 °C 30 sec, 72 °C 30 sec); 72 °C 10 min.

CASP9 Other regions: 95 °C 10 min; 5x (95 °C 45 sec, 59 °C 30 sec, 72 °C 30 sec); 40x (95 °C 45 sec, 56 °C 30 sec, 72 °C 30 sec); 72 °C 10 min.

CYP26A1 RARE/Promoter region: 95 °C 10 min; 45x (95 °C 45 sec, 56 °C 30 sec, 72 °C 35 sec); 72 °C 10 min.

8-Oxo-7, 8-dihydro-2'-deoxyguanosine (8-oxo-dG) DNA Assay

For 8-oxodG detection, 10⁶ MCF-7 cells were seeded onto glass slides and treated with 200 or 500 nM RA for 15 or 30 min. Control cultures were treated with equivalent vehicle volumes and concentrations. After treatments, the cells were fixed 15 min with 4% paraformaldehyde in PBS. The slides were then washed with TBS/Tween-20 and permeabilized by serial washes in methanol solutions, prior to be washed with TBS/Tween-20, blocked for 1 h at 37°C and incubated with FITC-labeled protein, that binds 8-oxo-dG, for 15 h at 4°C (Biotrin OxyDNA Test, Biotrin, UK). Cover slips were mounted in Moviol and viewed by fluorescence. To obtain LSD1 knock down, cells were transfected with specific or control siRNAs. After 48 h, cells were subjected to different treatments, according to experimental needs, and processed for fluorescence microscopy. For single cell transfection assays, cells were co-transfected with BLOCK-iT Alexa Fluor® Red Fluorescent Control. The efficiency of transfection was 65%+10. All images were captured with AxioCam microscopy (Zeiss) with a 63x objective in the same conditions of brightness and contrast.

Chromosome conformation capture (3C)

The 3C assay was performed as described previously (Dekker J. et al., 2002) with minor adaptations. Briefly: the NcoI restriction enzyme was used (Roche Applied Science). The cells (2.5×10^6) were crosslinked in 12 ml of PBS with 1% formaldehyde for 10 min at room temperature. The reaction was quenched by the addition of glycine to a final concentration of 125 mM. Fixed cells were harvested and the pellet resuspended in 1 ml of ice-cold lysis buffer (the same used for ChIP experiments). Nuclei were washed with 0.5 ml of restriction enzyme buffer (100 mM NaCl, 50 mM Tris-HCl, 10 mM MgCl₂, 1mM Dithioerythritol, pH 7,5 at 37 °C), centrifuged and resuspended in 100 µl of restriction enzyme buffer. SDS was added to a final concentration of 0.1%, and nuclei were incubated at 37 °C for 15 min. Triton X-100 was added to the final concentration of 1% to sequester SDS. Digestion was performed with 100 U of the restriction enzyme at 37 °C for 16 h. The restriction enzyme was inactivated by the addition of SDS to 2% and incubation at 65 °C for 30 min. The reaction was diluted into 1 ml ligation reaction buffer (66 mM Tris-HCl, 5 mM MgCl₂, 5 mM DTT, 1 mM ATP, pH 7,5) and incubated at 16 °C for 18 h with 50 U of T4 DNA Ligase (Roche Applied Science). EDTA (10 mM) was added to stop the reactions. Samples were treated with Proteinase K (200 µg/ml) and incubated for 5 h at 55 °C, and then overnight at 65 °C to reverse the formaldehyde crosslinks. The following day, the DNA was purified by phenol/chloroform/isoamyl alcohol extraction and ethanol precipitation. Samples were redissolved in 20 µl of TE buffer. To prepare a control template, we used a pool of plasmids containing an equimolar amount of the *CASP9* or *CYP26A1* inserts spanning the genomic regions of interest. Five micrograms of plasmid DNA were digested with NcoI in 50 µl of 1x buffer for 8 h at 37 °C and then ligated in 20 µl with 5 U of T4 Ligase at 16 °C for 4 h. The efficiency of digestion at the end of 3C treatment was quantified by real time PCR, amplifying a fragment spanning two NcoI (uncut) in different 3C DNA preparations. Primer sequences are reported in Table 1. PCR were performed using FastStart Taq DNA Polymerase (Roche Applied Science) with cycle conditions as follows:

CASP9 oligo A-F1, F1-L, F2-L, A-L: 95 °C 5 min; 5x (95 °C 45 sec, 55 °C 30 sec, 72 °C 30 sec); 30x (95 °C 45 sec, 52 °C 30 sec, 72 °C 30 sec); 72 °C 10 min.

CASP9 oligo F2-L: 95 °C 5 min; 5x (95 °C 45 sec, 54 °C 30 sec, 72 °C 35 sec); 30x (95 °C 45 sec, 51 °C 30 sec, 72 °C 35 sec); 72 °C 10 min.

PCR products were run on 1.2% agarose gels, stained with ethidium bromide and quantified with the imageJ program (Rasband WS, ImageJ, National Institutes of Health, Bethesda, Maryland, USA, <http://rsb.info.nih.gov/ij/>). The amplified fragments at the end of the procedure were verified by DNA sequence analysis.

Antibodies used for the experiments

RAR α sc-551 (Santa Cruz Biotechnology); PolII 05-623 (Upstate) ; P-Pol II 04-1572 (Upstate); H3K4me2 ab32356 (Abcam); H3K4me3 ab1012 (Abcam); H3K9me2 ab1220 (Abcam); H3K9me3 ab8898 (Abcam); Total H3 ab1791 (Abcam); LSD1 sc-271720 and sc-67272 (Santa Cruz Biotechnology); JMJD2A sc-135065 (Santa Cruz Biotechnology); Anti-FLAG F7425 (Sigma-Aldrich); OGG1 sc-33181 (Santa Cruz Biotechnology); TDG sc-22845 (Santa Cruz Biotechnology); UNG sc-28719 (Santa Cruz Biotechnology); APE1 ab-194 (Abcam); RPA sc-14691 (Santa Cruz Biotechnology); XPG sc-73274 (Santa Cruz Biotechnology); SSU72 sc-69613 (Santa Cruz Biotechnology); RXR α sc-553 (Santa Cruz Biotechnology); Normal rabbit IgG sc-2027 (Santa Cruz Biotechnology); Normal mouse IgG sc-2025 (Santa Cruz Biotechnology).

LSD1 Activity/Inhibition Assay

MCF-7 cells were serum starved for 2 days and treated with RA for the indicated times. Untreated or treated MCF-7 cells were washed three times with ice-cold PBS pH 7.4, scraped and lysed in buffer 1 containing 20 mM Tris pH 7.5, 10 mM KCl, 2 mM EDTA, 2 mM MgCl₂. After 10 seconds at 12,000 xg at 4 °C, the pellets were resuspended in a buffer 2 containing 20 mM Tris pH 7.5, 400 mM NaCl, 2 mM EDTA, 1 mM MgCl₂. After 10 minutes at 12,000 x g at 4 °C, the supernatants were assayed for LSD1 activity. The activity was measured by the

EpiQuik™ Histone Demethylase (H3K4 Specific) Activity/Inhibition Assay Kit, according to the manufacturer's instructions (Epigentek, P -3017; USA).

Statistical analysis

All data are presented as mean \pm standard deviation in at least three experiments in triplicate ($n \geq 9$). Statistical significance between groups was determined using Student's t test (matched pairs test or unmatched test were used as indicated in figure legends). All tests were performed using the JMP Statistical Discovery™ software by SAS, Statistical Analysis Software.

Table S1

	PRIMERS FOR mRNA	Locus
mRNA Fw	5' -CCATATGATCGAGGACATCCA-3'	<i>CASP9</i>
mRNA Rev	5' -GACTCCCTCGAGTCTCCAGAT-3'	<i>CASP9</i>
mRNA Fw	5' -GCAGCCACATCTCTGATCACT-3'	<i>CYP26A1</i>
mRNA Rev	5' -TGTTGTCTTGATTGCTCTTGC-3'	<i>CYP26A1</i>
mRNA Fw	5' -ATGGAAGGGGTGGAGGAT-3'	<i>JMJD2A</i>
mRNA Rev	5' -TTTGTCCCTATTCGGTGCTT-3'	<i>JMJD2A</i>
18SF	5' -TCCCCATGAACGAGGAATTC-3'	<i>h18S</i>
18SR	5' -GGCCTCACTAAACCATCCAA-3'	<i>h18S</i>
mRNA Fw	5' -CGTCATGGTCTTATCAACTTCG-3'	<i>LSD1</i>
mRNA Rev	5' -CAGCTACATAGTTTCCTTTGCG-3'	<i>LSD1</i>
mRNA Fw	5' -TTTGATGACCCGCAAAGG-3'	<i>OGG1</i>
mRNA Rev	5' -AACAACTTCCCTGTTTCACTT-3'	<i>OGG1</i>
mRNA Fw	5' -ACACTCAAGATCTGCTCTTGG-3'	<i>APE1</i>
mRNA Rev	5' -TGCCGTAAGAACTTTGAGTGG-3'	<i>APE1</i>
	PRIMERS FOR ChIP	Locus
ChIP Prom Fw	5' -CGCGCCGCCCCAGAACACG-3'	<i>CASP9</i>
ChIP Prom Rv	5' -GCGGGCGGGACGGTAACG-3'	<i>CASP9</i>
ChIP RARE Fw	5' -GGAGGCTGAGGCGGGTTTATC-3'	<i>CASP9</i>
ChIP RARE Rev	5' -GGTGGAGTGCGGTGGTGTG-3'	<i>CASP9</i>
ChIP RARE 5' Fw	5' -TTTTTCTTGCCGCTTTTCTCATC-3'	<i>CASP9</i>
ChIP RARE 5' Rev	5' -TGCTCAGCCTCCCCAGTAGC-3'	<i>CASP9</i>
ChIP PolyA1 Fw	5' -CTAGTAGGCCCCGGTTTGCTGAT-3'	<i>CASP9</i>
ChIP PolyA1 Rev	5' -AGTAACGCGTCTTCCTGAGTGGTG-3'	<i>CASP9</i>
ChIP PolyA2 Fw	5' -GTTCTCACCTTGTGCTTCCCT-3'	<i>CASP9</i>
ChIP PolyA2 Rev	5' -TTGCCTATTTCTTCCCTCATTTTG-3'	<i>CASP9</i>
Intron 2 Fw/CasF1	5' -GCAAGTACTCAATAATGTTTACC-3'	<i>CASP9</i>
Intron 2 Rev/CasE2	5' -TCAACAAAAATTCACCAAACTCA-3'	<i>CASP9</i>
Exon 13 Fw	5' -CCTCCTTGACCAGGCTAATTAC-3'	<i>TGFBI</i>
Exon 13 Rev	5' -GGCTGCAACTTGAAGGTTGTG-3'	<i>TGFBI</i>
ChIP RARE Fw	5' -GCGGAACAAACGGTTAAAGA-3'	<i>CYP26A1</i>
ChIP RARE Rev/C	5' -TGATCGCAGCCAGGAAGA-3'	<i>CYP26A1</i>
	PRIMERS FOR 3C	Locus
Cas A	5' -GTCTGTACATGTTTCACTACAATGC-3'	<i>CASP9</i>
Cas B	5' -CCACCTGCAGCTCTTCCA-3'	<i>CASP9</i>
Cas C	5' -CAGTCATCCGGAGACCTAAACC-3'	<i>CASP9</i>
Cas D	5' -AGATGCCCAGCACTATGCTAAG-3'	<i>CASP9</i>
Cas E1	5' -GGTGCCTGGCAAATAGCAAT-3'	<i>CASP9</i>
Cas E2	5' -TCAACAAAAATTCACCAAACTCA-3'	<i>CASP9</i>
Cas F1	5' -GCAAGTACTCAATAATGTTTACC-3'	<i>CASP9</i>
Cas F2	5' -GGGGTGATTCCAGCATAGGTTC-3'	<i>CASP9</i>
Cas G	5' -CTTCTGCGTCTGAACTTGAACC-3'	<i>CASP9</i>
Cas H	5' -CTGAAAGAAAGAAGGCTGGATGC-3'	<i>CASP9</i>
Cas I	5' -AGGAAGGAAACTACCGCTTGCT-3'	<i>CASP9</i>
Cas L	5' -TCCCTTTCACCGAAACAGCA-3'	<i>CASP9</i>
A	5' -TCACCACTGGACATATTCTTACC-3'	<i>CYP26A1</i>
B	5' -GTGGGGTTTGAAGCGCTGG-3'	<i>CYP26A1</i>
C	5' -TGATCGCAGCCAGGAAGA-3'	<i>CYP26A1</i>
D	5' -TATTTACAGTGGAGCTGGCC-3'	<i>CYP26A1</i>
E	5' -CTCAGGTTTGAACATTCAAGC-3'	<i>CYP26A1</i>
F	5' -TGTTTGCCAAAGTTTTCAACC-3'	<i>CYP26A1</i>

Supplementary Table S1. Complete list of DNA oligonucleotides used for PCR. On the left is shown the primer identification tag; on the centre is shown the DNA sequence; on the right are shown the specific genes or loci corresponding to the specific primers.

Results

Recruitment of RA receptor and activation of RNA polymerase II at RA-target promoters

Transcription allows the DNA to be copied into an RNA molecule, to transmit its genetic information and it is regulated by many transcriptional factors that bind to the regulatory regions of genes. It can be induced by nuclear receptors, a group of transcriptional factors activated by lipophilic substances. We used a model of transcription induced by retinoic acid.

The biological activity of RA is mediated by its binding to retinoic acid receptors ($RAR\alpha$, $RAR\beta$ and $RAR\gamma$). These belong to the type II group of nuclear receptors: the receptor is already in the nucleus and after binding to RA it detaches itself from co-repressors and function as heterodimers with retinoid X receptors (RXRs), targeting specific sites on DNA known as retinoic acid responsive elements (RAREs). In *CASP9*, a RARE element is localized in intron II, 9.5 Kb downstream to the transcription start site (Fig. 22); 2. At *CYP26A1*, RARE is contiguous with the promoter (-150 bp from the TSS) (Fig. 22). These sites are essential for RA induction of transcription of the two genes (Donato L. J. & Noy N., 2005; Ray W. J. et al., 1997).

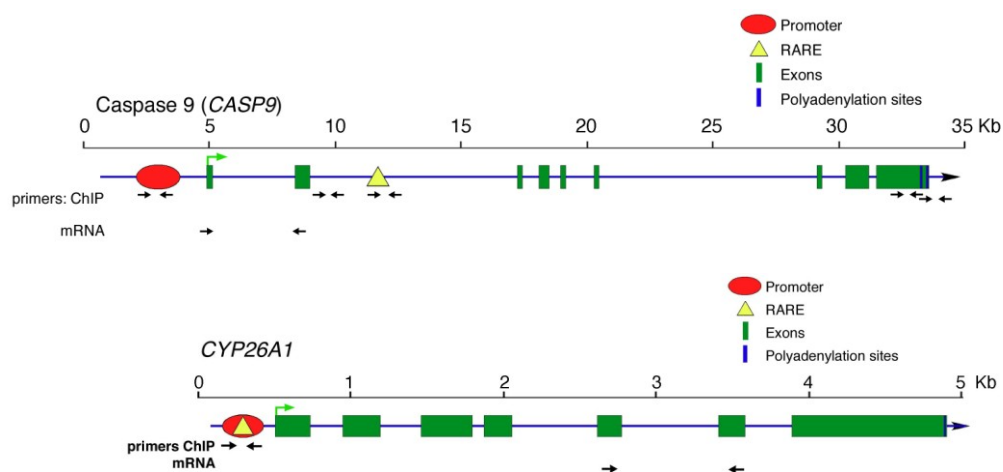


Figure 22 Structure of *CASP9* and *CYP26A1* genes. The direction of transcription is indicated by a green arrow. The green boxes indicate exons; the red, blue and yellow elements indicate the promoters, PolyA and RARE sequences respectively; the black arrows indicate the primers used for ChIP and mRNA experiments.

We studied induction by RA of *CASP9* and *CYP26A1* genes (Fig. 22) by exposing MCF-7 cells to RA and measuring mRNA levels at different times following stimulation. Fig 23 shows that both mRNAs accumulate 30 min after RA exposure. After a transient reduction at 60 min their levels reached a maximum 4-6 h after RA exposure.

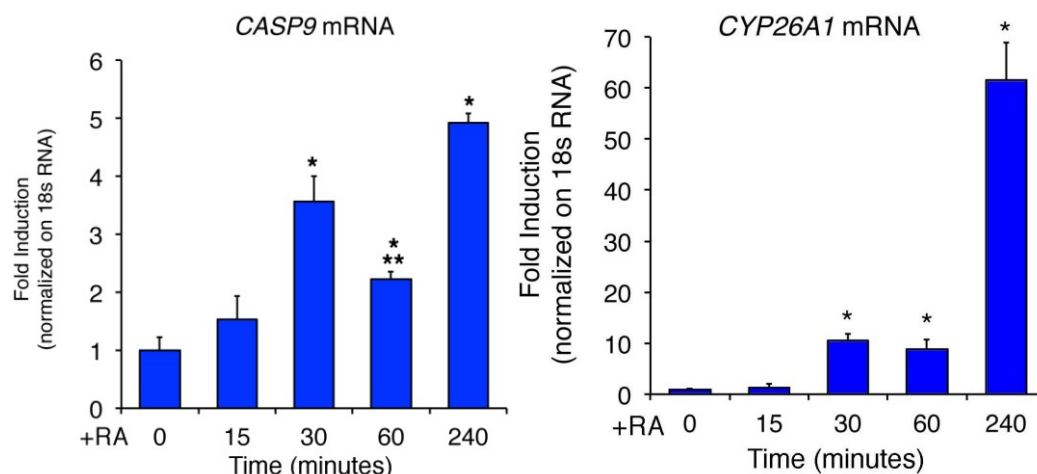


Figure 23 RA induction of *CASP9* and *CYP26A1* mRNA. Total RNA was prepared from MCF-7 hormone-starved or stimulated with 300 nM retinoic acid for 15, 30, 60 and 240 minutes and analyzed by qPCR with specific primers (Fig.1) to *CASP9* and *CYP26A1* mRNA normalized to 18s RNA levels. The statistical analysis derived from at least 3 experiments in triplicate ($n \geq 9$; Mean \pm SD); * $p < 0.01$ (matched pairs t test) compared to RA-unstimulated sample, ** $p < 0.01$ (matched pairs t test) comparing 30 to 60 min. of RA exposure.

To monitor recruitment of the retinoic acid receptor to the promoter and RARE elements after RA stimulation, we assessed the timing of association of retinoic acid receptor to the chromatin of *CASP9* and *CYP26A1* by ChIP. We included in our analysis very early times after RA induction (min) to detect the earliest chromatin changes induced by the hormone. Figure 24 shows that RAR α is rapidly (15 min) recruited to the RARE element and to the upstream promoter of *CASP9*. We noticed that the levels of RAR α recruited at the promoter and RARE chromatin were not stable, but oscillated between 15 and 60 min after RA exposure and stabilize after 4 hours.

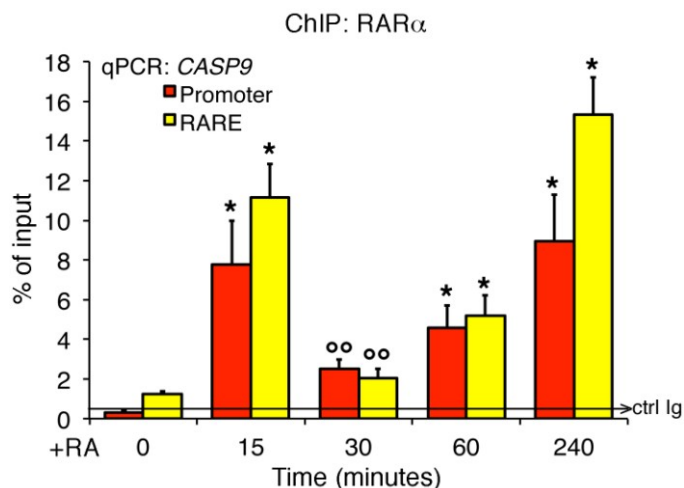


Figure 24 RA induction of RAR α accumulation on retinoic responsive elements (RARE) and promoter of *CASP9* gene. qChip analysis of RA dependent occupancy of retinoic acid receptor alpha (RAR α) on the promoter and retinoic responsive elements (RARE) of *CASP9*. MCF7 cells were stimulated with 300 nM retinoic acid (RA) for 15, 30, 60 and 240 minutes. The chromatin was immunoprecipitated with antibodies directed against RAR α . The black, horizontal, line indicates the percent of input from a control ChIP (Ab: non immune serum). The statistical analysis derived from at least 3 experiments in triplicate ($n \geq 9$; Mean \pm SD); * $p < 0.01$ (matched pairs t test) compared to RA-unstimulated sample; ° $p < 0.01$ (matched pairs t test) comparing 15 to 30 min.

The RAR α and RXR were also recruited to the RARE/promoter of *CYP26A1* (Fig. 25). RAR α accumulates on RARE/Promoter of *CYP26A1* with a peak after 30 min from induction and a subsequent oscillation up to 4 hours. RXR recruitment instead remains stable until 90 minutes after hormone exposure. Recruitment of RAR α and RXR was associated with accumulation of total and ser5-phosphorylated RNA polymerase II (Pol II and P-Pol II, respectively) at the RARE and promoter regions of *CASP9* and *CYP26A1* (Figs. 25, 26 and 27).

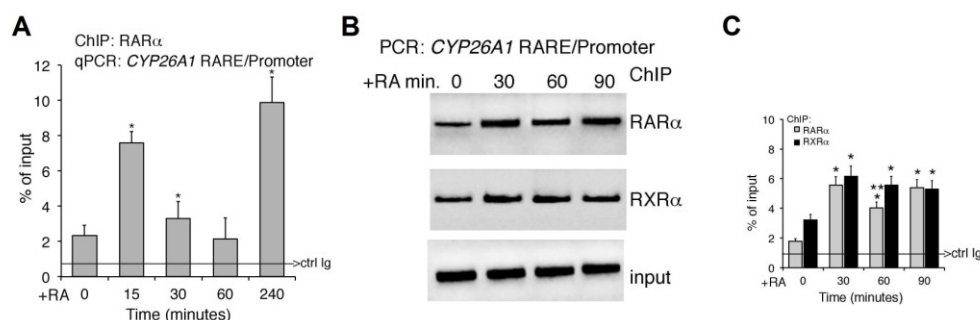


Figure 25 A, B, C. RA induction of accumulation of RAR α and RXR α on retinoic responsive elements (RARE) and promoter of *CYP26A1* gene. qChIP analysis of RA dependent occupancy of RAR α and RXR α on the promoter/RARE of *CYP26A1* gene. MCF7 cells were stimulated with 300 nM retinoic acid (RA) for 15, 30, 60 and 240 minutes. The chromatin was immunoprecipitated with antibodies directed against RAR α or RXR α . The panel **B** represents a semiquantitative PCR assay of the ChIP, the panel **C** the qPCR in an independent ChIP experiment. The black, horizontal, line in **A** and **C** indicates the percent of input from a control ChIP (Ab: non immune serum). The statistical analysis derived from at least 3 experiments in triplicate ($n \geq 9$); * $p < 0.01$ (matched pairs t test) compared to RA-unstimulated sample; ** $p < 0.01$ (matched pairs t test) comparing 30 to 60 min.

Figure 5 shows the results of a ChIP experiment to evaluate the levels of Pol II and P-Pol II on promoter and RARE regions of *CASP9*. As expected, Pol II and P-Pol II accumulated preferentially at the promoter relative to the RARE. Note that recruitment of total Pol II and P-Pol II to RARE oscillated synchronously with the recruitment of RAR α . P-Pol II progressively accumulated at the *CASP9* promoter with the time of RA stimulation. The levels of P-Pol II at the *CASP9* RARE dipped at 30 min and then increased over the next 210 min.

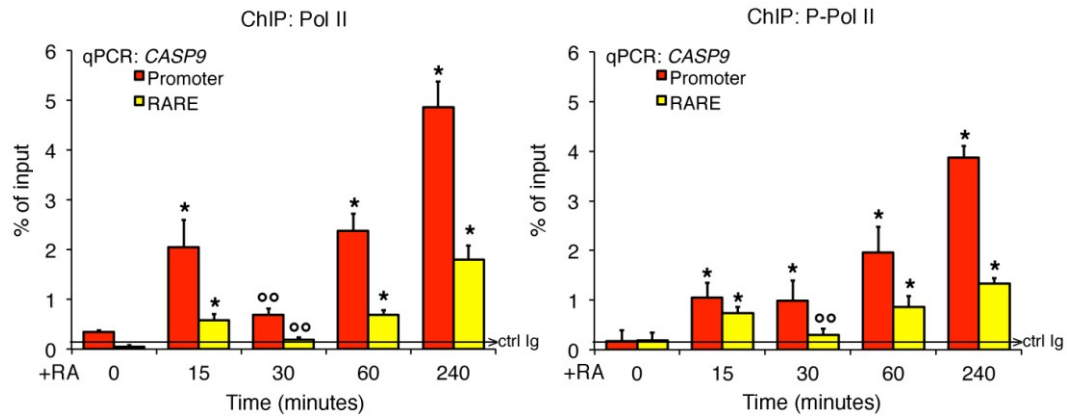


Figure 26 RA induction of accumulation of RNA polymerase II on retinoic responsive elements (RARE) and promoter of *CASP9* gene. qChip analysis of RA dependent occupancy of RNA polymerase II (Pol II) and phosphorylated RNA polymerase II (P-Pol II) on the promoter and retinoic responsive elements (RARE) of *CASP9*. MCF7 cells were stimulated with 300 nM retinoic acid (RA) for 15, 30, 60 and 240 minutes. The chromatin was immunoprecipitated with antibodies directed against Pol II or P-Pol II. The black, horizontal, lines indicate the percent of input from a control ChIP (Ab: non immune serum). The statistical analysis derived from at least 3 experiments in triplicate ($n \geq 9$; Mean \pm SD); * $p < 0.01$ (matched pairs t test) compared to RA-unstimulated sample; ° $p < 0.01$ (matched pairs t test) comparing 15 to 30 min.

ChIP analysis of Pol II and P-Pol II recruitment on RARE/Promoter chromatin of *CYP26A1* showed that Pol II and P-Pol II accumulates between 30 and 60 minutes after RA induction. Their levels tend to decrease at 90 min (Fig. 27).

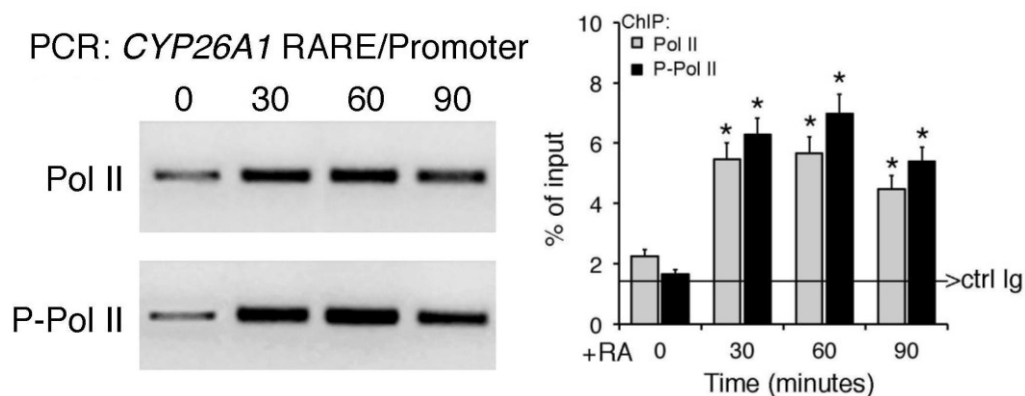


Figure 27 RA induction of accumulation of RNA polymerase II on retinoic responsive elements (RARE) and promoter of *CYP26A1* gene. ChIP analysis of RA dependent occupancy of RNA polymerase II (Pol II) and phosphorylated RNA polymerase II (P-Pol II) on the promoter/RARE of *CYP26A1* gene. MCF7 cells were stimulated with 300 nM retinoic acid (RA) for 30, 60 and 90 minutes. The chromatin was immunoprecipitated with antibodies directed against the large subunit of RNA polymerase II (Pol II and phosphorylated P-Pol II). The left panel represents a semiquantitative PCR assay of the ChIP, the right panel is the qPCR in an independent ChIP experiment. The black, horizontal, line indicates the percent of input from a control ChIP (Ab: non immune serum). The statistical analysis derived from at least 3 experiments in triplicate ($n \geq 9$); * $p < 0.01$ (matched pairs t test) compared to RA-unstimulated sample.

Histone H3 K4 and K9 methylation marks induced by RA

Histone methylation is an important type of chromatin modification that contributes to the control of gene expression through influencing chromatin compaction or signalling to other protein complexes. Histone lysine residues can be mono-, di-, or trimethylated, and different degrees of methylation on one particular site could be linked to different functional outcomes. Histone lysine methylation seems to be a quite stable modification and stably methylated histone lysine residues sustain the establishment and propagation of different patterns of gene expression in the same genome. Thus, methylated histone lysine residues have been considered “epigenetic marks” (Jenuwein T. & Allis C. D., 2001).

Methylation of lysine 4 in histone H3 (H3K4) marks transcribed loci, whereas dimethyl-lysine 9 in the same histone (H3K9me2) is associated with transcription silencing (Loh Y. H. et al., 2007; Vaute O. et al., 2002; Wang H. et al., 2001). To find the histone marks modified by RA exposure on chromatin, we analyzed the methylation profiles of lysines 4 (K4) and 9 (K9) in histone H3 in cells after treatment with RA. ChIP analysis was performed with H3 mK9- and mK4-specific antibodies at the promoter-start site, the RARE element, two polyA addition sites located at the 3' end of the *CASP9* and segments of the gene 2Kb distant from the TSS. As a result of methylation and de-methylation events induced by RA, the levels of methylated K9 or K4, normalized to total histone H3, were selectively modified shortly after RA exposure. Fig. 28 shows that promoter-associated H3K4 me2 and me3 were transiently de-methylated 15 min after RA challenge and then progressively re-methylated (Fig. 28A and B). Rapid de-methylation of H3K4 at RARE was also observed, but the rate of re-methylation was slower than at the promoter. RA induced a transient loss of H3K9me2 and H3K9me3, followed by accumulation of H3K9me2 but not of H3K9me3 (Fig. 28C and D). These cyclical methylation-demethylation events were strikingly synchronous on the RARE and promoter regions of *CASP9*. In contrast, H3 methylation was unaffected by RA at a site 2Kb downstream to RARE in intron II or in a non-RA induced gene, *TGFBI* (Fig. 28E).

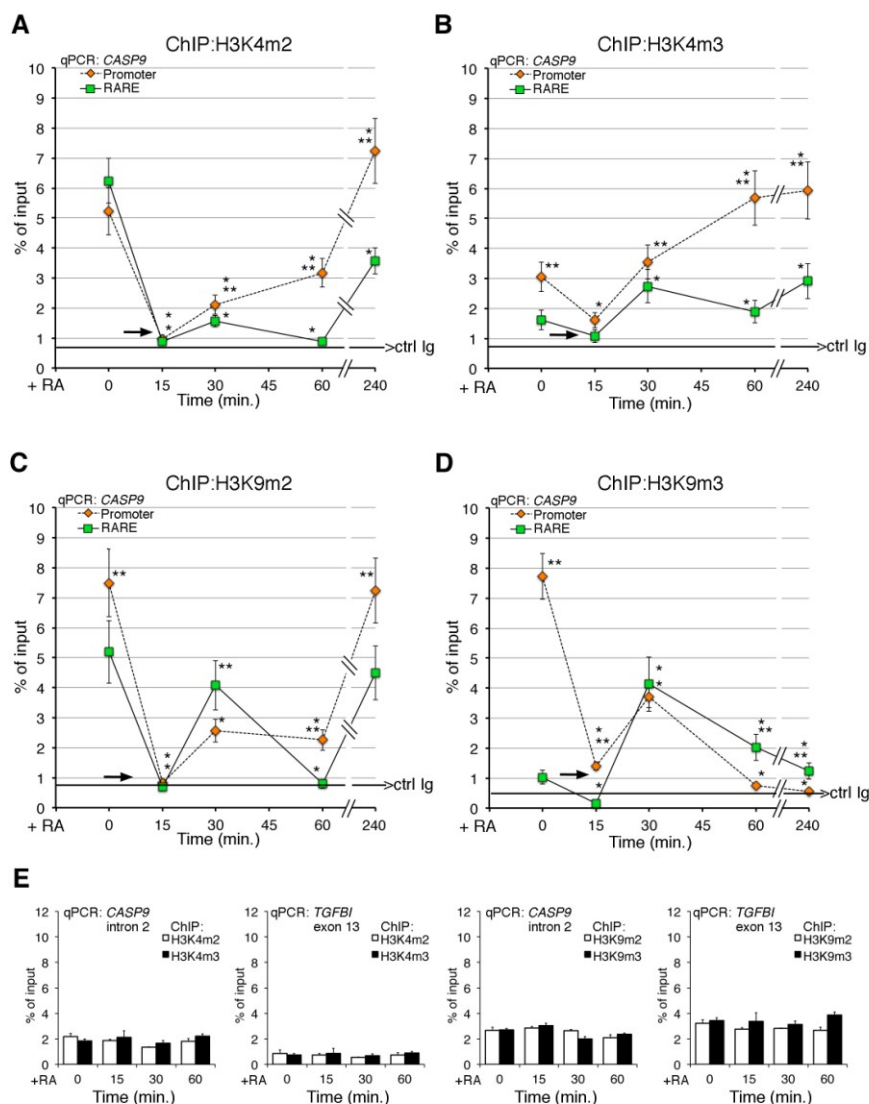


Figure 28 Methylation-demethylation cycles of histone H3K4/K9 induced by RA on *CASP9* promoter-RARE chromatin. MCF7 cells were serum starved and exposed to 300 nM RA at the indicated times (0, 15, 30, 60 and 240 min). qChIP was carried out using specific antibodies recognizing H3K4me3, H3K4me2, H3K9me3 and H3K9me2. **A, B.** H3K4me2 and H3K4me3 occupancy on *CASP9* promoter and RARE. **C, D.** H3K9me2 and H3K9me3 occupancy on *CASP9* promoter and RARE. H3K4me2 and H3K4me3 were transiently de-methylated 15 min upon RA (black arrows) and progressively methylated 30-60 min later. H3K9 was selectively demethylated as shown by loss of H3K9me3 and accumulation of H3-K9me2. The statistical analysis derived from at least 3 experiments in triplicate ($n \geq 9$; Mean \pm SD); * $p < 0.01$ (matched pairs t test): compared to the RA-unstimulated sample; ** $p < 0.01$ (student t test): comparison between each regions at same time. **E.** ChIP analysis of *CASP9* II intron and of *TGFBI* exon 13 (non RA-induced gene), in cells exposed to RA for 15, 30 and 60 minutes.

A similar methylation-de-methylation cycle was observed at the *CYP26A1* promoter–RARE chromatin (Fig. 29). The levels of H3K4m2 and H3K4m3 show an oscillation at 15 min from RA induction and a progressive increase up to 4 hours, while the levels of H3K9m2 and H3K9m3 have a peak at 30 min and then a strong reduction.

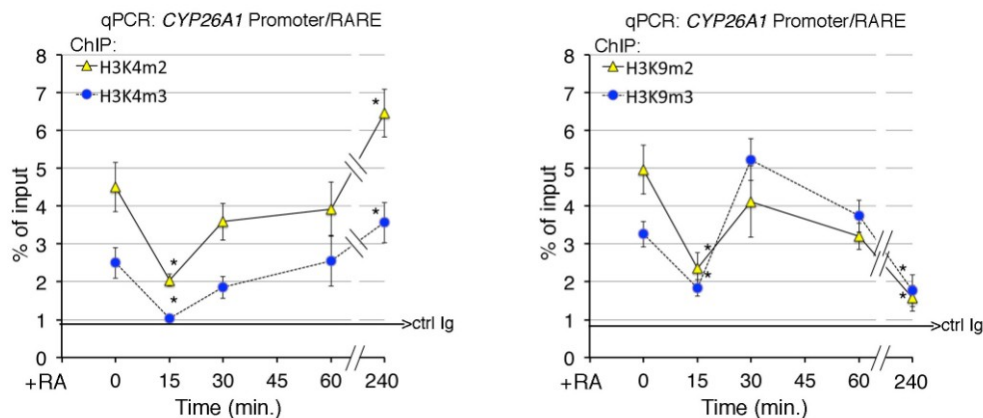


Figure 29 Methylation-demethylation cycles of histone H3K4/K9 induced by RA on *CYP26A1* promoter/RARE chromatin. MCF7 cells were serum starved and exposed to 300 nM RA at the indicated times (15, 30, 45, 60 and 240 min). qChIP was carried out using specific antibodies recognizing H3K4me3, H3K4me2, H3K9me3 and H3K9me2. The statistical analysis derived from at least 3 experiments in triplicate ($n \geq 9$); * $p < 0.01$ (matched pairs t test): compared to the RA-unstimulated sample.

We also probed the 3' end of *CASP9* gene, where two major polyA addition sites are located. H3K4me2 and me3 and H3K9me2 at the polyA1 and polyA2 sites also underwent transient de-methylation. H3K9me3 was permanently de-methylated at the polyA2 site, but was essentially unchanged at the polyA1 site (Fig. 30). We conclude that the polyA1 and polyA2 sites undergo methylation changes similar to those seen on the promoter and RARE, raising the possibility that these regions are functionally and physically associated in a unique chromatin domain.

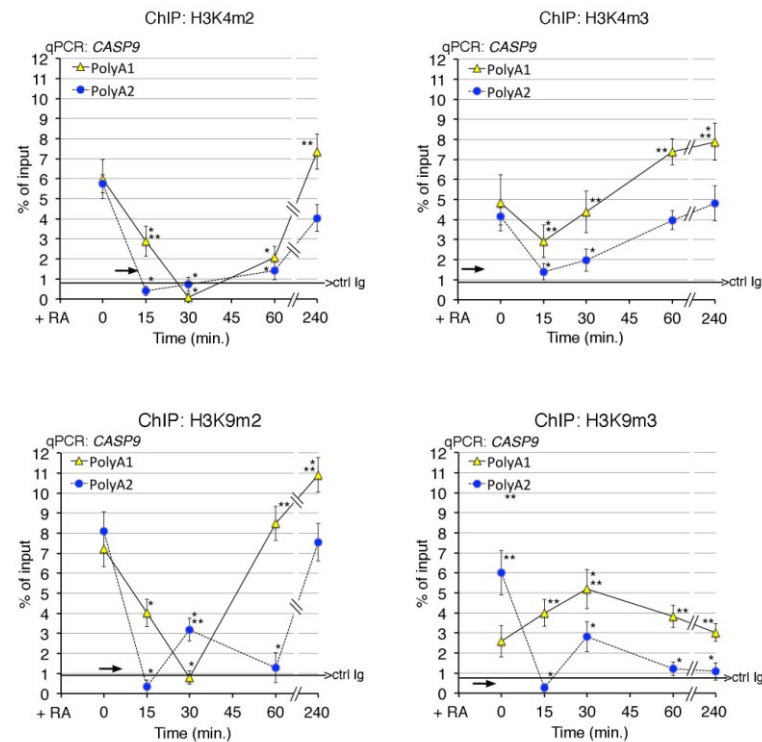


Figure 30 Methylation-demethylation cycles of histone H3K4/K9 induced by RA on *CASP9* polyA1 and polyA2 chromatin. MCF7 cells were serum starved and exposed to 300 nM RA at the indicated times (15, 30, 45, 60 and 240 min). qChIP was carried out using specific antibodies recognizing H3K4me3, H3K4me2, H3K9me3 and H3K9me2. The upper panels show the H3K4me2 and H3K4me3 occupancy on polyA1 and polyA2 of *CASP9* gene. The lower panels show the H3K9me2 and H3K9me3 occupancy on polyA1 and polyA2 of *CASP9* gene. The statistical analysis derived from at least 3 experiments in triplicate ($n \geq 9$); * $p < 0.01$ (matched pairs t test): compared to the RA-unstimulated sample; ** $p < 0.01$ (student t test): comparison between each region at same time.

These dynamic methylation changes of K4 and K9 induced by RA exposure suggest that K4 and K9 de-methylation enzymes are also recruited to *CASP9* and *CYP26A1* chromatin. H3K9me3 and H3K4me3 can be de-methylated by enzymes of the Jumonji C (JMJC) family (Loh Y. H. et al., 2007; Wissmann M. et al., 2007), whereas H3K9me2 and H3K4me2 are de-methylated by LSD1 (Forneris F. et al., 2005; Metzger E. et al., 2005). Figure 31A and B shows recruitment of both LSD1 and JMJD2A histone de-methylases to the RARE elements and promoters of *CASP9* and *CYP26A1* following RA treatment. Notably, the kinetics of

recruitment of LSD1 and JMJD2A parallels the kinetics of loss of the H3K4 and H3K9 methyl groups. The methylation changes (H3K9me2/3) and the kinetics of LSD1 and JMJD2A recruitment on *CYP26A1* chromatin are very similar to those seen at RARE. We believe that this similarity is due to the fact that the promoter and RARE are physically contiguous in the *CYP26A1* but are dissociated in *CASP9*.

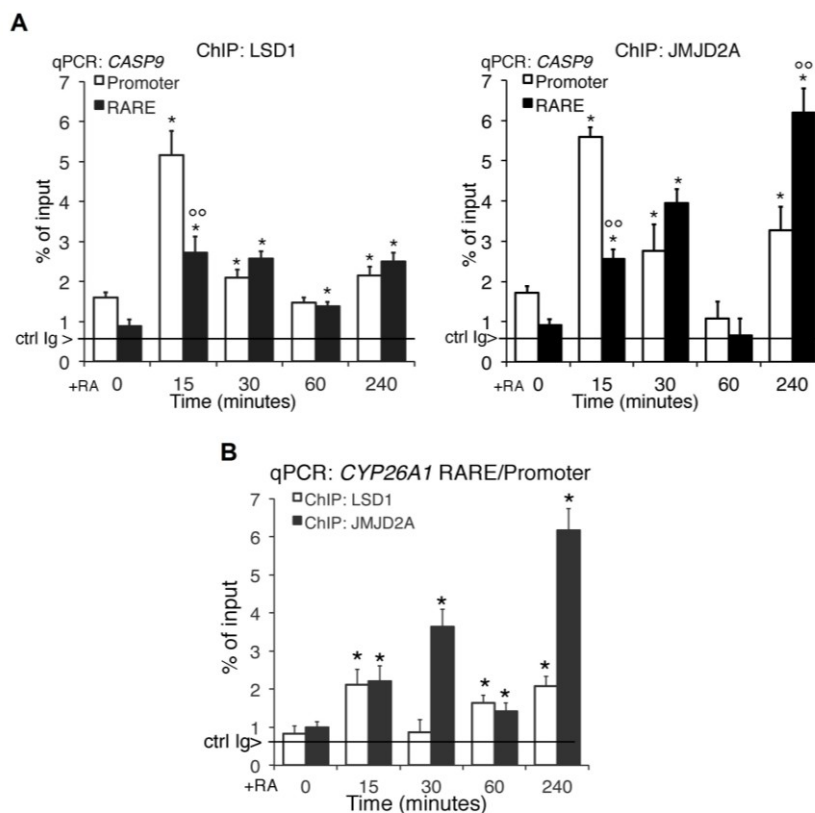


Figure 31 Recruitment of LSD1 and JMJD2A to the promoter and RARE of *CASP9* and *CYP26A1* genes. MCF7 cells were serum starved and exposed to 300 nM RA at the indicated times (15, 30, 60 and 240 min). qChIP was carried out using specific antibodies recognizing LSD1 and JMJD2A. The panel **A** shows the time course of the recruitment of LSD1 and JMJD2A on RARE and on the promoter sequences of *CASP9* while the panel **B** shows the recruitment of LSD1 (white) and JMJD2A (black) on the promoter/RARE sequence of *CYP26A1* analyzed by qPCR. The black, horizontal, line in each plot indicates the percent of input from a control ChIP (Ab: non immune serum). The statistical analysis derives from at least 3 experiments in triplicate ($n \geq 9$; Mean \pm SD); * $p < 0.01$ (matched pairs t test) compared to RA-unstimulated sample; $^{\circ\circ}p < 0.01$ (student t test): comparison between Promoter and RARE regions.

To demonstrate that both lysine de-methylases were necessary for RA-induced transcription, we knocked down LSD1 and JMJD2A with specific siRNAs, and induced the cells with RA (Fig. 32). Knock down of either demethylase significantly reduced expression of *CASP9* (Fig. 32B) and *CYP26A1* (Fig. 32C). Expression of the de-methylases in silenced cells restored *CASP9* and *CYP26A1* expression, indicating that RA induction of transcription requires the concerted action of both LSD1 and JMJD-domain containing de-methylases.

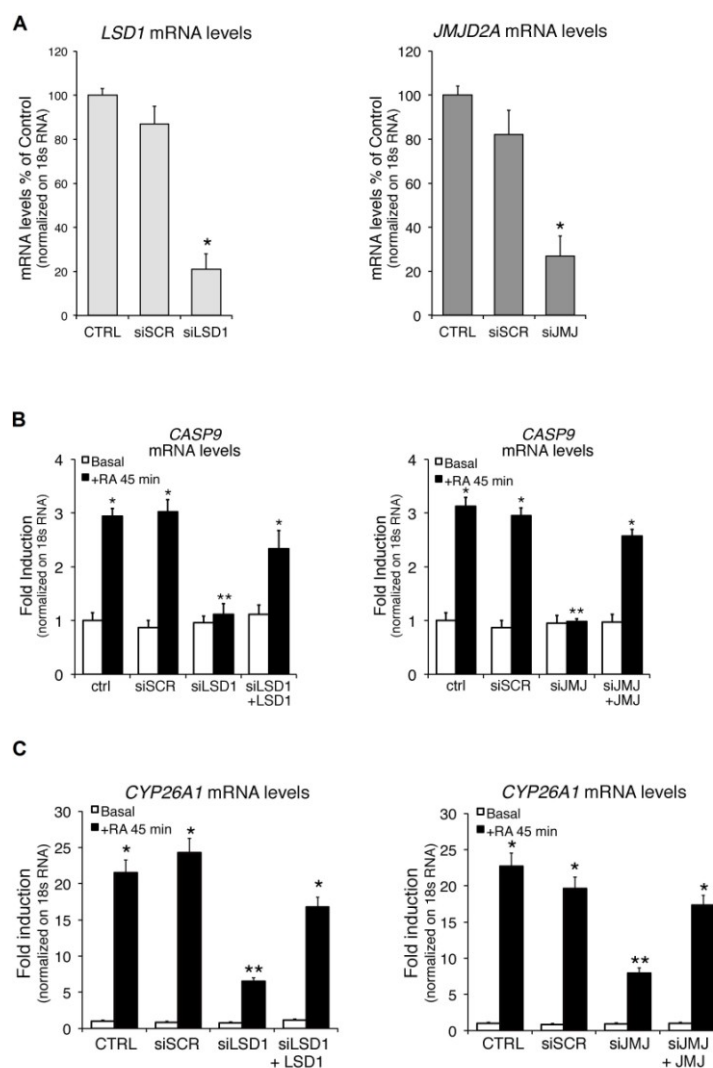


Figure 32 Depletion of the histone demethylases, LSD1 and JMJD2A blocks RA-induced transcription. MCF7 were transiently transfected with LSD1 siRNA or JMJD2A siRNA with or without the wild type protein expressing vectors. The efficiency of siRNA treatments was measured by qPCR using primers for LSD1 and JMJD2A

mRNAs (panel A). Transfection efficiency was monitored by FACS (Alexa Fluor or co-transfected pEGFP Vector). After 48 h, total RNA was prepared from control cells (starved) or RA induced cells (300 nM RA for 45 min) and analyzed by qPCR with specific primers to *CASP9* (panel B) or *CYP26A1* (panel C) mRNA. The statistical analysis derives from at least 3 experiments in triplicate ($n \geq 9$); * $p < 0.01$ (matched pairs t test) compared to RA-unstimulated sample; ** $p < 0.01$ (student t test): comparison between siSCR and specific siRNA.

Importantly, silencing of either LSD1 or JMJD2A inhibited RA-induced de-methylation, as shown by retention of di- and tri-methylated H3K4 and H3K9 at the *CASP9* promoter (Fig. 33). Thus we do not find different substrate specificities for LSD1 and JMJD2A. Instead, our data indicate that the two enzymes cooperate to de-methylate H3K4 and H3K9.

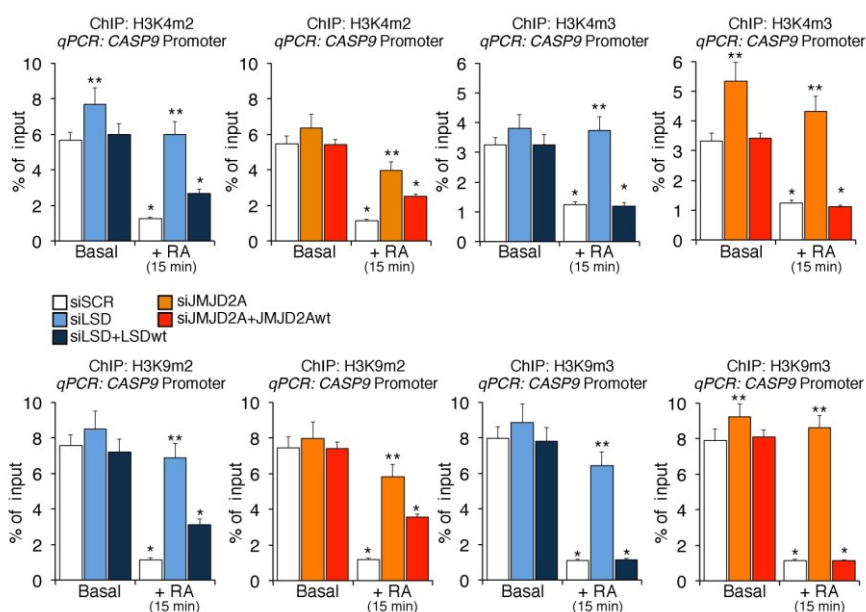


Figure 33 Depletion of LSD1 or JMJD2A inhibits the methylation changes induced by RA. qChIP was performed on cells transfected with LSD1 siRNA or JMJD2A siRNA and induced with 300 nM RA for 15 minutes. qChIP was carried out using specific antibodies recognizing H3K4me3, H3K4me2, H3K9me3 and H3K9me2. The statistical analysis derived from at least 3 experiments in triplicate ($n \geq 9$; Mean \pm SD); * $p < 0.01$ (matched pairs t test) compared to RA-unstimulated sample; ** $p < 0.01$ (student t test): comparison between SCR and specific siRNAs. Transfection efficiency was monitored by FACS (Alexa Fluor or co-transfected pEGFP Vector).

To explore further the relationship between H3K9 and H3K4 methylation and LSD1, we over-expressed an N-terminal dominant-negative mutant (T110A) of LSD1 (LSD1ALA). This mutant is still enzymatically active, but is unable to target transcription factors (Ambrosio R. et al., 2013; Amente S. et al., 2010a). The LSD1ALA mutant protein was defective in binding to the promoter or RARE elements of *CASP9* or *CYP26A1* (Fig. 34): In basal conditions, the amount of bound LSD1ALA is lower than LSD1WT and 30 min after RA induction LSD1WT is recruited on RARE/promoter regions of the two genes, while LSD1ALA is not.

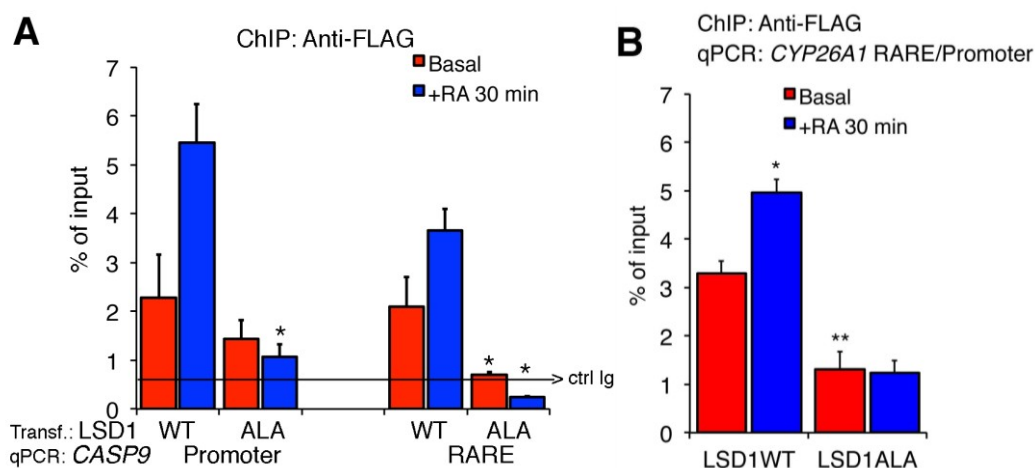


Figure 34 Recruitment of wild type and mutant LSD1 (LSD1ALA) to the *CASP9* and *CYP26A1* RARE/promoter. MCF7 were transiently transfected with LSD1 vectors (WT or mutant), starved (Basal) or treated with 300 nM RA for 30 minutes and were analyzed by qChIP using Anti-FLAG antibodies to recognize the recombinant LSD1. The panel **A** shows the recruitment of the flagged LSD1WT and LSD1ALA on RARE and on the promoter sequences of *CASP9*. The black, horizontal line indicates the percent of input from a control ChIP (Ab: non immune serum). The panel **B** shows the recruitment of the flagged LSD1 WT and LSD1ALA on promoter/RARE chromatin of *CYP26A1*. The statistical analysis derives from at least 3 experiments in triplicate ($n \geq 9$); * $p < 0.01$ (matched pairs t test) compared to RA-unstimulated sample; ** $p < 0.01$ (student t test): comparison between LSD1WT and LSD1ALA.

LSD1ALA also inhibited activation of *CASP9* or *CYP26A1* transcription upon RA induction (Fig. 35): the transcriptional response in the presence of LSD1ALA is strongly reduced.

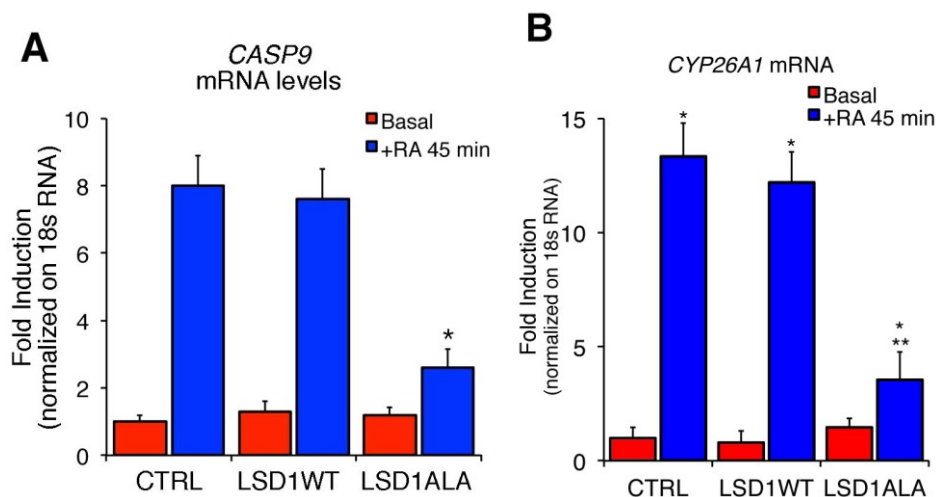


Figure 35 Targeted demethylation by LSD1 is essential for RA-dependent transcription. LSD1ALA inhibits RA induced *CASP9* and *CYP26A1* transcription. Total RNA was prepared from MCF7 transiently transfected with LSD1WT or LSD1ALA. After 48 h, mRNAs from control cells (starved) or RA induced cells (300 nM RA for 45 min) were analyzed by qPCR with specific primers to *CASP9* (A) and *CYP26A1* (B) mRNA.

The methylation levels both H3K4 and H3K9 me2/3 were already low in the absence of RA (Fig. 36) because LSD1ALA was constitutively active and not inducible by RA (Fig. 37). Although the basal methylation levels of H3K4 and H3K9 (me2/3) were high in LSD1 and JMJD2A depleted cells (Fig. 33) and low in LSD1ALA expressing cells (Fig. 36), the methylation-demethylation cycle was abolished in both cases. We do note that in LSD1ALA expressing cells late remethylation of H3K4 and H3K9 (me2/3) occurs (Fig. 36).

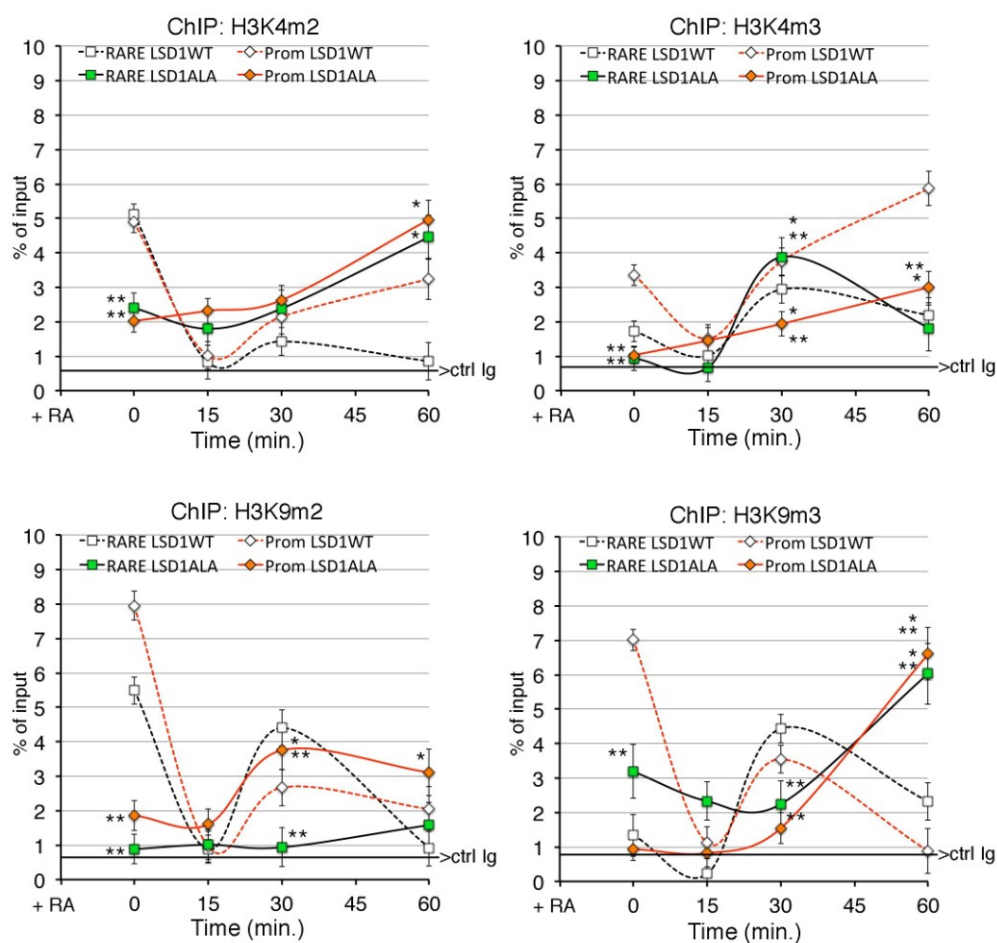


Figure 36 LSD1ALA inhibits H3K4 and H3K9 demethylation. H3K4m2, H3K4m3, H3K9m2 and H3K9m3 occupancy on promoter and RARE of *CASP9* in LSD1ALA expressing cells. The statistical analysis derived from at least 3 experiments in triplicate ($n \geq 9$; Mean \pm SD); * $p < 0.01$ (matched pairs t test): compared to the RA-unstimulated sample; ** $p < 0.01$ (student t test): comparison between control and the LSD1ALA expressing cells at same time.

The H3K4 and H3K9 methylation marks are very similar to those found in RA-exposed cells depleted of LSD1 or JMJD2A exposed to RA (Fig. 33). We noticed that both depletion of wild-type LSD1 and expression of the LSD1ALA eliminated H3K9 and H3K4 massive demethylation at the very early times of RA-induced transcription and inhibited mRNA accumulation, although the LSD1ALA mutant was still catalytically active (Fig. 37A). Note that the LSD1ALA mutant did not respond to RA (Fig. 37A).

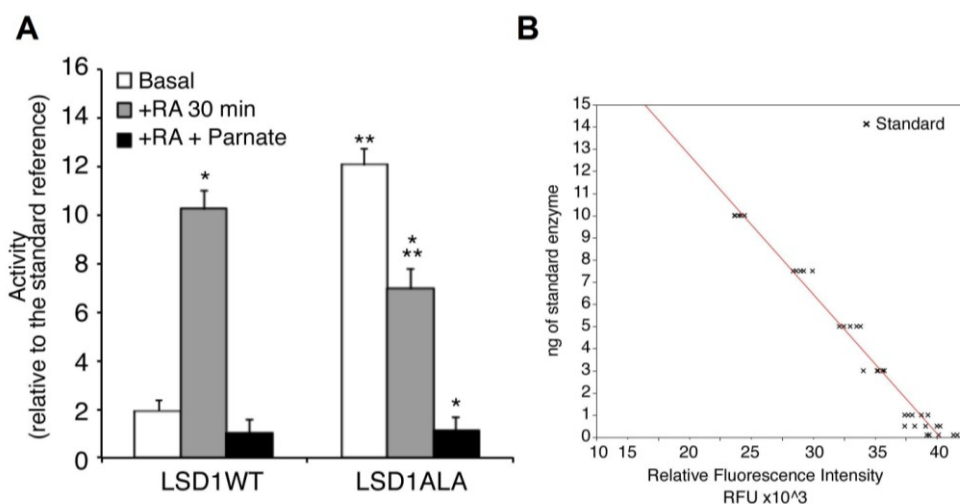


Figure 37 Activity of the LSD1ALA mutant. Cells transfected with LSD1WT-FLAG or the LSD1ALA-FLAG mutant were exposed to RA for 30 min in the presence or absence of parnate (1 μ M), a monoaminoxidase inhibitor. Total cell extracts were prepared and analyzed for LSD1 activity with a fluorescent H3K4me2 substrate (EpiQuik™ Histone Demethylase Activity/Inhibition Assay Kit). A standard curve was generated by diluting the purified enzyme (**B**). The statistical analysis derives from at least 3 experiments in triplicate ($n \geq 9$); * $p < 0.01$ (matched pairs t test) compared to RA-unstimulated sample; ** $p < 0.01$ (student t test): comparison between LSD1WT and LSD1ALA.

We also measured the H3K4 and H3K9 methylation changes on the polyA1 and polyA2 chromatin in cells expressing LSD1ALA. Fig. 38 shows that, although the basal levels of H3K4 and H3K9 me2 and me3 were lower in LSD1ALA expressing cells exposed to RA, demethylation of H3K4me3 and K9me2 and me3 was significantly inhibited similarly to the RARE promoter chromatins (Fig. 28). It is notably that methylation changes at the polyA2 site were mostly affected by expression of LSD1ALA.

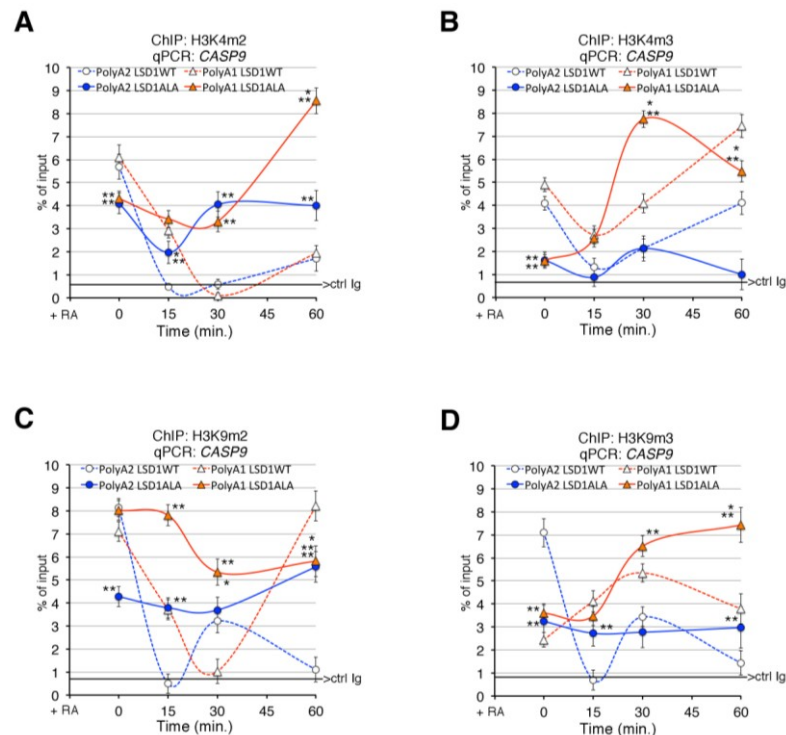


Figure 38 Expression of the LSD1ALA mutant inhibits H3K4 and H3K9 methylation cycles on *CASP9* polyA1 and polyA2 chromatin sites. Cells transfected with the LSD1WT-FLAG or the LSD1ALA-FLAG mutant were analyzed 36 h later by ChIP with specific antibodies to H3K4me2 (A), H3K4me3 (B) and H3K9me2 (C) and H3K9me3 (D). The statistical analysis derived from at least 3 experiments in triplicate ($n \geq 9$); * $p < 0.01$ (matched pairs t test): compared to the RA-unstimulated sample; ** $p < 0.01$ (student t test): comparison between control and the LSD1ALA expressing cells.

Collectively, these data indicate that the recruitment of both LSD1 and JMJD2A to the chromatin of *CASP9* and *CYP26A1* leads to demethylation of H3K4 and H3K9 me2 and me3 at the RARE, promoter and polyA2 sites (Figs. 28 and 30) in cells exposed to RA. These localized demethylation events by both demethylases are essential for the induction of transcription of *CASP9* and *CYP26A1* by RA (Fig. 35).

Recruitment of base (BER) or nucleotide (NER) excision repair enzymes to the RARE-promoter chromatin following RA induction

Transcription process exposes the DNA template to damage by genotoxic agents and generates potentially harmful DNA structures that are prone to mutagenesis and recombination. Gene activation sometimes also requires transient and localized DNA damage at promoters that must be repaired. It was recently reported that NER enzymes are recruited to promoter and RARE elements. NER is essential for the formation of discrete chromatin loops 6 h after RA exposure and RA-induced of transcription (Le May N. et al., 2012). Similarly, transcription-induced recruitment of BER enzymes (such as 8-oxoguanine-glycosylase, OGG1) to MYC E box target DNA or to estrogen responsive elements has been described (Amente S. et al., 2010b). Activation of a Fe⁺⁺ dioxygenase (JMJD2A) and a FAD oxidase (LSD1) at the same chromatin sites (ERE or E box-promoters) triggers local oxidation. Oxidized guanine (8-oxo-dG) is recognized by OGG1 (Amente S. et al., 2010a; Perillo B. et al., 2008). Therefore, we have searched for appearance of 8-oxo-G foci and recruitment of OGG1 on RARE chromatin after RA addition. That oxidation of guanine also occurs after RA induction of transcription is shown in Fig. 39A. We observed a rapid (15 min) nuclear accumulation of 8-oxo-G in discrete foci in MCF7 cells exposed to RA, demonstrating that RA can induce guanine oxidation in chromatin of target cells. As predicted, production of 8-oxo-G foci was inhibited by LSD1 knockdown (Fig. 39B).

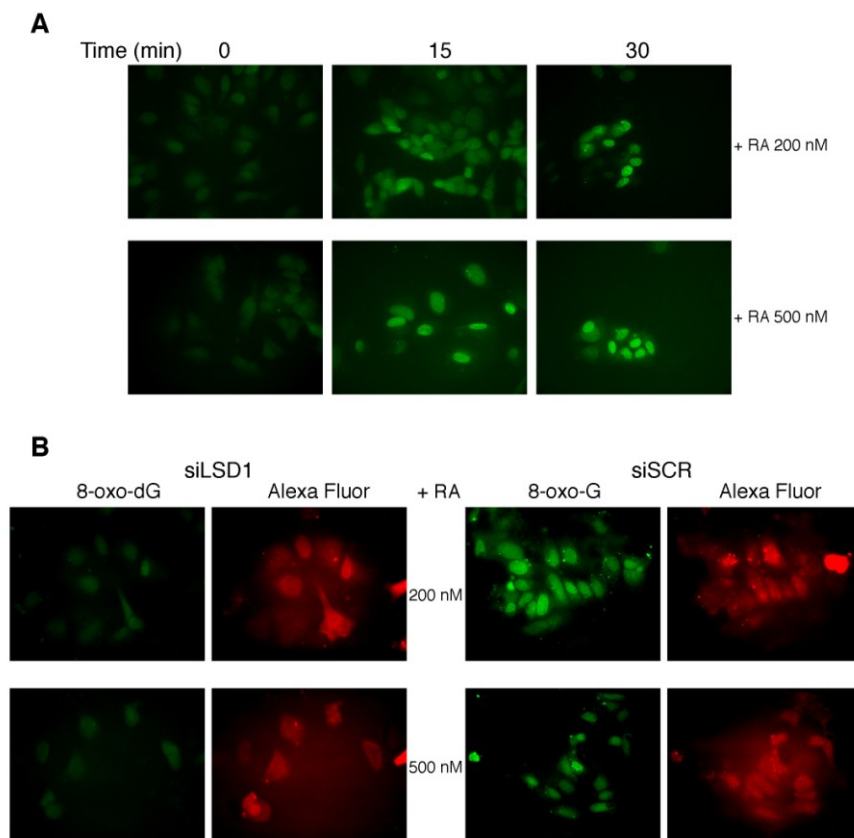


Figure 39 RA induces 8-oxo-dG foci through LSD1-dependent mechanism. A. Time course of 8-oxo-dG staining in cells exposed to RA. MCF7 cells seeded onto glass slides, were starved and treated with 200 nM or 500 nM RA for 15 and 30 min, fixed, and analyzed for the presence of 8-oxo-dG. The 8-oxo-dG signal (green fluorescence) was quantified by ImageJ 1.43 (NIH). Positive cells, containing a signal 2 S.D. higher than controls (CTRL), were 42 and 57% after 15 minutes and 67 and 72 % after 30 minutes of 200 or 500 nM of RA respectively. **B.** LSD1 knockdown inhibits RA-induced 8-oxo-dG accumulation. Treatment with 200 or 500 nM of RA for 30 minutes induce 8-oxo-dG foci only in the cells transfected with the control siRNA (right panels). Transfection efficiency was monitored using Alexa Fluor Red Fluorescent Control.

ChIP analysis showed that OGG1 was recruited to the promoter and RARE elements of *CASP9* (Fig. 40) 15 min following RA treatment. At 30 min, occupancy of these sites by OGG1 markedly decreased. At 240 min, OGG1 was detected at the RARE element but not at the promoter.

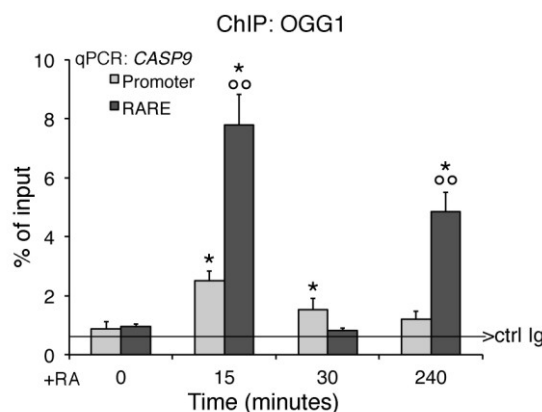


Figure 40 Recruitment of OGG1 to *CASP9* chromatin following RA induction. MCF7 cells, starved or treated with RA for 15, 30 and 240 min, were analyzed by qChIP using specific antibodies recognizing the 8-oxoguanine–DNA glycosylase-1 (OGG1). It is shown the recruitment of OGG1 to the *CASP9* promoter, and RARE sequences. The statistical analysis derived from at least 3 experiments in triplicate ($n \geq 9$; Mean \pm SD); * $p < 0.01$ (matched pairs t test) compared to RA-unstimulated sample; ^{oo} $p < 0.01$ (student t test): comparison between two amplicons.

Similar oscillations of OGG1 binding to the promoter/RARE of *CYP26A1* was also seen (Fig. 43). Complexes enucleated by OGG1 may be important not only for the repair of oxidized lesions but also for assembly of transcription initiation complexes at RA-, estrogen- or Myc-dependent promoters (Huffman J. L. et al., 2005; Perillo B. et al., 2008; Amente S. et al., 2010a; Fong Y. W. et al., 2013). To dissect the components of the OGG1 complex that could link repair and RA-induced transcription, we probed for other BER enzymes that associate with a RARE element or its cognate promoter after RA induction. Specifically, we monitored the recruitment of: 1. the APurinic site Endonuclease1, APE1, which recognizes the apurinic site generated by OGG1 and cleaves the phosphodiester backbone, immediately 5' to the site; 2. thymine DNA glycosylase (TDG), which is required for base excision repair of deaminated methylcytosine, a frequent product of base oxidation; and 3. Uracil Glycosylase (UNG), which removes uracil or oxidized cytosine. Figures 41 and 43 show that all these enzymes are recruited to the promoter and to RARE chromatin 15 min following RA induction

similar to the recruitment of RAR α and Pol II (compare Figs. 40 and 41 with Fig. 22).

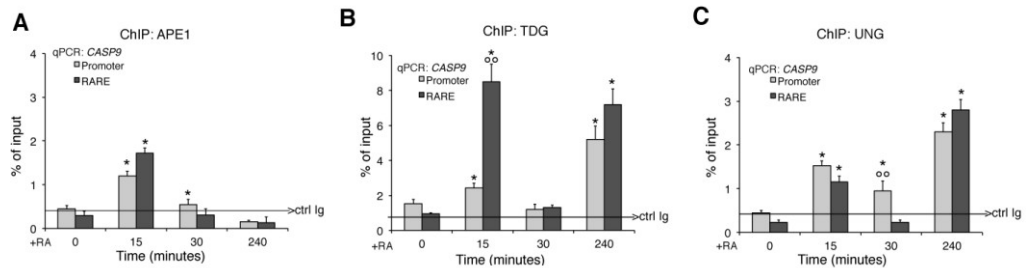


Figure 41 Recruitment of BER enzymes to *CASP9* chromatin following RA induction. MCF7 cells, starved or treated with RA for 15, 30 and 240 min, were analyzed by qChIP using specific antibodies recognizing the AP endonuclease (APE1), Thymine-DNA glycosylase (TDG), Uracil-DNA glycosylase (UNG). **A, B, C** show the recruitment of APE1, TDG and UNG to the *CASP9* promoter, and RARE sequences. The black, horizontal, line in each plot, indicates the percent of input from a control ChIP (Ab: non immune serum). The statistical analysis derived from at least 3 experiments in triplicate ($n \geq 9$; Mean \pm SD); * $p < 0.01$ (matched pairs t test) compared to RA-unstimulated sample; °° $p < 0.01$ (student t test): comparison between two amplicons.

We knocked down two of these enzymes (OGG1 and APE1; Fig. 42 A and B) and asked if this impacted on RA-induced transcription. Figs. 42 C, D, E and F summarize of these experiments. Our results clearly indicate that depletion of these BER enzymes significantly reduced RA-induced transcription.

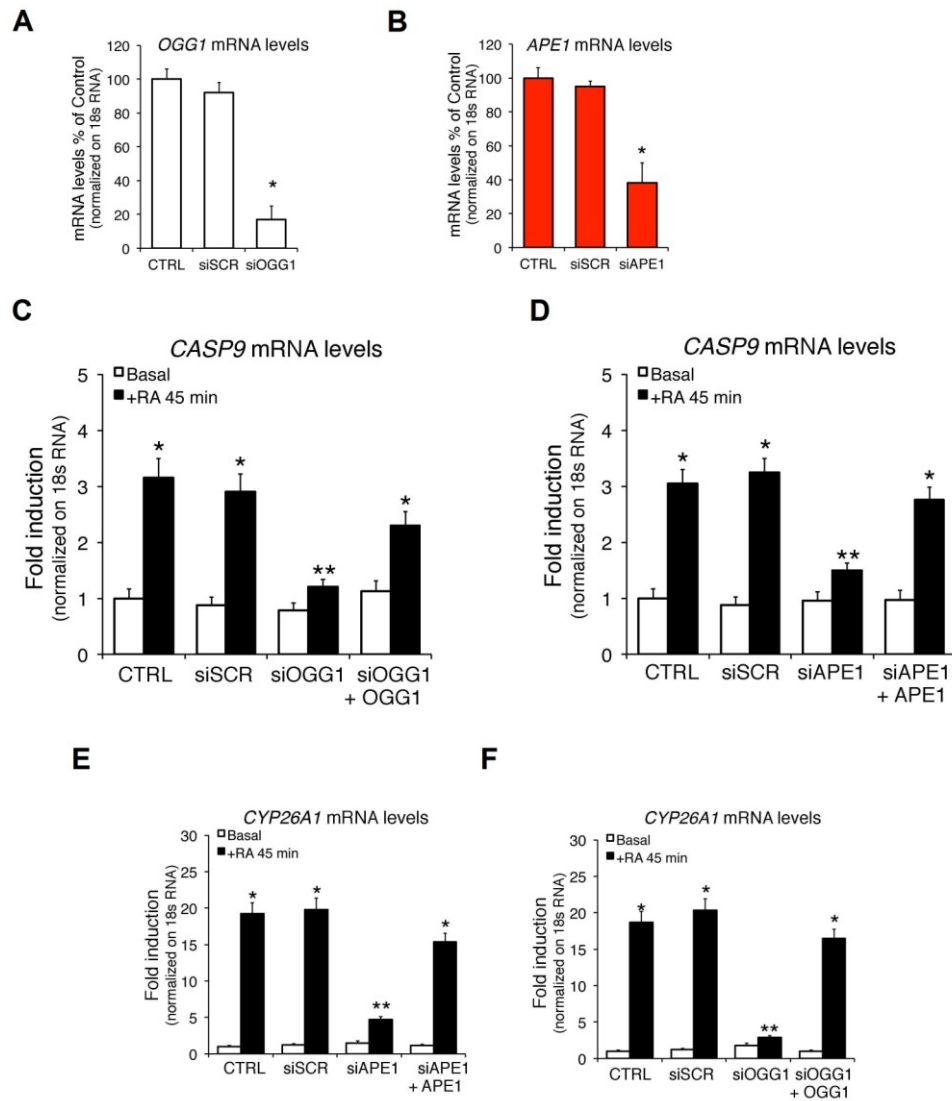


Figure 42 Knockdown of OGG1 and APE1 inhibits RA-induced transcription.

Panels **A** and **B** show the levels of OGG1 and APE1 mRNAs in cells exposed to specific targeting siRNAs, respectively. **C**, **D**, **E**, **F**. Serum-deprived MCF7 cells were treated for 45 minutes with 300 nM RA and specific siRNA targeting OGG1 or APE1; *CASP9* and *CYP26A1* expression levels were quantified by qPCR. To assess the transfection efficiency cells were co-transfected with pEGFP Vector (Clontech) and analyzed by FACS. The statistical analysis derived from at least 3 experiments in triplicate ($n \geq 9$; Mean \pm SD); * $p < 0.01$ (matched pairs t test) compared to RA-unstimulated sample; ** $p < 0.01$ (student t test): comparison between siSCR and specific siRNA.

Also for *CYP26A1* we evaluated the recruitment of BER and NER enzymes and we found that all enzymes were selectively recruited on *CYP26A1* promoter/RARE chromatin after 30 min from RA induction. Next there is an accumulation of NER enzymes XPG and RPA.

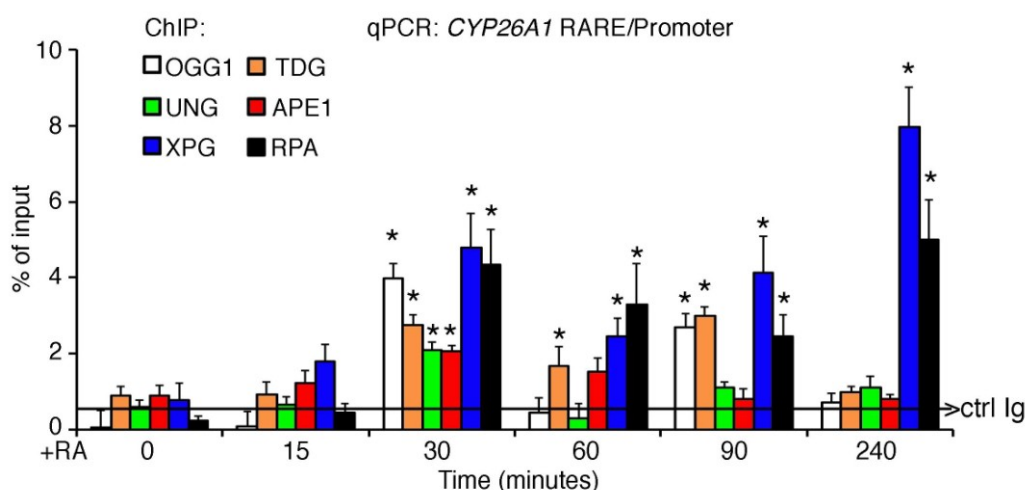


Figure 43 Recruitment of BER and NER enzymes to *CYP26A1* chromatin induced by RA. MCF7 cells, starved or treated with RA for 15, 30 and 240 min, were analyzed by qChIP using specific antibodies recognizing the 8-oxoguanine–DNA glycosylase-1 (OGG1), AP endonuclease (APE1), Thymine-DNA glycosylase (TDG), Uracil-DNA glycosylase (UNG), XPG and RPA. The black, horizontal, line indicates the percent of input from a control ChIP (Ab: non immune serum). The statistical analysis derived from at least 3 experiments in triplicate ($n \geq 9$); * $p < 0.01$ (matched pairs t test) compared to RA-unstimulated sample.

We also probed for a NER enzyme (XPG) and for Replication Protein A (RPA) on RARE by ChIP (Fig. 44). As previously shown by others (Le May N. et al., 2012) XPG and RPA selectively accumulated at RARE chromatin following 15 min of RA stimulation. These authors reported recruitment of NER and BER enzymes to the RARE or other inducible promoters 3 to 4 hours after hormonal induction, a period corresponding to maximal accumulation of specific mRNA levels (Le May et al., 2012). Our results show that BER and NER accumulate early at the RARE and promoter, shortly before we could detect mRNA

accumulation (Fig. 23). The modifications we describe mark the first productive transcription cycle of *CASP9* and *CYP26A1* induced by RA.

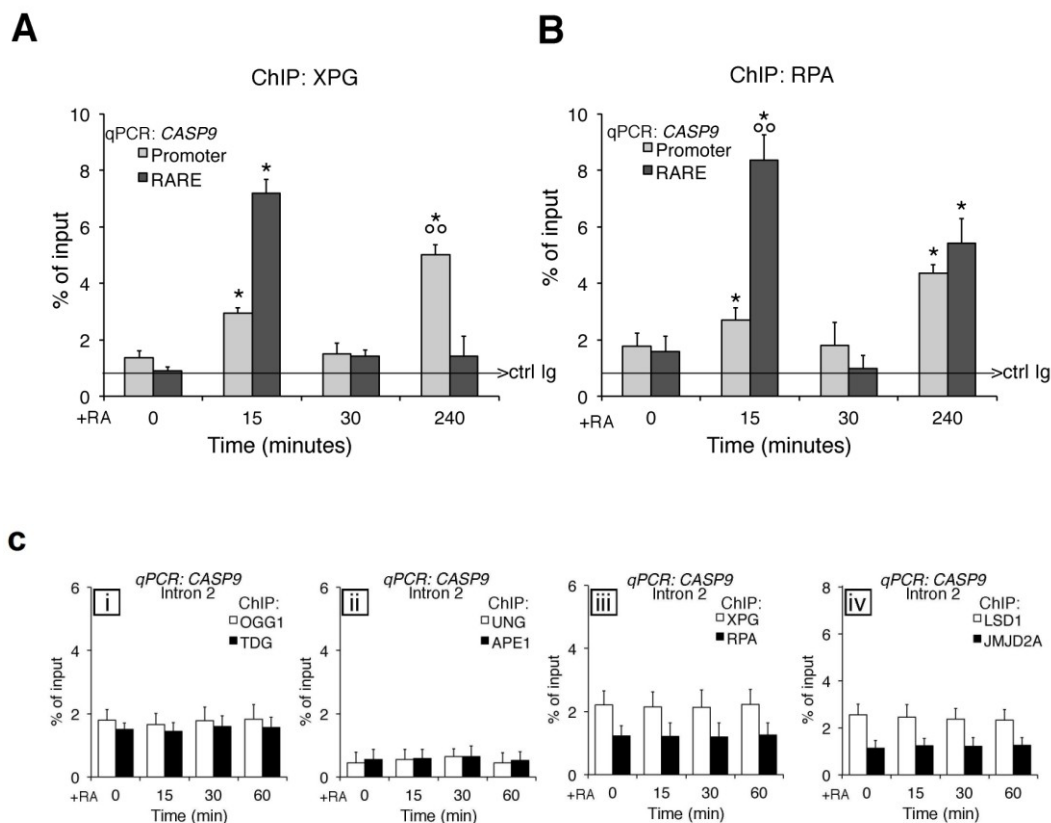


Figure 44 Recruitment of NER enzymes to *CASP9* chromatin following RA induction. **A, B.** MCF7 cells, starved or treated with RA for 15, 30 and 240 min, were analyzed by qChIP using specific antibodies recognizing XPG and RPA. **C** shows the ChIP analysis of *CASP9*, II intron, in cells exposed to RA for various periods with specific antibodies to OGG1, TDG (i); UNG, APE1 (ii); XPG, RPA (iii); LSD1, JMJD2A (iv). The statistical analysis derived from at least 3 experiments in triplicate ($n \geq 9$; Mean \pm SD); * $p < 0.01$ (matched pairs t test) compared to RA-unstimulated sample; $^{\circ\circ}p < 0.01$ (student t test): comparison between two amplicons.

Formation of dynamic chromatin loops governing the selection of 5' and 3' borders of RA-induced transcription units

The data shown above indicate that the *CASP9* and *CYP26A1* promoter, RARE and polyA addition sites undergo similar changes in histone H3 K4-K9 methylation changes and accumulate BER and NER enzymes after RA treatment. This coordination is consistent with the idea that these regions are physically associated after induction. Note that the *CASP9* RARE and 5' start sites are 9.5 Kb apart and the polyA site is 22 Kb to the 3' end of RARE (Fig. 22). Recall that the methylation status H3K4 and H3K9 was not modified at chromatin neighboring these sites (2 Kb at the 5' and 3' end); BER enzymes were not recruited to these sites in RA-treated cells (Figs. 28E and 44C). These data suggest that RA induces early changes of specific chromatin domains that bring the promoter transcription start site, the RARE and the polyA addition sites into close proximity.

To understand how the gene structure is organized during transcription reducing the complexity of the system it was necessary to work on synchronized transcriptions. To find the relevant chromatin domains assembled in response to RA, we systematically analyzed the structure of *CASP9* (Fig. 45) and *CYP26A1* (Fig. 46) chromatin by the 3C technique (see Methods). Briefly, fixed chromatin DNA was cleaved with a restriction enzyme (NcoI) and ligated after dilution. In these conditions the ligation between intra-molecular fragments is favoured respect to that between inter-molecular ones. Real Time-qPCR was then used to detect the ligated DNA segments. Fig. 45A shows the summary of such analysis by using several probes and “baits” centered on the transcription start site, RARE and 3' end of *CASP9*.

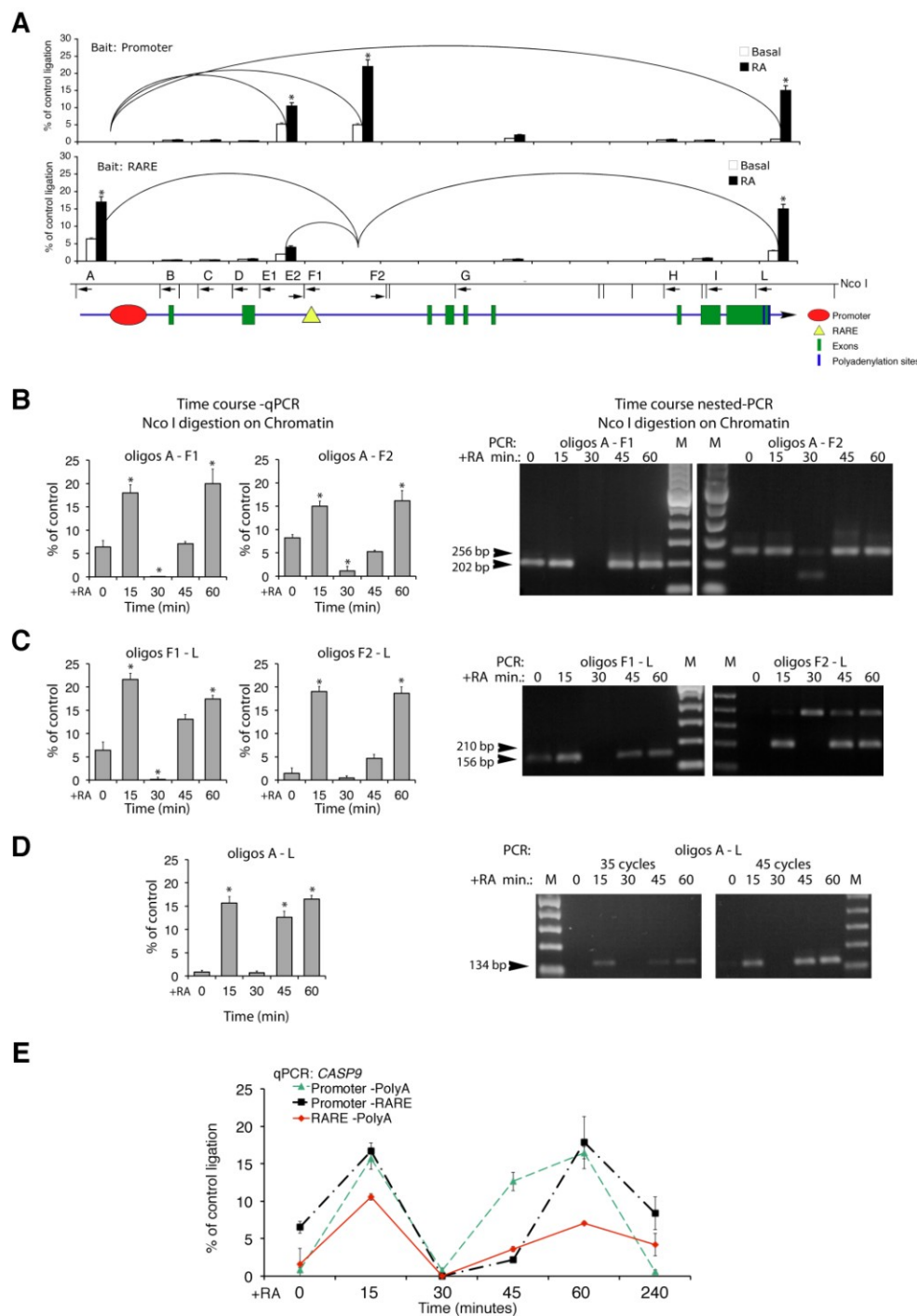


Figure 45 Formation of dynamic chromatin loops during early RA-induced transcription. 3C analysis of *CASP9* chromatin in MCF7 cells exposed to 300 nM of RA for various periods of time. **A.** The histograms show the frequency of ligation of the *CASP9* NcoI fragments amplified with primers indicated below the NcoI restriction map. All the combinations of primers indicated, were performed on ligated chromatin; the

histogram shows qPCR amplifications above 1%, relative to the control. Each loop was detected with different primers pairs and the two histograms show the analysis by using several probes and “baits” centered on the transcription start site, RARE and 3’ end of the *CASP9*. Differences between recombinant, Basal and RA treated chromatin were tested for statistical significance using Student’s t test: * $p < 0.01$ as compared to untreated control. **B, C, D.** Time course of chromatin looping during RA induction. 3C analysis was carried out as described in Methods and the loops shown in panels **B, C** and **D** were quantified by qPCR (left panels) and verified by gel electrophoresis (right panels) and DNA sequencing (data not shown). The results shown derive from a least 3 experiments in triplicate ($n \geq 9$; Mean \pm SD). * $p < 0.01$ as compared to untreated samples. **E.** The panel shows the time course of loop formation; data were collected from Real Time qPCR and from semi-quantitative, nested PCR experiments.

RA enhanced formation of two loops connecting the 5’ and the 3’ ends of the gene with the RARE element. Extensive quantitative analysis of these loops revealed: 1. a 5’ end loop connecting the RARE to the promoter (A-F1, A-F2) was induced by RA (Fig. 45B); 2. a loop (F1-L, F2-L) connecting the RARE region to the 3’ end of *CASP9*, where three different polyA addition sites generate 3 mRNA ends (NM_001229.2; CR613097; CN290432). Assembly of this loop was almost entirely dependent on RA treatment (Fig. 45C); 3. a loop connecting the 5’ and the 3’ ends of the gene, bridging the above mentioned loops (A-L) (Fig. 45D). Transcription of *CASP9* is promoted by many different stimuli. Thus the loops formed in the absence of RA may nevertheless reflect activated genes.

Strikingly, formation of all of these loops is cyclical. They first appear 15 min following RA exposure, disappear at 30 min, and reform by 60 min. This oscillation resembles that displayed in previous figures showing *CASP9* mRNA synthesis, promoter and RARE occupancy by protein factors, and histone modification. The *CYP26A1* gene also formed chromatin loops upon RA treatment. Since RARE and promoter are contiguous in *CYP26A1* we detected essentially one major loop connecting the 5’ (promoter-RARE) with the 3’ end of the gene (polyA). This loop peaked 15 min after RA and slowly disappeared (Figs. 46B and C) similar to the early loop induced by RA on *CASP9* chromatin (Fig. 45E). The physical association of the 5’ and 3’ ends of *CASP9* and *CYP26A1*

genes induced by RA implies that the same proteins are present on the chromatin of the promoter, RARE and 3' end sites. The physical contiguity (620 bp) of the 2 polyA sites (1 and 2) does not discriminate which polyA1 or 2 is included in the RA-induced loop of *CASP9* gene.

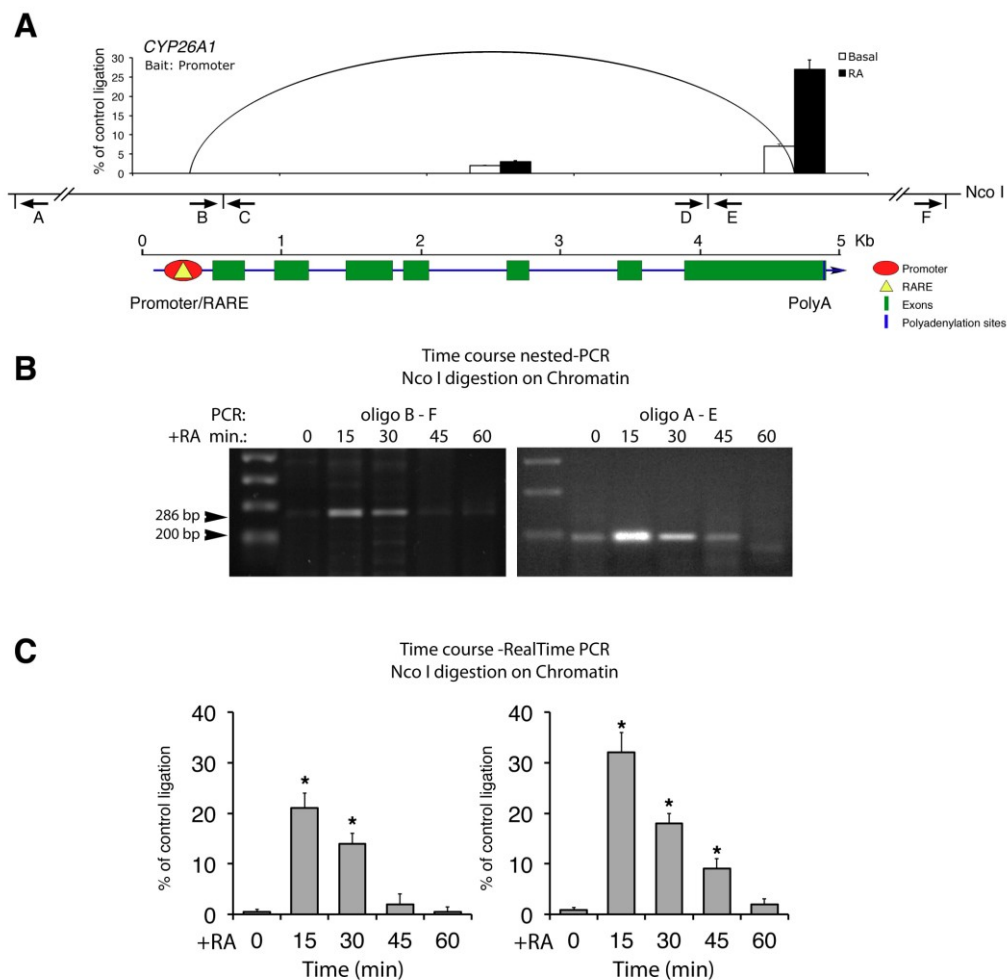


Figure 46 *CYP26A1* DNA chromatin loops induced by RA. **A.** Schematic diagram of *CYP26A1* gene regions. The curved lines indicate the 5' and 3' of the loops detected by 3C technique. The histograms show the frequency of the ligated fragments compared to ligation of the same cloned fragments from genomic DNA. The primers are indicated by arrows. **B.** Time course of loop formation following RA treatment. Semiquantitative nested PCR after digestion with NcoI and ligation of chromatin of cells exposed to RA for various periods of time. The specific products detected with forward B-F and reverse A-E primers are indicated by arrows. **C.** qPCR analysis of the 3C products in chromatin of cells exposed to RA for various periods of time. The results shown derive from a least 3 experiments in triplicate ($n \geq 9$). * $p < 0.01$ as compared to untreated samples.

To find a 3' end specific RA-dependent marker of *CASP9*, we investigated the localization on *CASP9* chromatin of SSU72, a protein which marks the 3' end of genes and interacts with the general transcription initiation factor, TFIIB (He X. et al., 2003). Recently it has been showed that the loss of gene-loop formation by inactivation of SSU72 leads to increased synthesis of promoter-associated divergent ncRNAs, thus SSU72 enforces promoter directionality (Tan-Wong S. M. et al., 2012). Figures 47A, B and C show that SSU72 binds the promoter and RARE with the same kinetics seen with RAR α and Pol II following RA exposure i.e. a peak at 15 min corresponding to the early RA induced loop (Fig. 45E). At 15 min after RA, SSU72 disappeared from the polyA2 and concentrated at the promoter and RARE (Figs. 47 A and B). Apparently, SSU72 was always present at the polyA2 site of *CASP9* gene except at 15' min when the receptor and the promoter were recruiting Pol II and RAR α .

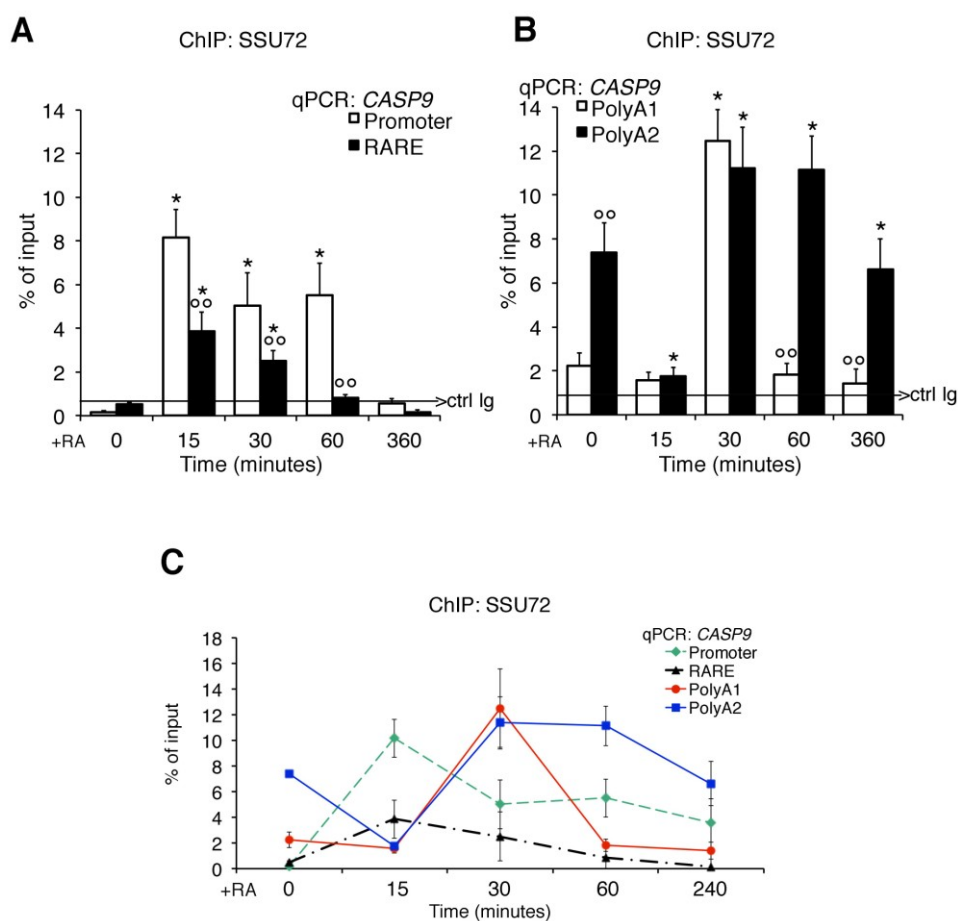


Figure 47 Recruitment of the termination protein SSU72 to the promoter, RARE, polyA1 and polyA2 of *CASP9* gene. Cells were exposed to RA for various periods of time and subjected to ChIP analysis with specific antibodies to SSU72 protein. **A, B** indicate the fraction of SSU72 bound to the promoter and RARE or to the polyA1 and polyA2 of *CASP9* gene, respectively. **C.** The panel shows the time course of SSU72 recruitment. The statistical analysis derive from at least 3 experiments in triplicate ($n \geq 9$); * $p < 0.01$ (matched pairs t test): compared to the RA-unstimulated sample; °° $p < 0.01$ (student t test): comparison between two amplicons.

How relevant are these loops to RA-induced transcription? To address this question we measured the loops involving RARE in cells expressing the LSD1ALA mutant. Expression of LSD1ALA inhibited RA-induced demethylation (Fig. 36) and RA-induced transcription (Fig. 35). Figure 48 shows that the formation of the 15 min loops connecting RARE to the polyA1/2 site or to the promoter upon RA exposure were inhibited: some were delayed (RARE-polyA1/2) and some others were completely eliminated (RARE-promoter). We conclude that the ordered formation of the loops induced by RA is essential for the assembly of transcription initiation complex induced by RA.

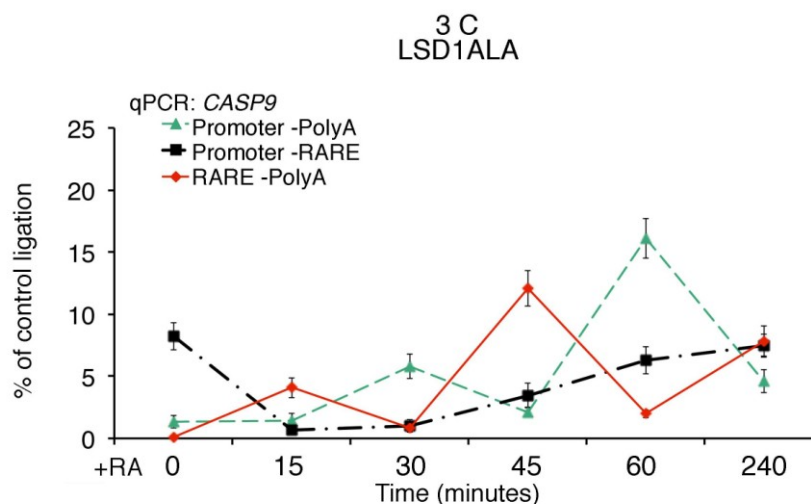


Figure 48 3C analysis of *CASP9* chromatin in the presence of LSD1ALA. MCF7 cells were transfected with the LSD1ALA mutant and exposed to 300 nM of RA for various periods of time; the figure shows the time course of loop formation; data were collected from Real Time qPCR and from semi-quantitative, nested PCR experiments.

Discussion and conclusions

The data reported here show for the first time that the methylation changes of K4 and K9 of histone H3 are linked to the recruitment of repair enzymes and, most importantly, to the formation of chromatin-DNA loops that juxtapose the 5' end transcription start site, the enhancer (RARE) and the 3' end of the transcribed gene (Figs. 45-48). Histone methylation-demethylation cycles (Shi L. et al., 2011) and the formation of loops connecting the 5' gene ends, 3' ends and enhancers have been described extensively in many genes induced by nuclear receptors (Li W. et al., 2013; Le May N. et al., 2012), but insofar these chromatin features, although all required for transcription induction by nuclear hormones, have never been mechanistically and temporally linked. Also, recruitment of NER enzymes to RA induced promoter(s) and the formation of loops bridging the 5', the RARE and the 3' end have been shown upon exposure to RA. In fact, depletion of these enzymes seriously compromises transcription and chromatin looping induced by RA (Le May N. et al., 2012). However, notwithstanding the plethora of data, the mechanism used by RA or other inducers to trigger the recruitment of NER enzymes and formation of chromatin loops is still not known. Our experiments show that demethylation of H3K4 and K9 (Fig. 28) on the RARE, promoter and 3'end (Fig. 30) is the primary event induced by the RA-RAR α complex recruited to the RARE. The timing and the kinetics of demethylation of H3K4 and K9 are similar on the promoter, RARE (Fig. 45D) and on the 3' end of the gene, suggesting that these sites, although not contiguous in the DNA, are included in the same complex, driven by active RAR α -RA. Inhibition of demethylation of H3K4 and H3K9 by depleting the demethylating enzymes (LSD1 or JMJD2A) or by expressing a dominant negative LSD1 variant (Amente S. et al., 2010a; Ambrosio R. et al., 2013) inhibits RA-induced transcription (Figs. 32B-C), nuclear dG oxidation (Fig. 39B), the recruitment of BER (Perillo B. et al., 2008) and NER enzymes and formation of loops induced by RA (Fig. 48). The formation of the chromatin loops with discrete 5' and 3' borders is facilitated by local DNA oxidation following demethylation by LSD1, which presumably releases supercoiling and rigidity of the helix and targets BER and NER enzymes to oxidized bases (Parlanti E. et al., 2012; Pezone A & A.P.; manuscript in preparation). In all cases, BER and NER enzymes are instrumental to localize and

repair DNA nicks and altered bases produced by dG and mdC oxidation (Perillo B. et al., 2008; Lin C. et al., 2009; Fong Y. W. et al., 2013).

A further complication in the comparative analysis of the data published thus far, is represented by the different temporal frames used in various studies to describe the molecular events induced by transcriptional activators. The majority of studies have been carried out at 1 h (Li W. et al., 2013) to several hours or days (Le May N. et al., 2012) after hormonal induction. At this time, synchronization is lost. Each cell is starting and restarting the transcription cycle and only the mature RNA accumulates exponentially and can be easily detected even in asynchronous transcribing cell populations. The H3 (K4-K9) methylation code hours and days after the initial RA induction is not informative since does not change in control and chronically stimulated cells: high of H3K4me2/3 and low H3K9me2/3 content (Shi L. et al., 2011).

An important feature of RA-induced transcription reported here is the timing and synchronous oscillation with a period of ca. 30 min of chromatin bound RAR α , Pol II, H3K4- H3K9 demethylation and looping involving the RARE and the 5' and 3' of *CASP9* and *CYP26A1* genes (Fig. 45). Also, RA-induced mRNA levels of *CASP9*, *CYP26A1* and 5 other genes (A. Pezone, unpublished observations) oscillate with a period of 60 min (30 later than chromatin markers reported in Fig. 45). With the time (2-4 h) this periodic oscillation is lost (Fig. 45). We do not know if this loss of synchrony with the time is due to our inability to track these chromatin changes in non-synchronized cell populations or whether the oscillation we observe is limited to the first transcription cycle. An early and similar cycle of transcription induced by estrogens has been reported to be unproductive in terms of RNA accumulation. This cycle is suggested to prepare the promoter for subsequent transcription followed by two different transcriptionally productive cycles (Métivier R. et al., 2003; Métivier R. et al., 2006; Gao H. & Dahlman-Wright K., 2011). However, our data suggest that: 1. the first cycle (unproductive in Métivier R. et al., 2003; Métivier R. et al., 2006; Gao H. & Dahlman-Wright K., 2011) is indeed the cycle (15-30 min) that sets and defines the physical borders of the transcription unit induced by the hormone (in our case RA); and 2. the oscillations in RAR α recruitment, H3K4 and H3K9 demethylation and

chromatin looping are caused by the physical interference with transcription complexes travelling on the same DNA molecule induced by other (non-RA) stimuli. Why the definition of the physical borders of a transcription unit should be important? Each gene in eukaryotes is indeed the target of many stimuli, which can independently induce transcription. The protein SSU72 has been identified in yeast as an element required for marking the 3' end of the chromatin loops and their stabilization (Hampsey M. et al., 2011). The accumulation of SSU72 in the various sites of *CASP9* gene before and after RA induction may indicate the changes of chromatin engaged in RA and non-RA induced transcription. This protein is present on 3' end of *CASP9* before RA induction, disappears from this site at 15 min and re-appears 30 min after RA treatment (Fig. 47). Fifteen minutes after RA stimulation, SSU72 moves from the 3' end to the TSS and RARE of *CASP9* gene (Fig. 47), where its concentration oscillates synchronously with methylation-demethylation cycles and recruitment of BER and NER enzymes. These data indicate that SSU72 specifies the 3' end of both RA dependent and independent transcription units (Fig. 47C), which are also marked by end- and time-selective specific chromatin loops. In fact, the loops connecting promoter-RARE and 3' end of *CASP9* gene are formed and stabilized by RA in a precise temporal window (15 and 60 min post RA exposure) (Fig. 45B). The synchrony of loop formation by RA is lost in cells expressing the LSD1ALA mutant and some loops are lost (promoter-RARE) or delayed (promoter-polyA or RARE-polyA) and resemble loops formed in the absence of RA (Fig. 48). It is worth noting that LSD1 has been recently shown to control the rhythmicity and the circadian clock and that a mutant in the residue adjacent (aa 111) to that of LSD1ALA (aa 110) is unable to reset the clock oscillations (Nam H. J. et al., 2014). We suggest that RA-induced synchronous demethylation-methylation cycles trigger the recruitment of BER and NER enzymes to the chromatin of promoter-RARE-3'end of *CASP9* gene and synchronize the formation of chromatin 5'end-3'end with RARE-driven transcriptional loops (Figs. 45-48). We believe that RA-induced synchronization of DNA chromatin loops is required for calibration and rapid re-induction of transcription, rather than for high levels of transcription. The massive accumulation of RA-specific mRNAs in fact, occurs

hours after the initial RA exposure (Okuno M. et al., 1995). This efficient re-induction of transcription may represent a “transcription memory”, which has been noted before and requires gene looping and SSU72 in yeast (Brickner J. H., 2010; Lainé J. P. et al., 2009).

The periodic variation of recruited receptor on the enhancer represents a simple mechanism to titrate and calibrate the concentration of RA molecules present in the environment, because any drop in the inducer levels reduces the concentration of active receptor on the enhancer leading to dissolution of the chromatin DNA loop enucleated initially by the active receptor (RARE-Promoter). Conversely, a rise or constant levels of the inducer rapidly reactivate transcription by stabilizing the dissolved loops (transcription memory).

References

- Ambrosio R., Damiano V., Sibilio A., De Stefano M. A., Avvedimento V. E., Salvatore D. & Dentice M. (2013). Epigenetic control of type 2 and 3 deiodinases in myogenesis: role of Lysine-specific Demethylase enzyme and FoxO3. *Nucleic Acids Res.* 41, 3551-3562.
- Amente S., Bertoni A., Morano A., Lania L., Avvedimento E. V. & Majello B. (2010a). LSD1-mediated demethylation of histone H3 lysine 4 triggers Myc-induced transcription. *Oncogene* 29, 3691-3702.
- Amente S., Lania L., Avvedimento E. V. & Majello B. (2010b). DNA oxidation drives Myc mediated transcription. *Cell Cycle* 9, 3002-3004.
- Angrisano T., Sacchetti S., Natale F., Cerrato A., Pero R., Keller S., Peluso S., Perillo B., Avvedimento V. E., Fusco A., Bruni C. B., Lembo F., Santoro M., & Chiariotti L. (2011). Chromatin and DNA methylation dynamics during retinoic acid-induced RET gene transcriptional activation in neuroblastoma cells. *Nucleic Acids Res.* 39, 1993-2006.
- Balajee A. S. & Bohr V. A. (2000). Genomic heterogeneity of nucleotide excision repair. *Gene* 250, 15-30.
- Balmer J. E. & Blomhoff R. (2002). Gene expression regulation by retinoic acid. *J. Lipid Res.* 43, 1773–1808.
- Banerji J., Rusconi S. & Schaffner W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* 27, 299-308.
- Bannister A. J. & Kouzarides T. (2011). Regulation of chromatin by histone modifications. *Cell Res.* 21, 381–395.
- Bartke T., Vermeulen M., Xhemalce B., Robson S. C., Mann M. & Kouzarides T. (2010). Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* 143, 470–484.

- Baylin S. B. & Ohm J. E. (2006). Epigenetic gene silencing in cancer – a mechanism for early oncogenic pathway addiction? *Nat Rev Cancer* 6, 107–116.
- Bell O., Tiwari V. K., Thoma N. H. & Schubeler D. (2011). Determinants and dynamics of genome accessibility. *Nature Rev. Genetics* 12, 554–564.
- Bernatavichute Y. V., Zhang X., Cokus S., Pellegrini M. & Jacobsen S. E. (2008). Genome-wide association of histone H3 lysine nine methylation with CHG DNA methylation in *Arabidopsis thaliana*. *PLoS ONE* 3, e3156.
- Bernstein B. E., Humphrey E. L., Erlich R. L., Schneider R., Bouman P., Liu J. S., Kouzarides T. & Schreiber S. L. (2002). Methylation of histone H3 Lys 4 in coding regions of active genes. *Proc. Natl Acad. Sci. USA* 99, 8695–8700.
- Bernstein B. E., Mikkelsen T. S., Xie X., Kamal M., Huebert D. J., Cuff J., Fry B., Meissner A., Wernig M., Plath K., Jaenisch R., Wagschal A., Feil R., Schreiber S. L. & Lander E. S. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315–326.
- Bird A. P. & Wolffe A. P. (1999). Methylation-induced repression—belts, braces, and chromatin. *Cell* 99, 451–454.
- Brickner J. H. (2010). Transcriptional memory: staying in the loop. *Curr Biol.* 20, R20-1.
- Borun T. W., Pearson D. & Paik W. K. (1972). Studies of histone methylation during the HeLa S-3 cell cycle. *J. Biol. Chem.* 247, 4288–4298.
- Budhu A., Gillilan R. & Noy N. (2001). Localization of the RAR interaction domain of cellular retinoic acid binding protein-II. *J Mol Biol* 305, 939–949.
- Byvoet P., Shepherd G. R., Hardin J. M. & Noland B. J. (1972). The distribution and turnover of labelled methyl groups in histone fractions of cultured mammalian cells. *Arch. Biochem. Biophys.* 148, 558–567.

- Caldecott K.W., Aoufouchi S., Johnson P. & Shall S. (1996). XRCC1 polypeptide interacts with DNA polymerase beta and possibly poly (ADP-ribose) polymerase, and DNA ligase III is a novel molecular 'nick-sensor' in vitro. *Nucleic Acids Res.* 24, 4387-4394.
- Campos E. I. & Reinberg D. (2009). Histones: annotating chromatin. *Annu. Rev. Genet.* 43, 559–599.
- Chambon P. (1996). A decade of molecular biology of retinoic acid receptors. *FASEB J.* 10, 940–954.
- Chang B., Chen Y., Zhao Y. & Bruick R. K. (2007). JMJD6 is a histone arginine demethylase. *Science* 318, 444–447.
- Cho Y. H., Yazici H., Wu H. C., Terry M. B., Gonzalez K., Qu M., Dalay N. & Santella R. M. (2010). Aberrant promoter hypermethylation and genomic hypomethylation in tumor, adjacent normal tissues and blood from breast cancer patients. *Anticancer Res.* 30, 2489-2496.
- Cirillo L. A., Lin F. R., Cuesta I., Friedman D., Jarnik M. & Zaret K. S. (2002). Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA4. *Mol. Cell* 9, 279–289.
- Compe E. & Egly J. M. (2012). TFIIH: when transcription met DNA repair. *Nat Rev Mol Cell Biol.* 13, 343-354. Erratum in: *Nat Rev Mol Cell Biol.* 2012; 13, 476.
- Cooper G. M. (2000). The cell: a molecular approach. 2nd edition. Sunderland (MA): Sinauer Associates.
- Cuozzo C., Porcellini A., Angrisano T., Morano A., Lee B., Di Pardo A., Messina S., Iuliano R., Fusco A., Santillo M. R., Muller M. T., Chiariotti L., Gottesman M. E. & Avvedimento E. V. (2007). DNA damage, homology-directed repair, and DNA methylation. *PLoS Genet.* 3, e110.
- Cuthbert G. L., Daujat S., Snowden A. W., Erdjument-Bromage H., Hagiwara T., Yamada M., Schneider R., Gregory P. D., Tempst P., Bannister A. J. &

- Kouzarides T. (2004). Histone deimination antagonizes arginine methylation. *Cell* 118, 545–553.
- Dantzer F., de La Rubia G., Ménissier-De Murcia J., Hostomsky Z., de Murcia G. & Schreiber V. (2000). Base excision repair is impaired in mammalian cells lacking Poly(ADP-ribose) polymerase-1. *Biochemistry* 39, 7559-7569.
- Dekker J., Rippe K., Dekker M. & Kleckner N. (2002). Capturing chromosome conformation. *Science* 295, 1306-1311.
- del Rincón S. V., Rousseau C., Samanta R. & Miller W. H. Jr. (2003). Retinoic acid-induced growth arrest of MCF-7 cells involves the selective regulation of the IRS-1/PI 3-kinase/AKT pathway. *Oncogene* 22, 3353-3360.
- Deng W. & Blobel G. A. (2010). Do chromatin loops provide epigenetic gene expression states? *Curr. Opin. Genet. Dev.* 20, 548–554.
- Dichtl B., Blank D., Ohnacker M., Friedlein A., Roeder D., Langen H. & Keller W. (2002). A role for SSU72 in balancing RNA polymerase II transcription elongation and termination. *Mol. Cell* 10, 1139–1150.
- Dilworth F. J. & Chambon P. (2001). Nuclear receptors coordinate the activities of chromatin remodeling complexes and coactivators to facilitate initiation of transcription. *Oncogene* 20, 3047–3054.
- Donato L. J. & Noy N. (2005). Suppression of mammary carcinoma growth by retinoic acid: proapoptotic genes are targets for retinoic acid receptor and cellular retinoic acid-binding protein II signaling. *Cancer Res* 65, 8193-8199.
- Dong D., Ruuska S. E., Levinthal D. J. & Noy N. (1999). Distinct roles for cellular retinoic acid-binding proteins I and II in regulating signaling by retinoic acid. *J Biol Chem* 274, 23695–23698.
- Eissenberg J. C. & Shilatifard A. (2010). Histone H3 lysine 4 (H3K4) methylation in development and differentiation. *Dev. Biol.* 339, 240–249.

- Earnshaw W.C., Martins L.M. & Kaufmann S.H. (1999) Mammalian caspases: structure, activation, substrates, and functions during apoptosis. *Annu. Rev. Biochem.* 68: 383-424.
- Escriva H., Delaunay F. & Laudet V. (2000). Ligand binding and nuclear receptor evolution. *Bioessays* 22, 717-727.
- Espada J., Ballestar E., Fraga M. F., Villar-Garea A., Juarranz A., Stockert J. C., Robertson K. D., Fuks F. & Esteller M. (2004). Human DNA methyltransferase 1 is required for maintenance of the histone H3 modification pattern. *J. Biol. Chem.* 279, 37175–37184.
- Feng Q. Wang H., Ng H.H., Erdjument-Bromage H., Tempst P., Struhl K. & Zhang Y. (2002). Methylation of H3 lysine 79 is mediated by a new family of HMTases without a SET domain. *Curr. Biol.* 12, 1052–1058.
- Fischle W., Franz H., Jacobs S. A., Allis C. D. & Khorasanizadeh S. (2008). Specificity of the chromodomain Y chromosome family of chromodomains for lysinemethylated ARK(S/T) motifs. *J. Biol. Chem.* 283, 19626–19635.
- Fong Y. W., Cattoglio C. & Tjian R. (2013). The intertwined roles of transcription and repair proteins. *Mol Cell.* 52, 291-302.
- Forneris F., Binda C., Vanoni M. A., Battaglioli E. & Mattevi A. (2005). Human histone demethylase LSD1 reads the histone code. *J Biol Chem.* 280, 41360-41365.
- Fortini P., Pascucci B., Parlanti E., Sobol R. W., Wilson S. H. & Dogliotti E. (1998). Different DNA polymerases are involved in the short- and long-patch base excision repair in mammalian cells. *Biochemistry* 37, 3575-3580.
- Fuda N. J., Ardehali M. B. & Lis J. T. (2009). Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* 461, 186–192.

- Fujii H., Sato T., Kaneko S., Gotoh O., Fujii-Kuriyama Y., Osawa K., Kato S. & Hamada H. (1997). Metabolic inactivation of retinoic acid by a novel P450 differentially expressed in developing mouse embryos. *EMBO J.* 16, 4163–4173.
- Gao H. & Dahlman-Wright K. (2011). The gene regulatory networks controlled by estrogens. *Mol Cell Endocrinol.* 334, 83-90.
- Gershey E. L., Haslett G. W., Vidali G. & Allfrey V. G. (1969). Chemical studies of histone methylation. Evidence for the occurrence of 3-methylhistidine in avian erythrocyte histone fractions. *J. Biol. Chem.* 244, 4871–4877.
- Goll M. G. & Bestor T. H. (2005). Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.* 74, 481–514.
- Gros L., Saparbaev M. K. & Laval J. (2002). Enzymology of the repair of free radicals-induced DNA damage. *Oncogene* 21, 8905-8925.
- Guccione E., Martinato F., Finocchiaro G., Luzi L., Tizzoni L., Dall' Olio V., Zardo G., Nervi C., Bernard L. & Amati B. (2006). Myc-binding-site recognition in the human genome is determined by chromatin context. *Nature Cell Biol.* 8, 764–770.
- Hadano S., Nasir J., Nichol K., Rasper D. M., Vaillancourt J. P., Sherer S. W., Beatty B. G., Ikeda J. E., Nicholson D. W. & Hayden M. R. (1999). Genomic organization of the human caspase-9 gene on Chromosome 1p36. 1-p36.3. *Mamm Genome* 10, 757-760.
- Haempel K., Lange H. W. & Birkofer L. (1968). E-N-trimethyllysine, a new amino acid in histones. *Die Naturwissenschaften* 55, 37.
- Hampsey M., Singh B. N., Ansari A., Lainč J. P. & Krishnamurthy S. (2011). Control of eukaryotic gene expression: gene loops and transcriptional memory. *Adv Enzyme Regul.* 51, 118-125.
- Hanawalt P. C. (2002). Subpathways of nucleotide excision repair and their regulation. *Oncogene* 21, 8949-8956.

- He X., Khan A. U., Cheng H., Pappas D. L. Jr., Hampsey M. & Moore C. L. (2003). Functional interactions between the transcription and mRNA 3' end processing machineries mediated by Ssu72 and Sub1. *Genes Dev.* 17, 1030-1042.
- Heintzman N. D., Stuart R. K., Hon G., Fu Y., Ching C. W. , Hawkins R. D., Barrera L. O., Van Calcar S., Qu C., Ching K. A., Wang W., Weng Z., Green R. D., Crawford G. E. & Ren B. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature Genet.* 39, 311–318.
- Hermann A., Goyal R. & Jeltsch A. (2004). The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J. Biol. Chem.* 279, 48350–48359.
- Hua S., Kittler R. & White K. P. (2009). Genomic antagonism between retinoic acid and estrogen signaling in breast cancer. *Cell* 137, 1259-1271.
- Huang J. & Berger S. L. (2008). The emerging field of dynamic lysine methylation of non-histone proteins. *Curr. Opin. Genet. Dev.* 18, 152–158.
- Huffman J. L., Sundheim O. & Tainer J. A. (2005). DNA base damage recognition and removal: new twists and grooves. *Mutat Res.* 577, 55-76.
- Jenuwein T. & Allis C.D. (2001). Translating the histone code. *Science* 293, 1074–1080.
- Kastner P., Mark M., Ghyselinck N., Krezel W., Dupe V., Grondona J. M. & Chambon P. (1997). Genetic evidence that the retinoid signal is transduced by heterodimeric RXR/RAR functional units during mouse development. *Development* 124, 313–326.
- Klungland A. & Lindahl T. (1997). Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). *EMBO J.* 16, 3341-3348.

- Kolasinska-Zwierz P., Down T., Latorre I., Liu T., Liu X. S. & Ahringer J. (2009). Differential chromatin marking of introns and expressed exons by H3K36me3. *Nature Genet.* 41, 376–381.
- Kooistra S. M. & Helin K. (2012). Molecular mechanisms and potential functions of histone demethylases. *Nat Rev Mol Cell Biol.* 13, 297-311.
- Kouzarides T. (2007). Chromatin modifications and their function. *Cell* 128, 693–705.
- Krishnamurthy S., He X., Reyes-Reyes M., Moore C. & Hampsey M. (2004). Ssu72 is an RNA polymerase II CTD phosphatase. *Mol. Cell* 14, 387–394.
- Kuida K. (2000). Caspase-9. *The international journal of biochemistry & cell biology* 32, 121-124.
- Lainé J. P., Singh B. N., Krishnamurthy S. & Hampsey M. (2009). A physiological role for gene loops in yeast. *Genes Dev.* 23, 2604-2609.
- Laird P. W., Jackson-Grusby L., Fazeli A., Dickinson S. L., Jung W. E., Li E., Weinberg R. A. & Jaenisch R. (1995). Suppression of intestinal neoplasia by DNA hypomethylation. *Cell* 81, 197–205
- Laudet V. & Gronemeyer H. (2002). The nuclear receptors factsbook. London: Academic Press.
- Law J. A. & Jacobsen S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Rev. Genet.* 11, 204–220.
- Le May N., Fradin D., Iltis I., Bougnères P. & Egly J. M. (2012). XPG and XPF endonucleases trigger chromatin looping and DNA demethylation for accurate expression of activated genes. *Mol Cell.* 47, 622-632.
- Lewis J. D., Meehan R. R., Henzel W. J., Maurer-Fogy I., Jeppesen P., Klein F. & Bird A. (1992) Purification, sequence and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* 69, 905-914.

- Li E., Bestor T. H. & Jaenisch R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69, 915–926.
- Li E., Beard C. & Jaenisch R. (1993). Role for DNA methylation in genomic imprinting. *Nature* 366, 362–365.
- Li P., Nijhawan D., Budihardjo I., Srinivasula S.M., Ahmad M., Alnemri E.S. & Wang X. (1997). Cytochrome c and dATP-dependent formation of Apaf-1/caspase-9 complex initiates an apoptotic protease cascade. *Cell* 91, 479-489.
- Li W., Notani D., Ma Q., Tanasa B., Nunez E., Chen A. Y., Merkurjev D., Zhang J., Ohgi K., Song X., Oh S., Kim H. S., Glass C. K. & Rosenfeld M. G. (2013). Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* 498, 516-520.
- Lin C., Yang L., Tanasa B., Hutt K., Ju B. G., Ohgi K., Zhang J., Rose D. W., Fu X. D., Glass C. K. & Rosenfeld M. G. (2009). Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. *Cell* 139, 1069-1083.
- Lindahl T. & Wood R. D. (1999). Quality control by DNA repair. *Science* 286, 1897-1905.
- Loh Y.H., Zhang W., Chen X., George J. & Ng H. H. (2007). Jmjd1a and Jmjd2c histone H3 Lys 9 demethylases regulate self-renewal in embryonic stem cells. *Genes Dev.* 21, 2545-2557.
- Loudig O., Babichuk C., White J., Abu-Abed S., Mueller C. & Petkovich M. (2000). Cytochrome P450RAI(CYP26) promoter: a distinct composite retinoic acid response element underlies the complex regulation of retinoic acid metabolism. *Mol. Endocrinol.* 14, 1483–1497
- Loudig O., Maclean G. A., Dore N. L., Luu L. & Petkovich M. (2005). Transcriptional co-operativity between distant retinoic acid response elements in regulation of Cyp26A1 inducibility. *Biochem J.* 15:241-248.

- 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., Mader A. W., Richmond R. K., Sargent D. F. & Richmond T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251–260.
- Mangelsdorf D. J. & Evans R. M. (1995). The RXR heterodimers and orphan receptors. *Cell* 83, 841–850.
- Margueron R. & Reinberg D. (2010). Chromatin structure and the inheritance of epigenetic information. *Nat. Rev. Genet.* 11, 285–296.
- Meissner A., Mikkelsen T. S., Gu H., Wernig M., Hanna J., Sivachenko A., Zhang X., Bernstein B. E., Nusbaum C., Jaffe D. B., Gnirke A., Jaenisch R. & Lander E. S. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 454, 766–770.
- Mellon I., Spivak G. & Hanawalt P. C. (1987). Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell* 51, 241–249.
- Mellon I. & Hanawalt P. C. (1989). Induction of the Escherichia coli lactose operon selectively increases repair of its transcribed DNA strand. *Nature* 342, 95–98.
- Métivier R., Penot G., Hübner M. R., Reid G., Brand H., Kos M. & Gannon F. (2003). Estrogen receptor- α directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell* 115, 751–763.
- Métivier R., Reid, G. & Gannon, F. (2006). Transcription in four dimensions: nuclear receptor-directed initiation of gene expression. *EMBO Rep.* 7, 161–167.
- Metzger E., Wissmann M., Yin N., Müller J. M., Schneider R., Peters A. H., Günther T., Buettner R. & Schüle R. (2005). LSD1 demethylates repressive

histone marks to promote androgen-receptor-dependent transcription. *Nature* 437, 436-439.

Mohan M., Herz H. M., Takahashi Y. H., Lin C., Lai K. C., Zhang Y., Washburn M. P., Florens L. & Shilatifard A. (2010). Linking H3K79 trimethylation to Wnt signalling through a novel Dot1-containing complex (DotCom). *Genes Dev.* 24, 574–589.

Morano A., Angrisano T., Russo G., Landi R., Pezone A., Bartollino S., Zuchegna C., Babbio F., Bonapace I. M., Allen B., Muller M. T., Chiariotti L., Gottesman M. E., Porcellini A. & Avvedimento E. V. (2014). Targeted DNA methylation by homology-directed repair in mammalian cells. Transcription reshapes methylation on the repaired gene. *Nucleic Acids Res.* 42, 804-821.

Murray K. (1964). The occurrence of ϵ -N-methyl lysine in histones. *Biochemistry* 3, 10–15.

Nam H. J., Boo K., Kim D., Han D. H., Choe H. K., Kim C. R., Sun W., Kim H., Kim K., Lee H., Metzger E., Schuele R., Yoo S. H., Takahashi J. S., Cho S., Son G. H. & Baek S. H. (2014). Phosphorylation of LSD1 by PKC α Is Crucial for Circadian Rhythmicity and Phase Resetting. *Mol Cell. pii: S1097-2765(14)00111-7*. doi: 10.1016/j.molcel.2014.01.028.

Okano M., Bell D. W., Haber D. A. & Li E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99, 247–257.

Okuno M., Caraveo V. E., Goodman D. S. & Blaner W. S. (1995). Regulation of adipocyte gene expression by retinoic acid and hormones: effects on the gene encoding cellular retinol-binding protein. *J Lipid Res.* 36, 137-147.

Paik W. K. & Kim S. (1967). ϵ -N-dimethyllysine in histones. *Biochem. Biophys. Res. Commun.* 27, 479–483.

Panning B. & Jaenisch R. (1998). RNA and the epigenetic regulation of X chromosome inactivation. *Cell* 93, 305–308.

- Pappas D. L. Jr & Hampsey M. (2000). Functional interaction between Ssu72 and the Rpb2 subunit of RNA polymerase II in *Saccharomyces cerevisiae*. *Mol Cell Biol.* 20, 8343-8351.
- Parlanti E., D'Errico M., Degan P., Calcagnile A., Zijno A., van der Pluijm I., van der Horst G. T., Biard D. S., Dogliotti E. (2012). The cross talk between pathways in the repair of 8-oxo-7,8-dihydroguanine in mouse and human cells. *Free Radic Biol Med.* 53, 2171- 2177.
- Perillo B., Ombra M. N., Bertoni A., Cuozzo C., Sacchetti S., Sasso A., Chiariotti L., Malorni A., Abbondanza C. & Avvedimento E. V. (2008). DNA oxidation as triggered by H3K9me2 demethylation drives estrogen-induced gene expression. *Science* 319, 202-206.
- Phillips T. (2008). Regulation of transcription and gene expression in eukaryotes. *Nature Education* 1 (1):199.
- Ptashne M. (2007). On the use of the word ‘epigenetic’. *Curr. Biol.* 17, R233–R236.
- Ranalli T. A., Tom S. & Bambara R. A. (2002). AP endonuclease 1 coordinates flap endonuclease 1 and DNA ligase I activity in long patch base excision repair. *J Biol Chem.* 277, 41715-41724.
- Rao K.S. (2007). Mechanisms of disease: DNA repair defects and neurological disease. *Nature clinical practice* 3, 162-172.
- Ray W. J., Bain G., Yao M. & Gottlieb D. I. (1997). CYP26, a novel mammalian cytochrome P450, is induced by retinoic acid and defines a new family. *J Biol Chem.* 272, 18702- 18708.
- Rea S., Eisenhaber F., O'Carroll D., Strahl B. D., Sun Z. W., Schmid M., Opravil S., Mechtler K., Ponting C. P., Allis C. D. & Jenuwein T. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406, 593–599.

- Reményi A., Schöler H. & Wilmanns M. (2004). Combinatorial control of gene expression. *Nature Structural and Molecular Biology* 11, 812–815.
- Robinson-Rechav, M., Carpentier, A. S., Duffraisse M. & Laudet V. (2001). How many nuclear hormone receptors in the human genome? *Trends Genet* 17, 554-556.
- Ruthenburg A. J., Li H., Patel D. J. & Allis C. D. (2007). Multivalent engagement of chromatin modifications by linked binding modules. *Nature Rev. Mol. Cell Biol.* 8, 983–994.
- Santos-Rosa H., Schneider R., Bannister A. J., Sherriff J., Bernstein B. E., Tolga Emre N. C., Schreiber S. L., Mellor J. & Kouzarides T. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature* 419, 407–411.
- Sattler U., Frit P., Salles B. & Calsou P. (2003). Long-patch DNA repair synthesis during base excision repair in mammalian cells. *EMBO Rep.* 4, 363-367.
- Sentenac A. (1985). Eukaryotic RNA polymerases. *CRC Crit Rev Biochem.* 18, 31-90.
- Serandour A. A., Avner S., Percevault F., Demay F., Bizot M., Lucchetti-Miganeh C., Barloy-Hubler F., Brown M., Lupien M., Métivier R., Salbert G. & Eeckhoutte J. (2011). Epigenetic switch involved in activation of pioneer factor FOXA1-dependent enhancers. *Genome Res.* 21, 555–565.
- Shahhoseini M., Taghizadeh Z., Hatami M. & Baharvand H. (2013). Retinoic acid dependent histone 3 demethylation of the clustered HOX genes during neural differentiation of human embryonic stem cells. *Biochem Cell Biol.* 91, 116-122.
- Shi Y., Lan F., Matson C., Mulligan P., Whetstone J. R., Cole P. A., Casero R.A. & Shi Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homologue LSD1. *Cell* 119, 941–953.
- Shi X., Hong T., Walter K. L., Ewalt M., Michishita E., Hung T., Carney D., Peña P., Lan F., Kaadige M. R., Lacoste N., Cayrou C., Davrazou F., Saha A., Cairns

- B. R., Ayer D. E., Kutateladze T. G., Shi Y., Côté J., Chua K. F. & Gozani O. (2006). ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* 442, 96–99.
- Shi L., Sun L., Li Q., Liang J., Yu W., Yi X., Yang X., Li Y., Han X., Zhang Y., Xuan C., Yao Z. & Shang Y. (2011). Histone demethylase JMJD2B coordinates H3K4/H3K9 methylation and promotes hormonally responsive breast carcinogenesis. *Proc Natl Acad Sci U S A*. 108, 7541- 7546.
- Shogren-Knaak M., Ishii H., Sun J. M., Pazin M. J., Davie J. R. & Peterson C. L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* 311, 844–847.
- Sims R. J., Millhouse S., Chen C. F., Lewis B. A., Erdjument-Bromage H., Tempst P., Manley J. L. & Reinberg D. (2007). Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Mol. Cell* 28, 665–676.
- Sonneveld E., van den Brink C. E., van der Leede B. M., Schulkes R. K., Petkovich M., van der Burg B. & van der Saag P. T. (1998). Human retinoic acid (RA) 4-hydroxylase (CYP26) is highly specific for all-trans-RA and can be induced through RA receptors in human breast and colon carcinoma cells. *Cell Growth Differ.* 9:629-637.
- Sproul D., Gilbert N. & Bickmore W. (2005). The role of chromatin structure in regulating the expression of clustered genes. *Nature Reviews Genetics* 6, 775–781.
- Srivastava D. K., Berg B. J., Prasad R., Molina J. T., Beard W. A., Tomkinson A. E. & Wilson S. H. (1998). Mammalian abasic site base excision repair. Identification of the reaction sequence and rate-determining steps. *J Biol Chem.* 273, 21203-21209.

- St-Pierre B., Liu X., Kha L. C., Zhu X., Ryan O., Jiang Z. & Zacksenhaus E. (2005). Conserved and specific functions of mammalian ssu72. *Nucleic Acids Research* 33, 464-477.
- Strahl B. D. & Allis C. D. (2000). The language of covalent histone modifications. *Nature* 403, 41–45.
- Struhl K. (1999). Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. *Cell* 98, 1–4.
- Suganuma T. & Workman J. L. (2011). Signals and combinatorial functions of histone modifications. *Annu. Rev. Biochem.* 80, 473–499.
- Sun Z.W. & Hampsey M. (1996). Synthetic enhancement of a TFIIB defect by a mutation in SSU72, an essential yeast gene encoding a novel protein that affects transcription start site selection in vivo. *Mol. Cell. Biol.* 16, 1557–1566.
- Sung J. S. & Demple B. (2006). Roles of base excision repair subpathways in correcting oxidized abasic sites in DNA. *FEBS J.* 273, 1620-1629.
- Tan-Wong S. M., Zaugg J. B., Camblong J., Xu Z., Zhang D. W., Mischo H. E., Ansari A. Z., Luscombe N. M., Steinmetz L. M. & Proudfoot N. J. (2012). Gene loops enhance transcriptional directionality. *Science* 338, 671-675.
- Taverna S. D., Li H., Ruthenburg A. J., Allis C. D. & Patel D. J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nature Struct. Mol. Biol.* 14, 1025–1040.
- Tsai W. W., Wang Z., Yiu T. T., Akdemir K. C., Xia W., Winter S., Tsai C. Y., Shi X., Schwarzer D., Plunkett W., Aronow B., Gozani O., Fischle W., Hung M. C., Patel D. J. & Barton M. C. (2010). TRIM24 links a non-canonical histone signature to breast cancer. *Nature* 468, 927–932.
- Tsukada Y., Fang J., Erdjument-Bromage H., Warren M. E., Borchers C. H., Tempst P. & Zhang Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. *Nature* 439, 811–816.

- Vaquero A., Loyola A. & Reinberg D. (2003). The constantly changing face of chromatin. *Sci. Aging Knowledge Environ.* 14, re4.
- Vaute O., Nicolas E., Vandel L. & Trouche D. (2002). Functional and physical interaction between the histone methyl transferase Suv39H1 and histone deacetylases. *Nucleic Acids Res.* 30, 475-481.
- Vermeulen M., Eberl H. C., Matarese F., Marks H., Denissov S., Butter F., Lee K. K., Olsen J. V., Hyman A. A., Stunnenberg H. G. & Mann M. (2010). Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* 142, 967–980.
- Waddington C. H. (2012). The epigenotype. *Endeavour* 1942 1, 18-20. Reprinted in *Int J Epidemiol* 41, 10-13.
- Wang H., Cao. R., Xia L., Erdjument-Bromage H., Borchers C., Tempst P. & Zhang Y. (2001). Purification and functional characterization of a histone H3-lysine 4-specific methyltransferase. *Mol Cell.* 8, 1207-1217.
- Webby C. J., Wolf A., Gromak N., Dreger M., Kramer H., Kessler B., Nielsen M. L., Schmitz C., Butler D. S., Yates J. R. 3rd, Delahunty C. M., Hahn P., Lengeling A., Mann M., Proudfoot N. J., Schofield C. J. & Böttger A. (2009). JMJD6 catalyses lysyl-hydroxylation of U2AF65, a protein associated with RNA splicing. *Science* 325, 90–93.
- Wernig M., Meissner A., Foreman R., Brambrink T., Ku M., Hochedlinger K., Bernstein B. E. & Jaenisch R. (2007). *In vitro* reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* 448, 318–324.
- Whetstine J. R., Nottke A., Lan F., Huarte M., Smolikov S., Chen Z., Spooner E., Li E., Zhang G., Colaiacovo M. & Shi Y. (2006). Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* 125, 467–481.
- White J. A., Guo Y. D., Baetz K., Beckett-Jones B., Bonasoro J., Hsu, K. E., Dilworth F. J., Jones G. & Petkovich M. (1996). Identification of the retinoic

acid-inducible all-trans-retinoic acid 4-hydroxylase. *J. Biol. Chem.* 271, 29922–29927.

Wissmann M., Yin N., Müller J. M., Greschik H., Fodor B. D., Jenuwein T., Vogler C., Schneider R., Günther T., Buettner R., Metzger E. & Schüle R. (2007). Cooperative demethylation by JMJD2C and LSD1 promotes androgen receptor-dependent gene expression. *Nat Cell Biol.* 9, 347-353.

Wu C. & Morris J. R. (2001). Genes, genetics, and epigenetics: a correspondence. *Science* 293, 1103-1105.

You J. S., Wang M. & Lee S. H. (2003). Biochemical analysis of the damage recognition process in nucleotide excision repair. *J Biol Chem.* 278, 7476-7485.

Young N. L., Dimaggio P. A. & Garcia B. A. (2010). The significance, development and progress of high-throughput combinatorial histone code analysis. *Cell. Mol. Life Sci.* 67, 3983–4000.

Yudkovsky N., Ranish J. A. & Hahn S. (2000). A transcription reinitiation intermediate that is stabilized by activator. *Nature* 408, 225-229.

Zee, B. M., Levin R. S., Xu B., LeRoy G., Wingreen N. S. & Garcia B.A. (2010). *In vivo* residue-specific histone methylation dynamics. *J. Biol. Chem.* 285, 3341–3350.

Zhang X. K., Hoffmann B., Tran P. B., Graupner G. & Pfahl M. (1992). Retinoid X receptor is an auxiliary protein for thyroid hormone and retinoic acid receptors. *Nature* 355, 441–446.

Zhu L., Zhang Y., Zhang W., Yang S., Chen J. Q. & Tian D. (2009). Patterns of exon–intron architecture variation of genes in eukaryotic genomes. *BMC Genomics* 10, 47.

Zilberman D., Coleman-Derr D., Ballinger T. & Henikoff S. (2008). Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. *Nature* 456, 125–129.

List of publications:

- Messina S, Frati L, Leonetti C, Zuchegna C, Di Zazzo E, Calogero A and Porcellini A. Dual-specificity phosphatase DUSP6 has tumor-promoting properties in human glioblastomas. (2011). *Oncogene* 30: 3813-3820. ISSN 0950-9232 DOI 10.1038/onc.2011.99.
- Donini C F, Di Zazzo E, Zuchegna C, Di Domenico M, D'Inzeo S, Nicolussi A, Avvedimento E V, Coppa A, Porcellini A. The p85 α regulatory subunit of PI3K mediates cAMP-PKA and retinoic acid biological effects on MCF7 cell growth and migration. (2012). *International Journal Of Oncology* 40(5): 1627-1635. ISSN 1019-6439 DOI 10.3892/ijo.2012.1383.
- Feola A, Cimini A, Migliucci F, Iorio R, Zuchegna C, Rothenberger R, Cito L, Porcellini A, Unteregger G, Tombolini V, Giordano A, Di Domenico M. The inhibition of p85 α PI3KSer83 phosphorylation prevents cell proliferation and invasion in prostate cancer cells. (2013). *J Cell Biochem.* Sep; 114(9): 2114-2119. ISSN 0730-2312 DOI 10.1002/jcb.24558.
- Morano A, Angrisano T, Russo G, Landi R, Pezone A, Bartollino S, Zuchegna C, Babbio F, Bonapace IM, Allen B, Muller MT, Chiariotti L, Gottesman ME, Porcellini A, Avvedimento EV. Targeted DNA methylation by homology-directed repair in mammalian cells. Transcription reshapes methylation on the repaired gene. (2013). *Nucleic Acids Research.* ISSN 0305-1048 DOI 10.1093/nar/gkt920.
- Di Zazzo E, Feola A, Zuchegna C, Romano A, Donini CF, Bartollino S, Costagliola C, Frunzio R, Laccetti P, Di Domenico M, Porcellini A. The p85 regulatory subunit of PI3K mediates cAMP–PKA and insulin biological effects on MCF-7 cell growth and motility. (2014). In submission to *Journal of Cellular Physiology*.

- Zuchegna C, Aceto F, Bertoni A, Romano A, Perillo B, Laccetti P, Gottesman ME, Avvedimento VE, Porcellini A. Mechanism of Retinoic acid-induced transcription: histone code, DNA oxidation and formation of chromatin loops. (2014). In submission to *Molecular Cell*.