Phosphorylation and distribution of High-Mobility Group protein HMGN1 in the context of Immediate-Early (IE) gene induction

Edgar Allan Pogna

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

46.204 words

Phosphorylation and distribution of High-Mobility Group protein HMGN1 in the context of Immediate-Early (IE) gene induction

Edgar A. Pogna
Trinity College, University of Oxford
A thesis submitted for the degree of Doctor of Philosophy
Trinity Term, 2012

Abstract

Eukaryotic genomes are highly organized and packaged into chromatin, a complex structure formed of proteins and DNA, in which the basic repeating unit is the nucleosome. Chromatin can be arranged in condensed or relaxed structures influencing accessibility of proteins that regulate transcription, replication, recombination and repair. One class of transiently chromatin-associated proteins is the High-Mobility Group (HMG) protein family. HMG proteins are subdivided into three subgroups: HMGA, HMGB and HMGN. HMGN1, the subject of this study, is a prominent member of the HMGN (High-Mobility Group Nucleosome-binding) protein family, the only HMG proteins that specifically binds to the nucleosomes. HMGNs are maintained in dynamic balance between nucleosome-associated and nucleosome-free pools. Regulation of chromatin involves several enzymatic activities that modify specific residues on chromatin proteins, which may influence these interactions. While associated with nucleosomes, HMGNs can interfere with some modifications of histone tails. Modification of HMGN1 on specific residues and post-translational modification (PTM) of histones are concomitantly regulated by the complex signalling networks associated with the induction of immediate early (IE) genes. Induction of IE genes is associated with phosphorylation of HMGN1 which has been suggested to increase the rate of dissociation of HMGN1 from the nucleosome, thus allowing access and modification of histone tails. My research has been focused on characterizing HMGN1 isoforms present in different cellular compartments and at different time-points during IE gene induction with various stimuli, including epidermal growth factor (EGF), anisomycin (An) and 12-O-tetradecanoylphorbol-13-acetate (TPA). Furthermore, I investigated the localization of HMGN1 within the nucleus and at specific IE gene loci, especially at sites where post-translationally modified histones are localised. In my analysis only the phosphorylation at serine 6 of HMGN1 shows a correlation with gene induction. Analysis of DNA sequences from chromatin immunoprecipitation (ChIP) has shown that HMGN1 is present at equal levels in active and inactive genes. It appears that HMGN1 localization on DNA is not dictated by a particular preference for any gene elements such as promoters, exons, introns or gene termination sequences.

Acknowledgments

The work presented here is the result of the unwavering support of many people throughout the course of my DPhil.

My gratitude goes first to Prof. Louis Mahadevan for welcoming me into the Nuclear Signalling lab and supporting me with his constant encouragement and positive attitude. Secondly, Dr. Alison Clayton's help during my first year in the lab has been fundamental for my thesis. She introduced me into the chromatin signalling field and kindly provided me with guidance every day. I am also incredibly grateful to Dr. Catherine Hazzalin, Dr. Nicholas Crump and Dr. Benjamin Lee for sharing their extensive experimental expertise and their time to discuss scientific ideas. Thanks to all the other members of the lab in these years; Duen-Wei Hsu, Caroline Schneider, Jonathan Neve and Kira Fischer, for their contribution to the pleasant atmosphere we shared at work.

I am thankful to Dr. Benedikt Kessler and Dr. Mukram MacKeen who performed the mass spectrometry analysis and guided me through the troubleshooting and the interpretation of my data. I am grateful to Dr. Penelope Mason and Dr. Anne-Marie Couto for their support in collecting the data from indirect immunofluorescence.

This work was generously supported by the Wellcome Trust programme for Chromatin and Developmental Biology. The programme also provided the invaluable help and the advices of many colleagues throughout the participating departments. I especially would like to thank Prof. Matthew Whitby for leaving his office door always open for me.

I am grateful to Trinity College for granting me awesome dinners, beautiful lawns and great company for four long years. Thanks also to the Oxford Sport Federation for their help and hospitality to the Oxford University Kung Fu club; the time spent teaching and performing martial arts kept my life healthy and balanced during my DPhil.

Several other people deserve my thanks for sharing the ups and downs during my time in Oxford. My hope is that future will grant me a chance to give back to them what I received.

List of publications

E. Pogna, A.L. Clayton, L.C. Mahadevan, Signalling to chromatin through post-translational modifications of HMGN. Biochim Biophys Acta. 2010 Jan-Feb;1799(1-2):93-100.

List of abbreviations

ACF:	*ATP-utilizing chromatin assembly and remodelling factor*
ACH:	*active chromatin hub*
An:	*anisomycin*
AP-1:	*activating protein 1*
APS:	*ammonium persulphate*
ASK:	*apoptosis signal-regulating kinase*
ATF:	*activating transcription factor*
ATP:	*adenosine-5'-triphosphate*
bFGF:	*basic fibroblast growth factor*
BDV:	*born a disease virus*
BMK:	*big MAPK*
BSA:	*bovine serum albumin*
bZIP:	*basic region leucine zipper*
cGK:	*cyclic GMP-dependent protein kinase*
CBP:	*CREB-binding protein*
ChIP:	*chromatin immunopurification*
X-ChIP:	*cross-linked chromatin immunopurification*
N-ChIP:	*native chromatin immunopurification*
cAMP:	*cyclic adenosine mono-phosphate*
CHUD:	*chromatin unfolding domain*
CIP:	*calf intestinal alkaline phosphatase*
CISH	*cytokine-inducible SH2-containing protein*
CRE:	*cAMP response element*
CTCF:	*CCCTC-binding factor*
CTD:	*carboxyl-terminal domain*
CK2:	*casein kinase 2*
DAPI:	*4', 6'-diamidino-2-phenylindole*
DMEM:	*Dulbecco's modified Eagle's medium*
DMP:	*dimethyl pimelimidate*
DNA:	*deoxyribonucleic acid*
DR:	*direct repeat*
DRB:	*5', 6'-dichloro-1-β-D-ribofuranosylbenzimidazole*
DSP1:	*dorsal switch protein 1*
ECL:	*enhanced chemiluminescence*
EGF:	*epidermal growth factor*
ERK:	*extracellular-signal-regulated kinases*
FCS:	*fetal calf serum*
FLIP:	*fluorescence loss in photobleaching*
FP:	*footprint*
FRAP:	*fluorescence recovery after photobleaching*
FRET:	*fluorescence resonance energy transfer*
Gapdh:	*glyceraldehyde-3-phosphate dehydrogenase*
Glyt1	*glycine transporter 1*
Glut2	*glucose transporter 2*
GTP:	*guanosine tri-phosphate*
GR:	*glucocorticoid receptor*
HAT:	*histone acetyltransferase*
Hbb:	*haemoglobin, beta or β -globin*

HDAC:	*histone deacetylase*
HMG:	*high-mobility group*
HMGN:	*high-mobility group nucleosome-binding protein*
HP-1:	*heterochromatin protein-1*
HS:	*hyper sensitive*
JNK:	*c-Jun N-terminal kinase*
ICR:	*imprinting control region*
IE:	*immediate early*
i-IF:	*indirect immunofluorescence*
iNOS:	*inducible form of nitric oxide synthase*
IL-1:	*interleukin-1*
ILV:	*isoleucine, leucine and valine*
IR:	*ionizing radiation*
IRF:	*interferon response factor*
LC-MS:	*liquid chromatography-mass spectrometry*
MAPK:	*mitogen-activated protein kinase*
MAPKK:	*MAPK kinase*
MAP2K:	*MAP second kinase*
MAPKKK:	*MAPKK kinase*
MAP3K:	*MAP third kinase*
MBT:	*mid-blastula transition*
MNase:	*micrococcal nuclease*
MSK:	*mitogen and stress activated protein kinase*
MEK:	*MAPK/ERK kinase*
MEF:	*mouse embryonic fibroblasts*
MKK:	*MAPK kinase*
MHC-II：	*major histocompatibility complex class II*
MLK:	*mixed lineage kinase*
mRNA:	*messenger ribonucleic acid*
MS:	*mass spectrometry*
NBD:	*nucleosome binding domain*
NER:	*nucleotide excision repair*
NFkB:	*nuclear factor-kappaB*
NCP:	*nucleosome core particle*
NLS:	*nuclear localisation signal*
NTD:	*amino (N)-terminal domain*
NURF:	*nucleosome remodelling factor*
OKA:	*okadaic acid*
ORF:	*open reading frame*
(o/n):	*over night*
PBS:	*phosphate buffer saline*
Pc:	*polytene chromosome*
PcG:	*polycomb group*
PCA:	*perchloric acid*
PCAF:	*p300/CBP-associated factor*
PDGF:	*platelet-derived growth factor*
PKA:	*protein kinase A*
PKC:	*protein kinase C*
PHO:	*pleiohomeotic protein*
PMSF:	*phenylmethylsulfonyl fluoride*
PRC:	*polycomb repressive complex*
PRE:	*polycomb response element*

PRLr	*prolactin receptor*
PTM:	*post-translational modification*
PVDF:	*polyvinyl difluoride*
RAF:	*rapidly accelerated fibrosarcoma*
RAPD:	*random amplified polymorphic DNA*
RD:	*regulatory domain*
rhEGF:	*recombinant human epidermal growth factor*
RNA:	*ribonucleic acid*
RSRF:	*related to-serum-response factor*
RSK2:	*ribosomal S6 kinase 2*
rT°	*room temperature*
RTA:	*replication and transcription activator*
SAPK:	*stress activated protein kinase*
SF:	*soluble fraction*
SICER:	*spatial clustering approach for the identification of ChIP enriched regions*
SIE:	*sis-inducible element*
S/MAR:	*scaffold-/matrix- associated region*
SRE:	*serum responsive element*
SREBP:	*sterol-regulatory element-binding protein*
STAT:	*signal transducer and activator of transcription*
TAK:	*transforming growth factor β–activated kinase 1*
TCA:	*trichloroacetic acid*
TBP:	*TATA-binding protein*
TEMED:	*N, N, N', N'-tetramethylethylenediamine*
TCF:	*ternary complex factor*
TF:	*transcription factor*
TFF:	*transcription factors fraction*
TLCK:	*tosyl lysyl chloromethyl ketone*
TNF-α:	*tumour necrosis factor α*
TPA:	*12-O-tetradecanoylphorbol-13-acetate*
TPCK:	*tosyl phenylalanyl chloromethyl ketone*
Tpl:	*tumour progression locus*
TRE:	*TPA response element*
TSS:	*transcriptional start site*
TSH:	*thyroid-stimulating hormone*
UTR:	*un-translated region*
UV:	*ultraviolet*
YY1:	*yin yang protein 1*
ZF:	*zinc finger*
4IPBA:	*4-iodophenylboronic acid*

Histone marks Brno nomenclature (Turner, 2005):

ac:	*acetylated*
me1/2/3:	*mono/di/tri methylated*
ub:	*ubiquitinated*
ph:	*phosphorylated*

Contents

List of figures

List of tables

Chapter 1. Introduction

1.1 The MAPK cascades: from extracellular stimuli to gene activation

Mammalian cells respond to various external stimuli, such as growth factors, cytokines and pharmacological compounds, and physical stresses, such as mechanical and osmotic stress or UV light and ionizing radiation (IR). Most of these diverse stimuli elicit intracellular signalling mechanisms that rapidly initiate transcription of subsets of genes. An extensively studied example of such a response is immediate early (IE) gene activation initiated via the mitogen-activated protein kinase (MAPK) cascades (Clayton and Mahadevan, 2003).

The mechanism of this signalling pathway is that a stimulus is received by a membrane receptor that triggers the activation of a series of successive kinases forming a cascade of sequential phosphorylation (Figure 1). Once activated, the most downstream cytoplasmic kinases translocate into the nucleus and phosphorylate transcription factors, transcriptional co-activators and nucleosomal proteins, resulting in IE gene transcription. IE gene induction is a primary response and does not require new transcription or translation of the signalling effectors (Edwards and Mahadevan, 1992). IE genes are a large family and I focus here on two of the best-studied IE genes, the *c-fos* and *c-jun* genes.

1.1.1 MAPK cascades transduce signals from a variety of stimuli

MAPK signalling pathways regulate proliferation, gene expression, differentiation, mitosis, cell survival and apoptosis (Cano and Mahadevan, 1995; Kyriakis and Avruch, 2001, 2012 and reference therein). MAPKs are present in all eukaryotic cell types and are highly evolutionarily conserved. They coordinate and integrate responses to a large array of stimuli, including hormones, growth factors, osmotic stress, heat shock and cytokines. Examples of MAPK agonists include epidermal growth factor (EGF), platelet-derived growth factor (PDGF), and basic fibroblast growth factor (bFGF) (Bravo, 1990; Kyriakis and Avruch, 2001). Inflammatory stimuli that activate MAPKs include tumour necrosis factor α (TNF-α), interleukin-1 (IL-1), double-stranded RNA, and lipopolysaccharide (Freshney et al., 1994; Sluss et al., 1994; Xia et al., 2000). MAPKs also transduce signals in response to pharmacological compounds, such as anisomycin (An), phorbol esters 12-O-tetradecanoyl-phorbol-13-acetate (TPA) and okadaic acid (OKA) and from physical and physiological stresses, like UV radiation, heavy metals and hyperosmotic and mechanical stresses (Edwards and Mahadevan, 1992; Han et al., 1994; Hazzalin et al., 1998; Hibi et al., 1993; Komuro et al., 1996; Kyriakis et al., 1994; Rosette and Karin, 1996; Rouse et al., 1994, and references therein). The best-studied MAPK cascades identified so far are schematically summarised in Figure 1. They include the extracellular-signal-regulated kinase (ERK), c-jun N-terminal kinase or stress-activated protein kinase (JNK/SAPK), big MAPK (BMK) and p38 MAP kinase cascades (Clayton and Mahadevan, 2003; Edmunds and Mahadevan, 2004).

1.1.1.1 MAPK cascades structure

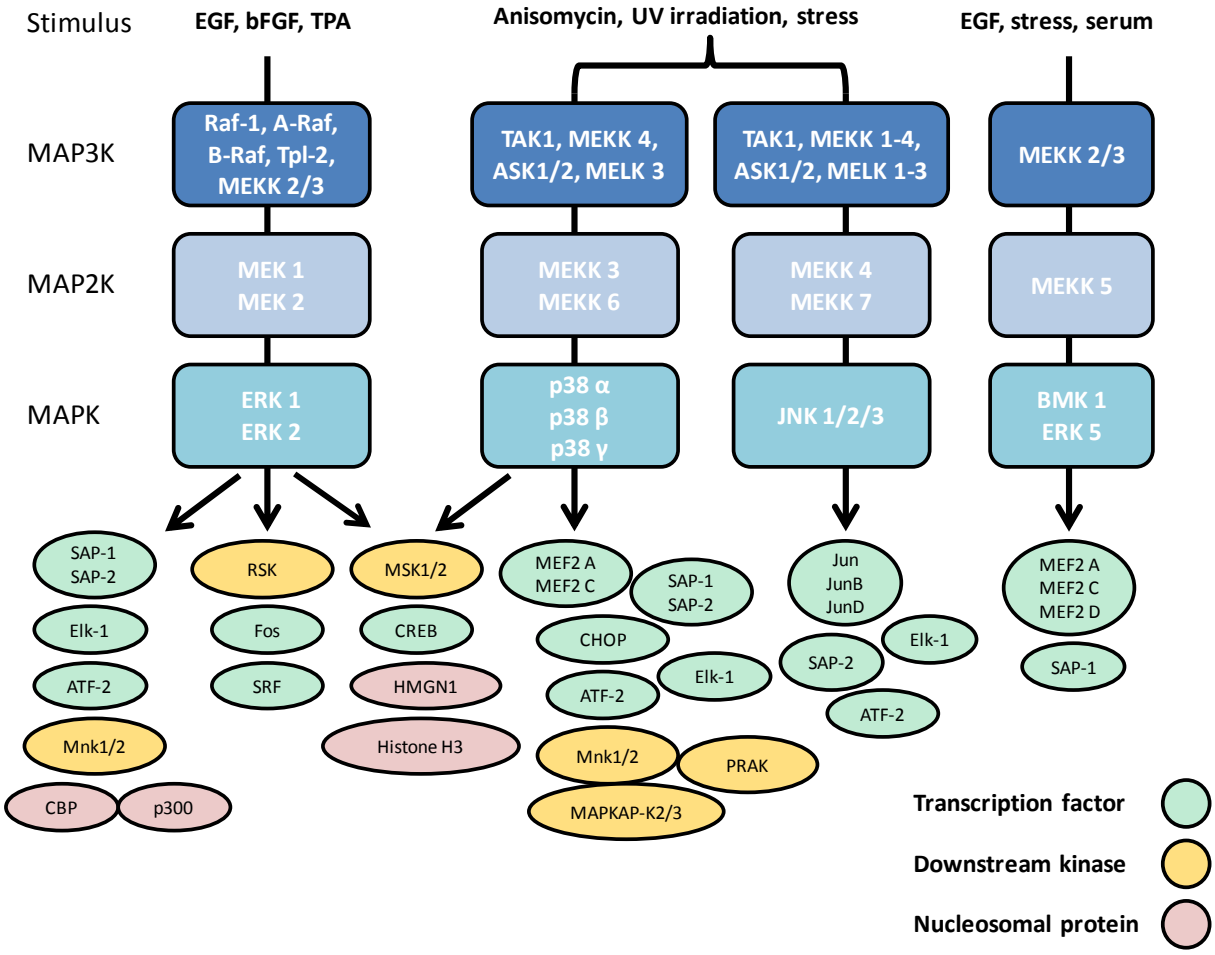
MAPK signalling may be subdivided into cytoplasmic and nuclear parts. The cytoplasmic components respond to the activation of cell membrane receptors and integrate the signals received. They modulate the activation of the final effectors that translocate from the cytoplasm to their nuclear targets. The transduction of the signal in the cytoplasm is based on three-step signal modules (Figure 1). In mammals, the downstream step is represented by the MAP-Kinases, ERKs, JNKs, p38s and BMK; they enter the nucleus to phosphorylate their substrates, including transcription factors, co-activators and nucleosomal proteins that promote gene transcription (Clayton and Mahadevan, 2003; Edmunds and Mahadevan, 2004; Hazzalin and Mahadevan, 2002; Kyriakis and Avruch, 2001).

MAPKs are activated by the phosphorylation of a tyrosine (Tyr) and a threonine (Thr) in their kinase domain activation loop catalyzed by dual specific kinases referred to as MAPKK (MAPK Kinases or MAP2Ks). In mammalian cells, MAPKK, MEK1/2, MKK3/6, MKK4/7 and MEK5 kinases are at the second level of cascade regulation. MAPKK are themselves substrates of serine (Ser)/Thr phosphorylation within a conserved motif in their kinase domain catalyzed by MAPKKKs (MAPK Kinase Kinases or MAP3Ks). There are several MAPKKKs in mammals: the RAF kinase family, RAF-1, B-RAF and A-RAF, involved in cellular growth, cancer and developmental diseases, the apoptosis signal-regulating kinases-1 and -2 (ASK1/2), the pro-inflammatory activators TAK1 and Tpl2, the mixed lineage kinases (MLKs), a small family of protein Ser/Thr kinases with a common structural configuration and MEKKs 1-4, promiscuous kinases that catalyze the activation of multiple different MAPKKs (Hagemann and Rapp, 1999; Kyriakis and Avruch, 2001, 2012; Pandit et al., 2007; Roberts and Der, 2007 and reference there in). Although MAPKKK activation has not been characterized in detail, some evidence suggests involvement of guanosine triphosphate (GTP)-binding proteins, membrane receptors and cytoskeletal proteins (Aikawa et al., 2002).

Several phosphatases modulate MAPK-mediated signal transduction allowing dynamic but tight quantitative control of the signal (Farooq and Zhou, 2004; Kyriakis and Avruch, 2012). MAPK cascades are commonly named after the category of stimulus inducing a specific group of kinases or the most downstream nuclear effectors of the signalling pathway (Figure 1). It is often the case that similar stimuli induce the activation of the same MAPK, and the correlation between stimuli and effectors is conserved in simpler eukaryotes like as yeast where MAPK cascades are less complex and cannot be activated in parallel by a singular stimulus (Edmunds and Mahadevan, 2004). Since the target of my analysis is signal transduction at nucleosomal level, I will focus on the events that link MAPKs with activation of IE genes

Figure 1. **MAPK cascades.**

Schematic representation of mammalian MAPK cascades and target substrates, including transcription factors that are involved in the regulation of *fos* and *jun* IE genes.



adapted from (Edmunds and Mahadevan, 2004; Hazzalin and Mahadevan, 2002; Roberts and Der, 2007)

1.1.1.2 MSK1/2 are the kinases that mediate the nucleosomal response

The final outcome of nuclear signalling is the targeted activation of a subset of genes in response to specific external stimuli (Figure 2). To achieve this, several DNA-binding proteins and transcription factors are involved in a complex network of activating or inhibitory interactions (Hazzalin and Mahadevan, 2002). As shown in Figure 1, targets of MAPK cascades include transcription factors, nuclear kinases and chromatin proteins. Among the targets of these cascades, histone H3 and HMGN1 (previously known as HMG-14) are notable for being targeted for phosphorylation by both p38 and ERK pathways; this phenomenon is called the nucleosomal response (Thomson et al., 1999). The concomitant phosphorylation of histone H3 Ser10 and Ser28 and HMGN1Ser6 is observed in the nucleus in response to several stimuli (Mahadevan et al., 1991; Thomson et al., 1999). The ERK and p38 MAPK kinase cascades respond to growth factors (EGF, bFGF), UV stress and pharmacological compounds (An, TPA), activating several nuclear kinases, MSKs, RSKs and Mnk1/2 (Arthur and Cohen, 2000; Frödin and Gammeltoft, 1999; Raingeaud et al., 1995; Waskiewicz et al., 1997). Studies in human fibroblasts from Coffin-Lowry patients (who carry mutations that inactivate the ribosomal kinase 2 [RSK2] gene) and mouse embryonic stem cells (in which the RSK2 gene had been knocked out by homologous recombination) observed loss of EGF-, TPA- and UV-inducible histone H3 phosphorylation (Trivier et al., 1996 ; Sassone-Corsi et al., 1999). However RSK2 is a kinase that lies downstream of only ERK1/2 and not p38. On the other hand, MSK1/2 were identified as nuclear targets of both ERK1/2 and p38 and also proved strong *in vitro* kinases for histone H3 and HMGN1 (Deak et al., 1998). The controversy on the role of RSK and MSK was resolved by later work on RSK2 and MSK mouse knockouts that produced conclusive *in vivo* evidence that the MSKs, particularly MSK2, but not RSK2, are the major histone H3 and HMGN1 kinases (Bruning et al., 2000; Soloaga et al., 2003; Wiggin et al., 2002). It has been proposed that histone

modifications may influence gene transcription by regulating the remodelling of chromatin architecture (Berger, 2007; Bernstein et al., 2007; Kouzarides, 2007; Li et al., 2007; Mizzen and Allis, 1998; Shilatifard, 2006). Inhibition of MSKs and studies in MSKs knockout cells suggested that the nucleosomal response may not be essential for IE gene induction, but may be involved in the efficiency of the response as variation of the rate/profiles of expression of these genes has been observed (Lim et al., 2004; Soloaga et al., 2003).

1.1.2 Signalling through MAPK/MSKs activates IE gene transcription

The best-studied system of IE gene induction is that of the *fos* and *jun* families (*c-fos*, *fos* B, *c-jun*, *jun* B and *jun* D) by MAPK/MSKs (Figures 1, 2). In general, the strategies adopted to ensure transient expression are rapid transcriptional shut-down following induction and destabilization of the gene products (Edwards and Mahadevan, 1992). Cultured cells represent one of the few experimental systems in which quantitative influences on gene transcription, messenger RNA (mRNA) accumulation and degradation can be monitored. Different stimuli produce clearly distinguishable profiles of gene induction (Figure 2). These profiles are highly reproducible and are specific for the stimulus. For example both bFGF and EGF activate the ERK pathway but they induce different profiles of activation of five IE genes (Figure 2). With some growth factors, full activation of ERKs is seen within 1 minute of stimulation in every single cell in a synchronized population (Cano et al., 1995). If the transcriptional response is almost instantaneous and the dynamics of mRNA transcription and degradation are fairly constant then the differing expression profiles observed are most likely generated by variable rates of transcriptional initiation. Transcription factors regulate transcriptional initiation and it is the strength of intracellular signal transduction that determines their activation, with

chromatin structures modulating the process. Auto-repression and degradation are obvious ways to ensure the discrete availability of a particular transcription factor, which may explain tissue- and cell type-specific differences in IE gene expression.

1.1.2.1 IE genes; c-fos and c-jun

The IE genes *c-fos* and *c-jun* were first identified as proto-oncogenes and have become the canonical example of transcription factors activated by external stimuli. A major advantage of the use of these genes in a mouse model system is that they are poised for transcription and can be induced within minutes. Transcription of *c-fos* and *c-jun* is switched off at about 30 min after stimulation, and their mRNAs are rapidly degraded, resulting in the transient presence of transcripts in the cell (Edwards and Mahadevan, 1992; Hazzalin and Mahadevan, 2002). The rapid turnover of *c-fos* and *c-jun* mRNA depends on the presence of a conserved AU-rich sequence in the 3' untranslated sequence, found in other unstable transcripts (Angel and Karin, 1991; Shaw and Kamen, 1986; Wilson and Treisman, 1988). *c-fos* and *c-jun* genes encode basic region-leucine zipper (bZIP) DNA-binding transcription factors. bZIP proteins are composed of two functional domains, a basic DNA-binding domain, and a leucine-zipper domain used for dimerisation (Angel et al., 1988; Bohmann and Tjian, 1989; Smeal et al., 1989). Dimerisation occurs through the bZIP domain bringing together the basic regions of the dimer, such that a contiguous DNA-contact interface is produced. This interface interacts with specific DNA sequences (Schumacher et al., 2000). It has been established by site-directed mutagenesis which of the leucines and other residues are essential for the interaction of these proteins (Kouzarides and Ziff, 1988; Ryseck et al., 1990). Members of the Fos and Jun protein families dimerise in homo- and/or heterodimeric combinations forming the activating protein 1 (AP-1) early response transcription factor (Angel and Karin, 1991). AP-1 upregulates transcription of genes containing the TPA response element (TRE; 5'-TGAG /

CTCA-3') or cyclic adenosine mono-phosphate (cAMP) response elements (CRE; 5′-TGACGTCA-3′).

c-jun is an intronless gene and in humans is mapped to 1p32-p31, a chromosomal region involved in both translocations and deletions in human cancer. JNK kinases activate c-Jun through phosphorylation of two sites, Ser63 and Ser73, located within the trans-activation domain, increasing the ability of c-Jun to activate transcription (Binétruy et al., 1991; Smeal et al., 1991). Phosphorylation of sites clustered next to the basic region of c-Jun inhibits DNA binding (Boyle et al., 1991; de Groot et al., 1993; Lin et al., 1992). Three genes related to the oncogene *c-jun* have been identified in mouse: c-*jun* , *jun B* and *jun D* (Almendral et al., 1988; Cochran et al., 1983; Lamph et al., 1988; Ryder et al., 1988; Ryseck et al., 1988). Only two of them, *c-jun* and *jun B*, are significantly induced by mitogens in quiescent fibroblasts. The expression of *jun D* is high in non-proliferating cells and slightly increases after serum stimulation. The *c-jun* mouse gene can be transcribed in two forms, one of 2.7 kb and another of 3.2 kb (Figure 9a) (Ryseck et al., 1988). The mRNAs of *jun B* and *jun D* are shorter, being 2.1 and 2.0 kb, respectively. The size of the three JUN proteins is very similar: c-JUN has 334 amino acids, JUN B 344 amino acids, and JUN D 341 amino acids. These proteins show highest similarity (≈ 75%) in the last 100 amino acids, that includes the DNA-binding domain and five leucines potentially involved in a zipper (Hirai et al., 1989; Landschulz et al., 1988). It is clearly established that each of the three JUN proteins can form a complex with each of the FOS proteins (Bravo, 1990). In all cases, the complex formed significantly increases the binding affinity to an AP-1 binding site. The three JUN proteins and FOS proteins can be simultaneously present in the nucleus of stimulated cells. Thus, by modulating their expression, a stimulus can generate the formation of different complexes of AP-1 factors with different affinities for DNA.

c-fos is a canonical gene; in mouse it encodes a mRNA size of 2.2 kb and a protein of 380 amino acid and 40.8 kDa. The c-FOS protein is unstable and undergoes rapid degradation but upon phosphorylation at Ser374, catalyzed by ERK1/2, and at Ser362, catalyzed by RSK, the life of c-FOS is prolonged by several hours (Chen et al., 1993; Murphy et al., 2002; Okazaki and Sagata, 1995). This double phosphorylation occurs within the nucleus. The Fos family include also three more genes *fos B, fra-1* and *fra-2*. The *fra-1* mRNA (1.8 kb) and the *fos B* mRNA (5.1 kb) encode proteins of 275 amino acids and 338 amino acids respectively. Comparison of the c-FOS protein sequence with that of FOS B and FRA-1 shows a 75% identity in a region of 78 amino acids which includes the DNA-binding domain and the five leucines potentially involved in the zipper (Bravo, 1990). c-FOS and FOS B contain transcriptional activation domains while FRA1 and FRA2 do not (Shaulian and Karin, 2002). The FOS proteins cannot homodimerize and instead form stable heterodimers in cooperation with JUN proteins that enhance their DNA binding activity. Fos family genes respond differently to mitogenic stimulation and have different 5' upstream regulatory elements (Bravo, 1990). cAMP is a strong inducer of *fra-1* but a very weak inducer of *c-fos* and *fos B* expression. TPA slightly induces *fos B* but it is a strong inducer of *fra-1* and *c-fos* expression. TPA-elicited transcription appears 15 min after induction and persists until 45 min. EGF and bFGF also induce *c-fos*. bFGF elicits *c-fos* transcription at early time points and transcripts remain present also at 60 min after stimulation. EGF-dependent transcription is activated earlier than that of bFGF but it is shut off between 30-45 min after induction. Anisomycin both induces mRNA production and inhibits translation. Thus *c-fos* transcripts accumulate at later time points and signal persists 60 min after stimulation (Figure 2, Hazzalin and Mahadevan, 2002).

Figure 2. **Expression of IE genes in response to TNF-α, bFGF, EGF, TPA and anisomycin.**
Northern blot analysis of IE genes RNA isolated from quiescent C3H 10T1/2 cells stimulated as indicated: TNF-α (5 ng/ml); bFGF (20 ng/ml); EGF (50 ng/ml); TPA (100 nM); or An (10 μg/ml), for 15–60 minutes. C, control (unstimulated). The housekeeping gene *gapdh* was used as a control for RNA loading.

from (Hazzalin and Mahadevan, 2002)

1.1.2.2 Regulation of c-fos and c-jun

AP-1, a transcription complex formed by homo- or heterodimerisation of FOS, JUN and other transcription factors, is an early response transcription factor responsible for the regulation of multiple genes with implications in a large variety of biological processes, proliferation, apoptosis, differentiation, epidermal and neuronal development, immune and inflammatory responses, and tumorigenesis (Meng and Xia, 2011; Shaulian and Karin, 2002). Regulation of IE genes *fos* and *jun* is critical for determination of the balance in the composition of the homo- or heterodimer AP-1, allowing qualitatively different types of AP-1 complexes to emerge. IE gene activation is independent of *de novo* transcription or translation of the signalling effectors and so transcription completely relies on the existing levels of the nuclear targets (Mahadevan et al., 1991) . Complete understanding of the regulation of *c-fos* and *c-jun* has been a major challenge, because these genes possess various control elements and they engage in self-regulatory positive and/or negative feedback loop. Work from various laboratories describe multiple upstream elements responsible for *c-fos* and *c-jun* transcriptional activation (Figure 3, reviewed in Hazzalin and Mahadevan, 2002).

Identified control elements for *c-fos* include: a SIE, a SRE, a TCF site, a AP-1 site and AP-1/CRE, a DR and a CRE elements (Fisch et al., 1989; Hayes et al., 1996; Shaw et al., 1989; Sheng et al., 1988; Treisman, 1992; Wagner et al., 1990). The *c-jun* promoter is controlled by the jun2 AP-1 site, a FP, a NF-jun site, two overlapping SP-1 sites, a CCAAT box, the jun1 AP-1 site, a RSRF site (Clarke et al., 1998; Han and Prywes, 1995; Han et al., 1997) and in the 5' UTR, two AP-2 sites and a weak AP-1 site (Angel et al., 1988; Herr et al., 1994; Rozek and Pfeifer, 1995). The Activating Transcription Factor-2 (ATF) is also implicated in regulating the *c-jun* gene (van Dam et al., 1995).

Signalling to these elements operates in two different ways. In one case, the DNA elements are constitutively occupied by sequence-specific transcription factors or complexes of such

factors (reviewed in Karin et al., 1997). Activation via these sites requires the arrival of a kinase and subsequent phosphorylation of these factors (reviewed in Karin and Hunter, 1995; Treisman, 1996). In the other case, an upstream DNA element is unoccupied in quiescent cells, and upon stimulation becomes occupied by transcription factors. For example, for STAT- and NFkB- binding sites, the inactive transcription factor is located in the cytoplasm and is activated via specific kinase pathways upon stimulation and then translocates into the nucleus to effect transcription via their respective upstream regulatory elements (reviewed in Baeuerle and Baltimore, 1996; Darnell, 1997). Similar complexity has been observed not only for *c-fos* and *c-jun* but also for other IE genes, *fos B, jun B* and *jun D* (Figure 3, Hazzalin and Mahadevan, 2002). All five genes are controlled by several elements, so that from a single stimulus each gene could potentially require the simultaneous and parallel activation of multiple kinases, the phosphorylation of more than one transcription factor and the delivery of signal(s) to the gene through several sequence elements. Although these analyses have been carried out in different systems, they do not exclude the possibility that the different factors simultaneously co-regulate the same gene in a particular cell type (Hazzalin and Mahadevan, 2002).

In addition to signalling mediated by transcription factors, there has been constant interest in the role that chromatin structure plays in IE gene induction. The early studies reported transient relaxation in chromatin structure conformation that correlates well with IE gene induction and shut down (Feng and Villeponteau, 1992). Previous approaches aimed at isolating active genes, which include the active *c-fos* gene, already noticed that nucleosomes at these loci show an altered conformation (Allegra et al., 1987; Chen and Allfrey, 1987). By targeting phosphorylation to transcription factors controlling *c-fos* and *c-jun* , MAPK signalling is implicated in the recruitment of coactivators such as p300/CBP and pCAF, which are histone acetyltransferases (HATs), to these promoters (reviewed in Bannister et al., 1995;

Janknecht and Nordheim, 1996; Karin, 1996; Kouzarides, 1999; Kuo and Allis, 1998; Treisman, 1996 and references therein). As previously mentioned, MAPKs also target chromatin at *c-fos* and *c-jun* loci through MSK1/2-dependent histone H3 and HMGN1 phosphorylation (Clayton and Mahadevan, 2003). Moreover, high-resolution comparative maps of the distribution of chromatin modifications across *c-fos* and *c-jun* by chromatin immunoprecipitation (ChIP) assays established a direct correlation between gene induction and localised histone modification (Figure 9) (Edmunds et al., 2008). Modifications at these sites include H3K4me3, H3K36me3, H3K79me2 and H3K9ac, and phosphoacetylated H3 (Barratt et al., 1994; Cheung et al., 2000; Clayton et al., 2000; Edmunds et al., 2008).

Figure 3. **Binding sites upstream of *jun* and *fos* genes.**

Schematic of the transcription factor binding-sites involved in IE gene expression. The relative positions of identified and putative regulatory elements involved in the expression of the IE genes *c-fos*, *fos B*, *c-jun*, *jun B* and *jun D* genes are presented diagrammatically. Note that the scale at which these regions are presented is different for each gene. The regulatory elements are in their relative positions with respect to the transcriptional start site at +1.

from (Hazzalin and Mahadevan, 2002)

1.2 Chromatin dynamics and response to nuclear signalling

The DNA of eukaryotic genomes is organised into an ordered repeat of units called nucleosomes. The nucleosome is formed by an octamer core of histone proteins [H2A, H2B, H3 and H4] encircled by ~147 base pairs of DNA. The nucleosome is responsible for one level of the packaging of genomes and it has been proposed that DNA transcription requires site-specific regulation of the level of chromatin compaction (Campos and Reinberg, 2009). Extensive work has identified multiple players involved in both direct interaction with DNA and the regulation of the compaction-relaxation balance. Histones exhibit considerable structural and compositional diversity, that is achieved through histone variant incorporation, alternative nucleosomal structures and a vast array of post-translational histone modifications (Klose and Bird, 2006; Kouzarides, 2007; Sarma and Reinberg, 2005; Zlatanova et al., 2009). A large number of proteins is involved with direct nucleosome interaction that influence histone modification (HATs, HDACs, MSKs etc.), positioning (ATP-dependent chromatin remodelling complexes such as Swi/Snf), compaction (histone H1) and the architecture of chromatin domains and regions (CTCF and HMGs) (Berger, 2007; Campos and Reinberg, 2009; Happel and Doenecke, 2009; Phillips and Corces, 2009; Postnikov and Bustin, 2010). As explained previously, signalling pathways, particularly the MAPK cascades, directly act on chromatin proteins such as histones and HMGNs, and association of specific modifications with particular patterns of gene expression suggests MAPK involvement in regulation of the overall efficiency of transcription (Clayton and Mahadevan, 2003).

1.2.1 The nucleosome and core histones

Histone proteins have been studied for more than a century since they were identified by Kossel in 1884. Histones are universal components of eukaryote chromosomes and because they have abundance similar to that of the DNA, they were long believed to be the genetic material itself. Early studies took advantage of acid purification methods to identify five histone types, now designated H1, H2A, H2B, H3, and H4 (Phillips and Johns, 1965). Histones are small protein molecules of about 10-20 kDa that proved to be among the most invariant proteins known with near perfect conservation across species (De Lange et al., 1969). Unfortunately, for many years the denaturing nature of acid extraction prevented any investigation of their molecular composition, structure and biochemistry within nucleosomes, and only the advent of powerful proteinase inhibitors allowed further investigation leading to the discovery of the $(H3)_2(H4)_2$ tetramer, as well as the H2A-H2B dimer (Kelley, 1973; Kornberg and Thomas, 1974). Structural studies now show that two H3-H4 histone dimers are bridged together as a stable tetramer that is flanked by two separate H2A-H2B dimers forming the core histone octamer (Luger et al., 1997). Each octamer is encircled by ~147 DNA base pairs and, depending on compaction level, flanked by a linker histone, H1 that completes the chromatosome by protecting internucleosomal linker DNA near the nucleosome entry-exit point. Once the nucleosome structure is assembled, a hydrophobic core, the globular domain (also referred as the histone fold domain) resides within the nucleosome core particle (NCP) while the amino-terminal domain (NTD) and the short proteinase-accessible carboxy-terminal domain (CTD) of all eight histones protrude out from the surface of the particle (Davey et al., 2002; Hacques et al., 1990; Luger et al., 1997). The NTDs and CTDs are rich in basic residues and subjected to multiple posttranslational modifications (PTMs) including methylation and acetylation, phosphorylation,

ubiquitinylation, sumoylation, biotinylation, ADP-ribosylation, prolyl isomerisation and tail clipping (Kouzarides, 2007).

While histone H4 and H2B are largely invariant, diversity is more evident within histone H3 and H2A. Histone variants can affect nucleosome stability and some have been implicated in gene transcription (H3.1- and H3.3-H2A.Z) while others are specifically positioned throughout chromosome regions like the centromeric specific histone H3 (CenH3 or CENP-A) or specifically targeted by nuclear processes such as H2A.X in DNA repair (Campos and Reinberg, 2009 and reference therein).

Nucleosomes can influence chromatin in three possible ways: (i) they are directly responsible for biophysical properties of chromatin and nucleosome density can affect chromatin compaction; (ii) the deposition of combinatorial layers of PTMs can act as a histone code, recognized by specific effector proteins; (iii) histones can be part of signal transduction pathways and participate in the redundancy and reinforcement of signalling feedback loops (Schreiber and Bernstein, 2002; Strahl and Allis, 2000; Taverna et al., 2007).

1.2.1.1 Chromatin compaction and relaxation

The discovery of chromatin by Flemming (1879, 1882) and Boveri (1904) revealed a basic dual conformational state. This macromolecular complex of DNA and proteins was detected either as condensed and brightly stained chromatin named heterochromatin or a decondensed scarcely stained complex referred as euchromatin. It was suggested that euchromatin was the active part of the genome while heterochromatic chromosomes and heterochromatin in general were transcriptionally inactive. Although, these assumptions have not always proved correct, it is generally accepted that chromatin compaction is a mark of transcriptional silencing and, *vice versa*, actively transcribed regions required relaxation of chromatin for optimal functioning (Trojer and Reinberg, 2007). This assumption is largely supported by the

evidence that DNA requires direct physical interaction with a large variety of proteins to be properly transcribed, copied and maintained (Groth et al., 2007; Li et al., 2007). The regulation of chromatin compaction involve three main players: nucleosomes, including histone variants and histone tail modifications, chromatin-associated proteins, including architectural and remodelling proteins, and finally DNA methylation (Trojer and Reinberg, 2007). Each of these levels requires specific regulation and is involved in cross interactions that are poorly understood. I will describe some generally accepted examples which may represent a model for similar events. The first element associated with compaction of chromatin is the presence of the histone linker H1. Chromatin devoid of H1 present a "beads-on-a-string" structure of about 11 nm that is usually considered the basic structure of euchromatin, a state in which DNA can be accessed by transcription factors and nucleosomes can be modified, replaced or even removed. Euchromatin can be compacted into a 30 nm heterochromatic fiber by the repositioning of histone H1, the deacetylation of histones H3 and H4, the methylation of H3K9, H4K20, the methylation of DNA, the incorporation of histone variant (macro H2A and/or H3.2) and the activity of chromatin components and *trans*-acting factors (Campos and Reinberg, 2009; Trojer and Reinberg, 2007, and reference therein). Depending on the degree of modification and the presence of specific factors, it is possible to distinguish several type of heterochromatin. Constitutive heterochromatin is defined as a compacted status of chromatin that once established is mostly regarded as permanently inactive. This chromatin is usually found at specific region of chromosomes, such as centromeres and telomeres and these loci are associated with gene-poor, highly repetitive DNA sequences (Grigoryev et al., 2006; Richards and Elgin, 2002). The inactive status of constitutive heterochromatin has been correlated with specific status of histone methylation (H3K9me3, H4K20me3) and hypoacetylation. Heterochromatin protein-1 (HP-1) is enriched at these loci, targeted by its chromodomain that binds H3K9me3. HP-1 is also a major

scaffold for recruitment of methylases and it is directly involved in maintaining histone and DNA methylation (Campos and Reinberg, 2009).

Alternatively, facultative heterochromatin might be established as in the case of X chromosome inactivation in female mammals, autosomal imprinted genomic loci, long-range silencing (e.g., HOX gene clusters) or local gene silencing. X chromosome inactivation requires the long non-coding RNA Xist, guiding the silencing via methylation and ubiquitination of histones tails, (H4K20me1, H3K9me2, H3K27me3, H2AK119ub1), and incorporation of macroH2A (Campos and Reinberg, 2009; Trojer and Reinberg, 2007). Chromatin factors such as PRC complexes and polycomb group (PcG) proteins are major players in directing histone ubiquitin ligases and histones and DNA methylases to their specific targets (Campos and Reinberg, 2009). Autosomal imprinted genomic loci silencing works in a fashion similar to X chromosome inactivation, but the RNA components are not so well characterized and histone marks are slightly different; ubiquitination is not present, H4K20 is trimethylated and so is H3K9, CTCF and PRC2 proteins are involved in the delimitation and maintenance of the silenced loci (Trojer and Reinberg, 2007). Notably, histone methylation plays a dual role in the silencing of autosomal loci, with H3K4me3 coinciding with the active regions while H3K9me2/3, H3K27me3 and H4K20me3 are present in the silenced regions. Long-range silencing involves histone marks that recapitulate X-chromosome silencing without the incorporation of macroH2A. However this kind of silencing seems to rely mainly on looping of chromatin and trans regulation by PcG proteins, via contact between regions on different chromosomes (Campos and Reinberg, 2009; Trojer and Reinberg, 2007 and reference therein). On the other hand, local gene silencing relies more on the specific incorporation of histone linker H1 that triggers a local compaction of the chromatin that elicits H3K9me2, H4K20me1, H2AK119ub1 and recruitment of other PcG

proteins and HP1. No RNA involvement has been reported in the process, and neither is DNA methylation observed.

It is clear that the processes involved in various states of heterochromatization are redundant, differences between these events being the relative abundance of specific histone modifications, the presence of co-factors and the driving force that establishes the silencing.

Apart from the absence of heterochromatin marks, euchromatin can differ in degrees of accessibility and nucleosome density, with histone variants and PTMs playing a major role (Groth et al., 2007; Li et al., 2007). Most common euchromatin is enriched in hyperacetylated histones, H3K4me2/3 and H3K36me3. Incorporation of histone variant H3.3, H2A.Z and presence of ATP-dependent chromatin remodelers is also frequent (Berger, 2007; Campos and Reinberg, 2009; Kouzarides, 2007; Li et al., 2007 and reference therein). Boundaries between regions of actively transcribed chromatin are often enforced by factors such as CTCF that bind DNA and physically organize its architecture (Phillips and Corces, 2009).

1.2.1.2 Histone tails and PTMs, distribution on the genome

Histones are modified at over 60 different residues, mostly arginine (Arg, R), lysines (Lys, K), serines (Ser, S) and threonines (Thr, T). Further complexity arises from the fact that methylation at lysines or arginines is observed in three different forms: mono-, di-, or trimethyl- for lysines and mono- or di- (asymmetric or symmetric) for arginines (Kouzarides, 2007; Lee and Mahadevan, 2009; Li et al., 2007).

Specific patterns of modification have been observed within transcribed chromatin (Figure 4). Acetylation of histone H2, H3 and H4 is enriched at specific sites in the promoter and 5’ end of the coding regions; H3 acetylation is observed in enhancer regions. Acetylation is consistently linked to gene expression, but no single site-specific modification correlates exclusively with increasing gene activation (Wang et al., 2008). H2BK12ac, H2BK20ac,

H2BK120ac, H3K4ac, H4K5ac, H4K8ac, H4K12ac and H4K16ac are elevated in the promoter and transcribed regions of active genes. H2AK9ac, H2BK5ac, H3K9ac, H3K18ac, H3K27ac, H3K36ac and H4K91ac are mainly located in the region surrounding the transcriptional start sites (TSS) (Wang et al., 2008). H3K18ac and H3K27ac are frequently observed also at enhancers regions (Heintzman et al., 2009). H3K9ac distribution is skewed in a symmetric double peak, a first enrichment is observed within the promoter regions (within 1 kb from the TSS), greater depletion occurs directly over the TSS, due to nucleosome depletion or RNA polymerase II (Pol II) binding, and a second peak is present downstream TSS (Barski et al., 2007; Bernstein et al., 2005; Edmunds et al., 2008; Guenther et al., 2007; Heintzman et al., 2009; Koch et al., 2007).

Targeted lysine methylation is present at enhancers, promoters and protein coding regions. Specific H3 methylation events have particular distribution patterns. Enhancers are enriched in H3K4me1/2/3 (Heintzman et al., 2009; Lee and Mahadevan, 2009). Presence of H3K4me1 has been suggested to correlate with H3K4me3 depletion, but it is currently contentious due to conflicting ChIP-chip and ChIP-Seq results (Barski et al., 2007; Heintzman et al., 2007, 2009). H3K9me1/2/3 and H4K20me3 are also occasionally observed at enhancer regions (Lee and Mahadevan, 2009). H3K9me3 is widely reported as a repressive and heterochromatic modification but recent detection in active chromatin domains suggests a role in transcription elongation and HP1 protein targeting into coding regions of active genes (Barski et al., 2007; Kim et al., 2007a; Vakoc et al., 2005, 2006). H3K4me1/2/3 are present at promoter and 5'end coding regions of actively transcribed genes. H3K4me3 is present as an asymmetric double peak; similar to H3K9ac, a major peak of enrichment is observed downstream of the TSSs and a minor peak is found within the promoter regions, 1 kb from TSSs, with depletion directly over the TSS (Barski et al., 2007; Bernstein et al., 2005, 2007; Edmunds et al., 2008; Guenther et al., 2007; Heintzman et al., 2007; Koch et al., 2007). H3K4me1 and H3K4me2

are also associated with TSSs and often co-associated with H3K4me3, they are more broadly distributed downstream of TSSs (Barski et al., 2007; Koch et al., 2007). Highly expressed genes present multiple methylations in the open reading frame (ORF), H4K20me1, H3K9me2/3, H3K27me2/3, H3K79me1/2/3 and H2BK5me1 (Wang et al., 2008). H2BK5me1, H3K4me1/2/3, H3K9me1, H3K27me1, H4K20me1 and H3K79me1/2/3 predominantly occur over the 5' coding region while H3K36me3 peaks at the 3' end while H3K9me2/3 and H3K27me2/3 mark the transcriptional termination site (Barski et al., 2007; Lee and Mahadevan, 2009; Steger et al., 2008; Vakoc et al., 2006; Wang et al., 2008). H3K36me3 and H4K20me1 in coding regions of active genes have been shown to be dependent on transcription elongation (Edmunds et al., 2008; Vakoc et al., 2006). Whether modifications like H2BK5me1, H3K9me1 and H3K27me1 exist prior to transcription or are deposited co- or post-transcriptionally is currently unclear. Recent ChIP-chip data suggests that H3K36me3 preferentially marks exons, yet more varied patterns of H3K36me2 and H3K36me3 have been observed across specific loci (Kolasinska-Zwierz et al., 2009; Lian et al., 2008). H3K4me3-H3K36me3 enrichment at specific domains has been used to identify and improve TSS and termination site annotation. The strong correlations of H3K4me3-H3K36me3 enrichment and active transcription have led to the identification of new transcription units for non-coding RNA (Barski et al., 2007; Guenther et al., 2007; Guttman et al., 2009; Mikkelsen et al., 2007).

Methylation and acetylation marks are often observed at the same location in specific combinations as follows: enhancers are frequently associated with H3K4me1/2/3, K9me1, H3K18ac and H3K27ac (Heintzman et al., 2009; Wang et al., 2008). Histone modifications at promoters are highly similar across cell types while enhancer modification profiles are highly cell-type specific despite similar global gene expression patterns (Heintzman et al., 2009). H3K4me3-enriched promoters are acetylated preferentially at H3K9ac, H3K18ac and

H3K27ac. H3K4me3 and H3K9ac occur at poised IE genes independently of transcription and almost 60% of silent promoters contain H3K4me3 (Barski et al., 2007; Edmunds et al., 2008; Guenther et al., 2007; Hazzalin and Mahadevan, 2005). Promoters with H3K4me3 in the absence of H3K27me3 positively correlate with gene expression and are over represented in housekeeping genes. On the other hand, the H3K4me3 and H3K27me3 combination correlates with gene repression (Mikkelsen et al., 2007). Curiously, several residues (H2BK5, H3K9, H3K27 and H3K36) are present in an acetylated version in gene promoters and a methylated version in transcribed regions (Wang et al., 2008).

Ser/Thr (ST) phosphorylation is also involved in transcription and H3 phosphorylation (ph) is the best characterized. H3S10ph and H3S28ph localize within actively transcribed regions and mitotic chromosomes, and are thus associated with both relaxed and condensed chromatin respectively. Mitotic H3 phosphorylation involves the majority of H3 in the nucleus and is mediated by Aurora B kinase (Goto et al., 2002). It appears in pericentric heterochromatin during late G2 phase, is widespread by prophase and remain detectable until telophase (Prigent and Dimitrov, 2003). Induced H3 phosphorylation, by contrast, involves a small fraction of nucleosomes and is localized at active genes. H3 phosphorylation occurs in response to ERK and p38 cascade activation on a subfraction of nucleosomes that is also sensitive to hyperacetylation (Thomson et al., 1999, 2001). A background level of phosphorylation is occasionally detected in quiesced cells but IE induction increases the presence of both Ser10ph and Ser28ph within 10 minutes of stimulation (Soloaga et al., 2003). The dynamic acetylation of H3ph occurs at lysine 9 and both H3S10ph/H3K9acS10ph have been detected at IE genes (Thomson et al., 1999, 2001). While characterization of H3 phosphorylation dynamics has been successful, the attempts to carefully map this modification with ChIP lead to scarce (<5% input) recovery of DNA fragments associated with H3S10ph or H3K9acS10ph (Lee and Mahadevan, 2009; Thomson et al., 2001). Such

low recoveries indicate that only a small proportion of nucleosomes at a given position carry H3S10ph. This also suggests that mitogen-inducible H3S10ph is a highly dynamic modification either turned over continuously or targeted asynchronously as a single pulse to individual genes. Limited evidence from *c-fos* and *c-jun* ChIP profiling with H3K9acS10ph-specific antibody indicate similarity with H3K9ac distribution (A.Clayton unpublished data). This vast array of modifications gives potential for functional responses, but the interpretation of data is complicated by the existence of microheterogeneity of histone modification within homogeneous, synchronised cell populations and the dynamics and reversibility of almost all modifications.

Figure 4. **Histone modifications across a typical mammalian chromosome.**

a) Histone modifications across a typical mammalian chromosome. **b)** Composite profile of typical histone modifications associated with active, poised, silenced genes and associated regulatory elements. **c)** Composite profile of typical histone modifications associated with an actively transcribed gene.

a)

b)

c)

adapted from (Lee and Mahadevan, 2009)

1.2.1.3 The histone code hypothesis

The large number of histone modifications identified has lead researchers to propose the existence of organized combinations of histone marks forming a ‘‘histone code”. It has been postulated that this code might define the specific set of modifications required for a specific function or ensures that a given modification on a specific histone residue determines subsequent modifications of the same histone or another histone molecule (Jenuwein and Allis, 2001; Kouzarides, 2007; Liu et al., 2005; Margueron et al., 2005; Rice and Allis, 2001). It is clear from multiple reports that certain combinations of histone marks have been preferentially associated with global chromatin environments, such as heterochromatin (see chapter 1.1.3.1) and euchromatin. The fact that protein domains have been preserved by evolution to recognize distinct modifications (for example: chromodomains and plant homeodomain recognize methylation, bromodomains target acetylation and 14-3-3 proteins bind phosphorylated residues) can potentially explain how the cell could interpret the code on histone tails (Campos and Reinberg, 2009; Kouzarides, 2007; Strahl and Allis, 2000; Turner, 2005). However, the “histone code” is still considered a controversial hypothesis because histone marks are redundant and mutagenesis has demonstrated that not all combinations of histone modifications lead to distinct transcriptional states (Dion et al., 2005). Genome-wide localization studies of histone modifications in yeast, flies, and mammals have also demonstrated that not all possible histone-modification patterns occur *in vivo* (Bernstein et al., 2005; Kurdistani et al., 2004; Schübeler et al., 2004). The recent debate on the dynamic nature of most histone modifications complicate the interpretation of the interaction between histone marks and their “readers” (Clayton et al., 2006; Lee and Mahadevan, 2009).

It may be that global chromatin environments are established by a rather permissive marking system characterised by particular levels of acetylation and methylation of histones. This does not exclude the presence at similar positions of different marks at the same residues (as

H3K4me1/2/3) or different combinations of residues (like multiple H3 lysine acetylation and methylation) (Campos and Reinberg, 2009; Kouzarides, 2007; Lee and Mahadevan, 2009). At the nucleosomal level, there is evidence that some sets of remarkably stable PTMs may be targeted by enzymes responsible for the deposition of a second or even third level of modifications (discussed in Crump et al., 2011; Edmunds et al., 2008; Hazzalin and Mahadevan, 2005) (see also chapter 1.4.1).

1.2.1.4 The function of histone PTMs, a general perspective

Acetylation results in a change in the net basic charge of histones that has the potential to loosen inter- or intranucleosomal DNA-histone interactions. It has been observed that acetylated histones are easier to displace from DNA both *in vivo* and *in vitro* (Chandy et al., 2006; Hassan et al., 2006; Ito et al., 2000; Reinke and Hörz, 2003; Zhao et al., 2005). Histone acetylation occurs at multiple lysine residues and is usually carried out by a variety of HATs (Brown et al., 2000). However dynamic acetylation turnover is targeted to specific residues, occurring at lysine 9 on H3 tails carrying H3K4me3 (Clayton et al., 2006; Crump et al., 2011; Hazzalin and Mahadevan, 2005). Another example of acetylation at a specific lysine is that of H4K16ac (Dion et al., 2005). H4K16ac inhibits the formation of compact 30 nm fibers but also impairs the efficiency of ATP-dependent chromatin assembly and mononucleosome mobilization by the ACF histone chaperone (Shogren-Knaak et al., 2006).

Phosphorylation is another modification that has the potential of influencing chromatin compaction via charge changes. There is no rigorous demonstration of such an effect *in vitro* but the role of phosphorylation in mitosis, apoptosis, and gametogenesis strongly suggest a function for the negative charges that phosphorylation introduces (Ahn et al., 2005; Fischle et al., 2005; Krishnamoorthy et al., 2006). The phosphorylation of histone H3 observed during the induction of IE genes has been linked to both gene transcription and histone H3

acetylation (Thomson et al., 2001; Soloaga et al., 2003). However, phosphorylation seems to act on a minute fraction of nucleosomes and its effects on efficiency of transcription are somehow limited (Soloaga et al., 2003). There is evidence that phosphorylation plays a role in modulating histone acetylation in yeast; phosphorylation of Ser10 promotes Lys14 acetylation while Lys9 appears to be inhibited (Edmondson et al., 2002; Lo et al., 2000). In mammalian cells, although it has been shown that the two modifications can be targeted to the same histone H3 tail, inhibition of histone phosphorylation does not affect the acetylation of nucleosomes demonstrable both by inhibition of signalling or by genetic knock out of MSKs, indicating that phosphorylation and acetylation are independently delivered to histone H3 at IE genes (Clayton and Mahadevan, 2003; Soloaga et al., 2003; Thomson et al., 2001). Methylation does not directly alter nucleosome charge balance but appears to work as a recognisable epitope for protein binding. Methylation is recognized by chromo-like domains of the Royal family (chromo, tudor, MBT) and nonrelated PHD domains. H3K4me is one of the major targets of these proteins. H3K4me3 recruits nucleosome remodelling factor (NURF) complex, histone deacetylases complex (mSin3a-HDAC1), histone lysine demethylase (JMJD2A) and ATPase (CHD1) (Huang et al., 2006; Peña et al., 2006; Pray-Grant et al., 2005; Shi et al., 2006; Sims et al., 2005; Wysocka et al., 2006). Further methylation at other residues promotes the recruitment of enzymes. For example H3K27me is recognized by polycomb protein PC2, associated with ubiquitin ligase activity specific for H2A. HP1 protein binds H3K9me and is associated with deacetylase activity and methyltransferase activity (Kouzarides, 2007). As well as the promotion of binding, histone modification can prevent docking of proteins onto chromatin. H3K4me disrupts the binding of the NuRD complex and H3T3ph prevents the binding of the INHAT complex. Both complexes have a repressive effect on transcription (Margueron et al., 2005). Furthermore, acetylation is recognized by bromodomains, and phosphorylation is recognized by 14-3-3 proteins. Bromodomains are

found in more than 30 human proteins but these proteins are not equally targeted to acetylated histones. Bromodomain recognition of acetyl groups largely differ in their sequence dependence binding capabilities often being affected by the presence of additional protein subunits and by contingent residues flanking the acetylated lysine (Mujtaba et al., 2007). Phospho-binding protein 14-3-3 has been observed *in vivo* to bind phosphorylated and phospho-acetylated histones in a strictly phosphorylation-dependent manner. 14-3-3 proteins are inducibly recruited to *c-fos* and *c-jun* nucleosomes upon gene activation (Macdonald et al., 2005).

In addition to the functions of a single type of modification, it is important to consider the crosstalk between different marks. This communication can operate at different levels. At the same position, distinct types of modifications on lysines are mutually exclusive and between adjacent modifications, active competition or disruption may occur. For example, H3S10ph affect the binding of HP1 to methylated H3K9 (Fischle et al., 2005). Modification can compromise or increase substrate recognition by enzyme activities. Isomerisation of H3P38 affects methylation of H3K36 by Set2 (Nelson et al., 2006). The GCN5 acetyltransferase recognizes H3 more effectively when it is phosphorylated on Ser10 (Clements et al., 2003). Occasionally, modifications can also interact with different histone tails; for example, methylation of H3K4me3 seems to require ubiquitinylation of H2B (Kouzarides, 2007).

1.2.1.5 Histone PTMs, a dynamic perspective

Most histone PTMs have been found to be reversible and dynamic, with several degrees of variability, silencing modifications proving to be far more stable than PTMs associated with transcription. Several enzymes that remove histone modifications have been identified. Histone deacetylases (HDACs) remove acetyl groups and Ser/Thr phosphatases remove phosphate groups. Ubiquitin proteinases remove mono-ubiquitin from H2B. Arginine

methylation is altered by deiminases, which convert the side chain to citrulline. So far there are two classes of lysine demethylase: the LSD1 domain class (which removes H3K4me1/2) and the JmjC domain class (which removes H3K4me2/3, H3K9me2/3, and H3K36me2/3) (Berger, 2007; Lee et al., 2005; Shi et al., 2004 and reference therein). Of all the enzymes that modify histones, the methyltransferases and kinases are the most specific. The specificity of enzymes that modify histones can be influenced by chromatin factor complexes. Some enzymes show a preference for nucleosomal over free histones and proteins that associate with the enzyme may affect its selection of substrate or the level of methylation (mono-, di-, or tri-) at a specific residue (Lee et al., 2005; Metzger et al., 2005; Steward et al., 2006). Recent advances of ChIP techniques allow us to observe localization of multiple marks throughout the genome. The enormous data sets generated have shown that modification patterns can distinguish nucleosomes at one location (e.g., promoter nucleosomes) from those at another location, and modifications involved in transient function, such as transcription and DNA repair are highly dynamic and transient (Barski et al., 2007; Clayton et al., 2006; Lee and Mahadevan, 2009; Wang et al., 2008). Also, high resolution studies at individual mammalian loci provide evidence that histone modification in transcription acts not as static landmark, but more as a dynamic and transient operational mark (Edmunds et al., 2008). Histone H3 methylation, acetylation and phosphorylation and H4 acetylation at IE genes appears to be a dynamic process that is continuously turning over, even in unstimulated cells (Clayton et al., 2006; Hazzalin and Mahadevan, 2005; Lee and Mahadevan, 2009; Thomson et al., 2001). In the case of H3 acetylation, it has been shown that HAT and HDAC activity at IE genes is in a regulated balance. Induction of MAPK signalling shifts the balance in favour of acetylation. The situation is highly dynamic and not all nucleosomes are uniformly and stably modified along the gene (Thomson et al., 2001). Studies of these genes have also demonstrated the simultaneous presence at equivalent positions within homogeneous

synchronised cells, of distinct populations of nucleosomes with different set of modifications (Lee and Mahadevan, 2009). Nevertheless, it is still possible to distinguish sets of modifications that are unlikely to be uniformly present on equivalent nucleosomes of a specific IE gene across an entire cell population. H3K4me3, H3K36me3 and H3K79me2 might not be present at the same position in every cell but their distribution is still limited within distinct boundaries (Edmunds et al., 2008). Thus along the length of a single gene, the PTMs on nucleosomes distinguish histone populations that are mutually exclusive (not present at the same locus), partially overlapping (distributed in contiguous areas) or completely overlapping (Lee and Mahadevan, 2009).

1.2.2 The architectural proteins: HMGs, CTCF, H1

The intense work on histone modification has relied on the advantage of analysing the abundant and easily extractable core histones. An additional large class of architectural proteins is present associated with chromatin, proteins that bind to DNA and nucleosomes and induce structural changes in the chromatin fiber. Architectural proteins include linker histone H1, CTCF and members of the high-mobility group (HMG) protein superfamily.

H1 helps establish the higher order compact chromatin structure by promoting chromatin compaction. H1 stabilizes nucleosomal structures and obstructs nucleosomal access to regulatory factors, nucleosome remodelling complexes, and histone modifiers to their chromatin binding sites (Brown, 2003; Cheung et al., 2002; Herrera et al., 2000; Hill and Imbalzano, 2000; Horn et al., 2002; Koop et al., 2003; Thomas, 1999; Vicent et al., 2002b). H1 depletion has been shown to result in both up-regulation and down-regulation of specific genes, implying a role for H1 in transcriptional regulation (Alami et al., 2003; Hellauer et al., 2001; Shen and Gorovsky, 1996; Takami et al., 2000).

HMG proteins were characterized in 1973 by Goodwin et al., as a new group of chromatin-associated proteins, soluble in dilute acids and having a remarkable high mobility in polyacrylamide gel electrophoresis. HMGs were originally classified by acid-extractability and their mobility but better understanding of their properties has lead to a new nomenclature (Bustin, 2001). The HMG superfamily is composed of three families: HMGB, HMGA, and HMGN, each family being defined by a distinct DNA- or chromatin-binding motif (Bustin, 1999). HMGs promote decompaction of higher-order chromatin structures, maintaining a dynamic balance between compaction and decompaction of chromatin (Agresti and Bianchi, 2003; Bustin, 1999; Reeves, 2001; Thomas and Travers, 2001). HMG proteins bind chromatin without DNA sequence specificity and seem to play a role on genome wide rather than locus specific chromatin dynamics (Bianchi and Agresti, 2005; Hock et al., 2007).

CTCF (CCCTC-binding factor) is one of the best characterized human transcription factors that can act as an insulator (Bell et al., 1999; Filippova, 2008; Kim et al., 2007b; Ohlsson et al., 2001; Wallace and Felsenfeld, 2007). It is an essential 11-zinc finger (ZF) DNA binding domain protein, ubiquitously expressed and highly conserved from Drosophila to mice and man (Fedoriw et al., 2004; Moon et al., 2005). CTCF plays a major role in mediating long-range chromatin interactions in transcription and/or epigenetic regulation, including regulation of genomic imprinting and X-chromosome inactivation. CTCF is involved also in noncoding transcription and establishing local chromatin structure at the repetitive elements in mammalian genomes. The role of an insulator protein such as CTCF is to contain the spread of heterochromatin and restrict transcriptional enhancers preventing activation of unrelated downstream promoters (Bulger and Groudine, 1999). Insulators have also been suggested to facilitate the formation of loops, via attachment of chromosomal regions to the nuclear membrane, keeping the intermediate regions exposed for only local interactions between enhancers and promoters (Yusufzai et al., 2004).

1.2.3 HMG-As, HMG-Bs, HMG-Ns

High mobility group proteins (HMGs) are, after histones, the second most abundant chromatin proteins (Bianchi and Agresti, 2005; Johns, 1982). HMGs (11 proteins, including the splice variants) are small nuclear proteins that organize chromatin by appropriately bending and modelling DNA, establishing active or inactive chromatin domains. They are involved in the response to both rapid environmental changes and developmental processes. HMGs interact with nucleosomes, transcription factors, nucleosome-remodelling machines, and with histone H1 (Bianchi and Agresti, 2005; Gerlitz et al., 2009). They compete with histone H1 for chromatin binding sites and affect its dynamics providing flexibility to chromatin structure and a fine tuning of gene transcription (Catez et al., 2004).

HMG proteins are classified into three families (HMGA, HMGB and HMGN) and despite sharing similar function they have very different structures and specific post-translational modification pathways. HMG proteins interaction with chromatin is highly dynamic and reversible (Figure 5). HMGAs interact mainly with multiprotein complexes of transcription factors and cofactors, HMGNs bind to the nucleosome surface and HMGBs to both nucleosomes and transcription factors (Bianchi and Agresti, 2005; Gerlitz et al., 2009).

HMGA proteins are named after their three AT-hook domains, protein segments of nine amino acid that interact with AT-rich DNA stretches in the DNA minor groove (Reeves and Nissen, 1990; Solomon et al., 1986). HMGB proteins also bind into the minor groove of DNA, using their HMG boxes, two tandem 80 amino acid domains, but with limited or no sequence specificity (Thomas and Travers, 2001). HMGN proteins bind specifically to nucleosomes, independently of DNA sequence (Shirakawa et al., 2000a; Ueda et al., 2008).

1.2.3.1 HMGAs

HMGA proteins are strongly expressed during early developmental stages and are less abundant in adult cells (Sgarra et al., 2004). HMGAs play a key role in the assemblages of enhanceosomes, multiprotein complexes of transcription factors and cofactors on nucleosome-free control regions of genes. HMGA1a, the transcription factors NF-kB, IRF (interferon response factor), and ATF2/c-Jun form a well known enhanceosome on the interferon-b (IFN-β) promoter (Bianchi and Agresti, 2005). HMGA1a distorts the DNA, forcing it into a more accessible state for the transcription factor binding sites. Acetylation mediated by GCN5 (general control of amino acid synthesis 5) occurs at nucleosomal histones and at HMGA1aK64. This is supposed to stabilize the enhanceosome and favour the recruitment of the SWI–SNF nucleosome- remodelling machinery, responsible for nucleosome displacement at the TATA box, thus promoting the recruitment of TBP (TATA-binding protein) and the transcription machinery (Munshi et al., 1998). CBP (CREB-binding protein) mediates the disassembly of the enhanceosome by acetylation of HMGA1aK70 (Munshi et al., 2001) . HMGA1 possibly regulates many other genes in a similar fashion; these include interferon-g, E-selectin, interleukin-2 receptor-α, the inducible form of nitric oxide synthase (iNOS), and the insulin receptor (Attema et al., 2002; Chau et al., 2005; Darville et al., 2004; Foti et al., 2005; Lewis et al., 1994; Martinez Hoyos et al., 2004; Treff et al., 2004; Whitley et al., 1994). HMGA1 also negatively regulates BRCA1 and has been suggested to play a role in defining cellular nuclear matrix domains by binding segments of AT-rich DNA at S/MARs (scaffold-/matrix- associated regions) (Sgarra et al., 2004). HMGA1 modification is not limited to acetylation, as phosphorylation and methylation are also frequently observed (Gerlitz et al., 2009). Phosphorylation results in decreased binding of DNA while methylation has been strictly related to apoptosis (for details see Zhang and Wang, 2008). HMGA1a is highly mobile and present both in euchromatin and in

heterochromatin. It is localized preferentially in heterochromatin *in vivo* due to significantly longer residence time on the chromatin of condensed chromosomes (Harrer et al., 2004). *Hmga1* gene is not essential but impaired expression affects sperm production in heterozygote male mice and *Hmga1*-/- mice are diabetic as a result of reduced expression of the insulin receptor in all tissues (Foti et al., 2005; Liu et al., 2003). In human, reduced HMGA1 expression is associated with reduced expression of the insulin receptor as well and patients show insulin resistance and type 2 diabetes. Mutation in the mouse *Hmga2* gene correspond to reduced growth and the near absence of fat (Anand and Chada, 2000). HMGA overexpression is a hallmark of several malignant and benign tumours (Fedele et al., 2001; Reeves, 2001; Sgarra et al., 2004).

1.2.3.2 HMGBs

HMGBs have been intensively studied in mammals, Drosophila and yeast. HMG boxes can be present alone, in tandem or embedded into large protein domains, and in all of these configurations they can interact with and bend DNA (Thomas and Travers, 2001). While HMGA AT-hooks cause modest distortions, HMG boxes can bend DNA to right angles or even more sharply. HMGBs in mammals possess two tandem HMG box domains followed by a long acidic tail. The two boxes bind DNA independently and the tail appears to remain unstructured (Knapp et al., 2004). HMGBs are supposed to influence gene transcription by interacting directly with nucleosomes, enhancing accessibility by loosening the DNA wrapping or by constraining a tight loop of DNA at the edge of a nucleosome (Agresti and Bianchi, 2003; Bonaldi et al., 2002; Travers, 2003; Wu and Travers, 2004). Often HMGB1 DNA-bending stabilizes a partner protein on its DNA target sequence. HMGBs do not alter the sequence specificity of their partners and increase binding affinities for DNA by a similar magnitude for both primary and secondary binding sites (Das and Scovell, 2001; Melvin et

al., 2004). HMGB protein partners are: TBP, steroid (class I) nuclear receptors, HOX and POU proteins, p53 and p73, and several NF-kB subunits (Agresti and Bianchi, 2003; Das and Scovell, 2001; Dasgupta and Scovell, 2003; Najima et al., 2005). The interaction of HMGB with chromatin is dynamic but sufficiently stable to be detected with the use of fluorescence resonance energy transfer (FRET) in living cells (Agresti et al., 2005). Studies on HMGB1 interaction with glucocorticoid receptors (GR) suggest that binding occurs only within chromatin. HMGB1–GR–chromatin complexes are stable but the complex is kept in a dynamic balance by ATP-dependent active processes (Agresti et al., 2005; Phair et al., 2004; Sapojnikova et al., 2005). HMGBs are also found in Epstein-Barr virus enhanceosomes and Gamma-2 herpes viruses, where HMGB1 stabilizes binding of their RTA (replication and transcription activator) protein to different promoters (Song et al., 2004). HMGBs are also involved in maintenance of genome integrity in DNA repair, transposition and V(D)J recombination. They recognize distorted and damaged DNA structures and contribute to the correct disposition of chromatin structures (Agresti and Bianchi, 2003). HMGB protein expression is high in the mouse embryos and HMGB1 remains ubiquitous in the adult, while HMGB2 and HMGB3 are only expressed in the testis and lymphoid organs (Müller et al., 2004). The three mammalian HMGBs have partially redundant functions, *Hmgb1* is essential but *Hmgb2*-/- mice are viable, *Hmgb3*-/- deficient mice are viable but erythrocythemic (Hock et al., 2007; Nemeth et al., 2005). In inflammatory cells, acetylation of HMGBs promotes their translocation into the cytoplasm and their secretion (Bonaldi et al., 2003). Secreted HMGB1 has been reported to act as a cytokine and has pathological effects in sepsis, arthritis, cancer and other diseases (Lotze and Tracey, 2005). *In vivo*, HMGB1 and HMGB2 bind equally well, and with similar dynamics, to both euchromatin and heterochromatin (Pallier et al., 2003).

Mammals have a larger repertoire of HMG boxes whereas plants, fungi and unicellular eukaryotes mostly have short proteins with a single HMG box and rarely have proteins containing tandem HMGB boxes (Cetkovic H, Lukic-Bilela L, 2003). Drosophila possesses HMGB-like proteins with a single HMG box, HMG-D (HMG Drosophila) and HMG-Z (HMG zygotic), and DSP1 (dorsal switch protein 1) that has two tandem HMG boxes and is a true orthologue of HMGB. DSP1 mediates recruitment of polycomb group (PcG) complexes to polycomb response elements (PREs). This is suggesting a role for HMGB proteins in chromatin repression (Déjardin et al., 2005). It was also shown that HMGB2 forms a protein complex with nucleolin and YY1 (Yin Yang 1) proteins. HMGB2-YY1 complex binds the D4Z4 repeats located in the subtelomeric region on the long arm of human chromosome 4 (Gabellini et al., 2002). Knockdown of HMGB2, nucleolin or YY1 relieves D4Z4-dependent silencing of the subtelomeric domain.

Figure 5. **Architectural proteins: HMGAs and HMGBs.**

a) HMGAs contain three AT hooks (green) and an acidic C-terminal part (blue), except HMGA1c, which contains only two hooks. HMGB proteins contain two HMG boxes (yellow) and an extended acidic C-terminal (blue). **b)** HMG proteins alter chromatin structure by a variety of mechanisms. **(1)** HMGA and HMGB proteins DNA bending activity alter the chromatin structure and facilitate the binding of additional factors. **(2)** HMGs are part of the binding module of multiprotein complexes. For example, HMGA contributes to the formation of enhanceosomes. **(3)** HMGs are either preventing or facilitating access for modulating factors to chromatin. **(4)** HMG proteins compete with other nuclear proteins for binding sites. **(5)** HMGs alter the chromatin structure unfolding or compacting chromatin.

a)

b-1)

b-2)

b-3)

b-4)

b-5)

adapted from (Hock et al., 2007)

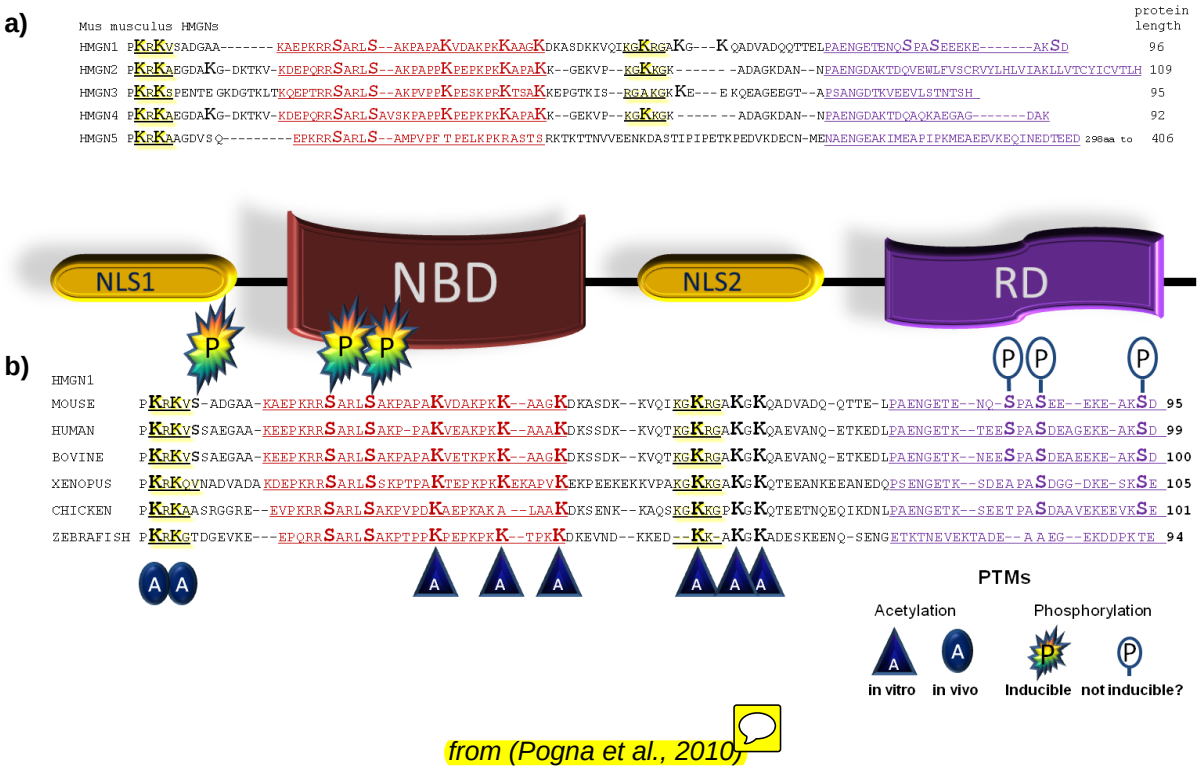
1.2.3.3 HMGNs

HMGNs are the only non-histone proteins known to bind specifically to nucleosomes and they are found in the nuclei of most vertebrates cells (Bustin, 2001b; Bustin and Reeves, 1996; Gerlitz et al., 2009; West, 2004). There are five closely related proteins currently classified in the HMGN family: HMGN1 (HMG-14), HMGN2 (HMG-17), HMGN3a, HMGN3b, HMGN4 and HMGN5 (NBP-45) (Figure 6). HMGNs are all of approximately 10 kDa mass (except HMGN5 that posseses a long C-terminus) and the conserved hallmark of this class of proteins, the nucleosome-binding domain (NBD), consists of a positively charged domain of 30 amino acids. Among the members of the HMGN family, HMGN1 and HMGN2 are the best described. They specifically bind the nucleosome core particle (CP) on the nucleosomal dyad axial surface (Figure 7a) occupying a position located ~25 base pairs from the end of the nucleosomal DNA (Alfonso et al., 1994; Bustin, 2001b; Kugler et al., 2012). The H3 histone tail and H2B are involved in the interaction. Nucleosomes predominantly contain homodimers of HMGNs, either HMGN1 or HMGN2, but not mixed HMGN1/2 populations (Bustin, 2001b; Postnikov et al., 1995). HMGN positioning on DNA is not sequence specific and it is highly dynamic (residence-time ranges from 4 to 25 seconds) (Phair and Misteli, 2000; Shirakawa et al., 2000a). HMGNs unfold higher order chromatin structure, increase the accessibility of the nucleosomal DNA and facilitate DNA transcription, repair and replication. These proteins also compete with H1 interaction with chromatin as part of their unfolding activities (Figure 7b) (Catez et al., 2004). Levels of expression of these proteins vary between different tissues and developmental stages (reviewed in Gerlitz et al., 2009). However, the general consensus in the HMGN field is that HMGN/nucleosome ratio varies from 0.01 to 0.1. Thus, the availability of HMGN molecules for binding to nucleosomes is the limiting factor in their interaction. In interphase cells, the majority (>80% or >90% depending on reports) of HMGNs are dispersed through the nucleus

with the remainder being found in the cytoplasm (Louie et al., 2000; Phair et al., 2004). Over expression and knock-down of different HMGNs affects transcription of specific genes (Rochman et al., 2011). Transcriptional modulation requires the combination of a functional NBD with the specific C-terminal domain of each HMGN. Changes of expression are subtle and it is not possible to recognize any common features among the genes affected by each HMGN. However it is clear that HMGN subtypes are not fully redundant. These results also suggest that HMGNs contribute more to the optimization and fine tuning of transcription. HMGNs undergo several posttranslational modifications (Hock et al., 1998a; Pogna et al., 2010; Zhang and Wang, 2008). Stimulation of the ERK and p38 pathways leads to activation of the MSK kinases that concomitantly phosphorylate H3Ser10 and Ser28, and HMGN1Ser 6, the so-called "nucleosomal response" (Soloaga et al., 2003; Thomson et al., 1999). The nucleosomal response is invariably correlated with transcriptional induction of immediate-early (IE) genes and part of the aim in this thesis is to elucidate the role of HMGN1 in IE gene activation. HMGN1 phosphorylation is supposed to weaken its interaction with nucleosomes (Lim et al., 2004, Figure 8) but it is unclear to which extent phosphorylation occurs and which specific residues are involved in gene activation through MAPK signalling (Pogna et al., 2010). HMGN1 detachment from mitotic chromatin concomitantly with H3Ser10 phosphorylation during mitosis suggest that a similar situation might occurs in transcriptional activation (Bianchi and Agresti, 2005). Being the subject of this study, HMGN1 properties are discussed in detail in chapter 1.3.

Figure 6. **Sequence conservation of HMGNs.**

Sequence information is derived from the Swiss-prot/NCBI database (accession numbers in brackets). **a)** Mus musculus HMGNs: HMGN1 (P18608), HMGN2 (Q5XK38), HMGN3a (Q9DCB1), HMGN4 (XP_001473154.1) and HMGN5 (Q9JL35). For HMGN5, only amino acid residues 1–108 are shown for clarity and due to lack of homology of the protein (residues 109–406). **b)** HMGN protein domains and alignment of vertebrate HMGN1. Mouse (P18608), human (P05114), bovine (P02316), xenopus (Q9I988), chicken (P12274), and zebrafish (A1A5Y7). Residues that are post-translationally modified are shown in bold. Phosphorylation of all six residues has been reported in *vivo* and in *vitro*. Phosphorylation on Ser6–20–24 has been observed to be inducible by specific stimuli and is marked by star symbols. Phosphorylation of Ser85–88–98 has not been related with any specific induction and is marked with balloon symbols. Acetylation has been observed at different residues. Major sites identified in *vitro* (Lys30, 37, 41, 54, 58 and 60) are marked with triangles. Major sites identified in *vivo* (Lys2, 4) are marked with ovals.



from (Pogna et al., 2010)

Figure 7. **Schematics of HMGN binding to the nucleosomes.**

a) Distribution of assigned histone methyl groups in isoleucine, leucine, and valine (ILV) in the nucleosome showing top **(i-iv)** and side views **(ii)**. The backbone of the histones is shown with ribbons, ILV methyl groups as spheres. Colour coding: H2A, orange; H2B, red; H3, blue; H4, green; DNA, gray. **(iii)** schematic model. **b)** Dynamic modulation of chromatin structure by HMGNs. HMGN proteins compete with histone H1 for nucleosomes. HMGNs interact with histone H3 and H4 tails. The C-terminal domain of HMGNs interacts with the N-terminal of histone H3, and the binding of the NBD with the acidic patch in the histone octamer competes with the binding of the N-terminal of H4 from a neighbouring nucleosome.

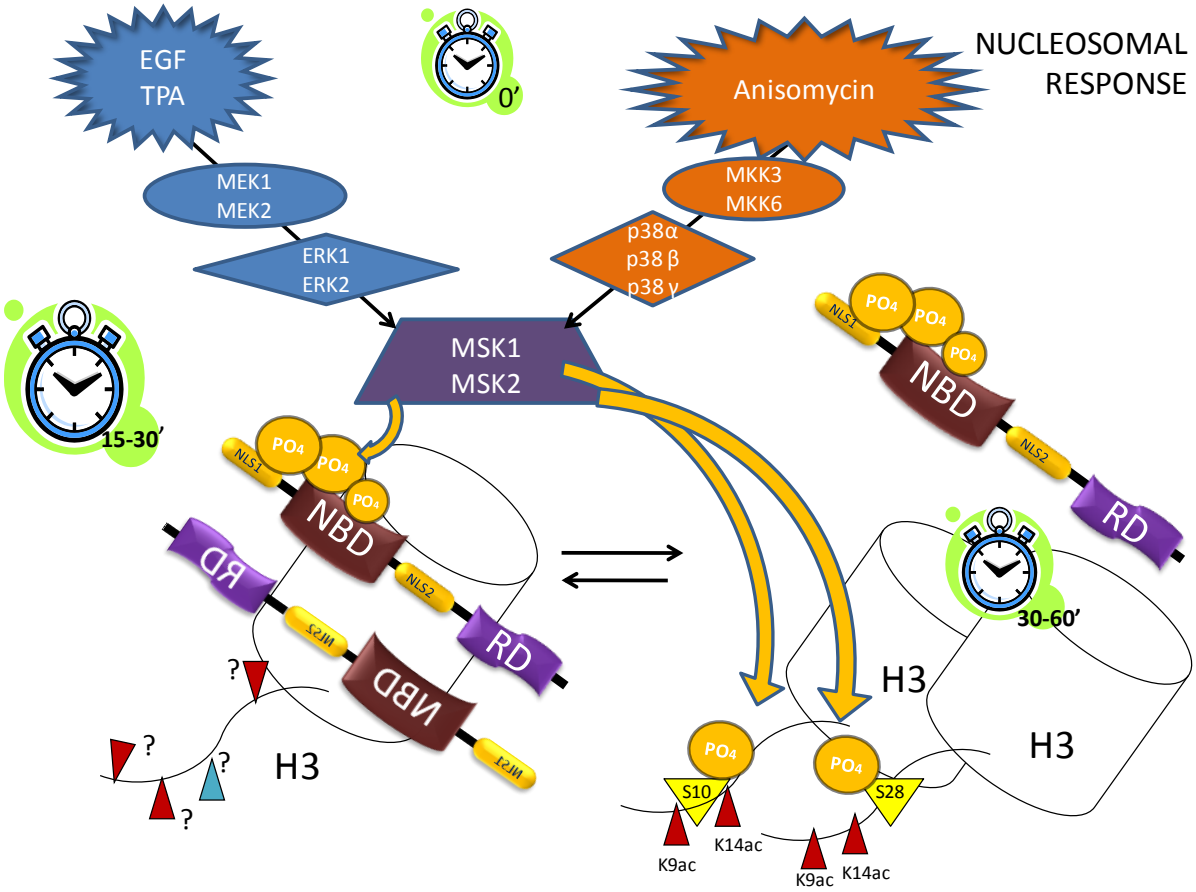
a)

adapted from (Alfonso et al., 1994; Kato et al., 2011)

b)

adapted from (Kugler et al., 2012)

Figure 8. **A proposed model for the signalling to chromatin through HMGN1.**
Induction of MAPK pathways (ERK and p38 kinases) leads to activation of MSKs. HMGN1 is phosphorylated at several residues (Ser6, Ser19 and/or Ser23) promoting HMGN1 dissociation from nucleosomes. Histone tails of nucleosomes devoid of HMGN1 are more accessible for interaction with activated MSKs that can phosphorylate either H3 Ser10 or Ser28. Additional modifications on H3 tails might include Lys9 and/or Lys 14 acetylation.



from (Pogna et al., 2010)

1.2.4 H1

The fifth histone, termed H1 or linker histone, comprises the most dynamic and divergent class of histone proteins. H1 stabilizes condensed chromatin structures and regulates chromatin accessibility of regulatory proteins, chromatin remodelling factors and histone modification enzymes to their target sites (Happel and Doenecke, 2009). H1 associates with the linker DNA (about 20 nucleotide pairs) and seals the 147 base pairs of the two rounds of DNA wrapped around the nucleosome by sitting at the DNA entry/exit site on the surface of the nucleosome core (Simpson, 1978). H1 is a lysine-rich protein family counting in mammals eleven different subtypes. H1 variants differentiate between subtypes expressed in germ cells, the testicular isoform genes H1t, H1T2 and HILS1, and the gene encoding the oocytic H1oo, and genes expressed in somatic cells, in a variant mode, H1.0 (or H1°) and H1x, or specifically during the S phase of the cell cycle, H1.1 to H1.5 (Happel and Doenecke, 2009). H1 main subtype genes are constitutively expressed, apparently in all cells (H1.1 to H1.5) (Happel and Doenecke, 2009). The H1 histones have a three domain structure, a short N-terminal domain with a length of about 45 amino acids and enriched in basic amino acids, a central globular domain of about 75 amino acids, highly conserved among all H1 subtypes and a C-terminal domain of about 100 amino acids, highly enriched in lysine, serine and proline. The C-terminal portion of the protein is unstructured in aqueous phase and DNA binding directs the acquisition of its specific secondary structure (Roque et al., 2005). The central globular H1 domain asymmetrically interacts with DNA at the nucleosome. H1 is implicated in the formation of a higher order chromatin structure, the 30 nm fiber of nucleosomes arranged in a solenoidal shape (Thoma et al., 1979). The exact disposition of nucleosomes in the solenoidal fiber is unclear (Bednar et al., 1998; Robinson and Rhodes, 2006; Robinson et al., 2006). Two models describe binding of H1 to DNA wrapped on nucleosomes. One model proposes that the globular H1 domain bridges two DNA duplexes

between the dyad axis of the nucleosome core between the dyad and one of the ends of the chromatosomal DNA, while a second one suggests that the globular H1 domain is positioned 65 base pairs away from the nucleosome dyad inside of one of the two chromatosomal DNA turns (Hayes et al., 1996; Pruss et al., 1996; Travers, 1999; Zhou et al., 1998).

What determines the difference between H1 subtypes binding to chromatin is the C-terminal tail which is the most divergent of the three domains of linker histones (De et al., 2002; Hendzel et al., 2004; Lever et al., 2000). The C-terminal domain makes up about half of the H1 molecule and can be used to classify the H1 subtypes into three major binding groups: highly mobile subtypes, H1.1 and H1.2, with low affinity for chromatin, H1.0 and H1.3 with moderate affinity and H1.4 and H1.5, the least mobile, with high affinity for chromatin (Th'ng et al., 2005). A correlation exists between the binding affinity and the length of the C-terminal domain of the different H1 subtypes, but H1.0 makes an exception since it has a moderate affinity for chromatin, but it is the H1 subtype that has the shortest C-terminal tail.

Once regarded just as a structural component of chromatin, the linker histone is now considered a key player in the highly dynamic network of interactions between chromatin and regulatory factors (Brown, 2003; Bustin et al., 2005). Interaction of H1 with a nucleosome is transient, with a residence time of 3 to 4 min and rapid interchanging between binding sites (Lever et al., 2000; Misteli et al., 2000; Phair et al., 2004). Such a residence time is short for a histone protein but far longer than a transcription factor or HMG protein that have mean residence times of 5 to 25 s (Catez et al., 2006; Phair et al., 2004). H1 histones are subjected to competition by other chromatin-binding proteins. Each of the three subfamily of the HMGs proteins (HMGA, HMGB, HMGN) can increase H1 mobility, displacing H1 from chromatin and the competition kinetics observed with different HMG subtypes suggests that different binding sites may be affected through different HMG proteins (details on HMGN1 dynamic interactions with histone H1 in chapter 1.3.3.2) (Catez et al., 2004).

Phosphorylation is the major PTM of histone H1 (Lever et al., 2000), it progressively increases during the cell cycle and is strongest during G2 and mitosis, then decreases sharply at the end of mitosis in telophase (Boggs et al., 2000; Bradbury et al., 1974; Gurley et al., 1978; Hohmann et al., 1976; Langan et al., 1989). H1 phosphorylation appears to be involved in chromatin decondensation during the S phase, but is also required for chromatin condensation during M phase (Alexandrow and Hamlin, 2005, reviewed in Baatout and Derradji, 2006). The phosphorylation of specific sites in the C-terminal domain of H1 is mediated by cyclin-dependent kinase and causes a conformational switch in the structure of the C-terminal domain, decreasing the proportion of α-helix and favouring the β-sheet content (Roque et al., 2008). The effect of phosphorylation is cumulative and the fully phosphorylated C-terminal domain possesses higher aggregation capacity for DNA than the partially phosphorylated form.

Phosphorylation of H1 is involved also in regulating gene expression. Phosphorylated H1 is enriched at transcriptionally active chromatin domains (Chadee et al., 1995; Lu et al., 1995). Experiments on minichromosomes have shown that incorporation of H1 histones improves gene regulation by stabilizing nucleosome positioning such that basal transcription is decreased, but upon induction transcription is higher in H1 enriched chromatin. Phosphorylation of H1 promotes ATP-dependent chromatin remodelling from complex NURF that opens the nucleosome structure. H1 is reported to leave the promoter after the formation of the transcription initiation complex (Koop et al., 2003; Vicent et al., 2002b). Phosphorylation of H1 is also reported to reduce H1 binding to nucleosomes resulting either in the complete removal of these proteins or at least of their phosphorylated C- and/or N-terminal tails from chromatin (Contreras et al., 2003; Dou et al., 2002; Green et al., 1993; Hendzel et al., 2004; Lever et al., 2000). H1 is subjected to several additional modifications more recently discovered, including acetylation, methylation, N-formylation of lysine, and

ubiquitination (Deterding et al., 2008; Garcia et al., 2004, 2006; Jiang et al., 2007; Sarg et al., 2006; Wisniewski et al., 2007, 2008). Most of the acetylated lysines identified in H1 are located at the globular domain at positions involved in DNA binding (Goytisolo et al., 1996; Wisniewski et al., 2007). Thus, if acetylation interferes with H1 binding to DNA, it is not surprising that deacetylation of the lysine-26 (K26) in H1 by SirT1 promotes facultative heterochromatin similarly to the deacetylation of H3 and H4 (Vaquero et al., 2004). SirT1is part of the multisubunit methyltransferase complex PRC4 (Polycomb repressive complex 4), that not only removes acetylation from K26, a highly conserved residue, present in all somatic H1 subtypes, but methylates H1 at this same residue. PRC4 preferentially methylates K26 of H1.4, while H1.2 K26 is targeted by PRC2 (Kuzmichev et al., 2004, 2005). Methylated K26 of H1.4 is recognized and bound by HP1, a previously mentioned element of heterochromatin and involved in transcriptional repression (Daujat et al., 2005). H1 is the most frequently formylated histone class with thirteen formylated lysine residues, distributed on all three structural domains of H1. Most of these sites are lysine residues important for DNA binding where acetylation or methylation have also been observed (Wisniewski et al., 2008).

1.3 HMGN1 role in chromatin remodelling

Most of the interactions between DNA and proteins inside chromatin have been long considered as relatively static and structural. Nowadays, it is evident that chromatin interactions occur either at slow dynamics, like the nucleosome core–DNA interaction, responsible for the packaging and physical organization of the genome, or at fast dynamics like the binding of architectural proteins, enzymes and regulatory proteins (Misteli, 2005; Phair et al., 2004). Among the proteins that interact with chromatin in a highly dynamic fashion, HMGs are well characterized examples. The intrinsically highly dynamic interaction between HMGN1 and the nucleosome has obstructed the definition of the role of this protein in chromatin regulation. Thus, research has focused on defining HMGN1 structure, its protein domains and modifications, HMGN1 expression and more recently, HMGN1 localization in chromatin.

1.3.1 HMGN1 structure

The HMGN family consists of five closely related proteins, HMGN1 (HMG-14), HMGN2 (HMG-17), HMGN3a, HMGN3b (missing C-terminal domain), HMGN4 and HMGN5 (formerly NSBP1, or NBP-45) (Figure 6 from Pogna et al., 2010; Gerlitz et al., 2009; Rochman et al., 2009; Shirakawa et al., 2000b). The most conserved domain of this class of proteins is their 30-amino acid long nucleosome-binding domain (NBD) (Figure 6). This domain is flanked by two highly conserved nuclear localisation signal (NLS) domains that are both necessary for active transport of HMGNs into the nucleus (Hock et al., 1998). A negatively charged regulatory domain (RD), previously named the chromatin unfolding domain (CHUD) is found at the C-terminal of HMGNs (Gerlitz et al., 2009). HMGN proteins are highly conserved through evolution and in many respects also amongst the five different

members of the family, although each HMGN subtype belongs to clearly distinct subgroups (Bustin and Reeves, 1996). The highest degree of conservation is observed within their functional domains, particularly the NBD, and many key residues are identical even amongst different classes of HMGNs. To date, phosphorylation and acetylation are the only PTMs identified on HMGN proteins (reviewed in Pogna et al., 2010; Zhang and Wang, 2008). Sites of *in vitro* and *in vivo* phosphorylation and acetylation of HMGN1 are shown in Figure 6. Acetylation sites are very highly conserved between all organisms analysed so far, while only those serine phosphorylation sites lying within the NBD are similarly conserved. Phosphorylation sites at the C-terminal are present in mammals HMGN1 and they are not present in chicken, zebrafish or Xenopus HMGN1. In what follows, I first provide a discussion of the general features and functions of each HMGN1 domain followed by a more extensive discussion of modification sites and their roles.

1.3.1.1 NBD, NLS domains and HMGN1 cellular localization

HMGN proteins recognize and bind the generic structure of the 147-bp nucleosome core via their nucleosome-binding domain (NBD). HMGN1 can also interact with purified DNA but not with isolated histones (Abercrombie et al., 1978; Ueda et al., 2008). The NBD is defined by its property to be the minimum peptide that binds specifically to nucleosome core particles (CP) independently of DNA sequence or histone modification (Crippa et al., 1992; Shirakawa et al., 2000a; Ueda et al., 2008). This domain is able to fold autonomously and incubation of permeabilized cells with a peptide corresponding to the domain sequence causes specific displacement of HMGN proteins from chromatin and interestingly arrests polymerase II transcription (Hock et al., 1998). Within this domain lies the RRSARLSA motif, an eight-amino-acid motif that is absolutely conserved in all HMGNs. This motif is interchangeable between HMGN1 and HMGN2 without any effect on the chromatin residence time making it

functionally independent from the context of the whole protein (Ueda et al., 2008). Point mutation of any of the underlined R, S, R or L residues is sufficient to disrupt its function, also causing mislocalization of the proteins into the nucleolus, an organelle from which the wild-type protein is excluded (Prymakowska-Bosak et al., 2002; Ueda et al., 2008). The first NBD-serine (in human, Ser20 in HMGN1, Ser24 in HMGN2) is absolutely necessary for the domain function; even conservative replacements by structurally similar residues such as cysteine, alanine, threonine, or asparagine abolished the interaction of HMGNs with chromatin. Deletion mutants have shown that the C-terminal domain strengthens the affinity of the protein for chromatin and mutants lacking the C-terminal amino acids possess an affinity that is two-fold lower than that of the wild-type proteins (Ueda et al., 2006). However the specificity for nucleosome recognition resides only in the NBD (Ueda et al., 2008). HMGNs bind to nucleosomes in a position very similar to that of histone H1, on the dyad axial surface, with the C-terminal of the protein potentially interacting with DNA linker. HMGN1 binding to nucleosomes also promotes interaction between the N-terminal domain and a restricted region in histone H2B, whereas the C-terminal domain targets a restricted region in the N terminal of histone H3 (Kato et al., 2011; Kugler et al., 2012; Trieschmann et al., 1998) (Figure 7). Studies on hydroxyl radical DNA cleavage suggest that HMGNs also protect the ends of DNA possibly by bridging the two adjacent DNA strands on the surface of the particles (Alfonso et al., 1994). *In vitro* observations suggest that HMGN1 or HMGN2 binds to nucleosomes as homodimeric complexes and it appears also to be cooperative, with dimerisation occurring via cross-talk between HMGNs, leading to stretches of contiguous nucleosomes containing a uniform population of HMGN homodimers, either HMGN1 or HMGN2, but not mixed HMGN1/2 populations (Postnikov et al., 1995). All HMGNs bind to chromatin with very similar affinities both *in vitro* and *in vivo* (Postnikov et al., 1995; Ueda et al., 2008). The interaction of HMGN proteins with nucleosomes is reported to stabilize the

structure of the isolated CP, reducing the "compaction" of the higher-order chromatin structure and possibly altering the levels of post-translational modifications in the tail of the nucleosomal histones (Bustin, 2001b; Lim et al., 2004, 2005; Postnikov et al., 2006; Ueda et al., 2006).

Phosphorylation and acetylation of residues within the NBD have been reported. Acetylation marks have been detected mostly *in vitro* using recombinant enzymes or enzymes from HeLa nuclear extracts and, as substrate, either recombinant HMGN proteins or reconstituted nucleosome cores complexed with HMGNs. In these conditions, two enzymes have been shown to acetylate HMGNs, PCAF and p300 (Bergel et al., 2000). PCAF seems to exclusively target HMGN2 even under conditions when HMGN1 is in excess (Herrera et al., 1999). p300 acetylates both HMGN1 and HMGN2 at multiple sites, lysines 30, 37 and 41 in the NBD (Bergel et al., 2000). However it is evident from these experiments that even after extensive forced *in vitro* acetylation, only a small proportion of these proteins becomes extensively acetylated on all potential sites. Interestingly, acetylation of nucleosome-bound HMGNs is less efficient than free HMGNs suggesting the presence of additional cofactors for efficient acetylation of nucleosome-bound HMGN1/2 *in vivo*. [^{3}H]-acetate labelling of HeLa cells indicate that the NBD is acetylated *in vivo*, but unfortunately, the identity of the acetylated sites is still unclear (Bergel et al., 2000). However even *in vivo* experiments reported that a very small percentage of HMGNs is highly acetylated, approximately 3%. Such a small population is very elusive and this may explain why attempts to detect with western blotting multiply modified (acetylated) HMGN1 have failed to date (Barratt et al., 1994, Figure 16). Acetylation marks are likely to lower HMGNs binding affinity for nucleosomes and possibly this may facilitate enzyme access to core histone tails. p300/CBP and PCAF/GCN5 have an important role in histone acetylation during IE gene induction and their activity on HMGNs might be part of an acetylation cascade occurring at

these genes (Thomson et al., 1999). Nevertheless, it is important to highlight that at present, there is no direct *in vivo* evidence that p300 or PCAF are the enzymes that acetylate HMGN1/2 in living cells. The most striking observation remains that despite the high degree of sequence identity between their NBDs, HMGN1 and HMGN2 show significant differences with respect to the sites involved in acetylation and the enzymes responsible for it.

Phosphorylation of HMGN NBD is an equally controversial issue. A series of *in vitro* experiments have reported phosphorylation of HMGN NBD by PKC and PKA and antibodies have been generated against the peptide PKRRS(P)ARLS (P)AKP, containing the phosphorylated form of two conserved serines (Ser20-24 in human HMGN1) in the core NBD of HMGNs (Ser24 and Ser28 in HMGN2) (Prymakowska-Bosak et al., 2001; Ramachandran et al., 1984). The lack of information on the specificity of these antibodies leaves open the possibility that they may recognize a single phosphorylated serine instead of both Ser20 and Ser24 in phosphorylated form (see Pogna et al., 2010). Using these reagents, stimulus-induced phosphorylation and mitotic phosphorylation during metaphase have been detected. Mitotic phosphorylation of the NBD prevents formation of HMGN-nucleosomes complexes suggesting a role for HMGN1 in chromosome compaction. However, it has also been observed that HMGN1 still associates with condensed chromatin, but as monomers and with lower affinity (Cherukuri et al., 2008; Gerlitz et al., 2009). *In vivo*, NBD phosphorylation induced by stimuli has been postulated to be part of the phosphorylation cascade mediated by MSKs (Lim et al., 2004; Louie et al., 2000). According to the hypothesis of Lim et al., NBD phosphorylation of HMGNs by MSKs promotes exclusion from the nucleosome, allowing MSKs to phosphorylate H3 tails (Lim et al., 2004; Pogna et al., 2010).

As previously noted, reported cases of NDB phosphorylation investigated with NBD phospho-specific antibodies do not reveal the exact residues involved. Nevertheless, Ser20-24 residues are embedded in very different sequences, and while Ser20 has an MSK-substrate

motif whilst Ser24 does not. It is also known that Ser20 possesses additional characteristics. It is one of four critical residues for nucleosome binding and single Ser20 mutants behave identically to Ser 20-24 mutants regarding chromatin association, mobility and cellular localization (Prymakowska-Bosak et al., 2002; Ueda et al., 2008). Additional uncertainty regarding NBD phosphorylation arises from two independent [^{32}P] labelling studies under conditions of maximal MAP kinase and MSK activation (Barratt et al., 1994; Walton and Gill, 1983). Peptide mapping with [^{32}P] radiolabelled HMGN1 identified a single radiolabelled N-terminal peptide extending up to serine 19 (Ser20 in human), but excluding Ser23 (Ser24 in human).

The HMGN1 NBD contains a recognition motif for 14-3-3 protein binding. 14-3-3 are a family of seven different proteins β, γ, ε, σ, ζ, τ, and η involved in regulation of metabolism, protein trafficking, signal transduction, apoptosis and cell-cycle-specific regulation. HMGN NDB interactions with 14-3-3 ζ are phosphorylation-dependent and have been proposed to maintain HMGN1 outside the nucleus during mitosis, so that only unphosphorylated HMGN1 can re-enter the nucleus after mitosis (Morrison, 2009; Prymakowska-Bosak et al., 2002). The NBD domain is flanked by two highly conserved nuclear localisation signals (NLS) named after their involvement in active transport of HMGNs into the nucleus (Hock et al., 1998). These domains have a high concentration of basic residues similar to that found in NLSs of other proteins and despite being separated by a 40 residue sequence, they contain all the information required for efficient transport, since peptides containing the fusion of both NLS are fully functional. However a single domain is not sufficient and nuclear import relies on active ATP-dependent transport, possibly mediated by cytoplasmic factors commonly used by other nuclear proteins. NLS2 has also been implicated in specific functions for different HMGNs. HMGN1 NLS2 has been reported to be necessary and sufficient to regulate phosphorylation of H3 (Ueda et al., 2006). This study remarkably proved that HMGN

nucleosome binding itself is not sufficient to elicit HMGN functions, but other sequences within NLS and RD are involved. NLSs are also targeted by modifying enzymes. Acetylation has been observed at Lys 2 and 4 in the NLS1 and Lys 54, 58 and 60 at or near the NLS2 (Bergel et al., 2000). Acetylation of Lys 2 has been reported *in vitro* and *in vivo*, in the presence and absence of butyrate (a deacetylase inhibitor, frequently used to promote extensive acetylation), and is among the acetylation marks identified in HMGNs that are supported by the most solid evidence (Bergel et al., 2000; Herrera et al., 1999; Sterner et al., 1981). PCAF has been suggested to be responsible for Lys 2 acetylation but since this enzyme only acetylates HMGN2 *in vitro*, alternative enzymes must be involved in HMGN1 acetylation at NLS1 (Sterner et al., 1981). Phosphorylation of NLS residues does not occur. However, HMGN1 possesses a Ser at position 6, so immediately flanking NLS1, whose phosphorylation is part of the nucleosomal response induced by IE gene activation and a key residue in HMGN1 signalling (Thomson et al., 1999).

1.3.1.2 HMGN1 N-terminal and nuclear signalling

The serine residue at position 6 is characteristic of mammalian HMGN1 and it is the first phosphorylation mark on HMGN1 associated with nuclear signalling (Barratt et al., 1994; Soloaga et al., 2003; Thomson et al., 1999). HMGN1Ser6 is flanked by the basic residues of NLS1 forming a sequence targeted by several protein kinases, including PKC, PKA, RSK2, CK2, cGK (Cooper et al., 1982; Jiang and Wang, 2006; Prymakowska-Bosak et al., 2001; Walton et al., 1984). However only MSK1/2 proved to be conclusively responsible for MAP kinase-mediated HMGN1 phosphorylation *in vivo* at this position (Soloaga et al., 2003). Many different stimuli can induce phosphorylation of Ser6, including TSH (Thyroid-Stimulating Hormone), EGF, anisomycin, TPA, UV-light, Okadaic acid, but all evidence indicates completely dependency on activation of ERK or p38 MAP kinases (Barratt et al.,

1994; Cooper et al., 1982; Louie et al., 2000; Soloaga et al., 2003; Thomson et al., 1999, Figures 14, 15, 17). MSKs also phosphorylate Ser10 and Ser28 on histone H3. Despite the fact that both radiolabelling studies and phospho-specific antibodies independently confirmed Ser6 site as a major site of inducible phosphorylation, the lack of understanding of the function of HMGN1 N-terminal favoured the focusing of attention on phosphorylation occurring at the nucleosomal binding domain. Regarding the issue of which sites of HMGN1, NBD Ser20-24 or N-terminal Ser6, is more involved in responses to external stimuli, two major points require clarification. First, in order to be considered a stimulus-dependent mark, evidence must be collected on which stimulus triggers the modification and via which kinase pathway. While this is well characterized for HMGN1Ser6ph, all the work done on Ser20-24ph is based on the assumption that all phosphorylations are uniquely added by one kinase. Although *in vitro* experiments, with excess of kinase and protein substrate, revealed several enzymes able to deposit these marks, MSKs are presented as responsible for all HMGN1 phosphorylation during gene induction (Barratt et al., 1994; Lim et al., 2004; Postnikov and Bustin, 2010; Prymakowska-Bosak et al., 2001; Soloaga et al., 2003; Thomson et al., 1999). Secondly, it is expected from a signalling mark, especially if linked with rapid and transient activation of genes, to be following a curve of activation and deactivation, possibly within the time frame of the reported biological function. Data on HMGN1Ser6ph indicate clearly the rapid phosphorylation of Ser6 in response to various stimuli while very little is known on the dynamics of Ser20-24 phosphorylation (Lim et al., 2004 supp. data; Soloaga et al., 2003; Figures 14, 15, 17). However, it is remarkable that HMGN1 phosphorylation on Ser6 is very stable, showing signal even hours after the turnoff of the stimuli induced IE genes (Lim et al., 2004 supp. data; Soloaga et al., 2003). Finally, the role of HMGN1 and H3 phosphorylation in mitosis has often been regarded as a parallel to the stimulus-induced phosphorylation, due to the assumption that similar mechanisms might be used by the cell for different purposes.

However, detailed data on HMGN1 phosphorylation at other sites is lacking. Functionally, only NBD-phosphorylation has been examined even if HMGN1Ser6ph levels also increase during mitosis and the most abundant phosphopeptide observed is mono-phosphorylated-HMGN1, followed by the diphosphorylated-HMGN1 (Cherukuri et al., 2008; Prymakowska-Bosak et al., 2001, 2002). It is interesting to note that while most early work on HMGN functions mostly focussed on characterizing its chromatin-binding properties, more recent work has established that many of the protein functions rely on other domains of the proteins, such as the previously described NLSs and the RD at the C-terminal (Ueda et al., 2006). It is known that concomitantly to HMGN1Ser6 phosphorylation, H3 tails are phosphorylated and IE genes are activated. At present, unanswered questions remain about whether Ser6, Ser20 and Ser24 phosphorylation all occur during IE gene activation and if they might be on the same HMGN1 molecule.

Lim et al (2004) observed that upon anisomycin treatment HMGN1 and H3 are sequentially phosphorylated. HMGN1Ser20 and/or Ser24 appear to be phosphorylated first, followed by HMGN1Ser6 and then H3Ser10. HMGN1 phosphorylated in the NBD has a reduced affinity for the nucleosome, a faster rate of exchange and shorter time of residence (Phair et al., 2004; Prymakowska-Bosak et al., 2001). In this model, shown schematically in Figure 8 (from Pogna et al., 2010), after 10–15 min of anisomycin treatment, activated MSKs phosphorylate HMGN1Ser20 and Ser24, loosening the interaction between HMGN1 and nucleosomes. This facilitates MSKs access to the H3 tails and within 30–60 min allows phosphorylation of H3Ser10 and Ser28, promoting IE gene transcription. The role of Ser6 appears to be marginal while the NLS2 of HMGN1 is absolutely necessary for the HMGN1-dependent inhibition of H3 phosphorylation by MSK1 (Ueda 2006). This model of signal transduction to chromatin lacks solid proofs and provides no explanation of the targeting of HMGN1Ser6 but is at present the only one attempting to reconcile all collected evidence (Pogna et al., 2010).

1.3.1.3 HMGN1 C-terminal and chromatin remodelling

The HMGN1 regulatory domain (RD) is the C-terminal portion of the protein. It represents approximately the last 23 residues of the protein. While most of HMGN domains are conserved among the different members of the family, the C-terminal is very variable and distinct for each member. It has been suggested that this peculiarity might be responsible for specific functions of different HMGNs and the fact the C-terminal sequence is mostly invariant throughout mammalian HMGN1s suggest the existence of conservative pressure to preserve its function through evolution (Rochman et al., 2011). The HMGN1-RD was originally defined as chromatin unfolding domain (CHUD), due to the fact that this negatively charged domain was necessary to promote chromatin relaxation and was responsible for enhancing transcription by both RNA pol II and pol III *in vitro* (Trieschmann et al., 1995). It was suggested that part of this function is mediated by its interaction with H3 tails and interference with histone H1 binding (Ding et al., 1997; Trieschmann et al., 1998). Most recent studies reported additional properties of the HMGN1 C-terminus. It has been discovered that binding to mitotic chromatin occurs via the C-terminus despite displacement from chromatin of the majority of phosphorylated HMGN1 (Cherukuri et al., 2008; Pallier et al., 2003). Moreover when HMGNs are associated with nucleosomes, both HMGN1 and HMGN2 are capable of causing enhanced acetylation of H3K14 through their RD domain (Ueda et al., 2006). Interestingly, in this effect HMGN2 shows a much stronger potential (3.5 fold increase) than HMGN1 (2-fold increase). The difference in the amino acid sequence of the RD domain of HMGN1 and HMGN2 seems to be the determinant of HMGN2 superiority, most likely due to the specific interaction of HMGN2 with acetylase PCAF (Ueda et al., 2006). HMGN1 also seems to modulate modifications on other residues of H3 tails (H3K9me, H3K9ac) and other histone such as H2A (Ser1ph) (Lim et al., 2004, 2005; Postnikov et al.,

2006; Ueda et al., 2006). However, only in the case of H3K14ac has a rigorous difference in HMGN1/2 effects been established. More solid evidence of the RD specificity comes from recent a study on transcriptional profiles of HMGN knockouts and overexpression mutants. Alteration of the abundance of different HMGNs caused the impaired transcription of genes specific for each HMGN (Rochman et al., 2011).

Phosphorylation within the CHUD/RD was first reported in radiolabelling studies *in vivo* and *in vitro*, followed by peptide mapping and sequence analysis of peptides deriving from HMGN1 (Walton and Gill, 1983; Walton et al., 1985). Targeted sites (Ser85, Ser88, and Ser98) are not conserved in HMGN2 and are seen in mammals, Xenopus and chicken, but not in the zebrafish sequence (Figure 6). Casein kinase 2 (CK2) is capable of phosphorylating serines 88 and 98 within this region *in vitro*, but the identity of the residue that is phosphorylated *in vivo* is unknown. The Ser88 flanking motif is a much stronger target site for CK2 than Ser98, and the results of the kinase assay confirmed the difference (Walton et al., 1985). *In vitro* studies show that CK2 may also phosphorylate Ser85 (Jiang and Wang, 2006). However Ser85 is followed by proline and is thus a better target for proline-directed kinases such as ERKs, JNK/ SAPK and p38 MAP kinases than for CK2. Zou et al (2004) reported phosphorylation on Ser85, Ser88, and Ser98 in MCF-7 cells and suggested that the lack of results at these sites in previous mass spectrometric analysis may be due to low levels of phosphorylation below the sensitivity of the technique (Louie et al., 2000).

The recent observation that HMGN1 C-terminal can mediate the association of the protein with chromatin regardless of a functional NBD in mitosis may suggest a function for C-terminal phosphorylation in reducing chromatin binding or in cytoplasmic retention of phosphorylated HMGN1 (Cherukuri et al., 2008; Prymakowska-Bosak et al., 2001). No phosphorylation at the C-terminal has been reported as stimulus inducible.

1.3.2 HMGN1 expression in tissues and during development

HMGN proteins are unique to higher eukaryotes (West, 2004). The five different HMGNs show distinct patterns of expression in tissues and HMGN knockout mutants in some cases present specific phenotypes. However, HMGN knockout phenotypes are not severe and a high degree of compensation is observed (Birger et al., 2003; Furusawa and Cherukuri, 2010; Hock et al., 2007). The compensation occurring in the knockout mutants and the fact that the most conserved functional domain among HMGNs is the NBD suggests that most of the functions of these proteins rely on their capability to bind to nucleosomes. Details of the phenotype of the *Hmgn1−/−* mice will be discussed in the following chapter.

The developmental roles of HMGN proteins are mainly based on *in vitro* cellular differentiation of mouse, chicken, or human cells. Although HMGN1 and HMGN2 are ubiquitously expressed in adult and embryonic tissues, they are most strongly expressed during embryogenesis and in continuously renewing cells (Furusawa and Cherukuri, 2010). HMGN proteins are progressively down-regulated during embryogenesis and differentiation, HMGN1 and HMGN2 expression is reduced during myogenesis, erythrogenesis, chondrogenesis and kidney development (Crippa et al., 1991; Furusawa et al., 2006; Hock et al., 2007; Lehtonen and Lehtonen, 2001; Pash et al., 1990, 1993). HMGN1/2 genes show specific expression patterns in the hair follicle. HMGN1 is expressed in the stem cell of hair follicles, whereas both HMGN1 and HMGN2 are expressed in the less differentiated outer root sheath cells and basal layer cells, but not in the terminally differentiated inner root sheath cells, hair shaft, or suprabasal layer (Furusawa et al., 2009). HMGN1 and HMGN2 are also involved in blastocyst development. HMGN1/2 proteins are detected throughout oogenesis, and their embryonic transcripts accumulate beyond the two-cell stage. Interestingly injection of HMGN1/2 antisense oligonucleotides into mouse oocytes delays cell cleavage and the

onset of the blastocyst stage (Mohamed et al., 2001). Studies from bovine embryo early-stage development reported downregulation of HMGN1/2 transcripts after fertilization and they almost disappear by the 8-cell stage (Vigneault et al., 2004). Moreover, embryos that are incapable of downregulating HMGN2 failed to develop into blastocysts, possibly due to the altered chromatin remodelling (Bastos et al., 2008). In Xenopus, Hmgn1 and Hmgn2 expression in the three germinal layers starts after mid-blastula transition (MBT) and continues throughout the neurula and tadpole embryos (Körner et al., 2003). HMGN1 and HMGN2 do not affect embryonic stages prior to the MBT but after the MBT, overexpression of either HMGN1 or HMGN2 causes imperfectly closed blastopores, distorted body axis, and abnormal head structure in these embryos (Körner et al., 2003). In humans, due to HMGN1 gene location on chromosome 21, a possible role for HMGN1 in the etiology of Down's syndrome has been suggested (Potier et al., 2006). Indeed, HMGN1 is overexpressed in Down's syndrome patients and also in the Ts1Cje mouse, which is used as a mouse model for Down's syndrome (Pash et al., 1990b; Potier et al., 2006). The collected evidence from different organisms suggest that HMGN's downregulation could be part of the changes in chromatin organization and in the cellular transcription potential that occurs during differentiation. Despite the fact that HMGN1 and HMGN2 showed similar expression patterns and molecular structure, HMGN2 does not always compensate for the loss of HMGN1, and more interestingly, experiments in progress suggest that loss of HMGN2 leads to embryonic lethality (S. Cherukuri, unpublished data, Furusawa and Cherukuri, 2010). Regarding the competition for nucleosome-binding between HMGNs and H1, it is notable that in Xenopus animal cap assays, depletion of histone H1, overexpression of either HMGN1 or HMGN2 all have a similar effect in prolonging mesoderm induction by activin, (Catez et al., 2004; Körner et al., 2003; Steinbach et al., 1997).

The studies on HMGN1 post-translational modifications and on its role in chromatin modifications and compaction have been mostly done in primary mouse embryonic fibroblasts (MEF), C3H 10T1/2 fibroblasts, human HeLa and MCF7 cancer cells model systems. The level of HMGN proteins in the nuclei of these mammalian cells is generally considered limiting and the general consensus in the HMGN field is that HMGN/nucleosome ratio varies from 0.01 to 0.1 (Johns, 1982; Pogna et al., 2010; Shirakawa et al., 2000a). However there are report of HMGN1/nucleosome ratios of 1:10, 0.5:1, and 1.5:1 in specific tissues (Boumba et al., 2004; G.H. Goodwin and E.W. Johns, 1978; Kuehl et al., 1984).

1.3.2.1 HMGN1 depletion in the mouse model

Depletion of HMGN1 in mice produces a viable mutant. *Hmgn1*−/− mice appear normal but are subfertile and have slight defects in the development of corneal epithelium. Tissue analysis discovered increased sensitivity to various stress conditions, such as exposure to UV light or ionizing irradiation, impaired gene expression and histone post translational modification (Birger et al., 2003, 2005, 2006; Lim et al., 2004, 2005; West, 2004). Abnormalities in the development of corneal epithelium could be linked to the high expression of HMGN1 in the basal layer of this tissue where it colocalizes with p63, a protein involved in the regulation of epithelial differentiation (Birger et al., 2006). *Hmgn1*−/− mice are almost twice more susceptible to tumors than wild-type mice and *Hmgn1*−/− cells have an increased tumorigenic potential, measured both by colony formation in soft agar and generation of tumors in nude mice (Birger et al., 2005). *Hmgn1*−/− cells are hypersensitive to UV and IR, possibly due to the increased compaction of damaged chromatin and reduced accessibility of the repair machinery (Birger et al., 2003). Direct recruitment of HMGN1 by Cockayne syndrome protein A to the RNA polymerase stalled at the UV-damaged sites has been reported. *Hmgn1*−/− cells might then have reduced transcription-coupled repair

(Fousteri et al., 2006). In the other hand, hypersensitivity to ionizing radiation has been linked to the impaired ability of *Hmgn1*−/− cells to activate the G2-M checkpoint. Loss of HMGN1 reduces IR-induced ATM autophosphorylation and the activation of several ATM targets leading to increased genomic instability (Birger et al., 2005, Kim et al., 2009). Additionally loss of HMGN1 affects chromatin modification. Knockout of HMGN1 influences global levels of histone H3 acetylation on lysine 14 (Lim et al., 2005). *Hmgn1* −/− cells show reduced H3K14ac compared to wild type and interestingly using the deacetylase inhibitor TSA it is possible to mostly restore the kinetics of H3K14ac (Birger et al., 2003). Use of TSA to compensate HMGN1 loss in H3K14ac regulation rescues the IR sensitivity of *Hmgn1* −/− mutants, proving the importance of this histone modification in DNA repair and suggesting a possible link between HMGN1, histones and ATM (Kim et al., 2009). It has been proposed that HMGN1 influences H3K14ac by shielding H3 tails and preventing the action of HDACs (Lim et al., 2005). Consistent with previous reports, the global acetylation and phospho-acetylation of H3 during anisomycin induction of IE genes is not altered in *Hmgn1*−/− cells (Lim et al., 2005; Soloaga et al., 2003; Thomson et al., 2001). Evidence from Clayton et al. (2000) indicated the presence of lysine 9 acetylation on phosphoacetylated H3 at IE genes (*c-jun* and *c-fos*). Since H3K9 and K14 acetylation can coexist on phosphoacetylated histones it is possible that HMGN1 affect also H3K9ac but this remains unresolved (Clayton et al., 2000; Macdonald et al., 2005). In *Hmgn1*−/− cells, loss of HMGN1 elevates the steady-state levels and enhances the rate of stress-induced phosphorylation of histone H3 on serine 10 (Lim et al., 2004). Anisomycin induction of IE genes has been analyzed in these cells. Expression of *fos B* is increased while that of jun D appears to be reduced; *c-fos* remained unchanged (Lim et al., 2004, 2005). *c-jun* and *jun B* show impaired activation at early time points but expression overall is unchanged especially at time points when the activation is maximal (Lim et al., 2004). These results conflict with previous report of IE gene expression

upon anisomycin stimulation where no expression of *fos B* has been observed (Hazzalin and Mahadevan, 2002). It is possible that the reason why *c-fos* and *c-jun* expression remains mostly unchanged in *Hmgn1*−/− cells is due to the strong inducing effects of anisomycin, whereas *jun D* activation is weaker and possibly more sensitive to the removal of HMGN1. However *jun B* activation by anisomycin is even weaker than *jun D* and HMGN1 depletion does not affect it (Hazzalin and Mahadevan, 2002; Lim et al., 2004). Thus, it seems that the data available at the moment are too limited to derive a final judgment. It would be interesting to test the effect of the activation of other MAPK cascades than the JNK/p38 pathway in *Hmgn1*−/− cells.

The altered phenotype of chromatin modification in *Hmgn1*− /− cells can be rescued by introduction of wild type HMGN1 but not by HMGN1 mutants possessing a dysfunctional NBD. At present no data is available on rescue of this phenotype by HMGN2, therefore it is not clear if this phenotype is caused by more chromatin compaction or if other associations are impaired. In any case it is likely that the functional redundancy between HMGN1and other HMGNs are mitigating the loss of HMGN1 during embryogenesis. The alteration of the expression of other HMG proteins such as HMGN2, HMGAs or HMGBs results in more severe phenotype during mouse development (S. Cherukuri, unpublished data, Furusawa and Cherukuri, 2010; Hock et al., 2007).

1.3.3 Dynamic association of HMGN1 with chromatin

One of the major challenges of modern chromatin research is the analysis of highly dynamic biological processes. The development of several fluorescent imaging techniques such as fluorescence recovery after photobleaching (FRAP) and fluorescence loss in photobleaching (FLIP) has allowed collection of detailed information on the dynamics of binding of HMGNs

and several chromatin-associated proteins (Phair and Misteli, 2000; Phair et al., 2004). HMGN1/2 interaction with nucleosomal histone *in vivo* has an average time of residence of less than 30 sec and a diffusion speed that allows travel through the nucleus in less than 1 minute (Phair and Misteli, 2000). HMGN1 seems to be divided into a fast fraction (≈ 20%) with residence times on nucleosomes of about 4 sec, and a slow fraction (≈ 80%) with residence times of about 25 sec. Considering both populations, at any given moment an overall 99.5% of the HMGN1 molecules are associated with chromatin (Phair et al., 2004). All the evidence collected so far indicates that HMGN2 binding has similar dynamics (Phair and Misteli, 2000). Disruption of the NBD alters the balance between the slow and fast fraction of HMGN1 and 96.5 % of the protein is resident on nucleosome for about 3 sec (Phair et al., 2004). These dysfunctional HMGN1 molecules still associate with chromatin, suggesting that other domains of the protein maintain the nuclear localization. Phair et al., (2004) observed that almost all HMGN1 molecules are localized in the nucleus, in contrast with other reports that suggested the presence of about 20% of HMGN1 molecules dispersed through the cytoplasm in interphase cells (Louie et al., 2000).

To compare with other chromatin factors, H1 has a residence time of 180 sec and histone H2B has a residence time of more than 1h. Other dynamically chromatin-associated factors such as transcription factors Jun and Fos, have respectively 5.9 and 2.4 sec residence times for their fast fractions (19.5% and 32.4%) and 27.3 and 14.6 sec for their slow fractions (75.2% and 54.2%), indicating that several molecules are exchanging at high speed and frequency with chromatin (Phair et al., 2004). If compared with the other chromatin-associated proteins considered by these studies, the relatively long mean residence time (24.8 s) of almost 80% of HMGN1 molecules, makes HMGN1 one of the more stably-associated proteins within this group (Phair 2004).

1.3.3.1 HMGN1 genomic profiling

Despite the fact that HMGN molecules have such a dynamic interaction with chromatin, several attempts have been made to address their localization throughout the genome and at specific loci. It has been known for a decade that both HMGN1/2 positioning within chromatin is independent of underlying DNA sequence (Shirakawa et al., 2000). ChIP assays reported that the presence of HMGN2 at active transcribed genes is only 1.5 times higher than in silent chromatin regions. When random amplified polymorphic DNA (RAPD) was used to screen the nucleosomes obtained from the entire genome, HMGN1/2 proteins were found preferentially at DNA regions depleted of CG and enriched in AT and TA but no specific sequence nor regions that serve as HMGN1/2-binding sites where identified (Shirakawa et al., 2000a; Williams et al., 1990). Recently Cuddapah et al. (2011) have published results of genomic profiling of HMGN1 in CD4 cells, obtained by ChIP-Seq on cross-linked chromatin immunopurified with an antibody recognizing both modified and un-modified HMGN1. Binding sites for HMGN1 were identified using the spatial clustering approach for the identification of ChIP enriched regions (SICER) algorithm to identify domains enriched for HMGN1 ChIP-Seq signals which were considered to be binding sites. This procedure produces 146,127 HMGN1 binding regions in the human genome of CD4 cells (Cuddapah et al., 2011; Zang et al., 2009). Comparison of the position of these HMGN1-enriched regions with genome-wide distribution of regulatory chromatin marks shows that HMGN1 preferentially localizes to DNase I hypersensitive (HS) sites, promoters, functional enhancers, and transcription factor binding sites. The extensive colocalization of DNase I HS sites and HMGN1 binding sites seems to reflect the observation that lack of HMGN1 caused increased rate of digestion by micrococcal nuclease at the *HSP70* promoter locus in *Hmgn1*-/- cells (Belova et al., 2008). It is known that sensitivity to digestion is a sign of chromatin

decompaction and accessibility often correlated with gene transcription and regulatory regions (Edmunds et al., 2008).

According to Cuddapah et al. (2011) HMGN1 ChIP is enriched with nucleosomes surrounding binding sites of the architectural protein CTCF and zinc-finger transcription factor YY1 that interacts with both CTCF and HMGB2. This suggests that HMGN1 plays a role in maintaining a nucleosome occupancy profile that enhances the accessibility of regulatory factors to their targets. The involvement of HMGN1 chromatin remodelling activity of ATP-dependent enzymes is controversial but these findings support a role for HMGN1 in regulating genome-wide nucleosome positioning (Hill et al., 2005; Rattner et al., 2009). Previous results reported that in CD4 and HeLa cells, specific CTCF binding sites are occluded by nucleosomes and devoid of CTCF (Cuddapah et al., 2009). Interestingly HMGN1 in CD4 cells is absent at these sites, suggesting that HMGN1 specifically influences the maintenance of nucleosomes at regions were CTCF is present. Further investigation is necessary to test whether HMGN1 directly influences specific chromatin remodelling activity such as the INO80 complex, prerequisite for YY1 binding to its target sites, or ATP-dependent enzymes (Cai et al., 2007; Rattner et al., 2009). Loss of the canonical -1 nucleosome upstream of the TSS and precise positioning of other surrounding nucleosomes is a common feature of promoters of actively transcribed genes (Schones et al., 2008). It has been proposed that such precision might be enforced by one or two nucleosomes being strictly regulated with the neighbouring nucleosomes adjusting their position relative to these nucleosomes (Mavrich et al., 2008). Thus, the hypothetical role of HMGN1 in anchoring nucleosomes at the boundaries of the CTCF and YY1 binding sites may also explain the significant enrichment of HMGN1 in nucleosome-depleted regions in active promoters. Nucleosome-depleted regions are of variable length in different promoters, and this seems to reflect the broader distribution of HMGN1 binding at promoters than at CTCF and YY1

binding sites. Overall, this profiling shows enrichment of HMGN1 at active promoters and only very low levels of HMGN1 binding at silent gene promoters (Cuddapah et al., 2011). In addition to evidence collected by ChIP-Seq, previous work on *Hmgn1* -/- cells and anisomycin stimulation reported that levels of H3K14ac, and to some extent also phosphoacetylation of H3 are higher at both the promoter and transcribed portion of *jun D* in *wild type* cells (Lim et al., 2005). This might suggest that presence of HMGN1 at these loci is required for proper deposition of histone marks. Anisomycin-induced depletion of HMGN1 was also reported at *fos B* chromatin, while the levels of *c-jun* and *c-fos* sequences in HMGN1 ChIP show little variation through time after stimulation (Lim et al., 2004).

In contrast with the results of the genome wide analysis, targeted investigations at several genes presents a more unclear picture. HMGN1 distribution across the *Glyt1* (Glycine transporter 1) locus is homogenous in hepatoma line Hepa-1 (Barkess et al., 2012). There is no notable enrichment of HMGN1 at any genomic feature and no correlation between HMGN binding and presence of histone modifications H3K4me3 and H3K9ac that mark the promoter and putative enhancer. HMGN1 distribution throughout various genomic regions is also even at the HSP70 locus in MEF (Belova et al., 2008). However in limb bud cells (which contain both prechondrogenic cells and chondrocytes) three regions in the *Sox9* gene were identified as enriched in HMGN1, a region 2 kb upstream of the promoter, exon 2, and exon 3 (Furusawa et al., 2006). *Sox9* expression in these cells occurs at later developmental stages when HMGN1 is significantly down-regulated. The three regions do not contain any common motifs but the HMGN1-containing gene is more susceptible to DNase I digestion as previously mentioned for HSP70 (Belova et al., 2008; Furusawa et al., 2006). In MCF-7 cells, the role of HMGN1 in the induction of estrogen regulated genes, including *TFF1* and *FOS*, is another clear example of how HMGN1 distribution can appear targeted or evenly distributed depending on the gene. In this system the presence of HMGN1 is evidently different at TFF1

promoter and exon 1. However at the other analysed genes, HMGN1 levels are homogenous both across different genes and at different loci. Moreover, among the four gene promoters investigated, only TFF1 promoter shows a significant variation of HMGN1 presence during estrogen stimulation while the other gene promoters (including FOS) have mostly invariable profiles (Zhu and Hansen, 2007). Nevertheless HMGN1 specifically interacts with estrogen receptor α, significantly inhibiting the estrogen-driven activation of the *TFF1* and *FOS* gene (Zhu and Hansen, 2007). Interestingly, other HMGN proteins show variety of genome binding profiles. HMGN2, HMGN3a and HMGN3b are bound evenly across the *Glyt1* gene locus but HMGN3 is enriched in a 3 kb region at the promoter of *Glut2* (glucose transporter 2), and HMGN2 is enriched at the *CISH* (cytokine-inducible SH2-containing protein) gene proximal promoter region where it mediates prolactin receptor (PRLr) and transcription factor Stat5a promoter binding (Barkess et al., 2012; Fiorillo et al., 2011; Ueda et al., 2009).

In summary, data collected to date suggest that HMGN1 and other HMGN proteins can show enrichment at specific loci and directly interact with several transcription factors that regulate those loci. However it is still unclear whether HMGN1 is involved in the establishment or maintenance of these regulatory sites or if transcription factors and/or chromatin structural elements are driving the enhanced HMGN recruitment occasionally observed. It also remains unclear to which extent HMGN1 localization affects the levels of histone modifications, or if on top of a genome-wide function of anchoring regulatory nucleosomes, HMGN1 plays a specific role in a subset of the anisomycin-induced IE genes (Lim et al., 2005).

1.3.3.2 Dynamic interactions of HMGN1 with histone H1 and methyl CpG-binding protein 2 (MeCP2)

Additional insight on the dynamic association of HMGN1 with chromatin comes from studies on the interaction between H1 and HMG proteins and from a recent report that observed

HMGN1-dependent regulation of methyl CpG-binding protein 2 (*MeCP2*), a DNA-binding protein known to affect neurological functions including autism spectrum disorders (Abuhatzira et al., 2011; Catez et al., 2002, 2004).

Histone H1 is known for its function in chromatin compaction and stabilization while the HMGN proteins are supposed to decrease the compactness of the chromatin fiber and enhance the accessibility of chromatin targets to regulatory factors. It also evident that while H1 has a relatively continuous interaction with chromatin, HMGNs are much more dynamic (Phair et al., 2004). To prove direct competition at binding sites, photobleaching experiments were conducted in living cells (Catez et al., 2002, 2004). In an average mouse cell nucleus, there are 10^7 nucleosomes, an equal number of H1 molecules and about 10^6 HMGB, 10^5 HMGN, and 10^4 HMGA molecules (Bustin, 1999; van Holde, 1988; Johns, 1982). The amount of H1 varies little among cells but the various HMG proteins fluctuate significantly, especially during development (Bianchi and Agresti, 2005; Furusawa and Cherukuri, 2010; Hock et al., 2007). Upon injecting 4 x10^5 to 3.6 x10^6 HMG molecules into the nucleus, corresponding to 4-35% of the potential nucleosomal and H1 binding targets, detectable alteration in the chromatin binding of H1 is observed (Catez et al., 2004). Wild-type HMGN1, but not HMGN1 point mutants that do not bind to nucleosomes or a deletion mutant that lacks the C-terminal region, inhibits the binding of H1 to nucleosomes and increases their mobility and rate of recovery from photobleaching. Injected HMGN1 colocalizes with the stably expressed $H1^0$–GFP within 15 min after treatment. Competition was observed at both euchromatin and heterochromatin regions. Control and injected cells reach a similar recovery plateau from photobleaching suggesting that HMGN1 enhances the mobility of the rapidly exchanging H1 to a larger extent than the less mobile fraction (Catez et al., 2002). Further, other HMG proteins weaken H1 binding synergistically and without competing among each other, indicating specific targeting of distinct H1 binding sites. HMGs affect H1 binding throughout

the entire nucleus, regardless of the state of chromatin compaction and histone modification (Catez et al., 2004). Considering the difference in chromatin association dynamics between H1 and most HMGs, it is likely that HMGs moving very rapidly throughout the nucleus can access the nucleosomal sites temporarily vacated by H1. These observations reinforce evidence of genome wide functions for HMGs and specifically for HMGN1, consistent with the observations that HMGN1 proteins bind chromatin without DNA sequence specificity and that HMGNs are involved in a wide range of nuclear activities (Gerlitz, 2010; Postnikov and Bustin, 2010; Shirakawa et al., 2000a; Zhu and Hansen, 2010). HMG competition with H1 for chromatin targets can potentially affect H1 function in up- and down- regulation of gene expression, promotion of chromatin compaction, stabilization of nucleosomal structures and reduction of nucleosomal access (Alami et al., 2003; Brown, 2003; Herrera et al., 2000b; Shen and Gorovsky, 1996; Vicent et al., 2002a). It is important to stress that the temporary vacancy caused by competition between HMGs and H1 does not necessarily promote gene activation. HMG binding increases the ability of both positive and negative regulatory factors to access their nucleosomal targets. Thus competition between H1 and HMGN1 may be a pivotal mechanism of optimized cellular response to external or internal stimuli.

The dynamic and global nature of H1-HMGN1 interaction within chromatin indicates a role for HMGN1 in genome wide regulation. Nevertheless, studies from the IE gene transcription profile in the *Hmgn1* -/- mutant suggest distinct regulation of specific genes. Another example of specific influence of HMGN1 on transcription of a distinct gene has been reported in human and mouse brain cells. Impaired expression of HMGN1 affects the transcription profile of methyl CpG-binding protein 2 (*MeCP2*), a DNA-binding protein known to regulate neurological functions (Abuhatzira et al., 2011). *MeCP2* is a gene highly expressed in brain and its mis-expression or mutation are strongly linked to behavioural abnormalities in human and mouse, including autism spectrum disorders (Bird, 2008; Gemelli et al., 2006; Guy et al.,

2001; Moretti et al., 2005; Skene et al., 2010). Quantitative PCR and western blot analyses of cell lines and brain tissues from mice indicate that HMGN1 is a negative regulator of *MeCP2* expression. HMGN1 appears to affect accessibility to the *MeCP2* promoter by downregulating H3K9ac and promoting H3K9me2 (Abuhatzira et al., 2011). Behaviour analyses noted abnormalities in activity, anxiety and social deficits in mice with altered HMGN1 levels. Overexpression of HMGN1 leads to hyperactivity, reduced anxiety, and abnormal social interest, some of the behavioural phenotypes detected in mouse models for Down syndrome (Coussons-Read and Crnic, 1996). *Vice versa*, *Hmgn1-/-* mice display hypoactivity and abnormal social behaviour and resemble neurological abnormalities seen in models of autistic spectrum disorders (Silverman et al., 2010). Despite the lack of evidence that HMGN1 influences transcription of other genes regulating neuronal physiology, the discovery that HMGN1 depletion or over-expression elicits several neurological abnormalities, not solely due to its effect on MeCP2, suggests an important role for HMGN1 in neuronal function. The widespread effects of HMGN1 on gene expression, with most of the changes relatively small, and its association with regulatory sites throughout chromatin, impedes unequivocal identification of all the genes influenced by HMGN1 that may be responsible for the phenotypes observed (Cuddapah et al., 2011; Rochman et al., 2011). It is likely that with more detailed genome profiling in the future, candidate genes with a more distinct HMGN1 influence on their transcription will emerge. At the present time it is not possible to indicate whether HMGN1 has evolved to specifically regulate these genes or if chromatin features at these genes make them more sensitive to the presence of HMGN1.

1.4 Cross-interaction between architectural and nucleosomal proteins

The chromatin environments at gene loci are defined by complex and dynamic networks of architectural and nucleosomal proteins. The general concepts of histone modification and genome-wide distribution of several architectural proteins have been discussed in the previous chapters. However the overall picture derived from these analyses is still very undefined and controversial. It is helpful therefore to focus on those gene systems that are well characterized. IE gene activation has been analyzed deeply to reveal how histone modification and possibly HMGN1 might function in their regulation.

1.4.1 Timing and distribution of PTM of histones at *c-fos* and *c-jun*

A clear understanding of the function of histone modifications present across inducible genes requires a quantitative and comparative analysis of their dynamics during gene activation and the identification of the enzymes responsible. I previously discussed how histone modifications, which include acetylation, phosphorylation, methylation and ubiquitination, mark and define specific regions of the genome and specific domains of gene templates (reviewed in Berger, 2007; Bernstein et al., 2007; Kouzarides, 2007; Li et al., 2007; Shilatifard, 2006). However, histone PTMs do not only play key roles in global genome arrangement into heterochromatin and euchromatin but also in gene regulation at very specific sequences and well-defined stretches of nucleosomes.

Very little investigation had been carried out on the changes in histone modifications and position induced by gene activation at specific loci. In 2008, Edmunds et al, produced high-

resolution comparative maps of the distribution and dynamics of H3K4me3, H3K36me3, H3K79me2 and H3K9ac across *c-fos* and *c-jun* upon gene induction in murine fibroblasts. In addition to the previously discussed histone H3 and HMGN1 phosphorylation, dynamic acetylation occurs at *c-fos* and *c-jun* upon gene induction, creating phosphoacetylated H3 (Barratt et al., 1994; Cheung et al., 2000; Clayton and Mahadevan, 2003; Clayton et al., 2000; Thomson et al., 1999, 2001). These two modification, although they colocalize, appear to be independently regulated by distinct mechanisms. Unfortunately attempts to perform detailed and quantitative ChIP analysis of the distribution of these phosphorylated-H3 and phosphoacetylated H3 marks have been unsatisfactory due to very low recovery (less than 5% of input) (Lee and Mahadevan, 2009; Thomson et al., 2001). Phosphoacetylated histone H3 has been reported at the start site of IE genes, -733 to -572 for *c-jun* and -131 to +93 for *c-fos*, (numbers are with respect to the transcription start sites; Clayton et al., 2000). However, it was possible to raise high quality antibodies to perform ChIP on H3K4me3, H3K36me3, H3K79me2 and H3K9ac histone modifications. This analysis supports a model of four distinct layers of histone modification across these inducible genes (Figure 9). The first layer is composed of pre-existing, possibly epigenetic, histone modifications present also in quiescent G0 cells. H3K4me3 across the start site and K79me2 and K36me3 in the coding region are detectable in unstimulated cells. The second layer of histone acetylation is dynamically deposited and removed by the action of HATS and HDACs. Dynamic acetylation occurs at all K4me3-modified H3 in the mouse nucleus (reviewed in Clayton et al., 2006; Hazzalin and Mahadevan, 2005). This second layer is also independent of stimulation. The third layer is composed of stimulation-dependent marks and includes H3 phosphorylation which is seen mostly across the start site. The last layer is dependent on gene activation and is associated with the passage of Pol II; these marks are sensitive to treatment with 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB), an RNA polymerase II inhibitor

and are restricted to coding regions of these genes. Pol II elongation enhances H3K4me3 and H3K36me3 levels, which peak respectively at the 5'end and 3'end of the coding region of *c-fos* and *c-jun* (Edmunds et al., 2008).

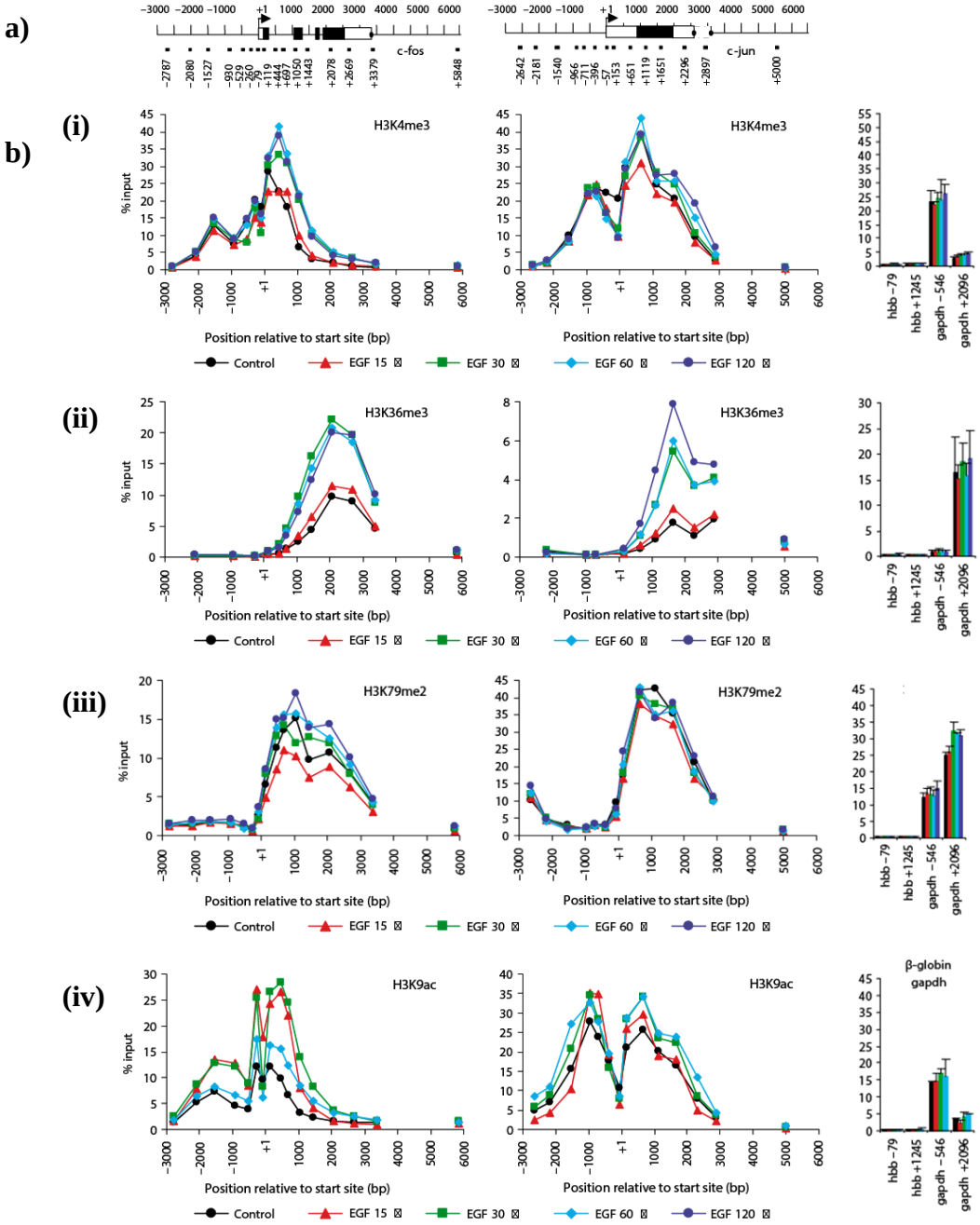
In summary, the full histone modification pattern is imposed by a combination of epigenetic, dynamic, kinase- and Pol II-dependent enzymes. ChIP with phosphoacetylated (S10phK9ac) or H3K4me3-specific antibodies and re-ChIP experiments showed that all three modifications can occur on the same nucleosomes (Hazzalin and Mahadevan, 2005). Because H3K4me3 was continuously detectable, whereas H3K9ac and H3S10ph were transient, it is most likely that H3K4me3 is the key pre-existing modification, which attracts both dynamic acetylation and stimulus-dependent H3Ser10 phosphorylation (reviewed in Clayton et al., 2006). Notably acetylation by itself does not correlate with transcription; inhibition of turnover of acetylation at these sites using TSA (deacetylases inhibitor) enhances histone acetylation at these loci but inhibits *c-fos* and *c-jun* induction (Edmunds et al., 2008).

The complex layering of different modifications at IE genes *c-fos* and *c-jun* is produced by the action of distinct enzymes specifically activated by gene induction or continuously operating regardless of stimuli. It is clear how phosphorylation arises on H3 tails *via* the action of MSKs, that p300/CBP mediates acetylation of all K4me3 histone H3 and siRNA knockout proved that Huntingtin-interacting protein HYPB/Setd2 is responsible for virtually all global and transcription dependent H3K36 trimethylation, but not H3K36-mono- or dimethylation (Crump et al., 2011; Edmunds et al., 2008). Interestingly, less complex eukaryotes, such a yeast, posses single methylation enzymes that are responsible for all H3K4 or K36 methylation, Set1 and Set2 respectively, suggesting that fine regulation of specific histone marks is an evolutionary acquired characteristic of higher eukaryotes (Briggs et al., 2001; Krogan et al., 2002; Strahl et al., 2002). The major caveat to the interpretation of these studies is the possibility of microheterogeneity. At some positions, different sets of histone

modifications, deriving from different cells within a population, are detectable at the same time. An obvious source of microheterogeneity is the transient dynamic nature of most histone modifications, for example, H3 acetylation described above. A second relates to the fact that some histone-modifying enzymes travel with elongating RNA Pol II, producing elongation-dependent variations that can be lost in the analysis of mixed population of asynchronous cells.

Figure 9. **EGF-stimulated distribution of acetyl- and methyl-histone H3 across *c-fos* and *c-jun*.**

a) Schematic representation of *c-fos* and *c-jun* showing positions analysed by real-time PCR. Exons are represented by boxes, unfilled for untranslated regions and filled for translated regions. Termination sites are shown as filled circles. **b)** Quiescent cells were unstimulated (control) or stimulated with EGF (50 ng/ml) for 15, 30, 60 or 120 min and formaldehyde cross-linked; then mononucleosomes were prepared for ChIP. Aliquots of chromatin were immunoprecipitated with **(i)** H3K4me3-, **(ii)** H3K36me3-, **(iii)** H3K79me2-, **(iv)** H3K9ac-specific antibodies. Recovered DNA sequences were quantified by real-time PCR. For each ChIP, primers spanning two regions of the inactive b-globin (*hbb*) and the constitutively expressed *gapdh* genes were also analysed (right-hand side graphs) for comparison.



adapted from (Edmunds et al., 2008)

1.4.2 Proposed role of HMGN1 in signalling to chromatin

The evidence discussed in the previous chapters describes a system in which many players involved in gene regulation have both genome wide and loci-specific functions, but with very different dynamics. Global functions that have a key role in marking specific genomic regions require a stable presence, such as histone methylation at promoters, enhancers and heterochromatin, or the presence of CTCF in loops and promoters. On the other hand, loci-specific functions are transient and dynamic, like histone acetylation and phosphorylation at IE genes, or involve the rapid and interchangeable association with nucleosomes of many chromatin activating proteins, such as HMGs, H1 and several transcription factors (TFs). The distinction between the two categories is subtle and only very few chromatin proteins or protein marks can be considered completely stable. Nevertheless, it is clear that the old vision of *"a stable on/off switch"* for transcriptional activation fails to describe the intense dynamics undergone by the several components of the switch. An alternative model proposed to explain the events occurring during IE gene activation, is the *dynamic coupled mechanism for a continuously variable gene switch"* (Hazzalin and Mahadevan, 2002). This model proposes that the rate of transcriptional initiation directly reflects the activation of kinases, so that a single phosphorylation event on a transcription factor supports a single round of transcription, or more generally, that a single signal firing through a signalling pathway is responsible for the kinase cascade activation and the enforcement of a single wave of histone modification marks that influences protein binding to specific gene loci. Since most of the dynamic histone marks and the phosphate groups on these kinases and TFs are continually removed, the balance between kinetics of modifying enzymes and removing activities influence the frequency of transcription initiation. At early time points the balance between the kinetics of enzyme keeps the frequency of initiation low, then the repetitive firing of the signalling pathway gradually changes the balance increasing the frequency of initiation, leading to the

accumulation of transcripts at later time points. Thus the persistence of a signal, either a stress or a chemical agonist, can be imposed at the level at which a gene is regulated by the modulation of frequency of transcriptional initiation. This model is energetically consuming but allows fine regulation of the signal-transcription rate at many levels, during the firing of the signal through kinase cascades and at the level of chromatin binding, at which diffusible molecules present in large numbers modulate an effect on a singular entity, the regulatory element of a specific gene. This frequency-modulated signal may have a major implication in the regulation of genes such as IE genes, for which continuously variable regulatory mechanisms coupled to the intensity of signalling has obvious biological significance. In the context of a "*continuously variable gene switch",* HMGN1 is likely to play a dual role. First, it may act as a barrier to chromatin compaction, enforcing an accessible state of chromatin due to its dynamic competition for histone H1 binding sites and its function as barrier to H3 de-acetylation (Catez et al., 2004; Lim et al., 2005; Ueda et al., 2006). Secondly, it can act as a signalling molecule via phosphorylation of Ser6 and possibly regulating the binding of MSKs to H3 tails (Barratt et al., 1994; Lim et al., 2004). It is unclear if the function of HMGN1 could be extended to a global genomic level or if it is instead restricted to a specific class of genes and also if HMGN1 phosphorylation is a primary event occurring to promote gene activation or if it is enforced secondarily after early transcription has begun to maintain an adequate efficiency of signal-transcription rate (Birger et al., 2003; Cuddapah et al., 2011; Lim et al., 2004; Rochman et al., 2011).

1.5 Thesis aim and summary

The aim of this thesis is to analyse the modifications of HMGN1 and its possible role in the regulation of immediate early genes, *c-fos* and *c-jun*. The most important evidence that associates HMGN1 with IE gene induction is the phosphorylation of HMGN1Ser6 concomitantly with histone H3 phosphorylation. The discovery of other phosphorylation marks on HMGN1 NBD and RD, together with the report of acetylation of HMGN1 at several residues, necessitates revisiting the role of HMGN1 and its modification in signalling, with more detailed evaluation of the different subpopulation of HMGN1 molecules present in the mouse fibroblasts model during induction. Therefore, I have investigated the state of phosphorylation and acetylation of HMGN1 molecules in different cellular extracts and the dynamic of appearance of different modification profiles in response to various stimuli. Using multiple experimental approaches, including mass spectrometry, PAGE electrophoresis, western blot and immunofluorescence, I have identified a single modification which responds to stimulation, phosphorylation of HMGN1Ser6. Phosphorylation at NBD and RD, as well as acetylation have been detected, but at remarkably low levels and/or without any relation to cell stimulation. This investigation identifies a new nuclear and acid extractable fraction of HMGN1 molecules carrying a distinct profile of PTMs in *non-quiesced* cells (chapter 3).
In parallel to the characterization of the modification profile of HMGN1, I have optimized the condition for chromatin immunoprecipitation of HMGN1 with two distinct antibodies, Ab-R371 and Ab-S520, recognizing respectively HMGN1 C-terminal and HMGN1Ser6ph respectively. The distribution of HMGN1 at the *c-fos* and *c-jun* loci has been investigated under different conditions, including variation of the cross-linking protocol for chromatin preparation and induction with various stimuli. The recovery of DNA sequences cross-linked with HMGN1 molecules is significantly lower than recovery of histones by ChIP, as expected

considering the dynamic interaction and the relative abundance of the proteins, but remarkably higher than for other transcription factors and similar to that of RNA polymerase II. Distribution does not vary between silent and active genes, in both control samples and induced samples. At some specific loci there is a slight variation of abundance upon induction while at other positions almost no change is observed. Overall a reduction of abundance is detected upon stimulation after 15-30 min, and recovery of basal level seems to occur after 60-90 min. Sequential ChIP using HMGN1 and antibodies targeting histones shows that the procedure is indeed capturing HMGN1 and nucleosomes in a direct interaction and their relative abundance suggests that HMGN1 does not preferentially associate with either H3K9ac or H3K4me3 (chapter 4).

Chapter 2. Materials and methods

2.1 Nomenclature

Modified histones and modified HMGN1 are defined using the Brno nomenclature (Turner, 2005) (see also list of abbreviations). For example H3S10ph indicate a histone H3 modified on serine 10 by the addition of a phosphate group and similarly HMGN1Ser6ph indicate a HMGN1 protein carrying a phosphorylation mark on serine 6. Multiple modifications are combined such as H3K9acS10ph indicates a histone H3 with acetylation at lysine 9 and phosphorylation at serine 10.

2.2 Materials

Unless stated otherwise, chemicals were purchased from Greiner Bio One or SIGMA, and cell culture reagents from Invitrogen. Recombinant human epidermal growth factor (rhEGF) was obtained from PROMEGA and microcystin-LR, H89, U0126 and SB203580 were purchased from Alexis Biochemicals (Nottingham, UK). PCR primers were synthesised by Eurogentec.

2.2.1 Antibodies

The antibodies used in this work have been produced and characterized in the lab and have been described previously (Thomson et al., 1999; Clayton et al., 2000; Dyson et al., 2005; Edmunds et al., 2008). Table 2.1 lists the conditions used for ChIP, western blot and immunofluorescence.

Table 2.1 **Antibodies.** Used in western blotting, immunofluorescence and immunoprecipitations.

Antibody	Source	Species	Antibody working dilution		Ammount used for ChIP [a]
			Western Blot	Immunoflourensce	
HMGN1-C-term [R371]	lab antibody	rabbit polyclonal AP	**1:2000**	**1:1000**	**120 ul**
HMGN1-Ser6ph [S520]	lab antibody	sheep polyclonal AP	**1:3000**	**1:1000**	**-**
H3K4me3	lab antibody	rabbit polyclonal serum	**1:10000**	**-**	**40 ul**
H3K9ac	lab antibody	rabbit polyclonal serum	**1:10000**	**-**	**40 ul**
H3S10ph	lab antibody	sheep polyclonal AP	**1:5000**	**-**	**-**
H3K9acS10ph	lab antibody	sheep polyclonal AP	**1:5000**	**-**	**-**
anti-rabbit-HRP	GE Healthcare	donkey polyclonal AP	**1:5000-1:10000**	**-**	**-**
anti-sheep/goat-HRP	AbD Serotec	donkey polyclonal AP	**1:5000-1:10000**	**-**	**-**
anti-rabbit-Cy3.5	Jackson Immunoresearch	donkey polyclonal AP	**-**	**1:100**	**-**
anti-sheep/goat-FITC	Jackson Immunoresearch	donkey polyclonal AP	**-**	**1:50**	**-**

AP: affinity purified, HRP: horse radish peroxidase, [a] used for 1 x 145 x 22 mm plate of confluent cells

2.3 Cell culture and stimulation

C3H 10T ½ mouse embryonic fibroblast cells were cultured in Dulbecco's modified Eagle's medium (DMEM) containing fetal calf serum (FCS) 10% (v/v) and 2 mM glutamine, maintained at 37°C and 5% CO_2. Cells were grown to confluency before splitting of a ratio 1:18-1:20 into fresh flasks. Quiescence was achieved by incubation of confluent dishes with DMEM containing 0.5% FCS and 2 mM glutamine for 18-24 h. Induction was performed adding EGF 50 ng/ml and/or anisomycin 10 µg/ml (inhibiting concentration, An), 25 ng/µl (sub-inhibiting concentration, sAn), or TPA 100 nM for the indicated time. Kinase inhibitors were added prior to stimulation for the indicated times (Figures 20-22). Inhibitors were used at the following concentrations: H89 (10 µM), U0126 (10 µM), SB203580 (10 µM) (see Bain et al., 2003, 2007; Davies et al., 2000 for references). UV irradiation was performed exposing cells in culturing dishes to 200 J/m^2 of UV radiation (254 nm), delivered via a Spectrolinker XL-1000 (Spectronics Corp.).

2.4 Protein extraction

Cells lysis was performed in different buffers, a histone and nuclear protein extraction buffer and two chromatin preparation buffers (Clayton et al., 2000; Edmunds et al., 2008; O'Neill and Turner, 2003). Details on the chemicals and concentration used are presented in table 2.2. Ice-cold lysis buffer was added directly to culturing dish after aspiration of the medium and cells were scraped off, moved to sample tubes and incubated on ice for 10 min. Further extraction procedures differ from sample to sample (see following chapters). Proteinase inhibitors added to the lysis buffers, consisted of (final concentration) 2 µg/ml aprotinin, 5 µg/ml leupeptin, 50 µg/ml soybean trypsin inhibitor, 100 nM tosyl lysyl chloromethyl ketone

(TLCK), 100nM tosyl phenylalanyl chloromethyl ketone (TPCK) and 2 mM phenylmethyl sulphonyl fluoride (PMSF). Phosphatases inhibitors consisted of 100 μm ortovanadate ($NaVO_4$), 10 mM sodium butyrate, 20 mM β glycerol-phosphate and 1 μM microcystin-LR.

2.4.1 Acid extraction of histones and HMGNs

The lysates prepared with histone and nuclear protein extraction buffer were centrifuged at 800 g for 5 min at 4°C [(Heraeus) Biofuge Fresco MicroFuge] and supernatant fraction was collected, designated as soluble fraction.

Nuclear total HCl extraction

Nuclear pellets were dissolved in 0.4 M HCl. After 1 hour incubation on ice, insoluble proteins were pelleted by centrifugation at 16.000 g for 15 min at 4°C. Nuclear samples supernatant was collected and incubated overnight (o/n) at -20°C with acetone.

Nuclear PCA and HCl extraction

Nuclear pellets were dissolved in 5% perchloric acid (PCA). After 1 hour incubation on ice, insoluble proteins were pelleted by centrifugation at 16.000 g for 15 min at 4°C. Nuclear PCA supernatant was collected, adjusted to 0.4 M HCl and incubated (o/n) at -20°C with 10 volumes of acetone. Pellets were dissolved in 0.4 M HCl for 1 hour on ice. Nuclear HCl supernatant was collected after centrifugation at 16.000 g for 15 min at 4°C and incubated (o/n) at -20°C with acetone.

Soluble fraction, PCA/TCA extraction

Soluble fractions were adjusted to 5% PCA. After 1 hour incubation on ice, insoluble proteins were pelleted by centrifugation at 16.000 g for 15 min at 4°C. Samples supernatant was collected and adjusted to trichloroacetic acid (TCA) 25% and after 30 min incubation on ice,

insoluble proteins were pelleted by centrifugation at 16.000 g for 15 min at 4°C. Supernatant was discarded and the pellet was incubated (o/n) at -20°C with acetone.
Samples were resuspended in 4 M or 8 M urea (see chapter 2.5.2 AU-PAGE).

2.4.2 Calf intestinal alkaline phosphatase (CIP) treatment

PCA or HCl extracted HMGN1 samples were solubilised in 300 µl of NEB buffer 2 [50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM dithiothreitol] (New England Biolabs) instead of urea. One third of the samples were kept on ice 1 hour, to avoid protein degradation. One third was incubated at 37°C for 1 hour, to test for possible protein degradation. One third was incubated with 5 µl alkaline phosphatase (10,000 U/ml) at 37°C for 1 hour. After digestion all samples were adjusted to 25% (w/v) TCA, pelleted by centrifugation at 16.000 g for 15 min at 4°C and washed with acetone. Samples were solubilised in 8 M urea and analyzed by AU PAGE (see chapter 2.5.2).

2.5 Polyacrylamide gel electrophoresis (PAGE)

2.5.1 Sodium dodecyl sulphate (SDS) PAGE

2.5.1.1 Sample preparation

Sample extracted with any of the mentioned buffers (chapter 2.4) after been dried of the acetone were resuspended in 4 M urea and adjusted to Laemmli buffer [10 mM Tris-HCl pH 6.8, glycerol 10% (v/v), SDS 2% (w/v), β-mercaptoethanol 5% (v/v)and bromophenol blue 0.01% (w/v)] and incubated for 1 min at 95°C.

2.5.1.2 Electrophoresis

SDS-PAGE was carried following standard protocols (Laemmli, 1970). Acrylamide gels were prepared using Protogel solutions (National Diagnostics). Unless otherwise stated, standard SDS gel contained in the stacking portion, 125 mM Tris-HCl pH 6.8, SDS 0.1% (w/v), acrylamide 5% (w/v) and bisacrylamide 0.13% (w/v). The resolving portion contained, 375 mM Tris-HCl pH 8.8, SDS 0.1% (w/v), acrylamide 15% (w/v) and bisacrylamide 0.4% (w/v). Gels were polymerised by addition of N,N,N',N'-tetramethylethylenediamine (TEMED) (v/v), 0.1% (resolving) or 0.3% (stacking), and ammonium persulphate (APS) (w/v) 0.1% (resolving) or 0.05% (stacking).

Samples were run at 150 Volt in Laemmli running buffer [25 mM Tris, 192 mM glycine, SDS 0.1% (w/v)] and transferred to polyvinyl difluoride (PVDF) membranes for immunoblotting (chapter 2.6.3).

2.5.2 Acid-urea (AU) PAGE

2.5.2.1 Sample preparation

For proper AU PAGE analysis samples require solubilisation in urea. Samples were solubilised with 8 M urea, β-mercaptoethanol 5% (v/v) and incubated at room temperature for about 10 min before addition of acetic acid 5% (v/v) and methyl green 0.015 % (w/v) (final concentration). Addition of acetic acid prevent the chemical modification of cationic lysine residues by urea (Shechter et al., 2007).

2.5.2.2 Gel preparation

AU electrophoresis was conducted as previously described (Barratt et al., 1994). The resolving portion of the gel contained acrylamide 20-25% (w/v), bisacrylamide 0.2% (w/v),

8 M urea and 5% acetic acid. The stacking portion contained acrylamide 7.5% (w/v), bisacrylamide 0.1% (w/v), 8 M urea and 5% acetic acid. Urea was dissolved in gel solutions at 45-50°C and then cooled to about room temperature before pouring of the gel. Gel were polymerised by addition of N,N,N',N'-tetramethylethylenediamine (TEMED) (v/v), 0.4% (resolving) or 0.8% (stacking), and ammonium persulphate (APS) (w/v) 0.2% (resolving) or 0.4% (stacking).

2.5.2.3 Electrophoresis

Gel were run in a AU buffer containing acetic acid 5% (v/v) and electrodes were reversed (compared to SDS-ERK PAGE) to allow the lower reservoir to become the cathode and *vice versa* the higher reservoir to become the anode. AU gels require a pre-run to elute ionic contaminants, usually (o/n) at 150 Volts. Running buffer was replaced with fresh one prior to electrophoresis of samples. Efficient separation of HMGN1 molecules in 20-25% acrylamide gels require about 8-10 hours at 15 mA. Transfer of sample was performed following the AU procedures (chapter 2.6.3)

2.5.3 Coomassie staining

Proteins in gels were visualized by staining with Coomassie blue solution [Coomassie brilliant blue 0.1% (w/v), isopropanol 25% (v/v), and acetic acid 10% (v/v)] (o/n), and adequate destaining in methanol 25% (v/v) and acetic acid 7.5% (v/v).

2.6 Western blotting

2.6.1 Protein transfer

All protein transfers were performed following a wet-transfer procedure, assemblage of the transfer sandwich and transfer buffer differs depending on the PAGE system. Samples were transferred (o/n) at 200 mA at 4°C with continuous stirring of the transfer buffer.

2.6.1.1 SDS transfer

PVDF membranes (Immobilion-P; Millipore) required pre-wetting in methanol for 1 min before assembly of the transfer sandwich. Transfer was performed in 0.5x Laemmli running buffer [12.5 mM Tris, 96 mM glycine, SDS 0.05% (w/v)] containing methanol 10%.

2.6.1.2 Towbin transfer

The Towbin transfer procedure differs from SDS transfer. Towbin transfer buffer contain 25 mM Tris, 192 mM glycine, methanol 20% (v/v). Since SDS PAGE is performed in running buffer containing SDS, gels were rinsed in water and then in transfer buffer before assembly of the transfer sandwich.

2.6.1.3 AU transfer

AU gel requires to be equilibrated for 30 min in methanol 10% (v/v) supplemented with acetic acid 5% (v/v). The PVDF membranes were pre-wetted in methanol for 1 min. Then, PVDF and AU gels were incubated in AU transfer buffer [methanol 10% (v/v) and acetic acid 0.1% (v/v)] for about 5 min. Finally they were assembled in the transfer sandwich with fresh AU transfer buffer. AU transfer requires reversed electrodes.

2.6.2 Ponceau staining

Efficient and correct transfer was assayed by staining of PVDF membranes with Ponceau S solution [Ponceau 0.4% (w/v), acetic acid 5% (v/v)] for 2-3 min. Adequate destaining was obtained with methanol washes. Then membrane were allowed to dry and either stored or subjected to immunoblotting. Migration of histones was labelled with pencil on blank regions of the membranes to allow precise identification of western blot signals.

2.6.3 Immunoblotting

Membranes were directly probed with primary antibodies (see table 2.1 for dilutions) in blocking solution [skimmed milk powder 5% (w/v), Tween-20 0.1% (v/v), phosphate buffered saline (PBS): 137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4] for 1 hour at room temperature. Membranes were then washed three times for about 10min with PBS-Tween-20 0.1% (v/v), and then incubated with secondary antibody in blocking solution. Three additional washes were performed prior to detection. Samples signal was detected by enhanced chemiluminescence (ECL) using either the Immobilion western chemiluminescence HRP substrate (Millipore) or two solution of ECL reagents, Cumaric-Luminol [100 mM Tris-HCl pH 8.5, 1.25 mM 3-aminophthalhydrazide (Luminol), 200 µM p-Cumaric acid and H_2O_2 0.01% (v/v)] or 4IPBA-Luminol [100 mM Tris-HCl pH 8.5, 1.25 mM 3-aminophthal-hydrazide (Luminol), 2 mM 4-iodophenylboronic acid (4IPBA) and H_2O_2 0.01% (v/v)]. Membranes were incubated in these ECL solutions for 3-5 min (Millipore, 4IPBA-Luminol) or 1 min (Cumaric-Luminol) and then exposed to X-ray films (Kodak) to visualize the antibody-HRP signal. X-ray films were then scanned with HP ScanJet G4050 obtaining Jpeg images that were analysed for quantification of the immunoblot signal with Image J software (National Institutes of Health).

2.7 Indirect immunofluorescence (i-IF)

2.7.1 Preparation of samples for (i-IF)

Dividing C3H 10T ½ cells were plated on glass coverslips, moved into 24-well plates ($6x10^4$ cells per well) and quiesced by incubation for 24 hours in DMEM supplemented with FCS 0.5% (v/v) and 2 mM glutamine. Cells were treated as stated, then washed once with ice-cold PBS and fixed in 4% paraformaldehyde in PBS for 20 min at room temperature. Cells were washed three times in PBS then permeabilised for 5 min in ice-cold PBS with Triton X-100 0.5% (w/v), and then washed into PBS.

2.7.2 Antibody binding

Blocking solution [10% bovine serum albumin (BSA)] was added to the fixed coverslips for 1 hour, then it was replaced with a blocking solution containing the appropriate primary antibody (see table 2.1 for dilutions) for 1 hour. Coverslips were then washed three times with PBS and then incubated for 1hour in the dark with blocking solution containing the fluorophore-coupled secondary antibodies. Coverslips were then washed three times with PBS and mounted on microscope slides using VECTASHIELD mounting medium (Vector Labs), containing DAPI (4',6'-diamidino-2-phenylindole) counterstain to label DNA. Immunofluorescence was detected using an Axioscope 2 fluorescence microscope with Axiovision imaging software (Zeiss). Settings and exposure times for each channel were optimized and kept constant for each experiment.

2.8 Chromatin immunoprecipitation (ChIP)

2.8.1 Preparation of mononucleosomes

Mononucleosomes were prepared by micrococcal nuclease digestion (MNase) of chromatin from either cross-linked or native preparations.

2.8.1.1 Cross-linked chromatin

Cross-linked chromatin was prepared as previously described (Macdonald et al., 2005; Crump et al., 2011). Cells plates (3x 145 mm cell dishes for each treatment) were quiesced for 18 hours in DMEM supplemented with FCS 0.5% (v/v) and 2 mM glutamine. Cells were treated as stated, then cross-linked by incubation with formaldehyde 1.1% (v/v) for 10 min at either 37°C or room temperature. Dishes were washed two time with ice cold PBS and lysed in chromatin preparation buffer (see table 2.2) and kept on ice for 10 min. Nuclei were then centrifuged at 800 g for 4 min at 4° and washed once with ice cold wash buffer (10 mM Tris HCl pH 8.0, 1 mM EDTA, 0.5 mM EGTA, 200 mM NaCl, 10 mM sodium butyrate, 1 mM sodium orthovanadate, 20 mM sodium β-glycerophosphate,1 mM microcystin-LR and complete proteinase inhibitor cocktail) and once with MNase buffer (10 mM Tris-HCl pH 7.5, 250 mM sucrose, 75 mM NaCl, 10 mM sodium butyrate, 1 mM sodium orthovanadate, 20 mM sodium β-glycerophosphate, 1 mM microcystin-LR and complete proteinase inhibitor cocktail). Samples from three plates were pooled together in MNase buffer adjusted to 3 mM $CaCl_2$ and digested with 100 U of MNase for 10 min at 37°C. Digestion was stopped by the addition of EDTA (5.8 mM final concentration). Samples were then adjusted to 0.5% SDS and rotated at room temperature for 1 hour. Insoluble material was pellet by centrifugation at 16.000 g for 10 min at room temperature. A portion (1/20) of soluble material was analyzed for its DNA contents to ensure proper digestion (see following chapters). Soluble chromatin

was adjusted to RIPA buffer [10 mM Tris HCl pH 8.0, Triton X-100 1%(v/v), SDS 0.1% (w/v), sodium deoxycholate 0.1%, 1 mM EDTA, 0.5 mM EGTA, 140 mM NaCl] for ChIP analysis.

2.8.1.2 Native chromatin

Native chromatin was prepared as previously described (O'Neill and Turner, 2003). Cells plates (3x 145 mm cell dishes for each treatment) were quiesced for 18 hours in DMEM supplemented with FCS 0.5% (v/v) and 2 mM glutamine. Cells were treated as stated, then were lysed with 10 mM Tris-HCl pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% NP-40 (v/v), 320 mM sucrose, 10 mM sodium butyrate, 1 mM sodium orthovanadate, 20 mM sodium β-glycerophosphate, 1 mM microcystin-LR and complete proteinase inhibitor cocktail. The lysates were centrifuged at 800 g for 5 min at 4°C and supernatant fraction was collected. Nuclear pellets were washed with wash buffer [lysis buffer without 0.1% NP-40] and centrifuged at 800 g for 5 min at 4°C. Washed pellets were resuspended in wash buffer and three samples pooled together and layered on a 30% sucrose cushions. Centrifugation 1000 g for 15 min at 4°C of the nuclei was followed by resuspension in wash buffer and glycerol 40%. Samples were stored at -80°C. On a different day MNase digestion was performed. Samples were centrifuged at 800 g for 5 min at 4°C and washed three times with nuclei wash buffer [10 mM Tris HCl pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 250 mM sucrose, 10 mM sodium butyrate, 1 mM sodium orthovanadate, 20 mM sodium β-glycerophosphate, 1 mM microcystin-LR and complete proteinase inhibitor cocktail]. Pellets were resuspended in nuclei wash buffer and quantified with spectrophotometer. Samples were adjusted to a final DNA concentration of 1 mg/ml and 1 mM $CaCl_2$ and incubated for 10 min 37°C, with 50 U/ml of MNase. Digestion was stopped adjusting the sample to 5 mM EDTA and living them on ice for 10 min. Samples were centrifuged at 16.000 g for 15 min at 4°C and supernatant

collected as S1 fraction. Pellets were resuspended in the same previous volume with swelling buffer [10 mM Tris HCl pH 7.5, 5 mM EDTA, 10 mM sodium butyrate, 1 mM sodium orthovanadate, 20 mM sodium β-glycerophosphate, 1 mM microcystin-LR and complete proteinase inhibitor cocktail]. Samples were vortexed and kept on ice for 30 min and then centrifuged at 16.000 g for 20 min at 4°C. The supernatant was collected as S2 fraction. Pellets were resuspended in same previous volume with swelling buffer adjusted to 0.5% SDS. A portion (1/20) of each fraction was analyzed for its protein and DNA contents to ensure proper digestion and to identify the fraction HMGN1 enriched. S1 and S2 usually contain 100% of HMGN1 (data not showed). These two fractions were pooled and adjusted to RIPA buffer for ChIP analysis.

Table 2.2. **Lysis buffers.** Used in preparation of nuclear extracts and chromatin preparation for immunoprecipitation

Lysis buffers		
Nuclear proteins	**N-Chromatin preparation**	**X-Chromatin preparation**
0.2% Triton X-100 (v/v)	Tris-HCl pH 7.5 10 mM	Tris-HCl pH 8.0 10 mM
10 mM HEPES pH 7.6	10 mM NaCl	Triton X-100 0.25% (v/v)
1.5 mM $MgCl_2$	3 mM $MgCl_2$	EDTA 10 mM
10 mM KCl	NP-40 0.1% (v/v)	EGTA 0.5 mM
250 mM sucrose (optional)	320 mM sucrose	-
100/200 µm $NaVO_4$	100/200 µm $NaVO_4$	100/200 µm $NaVO_4$
10 mM sodium butyrate	10 mM sodium butyrate	10 mM sodium butyrate
20 mM β glycerophosphate	20 mM β glycerophosphate	20 mM β glycerophosphate
1 µM microcystin LR	1 µM microcystin LR	1 µM microcystin LR
1x proteinase inhibitor A	1x proteinase inhibitor A	1x proteinase inhibitor A
1x proteinase inhibitor B	1x proteinase inhibitor B	1x proteinase inhibitor B
H_20	H_20	H_20

Proteinases inhibitors A/B described in Chapter 2.4

2.8.1.3 Analysis of DNA size

To ensure proper MNase digestion the 1/20 aliquot of cross-linked chromatin were analyzed. Samples were diluted four-fold with water, and then treated with RNAse A (100 mg/ml, 1 hour, 37°C) and SDS added to 0.25% (w/v) before treatment with proteinase K (100 mg/ml, 65°C (o/n), for reversal of cross-links. DNA was extracted from samples by addition of equal volume of phenol: chloroform and centrifugation at 16.000 g for 10 min at room temperature. The supernatant aqueous phase containing DNA was then loaded to a 1% agarose-TAE (40 mM Tris-acetate, 1 mM EDTA pH 8.0) gel containing 0.5 μg/ml ethidium bromide. Usually 16 μl of the aqueous phase were mixed with 4 μl of 5x gel-loading buffer [glycerol 50% (v/v), 100 mM EDTA pH 8.0, Orange G 0.25% (w/v)]. Samples were electrophoresed at 100 Volts for the required time in TAE buffer and visualised by UV transillumination.

2.8.2 Immunoprecipitation (ChIP)

ChIPs were performed as described previously (Edmunds et al., 2008). Antibodies (see table 2.1) were added to chromatin aliquots and rotated for 2 hours at 4°C. Protein A-Sepharose-beads (≈5 mg) pre-blocked for 10 min with 10 mg sonicated herring sperm DNA and BSA (0.02% final concentration after addition to chromatin) were incubated with the chromatin-antibodies and rotated for additional 2 hours at 4°C. Beads were recovered by centrifugation (2000 rpm for 3 min at 4°C), and unbound supernatant chromatin was retained and analysed for epitope depletion. Beads were washed sequentially five time with RIPA buffer then washed twice with 1x TE [10 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0] then resuspended in 100 μl of 1x TE. Bound ChIP DNA was then extracted in parallel with input DNA. (see chapter 2.8.4)

2.8.3 Sequential immunoprecipitation

2.8.3.1 Preparation of antibody cross-linked beads

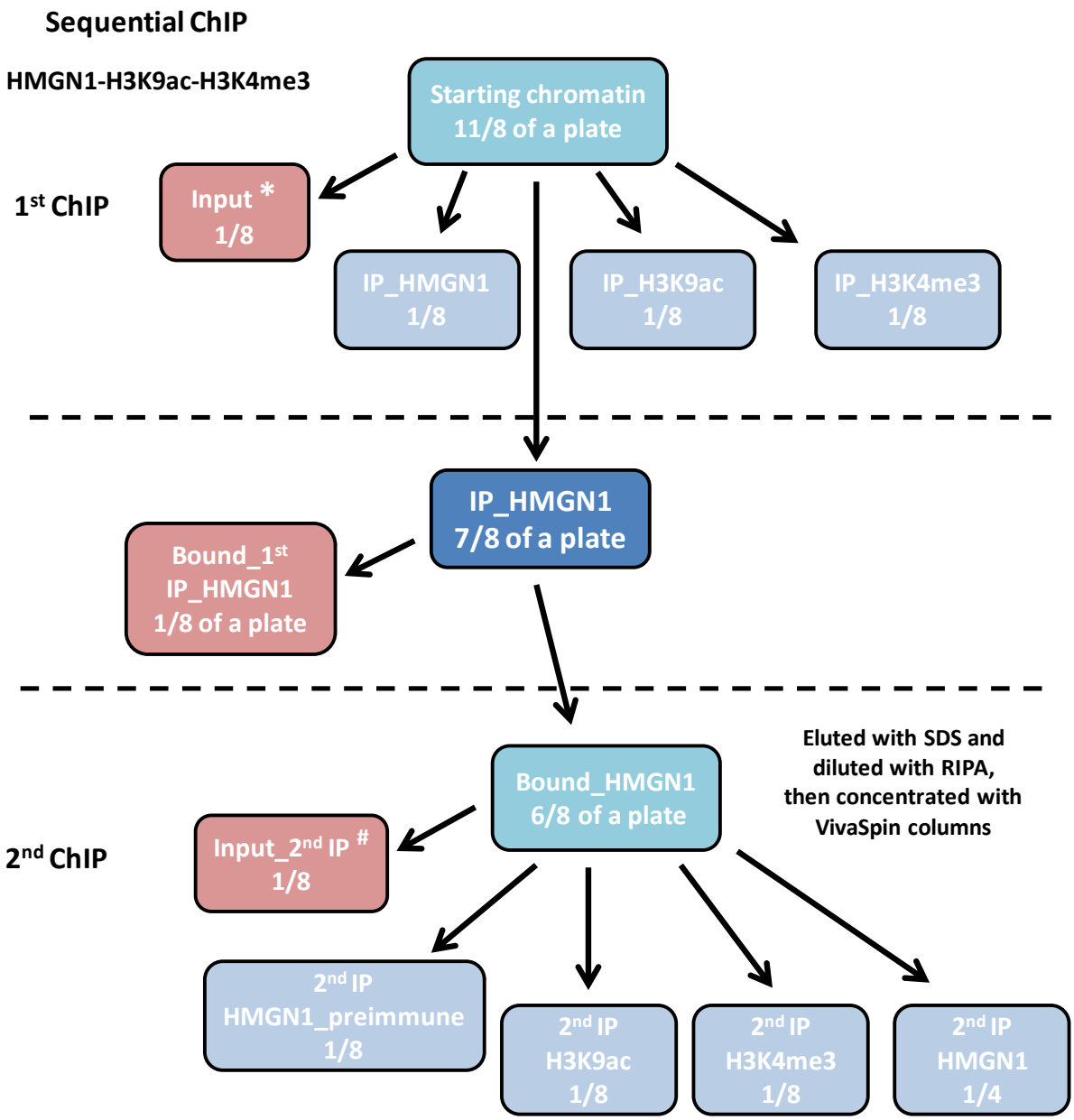
Sequential ChIP was performed using HMGN1 C-terminal, H3K9ac and H3K4me3 antibodies (or pre-immune sera) covalently bound to protein-A Sepharose beads (Figure 10). Antibodies (see table 2.2) were incubated with RIPA buffer-hydrated beads, for 2 hours at 4°C. Bound beads were washed three times with RIPA buffer and twice with 0.2 M sodium borate pH 9.0. Beads were left in 1 ml 0.2 M sodium borate pH 9.0 and dimethyl pimelimidate (DMP) powder was added to 20 mM final concentration. Beads were rotated for 30 min a room temperature. DMP cross-linking was stopped by washing the beads twice with 0.2 M ethanolamine pH 8.0. Contamination from not cross-linked antibody was removed by incubation in 0.2 M glycine pH 2.0 with rotation for 30 min at room temperature. Beads were then washed once with 1 M Tris-HCl pH 8.0 and twice with RIPA buffer.

2.8.3.2 Immunoprecipitation

Immunoprecipitation (Figure 10) was performed as explained previously (chapter 2.8.2) except that cross-linked beads were pre-blocked for 10 min with 10 mg sonicated herring sperm DNA and BSA (0.02% final concentration) before the incubation with chromatin. Unbound chromatin was discarded after pelletting beads (800 g, 4 min, at 4°C) and beads were washed as previously described (chapter 2.8.2). Bound chromatin was eluted by addition of 1x TE, SDS 1% (w/v) and rotation for 30 min at room temperature. An aliquot of this chromatin was collected as control and the rest was then diluted ten-fold with RIPA buffer. Bound chromatin was then concentrated using VivaSpin 20 columns (50.000 kDa cut-off PES membrane, Sartorius), and divided into aliquots for the second round of IP and for analysis of the first round IP-DNA. Second IP was performed as usual. Besides the antibodies, pre-immunized serum was used as internal control to test unspecific purification.

Figure 10. **Sequential ChIP.**

Schematics of a sequential ChIP using HMGN1 C-terminal antibody as 1st round antibody for ChIP. The same procedure was followed for the reverse sequential ChIP having H3K9ac antibody as 1st round antibody for ChIP. Notice that due to the low recovery of HMGN1-bound DNA sequences and the loss of signal during a sequential ChIP, it is necessary to involve a double amount of chromatin for the 2nd round HMGN1 ChIP. DNA recoveries were quantified by real-time PCR. Bound DNA from 1st round ChIPs was related to **input *** from starting chromatin while bound DNA from 2nd round ChIPs was related to **input #** generated during the 1st round ChIPs.



2.8.4 Purification of DNA from ChIP fractions

Input DNA was extracted from an aliquot of chromatin equivalent to those used in ChIP. Chromatin was pelleted adding 300 mM sodium acetate pH 5.2 and 2-3 volumes of ethanol and then incubating (o/n) at -20°C. The next day followed a centrifugation at 16.000 g for 30 min at 4°C and resuspension in TE buffer.

The input DNA and the bound ChIP DNA were released from chromatin by treatment with RNAse A (100 mg/ml, 1 h, 37°C) then addition of SDS to 0.25% (w/v) before treatment with Proteinase K (100 mg/ml, 65°C o/n, for reversal of cross-links). DNA was then extracted using an equal volume of phenol, followed by centrifugation at 16.000 g for 10 min room temperature. The aqueous phase was then isolated and the phenol phase was re-extracted using additional 100 µl of water and by centrifugation at the same conditions. The two aqueous phases combined were then purified with an equal volume of phenol: chloroform (1:1), then chloroform and centrifugation at the same conditions. The DNA in final aqueous phase was precipitated adding 300 mM sodium acetate pH 5.2 and 1 ml of ethanol and then incubating (o/n) at -20°C in the presence of 20 µg of glycogen (Roche). The next day followed a centrifugation at 16.000 g for 30 min at 4°C then washes with ethanol 80% (v/v). DNA was resuspended in 1x TE pH 7.5 [10 mM Tris-HCl, 0.1 mM EDTA] and analyzed by real-time PCR.

2.9 Histones and HMGN1 immunoprecipitation

In order to analyse bound protein and unbound protein fractions during ChIP, immunopurifications were performed as previously described (chapter 2.8.2) except that unbound fractions were collected and after washes with RIPA buffer, chromatin was extracted adding 200 µl of immuno-elution buffer A [SDS 0.5% (w/v), 10 mM Tris-HCl pH 8.0, 1 mM

EDTA, pH 7.5, 10 mM sodium butyrate, 20 mM sodium β-glycerophosphate, 30 mM NaCl] followed by rotation for 15 min at room temperature and centrifuged at 800 g for 3 min at room temperature. Supernatant was collected as bound fraction. Beads pellets were resuspended in immuno-elution buffer B [1.5% SDS, 10 mM Tris-HCl pH 8.0, 1 mM EDTA, pH 7.5, 10 mM sodium butyrate, 20 mM sodium β-glycerophosphate, 30 mM NaCl] and after incubation for 10 min at room temperature, centrifuged at 800 g for 3 min at room temperature. Supernatant was collected and added to bound fraction. The unbound and bound fractions were treated with phenol:chloroform to extract DNA. The aqueous phase was stored at -20°C. The phenol phase was adjusted to 100 mM HCl and incubated (o/n) at -20°C with 10 volumes of ice cold acetone. The following day after three acetone washes samples were dried. To reverse cross-links, input, unbound and bound samples were heated at 95°C for 30 min before loading to gel.

Alternatively if only protein from unbound fractions and input are analyzed via SDS-PAGE, samples can be adjusted to 1x Laemmli buffer and heated at 95°C for 30 min to reverse cross-links before loading to gel.

2.10 Real-time PCR

Real-time PCR was performed using the ABI Prism 7000 sequence detection system (Applied Biosystems). Primers (table 2.3) were designed by A. Clayton and J. Edmunds (Edmunds et al.,2008) using Primer Express 2.0 software (Applied Biosystems). Sequences of *c-fos*, *c-jun*, *gapdh* and *hbb* were detected using the 5'nuclease (Taqman) assay, with 5'FAM 3'TAMRA-labelled probes and qPCR Mastermix Plus (Eurogentec). Primers were tested for specificity by BLAST search. PCR reactions were carried out in duplicate for each biological sample.

Table 2.3. **Real-time PCR primers.**

Primer name	Primer sequence		5′FAM 3′TAMRA probe	amplicon (bp)	Final concentration (nM)	
	forward	reverse			forward	reverse
***c-fos* -1527**	CGAGGCAGGTTTGTGGTTCT	TTCCAAGACGGACTAGAGAGGATTA	CGCCCCTAAAGATCCATCCGCTGT	81	300	900
***c-fos* -260**	GGGACCATCTCCGAAATCCT	TCTTTCACCTATGAGTGTTCACATTTG	CACGCGGAAGGTCTAGGAGACCCC	81	900	900
***c-fos* +119**	GGCTTTCCCCAAACTTCGA	GCTACTGCAGCGGGAGGAT	CTCGGGTTTCAACGCCGACTACGA	78	900	900
***c-fos* +444**	GCGTGCACACTTGTCATAGTAAGA	CGGCCGAGAGAAGAATGC	TTAGTGACCCCTTCCCGCGCG	90	900	300
***c-fos* +1050**	CTGTCCGTCTCTAGTGCCAACTT	GAGACCAGAGTGGGCTGCA	CCCCACGGTGACAGCCATCTCC	92	900	900
***c-fos* +2078**	GGATTTGACTGGAGGTCTGCC	TGGGCTCAGGGTCGTTGA	CTGAGGAGGCCTTCACCCTGCCC	86	900	900
***c-jun* -966**	TTCGAACCTTTATATGTTGCTTTCG	GGACCCCTCCTCTAGGTGTGA	ATGCTAGCCCAGCGCCCATGG	75	300	300
***c-jun* -396**	AGACTCCGCAAGGGTCCTG	TTGTGGGCTGACGTCTTGG	CGACCGGCAGCCTCCGTCA	71	900	900
***c-jun* +153**	CTCCTGCACCGGTATTTTGG	AGCAGCGCCCGGACTT	AGTCCCTTCTCCCGCCTTCCACG	86	300	900
***c-jun* +651**	GAGTGACCGCGACTTTTCAAA	AGGGAGCGCAGGGTTAAGAC	CTAGCACTCACGTTGGTAGGCTCCCG	84	300	300
***c-jun* +1119**	TGCTCAAGCTGGCGTCG	GTGTAGTGGTGATGTGCCCATT	CTGGAGCGCCTGATCATCCAGTCC	72	300	300
***c-jun* +2296**	GATTGCTTCTGTAGTGCTCCTTAACA	GCAGTCTAGCCTGGCACTTACAA	CCCTTCCCAGCCCTCCCTGCTT	81	900	900
***gapdh* -546**	GCTGTGTCACTACCGAAGAACAA	CGAACCTCTCCCCATTATTGAA	AGGAGAAGATCCTCAACTTTTCCGCAGC	78	900	900
***gapdh* +751**	CCGCCATGTTGCAAACG	AGCTAGGAAGAAGGAAGGCCTAAG	TTTCATAACGGCGGTTCATTCATTTCCTT	77	300	300
***gapdh* +3390**	GGATGGGAAACCTGACTTTTAAGA	ATTGGGACAACTGTTAGTTTGGGT	TGAGCTAGCCACCAGAGGGCACCA	93	900	900
***hbb* -79**	GGGCCAATCTGCTCACACA	GAGCAACTGATCCTACCTCACCTT	TGCCCTGGCTCCTGCCCTCTCTAT	77	900	900
***hbb* +1245**	CCCTGGCTCACAAGTACCACTAA	GCTCTCTTGGGAACAATTAACCA	CCCTTTCCTGCTCTTGCCTGTGAACA	76	900	900

2.10.1 Preparation of standards with genomic DNA

DNA genomic standards were generated from C3H 10T ½ mouse cells, washed and harvested in PBS. After centrifugation at 800 g for 4 min at room temperature, cells were resuspended in lysis buffer [10 mM Tris-HCl pH 8.0, 400 mM NaCl, 2 mM EDTA pH 8.0], after adding SDS to 0.5% (w/v) samples were treated o/n with Proteinase K (250 mg/ml, 37°C). Proteins were then precipitated adding NaCl up to 1.5 M (final concentration) and centrifuged at 1000 g for 15 min at room temperature. DNA contained in the supernatant was then precipitated by addition of 2-3 volumes of ethanol. DNA was then resuspended in TE buffer. In order to allow a fair comparison with mononucleosomal DNA purified by ChIP, aliquots of standards were sonicated to 200-400 base pairs fragments using a Sonic Vibracell VC130 sonicator with a 3 mm probe.

2.10.2 Real-time PCR data analysis

Input, bound and unbound from ChIP assay were analysed including linear dilution of full length genomic standard, sonicated standards and blank control. If standards curve dilutions did not give a linear response, the plate was not considered valid. Data were analysed by standard curve absolute quantification method, comparing unknown samples to standard quantities of DNA. Data are represented as percentage of input: alias the ratio between the quantities of DNA from the same sequences, detected in the bound or unbound and in the input [bound / input x100]. Since several dilution occurs during the ChIP procedures all data are referred to a full chromatin plate before final calculation of percentage of input.

2.11 Mass spectrometry analysis (MS)

Mass spectrometry was performed in collaboration with Dr Benedikt Kessler and Dr Mukram MacKeen (Nuffield Department of Clinical Medicine, Oxford).

2.11.1 MS samples preparation

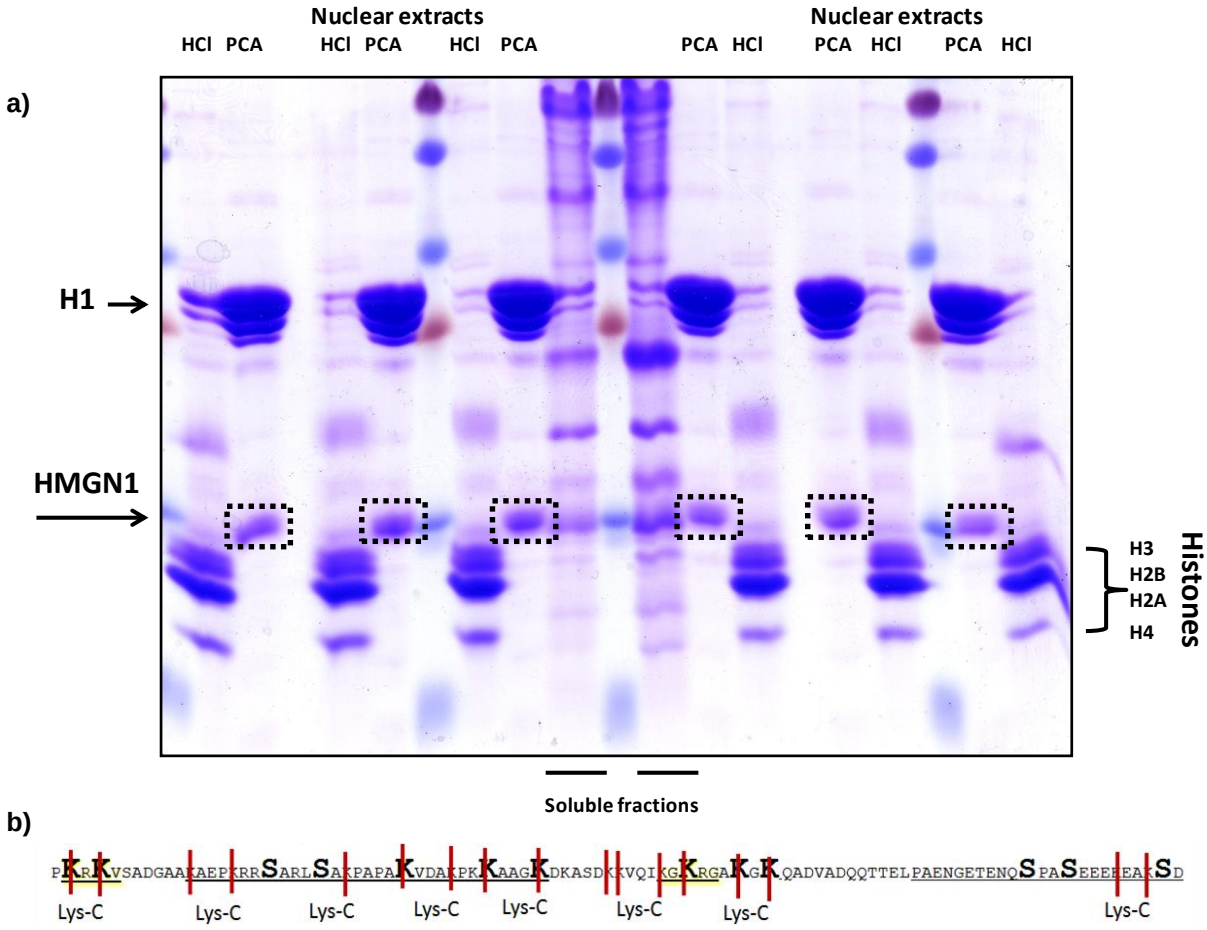
To ensure an adequate yield of HMGN1 proteins for MS detection, lysate using nuclear protein buffer from three confluent plates were merged and acid extracted (see chapter 2.4.1). Samples were then separated by SDS–PAGE and specific HMGN1 gel bands were cut (Figure 11). Samples within the gel were reduced with dithiothreitol and alkylated with iodacetamide before proteolysis with Lys-C enzyme sequencing grade (Roche) as previously described in (Batycka et al., 2006; Kinter and Sherman, 2000). Digests were desalted and subjected to LC-MS/MS (liquid chromatography-mass spectrometry).

2.11.2 Spectrophotometers, Walters Q-ToF, Agilent Ion Trap, OrbiTrap

Mass spectrometry analysis was performed with three different instruments, (Waters) Q-ToF, (Agilent) Ion Trap and Orbi Trap Velos (Thermo Scientific). Most of the data presented are exclusively qualitative, due to several difficulties in achieving an optimal detection of HMGN1 peptides. Only data from the Orbi Trap spectrophotometer have been sufficiently accurate to allow relative quantification of the peptides.

Figure 11. **Preparation of samples for MS.**

a) SDS-PAGE of nuclear protein samples. PCA extraction allows separation of HMGN1 and H1 from the nucleosome core histones (H3, H2A, H2B and H4). **b)** Schematic of HMGN1 amino acid sequence and predicted digestion sites with Lys-C enzyme.



2.11.3 Data analysis and quantification

Peptides were detected with Progenesis LC-MS software (Nonlinear Dynamics) and Agilent Masshunter Qualitative Analysis software using default settings. The peak list generated by Progenesis LC-MS was searched against the Swiss UniProt mouse database, using Mascot (http://www. matrixscience.com/), allowing three missed cleavages and 20 ppm/0.5 Da mass deviations in MS and MS/MS. Carbamidomethylation of cysteine was a fixed modification. Oxidation of methionine, deamidation of asparagine and glutamine, phosphorylation of serine, threonine and tyrosine and acetylation of lysine were used as variable modifications. Data on relative abundance of modifications were displayed as well as unmodified peptides, also overall recovery of all HMGN1 peptides is shown for comparison (see Figure 28, Table 3.1).

Chapter 3.

Signalling through the MAPK pathway induces phosphorylation of HMGN1 mainly at serine 6

3.1 Introduction

The aim of this work is to clarify the extent and role of post-translational modifications of HMGN1 during IE gene induction. Evidence from different laboratories suggests that HMGN1 is subjected to multiple modifications. To interpret the distribution of modified forms of HMGN1 across IE gene loci, it is first necessary to analyse the complexity of PTMs on HMGN1 molecules. HMGN1 protein extracted from C3H 10T1/2 mouse fibroblast lysates have been characterized using HMGN1Ser6ph [S520] and C-terminal [R371] antibodies, SDS and acid-urea gel electrophoresis, immunofluorescence and mass spectrometry.

3.1.1 HMGN1 phosphorylation, unresolved issues

Phosphorylation of HMGN1 occurs on several residues: Ser6 at the N-terminal, within the NBD (Ser20-24 in human, Ser19-23 in mouse) and at the C-terminal (Ser85-88-98 in human, Ser83-86-94 in mouse) (Barratt et al., 1994; Lim et al., 2004; Prymakowska-Bosak et al., 2001; Soloaga et al., 2003). Phosphorylation at these positions has not been detected in all cell systems. Specific modifications have been identified only in defined cell lines (see chapter 1.3.1). In my experimental model (mouse C3H 10T1/2 cells), phosphorylation at NBD has never been identified. C3H 10T1/2 cells have been widely used to characterise very precisely both background and inducible phosphorylation for many phosphoproteins, including histone H3 and HMGN1 itself (Barratt et al., 1994; Clayton and Mahadevan, 2003; Soloaga et al., 2003; Thomson et al., 1999, 2001). There can be various reasons why NBD phosphorylation has not been detected in C3H 10T1/2 cells. First of all, there might have been detection issues in previous reports. Peptide mapping of radiolabelled (^{32}P) HMGN1 in C3H 10T1/2 cells detected a radiolabelled HMGN1 N-terminal (up to Ser19). Only Ser6ph residues were identified by sequencing (Barratt et al., 1994). In this study Ser19ph could have been missed due to the lack of resolution of the sequencing methods performed (Pogna et al., 2010). However it is clear that Ser23ph was totally absent (Barratt et al., 1994). Secondly, it is possible that phosphorylation on NBD is prominent during cell mitosis and that signal is occasionally carried over from mitotic phosphorylation of HMGN1, as similarly happens for H3 mitotic phosphorylation when quiescing is not complete (Dyson et al., 2005b). This might be the case for most of the phospho-NBD results available to date that are based on rapidly dividing cancer cell models such as HeLa and K562 (human erythroleukemia) cells, intrinsically showing disrupted control of the cell cycle, or mouse embryonic cells, that might insufficiently respond to quiescing procedures (Lim et al., 2004; Louie et al., 2000;

Prymakowska-Bosak et al., 2001, 2002; Ueda et al., 2008). HMGN1Ser6ph is not detectable during mitosis in C3H 10T1/2 cells (Dyson et al., 2005b).

3.1.2 Sub-cellular localization of HMGN1

Previous studies of HMGN1 localization during mitosis suggested that phosphorylation could influence subcellular localization of HMGN1 during IE gene induction (Louie et al., 2000; Prymakowska-Bosak et al., 2001, 2002). Mitotic phosphorylation of the NBD promotes the exclusion of HMGN1 from chromatin and from the nucleus (Prymakowska-Bosak et al., 2001, 2002). Similarly, Louie et.al, (2000) reported that phosphorylation induced by TPA or OKA is detected at higher frequency in "cytoplasmic" populations of HMGN1. These data on mitotic trafficking seems however to refer more to the exclusion of HMGN1 from chromatin rather than the nucleus itself. In fact the report from Louie et.al, (2000) relies on a purification method that merely separates a crude nuclear pellet from a soluble fraction that contains cytoplasmic proteins but also nuclear protein that are not strongly associated with the nuclear pellet (Barratt et al., 1994; Louie et al., 2000). Considering that phosphorylation of HMGN1 is expected to decrease the protein affinity for chromatin and that according to FRAP and MS data, about 99% of the protein is associated with chromatin, it is equally possible that phosphorylated HMGN1 loses its association with nuclei during the purification procedure as that phosphorylated HMGN1 is a genuine cytoplasmic protein (Louie et al., 2000; Phair et al., 2004). Since no phosphorylation of HMGN1, apart from Ser6ph, has ever been reported in the C3H 10T1/2 cells nuclear extracts, I investigated the presence of HMGN1 in soluble fractions and in acid-extracted crude nuclear pellets to define HMGN1 PTM profile within these fractions.

3.1.3 Inhibitors of MAPKs

Little is known about the kinase responsible for HMGN1 phosphorylation at the NBD. Several kinases can phosphorylate HMGN1 NBD *in vitro*, including PKA, PKC and MSK1, but no agreement has been reached on the identity of the *in vivo* kinase responsible for NBD phosphorylation (Lim et al., 2004; Prymakowska-Bosak et al., 2001). The concomitant phosphorylation of Ser6 and Ser20-24 suggested by Lim et al (2004) seems to favour the idea that MSKs are the major HMGN1 NBD kinases.

Previous work in the lab has taken advantage of several MAPK inhibitors to dissect the signalling pathways and identify the kinases responsible for HMGN1 and H3 phosphorylation. I have tested the effects on HMGN1 phosphorylation of four inhibitors: U0126, inhibitor of MEK1 activation and activity, SB203580, inhibitor of p38 α/β activity, H89, a relatively non-specific inhibitor of MSK1, PKA, PKB, SGK, AMPK, CHK1 activity and SP600125 inhibiting Aurora B kinase and JNK kinases (Thomson et al., 1999, 2001, Davies et al., 2000; Bain et al., 2003, 2007).

3.2 Results

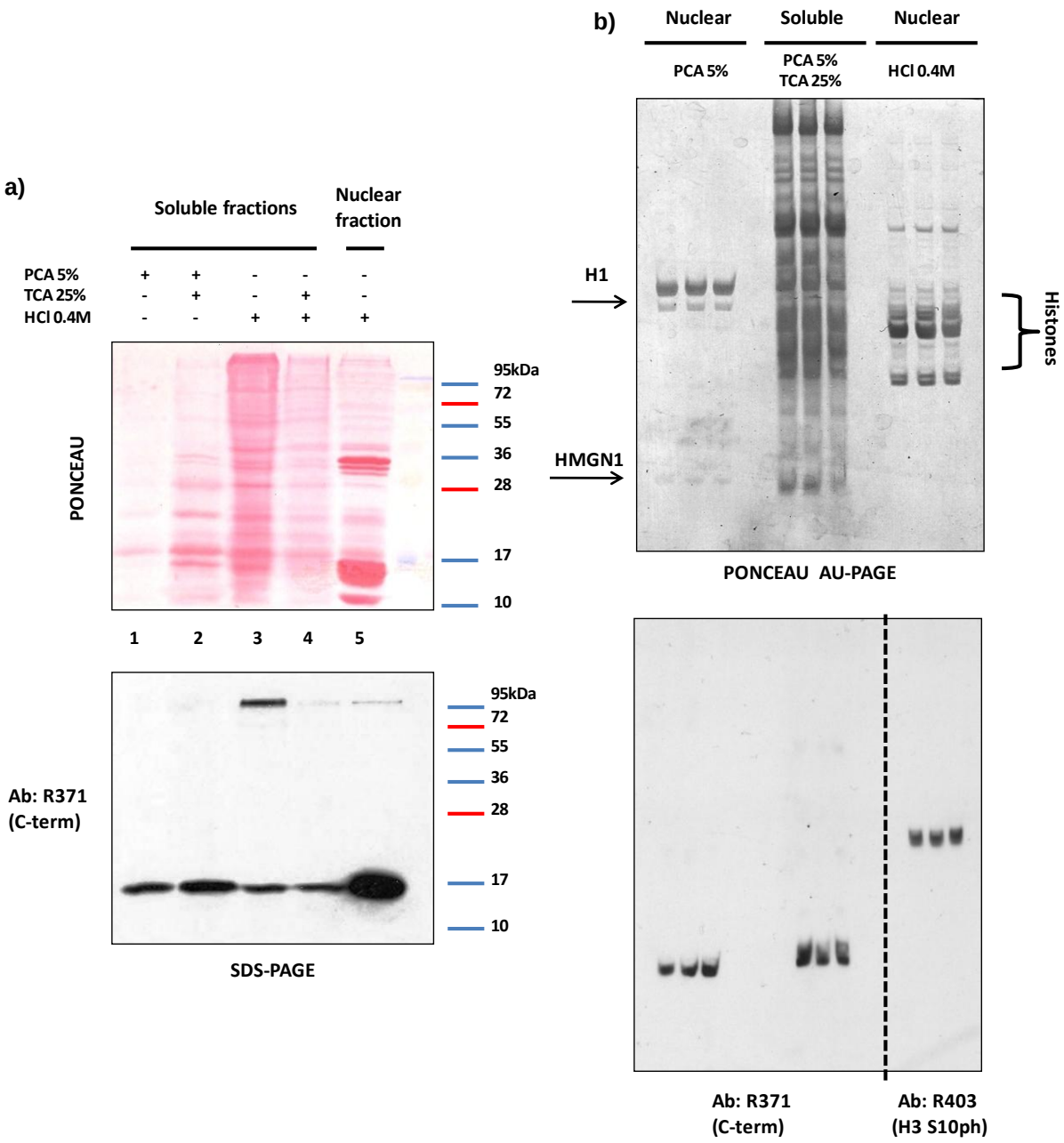
HMGN1 is present in the crude nuclear pellet and in the supernatant of C3H 10T1/2 cells extracts (Figure 12) generated with nuclear protein lysis buffer and extracted with HCl or PCA (see chapter 2.4.1). HMGN1 mobility in SDS-PAGE is similar to a 17kDa marker protein (PageRuler Plus, ThermoScientific) and is located right above histone H3 (Figure 12a). HMGN1 mobility in AU-PAGE is far greater than that of histones (Figure 12b). Soluble fractions were not routinely analysed in the lab previously and there was concern about the right procedure to separate HMGN1 from the total lysate and to precipitate HMGN1 proteins to allow loading on acrylamide gels. It is well known that HMGN1 is soluble in acids like PCA and HCl, but insoluble in 25% TCA solutions. Unfortunately several cytoskeletal proteins share similar properties and TCA precipitation can generate thick pellets of denatured proteins that are refractory to extraction or solubilisation in gel loading buffer (see chapter 2.5.1.1). I have tested four different procedures for the extraction of HMGN1 from soluble fractions in these subfractionations. **(1)** PCA extraction and direct loading on SDS gel after adjustment to 1x Laemmli buffer. **(2)** PCA extraction followed by a TCA precipitation then resuspension of the sample in 4 M urea x1 Laemmli buffer. **(3)** HCl extraction and direct loading on SDS gel after adjustment to 1x Laemmli buffer. **(4)** HCl extraction followed by a TCA precipitation then resuspension of the sample in 4 M urea x1 Laemmli buffer. The procedure providing the best signals in a HMGN1 C-terminal antibody immunoblot is **(2)** a PCA extraction followed by a TCA precipitation (Figure 12a). This procedure takes advantage of the fact that few proteins are soluble in PCA and thus the subsequent TCA precipitation resulted in a pellet completely soluble in gel loading buffer. The amount of soluble fraction loaded was equivalent to one confluent 145 x 22 mm dish, while the nuclear extracts represent 1/10 of such a plate i.e. a 10:1 soluble to pellet ratio is loaded on these gels.

The loading ratio in acrylamide gels between HMGN1 present in the soluble fractions and in the nuclear extracts has been kept constant in all the following experiments. Soluble fractions are very sensitive to experimental variability, mostly because of the low abundance of HMGN1 protein in these fractions. The signal from HMGN1 in the soluble fraction in comparison to the nuclear fraction counterpart is therefore hard to quantify, but a rough valuation is that HMGN1 in nuclear fractions is about 10-15 times the HMGN1 present in the soluble fractions. There was also sporadic loss of soluble fraction signal in single samples in large scale experiments.

To investigate the PTM profile of HMGN1, the phosphorylation of HMGN1 at Ser 6 was induced by EGF, TPA, sub-inhibitory concentration of anisomycin (sAn) and inhibiting concentration of anisomycin (inhibitor of protein synthesis, An). Irradiation of cells with 200 J/m^2 UV light was also used to elicit HMGN1 phosphorylation. To distinguish between HMGN1 modifications occurring in the context of IE gene induction and PTMs dependent on cell cycle regulation, samples presented in section are from quiesced C3H 10T1/2 cells (if not otherwise stated).

Figure 12. **Soluble fraction extraction methods.**

C3H 10 T1/2 cells were quiesced 18 h (o/n) and lysed using nuclear protein lysis buffer. Nuclear [1/10 of a 145 mm dish] and soluble [1x145 mm dish] fractions were separated by centrifugation and then subjected to SDS or AU-PAGE. **a)** Comparison of the yield from four different combinations of acid extractions and precipitations (see methods chapter 2.4.1). **b)** Mobility of HMGN1 from nuclear and soluble fractions in AU-PAGE as compared to histones. Western blots were performed with HMGN1 C-terminal antibody [R371] and histone H3S10ph [R403]. Secondary anti-Rabbit antibody was used for ECL visualisation.



3.2.1.1 Induction of HMGN1Ser6 phosphorylation with TPA, EGF and anisomycin

HMGN1 proteins from quiesced cells run in AU PAGE as a single band, designated band **0**, representing HMGN1 molecules that do not contain any Ser6ph (Figure 13, lane 1). AU-PAGE separates proteins by mass and charge and even a single acetylation or phosphorylation mark is able to generate a shift in the electrophoretic profile. Histone H3 for example is distributed in up to 6 different bands when acetylation is maximal (Crump et al., 2011; Hazzalin and Mahadevan, 2005). The presence of a single population of HMGN1 does not exclude that PTMs are present, but indicates that they are equally distributed among the entire population of HMGN1 molecules. Induction of HMGN1 phosphorylation with EGF and/or anisomycin generates a shift of almost all HMGN1 population from band **0** to band **1** within 60 minutes (Figure 13a, lane 3; 7; 13c, lane 4; 8). This band shift is due to the increase in negative charge on HMGN1 molecules. HMGN1 band **1** is recognized by the Ser6ph-specific antibody thus indicating that it is this phospho-mark that reduces HMGN1 mobility in AU-PAGE. Induction with EGF and anisomycin generates the maximum response, a state previously described as "super-induction" (Chen and Allfrey, 1987; Thomson et al., 1999). Induction of Ser6 phosphorylation with sub-inhibitory concentration of anisomycin is almost as effective as anisomycin stimulation (Figure 13c). Induction with EGF alone generates a lighter response than TPA or sAn. However this response is faster (Figure 14, lane 3) and signal of Ser6ph is detectable after 15min of induction. TPA and sAn induced responses are stronger but appear at later time points (Figure 14, lane 7; 10). It has been previously established that HMGN1Ser6 phosphorylation is part of the response to these agonist (EGF, An, TPA) but only by using AU-PAGE separation is it possible to evaluate the proportion of the total protein involved in the response.

Quantification of HMGN1 bands showed that upon stimulation, HMGN1Ser6 phosphorylation occurs at early time points in about 50% of the total population (Figure 13d). Prolonged stimulation and super-induction is able to elicit a shift of 80-90% HMGN1. Notably both soluble and nuclear fractions respond to the stimuli and with similar intensity (Figure 13d).

Figure 13. **EGF and anisomycin stimulation promotes HMGN1 splitting into two distinct populations.** C3H 10 T1/2 cells were quiesced 18 h (o/n) and then **a) b)** stimulated for the stated time with EGF (50 ng/ml) and anisomycin (An 10 µg/ml, or sAn 25 ng/ml) [EGF_An, EGF_sAn]. Cells were lysed with nuclear protein lysis buffer and nuclear [1/10 of a 145 mm dish] and soluble [1x145 mm dish] fractions were separated by centrifugation and then subjected to AU-PAGE. Western blots were performed with HMGN1 C-terminal antibody [R371] and HMGN1Ser6ph antibody [S520]. Secondary anti-Rabbit [for R371] and anti-Sheep [for S520] antibodies were used for ECL visualisation. **c)** Induction with inhibitory (An) or sub inhibitory (sAn) concentration of anisomycin. **d)** Quantification of HMGN1 bands signal with Image J software.

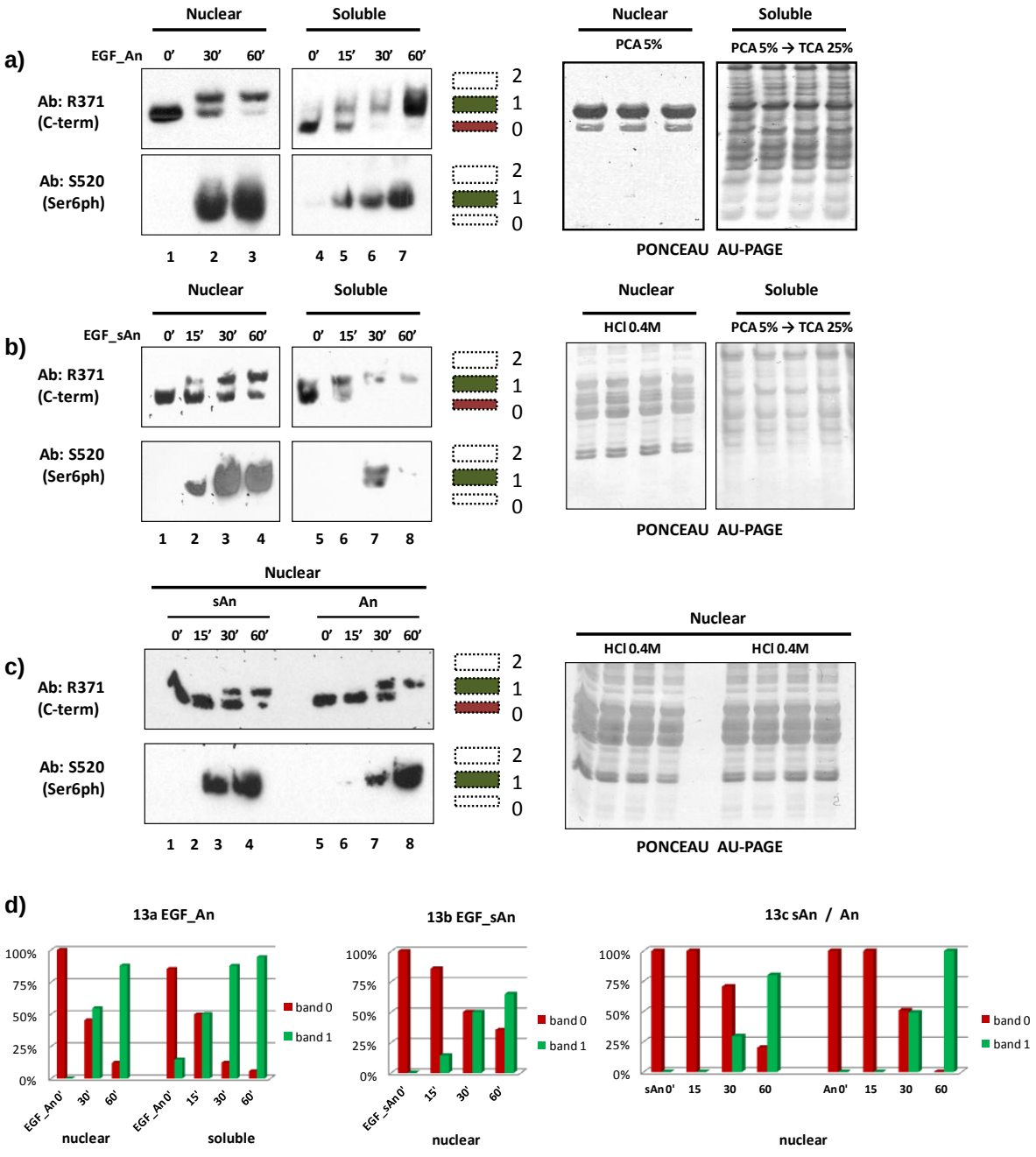
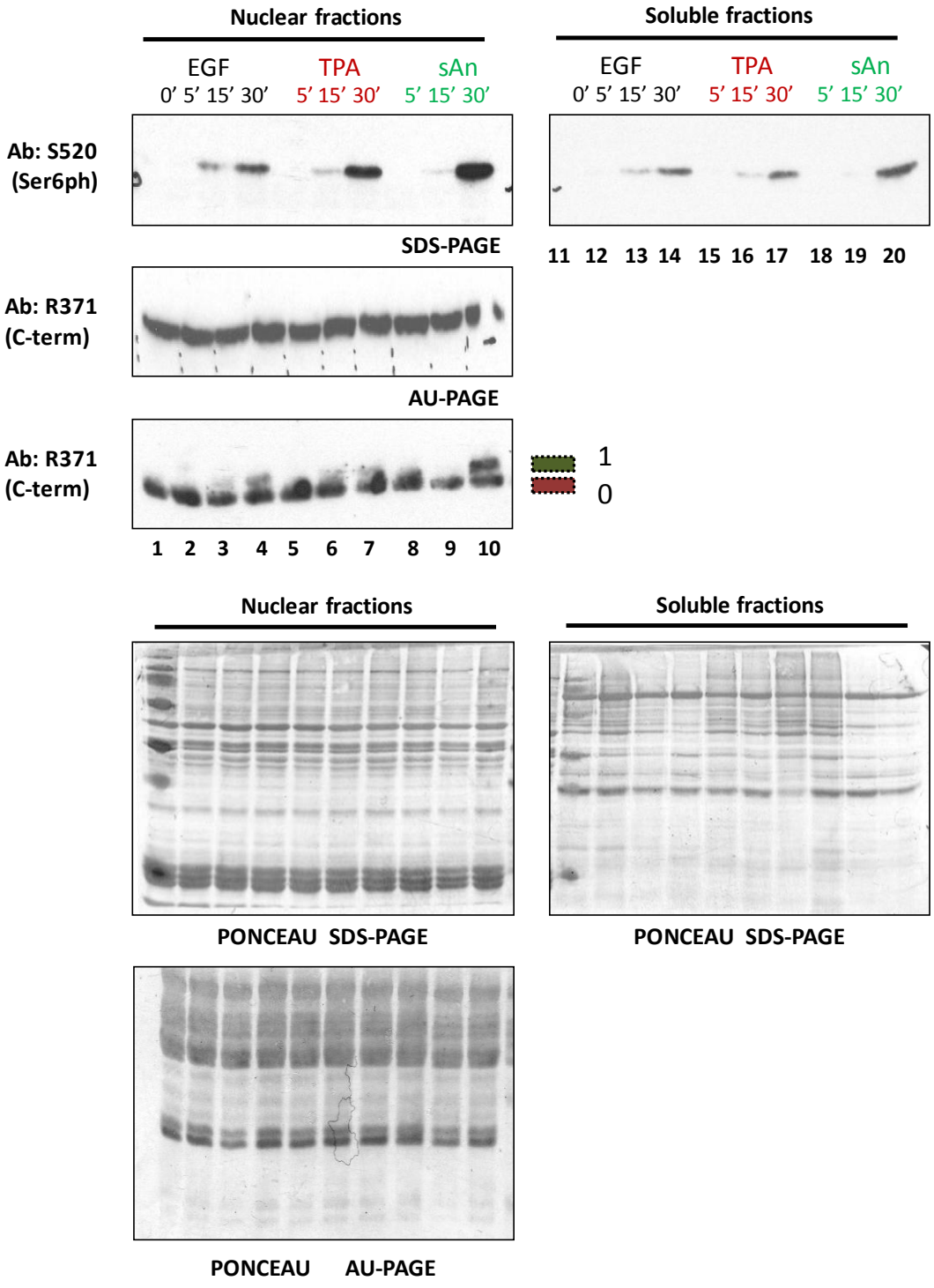


Figure 14. **HMGN1Ser6 phosphorylation is induced by various stimuli.**

C3H 10 T1/2 cells were quiesced 18 h (o/n) and then stimulated for the stated time with EGF (50 ng/ml), TPA (100 nM) or sAn (25 ng/ml). Cells were lysed with nuclear protein lysis buffer and nuclear [1/10 of a 145 mm dish] fractions were subjected to SDS or AU-PAGE. Western blots were performed with HMGN1 C-terminal antibody [R371] and HMGN1Ser6ph antibody [S520]. Secondary anti-Rabbit [for R371] and anti-Sheep [for S520] antibodies were used for ECL visualisation.



3.2.1.2 Effects of TSA treatment and UV irradiation on HMGN1

Treatment with the deacetylase inhibitor TSA perturbs the balance between acetylation and deacetylation and enhances modification of histone H3 (Crump et al., 2011; Hazzalin and Mahadevan, 2005). The TSA effect is evident especially when acetylation is highly dynamic. Previous results indicates that HMGN1 acetylation is not sensitive to TSA treatments alone but it is possible that acetylation of HMGN1 might be coupled to its phosphorylation. It has been proposed that acetylation occurring in the NBD reduces HMGN1 binding to nucleosome (Bergel et al., 2000). Thus, treatment with TSA was performed prior to stimulation to induce a state where inducible acetylation might be maintained stably to produce a shift in AU-PAGE of modified HMGN1 molecules. This showed that TSA does not produce any effect on HMGN1 separation in AU-PAGE (Figure 15a, b). Soluble and nuclear fractions respond to the treatment similarly and no additional band shift is visible. The only detectable alteration is due to Ser6ph elicited by the EGF and anisomycin stimulation. Interestingly, soluble fractions from control samples show a much higher background modification state in some of these experiments. HMGN1 profiles are already split into band **0** and band **1** (Figure 15b, lane 6, 15c, lane 7) indicating the presence of either a phosphorylation or an acetylation mark. Since both TSA-treated and untreated samples (Figure 15b, lane 6 and 8) have similar profiles, it is most likely that acetylation is not responsible for the shift. To clarify which modification causes HMGN1 retardation in AU-PAGE, samples were treated with calf intestinal phosphatases (CIP), a very aggressive enzyme that can virtually deplete all phosphorylation from HMGN1 protein (Figure 15c, lane 6). Such treatment removes the shift to band **1** in stimulated samples. This indicates that the difference in charge between HMGN1 samples running at position band **0** and band **1** is due to phosphorylation. Interestingly, the modifications present in soluble fraction samples are also reduced to band **0** after phosphatase

treatment, implying that phosphorylation is also the PTM occurring on HMGN1 molecules in these fractions (Figure 15c, lane 8; 10; 12).

Irradiation with UV activates p38 MAPK signalling and induces HMGN1 and H3 phosphorylation. It is unclear whether and to what extent UV treatment in these experiments causes DNA damage. MAPK activation and DNA damage response appears mostly independent (Cano and Mahadevan, 1995). HMGN1 depletion affects DNA damage response in *Hmgn1 -/-* knockout cells (Birger et al., 2003), although HMGN1 role in histone acetylation regulation appears to be responsible for the DNA damage sensitive phenotype (Kim et al., 2009). The soluble fraction of HMGN1 was never analysed previously and it was also unclear which portion of the total HMGN1 population responds to UV stress. The dynamics of UV response was similar to An induction. Stimulation for 30min with An elicits the same signal observed 30 min after UV irradiation (Figure 16, lane 2; 4).and leads to a shift from band **0** to band **1** (Figure 16, lane 4-6, 10-12). The soluble fraction HMGN1Ser6ph responds to UV irradiation with similar dynamics. Note that the amount of UV irradiation is identical for each time point.

Figure 15. **HMGN1 profile of PTMs is sensitive to CIP treatment but not to TSA treatment.**

C3H 10 T1/2 cells were quiesced 18 h (o/n) and then incubated for 4 h with 500 ng/ml of TSA and then stimulated for **a)** 30-60 min or **b)** 15-30 min with EGF (50 ng/ml) and An (10 µg/ml) [EGF_An]. Cells were lysed with nuclear protein lysis buffer and nuclear [1/10 of a 145 mm dish] and soluble [1x145 mm dish] fractions were separated by centrifugation and then subjected to AU-PAGE. **c)** Nuclear and soluble fraction samples were split and half of each sample was treated with CIP phosphatase and then subjected to AU-PAGE. Western blots were performed with HMGN1 C-terminal antibody [R371] and HMGN1Ser6ph antibody [S520]. Secondary anti-Rabbit [for R371] and anti-Sheep [for S520] antibodies were used for ECL visualisation.

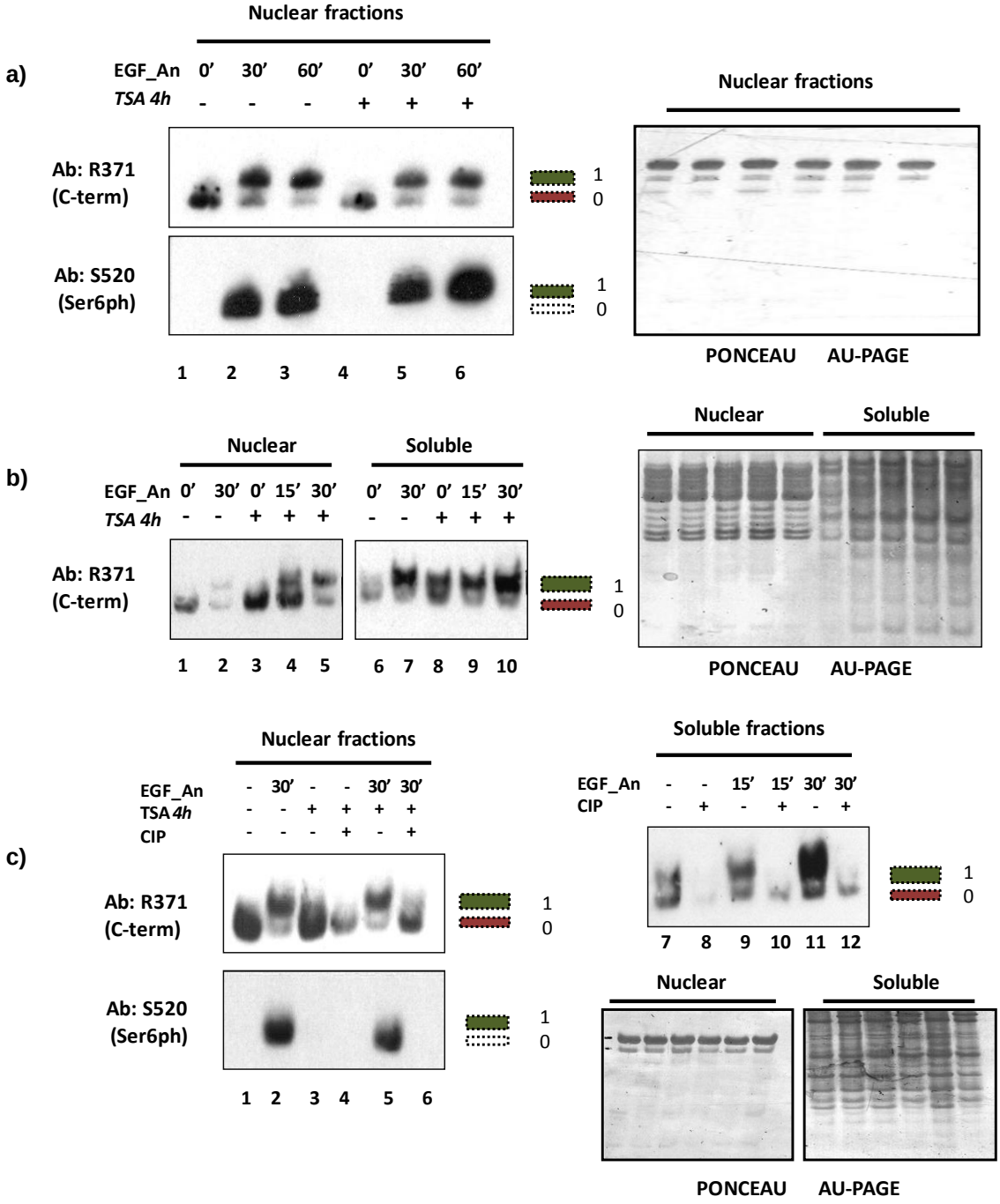
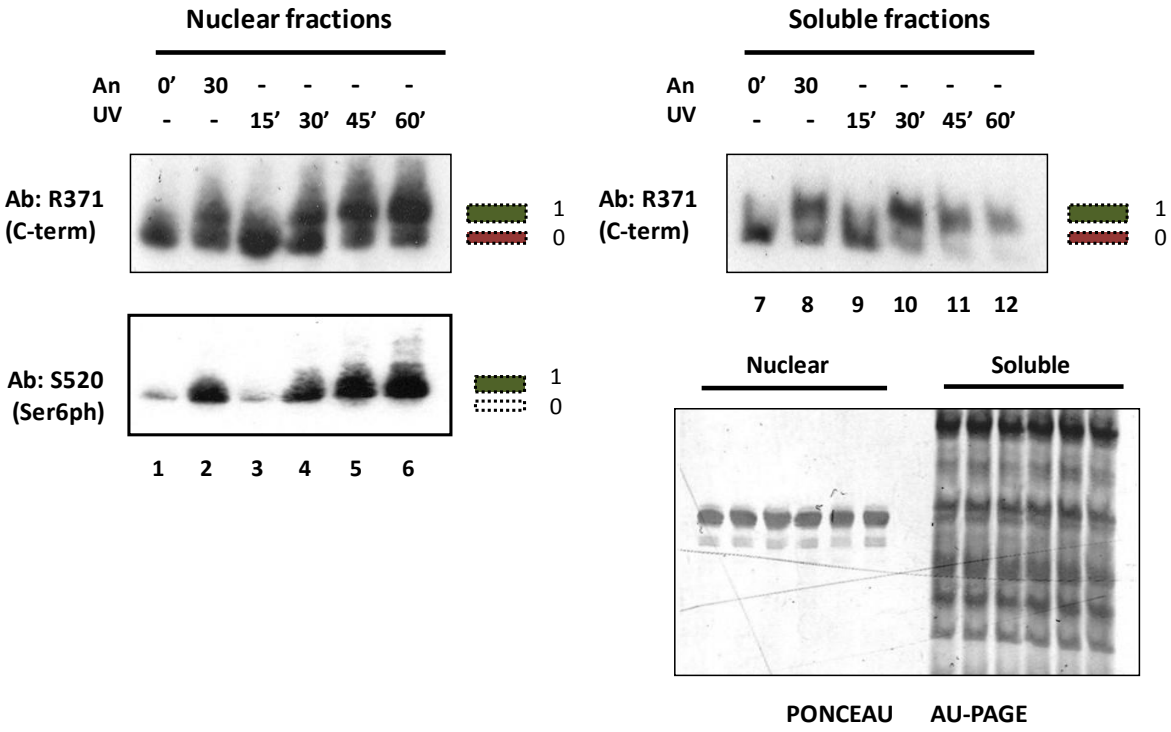


Figure 16. **HMGN1Ser6 phosphorylation is induced by UV irradiation.**

C3H 10 T1/2 cells were quiesced 18 h (o/n) and then irradiated with 200 J/m^2 UV-light. Thereafter cells were left at 37°C for the stated time. Cells were lysed with nuclear protein lysis buffer. Non-stimulated samples and cells treated for 30 min with An (10 µg/ml) were included as controls. Nuclear [1/10 of a 145 mm dish] and soluble [1x145 mm dish] fractions were separated by centrifugation and then subjected to AU-PAGE. Western blots were performed with HMGN1 C-terminal antibody [R371] and HMGN1Ser6ph antibody [S520]. Secondary anti-Rabbit [for R371] and anti-Sheep [for S520] antibodies were used for ECL visualisation.



3.2.2 Additional HMGN1 PTMs in soluble fractions

HMGN1 proteins present in both soluble and nuclear fractions seem to be targeted by the same enzymes. The only modification that is influenced by stimulation is phosphorylation at Ser6. Phosphatase treatments and TSA treatments proved that HMGN1 is not dynamically acetylated and that proteins present in band **0** are most likely unmodified. There are two major explanations for the absence of modification at NBD residues. It is plausible that only HMGN1 molecules directly associated with IE genes sequences are targeted for modification in the NBD. Considering the proportion of IE genes in the total genome, it is likely that such a small population of HMGN1 remains undetectable with immunoblotting. Alternatively it is possible that NBD phosphorylation is more strictly confined to mitosis in C3H 10T1/2 cells. To be able to test these hypothesis, I investigated the AU-PAGE profiles of HMGN1 proteins from nuclear extracts and soluble fractions from stimulated C3H 10T1/2 cells that did not undergo the standard quiescing procedure (see chapter 2.3), and I arranged a collaboration for analysing HMGN1 samples with mass spectrometry (see chapter 3.2.2).

3.2.2.1 Soluble fraction from non-quiesced cells contains HMGN1 molecules with additional PTMs

Proper quiescing of cell cycle-related signalling is achieved by deprivation of serum in cultured cells. Very few cell types efficiently respond to quiescing procedures and many laboratories invested a considerate effort identifying the ideal model for studying specific cell signalling. C3H 10T1/2 cells are among the most manageable system in the cell signalling field with low background phosphorylation of several proteins, for example H3S10ph or H3S28ph (Dyson et al., 2005a; Soloaga et al., 2003). Conversely, the mere absence of quiescing treatments does not imply that all cells on culturing plates are dividing while

experiments are conducted. For a rough estimate of the dividing population, it has been reported that about 20% of the C3H 10T1/2 cell are in late G2 phase in a non-confluent culturing population (Dyson et al., 2005b). Contact inhibition might also play a role and cells used in the following experiment were treated and lysed before reaching complete confluence. At this stage of my investigation it was not clear if the subfractionation by centrifugation of lysate was indeed selecting distinct populations of HMGN1 molecules. All the evidence collected did not exclude the possibility that HMGN1 in soluble fractions may represent a leakage from the cell nucleus, affecting proteins like HMGN1 due to their more unstable interaction with chromatin.

Nuclear fractions from C3H 10T1/2 cells did not differ in dynamics of induction, nor separation in AU-PAGE, from the standard quiesced C3H 10T1/2 cells extracts (Figure 17, *shorter exposure* lane 1-4, 13-16). However, soluble fractions from *non-quiesced* C3H 10T1/2 cells showed a remarkably higher shift of HMGN1 to band **2**. This shift is even more evident taking into consideration the signal from HMGN1Ser6ph antibody (Figure 17, lane 10-13). The majority of HMGN1Ser6ph runs within band **2** upon stimulation (Figure 17, lane 10-12) and background shift between band **0-1** is clearer in the control soluble fraction from *non-quiesced* cells (Figure 17, lane 9). These data indicate that a post-translational modification is present in soluble fractions but it is not detectable in the comparable nuclear extracts. This mark is sensitive to quiescing procedures and does not interfere with Ser6 phosphorylation. Notably only the fractions of HMGN1 carrying this mark are susceptible to Ser6 phosphorylation (all Ser6ph signal is detectable in band **2,** Figure 17, lane 10-12). Since all induced samples detected by Ser6ph-specific antibody run at the same position in AU-PAGE it is unlikely that this additional mark on the soluble fractions from *non-quiesced* cells is responsive to the stimulation of MAPKs.

Due to the presence of serum in *non-quiesced* cell's medium during the stimulation and the preparation of lysate, I tested the effect of serum on *non-quiesced* cells (Figure 18). The induction with EGF and An was not influenced by the presence or absence of serum. This means that the quiescing medium was not altering the distribution of HMGN1 proteins in AU-PAGE. As observed previously, HMGN1 in soluble fractions from *non-quiesced* cells was carrying an additional mark that was reducing HMGN1 mobility in AU-PAGE (Figure 18). Nuclear fractions were divided in two populations running at position band **0-1**, while soluble fractions were occupying band **0-1-2**, exclusively band **0-1** in control cells and mostly band **2** in induced samples (Figure 18, lane 10-18).

These unexpected results from *non-quiesced* cells strongly suggest that HMGN1 contained in the soluble fraction is different from HMGN1 present in nuclear extracts. The additional mark on soluble fractions HMGN1 from *non-quiesced* cells is probably a NBD phosphorylation event. No acetylation event has ever been directly linked to mitotic HMGN1 and phosphorylation at Ser6-20-24 has been reported in different studies (Louie et al., 2000; Prymakowska-Bosak et al., 2001). My attempts to treat these soluble fractions with alkaline phosphatase or to analyse their PTMs profile with mass spectrometry have failed to bring conclusive results. Thus, it remains uncertain which modification is causing the additional HMGN1 shift in soluble fractions of *non-quiesced* cells. Nevertheless it is intriguing to note the total absence of Ser6ph in the control sample and in band **1** of soluble fractions of *non-quiesced* cells. Ser6ph appears to be a strictly inducible modification and it is imposed only on those HMGN1 soluble fractions of *non-quiesced* cells that already carry the unknown modification reported here.

Figure 17. **Additional HMGN1 mark in *non-quiesced* cell soluble fractions.**

C3H 10 T1/2 *non-quiesced* cells were incubated for 30 min with EGF (50 ng/ml) and An (10 µg/ml) [EGF_An] in the presence of growing medium (DMEM, 10% FCS). Cells were lysed with nuclear protein lysis buffer and nuclear [1/10 of a 145 mm dish] and soluble [1x145 mm dish] fractions were separated by centrifugation and then subjected to AU-PAGE. Western blots were performed with HMGN1 C-terminal antibody [R371] and HMGN1Ser6ph antibody [S520]. Secondary anti-Rabbit [for R371] and anti-Sheep [for S520] antibodies were used for ECL visualisation.

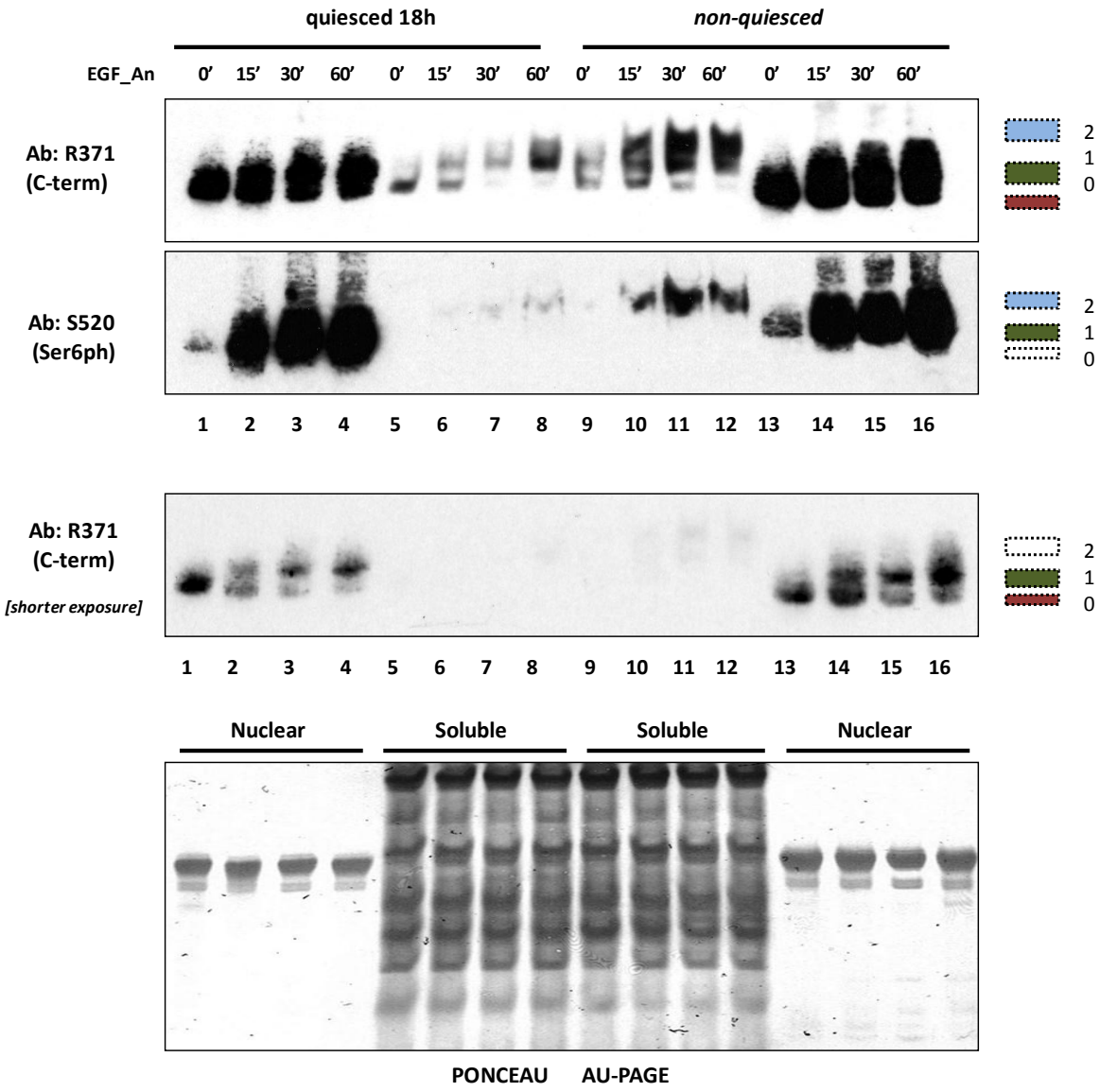
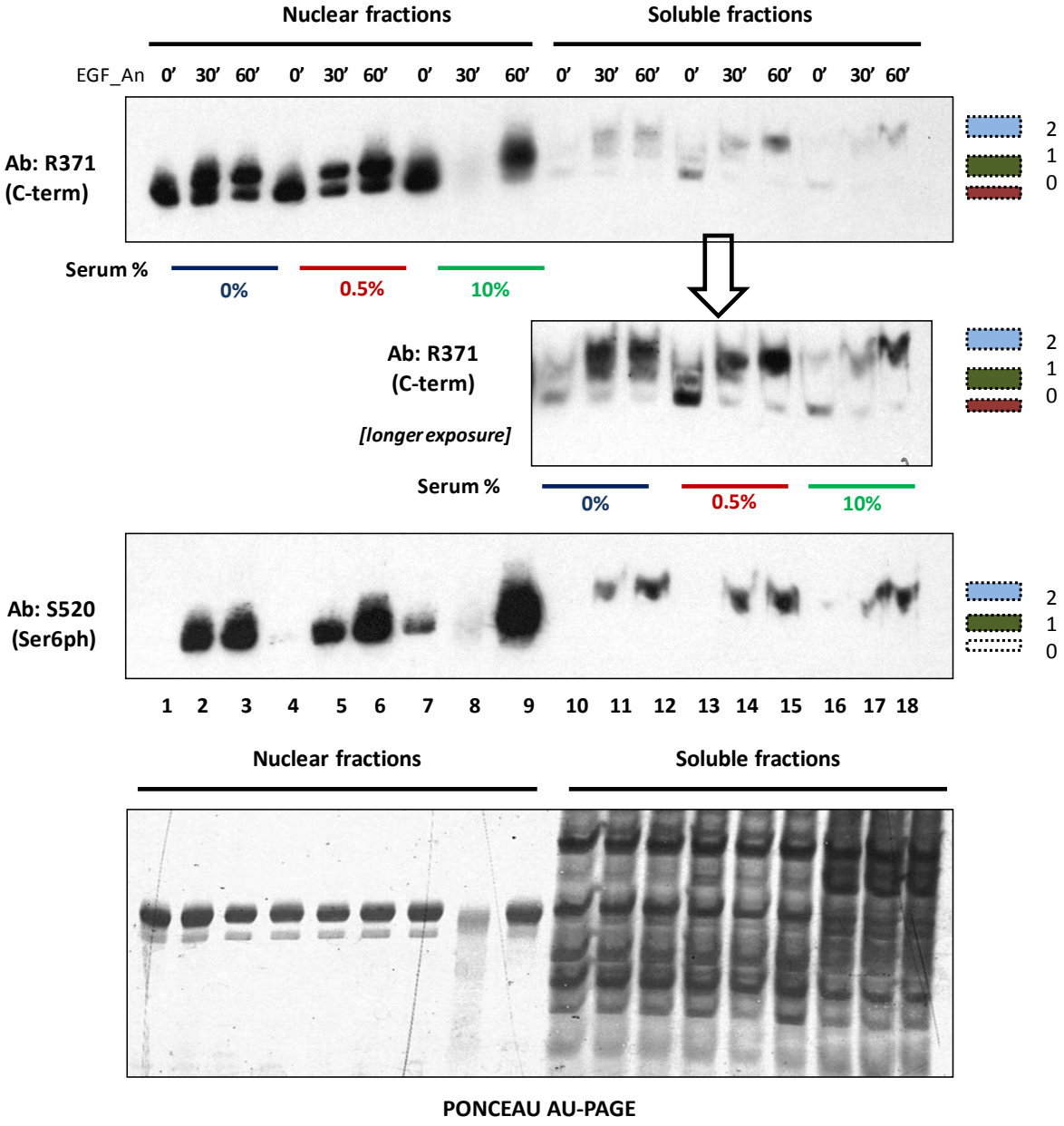


Figure 18. **Additional HMGN1 mark in soluble fractions: effect of fetal calf serum.**

C3H 10 T1/2 *non-quiesced* cells were incubated for 30 min with EGF (50 ng/ml) and An (10 µg/ml) [EGF_An] in the presence of either 0%, 0.5%, 10% FCS in the growth medium. Cells were lysed with nuclear protein lysis buffer and nuclear [1/10 of a 145 mm dish] and soluble [1x145 mm dish] fractions were separated by centrifugation and then subjected to AU-PAGE. Western blots were performed with HMGN1 C-terminal antibody [R371] and HMGN1Ser6ph antibody [S520]. Secondary anti-Rabbit [for R371] and anti-Sheep [for S520] antibodies were used for ECL visualisation. Arrow indicates signal achieved by longer exposure of film during the same experiment.



3.2.3 MAPK inhibitors; U0126, SB203580, H89, SP600128

Having identified a qualitatively different sub-population of HMGN1 carrying additional marks and potentially representing a partitioning of the HMGN1 population, I investigated the identity of the MAPK that lay upstream of and led to HMGN1 phosphorylation in soluble fractions at Ser6. Signalling to HMGN1 involves two principal pathways: the MEK/ERK pathway strongly activated by EGF and/or TPA, and the MKK/p38 pathway activated by An and UV light. I tested the effect of U0126, an inhibitor of MEK1 activation and activity, SB203580, an inhibitor of p38 α/β activity and H89, an inhibitor of MSK1 activity.
Mitotic phosphorylation of Histone H3 is mediated by Aurora B kinase (Crosio et al., 2002). Considering that H3 and HMGN1 are already targeted by the same kinase in the nucleosomal response, I investigated the possibility that the additional mark in *non-quiesced* cells might be imposed by Aurora B kinase. Thus, I treated cells with SP600125, an inhibitor that blocks Aurora B and JNK kinase activity.

3.2.3.1 Inhibition of the MAPK pathway

Kinase inhibitors are an important chemical tool to indentify the kinases responsible for mediating phosphorylation. The interpretation of their effects depends on the selectivity of a specific inhibitor and the dynamics of a phosphorylation mark. The development of more selective inhibitors is a major effort within therapeutic drug design and considerable effort has been spent in the last decade to define the specificity of most common kinase inhibitors (Bain et al., 2003, 2007; Davies et al., 2000). The effects of kinase inhibition are strongly dependent on the dynamics of phosphorylation. Those marks that are imposed by signal-dependent kinases and continuously removed by phosphatases are directly affected by kinase inhibition.

In the other hand, cell cycle mark might be less sensitive to kinase inhibition if phosphatases and kinases are not continuously targeting the modified residue.

Treatment with U0126 was not sufficient to abrogate EGF signal (Figure 19, lane 5, 6). A slight reduction of the shift to band **1** was observed and also Ser6ph signal decreased in EGF_An induced samples (Figure 19, lane 15; 16). SB203580 was able to abrogate An induction, restricting all HMGN1 to band **0**, but not EGF induction (Figure 19, lane 8; 14; 18). H89 treatment abrogated even the "super-induction" induced by EGF_An (Figure 19, lane18). Soluble and nuclear fractions responded similarly to inhibition of kinases (Figure 20).

Inhibition of Ser6ph in the nuclear extracts was mirrored by their soluble fraction counterpart, the abrogation of MAPK signalling causing redistribution of all HMGN1 band **1** population to the faster band **0** (Figure 20). Treatment with SP600125 (JNK and Aurora B kinase inhibitor) did not influence the activation of the MAPK pathway inducing HMGN1Ser6ph, nor the shift to band **2** in soluble fractions sampled from *non-quiesced* cells (Figure 21).

These results support the findings of the CIP experiment and confirmed that HMGN1 in soluble and nuclear fraction carry a phosphorylation mark that is detectable as a clear shift in AU-PAGE. Ser6ph is imposed by MAPK signalling in both fractions and it is most likely that MSKs are responsible of both events. The lack of effect of SP600128 inhibitor suggests that the additional HMGN1 shift in soluble fractions is independent of Aurora B or JNK kinases activities.

Figure 19. **Treatment with MAPK inhibitors: nuclear fractions.**

C3H 10 T1/2 cells were quiesced 18 h (o/n) and then incubated for 30 min with either U0126 (10 µM), SB203580 (10 µM) or H89 (10 µM) and then stimulated for 30 min with EGF (50 ng/ml) or An (10 µg/ml). Cells were lysed with nuclear protein lysis buffer and nuclear [1/10 of a 145 mm dish] fractions were subjected to AU-PAGE. Western blots were performed with HMGN1 C-terminal antibody [R371] and HMGN1Ser6ph antibody [S520]. Secondary anti-Rabbit [for R371] and anti-Sheep [for S520] antibodies were used for ECL visualisation.

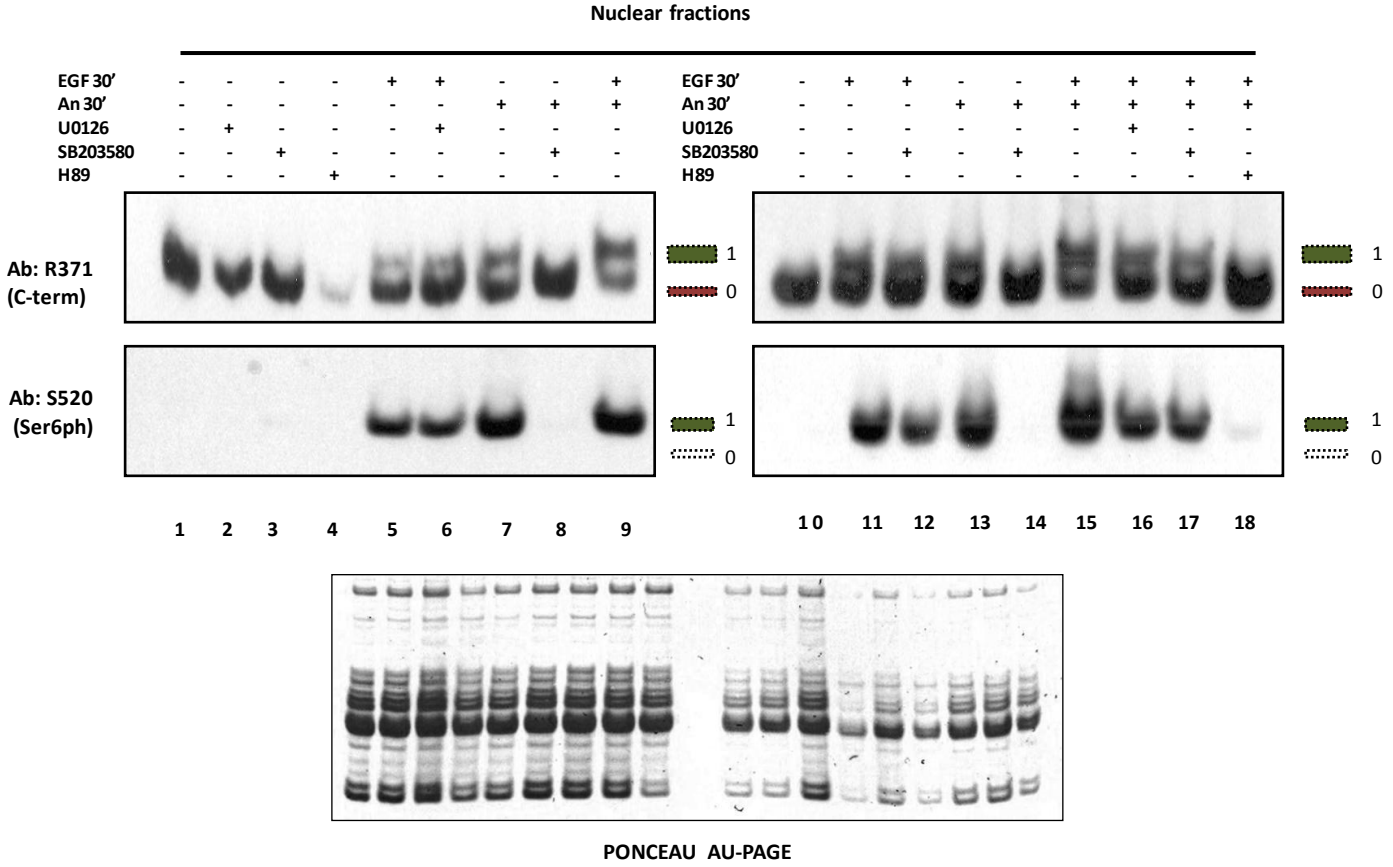


Figure 20. **Treatment with MAPK inhibitors: similar results in soluble and nuclear fractions.**

C3H 10 T1/2 cells were quiesced 18 h (o/n) and then incubated for 30 min with either U0126 (10 µM) or SB203580 (10 µM) and then stimulated for 30 min with EGF (50 ng/ml) or An (10 µg/ml). Cells were lysed with nuclear protein lysis buffer and nuclear [1/10 of a 145 mm dish] and soluble [1x145 mm dish] fractions were separated by centrifugation and then subjected to AU-PAGE. Western blots were performed with HMGN1 C-terminal antibody [R371] and HMGN1Ser6ph antibody [S520]. Secondary anti-Rabbit [for R371] and anti-Sheep [for S520] antibodies were used for ECL visualisation.

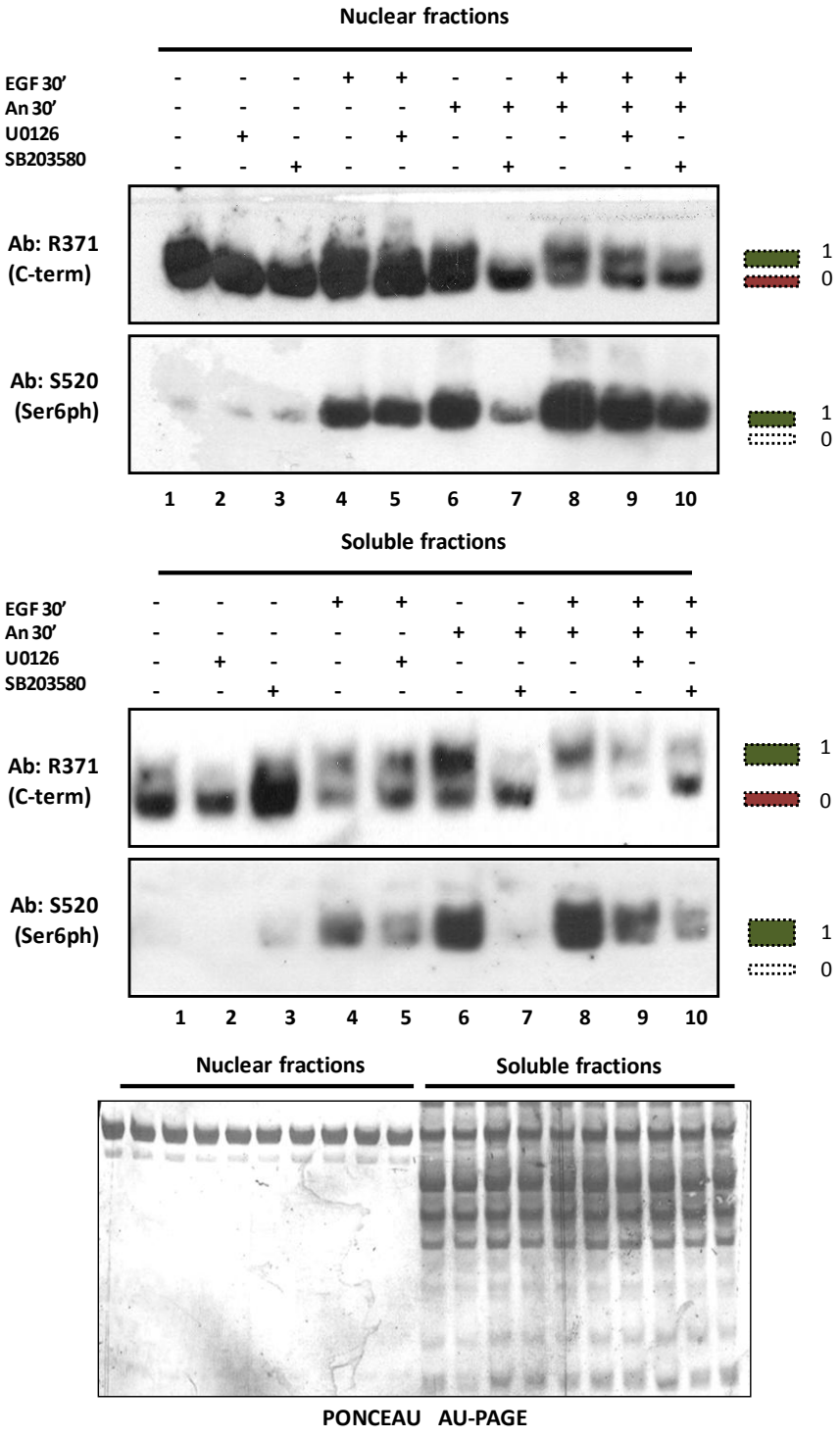
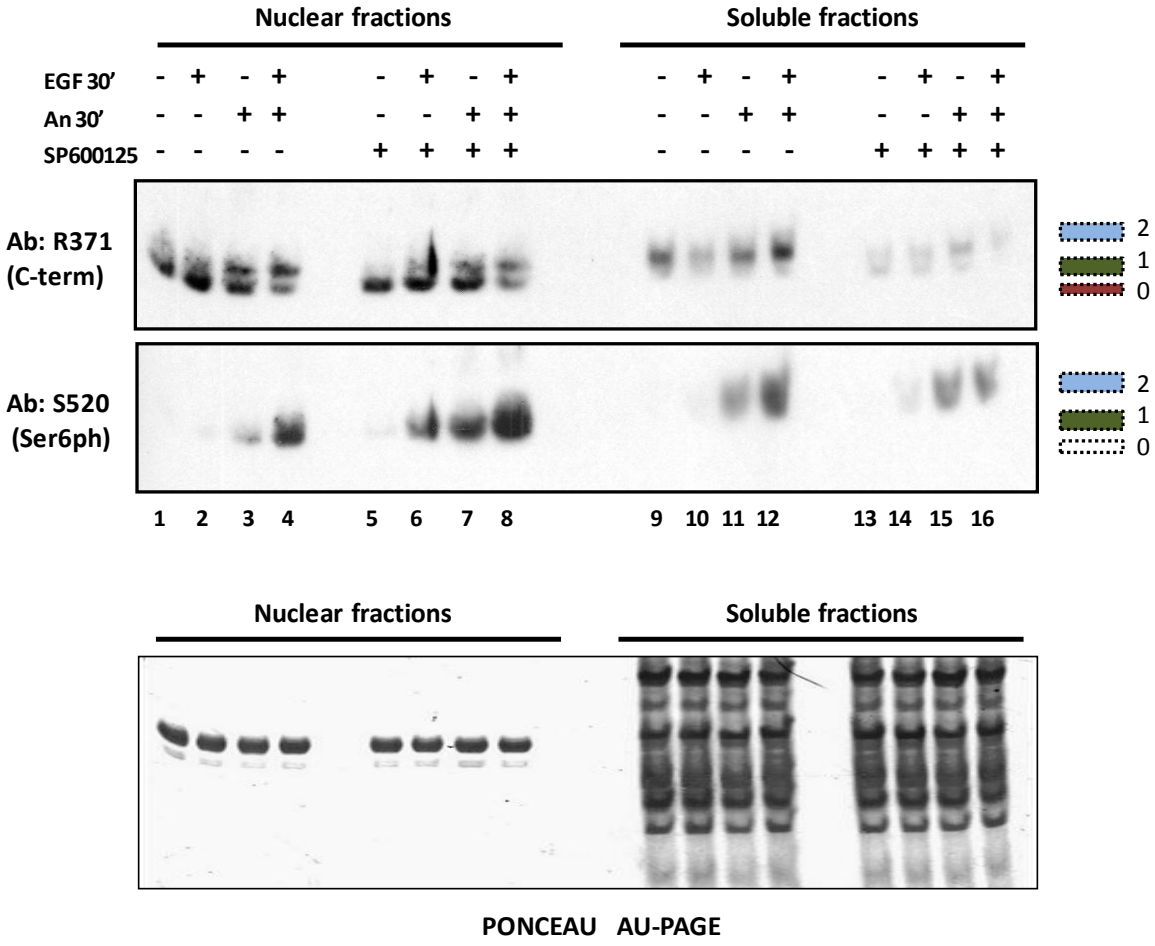


Figure 21. **SP600125 treatment does not perturb HMGN1 shift in nuclear or soluble fractions.**
C3H 10 T1/2 *non-quiesced* cells were incubated for 30 min with SP600125 (1 µM) and then stimulated for 30 min with EGF (50 ng/ml) or An (10 µg/ml). Cells were lysed with nuclear protein lysis buffer and nuclear [1/10 of a 145 mm dish] and soluble [1x145 mm dish] fractions were separated by centrifugation and then subjected to AU-PAGE. Western blots were performed with HMGN1 C-terminal antibody [R371] and HMGN1Ser6ph antibody [S520]. Secondary anti-Rabbit [for R371] and anti-Sheep [for S520] antibodies were used for ECL visualisation.



3.2.4 Immunofluorescence of C3H 10T1/2 cells with HMGN1 antibodies

Interpretation of the analysis of cellular extracts is limited by the disruptive procedures that can affect the distribution between fractions. The use of different salt concentrations and detergents allow separation of specific fractions of various proteins but the attribution of a cytoplasmic origin for a particular protein is inappropriate without direct investigation in intact cells. Especially in the case of HMGN1 in soluble fractions, it is necessary to attempt a parallel characterization of the distribution *in vivo* by indirect-immunofluorescence (i-IF) (chapter 2.7). Previous reports have shown that more than 95% of the total cellular HMGN1 molecules are present in the nucleus but also reported that upon induction of MAPK, the HMGN1 cytoplasmic fraction increases to about 10-15% (Louie et al., 2000; Phair et al., 2004). I discussed in the introduction the limitations in these reports (chapter 1.3.3) and decided to approach the matter independently.

3.2.4.1 HMGN1 localization during the nucleosomal response is nuclear

Immunofluorescence with HMGN1 C-terminal and HMGN1Ser6ph antibodies detected signal within the nucleus of C3H 10T1/2 cells, but not in the cytoplasm (Figure 22). Several dilutions of antibodies and various detection times were tested to optimize detection. In no case was it possible to detect a clear cytoplasmic signal that was distinguishable from background signal. Induction with EGF and An increases the signal from the HMGN1Ser6ph antibody that co-localizes with the signal from the HMGN1 C-terminal antibody (Figure 23). Images have been collected with the same exposure time. The settings for the time-course experiment were chosen based on optimal detection time for induced samples. Stimulation with MAPK agonists, like TPA or An, increases the signal intensity from the HMGN1Ser6ph

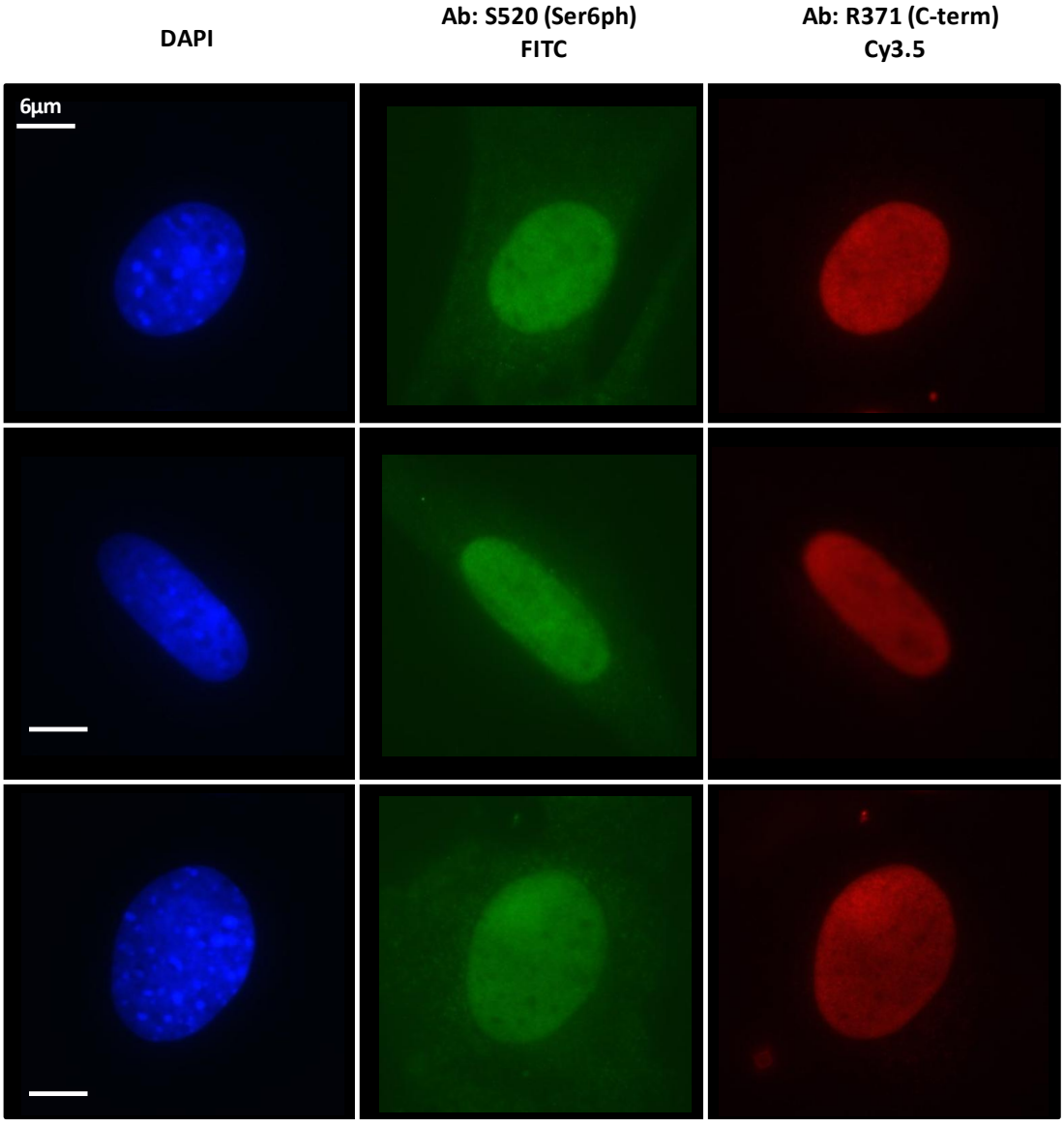
antibody as expected from previous SDS-AU-PAGE results (Figure 24). Settings for time-course experiment in this case were chosen based on optimal detection for the latest time points of the induction test.

The level of background signal from non-specific or auto-fluorescence of cellular material, when i-IF was performed in the presence of specific antibody competing peptides, is shown in Figure 25.

Overall, under the different conditions analyzed, no signal of HMGN1 was detected in the cytoplasm. I cannot exclude that a population of HMGN1 might exist in the cytoplasm that is so small that it remains undetectable by i-IF. Considering the heterogeneity of a cell population it is conceivable that the additional mark observed in soluble fractions from *non-quiesced* cells is present on nuclear HMGN1 that was extracted from the proportion of cells undergoing mitosis and therefore having HMGN1 already detached from mitotic chromatin. This mark is present only in the less tightly-associated HMGN1 and may affect chromatin binding.

Figure 22. **Immunofluorescence: distribution of HMGN1 in control cells.**

C3H 10T1/2 cells were quiesced 18 h (o/n) and then fixed with paraformaldehyde. I-IF was performed with HMGN1 C-terminal antibody [R371] and HMGN1Ser6ph antibody [S520]. Secondary anti-Rabbit Cy3.5 (RED) [for R371] and anti-Sheep FITC (GREEN) [for S520] antibodies were used to visualise fluorescence.



CTRL

Figure 23. **Immunofluorescence: distribution of HMGN1 upon stimulation.**

C3H 10T1/2 cells were quiesced 18 h (o/n) and then treated with EGF (50 ng/ml) and An (10 µg/ml) [EGF_An] for the indicated time. I-IF was performed with HMGN1 C-terminal antibody [R371] and HMGN1Ser6ph antibody [S520]. Secondary anti-Rabbit Cy3.5 (RED) [for R371] and anti-Sheep FITC (GREEN) [for S520] antibodies were used to visualise fluorescence.

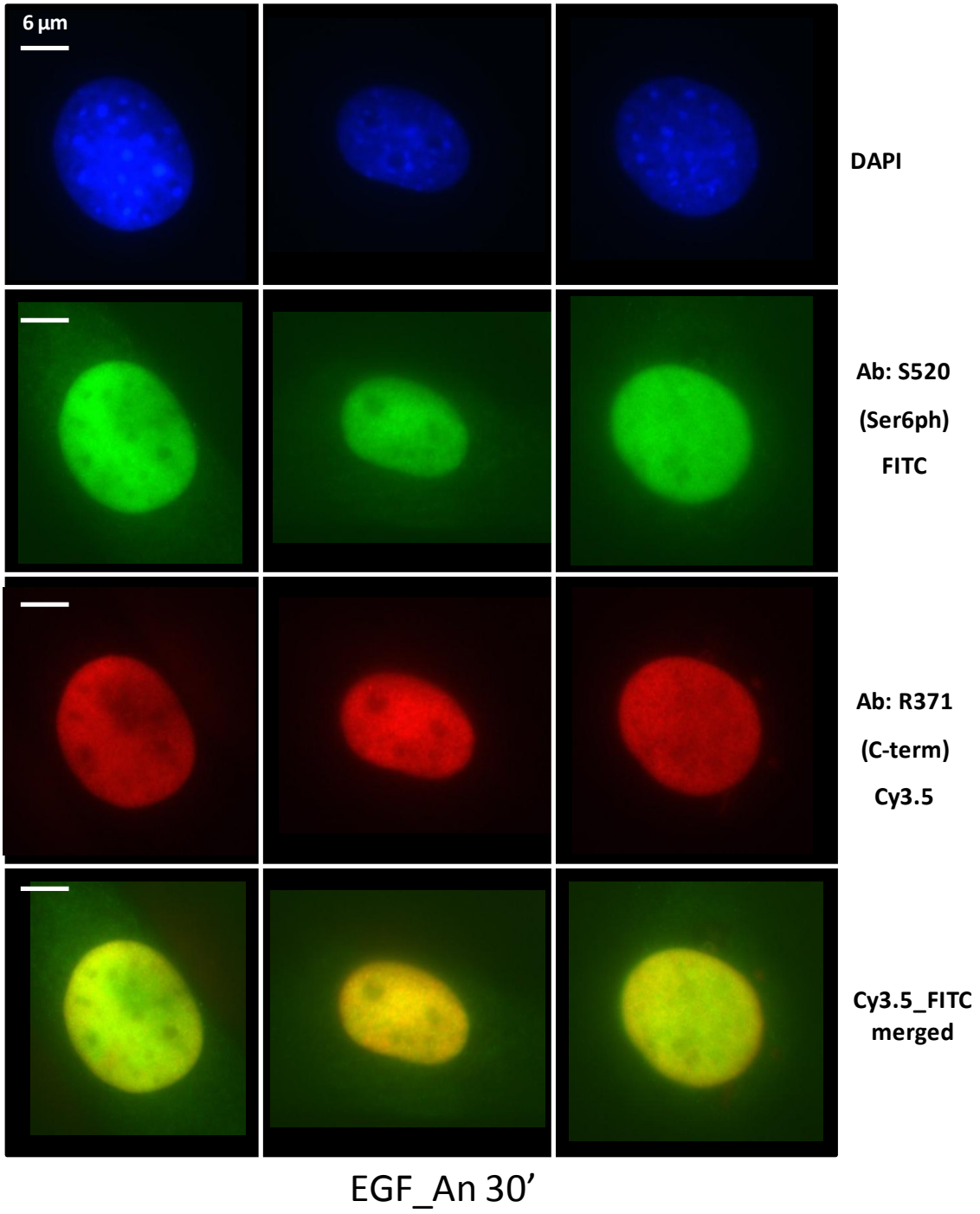


Figure 24. **Immunofluorescence: MAPK induction.**

C3H 10T1/2 cells were quiesced 18 h (o/n) and then treated with TPA (100 nM) and An (10 µg/ml) for the indicated time. I-IF was performed with HMGN1 C-terminal antibody [R371] and HMGN1Ser6ph antibody [S520]. Secondary anti-Rabbit Cy3.5 (RED) [for R371] and anti-Sheep FITC (GREEN) [for S520] antibodies were used to visualise fluorescence.

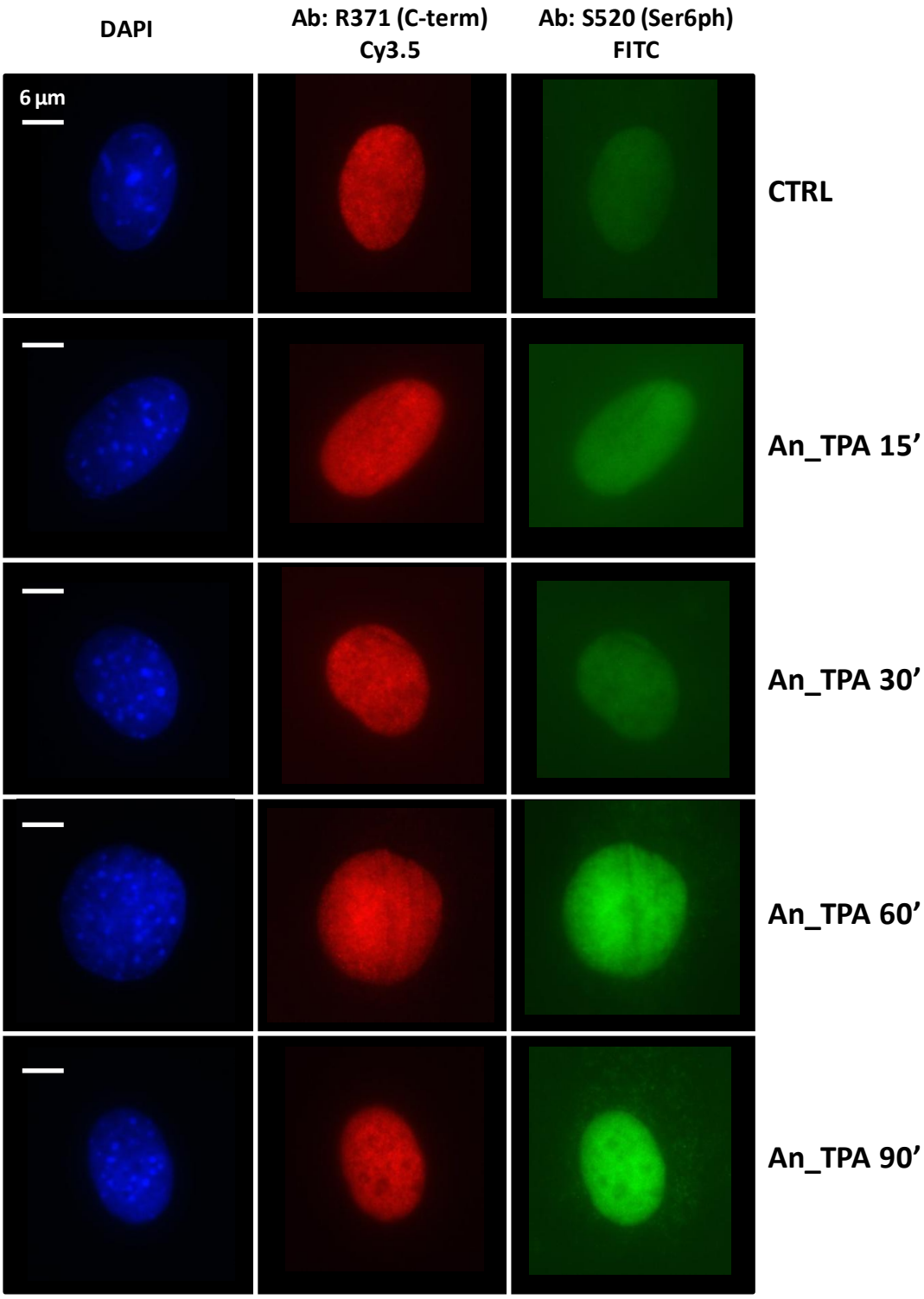
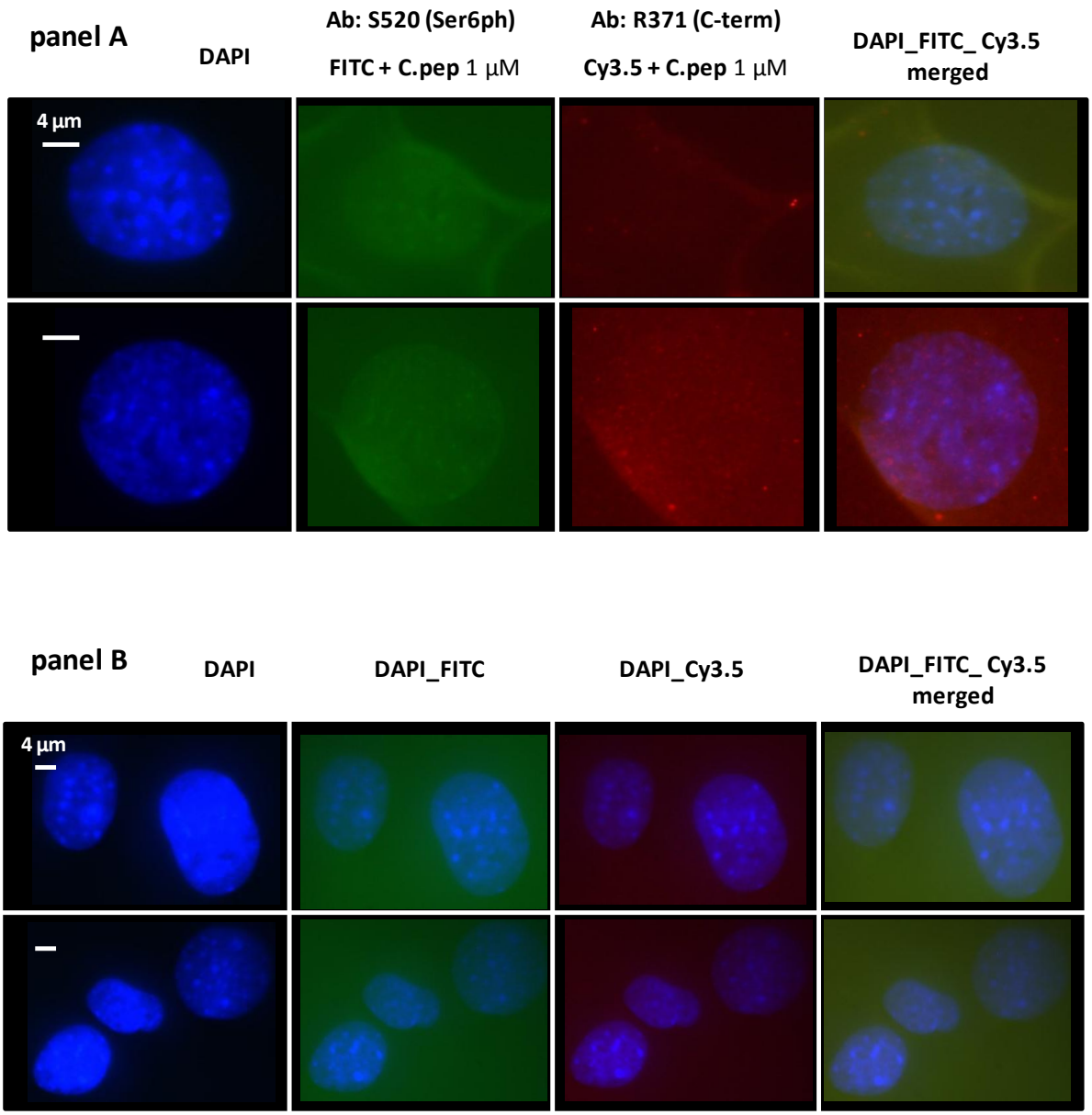


Figure 25. **Immunofluorescence: peptide competition and background signals.**

Panel **(A)**, i-IF with HMGN1 C-terminal antibody [R371] and HMGN1Ser6ph antibody [S520] competed with specific peptides containing epitopes sequences. Secondary anti-Rabbit Cy3.5 (RED) [for R371] and anti-Sheep FITC (GREEN) [for S520] were used to visualise fluorescence. Panel **(B)**, i-IF without HMGN1 C-terminal and Ser6ph-specific antibodies. Background signal is emerging from either unspecific binding of secondary antibody, anti-Rabbit Cy3.5 (RED) and anti-Sheep FITC (GREEN), or autofluorescence.



3.2.5 Mass spectrometry of HMGN1 protein from nuclear extracts

To overcome limitations of immunoblotting and AU-PAGE analysis in unequivocal identification of the complete modification profiles of HMGN1, a collaboration was started for the analysis of HMGN1 proteins using mass spectrometry (MS; with Dr. Benedikt Kessler and Dr Mukram MacKeen, (Nuffield Department of Clinical Medicine, Oxford). Samples were prepared (see chapter 2.11.1) from both nuclear and soluble fractions. Unfortunately, recovery from soluble fractions was scarce and the identification of peptides was limited to a small portion of the HMGN1 C-terminal. Therefore, only data obtained from nuclear fractions is presented.

3.2.5.1 HMGN1 is subject to multiple post-translational modifications

Mass spectrometry is based on the trajectory of small peptides, flying across magnetic fields, which is dependent upon the ratio between their charge and mass (m/z). MS of whole proteins is less sensitive than peptide MS and therefore proteins are degraded enzymatically to peptides by various enzymes, improving subsequent peptide identification (Aebersold and Mann, 2003) (see methods 2.11.1). Protein digestion differently influences the recovery of specific sequences and optimization of HMGN1 MS analysis required the use of Lys-C, Arg-C and Trypsin enzymes. Lys-C-digested samples gave the best recovery of HMGN1 peptides. HMGN1 peptides were detected covering almost all the amino acid sequence of the protein (table 3.1). Not every MS run yielded the same extent of protein sequence. Thus, to quantify HMGN1 PTMs, I have chosen the set of data that obtained the largest coverage and the best scores among the MS analyses performed. This set has been analysed with the Orbi-Trap MS and identified 78% of the amino acid sequence. Best recovery of peptides was from the C-

terminal of HMGN1, while peptides from the N-terminal and NBD were less abundant (Figure 26). Detection of the HMGN1Ser6ph peptide was higher in stimulated samples, but also detectable in control quiesced samples. The ratio of Ser6ph peptides detected was dramatically increasing during stimulation, leaving only 20% of the total in the unmodified state (Figure 27). Phosphorylation at the NBD was identified at both Ser19 and Ser23 in different samples. The most frequently detected mark was Ser23ph and the most abundant within the quantified runs is Ser19ph (Figure 27). The detection of di-phosphorylated Ser19-23ph NBD was scarce and overall phosphorylated NBD was never more than 3% of the total detected NBD. Regarding the quantification of the modified and unmodified peaks, it is important to consider that the data presented is a sum of the relative abundance of similar peptides partially covering the same sequence. The addition or absence of charge influences the detection of a peptide; for example, N-term peptides containing un-modified Ser6 are detected much better in peptides shorter than the most abundant Ser6ph peptide. Thus, direct comparison of peptides of identical length could bias the analysis in favour of those peptides having a more detectable mass/charge ratio (m/z). Table 3.1 lists all detected peptides and the relative Mascot score (see chapter 2.11.3) to allow evaluation of the identity assignment of each of the quantified peptides. In fact a much higher selection of peptides was found covering the sequence around Ser6 and these peptides showed stimulus-dependent phosphorylation, while NBD peptides are mostly detected un-modified (table 3.1). Phosphorylation was detected at Ser83-86-94 within the C-terminal (Figure 27). Best score peptides carry either Ser86ph or Ser94ph, but peptides carrying both phosphorylations have been detected. Phosphorylation at the NBD and at the C-terminal did not increase dramatically upon induction and their abundance in control samples is very similar to the abundance in induced samples (Figure 27).

The results of MS analysis of nuclear fraction HMGN1 support the data from AU-PAGE and immunoblotting. Induction with sAn or EGF_An generates a division into two distinct populations visually identified in AU-PAGE and characterized by the presence of Ser6ph. Any other additional mark on HMGN1 remains unchanged during induction. While being almost completely absent in control samples in AU-PAGE and immunoblotting, Ser6ph is detected by MS in about 5%of total HMGN1 in control samples.

Figure 26. **Most abundant HMGN1 peptides identified by mass spectrometry.**

Ions count of modified and unmodified HMGN1 peptides. Peptides were detected by an Orbi-Trap mass spectrometer and data were analysed by Progenesis LC-MS software. Data from peptides covering the same sequence were pooled into three distinct categories, N-terminal peptides (from a.a. 1-16), NBD peptides (a.a. 17-36) and C-terminal peptides (a.a. 61-95).

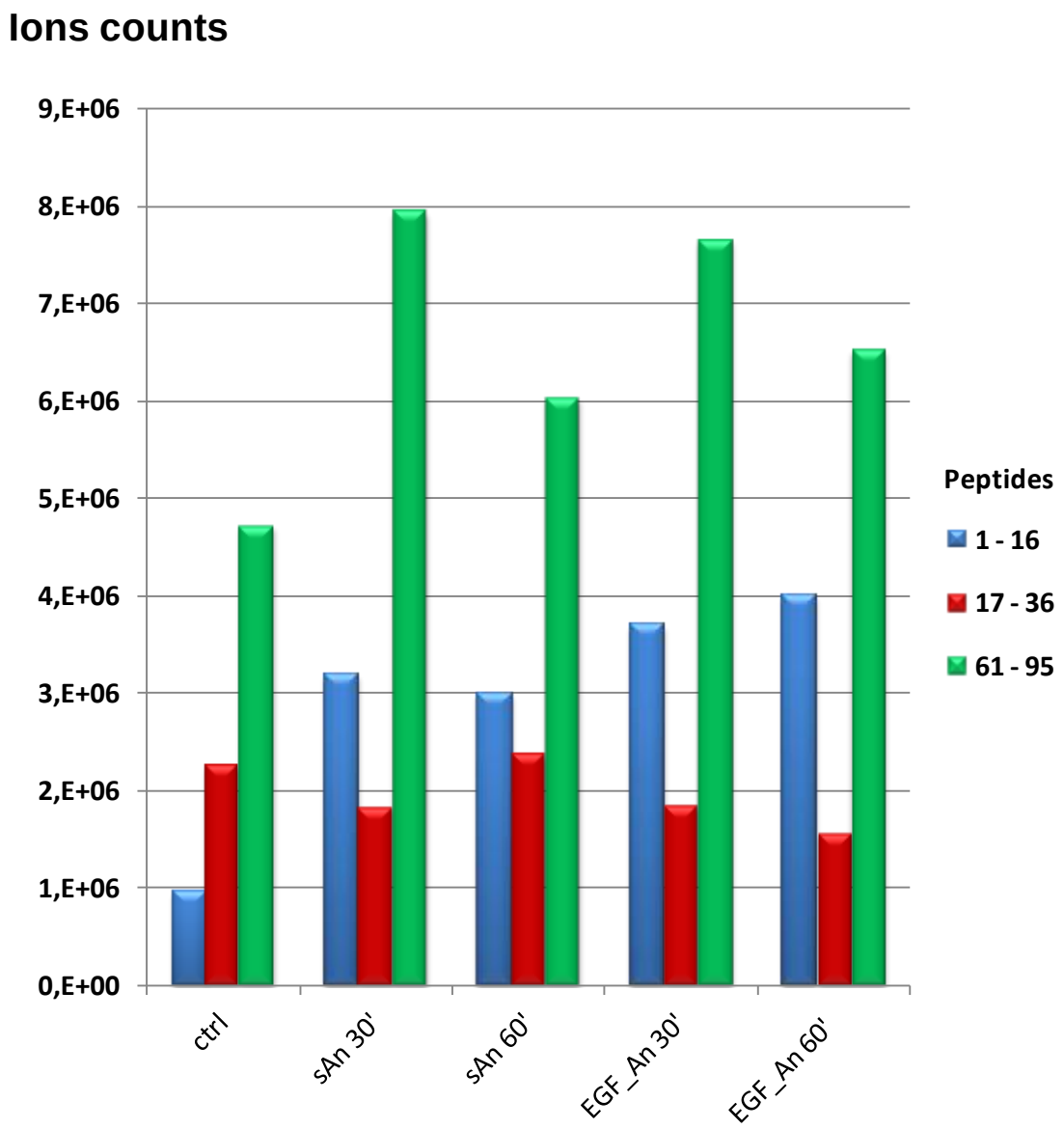


Figure 27. **Phosphorylated HMGN1 peptides identified by mass spectrometry.**

Schematics of the abundance of modified and unmodified HMGN1 peptides. Peptides were detected by an Orbi-Trap mass spectrometer and data were analysed by Progenesis LC-MS software. Data from peptide covering the same sequence were pooled into three distinct categories, N-terminal peptides (from a.a. 1-16), NBD peptides (a.a. 17-36) and C-terminal peptides (a.a. 61-95).

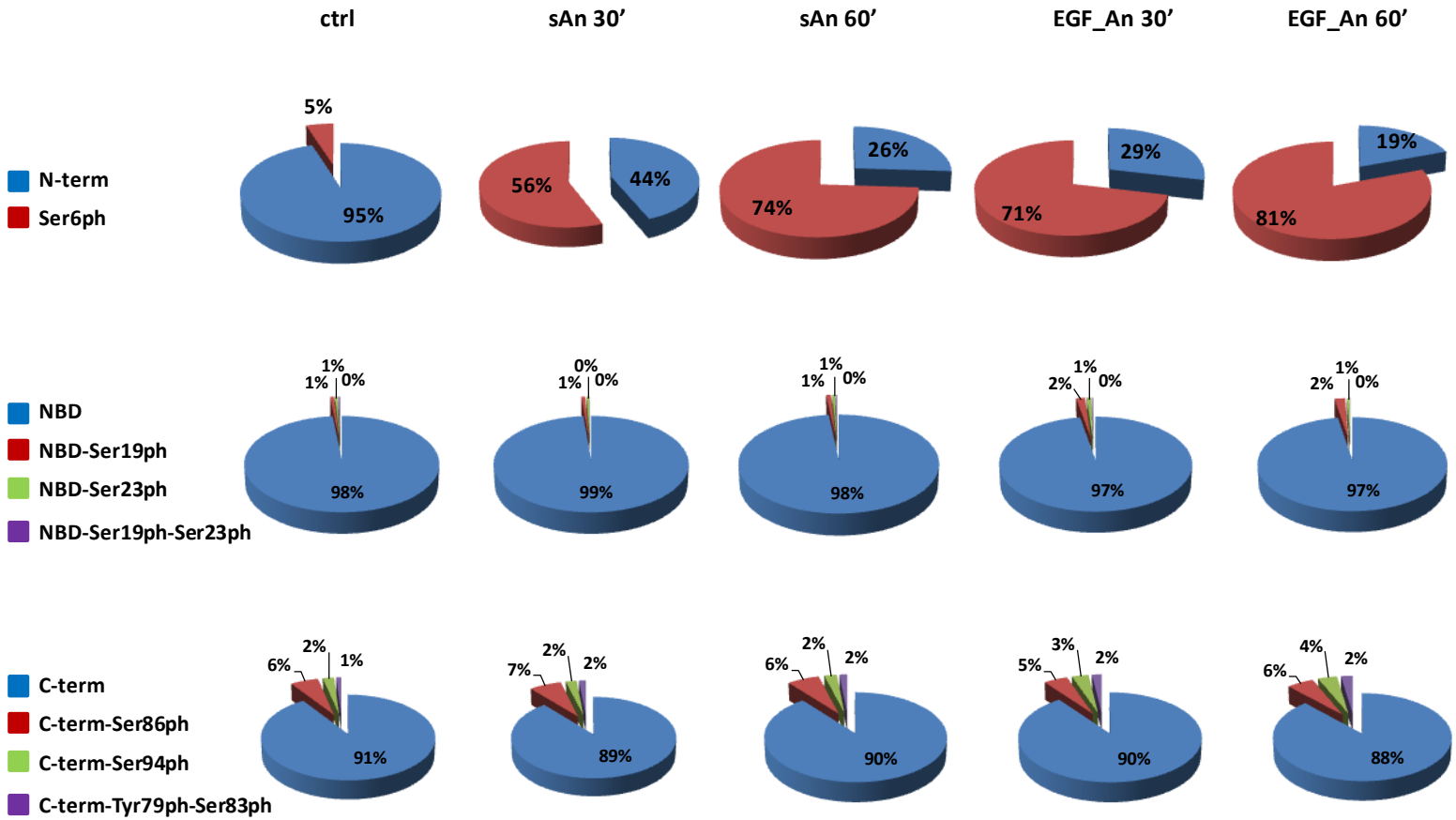


Table 3.1. **Complete list of HMGN1 peptides identified by mass spectrometry.**

Peptide length and sequence of the HMGN1 peptides detected with MS, including data from three different instruments, (Waters) Q-ToF, (Agilent) Ion Trap and Orbi Trap Velos (Thermo Scientific). Mascot ion scores are shown only for peptides included in the quantitative analysis (data acquired with Orbi Trap MS). Data refer to Swiss Uni_prot mouse database HMGN1.

HMGN1, peptides coverage:

1 PK RKVSADGAAKAEPK RRSARLSAKPAPAK VDAKPKKAA GKDKASDKKV
50 QIKGKRGAKG KQADVADQQT TELPAENGET ENQSPASEEE KEAKSD

Peptide length a.a.	Peptide sequence	m/z	Masscott ion score
3 – 16	K.RKVSADGAAKAEPK.R	357,70	25,15
	K.RKVSADGAAKAEPK.R Phospho (ST)	503,26	47,71
		754,38	60,11
		377,70	29,12
5 – 12	K.VSADGAAK.A	-	-
5 – 16	K.VSADGAAKAEPK.R	-	-
17 – 30	K.RRSARLSAKPAPAK.V	377,98	9,85
		503,64	35,32
		302,59	14,28
		302,59	21,36
		754,96	9,66
	K.RRSARLSAKPAPAK.V Phospho (ST)	397,98	13,5
17 – 36	K.RRSARLSAKPAPAKVDAKPK.K	430,26	16,78
		537,58	21,5
		716,43	6,8
	K.RRSARLSAKPAPAKVDAKPK.K Phospho (ST)	446,26	13,52
	K.RRSARLSAKPAPAKVDAKPK.K Phospho (ST) Acetyl (K)	562,52	1,77
	K.RRSARLSAKPAPAKVDAKPK.K Phospho (ST)	-	-
	K.RRSARLSAKPAPAKVDAKPK.K Phospho (ST) Acetyl (K)	620,09	2,77
31 - 36	K.VDAKPK.K Acetyl(K)	350,21	9,31
36 – 47	K.KAAGKDKASDK.K	-	-
44 – 54	K.ASDKKVQIKGK.R Acetyl(K)	429,59	2,04
48 – 52	K.KVQIK.G	-	-
50 – 59	V.QIKGKRGAKG.K	-	-
61 – 90	K.QADVADQQTTELPAENGETENQSPASEEEK.E	1082,81	7,19
		1082,81	95,10
61 – 95	K.QADVADQQTTELPAENGETENQSPASEEEKEAKSD	1259,55	26,27
		1259,55	149,16
		944,92	56,24
	K.QADVADQQTTELPAENGETENQSPASEEEKEAKSD Phospho (ST)	1286,21	55,69
	K.QADVADQQTTELPAENGETENQSPASEEEKEAKSD Phospho (ST)	964,91	47,80
	K.QADVADQQTTELPAENGETENQSPASEEEKEAKSD Phospho (ST)	984,90	8,16

3.3 Discussion

HMGN1 in C3H 10T1/2 cells is almost exclusively a nuclear protein that upon activation of MAPK signalling with several stimuli is subjected to phosphorylation at the Ser6 residue. During mitosis there is the possibility that additional marks are imposed on HMGN1 molecules. HMGN1 samples usually split in two distinct populations during AU-PAGE but in *non quiesced* samples an additional shift has been detected. The additional phosphorylation marks detected by MS analysis at the C-terminal and NBD could be responsible for affecting HMGN1 mobility in AU-PAGE.

The dislocation from chromatin of HMGN1 during mitosis that has been previously reported could explain the increase of signal in the soluble fraction of HMGN1 in *non-quiesced* cells. Moreover, the presence of Ser6ph exclusively on band **2** in these soluble fractions could be the result of the detection limits of AU-PAGE immunoblotting. If HMGN1 in dividing cells is completely excluded by the nucleus and modified, all mitotic HMGN1 molecules would then run at band **1**. Activation of MAPK may promote Ser6ph on these HMGN1 molecules, which will therefore run at band **2** position. The amount of dividing cells in a *non-quiesced* cells experiment has been estimated to be about 20% of total cell population (Dyson et al., 2005a). In these cells all HMGN1 is probably detached from chromatin and therefore purified in the soluble fraction instead of the nuclear fraction. At this point, soluble fractions in *non-quiesced* cell experiments will then be composed for large part of HMGN1 molecules extracted from the 20% of total cell population that is dividing and has no nuclear HMGN1.

Soluble HMGN1 from quiesced cells appears indistinguishable from nuclear HMGN1 in respect of both phosphorylation dynamics and separation in AU-PAGE immunoblotting. The absence of a cytoplasmic population of HMGN1 detectable by i-IF suggests that soluble HMGN1 represents a less stably chromatin-associated fraction of total HMGN1 that is

separated from chromatin during cell lysis. This soluble fraction in living cells could be nuclear in origin.

According to the literature, a very small proportion of HMGN1, approximately 3%, is highly acetylated (Bergel et al., 2000). Such a small population is, as seen for C-terminal and NBD phosphorylation, very elusive and undetectable in immunoblotting analysis on acid-urea gels. The comparison of results from MS analysis and the detection of shifts in AU-PAGE immunoblotting suggest that PTMs present on about 5-10% of the total protein are not distinguishable as a shifted population in AU-PAGE immunoblotting. It is impossible to detect by MS long peptides of HMGN1 that contain Ser6-19-23-83-86-94, thus it is hard to estimate the percentage of HMGN1 molecule carrying both N-terminal and C-terminal phosphorylation or any combined NBD-C/N-terminal phosphorylation. It is then plausible that part of the HMGN1 present at band **1** in AU-PAGE is carrying phosphorylation marks other than Ser6ph. Fortunately, MS analysis clarified that only Ser6 phosphorylation is concomitant with IE gene induction. All the other marks are generally unaffected by treatment and their contribution to IE gene signalling appear negligible. This investigation of HMGN1 PTMs profile is not conclusive; technical difficulties experienced with MS profiling of soluble fractions might be overcome by more recent advances of spectrometry technology and distribution of NBD phosphorylation in dividing C3H 10T1/2 cells requires further investigation. Nevertheless, with respect to signalling and chromatin protein distribution studies, it is evident that the C3H 10T1/2 cell model offers a more defined and efficient system, where it is possible to distinguish events related to IE gene induction by proper quiescing of the cells. In these conditions, stimulation of IE gene is associated with Ser6ph on HMGN1 in a gradually increasing portion of the population (from 5% to 80%). Phosphorylation dynamics depend on the specific stimulus, but the distribution of the protein remains nuclear and chromatin-associated. The studies that sought to ascribe a role to

HMGN1 in chromatin regulation focusing on phosphorylation of NBD residues might have been misleading due to the interpretation of NBD phosphorylation as a genuine stimulus-induced signalling phenomenon. Instead, the results from the C3H 10T1/2 cells model presented in this thesis clearly show that HMGN1 NBD phosphorylation is overall unaffected by stimulation.

Chapter 4. Distribution and dynamics of HMGN1 across *c-fos* and *c-jun*

4.1 Introduction

Chromatin immunoprecipitation (ChIP) analyses of the distribution of HMGN1 across IE genes *c-fos* and *c-jun* and the dynamics of HMGN1 at these loci during their induction are presented in this chapter. Native and cross-linked chromatin was prepared from C3H 10T1/2 mouse fibroblasts and immunoprecipitated with HMGN1 C-terminal antibody [Ab: R371]. Association of HMGN1 with specific histone H3 PTMs was tested by sequential ChIP with H3K9ac- and H3K4me3-specific antibodies. The presence of HMGN1 at housekeeping gene *gapdh* (glyceraldehyde-3-phosphate dehydrogenase) and at silenced gene *hbb* (hemoglobin-beta or β –globin) was investigated to verify association of HMGN1 with genes whose transcription is not influenced by MAPK signalling. DNA purified with HMGN1 C-terminal antibody [Ab: R371] was amplified by real-time PCR using primers specific for *c-fos* and *c-jun*, *gapdh* and *hbb* (table 2.3) (Crump et al., 2011; Edmunds et al., 2008). ChIP data are displayed as percentage of recovery in immunopurified DNA over total DNA present in input samples (% of input).

4.1.1 Distribution of H3 PTMs across *c-fos* and *c-jun*

In 2008, Edmunds et al., published a quantitative and comparative analysis of histone H3 PTMs at *c-fos* and *c-jun* and their dynamics during gene activation. The work included high-resolution maps of the distribution of H3K4me3, H3K36me3, H3K79me2 and H3K9ac. They identified four distinct layers of histone modification across *c-fos* and *c-jun* (Figure 9). The first layer is composed of H3K4me3 across the start site and H3K79me2 and H3K36me3 in the coding region. These modifications are present in quiesced cells as well. The second layer of histone modification comprises dynamic acetylation, specifically at H3K4me3 nucleosomes; this layer is independent of stimulation. The third layer is stimulation-dependent and includes H3 phosphorylation at Ser10 and/or Ser28 which is seen mostly across the start site (Soloaga et al., 2003). The last layer is imposed by the passage of Pol II associated enzymes. Pol II elongation enhances K4me3 and K36me3 levels, respectively at the 5'end and 3'end of the coding region of *c-fos* and *c-jun*. Therefore, the full histone modification pattern is imposed by a combination of epigenetic, dynamic, stimulus-induced and Pol II-dependent enzymes. Taking advantage of the information contained in these maps, I investigated the distribution of HMGN1 at *c-fos* and *c-jun* loci focusing on the possible correlation between the presence of HMGN1 and histone modifications. I also investigated the possibility that HMGN1 might associate preferentially to nucleosomes possessing specific modification profiles.

Histone modification distribution maps were also important as references for the evaluation of the reproducibility of ChIP procedures. ChIPs against histone modifications were performed alongside HMGN1 ChIPs and DNA sequences were analysed and compared to the previous reports.

4.1.2 Predicted interaction of HMGN1 and core histones

According to several reports, HMGN1 interacts with histone tails at different positions. HMGN1 N-terminal domain interacts with histone H2B, whereas the C-terminal domain targets the N-terminal of histone H3 (Trieschmann et al., 1998; Kato et al., 2011; Kugler et al., 2012) (Figure 7). When HMGN1 is associated with nucleosomes, it is capable of modulating H3 phosphorylation and enhancing acetylation of H3K14 (Lim et al., 2004, 2005). Thus, it is plausible that ChIP could identify specific gene loci within IE genes where HMGN1 is enriched or depleted and histone modifications are targeted, despite the lack of preference of HMGN1 for any specific DNA sequence (Shirakawa et al., 2000a). The lack of direct stable interactions between DNA and HMGN1 also suggests that nucleosomes are essential for the purification of HMGN1-bound DNA sequences. Therefore, cross-linking of chromatin was performed to immobilize HMGN1 on nucleosomes and DNA, allowing immunopurification (X-ChIP) of HMGN1-bound DNA sequences.

Non-cross-linked native ChIP (N-ChIP) was also performed to test HMGN1 depletion during immunopurification and to determine the correct amount of antibody required for ChIP of specific amounts of chromatin. Comparison of HMGN1-depleted fractions from N-ChIP and X-ChIP experiments (Appendix, Figure 37) provides an indication of any possible occlusion of HMGN1 antibodies due to cross-linking of chromatin (O'Neill and Turner, 2003).

4.1.3 Discrepancy of ChIP protocols

Several reports of ChIP genomic profiles of chromatin proteins have recently become available, including for HMGN1 (Barski et al., 2007; Jothi et al., 2008; Cuddapah et al., 2009, Cuddapah et al., 2011). Comparing the ChIP protocols used in these experiments and the protocol currently followed in the lab, the most significant difference appears to be the

temperature at which cross-linking of chromatin is performed (Crump et al., 2011; Edmunds et al., 2008). For this reason the cross-linking protocols in the present work were performed at both room temperature and 37°C. Comparison of the profiles obtained with these two procedures is discussed in this chapter and in the appendix.

4.2 Results

Chromatin immunoprecipitations of HMGN1 recovered DNA sequences from IE genes *c-fos* and *c-jun*, housekeeping gene *gapdh* and silenced gene *hbb* (Figures 28-36). IE genes *c-fos* and *c-jun* were chosen as targets of HMGN1 ChIP analysis due to the extensive description of histone PTMs at these genes and the tight control of their expression by MAPK signalling. The *gapdh* and *hbb* gene loci were included in the analysis as examples of a constitutively expressed gene and a silenced gene. ChIP recoveries for histone PTMs (H3K4me3, H3K9ac, H3K36me3, H3K79me2) at these loci are within the expected range for transcribed genes for *gapdh* sequences and as expected, remarkably low for *hbb* gene loci (Figure 9) (Edmunds et al., 2008).

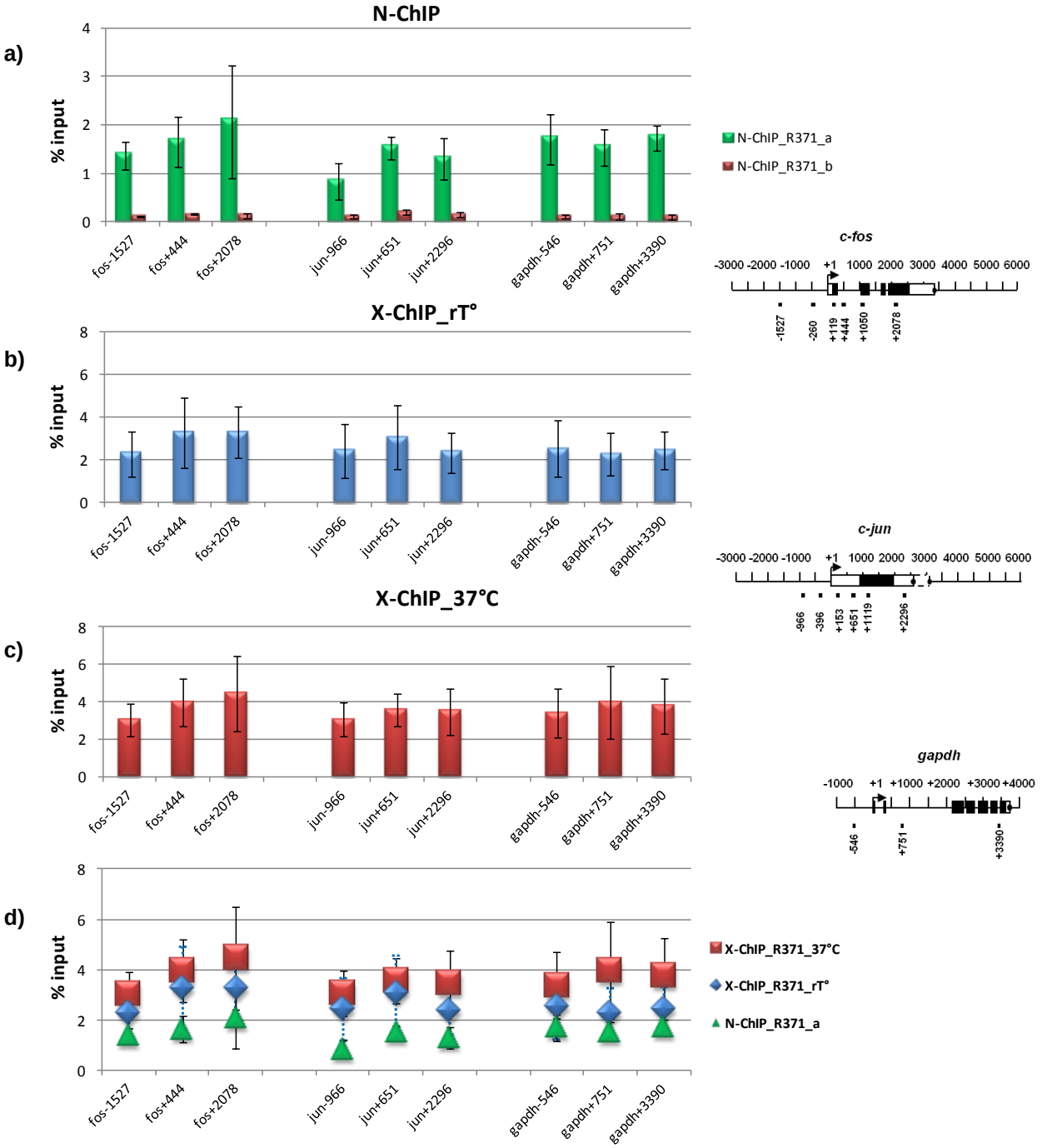
Recovery of HMGN1 immunopurified DNA is lower in N-ChIPs but remarkably similar to X-ChIPs (Figure 28d). This is quite surprising considering the lack of direct interaction of HMGN1 with DNA. The recovery of DNA in N-ChIPs might be then a reflection of the strength of HMGN1 interaction with nucleosomes in native conditions. However, native chromatin is unstable during storage; N-ChIPs performed on identical sets of chromatin after 3 weeks storage at -80°C showed dramatic loss of recovery of DNA (Figure 28a). Thus, even if interaction of HMGN1 with nucleosomes is sufficiently strong to persist through immunopurification, the freezing of samples compromises these assays. Problems with the reliability of the native chromatin preparation procedure lead to its discontinuation, and the exclusive use of X-ChIP. The lack of information on histone PTMs in native conditions

contributed to this choice. Cross-linking of chromatin at room temperature or 37°C did not influence the recovery of HMGN1-bound DNA (Figure 28d). ChIP results are almost identical or are contained within the standard deviation of the samples. However, the effect of temperature on cross-linking was monitored in all the following experiments (see appendix). Due to the lack of any clear effect produced by the 7-8° difference between cross-linking conditions, data from X-ChIP is presented as a whole.

4.2.1.1 Distribution of HMGN1 is not dependent on gene transcription

HMGN1 is present at IE gene *c-fos* and *c-jun* in quiesced cells (Figure 28) despite the lack of transcription of these genes in such conditions (Figure 2). This is also the case for histone PTMs studied, and the presence of HMGN1 at these loci could also be an indication of an open state of chromatin, in light of the effect that HMGN1 is reported to have on chromatin compaction (Kugler et al., 2012). The recovery of HMGN1-bound DNA sequences is as high at IE genes as at the continuously transcribed housekeeping gene *gapdh* (Figure 28). Overall, HMGN1 ChIPs recover 2-6% of input sequences. Such recoveries can be considered quite low for histones but are similar to those observed with ChIP of scarcely abundant chromatin proteins such as RNA Pol II (Crump et al., 2011; Edmunds et al., 2008). HMGN1 is present evenly at promoter regions (*c-fos* -1527, *c-jun -966, gapdh -546*), introns (*c-fos +444, gapdh +751*), exons (*c-fos* +2078, *gapdh +3390*), 5' UTR (un-translated region) (*c-jun* +651) and 3' UTR (*c-jun* +2296). Promoter regions are the less represented in HMGN1 ChIPs, while interaction with 3' sequences appear stronger (Figure 28). Nevertheless, the difference between percent of recovery is very small and HMGN1 distribution appears homogenous.

Figure 28. **HMGN1 distribution in unstimulated samples and reproducibility of N-ChIP and X-ChIP.**
C3H 10 T1/2 cells were quiesced 18 h (o/n) and then native and cross(X)-linked chromatin samples were prepared. ChIPs were performed with HMGN1 C-terminal antibody [R371]. **a)** Native (N)-ChIP was performed on three separate chromatin set [R371a]. A repeat of N-ChIP was made after 3 weeks of storage at -80°C [R371b]. **b) c)** X-ChIP was performed on three separate chromatin set of cross-linked chromatins prepared at **(b)** room temperature (rT°) or **(c)** at 37°C. A repeat of X-ChIP was made after 3 weeks of storage at -80°C, DNA recoveries were similar (data not shown) and have been pooled together with previous sets. **d)** Direct comparison of DNA recoveries of N-ChIPs and X-ChIPs.



4.2.2 Influence of IE gene activation on HMGN1 distribution

Inducible histone PTMs are enriched at loci of transcribed genes (Figure 9). HMGN1, by contrast, is evenly distributed across all loci examined (Figure 28). Induction of MAPK signalling with EGF or sAn does not significantly alter the presence of HMGN1 at the gene loci that are activated by stimulation (Figures 29-30). HMGN1 is present at *c-fos, c-jun, gapdh* and *hbb* (Figure 29), showing a slight preference for coding region of *c-fos* (*c-fos* +1050, *c-fos +2078*), *gapdh* and *hbb* promoter (*hbb* -79) (Figure 29). Lower recoveries are detected at *c-fos* and *c-jun* promoter regions (*c-fos* -1527, *c-fos -260, c-jun -966, c-jun -396*), and *hbb* 3' coding region (*hbb* +1247).

Induction of IE genes with EGF or sAn hardly perturbs HMGN1 distribution at early time points (Figure 30). Upon stimulation with sAn, HMGN1 presence slightly decreases at *c-fos* loci (Figure 30a) and (with the exception of *c-jun* +153 and *c-jun* +2296) remains mostly unchanged at *c-jun* loci (Figure 30b). Stimulation with EGF appears to reduce HMGN1 levels at some specific loci (*c-fos +119, c-fos +444, c-fos* +1050 *c-fos +2078, c-jun* +153, *c-jun* +1119, *c-jun* +2296 and *hbb* +1247), whereas at other loci (*c-fos* -1527, *c-jun +651, gapdh -546, gapdh +3390 and hbb* -79) an increase of HMGN1 is observed. Overall, IE gene induction at early time points tends to reduce the presence of HMGN1 at these loci although the variation of occupancy is very subtle. Interestingly, the loci where the presence of HMGN1 is most sensitive to induction are also the most enriched in HMGN1. This might suggest that the association of HMGN1 at these positions is counteracted by an induction-dependent mechanism dislocating the protein from these sites.

The distribution of HMGN1 at later time points of IE gene induction is also slightly different from basal level. At loci (*c-fos +119, c-fos +444, c-fos* +1050 *c-fos +2078, c-jun* +2296, *gapdh -546*), where there was depletion at previous time points (Figure 30), the basal level is

restored (Figure 32). Curiously, HMGN1 levels at *c-fos* loci are higher than basal after 90 min of sAn stimulation (Figure 32a). Overall, the changes in HMGN1 distribution at later time points of IE gene induction are even smaller than previously observed for the early time points.

In conclusion, these analyses mainly show that HMGN1 is present at low level ubiquitously across the examined loci and its distribution is only slightly affected by induction. Variations are observed at IE genes, constitutively transcribed and silenced genes.

Figure 29. **HMGN1 is present at active and inactive gene loci (EGF or sAn, 0-15-30 min).**
C3H 10 T1/2 cells were quiesced 18 h (o/n), stimulated for the stated time with EGF (50 ng/ml) or sAn (25 ng/ml) and then cross-linked chromatin samples were prepared. ChIPs were performed with HMGN1 C-terminal antibody [R371]. Recovery of HMGN1-bound DNA sequences at: **a)** *c-fos* loci, **b)** *c-jun* loci, **c)** *gapdh* and *hbb* loci. Error bars represent standard deviation of four different sets of chromatin samples for each time point and stimulus.

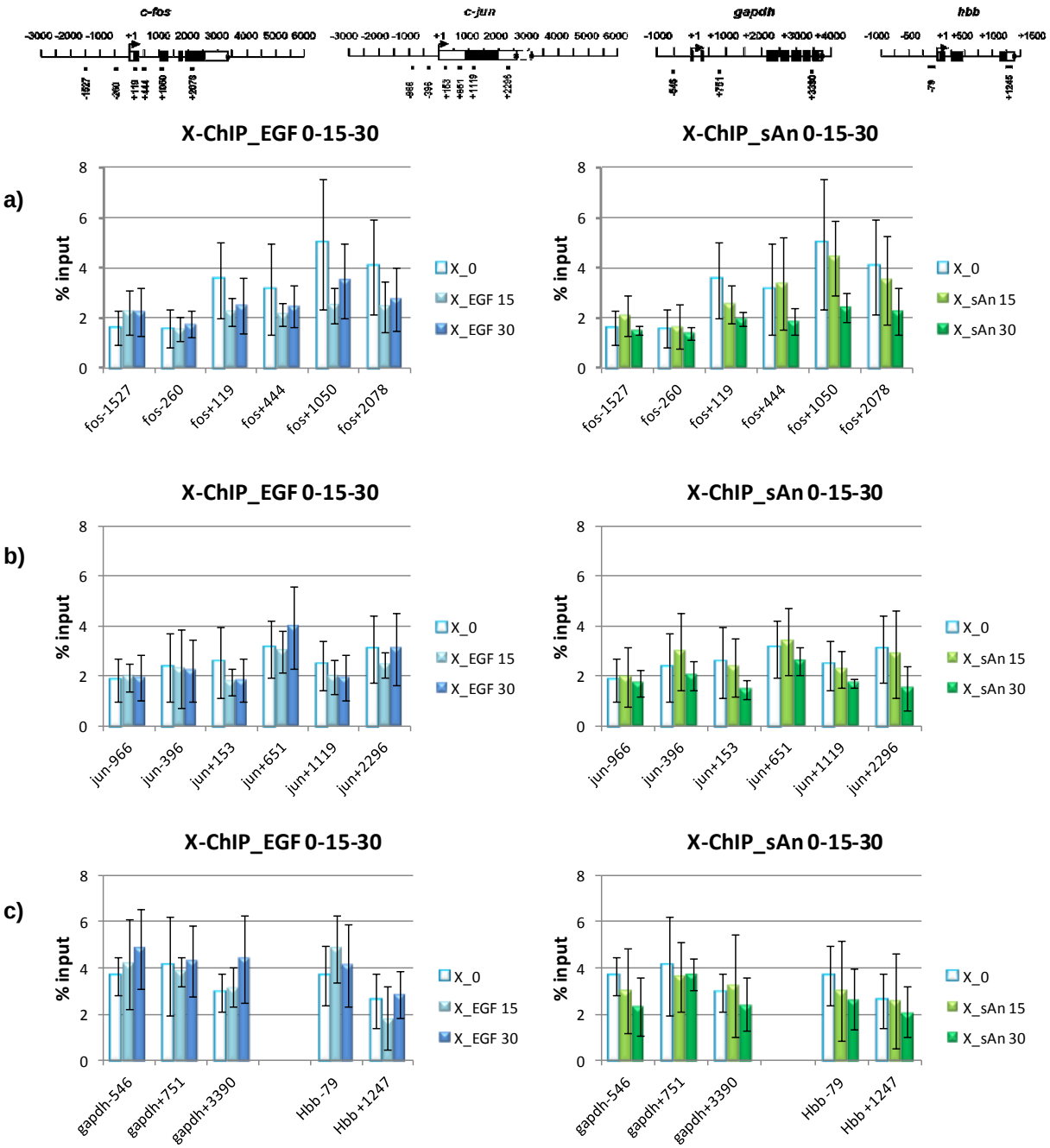


Figure 30. **HMGN1 distribution is not influenced by induction (EGF or sAn, 0-15-30 min).**
C3H 10 T1/2 cells were quiesced 18 h (o/n), stimulated for the stated time with EGF (50 ng/ml) or sAn (25 ng/ml) and then cross-linked chromatin samples were prepared. ChIPs were performed with HMGN1 C-terminal antibody [R371]. Recovery of HMGN1-bound DNA sequences at: **a)** *c-fos* loci, **b)** *c-jun* loci, **c)** *gapdh* and *hbb* loci. Re-plot of data from Figure 29.

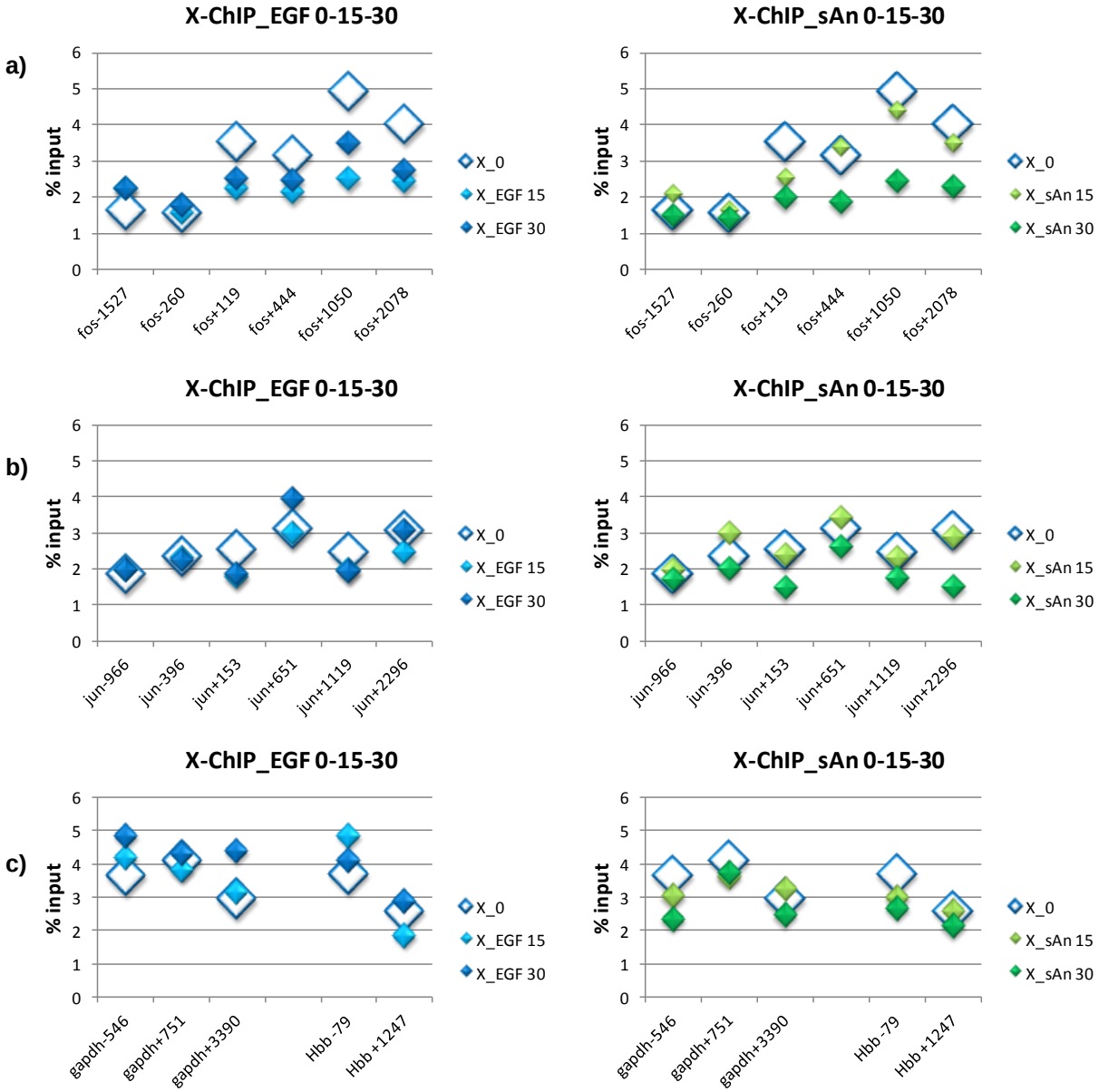


Figure 31. **HMGN1 is present at active and inactive gene loci (EGF or sAn, 0-45-60-90 min).**
C3H 10 T1/2 cells were quiesced 18 h (o/n), stimulated for the stated time with EGF (50 ng/ml) or sAn (25 ng/ml) and then cross-linked chromatin samples were prepared. ChIPs were performed with HMGN1 C-terminal antibody [R371]. Recovery of HMGN1-bound DNA sequences at: **a)** *c-fos* loci, **b)** *c-jun* loci, **c)** *gapdh* and *hbb* loci. Error bars represent standard deviation of four different sets of chromatin samples for each time point and stimulus.

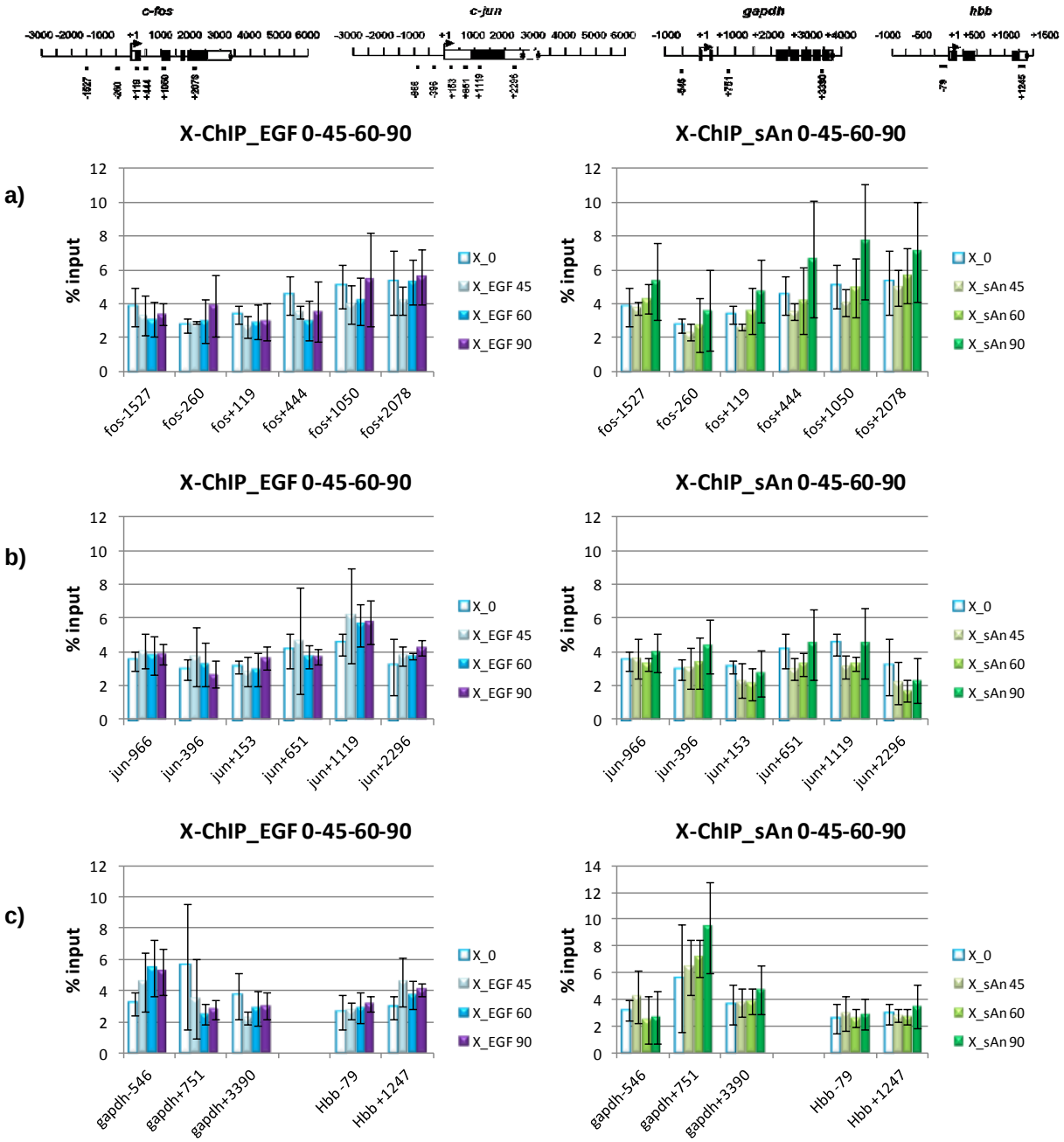
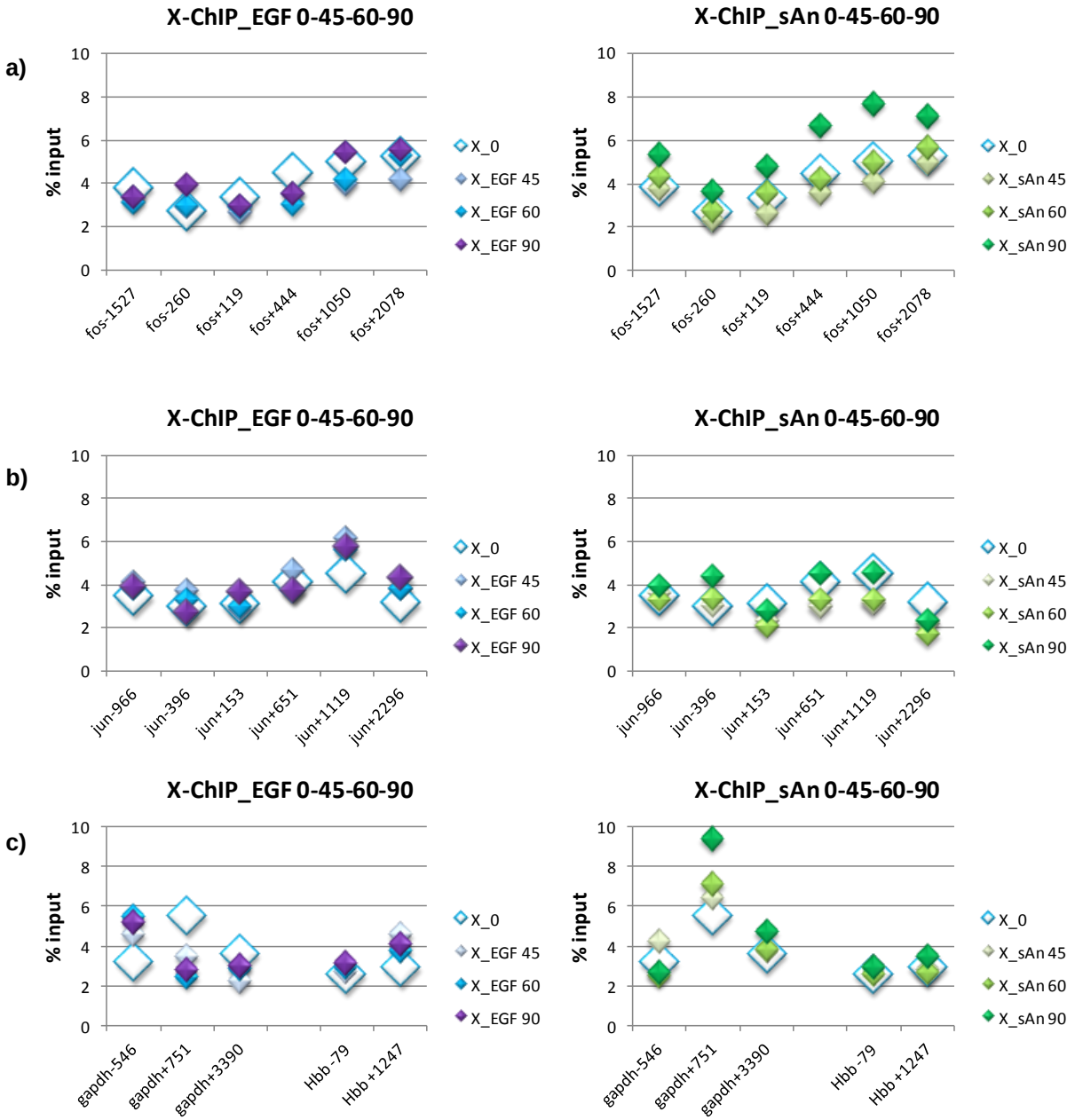


Figure 32. **HMGN1 distribution is not influenced by induction (EGF or sAn, 0-45-60-90 min).**
C3H 10 T1/2 cells were quiesced 18 h (o/n), stimulated for the stated time with EGF (50 ng/ml) or sAn (25 ng/ml) and then cross-linked chromatin samples were prepared. ChIPs were performed with HMGN1 C-terminal antibody [R371]. Recovery of HMGN1-bound DNA sequences at: **a)** *c-fos* loci, **b)** *c-jun* loci, **c)** *gapdh* and *hbb* loci. Re-plot of data from Figure 31.



4.2.3 Association of HMGN1 with histone H3 PTMs

Sequential ChIP of quiesced cross-linked chromatin was performed using HMGN1 C-terminal and histone H3K9ac-specific antibodies (Figures 33-34). Recovery of DNA in the second ChIP is calculated against the amount of DNA recovered in the first ChIP, and not the original input DNA (Figure 12). As control, a standard ChIP with the same antibody was performed at the same time (Figure 33a, c) and pre-immune sera were used as negative controls (Figure 33b, d). HMGN1 recovery in sequential ChIP is 40-70% of HMGN1-bound DNA from the first ChIP, whereas H3K9ac is 60-100% of first ChIP (Figure 33b, d). These values indicate the ability of the antibody to recognize the same DNA-protein complex in both ChIPs, with successful re-binding to the epitopes during the second ChIP. Not surprisingly, histone association with DNA is more stable and resistant than that of HMGN1. At *hbb* loci the scarce presence of H3K9ac marks might explain the unusually low recovery of histone-bound DNA (Figure 33d).

At *c-fos* loci (*c-fos* -1527*, c-fos -260, c-fos +119, c-fos +444, c-fos* +1050), HMGN1-bound sequences contain 10-20% of histone H3 tails carrying K9ac (Figure 34a). At *c-jun* loci (*c-jun* -966, *c-jun* +651, *c-jun* +2296), HMGN1-bound sequences contain 20-40% of histone H3 tails carrying K9ac (Figure 34a). These values are remarkably high, especially if compared to the recoveries at the same loci of HMGN1 ChIP on HMGN1-bound chromatin (Figure 33b). DNA sequence recoveries at *gapdh* loci (*gapdh -546, gapdh +751* and *gapdh +3390*) resemble the situation observed at *c-fos* loci; about 10-20% of HMGN1-bound DNA was recovered by the H3K9ac antibody (Figure 34a).

H3K9ac-bound DNA was not preferentially associated with HMGN1 at any distinct loci (Figure 34c). Secondary IP with HMGN1 antibody recovered only 5-10% of H3K9ac-bound sequences, with no distinction between loci at *c-fos, c-jun* or *gapdh*.

Overall, comparing the DNA recoveries from single ChIP (Figure 34b, d) with the result from sequential ChIP (Figure 34a, c), it appears that the distribution of HMGN1 and that of H3K9ac are independent. For example, HMGN1 is bound to no more than 2-4% of total input sequence at *c-jun* (Figure 34d) and of this DNA about 40% is recognized by H3K9ac antibody (Figure 34a). This is about the same percentage of chromatin that would be ChIPed by H3K9ac from a standard input (Figure 34b). Further, at other loci, the distribution of H3K9ac in HMGN1-bound chromatin is similar to that of a standard H3K9ac ChIP (Figure 34a, b) and conversely HMGN1 distribution in H3K9ac-bound chromatin is similar to that of a standard HMGN1 ChIP (Figure 34c, d).

Figure 33. **Efficiency of sequential HMGN1 and H3K9ac ChIPs.**

C3H 10 T1/2 cells were quiesced 18 h (o/n) and then cross-linked chromatin samples were prepared. Sequential ChIPs were performed with HMGN1 C-terminal antibody [R371] and H3K9ac antibody. Standard ChIPs were concomitantly performed with both antibodies. **a)** HMGN1 ChIPs, **b)** HMGN1 ChIPs on HMGN1-bound chromatin **c)** H3K9ac ChIPs, **d)** H3K9ac ChIPs on H3K9ac-bound chromatin. Error bars represent standard deviation of four different sets of chromatin samples.

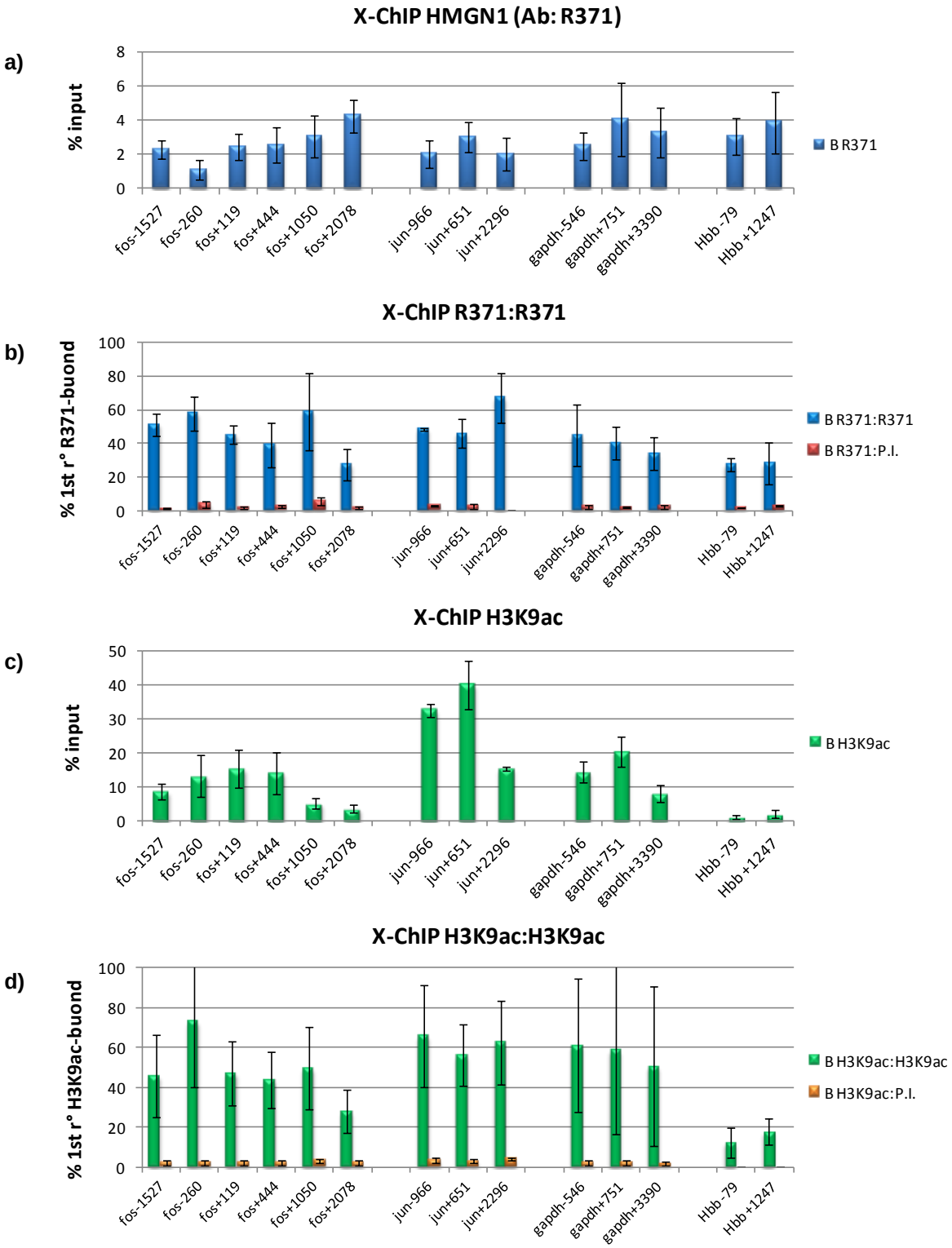
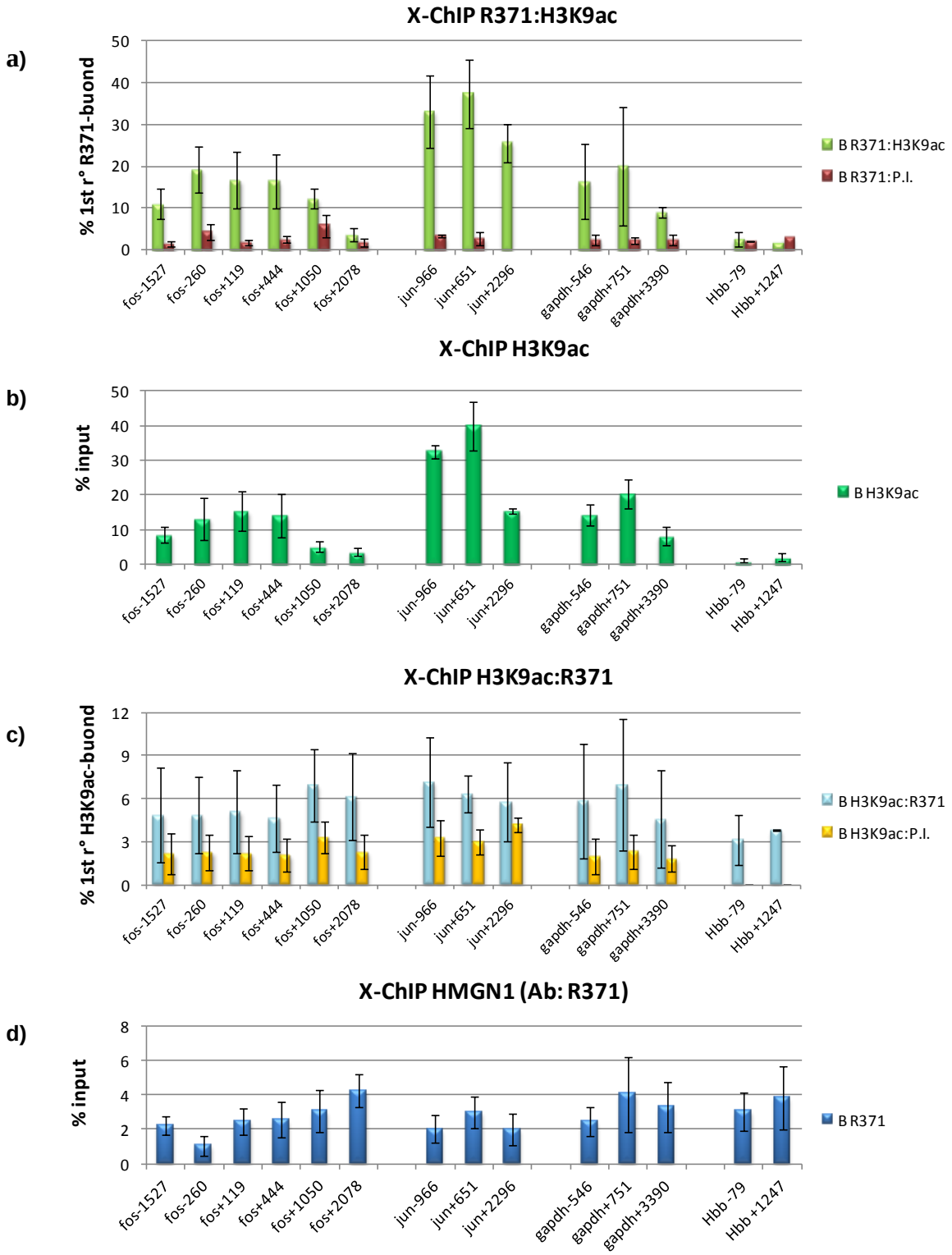


Figure 34. **H3K9ac and HMGN1 distributions on chromatin are independent.**

C3H 10 T1/2 cells were quiesced 18h (o/n) and then cross-linked chromatin samples were prepared. Sequential ChIPs were performed with HMGN1 C-terminal antibody [R371] followed by ChIP with H3K9ac antibody and *vice versa*. 2nd round ChIPs were also performed with HMGN1 [R371] and H3K9c pre-immune sera to test unspecific binding. Standard ChIPs were concomitantly performed with both antibodies. **a)** H3K9ac ChIPs on HMGN1-bound chromatin, **b)** H3K9ac ChIPs **c)** HMGN1 ChIPs on H3K9ac-bound chromatin, **d)** HMGN1 ChIPs. Error bars represent standard deviation of four different sets of chromatin samples.



4.2.3.1 Influence of induction of IE genes on HMGN1-H3 association

The interaction between HMGN1 and histones H3 carrying specific modifications was also investigated in the context of IE induction with sAn. I previously showed that HMGN1 distribution across *c-fos* and *c-jun*, and housekeeping gene *gapdh* and silenced gene *hbb*, is not influenced by IE gene induction (Figures 29, 31). Similarly, in the chromatin used for sequential ChIPs, distribution of HMGN1 does not show any remarkable variation upon IE induction (Figure 35a) and neither efficiency of sequential ChIP is influenced (Figure 35b). Association of HMGN1 with H3K4me3 and H3K9ac is detected at IE gene loci and at *gapdh* loci, while *hbb* loci and *c-fos* +2078 sequences are rather scarce (Figure 36a, c). HMGN1-bound sequences contain 20-40% of histone H3 tails carrying K9ac (Figure 36a) and 30-40% of H3K4me3 (Figure 36c). H3K4me3- and H3K9ac-bound DNA from ChIPs performed on HMGN1-bound chromatin are also in this case similar to the recoveries from a standard input (Figure 36b, d). IE gene induction does not alter H3K4me3 and H3K9ac DNA recovery from HMGN1-bound chromatin suggesting once again that HMGN1 presence at these loci is not altered by stimulation. Levels of H3 PTMs are very similar in standard and sequential ChIP, even though the stimulation with sAn did not induce H3 PTMs levels exactly as previously reported (Figure 9). The comparison of H3K9ac levels at *c-fos* -260, *c-fos* +444, *c-jun* +651 (Figure 36b) and HMGN1-H3K9ac levels at the same loci (Figure 36a), shows that rising levels of this H3 PTM in the whole chromatin are reflected in the HMGN1-bound chromatin as expected (Figure 9). The response of H3K4me3 to stimulation is less than expected in these experiments; however, H3K4me3 levels in the HMGN1-bound chromatin (Figure 36c) behaves similarly to H3K4me3 levels in the corresponding ChIPs (Figure 36d). It is unclear why the *jun* -396 locus behaves anomalously in all ChIPs. Despite the fact that the presence of HMGN1 at this locus does not change upon IE gene induction (Figure 35a, b), a major loss is evident in both H3K9ac and H3K4me3 levels (Figure 36).

Figure 35. **Efficiency of sequential HMGN1 ChIPs.**

C3H 10 T1/2 cells were quiesced 18h (o/n) stimulated for the stated time with sAn (25ng/ml) and then cross-linked chromatin samples were prepared. ChIPs were performed with: **a)** HMGN1 C-terminal antibody [R371] **(b)** followed by ChIP with HMGN1 C-terminal antibody [R371] **(c)** or pre-immune sera. Error bars represent standard deviation of four different sets of chromatin samples for each time point.

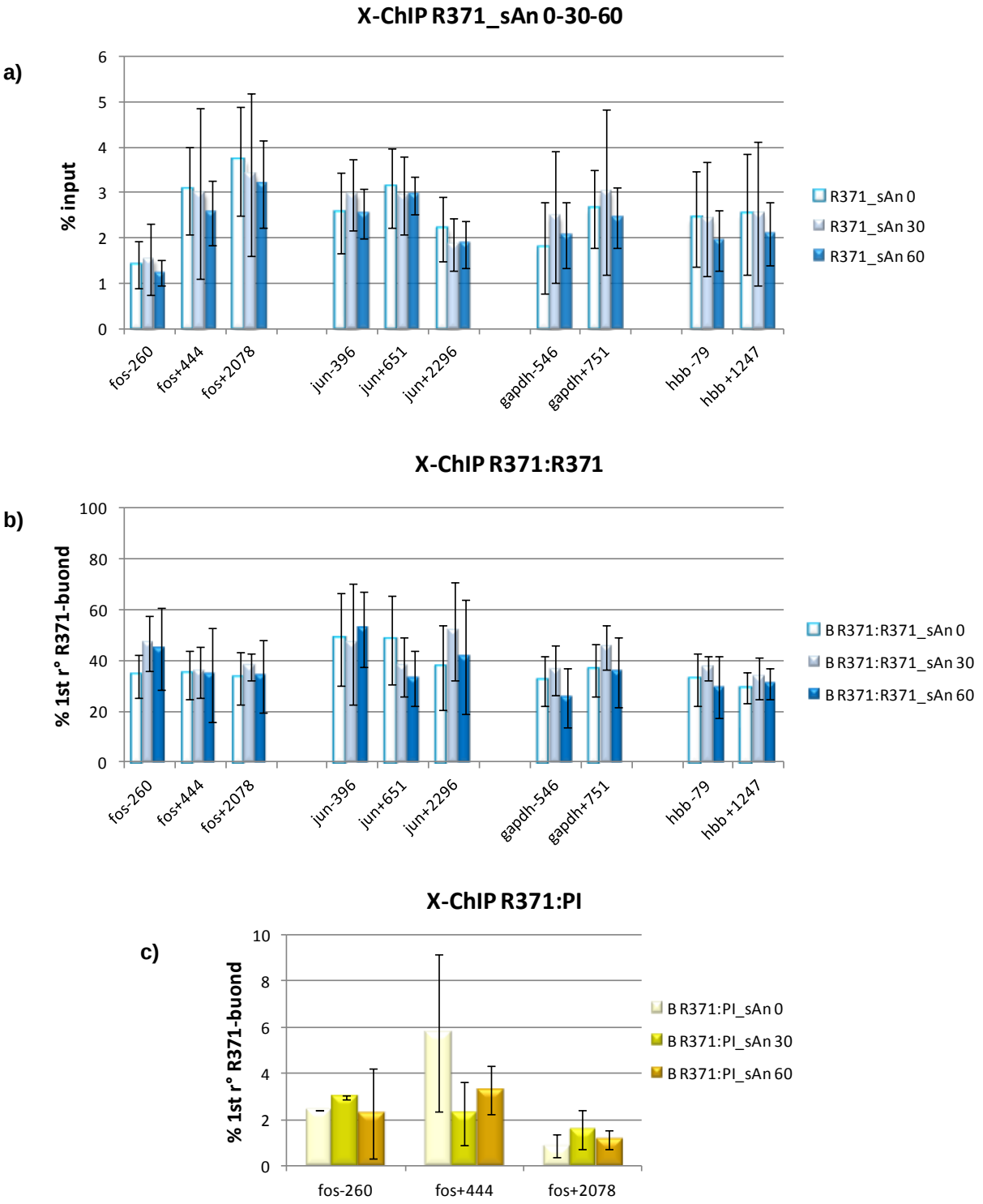
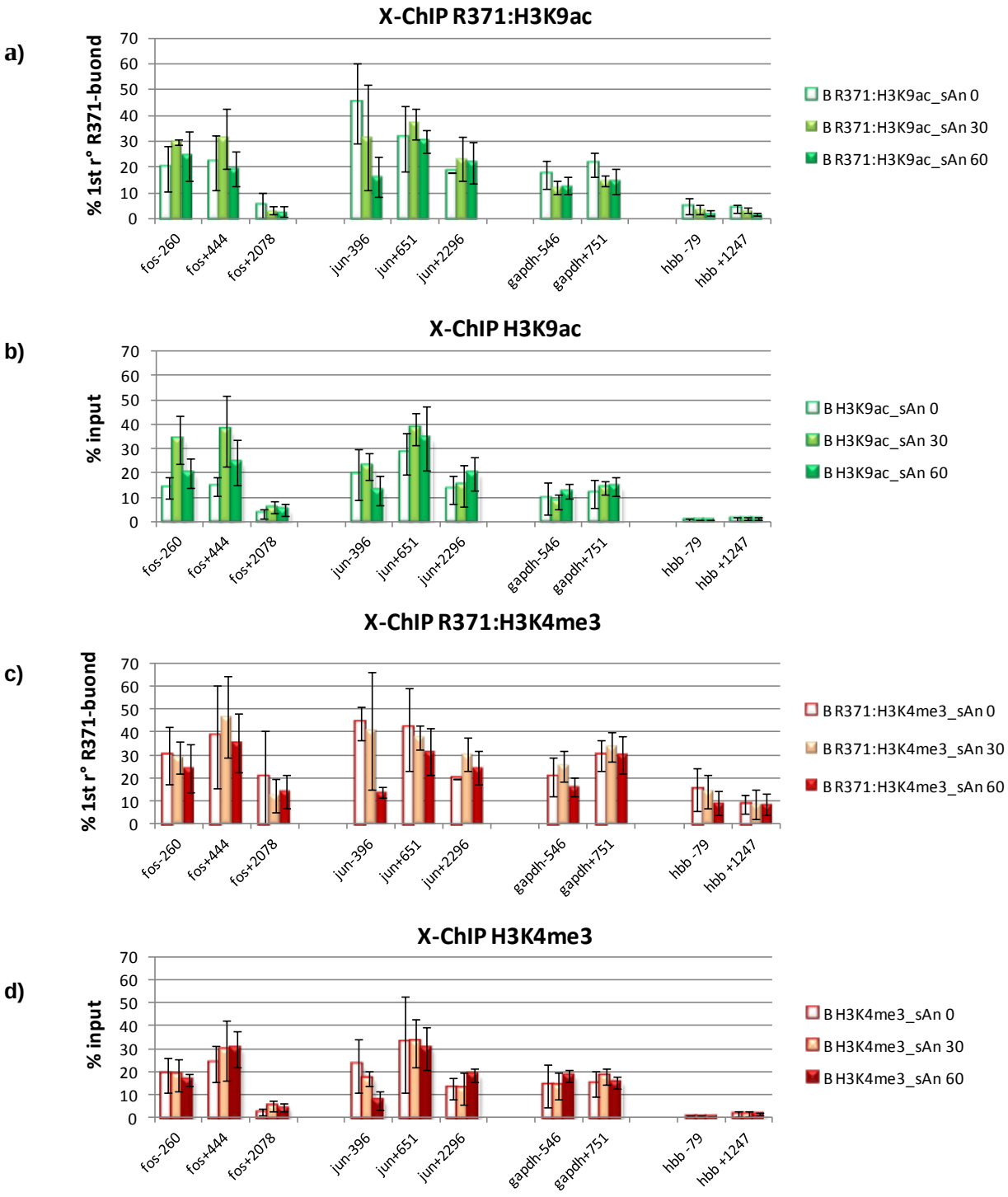


Figure 36. **H3 ChIP profiles (K9ac and K4me3) in the HMGN1-bound chromatin are similar to the general distribution.** C3H 10 T1/2 cells were quiesced 18h (o/n) stimulated for the stated time with sAn (25ng/ml) and then cross-linked chromatin samples were prepared. ChIPs were performed with HMGN1 C-terminal antibody [R371], H3K9ac or H3K4me3 antibodies. **a)** H3K9ac ChIPs on HMGN1-bound chromatin, **b)** H3K9ac ChIPs **c)** H3K4me3 ChIPs on HMGN1-bound chromatin, **d)** H3K4me3 ChIPs. Error bars represent standard deviation of four different sets of chromatin samples for each time point.



4.3 Discussion

The abundance of HMGN1 at the genomic loci of IE genes and control genes *gapdh* and *hbb* was investigated by ChIP. Recovery of sequences is of the order of 2-7% of total input, with variation observed among specific loci. Contemporary to this analysis, independent research groups published comparable results with different reagents and within different cell models, verifying the findings presented here (Barkess et al., 2012; Cuddapah et al., 2011).

Surprisingly HMGN1 is present at silent sequences of *hbb* as much as at continuously transcribed gene *gapdh*. HMGN1 is also present at IE genes before induction, suggesting a role in maintaining accessible chromatin structures. It is unclear what the function of HMGN1 might be at silent sequences.

The distribution of HMGN1 across the analysed loci appeared homogenous with a slight enrichment at 3' region of genes and exonic sequences and a small decrease of occupancy at promoters. These data conflict with the preferential localization at promoters and enhancers reported in the genome wide analysis of Cuddapah et al., (2011), but they are consistent with evidence on the Glyt1 locus in Hepa-1 cells and the FOS locus in MCF-7 cells (Barkess et al., 2012; Zhu and Hansen, 2007). However, variations are so small that it is likely that differences between the cellular models used or the parameters selected for the SICER algorithm in the genome wide analysis can account for them.

Induction with agonists of the IE gene response does not strongly perturb HMGN1 distribution. A dislocation of HMGN1 at specific sites was seen at early time points whilst a recovery of occupancy was observed at later time points. However, the variations were so small that it is hard to support the hypothesis that HMGN1 distribution responds to stimulation. This is quite unexpected considering the strong variation of histone PTMs occurring at the same positions and the strong correlation of transcription of these genes and induction with these agonists. It has been observed in the previous chapter that IE gene

induction does not produce phosphorylation of the NBD domain and that might suggest that HMGN1 binding is not perturbed. Attempts to perform ChIP with Ser6ph-specific antibodies seemed to recover additional proteins (possibly histone H1ph epitopes, data not shown). Thus, interpretation of HMGN1Ser6ph-bound DNA recovery was not possible.

The sequential purification of histone H3- and HMGN1-bound DNA sequences showed that H3 PTMs (H3K4me3 and H3K9ac) and HMGN1 interaction is captured by cross-linking and is resistant enough to allow chromatin purification with either HMGN1- or H3-specific antibodies. The distribution of HMGN1 and H3 PTMs at *c-fos*, *c-jun*, *gapdh* and *hbb* loci in preselected chromatin (either by HMGN1 or H3 ChIP) resembles the distribution in standard chromatin. In H3K9ac-bound chromatin, HMGN1 is as homogenously distributed among the loci examined as it is in HMGN1 ChIP. Similarly, peaks of H3K9ac enrichment are observed in HMGN1-bound chromatin as reported previously (Figure 9, Edmunds et al., 2008).

If the histone modification state at specific loci is not related to HMGN1 distribution, HMGN1 might be targeting specific sites due to other characteristics of chromatin. It is conceivable that HMGN1 is involved in regulating higher order chromatin organization and compaction, influencing the dynamics of genome organization. Conversely, HMGN1 distribution could also be itself influenced by the organization of chromatin and simply directed to those sites that are already more open and likely to be accessible to this fast moving molecule. Data from the literature suggest that HMGN1 has a preference for anchoring nucleosomes, particularly those nucleosomes that define the boundaries of the CTCF- and YY1-binding sites across promoters and enhancers. Instead, in C3H 10T1/2 cells the highest recoveries are observed at the 3’ end and coding region of IE genes. In both reports, no direct relation of active transcription and HMGN1 positioning has been noted. The conclusion from these data and the evidence from the literature is that the role of HMGN1 in chromatin regulation remains evasive and hard to define. There are insights that

events at specific loci and architectural structure of chromatin can influence its distribution but there is no indication of the biological principles behind this preference, nor is there any proof that HMGN1 itself establishes the chromatin environment instead of being attracted by specific chromatin structures. The minimal response to activation of transcription observed in this study, the subtle variation of the transcription profile in the knockout mutants and observations from this study of the distribution of HMGN1 in the context of IE genes all favour the idea that HMGN1 functions more as part of a dynamic network of proteins counteracting chromatin compaction rather than as a pivotal player in gene regulation (Catez et al., 2004; Cuddapah et al., 2011). It is also clear that the signalling cascade activated during IE gene induction is not fundamental in regulating the distribution of HMGN1. This is not necessarily surprising considering also that the proportion of HMGN1 molecules undergoing Ser6ph is about 50-80% of the total HMGN1 population, a fraction several times larger than the HMGN1 molecules likely present at a specific time on actively transcribed genes. This might also mean that phosphorylation itself has a different role that for the moment remains unclear.

Chapter 5. Discussion and future directions

5.1 Summary of results

Regulation of chromatin involves the combination of structural elements and dynamic elements. Nucleosomal histones, due to their stable interaction with DNA and their abundance, represent the cornerstone of chromatin organization. Multiple enzymatic activities are targeted to nucleosomal proteins that can affect accessibility to DNA, with effects both globally and locally. Histones themselves are subjected to the regulation of their accessibility by several architectural proteins, like HMGN1. The fast dynamics of association of architectural factors with chromatin often limits the level of insight that experiments can reveal.

5.1.1 HMGN1, a signalling protein

The most interesting characteristic of HMGN1 is the dual role that it plays in chromatin. In fact, HMGN1 acts both as a nucleosome-bound, chromatin architectural protein and as a signalling molecule. A lot of information is available regarding the role of different amino acids in HMGN1 in its interaction with nucleosomes. The dynamics of HMGN1 diffusion in the nucleus is well studied, even though a refined structure of HMGN1 is still unavailable (Phair et al., 2004; Ueda et al., 2008). The knockout genetic model provided intriguing information on the relation of HMGN1 with histone PTM regulation (Lim et al., 2004, 2005). On the signalling side, it is clear that HMGN1 is specifically targeted for phosphorylation by

MAPK cascades and HMGN1Ser6ph is absolutely dependent on MSKs (Soloaga et al., 2003). This characteristic of the nucleosomal response is a rare feature in comparison to the common promiscuity observed among MAP-kinases and their substrates. The most recent attempt to explain the two roles of HMGN1 provide a model that seemed convincing (Lim et al., 2004). In this model (Figure 8) HMGN1 phosphorylation at the NBD is a crucial part of the response to signalling. NBD phosphorylation interferes directly with HMGN1 binding to nucleosomes. There are no reasons to doubt the proposed function of NBD or its phosphorylation, but it is unclear to what extent NBD phosphorylation occurs during gene induction.

In this study HMGN1 phosphorylation during nucleosomal response was analysed searching for NBD phosphorylation that might influence HMGN1 nucleosome binding.

5.1.1.1 Search for HMGN1 NBD phosphorylation

NBD phosphorylation upon stimulation has been reported by Lim et al. (2004) and Louie et al (2000). The latter report involved the fractionation of lysates with a nuclear protein lysis buffer developed in the Mahadevan lab and not only suggested the presence of NBD phosphorylation in different fractions but also reported the increase of cytoplasmic fraction of HMGN1 upon stimulation. The lack of supporting evidence for these claims lead me to further investigate both nuclear and soluble fractions, to clarify the localization of HMGN1 phosphorylation in the mouse fibroblast model.

There are several signalling proteins distributed between cytoplasm and nucleus that required the development of various methods of protein purification for adequate characterization. High salt concentration, use of detergents, purification with acids and centrifugations are some of the most common methods used to separate fractions. Curiously, HMGs themselves were originally defined by their solubility in acid and high salt solutions (Johns, 1982). To the extent of my analysis, no phosphorylation at HMGN1 NBD was identified in direct response

to signalling in the mouse fibroblast model. Nuclear fractions were analysed with MS and phosphorylated NBD is almost undetectable and definitely not sensitive to signalling induction (Figure 27). HMGN1Ser6ph is instead strongly induced by MAPK activation and detected in several peptides (table 3.1). Resolution of HMGN1 in AU-PAGE was optimised and HMGN1 has been detected as a dual population, separated in two distinct bands due to phosphorylation (Figure 13). The slower population of HMGN1 is specifically recognized by HMGN1Ser6ph antibody.

The large majority of HMGN1 protein is extracted in association with histone proteins, either only histone H1 in case of PCA extractions, or with the core histones, in case of HCl extractions (Figures 11, 12). The presence of HMGN1 in soluble fractions is scarce and HMGN1 was almost impossible to detect by mass spectrometry in these fractions. The effects of treatments with MAPK inhibitors suggest that nuclear HMGN1 and soluble HMGN1 are targeted by the same kinases (Figures 19-21) and favour the hypothesis that soluble fractions represent leakage of protein from nuclear pellets during protein extractions.

To confute the hypothesis that NBD phosphorylation does not occur at all in mouse fibroblasts, the effects of quiescing protocols on HMGN1 modifications were investigated (Figures 17-18). HMGN1 in soluble fractions extracted from *non-quiesced* mouse fibroblast interestingly showed a novel degree of modification. These fractions split into three bands on AU-PAGE and also unstimulated samples split into two bands indicating the presence of signalling-independent PTMs. MAPK signalling still induces HMGN1Ser6 phosphorylation but specifically on HMGN1 molecules carrying the additional mark. It is known that HMGN1 phosphorylation at the NBD occurs during mitosis and the loosening of nucleosome binding that NBD phosphorylation causes could well explain the presence of this mark in the soluble fraction, but not in the nuclear fraction of *non-quiesced* fibroblasts. It is known that in a cultured cell population, only a fraction of cells undergo mitosis at any given time. Thus, the

lack of HMGN1 carrying NBD phosphorylation in soluble fractions extracted from quiesced cells could be interpreted as the result of a limited proportion of cells dividing. As for the restriction of Ser6ph on the most modified population of soluble HMGN1, it has been reported that when phosphorylation occurs at the NBD, Ser6ph is also detected (Lim et al., 2004; Louie et al., 2000; Prymakowska-Bosak et al., 2001). Ser6 and Ser20 both lie within a similar kinase substrate motif and it is likely that a kinase recognising Ser20 (and to some extent Ser24) might have similar activity towards Ser6. An alternative explanation could be that the proportion of soluble HMGN1 carrying non-phosphorylated NBD in cultured fibroblasts is below the detection threshold of immunoblotting. Nevertheless, Ser6ph in these fractions is still directly linked with MAPK induction and unstimulated samples do not carry this modification.

In parallel with analyses of cell lysates, indirect immunofluorescence allows some insights on HMGN1 distribution in the cell. HMGN1 antibodies and fluorescing secondary antibodies used in these experiments showed low levels of background and nonspecific signals when competed with the appropriate peptides (Figure 25). In these experimental conditions, HMGN1 appears as an exclusively nuclear protein with homogenous diffusion and partial depletion from nucleoli and DAPI-depleted regions (Figures 22-24). The signal from Ser6ph antibody is increased upon MAPK induction with dynamics in accordance with data from cell lysates (Figure 24). No signal is detected in the cytoplasm upon stimulation, excluding the possibility that HMGN1 cellular compartmentalization is influenced by Ser6 phosphorylation.

A major aim of this work was to define the distribution of HMGN1 at the IE genes *c-fos* and *c-jun*. The effort of investigating the dynamics of HMGN1 phosphorylation was directed towards establishing the level of PTMs in chromatin-associated HMGN1, in order to be able to interpret ChIP recoveries in relation to signalling. It is evident that Ser6ph is not the only

modification present on HMGN1 molecules, but it is nevertheless the only one directly linked to MAPK signalling in the mouse fibroblast model.

In conclusion, HMGN1 is a dynamic protein and it is not surprising that a portion of it is less strongly bound to chromatin. Despite claims from Louie et al (2000), it is inappropriate to define as cytoplasmic a protein whose distribution is more than 90% nuclear on the basis of lysate fractionations. Additional marks were identified on HMGN1and are discussed in the following section. Their dynamics are unchanged by IE gene induction and their abundance is negligible in the context of chromatin-associated HMGN1. Thus, ChIP data will be discussed only with respect to Ser6 phosphorylation in the chromatin of mouse fibroblast cells.

5.1.1.2 The post-translational modification profile of HMGN1

A combination of mass spectrometry and immunoblotting was used to define HMGN1 PTMs. HMGN1 phosphorylation occurs at serines 6, 19, 23, 83, 86 and 94 (residue positions in mouse HMGN1). Phosphorylation at NBD and the C-terminal is detected in control and stimulated samples at equal proportions. NBD phosphorylation has been linked to both mitosis and signalling while the significance of C-terminal phosphorylation remains unclear (Pogna et al., 2010). It has been proposed that casein kinase II might be responsible for HMGN1 phosphorylation but the evidence available comes only from *in vitro* experiments and with excessive amount of kinase used in the assays (Jiang and Wang, 2006). Ser83 is followed by proline and is thus a better target for proline-directed kinases such as ERKs, JNK/SAPK and p38 MAP kinases than for CK2. Ser86 falls within a CK2 consensus motif and seems to be the major site of phosphorylation whilst Ser94 is considered to be a minor site. C-terminal phosphorylation was detected *in vivo* in MCF-7 cells (Zou et al., 2004). The discovery of C-terminal-mediated HMGN1 chromatin-binding in mitosis might imply that C-terminal phosphorylation is also part of the mitotic regulation of HMGN1-nucleosome

association dynamics (Cherukuri et al., 2008). Thus, the additional modification observed in soluble fractions (Figures 17-18) might also occur at the C-terminal rather than the NBD. My attempts to analyse HMGN1 PTM profile via MS detected phosphorylation at the C-terminal in both of nuclear and soluble fractions. However HMGN1, contained in soluble fractions was so scarce that direct analysis of raw data was required. The peaks identified have almost undoubtedly an HMGN1 origin, as they elute during chromatography together with validated HMGN1 peptides and the isotope scattering of their peaks is identical to genuine C-terminal peptides (data not shown). HMGN1 C-terminal phosphorylation was detected in soluble fractions extracted from both stimulated and quiesced samples. C-terminal phosphorylation detected in the nuclear fraction was sufficiently abundant to be properly annotated and data suggested that these marks are not influenced by cell stimulation and IE gene induction (Figure 27). Levels of Ser86ph are stable in induced and super-induced samples, when Ser6ph is almost maximal (≈80%). Phosphorylation at NBD is similarly not influenced by MAPK signalling, and never rises to more than 1-2% of the total HMGN1 population.

Acetylation was detected on specific peptides within the NLS2 of HMGN1, pep31-36 and 44-54, and at the NBD (table 3.1). The absence of unmodified NLS2-peptides within the same cohort of samples impeded any attempt to quantify the relation between IE gene induction and HMGN1 acetylation. The low peptide score assigned by Mascot to acetyl-phosphorylated NBD peptide limits the interpretation of these data. It is very likely that the constant rise of sensitivity of mass spectrometry instrumentation will soon allow further characterization of HMGN1 modifications with a resolution similar to that achieved so far for histones.

In addition to MS data, attempts to increase acetylation of HMGN1 by inhibition of deacetylases with TSA have produced no additional shift on HMGN1 in AU-PAGE (Figure 15). Moreover, treatment with phosphatases has clearly indicated that the only mark

influencing HMGN1 mobility in nuclear and soluble fraction is phosphorylation (Figure 15). These observations do not exclude the possibility that acetylation might be present on HMGN1 at a background level that is neither dynamic nor influenced by MAPK signalling activation.

5.1.2 Dynamic association of HMGN1 with chromatin

The genomic distribution of HMGN1 in unstimulated samples has been shown to be independent of underlying DNA sequence. It has been reported that HMGN1 seemed to preferentially localize to DNase I hypersensitive (HS) sites, promoters, functional enhancers, and transcription factor binding sites (Cuddapah et al., 2011). The enrichment of HMGN1 in the nucleosome-depleted region in active promoters was suggested to be a consequence of a hypothetical role of HMGN1 in anchoring nucleosomes at the boundaries of transcription factor binding sites. Enrichment of HMGN1 was detected at active promoters while HMGN1 binding at silent gene promoters was minor. My analysis of the local distribution of HMGN1 at IE genes *c-fos* and *c-jun* as well as the housekeeping gene *gapdh* and silenced gene *hbb* showed a different situation, more similar to that observed at specific loci (*Glyt-1* in Hepa-1 cells and *FOS* in MCF-7 cells) with targeted ChIP analysis (Barkess et al., 2012; Zhu and Hansen, 2007).

5.1.2.1 HMGN1 at c-jun, c-fos, gapdh and hbb

The analysis of local distribution of HMGN1 at IE genes, *gapdh* and *hbb* led to interesting observations. First, HMGN1 distribution is remarkably uniform among genes with very different genomic structure and that are very differently regulated (Figures 28-29, 31). While it could be plausible that distribution of such a dynamic protein was similar among IE genes,

the presence of HMGN1 at *hbb* in mouse fibroblast indicates that HMGN1 is not exclusively found at active genes. HMGN1 is not only present at *hbb* but its dynamics are almost undistinguishable from that at other gene loci.

The only regions that showed a slight variability upon stimulation lay within the *c-fos* loci (Figure 30a). At *c-fos*, HMGN1 partially behaves according to the Lim et al. (2004) model. Particularly at the 3'end of *c-fos* gene, HMGN1 occupancy is reduced upon stimulation with EGF or sAn at early time points, to then revert to basal levels, and in some cases above basal level, at later time points (Figure 32a). The situation at the other genes examined is different. Occupancy at *c-jun* loci is less variable and that might be a consequence of the structure of the gene that does not contain introns (Figures 30b, 32b). HMGN1 occupancy at *gapdh* is again identical to anywhere else examined, with no difference clearly emerging between promoter, intron or coding regions (Figures 30c, 32c).

Despite the lack of a local distribution with distinct profiles, the presence of HMGN1 at these genes is still relatively high considering the abundance of HMGN1 in the cell, the importance of nucleosome-binding for its recovery in ChIP and the recovery of DNA sequences obtained by ChIP against histone PTMs. Assuming that the amount of HMGN1 available in the nucleus is enough to statistically allow one nucleosome in every ten to be bound by a molecule of HMGN1, a recovery of 2-9% of input at loci where ChIP against histone modifications recover 20-40% implies that in most cases HMGN1 molecules are interacting with these nucleosomes more often than predicted by mere stoichiometry. Moreover, if we consider the relatively low recovery (<5% input) of DNA fragments at IE genes associated with H3S10ph or H3K9acS10ph, which are highly dynamic histone modifications produced by MAPK signalling, we notice that HMGN1 recovery is reasonably close to these levels. The even distribution of HMGN1 might also be a consequence of microheterogeneity deriving from multiple factors. The antibody utilised in this study produced virtually complete

immunodepletion of HMGN1 (appendix Figure 37) and so it is expected that changes in percentage of input DNA purified by ChIP reflects changes in absolute levels of occupancy of HMGN1 at that loci. It has been shown that quiesced mouse fibroblast cultures provide homogeneous, synchronised cell populations to analyse histone modifications during IE gene induction, but the stability of nucleosome positioning and histone marks are far greater than nucleosomal interactions with HMGN1. Thus, it is possible that within the time required for proper cross-linking or for the diffusion of the agonist molecules (EGF or sAn), different HMGN1 distributions are preserved in the cross-linked cell populations. The overall amount of HMGN1 molecules is limited as previously noted and variation of such a small proportion of HMGN1-bound nucleosome can strongly influence the final outcome of ChIP. Variation of HMGN1 occupancy may also be caused by the association of other proteins at the same loci. Transcription is not initiated simultaneously within all cells and polymerase passage may influence HMGN1 local deposition.

In conclusion, HMGN1 ChIP at specific loci with C-terminal antibody has revealed the lack of preference for specific gene structures or for actively transcribed gene instead of silenced gene. The homogenous distribution might be the genuine result of HMGN1 functions as much as an unfortunate limitation of X-ChIP analysis of highly dynamic processes. In this context it is inappropriate to discuss the possible influences on IE gene induction. It is clear that the data acquired from both genomic profiling and local distribution requires further investigation (see future directions).

5.1.2.2 The association of HMGN1 with nucleosomes

To date, the shielding of histone H3 tails from deacetylases and kinases has been the only explanation that has been proposed for HMGN1 function in chromatin regulation. This hypothesis is supported by evidence from the study of HMGN1 -/- knockout mice. Since ChIP analysis of HMGN1 showed a very uniform scenario I considered the possibility of dissecting HMGN1 nucleosome interaction by specifically selecting those associated with modified nucleosomes. Several histone modifications occur dynamically at IE genes during their induction. Histone phosphorylation is itself a very transient modification and performing sequential ChIP with HMGN1 could have been very ineffective due to very low recoveries of DNA sequences. Histone methylation and acetylation are among the best described modifications occurring at IE genes. Dynamics of these marks have been especially well characterized with ChIP for H3K4me3 and H3K9ac. Sequential ChIP has already been successfully performed with antibodies recognizing both modifications. Moreover, sequential ChIP and shifts on acid-urea gels prove that these two modifications often coincide on the same histone tails and within the same nucleosomes (Crump et al., 2011). Thus, they were expected to perform similarly in ChIP.

Data from H3K9ac-HMGN1 (and *vice versa*) sequential ChIP showed that nucleosomes carrying H3K9ac interact with HMGN1 at different loci of IE genes, at *gapdh* and *hbb*. Efficiency of HMGN1 antibody in sequential ChIP is unsurprisingly lower than that of antibodies against histone modifications, mainly due to the weaker interaction of HMGN1 with DNA in comparison to histones (Figure 33b, d). The recovery of *c-jun, c-fos, gapdh* and *hbb* sequences suggests the unspecific association of HMGN1 with H3K9ac nucleosomes. The distribution of HMGN1-bound DNA sequences in H3K9ac purified chromatin was as homogenous as HMGN-bound DNA sequences in the general chromatin population (Figure 34c, d). H3K9ac-bound DNA sequences in the HMGN1 purified chromatin were enriched at

c-jun, *c-fos* and *gapdh* specific loci and depleted at *hbb* to an extent similar to previously reported analysis of unselected chromatin population (Figures 9, 34a, b). The repetition of a similar sequential ChIP including H3K4me3 and a new set of chromatin samples has confirmed this preliminary observation (Figure 36).

Thus, HMGN1-purified chromatin is indistinguishable from the general chromatin population with respect to the presence of nucleosomes containing histone H3K4me3 and H3K9ac. This not only implies that the general distribution of HMGN1 might not be affected at all by histone PTMs but also indicates that its proposed function as a regulator of accessibility to histone tails is not targeted, but may be the result of the abundance of the protein or its ability to bind any nucleosome. This interpretation does not conflict with observations collected so far on the global distribution of HMGN1 and effects of its depletion in mouse. Moreover, the association of HMGN1 with nucleosomes is not influenced by IE gene induction in these sequential ChIPs. Recovery of H3K9ac-bound DNA from HMGN1-bound chromatin mirrors the distribution of H3K9ac in the respective unselected chromatin and that previously reported on IE genes during induction (Figures 9, 36a). H3K4me3 profiles in these experiments correlated less well with previous observations, even though H3K4me3 distribution was similar in both HMGN1-bound chromatin and in the respective unselected chromatin (Figures 9, 36c, d). Curiously, recoveries of HMGN1-H3-bound DNA sequences at *c-jun* -396 are especially low compared to previously reported IE gene ChIP profiles. *c-jun* -396 is a locus very sensitive to MNase digestion and this usually indicates sites where nucleosomal displacement occurs (Edmunds et al., 2008). Transcription-dependent nucleosome displacement might explain the sensitivity to MNase digestion and the depletion observed in HMGN1-H3-bound DNA sequences. However HMGN1-HMGN1 recovery in sequential ChIP and HMGN1 recovery in standard ChIP do not show any comparable depletion of *c-jun* -396 sequences upon induction. The apparent lack of correlation between

HMGN1 DNA recoveries and MNase sensitivity might be influenced by the effects of microheterogeneity that is more evident for ChIP of highly dynamic molecules. As mentioned previously, transcriptional initiation that is responsible for nucleosomal displacement is most likely asynchronous in these cell populations. The fast dynamics of HMGN1 might allow binding of newly displaced nucleosomes much faster than methylases or acetylase. Thus, while HMGN1 ChIP purifies both modified and unmodified nucleosome-bound DNA sequences, ChIP with H3 PTM-specific antibodies is influenced by the dynamics of their deposition. Within the time window of cross-linking, it is more likely that ChIP against H3 PTMs will miss those DNA sequences wrapped around nucleosomes not yet modified that HMGN1 ChIP is still able to purify, averaging the effect of nucleosomal depletion on the recovery of HMGN1 ChIP.

In this study, HMGN1 has been shown to be uniformly distributed among different gene loci and specific interactions with modified histones have been investigated by sequential ChIP. The unchanged distribution of H3K9ac and H3K4me3-bound DNA sequence previously purified with HMGN1 antibody strongly suggest that HMGN1 interaction with histone H3 is independent of histone modification.

5.2 HMGN1 in the context of nuclear dynamics.

The field of chromatin biology has recently seen dramatic changes of its most fundamental principles. Models depicting DNA as a fixed fibre wrapped around a nucleosome, waiting to be bound by factors and traversed by polymerases have now to accommodate the knowledge that the cell nucleus is today regarded as a highly dynamic organelle and all aspects of nuclear function and organization are dynamic (Misteli, 2005). The extensive use of photobleaching experiments (FRAP, FLIP) has shown that chromatin proteins travel across the nucleus in search of high affinity binding sites (Phair and Misteli, 2000). The typical estimated residence times for non-histone proteins on chromatin is of the order of 2–30 seconds and the time between binding events is of the order of 50–200 ms (Phair et al., 2004). These values imply that once dissociation from a binding site occurs, chromatin proteins are immediately captured by a non-specific or a specific binding site in its proximity. These fast dynamics and the density of chromatin structures within the nucleus also suggest that at any given time, a large fraction of chromatin proteins have a higher chance to bind to non-specific sites, with only a small fraction of the protein actually unbound. This also implies that chromatin-binding proteins are often involved in non-specific transient interactions with chromatin. Random interactions amongst proteins are still consistent with stochastic gene expression events (Elowitz et al., 2002; Levsky and Singer, 2003). Histone modifications in this model may play the role of increasing the potential of gene loci for transcription or repression. Different modifications might influence the formation of chromatin structures that can more freely associate with domains that are functionally equivalent. Thus, the cumulation of largely random events may be driven by statistical preference for specific conformation, or binding may result in non-random preferential positioning of protein and gene loci (Misteli, 2004). HMGN1 has already been proven to be part of a dynamic, largely random interaction network involving other HMGs and histone H1, and data within this thesis confirms the largely

random and uniform distribution of the protein. The transient nature of protein and protein–chromatin interactions can still effectively poise the system for rapid changes in response to external stimuli. The mere fact that the dynamics of histone and HMGN1 are largely different suggest that within this model could coexist a more stable component responsible for the enforcement of critical decisions, and more dynamic components responsible for fine tuning of responses to various stimuli.

5.2.1 The role of HMGN1 in signalling at transcription and replication factories

Transcription and replication factories are one of the most controversial themes of chromatin biology (Cook, 2002; Jackson et al., 1990). After years of debate, the collection of gradually stronger evidence now supports the idea that preassembled transcription and replication factories exist in the nucleus and that rather being constituted around specific DNA sites, it is the DNA itself that moves into factories either stochastically or driven by active transport (Chakalova et al., 2005; Janga et al., 2008; Xu and Cook, 2008). One of the major characteristics of transcription factories is the local concentration of multiple functions. In a highly dynamic nuclear system, factories provide stability and local recruitment for all those chromatin proteins that roam in the nucleus. It is possible then that concomitant targeted phosphorylation of HMGN1Ser6 and H3Ser10 is a result of the physical co-localization of H3 and HMGN1 in a transcription factory and that the uniform distribution of HMGN1 is due to fast exclusion of the molecule from the factory. This could explain why activation of a restricted set of genes induces the phosphorylation of 80% of the HMGN1 protein. It is really unlikely that 80% of HMGN1 protein interacts specifically with IE genes that are a tiny proportion of the whole accessible genome in the time window of IE gene activation.

Following the same hypothesis, during mitosis HMGN1 and histone H3 could be simultaneously phosphorylated at replication factories. Thus, the similarity between induced and mitotic phosphorylation events could be both explained by physical concentration of chromatin factors thought to reside in different type of factories. The extensive NBD phosphorylation observed on mitotic HMGN1 could then be the result of different kinases operating prior to mitosis, but not during gene induction. In such a system, HMGN1 function may be thought to promote chromatin unfolding and inclusion into factories by targeting two main elements maintaining chromatin compaction, linker histone H1 and the nucleosomal histones. Phosphorylation at the NBD could be responsible for increased detachment from nucleosomes or compartmentalization due to protein binding (14-3-3).

This is a mere speculation; unfortunately, the almost pan-nuclear distribution of HMGN1 and the short residence time on chromatin might still limit the possibility of directly detecting HMGN1 molecules inside small territories, or to co localize the nucleosomal response to transcription factories or mitotic HMGN1 phosphorylation to replication factories.

Future directions

Improvement of our understanding of the role of HMGN1 in chromatin regulation depends on the clarification of the molecular mechanisms underlying its function. The model proposed by Lim et al. (2004) indicates a possible explanation for the role of HMGN1 in the nucleosomal response but relies completely on PTM profiling of HMGN1 and this profile is not replicated in all cells. It would be interesting to explore the function of protein phosphatases in the regulation of HMGN1 phosphorylation. Preliminary data on treatment with okadaic acid, an inhibitor of phosphatases PP1 and PP2A, suggested an increase of HMGN1 phosphorylation but the investigation has not been carried over to clarify the mechanism and kinases involved (Louie et al., 2000). Characterization of the PTM profile of HMGN1 in soluble fractions

might provide insight on the regulation of the retention on chromatin of HMGN1, while taking into consideration the limits of such crude separation of fractions. ChIP profiling of HMGN1 at the genomic level might be performed with chromatin from stimulated cells, addressing the issue of the effects of MAPK-induced phosphorylation and transcriptional activation on the mobility of HMGN1 genome wide. Investigating the distribution of HMGN1 in MSK knockout mutants could also help to clarify if phosphorylation targeted to HMGN1 has any effect at all on its distribution. Sequential ChIP with alternative H3 PTM-specific antibodies might provide a different picture of HMGN1 interaction with histones. Inhibition of transcription elongation with DRB could interfere with HMGN1 distribution at specific loci, even though there is a good chance that heterogeneity will affect the local distribution of HMGN1 ChIP recovery at a local level.

Producing a crystal structure of HMGN1 might also improve the understanding of the contribution of different HMGN1 domains to nucleosome binding, which remains the prominent function of HMGN1 on which all the additional roles depend.

A last interesting direction of investigation is the characterization of the mitotic kinase responsible of HMGN1 mitotic phosphorylation. Due to its role in H3 mitotic phosphorylation, Aurora B is a prime candidate. My attempts to use inhibitors to block Aurora B kinase failed to interfere with the additional shift in HMGN1 soluble fractions from *non-quiesced* cells. However, this analysis was only preliminary, and more investigation should be carried on properly mitotic cells, rather than a mixed population.

Appendix

This section contains additional data regarding HMGN1 ChIP.

Preliminary work on HMGN1 ChIP efficiency.

Depletion of antibody-targeted epitopes in unbound fractions was measured by western blot analysis to monitor efficiency of purification of chromatin (Figure 37). A loss of 70-80% of the epitope is regarded as an adequate depletion. HMGN1 ChIP with Ab: R371 (C-terminal) can recover almost 100% of the epitope from input chromatin. Purification of ChIP-bound proteins can be inefficient, complications arising from cross-linking methods and from the chemical composition of ChIP buffers. Nevertheless, it was possible to recover a bound fraction of HMGN1 from both native and cross-linked chromatin.

Additional data from HMGN1 ChIPs reported in chapter 4

In these panels, data from control samples from unstimulated chromatin were plotted separating the data from ChIPs on cross-linked chromatin prepared at room temperature and at 37°C (Figures 38-40). In the vast majority of cases, the variation observed between ChIP data falls within the range of standard deviation of each sample. Overall, cross-linking at 37°C seems to slightly increase retention of DNA sequences. Temperature affects cross-linking and ChIP of histone H3 (H3K4me3 and H3K9ac) less than that of HMGN1.

To clarify, cross-linking at room temperature occurs on chromatin samples immediately after dishes are moved from a 37°C incubator to tissue culture hood. Therefore, the actual temperature of cross-linking is likely to be higher than the standard (approx. 25°C) room temperature.

The differences between these two protocols were analysed to ensure compatibility with previous reports in the literature that utilised either cross-linking at room temperature (Edmunds et al., 2008) or at 37°C (Cuddapah et al., 2011) .

Preliminary work on efficiency of formaldehyde cross-linking and MNase chromatin digestion in comparison with dimethyl adipimidate (DMA) cross-linking and sonication.

DMA cross-linking was performed in combination with formaldehyde cross-linking to trap long range interaction between protein and DNA (Kurdistani and Grunstein, 2003; Zeng et al., 2006). DMA cross-linking does not affect DNA but can link proteins within a distance of 8.6 Å, compared to 2.0 Å for formaldehyde. Sonication and MNase digestion are two standard procedures of chromatin release (O'Neill and Turner, 2003). Previous work in the lab has established MNase sensitivity of DNA sequences at IE genes, *gapdh* and *hbb*, while their relative sensitivity to sonication is unknown (Edmunds et al., 2008). Nevertheless, several histone ChIP experiments in the lab have suggested that sonicated chromatin and MN digested chromatin perform almost identically in ChIP. HMGN1 and H3 ChIPs in this thesis have been performed with formaldehyde cross-linking and MNase digestion due to the observation that this procedure yields the best recovery of DNA sequences (table 6.1).

Figure 37. **Immunodepletion of HMGN1 from input chromatin in N-ChIP and X-ChIP.**
N-ChIPs and X-ChIPs were performed on chromatin prepared from one 145mm plate of confluent C3H 10 T1/2 cells. Increasing quantities of HMGN-1 C-terminal antibody [R371] were used to establish the amount of antibody required to immunoprecipitate HMGN1. Western blots were performed with HMGN1 C-terminal antibody [R371] and secondary anti-Rabbit antibodies were used for ECL visualization.

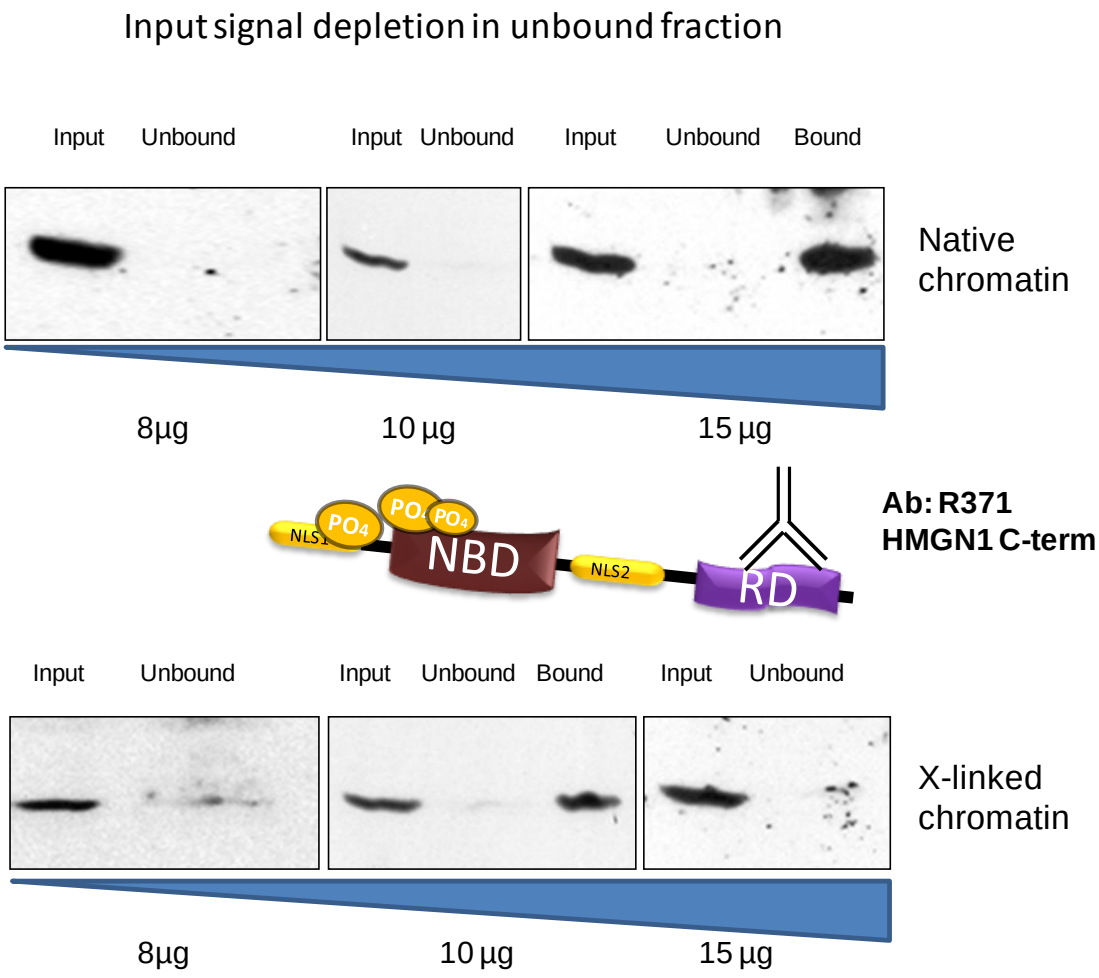


Figure 38. **Comparison of recoveries of ChIPs performed on chromatin X-linked at rT° or 37°C.**
C3H 10 T1/2 cells were quiesced 18h (o/n) and cross-linked either at rT° or 37°C. **a)** ChIP data from Figure 29. **b)** ChIP data from Figure 31.

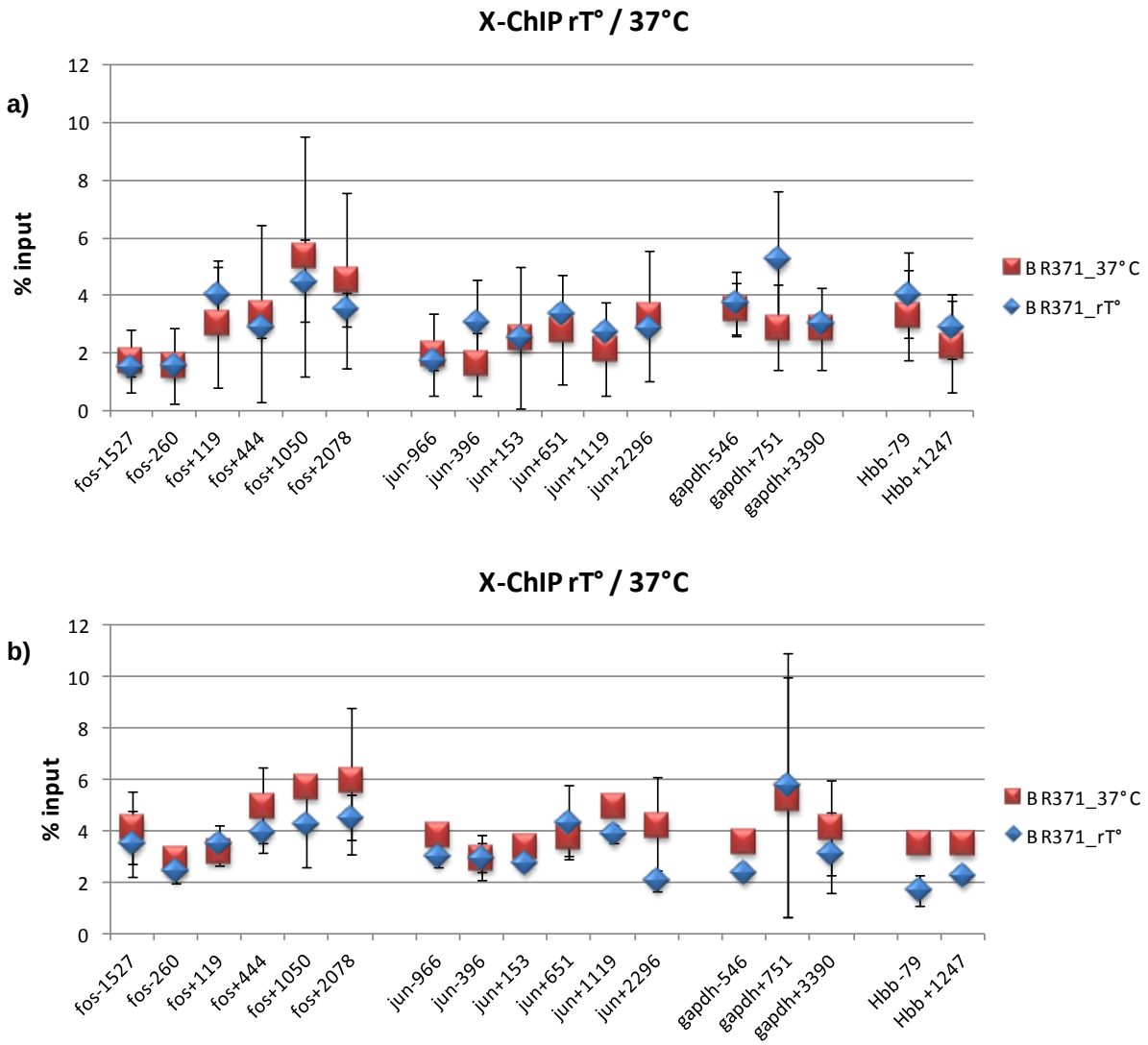


Figure 39. **Comparison of recoveries of ChIPs performed on chromatin X-linked at rT° or 37°C.**
C3H 10 T1/2 cells were quiesced 18h (o/n) and cross-linked either at rT° or 37°C. ChIP data from Figures 33-34. **a)** ChIPs performed with HMGN1-Cterminal antibody [R371] **b)** ChIPs performed with H3K9ac antibody.

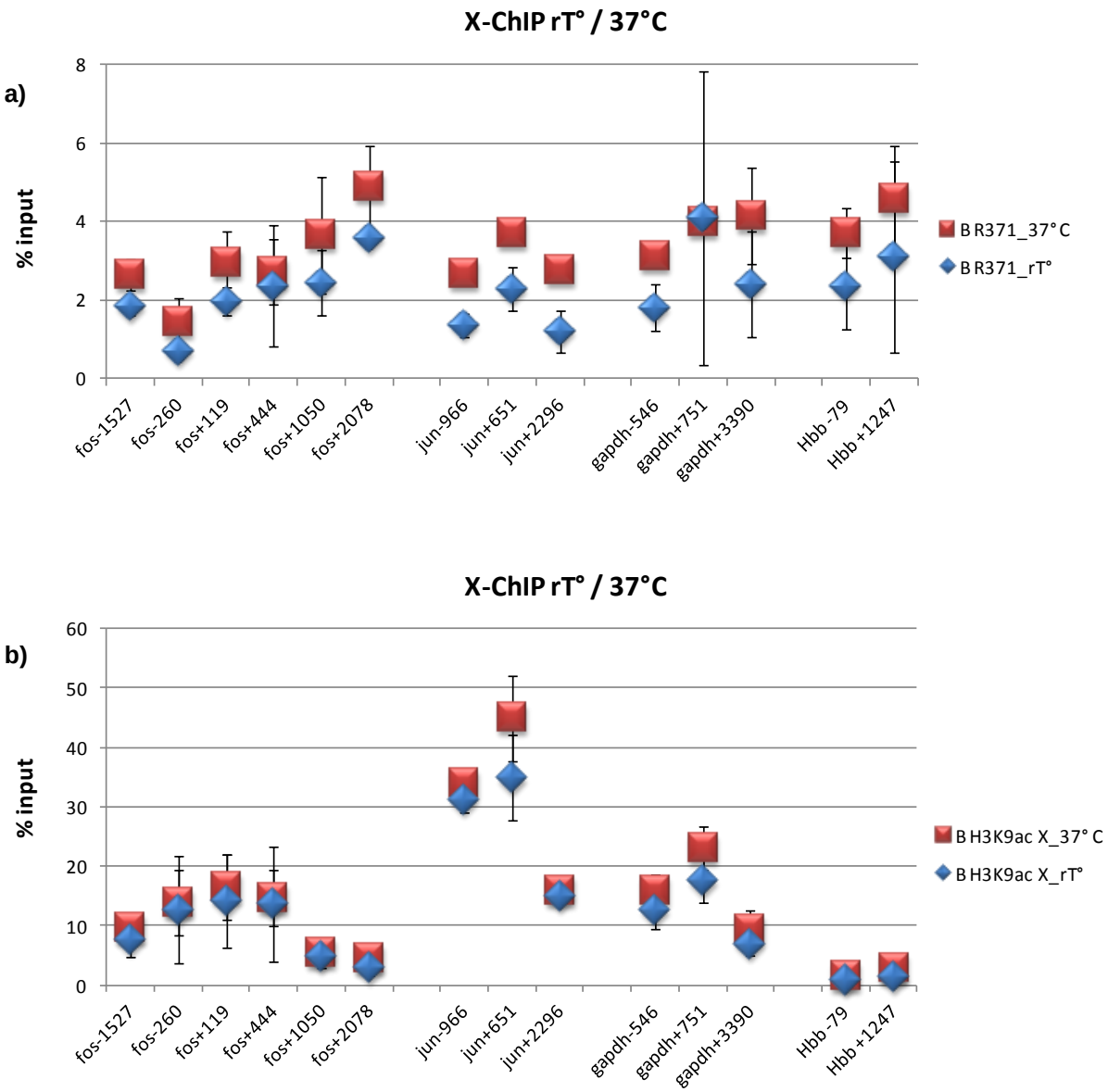


Figure 40. **Comparison of recoveries of ChIPs performed on chromatin X-linked at rT° or 37°C.**
C3H 10 T1/2 cells were quiesced 18h (o/n) and cross-linked either at rT° or 37°C. ChIP data from Figures 35-36. **a)** ChIPs performed with HMGN1-Cterminal antibody [R371] **b)** ChIPs performed with H3K9ac antibody **c)** ChIPs performed with H3K4me3 antibody.

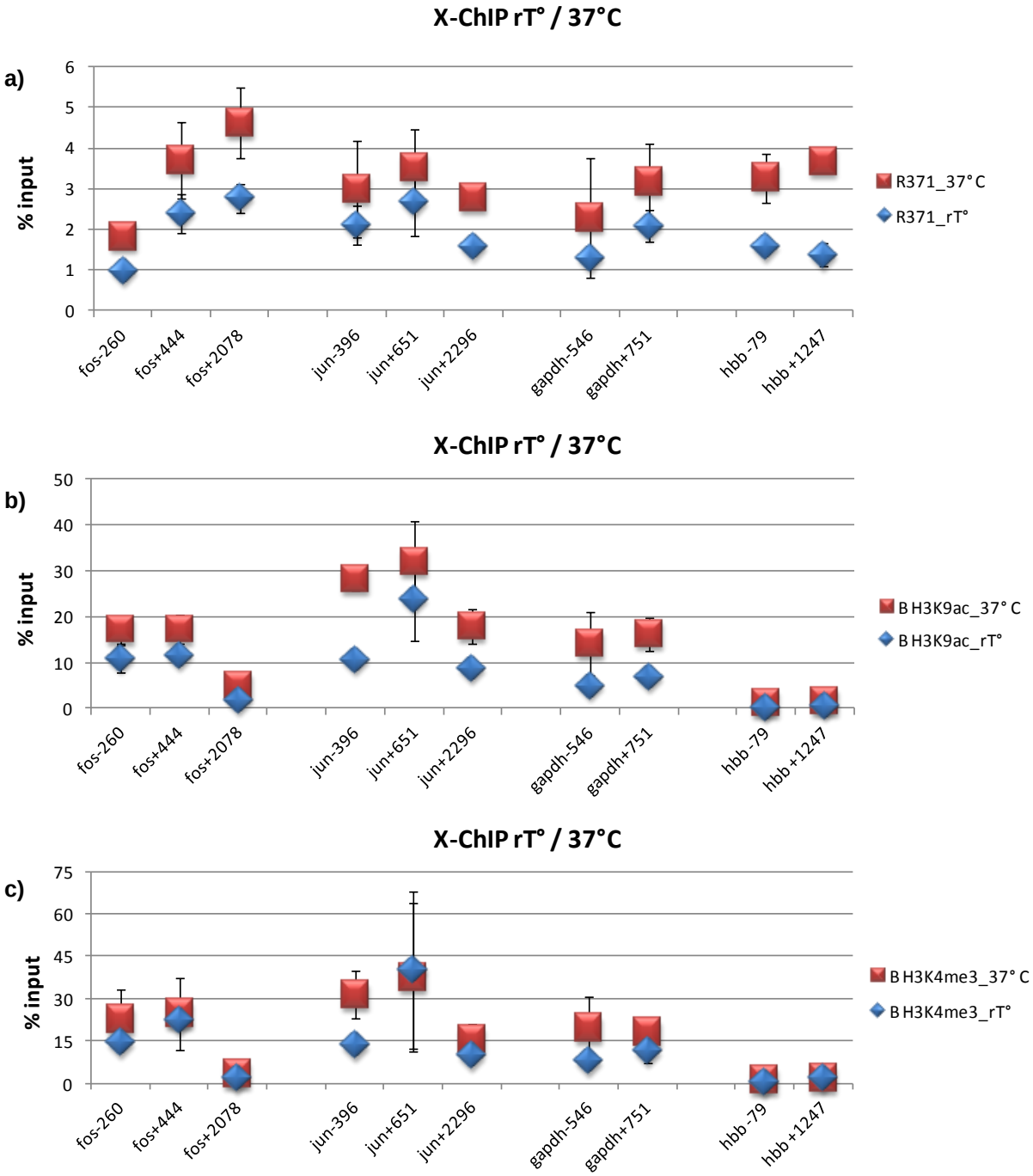


Table 6.1. **DMA and formaldehyde X-linking of MNase digested or sonicated chromatin.**

C3H 10 T1/2 cells were quiesced 18h (o/n) and stimulated for the stated time with An (10mg/ml). Cells were then cross-linked at rT° with either formaldehyde 1% or dimethyl adipimidate (DMA) 10mM. ChIPs were performed with HMGN1 C-terminal antibody [R371]. Recoveries of HMGN1-bound DNA sequences from: **a)** X-ChIP of formaldehyde (Form) cross-linked chromatin, digested with micrococcal nuclease (Mn) or sonicated (S); **b)** X-ChIP of formaldehyde (Form.) or DMA cross-linked chromatin, digested with MNase or sonicated [DMA_Form chromatin has been cross-linked first with DMA and then with formaldehyde].

a)	fos -1527	fos+444	fos+2078	jun-966	jun +651	jun +2296
	B/input %	*B/input %*	*B/input %*	*B/input %*	*B/input %*	*B/input %*
BR371 Mn_Form 0	*4,78*	*6,36*	*6,23*	*5,7*	*6,83*	*3,77*
BR371 Mn_Form 30 An	*4,81*	*6,26*	*5,56*	*5,28*	*5,1*	*3,75*
BR371 S_Form 0	*2,40*	*2,75*	*2,26*	*2,34*	*2,85*	*3,40*
BR371 S_Form 30 An	*4,19*	*3,23*	*3,40*	*3,63*	*3,65*	*4,18*

b)	fos -1527	fos+444	fos+2078	jun-966	jun +651	jun +2296
	B/input %	B/input %	B/input %	B/input %	B/input %	B/input %
BR371 Mn_Form 0	*4,78*	*6,36*	*6,23*	*5,7*	*6,83*	*3,77*
BR371 Mn_Form30 An	*4,81*	*6,26*	*5,56*	*5,28*	*5,1*	*3,75*
BR371 Mn_DMA 0	*0,08*	*0,19*	*0,27*	*0,08*	*0,13*	*0,51*
BR371 Mn_DMA 30 An	*0,11*	*0,26*	*0,29*	*0,14*	*0,18*	*0,48*
BR371 Mn_DMA_Form 0	*3,66*	*4,65*	*4,84*	*4,01*	*3,97*	*3,63*
BR371 Mn_DMA_Form 30 An	*2,96*	*3,36*	*3,56*	*3,12*	*3,03*	*1,96*
BR371 S_Form 0	*2,40*	*2,75*	*2,26*	*2,34*	*2,85*	*3,40*
BR371 S_Form 30 An	*4,19*	*3,23*	*3,40*	*3,63*	*3,65*	*4,18*
BR371 S_DMA_Form 0	*2,37*	*2,78*	*1,94*	*2,34*	*2,43*	*3,08*
BR371 S_DMA_Form 30 An	*1,60*	*1,92*	*1,39*	*1,46*	*1,70*	*1,85*

[Higher recovery ; Lower recovery ; Very low recovery]

Bibliography

Abercrombie, B.D., Kneale, G.G., Crane-Robinson, C., Bradbury, E.M., Goodwin, G.H., Walker, J.M., and Johns, E.W. (1978). Studies on the conformational properties of the high-mobility-group chromosomal protein HMG 17 and its interaction with DNA. Eur. J. Biochem. *84*, 173–177.

Abuhatzira, L., Shamir, A., Schones, D.E., Schäffer, A.A., and Bustin, M. (2011). The chromatin-binding protein HMGN1 regulates the expression of methyl CpG-binding protein 2 (MECP2) and affects the behavior of mice. J. Biol. Chem. *286*, 42051–42062.

Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. Nature *422*, 198–207.

Agresti, A., and Bianchi, M.E. (2003). HMGB proteins and gene expression. Curr. Opin. Genet. Dev. *13*, 170–178.

Agresti, A., Scaffidi, P., Riva, A., Caiolfa, V.R., and Bianchi, M.E. (2005). GR and HMGB1 interact only within chromatin and influence each other's residence time. Mol. Cell *18*, 109–121.

Ahn, S.-H., Cheung, W.L., Hsu, J.-Y., Diaz, R.L., Smith, M.M., and Allis, C.D. (2005). Sterile 20 kinase phosphorylates histone H2B at serine 10 during hydrogen peroxide-induced apoptosis in S. cerevisiae. Cell *120*, 25–36.

Aikawa, R., Nagai, T., Kudoh, S., Zou, Y., Tanaka, M., Tamura, M., Akazawa, H., Takano, H., Nagai, R., and Komuro, I. (2002). Integrins play a critical role in mechanical stress-induced p38 MAPK activation. Hypertension *39*, 233–238.

Alami, R., Fan, Y., Pack, S., Sonbuchner, T.M., Besse, A., Lin, Q., Greally, J.M., Skoultchi, A.I., and Bouhassira, E.E. (2003). Mammalian linker-histone subtypes differentially affect gene expression in vivo. Proc. Natl. Acad. Sci. U.S.A. *100*, 5920–5925.

Alexandrow, M.G., and Hamlin, J.L. (2005). Chromatin decondensation in S-phase involves recruitment of Cdk2 by Cdc45 and histone H1 phosphorylation. J. Cell Biol. *168*, 875–886.

Alfonso, P.J., Crippa, M.P., Hayes, J.J., and Bustin, M. (1994). The footprint of chromosomal proteins HMG-14 and HMG-17 on chromatin subunits. J. Mol. Biol. *236*, 189–198.

Allegra, P., Sterner, R., Clayton, D.F., and Allfrey, V.G. (1987). Affinity chromatographic purification of nucleosomes containing transcriptionally active DNA sequences. J. Mol. Biol. *196*, 379–388.

Almendral, J.M., Sommer, D., Macdonald-Bravo, H., Burckhardt, J., Perera, J., and Bravo, R. (1988). Complexity of the early genetic response to growth factors in mouse fibroblasts. Mol. Cell. Biol. *8*, 2140–2148.

Anand, A., and Chada, K. (2000). In vivo modulation of Hmgic reduces obesity. Nat. Genet. *24*, 377–380.

Angel, P., and Karin, M. (1991). The role of Jun, Fos and the AP-1 complex in cell-proliferation and transformation. Biochim. Biophys. Acta *1072*, 129–157.

Angel, P., Allegretto, E.A., Okino, S.T., Hattori, K., Boyle, W.J., Hunter, T., and Karin, M. (1988). Oncogene jun encodes a sequence-specific trans-activator similar to AP-1. Nature *332*, 166–171.

Arthur, J.S., and Cohen, P. (2000). MSK1 is required for CREB phosphorylation in response to mitogens in mouse embryonic stem cells. FEBS Lett. *482*, 44–48.

Attema, J.L., Reeves, R., Murray, V., Levichkin, I., Temple, M.D., Tremethick, D.J., and Shannon, M.F. (2002). The human IL-2 gene promoter can assemble a positioned nucleosome that becomes remodeled upon T cell activation. J. Immunol. *169*, 2466–2476.

Baatout, S., and Derradji, H. (2006). About histone H1 phosphorylation during mitosis. Cell Biochem. Funct. *24*, 93–94.

Baeuerle, P.A., and Baltimore, D. (1996). NF-kappa B: ten years after. Cell *87*, 13–20.

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

Bain, J., Plater, L., Elliott, M., Shpiro, N., Hastie, C.J., McLauchlan, H., Klevernic, I., Arthur, J.S.C., Alessi, D.R., and Cohen, P. (2007). The selectivity of protein kinase inhibitors: a further update. Biochem. J. *408*, 297–315.

Bannister, A.J., Oehler, T., Wilhelm, D., Angel, P., and Kouzarides, T. (1995). Stimulation of c-Jun activity by CBP: c-Jun residues Ser63/73 are required for CBP induced stimulation in vivo and CBP binding in vitro. Oncogene *11*, 2509–2514.

Barkess, G., Postnikov, Y., Campos, C.D., Mishra, S., Mohan, G., Verma, S., Bustin, M., and West, K.L. (2012). The chromatin-binding protein HMGN3 stimulates histone acetylation and transcription across the Glyt1 gene. Biochem. J. *442*, 495–505.

Barratt, M.J., Hazzalin, C.A., Zhelev, N., and Mahadevan, L.C. (1994). A mitogen- and anisomycin-stimulated kinase phosphorylates HMG-14 in its basic amino-terminal domain in vivo and on isolated mononucleosomes. EMBO J. *13*, 4524–4535.

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*, 823–837.

Bastos, G.M., Gonçalves, P.B.D., and Bordignon, V. (2008). Immunolocalization of the high-mobility group N2 protein and acetylated histone H3K14 in early developing parthenogenetic bovine embryos derived from oocytes of high and low developmental competence. Mol. Reprod. Dev. *75*, 282–290.

Batycka, M., Inglis, N.F., Cook, K., Adam, A., Fraser-Pitt, D., Smith, D.G.E., Main, L., Lubben, A., and Kessler, B.M. (2006). Ultra-fast tandem mass spectrometry scanning combined with monolithic column liquid chromatography increases throughput in proteomic analysis. Rapid Commun. Mass Spectrom. *20*, 2074–2080.

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*, 14173–14178.

Bell, A.C., West, A.G., and Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell *98*, 387–396.

Belova, G.I., Postnikov, Y.V., Furusawa, T., Birger, Y., and Bustin, M. (2008). Chromosomal protein HMGN1 enhances the heat shock-induced remodeling of Hsp70 chromatin. J. Biol. Chem. *283*, 8080–8088.

Bergel, M., Herrera, J.E., Thatcher, B.J., Prymakowska-Bosak, M., Vassilev, A., Nakatani, Y., Martin, B., and Bustin, M. (2000). Acetylation of novel sites in the nucleosomal binding domain of chromosomal protein HMG-14 by p300 alters its interaction with nucleosomes. J. Biol. Chem. *275*, 11514–11520.

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

Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas, E.J., 3rd, Gingeras, T.R., et al. (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. Cell *120*, 169–181.

Bernstein, B.E., Meissner, A., and Lander, E.S. (2007). The mammalian epigenome. Cell *128*, 669–681.

Bianchi, M.E., and Agresti, A. (2005). HMG proteins: dynamic players in gene regulation and differentiation. Curr. Opin. Genet. Dev. *15*, 496–506.

Binétruy, B., Smeal, T., and Karin, M. (1991). Ha-Ras augments c-Jun activity and stimulates phosphorylation of its activation domain. Nature *351*, 122–127.

Bird, A. (2008). The methyl-CpG-binding protein MeCP2 and neurological disease. Biochem. Soc. Trans. *36*, 575–583.

Birger, Y., West, K.L., Postnikov, Y.V., Lim, J.-H., Furusawa, T., Wagner, J.P., Laufer, C.S., Kraemer, K.H., and Bustin, M. (2003). Chromosomal protein HMGN1 enhances the rate of DNA repair in chromatin. EMBO J. *22*, 1665–1675.

Birger, Y., Catez, F., Furusawa, T., Lim, J.-H., Prymakowska-Bosak, M., West, K.L., Postnikov, Y.V., Haines, D.C., and Bustin, M. (2005). Increased tumorigenicity and sensitivity to ionizing radiation upon loss of chromosomal protein HMGN1. Cancer Res. *65*, 6711–6718.

Birger, Y., Davis, J., Furusawa, T., Rand, E., Piatigorsky, J., and Bustin, M. (2006). A role for chromosomal protein HMGN1 in corneal maturation. Differentiation *74*, 19–29.

Boggs, B.A., Allis, C.D., and Chinault, A.C. (2000). Immunofluorescent studies of human chromosomes with antibodies against phosphorylated H1 histone. Chromosoma *108*, 485–490.

Bohmann, D., and Tjian, R. (1989). Biochemical analysis of transcriptional activation by Jun: differential activity of c- and v-Jun. Cell *59*, 709–717.

Bonaldi, T., Längst, G., Strohner, R., Becker, P.B., and Bianchi, M.E. (2002). The DNA chaperone HMGB1 facilitates ACF/CHRAC-dependent nucleosome sliding. EMBO J. *21*, 6865–6873.

Bonaldi, T., Talamo, F., Scaffidi, P., Ferrera, D., Porto, A., Bachi, A., Rubartelli, A., Agresti, A., and Bianchi, M.E. (2003). Monocytic cells hyperacetylate chromatin protein HMGB1 to redirect it towards secretion. EMBO J. *22*, 5551–5560.

Boumba, V.A., Seferiadis, K., and Vougiouklakis, T. (2004). Content of the HMG-17 chromosomal protein in porcine tissues. Protein Pept. Lett. *11*, 515–519.

Boveri, T. (1904). Ergebnisse über die Konstitution der chromatischen Substanz des Zellkerns (Jena, Germany: G. Fischer).

Boyle, W.J., Smeal, T., Defize, L.H., Angel, P., Woodgett, J.R., Karin, M., and Hunter, T. (1991). Activation of protein kinase C decreases phosphorylation of c-Jun at sites that negatively regulate its DNA-binding activity. Cell *64*, 573–584.

Bradbury, E.M., Inglis, R.J., and Matthews, H.R. (1974). Control of cell division by very lysine rich histone (F1) phosphorylation. Nature *247*, 257–261.

Bravo, R. (1990). Growth factor-responsive genes in fibroblasts. Cell Growth Differ. *1*, 305–309.

Briggs, S.D., Bryk, M., Strahl, B.D., Cheung, W.L., Davie, J.K., Dent, S.Y., Winston, F., and Allis, C.D. (2001). Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in Saccharomyces cerevisiae. Genes Dev. *15*, 3286–3295.

Brown, D.T. (2003). Histone H1 and the dynamic regulation of chromatin function. Biochem. Cell Biol. *81*, 221–227.

Brown, C.E., Lechner, T., Howe, L., and Workman, J.L. (2000). The many HATs of transcription coactivators. Trends Biochem. Sci. *25*, 15–19.

Bruning, J.C., Gillette, J.A., Zhao, Y., Bjorbaeck, C., Kotzka, J., Knebel, B., Avci, H., Hanstein, B., Lingohr, P., Moller, D.E., et al. (2000). Ribosomal subunit kinase-2 is required for growth factor-stimulated transcription of the c-Fos gene. Proc. Natl. Acad. Sci. U.S.A. *97*, 2462–2467.

Bulger, M., and Groudine, M. (1999). Looping versus linking: toward a model for long-distance gene activation. Genes Dev. *13*, 2465–2477.

Bustin, M. (1999). Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. Mol. Cell. Biol. *19*, 5237–5246.

Bustin, M. (2001a). Revised nomenclature for high mobility group (HMG) chromosomal proteins. Trends Biochem. Sci. *26*, 152–153.

Bustin, M. (2001b). Chromatin unfolding and activation by HMGN(*) chromosomal proteins. Trends Biochem. Sci. *26*, 431–437.

Bustin, M., and Reeves, R. (1996). High-mobility-group chromosomal proteins: architectural components that facilitate chromatin function. Prog. Nucleic Acid Res. Mol. Biol. *54*, 35–100.

Bustin, M., Catez, F., and Lim, J.-H. (2005). The dynamics of histone H1 function in chromatin. Mol. Cell *17*, 617–620.

Cai, Y., Jin, J., Yao, T., Gottschalk, A.J., Swanson, S.K., Wu, S., Shi, Y., Washburn, M.P., Florens, L., Conaway, R.C., et al. (2007). YY1 functions with INO80 to activate transcription. Nat. Struct. Mol. Biol. *14*, 872–874.

Campos, E.I., and Reinberg, D. (2009). Histones: annotating chromatin. Annu. Rev. Genet. *43*, 559–599.

Cano, E., and Mahadevan, L.C. (1995). Parallel signal processing among mammalian MAPKs. Trends Biochem. Sci. *20*, 117–122.

Cano, E., Hazzalin, C.A., Kardalinou, E., Buckle, R.S., and Mahadevan, L.C. (1995). Neither ERK nor JNK/SAPK MAP kinase subtypes are essential for histone H3/HMG-14 phosphorylation or c-fos and c-jun induction. J. Cell. Sci. *108 (Pt 11)*, 3599–3609.

Catez, F., Brown, D.T., Misteli, T., and Bustin, M. (2002). Competition between histone H1 and HMGN proteins for chromatin binding sites. EMBO Rep. *3*, 760–766.

Catez, F., Yang, H., Tracey, K.J., Reeves, R., Misteli, T., and Bustin, M. (2004). Network of dynamic interactions between histone H1 and high-mobility-group proteins in chromatin. Mol. Cell. Biol. *24*, 4321–4328.

Catez, F., Ueda, T., and Bustin, M. (2006). Determinants of histone H1 mobility and chromatin binding in living cells. Nat. Struct. Mol. Biol. *13*, 305–310.

Cetkovic H, Lukic-Bilela L (2003). HMGB2 protein from the marine sponge Suberites domuncula. Journal of the Faculty of Food Technology and Biotechnology, University of Zagreb, Zagreb, Croatia *41*, 361–365.

Chadee, D.N., Taylor, W.R., Hurta, R.A., Allis, C.D., Wright, J.A., and Davie, J.R. (1995). Increased phosphorylation of histone H1 in mouse fibroblasts transformed with oncogenes or constitutively active mitogen-activated protein kinase kinase. J. Biol. Chem. *270*, 20098–20105.

Chakalova, L., Debrand, E., Mitchell, J.A., Osborne, C.S., and Fraser, P. (2005). Replication and transcription: shaping the landscape of the genome. Nat. Rev. Genet. *6*, 669–677.

Chandy, M., Gutiérrez, J.L., Prochasson, P., and Workman, J.L. (2006). SWI/SNF displaces SAGA-acetylated nucleosomes. Eukaryotic Cell *5*, 1738–1747.

Chau, K.-Y., Keane-Myers, A.M., Fedele, M., Ikeda, Y., Creusot, R.J., Menozzi, L., Cousins, D.J., Manfioletti, G., Feigenbaum, L., Fusco, A., et al. (2005). IFN-gamma gene expression is controlled by the architectural transcription factor HMGA1. Int. Immunol. *17*, 297–306.

Chen, T.A., and Allfrey, V.G. (1987). Rapid and reversible changes in nucleosome structure accompany the activation, repression, and superinduction of murine fibroblast protooncogenes c-fos and c-myc. Proc. Natl. Acad. Sci. U.S.A. *84*, 5252–5256.

Chen, R.H., Abate, C., and Blenis, J. (1993). Phosphorylation of the c-Fos transrepression domain by mitogen-activated protein kinase and 90-kDa ribosomal S6 kinase. Proc. Natl. Acad. Sci. U.S.A. *90*, 10952–10956.

Cherukuri, S., Hock, R., Ueda, T., Catez, F., Rochman, M., and Bustin, M. (2008). Cell cycle-dependent binding of HMGN proteins to chromatin. Mol. Biol. Cell *19*, 1816–1824.

Cheung, E., Zarifyan, A.S., and Kraus, W.L. (2002). Histone H1 represses estrogen receptor alpha transcriptional activity by selectively inhibiting receptor-mediated transcription initiation. Mol. Cell. Biol. *22*, 2463–2471.

Cheung, P., Tanner, K.G., Cheung, W.L., Sassone-Corsi, P., Denu, J.M., and Allis, C.D. (2000). Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation. Mol. Cell *5*, 905–915.

Clarke, N., Arenzana, N., Hai, T., Minden, A., and Prywes, R. (1998). Epidermal growth factor induction of the c-jun promoter by a Rac pathway. Mol. Cell. Biol. *18*, 1065–1073.

Clayton, A.L., and Mahadevan, L.C. (2003). MAP kinase-mediated phosphoacetylation of histone H3 and inducible gene regulation. FEBS Lett. *546*, 51–58.

Clayton, A.L., Rose, S., Barratt, M.J., and Mahadevan, L.C. (2000). Phosphoacetylation of histone H3 on c-fos- and c-jun-associated nucleosomes upon gene activation. EMBO J. *19*, 3714–3726.

Clayton, A.L., Hazzalin, C.A., and Mahadevan, L.C. (2006). Enhanced histone acetylation and transcription: a dynamic perspective. Mol. Cell *23*, 289–296.

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*, 461–473.

Cochran, B.H., Reffel, A.C., and Stiles, C.D. (1983). Molecular cloning of gene sequences regulated by platelet-derived growth factor. Cell *33*, 939–947.

Contreras, A., Hale, T.K., Stenoien, D.L., Rosen, J.M., Mancini, M.A., and Herrera, R.E. (2003). The dynamic mobility of histone H1 is regulated by cyclin/CDK phosphorylation. Mol. Cell. Biol. *23*, 8626–8636.

Cook, P.R. (2002). Predicting three-dimensional genome structure from transcriptional activity. Nat. Genet. *32*, 347–352.

Cooper, E., Palmer, R.J., and Spaulding, S.W. (1982). TSH stimulates 32P-labeling of thyroid nuclear HMG 14, a protein associated with actively transcribed chromatin. Endocrinology *110*, 1459–1461.

Coussons-Read, M.E., and Crnic, L.S. (1996). Behavioral assessment of the Ts65Dn mouse, a model for Down syndrome: altered behavior in the elevated plus maze and open field. Behav. Genet. *26*, 7–13.

Crippa, M.P., Nickol, J.M., and Bustin, M. (1991). Developmental changes in the expression of high mobility group chromosomal proteins. J. Biol. Chem. *266*, 2712–2714.

Crippa, M.P., Alfonso, P.J., and Bustin, M. (1992). Nucleosome core binding region of chromosomal protein HMG-17 acts as an independent functional domain. J. Mol. Biol. *228*, 442–449.

Crosio, C., Fimia, G.M., Loury, R., Kimura, M., Okano, Y., Zhou, H., Sen, S., Allis, C.D., and Sassone-Corsi, P. (2002). Mitotic phosphorylation of histone H3: spatio-temporal regulation by mammalian Aurora kinases. Mol. Cell. Biol. *22*, 874–885.

Crump, N.T., Hazzalin, C.A., Bowers, E.M., Alani, R.M., Cole, P.A., and Mahadevan, L.C. (2011). Dynamic acetylation of all lysine-4 trimethylated histone H3 is evolutionarily conserved and mediated by p300/CBP. Proc. Natl. Acad. Sci. U.S.A. *108*, 7814–7819.

Cuddapah, S., Jothi, R., Schones, D.E., Roh, T.-Y., Cui, K., and Zhao, K. (2009). Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. Genome Res. *19*, 24–32.

Cuddapah, S., Schones, D.E., Cui, K., Roh, T.-Y., Barski, A., Wei, G., Rochman, M., Bustin, M., and Zhao, K. (2011). Genomic profiling of HMGN1 reveals an association with chromatin at regulatory regions. Mol. Cell. Biol. *31*, 700–709.

Van Dam, H., Wilhelm, D., Herr, I., Steffen, A., Herrlich, P., and Angel, P. (1995). ATF-2 is preferentially activated by stress-activated protein kinases to mediate c-jun induction in response to genotoxic agents. EMBO J. *14*, 1798–1811.

Darnell, J.E., Jr (1997). STATs and gene regulation. Science *277*, 1630–1635.

Darville, M.I., Terryn, S., and Eizirik, D.L. (2004). An octamer motif is required for activation of the inducible nitric oxide synthase promoter in pancreatic beta-cells. Endocrinology *145*, 1130–1136.

Das, D., and Scovell, W.M. (2001). The binding interaction of HMG-1 with the TATA-binding protein/TATA complex. J. Biol. Chem. *276*, 32597–32605.

Dasgupta, A., and Scovell, W.M. (2003). TFIIA abrogates the effects of inhibition by HMGB1 but not E1A during the early stages of assembly of the transcriptional preinitiation complex. Biochim. Biophys. Acta *1627*, 101–110.

Daujat, S., Zeissler, U., Waldmann, T., Happel, N., and Schneider, R. (2005). HP1 binds specifically to Lys26-methylated histone H1.4, whereas simultaneous Ser27 phosphorylation blocks HP1 binding. J. Biol. Chem. *280*, 38090–38095.

Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W., and Richmond, T.J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. J. Mol. Biol. *319*, 1097–1113.

Davies, S.P., Reddy, H., Caivano, M., and Cohen, P. (2000). Specificity and mechanism of action of some commonly used protein kinase inhibitors. Biochem. J. *351*, 95–105.

De, S., Brown, D.T., Lu, Z.H., Leno, G.H., Wellman, S.E., and Sittman, D.B. (2002). Histone H1 variants differentially inhibit DNA replication through an affinity for chromatin mediated by their carboxyl-terminal domains. Gene *292*, 173–181.

Deak, M., Clifton, A.D., Lucocq, L.M., and Alessi, D.R. (1998). Mitogen- and stress-activated protein kinase-1 (MSK1) is directly activated by MAPK and SAPK2/p38, and may mediate activation of CREB. EMBO J. *17*, 4426–4441.

Déjardin, J., Rappailles, A., Cuvier, O., Grimaud, C., Decoville, M., Locker, D., and Cavalli, G. (2005). Recruitment of Drosophila Polycomb group proteins to chromatin by DSP1. Nature *434*, 533–538.

Deterding, L.J., Bunger, M.K., Banks, G.C., Tomer, K.B., and Archer, T.K. (2008). Global changes in and characterization of specific sites of phosphorylation in mouse and human histone H1 Isoforms upon CDK inhibitor treatment using mass spectrometry. J. Proteome Res. *7*, 2368–2379.

Ding, H.F., Bustin, M., and Hansen, U. (1997). Alleviation of histone H1-mediated transcriptional repression and chromatin compaction by the acidic activation region in chromosomal protein HMG-14. Mol. Cell. Biol. *17*, 5843–5855.

Dion, M.F., Altschuler, S.J., Wu, L.F., and Rando, O.J. (2005). Genomic characterization reveals a simple histone H4 acetylation code. Proc. Natl. Acad. Sci. U.S.A. *102*, 5501–5506.

Dou, Y., Bowen, J., Liu, Y., and Gorovsky, M.A. (2002). Phosphorylation and an ATP-dependent process increase the dynamic exchange of H1 in chromatin. J. Cell Biol. *158*, 1161–1170.

Dyson, M.H., Thomson, S., Inagaki, M., Goto, H., Arthur, S.J., Nightingale, K., Iborra, F.J., and Mahadevan, L.C. (2005a). MAP kinase-mediated phosphorylation of distinct pools of histone H3 at S10 or S28 via mitogen- and stress-activated kinase 1/2. J. Cell. Sci. *118*, 2247–2259.

Dyson, M.H., Thomson, S., and Mahadevan, L.C. (2005b). Heat shock, histone H3 phosphorylation and the cell cycle. Cell Cycle *4*, 13–17.

Edmondson, D.G., Davie, J.K., Zhou, J., Mirnikjoo, B., Tatchell, K., and Dent, S.Y.R. (2002). Site-specific loss of acetylation upon phosphorylation of histone H3. J. Biol. Chem. *277*, 29496–29502.

Edmunds, J.W., and Mahadevan, L.C. (2004). MAP kinases as structural adaptors and enzymatic activators in transcription complexes. J. Cell. Sci. *117*, 3715–3723.

Edmunds, J.W., Mahadevan, L.C., and Clayton, A.L. (2008). Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. EMBO J. *27*, 406–420.

Edwards, D.R., and Mahadevan, L.C. (1992). Protein synthesis inhibitors differentially superinduce c-fos and c-jun by three distinct mechanisms: lack of evidence for labile repressors. EMBO J. *11*, 2415–2424.

Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. (2002). Stochastic gene expression in a single cell. Science *297*, 1183–1186.

Farooq, A., and Zhou, M.-M. (2004). Structure and regulation of MAPK phosphatases. Cell. Signal. *16*, 769–779.

Fedele, M., Battista, S., Manfioletti, G., Croce, C.M., Giancotti, V., and Fusco, A. (2001). Role of the high mobility group A proteins in human lipomas. Carcinogenesis *22*, 1583–1591.

Fedoriw, A.M., Stein, P., Svoboda, P., Schultz, R.M., and Bartolomei, M.S. (2004). Transgenic RNAi reveals essential function for CTCF in H19 gene imprinting. Science *303*, 238–240.

Feng, J., and Villeponteau, B. (1992). High-resolution analysis of c-fos chromatin accessibility using a novel DNase I-PCR assay. Biochim. Biophys. Acta *1130*, 253–258.

Filippova, G.N. (2008). Genetics and epigenetics of the multifunctional protein CTCF. Curr. Top. Dev. Biol. *80*, 337–360.

Fiorillo, A.A., Medler, T.R., Feeney, Y.B., Liu, Y., Tommerdahl, K.L., and Clevenger, C.V. (2011). HMGN2 inducibly binds a novel transactivation domain in nuclear PRLr to coordinate Stat5a-mediated transcription. Mol. Endocrinol. *25*, 1550–1564.

Fisch, T.M., Prywes, R., and Roeder, R.G. (1989). An AP1-binding site in the c-fos gene can mediate induction by epidermal growth factor and 12-O-tetradecanoyl phorbol-13-acetate. Mol. Cell. Biol. *9*, 1327–1331.

Fischle, W., Tseng, B.S., Dormann, H.L., Ueberheide, B.M., Garcia, B.A., Shabanowitz, J., Hunt, D.F., Funabiki, H., and Allis, C.D. (2005). Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. Nature *438*, 1116–1122.

Flemming, W. (1879). Beitrage zur Kenntniss der Zelle und ihrer Lebenserscheinungen. Archiv Für Mikroskopische Anatomie *16*, 302–436.

Flemming, W. (1882). Zellsubstanz, Kern und Zelltheilung (Leipzig, Germany:: Verlag von F.C.W. Vogel).

Foti, D., Chiefari, E., Fedele, M., Iuliano, R., Brunetti, L., Paonessa, F., Manfioletti, G., Barbetti, F., Brunetti, A., Croce, C.M., et al. (2005). Lack of the architectural factor HMGA1 causes insulin resistance and diabetes in humans and mice. Nat. Med. *11*, 765–773.

Fousteri, M., Vermeulen, W., van Zeeland, A.A., and Mullenders, L.H.F. (2006). Cockayne syndrome A and B proteins differentially regulate recruitment of chromatin remodeling and repair factors to stalled RNA polymerase II in vivo. Mol. Cell *23*, 471–482.

Freshney, N.W., Rawlinson, L., Guesdon, F., Jones, E., Cowley, S., Hsuan, J., and Saklatvala, J. (1994). Interleukin-1 activates a novel protein kinase cascade that results in the phosphorylation of Hsp27. Cell *78*, 1039–1049.

Frödin, M., and Gammeltoft, S. (1999). Role and regulation of 90 kDa ribosomal S6 kinase (RSK) in signal transduction. Mol. Cell. Endocrinol. *151*, 65–77.

Furusawa, T., and Cherukuri, S. (2010). Developmental function of HMGN proteins. Biochim. Biophys. Acta *1799*, 69–73.

Furusawa, T., Lim, J.-H., Catez, F., Birger, Y., Mackem, S., and Bustin, M. (2006). Down-regulation of nucleosomal binding protein HMGN1 expression during embryogenesis modulates Sox9 expression in chondrocytes. Mol. Cell. Biol. *26*, 592–604.

Furusawa, T., Ko, J.-H., Birger, Y., and Bustin, M. (2009). Expression of nucleosomal protein HMGN1 in the cycling mouse hair follicle. Gene Expr. Patterns *9*, 289–295.

Gabellini, D., Green, M.R., and Tupler, R. (2002). Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. Cell *110*, 339–348.

Garcia, B.A., Busby, S.A., Barber, C.M., Shabanowitz, J., Allis, C.D., and Hunt, D.F. (2004). Characterization of phosphorylation sites on histone H1 isoforms by tandem mass spectrometry. J. Proteome Res. *3*, 1219–1227.

Garcia, B.A., Joshi, S., Thomas, C.E., Chitta, R.K., Diaz, R.L., Busby, S.A., Andrews, P.C., Ogorzalek Loo, R.R., Shabanowitz, J., Kelleher, N.L., et al. (2006). Comprehensive phosphoprotein analysis of linker histone H1 from Tetrahymena thermophila. Mol. Cell Proteomics *5*, 1593–1609.

Gemelli, T., Berton, O., Nelson, E.D., Perrotti, L.I., Jaenisch, R., and Monteggia, L.M. (2006). Postnatal loss of methyl-CpG binding protein 2 in the forebrain is sufficient to mediate behavioral aspects of Rett syndrome in mice. Biol. Psychiatry *59*, 468–476.

Gerlitz, G. (2010). HMGNs, DNA repair and cancer. Biochim. Biophys. Acta *1799*, 80–85.

Gerlitz, G., Hock, R., Ueda, T., and Bustin, M. (2009). The dynamics of HMG protein-chromatin interactions in living cells. Biochem. Cell Biol. *87*, 127–137.

Goodwin, G.H., Sanders, C., and Johns, E.W. (1973). A new group of chromatin-associated proteins with a high content of acidic and basic amino acids. Eur. J. Biochem. *38*, 14–19.

Goto, H., Yasui, Y., Nigg, E.A., and Inagaki, M. (2002). Aurora-B phosphorylates Histone H3 at serine28 with regard to the mitotic chromosome condensation. Genes Cells *7*, 11–17.

Goytisolo, F.A., Gerchman, S.E., Yu, X., Rees, C., Graziano, V., Ramakrishnan, V., and Thomas, J.O. (1996). Identification of two DNA-binding sites on the globular domain of histone H5. EMBO J. *15*, 3421–3429.

Green, G.R., Lee, H.J., and Poccia, D.L. (1993). Phosphorylation weakens DNA binding by peptides containing multiple "SPKK" sequences. J. Biol. Chem. *268*, 11247–11255.

Grigoryev, S.A., Bulynko, Y.A., and Popova, E.Y. (2006). The end adjusts the means: heterochromatin remodelling during terminal cell differentiation. Chromosome Res. *14*, 53–69.

De Groot, R.P., Auwerx, J., Bourouis, M., and Sassone-Corsi, P. (1993). Negative regulation of Jun/AP-1: conserved function of glycogen synthase kinase 3 and the Drosophila kinase shaggy. Oncogene *8*, 841–847.

Groth, A., Rocha, W., Verreault, A., and Almouzni, G. (2007). Chromatin challenges during DNA replication and repair. Cell *128*, 721–733.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130*, 77–88.

Gurley, L.R., D'Anna, J.A., Barham, S.S., Deaven, L.L., and Tobey, R.A. (1978). Histone phosphorylation and chromatin structure during mitosis in Chinese hamster cells. Eur. J. Biochem. *84*, 1–15.

Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte, M., Zuk, O., Carey, B.W., Cassady, J.P., et al. (2009). Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature *458*, 223–227.

Guy, J., Hendrich, B., Holmes, M., Martin, J.E., and Bird, A. (2001). A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. Nat. Genet. *27*, 322–326.

Hacques, M.F., Muller, S., De Murcia, G., Van Regenmortel, M.H., and Marion, C. (1990). Use of an immobilized enzyme and specific antibodies to analyse the accessibility and role of histone tails in chromatin structure. Biochem. Biophys. Res. Commun. *168*, 637–643.

Hagemann, C., and Rapp, U.R. (1999). Isotype-specific functions of Raf kinases. Exp. Cell Res. *253*, 34–46.

Han, T.H., and Prywes, R. (1995). Regulatory role of MEF2D in serum induction of the c-jun promoter. Mol. Cell. Biol. *15*, 2907–2915.

Han, J., Lee, J.D., Bibbs, L., and Ulevitch, R.J. (1994). A MAP kinase targeted by endotoxin and hyperosmolarity in mammalian cells. Science *265*, 808–811.

Han, J., Jiang, Y., Li, Z., Kravchenko, V.V., and Ulevitch, R.J. (1997). Activation of the transcription factor MEF2C by the MAP kinase p38 in inflammation. Nature *386*, 296–299.

Happel, N., and Doenecke, D. (2009). Histone H1 and its isoforms: contribution to chromatin structure and function. Gene *431*, 1–12.

Harrer, M., Lührs, H., Bustin, M., Scheer, U., and Hock, R. (2004). Dynamic interaction of HMGA1a proteins with chromatin. J. Cell. Sci. *117*, 3459–3471.

Hassan, A.H., Awad, S., and Prochasson, P. (2006). The Swi2/Snf2 bromodomain is required for the displacement of SAGA and the octamer transfer of SAGA-acetylated nucleosomes. J. Biol. Chem. *281*, 18126–18134.

Hayes, J.J., Kaplan, R., Ura, K., Pruss, D., and Wolffe, A. (1996). A putative DNA binding surface in the globular domain of a linker histone is not essential for specific binding to the nucleosome. J. Biol. Chem. *271*, 25817–25822.

Hazzalin, C.A., and Mahadevan, L.C. (2002). MAPK-regulated transcription: a continuously variable gene switch? Nat. Rev. Mol. Cell Biol. *3*, 30–40.

Hazzalin, C.A., and Mahadevan, L.C. (2005). Dynamic acetylation of all lysine 4-methylated histone H3 in the mouse nucleus: analysis at c-fos and c-jun. PLoS Biol. *3*, e393.

Hazzalin, C.A., Le Panse, R., Cano, E., and Mahadevan, L.C. (1998). Anisomycin selectively desensitizes signalling components involved in stress kinase activation and fos and jun induction. Mol. Cell. Biol. *18*, 1844–1854.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat. Genet. *39*, 311–318.

Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W., et al. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature *459*, 108–112.

Hellauer, K., Sirard, E., and Turcotte, B. (2001). Decreased expression of specific genes in yeast cells lacking histone H1. J. Biol. Chem. *276*, 13587–13592.

Hendzel, M.J., Lever, M.A., Crawford, E., and Th'ng, J.P.H. (2004). The C-terminal domain is the primary determinant of histone H1 binding to chromatin in vivo. J. Biol. Chem. *279*, 20028–20034.

Herr, I., van Dam, H., and Angel, P. (1994). Binding of promoter-associated AP-1 is not altered during induction and subsequent repression of the c-jun promoter by TPA and UV irradiation. Carcinogenesis *15*, 1105–1113.

Herrera, J.E., Sakaguchi, K., Bergel, M., Trieschmann, L., Nakatani, Y., and Bustin, M. (1999). Specific acetylation of chromosomal protein HMG-17 by PCAF alters its interaction with nucleosomes. Mol. Cell. Biol. *19*, 3466–3473.

Herrera, J.E., West, K.L., Schiltz, R.L., Nakatani, Y., and Bustin, M. (2000a). Histone H1 is a specific repressor of core histone acetylation in chromatin. Mol. Cell. Biol. *20*, 523–529.

Herrera, J.E., Schiltz, R.L., and Bustin, M. (2000b). The accessibility of histone H3 tails in chromatin modulates their acetylation by P300/CBP-associated factor. J. Biol. Chem. *275*, 12994–12999.

Hibi, M., Lin, A., Smeal, T., Minden, A., and Karin, M. (1993). Identification of an oncoprotein- and UV-responsive protein kinase that binds and potentiates the c-Jun activation domain. Genes Dev. *7*, 2135–2148.

Hill, D.A., and Imbalzano, A.N. (2000). Human SWI/SNF nucleosome remodeling activity is partially inhibited by linker histone H1. Biochemistry *39*, 11649–11656.

Hill, D.A., Peterson, C.L., and Imbalzano, A.N. (2005). Effects of HMGN1 on chromatin structure and SWI/SNF-mediated chromatin remodeling. J. Biol. Chem. *280*, 41777–41783.

Hirai, S.I., Ryseck, R.P., Mechta, F., Bravo, R., and Yaniv, M. (1989). Characterization of junD: a new member of the jun proto-oncogene family. EMBO J. *8*, 1433–1439.

Hock, R., Scheer, U., and Bustin, M. (1998a). Chromosomal proteins HMG-14 and HMG-17 are released from mitotic chromosomes and imported into the nucleus by active transport. J. Cell Biol. *143*, 1427–1436.

Hock, R., Wilde, F., Scheer, U., and Bustin, M. (1998b). Dynamic relocation of chromosomal protein HMG-17 in the nucleus is dependent on transcriptional activity. EMBO J. *17*, 6992–7001.

Hock, R., Furusawa, T., Ueda, T., and Bustin, M. (2007). HMG chromosomal proteins in development and disease. Trends Cell Biol. *17*, 72–79.

Hohmann, P., Tobey, R.A., and Gurley, L.R. (1976). Phosphorylation of distinct regions of f1 histone. Relationship to the cell cycle. J. Biol. Chem. *251*, 3685–3692.

Van Holde (1988). Chromatin. *Springer Verlag, New York, N.Y.*

Horn, P.J., Carruthers, L.M., Logie, C., Hill, D.A., Solomon, M.J., Wade, P.A., Imbalzano, A.N., Hansen, J.C., and Peterson, C.L. (2002). Phosphorylation of linker histones regulates ATP-dependent chromatin remodeling enzymes. Nat. Struct. Biol. *9*, 263–267.

Huang, Y., Fang, J., Bedford, M.T., Zhang, Y., and Xu, R.-M. (2006). Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. Science *312*, 748–751.

Ito, T., Ikehara, T., Nakagawa, T., Kraus, W.L., and Muramatsu, M. (2000). p300-mediated acetylation facilitates the transfer of histone H2A-H2B dimers from nucleosomes to a histone chaperone. Genes Dev. *14*, 1899–1907.

Jackson, D.A., Dickinson, P., and Cook, P.R. (1990). The size of chromatin loops in HeLa cells. EMBO J. *9*, 567–571.

Janga, S.C., Collado-Vides, J., and Babu, M.M. (2008). Transcriptional regulation constrains the organization of genes on eukaryotic chromosomes. Proc. Natl. Acad. Sci. U.S.A. *105*, 15761–15766.

Janknecht, R., and Nordheim, A. (1996). MAP kinase-dependent transcriptional coactivation by Elk-1 and its cofactor CBP. Biochem. Biophys. Res. Commun. *228*, 831–837.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074–1080.

Jiang, X.G., and Wang, Y. (2006). Phosphorylation of human high mobility group N1 protein by protein kinase CK2. Biochem. Biophys. Res. Commun. *345*, 1497–1503.

Jiang, T., Zhou, X., Taghizadeh, K., Dong, M., and Dedon, P.C. (2007). N-formylation of lysine in histone proteins as a secondary modification arising from oxidative DNA damage. Proc. Natl. Acad. Sci. U.S.A. *104*, 60–65.

Johns (1982). The HMG Chromosomal Proteins. Academic Press, London, 1982.

Karin, M. (1996). The regulation of AP-1 activity by mitogen-activated protein kinases. Philos. Trans. R. Soc. Lond., B, Biol. Sci. *351*, 127–134.

Karin, M., and Hunter, T. (1995). Transcriptional control by protein phosphorylation: signal transmission from the cell surface to the nucleus. Curr. Biol. *5*, 747–757.

Karin, M., Liu, Z. g, and Zandi, E. (1997). AP-1 function and regulation. Curr. Opin. Cell Biol. *9*, 240–246.

Kato, H., van Ingen, H., Zhou, B.-R., Feng, H., Bustin, M., Kay, L.E., and Bai, Y. (2011). Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR. Proc. Natl. Acad. Sci. U.S.A. *108*, 12283–12288.

Kelley, R.I. (1973). Isolation of a histone IIb1-IIb2 complex. Biochem. Biophys. Res. Commun. *54*, 1588–1594.

Kim, A., Kiefer, C.M., and Dean, A. (2007a). Distinctive signatures of histone methylation in transcribed coding and noncoding human beta-globin sequences. Mol. Cell. Biol. *27*, 1271–1279.

Kim, T.H., Abdullaev, Z.K., Smith, A.D., Ching, K.A., Loukinov, D.I., Green, R.D., Zhang, M.Q., Lobanenkov, V.V., and Ren, B. (2007b). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell *128*, 1231–1245.

Kim, Y.-C., Gerlitz, G., Furusawa, T., Catez, F., Nussenzweig, A., Oh, K.-S., Kraemer, K.H., Shiloh, Y., and Bustin, M. (2009). Activation of ATM depends on chromatin interactions occurring before induction of DNA damage. Nat. Cell Biol. *11*, 92–96.

Kinter, M., and Sherman, N.E. (2000). Protein Sequencing and Identification Using Tandem Mass Spectrometry (John Wiley & Sons).

Klose, R.J., and Bird, A.P. (2006). Genomic DNA methylation: the mark and its mediators. Trends Biochem. Sci. *31*, 89–97.

Knapp, S., Müller, S., Digilio, G., Bonaldi, T., Bianchi, M.E., and Musco, G. (2004). The long acidic tail of high mobility group box 1 (HMGB1) protein forms an extended and flexible structure that interacts with specific residues within and between the HMG boxes. Biochemistry *43*, 11992–11997.

Koch, C.M., Andrews, R.M., Flicek, P., Dillon, S.C., Karaöz, U., Clelland, G.K., Wilcox, S., Beare, D.M., Fowler, J.C., Couttet, P., et al. (2007). The landscape of histone modifications across 1% of the human genome in five human cell lines. Genome Res. *17*, 691–707.

Kolasinska-Zwierz, P., Down, T., Latorre, I., Liu, T., Liu, X.S., and Ahringer, J. (2009). Differential chromatin marking of introns and expressed exons by H3K36me3. Nat. Genet. *41*, 376–381.

Komuro, I., Kudo, S., Yamazaki, T., Zou, Y., Shiojima, I., and Yazaki, Y. (1996). Mechanical stretch activates the stress-activated protein kinases in cardiac myocytes. FASEB J. *10*, 631–636.

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.

Kornberg, R.D., and Thomas, J.O. (1974). Chromatin structure; oligomers of the histones. Science *184*, 865–868.

Körner, U., Bustin, M., Scheer, U., and Hock, R. (2003). Developmental role of HMGN proteins in Xenopus laevis. Mech. Dev. *120*, 1177–1192.

Kouzarides, T. (1999). Histone acetylases and deacetylases in cell proliferation. Curr. Opin. Genet. Dev. *9*, 40–48.

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

Kouzarides, T., and Ziff, E. (1988). The role of the leucine zipper in the fos-jun interaction. Nature *336*, 646–651.

Krishnamoorthy, T., Chen, X., Govin, J., Cheung, W.L., Dorsey, J., Schindler, K., Winter, E., Allis, C.D., Guacci, V., Khochbin, S., et al. (2006). Phosphorylation of histone H4 Ser1 regulates sporulation in yeast and is conserved in fly and mouse spermatogenesis. Genes Dev. *20*, 2580–2592.

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

Kuehl, L., Salmond, B., and Tran, L. (1984). Concentrations of high-mobility-group proteins in the nucleus and cytoplasm of several rat tissues. J. Cell Biol. *99*, 648–654.

Kugler, J.E., Deng, T., and Bustin, M. (2012). The HMGN family of chromatin-binding proteins: Dynamic modulators of epigenetic processes. Biochim. Biophys. Acta *1819*, 652–656.

Kuo, M.H., and Allis, C.D. (1998). Roles of histone acetyltransferases and deacetylases in gene regulation. Bioessays *20*, 615–626.

Kurdistani, S.K., and Grunstein, M. (2003). In vivo protein-protein and protein-DNA crosslinking for genomewide binding microarray. Methods *31*, 90–95.

Kurdistani, S.K., Tavazoie, S., and Grunstein, M. (2004). Mapping global histone acetylation patterns to gene expression. Cell *117*, 721–733.

Kuzmichev, A., Jenuwein, T., Tempst, P., and Reinberg, D. (2004). Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. Mol. Cell *14*, 183–193.

Kuzmichev, A., Margueron, R., Vaquero, A., Preissner, T.S., Scher, M., Kirmizis, A., Ouyang, X., Brockdorff, N., Abate-Shen, C., Farnham, P., et al. (2005). Composition and histone substrates of polycomb repressive group complexes change during cellular differentiation. Proc. Natl. Acad. Sci. U.S.A. *102*, 1859–1864.

Kyriakis, J.M., and Avruch, J. (2001). Mammalian mitogen-activated protein kinase signal transduction pathways activated by stress and inflammation. Physiol. Rev. *81*, 807–869.

Kyriakis, J.M., and Avruch, J. (2012). Mammalian MAPK signal transduction pathways activated by stress and inflammation: a 10-year update. Physiol. Rev. *92*, 689–737.

Kyriakis, J.M., Banerjee, P., Nikolakaki, E., Dai, T., Rubie, E.A., Ahmad, M.F., Avruch, J., and Woodgett, J.R. (1994). The stress-activated protein kinase subfamily of c-Jun kinases. Nature *369*, 156–160.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680–685.

Lamph, W.W., Wamsley, P., Sassone-Corsi, P., and Verma, I.M. (1988). Induction of proto-oncogene JUN/AP-1 by serum and TPA. Nature *334*, 629–631.

Landschulz, W.H., Johnson, P.F., and McKnight, S.L. (1988). The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science *240*, 1759–1764.

Langan, T.A., Gautier, J., Lohka, M., Hollingsworth, R., Moreno, S., Nurse, P., Maller, J., and Sclafani, R.A. (1989). Mammalian growth-associated H1 histone kinase: a homolog of cdc2+/CDC28 protein kinases controlling mitotic entry in yeast and frog cells. Mol. Cell. Biol. *9*, 3860–3868.

De Lange, R.J., Fambrough, D.M., Smith, E.L., and Bonner, J. (1969). Calf and pea histone IV. 3. Complete amino acid sequence of pea seedling histone IV; comparison with the homologous calf thymus histone. J. Biol. Chem. *244*, 5669–5679.

Lee, B.M., and Mahadevan, L.C. (2009). Stability of histone modifications across mammalian genomes: implications for "epigenetic" marking. J. Cell. Biochem. *108*, 22–34.

Lee, M.G., Wynder, C., Cooch, N., and Shiekhattar, R. (2005). An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. Nature *437*, 432–435.

Lehtonen, S., and Lehtonen, E. (2001). HMG-17 is an early marker of inductive interactions in the developing mouse kidney. Differentiation *67*, 154–163.

Lever, M.A., Th'ng, J.P., Sun, X., and Hendzel, M.J. (2000). Rapid exchange of histone H1.1 on chromatin in living human cells. Nature *408*, 873–876.

Levsky, J.M., and Singer, R.H. (2003). Gene expression and the myth of the average cell. Trends Cell Biol. *13*, 4–6.

Lewis, H., Kaszubska, W., DeLamarter, J.F., and Whelan, J. (1994). Cooperativity between two NF-kappa B complexes, mediated by high-mobility-group protein I(Y), is essential for cytokine-induced expression of the E-selectin promoter. Mol. Cell. Biol. *14*, 5701–5709.

Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. Cell *128*, 707–719.

Lian, Z., Karpikov, A., Lian, J., Mahajan, M.C., Hartman, S., Gerstein, M., Snyder, M., and Weissman, S.M. (2008). A genomic analysis of RNA polymerase II modification and chromatin architecture related to 3' end RNA polyadenylation. Genome Res. *18*, 1224–1237.

Lim, J.-H., Catez, F., Birger, Y., West, K.L., Prymakowska-Bosak, M., Postnikov, Y.V., and Bustin, M. (2004). Chromosomal protein HMGN1 modulates histone H3 phosphorylation. Mol. Cell *15*, 573–584.

Lim, J.-H., West, K.L., Rubinstein, Y., Bergel, M., Postnikov, Y.V., and Bustin, M. (2005). Chromosomal protein HMGN1 enhances the acetylation of lysine 14 in histone H3. EMBO J. *24*, 3038–3048.

Lin, A., Frost, J., Deng, T., Smeal, T., al-Alawi, N., Kikkawa, U., Hunter, T., Brenner, D., and Karin, M. (1992). Casein kinase II is a negative regulator of c-Jun DNA binding and AP-1 activity. Cell *70*, 777–789.

Liu, C.L., Kaplan, T., Kim, M., Buratowski, S., Schreiber, S.L., Friedman, N., and Rando, O.J. (2005). Single-nucleosome mapping of histone modifications in S. cerevisiae. PLoS Biol. *3*, e328.

Liu, J., Schiltz, J.F., Ashar, H.R., and Chada, K.K. (2003). Hmga1 is required for normal sperm development. Mol. Reprod. Dev. *66*, 81–89.

Lo, W.S., Trievel, R.C., Rojas, J.R., Duggan, L., Hsu, J.Y., Allis, C.D., Marmorstein, R., and Berger, S.L. (2000). Phosphorylation of serine 10 in histone H3 is functionally linked in vitro and in vivo to Gcn5-mediated acetylation at lysine 14. Mol. Cell *5*, 917–926.

Lotze, M.T., and Tracey, K.J. (2005). High-mobility group box 1 protein (HMGB1): nuclear weapon in the immune arsenal. Nat. Rev. Immunol. *5*, 331–342.

Louie, D.F., Gloor, K.K., Galasinski, S.C., Resing, K.A., and Ahn, N.G. (2000). Phosphorylation and subcellular redistribution of high mobility group proteins 14 and 17, analyzed by mass spectrometry. Protein Sci. *9*, 170–179.

Lu, M.J., Mpoke, S.S., Dadd, C.A., and Allis, C.D. (1995). Phosphorylated and dephosphorylated linker histone H1 reside in distinct chromatin domains in Tetrahymena macronuclei. Mol. Biol. Cell *6*, 1077–1087.

Luger, K., Mäder, 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*, 251–260.

Macdonald, N., Welburn, J.P.I., Noble, M.E.M., Nguyen, A., Yaffe, M.B., Clynes, D., Moggs, J.G., Orphanides, G., Thomson, S., Edmunds, J.W., et al. (2005). Molecular basis for the recognition of phosphorylated and phosphoacetylated histone h3 by 14-3-3. Mol. Cell *20*, 199–211.

Mahadevan, L.C., Willis, A.C., and Barratt, M.J. (1991). Rapid histone H3 phosphorylation in response to growth factors, phorbol esters, okadaic acid, and protein synthesis inhibitors. Cell *65*, 775–783.

Margueron, R., Trojer, P., and Reinberg, D. (2005). The key to development: interpreting the histone code? Curr. Opin. Genet. Dev. *15*, 163–176.

Martinez Hoyos, J., Fedele, M., Battista, S., Pentimalli, F., Kruhoffer, M., Arra, C., Orntoft, T.F., Croce, C.M., and Fusco, A. (2004). Identification of the genes up- and down-regulated by the high mobility group A1 (HMGA1) proteins: tissue specificity of the HMGA1-dependent gene regulation. Cancer Res. *64*, 5728–5735.

Mavrich, T.N., Ioshikhes, I.P., Venters, B.J., Jiang, C., Tomsho, L.P., Qi, J., Schuster, S.C., Albert, I., and Pugh, B.F. (2008). A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. Genome Res. *18*, 1073–1083.

Melvin, V.S., Harrell, C., Adelman, J.S., Kraus, W.L., Churchill, M., and Edwards, D.P. (2004). The role of the C-terminal extension (CTE) of the estrogen receptor alpha and beta DNA binding domain in DNA binding and interaction with HMGB. J. Biol. Chem. *279*, 14763–14771.

Meng, Q., and Xia, Y. (2011). c-Jun, at the crossroad of the signaling network. Protein Cell *2*, 889–898.

Metzger, E., Wissmann, M., Yin, N., Müller, J.M., Schneider, R., Peters, A.H.F.M., Günther, T., Buettner, R., and Schüle, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature *437*, 436–439.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R.P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553–560.

Misteli, T. (2004). Spatial positioning; a new dimension in genome function. Cell *119*, 153–156.

Misteli, T. (2005). Concepts in nuclear architecture. Bioessays *27*, 477–487.

Misteli, T., Gunjan, A., Hock, R., Bustin, M., and Brown, D.T. (2000). Dynamic binding of histone H1 to chromatin in living cells. Nature *408*, 877–881.

Mizzen, C.A., and Allis, C.D. (1998). Linking histone acetylation to transcriptional regulation. Cell. Mol. Life Sci. *54*, 6–20.

Mohamed, O.A., Bustin, M., and Clarke, H.J. (2001). High-mobility group proteins 14 and 17 maintain the timing of early embryonic development in the mouse. Dev. Biol. *229*, 237–249.

Moon, H., Filippova, G., Loukinov, D., Pugacheva, E., Chen, Q., Smith, S.T., Munhall, A., Grewe, B., Bartkuhn, M., Arnold, R., et al. (2005). CTCF is conserved from Drosophila to humans and confers enhancer blocking of the Fab-8 insulator. EMBO Rep. *6*, 165–170.

Moretti, P., Bouwknecht, J.A., Teague, R., Paylor, R., and Zoghbi, H.Y. (2005). Abnormalities of social interactions and home-cage behavior in a mouse model of Rett syndrome. Hum. Mol. Genet. *14*, 205–220.

Morrison, D.K. (2009). The 14-3-3 proteins: integrators of diverse signaling cues that impact cell fate and cancer development. Trends Cell Biol. *19*, 16–23.

Müller, S., Ronfani, L., and Bianchi, M.E. (2004). Regulated expression and subcellular localization of HMGB1, a chromatin protein with a cytokine function. J. Intern. Med. *255*, 332–343.

Mujtaba, S., Zeng, L., and Zhou, M.-M. (2007). Structure and acetyl-lysine recognition of the bromodomain. Oncogene *26*, 5521–5527.

Munshi, N., Merika, M., Yie, J., Senger, K., Chen, G., and Thanos, D. (1998). Acetylation of HMG I(Y) by CBP turns off IFN beta expression by disrupting the enhanceosome. Mol. Cell *2*, 457–467.

Munshi, N., Agalioti, T., Lomvardas, S., Merika, M., Chen, G., and Thanos, D. (2001). Coordination of a transcriptional switch by HMGI(Y) acetylation. Science *293*, 1133–1136.

Murphy, L.O., Smith, S., Chen, R.-H., Fingar, D.C., and Blenis, J. (2002). Molecular interpretation of ERK signal duration by immediate early gene products. Nat. Cell Biol. *4*, 556–564.

Najima, Y., Yahagi, N., Takeuchi, Y., Matsuzaka, T., Sekiya, M., Nakagawa, Y., Amemiya-Kudo, M., Okazaki, H., Okazaki, S., Tamura, Y., et al. (2005). High mobility group protein-B1 interacts with

sterol regulatory element-binding proteins to enhance their DNA binding. J. Biol. Chem. *280*, 27523–27532.

Nelson, C.J., Santos-Rosa, H., and Kouzarides, T. (2006). Proline isomerization of histone H3 regulates lysine methylation and gene expression. Cell *126*, 905–916.

Nemeth, M.J., Cline, A.P., Anderson, S.M., Garrett-Beal, L.J., and Bodine, D.M. (2005). Hmgb3 deficiency deregulates proliferation and differentiation of common lymphoid and myeloid progenitors. Blood *105*, 627–634.

O'Neill, L.P., and Turner, B.M. (2003). Immunoprecipitation of native chromatin: NChIP. Methods *31*, 76–82.

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

Okazaki, K., and Sagata, N. (1995). The Mos/MAP kinase pathway stabilizes c-Fos by phosphorylation and augments its transforming activity in NIH 3T3 cells. EMBO J. *14*, 5048–5059.

Pallier, C., Scaffidi, P., Chopineau-Proust, S., Agresti, A., Nordmann, P., Bianchi, M.E., and Marechal, V. (2003). Association of chromatin proteins high mobility group box (HMGB) 1 and HMGB2 with mitotic chromosomes. Mol. Biol. Cell *14*, 3414–3426.

Pandit, B., Sarkozy, A., Pennacchio, L.A., Carta, C., Oishi, K., Martinelli, S., Pogna, E.A., Schackwitz, W., Ustaszewska, A., Landstrom, A., et al. (2007). Gain-of-function RAF1 mutations cause Noonan and LEOPARD syndromes with hypertrophic cardiomyopathy. Nat. Genet. *39*, 1007–1012.

Pash, J., Popescu, N., Matocha, M., Rapoport, S., and Bustin, M. (1990a). Chromosomal protein HMG-14 gene maps to the Down syndrome region of human chromosome 21 and is overexpressed in mouse trisomy 16. Proc. Natl. Acad. Sci. U.S.A. *87*, 3836–3840.

Pash, J.M., Bhorjee, J.S., Patterson, B.M., and Bustin, M. (1990b). Persistence of chromosomal proteins HMG-14/-17 in myotubes following differentiation-dependent reduction of HMG mRNA. J. Biol. Chem. *265*, 4197–4199.

Pash, J.M., Alfonso, P.J., and Bustin, M. (1993). Aberrant expression of high mobility group chromosomal protein 14 affects cellular differentiation. J. Biol. Chem. *268*, 13632–13638.

Peña, P.V., Davrazou, F., Shi, X., Walter, K.L., Verkhusha, V.V., Gozani, O., Zhao, R., and Kutateladze, T.G. (2006). Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. Nature *442*, 100–103.

Phair, R.D., and Misteli, T. (2000). High mobility of proteins in the mammalian cell nucleus. Nature *404*, 604–609.

Phair, R.D., Scaffidi, P., Elbi, C., Vecerová, J., Dey, A., Ozato, K., Brown, D.T., Hager, G., Bustin, M., and Misteli, T. (2004). Global nature of dynamic protein-chromatin interactions in vivo: three-dimensional genome scanning and dynamic interaction networks of chromatin proteins. Mol. Cell. Biol. *24*, 6393–6402.

Phillips, D.M., and Johns, E.W. (1965). A fractionation of the histones of group F2A from calf thymus. Biochem. J. *94*, 127–130.

Phillips, J.E., and Corces, V.G. (2009). CTCF: master weaver of the genome. Cell *137*, 1194–1211.

Pogna, E.A., Clayton, A.L., and Mahadevan, L.C. (2010). Signalling to chromatin through post-translational modifications of HMGN. Biochim. Biophys. Acta *1799*, 93–100.

Postnikov, Y., and Bustin, M. (2010). Regulation of chromatin structure and function by HMGN proteins. Biochim. Biophys. Acta *1799*, 62–68.

Postnikov, Y.V., Trieschmann, L., Rickers, A., and Bustin, M. (1995). Homodimers of chromosomal proteins HMG-14 and HMG-17 in nucleosome cores. J. Mol. Biol. *252*, 423–432.

Postnikov, Y.V., Belova, G.I., Lim, J.-H., and Bustin, M. (2006). Chromosomal protein HMGN1 modulates the phosphorylation of serine 1 in histone H2A. Biochemistry *45*, 15092–15099.

Potier, M.-C., Rivals, I., Mercier, G., Ettwiller, L., Moldrich, R.X., Laffaire, J., Personnaz, L., Rossier, J., and Dauphinot, L. (2006). Transcriptional disruptions in Down syndrome: a case study in the Ts1Cje mouse cerebellum during post-natal development. J. Neurochem. *97 Suppl 1*, 104–109.

Pray-Grant, M.G., Daniel, J.A., Schieltz, D., Yates, J.R., 3rd, and Grant, P.A. (2005). Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. Nature *433*, 434–438.

Prigent, C., and Dimitrov, S. (2003). Phosphorylation of serine 10 in histone H3, what for? J. Cell. Sci. *116*, 3677–3685.

Pruss, D., Bartholomew, B., Persinger, J., Hayes, J., Arents, G., Moudrianakis, E.N., and Wolffe, A.P. (1996). An asymmetric model for the nucleosome: a binding site for linker histones inside the DNA gyres. Science *274*, 614–617.

Prymakowska-Bosak, M., Misteli, T., Herrera, J.E., Shirakawa, H., Birger, Y., Garfield, S., and Bustin, M. (2001). Mitotic phosphorylation prevents the binding of HMGN proteins to chromatin. Mol. Cell. Biol. *21*, 5169–5178.

Prymakowska-Bosak, M., Hock, R., Catez, F., Lim, J.-H., Birger, Y., Shirakawa, H., Lee, K., and Bustin, M. (2002). Mitotic phosphorylation of chromosomal protein HMGN1 inhibits nuclear import and promotes interaction with 14.3.3 proteins. Mol. Cell. Biol. *22*, 6809–6819.

Raingeaud, J., Gupta, S., Rogers, J.S., Dickens, M., Han, J., Ulevitch, R.J., and Davis, R.J. (1995). Pro-inflammatory cytokines and environmental stress cause p38 mitogen-activated protein kinase activation by dual phosphorylation on tyrosine and threonine. J. Biol. Chem. *270*, 7420–7426.

Ramachandran, C., Yau, P., Bradbury, E.M., Shyamala, G., Yasuda, H., and Walsh, D.A. (1984). Phosphorylation of high-mobility-group proteins by the calcium-phospholipid-dependent protein kinase and the cyclic AMP-dependent protein kinase. J. Biol. Chem. *259*, 13495–13503.

Rattner, B.P., Yusufzai, T., and Kadonaga, J.T. (2009). HMGN proteins act in opposition to ATP-dependent chromatin remodeling factors to restrict nucleosome mobility. Mol. Cell *34*, 620–626.

Reeves, R. (2001). Molecular biology of HMGA proteins: hubs of nuclear function. Gene *277*, 63–81.

Reeves, R., and Nissen, M.S. (1990). The A.T-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. J. Biol. Chem. *265*, 8573–8582.

Reinke, H., and Hörz, W. (2003). Histones are first hyperacetylated and then lose contact with the activated PHO5 promoter. Mol. Cell *11*, 1599–1607.

Rice, J.C., and Allis, C.D. (2001). Code of silence. Nature *414*, 258–261.

Richards, E.J., and Elgin, S.C.R. (2002). Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. Cell *108*, 489–500.

Roberts, P.J., and Der, C.J. (2007). Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. Oncogene *26*, 3291–3310.

Robinson, P.J.J., and Rhodes, D. (2006). Structure of the "30 nm" chromatin fibre: a key role for the linker histone. Curr. Opin. Struct. Biol. *16*, 336–343.

Robinson, P.J.J., Fairall, L., Huynh, V.A.T., and Rhodes, D. (2006). EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. Proc. Natl. Acad. Sci. U.S.A. *103*, 6506–6511.

Rochman, M., Postnikov, Y., Correll, S., Malicet, C., Wincovitch, S., Karpova, T.S., McNally, J.G., Wu, X., Bubunenko, N.A., Grigoryev, S., et al. (2009). The interaction of NSBP1/HMGN5 with nucleosomes in euchromatin counteracts linker histone-mediated chromatin compaction and modulates transcription. Mol. Cell *35*, 642–656.

Rochman, M., Taher, L., Kurahashi, T., Cherukuri, S., Uversky, V.N., Landsman, D., Ovcharenko, I., and Bustin, M. (2011). Effects of HMGN variants on the cellular transcription profile. Nucleic Acids Res. *39*, 4076–4087.

Roque, A., Iloro, I., Ponte, I., Arrondo, J.L.R., and Suau, P. (2005). DNA-induced secondary structure of the carboxyl-terminal domain of histone H1. J. Biol. Chem. *280*, 32141–32147.

Roque, A., Ponte, I., Arrondo, J.L.R., and Suau, P. (2008). Phosphorylation of the carboxy-terminal domain of histone H1: effects on secondary structure and DNA condensation. Nucleic Acids Res. *36*, 4719–4726.

Rosette, C., and Karin, M. (1996). Ultraviolet light and osmotic stress: activation of the JNK cascade through multiple growth factor and cytokine receptors. Science *274*, 1194–1197.

Rouse, J., Cohen, P., Trigon, S., Morange, M., Alonso-Llamazares, A., Zamanillo, D., Hunt, T., and Nebreda, A.R. (1994). A novel kinase cascade triggered by stress and heat shock that stimulates MAPKAP kinase-2 and phosphorylation of the small heat shock proteins. Cell *78*, 1027–1037.

Rozek, D., and Pfeifer, G.P. (1995). In vivo protein-DNA interactions at the c-jun promoter in quiescent and serum-stimulated fibroblasts. J. Cell. Biochem. *57*, 479–487.

Ryder, K., Lau, L.F., and Nathans, D. (1988). A gene activated by growth factors is related to the oncogene v-jun. Proc. Natl. Acad. Sci. U.S.A. *85*, 1487–1491.

Ryseck, R.P., Hirai, S.I., Yaniv, M., and Bravo, R. (1988). Transcriptional activation of c-jun during the G0/G1 transition in mouse fibroblasts. Nature *334*, 535–537.

Ryseck, R.P., Kovary, K., and Bravo, R. (1990). Integrity of FOS B leucine zipper is essential for its interaction with JUN proteins. Oncogene *5*, 1091–1093.

Sapojnikova, N., Maman, J., Myers, F.A., Thorne, A.W., Vorobyev, V.I., and Crane-Robinson, C. (2005). Biochemical observation of the rapid mobility of nuclear HMGB1. Biochim. Biophys. Acta *1729*, 57–63.

Sarg, B., Helliger, W., Talasz, H., Förg, B., and Lindner, H.H. (2006). Histone H1 phosphorylation occurs site-specifically during interphase and mitosis: identification of a novel phosphorylation site on histone H1. J. Biol. Chem. *281*, 6573–6580.

Sarma, K., and Reinberg, D. (2005). Histone variants meet their match. Nat. Rev. Mol. Cell Biol. *6*, 139–149.

Sassone-Corsi, P., Mizzen, C.A., Cheung, P., Crosio, C., Monaco, L., Jacquot, S., Hanauer, A., and Allis, C.D. (1999). Requirement of Rsk-2 for epidermal growth factor-activated phosphorylation of histone H3. Science *285*, 886–891.

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*, 887–898.

Schreiber, S.L., and Bernstein, B.E. (2002). Signaling network model of chromatin. Cell *111*, 771–778.

Schübeler, D., MacAlpine, D.M., Scalzo, D., Wirbelauer, C., Kooperberg, C., van Leeuwen, F., Gottschling, D.E., O'Neill, L.P., Turner, B.M., Delrow, J., et al. (2004). The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. Genes Dev. *18*, 1263–1271.

Schumacher, M.A., Goodman, R.H., and Brennan, R.G. (2000). The structure of a CREB bZIP.somatostatin CRE complex reveals the basis for selective dimerization and divalent cation-enhanced DNA binding. J. Biol. Chem. *275*, 35242–35247.

Sgarra, R., Rustighi, A., Tessari, M.A., Di Bernardo, J., Altamura, S., Fusco, A., Manfioletti, G., and Giancotti, V. (2004). Nuclear phosphoproteins HMGA and their relationship with chromatin structure and cancer. FEBS Lett. *574*, 1–8.

Shaulian, E., and Karin, M. (2002). AP-1 as a regulator of cell life and death. Nat. Cell Biol. *4*, E131–136.

Shaw, G., and Kamen, R. (1986). A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell *46*, 659–667.

Shaw, P.E., Schröter, H., and Nordheim, A. (1989). The ability of a ternary complex to form over the serum response element correlates with serum inducibility of the human c-fos promoter. Cell *56*, 563–572.

Shechter, D., Dormann, H.L., Allis, C.D., and Hake, S.B. (2007). Extraction, purification and analysis of histones. Nat Protoc *2*, 1445–1457.

Shen, X., and Gorovsky, M.A. (1996). Linker histone H1 regulates specific gene expression but not global transcription in vivo. Cell *86*, 475–483.

Sheng, M., Dougan, S.T., McFadden, G., and Greenberg, M.E. (1988). Calcium and growth factor pathways of c-fos transcriptional activation require distinct upstream regulatory sequences. Mol. Cell. Biol. *8*, 2787–2796.

Shi, X., Hong, T., Walter, K.L., Ewalt, M., Michishita, E., Hung, T., Carney, D., Peña, P., Lan, F., Kaadige, M.R., et al. (2006). ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. Nature *442*, 96–99.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A., and Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell *119*, 941–953.

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

Shirakawa, H., Herrera, J.E., Bustin, M., and Postnikov, Y. (2000a). Targeting of high mobility group-14/-17 proteins in chromatin is independent of DNA sequence. J. Biol. Chem. *275*, 37937–37944.

Shirakawa, H., Landsman, D., Postnikov, Y.V., and Bustin, M. (2000b). NBP-45, a novel nucleosomal binding protein with a tissue-specific and developmentally regulated expression. J. Biol. Chem. *275*, 6368–6374.

Shogren-Knaak, M., Ishii, H., Sun, J.-M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. Science *311*, 844–847.

Silverman, J.L., Yang, M., Lord, C., and Crawley, J.N. (2010). Behavioural phenotyping assays for mouse models of autism. Nat. Rev. Neurosci. *11*, 490–502.

Simpson, R.T. (1978). Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. Biochemistry *17*, 5524–5531.

Sims, R.J., 3rd, Chen, C.-F., Santos-Rosa, H., Kouzarides, T., Patel, S.S., and Reinberg, D. (2005). Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. J. Biol. Chem. *280*, 41789–41792.

Skene, P.J., Illingworth, R.S., Webb, S., Kerr, A.R.W., James, K.D., Turner, D.J., Andrews, R., and Bird, A.P. (2010). Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. Mol. Cell *37*, 457–468.

Sluss, H.K., Barrett, T., Dérijard, B., and Davis, R.J. (1994). Signal transduction by tumor necrosis factor mediated by JNK protein kinases. Mol. Cell. Biol. *14*, 8376–8384.

Smeal, T., Angel, P., Meek, J., and Karin, M. (1989). Different requirements for formation of Jun: Jun and Jun: Fos complexes. Genes Dev. *3*, 2091–2100.

Smeal, T., Binetruy, B., Mercola, D.A., Birrer, M., and Karin, M. (1991). Oncogenic and transcriptional cooperation with Ha-Ras requires phosphorylation of c-Jun on serines 63 and 73. Nature *354*, 494–496.

Soloaga, A., Thomson, S., Wiggin, G.R., Rampersaud, N., Dyson, M.H., Hazzalin, C.A., Mahadevan, L.C., and Arthur, J.S.C. (2003). MSK2 and MSK1 mediate the mitogen- and stress-induced phosphorylation of histone H3 and HMG-14. EMBO J. *22*, 2788–2797.

Solomon, M.J., Strauss, F., and Varshavsky, A. (1986). A mammalian high mobility group protein recognizes any stretch of six A.T base pairs in duplex DNA. Proc. Natl. Acad. Sci. U.S.A. *83*, 1276–1280.

Song, M.J., Hwang, S., Wong, W., Round, J., Martinez-Guzman, D., Turpaz, Y., Liang, J., Wong, B., Johnson, R.C., Carey, M., et al. (2004). The DNA architectural protein HMGB1 facilitates RTA-mediated viral gene expression in gamma-2 herpesviruses. J. Virol. *78*, 12940–12950.

Steger, D.J., Lefterova, M.I., Ying, L., Stonestrom, A.J., Schupp, M., Zhuo, D., Vakoc, A.L., Kim, J.-E., Chen, J., Lazar, M.A., et al. (2008). DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells. Mol. Cell. Biol. *28*, 2825–2839.

Steinbach, O.C., Wolffe, A.P., and Rupp, R.A. (1997). Somatic linker histones cause loss of mesodermal competence in Xenopus. Nature *389*, 395–399.

Sterner, R., Vidali, G., and Allfrey, V.G. (1981). Studies of acetylation and deacetylation in high mobility group proteins. Identification of the sites of acetylation in high mobility group proteins 14 and 17. J. Biol. Chem. *256*, 8892–8895.

Steward, M.M., Lee, J.-S., O'Donovan, A., Wyatt, M., Bernstein, B.E., and Shilatifard, A. (2006). Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. Nat. Struct. Mol. Biol. *13*, 852–854.

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

Strahl, B.D., Grant, P.A., Briggs, S.D., Sun, Z.-W., Bone, J.R., Caldwell, J.A., Mollah, S., Cook, R.G., Shabanowitz, J., Hunt, D.F., et al. (2002). Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. Mol. Cell. Biol. *22*, 1298–1306.

Takami, Y., Nishi, R., and Nakayama, T. (2000). Histone H1 variants play individual roles in transcription regulation in the DT40 chicken B cell line. Biochem. Biophys. Res. Commun. *268*, 501–508.

Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat. Struct. Mol. Biol. *14*, 1025–1040.

Th'ng, J.P.H., 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*, 27809–27814.

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*, 403–427.

Thomas, J.O. (1999). Histone H1: location and role. Curr. Opin. Cell Biol. *11*, 312–317.

Thomas, J.O., and Travers, A.A. (2001). HMG1 and 2, and related "architectural" DNA-binding proteins. Trends Biochem. Sci. *26*, 167–174.

Thomson, S., Clayton, A.L., Hazzalin, C.A., Rose, S., Barratt, M.J., and Mahadevan, L.C. (1999). The nucleosomal response associated with immediate-early gene induction is mediated via alternative MAP kinase cascades: MSK1 as a potential histone H3/HMG-14 kinase. EMBO J. *18*, 4779–4793.

Thomson, S., Clayton, A.L., and Mahadevan, L.C. (2001). Independent dynamic regulation of histone phosphorylation and acetylation during immediate-early gene induction. Mol. Cell *8*, 1231–1241.

Travers, A. (1999). The location of the linker histone on the nucleosome. Trends Biochem. Sci. *24*, 4–7.

Travers, A.A. (2003). Priming the nucleosome: a role for HMGB proteins? EMBO Rep. *4*, 131–136.

Treff, N.R., Dement, G.A., Adair, J.E., Britt, R.L., Nie, R., Shima, J.E., Taylor, W.E., and Reeves, R. (2004). Human KIT ligand promoter is positively regulated by HMGA1 in breast and ovarian cancer cells. Oncogene *23*, 8557–8562.

Treisman, R. (1992). The serum response element. Trends Biochem. Sci. *17*, 423–426.

Treisman, R. (1996). Regulation of transcription by MAP kinase cascades. Curr. Opin. Cell Biol. *8*, 205–215.

Trieschmann, L., Postnikov, Y.V., Rickers, A., and Bustin, M. (1995). Modular structure of chromosomal proteins HMG-14 and HMG-17: definition of a transcriptional enhancement domain distinct from the nucleosomal binding domain. Mol. Cell. Biol. *15*, 6663–6669.

Trieschmann, L., Martin, B., and Bustin, M. (1998). The chromatin unfolding domain of chromosomal protein HMG-14 targets the N-terminal tail of histone H3 in nucleosomes. Proc. Natl. Acad. Sci. U.S.A. *95*, 5468–5473.

Trivier, E., De Cesare, D., Jacquot, S., Pannetier, S., Zackai, E., Young, I., Mandel, J.L., Sassone-Corsi, P., and Hanauer, A. (1996). Mutations in the kinase Rsk-2 associated with Coffin-Lowry syndrome. Nature *384*, 567–570.

Trojer, P., and Reinberg, D. (2007). Facultative heterochromatin: is there a distinctive molecular signature? Mol. Cell *28*, 1–13.

Turner, B.M. (2005). Reading signals on the nucleosome with a new nomenclature for modified histones. Nat. Struct. Mol. Biol. *12*, 110–112.

Ueda, T., Postnikov, Y.V., and Bustin, M. (2006). Distinct domains in high mobility group N variants modulate specific chromatin modifications. J. Biol. Chem. *281*, 10182–10187.

Ueda, T., Catez, F., Gerlitz, G., and Bustin, M. (2008). Delineation of the protein module that anchors HMGN proteins to nucleosomes in the chromatin of living cells. Mol. Cell. Biol. *28*, 2872–2883.

Ueda, T., Furusawa, T., Kurahashi, T., Tessarollo, L., and Bustin, M. (2009). The nucleosome binding protein HMGN3 modulates the transcription profile of pancreatic beta cells and affects insulin secretion. Mol. Cell. Biol. *29*, 5264–5276.

Vakoc, C.R., Mandat, S.A., Olenchock, B.A., and Blobel, G.A. (2005). Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. Mol. Cell *19*, 381–391.

Vakoc, C.R., Sachdeva, M.M., Wang, H., and Blobel, G.A. (2006). Profile of histone lysine methylation across transcribed mammalian chromatin. Mol. Cell. Biol. *26*, 9185–9195.

Vaquero, A., Scher, M., Lee, D., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2004). Human SirT1 interacts with histone H1 and promotes formation of facultative heterochromatin. Mol. Cell *16*, 93–105.

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., Meliá, 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*, 501–517.

Vigneault, C., McGraw, S., Massicotte, L., and Sirard, M.-A. (2004). Transcription factor expression patterns in bovine in vitro-derived embryos prior to maternal-zygotic transition. Biol. Reprod. *70*, 1701–1709.

Wagner, B.J., Hayes, T.E., Hoban, C.J., and Cochran, B.H. (1990). The SIF binding element confers sis/PDGF inducibility onto the c-fos promoter. EMBO J. *9*, 4477–4484.

Wallace, J.A., and Felsenfeld, G. (2007). We gather together: insulators and genome organization. Curr. Opin. Genet. Dev. *17*, 400–407.

Walton, G.M., and Gill, G.N. (1983). Identity of the in vivo phosphorylation site in high mobility group 14 protein in HeLa cells with the site phosphorylated by casein kinase II in vitro. J. Biol. Chem. *258*, 4440–4446.

Walton, G.M., Gill, G.N., Cooper, E., and Spaulding, S.W. (1984). Thyrotropin-stimulated phosphorylation of high mobility group protein 14 in vivo at the site catalyzed by cyclic nucleotide-dependent protein kinases in vitro. J. Biol. Chem. *259*, 601–607.

Walton, G.M., Spiess, J., and Gill, G.N. (1985). Phosphorylation of high mobility group protein 14 by casein kinase II. J. Biol. Chem. *260*, 4745–4750.

Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Peng, W., Zhang, M.Q., et al. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. Nat. Genet. *40*, 897–903.

Waskiewicz, A.J., Flynn, A., Proud, C.G., and Cooper, J.A. (1997). Mitogen-activated protein kinases activate the serine/threonine kinases Mnk1 and Mnk2. EMBO J. *16*, 1909–1920.

West, K.L. (2004). HMGN proteins play roles in DNA repair and gene expression in mammalian cells. Biochem. Soc. Trans. *32*, 918–919.

Whitley, M.Z., Thanos, D., Read, M.A., Maniatis, T., and Collins, T. (1994). A striking similarity in the organization of the E-selectin and beta interferon gene promoters. Mol. Cell. Biol. *14*, 6464–6475.

Wiggin, G.R., Soloaga, A., Foster, J.M., Murray-Tait, V., Cohen, P., and Arthur, J.S.C. (2002). MSK1 and MSK2 are required for the mitogen- and stress-induced phosphorylation of CREB and ATF1 in fibroblasts. Mol. Cell. Biol. *22*, 2871–2881.

Williams, J.G., Kubelik, A.R., Livak, K.J., Rafalski, J.A., and Tingey, S.V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. *18*, 6531–6535.

Wilson, T., and Treisman, R. (1988). Removal of poly(A) and consequent degradation of c-fos mRNA facilitated by 3' AU-rich sequences. Nature *336*, 396–399.

Wisniewski, J.R., Zougman, A., Krüger, S., and Mann, M. (2007). Mass spectrometric mapping of linker histone H1 variants reveals multiple acetylations, methylations, and phosphorylation as well as differences between cell culture and tissue. Mol. Cell Proteomics *6*, 72–87.

Wisniewski, J.R., Zougman, A., and Mann, M. (2008). Nepsilon-formylation of lysine is a widespread post-translational modification of nuclear proteins occurring at residues involved in regulation of chromatin function. Nucleic Acids Res. *36*, 570–577.

Wu, C., and Travers, A. (2004). A "one-pot" assay for the accessibility of DNA in a nucleosome core particle. Nucleic Acids Res. *32*, e122.

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P., et al. (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature *442*, 86–90.

Xia, Y., Makris, C., Su, B., Li, E., Yang, J., Nemerow, G.R., and Karin, M. (2000). MEK kinase 1 is critically required for c-Jun N-terminal kinase activation by proinflammatory stimuli and growth factor-induced cell migration. Proc. Natl. Acad. Sci. U.S.A. *97*, 5243–5248.

Xu, M., and Cook, P.R. (2008). Similar active genes cluster in specialized transcription factories. J. Cell Biol. *181*, 615–623.

Yusufzai, T.M., Tagami, H., Nakatani, Y., and Felsenfeld, G. (2004). CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. Mol. Cell *13*, 291–298.

Zang, C., Schones, D.E., Zeng, C., Cui, K., Zhao, K., and Peng, W. (2009). A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. Bioinformatics *25*, 1952–1958.

Zeng, P.-Y., Vakoc, C.R., Chen, Z.-C., Blobel, G.A., and Berger, S.L. (2006). In vivo dual cross-linking for identification of indirect DNA-associated proteins by chromatin immunoprecipitation. BioTechniques *41*, 694, 696, 698.

Zhang, Q., and Wang, Y. (2008). High mobility group proteins and their post-translational modifications. Biochim. Biophys. Acta *1784*, 1159–1166.

Zhao, J., Herrera-Diaz, J., and Gross, D.S. (2005). Domain-wide displacement of histones by activated heat shock factor occurs independently of Swi/Snf and is not correlated with RNA polymerase II density. Mol. Cell. Biol. *25*, 8985–8999.

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*, 402–405.

Zhu, N., and Hansen, U. (2007). HMGN1 modulates estrogen-mediated transcriptional activation through interactions with specific DNA-binding transcription factors. Mol. Cell. Biol. *27*, 8859–8873.

Zhu, N., and Hansen, U. (2010). Transcriptional regulation by HMGN proteins. Biochim. Biophys. Acta *1799*, 74–79.

Zlatanova, J., Bishop, T.C., Victor, J.-M., Jackson, V., and van Holde, K. (2009). The nucleosome family: dynamic and growing. Structure *17*, 160–171.

Zou, Y., Jiang, X., and Wang, Y. (2004). Identification of novel in vivo phosphorylation sites in high mobility group N1 protein from the MCF-7 human breast cancer cells. Biochemistry *43*, 6322–6329.