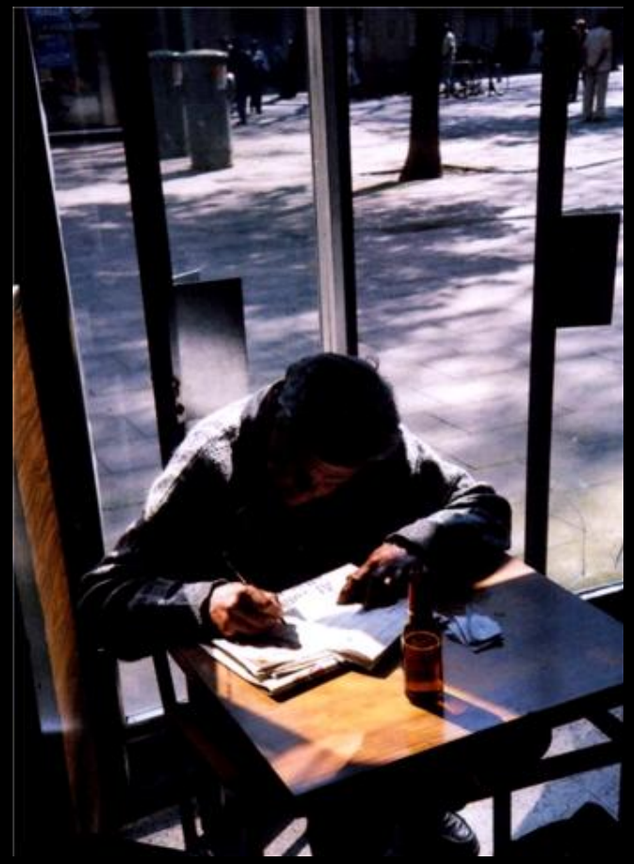
Roser Zaurin Quer

PhD Thesis - 2009

Sequence-Dependent Nucleosome Positioning



and **C**hromatin
Remodelling of Hormone Responsive Genes

Roser **Z**aurin **Q**uer

Foto de portada: “L’Americano”
Foto de contraportada: “Ceagabeau”

d'*Ennio Francavilla*

Sequence-Dependent Nucleosome Positioning and Chromatin Remodelling of Hormone-Responsive Genes

Roser Zaurin Quer

a la iaia i
... a l'abuelito

Sequence-Dependent Nucleosome Positioning and Chromatin Remodelling of Hormone-Responsive Genes

Roser Zaurin Quer

Memòria presentada per optar al Grau de Doctora per la Universitat Pompeu Fabra.

Aquesta tesi ha estat realitzada al laboratori de Cromatina i Expressió Gènica del programa de Regulació Gènica, al Centre de Regulació Genòmica (CRG) de Barcelona, sota la direcció del Dr. Miguel Beato del Rosal i el Dr. Guillermo Pablo Vicent, des de setembre de 2004 a juliol de 2009.

Dr. Miguel Beato

Dr. Guillermo P. Vicent

Roser Zaurin Quer

Barcelona, juliol de 2009

Table Of Contents

Summary

Evidence has been accumulating over the last few years pointing to the importance of chromatin structure and nucleosome positioning in cellular processes such as transcriptional regulation. Recent technological advances in the field have allowed the construction of detailed genome-scale maps of nucleosome positions, and there have been several attempts to define the sequence characteristics that guide the positioning of nucleosomes, the so-called “nucleosome code”. In this thesis I give experimental evidence for the existence of a subset of very well-positioned nucleosomes in the human genome based on nucleosome-resolution tilling arrays and on deep sequencing of MNase-digested chromatin using the Solexa-Illumina platform. I show that these nucleosomes, which we have named “key nucleosomes”, tend to occupy genomic locations of specific function, indicative of their special role. The DNA of these “key nucleosomes” exhibits a high symmetry of curvature, allowing their precise position on the human genome to be predicted *in silico* based on the structural attributes of the primary DNA sequence.

The second part of this thesis provides new insights into the importance of chromatin organization and dynamics in the context of gene regulation. Steroid hormones induce transcription of their target genes by a complex mechanism requiring binding of the hormone receptors to hormone responsive elements (HREs) and the recruitment of a variety of coregulators. The Mouse Mammary Tumor Virus (MMTV) promoter has long been used as a model for the study of hormone receptor-mediated gene activation. It is known that progesterone receptor (PR) binds the exposed HRE1 of the MMTV promoter chromatin and recruits chromatin remodellers that catalyse ATP-dependent histone H2A/H2B displacement. I show that the ATP-dependent chromatin remodelling complex BAF, but not PBAF, is recruited after hormone treatment and is necessary for MMTV promoter activation. Along with the previously reported phosphorylation of H3S10 by Msk, I show that an early PCAF-mediated acetylation of H3K14 is essential for the activation of the promoter by anchoring the BAF complex. Following transient displacement of H2A/H2B dimers, binding of NF1 is required for stabilizing the

remodelled conformation of the MMTVnucleosome. To further study the activation process I have used MMTVminichromosomes, mononucleosomes and H3/H4 tetramer particles reconstituted on wild type MMTV and MMTV promoter fragments with point mutations disrupting binding of PR and NF1. I show that only when MMTV sequences are assembled on H3/H4 tetramer particles can PR bind to all five HREs while allowing NF1 access to its cognate site. Furthermore, I found that binding of NF1 facilitates access of PR to the central HREs 2 and 3, thus contributing to the reciprocal synergism between PR and NF1.

Resum

Evidències recents han remarcat la importància del paper de l'estructura de la cromatina i el posicionament de nucleosomes en processos cel·lulars bàsics com és la regulació de la transcripció gènica. L'avenç de noves tecnologies en el camp ha permès l'estudi detallat de la disposició de nucleosomes en genomes sencers. Hi ha hagut també molts intents per definir les característiques de la seqüència de l'ADN que podrien arribar a guiar el posicionament dels nucleosomes; sent el conjunt d'aquestes característiques l'anomenat "codi nucleosòmic". En la present tesis doctoral, s'aporten evidències experimentals sobre l'existència d'un grup de nucleosomes molt ben posicionats en el genoma humà. Això ha estat possible mitjançant tècniques com els microarray i la seqüenciació massiva en paral·lel. Aquest treball mostra com aquests nucleosomes, que anomenem "nucleosomes clau", tenen tendència a ocupar regions del genoma amb funcions específiques, la qual cosa indica el seu paper especial. L'ADN dels "nucleosomes clau" resulta tenir una alta simetria de curvatura. Aquesta característica inherent a la seqüència fa que sigui possible la predicció *in silico* de les posicions de nucleosomes d'aquest tipus.

En la segona part de la present tesi doctoral, aporto noves evidències experimentals que fan avançar el camp de la organització de la cromatina i la seva dinàmica en el context de la regulació gènica. Les hormones esteroidees indueixen la transcripció dels seus gens diana a través de la unió dels receptors hormonals amb els seu corresponent motiu de reconeixement a l'ADN (HREs), així com el reclutament d'una gran varietat de co-reguladors. El promotor del Virus de Tumor Mamari de Ratolí (MMTV) ha estat un model molt usat per l'estudi dels efectes en l'activació gènica dels receptors hormonals. És conegut que el receptor de progesterona (PR) s'uniex a l'HRE1, accessible, i recluta maquinàries de remodelament de la cromatina que utilitzen l'ATP com a font energètica per expulsar els dimers d'histones H2A/H2B del

nucleosoma B del promotor de l'MMTV. Aquí demostro que la màquinaria de remodelament reclutada és específicament BAF. Treballs anteriors van demostrar que la kinasa Msk1 és la responsable de la fosforilació de la serina 10 de la histona H3. En aquesta tesi es demostra que l'acetilació de la Lysina 14 de la histona H3 és essencial per l'activació del promotor, així com per l'anclatge de BAF. Després de l'expulsió dels dimers d'H2A/H2B, la unió d'NF1 al promotor és indispensable per estabilitzar la forma remodelada del nucleosoma. Per l'estudi en més detall de l'activació del promotor de l'MMTV he utilitzat el sistema de minicromosomes, mononucleosomes i tetràmers d'histones H3/H4 reconstituïts en seqüències salvatges i mutants del promotor de l'MMTV. He demostrat que, només quan un fragment de la seqüencia de l'MMTV està reconstituïda en tetràmers, el PR i l'NF1 poden estar units simultàniament a la seva seqüència de reconeixement en el promotor. També aporto evidències on es demostra que la unió d'NF1 al promotor facilita la posterior unió de més mol·lècules de PR als HREs interns (llocs 2 i 3) caracterizant en més detall el sinergisme funcional que existeix entre el PR i l'NF1 en aquestes condicions.

Abbreviations

ACF : ATP-utilizing chromatin assembly and remodelling factor
AD: activation Domain
ATP : adenosine triphosphate
BAF: BRG1/hBRM-Associated Factors
BAP: Brahma-associated Protein
BP: base pair
BRG1 : brahma related gene1
CARM1: Coactivator-associated arginine methyltransferase
CTD : C terminal domain
CBP : CREB binding protein
CD: *C*hromo*d*o*m*ain
CHRAC: *C*hromatin Accessibility Complex
Co-Rest: RE-1 silencing transcription factor (REST) co-repressor
CENP-A : centromeric protein A
Chd1 : chromodomain helicase DNA binding 1
ChIP : chromatin immunoprecipitation
DNA : Deoxyribonucleic acid
DBD: Dna binding Domain
DREX: Drosophila embryo extract
ERK: Extracellular-signal regulated kinase
Gcn5 : general control nonderepressible 5
GNAT : Gcn5 N-acetyltransferase related
GST : glutathione S transferase
GR: glucocorticoid receptor
HAT : histone acetyltransferase
Hat1 : histone acetyltransferase 1
HDAC : histone désacétylase
HMT : histone methyltransferase
HP1 : heterochromatin protein 1
HMG: High Mobility Group
HRE: Hormone Resposnsive Element
HRR: Hormone Responsive Region
ISWI : imitation SWI/SNF

LTR: Long Terminal Repeat
LXXLL: L, Leu; X, any amino acid
MAPK: Mitogen-Activated Protein Kinase
MYST : MOZ, Ybf2/Sas3, Sas2 et TIP60
Mi-2: Dermatomyositis-specific autoantigen
MMTV: Mouse Mammary Tumor Virus
MSK1: Mitogen-and Stress-Activated Specific Kinase
NcoR: Nuclear co-repressor
NCP: nucleosome core particle
NAD : nicotinamide adenine dinucleotide
NuA4 : nucleosome acetyltransferase of H4
NuRD : nucleosome remodeling histone deacetylase complex
NURF: Nucleosome Remodelling Factor
ORF : open reading frame
Oct-1: Octamer binding transcription factor
PCAF : p300/CBP associated factor
PCR : polymerase chain reaction
PHD : plant homeodomain
PI3K : phosphatidylinositol 3 kinase
PIC : pre initiation complex
pRb : protein retinoblastoma
PRMT1 : proteine arginine methyltransferase 1
PBAF: Polybromo associated-BAF
Pol: polymerase
PR: progesterone receptor
RAR: retinoc aced receptor
RNA : Ribonucleic acid
RSC : remodel the structure of chromatin
SAGA : Spt, Ada, Gcn5 Acetyltransferase
SAM : S adenosyl methionine
SANT : Swi3, Ada2, N-Cor et TFIIIB
SET : Su(var3-9), enhancer of zeste, trithorax
Sin3 : swi-independent 3
Sir : silent information regulator protein
SNF : sucrose non fermentation
SWI : mating type switching
Swr1 : Swi2/Snf2-related ATPase 1
TAF : TBP associated factor
TBP : TATA box binding protein
TFTC : TATA-binding-protein-free TAFII containing complex
TIP60 : Tat interacting protein (60kDa)
TSA : trichostatin A
WT : wild type

Introduction

Section 1: The bases of chromatin organization

1. The nucleosome

Mammals package 3 billion base pairs of DNA, a total of 2 meters of human DNA, that encode about 30,000 genes into a cell nucleus which average diameter is approximately 6 micrometers. DNA is packaged using a protein scaffold, forming a complex tertiary structure that is referred as **chromatin**. The fundamental structure of chromatin is the **nucleosome**, which consists of 146 base pairs of DNA wrapped around an octamer of two molecules each of the histones H2A, H2B, H3 and H4 (Figure 1). These cells must therefore accomplish the difficult task of folding the nucleosomes in a highly compact manner while still allowing access to various nuclear factors. A fifth histone protein, the linker histone or histone H1, has an important function in the higher structure organization and its stability. The location of H1 in the nucleosome has been mapped by protein-DNA photo-crosslinking showing that the H1 globular domain forms interaction with the DNA at either the entry or exit strand of the nucleosomal DNA, but not with the core histones. The unit consisting of a nucleosome particle plus one H1 molecule was first observed by Simpson et al. (1978). Adjacent nucleosomes are connected by linker DNA and progressive coiling of nucleosomes leads to compact, higher-order chromatin structures.

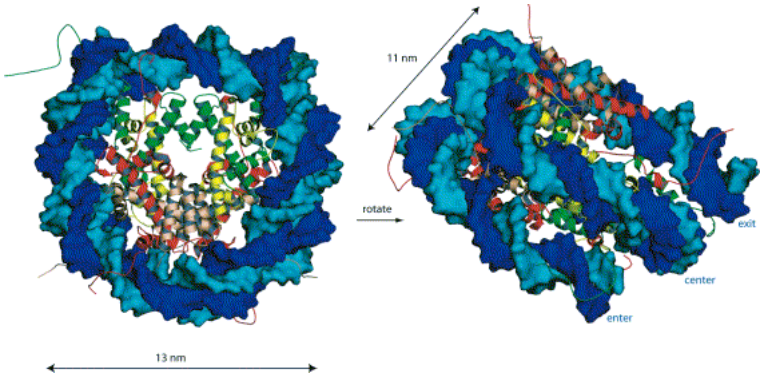


Figure 1. The atomic structure of the nucleosome core particle. Each strand of DNA is shown in differenr shade of blue. The DNA makes 1.7 turns around the histone octamer to form an overall particle with disk-like structure. Histones are colored as Figures 2 and 3. (from Khorasanizadeh, 2004)

Each core histone uses a protein fold, the histone fold, consisting of a three-helix core domain: a long central α-helix flanked on either side by a loop and a short α-helix. These domains form handshake arrangements to give rise to the heterodimer H2A/H2B (Figure 2) and the heterodimer H3/H4 (Figure 3) (Arents et al, 1991). Biochemical studies have shown that in solutions of moderate salt and in the absence of DNA, the H3/H4 complex forms a tetramer whereas H2A/H2B complex remains a stable dimer. However, the H3/H4 tetramer alone can form a stable DNA complex with about 120 bp (Hayes et al, 1991) and is enough to confer nucleosome positioning (Dong and van Holde, 1990).

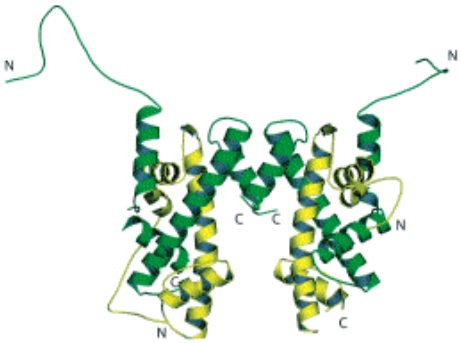


Figure 2. The atomic structure of core histones tetramer of H3 (green) **and H4** (yellow) (from Khorasanizadeh, 2004)

Figure 3. The atomic structure of core histones dimmer of H2A (red) **and H2B** (pink). (from Khorasanizadeh, 2004)

In addition to the structured histone fold core, each histone forms extensions consisting of N-terminal and C-terminal tails with no secondary structure that protrude from the nucleosome and are rich in basic amino acids, based on crystal structure assays (Luger et al, 1997). However, the tail segments of H3 and H2B protrude from the DNA gyres through the minor-groove channels. By extending beyond the disk-

shaped nucleosome surface, the tails form ideal surfaces for covalent modifications by enzyme machineries.

The DNA is wrapped around the histone octamer so that it forms 1.65 turns of a left-handed superhelix within the nucleosome core particle (see Figure 1, blue chain is DNA with different shades of blue for each strand). The helical periodicity around the nucleosome core is 10.2 base pair as compared to 10.6 base pair for the helical periodicity of a free B-DNA. This small adjustment between free and nucleosomal DNA is largely a result of the torsion during wrapping into a superhelix, which also allows the minor grooves to bind to histones. Approximately 50% of the 120 direct protein-DNA interactions are formed by hydrogen bonds between the protein main chain amide groups and the oxygen atoms of the phosphodiester backbone, each time the minor groove faces the histone octamer (Muthurajan et al, 2003).

1.1 Linker histones

Linker histone does not have a histone fold. The canonical metazoan linker histone molecule contains three domains: two highly basic and unstructured tails in solution and a non-polar central globular domain. Whereas the sequence of the central globular domain is relatively well conserved through evolution in animals, plants and funghi, the N and C terminal domains are extremely heterogenous, both in length and amino acid composition. An interesting work has mapped the intereaction surface of linker histone H1 with the nucleosome of native chromatin in vivo (Brown et al, 2005) (Figure 4).

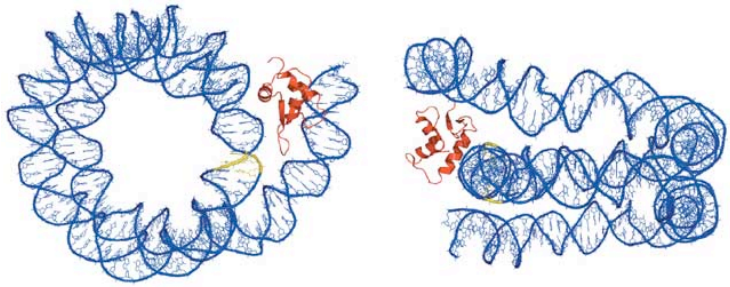


Figure 4. Molecular model of the location of H1 within the nucleosome. *Blue*, chromatosomal DNA; *yellow*, nucleosome dyad; *red*, the globular domain of H1. (from Brown et al, 2006)

2. The chromatin fiber

Nucleosomes are arranged as a linear array along the DNA polymer as “beads of a string”. This structure can be further compacted by linker histone H1 into higher-order transcriptionally inactive 30 nm fibres (Thoma et al, 1979 and Bednar et al, 1998). There are two models for the 30 nm fibre structure: the solenoid model and the “zigzag” model (Figure 5). Which one of the two models depicts the authentic structure is still not clear, though the form of 30 nm fibre might differ under different conditions. Subsequent stage of compaction involves looped domains of nucleosomal DNA, which are part of the most compact form of DNA found in the cell – the mitotic chromosomes, which reduces the linear length of DNA by 10,000-fold.

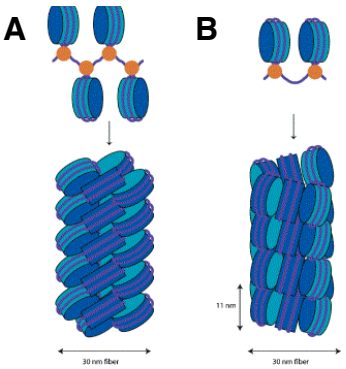


Figure 5. A schematic of the 30 nm fiber interpreted as a zigzag model of alternating nucleosomes vs a solenoid of adjacent nucleosomes. The histone octamer is shown in blue and the DNA is shown in magenta. **A**. The alternating aspect of adjacent nucleosome creates a zigzag pattern of packing. **B**. The consecutive arrangement of six nucleosomes in a turn of a helix can form a solenoid with 11 nm pitch. (from Khorasanizadeh, 2004)

Early studies interpreted this higher order organization of nucleosomes by invoking a **solenoid model** (Figure 5B). The solenoids involve six consecutive nucleosomes arranged in a turn of a helix that can condense into a supercoil structure with a pitch of 11 nm. A solenoid is a simple helix with one origin in which linker DNA has to be bended to connect each pair of consecutive nucleosome core. This structure is predicted to be held together by histone-histone interactions (Finch and Klug, 1976).

Biochemical and electron microscopy (EM) studies favor a **zigzag model** (Figure 5A) over the solenoid model for the arrangement of nucleosomes within the 30 nm fiber (Bednar et al., 1998). The zigzag formation promotes a stem-like organization of the entering and exiting linker DNA segments. In contrast to the solenoid model, the zigzag model uses the entry and exit paths of the DNA to establish the relative positioning of the nucleosomes, and not protein-protein interactions between nucleosomes. Moreover, in the zigzag model, alternate nucleosomes are physically closer than adjacent nucleosomes. A very recent work has provided more evidences towards the solenoid model as the underlying topology of the 30-nm fiber (Kruithof et al, 2009).

Section 2: Chromatin modification and gene regulation

The cell has developed mechanisms to modify in a temporal/spacial manner the chromatin organization and to ensure the maintenance of such an organization through mitotic and meiotic cell division:

a) Posttranslational covalent modification of histones within a nucleosome can either facilitate of hinder the accessibility of other transcriptional co-activators or co-repressors to chromatin.

b) Canonical histones in a nucleosome can be replaced by **histone variants** through a DNA-replication independent deposition mechanism.

c) Multisubunit complexes that use the energy of adenosine triphosphate (ATP), the so-called **ATP-dependent remodelling complexes**, can twist or slide nucleosomes exposing or occluding areas to interaction with regulator factors.

d) Methylation at the C-5 position of cytosine residues present in CpG dinucleotides by DNA methyltransferases (DNMTs) facilitates static long-term gene silencing. This is achieved through recognition of methyl-cytosine by specific methyl-DNA binding proteins that recruit transcriptional repressor complexes and histone modifying activities.

1. Histone modifications

Histone modifications have been associated with regulation of transcription, with some modifications seemingly correlating with a repressive state of chromatin (heterochromatin), while others seem to indicate transcriptionally active chromatin (euchromatin). These modifications can have two mechanistic functions. Certain histone modifications, such as incorporation of an acetyl group, will change electrostatic charge of histones' tails and consequently weaken DNA:histone interactions. On the other hand, epigenetic marks can function as anchoring sites for effectors proteins that can regulate transcription. There are at least eight distinct types of modifications of histones: acetylation, methylation, ribosylation, phosphorylation, ubiquitylation sumoylation, ADP ribosylation, deimination and proline isomerization

(Kouzarides, 2007; Lehnertz et al, 2003). All modifications with target residues as well as their related functions are shown in Table 1. Most modifications affect the N-terminal and C-terminal histone tails and few of them the globular domain.

Chromatin Modifications	Residues Modified	Functions Regulated
Acetylation	**K**-ac	Transcription, Repair, Replication, Condensation
Methylation (lysines)	**K**-me1 **K**-me2 **K**-me3	Transcription, Repair
Methylation (arginines)	**R**-me1 **R**-me2a **R**-me2s	Transcription
Phosphorylation	**S**-ph **T**-ph	Transcription, Repair Condensation
Ubiquitylation	**K**-ub	Transcription, Repair
Sumoylation	**K**-su	Transcription
ADP ribosylation	**E**-ar	Transcription
Deimination	**R > Cit**	Transcription
Proline Isomerization	**P-cis > P-trans**	Transcription

Table 1. **Covalent modifications of histones**. Table shows all known modifications and its related function (From Kouzarides, 2007)

The abundance of modifications on the histone tails makes "crosstalk" between modifications very likely (Figure 6). Firstly, many different types of modification occur on lysine residues (see Table 1 second column). This implies that distinct types of modifications on lysines are mutually exclusive. Secondly, the binding of a protein could be disrupted by an adjacent modification. The best example of this is the phosphorylation of H3S10 that affects binding of HP1 to methylated H3K9 (Fischle et al., 2005). Thirdly, an enzyme could recognize its substrate more effectively in the context of a second modification; for instance, GCN5 acetyltransferase, recognize H3 more effectively when it is phosphorylated at H3S10 (Clements et al, 2003). This crosstalk between histone modifications, named the "histone code" hypothesis, was first postulated by Jenuwein and Allis (2001).

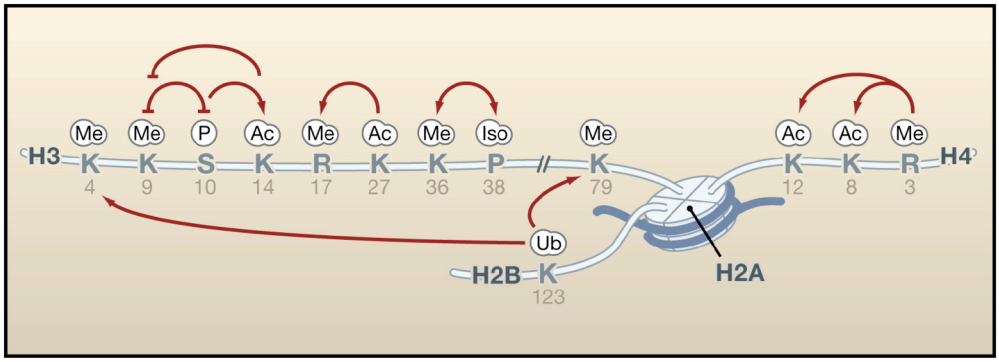


Figure 6. Crosstalk between Histone Modifications: the "histone code". Some examples of crosstalks between histone posttranslational modifications between residues from the same of different histone tail. The positive influence of one modification over another is shown by an arrow and the negative effect by a dish-line (from Kouzarides, 2007).

1.1 Enzymes that covalently modify the core histones

In addition to histones, transcription factors and other chromatin-associated proteins are themselve substrates of histone-modifying enzymes and are regulated by these posttranslational modifications in their binding affinity to DNA or interaction partners, transactivation capacity, protein stability and nuclear localisation (Berger et al, 2001 and Kouzarides et al, 2000).

From all possible histone posttranslational modifications we will focus on incorporation or removal of acetyl and methyl groups, as well as the histone modifier enzymes that catalyse these reactions.

1.1.1 Histone acetyltransferases and Histone deacetylases

The best characterized of the histone modifications is acetylation (Marmorstein et al, 2001 and Kouzarides et al, 1999), which is carried out by two families of enzymes, the histone acetyltransferases (HATs) and the histone deacetylases (HDACs). Acetylated histones always associate with transcriptionally active genes, whereas deacetylation is involved in transcriptional repression and heterochromatin formation. Consistent with this, many coactivators possess an intrinsic HAT activity.

Histone acetyltransferase are classified into two different classes (Figure 7), based on their functional localization:

1) Type A-HATs, which are nuclear
2) Type B-HATs, which are cytoplasmic and modify the newly synthesized histones before the assembly. HAT1 and HAT2 are the members of this family.

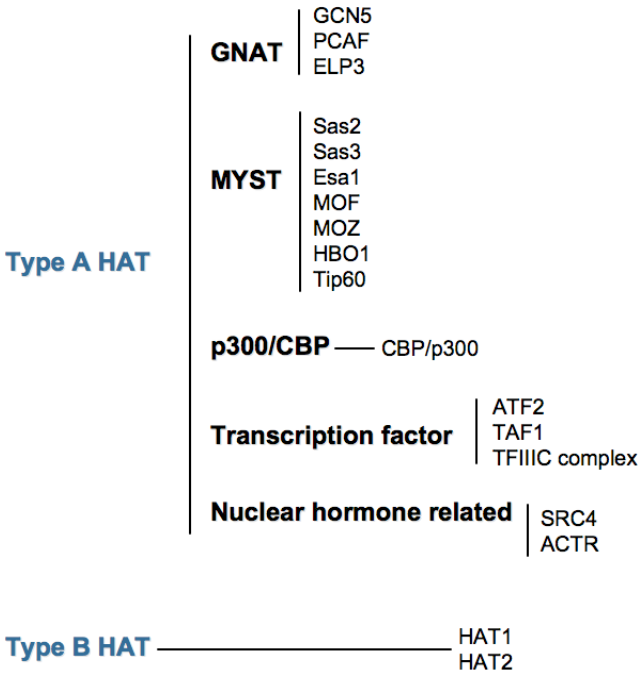


Figure 7. Classification of Histone acetyltransferases (HATs) depending on their localization. Type A is nuclear while type B is cytoplasmic.

The other five different classes of type A-HATs are classified based on structural and functional differences (Grant et al, 1999) (Figure 7). The **GNAT family** protein members are structured with a HAT domain of around 160 residues and a conserved bromodomain, which has been shown to recognize and bind acetyl-lysine residues. Gcn5 was the first histone acetyltransferase identified. This enzyme acts predominantly on histone substrates, but also on transcription activator factors (TAFs). Another member of this family that has a wide array of substrates is p300/CREB-binding protein (CBP)-associated factor (PCAF). It acetylates histones, nonhistone proteins like p53, E1A, and various other substrates with diverse physiological effects. This enzyme is an important transcriptional coactivator (Yang XJ et al, 1996). In both Gcn5 and PCAF, the N- and C-terminal regions play a direct role in histone substrate binding. ELP3 is another enzyme belonging to this class and acts only on histone substrates.

The **MYST family** of HAT proteins differ from the other groups from the catalytical mechanism involved in acetylation reaction. Members of the MYST family are involved in a wide range of regulatory functions including transcriptional activation, transcriptional silencing, dosage compensation and cell cycle progression.

The **p300/CBP family** is another major group of nuclear HATs that has been extensively characterized. Members of this family are more global regulators of transcription; contain a considerably larger HAT domain of about 500 residues, and other protein domains, including a bromodomain.

Some **transcription factors** have also been shown to possess acetyltransferase activity, thereby having a direct effect on transcriptional activation. The well-known members of this class are ATF2, TAF1, and TFIIIC90. The **nuclear hormone-related** HATs are relatively few; representative members are SRC-1, -3 and -4 and ACTR that act on the histone substrate. These enzymes were identified as nuclear receptor coactivators and then found to be endowed with histone acetyltransferase activity.

Histone acetyltransferase enzymes that show specificity for one of few sites are listed in the Table 2 along with the residues they modify in every histone:

Acetyltransferases	Residues Modified
HAT1	**H4** (K5, K12)
CBP/P300	**H3** (K14, K18) **H4** (K5, K8) **H2A** (K5) **H2B** (K12, K15)
PCAF/GCN5	**H3** (K9, K14, K18)
TIP60	**H4** (K5, K8, K12, K16) **H3**K14
HB01	**H4** (K5, K8, K12)

Table 2. Acetyltransferase enzymes and the modified residues. Enzymes with specificity for one or few sites are shown. (From Kouzarides, 2007)

Histone deacetylation (removal of an acetyl group from an acetylated lysine) is performed by histone deacetyltransferases (HDACs). This group of enzymes can be classified into three classes depending on their homology withy yeast HDAC proteins. Sin3 complex, NuRD and SIR2 complex are exemplary members of this family.

1.1.2 Histone methyltransferases and Histone demethylases

Methylation is associated either with activation or repression of transcription. Indeed, this modification can have different biological outcomes depending on the residue and on the timing on which it occurs on the same residue (mono-, di- and tri-methylation). Thus, for example, highly condensed heterochromatic regions show a high degree of trimethylated H3K9, whereas euchromatic regions are preferentially enriched in mono- and dimethylated H3K9 (Martin et al, 2005 and Jenuwein T, 2006).

Histones are methylated at arginine and lysine residues. Conversely to HATs, histone methyltransferases (HMTs) are quite specific enzymes that usually modify one single residue on a specific histone. Unlike histone acetylation, histone methylation does not change nucleosome charge, but usually it can be recognized by effector proteins displaying chromo-like domains.

Histone lysine methylation is mediated by HMTs, many of which contain a conserved SET [Su(var)3-9, Enhancer-of-zeste, Trithorax] domain, such as Suv39h1 (Suppressor of variegation 39h1) and G9a (Kouzarides T, 2007; Lachner et al, 2004; and Schneider et al, 2002).

Histone demethylases catalyze the removal of methyl groups from histone lysines and arginines. Since the half-life of histones and of their methylated lysines is identical, lysine methylation has been considered for years to be an irreversible mark (Byvoet et al, 1972, Duerre and Lee, 1974). However, recent progress in the field led to the identification of two enzyme-families capable of catalysing this reaction, the LSD-1 family and the jumonji-family.

2. Incorporation of canonical histone variants and associated functions

2.1 Core histone variants

In higher organisms, each histone is represented by a family of genes encoding multiple non-allelic primary-sequence variants. It is thought that histone variants serve two main purposes in the cell. First, the histone-exchange removes epigenetic marks on core histones and facilitates reprogramming of the particular gene. Second, their incorporation permits effectuation of various functions. Figure 8 shows all histone variants and the functions to which they are associated to that exist in mammalian cells. The most studied histone variants are the ones of histones H3 and H2A. These include H3.3, which functions in transcriptional activation (Ahmad and Henikoff, 2002) and CENPA - centromeric histone H3 (Sullivan et al., 1994), H2A.Z, which plays a role in gene expression and chromosome segregation (Allis et al., 1986; Rangasamy et al., 2004), H2AX, which acts in DNA repair and recombination (Rogakou et al., 1998; de la Barre et al., 2001), macro H2A – transcriptional repression and X chromosome

inactivation (Costanzi and Pehrson, 1998; Angelov et al., 2003), and H2ABBD possibly functioning in transcriptional activation (Bao et al., 2004; Gautier et al., 2004).

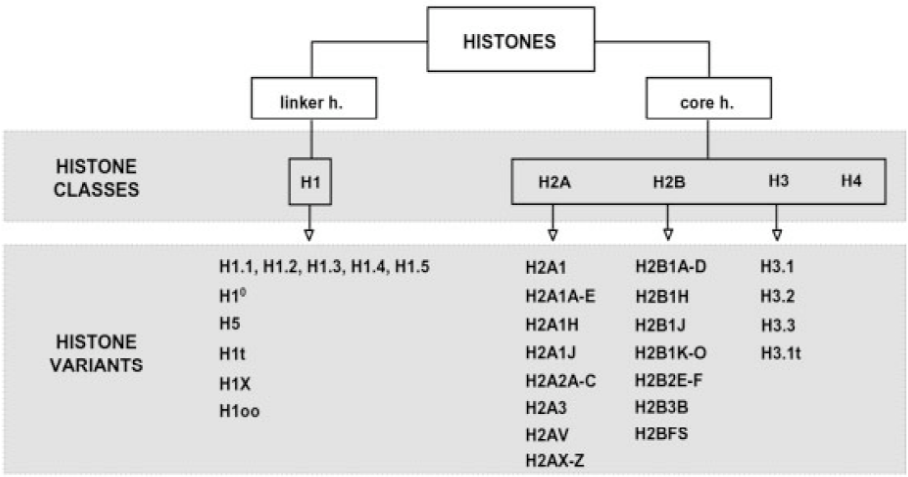


Figure 8. Synopsis of core histone and linker histone variants in human cells (From Lindner, 2008)

2.2 Linker histone variants

In superior eukaryote there are multiple isoforms or subtypes of linker histone H1 that are important for cell growth and proliferation (for review see, Khochbin, 2001). In mammals, six somatic variants (designated H1.0, H1.1, H1.2, H1.3, H1.4 and H1.5), a male germ line-specific subtype (H1t) and an oocyte-specific subtype (H1oo) have been identified (Panyim and Chakley, 1969; Bucci et al, 1982; Lennox et al, 1984; Tanaka et al, 2001) (Figure 8). The variants differ in timing of expression (Khochbin et al, 2001), extent of phosphorylation (Lennox et al, 1982), turnover rate (Hall et al, 1985; Dominguez et al, 1992), binding affinity (Th'ng et al, 2005), and evolutionary stability (Ponte et al, 1998). Recently, our lab has published an extensive work on H1 variants in T47D breast cancer cells in which H1.2 and H1.4 depletion specifically caused arrest of cell proliferation (Sancho et al. 2008) In particular, H1.2 caused decreased in global nucleosome spacing and repressed expression of a number of cell cycle genes, supporting the idea that distinct roles exist for the linker histone variants (Sancho et al, 2008).

3. ATP-dependent chromatin remodelling complexes

This type of chromatin remodelling complexes uses the energy of ATP to perform nucleosome alterations. The ATP-dependent chromatin remodelling complexes comprise a central helicase-related, ATP-consuming subunit, with various

numbers of associated subunits that modulate and target its activity. In humans, these complexes are divided into four classes depending on the structural domains contained in the ATPase protein subunit. The four family proteins are (Figure 9):

a) SWI2/SNF2 family, containing a bromodomain.

b) ISWI family, containing a SANT domain.

c) CHD family, containing a cromodomain and a DNA-binding motif.

d) INO80 and SWR1 families, without any additional domain.

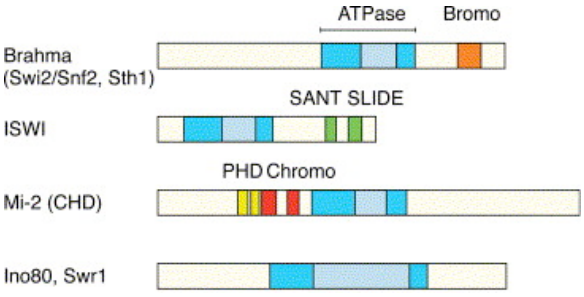


Figure 9. ATPases of the four main families of ATP-dependent Swi2/Snf2-related chromatin remodelling complexes: SWI/SNF, ISWI, CHD (or Mi-2) and INO80. Each family is defined by the presence of a distinct ATPase containing signature structural domains and a unique subunit composition. In addition, an extended split in their ATPase domain characterizes the INO80 and SWR1 class.

(From Mohrmann and Verrijzer, 2005)

These ATPases that define subfamilies are present in multicomponent complexes, having each of them numerous subunits (typical of SWI2/SNF2 subfamily) or few (two or four in ISWI subfamily).

3.1 The four classes of remodelling complexes

SWI2/SNF2 family

The components of the SWI/SNF chromatin-remodelling complex were initially identified in yeast in screens for genes that regulate mating-type switching (SWI) and sucrose non-fermenting (SNF) phenotypes (Carlson et al., 1981; Neigeborn and Carlson, 1984, 1987; Stern et al., 1984; Abrams et al., 1986; Nasmyth and Shore, 1987 and Carlson and Laurent, 1994). It was recognized that a subset of the SWI genes are identical to those identified in the SNF screen, and those genes that are involved both in mating-type switching and sucrose fermentation have come to be known as SWI/SNF genes (Peterson et al., 1994 and Wolffe, 1994). The ySWI/SNF remodelling complex is a multiprotein complex with a central ATPase subunit, known as Swi2/Snf2. Homologs of SWI/SNF complexes were found in *Drosophila* and humans.

The mammalian SWI/SNF complexes have two mutually exclusive ATPase subunits, named BRM and BRG1 (these are ortholog subunit of the yeast Swi2/Snf2), and 8–10 subunits, which are referred to as BRM- or BRG1-associated factors or BAFs (Wang et al., 1996). Table 3 lists all of the well-characterized mammalian subunits of the SWI/SNF complex and shows their non-vertebrate orthologs, the *Drosophila* BAPs and yeast SWI/SNF gene family members. The yeast SWI/SNF complex exhibits an apparent molecular mass of 1.14 MDa (Smith et al., 2003), whereas the mammalian SWI/SNF complex has an apparent molecular mass of approximately 2 MDa. Mammalian BAF proteins are conventionally identified by their molecular size; hence BAF47 refers to a BRG1-associated protein with an apparent molecular mass of 47 kDa. Human BRG1 is approximately 74% identical to human BRM (Khavari et al., 1993), 52% identical to *Drosophila* BRM and 33% identical to yeast Swi2/Snf2 (Fry and Peterson, 2001).

Yeast		*Drosophila*		Human	
SWI/SNF	**RSC**	**BAP**	**PBAP**	**BAF**	**PBAF**
Swi2/Snf2	Sth1	Brahma	Brahma	BRG1 or hBRM	BRG1
Swi1/Adr6		OSA		BAF250	
	Rsc1, Rsc2, Rsc4		Polybromo		Polybromo/BAF180
	Rsc9*		BAP170*		
Swi3	Rsc8	Moira	Moira	BAF170 & BAF155	BAF170 & BAF155
		BAP111	BAP111	BAF57	BAF57
Swp73	Rsc6	BAP60	BAP60	BAF60a	BAF60a or BAF60b
Swp61/Arp7 Swp59/Arp9	Rsc11/Arp7 Rsc12/Arp9	BAP55	BAP55	BAF53	BAF53
		actin	actin	actin	actin
Snf5	Sfh1	Snr1	Snr1	hSNF5/INI1	hSNF5/INI1
	Rsc5, 7, 10, 13-15				
	Rsc3, Rsc30				
Swp82					
Swp29/Tfg3/TAF30/Anc1					
Snf6, 11					

Table 3. Relationship between SWI2/SNF2-subfamily chromatin remodelling complexes from yeast, *Drosophila* and mammals. Conserved subunit are listed horizontally with the ATPase subunit listed first. SWI/SNF (Switching defective/Sucrose nonfermenting); RSC (Remodel the Structure of Chromatin); BAP (BRM-Associated Proteins); PBAP (Polybromo-associated BAP); BAF (BRG1/hBRM-Associated Factors; PBAF (Polybromo-associated BAF). Distinction between ySWI/SNF and RSC has been conserved throughout eukaryotic evolution. ySWI/SNF class remdodelers are referred as BAP and BAF, whereas PBAP and PBAF are RSC orthologues (From Mohrmann and Verrijzer, 2005)

Individual SWI/SNF complex contains either BRM or BRG1, but not both (Wang et al., 1996), such that BRM/BAF complexes are structurally distinct from BRG1/BAF complexes (see Table 3). The extent to which these complexes are functionally distinct is a topic of active investigation. The BRG1-containing complexes are further divided into those that contain the BAF250 (or OSA protein) or the BAF180 protein (Wang, 2003). BAF180 is the mammalian ortholog of *Drosophila* polybromo, and the BAF180-containing SWI/SNF complex has been designated as PBAF, to distinguish it from the BAF250-containing BRM/BAF and BRG1/BAF complexes (Table 3).

PBAF complex is thus defined by the presence of BAF180 and BAF200 whereas BAF complex is exclusively containing the BAF 250 subunit. BAF180 (polybromo, PB1, PBRM1)-containing SWI/SNF complexes have different properties from those that contain BAF250 subunits (Nie et al., 2000; Lemon et al., 2001 and Wang et al., 2004). BAF180 harbors a distinctive set of structural motifs, characteristic of three components of RSC (Xue et al., 2000 and Table). It also contains a Bromo domain that binds to acetylated histones. BAF200 is also believed to function as part of the PBAF complex, similar to BAF180 (Yan et al., 2005). BAF200, like BAF250, is a member of the ARID gene family and is encoded by ARID2.

ISWI subfamily

The ISWI (Imitation of SWI) based complexes were initially purified from *Drosophila*. Genetic studies in this organism suggest that ISWI based complexes can activate gene expression in both *in vitro* and *in vivo* (Dirscherl and Krebs, 2004). *In vitro*, most of the ISWI-based complexes have nucleosome spacing activity that results in the formation of regulatory ordered nucleosome arrays. These properties have been used to assemble chromatin *in vitro* using an ATP-utilizing chromatin assembly and remodelling factor (ACF) and a histone chaperone (such as NAP-1 or CAF-1) (Ito et al, 1997). *In vivo* it is thought that this ability is related with the chromatin assembly following DNA replication and chromosome organization (Corona and Tamkun, 2004).

CHD subfamily

Mi-2/NuRD complexes are highly heterogeneous in subunit composition but they all contain Mi-2 as the core ATPase. The complexes also contain HDACs, methylated DNA binding proteins (MBDs), histone H4-interacting proteins

(Retinoblastoma associated proteins PbAp 46/48) and members of metastasis-associated protein gene family (MTAs). Because both HDACs and methylated DNA are usually associated with gene repression and silencing, these complexes have been involved in transcriptional repression (Bowen et al, 2004).

INO80 and SRW1

They are the ATPase of distinct multisubunit complexes firstly found in yeast. INO80 complexes are involved in transcriptional activation and DNA damage repair (Shen et al, 2004). They also catalyze ATP-dependent sliding of nucleosomes along DNA. Recently it has been also characterized the human complex (Jin et al, 2005). The SWR1-based complex is responsible for the incorporation of the histone H2A.Z variant (Htz1 in yeast) into nucleosomes in a replication-independent manner (Mizuguchi et al, 2004). A human complex with many similarities has being discovered (Cai et al, 2005).

3.2 Mechanisms of chromatin remodelling

The energy for SWI/SNF-mediated chromatin remodelling is transduced by the catalytic subunit, BRM or BRG1, both of which have DNA-dependent ATPase activity (Muchardt and Yaniv, 1999 and Hassan et al., 2001). The function of SWI/SNF complex has been thoroughly studied *in vitro*. It was shown to be capable of several biochemical activities that lead to chromatin disruption events (Figure 10). It causes mobilization of histone core *in cis* (on the same DNA molecule), called nucleosome sliding (Schnitzler et al, 1998, Whitehouse et al, 1999 and Schnitzler et al, 2001). Sliding is nucleosome translocation in which DNA sequences previously occluded by core histones are exposed in internucleosomal regions. SWI/SNF complex also transfer histone octamer *in trans* to other DNA molecules (Lorch et al, 1999), cause formation of altered dinucleosome particles (Lorch et al, 1998 and Peterson, 2000), generation of negative superhelical torsion on linear DNA and nucleosomal templates (Havas et al, 2000) and partial or total displacement of core histones (Boeger et al, 2003; Bruno et al, 2003; Reinke and Horz, 2003; Vicent et al, 2004). Biochemical evidence and recent atomic force microscopy studies indicate that SWI/SNF remodelling involves both sliding and disruption of histone-DNA interactions. Despite extensive study, the mechanism by which SWI/SNF complexes remodel nucleosomes is not well understood.

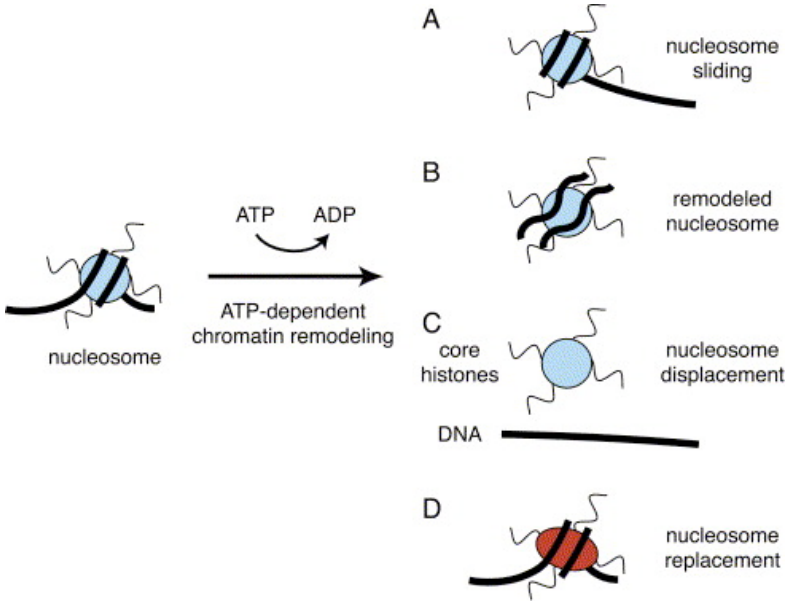


Figure 10. Schematic model depicting different modes of ATP-dependent chromatin remodelling. ATP-dependent chromatin remodelling complexes use the energy derived from ATP hydrolysis to alter histone-DNA contacts in such a way that (**A**) nucleosomes slide to another position, (**B**) a remodelled state is created in which the DNA becomes more accessible but histones remain bound, (**C**) DNA and histones dissociate completely, or (**D**) histone (variant) replacement. (From Mohrmann and Verrijzer, 2005).

These perturbations increase susceptibility of nucleosomal DNA to DNaseI and restriction endonucleases (Coté et al, 1998; Imbalzano et al, 1994 and 1996), change the crosslinking pattern between histones and nucleosomal DNA (Lee et al, 1999 and Sengupta et al, 2001) and alter the number of DNA super coils constrained by nucleosomes (Kwon et al, 1994; Bazett-Jones et al, 1999 and Jaskelioff et al, 2000).

3.3 SWI/SNF-related functions

SWI/SNF is a regulator of gene expression. In mammalian cells, SWI/SNF has been linked to a large number of transcription factors. The oncogenic transcription heterodimer activated protein-1 (AP-1) is known to be SWI/SNF dependent (Ito et al., 2001); similarly, EKLF, which regulates beta-hemoglobin synthesis, also requires this complex for its function (Armstrong et al., 1998 and Lee et al., 1999). All known steroid receptors are functionally linked to SWI/SNF (Yoshinaga et al., 1992; Sumi-Ichinose et al., 1997; Fryer and Archer, 1998; Belandia et al., 2002 and Vicent et al, 2004). Hence, SWI/SNF does not regulate an exclusive signalling pathway; instead, it serves as a fundamental component of various essential, and often unrelated, pathways.

SWI/SNF activity is essential for differentiation in yeasts, flies and mammals. Moreover, SWI/SNF factors are also involved in malignant transformation, an association most clearly demonstrated with the SWI/SNF subunit BAF47. The strongest evidence for the role of BAF47 in cancer development comes from studies on rhabdoid tumors showing that one BAF47 allele is consistently deleted, and the other allele is either mutated or silenced by methylation (Versteege et al., 1998; Rousseau-Merck et al., 1999; Sevenet et al., 1999; Biegel et al., 2000; Biegel and Pollack, 2004). Other than transcription regulation, SWI/SNF has been linked to different cellular mechanisms such as alternative splicing, with a complex that associates Brm with Sam68 (Batsché et al, 2006) and repression of transcription interacting with the methyl-CpG binding protein MeCP2 (Harikrishnan et al, 2005).

Section 3: Nucleosome positioning and gene regulation

1. The concept

Condensation of the genomic material into chromatin is carried out in an effective and precise way that allows all major cellular processes to take place. Furthermore, it has been shown that the effective packaging of DNA into nucleosomes can actively take part into the regulation of transcription, allowing the access of transcription factors at specific sites, while protecting others from being transcribed. Thus, the positioning of nucleosomes, their translocation or disruption may play a crucial role in gene expression and regulation.

2. A brief historical overview: searching DNA sequence patterns

Nucleosome positioning was first addressed as a structural more than a functional issue. First works done on the field were based on the alignment of 147 bp DNA sequence of hundreds of well-positioned nucleosomes and the identification of particular base pair combination commons in all of them. Early in the 90's, a number of works defined sequence prerequisites for nucleosome positioning (or exclusion), although a definite "nucleosomal signature" needs still to be described. Early characterisation of nucleosomal DNA revealed several nucleosome positioning signals: an abundant of AA/TT dinucleotides separated by roughly 10 bp (Drew HR et al, 1985), five repeated $(A/T)_3NN(G/C)_3NN$ motifs embedded in a DNA fragment, the so-called TG pentamer, and fragments having repeated TATAAACGCC motifs, referred to as TATA tetrads (Widlund HR et al, 1997). More extended periodicities, have also been noted. The trinucleotide pattern VWG (V = non-T, W = A or T) was observed to follow a ~11bp period in nucleosomal sequences (Baldi et al., 1996; Stein and Bina, 1999). As dinucleotide periodicity studies became more detailed due to the increase of available sequences (Widom et al, 1996) a more consistent periodical pattern was revealed, according to which two consecutive purines (RR) were to be observed with a period of ~10.5 counter-phase with a corresponding occurrence of two consecutive pyrimidines (YY) with the same period (Kogan and Trifonov, 2005). Finally, periodically occurring triplets were correlated to gene expression of yeast mini-chromosomes early in 2002 by Tomita and colleagues suggesting a link between chromatin structure and transcription.

3. Recent advances in nucleosome cartography: genome-wide studies

The first genome-scale identification of nucleosome positions was done for the whole chromosome 3 of *Saccharomyces cerevisiae* using the tilling microarray technique (Yuan et al, 2005). Based on statistical calculations, this study revealed that between 65 and 69% of the yeast DNA may be organized in well-positioned nucleosomes. Recently, the densities of printed probes on microarrays have increased dramatically, allowing millions of genomic loci to be interrogated by ChIP-chip analysis in a single experiment. The first two complete genomic maps of positioned nucleosomes were published for Lee et al (2007) and Mavrich et al (2008, Nature) for yeast and *Drosophila* genomes, respectively. However, a new genome-scale techniques based on massively parallel sequencing was developed to study more precisely nucleosome positioning in complete genomes. The first such high-resolution genome-wide ChIP-Seq nucleosome map was achieved for nucleosomes containing H2A.Z in *S.cerevisiae* (Albert et al, 2007). Nucleosome maps of a similar resolution in yeast (Whitehouse et al, 2007 and Shivaswamy et al, 2008), worm (Valouev et al, 2008) and humans (Ozsolak et al, 2007, Schones et al, 2008 and Barski et al, 2007) have now been published, helping to understand how nucleosome positions are exactly determined *in vivo*.

From this and other works on the field, we have learned some points that help to clarify the difficult goal of understanding nucleosome positioning:

1) *Nucleosome organization around transcription start sites*. Nucleosomes at most genes are organized around the beginning of genes in basically the same way: a nucleosome free-region (NFR) of about 140 bp flanked by two well-positioned nucleosomes (the -1 and the +1 nucleosome, up and downstream of the NFR, respectively), which is followed by a nucleosomal array that package the gene. This pattern has been basically described for yeast (Yuan et al, 2005), but also observed for some genes in metazoan genomes, such as *Drosophila* (Mavrich et al, 2008-nat) and human (Barski et al, 2007) genomes. (Figure 11)

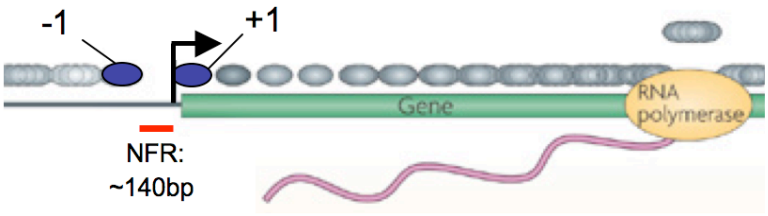


Figure 11. Nucleosome-free region around the transcription start sites in yeast genes. Nuclesomes -1 and +1 flancking the Nucleosome-Free Region (NFR) are coloured in blue. The arrow represents the TSS. (From Jiang et al, 2009).

2) *Nucleosome spacing thoughout the genome.* There have been observed differences among the species in distance between adjacent nucleosome midpoints. *S. cerevisiae* has approximately 165 bp distance between adjacent nucleosomes (that is a linker DNA of ~18 bp) (Lee et al, 2007, Mavrich et al 2008-gen research and Shivaswamy et al, 2008); for *D. melanogaster* (Mavrich et al, 2007) and *C. elegans* (Valouev et al, 2008) there is a 175 bp distance in average (that is linkers of ~28 bp); and for humans (Barski et al, 2007 and Schones et al, 2008) it has been described a 185 bp distance between neighbour nucleosomes (and it means a linker DNA of ~38 bp). This differences in linker length may be explain by absence and presence of linker histone H1 and other remodelling factors that have species-specific DNA length requirements for binding, as well as increasing complexity of genomes through evolution.

And some other general conclusions such as:

1) Nucleosomes positions are highly dynamic. (Mellor et al, 2005). Hogan et al. (2006) have reported cell cycle-specific fluctuation of nucleosome occupancy at gene promoters and Shivaswamy et al. (2008) have identified changes in individual nucleosome positions before and after subjecting cells to heat shock.

2) There exist other factors that in specific locations can contribute to nucleosome positioning (Saha et al, 2006, Whitehouse et al, 2007, Pechham et al, 2007 and Whitehouse et al, 2006), indicating that no only the DNA sequence is acting as a positional constraint in the organization of nucleosomes. A well example is the work done by Whitehouse and colleagues (2007). They have shown that the ATP-dependent chromatin remodelling complex Isw2 repositions nucleosomes onto promoter regions that are intrinsically designed to exclude them. This Isw2-dependent nucleosome repositioning enforces directionality on transcription by preventing transcription initiation from cryptic sites.

4. Yeast: the best characterized nucleosome occupancy map

Probably due to its low genomic complexity and reduced size, yeast has been the best well-characterized genome in terms of nucleosome organization. What we have learned from yeast works is summarized in Figure 12. Firstly, nucleosomes are well organized around the transcription start sites (TSS) and transcription termination sites (TTS) of genes. Secondly, there is a nucleosome free region (NFR) right upstream of the TSS and downstream of TTS. Nucleosome -1 is located approximately between nucleotide position -150 and -300 and it is very dynamic and usually suffers one or more of the following modifications: histone replacement, acetylation, methylation, translational repositioning and finally completely eviction. All these modifications facilitate the pre-initiation complex (PIC) formation and allow transcription to take place. After the NFR we found the nucleosome +1. It is usually defined by the presence of histone variants such as histone H2A.Z and H3.3 (Malik et al, 2003), as well as by methylation and acetylation of residues in its histone tails (Kouzarides, 2007 and Li et al, 2007). This nucleosome is most likely evicted when transcription takes place but rapidly return to its location after the Pol II has passed. Downstream of the nucleosome +1, we find nucleosome +2 and +3 that mostly share the same characteristics of the nucleosome +1 (content histone variants and histone acetylation and/or methylation) but in a less consistent way (Li et al, 2007 and Lieb et al, 2005). Constraints for nucleosome positioning disappear as we go inside the coding region of the gene. In contrast, nucleosome organization around the NFR of the TTS is not yet as well known as it is in the TSS.

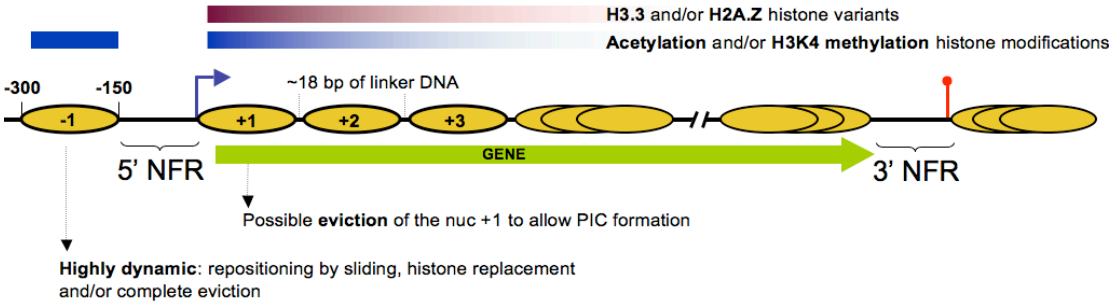


Figure 12. General roles for nucleosome organization in yeast genes. Nucleosomes are represented as yellow ovals. Numbers inside nucleosomes correspond to the consensus nucleosome numbering respect the NFR (Nucleosome-Free Region). Transcription start site (TTS) is marked with a blue arrow, while transcription termination site (TTS) is represented with a line and a small circle in red. The probability to find acetylation and H3K4 methylation as well as some histone variants is shown in different intensities of blue and dark red, respectively. Less intensity of these colours means less probability to find these histone marks and variants on consecutives nucleosomes (for instance, nucleosome+3 displays less histone variant H2A.Z than nucleosome +2).

5. Methods to predict nucleosome positions

Many studies have attempted to predict computationally nucleosome positioning *de novo* in different organisms based on properties of the DNA, since Segal et al. published in 2006 a nucleosome-DNA interaction model that highlight that ~50% of the nucleosomal positioning is governed by the sequence. Briefly, they used the AA/TT/TA dinucleotide frequency patterns obtained from a set of 199 positioned mononucleosomes to construct a probabilistic model, representative of the DNA sequence preferences of nucleosomes in *Saccharomyces cerevisiae*. A following study (Ioshikhes et al, 2006) used comparative analysis of six *Saccharomyces* species combined with the frequencies of AA/TT dinucleotides to observe conserved patterns of nucleosome positioning among yeast species, pointing out the existence of sequence constraints in nucleosomal positions. Both of the above studies identified positions of NFRs in the vicinity of promoter and gene-upstream regions, in agreement with experimentally defined ones (Yuan et al, 2005). A more recent method (Peckham et al, 2007) made use of the original raw experimental data of Yuan et al. (2005) to develop a method that distinguish nucleosome occupancy from depletion based on the GC content of the sequence, which constitutes a very general attribute to be assigned as a nucleosome-formation pattern. All these studies that have used DNA sequence features to predict genome-wide nucleosome positions (Yuan et al, 2008, Ioshikhes et al, 2006, Segal et al, 2006, Miele et al, 2008, Field et al, 2008, Pechham et al, 2007), confirm that nucleosome positioning is partially encoded in the genomic DNA sequence. The most accurate method that makes *de novo* predictions of nucleosome positions till now is probably Gupta et al. (2008).

Section 4: Progesterone-responsive gene regulation

1. Nuclear receptors

Nuclear receptors (NR) act as ligand inducible transcription factors by directly interacting as monomers, homodimers or heterodimers with DNA response elements of target genes, as well as by cross-talking to other signalling pathways. The effects of NR on transcription are mediated through interaction with other transcription factors and through recruitment of coregulators. The nuclear receptor superfamily is typically subdivided into three families:

1) The *steroid receptor family*,

2) the *thyroid/retinoid family*, that includes thyroid receptor (TR), vitamin D receptor (VDR), retinoic acid receptor (RAR), and peroxisime proliferators-activated receptor (PPAR) and,

3) the *orphan receptor family*, defined by a set of proteins identified by comparative sequence analysis as belonging to the nuclear receptor superfamily, but which ligands are unknown.

NR can be also grouped into a large superfamily depending on the way in which they bind to DNA, as homodimers (ER, GR, PR, AR and MR), as heterodimers (RAR, RXR, TR, VDR and EcR) or monomers (SF1) (Dilworth and Chambon, 2001).

2. Steroid hormones (SH) and steroid hormone receptors (SHR)

Hormones are chemical signalling molecules produced by endocrine glands of vertebrates and other organisms. Their functions vary depending on the tissue or cell type. Several hormones and their receptors have been shown to play an important role in numerous cellular functions, including metabolic regulation, immune system responses, cell growth and development. One of the four classes of hormones is steroid hormones, which are small cholesterol-derived hydrophobic molecules that can traverse the cell membrane and activate their specific steroid hormone receptors. Mammalian steroid hormones, produced by gonads and adrenal glands, can be grouped into five groups by the receptors to which they bind: glucocorticoids, mineralocorticoids, androgens, estrogens, and progestagens. It is through these receptors that the steroid hormones exert their regulatory effects in a wide variety of biological processes including reproduction, differentiation, development, cell

proliferation, apoptosis, inflammation, metabolism, homeostasis and brain function (Mangelsdorf et al, 1995). Unliganted SHRs are assossiated with a large multiprotein complex of chaperons, including heat shock protein 90 (Hsp 90) and immunophilin Hsp56.

Steroid hormone receptors (SHRs) contain two structural subunits (Figure 13): firstly, a C-terminal ligand binding subunit (LBD) and secondly, a centrally located DNA-binding domain (DBD). LBD contains an interior binding pocket specific for its corresponding hormone. It also contains a ligand-regulated transcriptional activity function (AF-2) needed for recruiting co-activating proteins. Finally, the LBD is the primary mediator of dimerization, necessary for DNA response elements binding (Kumar, Chambon, 1988). The DBD is the responsible of the binding to the Hormone Responsive Elements (HREs, explained below). The N-terminal of DBD contains a transcriptional activation function (AF-1) and shows weak conservation between the SHRs family members. Finally, the DBD is connected to the LBD via a short amino acid sequence termed the hinge.

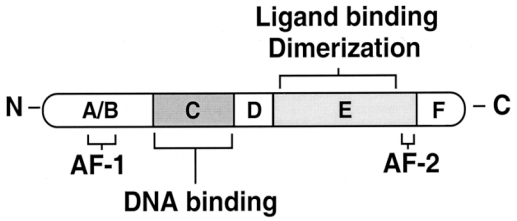


Figure 13. Schematic representation of a nuclear receptor. The variable N-terminal A/B region contains the ligand-independent AF-1 transactivation domain. The conserved DNA-binding domain (DBD), or region C, is responsible for the recognition of specific DNA sequences. A variable linker region D connects the DBD to the conserved E/F region that contains the ligand-binding domain (LBD) as well as the dimerization surface. The ligand-dependent AF-2 transactivation domain is the C-terminal portion of LBD (From Aranda and Pascual, 2001).

3. Progesterone receptor

3.1 Structure

Progesterone receptor (PR) belongs to the family of ligand-activated nuclear hormone receptors, and it mediates the physiologic effects of progesterone The progesterone receptor, like all steroid receptors, is composed of four domains: a modulating N-terminal domain, the hinge region, a DNA-binding domain (zinc finger) and a C-terminal steroid binding domain (Figure 14).

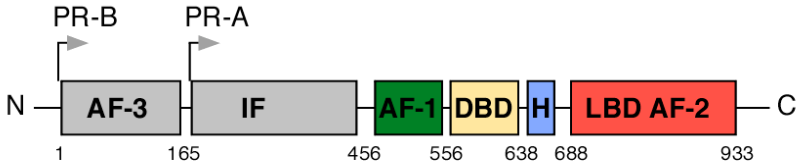


Figure 14. Structure of the progesterone receptor (PR_A and PR_B are the two isoforms). The numbers refer to amino acid positions. Arrows indicate the amico acid initation of PR isoforms. AF-3: activation function 3; IF: inhibition function; AF-1Activation Function1; DBD: DNA binding domain; H: hinge, short amino acid sequence that connects DBD with LBD; LBD AF-2: comprise ligand binding domain and activation function2.

The gene encoding the PR protein contains separate promoters and translational start sites to produce two isoforms – PR_A and PR_B, that differ in their molecular weight (additional 164 amino acids present only in the N terminus of PR_B) (Figure 14). Although PR_A and PR_B isoforms show high degree of sequence identity, they display significant different functional properties on the regulation of target promoter. PR_B is a much stronger transcriptional activator than PR_A (Sartorius et al, 1994). PR_A gene knock-out mice develop uterine dysplasia and abnormal ovaries, whereas PR_B gene knock-out have affected the mammary glands, causing incomplete lobular-alveolar differentiation (Mulac-Jericevic et al, 2000). Moreover, microarray studies showed that the two isoforms regulate different subset of genes (Richer et al, 2002).

3.2 PR effects on gene regulation

Progesterone receptor is a sequence-specific transcription factor. Upon binding of hormone to its carboxy terminal domain it undergoes homodimerization. The activated receptor plays a role in control of gene activation by either *1.* direct binding to hormone responsive elements (HREs) in regulated genes (pathway known as ***genomic*** or ***direct effects***) or *2.* induction of kinase cascades activated by cytoplasmic events (pathway known as ***non-genomic*** or ***signalling-mediated effects***).

The genomic pathway implicates direct binding to HREs in regulatory regions, where it recruits modifiers which remodel the nucleosomes and these in turn recruit other co-regulators. The critical interactions with chromatin and its complement of associated proteins are therefore crucial for progesterone receptor action. The non-

genomic pathway often occurs via second messenger cascades, which in turn originate from signalling complexes located at membranes. For instance, estrogens activate the Src/Ras/Erk and PI3K/Akt pathways via direct interaction of ER with c-Src and the regulatory subunit of PI3K, respectively (Castoria et al, 2001 and Migliaccio et al, 1996). Progesterone receptor also cross-talk to kinases cascades through an interaction of PR with SH3 domain of c-Src (Ballaré et al, 2003 and Boonyaratanakornkit et al, 2001). The targets of the activated kinases cascades might be transcription factors and co-regulators involved in DNA synthesis and cell proliferation.

3.3 Co-activators

Co-activators recruited by ligand-bound nuclear receptors include chromatin remodelling complexes, such as SWI/SNF, as well as members of the steroid receptor coactivators family (SRC) (Beato, Klug 2000), which serve as adaptors to mediate interactions with histone acetyltransferases (HATs) (Torchia et al, 1997 and Voegel et al, 1998). SRC family consists of three members: SRC-1 (or NCoA-1), SRC-2 (or NCoA-2, GRIP-1 or TIF-2) and SRC-3. C-terminal domains of SRC-1 and -3 contain histone acetyltransferase (HAT) activity although it is much weaker than those in CBP, p300 or PCAF enzymes (Spencer et al, 1997). SRC pre-existing complex with CBP, p300, PCAF, CARM-1 and PRMT-1 are recruited to chromatin by ligand-triggered interaction between SHRs and SRCs (Li et al, 2003).

In addition to co-activators that interact directly with transcriptional activation domains of SHRs, proteins that modulate sequence specific DNA binding SHRs have been identify. The best characterized of these factors are the nuclear high mobility group of proteins HMG-1 and related HMG-2, that enhance the affinity of the SHRs for their specific target HREs. HMG-1 and -2 in mammalian cells enhance the hormone-dependent transcriptional activity of several different SHRs including PR, AR, GR and ER, contributing to the stability of the receptor-DNA complex (Onate et al, 1994).

3.4 Co-repressors

Several groups have discovered that antagonists of ER and PR promote receptor association with the co-repressor NCoR (Nuclear receptor co-repressor) and

SMRT (silencer mediator or retinoc and thyroid receptor) (Wagner et al, 1998). NCoR and SMRT are multiprotein complexes that exhibit deacetylase activity (HDAC), indicating that targeted deacetylation of core histones is necessary for the silencing activity of these proteins. Thus, NCoR and SMRT have opposite effect on chromatin structure to HATs and repress access of general transcription factors to DNA.

4. Hormone Responsive Elements (HREs)

Hormone Responsive Elements were initially described as glucocorticoid responsive elements (GREs) (Karin et al, 1984 and Scheidereit et al 1983). They have been classically described in the 5' regulatory region of a given gene but recent evidences show HREs can be located in coding regions (exons and introns) and intergenic regions. HREs are generally composed of two hexameric core motifs. Two such consensus motifs have been identified (AGAACA and AGG/TTCA), which are recognized by different members of the nuclear receptors superfamily, though usually HREs can show significant variations from the consensus. Steroid hormone receptors including glucocorticoids, mineralocorticoids, progesterone and androgens typically bind to palindromes of AGAACA sequences separated by three non-conserved base pairs (Beato, 1989), with the exception of estrogen receptor, which recognizes AGGTCA motif. Modification of location of the half-sites relative to one another contributes to diversity of HREs. For dimeric HREs, the half-sites can be configured as palindromes, inverted palindromes or direct repeats. HREs additionally vary in a number of neutral base pairs separating the half-site repeats (Zhao et al, 2000). It is the separating sequence that contributes to the binding specificity of different heterodimer molecules. It also provides the geometry that is needed for two subunits to interact specifically. The DNA nucleotide sequence and its specific packaging into chromatin determine the interaction of hormone regulatory sites with their correspondent SHRs (Beato and Klug, 2000). A number of studied promoters contain HREs in their DNA sequences. One of the most studied ones are TAT promoter (Rahman et al, 1995), 11βHSD promoter, Gluco/P_4, metallothionein promoter, genital HPV promoter, E2, PS_2 and MMTV promoter. All of the mentioned promoters are organized into positioned nucleosomes.

Section 5: MMTV promoter

The promoter of the mouse mammary tumor virus (MMTV) provirus is probably the best characterized example of transcriptional control by steroid hormones in which the chromatin organization plays an important role (Richard-Foy and Hager, 1987).

1. Chromatin organization of the MMTV promoter

The mouse mammary tumour virus (MMTV) is a B-type retrovirus belonging to the genus beta-retroviruses. Discovered in 1936, it was found to induce adenocarcinomas and T-cell lymphomas in mice through a process called insertional mutagenesis, in which a part of the viral genome is inserted within or near an oncogene. There are two routes of transmission of the virus - exogenous MMTV is transmitted vertically via milk, whereas endogenous is inherited through the germ line. The genome of the MMTV is 8805bp in size and contains the retroviral structural genes and an additional gene *sag* coding for a superantigen. The MMTV long terminal repeat (LTR) is responsive to steroid hormones, thus it can infect a variety of cells but it is tumourigenic only in mammary epithelial cells. Some studies have shown that following integration of the exogenous MMTV, the mouse *int, wnt* and *fgf* genes are activated (Clausse et al., 1993; Durgam and Tekmal, 1994; Uyttendaele et al., 1996).

The MMTV promoter, located in the long terminal repeat (LTR) of the provirus, is induced by glucocorticoids and progestins. The hormone responsive region comprises five binding sites for the hormone receptors (Hormone Responsive Elements, HREs), a binding site for NF1, and two binding site for the Octamer transcription factor 1, Oct-1 (Truss and Beato, 1993). The MMTV promoter is organized into six positioned nucleosomes (Richard-Foy and Hager, 1987), named from A to F, with a nucleosome (nucleosome B) covering all the HREs, the binding site for NF1 and for Oct-1 (Figure 15).

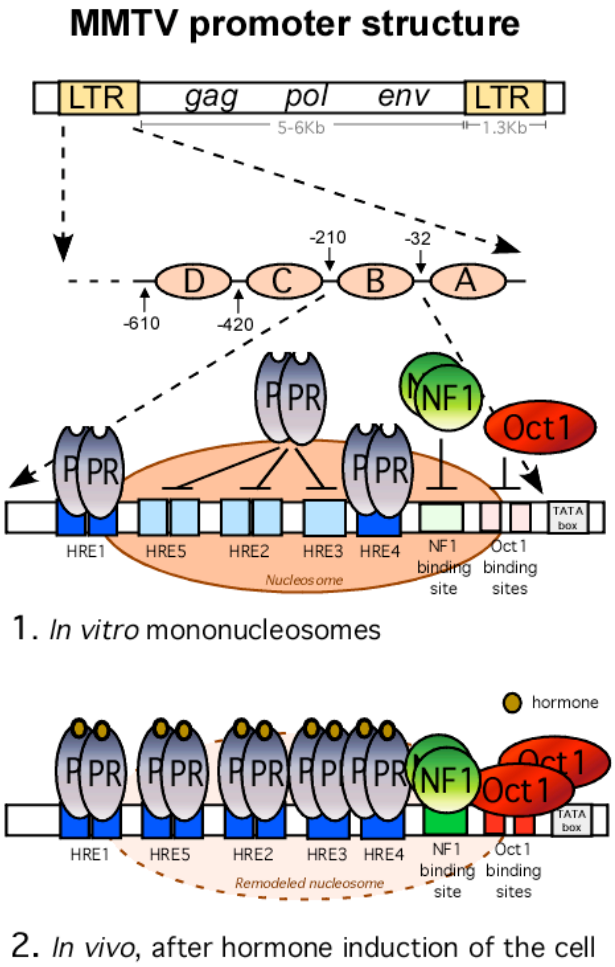


Figure 15. Chromatin organization of the LTR where the MMTV is located. *HRE*: Hormone Resposnive Element; *NF1*: Nuclear Factor-1*; Oct-1*: Octamer binding factor 1. (From Vicent et al, 2006)

In vivo setting of the nucleosomes in MMTV promoter regulatory region is similar to that obtained *in vitro* (Piña et al., 1990). Through *in vitro* nucleosome reconstitution experiments precise translational and rotational positioning of the MMTV promoter sequence on the surface of histone octamers was revealed. In the experiments for analysis of the chromatin structure of the non-induced MMTV promoter *in vivo* it was shown the nucleosome B covering the regulatory region extends from –210 to –32 in agreement with the reconstitution experiments *in vitro* (Piña et al., 1990). Glucocorticoid or progesterone treatment renders the nucleosome B region of MMTV hypersensitive to endonucleases, which is characteristic of remodelled, or “active” chromatin regions (Lee and Archer, 1994).

2. Hormonal induction of the MMTV promoter

2.1 Nuclear factor-1 (NF1)

Nuclear factor 1 (NF1) protein family is a group of sequence-specific transcription factors, required for the expression of many cellular and viral genes. Human NF1 was first isolated from HeLa cells, where it was found to function as a replication factor for adenovirus replication (Nagata et al, 1983). The family is derived from alternative splicing of transcripts of four genes – NF1-A, NF1-B, NF1-C and NF1-X, which are highly conserved from chickens to humans. NF1 proteins share a highly conserved N-terminal domain, which is a DNA binding domain. They can bind to DNA either as homodimers or heterodimers, recognizing a partially palindromic consensus DNA sequence TGGA/C(N_5)GCCA (de Vries et al, 1987). NF1 binding sites occur in a variety of tissue-specific and development-related genes, as well as in the regulatory regions of several viruses. The carboxy terminal is highly variable within the family and this implicates that NF1 proteins have diverse functions.

2.2 MMTV promoter activation process

The provirus integrated in the host cell chromatin is virtually silent in the absence of hormones but responds with rapid transcriptional activation to the addition of either glucocorticoids or progestins. The receptors for these hormones bind to a cluster of HREs and facilitate the interaction of other transcription factors with their target sites locate between the HREs and the TATA box. This results in a synergistic activation of transcription by the hormone receptors and Nuclear Factor 1 (NF1) (Di Croce et al., 1999) as well as the octamer transcription factor, Oct1/OTF1 (Bruggemeier et al., 1991) (for a review see (Beato et al., 1995)). How these factors synergize is a question that has attracted considerable attention, but the mechanism is not simply cooperative binding of the various proteins to the MMTV promoter (Bruggemeier et al., 1990).

When introduced in *Saccharomyces cerevisiae* engineered to express GR or PR and NF1, the MMTV promoter is silent in the absence of hormone, is organized into positioned nucleosomes, responds poorly to NF1 or to a NF1-VP16 fusion, but can be induced by hormone treatment (Chavez et al., 1995). Deletion of the HREs disturbs

nucleosome positioning and makes the promoter responsive to NF1-VP16, as if the HREs would repress access to the NFI site by positioning a nucleosome (Candau et al., 1996). Hormone induction was believed to cause a displacement of the nucleosome over the promoter, thus allowing free access of NF1 to its binding site and transcriptional activation (Richard-Foy and Hager, 1987). Moreover, in a cell-free system synergistic binding of receptors and NF1 to the MMTV promoters depends on its previous assembly into minichromosomes with positioned nucleosomes and on preincubation in the presence of ATP (Di Croce et al., 1999). The nature of the hormone induced nucleosomal change that permits simultaneous binding of transcription factors to the MMTV promoter remains obscure. Depletion of histone H1 from the MMTV promoter could be a possibility (Bresnick et al., 1992) and a role for histone H1 phosphorylation in modulating MMTV activation has been postulated (Lee and Archer, 1998). However, the MMTV promoter is regulated in budding yeast which lacks linker histones (Chavez et al., 1995) and also in minicromosomes assembled in the absence of histone H1 (Di Croce et al., 1999). Alternative possibilities are recruitment by the receptors of chromatin remodelling activities, either ATP-dependent, such as the SWI/SNF complex (Cote et al., 1994), or ATP-independent, such histone acetyl transferases (HATs) (Lee et al., 1993). This later possibility has received considerable attention following the discovery that several steroid receptors coactivators exhibit HAT activity, including members of the SRC-1 family, P/CAF and p300/CBP (Brown et al., 2000).

The role of histone acetylation on hormone induction of the MMTV promoter is not yet clear. High doses of histone deacetylases inhibitors, butyrate or TSA, lead to intense hiperacetylation of core histones and inhibit hormone induction (Bartsch et al., 1996; Bresnick et al., 1990) without altering nucleosome positioning (Bresnick et al., 1991). However, low dosis of the inhibitors activate the MMTV promoter both in the absence and in the presence of hormone (Bartsch et al., 1996). The extend of the hormonal induction is not affected by derepressing concentrations of the inhibitors of histone deacetylases (Bartsch et al., 1996). Similar results are obtained with these inhibitors in the cell-free transcription assay with MMTV minichromosomes, suggesting that inhibitor-sensitive acetylation is not involved in mediating the synergism between receptors and NF1.

There is evidence for a requirement of the SWI/SNF complex in glucocorticoid gene regulation in yeast (Yoshinaga et al., 1992) and in animal cells (Muchardt and Yaniv, 1993), and hSWI/SNF seems to be required for chromatin remodelling initiated by glucocorticoids (Fryer and Archer, 1998). The situation appears to be different for progestins, which act through the same HREs of the MMTV promoter (Fryer and Archer, 1998). We have reproduced the synergism between SHR and NF1 in a cell-free system driven by MMTV minichromosomes assembled in extracts from *Drosophila* embryos (Di Croce et al., 1999). These minichromosomes exhibit nucleosomes positioned in the same way as in cultured cells and the NF1 site in the promoter is not accessible (Venditti et al., 1998). Addition of hormone activated GR or PR leads to an ATP-dependent remodeling of these minichromosomes, as demonstrated by topoisomer analysis (Di Croce et al., 1999), and to simultaneous binding of receptors and NF1. This activity can be partly reproduced with purified MMTV minichromosomes with recombinant dISWI, but not with purified dCHRAC or ySWI/SNF complex (Di Croce et al., 1999). Moreover, incubation of the minichromosomes in the *Drosophila* extract in the presence of receptors and NF1 leads to recruitment from dISWI, and dNURF30, two subunits of the NURF complex (Gdula et al., 1998). Thus, in the *Drosophila* embryo extract, PR recruits NURF to the MMTV promoter chromatin, and NURF remodels the minichromosomes in an ATP-dependent fashion.

However, in mammalian cells with a single MMTV promoter integrated in chromatin we have recently found the hormone induced recruitment to the promoter of both Brg1 and ISWI containing complexes (Vicent et al., 2004). Moreover in experiments with MMTV mononucleosomes assembled *in vitro*, PR is able to recruit purified yeast Swi/Snf to the promoter and it catalyzes the ATP-dependent displacement of H2A/H2B dimmers (Vicent et al., 2004). Though we can not formally exclude a replacement of the H2A/H2B dimers by dimers containing histone variants, our results strongly suggest that the activated MMTV nucleosome containing the bound SHRs and NF1 is a an H3/H4 tetramer particle.

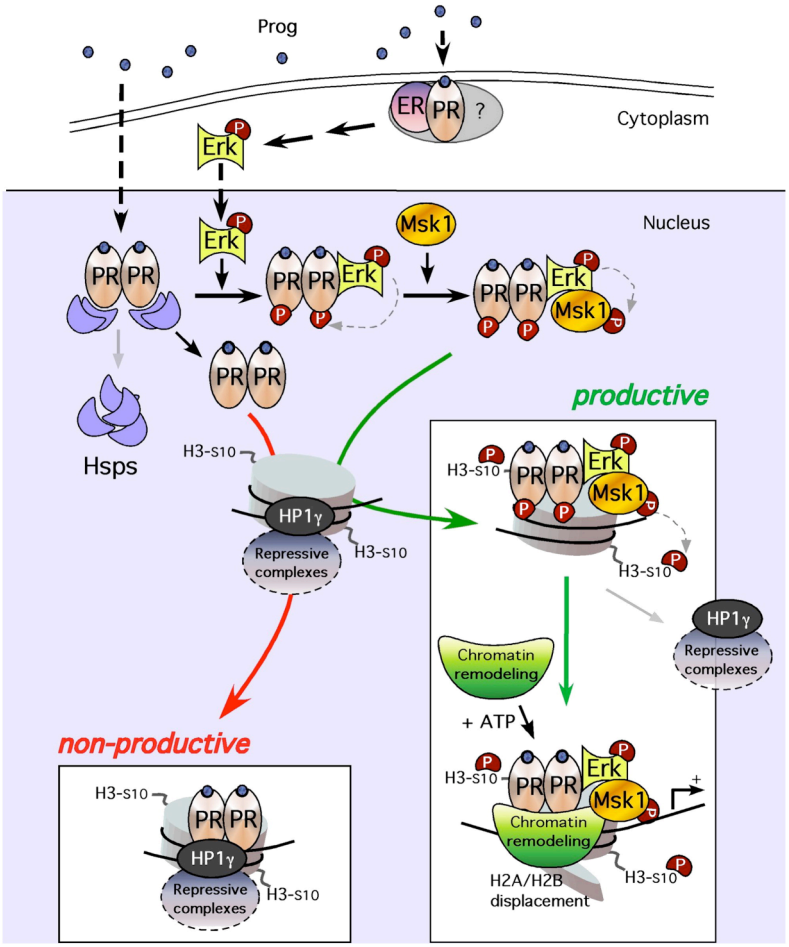


Figure16. Model of the crosstalk between Erk and Msk kinases and PR molecules upon hormone induction. (from Vicent et al, 2006)

We have investigated the function of the activation of Erk by progesterone, which is occurs via an interaction between PR and ERα (Migliaccio et al., 1998) (Figure 16). Recently we observed that progestin activation of the MMTV promoter can be blocked by inhibiting Erk activation either with antiestrogens, with small molecular kinase inhibitors (PD 98059), with dominant negative kinase mutant, or with RNAi against Erk (Vicent et al., 2006). As PR is phosphorylated by Erk at ser293 in response to progesterone, it is possible that the transcriptional inhibition of MMTV induction was due to the lack of PR phosphrylation. However, a similar inhibition of MMTV induction was observed when the activation of Msk1 was prevented by small molecular inhibitors (H89), a dominant negative mutant of Msk, or down-regulation of Msk1 with an specific RNAi (Vicent et al., 2006). Unexpectely we detect a ternary complex of phosphorylated PR and activated Erk and Msk1 already 5 minutes after hormone treatment and this complex is selectively recruited to the MMTV promoter nucleosome containing the HREs (Vicent et al., 2006). Concomitant with the recruitment of the ternary complex, histone H3 becomes phosphorylated at serine 10 and acetylated at lysine 14, only over

the nucleosome containing the HREs and not in adjacent nucleosomes (Vicent et al., 2006). The results with the MMTV promoter underline the significance of incorporating a precise knowledge of the structure of the native DNA sequences in chromatin into any attempt to understand the interactions of various transcription factors with their target sequences and with each other responsible for the physiological regulation of individual genes.

Materials and Methods

General Molecular Biology Protocols

Small-scale and Large-scale preparation of plasmid DNA

To isolate plasmid DNA from bacterial culture the large number of methods, including treatment of nonionic and ionic detergents, organic solvents, alkali, or heat can be employed. Both for mini- and maxi-scale DNA preparation a NucleoBond kit (BD Biosciences) were used. It employs a modified alkaline/SDS lysis procedure, potassium acetate (KAc) for precipitate formation and NucleoBond column which contains silica-based anion exchange resin. The procedure is as follows. First the bacterial culture is spinned down in a 4 °C centrifuge (4500 rpm for 4 min. for mini-prep, 6000g for 15 min. for maxi.prep). Then, the cells are resuspended in Buffer S1 (50 mM Tris-HCl, 10 mM EDTA, 100 μg/mL RNase A) and lysed when Buffer S2 is added (200 mM NaOH, 1% SDS). The lysate is neutralized upon addition of Buffer S3 (2.8 M Kac, pH 5.1) and incubation for 20 min. on ice. In the case of mini-prep after this step the phenol/chloroform extraction and ethanol precipitation was performed as already described before. For maxi-prep the protocol from NucleoBond was followed, using the AX 500 column to bind, wash and elute

the plasmid DNA. Precipitation was performed with isopropanol and subsequent centrifugation at 16000g for 30 min. The DNA pellet was washed with 70% cold ethanol, air-dried and resuspended in 400 μL of ddH_2O.

Transformation of DH5α competent cells

Bacterial transformation is the process by which bacterial cells take up naked DNA molecules. If the foreign DNA has an origin of replication recognized by the host cell DNA polymerases, the bacteria will replicate the foreign DNA along with their own DNA. Bacteria can be induced to uptake certain DNA molecules and the antibiotic selection techniques coupled with transformation are the way to select the correct transformants. The *E.Coli* DH5α competent cells (100 μL) were transformed with 7 μL of the ligation mix. The transformation method involved incubation on ice for 30 seconds, the heat shock step at 42°C for 40 sec. and a 5 min. incubation on ice. Subsequently 900 μL of Luria - Bertani (LB) medium was added and the transformed cells were grown at 37 °C for 45 min. in a shaker. After that the cells were spinned down for 4 min. at 8000 rpm and 900 μL of the supernatant was discarded. One hundred microlitres of resuspended cells were plated on LB plates with Ampicillin (Cf=100 μg/mL) (for selection of positive clones) and incubated at 37 °C overnight. A single colony was picked the next day and incubated overnight with agitation in 4 mL of LB medium with Ampicillin.

Phenol/chloroform extraction and ethanol precipitation

This method is a way to remove proteins from DNA solutions using equilibrated phenol:chloroform:isoamyl (25:24:1) (v/v/v) mixture. This method was used after each PCR reaction, restriction enzyme digestion experiment, alkaline phosphatase treatment or gel purification. Before phenol/chloroform extraction proteinase K to the final concentration of 0.075 μg/μL was added and incubated 1 hour at 45 ˚C. The volume in each sample was increased to 200 μL with H_2O and 1 volume of equilibrated phenol:chloroform:isoamyl alcohol mix was added. Following 15 seconds vortexing and centrifugation for 5 min. 13000 rpm the aqueous phase was transferred into a fresh tube containing 20 μg of glycogen. 3M NaAc was added to the final concentration of 300mM and 3 volumes of 100% ethanol. After short vortexing DNA precipitation was done at -80°C for 30 min.

Following centrifugation for 20 min. at 4 °C the pellet was washed with 900 µl of 70% ethanol and centrifuged at RT for 5 min. at max. speed. The pellet was shortly air-dried and resuspended in 20 µL of ddH_2O or TE buffer.

Sequencing reactions

The reaction was done using a ready reaction mix, which contained the dye terminators, deoxynucleoside triphosphates, AmpliTaq DNA Polymerase, FS, magnesium chloride, and buffer. The reaction components were as follows:

premix	4 µL
DMSO 50X	0.7 µL
dsDNA	300 ng-500 ng
RV3primer (Ci=1 pmol/µL)	3.2 µL
ddH_2O	until 10 µL
Total volume	10 µL

The following programme was used in GeneAmp® PCR System 9700:

94°C	3 min.
96°C	10 sec.
50°C	5 sec.
60°C	4 min.
4°C	∞

The extension products were purified by ethanol/sodium acetate precipitation. Eighty µL of the mix of 62.5 µL of ethanol 95% + 3 µL 3M NaAc (pH 4.6) + 24.5 µL of ddH_2O was added to each sample and incubated for 15 min. at RT. Following the centrifugation for 20 min. at 14000 rpm the pellet was washed with 250 µL of 70% ethanol, centrifuge for 5 min. at 14000 rpm. The pellet was air-dried and sent for sequencing.

DNA purification from agarose gel

The quick and efficient way to separate and extract DNA fragments is a silica-gel-based purification. There are many commercially available purification kits. In general the

procedure involves first solubilization of agarose gel slice (with DNA fragment in it). The DNA is bound to the silica-gel material in the presence of high concentration of chaotropic salts. A quick wash step removes impurities, and DNA is then eluted in low-salt buffer. GFX PCR DNA and Gel Band Purification Kit (GE Healthcare) was used to perform purifications. The procedure was followed according to manufacturer's instructions. Shortly, the digested mutated fragments and the pGAW222 plasmid were run on 1% agarose gel and the slices of agarose containing the DNA to be purified were excised. 10 µL of capture buffer was added for each 10 mg of gel slice, and the samples were incubated for around 15 min. at 60 °C. The samples was transferred to the GFX columns (MiscoSpinTM columns pre-packed with a glass fibre matrix), and centrifuged for 30 sec. at 13000 rpm. The columns were washed with 500 µL of wash buffer (Tris-EDTA buffer in 80% ethanol) and subsequently 15 µL of ddH_2O was added and the columns were centrifuged at full speed to recover the purified DNA.

Phosphatase treatment

Before ligation reaction calf intestinal alkaline phosphatase (CIAP) was used to remove 5' phosphate groups to suppress self-ligation and recircularization of the pGAW222 plasmid DNA. The reaction was performed in CIAP buffer using 2 µL (10U/µL) of CIAP (InvitrogenTM) for 1.5 hours at 37 °C. The reaction was stopped with 0.5 M EDTA and dephosphorylated DNA was then purified and extracted with phenol/chloroform method described earlier.

Quantification of the protein concentration (Bradford)

The Bradford solution (Bio-Rad Laboratories GmbH) was first diluted and the standard curve was prepared with bovine serum albumin (BSA). The octamer solution was diluted 1:50 in a separate tube and 900 µL of the prepared Bradford reagent was added to all the samples. The standard curve samples and the octamer were incubated at RT for 5 min. and the absorbance was measured at 595 nm. The concentration of the octamer in the solution was calculated based on the obtained graph. The concentration of the octamer obtained in two reconstitutions was on average 0.76 µg/µL.

End-labelling of DNA fragments (at 5' end) with ^{32}P

In order to analyse the 3' end of nucleosome B DNA the 5' end of the DNA was labelled. This is because in the DNase I digestion experiments the 3' end digestion products of nucleosome B DNA (where the NF1 site is located) could be observed at a higher resolution (running slower or "higher") than the 5' end digestion products, located at the lower parts of the gel (running faster), where the resolution was not as high.

50 μg of wt and HRE2$^-$/3$^-$ plasmid DNA (pGAW222 and pGAW222 ΔHRE2/3) were digested with EcoRI restriction enzyme for 2 hours at 37°C to create 5' overhang. Large fragment of E.coli DNA Polymerase I - Klenow fragment (2U/μL) (Roche diagnostics) was used to fill the recessed 5' terminus. The reaction contained a mix of dGTP, dCTP and dTTP, and $\alpha[^{32}P]$-ATP (GE Healthcare). The sample was incubated for 30 minutes at 30°C and then the reaction was stopped for 15 min. at 75°C. The second restriction enzyme BamH I was used to separate the 222 bp fragments from the plasmid. After that phenol/chloroform extraction and ethanol precipitation was performed and the pellet was resuspended in approximately 15 μL of ddH_2O.

5% polyacrylamide gel

Component	***Quantity***
H_2O	77.6 mL
10X TBE buffer	5 mL
30% acrylamide:Bisacrylamide	16.6 mL
10% APS	0.7 mL
Temed	70 μL

After setting up the gel in the support, the gel was pre-run for 30 min. at 90V in 0.5X TBE buffer. After addition of glycerol the samples were loaded and the gel was run at 170 V.

Purification of DNA fragments from agarose gel using "trap" method

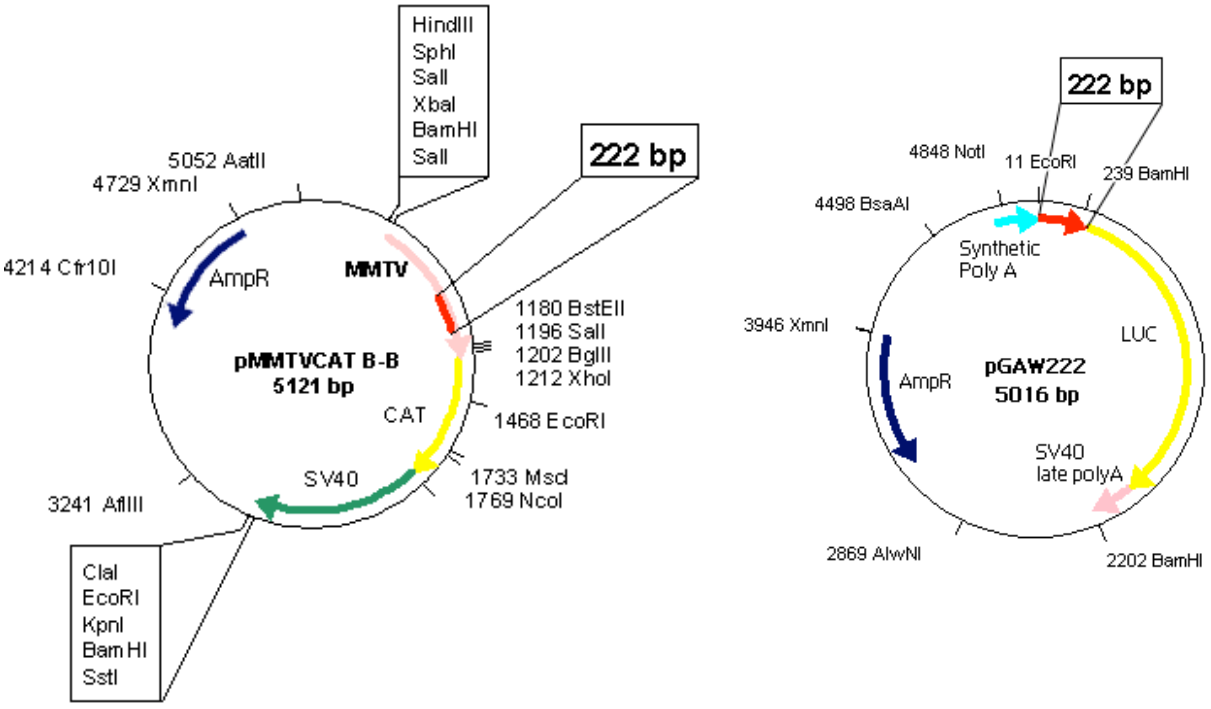
Following the digestion of 800 μg of plasmids pGAW222 ΔHRE 1, pGAW222 ΔHRE 2/3 and pGAW222 ΔNF1 or others with EcoR I and BamH I, the digestion products were run

on 1% agarose gel. To obtain the 232 bp fragments (222 bp of nucleosome B DNA + 10 bp remaining from restriction sites), the gel was stopped and the gel slice downstream the running 232 bp band was excised and the space was filled with 15% polyethylene glycol (PEG) in 1X TAE. The voltage was applied again and after several minutes the PEG solution containing the trapped DNA was collected into eppendorf tubes. The DNA was purified using phenol/chloroform extraction method. In order to acquire DNA of high purity the procedure was repeated. At least 20 μg of each mutated fragment was obtained.

Specific Molecular Biology Protocols

Generation of pGAW222 ΔHRE 1, pGAW222 ΔHRE 2/3 and pGAW222 ΔNF1.

The pMMTVCAT B-B plasmids containing MMTV promoter sequence from –640 to +126 with point mutations in HRE 1, HRE 2/3 and NF1 site were used to obtain 222bp long DNA sequences, which were used for subcloning into pGAW222 plasmid. The vector pGAW is a derivative plasmid from pGL3-basic (Promega).This plasmid carries a 222bp insert which corresponds to the wt DNA sequence of nucleosome B of MMTV promoter (kindly provided by Dr. Guntram Suske, IMT, Philipps Universität, Marburg, Germany). The advantage of using pGAW222 plasmid over pMMTV plasmids was the presence of EcoRI and BamHI sites present in the pGAW222 vector. These two restriction sites, flanking the 222bp fragments after insertion, enabled to easily obtain the 222bp mutated sequences for further experiments. The maps of pMMTVCAT B-B and pGAW222 plasmids are shown below:



The 222 bp DNA fragments of mutated HRE 1, HRE 2/3 and NF1 were amplified by polymerase chain reaction from pMMTVCAT B-B plasmids containing point mutations in:

1) HRE 1 – two point mutations in both hexameric core motifs (C→G, C→G)
2) HRE 2/3 – point mutation in HRE 2 (G→C) and HRE 3 (C→G) core consensus
3) NF1 – two point mutations (G→A, C→T)

50 ng of each plasmid was amplified using primers NucEco and NucBam (see Appendix for primer sequences) in the following reaction.

adequate plasmid template (C_i=10ng/μL)	5 μL
primer NucEco (10 μM)	10 μL
primer NucBam (10 μM)	10 μL
Taq polymerase (Invitrogen™) (5U/μL)	0.5 μL
10X PCR buffer	10 μL
dNTPs (20 mM)	1 μL
$MgCl_2$ (50 mM)	6 μL
ddH_2O	57.5 μL
Total	100 μL

The PCR was performed using GeneAmp® PCR System 9700 with the following programme.

94°C	5 min.
95°C	30 sec. (denaturing step)
56°C	30 sec. (annealing)
72°C	30 sec. (elongation)
72°C	7 min.

Expression and purification of histones in *E.Coli*

BL21(DE3)pLysS *Escherichia coli* strain was used for the expression of histone proteins H2A, H2B, H3 and H4. Cells grown in 2X TY medium containing 16 g/l Bacto Tryptone, 10g/l yeast extract, 5g/l NaCl, 100ug/l ampicillin and 25μg/l chloramphenicol. Expression

was induced at an A600 of 0.8 by addition of isopropyl-b,D-thiogalactopyranoside (IPTG) to a final concentration of 0.2mM and the culture was incubated for another 2 hours for all histone proteins. Cells were harvested by centrifugation for fifteen minutes at 4000rpm, room temperature. The total cell mass was ressuspended in 33ml of wash buffer containing 50mM Tris-HCl (pH 7.5), 100mM NaCl, 1mM benzamidine and 1mM Na-EDTA, and store at –20ºC.

The cell suspension was thawed at 37ºC in a water bath, adjusted to 40ml with wash buffer, and reduced the viscosity by shearing with an Ultraturrax homogeniser. The crude extract was spun for twenty minutes at 4ºC and the supernatant was discarded, because the pellets contained inclusion bodies of the corresponding histone protein. The pellet was then washed by the ressuspension and centrifugation three times in 35ml of wash buffer containing 1% (v/v) Triton X-100. The detergent was removed by repeating the washing two times with the wash buffer. The remaining pellet containing the inclusion bodies was stored at –80ºC.

Next day, the pellet was thaw with 333.5ul of dimethyl-sulfoxide (DMSO) for 30 minutes at room temperature. A 5.5 ml volume of a 7M guanidinium hydrochloride solution, containing 20mM Tris-HCl (pH 7.5), 10mM DTT, was added slowly and unfolding was allowed to proceed for one hour at room temperature under gentle mixing. Solubilized inclusion bodies were applied to a 26/60 Sephacryl S200 gel filtration column equilibrated with 1.5 columns volumes of S200-buffer containing 7M urea, 20mM sodium acetate (pH 5.2), 200mM NaCl and 5mM b-mercaptoethanol. The column was run at 22ºC at a flow rate of 1ml/min; the elution of protein was monitored by its absorbance at 276nm. Pooled protein fractions were dialyzed extensively at 4ºC against water containing 5mM b-mercaptoethanol, and stored in lyophilized form at -20ºC.

Octamer and tetramer reconstitution

Plasmids used

The pMMTVCAT B-B plasmids containing MMTV promoter sequence from –640 to +126 with point mutations in HRE 1, HRE 2/3 and NF1 site were used to obtain 222bp long DNA sequences, which were used for subcloning into pGAW222 plasmid. The vector pGAW is a

derivative plasmid from pGL3-basic (Promega).This plasmid carries a 222bp insert which corresponds to the wt DNA sequence of nucleosome B of MMTV promoter (kindly provided by Dr. Guntram Suske, IMT, Philipps Universität, Marburg, Germany). The advantage of using pGAW222 plasmid over pMMTV plasmids was the presence of EcoRI and BamHI sites present in the pGAW222 vector. These two restriction sites, flanking the 222bp fragments after insertion, enabled to easily obtain the 222bp mutated sequences for further experiments. The maps of pMMTVCAT B-B and pGAW222 plasmids are shown below.

The constructs were point mutated in the following sites: buscar número exacted dels residus) HRE 1 (two point mutations in both hexameric core motifs C→G, C→G); HRE 2/3 (point mutation in HRE 2 G→C and HRE 3 C→G); core consensus NF1 (two point mutations G→A, C→T). 50 ng of each plasmid was amplified using primers NucEco (5'-ATCTGGGAATTCCCAACCTTGCGGTTCCCAGGGTTTAAA-3') and NucBam (5'-ATACCGGGATTCAGTTTACTCAAAAAATCAGCACTCTTTT-3').
HRE -1 5'-ATCTGGGAATTCCCAACCTTGCGGTTCCCAGGGTTT-3'

Octamer/tetramer particles reconstitution by salt dyalisis

In order to avoid an interference of storage buffer (50% glycerol), prior to the reconstitution the octamer were dialysed against the refolding buffer (see below) in a Slide-A-Lyzer MINI dialysis unit (Pierce). The histone octamer solution (20 μg per each reconstitution sample) was placed in the dialysis vial and dialyzed against 2L of refolding buffer overnight at 4°C. The next day the protein concentration was determined using Bradford quantification assay (Bio-Rad Laboratories GmbH).

Refolding buffer:

NaCl (Cf 2M)	116.8 g
Tris- HCl 1M pH 7.5	10 mL
*βmercapto 14M (Cf 5M)	392 μL
EDTA 100mM	10 mL

* Added just before using

Mononucleosome templates and tetramer particles were generated with a 222 bp EcoRI-BamHI fragment containing the MMTV promoter sequence (from –221 to +1). The fragment was radiolabeled at the 5' ends with the Klenow fragment of DNA polymerase

and α[32P]-CTP. The mouse rDNA region from –232 to +16 relative to the transcription start site was synthesized by PCR with the plasmid pMR974 (Langst et al, 1999), as template and radiolabeled at the 5' end with polynucleotide kinase and γ-[32P]ATP. All labelled fragments were purified on a native 5% polyacrylamide gel. Mononucleosomes and tetramers were reconstituted by the salt dialysis method as described (Vicent et al., 2002), using recombinant X.laevis histones expressed in E.Coli (Luger K et al. 1999).

Four solutions of different salt concentrations were prepared for various steps of reconstitution. Solutions composition can be consulted below.

	Sol A (2M)	Sol B (0M)	Sol C (250mM)	Sol D (0M)
H2O	145 mL	734 mL	232 mL	490 mL
NaCl 5 M	100 mL	0 mL	12.5 mL	0 mL
Tris 1% (1 M) pH 8.0	2.5 mL	7.5 mL	2.5 mL	5 mL
EDTA (100 mM)	2.5 mL	7.5 mL	2.5 mL	0 mL
*βmercapto 14 M	39 μL	117 μL	39 μL	78 μL
TOTAL	250 mL	750 mL	250 mL	500 mL

* Added just before using

DNA (20 μg) together with the histone octamer (20 μg) and different components required for the reconstitution were placed in an eppendorf tube in the total volume of 100 μL (the components are listed in the order of addition).

H_2O	until 100 μL
NaCl 5 M	40 μL
Tris 1 M pH 8.0	1 μL
DTT 1 M	1 μL
EDTA 100 mM	1 μL
BSA (C 1.6 μg/μL)	2 μL
*βmercapto 14 M	1.66 μL
DNA (20 μg)	15 μL
Octamer (20 μg)	depends on conc.
TOTAL	100 μL

The mix was incubated for 30 min. at 4 °C. The incubated radioactive solution was then transferred into a dialysis vial and placed in a beaker with solution A (salt concentration 2 M). The dialysis lasted for two hours with steering. To decrease the salt concentration from 2 M to 250 mM using solution B (0 M) the pump was set overnight with flow rate 0.8 mL/min. The next day the vial was transferred into solution C (250 mM) for 2h at 4 °C with steering, and subsequently to solution D for another 2h. Finally the reconstituted nucleosome solution was collected in an eppendorf tube and stored overnight at 4 °C.

Purification of the reconstituted material

Purification of the reconstituted material was done by glycerol gradient ultracentrifugation using a linear gradient from 10-30% (v/v) glycerol in 50mM Tris-Hcl (pH 8.0), 5mM EDTA, 1mM DTT, 0.1mg/ml BSA. Solutions recipes:

10% glycerol			30% glycerol		
Quantity	Component	Final concentration	Quantity	Component	Final concentration
1 mL 0.1 mL 0.1 mL 10 μL 0.5 mL 8.25 mL	glycerol 100% EDTA 0.5M BSA 10 mg/mL DTT 1 M Tris 1 M pH 8.0 H_2O	 5 mM 0.1 mg/mL 1 mM 50 mM	3 mL 0.1 mL 0.1 mL 10 μL 0.5 mL 6.25 mL	glycerol 100% EDTA 0.5 M BSA 10 mg/mL DTT 1 M Tris 1 M pH 8.0 H_2O	 5 mM 0.1 mg/mL 1 mM 50 mM

Gradient master was used to create the gradient. The tube needs to be at least one hour at 4ºC before loading the sample. Centrifugation was performed in an SW60 rotor for 9h at 55.000rpm and 4ºC. Fractions of 100µl were manually collected from the bottom of the gradient and stored at 4ºC.

The reconstituted mononucleosomes samples were slowly added on the top of the tubes with the glycerol gradient. The tubes were centrifuged in SW60Ti Beckman rotor in the OptimaTM XL-100K Ultracentrifuge (Beckamn Coulter) at 55000 rpm for 9 hours at 4°C. The fractions were collected using manual sample collector (around 5 drops in each tube). The ones presumed to contain the mononucleosome populations and the free DNA were analyzed by 5% polyacrylamide gel (preparation as described before). Afterwards the gel was dried in the gel dryer model 583 (Bio-Rad) for at least 1 hour at 85 °C. The storage phosphor screen was placed on top of the dried gel for overnight exposure. The

radioactivity was read in phosphorimager and image processing was done in Image Gauge v 4.0 (Fujifilm). The fractions containing the mononucleosomes, as well as the ones with free DNA were used in further experiments.

Nucleosome Remodelling reaction: sliding

Reactions (10*µ*l) were done in 10mM HEPES (pH 7.9), 60mM KCl, 6mM MgCl2, 60uM EDTA, 2mM DTT, 13% glycerol containing 20nM of MMTV mononucleosomes and either 1.5 or 6 nM ySWI/SNF in the presence of 1mM ATP. Nucleosomes were incubated for 30 min at 30ºC followed by additional 30 min incubation with 250ng of poly-dIdC as competitor in order to remove SWI/SNF from the nucleosome. Nucleosomes remodelled were analyzed by electrophoresis in a 5% acrylamide/0,5X TBE gel at 4ºC, 5-6h at 170V. The gel was dried and autoradiographed.

Nucleosome Remodelling reaction: Sac I restriction enzyme accessibility

Reconstituted mononucleosomes were remodelled as previously described and digested at 37ºC, for 40 minutes, with 80U of SacI in a total volume of 160*µ*l. At the indicated time, 40ul of STOP buffer and 40ug of proteinase K were added and incubated at 50ºC for 1h. The samples were then extracted with phenol, phenol/chloroform and chloroform/isoamyl alcohol, and precipitated with three volumes of ethanol. After washing with 70% (v/v) ethanol and drying, the samples were analyzed on 5% acrylamide gel.

The Sac I restriction digestion enzyme was used as a determinant for SWI/SNF remodelling activity on mononucleosomes. Therefore first the SWI/SNF remodelling of reconstituted mononucleosomes was performed in a SWI/SNF buffer (see below), in the presence of ATP for 30 min. at 30 ºC. After that the poly dIdC was added as a competitor to withdraw the remodelling complex from the nucleosomes, and the reaction was incubated for 30 min at 30 ºC. Following that Sac I enzyme was added (80 U/reaction) to the final volume of 160 *µ*L. The reaction was stopped the EDTA and DNA was purified using phenol/chloroform extraction and ethanol precipitation method. The DNA fragments were analysed on 5 % polyacrylamide gel.

SWI/SNF buffer composition

HEPES 1M pH 8.0	200 µL
KCl 1M	1.2 mL
MgCl2 1M	120 µL
EDTA 100 mM	12 µL
DTT 1M	40 µL
Glycerol 50 %	5.2 mL
H_2O	3.228 mL

Chromatin assembly with pre-blastoderm *Drosophila Melanogaster* embryo extract

Mix:

95 µl Extraction buffer 50 (Exb 50)

1.2 µl DNA at 0.5mg/ml (supercoiled plasmid)

14 µl McNAP

30 µl DREX postblastodermic

total = 140µl

Incubate for 4-6h at 26ºC in a regulated bath

Exb 50: (50 refers to the mM concentration of KCl); prepare and store at -20ºC

838.5 µl H_2O

10 µl HEPES 1M pH 7.6

25 µl KCl 2M

1.5 µl MgCl2 1M

5 µl EGTA 0.1M pH8

20 µl β-Glycerophosphate 0.5 M

100 µl Glycerol

1ml

McNAP: (prepare each time)

16.5 µl H_2O

1.5 µl $MgCl_2$ 1M

1 µl 0.1M DTT

15 µl 100mM ATP

15 µl Creatine Phosphate 1M

1 µl Creatine Kinase 1ug/ul (it can be stored with Exb 50 with 50% glycerol)

total= 50 µl

Assembled minichromosomes are analyzed by Micrococcacl nuclease digestion. Reconstitutions can be checked along time by a supercolinig assay or performing a MNase ladder.

***In vitro* transcription of Chromatin templates or naked DNA**

Mix: 20 µl reaction volume

Rxn mix:

220 µl H_2O

8 µl Tris-HCl pH 7.8

40 µl Calf Thymus DNA (200ng/µl)

2 µl $MgCl_2$ (1M)

12 µl DTT (100mM)

24 µl rNTP mix (25mM each)

100 µl PEG 8000 (25%)

10 µl transcription factors + C90 (e.g. 3 µl PR + 1 µl NF1 + 6 µl C90)

C90 buffer:

734 µl H_2O

20 µl Tris-HCl

2 µl EDTA 0.5M

10 µl DTE 100mM

115 µl glycerol 87%

18 µl NaCl 5M

TF must be titrated in a wide range that can vary between dilutions until direct addition of several µl with the less concentrated preparation of factors. A proper synergism between factors is the objective of the titrations, as well, as a good signal for total transcription.

12 µl chromatin assembly reaction (25ng pMCBB) that has proceed during 4h at 26ºC for transcription of naked DNA and assembly reaction without DNA has been incubated in parallel with the chromatin reconstitution. At the moment before the addition to the transcription, the corresponding amount of DNA is added, without leaving enough time to

organize in chromatin. The control of naked DNA is used to know which are the chromatin specific effects.

- Incubate for 30min at 30ºC
- Add 12 µl HeLa nuclear extract (the extract is depleted of NF1 by incubation with a DNA affinity matrix)
- Incubation at 30ºC for 1h
- Add 100ul of 2xStop mix (before use add 12 µl 20mg/ml tRNA and 24 µl 10mg/ml protein K per 1ml of 2xStop mix)

    2xstop mix: 200mM NaCl, 20mM EDTA, 1%SDS

- Incubate for 15 min at 39ºC
- Extract once with 1x vol (165 µl) Phenol/Chloroform/Isoamyl Alcohol (25:24:1)
- Precipitate by adding 550 µl cold ethanol
- 30min at -80ºC and spin 25 min at 4ºC
- Discard supernatant and dry pellet for 5min in speed-vac
- Add 24 µl annealing mix. Leave 5 min at RT. Resuspend.

    Annealing mix:

    365 µl H_2O

    25µl KCl 2M

    10µl Tris-HCl 1M pH8.3

    2µl ^{32}P-labelled CAT probes

- 10 min at -80ºC
- 30 min at 39ºC
- Spin down briefly and put back at 39ºC
- Add 66 µl of Primer Extension mix

    PE mix:

    830 µl H_2O

    57 µl Tris-HCl 1M pH 8.3

    8 µl KCl 2M

    14 µl DTT 100mM

    24 µl dNTP mis (25mM each)

    12 µl M-MLV Rev.Transcriptase 200U/µl

- Incubate at 39ºC for 25 min
- Add 165ul of (1ml 2xStop mix+300 µl H_2O)

- Extract once with 260 µl pf Phenol/Chl/IA
- Ethanol precipitation (820 µl cold ethanol). 30min at -80ºC and spin 25 min at 4ºC
- Pipet off supernmatant and dry pellet for 2min speed-vac
- Add 8 µl formamide loading Dye: 100% Formamide + 0,025% Xylene cyanol FF + 0,025% Bromophenol blue
- Leave for 3min at RT to dissolve better
- Incubate for 2 min at 100ºC
- Resuspend (by pipeting up and down)
- Incubate 3min at 100ºC
- Load on a denaturing 10% PA gel:

    10% poluacrilamide

    8.3M Urea

    1xTBE

  Tip 30 ml mix, add 12,5 µl TEMED and 225 µl 10% APS

  Pour in a 26x18 cm gel, 0,25 mm thick

  Pre-run at 32 Watts for 15-30 min

  Run at 25-30 Watt in 1xTBE for aprox. 1h
- Transfer the gel to whatmann filter paper
- Dry under vacuum at 85ºC for 35 min
- Autoradiograph over night (up to 5 days)

In vitro ChIPs in mononucleosomes and minichromosomes

Nucleosome remodelling reactions (10 µl) were done in 10mM HEPES (pH 7.9), 60mM KCl, 6mM MgCl2, 60µM EDTA, 2mM DTT, 13% glycerol containing 20nM of MMTV nucleosomes and 6 or 12 nM of WT-SWI/SNF or Minimal SWI/SNF complex, in the presence of 1mM ATP. Nucleosomes were incubated for 30 minutes at 30 ºC followed by additional 30 minutes incubation with 250ng of poly-dIdC as competitor. Remodelled nucleosomes were cross-linked with 2.5% HCHO for 10 minutes at 37ºC, and the reaction was stopped by 0.1M glycine (pH 7.5) for 5 minutes at room temperature. 1ml of ChIP Immunoprecipitation (IP) buffer (0.01% SDS, 1.1% Triton X-100, 1.2mM EDTA, 16.7mM Tris-pH8.1, 167mM NaCl, 1mM PMSF, 1µg/ml aprotinin and 1µg/ml pepstatinA) was added and, after removal of an aliquot for input control (10% total volume),

mononucleosomes were subjected to immunoprecipitation with antibodies against histones H2A and H4 (Angelov et al., 2000). Before extraction with phenol/chloroform and ethanol precipitation, the samples were decross-linked at 65ºC. The PCRs were carried out with Taq DNA polymerase under standard conditions. The specific primers generate a 232 bp fragment of the nucleosome B of the MMTV promoter. PCR products were resolved on 1% agarose gels and stained with ethidium bromide.

Chromatin immunoprecipitation was performed in order to analyse NF1 and PR binding to both free DNA and mononucleosomes. The protocol used was a modified protocol from Upstate Biotechnologies. Four different rabbit polyclonal antibodies against NF1 were used – N20, B1, C1 (Santa Cruz) and self-made rabbit polyclonal antibody from Naoko Tanese's laboratory (here called TNF1). After addition of desired factors to the mononucleosomes the reactants were crosslinked with 2 μL of diluted (11%) HCHO and incubated for 10 min. at 37 °C. The reaction was stopped with 0.12 M glycine and incubated for 5 min. at RT. One mL of immunoprecipitation buffer containing protease inhibitors (PMSF 1 mM, aprotinin 1 μg/mL, pepstatin A 1 μg/mL) was added to each sample. Aliquots of 100 μL were taken from each sample as input material. The specific antibody was added and the samples were incubated overnight at 4 °C on a nutator. The next day 8 μL of Dynabeads® Protein A polystyrene magnetizable beads (Dynal® biotech) were added and incubated nutating for 2 hours at 4 °C. After that three washes of beads with Wash buffer I, II and III and two washes with 1X TE buffer were performed. Each washing step involved 5 min. nutating at 4 °C. Elution was performed with elution buffer (250 μL) for 30 min. at RT. The supernatant was transferred into a new eppendorf tube. The elution step was repeated and the supernatants pooled. Uncrosslinking was carried out overnight at 65 °C by addition of 20 μL of 5 M NaCl both to the samples and the inputs. The mix of:

EDTA 0.5 M	-	10 μL
Tris 1 M (pH 6.5)	-	20 μL
Proteinase K (10 μg/μL)	-	2 μL

was added to each sample (incubation for 1h at 45 °C). Phenol/chloroform extraction and ethanol precipitation was used to purify the DNA. For each ChIP experiment, PCR reactions were performed with a different number of cycles to determine the linear range of

the amplification. The specific primers generated a 222 bp fragment of the nucleosome B of the MMTV promoter sequence.

DNA	1 μL
PCR buffer 10x	2.5 μL
Mg2+ (50 mM)	1.5 μL
dNTPs (25 mM)	0.4 μL
NucEco up (10 μM)	0.5 μL
NucBam down (10 μM)	0.5 μL
Taq polymerase (5U/μL)	0.25 μL
ddH2O	18.35 μL
Total	25μL

The following programme was utilized:

94 °C 3min.
95 °C 1min.
59 °C 1min.
72 °C 1min.
72 °C 7min.
4 °C ∝

The PCR products were resolved on 1% agarose gels stained with ethidium bromide

Electrophoretic Mobility Shift Assay (EMSA)

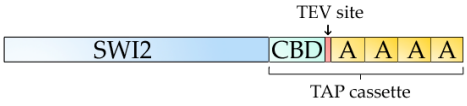
Electrophoretic mobility shift assay (EMSA), also called gel mobility shift assay, is a method for studying DNA-protein interactions in vitro. The method is based on the observation, that the DNA-protein complexes run slower on a gel than the free DNA when subjected to non-denaturing polyacrylamide or agarose gel electrophoresis.
Recombinant human PR, isoform B (PRB), and Nuclear factor 1 were expressed in baculovirus. Nucleosomes and naked DNA were incubated with different amounts of PR or NF1 in TGA-buffer (10 mM Tris-HCl [pH 8.0], 0.5 mM EDTA, 5% glycerol, 0.5 mM 2-

mercaptoethanol, 90 mM NaCl, 1 µg poly-dI-dC, 100 ng of calf thymus DNA, and 60 µg bovine serum albumin) in a 20 µl final volume. Binding reactions were incubated for 20 min at room temperature and analyzed by electrophoresis at 4°C in a 3.5% acrylamide/20% glycerol/0.5% agarose/0.3x TBE gel. For PR and NF1 binding to tetramer particles a 3.5% acrylamide gel was used.

Electrophoretic mobility shift assay was used to determine NF1-C isoform (from pig liver, purified using baculovirus expression system) and PRB [human recombinant isoform B expressed in baculovirus (Di Croce et al., 1999)] binding both to free DNA and to mononucleosomes in TGA buffer (10 mM Tris-HCl pH 8.0, 0.5 mM EDTA, 5% glycerol, 0.5 mM 2-mercaptoethanol, 90 mM NaCl). Some reactions with PR contained 250 ng of poly-dI-dC and 60 *µ*g of bovine serum albumin (BSA) in a final volume of 20 *µ*L. Binding reactions were incubated at 30 min. at RT and analysed by electrophoresis at 4 ºC in a 3.5 % acrylamide/20 % glycerol/0.5 % agarose/0.3X TBE gel, or alternatively in a 5 % acrylamide gel. The gel was then dried and radioactivity was detected as described before.

Purification of recombinant SWI/SNF from yeast cells using TAP purification procedure

The purification procedure was adopted from Smith et al., and the yeast strain CY944 Tandem Tap-tagged SWI2 was a gift from C.L. Peterson's laboratory. The strain contains the Tandem Purification Affinity (TAP) cassette, which encodes a four tandem repeats of protein A and a calmodulin binding domain (CBD) separated from the repeats by a TEV protease cleavage site. The cassette is located downstream the SWI2/SNF2 gene which codes for the ATPase subunit of yeast SWI/SNF. The cassette is expressed at the C terminus of the Swi2p/Snf2p. It consists of four tandem repeats of protein A, calmodulin binding domain (CBD) and a TEV site located in between.



The purification was performed as follows. A colony was picked from a YPD (Yeast extreact/Peptone/Dextrose) plate with the yeast strain incubated overnight at 30ºC, and grown in 100 mL of YPD medium overnight at 30 ºC at 180 rpm. The next day 10 ml of the starter culture was transferred into three 5 L Erlenmeyer flasks containing 2 L of YPD

medium, and grown at 30 ºC at 180 rpm until it reached O.D = 2.5-3 (~16h). After the incubation, the cultures were centrifuged at 4550g for 10 minutes at 4ºC. The pellet was transferred into a 20 mL syringe and emptied into liquid nitrogen. The frozen pellet was mixed with dry ice 1:1 weight ratio and ruptured in a home coffee mill for 3 minutes at maximum speed. The mix was transferred into a plastic tube and 1 volume of Elution buffer (E buffer) was added (at this step CO2 evaporation takes place). After subsequent centrifugation for 45 min. at 40.000 rpm at 4ºC the supernatant was collected in a 50 mL tube and pre-washed IgG beads (600 μL) were added to the lysate and incubated for 2h at 4ºC rotating in an Econocolumn (Bio-Rad). The flow through was discarded, and after washing with 15 mL of E buffer and 10 mL of TEV cleavage buffer, 4 mL of TEV cleavage buffer with previously added 300 units of TEV protease (30 μL) was added and incubated rocking o/n at 4 ºC (or at least 3h). The next day the flow through was collected, together with additional 1 mL of TEV cleavage buffer. To the 5 mL 20 μL of CaCl2 (Cf=2 mM) was added and transferred to a previously prepared Econocolumn with Calmodulin affinity resin and the column was rotated for 2.5 hours at 4 ºC. The unbound material was discarded and the resin was washed with 10 mL of calmodulin binding buffer (CBB). SWI/SNF was eluted from the calmodulin resin with 1.2 mL of calmodulin elution buffer (CEB). The protein solution was collected using the Vivaspin 2 mL tube and concentrated to 100 μL in a 4 ºC centrifuge (8000 rpm, around 30 min.). After that the protein solution was dialysed for 3 hours in a dialysis vial (Pierce) against 400 mL of E buffer containing 50 μM ZnCl2. SWI/SNF was again concentrated in 4 ºC centrifuge to around 30 μL using 500 μL Vivaspin tube. The solution was then aliquoted and rapidly frozen in liquid N2 and stored at -80ºC. Fractions collected at several steps and the obtained protein were analyzed by SDS Tris-HCl 5-20 % gradient gel.

Solutions used

E buffer (total 1L)

Hepes (pH 7.4) 1 M	20 mL
NaCl 5 M	70 mL
Glycerol 100%	100 mL
Tween 100%	1 mL
PMSF 0.1 M	3 mL
PIC	20 tablets

TEV cleavage buffer - E buffer + DTT (Cf=1 mM)

Calmodulin binding buffer (CBB) - E buffer + CaCl2 (Cf=2 mM)

Calmodulin Elution buffer (CEB) - E buffer + EGTA (Cf=10 mM)

SDS 5-20 % gradient gel and silver staining

The gradient gel for analysis of the purified yeast SWI/SNF subunits was prepared as follows:

5 %		20 %	
Quantity	Component	Quantity	Component
6.65 mL	30 % acrylamide	26.6 mL	30 % acrylamide
0.4 mL	10 % SDS	0.4 mL	10 % SDS
10 mL	Tris 1.5 M pH 8.8	10 mL	Tris 1.5 M pH 8.8
22.65 mL	H2O	2.75 mL	H2O
0.2 mL	APS 10 %	0.15 mL	APS 10 %
16 μL	TEMED	16 μL	TEMED

The gradient was prepared with a gradient maker with steering. The gel was poured using a pump at flow rate of 4 mL/min. After loading of the samples, the gel was run at 120V, and then silver stained to analyse the purified yeast SWI/SNF subunits. The procedure involved first fixing for 20 min. fixing solution (see below for solution contents), then washing with 50 % methanol in H_2O and another washing with H2O alone, both for 10 min. Following sensitisation for 1 minute, the gel was washed twice with water. The silver reaction was performed for 20 min. at 4 ºC. After the washing step (2x) developing was done in developing solution for around 40 min. Subsequently washing was performed for 1 min. in 5 % acetic acid.

Solutions used

Sensitization solution – 0.02 % sodium thiosulfate in H_2O

Silver reaction solution – 0.1 % silver nitrate in H_2O

Developing solution – 0.04 % formalin (35 % formaldehyde) 2 % sodium carbonate

Specific Protocols for the study of nucleosome positioning

Purification and Sequencing of Nucleosomal DNA

T47D-MTVL breast cancer cells carrying one stably integrated copy of the luciferase reporter gene driven by the MMTV promoter (Truss et al., 1995) were routinely grown in RPMI 1640 medium supplemented with 10% FBS, 2 mM L-glutamine, 100 U/ml penicillin, and 100 µg/ml streptomycin. For DNA preparation, cells were plated in RPMI medium without phenol red supplemented with 10% dextran-coated charcoal-treated FBS (DCC/FBS), and 48 hr later medium was replaced by fresh medium without serum. After 1 day in serum-free conditions, cells were treated for 5 minutes at 37ºC with 2ml of buffer A (150mM sucrose, 15mM tris pH 7.5, 60mM KCl, 15mM NaCl, 2mM $CaCl_2$) containing 0,5 mg/ml of lysolecithin. Micrococcal nuclease digestion was performed by adding 2ml of buffer A, supplemented with 720U of MNase (Worthington) for 3 minutes at 37ºC. The reaction was stopped by adding 40mM EDTA. Cells were scrapped and lysated with 1% SDS, 10mM EDTA and 50mM Tris pH 8.1 buffer, during 10 minutes on ice. Smaples were then treated with 0,1mg/ml RNAse A for 30 minutes at 37ºC and 0,1 mg/ml of proteinase K for 1 hour at 45ºC. DNA was purified by phenol/chloroform extractions and etanol precipitation. The digested DNA was run on 1% agarose gel, and band corresponding to mononucleosomes was gel-extracted.

Short linker sequences (5'-GATCCGCGGCCGCCCGAT-3' and 5'-ATCGGGCGGCCGCG-3') were ligated to both ends of the mononucleosomal DNA and fragments were clonned into the pBluescript SK- vector using the NotI restriction enzyme site. Sequenciation was carried out by using either T7promoter or T3promoter universal primers.

Microarray hybridization

Labeling was done using Bioprime genomic labeling kit (Invitrogen, cat.no. 18095-011 and 18095-012) following manufacturer's instructions. Briefly, 500 ng isolated nucleosomal DNA or total genomic DNA partially digested with MNase was labeled with either dUTP-Cy5 or dUTP-Cy3 fluorescent mononucleotides (Amersham GE PA55322), respectively. 2 mg of the obtained labeled sample were co-hybridized to the Agilent customized oligo microarray in the presence of 50 mg Human Cot-1 DNA and Agilent oligo aCGH hybridization kit (cat.no. 5188-5220) with includes blocking agents for unspecific hybridization following manufacturer's instructions. Hybridization was performed in a rotating oven at 65ºC for 40 hours. Arrays were washed and fluorescent images were acquire using G2565BA Microarray Scanner System (Agilent) with 100% laser power and 100% PMT settings. 16-bit TIFF images, one for each channel, were quantified using GenePix Pro 6.0 microarray analysis software (Axon Instruments Inc.). Microarray data was normalized applying a global loess algorithm using a smoothing factor of 0.2 (Yang *et al.* 2002) obtaining normalized $\log_2$[intensity Cy5]/[Intensity Cy3] mean values for each spot on the array.

Inferring nucleosome positions from tiling arrays

We use the derivative of the signal ($d(Signal)_n$) to filter out such fluctuations and therefore to assign nucleosome positions depending on local differences.

$d(Signal)_n = Signal_{n+1} - Signal_n$

Since the singnal pattern of a well-positioned nucleosome would be a peak of a given width (147nts), we search for how expressed this pattern is at every position and score its intensity. For every position *n* in the sequence we calculate the following score:

$Pnuc_n = - (d(Signal)_{n-70} . d(Signal)_{n+70})$

The *Pnuc* is higher when a local rise (positive derivative) in the signal values is followed by a local decrease (negative derivative) at a distance close to the nucleosomal length, thus corresponding to the assumed nucleosomal pattern. (We use a window of 140 nucleotides since the resolution of the chip does not allow a more detailed calculation). This sort of detrending allows us to capture the local fluctuations in the assumed nucleosome forming potential regardless of variable features of the sequence such as GC content, repeat sequences etc. Moreover, we assign a nucleosome formation score to all predicted

positions depending on these local differences, meaning steep changes in the values of symmetry of curvature would result into high scores and vice versa. A set of all possible nucleosome positions, without excluding overlapping ones, each one with a specific score is the final output of this process.

In the next step we apply dynamic programming (DP) algorithm to parse the query sequence in bound and unbound segments. Bound segments are of fixed 147 bp length, and unbound (spacer) segments are at least 20 bp long. The DP algorithm finds the partition Nuc* in successive non-overlaping nucleosome bound and unbound segments which maximizes the overall nucleosome-formation score *Pnuc*. This is explained in the following:

Let *S* be a DNA sequence $s_1,\ldots,s_l$ of length *l*, and let *P* be a distribution of scores along the sequence:

$$P(i)\in\square\ ,i = j,...,l - j(j = 25)$$

Let W_x be a 147bp sequence window starting at position *x* within *S*, that is, $W_x = s_x, s_{x+1},\ldots,$ s_{x+146}.

Let, the *SC* score of the window W_x be:

$$SC(W_x) = \sum_{i=0}^{146} SC_{(x+i)}$$

A nucleosome parsing of *S* is a series of 147bp windows, $W_{l(1)}, \cdots, W_{l(k)}$ $(k < l)$ such that for all

$i = 1,\ldots, k - 1, I(i + 1) > I(i) + 147 + 20$. That is, a nucleosome parsing is a series of 147bp windows separated by at least 20bp from each other.

Given a nucleosome parsing $N = W_{l(1)},\ldots,W_{l(k)}$ the *SC* score of *N* is:

$$P(N) = \sum_{i=1}^{k} SC(W_{I(i)})$$

The optimal nucleosome parsing of the sequence *S,* N^*, is the parsing that maximizes the *P* score.

That is, from all possible parsings, the optimal one is the one that has the maximum *P* score. Dynamic Programming can be used to find N^* very efficiently. Indeed, let $P(N_k^*)$ the score of the optimal nucleosome parse that can be constructed in the subsequence $s_1,\ldots,$ s_{k+146}. It is easy to see that

$$P(N_{k+1}^*) = \begin{cases} \max \begin{cases} P(N_k^*) \\ P(W_k) \end{cases} & \text{for k=1,...,167} \\ \max \begin{cases} P(N_k^*) \\ P(N_{k-167}^*) + P(W_{k+1}) \end{cases} & \text{for k=168,...,l-167} \end{cases}$$

with $P(N_0^*) = 0$. Then $P(N_{l-147}^*)$ is $P(N^*)$. The core of the optimal nucleosome parsing can thus be obtained recursively, by computing the score of the optimal parsing within each subsequence $s_1,\ldots,s_k$, $k = 1, l-146$. It is easy to see, therefore, that because the score of the optimal parsing at each position depends only on the score of one optimal parsing previously computed, the running time of the Dynamic Programming algorithm grows linearly with the length of the sequence. To recover the optimal nucleosome parsing, at each step *k* during the computation, an additional pointer variable *I* is updated. *I(k)* simply points to the terminal nuclesome in the optimal parsing within $s_1,\ldots, s_k$. That is, for $k = 1,\ldots, 167$, $I(k + 1) = I(k)$ if $P(N_k^*) > P(W_k)$, otherwise $I(k + 1) = 0$; For $k = 168,\ldots, l-147$, $I(k+1) = I(k)$ if $P(A_k^*) > P(N_{k-169}^*) + P(W_{k+1})$ otherwise $I(k+1) = k-167$;

We used the chaining program GenAmic (Guigo, 1998), as implemented within the program geneid (Parra et al., 2006) to obtain the optimal nucleosome parsing. The input to GenAmic is a file in General Feature Format (GFF, refs) with the coordinates of all 147bp windows in the query sequence *S*, and their *P* score. The final output is a GFF file representing the non-overlapping regions of a fixed length of 147nts with a minimum 20nts spacer between them, each one still assigned with its initial nucleosome-formation score. This allows us at the same time to have an overall view of the nucleosomal landscape for the sequence under examination while maintaining a means of filtering of the strength of each one of calls.

Objectives

1. Identification of special features on the nucleosomal DNA sequence that determine nucleosome positioning and studies on the potential use of these findings for *in silico* nucleosome positioning predictions.

2. Characterization of the early steps of hormone-dependent activation process of the MMTV promoter.

Results and Discussion

Chapter 1

"Studies on Sequence-Dependent Nucleosome Positioning and Predictability"

Part of the results presented on this chapter has been send for publication to *PloS Computational Biology* in the following format:

Nikolaou C., Zaurin R., Althammer S., Rué P., Beato M. and Guigó R. "A novel *ab initio* method for predicting nucleosome positioning gives insights into the chromatin organization of the human genome".

Results

1. Nucleosome positioning and the DNA sequence

As shown in the *Introduction* section, most of the work done on nucleosome positioning till 2004, focused on finding DNA motifs or nucleotide periodicity that defined nucleosomal DNA sequences. Despite all new findings in the field, a clear sequence-dependent positional code for nucleosomes still needed to be described. Moreover, our lab had discovered evidences that pointed towards the existence of features in the nucleotide sequence essential to decide the outcome of SWI/SNF-mediated nucleosome remodelling (Vicent et al, 2004). Early on 2006, we started to collaborate with C. Nikolaou and R. Guigó from the bioinformatic program of the CRG in order to further investigate the contribution of the nucleotide sequence to the positioning of nucleosomes, as well as to remodelling. The importance of curvature in the formation of nucleosomes has been pointed out by a number of studies before (Drew and Travers, 1985; Belikov et al, 1997; Richmond and Davey, 2003 and Edayathumangalam et al, 2004). Moreover, the nucleosomal structure is highly symmetrical, since the histone octamer comprises a symmetrical core tetramer $(H3/H4)_2$ flanked by two identical H2A/H2B heterodimers (Arents et al, 1991). It is also known that 147bp long nucleosomal double helix is wrapped around the octamer in a symmetrical manner with a central distorted part, which is associated with the octamer's dyad axis. Based on the nucleosome structural symmetry and the need for the DNA sequence spooling the histone octamer to be curved, it is plausible to assume that a tendency for curvature will be advantageous for a nucleosome-forming sequence. These features are more likely to be structural as the core histone octamer does not recognize specific nucleotide sequences via direct interaction with the DNA bases but rather their conformation in the three dimensional space. For this reason, the symmetric constraints are likely to be detected in the structure rather than the primary sequence itself.

The basic assumption of C. Nikolaou and R. Guigó's work is that sequences selected for nucleosome binding will tend to accumulate special features facilitating nucleosome information through evolution, being this signal a combination of curvature tendencies and symmetry. Thus, they postulated that a well-positioned nucleosome may present symmetry of curvature on both sides of the dyad axis and that the stronger the position, the higher the scores of these symmetry. On these bases, an algorithm hat calculates curvature of a given sequence and its intrinsic symmetry has been delineated. The so-called SymCurv method uses the BENDS algorithm (Goodsell and Dickerson, 1994) as extended by Munteaunu et al. (1998) to compute the curvature at each nucleotide in the nucleosome sequences and results in a SymCurv scores per each DNA sequence (Nikolaou et al., submitted).

The results presented in the following section are fruit of a tight collaboration with C.Nikolaou and R.Guigó. My contribution on this collaboration was in the design and perfoming of the experimental experiments and the discussion and experimental design of the rest of the work.

1.1 Symmetry of Curvature in yeast nucleosomal sequences

The concept of Symmetry of Curvature was first tested against the already published dataset of nucleosomal DNA from yeast, described by Segal et al (2006). The 199 *S. cerevisiae* mononucleosome DNA sequences of length between 142 and 152 bp were obtained through MNase digestion by Segal et al. (2006). The yeast dataset was pre-aligned at the center of the sequences, whose reverse complement was also included in the dataset. Then we computed the SymCurv value at each position along the complete length of each sequence, using a window of 50 nucleotides and computed the average values across the entire dataset. This produced a strong symmetrical pattern, which is reflected on the average SymCurv values (**Figure 17**) with a clear maximum at the center of the sequence alignment, coinciding with the dyad axis of the studied sequences. Thus these set of nucleosomal sequences presented symmetry of curvature confirming our previous hypothesis.

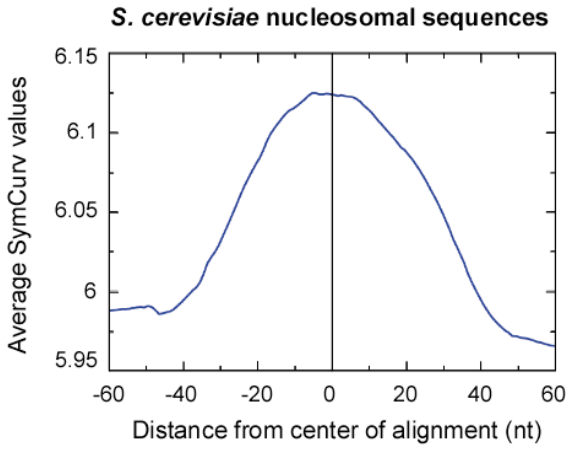


Figure 17. **Symmetry of Curvature in yeast nucleosomal sequences**. Observed curvature pattern obtained from 199 aligned yeast nucleosomes (Segal et al, 2006).

1.2 Preparation of a human nucleosomal dataset

We next wanted to test whether this feature was also present in the human nucleosomal sequences. Since there were no available data of human nucleosome sequences, I used a genome-wide assay to isolate DNA regions that were stably wrapped in nucleosomes. My experimental method maps nucleosomes on the human genome with great accuracy using *in vivo* Micrococcal nuclease (MNase) digestion followed by pyrosequencing (**Figure 18A**). Micrococcal nuclease was the enzyme that first biochemically defined the repetitive subunit structure of chromatin (Van Holde KE, 1988) and has increasingly been employed in mapping positioned nucleosomes due to the susceptibility of linker DNA to this nuclease (Simpson RT, 1998 and Simpson RT, 1999). MNase preferentially cuts chromatin in the linker region between nucleosome core particles, creating a ladder of oligonucleosomes in partial digestions. The principle of using MNase for studying nucleosome positioning is the following (**Figure 18B**). In regions where nucleosomes are randomly placed, nucleosomal DNA fragments obtained after MNase digestion mapped the entire region, as shown in Figure 18B (top panel). In contrast, when there exist constrains for specific nucleosome location, the MNase digestion results in an enrichment of some DNA fragments that correspond to the well-positioned nucleosomes protecting the DNA against nuclease attack while the nucleosome-free regions are absent (Figure 18B, bottom panel).

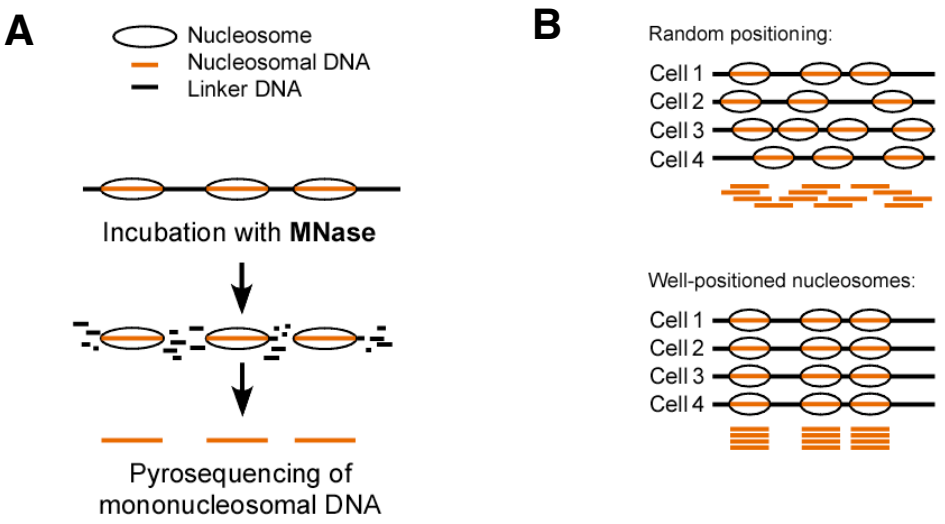


Figure 18. Preparation of nucleosomal sequences. **A**. Schematic representation of the chromatin digestion with MNase. **B**. Principle of the MNase digestion as a good method to identify well-positioned nucleosomes. Orange bar represents nucleosomal DNA.

Human nucleosome sequences were obtained as following. Briefly, T47D-MTVL breast cancer cells carrying one stably integrated copy of the luciferase reporter gene driven by the MMTV promoter were treated *in vivo* with limited amounts of MNase. Prior to the nuclease digestion, cells are permeabilized to allow access of the enzyme to the nucleus. After DNA extraction, samples were loaded in a 1% agarose gel. **Figure 19A** shows a well-defined micrococcal nuclease ladder on ethidium-stained agarose gel with individual nucleosome bands visible up to heptamer. Mononucleosome length DNA was extracted from the gel and short adaptor sequences were ligated to both ends. I used single-stranded gels to verify that there were no nicks on the DNA (data not shown). DNA sequences were cloned into the pBluescrit SK- vector using the NotI restriction enzyme site to create the first human nucleosomal DNA library for T47D breast cancer cell line. Sequencing was carried out using either T7 promoter or T3 promoter universal primers followed by pyrosequencing technique (**Figure 19B**, left column). I first mapped the obtained sequences on the human genome and retained those having at least BLAST hit with greater than 95% identity and a length range between 100 and 200 nucleotides (~85% of the total). From an initial collection of 250 human nucleosomal sequences, 213 were retained after imposing the above restrictions of size and quality (Table 4, collection A).

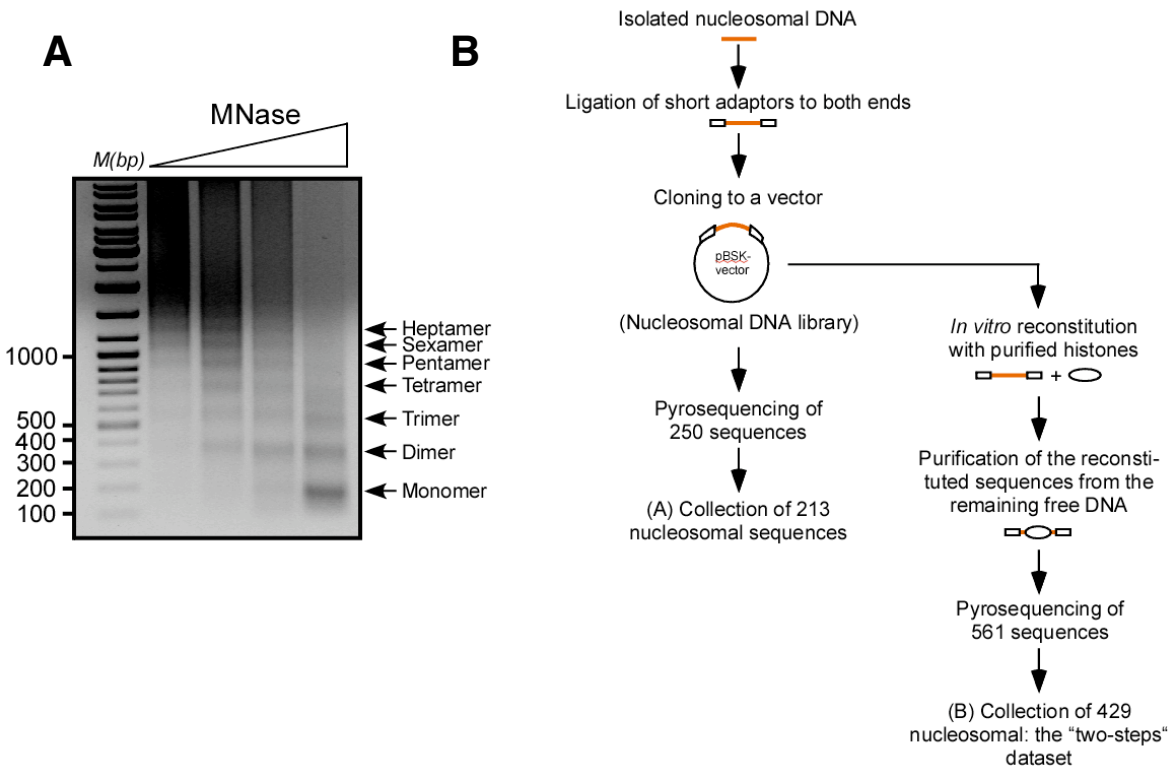


Figure 19. Preparation of nucleosomal sequences. A. *In vivo* permealized T47D cells were treated with different amounts of MNase for 5 minutes. After digestion cells samples were analyzed in a 5% acrylamide gel and stained by Ethidium Bromide. **B**. Scheme of the construction of the nucleosomal DNA libraries from T47D .

To obtained higher-quality nucleosomal sequences I added an additional *in vitro* reconstitution step (see scheme on Figure 19B, right column). Purified nucleosomal DNA from the library was reconstituted in nucleosomes with recombinant and purified *X. laevis* histones by salt-dialysis (see *Materials and Methods*). The *in vitro* reconstituted mononucleosomes were separated from the remaining free DNA by a 10%-30% glycerol gradient and then, sequenced as above. By this method, I obtained 561 sequences, from which 429 were retained after imposing the restriction of size and quality previously explained. This new library was called "two-steps" dataset or collection B (**Table 4**, collection B).

	In collection	
	A	B
Extracted or sequenced	250	561
BLAST Unique best hits Restrictions: identity >95% and 100-200nt length	213 (85,2%)	429 (76,5%)

Table 4. Nucleosomal bound sequences

1.3 Symmetry of Curvature in human nucleosomal sequences

We then applied the curvature calculation and symmetry to the new human nucleosomal dataset considering the central nucleotide as the dyad axis. As observed for the yeast nucleosome sequences, a periodical SymCurv pattern is present for the 213 human nucleosomal sequences of the collection A (**Figure 20A**) although lower scores are obtained compared to the yeast nucleosomes (Figure 17). We already expected that the performance of SymCurv in human sequences was lower than the one obtained from yeast due to the greater complexity of the former (presence of linker histones, chromatin remodelling complexes among other factors). However, differences could be explained as well by the insufficient quality of the human nucleosomal sequences. For this reason, I worked to produce a higher-quality dataset by introducing an *in vitro* mononucleosome reconstitution step before the final sequencing, as explained in Figure 19B. The 429 sequences of our "two-steps" *H. sapiens* nucleosomal DNA dataset were used in constructing artificial nucleosomal arrays (**Figure 20B**). The artificial arrays were done by juxtaposing the nucleosomal sequences with the use of random sequence spacers, whose length was selected each time so that a dyad-to-dyad distance of 300 nucleotides is maintained between consecutive sequences (see scheme on Figure 20B). That means that these artificial nucleosomal arrays were expected to have canonically spaced nucleosomes with a period of 300 nucleotides. In Figure 20B (red line) the SymCurv pattern for the alignment of all these artificial arrays is observed, which is in fact shown to be periodical with the expected periodicity, compared to the flat pattern obtained with random sequences (black line). The pattern obtained if one uses the reverse complementary sequence instead of the original one is practically identical (Figure 20 blue line).

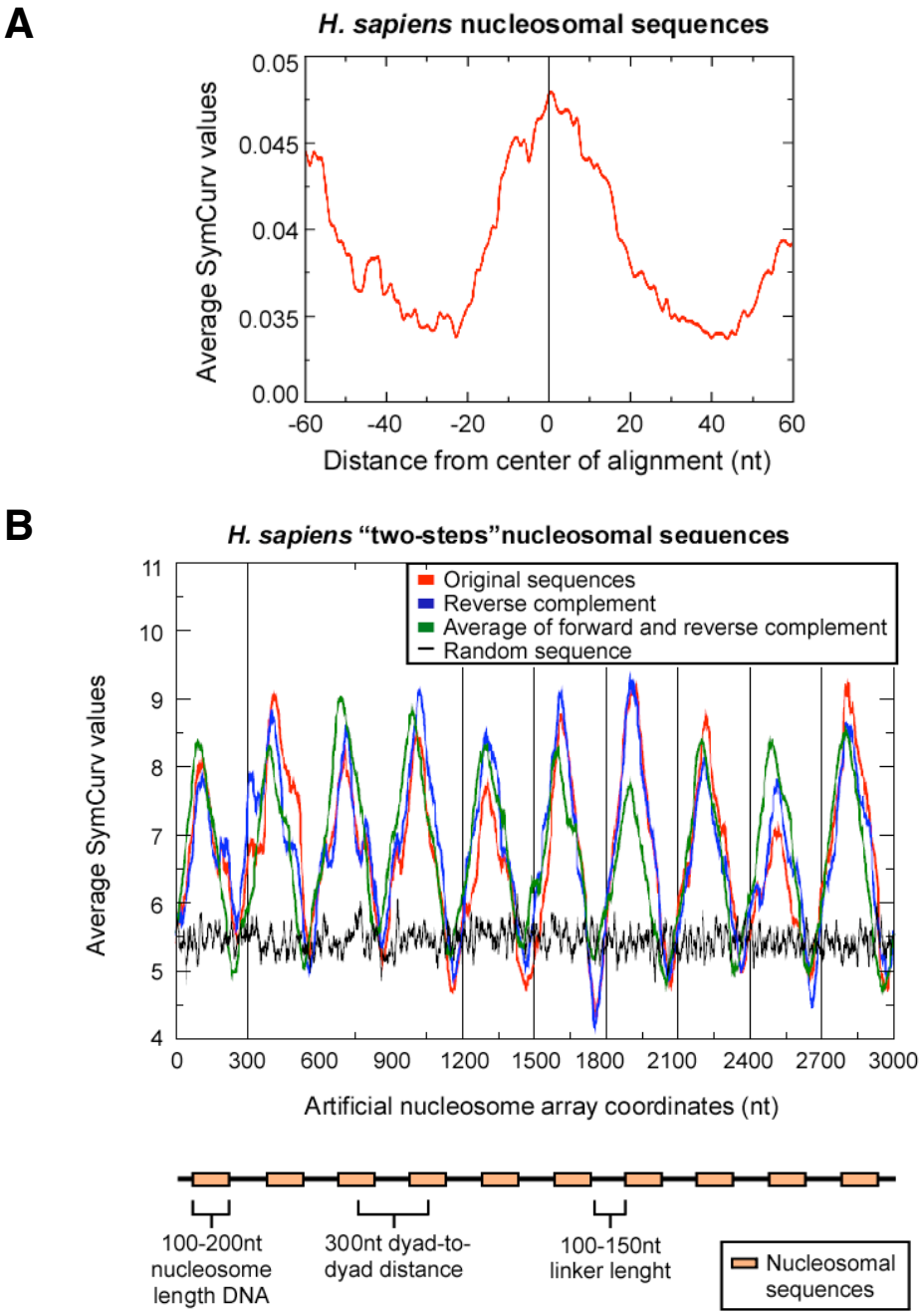


Figure 20. Symmetry of Curvature in human nucleosomal sequences. **A**. Collection A of nucleosomal DNA (213 sequences) were aligned by the central nucleotides and SymCurv values were calculated. **B**. 429 human sequences obtained through MNase digestion followed by *in vitro* histone reconstitution ("two-steps" collection) were juxtaposed in an artificial array interrupted by random sequence linkers. Each nucleosomal sequence was placed in the centre of a 300 nucleotide segment with the flanking remaining space being covered by random sequence spacers obtained through a random number generator using average nucleotide composition of the human genome. The resulting artificial sequence was scanned and its SymCurv values calculated.

1.4 Other features of the human nucleosomal sequences: specific sequence tendencies

After confirming the presence of symmetry tendencies on curvature features, we wanted to further characterize the nucleotide sequence of human nucleosomes. We first tested the percentage of repetitive sequences in the two collections of nucleosomal DNA. The repeated sequences content in both databases (in terms of percentage of total length) only slightly exceed the corresponding average for the human genome (46,3% against 45%, respectively). The same was observed for specific Alu elements, LINE-1 and 2 and LTR retrotransposons when comparing the average from both datasets against the expected value for the whole human genome (**Table 5**). Mapping of the sequences on the human genome suggested no significant preference for specific genomic regions, while the distribution of the positions throughout the human genome does not reveal any clustering tendencies among them (data not shown). This may be expected given the small number of examined sequences that does not permit sufficient coverage to reveal nucleosome positional preferences.

	In collection A+B	Expected
SINE (Alu elements)	10	10,6
LINE-1	17,7	16,9
LINE-2	2,1	3,2
LTR retrotransposons	6,8	8,3
Total bases % **(seqs belonging to repeats %)**	**46,3% (57,1%)**	**45%**

Table 5. Distribution of nucleosome bound sequences

We then tried to characterize in depth the human nucleosomal DNA focusing on the frequencies of occurrence of oligonucleotide groups. Frequencies of occurrence for di-, tri-, tetra- and penta-nucleotides were calculated and compared with the corresponding ratios obtained from the complete human genome. The results suggested considerable over-representation of CG-containing oligonucleotides. The CG dinucleotide is 3-fold overrepresented in nucleosomal sequences when compared to the complete genome. The 8 CG-containing tri-nucleotides are the most overrepresented triplets with ratios that go from 2-fold (CGC) to 6-fold (CGT). These results motivated a more careful look into the distribution of length-nucleosome dataset in the genome. The human nucleosomal sequences overlap with a CpG island with a frequency an order of

magnitude higher (10^{-4}) than expected by random distribution ($9x10^{-6}$), given the size of our dataset.

The average conservation values for all sequences across species were also calculated. As shown in **Figure 21**, most of the sequences have a conservation score over 0.2 being the average conservation value 0.181 (with a sd=0.147). It suggested that our nucleosomal sequences are not specially conserved across species. These results confirmed what was previously published by Johnson et al (2006) in *C. elegans* nucleosomal DNA studies were they observed no significant evolutionary constraint for a set of more than 160.000 nucleosomal sequences. This illustrates what we previously proposed of a nucleosome code based on structural features embedded on the DNA sequences rather than on nucleotide periodical patterns.

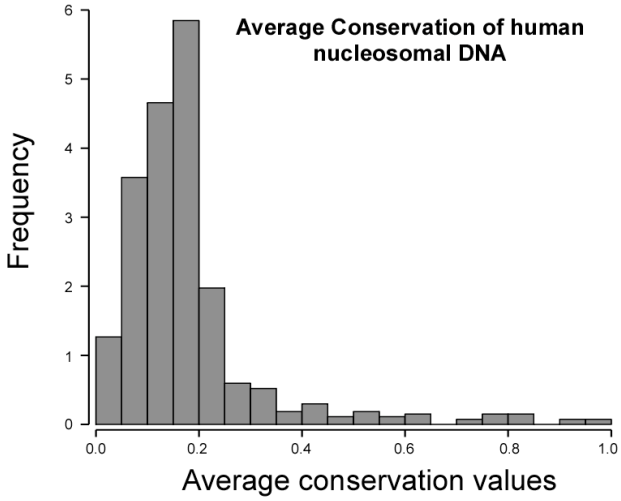


Figure 21. **Sequence conservation of human nucleosomal sequences.** 642 nucleosomal sequences were subjected to calculation of the sequence conservation scores. The histogram shows the frequencies in which every sequence conservation value appears.

From all these analysis I concluded that symmetry of curvature among the nucleosomal sequences exists, both in yeast and in human nucleosomes. SymCurv values were higher for the "two-steps" collection than for the collection A of nucleosomal sequences, suggesting that the more strongly positioned a nucleosome is, the higher symmetry of curvature degree presents. Relative enrichment of nucleosomal sequences on CpG islands is also an interesting result that would need further experiments to elucidate its biological function. However, studies of nucleosome positioning on a set of sequences has some limitations and I also wanted to study nucleosome prediction in longer genomic regions. For all these reasons, we decided to design a tiled nucleosome-resolution microarray approach for human genes in order to first validate the SymCurv method on large consecutive human genomic regions and study other features of the nucleosome positioning event.

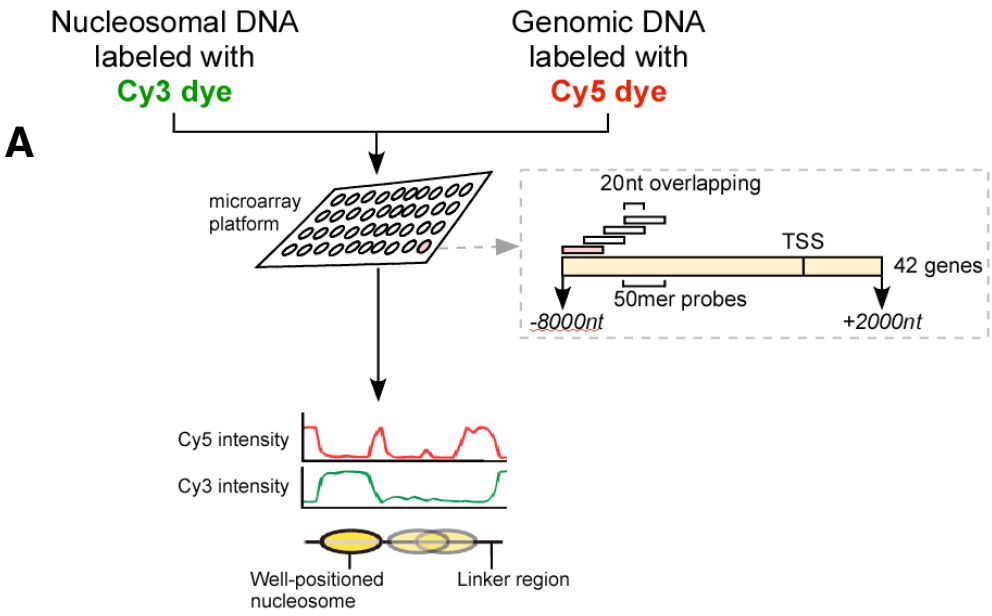
2. Studying the role of nucleosome disposition in promoters and coding regions of hormone-regulated genes

2.1 Description of the nucleosome-resolution tilling array platform

I used a DNA microarray method to identify nucleosomal and linker DNA sequences on the basis of susceptibility of linker DNA to Micrococcal nuclease (Rando et al, 2005, and **Figure 22A**). Our lab has designed the platform according to previous results we obtained from expression arrays performed in T47D breast cancer cells treated with progesterone or estradiol (Dr. Cecilia Ballaré, unpublished work). The most interesting cases (a total of 40 genes) were included on the nucleosome-resolution tilling array (**Figure 22B**) as well as two hormone-independent genes as negative controls (IFNB1 and HBE1 genes).

Nucleosomal DNA was isolated from T47D cells, as explained before, labeled with Cy3 fluorescent dye (green), and mixed with Cy5-labeled partially MNase-digested genomic DNA (red). MNase has sequence preference in digestion of protein-free DNA and therefore control digestions of naked nucleic acid are required when assessing the susceptibility of cutting sites in chromatin. In our case, the DNA control was total DNA isolated from T47D cells subjected to limited MNase digestion. This mixture was hybridized to the microarray platform made without removing repeated sequences and with oligonucleotide probes of around 50 nt, and similar melting profiles, spaced every 20 bp. In average each base pair was covered by three oligonucleotide probes (Figure 22A). The probes encompassed the immediate gene upstream region (from -8000 to +1nt) and part of the coding region (from +1 to +2000nts) of these 42 human genes. A graph of green:red ratio values for spots along the represented regions is expected to show nucleosomes as 120-160 base pairs broad peaks, or six to eight microarray spots surrounded by lower ratio values corresponding to linker regions (Figure 22A).

The results of all experiments done on this self-designed microarray platform and next detailed were submitted to Array Express (http://www.ebi.ac.uk/microarray-as/ae/), a public archive for microarray experiments, with the accession number: E-MEXP-1606 and name: “Nucleosome Mapping MNase”. Release date: the 1st of July 2009.



B

Induced by R5020:	Inhibited by R5020:	Induced by E2:
MMTV	IGFBP5	WISP2
EGFR	IGFBP4	MYB
STAT45A	NCOA3	ATF3
Dusp1	AR	CTSD
HSD11B2	GATA3	WNT5A
FOS	HDAC9	GDF15
CCNE2	PGR	EGR1
EGF	TNF	TFF1
CCND1		STAT5
SOS1	**Control genes:**	Snail1
SERPINA3	IFNB1	
BIRC3	HBE1	
MYC		
JUN		
TGFA		
RPS6KA5		
RPS6KA1		
PIK3CB		
PRL		
BCL2L1		
CYR61		
IL18		
CDKN1A		

Figure 22. **Nucleosome-resolution tilling microarray platform**. **A**. Scheme of the hybridazation process of the platform. Briefly, Cy3-labelled nucleosomal DNA is mixed with Cy5-labelled partially digested total genomic DNA and hibridized against a microarray platform. 40-mer spots overlap by 20 nt covering a total of 420Kb (10Kb per gene represented). Interpretation of the raw data obtained is shown on the Cy3/5 intensity plot. **B**. List of genes represented on the array classified by either activation/inhibition with progesterone (R5020) or activation/inhibition with estrogens (E2).

2.2 Characterization of nucleosome positioning in promoter and coding regions

Three biological and three technical replicas of the MNase treatment and mononucleosomal DNA isolation were mixed with partially-digested genomic DNA and hybridized against the platform. As observed in **Table 6**, Pearson correlation coefficients for all replicas are over 0.9 (except for the case of MNase_RepF vs MNase_RepD). This coefficient is used to determine the degree to which two or more variables vary together. Table 6 shows high coefficient numbers meaning that there is a high degree of parallelism between all replicas and thus, reliable data for the following analysis.

	MNAse_RepA	MNAse_RepB	MNAse_RepC	MNAse_RepD	MNAse_RepE	MNAse_RepF
MNAse_RepA	1,000	0,963	0,958	0,923	0,925	0,901
MNAse_RepB	0,963	1,000	0,955	0,920	0,965	0,927
MNAse_RepC	0,958	0,955	1,000	0,930	0,940	0,915
MNAse_RepD	0,923	0,920	0,930	1,000	0,938	0,881
MNAse_RepE	0,925	0,965	0,940	0,938	1,000	0,942
MNAse_RepF	0,901	0,927	0,915	0,881	0,942	1,000

Table 6. Pearson correlation coefficients of replicas

As an example, the raw signal, corresponding to the average fold-change green:red ratio obtained from all replicas, for the AR, SOS1 and PRL gene regions is plotted against gene coordinates in **Figure 23** (blue line). The blue line plot shows the ratio between Cy3 and Cy5 intensity dyes, that is, nucleosomal and genomic DNA, respectively. Experimental nucleosome calls were obtained after application of the derivative-dynamic programming approach on this raw data (see *Methods and Materials*). Peaks of about 120-160 base pair long surrounded by lower ratios values indicated positions of nucleosomes that we represented as brown boxes of 147-base pairs size. Different intensities of brown correspond to different experimental nucleosomes calls scores, related to stronger or weaker nucleosome positions (explained in more detail in section 2.3).

We have experimentally identified a total of 1653 nucleosomes by the tilling array methodology over 420 Kb of genomic human DNA, that is a 57,85% coverage of all genomic region represented. If these 1653 nucleosomes were periodically distributes, we would find a nucleosome every a hundred nucleotides in average. Inferring nucleosome positions from the microarray raw signal, allowed us to describe the first level of chromatin structure of the regions represented at the platform.

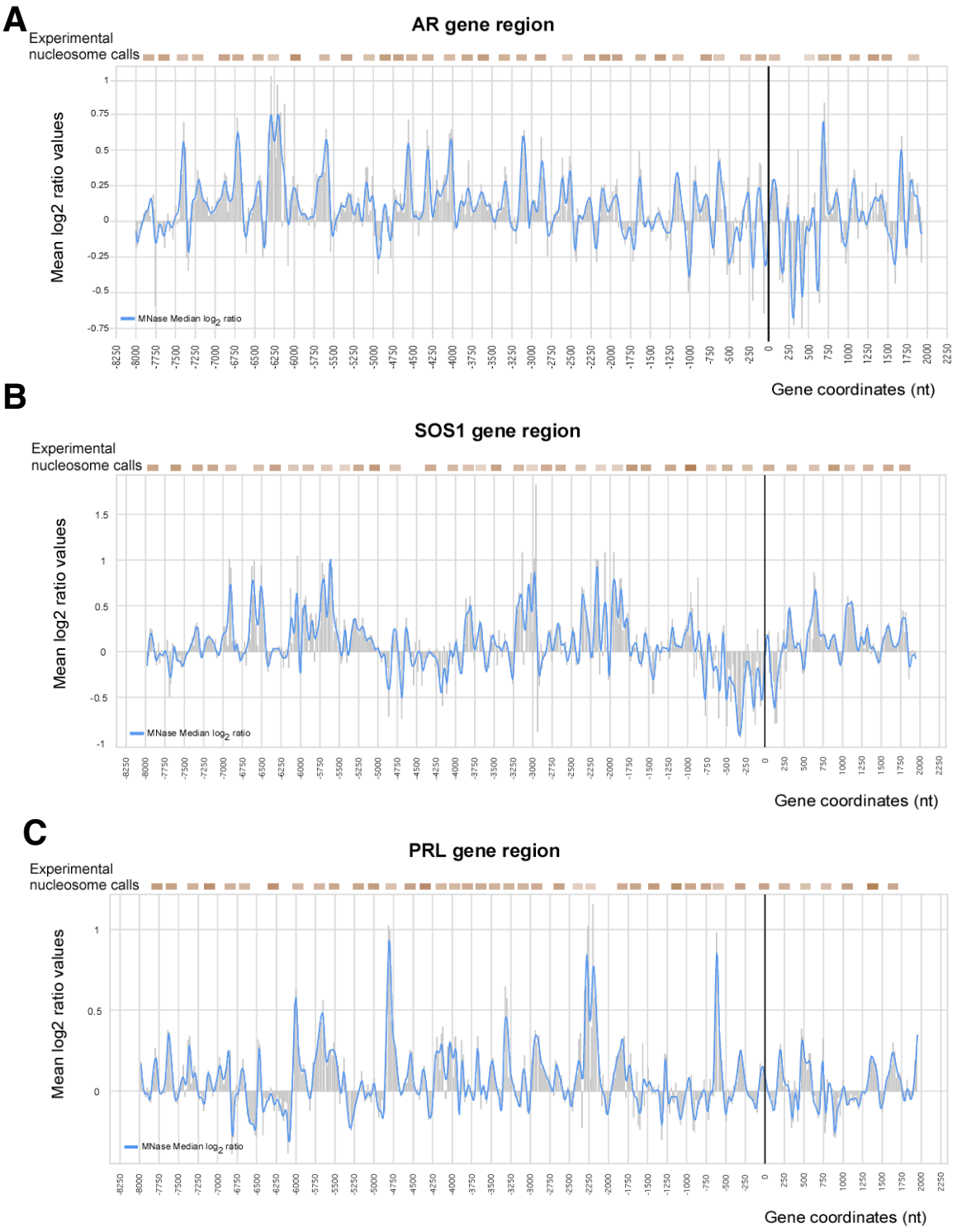


Figure 23. Nucleosome positioning in uninduced hormone-dependent genes. A. Nucleosome occupancy in the human DUSP1 gene region (dual specificity phosphatase 1, MKP1) (**A**), SOS1 gene region (Son of sevenless homolog 1 (Drosophila)) (**B**) and PRL gene region (Prolactin) (**C**). *Grey bars*: Median fold-change value from 6 replicate experiments, *Blue line*: Smoothed data. Brown bars represent nucleosomes and color intensity corresponds to relative scores (darker bars indicating higher scores).

2.2.1 Features of linker lengths and nucleosome distribution among coding and non-coding regions of hormone-dependent genes

I had also analyzed the general disposition of nucleosomes over 8Kb upstream and 2Kb downstream from the TSS on the hormone-dependent genes represented by measuring the linkers DNA length (**Figure 24A**). The linker size distribution showed a predominant peak of around 40-base pairs long linker DNA consistent with what has been seen by others (Barski et al, 2007, Schones et al, 2008), although a linker average size is of 98nt (with an sd = +65). Longer linkers are also observed, confirming the existence of extensive nucleosome-free regions, with special significance for two peaks: at 120 and 180 nt linker sizes. Regions with DNA linkers of around 180 nt length indicate nucleosome depletion or exclusion, while areas with linkers around 120 nt may have other meanings. Actually the nucleosome-free region (NFR) of around 140nt defined and well accepted in yeast (Yuan et al, 2005) here appears to be 40 nt longer.

I calculated the linker size average for two 2Kbp regions: flanking the transcription start site (**Figure 24B**). This allowed us to study the differences in nucleosome occupancy between the immediate promoter and coding regions in average for all genes of the platform. A tendency of longer linker lengths on promoter regions (2kb upstream of the TSS) compared to the coding regions (2kb downstream of the TSS) is observed. While in promoter regions the average linker is about 100 base pair long, in the coding region it is about 70 base pairs. The tendency of shorter linker lengths in coding regions is clear, although it exists a wide distribution of different linker DNA length in both cases.

I next wonder whether this pattern was observed in all the genes. For that the mean linker length of the two 2kb regions previously described (-2000…+1nt and +1…+2000nts) for each gene was plotted (**Figure 24C**). 85% of the genes showed longer linker size in average on promoter than in coding regions (Figure 24C, left part). The remaining 25% of the genes (Figure 24C, right part) showed either an inverted pattern or equal linker length in both regions, indicating that this pattern is not observed in all the genes, although there are no enough examples to define a biological meaning for this observation. Interestingly, all genes with longer linker length in coding regions appear to be progesterone-dependent.

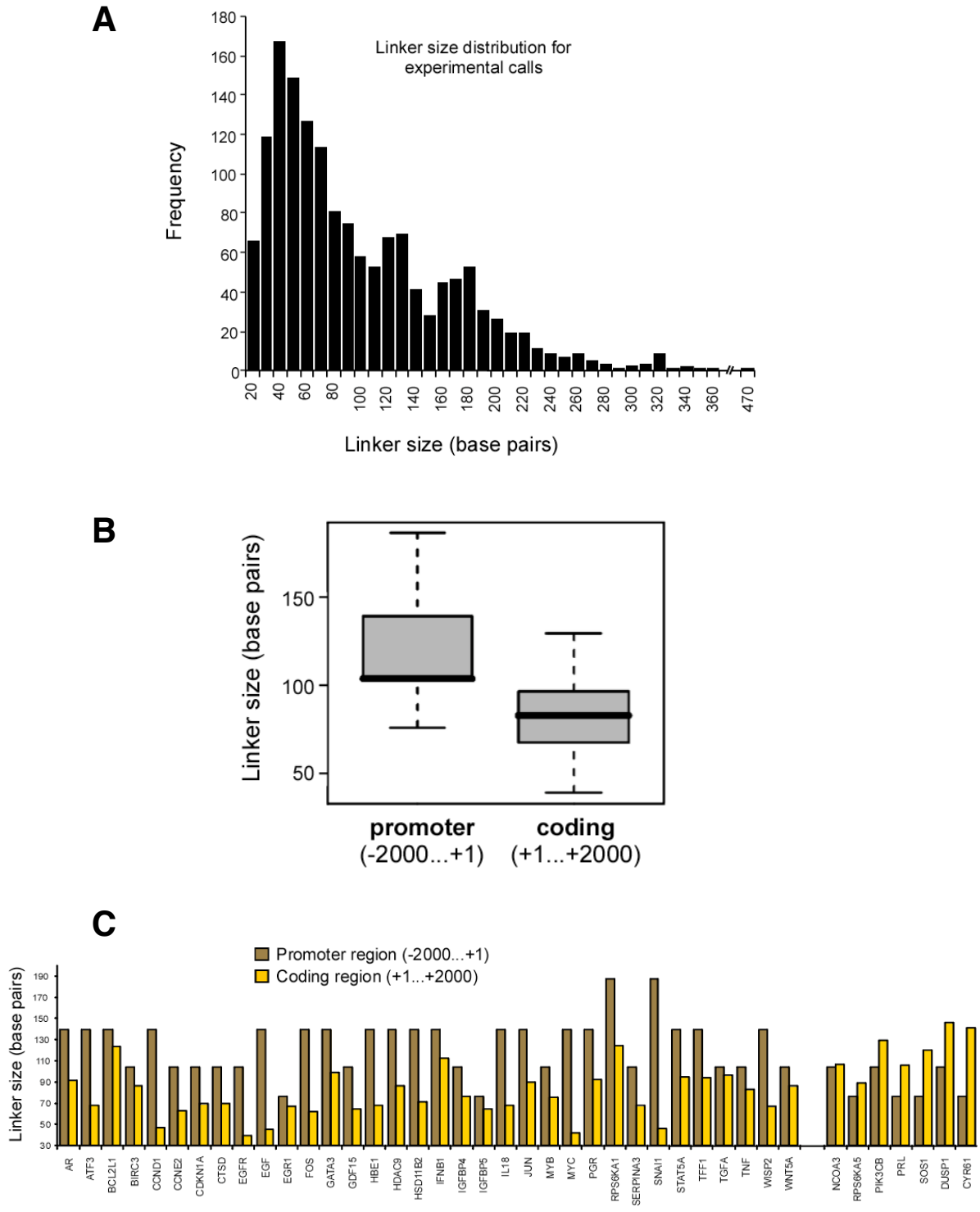


Figure 24. Linker DNA distribution from nucleosome occupancy on genes represented on the tilling array. **A**. Histogram of linker size distribution for experimental calls. The proportion of cases that fall into each linker size category. **B**. Nucleosome linker sizes in proximal promoter (2000nts upstream) and CDS regions (2000nts downstream) calculated from experimentally defined nucleosome positions in 42 human genes. Significantly smaller linker lengths in CDS regions suggest a more dense nucleosome occupancy compared to the immediate upstream region. Black bold bar is the median (Q2). The upper and lower limits of the boxes represent the upper quartile (Q3) and the lower one (Q1), respectively. Whiskers represent the minimum and maximum values of all the data. **C**. For all genes the average linker DNA size (in number of base pairs) in both promoter (from +1 to -2000nts) and coding (from +1 to +2000nts) regions is shown in the graph.

2.2.2 Nucleosome occupancy patterns around the Transcription Start Sites (TSS) of hormone-regulated genes

The first large-scale study in humans (Oszolak et al, 2007) detected a relative depletion of nucleosome in the immediate upstream of the TSS that was first described in yeast by Yuan et al. (2005). In order to study the chromatin structure around the TSS on our selected genes, we aligned a region of 600 base pairs covering de TSS (-300...+300) for all genes and calculated the average signal pattern of the mean fold-change values. 73,7% of all genes in the microarray platform shows a nucleosome covering the TSS. Interestingly, in most cases this nucleosome is slightly shifted towards the coding regions. Specifically, more than a third of the total nucleosome length (42nt) is covering the proximal promoter region while the rest lyies on the coding site. It is the first time that experimental evidences have proven a tendency of nucleosomes covering the TSS, since genome-wide analysis has described that TSS in humans resides generally in nucleosome-free regions (Barski et al, 2007 and Tirosh et al, 2008). My hypothesis is that it is due to the fact that we are studying a particular subset of genes, which may imply that a nucleosome covering the TSS is a particular feature of hormone-dependent genes. To better understand this observation further experiments need to be done. I next tried to cluster genes with similar nucleosome disposition pattern around the TSS with the same range +300 to -300 nts, but taking now into account the neighboring nucleosomes. Interestingly, the clustering resulted in a group of 20 genes with similar fold change ratio patterns around the TSS (**Figure 25**). For this group of genes a common nucleosome disposition tendency is observed within the region of interest. As expected, they showed a well-positioned nucleosome (Nuc+1) covering the transcription start site with dyad axis at around position +10 as well as a nucleosome-free region (NFR) right upstream of this first nucleosome of approximately 130 nts length. NFR total length is consistent with the second largest peak that appear in the general linker size distribution showed in Figure 24A. Moreover, another well-positioned nucleosome boundary is found around position -180 (Nuc-1). A third nucleosome (Nuc+2) appears downstream of the Nuc+1 although its position is less clear. Nucleosome +1 and +2 are separated by a short linker of about 30nt length. The genes belonging to this group are: JUN, TNF, ATF3, BCL2L1, TFF1, FOS, GDF15, AR, NCOA3, PRL, PIK3CB, WNT5A, CCND1, CDKN1A, CCNE2, TGFA, IGFBP5, IL18, SERPINA3 and SNAI1. 85% of them are either activated or inhibited by progesterone treatment.

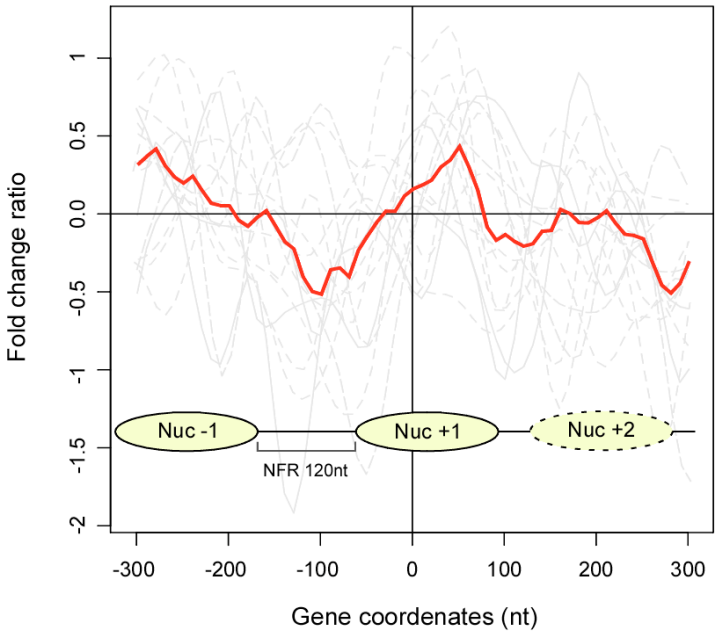


Figure 25. Nucleosome occupancy at the TSS. The mean of the log2 ratio values for a subset of genes for every nucleotide in a 600 base pair-region around the TSS (from -300 to +300) was calculated. Red line corresponds to the mean values for every nucleotide position. Grey lines on the back refers to the log2 ratio mean values for each of the genes of the subset separately. Yellow ovals represent the estimated position of nucleosomes inferred from the red-line curve.

2.3 The strongest positioned nucleosomes coincide with the highest SymCurv score

In order to be able to differentiate strong-positioned nucleosomes from weaker ones, we assigned an intensity score to each experimental call (Figure 23, different intensities of brown-coloured boxes). A high nucleosome call score is a direct indication of a fold-change pattern that fits well with our notion of a nucleosome peak (peaks of about 140 base pair long surrounded by lower ratios values). Along all regions represented on the array we have observed the existence of well-positioned nucleosomes alternating with less ones. Since I was particularly interested in understanding the biological function of the existence of a subset of strong nucleosomes, I looked in more detail at their position in the six replicas separately. Surprisingly, I have found that some nucleosomes were almost exactly reproduced (more than 90% overlap) in all replicas and that it coincided with the highest call scores values. Our analysis consisted basically on calculating the average call scores of nucleosome divided into six groups depending on their reproducibility among all six replicate experiments (**Figure 26**). As shown in the figure, this subpopulation of

reproducible positioned nucleosomes (class 6) exhibited also the highest average scores in the SymCurv algorithm. In particular, SymCurv scores for nucleosomes appearing in all 6 tiling array replicas were on average four-fold higher than the corresponding ones for nucleosomes appearing in a single replicate ($p<e-4$). This subset of nucleosomal positions that were consistent among all six experiments constituted a ~15% over the total number of nucleosomes. These parameters, reproducibility and high score, identify a subset of "well-positioned nucleosomes", pointing to a determining role of the underlying nucleotide sequence for defining its positions. We will refer to these nucleosomes as "key nucleosomes".

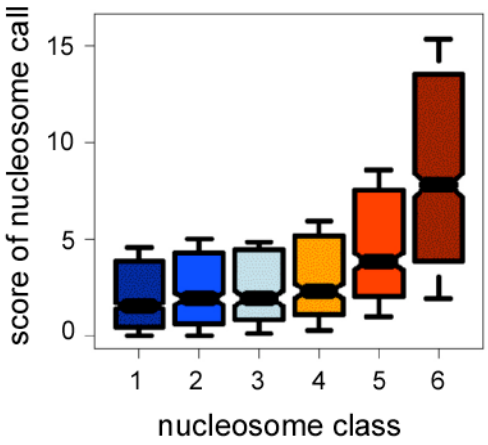


Figure 26. Experimentally well-defined nucleosomes with high call scores. Experimentally defined nucleosome positions are classified in six distinct classes depending on the number of experiments where they occur. The graph shows boxplots for all 6 classes on nucleosome against the experimental call scores assigned by a derivative-dynamic programming algorithm. Horizontal black bar for every box represent the mean of score values.

2.3.1 The concept of "key nucleosomes"

Higher experimental call score implies higher probability for a nucleosome to be always positioned in the same place. Sequence tendencies, such as symmetry of curvature, are clearer in this subset of consistent and reproducible nucleosome positions. These observations point to the existence of very well-defined nucleosome positions that are mainly governed by the sequence features, including the symmetry of curvature. Due to its dependent on the DNA sequence, it is feasible to hypothesis that the "key nucleosomes" will be consistent among cell types and different conditions. Whether they play relevant biological functions on the organization of chromatin needs to be further investigated. Moreover, given the number of unsuccessful attempts to define a universal sequence-based nucleosome positional code it is not surprising that

the nucleosomes which position is based on sequence features are merely a small subset of the total number of nucleosomes.

2.3.2 Symmetry of Curvature as a nucleosome positioning predictive tool

Based on the observation of high degrees of symmetry of curvature in “key nucleosomes”, we proposed the SymCurv algorithm as a predictive tool method for a potential nucleosomal DNA sequence. Since “key nucleosomes” are well defined by their strong degree of symmetry of curvature, this parameter could be enough to identify their positions.

We used the genomic regions represented in the array to validate the SymCurv as a predictive method for nucleosome positioning. We first calculated the average SymCurv scores, and from the symmetry of curvature values the positions of predicted nucleosomes were inferred. Next, we compare both experimental and predicted nucleosome calls and calculate the distance of each predicted position to its proximal experimentally defined one. The highest the overlapping between nucleosome calls, the better performance of SymCurv algorithm to predict nucleosome positions. Overlapping nucleosomes were classified into different groups depending on the number of nucleotides that overlaps (within a range from 146 to 112nts) (**Figure 27**). The relative scores (yellow bars for the experimental one, magenta for the SymCurv scores) increased with increase overlapping between predictions and experimental calls (grey curve). In addition there is a coincidence of high scores for both experimental calls and SymCurv-based predictions as the accuracy of the prediction increases, in other words, when there are 146nts of overlapping the highest SymCurv and experimental call scores are observed (Figure 27, firsts bars from the left). This observation implies that regions with high symmetry of curvature are also likely to be strongly occupied by nucleosomes and that this relation is quantitative. Thus, my conclusion is that stronger predictions tend to coincide better with strongly defined experimental positions and because of that, SymCurv calculation seems to be a useful tool to precisely identify “key nucleosomes”.

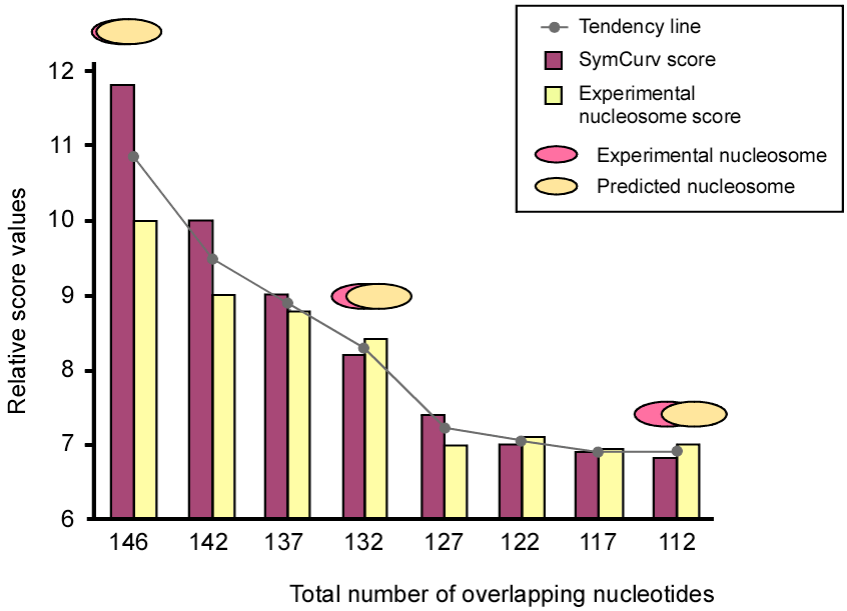


Figure 27. SymCurv algorithm as nucleosome positioning predictive method. Experimental nucleosome call scores and SymCurv scores of the predicted nucleosomes were calculated. The graph shows both scores for those experimental and predicted nucleosome that has different number of nucleotides overlapping . SymCurv score bars are magenta. Experimental nucleosome score bars are yellow. Grey line represents the tendency line between different distances categories. Yellow and magenta ovals represent predicted and experimentally defined nucleosomes, respectively.

3. Genome-wide studies of nucleosome positioning in a human breast cancer cell line by high-throughput sequencing

3.1 High-throughput sequencing as a tool to study positioning of nucleosomes for the whole genome

To expand the results obtained with the tilling array to the complete human genome I used massive parallel sequencing (technology provided by the Illumina-Solexa Genome Analyzer and Applied Biosystem SOLiD). This technique generated only 30-35 bp sequence tags, requiring both nucleosome borders to be inferred (**Figure 28**). Nonetheless, these short-read technology produces >100 times the number of sequence tags at a similar cost as the long-read technologies. The higher tag count of the short-read technology enhances mapping accurancy and thus provides a practical way of mapping nucleosomes in large genomes. Our analysis yielded ~70 million sequence reads which mapped uniquely on the human genome Figure 28, point 2). After extending these reads to the size of a complete nucleosome (147nts) towards the direction of the assigned strand (point 3), we implemented a sliding window

strategy at maximum resolution (1nt) to obtain a read-coverage score per nucleotide (point 4). Thus every nucleotide was attributed a score which was equivalent to the reads mapping on the corresponding genomic position. We then used these scores to assign non-overlapping nucleosomal positions along the human genome (Figure 28 point 5 and see *Material and Methods* for more detail). Finally, every nucleosome inferred has a score relative to the number of reads that defined it (from now, called the read-count score).The resulting dataset consisted of 8.89 million of non-overlapping nucleosomes. The distribution of nucleosome scores is highly non-random when compared to an exponential distribution with the same mean ($p<e-15$), suggesting a greater number of nucleosomes with high score than expected by random distribution.

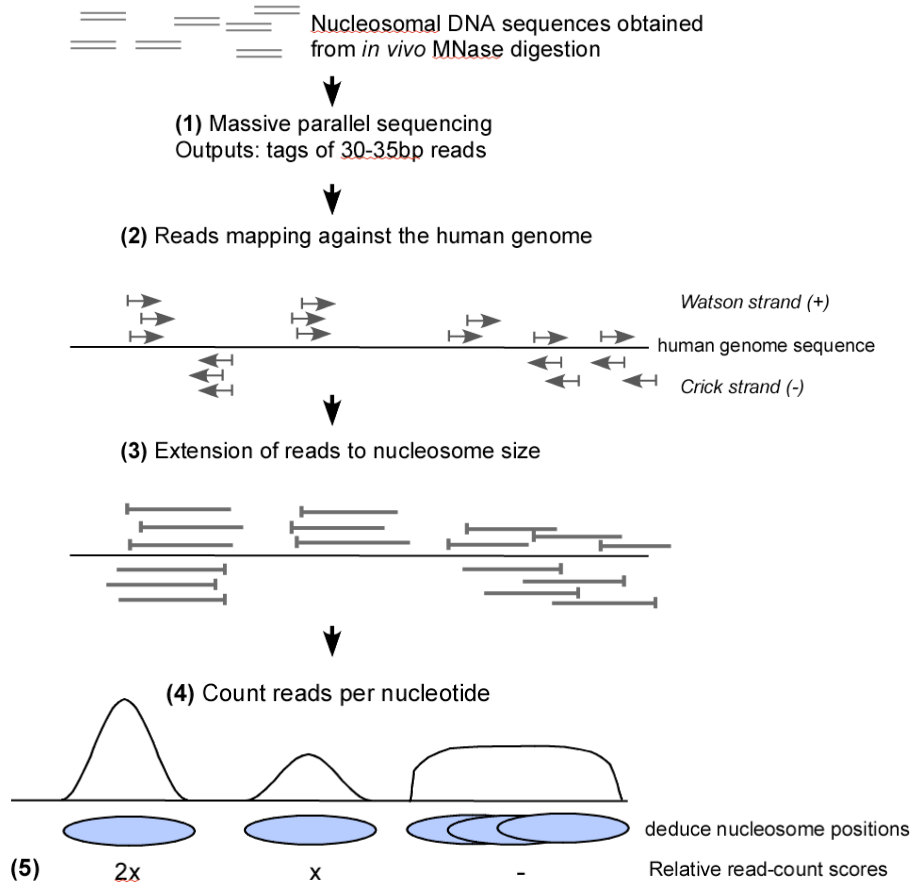


Figure 28. ChIP-Seq nucleosome mapping technology.(**1**) From 30 to 35 base pairs of each edge of nucleosomal sequences are sequenced (**2**) and mapped against the human genome. Some reads are from the positive strand (Watson) and others from the negative (Crick). (**3**) The reads are extended till a size of 147nt fragment in the direction of the strand. (**4**) We calculate a read-coverage score per nucleotide, counting reads from both strands. Positions of nucleosomes are deduced depending on the wide of the peak. Peaks of more than 147nt wide are considered fuzzy nucleosomes, as represented by overlapping of different ovals. (**5**) For those that have a well-defined position, we attribute a score depending on the number of reads that defines each deduced nucleosome. For instance, first nucleosome to the left is defined by 3 reads from positive and 3 reads from negative strand, while middle nucleosome position is defined by just 3 reads from one strand. Score of the first nucleosome will be twice the one for the second.

We first tested how many of the nucleosomes defined by deep sequencing overlapped with the "reproducible" subset of the tilling array. We found that 35% of the "Solexa" nucleosomes falling in the regions contained in the tiling array had significant (>50%) overlap with the "key nucleosome" set of the array. This modest representation of the "key nucleosomes" defined on the array in the newly defined nucleosome positions could be explained by undersampling. The 8.89 millions of identified nucleosomes cover just 43,56% of the whole human genome. We then focus on the "key nucleosomes" inferred from the microarray experiment and contained on the massive sequencing defined 8.89M of nucleosomes. We calculated the number of reads (read-count score) that defined this subset of "key nucleosomes" in the massive sequencing experiment. The average read-count score for the deep sequencing mapping the "key nucleosomes" were two-fold higher than those mapping the less reproducible nucleosomes of the tilling array, confirming the strong positional signals of these nucleosomes (**Figure 29A**).

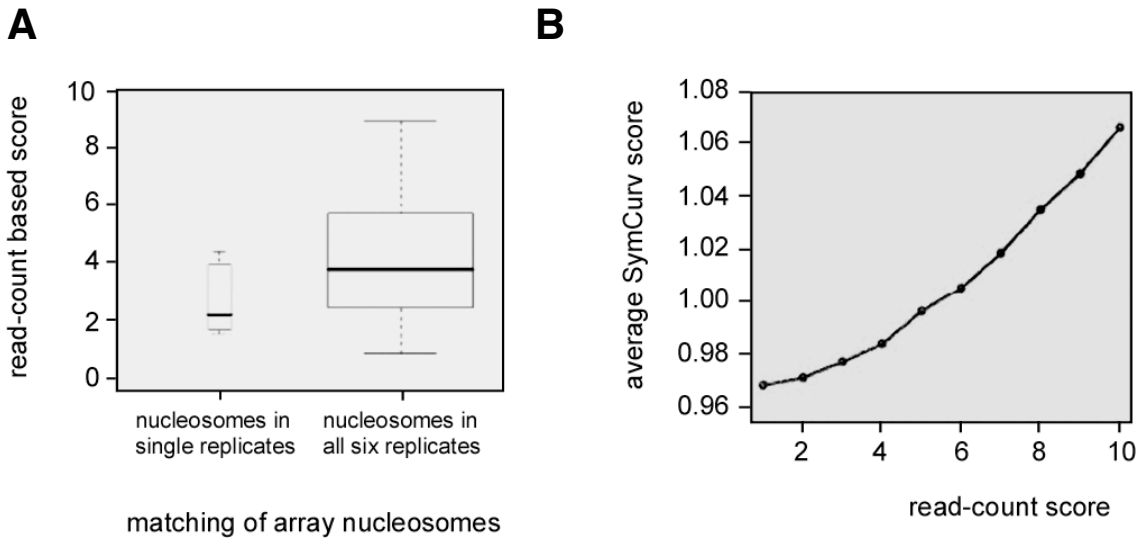


Figure 29. High read-count nucleosomes have strong symmetry of curvature. A. A 4-fold enrichment in overlap was observed for the "reproducible" 6-replicate nucleosomes (width of boxplots is proportionate to number of nucleosomes). In addition the mean read-count based score of the chipseq nucleosomes is significantly higher for the ones overlapping reproducible tiling-array nucleosomes (p< e-4). **B**. SymCurv prediction scores are correlated with read-count based scores of chipseq nucleosomes. Overlaps between SymCurv predicted nucleosomes and chipseq called nucleosomes were calculated genome-wide, for different thresholds of read-count based scores (2, 4, 6, 8 and 10). Comparison of the average SymCurv scores of the upper vs lower quartile in terms of overlap percentage, reveals a gradual enrichment in SymCurv scores in the upper quartile for increasing read-count score threshold, suggesting that the stronger the chipseq predictions the higher the SymCurv scores. (Calculated Z-scores of overlap between calls and predictions give: Correlation between log(count-based scores) and log(Z-score)=0.98, pval<0.0001. Correlation between log(Symcurv based scores) and log(Z-score)=0.52, pval < 0.001)

3.2 High read-count nucleosomes have strong symmetry of curvature

We next asked whether the highest read-count nucleosomes showed also high degree of symmetry of curvature. Indeed, a correlation between SymCurv scores and read-count scores nucleosome was observed (**Figure 29B**). Results presented in Figure 29B, show that high read-count nucleosomes also have high SymCurv scores as we previously observed with the subset of "key nucleosomes" on the tilling array. Thus, we concluded that deep sequencing identifies a particular set of well-positioned "key nucleosomes" now defined by high number of reads that also show high degrees of symmetry of curvature.

4. Properties of key nucleosomes

4.1 Key nucleosomes in biologically relevant locations: TSS, TTS and splicing sites

If "key nucleosomes" are important for chromatin organization, and function presumably they will be located in biologically relevant locations for gene expressions. To test this hypothesis we studied the distribution of "key nucleosomes" in transcription start and termination sites (TSS and TTS, respectively) as well as alternative splicing junctions. We used the complete set of transcription start sites from the ENSEMBL human genes to analyze the nucleosome occupancy. We observed an increasing enrichment in occupancy immediately downstream of the transcription initiation (**Figure 30A**, red and brown lines), indicating that "key" nucleosomes tend to be constitutively positioned at the beginning of transcripts. It also shows a tendency for a nucleosome-free region immediately upstream of the TSS. Interestingly, the pattern observed when nucleosomes are defined by a lower number of reads (less than 7) is the opposed (Figure 30A green and yellow lines). These lower read-counts nucleosomes present a nucleosome covering the TSS and two clear nucleosme-free regions at both sites. These results suggest that key nucleosomes are preferentially positioned in the vicinity of TSS region, more precisely right downstream of the TSS. A possible explanation for which these "organizing" nucleosomes are positioned at the initiation sites of transcription would be to facilitate transcription, imposing a physical "barrier" on the

correct position close to which the transcriptional machinery is assembled. Alternatively, they may serve to preclude “uncontrolled” transcription, being the silencing of genes the main function of chromatin.

An analysis identical to the one of the initiation sites was performed for the transcription termination sites (TTSs) (**Figure 30B**). Here the situation is inverted, as a nucleosome-free region is observed exactly at the end of the transcripts. In this case the pattern becomes clearer for high reads-count nucleosomes key nucleosomes (brown and red lines).

A similar sort of analysis performed on the splice sites of genes encompassing the ENCODE regions reveals strong nucleosomal patterns at the splice junctions (**Figure 30C**). In this case, just nucleosomes defined by 10 or more reads were plotted. The analysis revealed key nucleosome positioning at both exon-intron and intron-exon boundaries with a tendency for a positioned nucleosome at the side of the intron (**Figure 30D**). It is also worth to emphasize the valley of nucleosome occupancy located right at the beginning and end of exons. Recent works have pointed out the temporal-spatial connection between transcription and splicing. As in the case of transcription initiation nucleosome-free regions followed by a strong nucleosome patterns at the splice sites would act as physical barriers that might delay the RNA polymerase, giving time for the splicing machinery to be recruited. Thus, key nucleosomes seem to be functioning as chromatin barriers in genomic positions of interest to the cell's proper functioning.

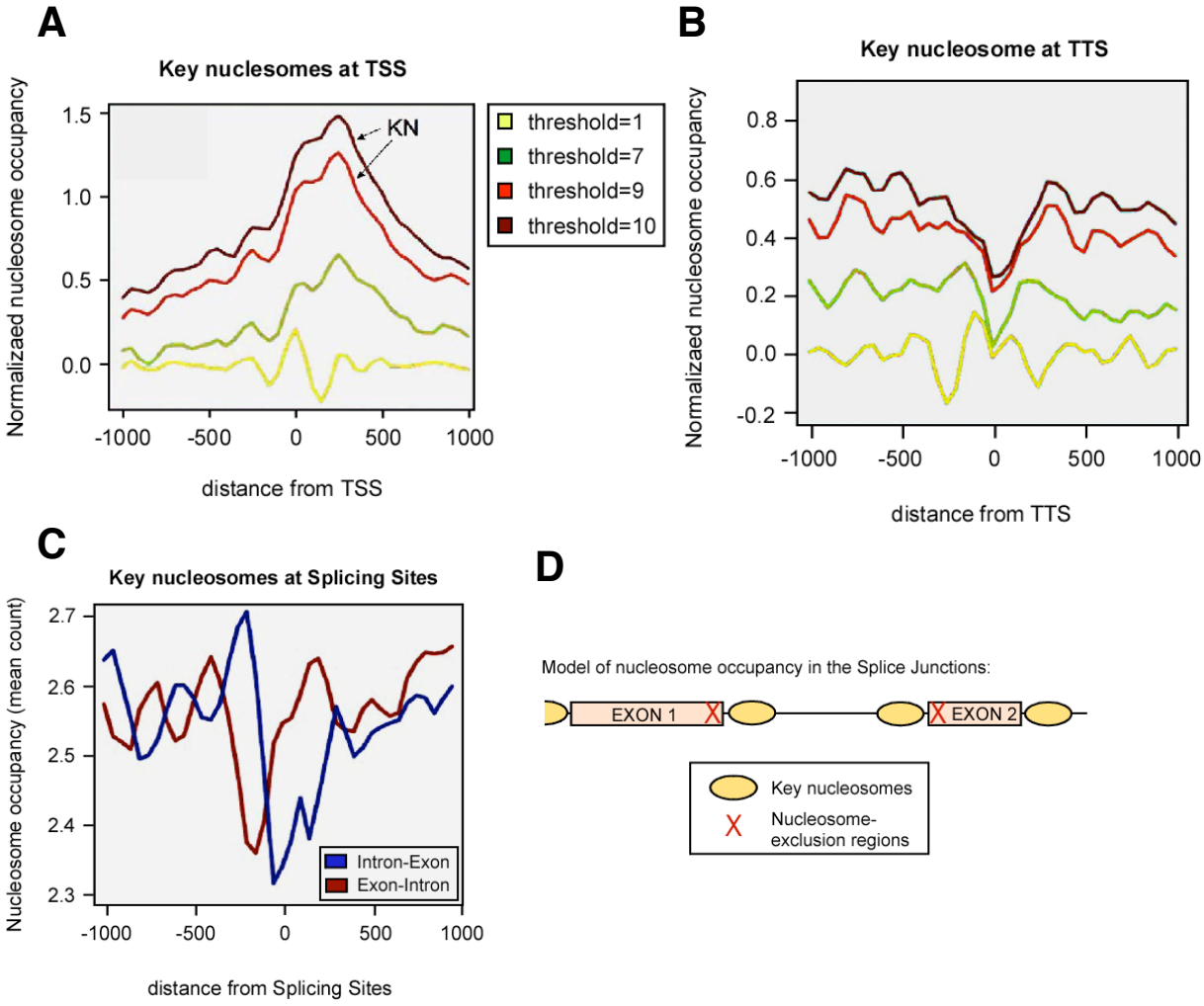


Figure 30. Key nucleosomes in biologically relevant locations: TSS, TTS and splicing sites. A. Normalized nucleosome occupancy of chipseq nucleosomes at aligned transcription start sites of ENSEMBL genes show a significant enrichment at the point of transcription initiation, that is more intense for stricter read-count based scores criteria (different thresholds). Normalized occupancies calculated as Z-score of overlaps between nucleosomes and windows of 50 nucleotides flanking the TSS. **B**. Normalized nucleosome occupancy of chipseq nucleosomes at aligned transcription termination sites of ENSEMBL genes show a significant depletion at the point of transcriptional termination. As in the case of TSS, the depletion becomes more intense as stricter read-count based scores criteria are applied. Normalized occupancies calculated as Z-score of overlaps between nucleosomes and windows of 50 nucleotides flanking the TTS. **C**. Chipseq nucleosomes exhibit strong patterns in splice sites. Splice sites from the ENCODE regions were divided in windows of 50 nucleotides and the nucleosome occupancy was calculated on the basis of read-count based scores at nucleotide resolution. Strong signals are observed at both splice site boundaries with stronger positioning at the side of the introns and a strong depletion at the side of the exon. **D**. Model of nucleosome occupancy for Splicing sites derived from results shown in (*C*).

4.2 Similar patterns on TSS, TTS and splice junctions are obtained with SymCurv values-based predictions

We tested transcription initiation and termination sites, as well as splice-sites from the ENCODE regions with the use of the SymCurv algorithm (Figure 31). For both TSS and TTS, SymCurv predictions yield a very similar pattern, predicting high nucleosome occupancy downstream of the start site and a nucleosome-free region on the termination site (**Figure 31A** red and blue lines, respectively).

Similarly, for the splice sites, strong nucleosomal patterns are predicted at the vicinity of both splice junctions with stronger nucleosome occupancy at the side of the introns (**Figures 31B**, intron-exon represented in blue and exon-intron in red). Similarities between the predicted nucleosomal patterns and the ones obtained with strict read-count score thresholds aiming at key nucleosomes, are indicative of the fact that chromatin structure of functional regions of the genome is reflected upon structural properties of the DNA primary sequence.

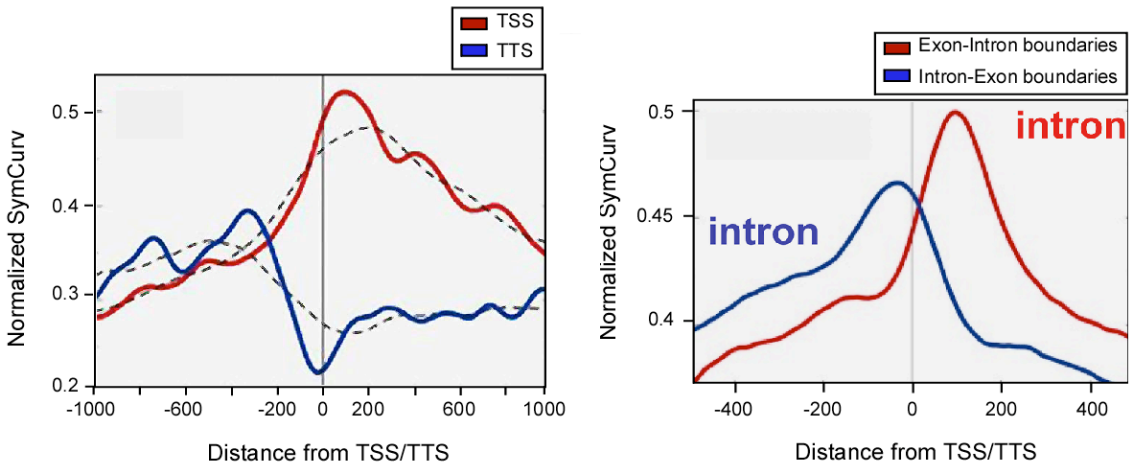


Figure 31. SymCurv-based predictions in TSS, TTS and Splicing Sites. A. SymCurv-predicted nucleosomes exhibit strong patterns at the starts and ends of genes. TSS and TTS sites obtained from the ENCODE regions were divided in windows of 50 nucleotides and the nucleosome occupancy was calculated on the basis of SymCurv scores at nucleotide resolution. Strong positioning is observed at the initiation and strong nucleosome-free regions are predicted for the termination sites in accordance with experimental results. **B**. SymCurv-predicted nucleosomes exhibit strong patterns around splice sites. Splice sites from the ENCODE regions were divided in windows of 50 nucleotides and nucleosome occupancy was calculated on the basis of SymCurv scores at nucleotide resolution.

Discussion

We introduce a new theoretical concept, SymCurv, that allows precise position of a subset of nucleosomes on the human genome to be predicted *in silico* based on the structural attributes of the primary DNA sequence. Nucleosomal DNA of these subset of nucleosomes exhibited the highest symmetry of curvature, suggesting that their positions are strongly determined by the conformation of the underlying DNA sequence. In this work I give experimental evidence for the existence of a subset of very well-positioned nucleosomes in the human genome. Using a tilling array encompassing 420 kb around the 5´-end of 42 human genes and a novel dynamic programming based algorithm, we found that ~15 % of the nucleosome calls mapped reproducibly to the same positions (more than 90 % overlapping) in six experimental replica, while the positions of the other nucleosomes were less consistent. Analysis of the whole human genome by massive nucleosome sequencing confirmed the existence of a subset of well-positioned nucleosomes, named “key nucleosomes”, which exhibit interesting sequence properties and a peculiar distribution along the genome, tending to occupy genomic space of specific functionalities.

The Symmetry of Curvature on nucleosomal DNA

The approach described to identify well-positioned nucleosomes, is based on the hypothesis for the existence of a genomic signature for nucleosome positioning. The basic assumption is that sequences selected for nucleosome binding will tend to accumulate special features facilitating nucleosome formation through evolution. These features are more likely to be structural and expected to be more extended, meaning covering a longer sequence string, and more degenerate than a usual protein binding sequence. Nucleosome formation involves wrapping of a ~146bp sequence around a protein octamer, thus DNA curvature is expected to facilitate this process. In addition, the symmetric structure of this octamer is likely to impose corresponding symmetry constraints on the sequence, as in a manner similar to the majority of symmetric DNA-

binding proteins. Combining the above prerequisites we have proposed the symmetry of curvature as a possible signal facilitating nucleosome formation. The results presented in this thesis show that there exists a correlation between this property and the tendency for nucleosome positioning. Moreover, symmetry of curvature as a property of nucleosomal DNA was a found in nucleosomal sequences analyzed by very different experimental techniques (collections of nucleosomal DNA, tilling microarray hybridization and massive parallel sequencing), proving the reliability of the concept.

SymCurv algorithm as a nucleosome predictive tool

Around 15% of the total number of experimental nucleosome calls showed a high tendency to occupy the same place among different experimental replicas in microarrays' experiments. In addition, SymCurv performed with increased efficiency on this same subset. Larger scale nucleosome mapping on human genome showed nucleosomes defined by a high number of reads that appear to have also the highest SymCurv scores. Consistent with our results, other works on the field have also suggested that only a subset of the predicted nucleosomes is actually guided by the sequence (Yuan et al., 2005; Segal et al., 2006; Kaplan et al., 2008), but no attempt was done on defining this subset of nucleosomes. Moreover, when testing various datasets of nucleosomal sequences, we observed a gradient increase of the SymCurv pattern intensity in the datasets obtained through more rigorous and step-filtering experimental processes, suggesting that SymCurv algorithm is also a quantitative method to predict nucleosomes (high scores correlate with strong signal of nucleosome positioning for a given sequence).

The first genome-wide nucleosomal sequence set published, originating from *C. elegans* (Johnson et al., 2006) also suggests that only a small part of the sequences is actually under constraint of nucleosome positioning. Later works by Valouev et al. (2008) and Schones et al. (2007) further supported this finding, pointing towards a model of statistical positioning of nucleosomes, according to which, a small set of well-positioned nucleosomes (or alternatively, a set of regions that strongly exclude them) would guide the subsequent binding of the large mass of nucleosome particles, which would occupy the remaining space. The presented results indicate that the SymCurv method could help in identifying regions of high nuclesome forming potential in many different species.

Nucleosome organization around hormone-dependent genes

Here I have used a tilling array to study the general patterns of nucleosome organization at promoter and coding regions, and specially around the TSS, of a subset of hormone responsive genes. One of the limitations of the method used to infer nucleosomes from the raw microarray data is that regions defined by fuzzy nucleosomes are, instead, occupied by experimental nucleosomes with low scores. Dynamic programming will always position nucleosomes even if it is a wide region with many different possible overlapping positions. However, our analysis on linker length and nucleosome occupancy of the TSS will not be modified by this point. Different linker lengths observed all over the regions represented on the array might reflect the implication of extra factors other than the histone core that modulate spacing and general positioning of the nucleosomes, such as presence of different linker histone variants and remodelling complex among others (Whitehouse et al, 2007). Alternatively, in some cases linker DNA would contain nucleosome exclusion signals that avoid positioning of core histone particles in specific locations. In yeast the meaning of nucleosome-free regions (NFR) around TSS has been widely studied (Petesch et al, 2008 and Venters et al, 2009). It has been described that although a NFR is permissive for transcription, it is not sufficient to activate genes. Thus, NFRs might allow low basal levels of leaky transcription, which could be interpreted as meaningless biological noise, particularly if the transcripts are rapidly degraded. NFRs in the 5' and the 3' of yeast genes has opposed functions: 5'NFRs are likely to be sites for the assembly of the transcription machinery, whereas 3'NFRs are likely to be sites for the disassembly of the transcription machinery. The meaning of NFRs in human genes remains still an open question.

Till now we have characterized the nucleosome occupancy on 40 hormone-responsive genes. Next step will be to study how nucleosome positioning is changed after induction/inhibition of transcription upon treating the cells with hormone. Our system is limited to 40 genes but it allows us to presicely study nucleosome-resolution changes after promoter regulations. It is an interesting assay, since in human genes it is still unclear the pattern of nucleosome occupancy around inducible genes, since most of the studies focus on constitutively activated genes.

Genomic-scale studies on nucleosome position

Recent technological advances in highly parallel sequencing allow the conduction of large-scale nucleosome positioning studies. In a recent study, Dennis et al. (2007) discuss experimental methods for large-scale structural analysis of mammalian chromatin, which, although elaborate, do not precisely replicate the nucleosomal pattern of well-studied cases. This is likely due to both the complexity of the positioning problem, (alternative positioning, flexibility of binding, influence of sequence-specific DNA-binding proteins, etc) and the number of information-processing steps needed from the raw experimental data to actually call a nucleosomal position. Thus it is not surprise that the concordance of computationally predicted positions with experimentally defined ones is relatively small in all published methods to date, whether they are based on probabilistic likelihood (Segal et al., 2006 and Kaplan et al, 2008), comparative genomics (Ioshikhes et al., 2006) or supervised learning (Peckham et al., 2007). Our aim was to focus on the subgroup of nucleosomes which position signal was strong and based on the DNA sequence features.

The meaning for a key nucleosome to be in specific functional locations

We have shown the significant enrichment of “key nucleosomes” upstream of the TSS and depletion right on the TTS, as well as specific positioning around alternative splicing sites. The biological meaning of these observations can be explained generally by the need to control of basic cellular mechanisms. “Key nucleosomes” at splice sites would act as physical barriers that might delay the RNA polymerase, giving time for the splicing machinery to be recruited, while the meaning for the TSS locations could be either a physical barrier or a platform for both specific histone posttranslational modifications and histone variant incorporations that may modulate the active state of the promoter.

However, future analysis is needed in order to further elucidate the functional role of “key nucleosomes” in specific locations. Future experiments could go, for instance, towards the study of the correlation of gene expression and “key nucleosome” occupancy or sequence conservation of these subset of nucleosomes compared to neigbouring ones. To test whether “key nucleosomes” are also playing an important role on higher-order chromatin organization, it would be interesting to study whether they are

periodically spaced on the human genome. One hypothesis could be that “key nucleosomes” are multi-functional entities that may at the same time regulate cellular processes and organize chromatin compaction. In this way, nucleosomal sequences may simultaneously achieve a number of different tasks.

Chapter 2

"Two chromatin remodeling activities cooperate during activation of hormone responsive promoters"

This manuscript has been send for publication to *PloS Genetics* (No. 08-PLGE-RA-1702R3). It is currently under the third revision.

Two chromatin remodeling activities cooperate during activation of hormone responsive promoters

Guillermo Pablo Vicent[1,3], Roser Zaurin[1,3], A. Silvina Nacht[1], Ang Li[1], Jofre Font-Mateu[1], Francois Le Dily[1], Michiel Vermeulen[2], Matthias Mann[2] and Miguel Beato[1]

[1]Centre de Regulació Genòmica (CRG), Universitat Pompeu Fabra, Parc de Recerca Biomèdica (PRBB), Dr Aiguader 88, E-08003, Barcelona, Spain.

[2]Max Planck Institute for Biochemistry, Department of Proteomics and Signal transduction, Am Klopferspitz 18, D-82152 Martinsried, Germany.

[3] both authors contribute equally to this work

Correspondence should be addressed to M.B.: *e-mail: miguel.beato@crg.es*

Running Title: BAF and PCAF cooperate in MMTV induction

Abstract

Steroid hormones regulate gene expression by interaction of their receptors with hormone responsive elements (HREs) and recruitment of kinases, chromatin remodeling complexes and coregulators to their target promoters. Here we show that in breast cancer cells the BAF, but not the closely related PBAF complex is required for progesterone induction of several target genes including MMTV, where it catalyzes localized displacement of histones H2A and H2B and subsequent NF1 binding. PCAF is also needed for induction of progesterone target genes and acetylates histone H3 at K14, an epigenetic mark that interacts with the BAF subunits anchoring the complex to chromatin. In the absence of PCAF, full loading of target promoters with hormone receptors and BAF is precluded and induction is compromised. Thus activation of hormone-responsive promoters requires cooperation of at least two chromatin remodeling activities, BAF and PCAF.

Author Summary

In order to adapt its gene expression program to the needs of the environment the cell must access the information stored in the DNA sequence that is tightly packaged into chromatin in the cell nucleus. How the cell manages to do it in a selective maner is still unclear. Here we show that in breast cancer cells treated with the ovarian hormone progesterone the hormone receptor recruits to the regulated genes two chromatin remodeling complexes that cooperate in opening the chromatin structure. One of the complexes puts a mark in a chromatin protein that anchors the other complex enabling full gene activation. The present discovery highlights the importance of the concerted order of events for access to genomic information during activation of gene expression and reveals the intricacies of hormonal gene regulation.

Introduction

Regulation of eukaryotic gene expression implies mechanisms that permit transcription factors to gain access to chromatin packaged DNA sequences. The basic unit of chromatin, the nucleosome, consists of an octamer formed of two copies of each of the four core histones (H2A, H2B, H3, and H4) around which 147 bp DNA is wrapped in 1.65 left-handed superhelical turns [1]. Modulation of the structure and dynamics of nucleosomes is an important regulatory mechanism in all DNA-based processes and is catalyzed by chromatin remodeling complexes. Such complexes can either modify histone residues or use the energy of ATP hydrolysis to alter the relationship between histones and DNA [2, 3].

The yeast SWI/SNF complex, the first ATP-dependent chromatin remodeling complex to be identified, is a 2-MDa complex of 11 subunits that regulates gene expression by catalyzing octamer transfer, nucleosome sliding, dinucleosome formation, and H2A/H2B displacement [4-6]. RSC, (Remodel the Structure of Chromatin) [7], is a closely related yeast chromatin remodeling complex of 15 subunits, that shares two identical and at least four homologous subunits with the ySWI/SNF complex [8]. There are two human SWI/SNF-like complexes both containing ATPase subunits similar to yeast Swi2/Snf2, hBRM (human Brahma) or BRG1 (Brahma-Related Gene 1), as well as a series of other subunits, some of which differ in various cell types [8]. The hSWI/SNF-α complexes, also called BAF for BRG1/hBRM-Associated Factor, contain either hBRM or BRG1 as ATPase and is orthologue to yeast SWI/SNF. The hSWI/SNF-β complex, also called PBAF (Polybromo-associated BAF) contains only hBRG1 and is orthologue to yeast RSC complex [4, 9]. The BAF and PBAF complexes share many subunits but have also subtype specific subunits: BAF250 and hBRM are only found in BAF, whrereas BAF180 and BAF 200 are only found in PBAF [4, 10]. BAF57 has been reported as a common subunit for BAF and PBAF complexes [4, 9].

Histone acetyltransferases (HATs) and histone deacetylases (HDACs) represent other group of chromatin remodeling complexes that regulate the level of acetylation on the N-terminal tails of core histone proteins and other protein substrates [11, 12]. The HATs are divided into five families, including the GCN5-related N-acetyltransferases (GNATs) with GCN5 and PCAF as the best characterized members; the MYST ('MOZ, Ybf2/Sas3, Sas2 and Tip60')-related HATs; p300/CREB-binding protein (CBP) HATs; the general transcription factor HATs including the TFIID subunit TBP-associated factor-1 (TAF1); and the nuclear hormone-related HATs SRC1 and ACTR (SRC3) [13, 14]. Recombinant PCAF, like full-lenght GCN5, acetylates either

free histones or nucleosomes [15], primarily on lysine -14 of histone H3 [16]. The role of PCAF in transcription has been investigated in multiple studies, and its requirement as a HAT and coactivator has been described in several processes including nuclear receptor mediated events [17, 18], but the precise mechanism of action has not yet been elucidated .

The functional relationship between different chromatin remodeling enzymatic activities is of great interest. A remarkable interdependence has been described during transcriptional activation in *S. cerevisiae* between the SWI/SNF complex and the histone acetylation complex SAGA [19]. Bromodomains of the RSC complex have been shown to recognize acetylated H3K14 [20]. HAT activity stabilizes SWI/SNF binding to promoter nucleosomes providing a mechanistic basis for the ordered recruitment of these complexes [21]. In mammalian cells, transcriptional activation by nuclear receptors requires multiple cofactors including CBP/p300, SWI/SNF and Mediator. The ordered recruitment of these cofactors to the promoters depends not only on the direct interactions between nuclear receptors and cofactors but also on cofactor-cofactor interaction and on histone modifications [22].

Gene regulation by steroid hormones is mediated by intracellular receptors, which upon hormone binding can activate signal transduction cascades and interact in the nucleus with other transcription factors and/or with dedicated DNA sequences, called hormone responsive elements (HREs) [23]. When ligand-receptor complexes interact with HREs they alter the transcriptional state of the associated genes via the recruitment of chromatin remodeling complexes and other coregulators to the target promoters [24, 25]. The mouse mammary tumor virus (MMTV) long terminal-repeat region encompasses a promoter including within 140 bp five degenerated hormone responsive elements (HREs) and a binding site for nuclear factor 1 (NF1). In chromatin the MMTV promoter is organized into positioned nucleosomes [26], with a nucleosome located over the five HREs and the NF1 binding site [27]. On this promoter nucleosome, the binding site for NF1 is not accessible and only two of the five HREs, the strong palindromic HRE1 and the weak half palindrome HRE4, can be bound by hormone receptors, while the central HREs, in particular the palindromic HRE2 and the half palindrome HRE3, are not accessible for receptor binding [28]. Following hormone induction *in vivo* all HREs and the binding site for NF1 are occupied simultaneously on the surface of a nucleosome-like structure and a functional synergism is observed between progesterone receptor (PR) and NF1 [27].

Already 5 minutes after addition of hormone to breast cancer cells with an integrated copy of an MMTV-luc reporter, the cytoplasmic signaling cascade Src/Ras/Erk is activated via an interaction of PR with the estrogen receptor, which activates Src [29]. As a consequence of Erk activation, PR is phosphorylated, Msk1 is activated, and a ternary complex PR-Erk-Msk1 is recruited to nucleosome B [30]. Msk1 phosphorylates H3 at serine 10, a modification which is accompanied by acetylation at H3 lysine 14, displacement of HP1γ, and recruitment of RNA polymerase II [30]. Inhibition of Erk or Msk activation blocks H3S10 phosphorylation, H3K14 acetylation, and hormonal induction of MMTV and other progesterone target genes [30].

There have been several reports indicating a role for SWI/SNF, Brg1 and Brm in glucocorticoid regulation of MMTV transcription [31-33], but the situation with progesterone is less clear. The MMTV promoter assembled in minichromosomes with positioned nucleosomes [34] is activated by PR in a process involving a NURF-dependent mutual synergism between PR and NF1 [35]. Progesterone treatment of a breast cancer cell line carrying a single integrated copy of a MMTV transgene leads to recruitment of PR, SWI/SNF, and SNF2h-related complexes to the MMTV promoter, accompanied by selective displacement of histones H2A and H2B from nucleosome B [6, 25]. Recently, the acidic N terminus of the Swi3p subunit of yeast SWI/SNF was identified as a novel H2A-H2B-binding domain required for ATP-dependent H2A/H2B dimer displacement [36]. Moreover, a characteristic and sharp DNase I hypersensitive site appears near the symmetry axis of the nucleosome encompassing the HREs [27]. The same hypersensitive site can be induced in the absence of hormone by treatment with low concentrations of histone deacetylases inhibitors, such as sodium butyrate or trichostatin A [37], suggesting that it can be initiated by a moderate increased in histone acetylation. However, the relationship between histone acetylation and ATP-dependent nucleosome remodeling following hormone treatment is not clear. To approach this question we have performed knockdown and ChIP experiments targeting human SWI/SNF complexes and histone acetyl transferases (HATs) in breast cancer cells.

Results

BAF is required for progesterone gene activation in T47D-MTVL cells and is recruited to the MMTV promoter upon induction

The requirement of Brm and Brg1 for MMTV promoter activation was assessed in T47D-MTVL cells transfected with specific siRNAs. When siRNA against Brg1 was

used, the levels of Brm protein increased (Figure 1 A, left panels) and viceversa (Figure 1 A, middle panels). A significant reduction (80% and 55% for Brm and Brg1, respectively) of both ATPases was only observed when the two siRNAs were combined (Figure 1 A, right panels). The levels of the two PR isoforms, PR_B and PR_A, were unaffected by Brg1 or Brm siRNAs (Figure 1 A, third row). In cells transfected with control siRNA an eight-fold increase in MMTV-luc transcription was observed 8 hours after hormone treatment (Figure 1 B). In cells transfected with siRNAs against either Brg1 or Brm, the induction was reduced by 30% and 20% respectively, whereas transfection of both siRNA simultaneously reduced the induction of the MMTV promoter by 60%, confirming the essential role of hSWI/SNF on MMTV promoter activation (Figure 1 B). Progestin induction of MYC and FOS, two other progesterone target genes, was also impaired by the combined siRNAs to a level similar to that observed with MMTV-luc (Figure 1C). However, not all hormone-dependent genes required SWI/SNF, as the combined siRNAs did not inhibit the progestin induction of the cyclin D1 gene (Figure 1 C).

To identify the nature of the Brm/Brg1-containing complex in T47D-MTVL cells, we performed western blotting against subunits of BAF and PBAF, the two distinctive hSWI/SNF complexes. Subunits of both complexes exist in T47D-MTVL cells (Supplementary Fig.1), indicating that our cell line is capable of forming both BAF and PBAF. The specific subunits that distinguish the two complexes are Brm and BAF250 for BAF, and BAF180 and BAF200 for PBAF. Knocking down BAF57 and BAF250 levels by approximately 50% resulted in a similar reduction in the hormonal induction of the MMTV promoter without changes in Brg1 and Brm protein levels (Figure 1 D and E and data not shown). Co-immunoprecipitation experiments demonstrated that in T47D-MTVL cells BAF57 forms a complex with the core subunits BAF155, BAF170, BAF 180 and the ATPases Brg1 and Brm (Figure 1 F, lane 2 and data not shown). We conclude that BAF and PBAF complexes in the breast cancer cell lines used for these studies contain BAF57.

To further identify the specific complex recruited to the MMTV promoter, we performed ChIP experiments using antibodies recognizing specific components of the BAF and PBAF complexes. Simultaneously with the recruitment of PR to the MMTV promoter after hormone treatment, we observed recruitment of Brg1, Brm, BAF250 and BAF57 (Figure 1 G). In contrast, BAF180 and BAF200, two of the specific subunits of PBAF, were not recruited to the MMTV promoter after hormone treatment (Figure 1 G, third and fourth rows from the bottom), while they were recruited to the RARb2

promoter after retinoic acid treatment in U937 cells (Supplementary Fig.2, A). Moreover, siRNA mediated downregulation of BAF180 in T47D-MTVL cells did no influenced hormonal induction of the MMTV promoter (Supplementary Fig.2, B). Thus, BAF complexes containing either one of the two related ATPases together with BAF250 and BAF57 are recruited to the MMTV promoter along with PR and likely mediate promoter transactivation.

BAF57 and BAF250 interact with progesterone receptor in T47D-MTVL cells

We used co-immunoprecipitation experiments to test whether BAF57 forms a complex with PR in cultured cells. In the absence of hormone, a certain proportion of BAF57 already coprecipitated with PR probably due to the large proportion of PR molecules already present in the nucleus in the uninduced state; however 30 minutes after hormone addition the extent of coprecipitation was increased (Figure 2 A, lanes 4 and 5). In contrast, no complex of PR with the PBAF specific subunit, BAF180 was observed independently of the addition of the hormone (Figure 2 B). As a positive control for this experiment we used BAF250, a known BAF specific subunit [38]. BAF250 as BAF57 also showed a hormone-dependent interaction with PR (Figure 2 B, lanes 2 and 3).

The PR-BAF57 interaction is likely direct, as activated ligand-bound PR interacts with GST-BAF57 in pull-down studies (Figure 2 D, lane 2). Deletion mutants lacking either the C-terminal domain or the N-terminal domain, or containing only the central NHRLI domain (Figure 2 C, left panel), could not form a complex with PR in GST-pulldown experiments (Figure 2 D top row, lanes 3-5).

The fact that BAF57 is present in both BAF and PBAF complexes but only BAF is recruited to the target promoters indicate that either the PR-BAF57 interaction is not sufficient for recruitment of the complex or BAF57 can only interact with PR in the context of the BAF complex but not when integrated in the PBAF complex.

Acetylation on histone H3K14 by PCAF is required for hormonal induction

As the central ATPases of the BAF complex, Brg1 and Brm, both contain bromodomains that recognize acetylated lysines [39, 40] and have been shown to bind acetylated histone H3 and H4 tails [41, 42], we wondered about possible roles of histone acetyltransferases (HATs) in hormone induction. Using siRNAs we knocked down PCAF and/or SRC1, which are known to be recruited to the MMTV promoter upon induction [6, 30] (Figure 3 A). We do not have an explanation for the

enhancement of the unspecific band (marked with an asterisk in Figure 3A, first row, lanes 2 and 4). When siPCAF was used, there was a 50% reduction of the MMTV induction (Figure 3 A, right panel) paralleling the decrease in PCAF protein (Figure 3 A, left panel), whereas siSRC1 caused only a 15% decrease (Figure 3 A, right panel). The combination of the two siRNAs decreased induction by 62% (Figure 3 A, right panel), indicating that PCAF is the more relevant HAT in hormonal activation of MMTV. PCAF depletion also decreased induction of other progesterone target genes such as MYC and FOS (Figure 3 B), which were shown to depend on BAF.

The need for the catalytic activity of PCAF in regulating MMTV promoter was studied in T47D-MTVL cells transfected with wild type PCAF or an enzymatic deficient mutant (PCAFΔHAT). While transfection of wild type PCAF further increased the level of activity obtained after hormone addition, transfection of PCAFΔHAT did not further stimulate the MMTV promoter (Figure 3 C). Although PCAF and PCAFΔHAT are expressed at similar levels (data not shown), PCAFΔHAT does not act as a dominant negative mutant possibly because its levels are not sufficient to compete with endogenous PCAF in transient transfection. Similar results were obtained when MYC and FOS genes were analysed (data not shown). Thus, the catalytic activity of PCAF is required for transcriptional activation of those progestin target genes, which induction depends on BAF. Regarding other acetyl transferases, no decrease in hormone-dependent activation of the MMTV was observed after transfection with specific siRNAs against GCN5 (data not shown).

Acetylated H3 lysine 14 anchors the BAF complex on the promoter

We next tested the influence of histone acetylation on BAF recruitment to the MMTV promoter. In control cells hormone treatment induced increased acetylation of H3K14 accompanied by recruitment of BAF57 and BAF250, whereas knockdown of PCAF and SRC1 abrogated K14 acetylation and markedly reduced BAF57 and BAF250 loading on the promoter (Figure 3 D, compare lanes 1-2 vs 3-4; quantification by real time PCR is shown in the right panel). To test whether H3K14 acetylation and BAF binding take place on the same promoter, we performed sequential ChIP (re-ChIP) experiments in the MMTV, FOS and the MYC promoters. Hormone treatment induced increased acetylation of H3K14 accompanied by recruitment of BAF250 to the corresponding promoters (Figure 3 E, lanes 5 and 6), indicating that the H3K14 acetylation and BAF recruitment occur in the same genomic region. In contrast, control re-ChIPs performed

with irrelevant IgG as first antibody showed no amplification product (Figure 3 E, lanes 3 and 4).

To test whether the association of H3K14 acetylation and BAF is a general phenomenon, we performed co-immunoprecipitation of total hormone-treated chromatin followed by western blot. This seems to be the case, as immunoprecipitation with antibody against acetylated H3K14 co-precipitated BAF155 (Figure 3 F) and Brg1 (data not shown).

We next used pull-down experiments with T47D-MTVL nuclear extracts and biotinylated peptides of H3 and H4 tails containing various modifications to test whether H3K14 acetylation influences the binding affinity of the BAF complex for histone tails. Histone H1 was used as loading control (Figure 4 A, lower row). We detected binding to histone tails containing H3K14ac of nearly all subunits of the BAF and PBAF complexes, namely Brg1, Brm, BAF170, BAF155, BAF57, BAF180 and SNF5 (Figure 4 A lanes 4, 5, 7 and 8). Similar results were also obtained using HeLa nuclear extracts (data not shown). Moreover, binding of BAF and PBAF complexes to H3 peptides depends solely on H3K14ac and it is not affected by acetylation, phosphorylation or methylation of adjacent residues. Peptides carrying additional modifications, such as K4me3, K9ac, K9me3, or S10p, exhibited similar affinity for BAF subunits as those carrying only K14ac (Figure 4 A lanes 5, 7, 8 and data not shown). H3 peptides acetylated at K9 and pan-acetylated (K5, 8, 12, 16) H4 peptides did not exhibit preferential BAF affinity (Figure 4 A lanes 3, 10 and 11). Although H4K8 acetylation takes place very early after progesterone addition to T47D-MTVL cells (Figure 4 B, upper panel), we found here no binding of the BAF complex to acetylated H4 peptides. The PBAF specific subunit, BAF180 showed the same behaviour as the BAF subunits when incubated with the modified peptides (Figure 4 A, third row from the top). It is important to note, that we detected a faint and consistent binding of PR to all tested H3 and H4 peptides without any preference for particular epigenetic modification, indicating that PR is not acting as a bridging factor for the binding of BAF complex to the K14 acetylated peptides (Figure 4 A, second row from the bottom). A similar behaviour was observed for PCAF without any preference for the tested modifications (data not shown).

Next, we quantify the interaction of the BAF subunits to H3 acetylated peptides by *s*table *i*sotope *l*abeling by *a*mino acids in *c*ell culture (SILAC) [43]. Nuclear extracts derived from human HeLaS3 cells grown in light or heavy medium were incubated with immobilized histone peptides in the nonacetylated (H3) and acetylated form,

respectively (Supplementary Fig.3, A). We identified BAF180, BAF200, Brg1 and Brm as selective interactors for H3K9ac, H3K14ac and H3K9acK14ac. The SILAC ratio (Heavy/Light, that is modified/unmodified) for Brm was 1.8 for H3K9ac and 7.9 for H3K14ac (Figure 4 C and Supplementary Fig.3), indicating that BAF complex is attached preferentially to K14 acetylated peptides and confirming the results obtained with peptide pull downs (Figure 4 A). Higher SILAC ratios with H3K14ac peptide were also obtained for Brg1 and for the PBAF specific subunits, BAF180 and BAF200 (Figure 4 C). An additive effect on binding was observed when the double modification H3K9acK14ac was used (Figure 4 C). However, we did not find an increase in H3K9ac on the MMTV promoter following hormone induction (Figure 4 B, lower panel). These results highlight the role of K14 acetylation as the main modification responsible for BAF and PBAF binding.

Knocking down Brg1 and Brm reduces the interaction of BAF57 and BAF155 with H3K14ac peptides without affecting the intracellular levels of their complex with BAF170 (Supplementary Fig.4 and data not shown), indicating that the bromodomains of the ATPases contribute to the anchoring of the BAF complex at H3Kac containing sites.

The BAF complex facilitates NF1 binding by mediating histone H2A/H2B displacement

Next we used knockdown experiments to test whether BAF is involved in NF1 binding to the MMTV promoter, a critical step in hormonal induction [35]. The NF1 family of transcription factors in vertebrates is composed of four different genes: NF1A, NF1B, NF1C and NF1X [44], of which NF1C most abundantly expressed in mammary gland and is involved in reciprocal and sequential synergism with hormone receptors [45, 46]. In control cells, NF1C bound to the MMTV promoter in response to hormone, while binding was diminished upon Brg1 and Brm knockdown (Figure 4 D, first row, compare lanes 1-2 vs 3-4; quantification by real time PCR is shown in the right panel), suggesting that BAF complexes are necessary for NF-1 to access the promoter in response to hormone.

To investigate how BAF action facilitates NF1 binding we investigated its involvement in the localized displacement of histones H2A and H2B dimers, which is observed after hormone induction [6]. In cells transfected with control siRNA, recruitment of PR 30 minutes after hormone addition (Figure 4 D, second row, lanes 1-2) is accompanied by displacement of histone H2A (Figure 4 D, third row, lanes 1-2).

Depletion of Brm and Brg1 diminished PR binding and abolished H2A displacement (Figure 4 D, lanes 3-4; quantification by real time PCR is shown in the right panel). Thus, BAF complexes are required for PR binding and H2A/H2B histone displacement, a requisite for NF1 binding. It is likely that the residual PR bound in the absence of BAF corresponds to receptor interacting with the exposed HRE1 [27].

Discussion

The results of this study contribute to a better understanding of the molecular mechanisms of promoter activation by progesterone. We have shown previously that PR interacts with an exposed HRE on the surface of a nucleosome on the MMTV promoter [27] and recruits Brg1/Brm-containing complexes [6]. Here we demonstrate the need of the hSWI/SNFa complex, known as BAF, for activation of MMTV and other progesterone target genes. We show that BAF is recruited at least in part *via* an interaction between PR and the BAF57 subunit of the complex. In a previous report we have described that acetylation of H3K14 in response to progestins [30]. Now we show that H3K14ac, likely generated by PCAF, anchors the BAF complex, suggesting a mechanism for the cooperation between two types of chromatin remodeling activities. The recruited BAF catalyzes the ATP dependent displacement of histones H2A/H2B needed for NF1 to gain access to the promoter site. Thus synergism between the two transcription factors PR and NF1 is mediated by a cooperation between two chromatin remodeling machines, BAF and PCAF.

BAF interacts with PR and is necessary for induction of progesterone target genes in breast cancer cells

Activation of several hormone sensitive promoters exhibits SWI/SNF-dependence reflecting the requirement of chromatin remodeling [38, 47-49]. There is evidence that glucocorticoid receptor recruits a Brg1-containing complex to promoters via protein-protein interaction with BAF250 [38], whereas androgen receptor (AR) and ERa can directly bind BAF57 [49, 50]. We have described a progesterone dependent recruitment of Brg1 and Brm to the MMTV promoter in breast cancer cells and have shown that yeast SWI/SNF can displace H2A/H2B dimers from MMTV recombinant nucleosomes [6], but the nature of the complex recruited in intact cells and its function in gene activation was not known.

The human SWI/SNF complexes can have either Brm or Brg1 as ATPase subunits [51]. Although the two ATPases exhibit certain differences in their biological

activities [52, 53], they can partly compensate for each other in mouse cells [54]. We found that in T47D-MTVL cells depletion of Brm increases Brg1 levels and viceversa. This could be due to the existence of an excess of both free ATPases in equilibrium with the complex bound forms. If one of the ATPase is depleted the other will be incorporated to a higher extent in the SWI/SNF complex thus becoming more resistant to proteolytic degradation. A similar finding has been reported in mice lacking Brm [55].

The BAF complex is recruited to the MMTV promoter within minutes after progestin treatment, likely through a direct interaction with the activated PR. The fact that PR was unable to form a complex with PBAF indicates that the receptor can discriminate between these two related machineries and promotes BAF recruitment to progestin target promoters. Although BAF57 is present in BAF and PBAF complexes it is possible that hormone-dependent interaction with PR is only possible in the context of the BAF complex. Similar results have been previously reported for AR and ERa [49, 50, 56]. Depletion of BAF250 has a similar effect on MMTV induction as depletion of both Brg1 and Brm, confirming the importance of the BAF complex. BAF dependence is observed with other progesterone responsive genes, such as FOS and MYC, indicating that chromatin remodeling by BAF plays a more general role in progesterone gene regulation. However, as exemplified by the cyclin D1 gene, not all hormone responsive genes required BAF for induction, indicating the existence of alternative pathways for hormonal gene regulation. It is likely that other ATP-dependent remodeling complexes participate in regulation of MMTV and other progesterone regulated genes. In minichromosomes reconstituted in *Drosophila* embryo extracts NURF mediates synergism between recombinant PR and NF1 in a SWI/SNF independent manner [35]. The difference with our present results may reflect the different nature of chromatin in breast cancer cells and Drosophila embryo, such as lack of linker histones and transcriptional activity. We know that hSnf2H is recruited to the MMTV promoter 30 min after progesterone treatment of T47D-MTVL cells but its role in hormonal induction remains to be established [30].

Acetylation on histone H3K14 by PCAF is essential for hormonal induction by anchoring BAF to the promoter

The role of histone acetylation on MMTV activity has been controversial. We have observed that low concentration of HDAC inhibitors leading to moderate increase in acetylation, enhance MMTV transcription [37], whereas high concentrations completely abolish transcription [37, 57]. In previous reports, we have demonstrated the co-

recruitment of PCAF and SRC1 to the MMTV promoter following progestin treatment [6, 30]. The siRNA knockdown results presented here demonstrate that PCAF is one of the main HATs acting on the MMTV promoter, that its levels correlated with the extent of H3K14 acetylation, and that acetylation is necessary and important for progesterone activation of the MMTV promoter.

There have also been controversies over the nature of the histone modifications influencing binding of the SWI/SNF complex [58]. It was first claimed that acetylation of H4K8 recruits the SWI/SNF complex via Brg1, whereas acetylation of H3K9 and H3K14 is only important for the recruitment of TFIID general transcription factors [41]. Two independent recent reports demonstrated that the bromodomain-containing ATPase Brg1 has the highest binding affinity to acetylated H3K14 peptide or doubly acetylated H3K9/K14 peptide, whereas binding to acetylated H4K8 peptide is insignificant [59, 60]. Although H4 acetylation takes place very early after progesterone addition to T47D-MTVL cells (Figure 4 B), we found here no binding of the BAF complex to acetylated H4 peptides (Figure 4 A). In contrast, we show that components of the BAF complex bind to peptides containing acetylated H3K14, alone or in combination with other histone modifications. Furthermore, we quantify the interaction of components of BAF and PBAF to H3 acetylated peptides by SILAC. Our results indicate that compared to H3K9ac, H3K14ac is the major contributor to anchoring of these chromatin remodeling complexes to nucleosomes. Several subunits of BAF and PBAF contain bromodomains that can bind acetylated residues. Obvious candidates are Brg1 and Brm, but other subunits, such as BAF180, could cooperate on the recognition of H3K14ac by PBAF. Despite the fact that both BAF and PBAF can bind acetylated K14, only BAF is recruited to the MMTV promoter, indicating that H3K14ac participates in anchoring or retention of the recruited complex but is not sufficient for recruitment.

The colocalization of BAF and H3K14ac (Figure 3 F) could reflect a general role of this histone mark to BAF anchoring in chromatin, but other explanations are possible. For instance, a common transcription factor could recruit BAF and HATs to the same chromatin regions. Alternatively, components of the BAF complex could be important for HATs recruitment. A combination of these hypotheses is also possible.

BAF mediates nucleosome remodeling leading to H2A/H2B displacement that facilitates NF1 binding

Following hormone treatment, PR is phosphorylated, forms a complex with activated Erk and Msk1 and this ternary complex binds the exposed HRE1 on the surface of the MMTV promoter nucleosome leading to modification of the H3 tail and displacement of a repressive complex [30]. PCAF is also recruited to the promoter and acetylation of H3K14 is observed 5 min after hormone treatment [30]. The activated ternary complex pPR-pErk-pMsk1 recruits BAF that initiates nucleosome remodeling, since inhibiting Erk or Msk1 activation blocks BAF recruitment and the subsequent steps. We have not addressed the initial kinetics of acetylation and BAF recruitment in sufficient detail and cannot propose a precise order of events. Moreover, we cannot exclude that PCAF interacts with BAF and that both complexes are recruited together.

In T47D-MTVL cells the BAF mediated chromatin remodeling event is a localized displacement of dimers of H2A/H2B from promoter nucleosome B, a catalytic activity exhibited by the purified yeast SWI/SNF complex on recombinant MMTV nucleosomes [6]. An earlier study with recombinant nucleosomes and SWI/SNF complexes from yeast suggest that both BAF and PBAF complexes would have the capacity to displace histones H2A/H2B as they share the relevant subunit, Swi3p [36]. This remodeling event is a prerequisite for enabling access of NF1 to its binding site in the MMTV promoter following hormone treatment, as demonstrated by the lack of H2A displacement and NF1 binding in BAF-depleted cells. We have previously shown that NF1 binding stabilizes the open conformation of the nucleosome allowing binding of further PR molecules to the internal HREs and subsequent activation of transcription [35]. A model representing our present view of the initial steps in progesterone activation of the MMTV promoter is shown in Figure 5. The model underlines the feed-forward cyclical nature of the activation process and explains why interfering with any of the initial steps has consequences for binding of all factors involved, PR, NF1, BAF and PCAF.

Finally, it is worth noting that we have not mentioned in these studies the possible role of linker histones. Our previous results [61] and those of other groups [62-64], suggest that changes histone H1 stoichiometry and phosphorylation by CyclinA/Cdk2 take place at different time points during the hormonal induction and are important for transcriptional activation. Future studies will be required to clarify the relationship of these changes to those reported in this study.

Materials and methods

Cell Culture and hormone treatments

T47D-MTVL breast cancer cells carrying one stably integrated copy of the luciferase reporter gene driven by the MMTV promoter [27] were routinely grown in RPMI 1640 medium supplemented with 10% FBS, 2 mM L-glutamine, 100 U/ml penicillin and 100 μg/ml streptomycin. For the experiments, cells were plated in RPMI medium without phenol red supplemented with 10% dextran-coated charcoal treated FBS (DCC/FBS) and 48 h later medium was replaced by fresh medium without serum. After 24h in serum-free conditions, cells were incubated with R5020 (10 nM) or vehicle (ethanol) for different times at 37ºC.

Immunofluorescence Assays

Cells were grown on coverslips and treated as described above. Cells were then fixed with 4% paraformaldehyde in PBS and permeabilized with PBS 0.2% Triton X-100 at room temperature. Coverslips were blocked with 5% BSA at room temperature and incubated with anti-BAF57 (Bethyl) and pPR (Abcam) antibodies. After washes with PBS containing 0.05% Tween 20, samples were incubated with appropriate secondary antibodies. After washes with PBS Tween 0.05% and DNA staining with DAPI, samples were mounted with mowiol.

Transient transfections

Cells were cultured into 6-well plates at a density of 4 X 10^5 cells/well and treated as indicated above. Transient transfections were performed by using Lipofectamine 2000 (Invitrogen). cDNAs expressing PCAF and its deacetylase defective mutant (PCAF ΔHAT) were kindly provided by Tony Kouzarides (Cambridge, UK).

Chromatin Immunoprecipitation (ChIP) and re-ChIP assays in cultured cells

ChIP assays were performed as described [65] using the NF1 specific antibody (gift from Dr Naoko Tanese), the H2A antibody (gift from Stefan Dimitrov), anti-BAF180 (gift from Dr Weidong Wang), anti-PR (H190) and anti-Brg1 (H88), both from Santa Cruz, anti-SMARCA2/BRM (ab15597) from Abcam, anti-BAF57/SMARCE1 from Bethyl and anti-acetyl(Lys14)-Histone H3 and anti-BAF250 from Upstate. Quantification of chromatin immunoprecipitation was performed by real time PCR using Roche Lightcycler (Roche). The fold enrichment of target sequence in the immunoprecipitated (IP) compared to input (Ref) fractions was calculated using the comparative Ct (the number of cycles required to reach a threshold concentration) method with the equation

$2^{Ct(IP)-Ct(Ref).}$ Each of these values were corrected by the human b-globin gene and referred as relative abundance over time zero. Primers sequences are available on request. For re-ChIP assays, immunoprecipitations were sequentially washed as previously described [66]. Complexes were eluted with 10 mM DTT at 37 °C for 30 min, diluted 50 times with dilution buffer, and immunoprecipitated with the indicated antibodies. The antibodies used for ChIPs assays are listed in Supplementary Fig.5.

RNA interference experiments

All siRNAs were transfected into the T47D-MTVL cells using Lipofectamine 2000 (Invitrogen). After 48 h the medium was replaced by fresh medium without serum. After one day in serum-free conditions, cells were incubated with R5020 (10 nM) or vehicle (ethanol) for different times at 37ºC. The down-regulation of Brg1, Brm, NF1, BAF57, BAF250, PCAF and SRC1 expression was determined by Western blotting. The siRNAs used are listed in Supplementary Fig.5.

RNA extraction and RT-PCR

Total RNA was prepared and cDNA generated as previously described [30]. Quantification of LUC and GAPDH gene products was performed by real time PCR. Each value calculated using the standard curve method was corrected by the human GAPDH and expressed as relative RNA abundance over time zero. Primer sequences are available on request.

Coimmunoprecipitation assay

Cells were lysed and cell extracts (2 mg) were incubated with protein G/A agarose beads previously coupled with 6μg of the corresponding antibodies or an unspecific control antibody. The immunoprecipitated proteins (IP) were eluted by boiling in SDS sample buffer. Inputs and IPs were analyzed by western blot using Brg1, Brm, BAF155, BAF170 and H3K14ac specific antibodies.

Peptide Pull Down Assays

Nuclear extracts from T47DMTVL breast cancer cells were prepared as described [67]. Peptide pull down assays were performed as described previously [68], with the exception of using 100 μg of nuclear extract during incubation of peptide-bound beads. Synthetic biotinylated H3 and H4 peptides were either purchased from Upstate or were kind gifts from M. Vermeulen. For Western immunoblotting, antibodies against BRG1,

BAF170, BAF155, PR (Santa Cruz), BRM (Abcam), BAF57, SNF5 (Bethyl) and H1 (Upstate 05-457) were used.

Mass spectrometry of proteins

After trypsin digestion of gel slices peptides were extracted, desalted using stage tips [69] and analyzed using a nano-HPLC Agilent 1100 nanoflow system connected online to an LTQ-Orbitrap mass spectrometer (Thermo Fisher, Bremen). The mass spectrometer was operated in the data-dependent mode to automatically switch between MS and MS^2. The instrument was operated with 'lock mass option' as recently described [70]. Survey spectra were acquired with 60,000 resolution in the orbitrap, while acquiring up to five tandem mass spectra in the LTQ part of the instrument. The raw data files were analyzed with an in-house developed quantitative proteomics software MaxQuant, version 1.0.12.5 [71], in combination with the Mascot search engine (Matrix Science). The data was searched against a decoy human IPI database 3.37 including common contaminants. False discovery rates, both at the peptide and protein level were set to 1%. Minimal peptide length was set to 6 amino acids. False positive rates for peptides are calculated as described in [72].

Acknowledgements

We wish to thank C. Peterson, Worcester, USA, for purified SWI/SNF; Naoko Tanese, NYU, New York, USA, for anti NF1 antibody, Stefan Dimitrov, CNR Grenoble for histone H2A antibodies and Borja Belandia, IIB, Madrid, Spain, for BAF57 constructs. G.P.V. was a recipient of a fellowship of the Ramón y Cajal Programme. M.V. was supported by the Dutch Cancer Society (KWF/NKB). The experimental work was supported by grants from the Departament d´Innovació Universitat i Empresa (DIUiE), Ministerio de Educación y Ciencia (MEC) BMC 2003-02902, Consolider (CSD2006-00049) and Fondo de Investigación Sanitaria (FIS) PI0411605 and CP04/00087, and EU IP HEROIC.

References

1. Luger K, Mader AW, Richmond RK, Sargent DF and Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 389: 251-260.

2. Aalfs JD and Kingston RE (2000) What does 'chromatin remodeling' mean? Trends Biochem Sci 25: 548-555.

3. Whitehouse I, Rando OJ, Delrow J and Tsukiyama T (2007) Chromatin remodelling at promoters suppresses antisense transcription. Nature 450: 1031-1035.

4. Mohrmann L and Verrijzer CP (2005) Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes. Biochim Biophys Acta 1681: 59-73.

5. Bruno M, Flaus A, Stockdale C, Rencurel C, Ferreira H, et al. (2003) Histone H2A/H2B dimer exchange by ATP-dependent chromatin remodeling activities.
Mol Cell 12: 1599-1606.
6. Vicent GP, Nacht AS, Smith CL, Peterson CL, Dimitrov S, et al. (2004) DNA instructed displacement of histones H2A and H2B at an inducible promoter. Mol Cell 16: 439-452.

7. Cairns BR, Schlichter A, Erdjument-Bromage H, Tempst P, Kornberg RD, et al. (1999) Two functionally distinct forms of the RSC nucleosome-remodeling complex, containing essential AT hook, BAH, and bromodomains. Mol Cell 4: 715-723.

8. Wang W (2003) The SWI/SNF family of ATP-dependent chromatin remodelers: similar mechanisms for diverse functions. Curr Top Microbiol Immunol 274: 143-169.

9. Xue Y, Canman JC, Lee CS, Nie Z, Yang D, et al. (2000) The human SWI/SNF-B chromatin-remodeling complex is related to yeast rsc and localizes at kinetochores of mitotic chromosomes. Proc Natl Acad Sci U S A 97: 13015-13020.

10. Yan Z, Cui K, Murray DM, Ling C, Xue Y, et al. (2005) PBAF chromatin-remodeling complex requires a novel specificity subunit, BAF200, to regulate expression of selective interferon-responsive genes. Genes Dev 19: 1662-1667.

11. Cheung P, Allis CD and Sassone-Corsi P (2000) Signaling to chromatin through histone modifications. Cell 103: 263-271.

12. Narlikar GJ, Fan HY and Kingston RE (2002) Cooperation between complexes that regulate chromatin structure and transcription. Cell 108: 475-487.

13. Carrozza MJ, Utley RT, Workman JL and Cote J (2003) The diverse functions of histone acetyltransferase complexes. Trends Genet 19: 321-329.

14. Roth SY, Denu JM and Allis CD (2001) Histone acetyltransferases. Annu Rev Biochem 70: 81-120.

15. Yang XJ, Ogryzko VV, Nishikawa J, Howard BH and Nakatani Y (1996) A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. Nature 382: 319-324.

16. Schiltz RL, Mizzen CA, Vassilev A, Cook RG, Allis CD, et al. (1999) Overlapping but distinct patterns of histone acetylation by the human coactivators p300 and PCAF within nucleosomal substrates. J Biol Chem 274: 1189-1192.

17. Blanco JC, Minucci S, Lu J, Yang XJ, Walker KK, et al. (1998) The histone acetylase PCAF is a nuclear receptor coactivator. Genes Dev 12: 1638-1651.

18. Korzus E, Torchia J, Rose DW, Xu L, Kurokawa R, et al. (1998) Transcription factor-specific requirements for coactivators and their acetyltransferase functions. Science 279: 703-707.

19. Roberts SM and Winston F (1997) Essential functional interactions of SAGA, a Saccharomyces cerevisiae complex of Spt, Ada, and Gcn5 proteins, with the Snf/Swi and Srb/mediator complexes. Genetics 147: 451-465.

20. Kasten M, Szerlong H, Erdjument-Bromage H, Tempst P, Werner M, et al. (2004) Tandem bromodomains in the chromatin remodeler RSC recognize acetylated histone H3 Lys14. Embo J 23: 1348-1359.

21. Hassan AH, Neely KE and Workman JL (2001) Histone acetyltransferase complexes stabilize swi/snf binding to promoter nucleosomes. Cell 104: 817-827.

22. Huang ZQ, Li J, Sachs LM, Cole PA and Wong J (2003) A role for cofactor-cofactor and cofactor-histone interactions in targeting p300, SWI/SNF and Mediator for transcription. Embo J 22: 2146-2155.

23. Beato M, Herrlich P and Schutz G (1995) Steroid hormone receptors: many actors in search of a plot. Cell 83: 851-857.

24. Metivier R, Reid G and Gannon F (2006) Transcription in four dimensions: nuclear receptor-directed initiation of gene expression. EMBO Rep 7: 161-167.

25. Vicent GP, Ballare C, Zaurin R, Saragueta P and Beato M (2006) Chromatin remodeling and control of cell proliferation by progestins via cross talk of progesterone receptor with the estrogen receptors and kinase signaling pathways. Ann N Y Acad Sci 1089: 59-72.

26. Richard-Foy H and Hager GL (1987) Sequence-specific positioning of nucleosomes over the steroid-inducible MMTV promoter. Embo J 6: 2321-2328.

27. Truss M, Bartsch J, Schelbert A, Hache RJ and Beato M (1995) Hormone induces binding of receptors and transcription factors to a rearranged nucleosome on the MMTV promoter in vivo. Embo J 14: 1737-1751.

28. Pina B, Bruggemeier U and Beato M (1990) Nucleosome positioning modulates accessibility of regulatory proteins to the mouse mammary tumor virus promoter. Cell 60: 719-731.

29. Migliaccio A, Piccolo D, Castoria G, Di Domenico M, Bilancio A, et al. (1998) Activation of the Src/p21ras/Erk pathway by progesterone receptor via cross-talk with estrogen receptor. Embo J 17: 2008-2018.

30. Vicent GP, Ballare C, Nacht AS, Clausell J, Subtil-Rodriguez A, et al. (2006) Induction of progesterone target genes requires activation of Erk and Msk kinases and phosphorylation of histone H3. Mol Cell 24: 367-381.

31. Bhattacharjee RN and Archer TK (2006) Transcriptional silencing of the mouse mammary tumor virus promoter through chromatin remodeling is concomitant with histone H1 phosphorylation and histone H3 hyperphosphorylation at M phase. Virology 346: 1-6.

32. Hebbar PB and Archer TK (2003) Nuclear factor 1 is required for both hormone-dependent chromatin remodeling and transcriptional activation of the mouse mammary tumor virus promoter. Mol Cell Biol 23: 887-898.

33. Muchardt C and Yaniv M (1993) A human homologue of Saccharomyces cerevisiae SNF2/SWI2 and Drosophila brm genes potentiates transcriptional activation by the glucocorticoid receptor. Embo J 12: 4279-4290.

34. Venditti P, Di Croce L, Kauer M, Blank T, Becker PB, et al. (1998) Assembly of MMTV promoter minichromosomes with positioned nucleosomes precludes NF1 access but not restriction enzyme cleavage. Nucleic Acids Res 26: 3657-3666.

35. Di Croce L, Koop R, Venditti P, Westphal HM, Nightingale KP, et al. (1999) Two-step synergism between the progesterone receptor and the DNA-binding domain of nuclear factor 1 on MMTV minichromosomes. Mol Cell 4: 45-54.

36. Yang X, Zaurin R, Beato M and Peterson CL (2007) Swi3p controls SWI/SNF assembly and ATP-dependent H2A-H2B displacement. Nat Struct Mol Biol 14: 540-547.

37. Bartsch J, Truss M, Bode J and Beato M (1996) Moderate increase in histone acetylation activates the mouse mammary tumor virus promoter and remodels its nucleosome structure. Proc Natl Acad Sci U S A 93: 10741-10746.

38. Nie Z, Xue Y, Yang D, Zhou S, Deroo BJ, et al. (2000) A specificity and targeting subunit of a human SWI/SNF family-related chromatin-remodeling complex. Mol Cell Biol 20: 8879-8888.

39. Sif S (2004) ATP-dependent nucleosome remodeling complexes: enzymes tailored to deal with chromatin. J Cell Biochem 91: 1087-1098.

40. Hassan AH, Prochasson P, Neely KE, Galasinski SC, Chandy M, et al. (2002) Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. Cell 111: 369-379.

41. Agalioti T, Chen G and Thanos D (2002) Deciphering the transcriptional histone acetylation code for a human gene. Cell 111: 381-392.

42. Kouskouti A and Talianidis I (2005) Histone modifications defining active genes persist after transcriptional and mitotic inactivation. Embo J 24: 347-357.

43. Schulze WX and Mann M (2004) A novel proteomic screen for peptide-protein interactions. J Biol Chem 279: 10756-10764.
44. Gronostajski RM (2000) Roles of the NFI/CTF gene family in transcription and development. Gene 249: 31-45.

45. Kane R, Finlay D, Lamb T and Martin F (2000) Transcription factor NF 1 expression in involuting mammary gland. Adv Exp Med Biol 480: 117-122.

46. Hebbar PB and Archer TK (2007) Chromatin-dependent cooperativity between site-specific transcription factors in vivo. J Biol Chem 282: 8284-8291.

47. Marshall TW, Link KA, Petre-Draviam CE and Knudsen KE (2003) Differential requirement of SWI/SNF for androgen receptor activity. J Biol Chem 278: 30605-30613.

48. Fryer CJ and Archer TK (1998) Chromatin remodelling by the glucocorticoid receptor requires the BRG1 complex. Nature 393: 88-91.

49. Belandia B, Orford RL, Hurst HC and Parker MG (2002) Targeting of SWI/SNF chromatin remodelling complexes to estrogen-responsive genes. Embo J 21: 4094-4103.

50. Link KA, Burd CJ, Williams E, Marshall T, Rosson G, et al. (2005) BAF57 governs androgen receptor action and androgen-dependent proliferation through SWI/SNF. Mol Cell Biol 25: 2200-2215.

51. Wang W, Xue Y, Zhou S, Kuo A, Cairns BR, et al. (1996) Diversity and specialization of mammalian SWI/SNF complexes. Genes Dev 10: 2117-2130.

52. Mizutani T, Ito T, Nishina M, Yamamichi N, Watanabe A, et al. (2002) Maintenance of integrated proviral gene expression requires Brm, a catalytic subunit of SWI/SNF complex. J Biol Chem 277: 15859-15864.

53. Wang F, Zhang R, Beischlag TV, Muchardt C, Yaniv M, et al. (2004) Roles of Brahma and Brahma/SWI2-related gene 1 in hypoxic induction of the erythropoietin gene. J Biol Chem 279: 46733-46741.

54. Reyes JC, Barra J, Muchardt C, Camus A, Babinet C, et al. (1998) Altered control of cellular proliferation in the absence of mammalian brahma (SNF2alpha). Embo J 17: 6979-6991.

55. Peterson CL and Herskowitz I (1992) Characterization of the yeast SWI1, SWI2, and SWI3 genes, which encode a global activator of transcription. Cell 68: 573-583.

56. Garcia-Pedrero JM, Kiskinis E, Parker MG and Belandia B (2006) The SWI/SNF chromatin remodeling subunit BAF57 is a critical regulator of estrogen receptor function in breast cancer cells. J Biol Chem 281: 22656-22664.

57. Mulholland NM, Soeth E and Smith CL (2003) Inhibition of MMTV transcription by HDAC inhibitors occurs independent of changes in chromatin remodeling and increased histone acetylation. Oncogene 22: 4807-4818.

58. Strahl BD and Allis CD (2000) The language of covalent histone modifications. Nature 403: 41-45.

59. Shen W, Xu C, Huang W, Zhang J, Carlson JE, et al. (2007) Solution structure of human Brg1 bromodomain and its specific binding to acetylated histone tails. Biochemistry 46: 2100-2110.

60. Singh M, Popowicz GM, Krajewski M and Holak TA (2007) Structural ramification for acetyl-lysine recognition by the bromodomain of human BRG1 protein, a central ATPase of the SWI/SNF remodeling complex. Chembiochem 8: 1308-1316.

61. Koop R, Di Croce L and Beato M (2003) Histone H1 enhances synergistic activation of the MMTV promoter in chromatin. Embo J 22: 588-599.

62. Bhattacharjee RN, Banks GC, Trotter KW, Lee HL and Archer TK (2001) Histone H1 phosphorylation by Cdk2 selectively modulates mouse mammary tumor virus transcription through chromatin remodeling. Mol Cell Biol 21: 5417-5425.

63. Narayanan R, Adigun AA, Edwards DP and Weigel NL (2005) Cyclin-dependent kinase activity is required for progesterone receptor function: novel role for cyclin A/Cdk2 as a progesterone receptor coactivator. Mol Cell Biol 25: 264-277.

64. Bresnick EH, Bustin M, Marsaud V, Richard-Foy H and Hager GL (1992) The transcriptionally-active MMTV promoter is depleted of histone H1. Nucleic Acids Res 20: 273-278.

65. Strutt H and Paro R (1999) Mapping DNA target sites of chromatin proteins in vivo by formaldehyde crosslinking. Methods Mol Biol 119: 455-467.

66. Shang Y, Myers M and Brown M (2002) Formation of the androgen receptor transcription complex. Mol Cell 9: 601-610.

67. Dignam JD, Lebovitz RM and Roeder RG (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res 11: 1475-1489.

68. Vermeulen M, Mulder KW, Denissov S, Pijnappel WW, van Schaik FM, et al. (2007) Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. Cell 131: 58-69.

69. Rappsilber J, Ishihama Y and Mann M (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. Anal Chem 75: 663-670.

70. Olsen JV, de Godoy LM, Li G, Macek B, Mortensen P, et al. (2005) Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. Mol Cell Proteomics 4: 2010-2021.

71. Cox J and Mann M (2007) Is proteomics the new genomics? Cell 130: 395-398.

72. Nesvizhskii AI, Vitek O and Aebersold R (2007) Analysis and validation of proteomic data generated by tandem mass spectrometry. Nat Methods 4: 787-797.

Figure Legends

Figure 1. BAF is essential for MMTV promoter activity in T47D-MTVL cells

Panel a. Cells were transfected either with siRNA against Brg1, Brm , or both siRNAs combined as indicated. After 48 h the medium was replaced by fresh medium without serum. After one day in serum-free conditions, cells were lysed and the levels of Brm, Brg1, PR and tubulin were determined by Western blotting. C, control siRNA.

Panel b. Cells were transfected with Control, Brg1 and Brm siRNAs as described in (A). After one day in serum-free conditions, cells were incubated with 10 nM R5020 for 8 hs and total RNA was prepared, cDNA was generated and used as template for real time PCR with luciferase oligonucleotides. The values represent the mean and standard deviation from 3 experiments performed in duplicate.

Panel c. Cells were transfected with Control, Brg1 and Brm siRNAs as described in (A) and treated with 10 nM R5020. cDNA was generated and used as template for real time PCR using specific c-Fos, c-Myc, Cyclin D1 and GAPDH primers. The values represent the mean ± SD of 3 experiments performed in duplicate.

Panel d. Cells were transfected with Control and BAF57 siRNAs as described in (A) and treated with 10 nM R5020. Left: BAF57 levels were analyzed by Western blotting. Right: RNA was extracted, cDNA was generated and used as template for real time PCR with luciferase oligonucleotides. The values represent the mean ± SD of 2 experiments performed in duplicate.

Panel e. Experiments similar to those shown in (d), but cells were transfected with Control and BAF250 siRNAs as described in (a) and (d). The values represent the mean ± SD of 2 experiments performed in duplicate.

Panel f. Cells were lysed and immunoprecipitated either with a-BAF57 antibody or with normal rabbit IgG as a negative control (IgG). The immunoprecipitates (IP) were analyzed by western blotting with non-discriminating Brg1/Brm antibody, α-BAF155 or α-BAF170 specific antibodies.

Panel g. Cells were untreated (0) or treated for 5, 30 or 60 min with 10 nM R-5020 and subjected to ChIP assays with α-PR, α-Brg1, α-Brm, α-BAF57, α-BAF250, α-BAF180 and α-BAF200 specific antibodies. The precipitated DNA fragments were subjected to PCR analysis to test for the presence of sequences corresponding to the MMTV nucleosome B. Input material (1%) is shown for comparison. A representative of three independent experiments is shown.

Figure 2. PR forms a complex with BAF in T47D-MTVL cells and binds BAF57 *in vitro*

Panel a. Cells were untreated (-) or treated for 30 min with R5020, lysed and immunoprecipitated either with a-PR antibody or with normal rabbit IgG as a negative control (IgG). Inputs and IPs were analyzed by western blot using α-BAF57 and α-PR, as indicated.

Panel b. Cells were treated as in A. Inputs and IPs were analyzed by western blot using α-BAF180, α-BAF250 and α-PR, as indicated.

Panel c. Left: Scheme of wild type GST-BAF57 or GST-BAF57 deletion mutants used for pulldown assays. Right: Coomassie-stained gel with the level of expression of the BAF57-GST fusion proteins.

Panel d. Recombinant PR and either wild type GST-BAF57 or GST-BAF57 deletion mutants were used for pulldown assays. Binding reactions were incubated with glutathione sepharose beads and eluted by boiling in SDS sample buffer. Input and bound material were analyzed by western blot with α-PR.

Figure 3. Acetylation on histone H3K14 by PCAF is essential for hormonal transactivation in T47D-MTVL cells

Panel a. Left: Cells were transfected with the indicated siRNAs and the levels of PCAF, SRC1, and tubulin were determined by Western blotting. The asterisks indicate inespecific bands. Right: Cells were transfected with the indicated siRNAs and treated with 10 nM R5020 for 6 h. RNA was extracted, cDNA was generated and used as template for real time PCR with luciferase oligonucleotides. The values represent the mean and standard deviation from 3 experiments performed in duplicate.

Panel b. Cells were transfected with siRNAs, treated with 10 nM R5020 as indicated and the levels of c-Fos and c-Myc mRNAs were analyzed by RT-PCR. The values represent the mean and standard deviation from 2 experiments performed in duplicate.

Panel c. Cells were transfected with PCAF or PCAF HAT mutant (PCAFΔHAT), treated with hormone as indicated and transcription from the MMTV promoter was determined as in (B). An empty vector was used as control. The values represent the mean and standard deviation from 2 experiments performed in duplicate.

Panel d. Left: Cells transfected with the siRNAs and treated with hormone as indicated were subjected to ChIP assays with α-BAF57, α-BAF250 and α-K14acH3. The precipitated DNA fragments were subjected to PCR analysis to test for the presence of sequences corresponding to the MMTV nucleosome B. A representative of three independent experiments is shown Right: quantification of the results by real time PCR from two experiments performed induplicate.

Panel e. Cells were treated with R5020 as indicated and subjected to re-ChIP assays with α-H3K14ac and IgG (first IP) followed by α-BAF250 for the second IP. Precipiated DNA was analysed by PCR for the presence of sequences corresponding to the MMTV, c-Fos and c-Myc PR binding regions.

Panel f. Cells were treated for 30 min with R5020, lysed and chromatin was immunoprecipitated either with a-H3K14ac or with normal rabbit IgG as a negative control (IgG). IPs were analyzed by western blot using a-BAF155 and a-H3K14ac.

Figure 4. The BAF complex binds preferentially to histone H3K14ac and promotes H2A displacement.

Panel a. Nuclear extracts derived from T47D-MTVL cells were used for pulldown experiments with the indicated H3 and H4 tail peptides coupled to beads. Immunoblotting was performed for the presence of components of the BAF and PBAF complex, histone H1 and PR.

Panel b. Cells were untreated (0) or treated for 5 and 30 min with 10 nM R-5020 and subjected to ChIP assays with α-H4K8ac and α-H3K9ac specific antibodies. The precipitated DNA fragments were subjected to PCR analysis to test for the presence of sequences corresponding to the MMTV nucleosome B. Input material (1%) is shown for comparison. A representative of three independent experiments is shown.

Panel c. Quantification of the interaction of selected human SWI/SNF subunits with the indicated H3 tail acetylated peptides. SILAC ratios represent the relative abundance of the 'heavy' (modified) to the 'light' (unmodified) peptide. The variation (in %) and the number of quantified peptides are indicated.

Panel d. Left: T47D-MTVL cells were transfected with siRNAs and treated with hormone as indicated, subjected to ChIP assays using α-NF1, α-PR and α-H2A and primers for MMTV nucleosome B. Right: quantification of the results by real time PCR. A representative of three independent experiments is shown.

Figure 5. Model for the initial steps of MMTV promoter activation

Before hormone addition the MMTV promoter is silent and associated with a repressive complex that includes HP1g (step 1). After hormone addition the activated complex of pPR-pErk-pMsk, as well as PCAF and BAF, are recruited to the MMTV promoter. For simplicity PR is shown as a monomer, though the active form is a homodimer. Msk and PCAF phosphoacetylates H3 leading to H3S10phK14ac (step 2). This modification displaces the repressive complex and anchors the BAF complex, enabling ATP-dependent H2A/H2B displacement (step 3). The nucleosome opening facilitates NF1 binding generating a stable platform that exposes previously hidden HREs for the recruitment of additional PR and BAF complexes, coactivator and eventually promoter activation (step 4). Depletion of BAF prevents progression of the activation process to

step 2: no histone displacement is observed, NF1 cannot bind and consequently less PR/BAF complexes are bound to the promoter. Depletion of PCAF has a similar effect, most likely by labilization of BAF binding, blocking the activation process at step 2.

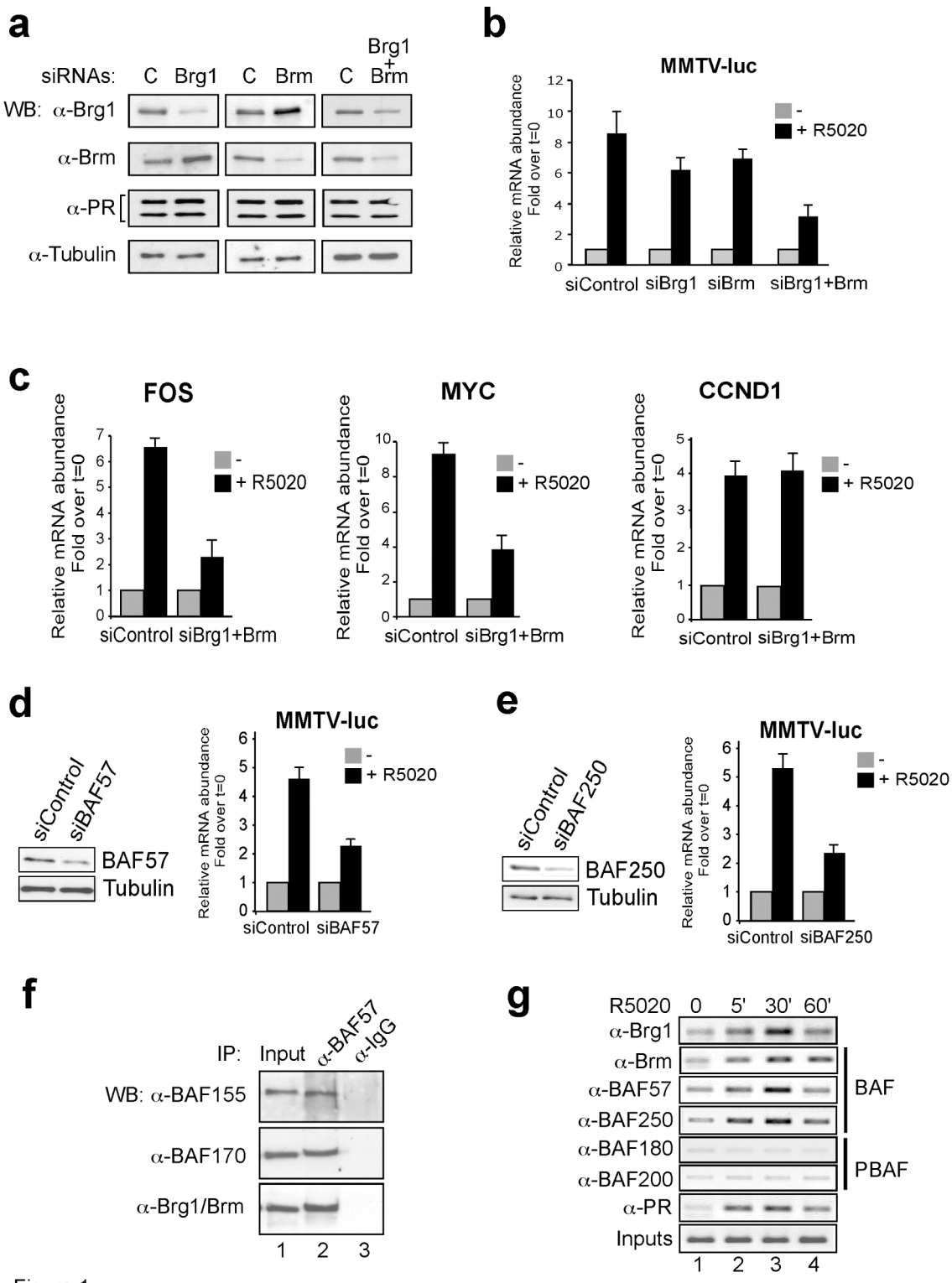


Figure 1
Vicent et al

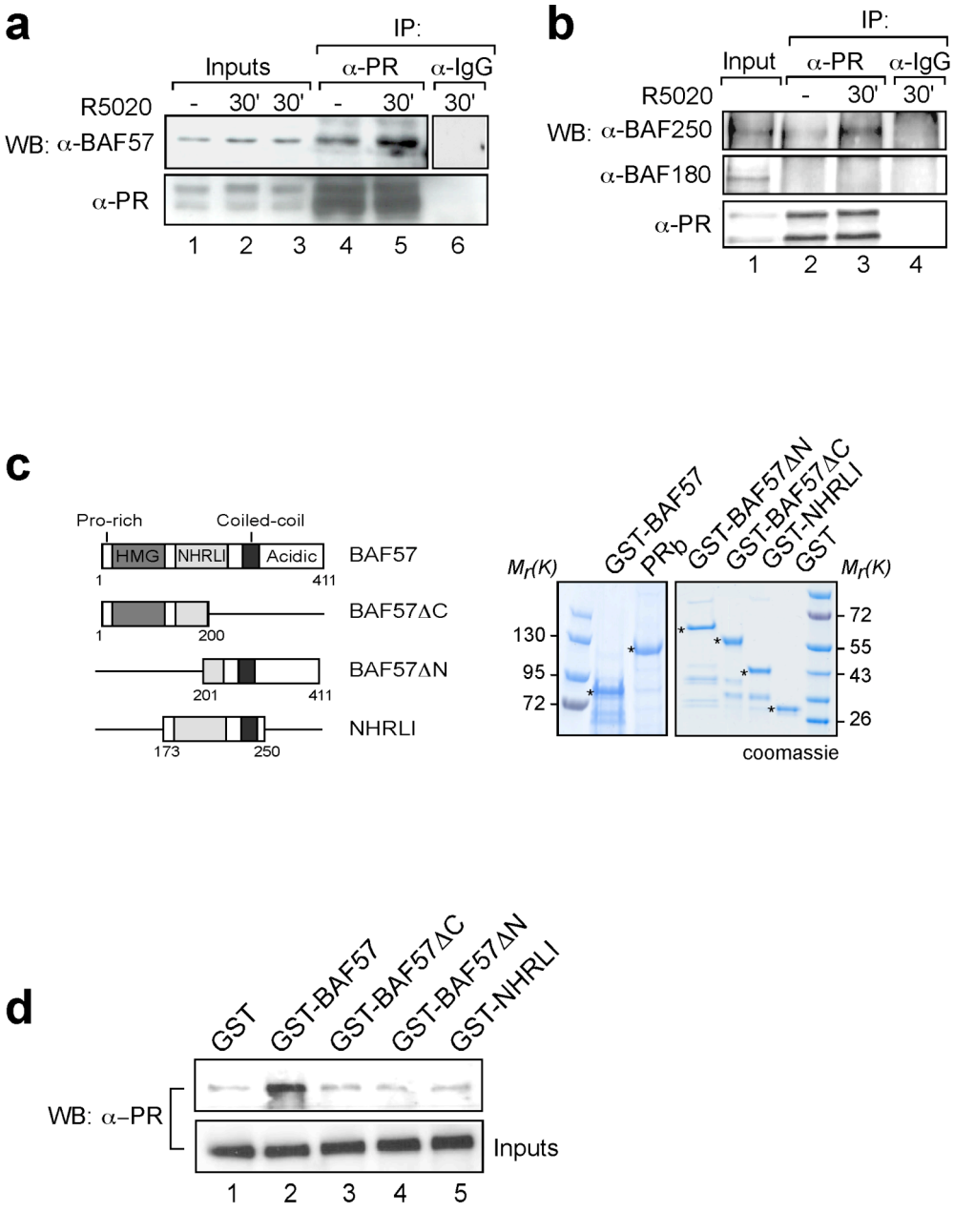


Figure 2
Vicent et al

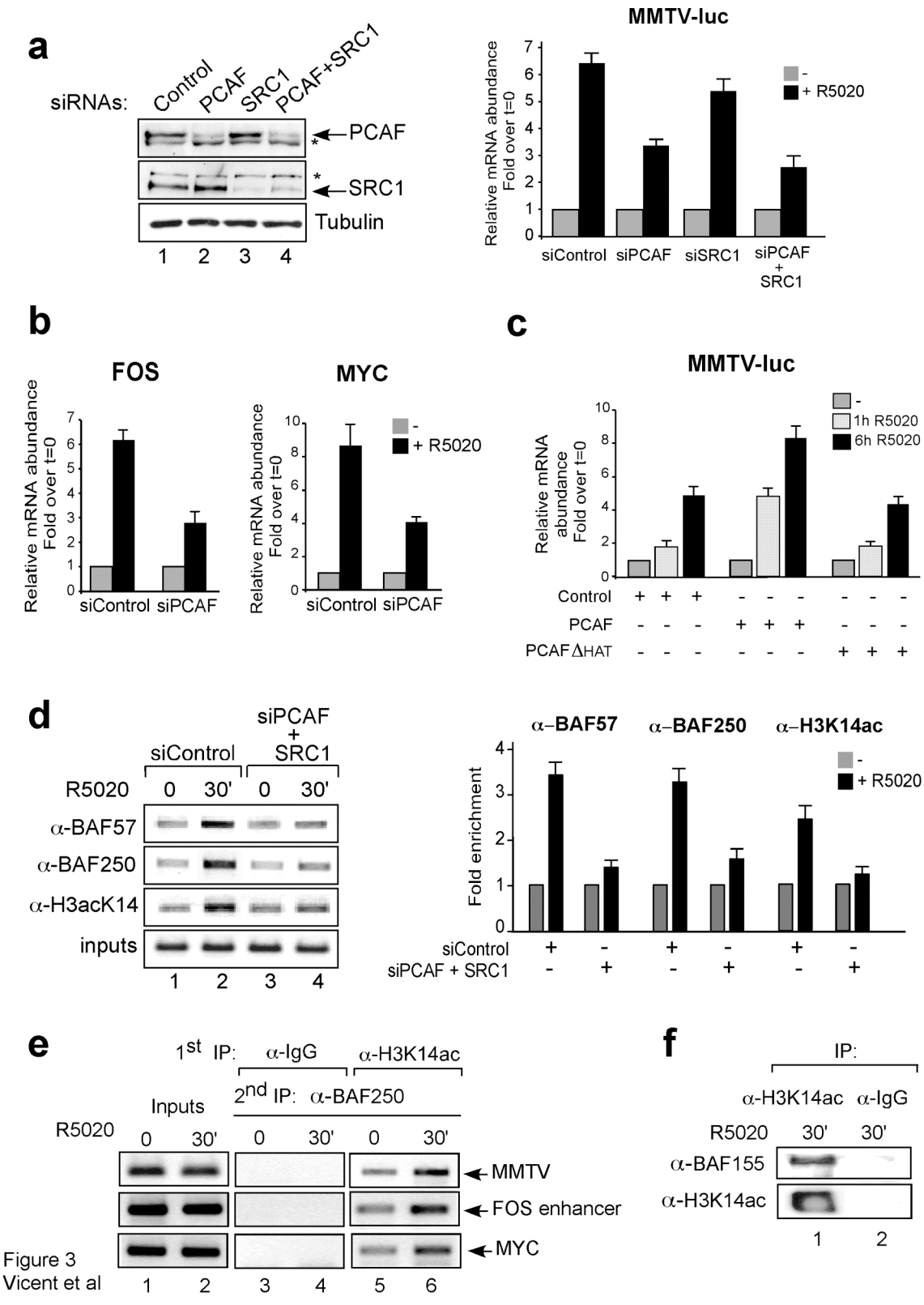
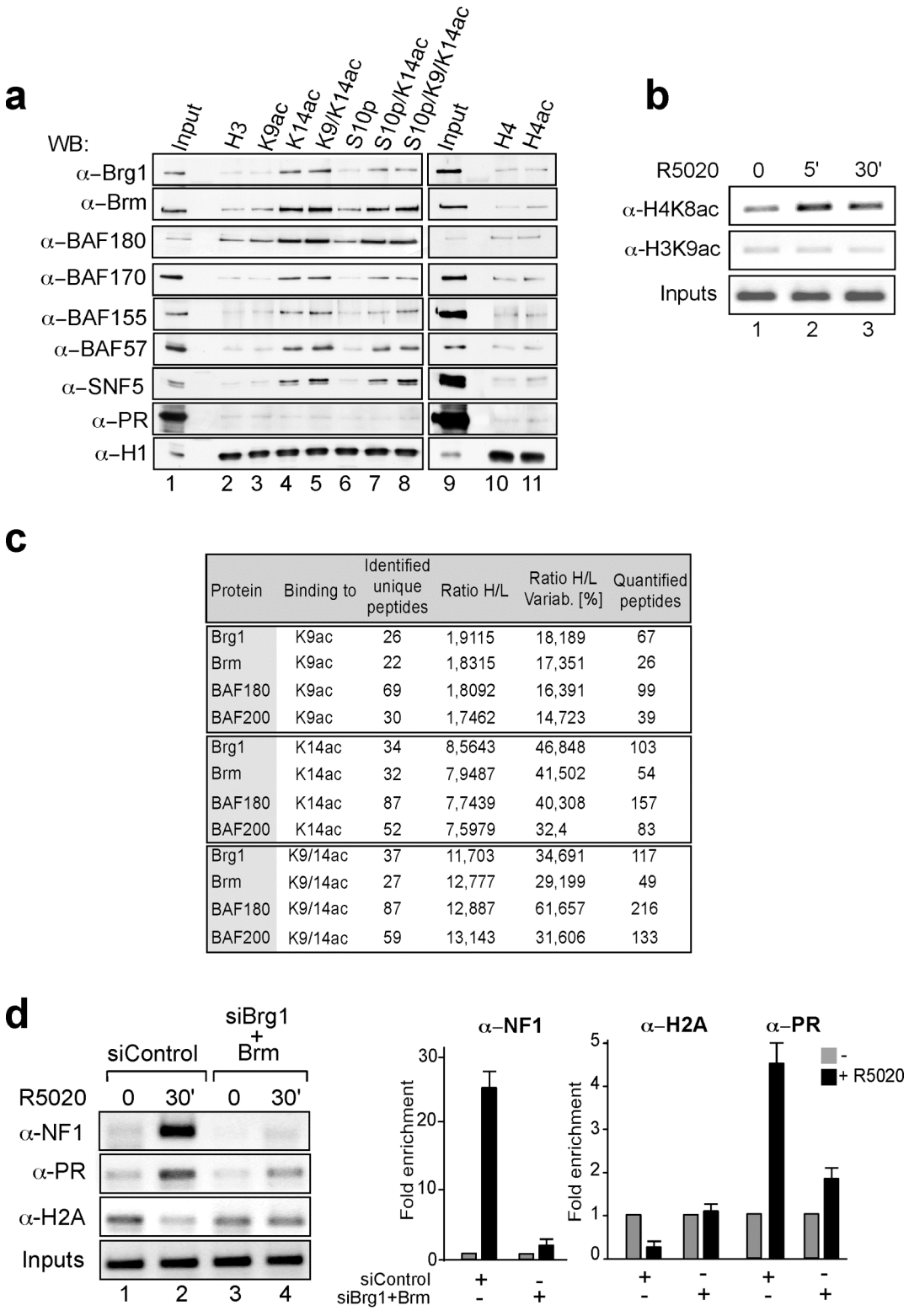


Figure 3
Vicent et al



Protein	Binding to	Identified unique peptides	Ratio H/L	Ratio H/L Variab. [%]	Quantified peptides
Brg1	K9ac	26	1,9115	18,189	67
Brm	K9ac	22	1,8315	17,351	26
BAF180	K9ac	69	1,8092	16,391	99
BAF200	K9ac	30	1,7462	14,723	39
Brg1	K14ac	34	8,5643	46,848	103
Brm	K14ac	32	7,9487	41,502	54
BAF180	K14ac	87	7,7439	40,308	157
BAF200	K14ac	52	7,5979	32,4	83
Brg1	K9/14ac	37	11,703	34,691	117
Brm	K9/14ac	27	12,777	29,199	49
BAF180	K9/14ac	87	12,887	61,657	216
BAF200	K9/14ac	59	13,143	31,606	133

Figure 4
Vicent et al

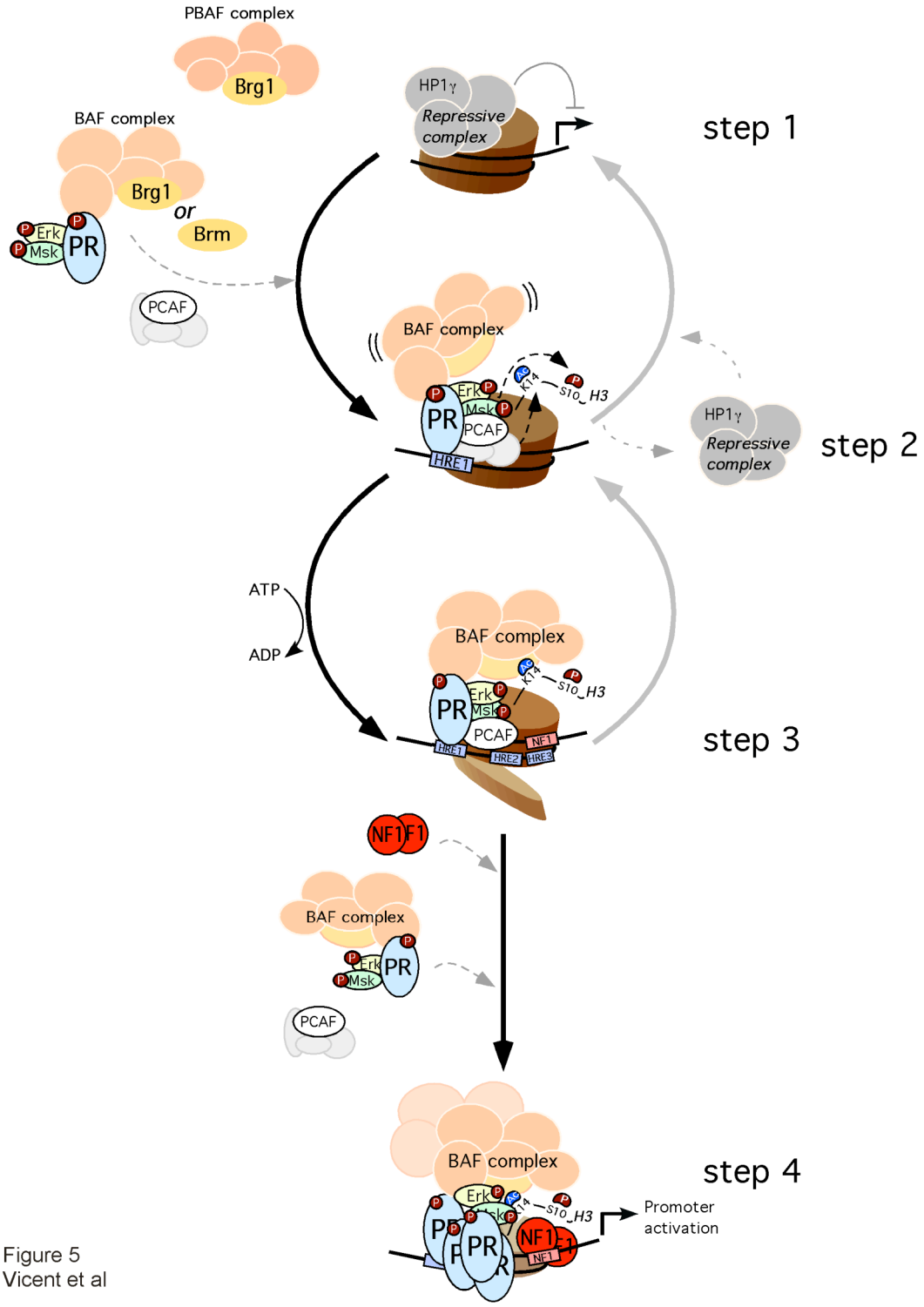


Figure 5
Vicent et al

Supplementary Information

Two chromatin remodeling activities cooperate during activation of hormone responsive promoters

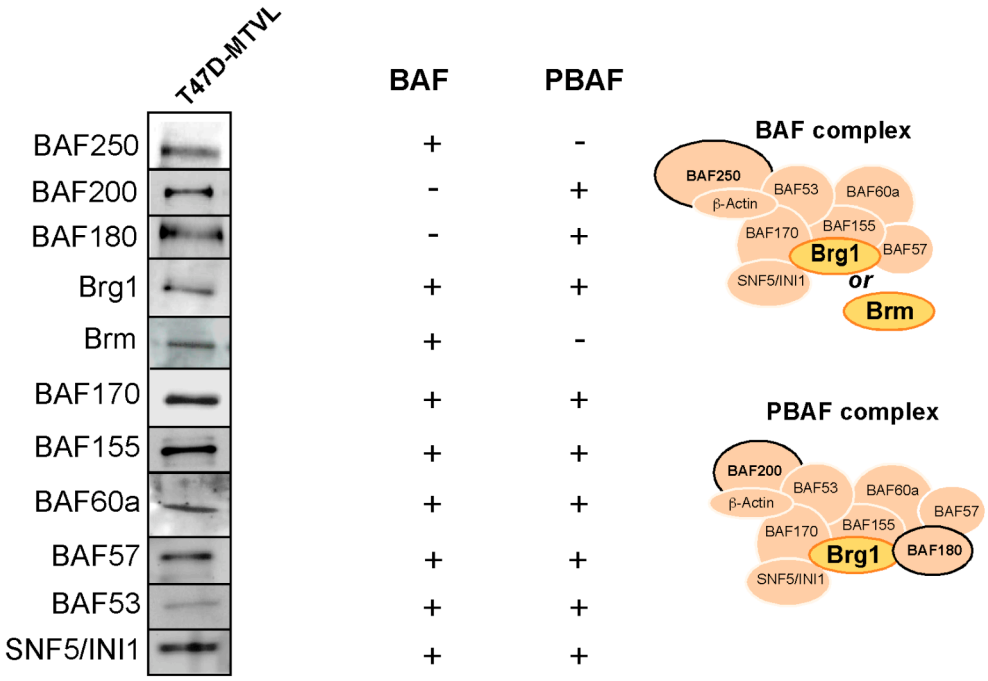
Guillermo Pablo Vicent[1,3], Roser Zaurin[1,3], A. Silvina Nacht[1], Ang Li[1], Jofre Font-Mateu[1], Francois Le Dily[1], Michiel Vermeulen[2], Matthias Mann[2] and Miguel Beato[1]

[1] Department of Gene Regulation
Centre de Regulació Genòmica (CRG)
Universitat Pompeu fabra
Parc de Recerca Biomèdica (PRBB)
Dr Aiguader 88, E-08003
Barcelona
SPAIN

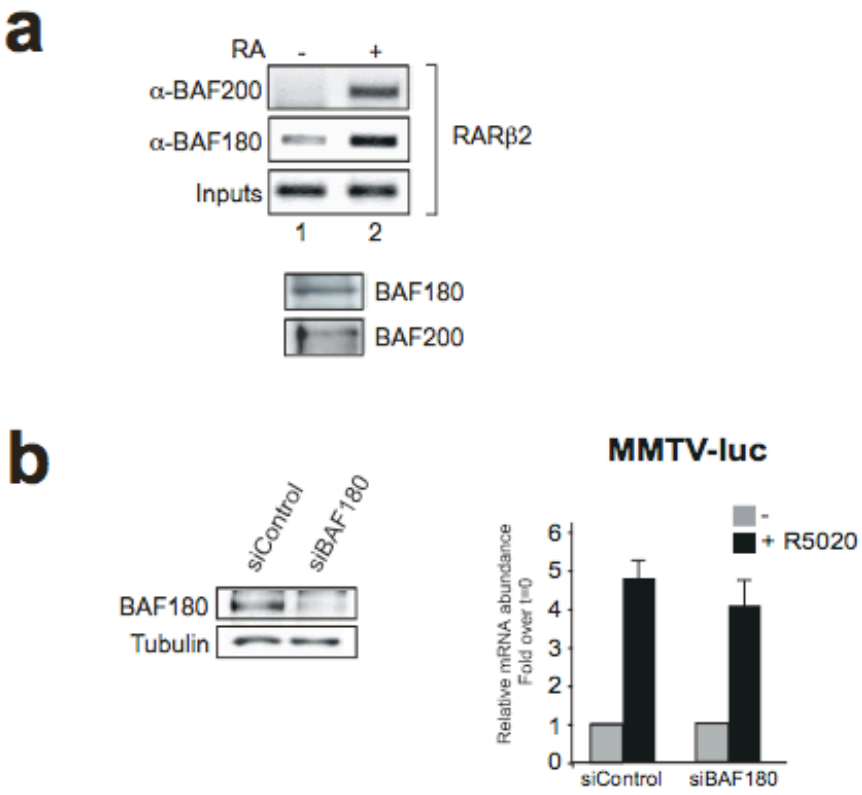
[2] Department of Proteomics and Signal transduction
Max Planck Institute for Biochemistry
Am Klopferspitz 18, D-82152 Martinsried
GERMANY

[3] These authors contributed equally to this work

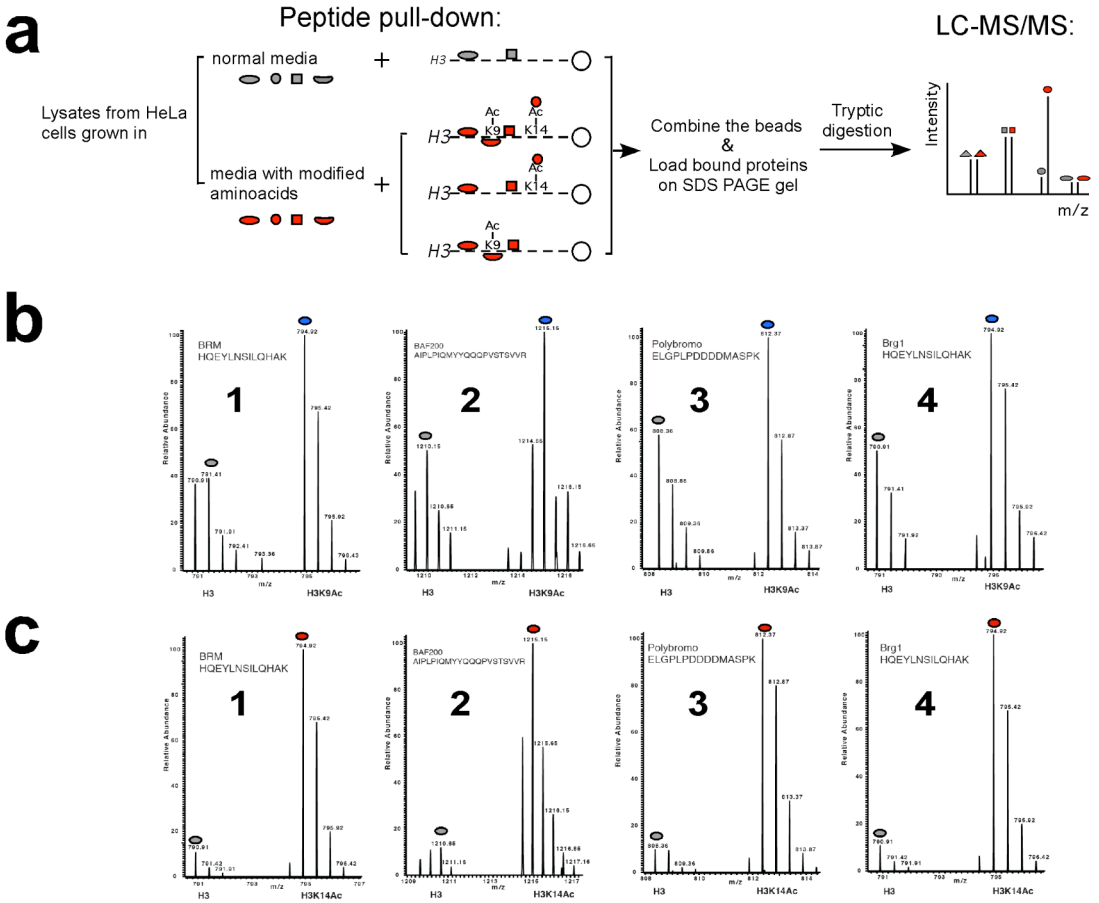
Correspondence should be addressed to M.B. ***email: miguel.beato@crg.es***



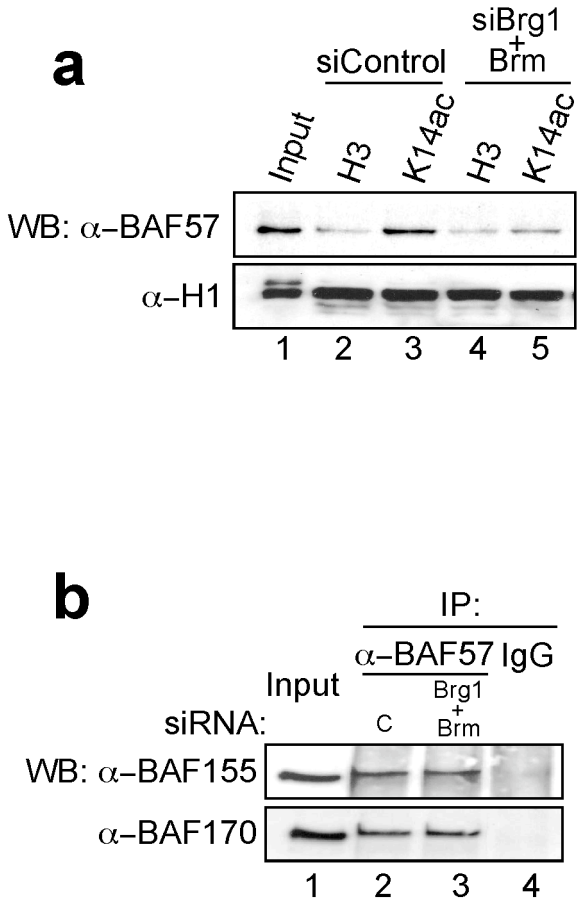
Supplementary Figure 1: Differential expression of BAF proteins in T47D-MTVL breast cancer cells.
Nuclear extracts from T47D-MTVL cells were analyzed by Western blotting with antibodies specific for the individual BAF proteins.



Supplementary Figure 2:
(a) BAF180 is recruited to the RARβ2 promoter. U937 cells (human promyelocitic leukemia cell line) were untreated (-) or treated (+) for 4hs with retinoic acid (RA) and subjected to ChIP assays with BAF200 and BAF180 specific antibodies. The precipitated DNA fragments were subjected to PCR analysis to test for the presence of sequences from -665 to -308 corresponding to the RARβ2 promoter. Lower Panel: expression of BAF180 and BAF200 proteins in U937 cells.
(b) Knock down of BAF180 does not affect MMTV transcriptional activity. Left: T47D-MTVL cells were transfected either with control siRNA or with siRNA against BAF180. After 48 h the medium was replaced by fresh medium without serum. After one day in serum-free conditions, cells were lysed and the levels of BAF 180 and tubulin were determined by Western blotting. Right: T47D-MTVL cells were transfected with Control or BAF 180 siRNAs in RPMI medium and cultured for 48 h. After one day in serum-free conditions, cells were incubated with 10 nM R5020 for 2 hs and total RNA was prepared, cDNA was generated and used as template for real time PCR with specific Luciferase and GAPDH primers. Each luciferase mRNA value was corrected by the GAPDH mRNA level and is expressed as relative RNA abundance over time zero. The values represent the mean and standard deviation from 3 experiments performed in duplicate.



Supplementary Figure 3: Schematic representation of the SILAC-based histone pull down approach
(**a**) SILAC peptide pull-down using unmodified H3 (grey ovals), H3K9ac (**b**, blue ovals), and H3K14ac (**c**, red ovals) peptide. The spectra show the relative binding of Brm (1), BAF200 (2), Polybromo/BAF180 (3) and Brg1 (4) to unmodified and modified peptides.



Supplementary Figure 4: Brg1 and Brm are required for the interaction of BAF and PBAF complexes with the H3K14ac mark.
(a) Nuclear extracts derived from T47D-MTVL cells transfected with control and Brg1 and Brm siRNAs were used for pulldown experiments with the indicated H3 tail peptides coupled to beads. Immunoblotting was performed for the presence of BAF57 and H1.
(b) Nuclear extracts derived from T47D-MTVL cells transfected with control and Brg1 and Brm siRNAs were used for immunoprecipitation either with BAF57 antibody or with normal rabbit IgG as a negative control (IgG). The immunoprecipitates (IP) were analyzed by western blotting with BAF155 and BAF170 specific antibodies.

siRNA against	Name	Company	Cat. number
Brg1	Human SMARCA4	Dharmacon	J-010431-07
Brm	Human SMARCA2	Dharmacon	J-017253-06
BAF57	Human BAF57	Dharmacon	J-017522-05
BAF250	Human BAF250	Dharmacon	L-017263-00
PCAF	Human PCAF	Dharmacon	L-005055
SRC-1	Human SRC1	Santa Cruz	SC-36555
BAF180	Human PB1	Dharmacon	L-008692

Commercial antibodies used for ChIP assays:

Ab against	Name	Company	Cat. Number
Brg1	Brg-1 (H88)	Santa Cruz	sc-10768
Brm	SMARCA2/BRM	Abcam	ab15597
BAF57	BAF57/SMARCE1	Bethyl	A300-810A
BAF250	BAF250a/ARID1a, clone PSG3	Upstate-Millipore	04-080 (Lot#DAM1418274)
BAF200	BAF200 (2036C1a)	Santa Cruz	sc-81050
PR	PR (H190)	Santa Cruz	sc-7208
H3K14ac	Acetyl-Histone H3 (Lys14)	Upstate-Millipore	07-353 (Lot#DAM1462567)
H3K9ac	Acetyl-Histone H3 (Lys9)	Abcam	ab4441
H4K8ac	Acetyl-Histone H4 (Lys8)	Upstate-Millipore	07-328

Supplementary Figure 5: siRNAs and antibodies used for ChIP assays in the present study.

Chapter 3

“Characterization of H3/H4 histone tetramer as the output of SWI/SNF-dependent remodelling of the nucleosome B within the MMTV promoter”

Results

1. H3/H4 histone tetramer as the output of the SWI/SNF-dependent remodelling of the nucleosome B within the MMTV promoter

The chromatin structure of the MMTV promoter is remodelled during hormone induction as indicated by the appearance of a DNAase I hypersensitive site at the dyad axis of nucleosome B (Truss et al, 1995). The remodelling mechanism following progesterone treatment was first described by our laboratory in 2004 (Vicent et al, 2004). In that study, we analyzed the recruitment of transcription factors and ATP-dependent remodelling complexes at the MMTV promoter of T47D-MTVL breast cancer cells, which carry a single copy of this promoter integrated in their nuclear genome (Truss et al, 1995). In a search for the molecular mechanism of chromatin remodelling, the stoichiometry of core histones in the promoter nucleosome was analysed by ChIP experiments using histone-specific antibodies and showed that H2A and H2B are selectively displaced from the nucleosome B (Vicent et al, 2004).

In **Figure 32**, I reproduced the previously published results showing that SWI/SNF-mediated remodelling of the nucleosome B results in ATP-dependent histones H2A and H2B displacement. In the study presented here, I used an *in vitro* system in which purified nucleosomes carrying the sequence of the nucleosome B were incubated with purified SWI/SNF complex and ATP. Remodelled products were subjected to ChIP assays using polyclonal antibodies against H2A and H4 followed by PCR amplification of the nucleosome B sequence. As expected, histone H4 content remained unchanged while the amount of histone H2A was 60% decreased (Figure 32A, lane 4 and 2, respectively and quantification in Figure 32B). These results reproduced and confirmed that dimmers of histone H2A and H2B are displaced from the promoter nucleosome B upon hormone induction, suggesting that the remodelled nucleosome would be an H3/H4 tetramer structure. To provide further evidence for this hypothesis, I used an *in vitro* system to reconstitute H3/H4 tetramer particles and SWI/SNF-remodelled nucleosomes in order to compare their structures and functional characteristics.

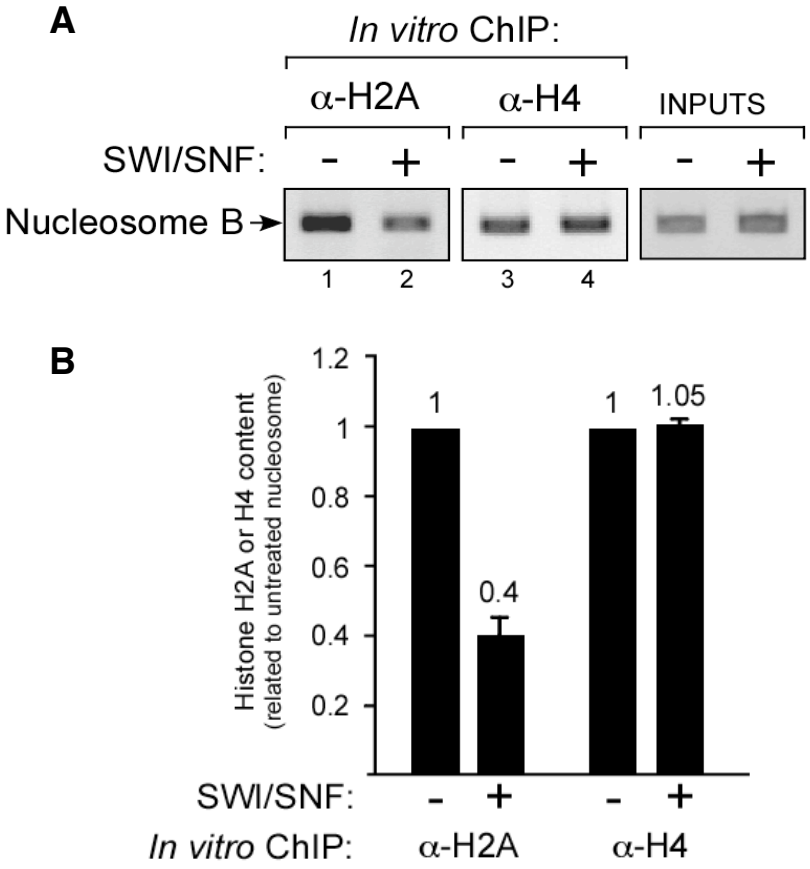


Figure 32. SWI/SNF-dependent histone dimmer H2A/H2B displacement in the MMTV promoter. A. *In vitro* ChIP assay done in monocleosomes treated and non-treated with ySWI/SNF and ATP. Polyclonal antibodies against histones H2A and H4 were used. The immunoprecipitated material was subjected to PCR amplification using specific primers for nucleosome B sequence. **B**. RealTime-PCR quantifications of experiment explained in (A).

1.1 Preparation of octamers and H3/H4 tetramer particles from recombinant histones

For the reconstitution of octamers and H3/H4 tetramers, I used a radioactively labelled DNA fragment comprising the sequence between -221 and +1 of the MMTV promoter, encompassing the region corresponding to nucleosome B (Figure 33). The positioned nucleosome B is located over the promoter region covering the five hormone responsive elements (HREs) and the binding site for nuclear factor 1 (NF1) (**Figure 33A**). Our laboratory has described the existence of two main translational positions of the nucleosome B shifted by 20 base pairs (Figure 33A, grey and yellow ovals). The dyad axis of these two nucleosome populations are located either at position -107 or at -127, relative to the transcription start site (+1) (Spangenberg et al, 1998). The predominant position after an *in vitro* reconstitution with the nucleosome B sequence is the one with dyad axis at -107 (Figure 33A, grey oval). The external HREs (HRE1 and

HRE4) have their major grooves exposed for receptor binding, while the central HREs (HREs 2, 3 and 5) are inaccessible (Piña et al, 1990). Moreover, NF1 is not able to bind to its site within the inactive MMTV promoter. **Figure 33B** shows an SDS gel of the purified recombinant *Xenopus laevis* core histone particles, octamers and tetramers, used for the reconstitution *in vitro*. In order to find the appropriate conditions for tetramer reconstitution, I evaluated five different histone/DNA ratios ranging from 0.5 to 2 (**Figure 33C**, lanes 2-6). I wanted to achieve the efficient formation of H3/H4 tetramers, while avoiding their further association into ditetramer particles. A band corresponding to ditetramer species was initially observed at histone/DNA ratio of 1.5 (lane 5), while the same ratio of 2 (lane 6) resulted in all DNA occupied by two H3/H4 tetramers. Therefore, I chose the histone/DNA ratio of 1 (molar ratio of 2.7) (lane 4) as under these conditions ditetramer particles were not formed yet and the amount of the reconstituted tetramers was comparable to the one observed at higher (1.5) histone/DNA ratio. We also utilized a histone/DNA ratio of 1 (molar ratio of 1.5) for the reconstitution of octamer particles (Figure 33C, lane 1).

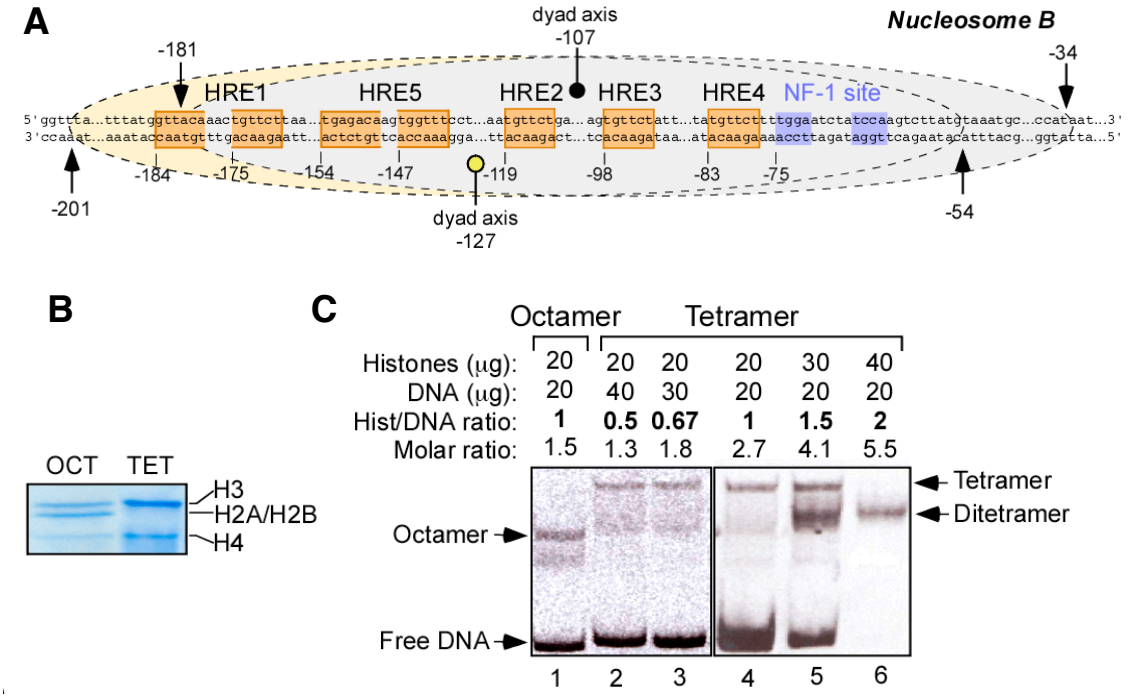


Figure 33. **Preparation of tetramer and octamer particles. A**. Schematic representation of the nucleosome B position relative to the *cis* regulatory elements of the MMTV promoter. *HRE*, Hormone Responsive Element; *NF-1*, Nuclear Factor-1. The two possible nucleosome positions are represented by grey (with dayd at position -107) and yellow ovals (with dyad at position -127) **B**. Composition of recombinant *X.laevis* histones H3, H4, H2A and H2B in a 18% SDS-PAGE gel followed by Commassie staining **C**. Different histone/DNA ratio to reconstitute tetramer particles. I assembled tetramers changing the histone/DNA ratio ranging from 0.5 to 2 (*lane 2-6*). An octamer particles reconstituted in a histone/DNA ratio of 1 (*lane 1*) is also shown. All reconstituted material were analyzed by electrophoresis in 5% polyacrylamide native gels (60:1 acrylamide/bisacrylamide ratio) and stained with Ethidium Bromdie

This figure also shows the different mobility of reconstituted tetramer and octamer particles on native acrylamide gels. Migration of these particles is determined by both molecular weight and the radius of rotation of the molecules. Tetramers have higher portion of DNA that is free of histone contacts at each end resulting in a longer rotation radius and slower running, while octamers comprise more DNA wrapped around the histone core particle generating a more compact shape allowing faster migration. This is consistent with previous studies (Pennings et al, 1991 and Meersseman et al, 1992). Purification of radioactively labelled reconstituted material was done by 10-30% (v/v) glycerol gradient (**Figure 34**). The fractions were collected from the bottom of the gradient and analyzed using 5% polyacrylamide gels (Figure 34, gel). I quantified the radioactivity of each fraction (Figure 34, plot). Purification of the octamer yielded two peaks. The first appearing peak (number 1) corresponded to octamer particles and the later peak (number 2) corresponded to free DNA. When only recombinant histones H3 and H4 were used for reconstitution, three peaks were detected in glycerol gradients. The first peak (number 3) corresponded to ditetramers of histones H3 and H4 assembled on the MMTV DNA fragment, the intermediate peak (number 4) corresponded to tetramer particles and the last peak (number 5) corresponded to free DNA. Similar findings have been reported for a shorter MMTV DNA fragment assembled into tetramers and octamers (Spangenberg et al, 1998). The histone composition of tetramer and octamer assembled particles were confirmed using SDS-PAGE gel electrophoresis (data not shown).

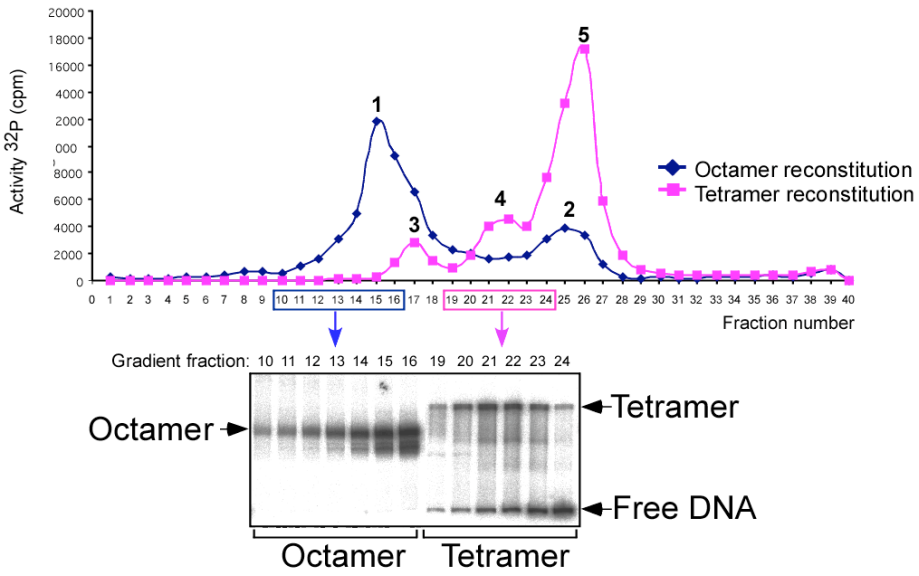


Figure 34. **Glycerol gradient profile of the reconstituted material.** Fractions of 100ul were collected from the bottom o f the gradient and radioactivity was measured with scintillation counter. Cpm for every fraction of the 10-30% glycerol gradient are plotted (upper part) and some fractions of both gradients were run in a 5% acrylamide gel and visualized using Phosphoimager.

2. The structure and binding properties of the reconstituted tetramer particles resemble the SWI/SNF-remodelled nucleosomes

2.1 SWI/SNF remodelling makes nucleosomal DNA more accessible for MNase digestion

I examined the structural organization of the SWI/SNF-remodelled MMTV nucleosomes and the reconstituted tetramer particles using digestion with Micrococcal nuclease (MNase). This method has been widely employed to study specific nucleosome positions due to the protection of DNA in contact with the histone octamer against cleavage by this nuclease (Simpson RT, 1998 and Simpson RT, 1999). The DNA products of MNase-digestions are the fragments that are protected by the histone core particle, since MNase is able to digest all remaining linker DNA. **Figure 35** shows the DNA digestion patterns for octamer (OCT), SWI/SNF-treated octamer particles (OCT+SWI/SNF) and tetramer (TET). Protected DNA fragments were extracted, labelled and separated on 6% polyacrylamide gels. MNase digestion of MMTV nucleosomes yielded a single 146 base pair fragment (Figure 35, lane 3), which corresponded to the core particle DNA. The 146-bp fragment was also observed following the MNase digestion of the SWI/SNF-remodelled nucleosomes, but was accompanied by shorter fragments with a main band of 96 base pairs (Figure 35, lanes 4-8). The 96-bp band was the dominant species detected when H3/H4 tetramers were subjected to MNase treatment (Figure 35, lane10).

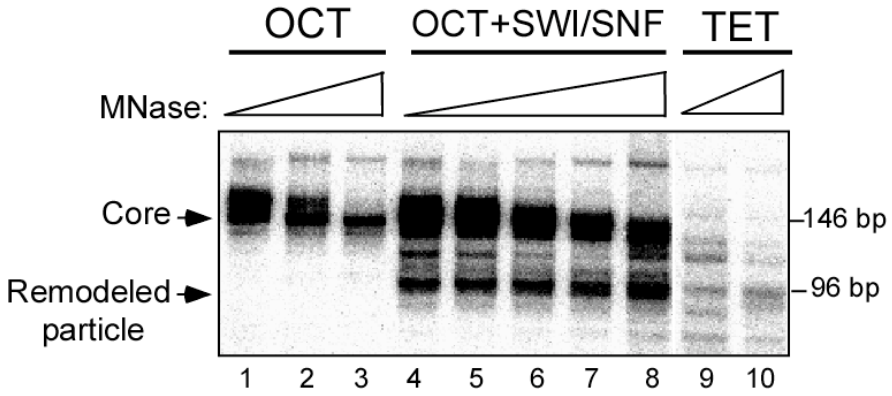


Figure 35. **Comparison of MNase-digestion of untreated ot SWI/SNF treated octamers and tetramers**. Reconstituted octamers (OCT), SWI/SNF-treated nucleosomes (OCT+SWI/SNF) and tetramers (TET) were digested with increasing concentrations of MNase (0.6, 1.2 and 2.4 unit/reaction in *lanes 1-3*; 0.6, 1.2, 2.4, 5 and 10 unit/reaction in *lanes 4-8*; 0.15 and 0.2 unit/reaction in *lanes 9 and 10*) for 5 minutes at 22°C. The digested particles were analyzed by electrophoresis in a 6% polyacrylamide native gel.

The remaining 146-bp band in SWI/SNF-treated octamers indicates that the used concentration of SWI/SNF was not sufficient to remodel all nucleosome particles (Figure 35, lanes 4-8). It seems likely that the reduction of total DNA observed after MNase digestion with SWI/SNF-remodelled octamers and with tetramers is due to the lack of H2A and H2B histones, that makes the DNA more accessible to the nuclease. The fact that SWI/SNF remodelled octamers generate an MNase product of the same size as H3/H4 tetramers suggests that a H3/H4 tetramer is the structure that results from SWI/SNF-mediated remodelling of the nucleosome B.

2.2 Simultaneous binding of PR and NF1 observed at *in vivo* activated MMTV promoters can be reproduced using reconstituted tetramers

The functional consequence of the observed changes in the nucleosome B structure during hormone induction is to allow the simultaneous binding of more molecules of PR to HREs 2, 3 and 4, together with the binding of NF1 to its recognition site, facilitating full activation of the MMTV promoter. As said above, only two of the five HREs can be bound by progesterone receptor in the inactive state of the MMTV promoter, in which the binding site of NF1 is also inaccessible. However, following hormone induction *in vivo* all HREs and the binding site for NF1 are occupied simultaneously on the surface of a nucleosome-like structure (Truss et al, 1995) and a functional synergism is observed between progesterone receptor (PR) and NF1 (Chalepakis et al, 1988).

2.2.1 NF1 and PR binding to reconstituted octamers and tetramers

Therefore, we next compared non-remodelled and remodelled nucleosome core particles, H3/H4 tetramer particles and free DNA in terms of NF1 binding. Tetramers and octamers were reconstituted *in vitro* on MMTV DNA as explained in Figure 33 and 34 (and *Materials and Methods*). The reconstituted particles and free DNA were incubated with increasing amounts of recombinant NF1 transcription factor (**Figure 36A**). Free DNA was used to determine the appropriate amount of NF1 required for the band shift experiment (Figure 36A lanes 2-4). As expected (Spangenberg et al, 1998), regardless of the protein concentration used, I have not detected binding of NF1 to DNA organized into the octamer core particle (Figure 36A lanes 6-8). However, NF1 was able to bind both the tetramer particle (Figure 36A lane 10-12) and the remodelled

octamer (Vicent et al, 2004). Based on these results, I concluded that the NF1 site of the MMTV promoter assembled into tetramer particles is accessible as in hormone-activated MMTV promoter chromatin *in vivo* (Truss et al, 1995).

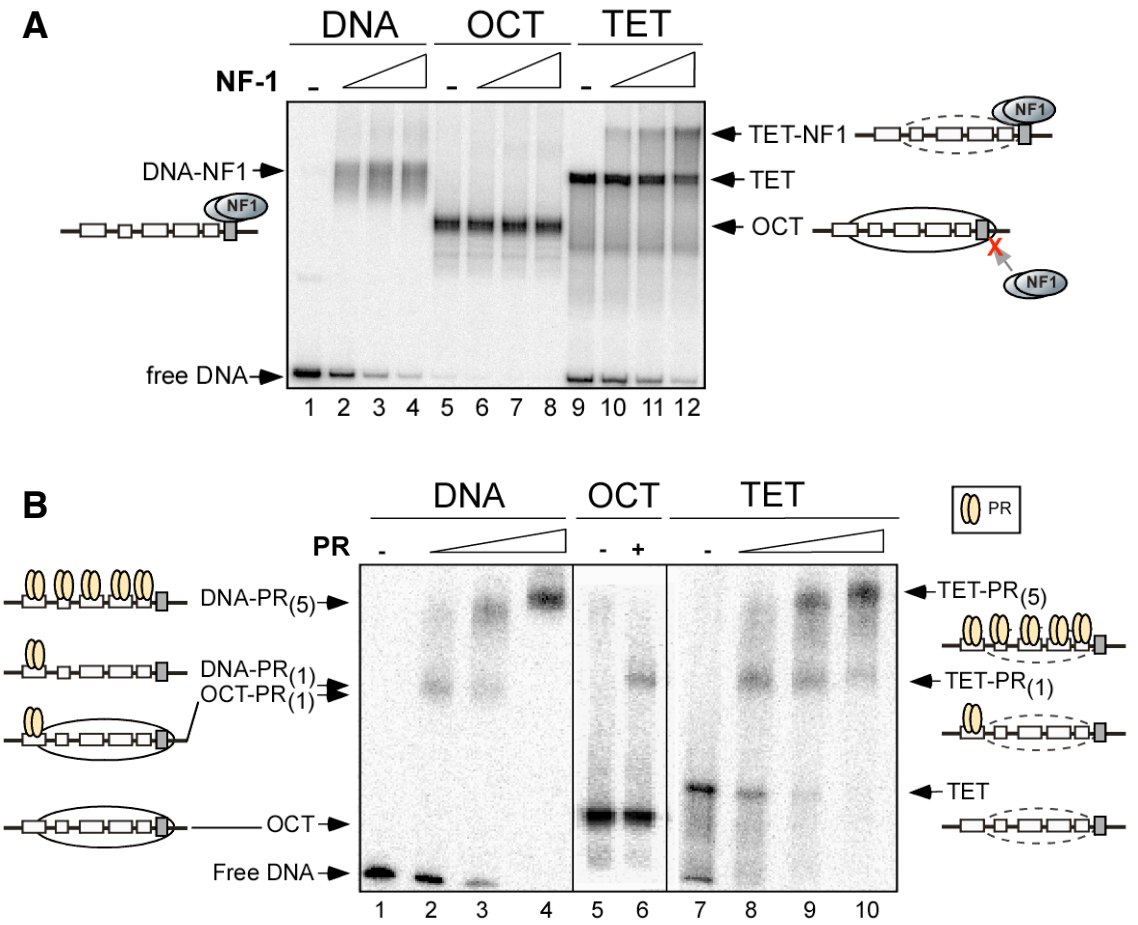


Figure 36. **Differential binding of NF1 and PR to assembled octamer and tetramer particles**. **A**. Binding of NF 1 to tetramer and octamer particles**.** Free DNA, octamer (OCT) and H3/H4 tetramer (TET) particles were incubated with increasing amounts of recombinant NF1 (1 ul, 2 ul and 4 ul). After allowing the binding reaction to proceed for 20 minutes at room temperature, the samples were analyzed on a 5% acrylamide gel. **B**. Binding of PR to tetramer and octamer particle. Free DNA, octamer (OCT) and tetramer (TET) particles were incubated for 20 minutes at room temperature with different amounts of PR (DNA: 1 ul, 2 ul and 5 ul; OCT: 8 ul; TET: 1 ul, 2 ul and 5 ul). The samples were analyzed in 3.5% acrylamide / 20%glycerol / 0.5% agarose / 0.3xTBE gel.

I next studied the progesterone receptor (PR) binding properties of an H3/H4 tetramer particle compared to an octamer particle and free DNA (**Figure 36B**). In the absence of remodelling only the exposed HRE 1 is occupied by PR (Figure 36B lane 6 and Vicent et al, 2002). The half palindromic HRE 4 has low affinity for PR and is not occupied under these conditions. In tetramers (Figure 36B lanes 7-10) and in remodelled nucleosomes (Vicent et al, 2004) access to the hidden HREs is visualized as two differentially migrating bands: one band corresponds to the occupancy of only

HRE 1 by PR (TET-$PR_{(1)}$), while a slower band corresponds to the situation where all HRE sites are occupied by PR molecules (TET-$PR_{(5)}$). The occupancy pattern of the HREs in free DNA is similar to that observed in reconstituted tetramer particles (Figure 36B, compare lanes 2-4 with 8-10). This band shift assay was visualized using an agarose-acrylamide gel since the TET-$PR_{(5)}$ complex is too large to be resolved in an acrylamide gel. The lower resolution of the gel, compared to the products of the NF1 band shift assay (Figure 36A), is due to the addition of agarose to the gel composition. In conclusion, all HRE can be occupied by PR in MMTV DNA assemble on H3/H4 tetramers.

2.2.2 Simultaneous binding of NF1 and PR to tetramers

In order to understand the properties of the nucleosome B in activated promoters, I decided to test if H3/H4 tetramer particles (TET) are able to simultaneously bind NF1 and PR. The binding of these two proteins to the MMTV sequences organized around an H3/H4 tetramer was studied using a band shift assay that detects very large complexes (**Figure 37A**). Incubation of tetramer particles with NF1 generated a single retarded complex (Figure 37A lane 2). Incubation of TET with PR generated a complex pattern of bands, reflecting occupancy of HRE 1 alone or in various combinations with other HREs (Figure 37A lane 3). Incubation of tetramer particles with NF1 and increasing amounts of PR produced two main slow-migrating complexes containing both proteins (Figure 37A lanes 4 and 5) and some intermediate complexes. The faster complex corresponds to tetramers carrying NF1 and one molecule of PR (TET-NF1-$PR_{(1)}$), while the slower migrating complex, dominant at higher concentrations of PR (lane 5), correspond to tetramers associated with NF1 and multiple PRs (TET-NF1-$PR_{(5)}$). The same slow-migrating complex was detected when either PR (Figure 37A lane 7) or NF1 (Figure 37A lane 8) were first bound to the tetramer and the second transcription factor (NF1 or PR, respectively) was added afterwards.

In order to confirm the nature of the observed complexes, I performed the same band shift experiments with the addition of excess of oligonucleotides carrying either the NF1 site sequence or the PRE element (**Figure 37B**). An excess of NF1 oligonucleotide competed with the TET-NF1 complex (Figure 37B, lane 3) leaving almost all tetramers in a free state (Figure 37B, compare lanes 2 and 3). The addition of NF1 oligo also abolished the appearance of the very slow migrating complex of TET-

NF1-$PR_{(5)}$ (Figure 37B, compare lanes 4 and 6). After addition of NF1 oligonucleotide to the TET-NF1-$PR_{(5)}$, the slowest migrating complex with NF1 and five PRs disappeared and we observed a faster migration band corresponding to the TET- $PR_{(5)}$ complex and, in lower proportion, free tetramers (Figure 37B, lane 6). Similarly, a PRE oligonucleotide competed with the formation of the slow migrating complex (TET-NF1-$PR_{(5)}$) resulting in the generation of the faster TET-NF1 complex (Figure 37B compare lane 4 and 5). The slowest migrating band observed in lane 5 could correspond to intermediate complexes of tetramer with NF1 and some PR molecules.

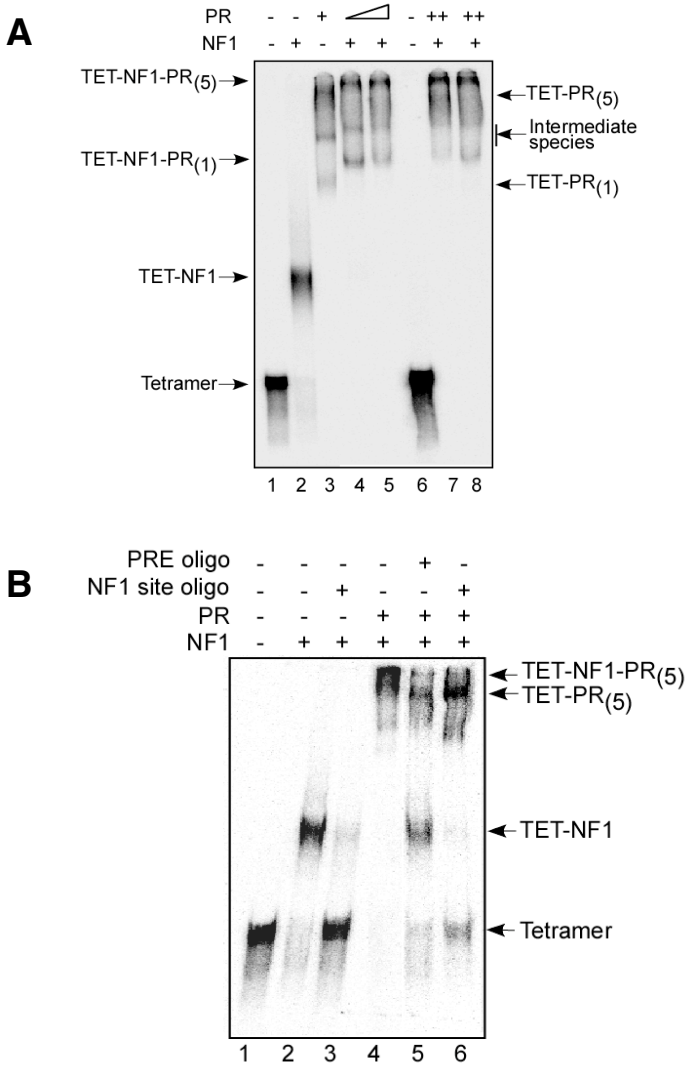


Figure 37. **The structure of reconstituted tetramer particles allow simultaneous binding of PR and NF1. A**. Tetramer particles were incubated at room temperature with a fixed concentration of recombinant NF1 or PR (*lanes 2-3*). Tetramers incubated with a fixed amount of NF1 and increasing amounts of PR is shown in *lanes 4-5*. The same complexes are observed when tetramer are pre-incubated with NF1 prior to PR binding (*lane 7*) or the contrary, pre-incubation with PR followed by binding of NF1 (*lane 8*). After 20 minutes of incubation samples were analyzed on a 3.5% acrylamide gel **B**. Tetramers were incubated with fixed amounts of either NF1 or PR with no remaining free tetramer form (*lane2 -3,* respectively). *PRE* (Progesterone Responsive Element) oligo is a short DNA fragment that contains DNA binding motif for PR. NF1 site oligo is a short DNA fragment containing DNA binding site sequence for NF1. Both oligos were added to the TET-NF1-PR mix separately (*lane 5-6*).

These results indicate that the tetramer particle, mimicking the product of the MMTV nucleosome remodelling generated by SWI/SNF *in vivo*, can accommodate the full loading of the promoter with PR and NF1 as observed in living cells after hormone induction (Truss et al, 1995).

3. H3/H4 tetramer is not a suitable substrate for SWI/SNF-dependent remodelling

Previous reports have claimed that SWI/SNF can remodel tetramers of histones H3 and H4, although these tetramers are poorer substrates for SWI/SNF-dependent remodelling compared to nucleosomal arrays (Boyer et al, 2000). To investigate this further with MMTV promoter sequences, I incubated MMTV octamer and tetramer particles with SWI/SNF and ATP and visualized the remodelled products in a 5% acrylamide gel (**Figure 38A**). SWI/SNF treatment resulted in faster migrating bands (Figure 38A lanes 2 and 3) that correspond to nucleosomes located to the ends of the sequence (as indicated in the scheme and lanes 2 and 3). In contrast, SWI/SNF-treated tetramer particles did not show any change in their migration pattern, even when the highest concentration of the remodelling complex was used (Figure 38A lane 5).

To further explore whether tetramers could be substrates for the remodelling activities of SWI/SNF, tetramers treated and non-treated with purified SWI/SNF were analyzed by centrifugation in 10-30% glycerol gradient and characterized by electrophoresis on 5% native polyacrylamide gels (**Figure 38B**). I found exactly the same distribution of fractions for treated (lanes 9-16) and non-treated (lanes 1-8) tetramer particles. The appearance of tetramer particles and free DNA occur exactly at the same fraction numbers (compare lanes 4-5 with 12-13 for the tetramer; and lanes 8 and 16 for free DNA) suggesting no change in either structure or histone composition in treated as compared to non-treated tetramers. Taken together, these results suggest that the tetramer particle is not a suitable substrate for purified SWI/SNF remodelling complex.

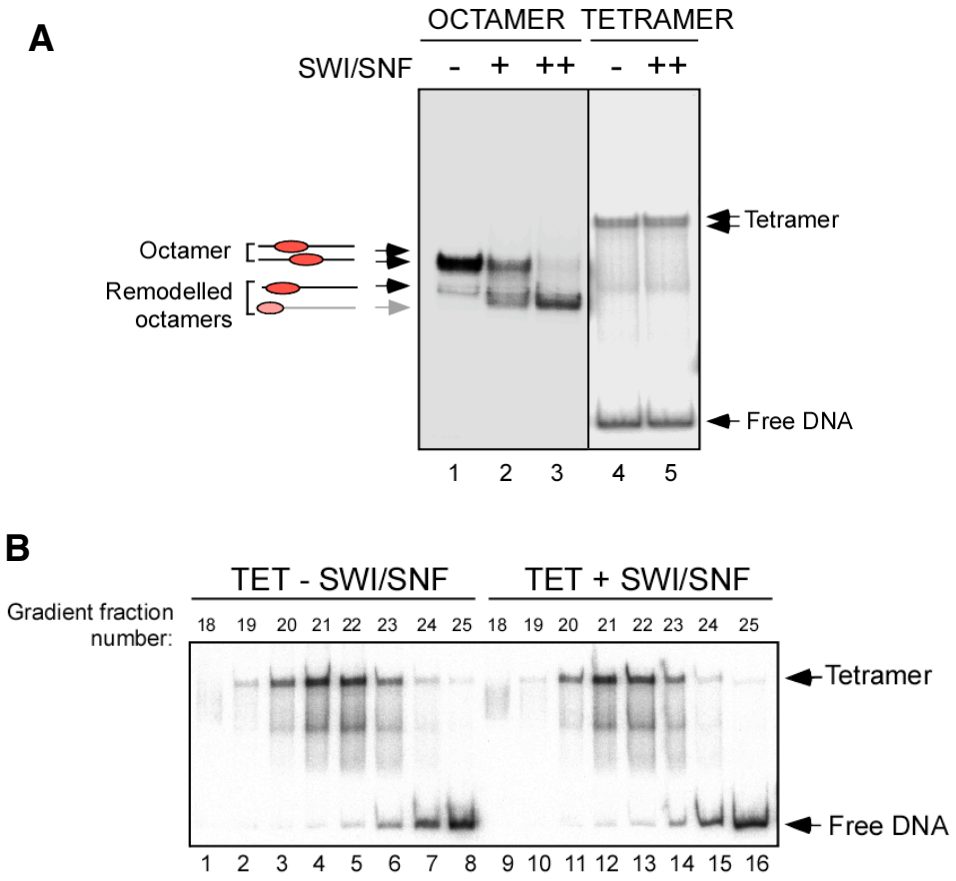


Figure 38. Tetramers are not suitable substrates for SWI/SNF-mediated remodelling. A. Reconstituted octamer and tetramer particles were incubated with different concentrations of SWI/SNF in the presence of ATP and analyzed on a 5% acrylamide gel **B.** Non-treated (TET-SWI/SNF) and treated (TET+SWI/SNF) tetramer particles were separated by ultracentrifugation in a 10-30% glycerol gradient and analyzed by electrophoresis on a 5% acrylamide gel.

One possible explanation is that SWI/SNF cannot bind to a tetramer particle. Alternatively SWI/SNF should be able to associate with a tetramer, but may be inactive in this complex. To distinguish between these two possibilities, I performed band shift analysis with octamer and tetramer particles pre-incubated with different amounts of purified SWI/SNF complex (**Figure 39A**). As expected, I observed a retarded band corresponding to the complex of SWI/SNF bound to reconstituted nucleosomes (Figure 39A lanes 2-4). However, this band was barely detectable when tetramer particles were used in the assay (Figure 39A lanes 6-8). Only with the highest SWI/SNF concentration, a weak band that corresponds to the complex TET-SWI/SNF was observed (Figure 39A lanes 8). Quantification of the slower migrating band corresponding to both complexes, OCT-SWI/SNF and TET-SWI/SNF, is shown in **Figure 39B**. Based on these calculations, only around 5% of the tetramers were bound to SWI/SNF when the highest concentration of the remodelling complex was used

(Figure 39A lane 8). On the other hand, at the same concentration, 80% of octamers form a complex with SWI/SNF (Figure 39A lane 4). These data, together with the results showing that SWI/SNF actively displaces H2A/H2B dimers from the nucleosome B upon hormone induction (Vicent et al, 2004), imply that H2A/H2B dimers are needed for the recognition and binding of the SWI/SNF to the nucleosome.

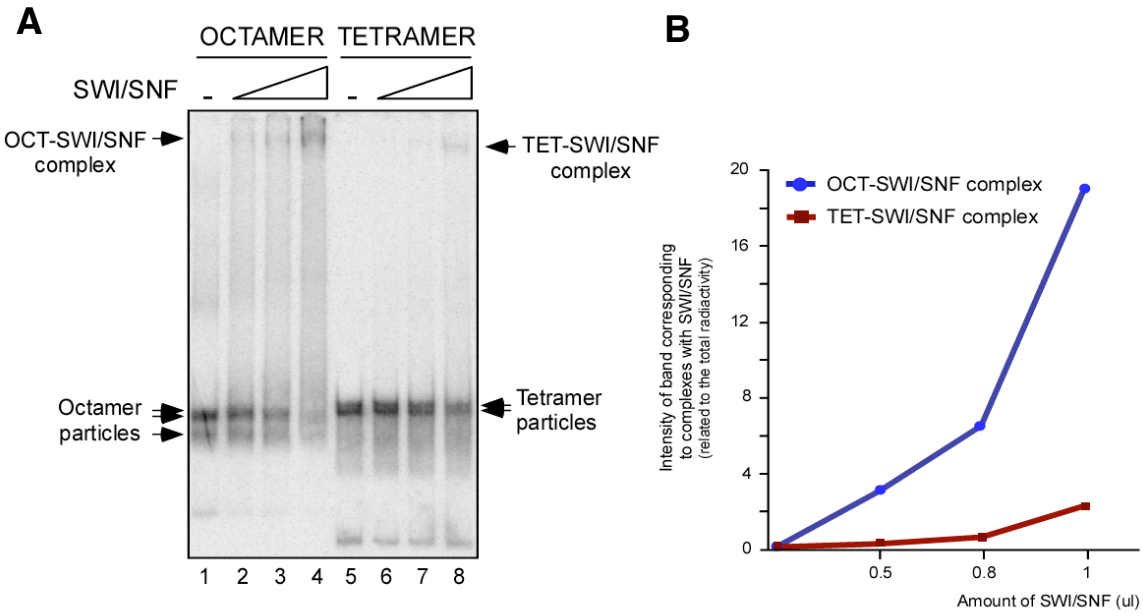


Figure 39. **SWI/SNF binding to tetramers and octamers. A**. Octamer and tetramer particles were incubated with increasing amounts of SWI/SNF in presence of ATP for 20 minutes at 30 ºC. After the incubation complexes were crosslinked and analysed in a 5% acrylamide gel **B**.Quantification of the appearing of a slow migrating band corresponding to either tetramer bound to SWI/SNF (TET-SWI/SNF complex and red line) or octamer bound to SWI/SNF (OCT-SWI/SNF complex and blue line).

4. Role of the internal HRE 2 and 3 in the activation of the MMTV promoter

Till now, I have shown that reconstituted tetramer particles resemble the SWI/SNF-remodelled nucleosomes in histone content (Figure 32 and Vicent et al, 2004), histone-DNA interactions (Figure 35) and PR and NF1 binding properties (Figures 36 and 37). We next examined the functional role of the inner HRE sites and the NF1 binding site, all of which become accessible in the MMTV nucleosome B during remodelling events.

4.1 Importance of the HREs and the NF1 site for the MMTV induction

We first performed luciferase assays in order to explore the importance of the different hormone responsive elements (HRE 1, 2 and 3), along with the NF1 binding site, for the transcriptional activation of the MMTV promoter. For these experiments, I introduced mutations in different HREs and NF1 binding site. The set of mutants include single mutants as well as the combination of mutations in HREs 2 and 3. In the HRE mutants, the HRE half sites have been changed in the following way: TGTTCT → TGTTGT. The NF1 binding site sequence has been changed as follows: TTGGAatctaTCCAA → TTGAAatctaTTCAA. I perform single point mutations to avoid an effect of the mutations on nucleosome positioning or stability. In fact several nuclease accessibility assay and salt resistant test demonstrated no differences between WT and mutant nucleosomes (data not shown). The data showed that the activity of the promoter decreased dramatically when constructs carrying point mutations either at both HRE1 (MMTV-HRE1$^-$) or at HRE2 and 3 sites (MMTV-HRE 2$^-$/3$^-$), or at and the NF1 (MMTV-NF1$^-$) site, were transiently transfected into T47D cells (**Figure 40**).

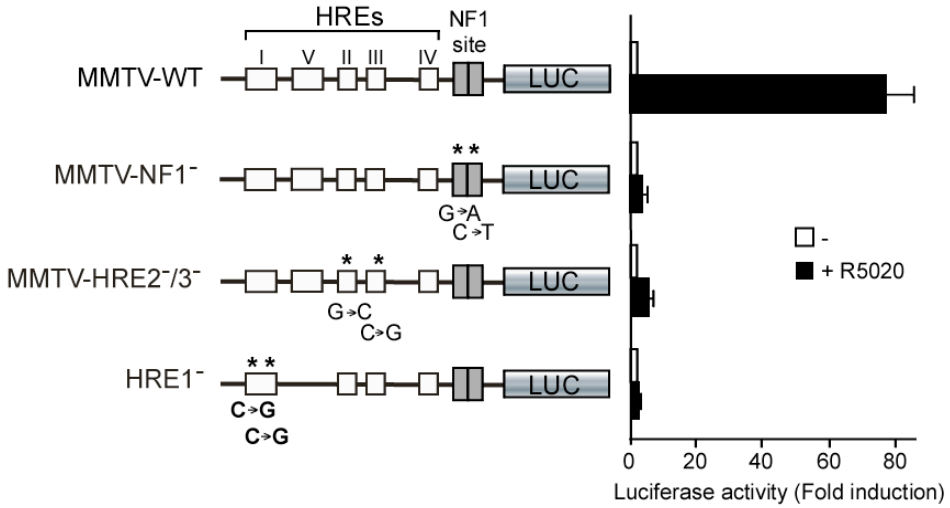


Figure 40. **Role of HREs 2 and 3 in activation and remodelling of the MMTV promoter**. T47D cells were transiently transfected with the luciferase reporter genes driven by the wild type MMTV promoter or by NF1- and HRE 2-/3- mutant promoters as indicated. After transfections cells were treated with 10 nM R5020 or vehicle for 16 h, lysed and luciferase activity was measured. The values represent the mean and standard deviation from 2 experiments performed in duplicate. This experiment was done by François Le Dily.

These results indicate that all these sites are required for the full PR-mediated transcriptional activation of the MMTV promoter. Surprisingly, point mutations at HREs 2 and 3 resulted in almost no activation of the promoter, although the binding of PR-SWI/SNF complex to HRE1, as well as NF1 binding to its cognate site, should not be affected in these mutants. Therefore, I have had particular interest in studying the role of the internal HREs 2 and 3 sites in the MMTV promoter activation, since our results suggest that the binding of PR to the exposed HRE 1 and NF1 recruitment are not sufficient for orchestrating the full transcriptional activation of the promoter (Figure 40).

4.2 Role of HREs 1 to 3 and NF1 site on remodelling of MMTV nucleosomeB.

In order to investigate this, mononucleosomes were reconstituted using recombinant and purified histones with DNA that carries point mutations on either both HREs 2 and 3 sites (HRE $2^-/3^-$), or at the individual HRE 1 site ($HRE1^-$), as described in Figure 40. I have first tested the SWI/SNF-dependent sliding of the WT, $HRE2^-/3^-$ and $HRE1^-$ nucleosomes by running a 5% acrylamide gel before and after incubation with SWI/SNF and ATP (**Figure 41A**). The remodelled products of these reactions can be observed as faster migration bands (Figure 41A, lanes 2, 4 and 6) in samples corresponding to WT, $HRE2^-/3^-$ and $HRE1^-$ mononucleosomes, respectively. SWI/SNF slides the nucleosome along the DNA sequence in a similar way for all three types of nucleosomes, independent of the point mutations they carry. This result confirms that HRE 2 and 3 are required for SWI/SNF remodelling.

In order to determine whether nucleosome remodelling could take place even when HREs 2 and 3 are not functional, I performed *in vitro* ChIP experiments using a specific antibody against the histone H2A to precipitate either the wild type (WT) and mutant mononucleosomes (**Figure 41B**). When mononucleosomes were incubated with recombinant SWI/SNF in the presence of ATP, content of histone H2A decrease extensively in both WT and $HRE2^-/3^-$ nucleosomes (lanes 2 and 6, respectively). When competitor DNA was added to the remodelling mixture, SWI/SNF was captured by the excess of competitor DNA and remodelling was not observed (lanes 3 and 7), as indicates the amount of histone H2A immunoprecipitated. However, incubation of mononucleosomes with SWI/SNF, ATP, competitor DNA and PR recovers the decrease on histone H2A content on WT and $HRE2^-/3^-$ mutant mononucleosomes

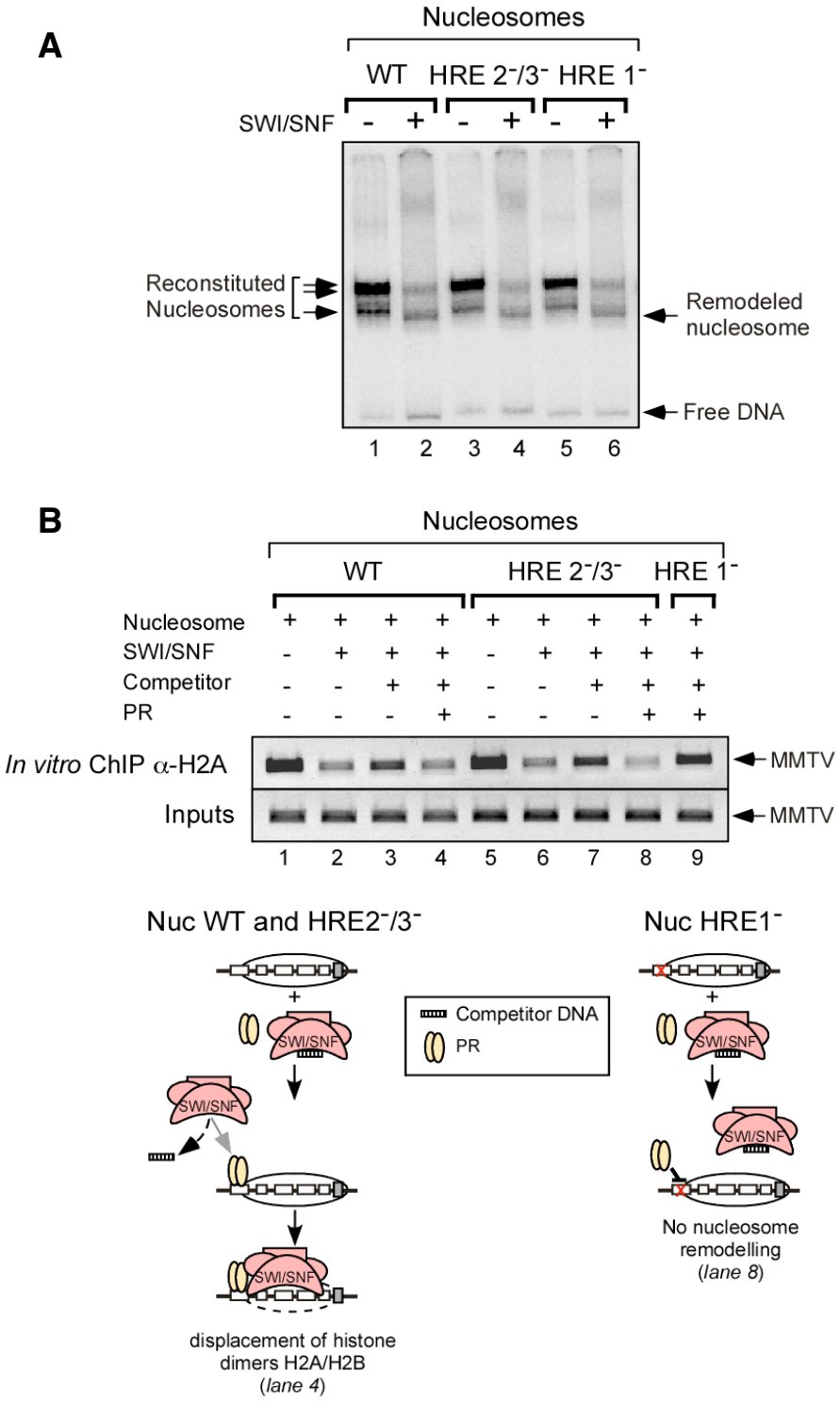


Figure 41. Role of HREs 2 and 3 in activation and remodelling of the MMTV promoter. A. Wild type, HRE 2-/3- and HRE1- mononucleosomes were treated with recombinant SWI/SNF as explained. Following incubation at 30ºC for 30 minutes excess of competitor DNA was added to the reaction for another 30 minutes at 30ºC. Remodelled products were analyzed by running a 5% acrylamide gel **B**. Nucleosomes assembled on wild type (WT) MMTV promoter DNA and on MMTV promoter mutated in HREs 2 and 3 (HRE 2-/3-) and HRE1 (HRE1-) were treated with SWI/SNF in the presence of ATP, competitor DNA and PR as indicated. Following incubation at 30ºC for 30 minutes, the nucleosomes were crosslinked in 0.25 % formaldehyde and immunoprecipitated with a polyclonal antibody against histone H2A. The precipitated DNA fragments were subjected to PCR analysis (20 cycles) with oligonucleotides corresponding to the MMTV promoter nucleosome B. 5% of the input DNA was used.

(lanes 4 and 8). The data imply that PR can outcompete the competitor DNA-SWI/SNF complex formation, resulting in the recruitment of SWI/SNF to the promoter with consequent H2A/H2B displacement. However, these results also indicate that the mutations that prevent PR binding to the HRE 2 and 3 sites do not interfere with the SWI/SNF-dependent nucleosome remodelling *in vitro.* I have used the mutant of the HRE1 site as a negative control. This site is exposed even in absence of remodelling, and its interaction with PR appears to be required for SWI/SNF recruitment. Accordingly, mononucleosomes carrying point mutations in the HRE1 are not able to bind PR, resulting in the lack of the SWI/SNF recruitment in the presence of competitor DNA and, consequently, no histone displacement is observed at the promoter (Figure 41B, lane 9).

Based on these results, I concluded that even in the absence of PR binding to the HREs 2 and 3, SWI/SNF is recruited to the nucleosomes and is able to displace the histones H2A/H2B. Our data until now suggest that the internal HREs 2 and 3 probably take part in the later process of the promoter activation.

4.3 NF1 binding is independent of PR binding in MMTV minichromosomes functional HREs 2 and 3

I next wanted to explore whether the binding of NF1 is dependent on the recruitment of PR to the internal HREs 2 and 3 sites. To address this question, I utilized a more complex system, which is the MMTV promoter assembled into minichromosomes using extracts from preblastodermic *Drosophila* embryos (Figure 42). This system closely mimics the *in vivo* situation; it carries positioned nucleosomes (Venditti et al, 1998) and is activated in a process involving a two-step synergism between PR and NF1 (Di Croce et al, 1999). Separate addition of either recombinant PR or NF1 had weak effects on transcription of wild type (WT) and HREs $2^-/3^-$ mutant MMTV minichromosomes (**Figure 42**, lanes 2, 3 and 6, 7 quantification is given below each lane). However, when both PR and NF1 were present a synergistic transcriptional activaton was observed with the wild type MMTV but not with the HREs $2^-/3^-$ mutant promoter (Figure 42, lanes 4 and 8).

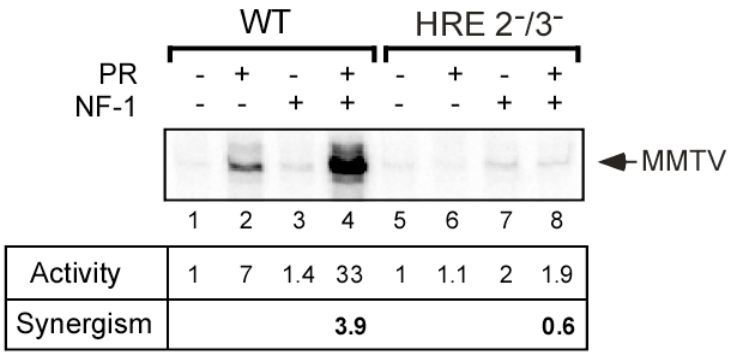


Figure 42.Central HREs are critical for transcriptional induction of the MMTV promoter. Wild type and HRE2-/3- mutant MMTV minichromosomes (25 ng of DNA in each reaction) assembled on post-blastodermic *Drosophila* embryo extracts were incubated with purified recombinant activated PR and NF1 and transcribed *in vitro* with HeLa nuclear extract. Products were visualized by primer extension analysis and sequencing gel electrophoresis. Quantification of intensities of the bands is shown below the gel. This experiment was performed by Jofre Font Mateu.

I next performed ChIP assays with the wild type and HRE $2^-/3^-$ mutant MMTV reconstituted minichromosomes. In the presence of PR and NF1, when the MMTV promoter is actively transcribed (Figure 42, lane 4), we detected the binding of PR (**Figure 43A**, lanes 7-8) and NF1 (**Figure 43B**, lanes7-8) to the wild type (WT) minichromosomes under the same conditions. I observed an eleven-fold decrease in PR binding to the HRE $2^-/3^-$ mutant minichromosomes, (Figure 43A lanes 11-12), which was in a good correlation with the low transcriptional rate displayed by this mutant (Figure 42, lane 8). Nevertheless, the HRE $2^-/3^-$ mutant binds NF1 to a similar extent as the wild type minichromosomes (Figure 43B, lanes 11-12). This HRE 2 and 3-independent binding of NF1 was also observed by *in vitro* ChIP assay with the MMTV promoter assembled in mononucleosomes (data not shown). These data lend further support to the hypothesis that SWI/SNF-mediated recruitment of NF1 is dependent on the binding of PR to the exposed HRE 1, but is independent of the PR binding to the central HREs.

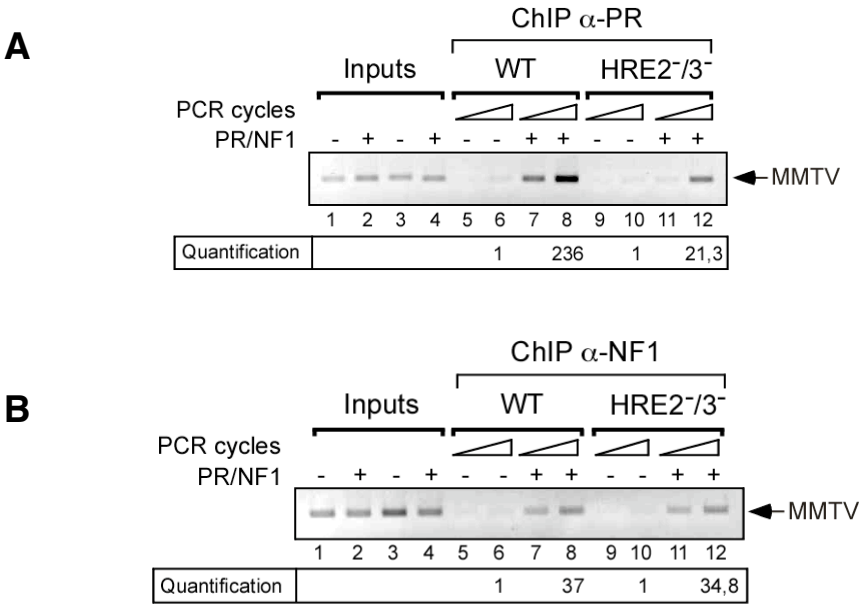


Figure 43. PR and NF1 binding on minichromosomes. Wild type and HRE2-/3- mutant MMTV minichromosomes were incubated with purified recombinant activated PR and NF1 and subjected to ChIP assays as described previously using specific antibodies against PR (***A***) and NF1 (***B***). The precipitated DNA fragments were subjected to PCR analysis (20 and 25 cycles) to test for the presence of the MMTV nucleosome B sequence. Input material (5%) is shown for comparison. Quantification was done by real time PCR as previously described (Vicent et al., 2006).

5. NF1 cannot recruit SWI/SNF through direct interaction with the MMTV promoter

I further wanted to more precisely define the role of NF1 binding in the activation of the MMTV promoter. One possibility is that NF1 could recruit more SWI/SNF complex to the promoter facilitating the final remodelling of the nucleosome B. To test this, I used a mutant MMTV promoter containing an insertion of 50bp downstream of the HREs. The 50-bp insertion (+50) maintains the rotational setting and moves the NF1 binding site away from the dyad axis and into the linker DNA (Eisfeld et al, 1997). I performed binding assays using both the reconstituted mutant nucleosomes (Nuc +50) and free DNA (DNA +50), which were incubated with increasing amounts of recombinant NF1 (**Figure 44A**). In contrast to wild type nucleosomes (Figure 36A lanes 6-8), NF1 binds very efficiently to +50 mutant nucleosomes and a single retarded complex is formed (Figure 44A lanes 6-8). The residual band includes a subpopulation of nucleosomes in which the histone octamer is located at the end of the DNA fragment partially covering the NF1 binding site (Eisfeld et al, 1997; see scheme on the right site). The control band shift experiments with free DNA (DNA +50) are depicted in

parallel (Figure 44A lanes 2-4). These data confirm that the mutant nucleosomes (Nuc +50) have an accessible NF1 binding site in reconstituted ocatmer core particles.

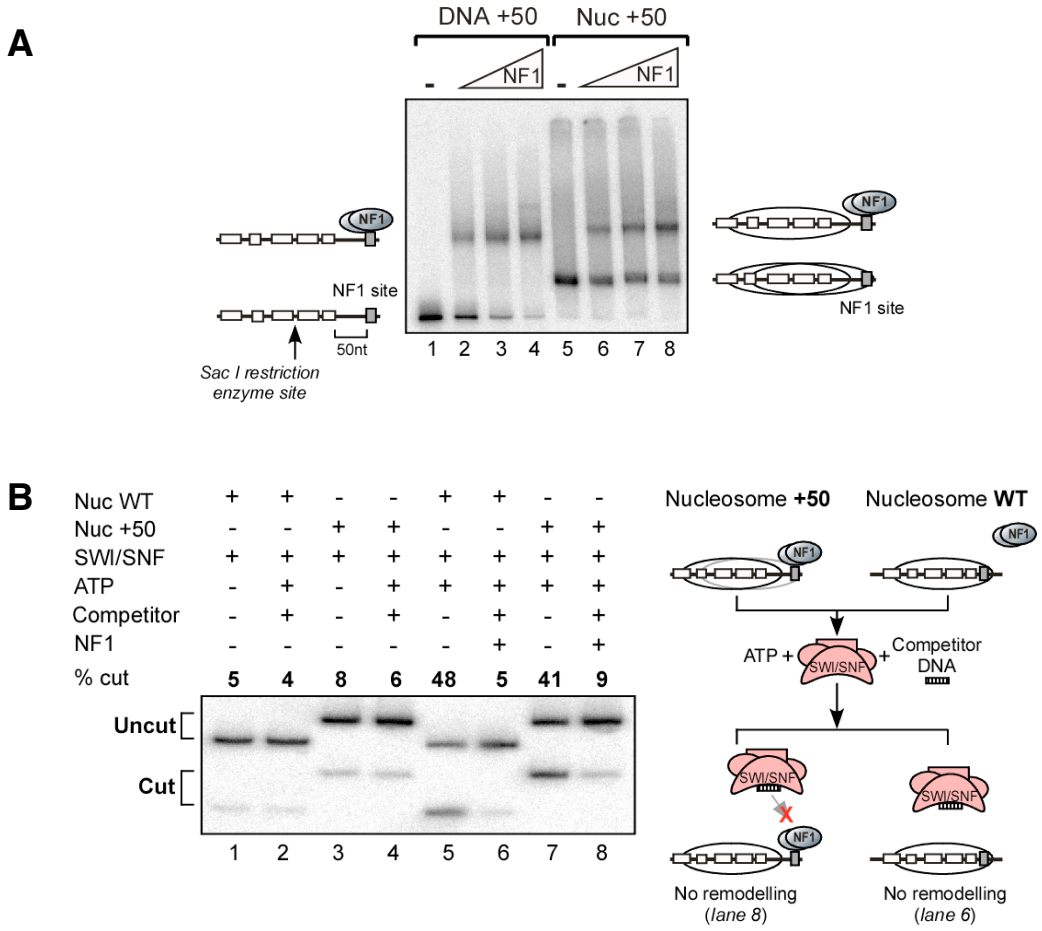


Figure 44. **NF1 cannot recruit SWI/SNF to MMTV nucleosome B. A.** Nucleosome B sequence carrying as insertion of 50 nucleotides (DNA +50) and reconstituted MMTV nucleosomes on this sequences of DNA (Nuc +50) were incubated at room temperature for 20 minutes with increasing amounts of recombinant NF1 in 20 ul reactions. The samples were electrophoresed on a 3.5% acrylamide/0.5% agarose/20% glycerol/0.3 x TBE gel. The interpretation of the bands is shown on the left for free DNA (DNA +50) and on the right for nucleosomes (Nuc +50). **B**. Reconstituted nucleosomes on the wild type nucleosome B sequence (WT) and nucleosomes reconstituted on the DNA+50 (Nuc +50) were treated with SWI/SNF in the presence of ATP, competitor DNA and NF1 as indicated. Following incubation at room temperature for 20 min, the nucleosomes were digested with SacI restriction enzyme. The DNA fragments were electrophoresed in 8% denaturing polyacrylamide gels. The % of cut DNA is shown on top of each lane in bold numbers.

To investigate whether NF1 is able to recruit SWI/SNF complex to the nucleosome B and promote histone displacement, I used SacI restriction enzyme accessibility assays with +50 nucleosomes compared to the wild type (WT) nucleosomes (**Figure 44B**). SacI target site is near the dyad axis (-104/-103; see scheme on Figure 44A) and is not accessible in the absence of SWI/SNF-dependent remodelling. Restriction endonuclease assay is used to determine structural changes

on nucleosome particles after specific remodelling reaction. Incubations with SWI/SNF and ATP (in the absence of competitor DNA and NF1) made the SacI restriction enzyme sites more accessible for cleavage similarly in WT and +50 nucleosomes (Figure 44B; lane 5 compared to lane 1 and lane 7 compared to lane 3, respectively), while no remodelling was observed in the presence of an excess of competitor DNA (Figure 44B lanes 2 and 4, respectively). When nucleosomes were preincubated with NF1, and in the presence of competitor DNA, no increase in SacI digestion was observed (Figure 44B lanes 6 and 8, respectively). Quantifications are shown in Figure 44B, line seven from the top. These results indicate that, in contrast to PR, NF1 transcription factor is not able to recruit SWI/SNF to MMTV nucleosomes *in vitro*, even when it is efficiently bound to the DNA.

Alternatively, NF1 could be bound too far from the nucleosome, since its site is now placed 50 nt more downstream in the linker DNA region. In order to role out this possibility I did the same experiment with a DNA carrying a shorter insertion: 30nt (data not shown). The addition of 30nt locates now the NF1 binding site right at the end of the nucleosomal DNA (while the edge of the nucleosome is in position -34, the new location for the NF1 binding site is between -44 and -31). Restriction enzyme accessibility assays performed in Nuc+30 gave similar results to those obtained with Nuc+50 (data not shown). Moreover, in the absence of competitor DNA, SWI/SNF was able to remodel the nucleosome even when NF1 was pre-bound to its binding site 30nt far from the original site, indicating that the presence of NF1 do not physically prevents SWI/SNF to remodel the nucleosome.

6. NF1 facilitates binding of PR to the HREs 2 and 3 after SWI/SNF remodelling

The reciprocal synergism between PR and NF1 has been interpreted in terms of NF1 stabilizing the open H2A/H2B depleted form of nucleosome B, but a more active role of NF1 cannot be excluded. I next sought to determine whether NF1 could have a role in PR binding to the inner HREs following SWI/SNF-mediated H2A/H2B displacement. I performed band shift assays in which we tested whether PR preferentially binds complexes of tetramer particles with NF1 rather than free tetramer particles (Figure 45). Reconstituted tetramer particles were pre-incubated with a fixed amount of NF1 that resulted in around 70% of the tetramers being in a free state and

roughly 30% in a complex with NF1 (TET-NF1) (**Figure 45A**, lane 2). Increasing amounts of PR were added in order to test its binding preference, which was followed as the disappearance of the bands corresponding to the complex of tetramer with NF1 and the tetramer alone (Figure 45A lanes 3-7). At low concentrations of PR, the TET-NF1 band decreased up to 70%, while the intensity of the free tetramer band was reduced only about 20% (Figure 45A, lanes 3-4, and quantification in **Figure 45B**). At high amounts of PR, both bands disappeared completely. These results imply that PR has higher binding affinity for an NF1-bound tetramer than for a free tetramer particle.

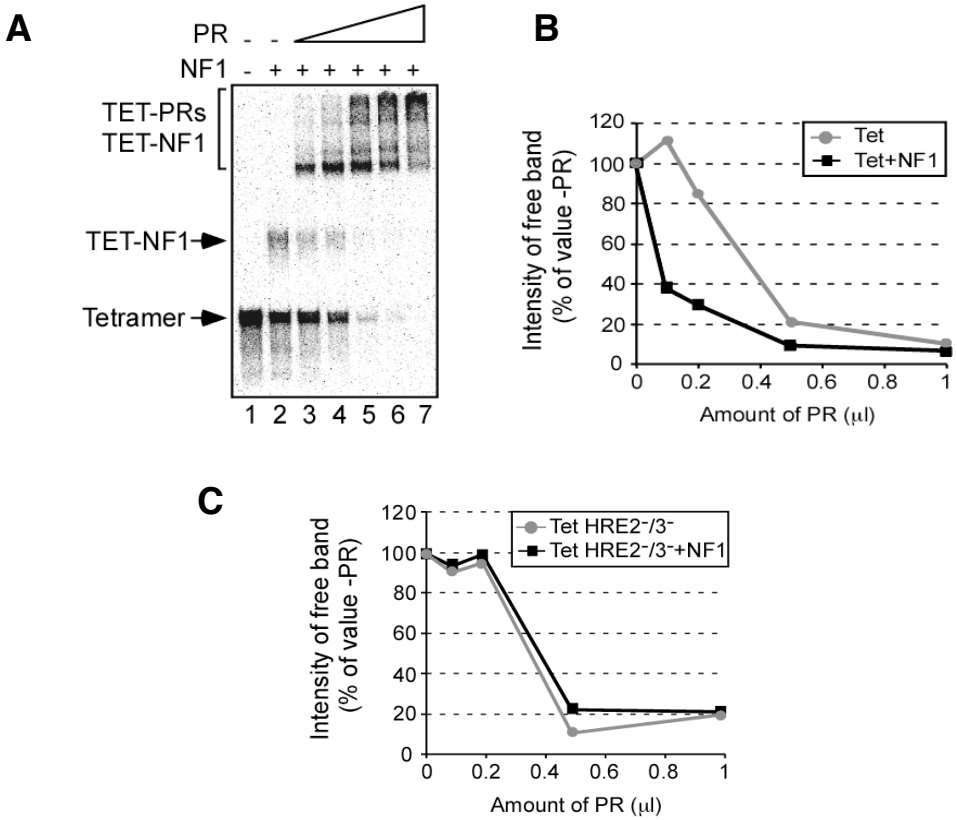


Figure 45. NF1 facilitates PR binding to the HREs 2 and 3. **A**. Reconstituted MMTV tetramer particles were incubated with a fixed amount of NF1 (*lane 2*) and increasing amounts of PR (*lane 3-7*). Samples were analyzed in a 5% acrylamide gel. **B**. Quantification of the bands corresponding to the free tetramers (TET) and the complex of tetramer and NF1 (TET-NF1), as percentage of the value in the absence of added PR, is plotted against the amount of added PR. **C**. Reconstituted HRE 2-/3- MMTV tetramer particles were subjected to the same experimenatl conditions explained in (*A* and *B*).

To test if the observed higher affinity of PR for TET-NF1 was due to PR binding to the internal HREs 2 and 3, we performed the same assay using tetramers reconstituted from a sequence carrying mutations in HREs 2 and 3 (**Figure 45C**). In this case, no preference of PR for TET-NF1 complex was found, the disappearance of the free tetramer and TET-NF1 complex bands followed the same trend starting from

the lowest to the highest concentration of PR. Thus, I reasoned that the previously detected difference in the affinity (Figure 45A and B) is due to the binding of PR to HREs 2 and 3 and not to HRE 1. Taken together, these results demonstrate that NF1 binding to its recognition site facilitates recruitment of PR to the inner HREs. This probably occurs through an NF1 dependent destabilization of the interaction of DNA with the H3/H4 tetramer that makes HREs 2 and 3 more accessible for PR binding. Accordingly, it is likely that NF1 plays a double structural role in the activation process of the MMTV promoter. First by preventing reforming of the octamer particle and second by labilizing histone-DNA interaction in the H3/H4 tetramer. Although I have observed that NF1 binding occurs prior to PR recruitment to HREs 2 and 3, I also tested if there is a significant difference in the affinity with which NF1 binds free tetramer particles and TET-PR complexes (**Figure 46A**). I incubated wild type tetramers with a PR concentration that results in 40% of the tetramer particles forming a complex with PR in which HRE 1 is occupied (Figure 46A lane 2). When increasing amounts of NF1 were added to the incubation mixture, I detected no significant difference between the binding affinity of NF1 for a TET-$PR_{(1)}$ complex and for a free tetramer particle (Figure 46A, lanes 3-6, and quantification in **Figure 46B**). These results imply that PR mediates NF1 binding through the induction of chromatin remodelling and histone displacement in the promoter region, but PR is not directly involved in the NF1 recruitment to its recognition site. Contrary to this, bound NF1 facilitates in a remodelling-independent manner an open conformation of the remodelled nucleosome optimal to PR recruitment and binding to HREs 2 and 3 (Figure 45).

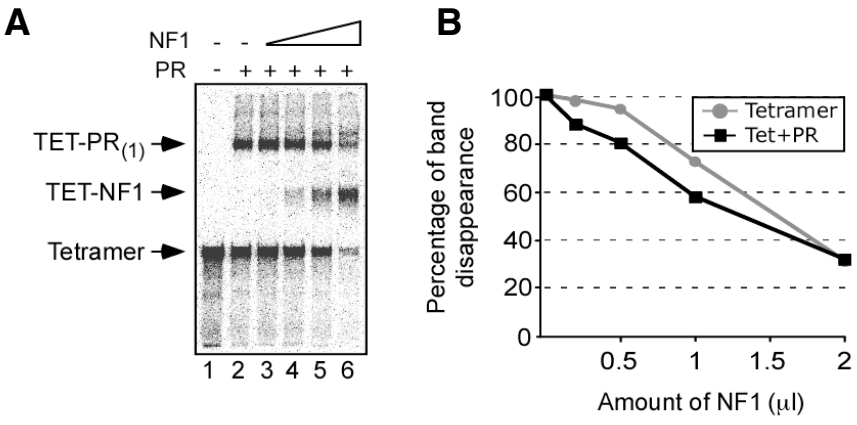


Figure 46. NF1 binds TET-PR complexes and free tetramers with similar affinity. A. Reconstituted MMTV tetramer particles were incubated with a fixed amount of PR (*lane 2*) and increasing amounts of NF1 (*lanes 3-6*) and the samples were analyzed in a 5% acrylamide gel. **B**. Quantification of the bands corresponding to free tetramer (TET) and the complex of tetramer with PR (TET-PR), as percentage of the value in the absence of added NF1, is plotted against the amount of added NF1.

Discussion

The results presented in this thesis contribute to a better understanding of the molecular mechanisms involved in activation of the MMTV promoter by progesterone. We have shown previously that PR interacts with an exposed HRE on the surface of a positioned nucleosome and initiates chromatin remodelling as a prerequisite for the subsequent steps of gene activation (Vicent et al, 2004). Here, I demonstrate the need of the hSWI/SNFα complex, known as BAF, for the initial chromatin remodelling and explore its mechanism of action. BAF is recruited at least in part via an interaction between PR and the BAF57 subunit of the complex and is anchored via an interaction with H3K14ac, likely generated by PCAF (Results Chapter 2). This suggests a mechanism for the cooperation between two types of chromatin remodelling activitites. The recruited BAF catalyzes the ATP dependent displacement of histones H2A/H2B needed for NF1 to gain access to its cognate site in the promoter. NF1 binding in turn stabilizes the remodelled nucleosome and facilitates recruitment of PR and associated BAF to the central HREs 2 and 3. Receptor binding to the central HREs 2 and 3 is necessary for full induction, but not for H2A/H2B displacement or NF1 binding. Finally, I show that PR and NF1 can bind synergistically to MMTV promoter sequence assembled on a tetramer of histones H3 and H4, a finding compatible with the idea that the hormone induced remodelling of MMTV chromatin generates a tetramer of H3/H4 that acts as a platform on which the activating complex of PR and associated kinases can orchestrate the assembly of co-activators and the basic transcriptional machinery.

The activated form of nucleosome B could be an H3/H4 tetramer particle

A reduction in the size of DNA protected against digestion with Micrococcal nuclease in mononucleosomes treated with recombinant SWI/SNF complex reflects structural changes on the histone composition of the core particle due to the remodelling (Truss et al, 1995). The appearance of a predominant subnucleosomal size fragment of 96 bp, that coincided with the one observed after MNase-treated tetramers, suggests similarities between the tetramer structure and the remodelled product of SWI/SNF-treated mononucleosomes. This observation is consistent with a

previous publication (Spangenberg et al, 1998), in which they studied tetramer positioning on the MMTV nucleosome B sequence. This work concluded, from DNA cleavage analysis, that the curvature of DNA is preserved between octamer and tetramer core particles except for the most proximal and distal two or three turns of the nucleosomal double helix, which are more accessible to nuclease in the tetramer than in the octamer. Consistent also with another previous reports where it was observed that assembly of an H3/H4 tetramer is sufficient to position at least 90 bp of DNA (Luger et al, 1997), I found that tetramers and SWI/SNF-treated octamers generate a 96 bp band in the MNase digestion. So, SWI/SNF-treated octamer particles protect the same length of DNA than H3/H4 tetramers and this is another indicative that the remodelled product could be a tetramer of histones H3 and H4.

Atomic Force Microscopy (AFM) studies of human SWI/SNF interacting with reconstituted MMTV promoter arrays have shown a decrease of 25-50% in nucleosome heights of some nucleosomes of the array after ATP-dependent remodelling, indicating formation of smaller structures probably due to histone H2A/H2B loss (Wang et al, 2005). Interestingly, they also showed hSWI/SNF complexes partially overlapping potentially histone-depleted nucleosomes (Wang et al, 2005). This result contradicts my observation of tetramer as weak substrates for SWI/SNF binding, although differences could be explained by the use of different systems and conditions. Some recent works has contributed to a better understanding on the SWI/SNF-nucleosome complex architecture and further remodelling actions (Sengupta et al, 2001, Wang et al, 2005, Dechassa et al, 2008) and specially a biochemical study on the yeast SWI/SNF complex has shown the acidic N terminus domain of the Swi3p subunit as a novel H2A/H2B-binding domain required for ATP-dependent H2A/H2B dimer loss (Yang et al, 2007). The direct interaction between Swi3p subunit and histones H2A and H2B may serve as well to anchor SWI/SNF to the nucleosome, thus, lack of both histone dimmer could results in weaker binding of the complex to the H3/H4 tetramer particle, as observed.

Whether active displacement of H2A/H2B dimmers is an exclusive phenomenon for the nucleosome B of the MMTV promoter or is a general effect of the SWI/SNF complex remained still to be further investigated. It will be very interesting to extent this study to other promoter or regulatory regions where SWI/SNF is acting and try to explain the biological function for which some nucleosomes suffer histone H2A/H2B

displacement. Our lab has proven that the contribution of the DNA sequence on the definition of the outcome of the remodelling process is essential (Vicent et al, 2006 and Sims et al, 2007). We have worked towards the finding of sequence features that lead SWI/SNF to one or other remodelling process, although results are still not conclusive. There are already some examples in the literature where SWI/SNF-mediated remodelling seems to be similar to the nucleosome B. The human ε-globin promoter region that is organized in well-positioned nucleosomes, with a nucleosome called N1 (the TATA-proximal nucleosome), is an interesting case. Transcriptional activation of the ε-globin gene is accompanied by chromatin structural changes limited to the N1 nucleosome. The altered structure is reflected in the increased sensitivity of N1 to nucleases and the generation of a DNase I Hypersensitive site (Gong et al, 1996). MNase digestion revealed that a subnucleosome-size N1 DNA fragment was released from actively transcribed promoters, while digestion of inactive promoters produced DNA fragments of nucleosomal length (Gui et al, 2001). Structural changes are accompanied by both H3 hyperacetylation and partially depletion of linker histone H1, as shown by *in vitro* ChIPs with reconstituted minichromosomes. It has been shown that recruitment of SWI/SNF to the nucleosome N1 is indispensable for remodelling and ε-globin promoter full activation (Gui et al, 2003). Thus, it would be interesting to study similar cases and understand the biological meaning of SWI/SNF-mediated displacement of histones H2A/H2B targeting specific nucleosome.

Synergistic binding of PR and NF1 on a H3/H4 tetramer structure

A tetramer of histones H3 and H4 positions MMTV promoter sequences in a similar way to histone octamers but NF1 can bind to an H3/H4 tetramer particle with relatively high affinity (Spangenberg et al 1998), although both octamer and tetramer populations cover the NF1 binding site (located at –75, respect to +1). Here I show that PR can acces all HREs in MMTV sequences positioned around an H3/H4 tetramer. More important, such particles can bind PR and NF1 simultaneously, reminiscent of what is observed in cells carrying a single copy of the MMTV promoter integrated in chromatin (Truss et al, 1995). Since NF1 recognition site and the HREs 3 are very closed in the DNA sequence (-75/-63 are the NF1 binding site boundaries while -98/-93 are the HRE 3 ones), it is plausible to believe that NF1 physically impedes PR binding on the HRE 2-4 cluster. Actually, it has been observed by DNAse I footprinting assays that no simultaneous binding of NF1 and PR in free DNA takes place (Di Croce et al,

1999). A physical impediment between the NF1 binding site and the HRE 2-4 cluster is the more feasible explanation. However, I observed that not only both factors bound simultaneously in reconstituted tetramer particles, but also pre-bound NF1 to the tetramer favours PR binding to inner HREs 2 and 3. The nature of this synergism is unknown but it does not result from a direct interaction between NF1 and PR (data not shown). Our interpretation is that the tridimentional structure of the DNA wrapped around an H3/H4 tetramer generates a new disposition of the NF1 binding site and the HRE 2-4, allowing the binding of both factors to its cognate sites on the DNA. Moreover, the presence of NF1 may cause a conformational change that favours the binding of more PR heterodimmers to the inner HRE 2 and 3, most likely as a prerequisite for full activation of the MMTV promoter.

Knocking down NF1c by siRNA in T47D-MTVL cells causes a drastic inhibition of the MMTV promoter activity, indicating that binding of NF1 is essential for promoter activation (Vicent et al, manuscript in preparation). The structural role of NF1 in the binding to the nucleosome B was first emphasized by Prado et al. (2002). They observed that the DNA binding domain of NF1, which is devoid of significant transactivation activity, still synergizes with glucocorticoid receptor (GR) in transactivation and chromatin remodelling, using both yeast system (with addition of recombinant GR and NF1) and *in vitro* transcription (with assembled minichromosomes in *Drosophila* embryo extracts and HeLa cell nuclear extract) (Prado et al, 2002). Thus, they showed that NF1 does not act like a conventional transcription factor in the context of the activated MMTV promoter, since removal of its proline-rich transactivation domain did not interact with the activation process of the MMTV promoter. As this structural effect is maintained when only the NF1 DNA binding domain is used with MMTV minichromosomes (Di Croce et al, 1999), it can be accounted for by at least two possible mechanism: direct competition between NF1 and histones H2A/H2B dimmers for DNA binding, or a deformation of the double helix upon NF1 binding favourable for PR access to HREs 2-4. Though we know that NF1 does not recruit BAF to the promoter, we cannot exclude that NF1 is being recruited to the promoter forming a complex with other co-activator factors that will help on the activation process.

For all these reasons, one of the main contributions of this work to the model we propose for MMTV activation is the essential role of NF1 binding prior to the second wave of PR recruitment to the internal HREs 2 and 3. Our hypothesis is that NF1

binding will act as a "switch" on the activation process. Once NF1 is bound, the transcriptional activation becomes irreversible. On the contrary, in cells lacking NF1, the nucleosome B undergoes the first steps of the activation process but, the open conformation is not stabilized by the binding of NF1 and thus neither PR binding to HREs 2 and 3 is possible nor full activation of the promoter. Thus, in the absence of NF1, the nucleosome B returns to the inactive state by deposition of H2A/H2B dimmers, binding of the HP1-γ repressive complex and displacement of BAF and other coactivators.

MMTV promoter activation: a model to better understand hormone-dependent promoters

Our present model for activation of the MMTV promoter is shown in Figure 47. Before hormone addition the MMTV promoter is silent due to its interaction with a repressive complex that includes HP1γ (**1**). The nature of this repressive complex is a current matter of investigation in the lab. Very rapidly after hormone addition (5-10 min), PR is phosphorylated by Erk at Ser293. Activated receptors bind to the exposed HRE1 as part of a ternary complex with activated Erk, activated Msk1, which has also been phosphorylated by Erk (Vicent et al, 2006). PR is also recruiting BAF complex to the nucleosome B through direct interaction of BAF57 subunit, as we demonstrated by GST pull down experiments. We cannot exclude an interaction between PR and other BAF complex. PCAF is present at short times after hormone induction on the promoter, but we still don't know if its recruitment is via a direct interaction with the progesterone receptor (**2**). The PR-Erk-Msk and PCAF complex phosphoacetylates H3, leading to H3S10PhosK14Ac, modifications that displace the repressive complex (**3**) and anchors the BAF complex which is stabilized specially by H3K14Ac (**4**). BAF catalyzes the ATP-dependent H2A/H2B displacement that facilitates NF1 binding (**5**). In the absence of NF1 the histone octamer particle is reformed and the HP1γ-containing repressive complex brings the promoter to the initial silenced state, thus preventing efficient activation. In this case, BAF is presumably displaced from the promoter as well as the other co-activators. How the nucleosome returns to a silent state is still unknown. Contrary, in the presence of NF1 bound to the H3/H4 tetramer particles (**6**) the reassociation of H2A/H2B dimmers is prevented and binding of further PR molecules and BAF complexes to the recently exposed HREs 2 and 3 is facilitated (**7**), leading to full promoter activation (**8**). It is worth to mention the key role on NF1 in cooperate with

the open conformation of the promoter to maintain accessible the inner HREs 2 and 3 for further PR binding. Whether this new PR is coming in a complex with more BAF and kinases or other factors it is still unclear. Actually, an open question that remains from this work is why the promoter still need BAF complex on the nucleosome after remodelling take place and is it anchor if the nucleosome has became an H3/H4 tetramer particle. How this full loaded promoter on an H3/H4 tetramer is further converted into a preinitiation complex remains to be established as well. We also do not know under which conditions the activated tetramer particle reverts to the inactive octamer state (**9**).

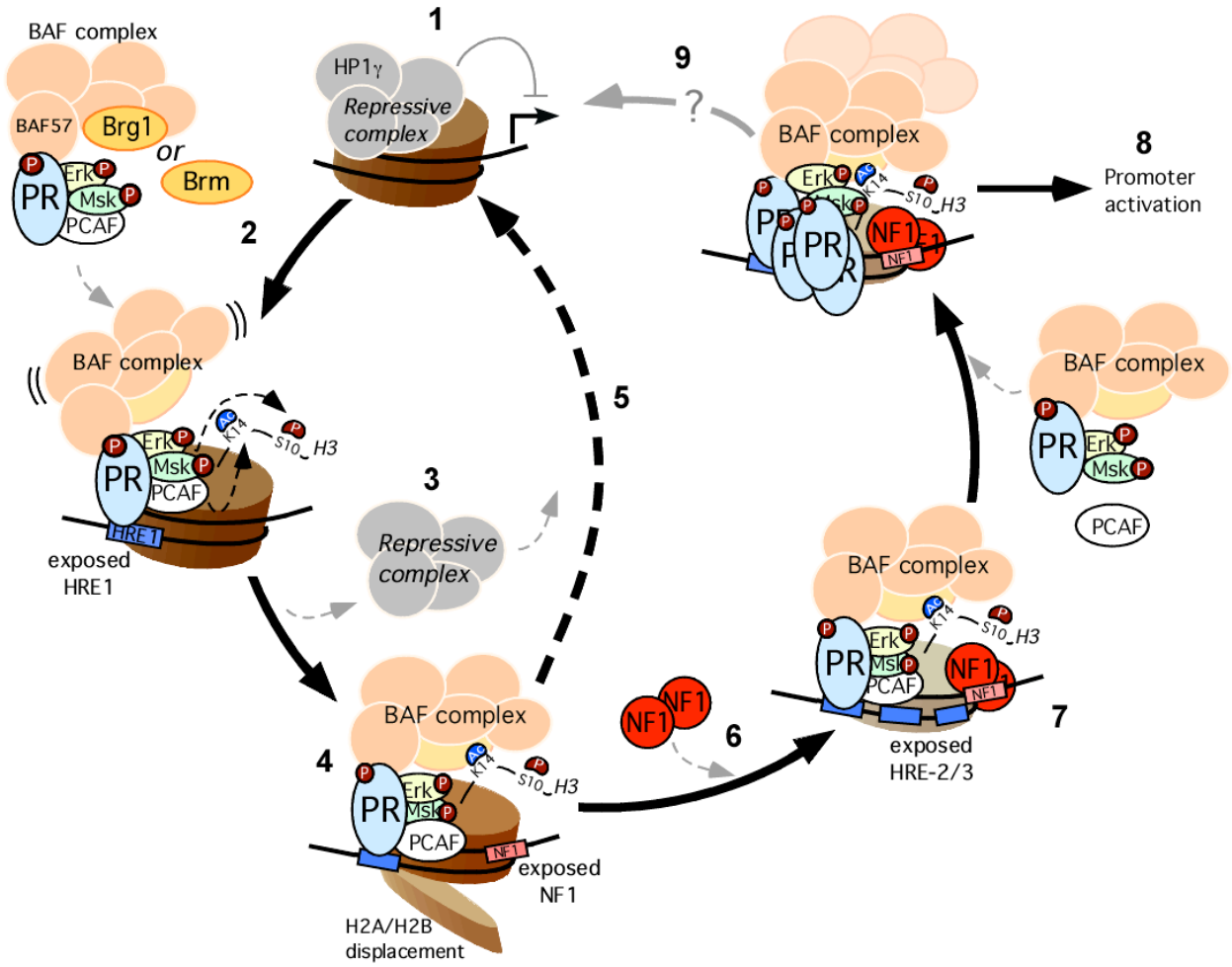


Figure 47. Model of the activation process of the MMTV promoter.

Conclusions

1. Both yeast and human nucleosomal DNA sequences exhibit symmetric patterns of curvature, while for human nucleosome sequences other features have been described as well: enrichment with CpG island regions, no contrains for sequence conservation among species and overrepresentation of certain triplet nucleotide patterns containing CG.

2. High SymCurv scores tend to coincide with well-positioned nucleosomes, suggesting the ability of the method to predict strong nucleosome positions with high accurancy.

3. “Key nucleosomes” are a subset of the whole nucleosome map with strong signal for positioning which is basically defined by the symmetry of curvature of its nucleosomal sequence.

4. “Key nucleosomes” are located in functional genomic regions such as Transcription Start Site and Splice junctions, suggesting its special role in cellular processes regulation.

5. BAF is required for progesterone gene activation in T47D-MTVL cells and is recruited to the MMTV promoter upon induction.

6. BAF57 and BAF250 interact with progesterone receptor in T47D-MTVL cells.

7. Acetylation on histone H3K14 by PCAF is required for hormonal induction and it anchors the BAF complex on the promoter.

8. The BAF complex facilitates NF1 binding by mediating histone H2A/H2B displacement.

9. H3/H4 histone tetramer is most likely the output of the SWI/SNF-dependent remodelling of the nucleosome B within the MMTV promoter.

10. The structure and binding properties of the reconstituted tetramer particles resemble the SWI/SNF-remodelled nucleosomes.

11. H3/H4 tetramer is not a suitable substrate for SWI/SNF-dependent remodelling.

12. NF1 binding is independent of PR binding in MMTV minichromosomes functional HREs 2 and 3.

13. NF1 cannot recruit SWI/SNF through direct interaction within the MMTV promoter.

14. NF1 facilitates binding of PR to the HREs 2 and 3 after SWI/SNF remodelling.

Supplementary articles

Yang X, Zaurin R, Beato M, Peterson CL.
Swi3p controls SWI/SNF assembly and ATP-dependent H2A-H2B displacement.
Nat Struct Mol Biol. 2007 Jun;14(6):540-7.

Vicent GP, Ballaré C, Zaurin R, Saragüeta P, Beato M.
Chromatin remodeling and control of cell proliferation by progestins via cross talk of progesterone receptor with the estrogen receptors and kinase signaling pathways.
Ann N Y Acad Sci. 2006 Nov;1089:59-72.

References

A

Abrams, E., Neigeborn, L. and Carlson, M. (1986). Molecular analysis of SNF2 and SNF5, genes required for expression of glucose-repressible genes in Saccharomyces cerevisiae. *Mol Cell Biol* **6**, 3643-51.

Ahmad, K. and Henikoff, S. (2002). The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell* **9**, 1191-200.

Albert, I., Mavrich, T. N., Tomsho, L. P., Qi, J., Zanton, S. J., Schuster, S. C. and Pugh, B. F. (2007). Translational and rotational settings of H2A.Z nucleosomes across the Saccharomyces cerevisiae genome. *Nature* **446**, 572-6.

Allis, C. D., Richman, R., Gorovsky, M. A., Ziegler, Y. S., Touchstone, B., Bradley, W. A. and Cook, R. G. (1986). hv1 is an evolutionarily conserved H2A variant that is preferentially associated with active genes. *J Biol Chem* **261**, 1941-8.

Angelov, D., Molla, A., Perche, P. Y., Hans, F., Cote, J., Khochbin, S., Bouvet, P. and Dimitrov, S. (2003). The histone variant macroH2A interferes with transcription factor binding and SWI/SNF nucleosome remodeling. *Mol Cell* **11**, 1033-41.

Aranda, A. and Pascual, A. (2001). Nuclear hormone receptors and gene expression. *Physiol Rev* **81**, 1269-304.

Arents, G., Burlingame, R. W., Wang, B. C., Love, W. E. and Moudrianakis, E. N. (1991). The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. *Proc Natl Acad Sci U S A* **88**, 10148-52.

Armstrong, J. A., Bieker, J. J. and Emerson, B. M. (1998). A SWI/SNF-related chromatin remodeling complex, E-RC1, is required for tissue-specific transcriptional regulation by EKLF in vitro. *Cell* **95**, 93-104.

B

Baldi, P., Brunak, S., Chauvin, Y. and Krogh, A. (1996). Naturally occurring nucleosome positioning signals in human exons and introns. *J Mol Biol* **263**, 503-10.

Ballare, C., Uhrig, M., Bechtold, T., Sancho, E., Di Domenico, M., Migliaccio, A., Auricchio, F. and Beato, M. (2003). Two domains of the progesterone receptor interact with the estrogen receptor and are required for progesterone activation of the c-Src/Erk pathway in mammalian cells. *Mol Cell Biol* **23**, 1994-2008.

Bao, Y., Konesky, K., Park, Y. J., Rosu, S., Dyer, P. N., Rangasamy, D., Tremethick, D. J., Laybourn, P. J. and Luger, K. (2004). Nucleosomes containing the histone variant H2A.Bbd organize only 118 base pairs of DNA. *Embo J* **23**, 3314-24.

Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I. and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823-37.

Bartsch, J., Truss, M., Bode, J. and Beato, M. (1996). Moderate increase in histone acetylation activates the mouse mammary tumor virus promoter and remodels its nucleosome structure. *Proc Natl Acad Sci U S A* **93**, 10741-6.
Batsche, E., Yaniv, M. and Muchardt, C. (2006). The human SWI/SNF subunit Brm is a regulator of alternative splicing. *Nat Struct Mol Biol* **13**, 22-9.
Bazett-Jones, D. P., Cote, J., Landel, C. C., Peterson, C. L. and Workman, J. L. (1999). The SWI/SNF complex creates loop domains in DNA and polynucleosome arrays and can disrupt DNA-histone contacts within these domains. *Mol Cell Biol* **19**, 1470-8.
Beato, M. (1989). Gene regulation by steroid hormones. *Cell* **56**, 335-44.
Beato, M., Herrlich, P. and Schutz, G. (1995). Steroid hormone receptors: many actors in search of a plot. *Cell* **83**, 851-7.
Beato, M. and Klug, J. (2000). Steroid hormone receptors: an update. *Hum Reprod Update* **6**, 225-36.
Becker, P. B., Gloss, B., Schmid, W., Strahle, U. and Schutz, G. (1986). In vivo protein-DNA interactions in a glucocorticoid response element require the presence of the hormone. *Nature* **324**, 686-8.
Bednar, J., Horowitz, R. A., Grigoryev, S. A., Carruthers, L. M., Hansen, J. C., Koster, A. J. and Woodcock, C. L. (1998). Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *Proc Natl Acad Sci U S A* **95**, 14173-8.
Belandia, B., Orford, R. L., Hurst, H. C. and Parker, M. G. (2002). Targeting of SWI/SNF chromatin remodelling complexes to estrogen-responsive genes. *Embo J* **21**, 4094-103.
Belikov, S. V. and Karpov, V. L. (1996). Mapping protein-DNA interactions with CIS-DDP: chromatin structure of promoter region of D. Melanogaster hsp 70 gene. *Biochem Mol Biol Int* **38**, 997-1002.
Biegel, J. A. and Pollack, I. F. (2004). Molecular analysis of pediatric brain tumors. *Curr Oncol Rep* **6**, 445-52.
Boeger, H., Griesenbeck, J., Strattan, J. S. and Kornberg, R. D. (2003). Nucleosomes unfold completely at a transcriptionally active promoter. *Mol Cell* **11**, 1587-98.
Boonyaratanakornkit, V., Scott, M. P., Ribon, V., Sherman, L., Anderson, S. M., Maller, J. L., Miller, W. T. and Edwards, D. P. (2001). Progesterone receptor contains a proline-rich motif that directly interacts with SH3 domains and activates c-Src family tyrosine kinases. *Mol Cell* **8**, 269-80.
Bowen, N. J., Fujita, N., Kajita, M. and Wade, P. A. (2004). Mi-2/NuRD: multiple complexes for many purposes. *Biochim Biophys Acta* **1677**, 52-7.
Bresnick, E. H., Bustin, M., Marsaud, V., Richard-Foy, H. and Hager, G. L. (1992). The transcriptionally-active MMTV promoter is depleted of histone H1. *Nucleic Acids Res* **20**, 273-8.
Bresnick, E. H., John, S., Berard, D. S., LeFebvre, P. and Hager, G. L. (1990). Glucocorticoid receptor-dependent disruption of a specific nucleosome on the mouse mammary tumor virus promoter is prevented by sodium butyrate. *Proc Natl Acad Sci U S A* **87**, 3977-81.
Bresnick, E. H., John, S. and Hager, G. L. (1991). Histone hyperacetylation does not alter the positioning or stability of phased nucleosomes on the mouse mammary tumor virus long terminal repeat. *Biochemistry* **30**, 3490-7.
Brown, C. E., Lechner, T., Howe, L. and Workman, J. L. (2000). The many HATs of transcription coactivators. *Trends Biochem Sci* **25**, 15-9.
Brown, D. T., Izard, T. and Misteli, T. (2006). Mapping the interaction surface of linker histone H1(0) with the nucleosome of native chromatin in vivo. *Nat Struct Mol Biol* **13**, 250-5.
Bruggemeier, U., Kalff, M., Franke, S., Scheidereit, C. and Beato, M. (1991). Ubiquitous transcription factor OTF-1 mediates induction of the MMTV promoter through synergistic interaction with hormone receptors. *Cell* **64**, 565-72.
Bruggemeier, U., Rogge, L., Winnacker, E. L. and Beato, M. (1990). Nuclear factor I acts as a transcription factor on the MMTV promoter but competes with steroid hormone receptors for DNA binding. *Embo J* **9**, 2233-9.
Bruno, M., Flaus, A., Stockdale, C., Rencurel, C., Ferreira, H. and Owen-Hughes, T. (2003). Histone H2A/H2B dimer exchange by ATP-dependent chromatin remodeling activities. *Mol Cell* **12**, 1599-606.
Bucci, L. R., Brock, W. A. and Meistrich, M. L. (1982). Distribution and synthesis of histone 1 subfractions during spermatogenesis in the rat. *Exp Cell Res* **140**, 111-8.

Byvoet, P., Shepherd, G. R., Hardin, J. M. and Noland, B. J. (1972). The distribution and turnover of labeled methyl groups in histone fractions of cultured mammalian cells. *Arch Biochem Biophys* **148**, 558-67.

C

Cai, Y., Jin, J., Florens, L., Swanson, S. K., Kusch, T., Li, B., Workman, J. L., Washburn, M. P., Conaway, R. C. and Conaway, J. W. (2005). The mammalian YL1 protein is a shared subunit of the TRRAP/TIP60 histone acetyltransferase and SRCAP complexes. *J Biol Chem* **280**, 13665-70.
Candau, R., Chavez, S. and Beato, M. (1996). The hormone responsive region of mouse mammary tumor virus positions a nucleosome and precludes access of nuclear factor I to the promoter. *J Steroid Biochem Mol Biol* **57**, 19-31.
Carlson, M. and Laurent, B. C. (1994). The SNF/SWI family of global transcriptional activators. *Curr Opin Cell Biol* **6**, 396-402.
Carlson, M., Osmond, B. C. and Botstein, D. (1981). Mutants of yeast defective in sucrose utilization. *Genetics* **98**, 25-40.
Castoria, G., Lombardi, M., Barone, M. V., Bilancio, A., Di Domenico, M., De Falco, A., Varricchio, L., Bottero, D., Nanayakkara, M., Migliaccio, A. et al. (2004). Rapid signalling pathway activation by androgens in epithelial and stromal cells. *Steroids* **69**, 517-22.
Chavez, S. and Beato, M. (1997). Nucleosome-mediated synergism between transcription factors on the mouse mammary tumor virus promoter. *Proc Natl Acad Sci U S A* **94**, 2885-90.
Chavez, S., Candau, R., Truss, M. and Beato, M. (1995). Constitutive repression and nuclear factor I-dependent hormone activation of the mouse mammary tumor virus promoter in Saccharomyces cerevisiae. *Mol Cell Biol* **15**, 6987-98.
Clausse, N., Baines, D., Moore, R., Brookes, S., Dickson, C. and Peters, G. (1993). Activation of both Wnt-1 and Fgf-3 by insertion of mouse mammary tumor virus downstream in the reverse orientation: a reappraisal of the enhancer insertion model. *Virology* **194**, 157-65.
Clements, A., Poux, A. N., Lo, W. S., Pillus, L., Berger, S. L. and Marmorstein, R. (2003). Structural basis for histone and phosphohistone binding by the GCN5 histone acetyltransferase. *Mol Cell* **12**, 461-73.
Corona, D. F. and Tamkun, J. W. (2004). Multiple roles for ISWI in transcription, chromosome organization and DNA replication. *Biochim Biophys Acta* **1677**, 113-9.
Costanzi, C. and Pehrson, J. R. (1998). Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* **393**, 599-601.
Cote, J., Peterson, C. L. and Workman, J. L. (1998). Perturbation of nucleosome core structure by the SWI/SNF complex persists after its detachment, enhancing subsequent transcription factor binding. *Proc Natl Acad Sci U S A* **95**, 4947-52.
Cote, J., Quinn, J., Workman, J. L. and Peterson, C. L. (1994). Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. *Science* **265**, 53-60.

D

de la Barre, A. E., Angelov, D., Molla, A. and Dimitrov, S. (2001). The N-terminus of histone H2B, but not that of histone H3 or its phosphorylation, is essential for chromosome condensation. *Embo J* **20**, 6383-93.

de Vries, E., van Driel, W., van den Heuvel, S. J. and van der Vliet, P. C. (1987). Contactpoint analysis of the HeLa nuclear factor I recognition site reveals symmetrical binding at one side of the DNA helix. *Embo J* **6**, 161-8.
Dechassa, M. L., Zhang, B., Horowitz-Scherer, R., Persinger, J., Woodcock, C. L., Peterson, C. L. and Bartholomew, B. (2008). Architecture of the SWI/SNF-nucleosome complex. *Mol Cell Biol* **28**, 6010-21.
Dennis, J. H., Fan, H. Y., Reynolds, S. M., Yuan, G., Meldrim, J. C., Richter, D. J., Peterson, D. G., Rando, O. J., Noble, W. S. and Kingston, R. E. (2007). Independent and complementary methods for large-scale structural analysis of mammalian chromatin. *Genome Res* **17**, 928-39.
Di Croce, L., Koop, R., Venditti, P., Westphal, H. M., Nightingale, K. P., Corona, D. F., Becker, P. B. and Beato, M. (1999). Two-step synergism between the progesterone receptor and the DNA-binding domain of nuclear factor 1 on MMTV minichromosomes. *Mol Cell* **4**, 45-54.
Dilworth, F. J. and Chambon, P. (2001). Nuclear receptors coordinate the activities of chromatin remodeling complexes and coactivators to facilitate initiation of transcription. *Oncogene* **20**, 3047-54.
Dirscherl, S. S. and Krebs, J. E. (2004). Functional diversity of ISWI complexes. *Biochem Cell Biol* **82**, 482-9.
Dominguez, V., Pina, B. and Suau, P. (1992). Histone H1 subtype synthesis in neurons and neuroblasts. *Development* **115**, 181-5.
Dong, F., Hansen, J. C. and van Holde, K. E. (1990). DNA and protein determinants of nucleosome positioning on sea urchin 5S rRNA gene sequences in vitro. *Proc Natl Acad Sci U S A* **87**, 5724-8.
Drew, H. R. and Travers, A. A. (1985). DNA bending and its relation to nucleosome positioning. *J Mol Biol* **186**, 773-90.
Duerre, J. A. and Lee, C. T. (1974). In vivo methylation and turnover of rat brain histones. *J Neurochem* **23**, 541-7.
Durgam, V. R. and Tekmal, R. R. (1994). The nature and expression of int-5, a novel MMTV integration locus gene in carcinogen-induced mammary tumors. *Cancer Lett* **87**, 179-86.

E

Edayathumangalam, R. S., Weyermann, P., Gottesfeld, J. M., Dervan, P. B. and Luger, K. (2004). Molecular recognition of the nucleosomal "supergroove". *Proc Natl Acad Sci U S A* **101**, 6864-9.
Espinas, M. L., Roux, J., Ghysdael, J., Pictet, R. and Grange, T. (1994). Participation of Ets transcription factors in the glucocorticoid response of the rat tyrosine aminotransferase gene. *Mol Cell Biol* **14**, 4116-25.

F

Field, Y., Kaplan, N., Fondufe-Mittendorf, Y., Moore, I. K., Sharon, E., Lubling, Y., Widom, J. and Segal, E. (2008). Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. *PLoS Comput Biol* **4**, e1000216.
Finch, J. T. and Klug, A. (1976). Solenoidal model for superstructure in chromatin. *Proc Natl Acad Sci U S A* **73**, 1897-901.
Fry, C. J. and Peterson, C. L. (2001). Chromatin remodeling enzymes: who's on first? *Curr Biol* **11**, R185-97.
Fryer, C. J. and Archer, T. K. (1998). Chromatin remodelling by the glucocorticoid receptor requires the BRG1 complex. *Nature* **393**, 88-91.

G

Gautier, T., Abbott, D. W., Molla, A., Verdel, A., Ausio, J. and Dimitrov, S. (2004). Histone variant H2ABbd confers lower stability to the nucleosome. *EMBO Rep* **5**, 715-20.
Gdula, D. A., Sandaltzopoulos, R., Tsukiyama, T., Ossipow, V. and Wu, C. (1998). Inorganic pyrophosphatase is a component of the Drosophila nucleosome remodeling factor complex. *Genes Dev* **12**, 3206-16.
Gong, Q. H., McDowell, J. C. and Dean, A. (1996). Essential role of NF-E2 in remodeling of chromatin structure and transcriptional activation of the epsilon-globin gene in vivo by 5' hypersensitive site 2 of the beta-globin locus control region. *Mol Cell Biol* **16**, 6055-64.
Goodsell, D. S. and Dickerson, R. E. (1994). Bending and curvature calculations in B-DNA. *Nucleic Acids Res* **22**, 5497-503.
Gui, C. Y. and Dean, A. (2001). Acetylation of a specific promoter nucleosome accompanies activation of the epsilon-globin gene by beta-globin locus control region HS2. *Mol Cell Biol* **21**, 1155-63.
Gui, C. Y. and Dean, A. (2003). A major role for the TATA box in recruitment of chromatin modifying complexes to a globin gene promoter. *Proc Natl Acad Sci U S A* **100**, 7009-14.
Gupta, S., Dennis, J., Thurman, R. E., Kingston, R., Stamatoyannopoulos, J. A. and Noble, W. S. (2008). Predicting human nucleosome occupancy from primary sequence. *PLoS Comput Biol* **4**, e1000134.

H

Hall, J. M. and Cole, R. D. (1985). Modulation in proportions of H1 histone subfractions by differential changes in synthesis and turnover during butyrate treatment of neuroblastoma cells. *Biochemistry* **24**, 7765-71.
Hassan, A. H., Neely, K. E. and Workman, J. L. (2001). Histone acetyltransferase complexes stabilize swi/snf binding to promoter nucleosomes. *Cell* **104**, 817-27.
Havas, K., Flaus, A., Phelan, M., Kingston, R., Wade, P. A., Lilley, D. M. and Owen-Hughes, T. (2000). Generation of superhelical torsion by ATP-dependent chromatin remodeling activities. *Cell* **103**, 1133-42.
Hayes, J. J., Clark, D. J. and Wolffe, A. P. (1991). Histone contributions to the structure of DNA in the nucleosome. *Proc Natl Acad Sci U S A* **88**, 6829-33.

I

Ioshikhes, I. P., Albert, I., Zanton, S. J. and Pugh, B. F. (2006). Nucleosome positions predicted through comparative genomics. *Nat Genet* **38**, 1210-5.

J

Jaskelioff, M., Gavin, I. M., Peterson, C. L. and Logie, C. (2000). SWI-SNF-mediated nucleosome remodeling: role of histone octamer mobility in the persistence of the remodeled state. *Mol Cell Biol* **20**, 3058-68.
Jenuwein, T. (2006). The epigenetic magic of histone lysine methylation. *Febs J* **273**, 3121-35.
Jenuwein, T. and Allis, C. D. (2001). Translating the histone code. *Science* **293**, 1074-80.
Jiang, C. and Pugh, B. F. (2009). Nucleosome positioning and gene regulation: advances through genomics. *Nat Rev Genet* **10**, 161-72.
Jin, J., Cai, Y., Yao, T., Gottschalk, A. J., Florens, L., Swanson, S. K., Gutierrez, J. L., Coleman, M. K., Workman, J. L., Mushegian, A. et al. (2005). A mammalian chromatin remodeling complex with similarities to the yeast INO80 complex. *J Biol Chem* **280**, 41207-12.

K

Kaplan, N., Moore, I. K., Fondufe-Mittendorf, Y., Gossett, A. J., Tillo, D., Field, Y., LeProust, E. M., Hughes, T. R., Lieb, J. D., Widom, J. et al. (2009). The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**, 362-6.
Karin, M., Haslinger, A., Holtgreve, H., Richards, R. I., Krauter, P., Westphal, H. M. and Beato, M. (1984). Characterization of DNA sequences through which cadmium and glucocorticoid hormones induce human metallothionein-IIA gene. *Nature* **308**, 513-9.
Katoh, M. (2002). WNT and FGF gene clusters (review). *Int J Oncol* **21**, 1269-73.
Khavari, P. A., Peterson, C. L., Tamkun, J. W., Mendel, D. B. and Crabtree, G. R. (1993). BRG1 contains a conserved domain of the SWI2/SNF2 family necessary for normal mitotic growth and transcription. *Nature* **366**, 170-4.
Khochbin, S. and Kao, H. Y. (2001). Histone deacetylase complexes: functional entities or molecular reservoirs. *FEBS Lett* **494**, 141-4.
Khorasanizadeh, S. (2004). The nucleosome: from genomic organization to genomic regulation. *Cell* **116**, 259-72.
Kogan, S. and Trifonov, E. N. (2005). Gene splice sites correlate with nucleosome positions. *Gene* **352**, 57-62.
Kouzarides, T. (1999). Histone acetylases and deacetylases in cell proliferation. *Curr Opin Genet Dev* **9**, 40-8.
Kouzarides, T. (2000). Acetylation: a regulatory modification to rival phosphorylation? *Embo J* **19**, 1176-9.
Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* **128**, 693-705.
Kruithof, M., Chien, F. T., Routh, A., Logie, C., Rhodes, D. and van Noort, J. (2009). Single-molecule force spectroscopy reveals a highly compliant helical folding for the 30-nm chromatin fiber. *Nat Struct Mol Biol* **16**, 534-40.
Kumar, V., Green, S., Stack, G., Berry, M., Jin, J. R. and Chambon, P. (1987). Functional domains of the human estrogen receptor. *Cell* **51**, 941-51.
Kwon, H., Imbalzano, A. N., Khavari, P. A., Kingston, R. E. and Green, M. R. (1994). Nucleosome disruption and enhancement of activator binding by a human SW1/SNF complex. *Nature* **370**, 477-81.

L

Lachner, M., Sengupta, R., Schotta, G. and Jenuwein, T. (2004). Trilogies of histone lysine methylation as epigenetic landmarks of the eukaryotic genome. *Cold Spring Harb Symp Quant Biol* **69**, 209-18.
Lee, C. H., Murphy, M. R., Lee, J. S. and Chung, J. H. (1999). Targeting a SWI/SNF-related chromatin remodeling complex to the beta-globin promoter in erythroid cells. *Proc Natl Acad Sci U S A* **96**, 12311-5.
Lee, D. Y., Hayes, J. J., Pruss, D. and Wolffe, A. P. (1993). A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* **72**, 73-84.
Lee, H. L. and Archer, T. K. (1994). Nucleosome-mediated disruption of transcription factor-chromatin initiation complexes at the mouse mammary tumor virus long terminal repeat in vivo. *Mol Cell Biol* **14**, 32-41.
Lee, H. L. and Archer, T. K. (1998). Prolonged glucocorticoid exposure dephosphorylates histone H1 and inactivates the MMTV promoter. *Embo J* **17**, 1454-66.
Lee, W., Tillo, D., Bray, N., Morse, R. H., Davis, R. W., Hughes, T. R. and Nislow, C. (2007). A high-resolution atlas of nucleosome occupancy in yeast. *Nat Genet* **39**, 1235-44.
Lehnertz, B., Ueda, Y., Derijck, A. A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T. and Peters, A. H. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* **13**, 1192-200.
Lennox, R. W. and Cohen, L. H. (1984). The alterations in H1 histone complement during mouse spermatogenesis and their significance for H1 subtype function. *Dev Biol* **103**, 80-4.
Lennox, R. W., Oshima, R. G. and Cohen, L. H. (1982). The H1 histones and their interphase phosphorylated states in differentiated and undifferentiated cell lines derived from murine teratocarcinomas. *J Biol Chem* **257**, 5183-9.
Li, B., Carey, M. and Workman, J. L. (2007). The role of chromatin during transcription. *Cell* **128**, 707-19.
Li, X., Wong, J., Tsai, S. Y., Tsai, M. J. and O'Malley, B. W. (2003). Progesterone and glucocorticoid receptors recruit distinct coactivator complexes and promote distinct patterns of local chromatin modification. *Mol Cell Biol* **23**, 3763-73.
Lieb, J. D. and Clarke, N. D. (2005). Control of transcription through intragenic patterns of nucleosome composition. *Cell* **123**, 1187-90.
Lindner, H. H. (2008). Analysis of histones, histone variants, and their post-translationally modified forms. *Electrophoresis* **29**, 2516-32.
Lorch, Y., Cairns, B. R., Zhang, M. and Kornberg, R. D. (1998). Activated RSC-nucleosome complex and persistently altered form of the nucleosome. *Cell* **94**, 29-34.
Lorch, Y., Zhang, M. and Kornberg, R. D. (1999). Histone octamer transfer by a chromatin-remodeling complex. *Cell* **96**, 389-92.
Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. and Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251-60.

M

Malik, H. S. and Henikoff, S. (2003). Phylogenomics of the nucleosome. *Nat Struct Biol* **10**, 882-91.
Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Schutz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P. et al. (1995). The nuclear receptor superfamily: the second decade. *Cell* **83**, 835-9.
Marmorstein, R. and Roth, S. Y. (2001). Histone acetyltransferases: function, structure, and catalysis. *Curr Opin Genet Dev* **11**, 155-61.

Mavrich, T. N., Jiang, C., Ioshikhes, I. P., Li, X., Venters, B. J., Zanton, S. J., Tomsho, L. P., Qi, J., Glaser, R. L., Schuster, S. C. et al. (2008). Nucleosome organization in the Drosophila genome. *Nature* **453**, 358-62.
McNally, J. G., Muller, W. G., Walker, D., Wolford, R. and Hager, G. L. (2000). The glucocorticoid receptor: rapid exchange with regulatory sites in living cells. *Science* **287**, 1262-5.
Mellor, J. (2005). The dynamics of chromatin remodeling at promoters. *Mol Cell* **19**, 147-57.
Miele, V., Vaillant, C., d'Aubenton-Carafa, Y., Thermes, C. and Grange, T. (2008). DNA physical properties determine nucleosome occupancy from yeast to fly. *Nucleic Acids Res* **36**, 3746-56.
Migliaccio, A., Di Domenico, M., Castoria, G., de Falco, A., Bontempo, P., Nola, E. and Auricchio, F. (1996). Tyrosine kinase/p21ras/MAP-kinase pathway activation by estradiol-receptor complex in MCF-7 cells. *Embo J* **15**, 1292-300.
Migliaccio, A., Piccolo, D., Castoria, G., Di Domenico, M., Bilancio, A., Lombardi, M., Gong, W., Beato, M. and Auricchio, F. (1998). Activation of the Src/p21ras/Erk pathway by progesterone receptor via cross-talk with estrogen receptor. *Embo J* **17**, 2008-18.
Mizuguchi, G., Shen, X., Landry, J., Wu, W. H., Sen, S. and Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* **303**, 343-8.
Mohrmann, L. and Verrijzer, C. P. (2005). Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes. *Biochim Biophys Acta* **1681**, 59-73.
Muchardt, C. and Yaniv, M. (1993). A human homologue of Saccharomyces cerevisiae SNF2/SWI2 and Drosophila brm genes potentiates transcriptional activation by the glucocorticoid receptor. *Embo J* **12**, 4279-90.
Muchardt, C. and Yaniv, M. (1999). The mammalian SWI/SNF complex and the control of cell growth. *Semin Cell Dev Biol* **10**, 189-95.
Mulac-Jericevic, B., Mullinax, R. A., DeMayo, F. J., Lydon, J. P. and Conneely, O. M. (2000). Subgroup of reproductive functions of progesterone mediated by progesterone receptor-B isoform. *Science* **289**, 1751-4.
Munteanu, M. G., Vlahovicek, K., Parthasarathy, S., Simon, I. and Pongor, S. (1998). Rod models of DNA: sequence-dependent anisotropic elastic modelling of local bending phenomena. *Trends Biochem Sci* **23**, 341-7.
Muthurajan, U. M., Park, Y. J., Edayathumangalam, R. S., Suto, R. K., Chakravarthy, S., Dyer, P. N. and Luger, K. (2003). Structure and dynamics of nucleosomal DNA. *Biopolymers* **68**, 547-56.

N

Nagata, K., Guggenheimer, R. A. and Hurwitz, J. (1983). Adenovirus DNA replication in vitro: synthesis of full-length DNA with purified proteins. *Proc Natl Acad Sci U S A* **80**, 4266-70.
Nasmyth, K. and Shore, D. (1987). Transcriptional regulation in the yeast life cycle. *Science* **237**, 1162-70.
Nie, Z., Xue, Y., Yang, D., Zhou, S., Deroo, B. J., Archer, T. K. and Wang, W. (2000). A specificity and targeting subunit of a human SWI/SNF family-related chromatin-remodeling complex. *Mol Cell Biol* **20**, 8879-88.

O

Onate, S. A., Prendergast, P., Wagner, J. P., Nissen, M., Reeves, R., Pettijohn, D. E. and Edwards, D. P. (1994). The DNA-bending protein HMG-1 enhances progesterone receptor binding to its target DNA sequences. *Mol Cell Biol* **14**, 3376-91.

Ozsolak, F., Song, J. S., Liu, X. S. and Fisher, D. E. (2007). High-throughput mapping of the chromatin structure of human promoters. *Nat Biotechnol* **25**, 244-8.

P

Panyim, S. and Chalkley, R. (1969). A new histone found only in mammalian tissues with little cell division. *Biochem Biophys Res Commun* **37**, 1042-9.
Peckham, H. E., Thurman, R. E., Fu, Y., Stamatoyannopoulos, J. A., Noble, W. S., Struhl, K. and Weng, Z. (2007). Nucleosome positioning signals in genomic DNA. *Genome Res* **17**, 1170-7.
Peterson, C. L. (2000). ATP-dependent chromatin remodeling: going mobile. *FEBS Lett* **476**, 68-72.
Peterson, C. L., Dingwall, A. and Scott, M. P. (1994). Five SWI/SNF gene products are components of a large multisubunit complex required for transcriptional enhancement. *Proc Natl Acad Sci U S A* **91**, 2905-8.
Pina, B., Bruggemeier, U. and Beato, M. (1990). Nucleosome positioning modulates accessibility of regulatory proteins to the mouse mammary tumor virus promoter. *Cell* **60**, 719-31.
Ponte, I., Vidal-Taboada, J. M. and Suau, P. (1998). Evolution of the vertebrate H1 histone class: evidence for the functional differentiation of the subtypes. *Mol Biol Evol* **15**, 702-8.
Prado, F., Koop, R. and Beato, M. (2002a). Accurate chromatin organization of the mouse mammary tumor virus promoter determines the nature of the synergism between transcription factors. *J Biol Chem* **277**, 4911-7.
Prado, F., Vicent, G., Cardalda, C. and Beato, M. (2002b). Differential role of the proline-rich domain of nuclear factor 1-C splice variants in DNA binding and transactivation. *J Biol Chem* **277**, 16383-90.

R

Rahman, A., Esmaili, A. and Saatcioglu, F. (1995). A unique thyroid hormone response element in the human immunodeficiency virus type 1 long terminal repeat that overlaps the Sp1 binding sites. *J Biol Chem* **270**, 31059-64.
Rangasamy, D., Greaves, I. and Tremethick, D. J. (2004). RNA interference demonstrates a novel role for H2A.Z in chromosome segregation. *Nat Struct Mol Biol* **11**, 650-5.
Reinke, H. and Horz, W. (2003). Histones are first hyperacetylated and then lose contact with the activated PHO5 promoter. *Mol Cell* **11**, 1599-607.
Richard-Foy, H. and Hager, G. L. (1987). Sequence-specific positioning of nucleosomes over the steroid-inducible MMTV promoter. *Embo J* **6**, 2321-8.
Richer, J. K., Jacobsen, B. M., Manning, N. G., Abel, M. G., Wolf, D. M. and Horwitz, K. B. (2002). Differential gene regulation by the two progesterone receptor isoforms in human breast cancer cells. *J Biol Chem* **277**, 5209-18.
Richmond, T. J. and Davey, C. A. (2003). The structure of DNA in the nucleosome core. *Nature* **423**, 145-50.
Rigaud, G., Roux, J., Pictet, R. and Grange, T. (1991). In vivo footprinting of rat TAT gene: dynamic interplay between the glucocorticoid receptor and a liver-specific factor. *Cell* **67**, 977-86.
Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S. and Bonner, W. M. (1998). DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem* **273**, 5858-68.

Rousseau-Merck, M. F., Versteege, I., Legrand, I., Couturier, J., Mairal, A., Delattre, O. and Aurias, A. (1999). hSNF5/INI1 inactivation is mainly associated with homozygous deletions and mitotic recombinations in rhabdoid tumors. *Cancer Res* **59**, 3152-6.

S

Saha, A., Wittmeyer, J. and Cairns, B. R. (2006). Chromatin remodelling: the industrial revolution of DNA around histones. *Nat Rev Mol Cell Biol* **7**, 437-47.

Sancho, M., Diani, E., Beato, M. and Jordan, A. (2008). Depletion of human histone H1 variants uncovers specific roles in gene expression and cell growth. *PLoS Genet* **4**, e1000227.

Sartorius, C. A., Melville, M. Y., Hovland, A. R., Tung, L., Takimoto, G. S. and Horwitz, K. B. (1994). A third transactivation function (AF3) of human progesterone receptors located in the unique N-terminal segment of the B-isoform. *Mol Endocrinol* **8**, 1347-60.

Scheidereit, C., Geisse, S., Westphal, H. M. and Beato, M. (1983). The glucocorticoid receptor binds to defined nucleotide sequences near the promoter of mouse mammary tumour virus. *Nature* **304**, 749-52.

Schneider, R., Bannister, A. J. and Kouzarides, T. (2002). Unsafe SETs: histone lysine methyltransferases and cancer. *Trends Biochem Sci* **27**, 396-402.

Schnitzler, G., Sif, S. and Kingston, R. E. (1998). Human SWI/SNF interconverts a nucleosome between its base state and a stable remodeled state. *Cell* **94**, 17-27.

Schnitzler, G. R., Cheung, C. L., Hafner, J. H., Saurin, A. J., Kingston, R. E. and Lieber, C. M. (2001). Direct imaging of human SWI/SNF-remodeled mono- and polynucleosomes by atomic force microscopy employing carbon nanotube tips. *Mol Cell Biol* **21**, 8504-11.

Schones, D. E., Cui, K., Cuddapah, S., Roh, T. Y., Barski, A., Wang, Z., Wei, G. and Zhao, K. (2008). Dynamic regulation of nucleosome positioning in the human genome. *Cell* **132**, 887-98.

Sengupta, S. M., VanKanegan, M., Persinger, J., Logie, C., Cairns, B. R., Peterson, C. L. and Bartholomew, B. (2001). The interactions of yeast SWI/SNF and RSC with the nucleosome before and after chromatin remodeling. *J Biol Chem* **276**, 12636-44.

Sevenet, N., Sheridan, E., Amram, D., Schneider, P., Handgretinger, R. and Delattre, O. (1999). Constitutional mutations of the hSNF5/INI1 gene predispose to a variety of cancers. *Am J Hum Genet* **65**, 1342-8.

Shen, X., Ranallo, R., Choi, E. and Wu, C. (2003). Involvement of actin-related proteins in ATP-dependent chromatin remodeling. *Mol Cell* **12**, 147-55.

Shivaswamy, S. and Iyer, V. R. (2008). Stress-dependent dynamics of global chromatin remodeling in yeast: dual role for SWI/SNF in the heat shock stress response. *Mol Cell Biol* **28**, 2221-34.

Simpson, R. T. (1978). Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. *Biochemistry* **17**, 5524-31.

Simpson, R. T. (1998). Chromatin structure and analysis of mechanisms of activators and repressors. *Methods* **15**, 283-94.

Sims, H. I., Lane, J. M., Ulyanova, N. P. and Schnitzler, G. R. (2007). Human SWI/SNF drives sequence-directed repositioning of nucleosomes on C-myc promoter DNA minicircles. *Biochemistry* **46**, 11377-88.

Smith, C. L., Archer, T. K., Hamlin-Green, G. and Hager, G. L. (1993). Newly expressed progesterone receptor cannot activate stable, replicated mouse mammary tumor virus templates but acquires transactivation potential upon continuous expression. *Proc Natl Acad Sci U S A* **90**, 11202-6.

Spangenberg, C., Eisfeld, K., Stunkel, W., Luger, K., Flaus, A., Richmond, T. J., Truss, M. and Beato, M. (1998). The mouse mammary tumour virus promoter positioned on a tetramer of histones H3 and H4 binds nuclear factor 1 and OTF1. *J Mol Biol* **278**, 725-39.

Spencer, T. E., Jenster, G., Burcin, M. M., Allis, C. D., Zhou, J., Mizzen, C. A., McKenna, N. J., Onate, S. A., Tsai, S. Y., Tsai, M. J. et al. (1997). Steroid receptor coactivator-1 is a histone acetyltransferase. *Nature* **389**, 194-8.

Stein, A. and Bina, M. (1999). A signal encoded in vertebrate DNA that influences nucleosome positioning and alignment. *Nucleic Acids Res* **27**, 848-53.
Stenoien, D. L., Nye, A. C., Mancini, M. G., Patel, K., Dutertre, M., O'Malley, B. W., Smith, C. L., Belmont, A. S. and Mancini, M. A. (2001a). Ligand-mediated assembly and real-time cellular dynamics of estrogen receptor alpha-coactivator complexes in living cells. *Mol Cell Biol* **21**, 4404-12.
Stenoien, D. L., Patel, K., Mancini, M. G., Dutertre, M., Smith, C. L., O'Malley, B. W. and Mancini, M. A. (2001b). FRAP reveals that mobility of oestrogen receptor-alpha is ligand- and proteasome-dependent. *Nat Cell Biol* **3**, 15-23.
Stern, M., Jensen, R. and Herskowitz, I. (1984). Five SWI genes are required for expression of the HO gene in yeast. *J Mol Biol* **178**, 853-68.

T

Tanaka, M., Hennebold, J. D., Macfarlane, J. and Adashi, E. Y. (2001). A mammalian oocyte-specific linker histone gene H1oo: homology with the genes for the oocyte-specific cleavage stage histone (cs-H1) of sea urchin and the B4/H1M histone of the frog. *Development* **128**, 655-64.
Th'ng, J. P., Sung, R., Ye, M. and Hendzel, M. J. (2005). H1 family histones in the nucleus. Control of binding and localization by the C-terminal domain. *J Biol Chem* **280**, 27809-14.
Thoma, F., Koller, T. and Klug, A. (1979). Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *J Cell Biol* **83**, 403-27.
Torchia, J., Rose, D. W., Inostroza, J., Kamei, Y., Westin, S., Glass, C. K. and Rosenfeld, M. G. (1997). The transcriptional co-activator p/CIP binds CBP and mediates nuclear-receptor function. *Nature* **387**, 677-84.
Truss, M., Bartsch, J., Hache, R. S. and Beato, M. (1993). Chromatin structure modulates transcription factor binding to the mouse mammary tumor virus (MMTV) promoter. *J Steroid Biochem Mol Biol* **47**, 1-10.
Truss, M., Bartsch, J., Schelbert, A., Hache, R. J. and Beato, M. (1995). Hormone induces binding of receptors and transcription factors to a rearranged nucleosome on the MMTV promoter in vivo. *Embo J* **14**, 1737-51.

U

Uyttendaele, H., Marazzi, G., Wu, G., Yan, Q., Sassoon, D. and Kitajewski, J. (1996). Notch4/int-3, a mammary proto-oncogene, is an endothelial cell-specific mammalian Notch gene. *Development* **122**, 2251-9.

V

Valouev, A., Ichikawa, J., Tonthat, T., Stuart, J., Ranade, S., Peckham, H., Zeng, K., Malek, J. A., Costa, G., McKernan, K. et al. (2008). A high-resolution, nucleosome position map of C. elegans reveals a lack of universal sequence-dictated positioning. *Genome Res* **18**, 1051-63.
Venditti, P., Di Croce, L., Kauer, M., Blank, T., Becker, P. B. and Beato, M. (1998). Assembly of MMTV promoter minichromosomes with positioned nucleosomes precludes NF1 access but not restriction enzyme cleavage. *Nucleic Acids Res* **26**, 3657-66.

Venters, B. J. and Pugh, B. F. (2009). A canonical promoter organization of the transcription machinery and its regulators in the Saccharomyces genome. *Genome Res* **19**, 360-71.
Versteege, I., Sevenet, N., Lange, J., Rousseau-Merck, M. F., Ambros, P., Handgretinger, R., Aurias, A. and Delattre, O. (1998). Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. *Nature* **394**, 203-6.
Vicent, G. P., Ballare, C., Nacht, A. S., Clausell, J., Subtil-Rodriguez, A., Quiles, I., Jordan, A. and Beato, M. (2006a). Induction of progesterone target genes requires activation of Erk and Msk kinases and phosphorylation of histone H3. *Mol Cell* **24**, 367-81.
Vicent, G. P., Ballare, C., Nacht, A. S., Clausell, J., Subtil-Rodriguez, A., Quiles, I., Jordan, A. and Beato, M. (2008). Convergence on chromatin of non-genomic and genomic pathways of hormone signaling. *J Steroid Biochem Mol Biol* **109**, 344-9.
Vicent, G. P., Ballare, C., Zaurin, R., Saragueta, P. and Beato, M. (2006b). Chromatin remodeling and control of cell proliferation by progestins via cross talk of progesterone receptor with the estrogen receptors and kinase signaling pathways. *Ann N Y Acad Sci* **1089**, 59-72.
Vicent, G. P., Koop, R. and Beato, M. (2002a). Complex role of histone H1 in transactivation of MMTV promoter chromatin by progesterone receptor. *J Steroid Biochem Mol Biol* **83**, 15-23.
Vicent, G. P., Melia, M. J. and Beato, M. (2002b). Asymmetric binding of histone H1 stabilizes MMTV nucleosomes and the interaction of progesterone receptor with the exposed HRE. *J Mol Biol* **324**, 501-17.
Vicent, G. P., Nacht, A. S., Smith, C. L., Peterson, C. L., Dimitrov, S. and Beato, M. (2004). DNA instructed displacement of histones H2A and H2B at an inducible promoter. *Mol Cell* **16**, 439-52.
Voegel, J. J., Heine, M. J., Tini, M., Vivat, V., Chambon, P. and Gronemeyer, H. (1998). The coactivator TIF2 contains three nuclear receptor-binding motifs and mediates transactivation through CBP binding-dependent and -independent pathways. *Embo J* **17**, 507-19.

W

Wagner, B. L., Norris, J. D., Knotts, T. A., Weigel, N. L. and McDonnell, D. P. (1998). The nuclear corepressors NCoR and SMRT are key regulators of both ligand- and 8-bromo-cyclic AMP-dependent transcriptional activity of the human progesterone receptor. *Mol Cell Biol* **18**, 1369-78.
Wang, F., Zhang, R., Beischlag, T. V., Muchardt, C., Yaniv, M. and Hankinson, O. (2004). Roles of Brahma and Brahma/SWI2-related gene 1 in hypoxic induction of the erythropoietin gene. *J Biol Chem* **279**, 46733-41.
Wang, H., Bash, R., Lindsay, S. M. and Lohr, D. (2005). Solution AFM studies of human Swi-Snf and its interactions with MMTV DNA and chromatin. *Biophys J* **89**, 3386-98.
Wang, W. (2003). The SWI/SNF family of ATP-dependent chromatin remodelers: similar mechanisms for diverse functions. *Curr Top Microbiol Immunol* **274**, 143-69.
Wang, W., Cote, J., Xue, Y., Zhou, S., Khavari, P. A., Biggar, S. R., Muchardt, C., Kalpana, G. V., Goff, S. P., Yaniv, M. et al. (1996a). Purification and biochemical heterogeneity of the mammalian SWI-SNF complex. *Embo J* **15**, 5370-82.
Wang, W., Xue, Y., Zhou, S., Kuo, A., Cairns, B. R. and Crabtree, G. R. (1996b). Diversity and specialization of mammalian SWI/SNF complexes. *Genes Dev* **10**, 2117-30.
Whitehouse, I., Flaus, A., Cairns, B. R., White, M. F., Workman, J. L. and Owen-Hughes, T. (1999). Nucleosome mobilization catalysed by the yeast SWI/SNF complex. *Nature* **400**, 784-7.
Whitehouse, I., Rando, O. J., Delrow, J. and Tsukiyama, T. (2007). Chromatin remodelling at promoters suppresses antisense transcription. *Nature* **450**, 1031-5.
Whitehouse, I. and Tsukiyama, T. (2006). Antagonistic forces that position nucleosomes in vivo. *Nat Struct Mol Biol* **13**, 633-40.
Widlund, H. R., Cao, H., Simonsson, S., Magnusson, E., Simonsson, T., Nielsen, P. E., Kahn, J. D., Crothers, D. M. and Kubista, M. (1997). Identification and characterization of genomic nucleosome-positioning sequences. *J Mol Biol* **267**, 807-17.
Wolffe, A. P. (1994). Transcriptional activation. Switched-on chromatin. *Curr Biol* **4**, 525-8.

X

Xue, Y., Canman, J. C., Lee, C. S., Nie, Z., Yang, D., Moreno, G. T., Young, M. K., Salmon, E. D. and Wang, W. (2000). The human SWI/SNF-B chromatin-remodeling complex is related to yeast rsc and localizes at kinetochores of mitotic chromosomes. *Proc Natl Acad Sci U S A* **97**, 13015-20.

Y

Yan, Z., Cui, K., Murray, D. M., Ling, C., Xue, Y., Gerstein, A., Parsons, R., Zhao, K. and Wang, W. (2005). PBAF chromatin-remodeling complex requires a novel specificity subunit, BAF200, to regulate expression of selective interferon-responsive genes. *Genes Dev* **19**, 1662-7.

Yang, X., Zaurin, R., Beato, M. and Peterson, C. L. (2007). Swi3p controls SWI/SNF assembly and ATP-dependent H2A-H2B displacement. *Nat Struct Mol Biol* **14**, 540-7.

Yang, X. J., Ogryzko, V. V., Nishikawa, J., Howard, B. H. and Nakatani, Y. (1996). A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature* **382**, 319-24.

Yoshinaga, S. K., Peterson, C. L., Herskowitz, I. and Yamamoto, K. R. (1992). Roles of SWI1, SWI2, and SWI3 proteins for transcriptional enhancement by steroid receptors. *Science* **258**, 1598-604.

Yuan, G. C., Liu, Y. J., Dion, M. F., Slack, M. D., Wu, L. F., Altschuler, S. J. and Rando, O. J. (2005). Genome-scale identification of nucleosome positions in S. cerevisiae. *Science* **309**, 626-30.

Z

Zhao, Q., Chasse, S. A., Devarakonda, S., Sierk, M. L., Ahvazi, B. and Rastinejad, F. (2000a). Structural basis of RXR-DNA interactions. *J Mol Biol* **296**, 509-20.

Zhao, X. Y., Malloy, P. J., Krishnan, A. V., Swami, S., Navone, N. M., Peehl, D. M. and Feldman, D. (2000b). Glucocorticoids can promote androgen-independent growth of prostate cancer cells through a mutated androgen receptor. *Nat Med* **6**, 703-6.

Zhou, Y. B., Gerchman, S. E., Ramakrishnan, V., Travers, A. and Muyldermans, S. (1998). Position and orientation of the globular domain of linker histone H5 on the nucleosome. *Nature* **395**, 402-5.

Agraïments

Gracias Miguel. La primera vez que entré al CRG era para tener una entrevista corta contigo, tenías que aceptarme como estudiante de prácticas por seis meses. Glòria me dijo que esperara en una sala en aquel pasillo de administración cuando el CRG era todavía de miniatura. Llegaste tarde y corriendo, vestido con camisa pero con calzado de deporte (!); uno de los cristales de tus gafas redonditas estaba partido en dos. La imagen que siempre había tenido del típico científico loco se hizo real de repente! Y después de cinco años puedo decirte que esa primera impresión no estaba muy lejos de la realidad…Y que siga siendo así por muchos años! Gracias por tu confianza des de muy al principio, de creer que podía hacer una buena tesis. Gracias también por creer tanto en mi proyecto, a veces más que yo misma incluso; por darle prioridad y mostrar mis resultados. Por darme la oportunidad de asistir a tantos congresos y meetings, que han ayudado tanto a mantener mi motivación, como a tener nuevas ideas, a conocer a gente interesante y también a tener una visión cada vez más clara de lo que es el mundo de la ciencia.
Pero sobretodo, quería darte las gracias por tu apoyo durante este último año, en que he dedicado un poco más de tiempo a otras cosas que no al doctorado y siempre he encontrado de tu parte una sonrisa y un “bienvenida otra vez, vamos a hablar de ciencia”.

Guille…me siento un poco ridícula escribiéndote… después de cinco años de verte (y escucharte), de tenerte al lado cada día. Escribirte es como si ya no estuvieras o estuvieras lejos o lo estuviera yo, que es lo que va a pasar. Porque va a pasar… qué raro imaginarlo, no. ¿Cómo se deshace una simbiosis? Pero ya que te escribo, por una vez que sean cosas serias. Que quede por escrito que estos años hubieran sido peores sin: las medias rosas y las minifaldas cortas, sin la Bersuit y los Piojos, sin las peleas y las reconciliaciones, sin los “das”, las agas y los angs, sin el boomerang que no vuelve, sin el muere lentamente que no es de Pablo, o la cabra que tenías y querías más que a tus hijos (Silvina, ¿sabías?), sin tantos que pasaron y fuimos superando (…), sin colarnos a los meetings des de Heidelberg al de Meritxell, sin el río de la plata que es en realidad un trozo de mar, sin tantas cosas que han ido pasando y quedarán…pero sobretodo sin ti. Esta tesis hubiera tenido otro color sin ti. Cuantas cosas aprendidas y desaprendidas… pero eso me lo quedo para mi. Gracias. Y que nunca se acaben estos tuyos microlitros para los (¿mejor las?) que van a seguir viniendo con ganas de conocer lo que es la cromatina...

Silvina, me prometí no decir nada de la pastafrola en los agradecimiento…pero no puedo resistir. Todos tiene que saber que haces la mejor pastafrola del mundo. Y Cecita lo sabe bien….(Jofre menos porque casi nunca llega a tiempo a pillar el trocito ¡qué pena!).
Qué bien haberte tenido cerca durante estos años, Silvina. Aguantándome, aguantándole, aguantándonos. Con tu aparente tranquilidad y buen humor, siempre allí. Un equipo… que incluye Cecita. Entre las dos, ¿cuántas preguntas de experimentos tontos habréis contestado con paciencia a estudiantes incapaces como

nosotros? Un monumento, Cecita. Te lo dije y ahora te lo escribo, a ver si Miguel lo lee y convierte tu poyata en un altar. Qué dúo hemos formado al final, eh! compartiendo la pasión por los bioinformáticos, esa pasión masoquista pero que crea dependencia... Con tablas que colapsan el ordenador, el Excel y el mundo entero. Pero nosotras seguimos allí con más y más reads. Gracias también por estar y que nos vaya bien en todo lo que emprendamos.

I tampoc em puc imaginar com hagués estat tot sense el Jofre, el Jaume i la seva radio. Quins temps els d'aquell estiu amb el Sinatra, Jaume… dies i nits. I amb el Vlad! Van ser uns bons inicis…érem un laboratori acollidor, a la nostra manera. I qué Jofre? …ara ja només quedaràs tu d'allò. Un plaer haver estat també amb tu aquests anys. Després de tot, casi com germans….Quina pena que s'acabi… Espavila que et vull veure defensar la TEVA tesis.

Gracias a tres investigadores que por suerte o por desgracia les tocó perder mucho tiempo con mi proyecto… Gracias a Elena, Pura y Juan... Y a Juan, especialmente!

A los que han llegado más tarde, hacia el final de este camino, quiero también dar las gracias por momentos bonitos, por querer conocer y compartir, por darme cada día un poquito más de vuestra amistad.
A: François (mi François), Elisa, Oscar, Gari, Mariana, Susana , los de inglés y Maños y demás… Adonis i Josep, por momentos bonitos, Roni, Marija, Michael, Diana, los del lab, Antoine, un momento fugaz, Marija, Laure, Ale, Imma, la "la mujé má bonita", i totes les secres guapes que ens fan la vida més fácil, Vincent y tu bioética, i a tots els SMSs amb els que he compartit reunions, cafès científics i opens days…

Iris. Sempre ens hem definit per no haver de posar paraules a res. I no és perquè no sàpiga el que vull dir, però trobo tant absurd donar-li una definició…Estem. Ens tenim. I no cal més. Gràcies, per un ser incondicional, atemporal.
Indefinible.

I l'Ana… on estàs últimament? O més ben dit on no estic jo… Bé, gràcies també per moure't aprop meu, per les converses interminables de passadís. Per una tesi que ens ha unit, i que duri!

Aquests anys han estat també els de més canvi personal, (i el que falta encara)…mai oblidaré els inicis amb festes de Sant Joan al pati de casa, amb les llokes, el Faren, el Tange, i després més tranquil·les ja, a Minyons i als Amics. Molta activitat i també molt suport encara que no entenieu res. I veu començar a entendre menys i menys quan ja no pujava cada cap de setmana, quan no arribava als 2x1 i marxava sense avisar dels locals (recordeu?). Jo si. I recordo també com em deieu "quan tornaràs a la ser la Rousy d'abans?" i fa poc encara t'ho vaig sentir dir, Txe... Però els processos es fan per avançar i no tornar o per avançar guardant trocets. Però no parar d'avançar. La tesi i les meves ganes de dedicar-li la vida han fet que el camí de 45 minuts en tren es fes més i més llarg, però també la distància dóna noves perspectives. I el que s'ha compartit no t'ho treu ningú. Ja sabrem retrobar el que, potser algun dia, es va quedar en un tren perdut entre Terrassa i Barcelona. Gràcies per haver entès tant bé, sense entendre res, que l'única manera de fer front a un tesi, és sent pacient.

També el Paral·lel ha estat part d'aquest primers anys… amb la Neus i la Trip. I… la Núria, l'Anna, el Ricard i l'Ana. Amb qui més he compartit perquè heu pogut i volgut entendre tant del que ha anat passant. Hem passat junts aquest procés de saltar sense xarxa i no ens ha anat gens malament. L'èxit més gran: haver conservat

l'amistat durant deu anys, que es diu ràpid. Quin orgull. Gràcies pel trocet que cada un de vosaltres ha donat per seguir, seguir aquí. I la colla es va fer gran i qué bé! Mateu. Martí, Maria, Edu, fins i tot la més petita estarà en aquesta tesi, la Mar…

Ferran, tu véns ara. Et toca, sentir-te dir que si vols caldo, aquí en tens 200 pàgines. Quants Iposes, i demés que han calgut per pujar els ànims dels moments més difícils. Però veus, poc a poc, treball de formigueta, i l'amistat es consolido (jo diria que és ja un totxo…), com s'han consolidat els dimecres. T'enrecordes quan ho vas propasar per primera vegada. Quina bona idea! Som de dimecres nosaltres…. Aquesta tesi té molt de tu, encara que no ho sembli. Qui sap quants viatges a Laos, Brasil, Buenos Aires hauré fet mentalment mentre corria un gel, esperava que acabés una PCR o feia temps per assecar el pellet de DNA! Regals d'imaginació. Mil i una gràcies. Pai. pai.

…El Raval, els Pescadors, Alaska, l'Americano… i caminant cap a la Barceloneta, carrer del Mar, la Soli, la Lancetta, il Botel, i els pieroghi de Sopot (con gocetto), la Voivodina (nyema problema), la senyora Rosa i moviments de cap, Danubi, i la Daisy de Camagüey, i més, molt més que m'has donat per fer que aquests anys no fossin només hores i hores de córrer gels, mirar bandes i preparar meetings. La tesi de vida que he fet, en paral·lel a aquesta, te la dec, i dedico, a tu, ho saps però t'ho torno a dir. Azzie…

Y llega el final y apareces tú. Qué mal momento! Pero que bueno…Nos encontramos en la Antártica, como dos pingüinos perdidos y qué bien estar perdido en la Antártica cuando allí fuera pasan cosas tan malas… ¿Te di las gracias? Seguro que no lo suficiente. Gracias por estar cerquita en los peores momentos, para seguir luchando y por entenderme tan bien en estos últimos días. Intensida. Qué ganas de volver a la Antártica! ¿Nos vamos?

Si durant cinc anys de tesis, hi ha moments en que es té poc temps pels que t'envolten, sense cap mena de dubte, els que sortiu més malparats sou vosaltres, família! Qué difícil és combinar un diumenge per dinar plegats o un divedres per un café, o pujar a Terrassa a veure la iaia, o estar aquí per nadal… quants poquets moments compartits, si miro enrera, però cap retret. Quina paciència…. Quan d'amor perquè si. Moltíssimes gràcies, de veritat. Especialment a mama, papa i Mar.

Solenn, que estarás ahora con tus abejorros, también hay un rinconcito para ti. Volando se ven mejor las cosa, no? Fue al final pero apareciste, y contigo tu pandilla: Ale, Karolos, Philipos, Davina, Judit,… otro mundo para mi. Y de Sol a…
Isita, como tú, yo igual. Nos define, verdad? Los últimos días sí. Qué placer haberte encontrado de paso y así poquito a poquito juntándonos para darnos soporte en momentos difíciles, para disfrutar de otros bonitos… poquet a poquet com es fan les coses ben fetes. Y qué experiencia la de escribir en paralelo. Para sufrir el doble pero cuando haya pasado todo, para disfrutar el doble! Seguimos…

El final, el final serà per les que acaben d'arribar… poc temps però intens, íntim i profund. A les binet filastin: Ari i Pili. Que també m'heu acompanyat….

….

Gràcies a tots per estar a prop en aquest camí que encara que no hagi resultat fàcil, em sento orgullosa d'haver-lo fet, i haver-lo fet a prop de vosaltres. I ara, a seguir **nadant**…

I si després de tot, no queda res?
Entonces…… ¿Qué?