ANALYSIS OF CHROMATIN TARGETING MODULES IN THE CHROMATIN REMODELLING ENZYME NURF

BY

BOYUN JANG

A thesis submitted to the
University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Institute of Biomedical Research
School of Immunity and Infection
College of Medical and Dental Sciences
University of Birmingham
October 2013

ABSTRACT

Drosophila nucleosome remodelling factor (NURF) is one of the founding members of the ISWI family of ATP-dependent chromatin remodelling enzymes and mediates energy-dependent nucleosome sliding leading to transcription regulation. In previous work (Wysocka et al., 2006), NURF was shown to be recruited to gene targets by binding specific histone modifications. The largest subunit of NURF, NURF301, contains a bromodomain and three PHD finger domains that have the ability to recognize specific histone modifications. Here we determine the histone binding-specificities of these domains, and how NURF histone binding is influenced by histone modification "cross-talk" - the influence of combinations of flanking histone modifications. This has been analyzed by histone peptide library array assays and our study shows that the PHD2 domain specifically recognizes the histone H3K4me3 mark. This binding can be inhibited by phosphorylation of H3 Thr 3 (H3T3p), while enhanced by acetylation of H3 Lys 9 (H3K9ac) and phosphorylation of Ser 10 (H3S10p). The bromodomain binds to histone H4 multiply acetylated at the K5, K8, K12 and K16 positions. The PHD1 domain binds to the H3K23me3 while the N-terminal PHD domain does not show specific histone modifications but binds generally to the H3 and H4 N-termini. This binding specificity was confirmed by peptide pull-down, Biacore and immunofluorescence microscopy assays. Moreover, two different NURF301-A/B and NURF301-C isoforms were CTAP-tagged by recombineering, and we used chromatin immunoprecipitation coupled sequencing (ChIP-Seq) to profile the genome-wide distribution of NURF. These data have allowed us to confirm the influence of these modifications on NURF recruitment to chromatin *in vivo*. Therefore, our results identify regulatory mechanisms of histone modifications directing recruitment of ATP-dependent chromatin remodelling enzymes.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

CHAPTER 2. MATERIALS AND METHODS

CHAPTER 3. IDENTIFICATION OF THE MODIFIED HISTONE BINDING SPECIFICITIES OF THE NURF301 PHD DOMAINS AND BROMODOMAIN *in vitro*

CHAPTER 4. IMMUNOFLUORESCENCE MICROSCOPY TO LOCALIZE THE HISTONE MODIFICATIONS: H3K4me3, H3T3p, H3K9ac, H3S10p, H3K23me3 and NURF301 in *Drosophila in vivo* AND ChIP SEQUENCING OF NURF301 ISOFORMS

CHAPTER 5. DISCUSSION

CHAPTER 6. APPENDIX

LIST OF TABLES

LIST OF DEFINITIONS

Ac, ac	acetyl
ACF	ATP-utilizing chromatin assembly and remodelling factor
ADP	adenosine di-phosphate
AIRE	human autoimmune regulator
Arg (R)	arginine
ATP	adenosine tri-phosphate
BAC	bacterial artificial chromosome
BHC80	benign hereditary chorea 80 protein
BLAST	basic local alignment search tool
bp	base pair(s)
BPTF	bromodomain PHD finger transcription factor
BSA	bovine serum albumin
CBP	CREB binding protein
cDNA	complementary deoxyribonucleic acid
CHD	chromodomain-helicase-DNA-binding protein
ChIP	chromatin immunoprecipitation
CHRAC	chromatin accessibility complex
Cy3	cyanine 3
Da	dalton
DAPI	4',6-diamidino-2-phenylindole
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EST	expressed sequence tag
FITC	fluorescein isothiocyanate
g	gram
Gal	galactose
GalK	galactokinase

GST	glutathione S-transferase
H3K4me3	Histone 3 lysine 4 tri-methylation
H3K4me3/K9ac	Histone 3 tri-methylated at lysine 4 and acetylated at lysine 9
H3K4me3/K9ac/S10p	Histone 3 tri-methylated at lysine 4, acetylated at lysine 9 and phosphorylated at serine 10
H3K23me3	Histone 3 lysine 23 tri-methylation
H3K36ac	Histone 3 lysine 36 acetylation
H3K36me1	Histone 3 lysine 36 mono-methylation
H3K36me2	Histone 3 lysine 36 di-methylation
H3K36me3	Histone 3 lysine 36 tri-methylation
H3T3p	Histone 3 threonine 3 phosphorylation
H3T3p/K4me3	Histone 3 phosphorylated at threonine 3 and tri-methylated lysine 4
H4K5,8,12,16ac	Histone 4 acetylated at lysine positions 5, 8, 12, 16
H4K16ac	Histone 4 lysine 16 acetylation
HAT	histone acetyltransferases
HDAC	histone deacetylase
HMG	high mobility group
HPTM	histone post translational modification
HRP	enzyme horseradish peroxidase
IgG	immunoglobulin G
ISWI	imitation switch protein
JAK/STAT	janus kinase/signal transducers and activators of transcription
kg	kilogram
kDa	kilodalton
L	litre, lambert
Lys (K)	lysine
M	molar (as in 0.1 M solution)
Me, me	methyl
min, mins	minute, minutes
mg	milligram

ml	millilitre
MNase	micrococcal nuclease
mol	mole
ng	nanogram
NMR	nuclear magnetic resonance
NURF	nucleosome remodelling factor
OD	optical density
ORF	open reading frame
p, pho	phospho
PCR	polymerase chain reaction
pH	negative log of hydrogen ion concentration
PHD	plant homeodomain
PTM	post translational modification
PVDF	polyvinylidene difluoride
RNA	ribonucleic acid
RNase A	ribonuclease A
rpm	revolutions per minute
SDS	sodium dodecyl sulfate
sec, secs	second, seconds
Ser (S)	serine
siRNA	small interfering RNA
TAP	tandem affinity purification
TFIID	transcription factor II D complex
Thr (T)	threonine
TRF	TBP-related factor
°C	degree, Celsius
µg	microgram

CHAPTER 1. INTRODUCTION

1.1 The basic structure of chromatin

DNA in the nucleus of eukaryotic cells does not exist in a free state but is found in repeated functional units called nucleosomes that form chromatin. Nucleosomes are the basic unit of chromatin and consist of a segment of DNA wrapped around the histone octamer (Marino-Ramirez et al., 2005, see Fig 1-1). The histone octamer is composed of two copies each of the four histone proteins H2A, H2B, H3 and H4 around which 147 bp of DNA are wound in 1.65 superhelical turns with five to ten fold compaction of DNA (Luger et al., 1997). In addition,

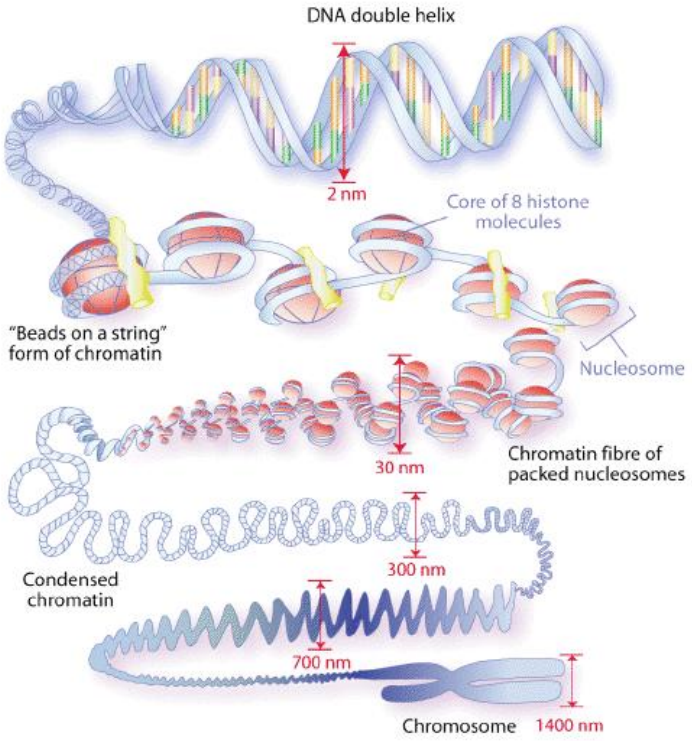


Figure 1-1. The model of chromatin structure. In eukaryotes, nucleosomes are composed of a core particle of eight histones around which147 bp of DNA are wrapped. This basic unit of chromatin forms 10 nm fibres the "beads on a sting". This is then compacted into the 30 nm fibres and condensed further to produce 300 nm and 700 nm fibres. Figure from Annunziato (2008).

histone H1 (or H5 in avian species), as linker histone, plays an important role providing increased protection of approximately 168 bp of DNA and is responsible for changing nucleosomal structure to make the more highly packaged chromatosome (Widom, 1998; Woodcock et al., 2006). These repeating nucleosomes called "beads on a string" are proposed to be arranged into 30 nm fibres, which result in 50 fold compaction (Felsenfeld G. and Groudine M., 2003). The 30 nm fibres are then proposed to be progressively condensed into 300 nm and 700 nm fibres ultimately to form chromosomes. However, the existence of the 30 nm fibres and the hierarchy of folding is still controversial (Staynov, 2008; Grigoryev and Woodcock, 2012; Luger et al., 2012). By modifying nucleosome dynamics and the interactions between DNA and histone proteins, these higher order structures can be disrupted to locally or globally decondense chromatin structure and allow access of the enzymes that process genetic information.

1.2 Chromatin and gene regulation

Chromatin has been categorised in two different functional forms which are euchromatin and heterochromain. These were originally defined by Heitz (1928) who distinguished them cytologically by the level of staining during interphase. Euchromatin stains lightly and appears less condensed, whereas heterochromatin stains darkly and appears highly condensed. Euchromatin is thought to be transcriptionally active chromatin, while heterochromatin as transcriptionally inactive chromatin. It was later revealed that heterochromatin could be divided into two further categories - facultative heterochromain, which is interchangeable between open or compact conformations, and constitutive heterochromain such as centromeres and near telomeres (Trojer and Reinberg, 2007).

The highly integrated nature of DNA in nucleosomes in chromatin means that DNA is not freely exposed to RNA polymerase or transcription factors. In this way, compaction of DNA in chromatin has the ability to block transcriptional regulatory elements and promoters leaving them inaccessible for transcription. Moreover, it had been assumed that nucleosomes on the body of genes would pose a barrier to processive RNA polymerases. Surprisingly, recent data has shown that nucleosome structure is maintained during the active transcription (Jiang and Pugh, 2009). Therefore, mechanisms must exist that allow RNA polymerase to overcome the barriers imposed by nucleosomes and higher-order chromatin structure. Many studies have shown that the nucleosome is not a simple static unit but has dynamic properties which are tightly regulated by various protein complexes (Li and Widom, 2004; White and Luger, 2004). Moreover, the dynamic properties of nucleosomes are regulated by a number of mechanisms including histone variant incorporation, histone eviction, histone post-translational modification and chromatin remodelling which affect all steps of transcription (Li et al., 2007).

1.3 Changes in chromatin structure

1.3.1 Histone variants

The initial concepts of nucleosome formation and composition were that nucleosomes are composed of core histone proteins (canonical histones) that are synthesized during S phase when DNA is replicated and chromatin structure would be duplicated (Kornberg et al., 1999). However, non-allelic variants of the major histone proteins have been found in diverse species. These include the histone H2A variants MacroH2A, H2A.Z, H2ABbd and H2A.X and the histone H3 variants H3.3 and centromere-specific histone CENP-A (Redon et al., 2002). These variants have diverse roles in epigenetic silencing, architectural changes in

chromosomes, gene expression and centromere function (Kamakaka and Biggins, 2005; Malik and Henikoff, 2003). Generally, canonical histones are expressed during S phase from clustered histone gene cassettes and deposited during DNA replication. In contrast, most histone variants are produced outside of S phase and incorporated in a process that is independent of DNA replication (Jin et al., 2005). They can be displaced in a transcription dependent manner and by protein complexes involved in transcription, replication and ATP-dependent nucleosome remodelling. For example, this histone exchange was firstly recognized by monitoring the exchange of fluorescently labelled histones after photobleaching of a small area of the nucleus. These experiments showed higher exchange levels of canonical histone H2A and H2B than histone H3 and H4 (Kimura and Cook, 2001). Later, Bruno et al. (2003) observed that several chromatin remodelling complexes, including the SWI/SNF, RSF and ISWIb complexes (Peterson et al., 1994; Peterson et al., 1995), can catalyze the exchange of H2A-H2B dimers between chromatin fragments in an ATP-dependent reaction.

It has been suggested that histone exchange reactions perform two functions *in vivo* (Sarma and Reinberg, 2005). Firstly, it allows the removal of all histone epigenetic marks and promotes reprogramming of genes. Secondly, the incorporated histone variants may have distinct biochemical properties that allow them to carry out alternative functions in cells.

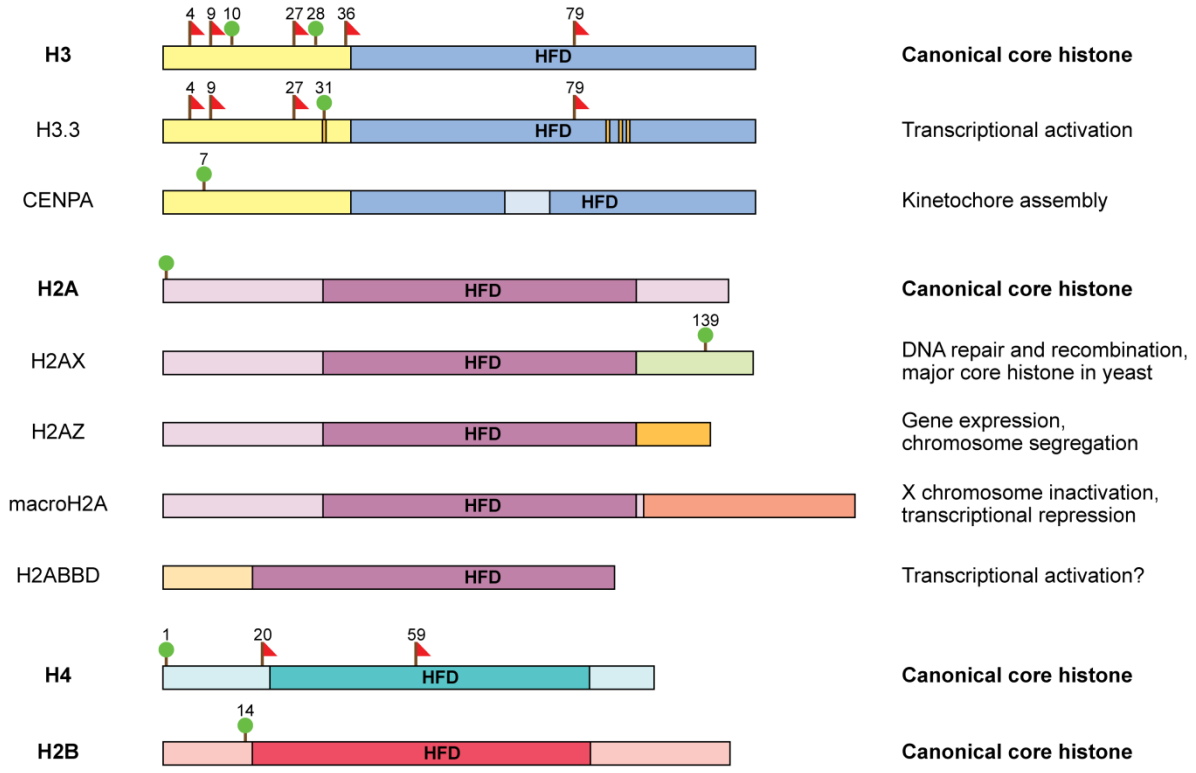


Figure 1-2. Domain structure of the canonical histones (H3, H2A, H4 and H2B) and their variants. The histone-fold domain (HFD) is highly conserved in both canonical histones and variants. The length and major histone modification positions of N- and C-terminal tails are slightly different. Well-established histone modifications, methylated lysines and phosphorylated serines are indicated by red flags and green flags respectively. Figure adapted from Sarma and Reinberg (2005).

Histone variants generally have substantial overall sequence identity within their canonical core histone-fold domain (HFD) and most post-translational modification sites are conserved between canonical histones and variant histones. As shown in figure 1-2, four histone H3, variants have been found: H3.1, H3.2, H3.3 (which are highly conserved and exhibit only a few amino acids differences) and CENPA, which is a centromeric histone H3 that has a divergent N-terminal region but similar C-terminal HFD with histone H3 (Marzluff et al., 2002; Hake et al., 2006). Histone H3.3 is found at transcriptionally active loci (Ahmad and Henikoff, 2002) and is incorporated during all stages of cell cycle (McKittrick et al., 2004). In contrast, cetromeric histone CENPA also called CenH3, is localized specifically at

centromeres but also spreads over the chromosome arms when overexpressed (Sullivan et al., 1994).

Histone H2A variants are also well studied. Four variants have been reported: H2A.Z, H2A.X, Macro H2A and H2A-bar-body-difficient (H2A.Bbd). The H2A.Z variant is highly conserved (Jackson et al., 1996), while Macro H2A and H2A.Bbd are found exclusively in vertebrates or mammals (Pehrson and Fuji, 1998; Chadwik and Willard, 2001). The H2A variants are distinguished from the canonical H2A histones by their C-terminal tails that vary in both length and sequence (Figure 1-2). The histone variant H2A.Z is the best studied H2A variant. Previous ChIP-chip data from yeast showed that H2A.Z is incorporated into specific promoter-bound nucleosomes and might confer changes in the chromatin structure that would allow transcriptional activation (Guillemette et al., 2005). Work in our laboratory has shown that the mutation of H2A.Z increases the hemocyte tumour incidence in concert with the nucleosome remodelling enzyme NURF in *Drosophila* (Badenhorst, unpublished data) suggesting that H2A.Z might be also involved in nucleosome positioning.

Little is known about H2B and H4 variants. Two testis specific H2B variants were identified in human (Zalensky et al., 2002; Churikov et al., 2004), but no H4 variant has been found so far.

1.3.2 Post-translational histone modifications

Another mechanism by which chromatin structure can be altered is through the incorporation of post-translational modifications on the histone proteins. Histone post translational modification (HPTM) targets both the histone tails and globular domains. It has been known

that post-translationally modified nucleosomes are distributed in distinct localized patterns within the upstream region of the transcription start site (TSS), enhancer, the core promoter, the 5' end of the open reading frame (ORF) and the 3' end of the ORF (Li et al., 2007; Wang, 2011; Zentner et al., 2011). These modifications can be divided into 4 different types: acetylation, methylation, phosphorylation and ubiquitylation/sumoylation. Either individually or in combination, these modifications can have a great effect on genome regulation and function (Figure 1-3).

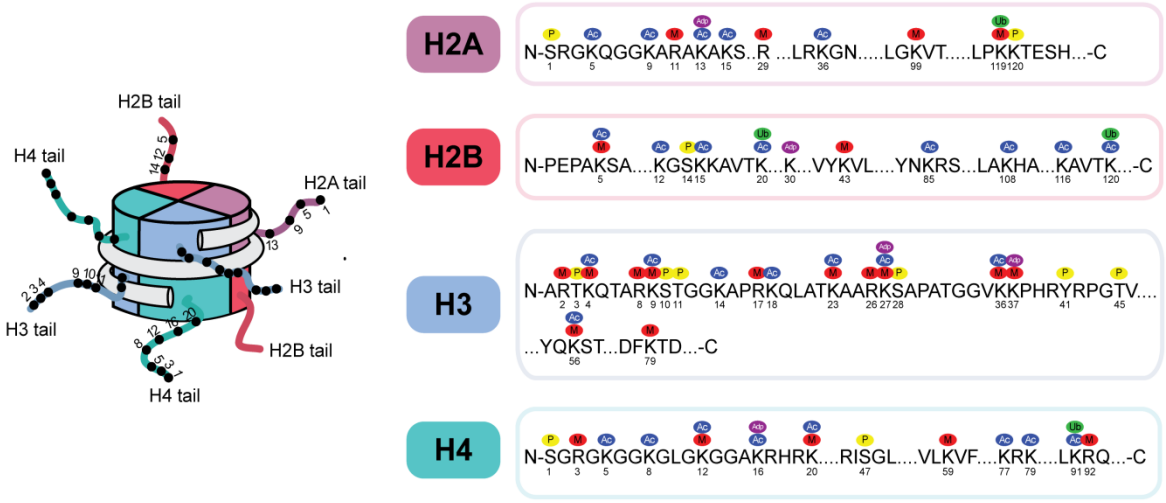


Figure 1-3. The main post translational modification sites on all canonical histones. In particular, histone N-terminal tails are subject to the modifications but some histone core residues are also modified. Modifications include: acetylation (blue), methylation (red), phosphorylation (yellow), ubiquitination (green) and ADP ribosylation (purple). Figure modified from Portela and Esteller (2010).

For each modification, enzyme pairs exist that either can put in place the modification or remove the modification. For example, histone acetyltransferases (HATs)/deacetylases (HDACs) for acetylation and methyltransferases/demethylases for methylation exist.

Acetylation was the first histone modification found in 1963 (Phillips, 1963). It has been

known that high level of acetylation and HAT activity is closely related to active transcription and also known as a major determinant of the euchromatin (Alfrey et al., 1964; Pogo et al., 1966; Hendzel et al., 1998). This function of acetylation can partly be explained by the neutralizing of the positive charge of lysine residues, which decreases the charge-interaction of the histone with the nucleosomal DNA, linker DNA or adjacent histones (Norton et al., 1989; Norton et al., 1990; Zentner and Henikoff, 2013). Therefore, accessibility of DNA to transcription factors can be increased. For example, tetra-acetylation of the histone H4 tail markedly reduces its binding to DNA *in vitro* (Hong et al., 1993). In addition, the acetylation modification can be also directly recognized by many proteins, so called "readers", which include transcription factors, activators and chromatin remodelers containing conserved chromatin binding domain like bromodomain (Struhl, 1998; Josling et al., 2012). In contrast, HDACs are more likely to work on transcriptionally silenced regions, which contain a low level of acetylation such as heterochromatin (Li et al., 2007). This means that sites of reversible acetylation can play an important role in gene regulation by turning on/off transcription (Shogren-Knaak et al., 2006).

Methylation of histones occurs at lysine residues in histone H3 (K4, 9, 27, 36, 79) and H4 (K20). Methylation plays a key role in both gene activation and repression. What determines the functional consequences of methylation is the identity of the residue that is methylated or demethylated. Methylated histone residues provide marks for the recruitment of effector complexes that can either activate or repress transcription. For example, methylation of histone H3 lysine 4 (H3K4) is associated with transcription activation and exerts its effect by recruiting the NURF complex that can mobilize nucleosomes at active genes in *Drosophila* and human (Li et al., 2006; Wysocka et al., 2006; Kwon et al., 2009). In contrast, methylation

of H3K9 has repressive effect by recruiting heterochromatin protein 1 (HP1) (Stewart et al., 2005), while methylation of H3K27 also represses transcription by mediating the recruitment of the Polycomb repressive complex (Martin and Zhang, 2005).

Phosphorylation is one of the most well studied HPTMs and the roles of kinases have been understood for a long time. For chromatin, kinases target phosphorylation sites in the H3 amino terminal tail, which have important functions in transcription activation and the control of the cell cycle. The main sites of phosphorylation include H3 Thr 3 (H3T3), H3 Ser 10 (H3S10) and H3 Ser 28 (H3S28). For example, phosphorylation of H3T3 by the kinase Haspin is able to affect mitotic chromosome regulation by maintaining cohesion (Dai et al., 2006). H3S10 phosphorylation is highly conserved and is known as an important phosphorylation site in humans, *Tetrahymena* and *Drosophila* (Nowak and Corces, 2004), while H3S28 phosphorylation is relatively less well studied but recently found to be involved in Polycomb silencing (Lau and Cheung, 2011). One of the consequences of histone phosphorylation is to make the site negatively charged so reduce the affinity of histones for DNA. In this manner, chromatin compaction which is closely related to transcription can be altered.

Ubiquitylation and sumoylation are distinct from acetylation, methylation and phosphorylation in that these modifications are large polypeptides which increase histone size by approximately 2-3 times. They also share similar sequences and three dimensional structures but are dissimilar in surface charge. Ubiquitylation acts on histone H2A and H2B and is associated with repression and activation of transcription respectively. For example, mono-ubiquitylation of H2B activates transcription by Rad6/Bre1 and leads to H3K4

methylation which is mark of active transcription as described above (Wood et al. 2003; Kim et al., 2005; Henry et al., 2003). In contrast, mono-ubiquitylation of H2A is repressive to transcription on Hox genes in mammals (Wei et al., 2006). Sumoylation was first found as a repressive histone modification in the yeast *Saccharomyces cerevisiae* (Nathan et al., 2006). Moreover, it interacts with other positive modifications like acetylation or ubiquitylation as it acts as a potential block to these modifications.

Recently ADP ribosylation and glycosylation were also identified as important histone modifications. The core histones and the linker histone H1 can potentially be ADP ribosylated. This modification has been shown to be involved in DNA repair, cell cycle, replication and transcription (Alvarez et al., 2011; Hassa et al., 2006; Messner and Hottiger, 2011). Glycosylation is present on several threonine and serine residues of the histones H2A, H2B and H4, but its function remains to be determined (Sakabe et al., 2010; Zhang et al., 2011).

To conclude, PTM could have three consensus functions. Firstly, with the exception of methylation, PTM results in a change in the net charge of the nucleosome and can modify nucleosomal interactions with DNA. Secondly, histone modifications can act as marks that are recognized by other effector proteins. Finally, some modifications directly influence higher-order chromatin structure.

1.3.3 ATP-dependent chromatin remodelling factors

To make transcription elements accessible in packaged DNA in a regulated manner, cells have evolved specialized famillies of chromatin remodelling complexes. These chromatin remodelling enzymes utilize ATP hydrolysis as an energy source, to move, destabilize, eject,

or restructure nucleosomes (Becker and Horz, 2002). There are four families of chromatin remodelling complexes: ISWI, INO80, Mi-2 and SWI/SNF families, based on the sequence homology of the associated catalytic ATPase subunit. They are conserved from yeast to human and all share some properties including binding affinity for the nucleosome, domains that can recognize histone modifications, ATPase domains and domains that can interact with chromatin or other transcription factors (Clapier and Cairns, 2009, Figure 1-4). However, they also have distinctive flanking domains making each family specific in its particular task (Flaus et al., 2006).

The ISWI (imitation switch) family of chromatin remodelling enzymes were initially identified in *Drosophila* (Elfring et al., 1994) and consist of the related NURF, CHRAC and ACF complexes. These complexes share the core ISWI catalytic subunit that contains a ATPase domain but also SANT and SLIDE domains that bind to unmodified histone tails and DNA (Boyer et al., 2004). The ISWI family complexes are composed of other subunits in addition to the ISWI subunit, and these impart many different domains including PHD domains, bromodomains and DNA-binding histone fold motifs. Most ISWI family complexes like ACF and CHRAC, promote chromatin assembly with regular spacing of nucleosomes and transcription repression (Längst and Becker, 2001), but they can have other functions depending on the attendant proteins (Corona and Tamkun, 2004b).

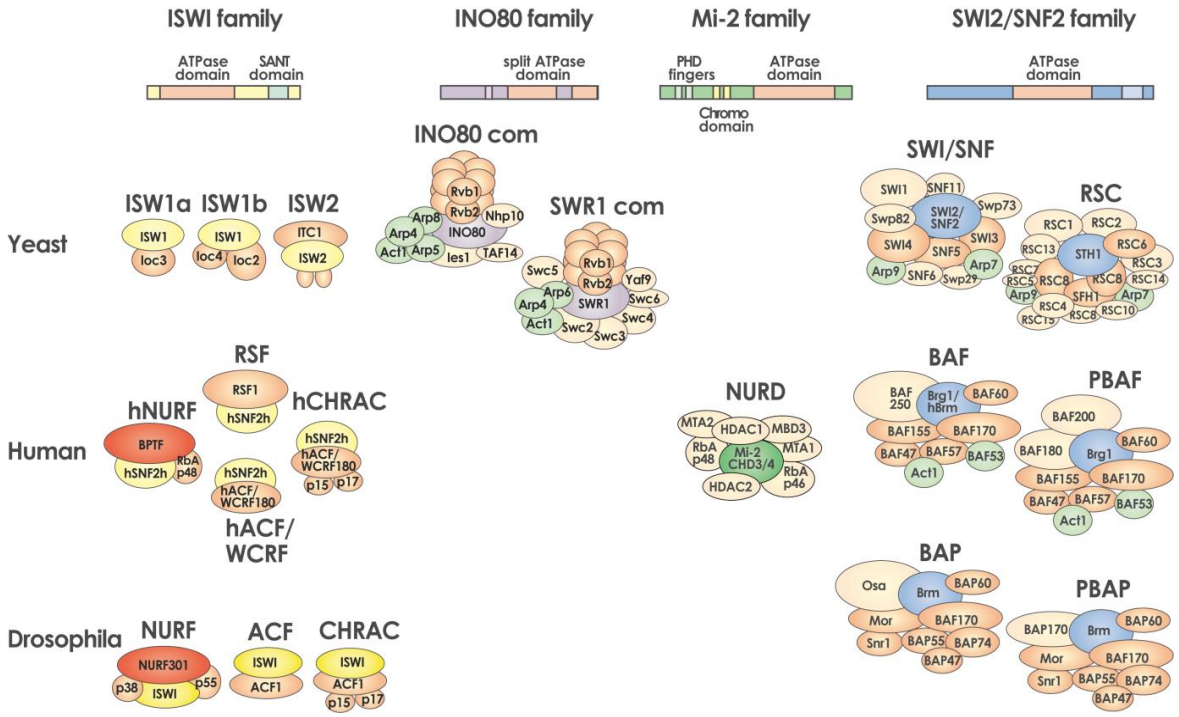


Figure 1-4. ATP dependent chromatin remodelling factor families : ISWI, INO80, Mi-2 and SWI2/SNF2 families. Figure adapted Badenhorst (2013).

The INO80/SWR1 (inositol requiring 80) complexes were initially purified from *S.cerevisiae* and have more than 10 subunits (Shen et al., 2000). The distinctive feature of these chromatin remodelling enzymes is that they have DNA strand displacement activity and, therefore, are ATP-dependent helicases different from other chromatin remodelling factors. This activity depends on the ATPase function of the INO80 protein (Bao and Shen, 2007). It is also involved in the facilitation of DNA repair (Jonsson et al., 2001). The SWR1 complex containing Swr1p subunit, a Swi2/Snf2-paralogue, specifically replaces histone H2A in nucleosomes for the H2AZ variant that could lead to structural or functional changes in the chromatin as described in chapter 1.3.1 (Kobor et al., 2004; Mizuguchi et al., 2004).

The Mi-2 family, also called CHD (chromodomain, helicase, DNA binding) family, of remodelers contain two tandemly arranged chromodomains (chromatin organization modifier)

and interact with subunits that contain DNA-binding domains and PHD, BRK, CR1-3 and SANT domains (Figure 1-4). Some members of the CHD family of chromatin remodelling enzymes are positive regulators of transcription that slide or eject nucleosomes (Bouazoune et al., 2002; Clapier and Cairns, 2009), but others are repressors of transcription, for example the Mi-2/NuRD complex represses transcription via its chromatin remodelling and histone deacetylase activities (Denslow and Wade, 2007) Moreover recently, *Drosophila* Mi-2 was found as a regulator of chromosome condensation and cohesin binding *in vivo* (Fasulo et al., 2012)

SWI/SNF (switching defective/sucrose nonfermenting) complex was originally found in *S. cerevisiae* as a transcription activator (Winston and Carlson, 1992) and is a large complex containing 8 to 14 subunits. Among these subunits, SWI2/SNF2 has a major role to remodel nucleosome arrays as an ATPase. The SWI/SNF complex is evolutionally conserved and in *Drosophila*, the closest relative of SWI2/SNF2 is Brahma (Tamkun et al., 1992). Brahma is at the core of the Brahma or BAP-associated complexes which both have ATPase activity and contain six subunits closely related to yeast SWI2/SNF2 (Dingwall et al., 1995; Kal et al., 2000). Mammalian SWI/SNF on the other hand, has two ATPases, Brg1 and Brm, and is known to be sufficient to remodel nucleosome arrays *in vitro* (Phelan et al., 1999; Euskirchen et al., 2011). The role of other subunits of these complexes has not been well characterized, but BAF155 and BAF170 were found to support that remodelling activity and function as scaffolds of other SWI/SNF subunits (Sohn et al., 2007). SWI/SNF complexes can disrupt nucleosomes either by repositioning or by dissociation, thus can have a major impact on transcriptional control, DNA repair, recombination and chromosome segregation (Clapier and Cairns, 2009; Liu et al., 2011). SWI/SNF chromatin remodelling is linked to tumorigenesis

with mutations in SWI/SNF subunits having been identified in a number of cancers (Wilson and Roberts, 2011)

.

1.4 The nucleosome remodelling factor (NURF)

1.4.1 Identification

The ATP-dependent nucleosome remodelling factor (NURF) is one of the ISWI family of chromatin remodelling enzymes. It was firstly identified and purified from *Drosophila* embryo extracts by assaying *Drosophila* embryo extracts (S150) for GAGA factor and ATP - dependent nucleosome disruption of *in vitro* chromatinized plasmids containing the *hsp70* promoter (Tsukiyama and Wu, 1995a; Tsukiyama et al., 1995b; Clapier et al., 2009). In these studies, two components were found to be critical for nucleosome disruption at the *hsp70* promoter: the GAGA factor and a sarkosyl-sensitive nucleosome remodelling factor in the S150 extract which utilizes ATP and was named NURF (Tsukiyama and Wu, 1995a). NURF was purified by fractionation of S150 *Drosophila* embryo extracts and ultimately revealed to be a four subunit protein complex consisting of the catalytic ISWI ATPase subunit and three additional subunits (Tsukiyama et al., 1995). As shown in figure 1-5, the NURF complex is composed of four subunits, NURF140 (or ISWI): ATPase (Tsukiyama et al., 1995b) which also exists in other ISWI members (ACF and CHRAC), NURF55: a WD-40 repeat protein which is found in many chromatin regulatory complexes (Martinez-Balbas et al., 1998; Hennig et al., 2005), NURF38: inorganic pyrophosphatase (Gdula et al., 1998), and the largest subunit of NURF named NURF301 (Xiao et al., 2001).

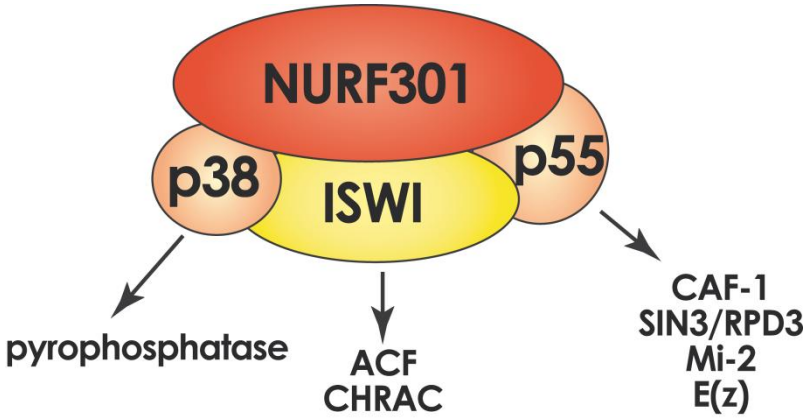


Figure 1-5. Structure of the NURF complex. NURF is composed of four subunits including ISWI, NURF55 (p55), NURF38 (p38) and the largest and unique subunit, NURF301. Figure adapted Badenhorst (2013).

Detailed biochemical experimentation has defined similarities but also important differences between the NURF and SWI/SNF chromatin remodelling complexes. NURF is similar to the yeast SWI/SNF complex as both hydrolyze ATP, contain catalytic ATPase domains with high sequence similarity, and structurally alter nucleosome particles (Hamiche et al., 1999). However NURF, as the founding member of the ISWI subfamily of chromatin remodelling enzymes, has a distinct size as it is composed of four polypeptides, whereas SWI/SNF is composed of more than 10 subunits. Moreover, the level of NURF expression is also higher (~ 0.08% of nuclear extract protein) than SWI/SNF protein which is only detected at ~100 molecules in a yeast cell (Cote et al., 1994). The ATPase activity of the NURF and SWI/SNF complexes is also stimulated in different ways, in that NURF ATPase activity is only stimulated by nucleosomes, while SWI/SNF ATPase activity is stimulated by free DNA (Georgel et al., 1997; Corona et al., 1999). Finally, although the yeast and human SWI/SNF complexes alter the DNaseI digestion patterns of reconstituted mononucleosome (Imbalzano et al., 1994; Kwon et al., 1994), NURF shows specific induction of DNaseI protection and hypersensitivity at discrete sites on nucleosomal DNA (Hamiche et al., 1999).

1.4.2 Homologues – human and yeast NURF like complexes

Since *Drosophila* NURF was purified, other complexes similar to NURF had been identified including the yeast ISW1 and ISW2 complexes (Tsukiyama et al., 1999) and the human NURF complex. Human NURF is almost identical to *Drosophila* NURF consisting of a large subunit BPTF (bromodomain and PHD domain transcription factor), a catalytic ISWI-like subunit SNF2L and CAF-1, a NURF55-like subunit (Barak et al., 2003). Importantly, there is high protein sequence identity (~35%) and sequence similarity (~50%) over the entire coding region between *Drosophila* NURF301 and human BPTF (Xiao et al., 2001, Figure 1-6). This allows *Drosophila* NURF and NURF301 to be used as good models to understand the function of human NURF.

1.4.3 *In vitro* biochemistry

The initial discovery of NURF was achieved using *in vitro* assays base on *in vitro* chromatinization of *hsp70* promoter sequences using a cell free *Drosophila* embryo extract (Tsukiyama and Wu, 1995a). Early characterizations of NURF functions were largely performed using the *hsp70* promoter although subsequently chromatin was reconstituted on this fragment using recombinant core histones and nucleosomes assembled by salt dialysis (Hamiche et al., 1999). Using this system, the nucleosome sliding ability, sliding position preference and comparison of remodelling efficiencies of ISWI alone or in complex with NURF were elucidated (Hamiche et al., 1999). Analysis of *in vitro* sliding have shown that only two subunits are absolutely required for nucleosome sliding, ISWI and NURF301 the largest subunit of NURF (Xiao et al., 2001). NURF301 is required for efficient nucleosome sliding and interacts with some sequence specific transcription factors, which allows it to target NURF recruitment to specific genes (Xiao et al., 2001). NURF301 is the only NURF-

specific subunit of NURF as other subunits are present in other chromatin remodelling complexes. In particular the ISWI subunit is present in three other complexes ACF (Ito et al., 1997a; Ito et al., 1997b), CHRAC (Varga-Weisz et al., 1997), and NORC (Strohner et al., 2001).

1.4.4 Biology and *in vivo* functions

Though the basic features of NURF were revealed by *in vitro* assays, it was essential to understand the physiological roles of NURF, in other words how NURF acts *in vivo*. Initial studies focused on mutants in the ISWI subunit of the NURF complex. These studies showed NURF potentially was involved in homeotic gene activation and chromatin condensation (Deuring et al., 2000). However, the weakness of these studies was that ISWI is also present in other chromatin remodelling complexes preventing these phenotypes being uniquely ascribed to NURF function. However, by analysing mutants in the NURF-specific NURF301 subunit it has been possible to show that NURF is required for homeotic gene expression and chromosome condensation (Badenhorst et al., 2002). *Nurf301* mutants also develop melanotic tumours in the third instar larvae suggesting that NURF is required for hematopoietic development (Badenhorst et al., 2002). By comparing whole genome expression profiles of *Nurf301* mutants with wild type larvae, target genes of NURF were also revealed to show functional involvement of NURF in ecdysteroid signalling and metamorphosis (Badenhorst et al., 2005). In addition these types of experiment also revealed that NURF is an important regulator of the JAK/STAT pathway, a signalling pathway that is closely involved in *Drosophila* innate immunity and spermatogenesis (Kwon et al., 2008; Kwon et al., 2009). Recent studies have shown that NURF function could be co-regulated by Putzig (Pzg) which is an integral component of the TRF2/DREF complex involved in regulation of replication

related gene activation such as Notch target genes (Kugler and Nagel, 2007; Kugler and Nagel, 2010; Kugler et al., 2011).

In contrast, the function of human NURF has not been well characterized, but the majority of functional studies indicate that BPTF is involved in neuronal development and neurodegenerative diseases such as Alzheimer's disease because of its altered expression and subcellular localization (Mu et al., 1997; Jordan-Sciutto et al., 2000; Strachan et al., 2004). As in *Drosophila*, it is also likely to have important role in adult germ line stem cell differentiation (Lazzaro et al., 2006).

1.4.5 Links between NURF chromatin remodelling and histone modifications

As described in the previous sections, *Drosophila* NURF interacts with transcription factors like GAGA, HSF (Heat Shock Factor), Ken, EcR (Ecdysone Receptor) and Pzg (Putzig). In addition, functional domains in the largest subunit NURF301 have the potential to bind to specific histone post-translational modifications. The interaction of NURF with histones during chromatin remodelling was primarily examined by determining the binding and sliding of NURF using reconstituted nucleosomes (Hamiche et al., 1999; Xiao et al., 2001). These nucleosomes are synthesized using bacterial over-expression systems and are devoid of histone post-translational modifications. However, recent studies have shown that NURF activity and recruitment can be regulated by histone post-translational modifications. In particular NURF ATPase (ISWI) activity and nucleosome binding are reduced by poly-ADP-ribosylation (Sala et al., 2008). Acetylation of ISWI by the acetyltransferase Gcn5 has also been observed, but the function of this modification has not been characterized (Ferreira et al., 2007). Other studies have shown that the NURF301 C-terminal PHD domain and

bromodomain can recognize specific histone protein modifications *in vitro* which allows targeted recruitment of NURF to sites containing the H3K4me3 and H4K16ac modifications (Wysoka et al., 2006; Kwon et al., 2009).

1.5 NURF isoforms and functional domains in NURF301

Three isoforms of *Drosophila* NURF301 have been identified: NURF301-A, NURF301-B and NURF301-C isoforms. NURF301-A and NURF301-B are full-length isoforms (~300 kDa) and contain several functional domains conserved in transcription factors and chromatin proteins. This includes a N-terminal HMG-A (High Mobility Group) domain, WAC/WACZ motifs, three PHD (Plant Homeo Domain) domains and a bromodomain as shown in figure 1-6. The HMG-A domain containing two AT hooks, which are short peptide motifs binding to the minor groove of AT-rich regions, as well as an acidic region (Reeves and Nissen, 1990; Reeves and Beckerbauer, 2001; Xiao et al., 2001). The function of NURF301 WAC/WACZ motifs has not been determined, but it was found important for the DNA-binding activity of ACF1 (Fyodorov and Kadonaga, 2002). PHD finger domains and bromodomain have been characterized in many chromatin remodelling proteins and potentially bind to specific histone modifications (which will be discussed in detail in the next chapter). Unlike the NURF301-A and NURF301-B isoforms, the NURF301-C is a C-terminally truncated isoform (~250 kDa) which terminates after the 7th exon of *Nurf301*. This isoform lacks the two C-terminal PHD domains and bromodomain. Our previous study showed that loss of those domains does not affect viability but does affect primary spermatocyte differentiation in *Drosophila* (Kwon et al., 2009). Further specific functions of NURF301-C isoform remain to be investigated. Therefore, it is important to understand the modified histone-binding specificity of these domains to determine precisely how NURF interacts with histone proteins and is recruited to

gene targets, and also to discriminate the functional differences between full-length NURF301-A/B and NURF301-C isoforms.

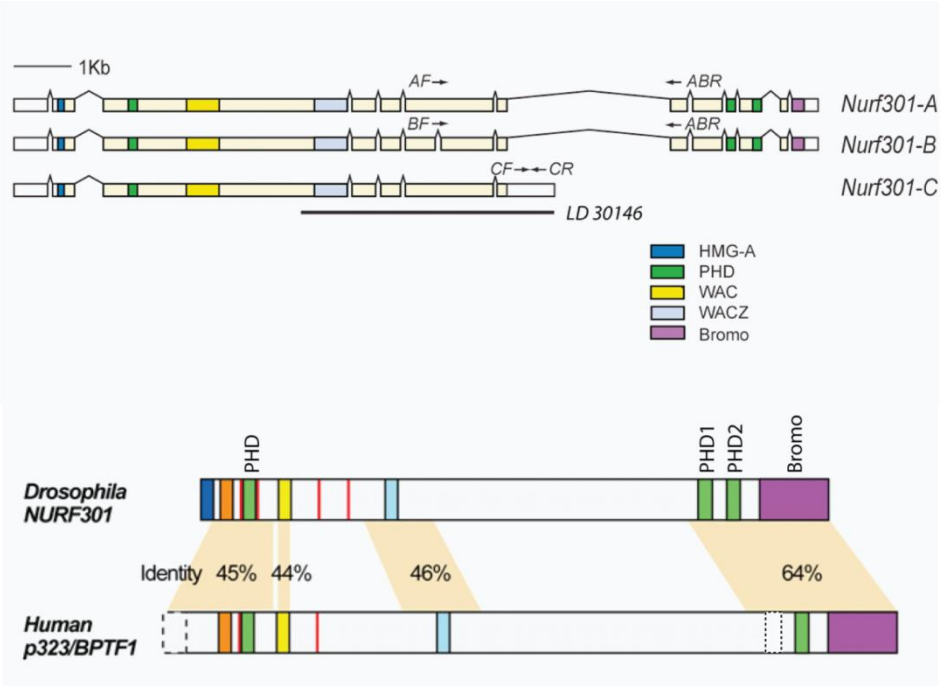


Figure 1-6. Domain structure and sequence comparison of *Drosophila* NURF301 with human BPTF. There are three NURF301 isoforms including full-length NURF301-A, NURF 301-B and C-terminal truncated NURF301-C isoforms. NURF301 contains a number of functional domains which are important for its chromatin binding (HMG-A, WAC/WACZ, three PHD domains and bromodomain). The PHD domains of NURF301 are numbered from the N-terminus to C-terminus PHD, PHD1 and PHD2 respectively. Figure adapted from Kwon et al. (2009).

1.5.1 PHD domain

Generally, PHD domains are divided into two major classes, one that is able to recognize methylated histone residues, and a second which binds to unmodified histone tails. Some PHD finger domains have a preference for H3K9me3 or H3K36me3, whereas other PHD fingers prefer to bind to H3K9ac or H3K14ac (Musselman and Kutateladze, 2009). These binding differences are structurally well characterized as described in figure 1-7. First, the PHD domain in complex with H3K4me3 has a highly conserved binding mechanism. For

example, structural analysis of the PHD domains of BPTF and ING2 (inhibitor of growth 2), which were the first identified H3K4me3 readers, in complex with histone H3K4me3 peptide revealed conserved features of their binding pockets (Figure 1-7 (a)). Both of BPTF and ING2 PHD domains contain a well defined aromatic cage composed of two to four aromatic residues where H3K4me3 is stably positioned. They also have a binding pocket often composed of acidic residues to bind to Arg 2 and backbone carbonyl nets generating hydrogen-bonding interactions with Ala 1 of the histone H3 peptide (Li et al., 2006; Peña et al., 2006; Musselman and Kutateladze, 2011).

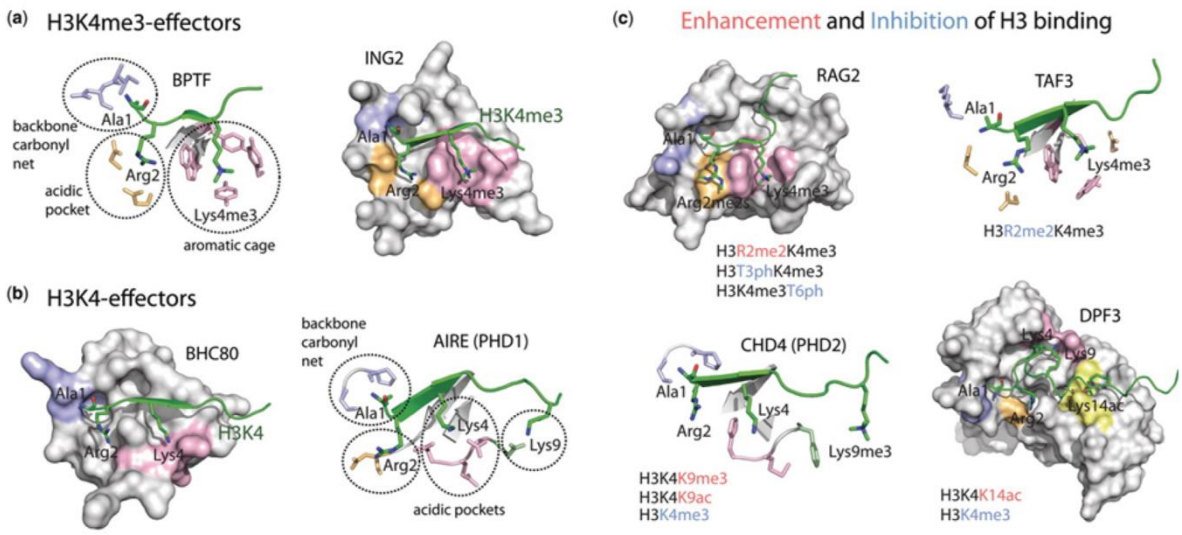


Figure 1-7. The structure of histone recognition by the PHD fingers. PHD fingers are specific for (a) H3K4me3 or (b) unmodified H3K4. (c) Enhancement and inhibition mechanisms of PHD finger binding to H3 by flanking modifications are also described. Figure was taken from Musselman and Kutateladze (2011).

PHD domains also bind to unmodified histone H3, such as BHC80 (Benign Hereditary Chorea 80) PHD domain or AIRE (human AutoImmune REgulator) PHD1 domain, also share some features that contain Ala 1 and Arg 2 binding pockets while they lack the aromatic cage (Figure 1-7 (b)). Instead, they have unique coordination with Lys 4 and some other basic

residues like Arg 8 and Lys 9 on the histone H3 tail, via the N-terminal acidic cluster forming salt bridges and hydrogen bonds with them. This interaction can be disrupted by methylation of Lys 4 (Lan et al., 2007; Chakravarty et al., 2009).

On the other hand, the PHD2 domain of CHD4 (chromodomain helicase DNA–binding protein 4) interacts with H3K9me3 by forming a cationic interaction with the tri-methy group of Lys 9 (Mansfield et al., 2011), while the PHD1 domain of DPF3 recognises H3K14ac in a binding pocket composed of hydrophobic and charged residues (Zeng et al., 2010, Figure 1-7 (c) below). Moreover, many studies are now showing the "cross-talk" between histone H3K4me3 and flanking modifications. For instance, a flanking di-methylated Arg 2 modification can enhance the binding of RAG2 (Ramón-Maiques et al., 2007), whereas it inhibits the binding of TAF3 (Van Ingen et al., 2008) to H3K4me3, indicating that additional modifications surrounding H3 Lys 4 differentially affects the binding of PHD domains to histone H3 tail (Figure 1-7 (c) above).

The PHD domains of NURF301 (from the N-terminus to C-terminus: PHD, PHD1 and PHD2) have sequence similarity, but it is possible that each domain has slightly different recognition specificities for histone modifications. Information of histone binding-specificity of NURF PHD domains was only available for the PHD2 domain, which has been shown to bind H3K4me3 (Li et al., 2006; Wysocka et al., 2006; Kwon et al., 2009). However, these experiments only addressed the binding of the PHD2 domain to a finite number of modified histone tails. The full spectrum of modified histone binding of the PHD2 domain has not been characterized. The binding properties of the PHD and PHD1 domains are unclear but BLAST comparison of the NURF301 PHD, PHD1 and PHD2 domains suggest that the NURF301

PHD domain is likely to be an unmodified histone H3 reader (Figure 1-8 A), while NURF301 PHD1 and PHD2 showed high sequence similarity to tri-methylated histone binders as expected (Figure 1-8 B and C).

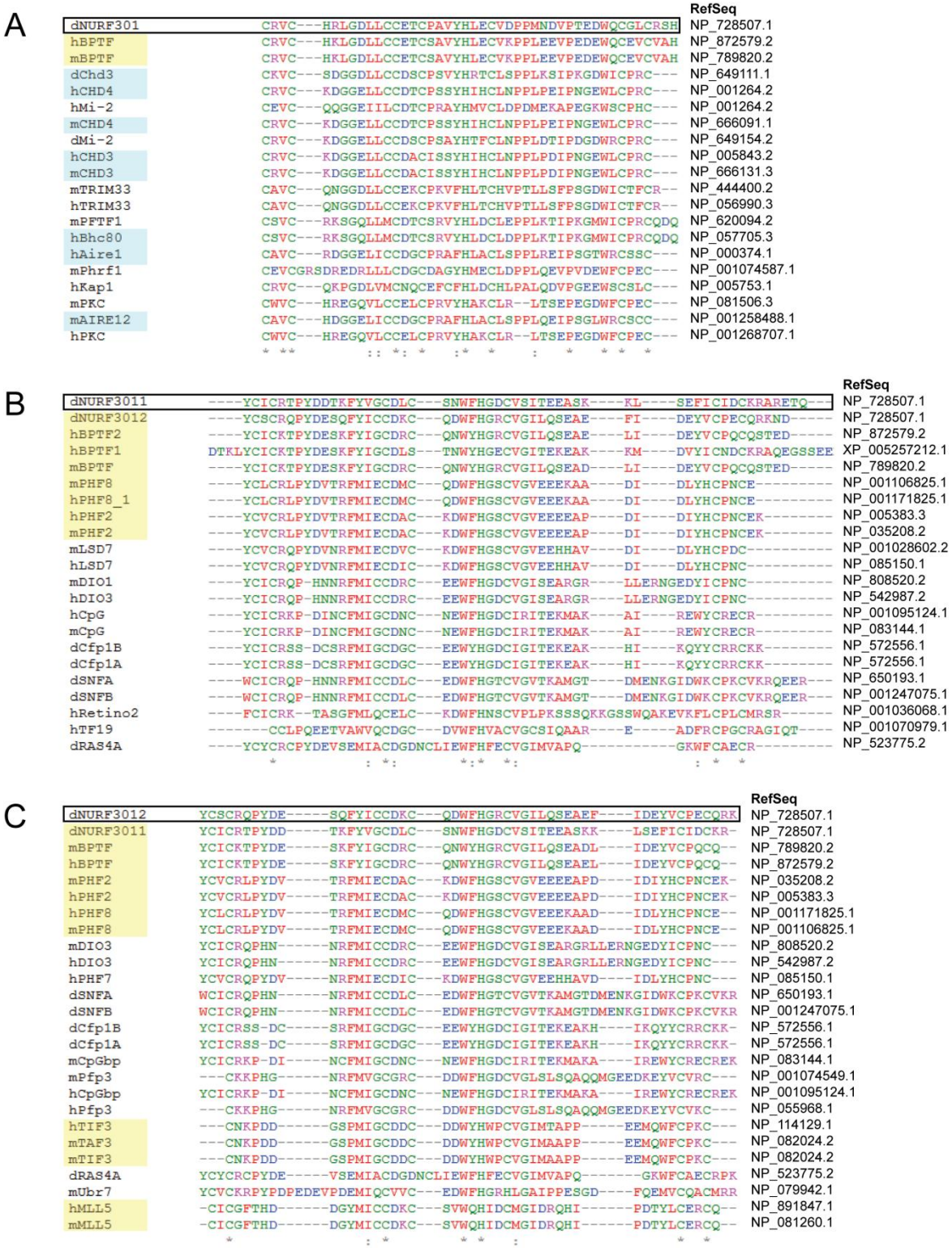


Figure 1-8. Protein sequence alignment of NURF301 (A) PHD, (B) PHD1 and (C) PHD2 domains. The top 20-25 proteins showing sequence similarity were selected from BLAST search results. Proteins known to recognize histone H3K4me3 are highlighted in yellow and unmodified histone H3 tail highlighted in blue. Reference sequence numbers are indicated.

1.5.2 Bromodomain

The bromodomain is commonly found in chromatin-associated proteins and specifically binds to acetylated lysine residues (Dhalluin et al., 1999). Many bromodomain structures have been determined, for example the bromodomain of the HATs P/CAF (Dhalluin et al., 1999; Mujtaba et al., 2002), Gcn5 (Hudson et al., 2000; Owen et al., 2000), Brd2,4 (Nakamura et al., 2007; Vollmuth et al., 2009) and BPTF (Ruthenburg et al., 2011). These analyses have shown that the structure of bromodomain is highly conserved. The bromodomain is composed of four left handed alpha helix bundles (helices $\alpha_{Z,}$ α_A, α_B and α_C) with long loops between helices α_B and α_C, and between α_Z and α_A (BC loop and ZA loop) which form a hydrophobic pocket that binds acetylated lysine residues (Filippakopoulos et al., 2012, Figure 1-9 A).

Most bromodomains can be categorized into three groups based on the functions of their associated proteins. The first bromodomain group are components of histone acetyltransferase (HAT) complexes and generally assist HAT complex recruitment to acetylated chromatin that could spread histone acetylations or recruit other transcription factors. For instance, GCN5 is a well known component of SAGA complex (Spt-Ada-Gcn5 acetyltransferase), contains a bromodomain and binds to acetylated H3 or H4K16ac (Barlev et al. 1998; Hassan et al., 2007). CBP(CREB binding protein) and p300 are also HAT domains containing bromodomain proteins, but the bromodomain of p300 interacts with all the core histones while the bromodomain of CBP binds to H4K20ac (Zeng et al., 2008; Josling et al., 2012). They are also found to interact with acetylated MyoD (Ploesskaya et al. 2001).

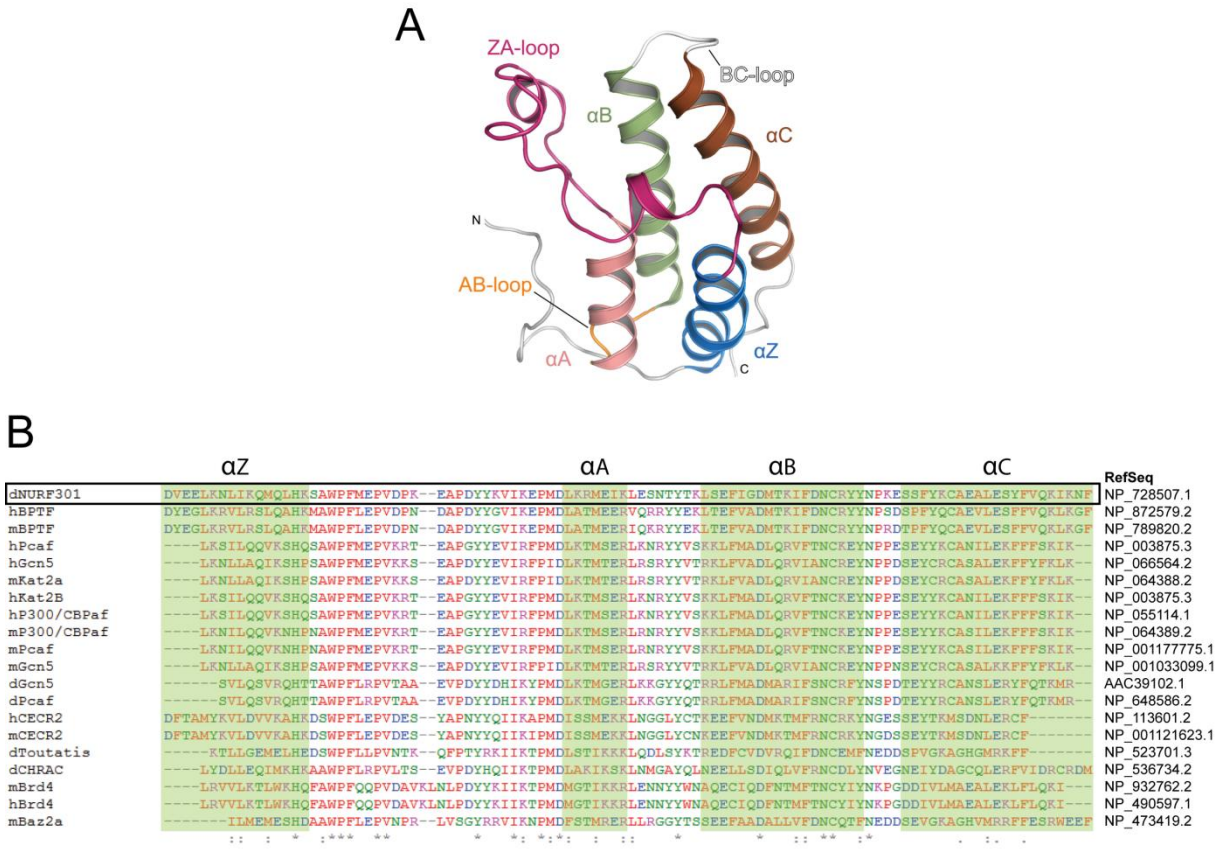


Figure 1-9. General structure of the bromodomain and protein sequence alignment of the NURF301 bromodomain. (A) Bromodomains generally share a conserved four α helix bundle (αZ, αA, αB and αC) linked by variable loop regions (BC and ZA loops). Structure is from Filippakopoulos et al. (2012), (B) the top 20-25 proteins showing sequence similarity to the NURF301 bromodomain were identified by BLAST search. The conserved four α helices (αZ, αA, αB and αC) of the bromodomains are highlighted in green. Reference sequence numbers are indicated.

Many other bromodomains are components of chromatin remodelling complexes. For example, Swi2/Snf2, the catalytic subunit of the SWI/SNF remodelling complex, contains a bromodomain that promotes its binding to chromatin containing H3K14ac (Hassan et al., 2002; Chatterjee et al., 2011). Deletion of the Swi2/Snf2 bromodomain decreases binding to acetylated histone and chromatin remodelling activity, indicating that the bromodomain is critical for recruitment of the complex as well as its function (Awad and Hassan, 2008). Moreover, NURF301, a component of the NURF complex, contains a bromodomain at the C-terminal region and binds to histone H4K16ac (Kwon et al., 2009). As expected, the

bromodomain of NURF301 also shows high sequence similarity with proteins known to recognize acetylated histones with conserved four alpha helix bundles (helices $\alpha_{Z,}$ α_A, α_B and α_C) in our protein sequence alignment result performed as described above (Figure 1-9 B).

The other bromodomains are the BET (Bromodomain Extra Terminal) family, which generally contains two bromodomains at the N-terminal end and extra terminal protein interaction domain at the C-terminal end. In the mammalian BET family, there are four proteins including Brd2, Brd3, Brd4 and Brdt, and they recognize acetylated histones, for example Brd 4 binds to di- or tetra- acetylated histone H4 (H4K5,12ac or H4K5,8,12,16ac) (Dey et al., 2003; Hargreaves et al., 2009). In addition, they also interact with other transcription factors such as P-TEFb (positive transcription elongation factor b), TBP (TATA binding protein) or the hematopoietic transcription factor GATA1 (Yang et al., 2005; Peng et al., 2007; Lamonica et al., 2011).

1.6 Aims and objectives

Epigenetics has been one of the fastest growing areas in science, and the understanding of epigenetic mechanisms related to chromatin remodelling and transcription regulation have made great impacts on developmental studies and the understanding of disease (Bernstein et al., 2007; Zahn and Travis, 2013). NURF is a highly conserved nucleosome remodelling enzyme, the recruitment of which has been speculated to be controlled by histone post-translational modifications. However, the full spectrum of modified histone interactions with "reader" domains (PHD fingers and bromodomain) on the largest NURF subunit NURF301 remains to be established. The main aim of this project is understanding the modified histone binding specificities of these functional modules in NURF301 and how chromatin localization of these specific histone modifications affects NURF recruitment. To do this we used a number of *in vitro* assays as well as *in vivo* immunostaining analysis.

Our first objective was to identify histone modifications recognized by putative reader domains on NURF301. To do this we overexpressed and purified GST-conjugated NURF301 PHD, PHD1, PHD2 and bromodomain respectively. Each purified domain was then applied to modified histone peptide library arrays to perform an unbiased screen to determine the full modified histone binding specificities, Binding was confirmed by peptide pull down assays, structure modelling analysis and Biacore analysis. The second objective was to characterize the *in vivo* localization of these modifications and NURF, and how histone modifying enzymes could affect NURF recruitment. This involved immunostaining of NURF and histone modifications in different cell types such as embryo, neuroblasts, polytene chromosomes, testes and ovaries from wild type and histone modifying enzyme mutant *Drosophila* animals. It also involved ChIP sequencing analysis to determine the genome-wide localization of NURF.

CHAPTER 2. MATERIALS AND METHODS

2.1 Medium

2.1.1 Bacterial medium

2.1.1.1 LB liquid and plate medium

For bacteria cell culture, we used LB medium (10 g/L bacto tryptone (BD), 5 g/L bacto yeast extract (BD), 10 g/L NaCl). 1.8% bacto agar (BD) was added for plate medium if required and poured into petri dishes. The concentration of antibiotics for selection of plasmid DNA or bacterial strain was: ampicillin (100 μg/ml), chloramphenicol (12.5 μg/ml) and tetracycline (12.5 μg/ml).

2.1.1.2 SOC medium

For the recovery of bacterial cells after transformation, SOC (Super Optimal broth with Catabolite repression) medium (0.5% yeast extract, 2% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgSO_4$, 10 mM $MgCl_2$, 20 mM glucose) was used. A solution containing the first four reagents was sterilized at 121℃ and cooled. Subsequently, $MgCl_2$ and glucose sterilized with a 0.2 μm filter were added to complete the medium.

2.1.1.3 M9 salts

M9 salts were used for washing recombination induced bacteria cells before they were spread on the plate medium. 1x M9 salts was prepared by mixing 6 g Na_2HPO_4, 3 g KH_2PO_4, 1 g NH_4Cl and 0.5 g NaCl in 1 L of distilled water as a total volume. Then the solution was autoclaved at 121℃ for 15 minutes.

2.1.1.4 M63 minimal medium

M63 minimal medium was used for *galK* selection in the recombineering system. 5x M63 salts stock solution was prepared by dissolving 10 g $(NH_4)_2SO_4$, 68 g KH_2PO_4 and 2.5 mg $FeSO_4 \cdot 7H_2O$ in 1 L distilled water and adjusting to pH 7.0 with KOH. The stock solution was autoclaved at 121℃ for 15 minutes before use. Sterile stock solutions of 0.2 mg/ml biotin, 10 mg/ml L-leucine, 1 M of $MgSO_4$, 20% galactose, 20% glycerol, 20% 2-deoxy-galactose and 25 mg/ml chloramphenicol were also prepared and filter sterilized separately. M63 minimal medium plates were prepared by adding 15 g bacto agar in 800 ml of H_2O and autoclaved at 121℃ for 15 minutes. The solution was then cooled to 80℃. 200 ml of autoclaved 5x M63 salts and 1 ml $MgSO_4$ solution were added. The medium was cooled to 50℃. 5 ml of biotin, 4.5 ml of L-leucine and 500 µl of chloramphenicol were mixed with the medium. For the *galK*+ selection 10 ml galactose was added, but for the *galK*- selection 10 ml of 2-deoxy-galactose and 10 ml glycerol were added at this step. Finally, 15 ~ 20 ml of M63 agar medium was poured into petri dishes and plates allowed to cool for 30 minutes.

2.1.1.5 McConkey agar medium

To remove *galK* negative contaminants in *galK* positive selection, a few selected clones were streaked on McConkey agar containing galactose as the sole source of carbon. McConkey agar powder (BD) 40 g was added to 1 L of water and autoclaved at 121℃ for 15 minutes. Before it was poured into plates, galactose and chloramphenicol were added to a final concentration of 1% and 12.5 µg/ml respectively.

2.1.2 Drosophila medium

2.1.2.1 Dextrose and yeast medium

For general *Drosophila* culture, we used dextrose and yeast medium. The medium was prepared by adding 4 L of cold water to stainless steel bucket. Then 550 g dextrose (Cargill) and 500 g yeast (Fermipan) were added and mixed thoroughly. 350 g wheat flour (W & H Marriage & Sons Ltd) was mixed with 2.5 L of water to make a paste in a separate beaker, and then added to the dextrose/yeast mixture. In a second stainless steel bucket, 4.5 L hot water and 80 g agar powder (B.T.P.) were mixed. Both mixtures were autoclaved at 121℃ for 15 minutes. After cooling to 70℃, the agar solution was added to dextrose/yeast/flour mixture. 40 ml propionic acid (Sigma) was added. 25 g methyl 4-hydroxybenzoate (Sigma) dissolved in 250 ml ethanol was added and mixed thoroughly. Medium was poured into vials 10 ml or bottles 25 ml. When the medium had set, the vials and bottles were capped and stored in a 4℃ cold room before use.

2.1.2.2 Apple medium

To collect fresh *Drosophila* eggs for injection experiments, we used apple medium plate which was prepared by adding 250 ml apple juice (Sainsburys) and 250 ml distilled water in 1 L plastic beaker. 10 g bacto agar was then added. The mixture was microwaved at full power for 5 minutes or longer if needed until the agar was melted completely. After it had cooled to 70℃, 4 ml of propionic acid was added to apple agar mixture. 2.5 g methyl 4-hydroxybenzoate dissolved in 25 ml ethanol was added and mixed thoroughly. The mixture was poured into 50 mm small petri dishes and stored at 4℃ after it set.

2.2 General molecular biology

All chemical reagents for buffers and media were molecular biology grade reagents obtained from Sigma-Aldrich.

2.2.1 PCR

2.2.1.1 Reaction conditions

We used Taq DNA Polymerase PCR system (Roche) for short DNA fragment amplification and Expand Long Template PCR System (Roche) to amplify long DNA fragments (5-20 kb) like NURF301 C-terminal deleted regions in our study. For reactions, 1x PCR reaction buffer, 0.5 mM dNTP, 25-50 ng of template DNA, 100 pmol forward/reverse primers and 1.25 units of DNA polymerase were used in 50 µl of reactions in a thin wall 0.2 ml PCR tube (Molecular BioProducts).

2.2.1.2 Thermocycler conditions for PCR

We used standard PCR conditions shown in Table 2-1, but the annealing temperature was altered depending on the primer's Tm value and GC-content. Extension time was varied depending on the length of DNA to be amplified, about 1 minute for each 1 kb of DNA.

Table 2-1. PCR programme for amplification.

Step	Temperature	Time
Initial denaturation	94℃	4 mins
30-35 cycles	94℃	30-50 secs
	50-65℃	30 secs - 1 min
	72℃	2 mins
Final extension	72℃	5 mins
Hold	4℃	∞

2.2.2 DNA purification

2.2.2.1 Agarose gel electrophoresis

We conducted 0.7% or 1.0% agarose gel electrophoresis (the percentage used depended on the size of the PCR fragment) to confirm our PCR reaction. We used 0.7 g or 1.0 g of agarose (Fisher Scientific) in 100 ml of 1x TAE buffer (40 mM tris-acetate, 1 mM EDTA). The agarose mixture was melted in a microwave oven for 50 seconds and poured into the gel tray after adding ethidium bromide (Invitrogen) to a 0.5 μg/ml final concentration. The gel was run in 1x TAE buffer. Each 5 μl PCR sample was mixed with 1 μl of 6x DNA loading buffer (0.25% bromophenol blue, 0.25% xylene, 30% glycerol) and loaded. The gel running condition was 120 Volts for 20 minutes.

2.2.2.2 Gel extraction

After the electrophoresis, the DNA was excised and purified by using a QIAquick Gel Extraction Kit (Qiagen). DNA fragments were excised from the agarose gel and visualized using the long-range UV light of a Compact UV Lamp (UVP) with a clean, sharp scalpel

blade. The gel slice was weighed in a 1.5 ml microcentrifuge tube and the DNA was extracted according to the supplier's protocol. At the final elution step, DNA was eluted in 50 µl of distilled water.

2.2.3 Plasmid DNA preparation

2.2.3.1 Plasmid DNA miniprep

Plasmid minipreps were performed using a QIAprep Spin Miniprep Kit (Qiagen). 1.5 ml of a overnight culture in LB medium (see 2.1.1.1) were pelleted by centrifugation at 21,000 *g* for one minute at room temperature, and plasmid DNA was purified according to the supplier's protocol. 50 µl of buffer EB was added to elute DNA at the last step.

2.2.3.2 Plasmid DNA midiprep

Plasmid midipreps were performed using a NucleoBond Xtra Midi EF kit (Machrey-Nagel). 100 ml of a overnight culture was pelleted at 7,900 *g* for five minutes at 4℃ in a Beckman Coulter Avanti centrifuge (Rotor JLA 10.5, Beckman Coulter) and the plasmid DNA was isolated according to the supplier's protocol. At the last step, DNA was dissolved in 150 µl TE buffer (10 mM tris-Cl at pH 8.0, 0.1 mM EDTA) or distilled water.

2.2.4 Restriction enzyme digestion

Table 2-2 lists restriction enzymes that we used in this study.

Table 2-2. Restriction enzyme list.

Name	Description	Suppiler
BamHI	GST fusion cloning	New England Biolabs (NEB)
EcoRI		
Spe I	CTAP tag check	

For each reaction, 10 units of restriction enzyme, 1 x NEB buffer, 1 µg of PCR fragment and plasmid DNA for GST fusion cloning or 2 µg fly genomic DNA for inverse PCR were used in 50 µl of reaction, and the sample was incubated for two hours at 37℃ in water bath.

2.2.5 DNA ligation

DNA ligation was conducted with 2,000 units of T4 DNA ligase (New England Biolabs), 1x NEB T4 DNA ligase buffer, 1:5 ratio of plasmid and insert DNA in a reaction volume of 20 µl overnight at 16℃.

2.2.6 Bacterial strains and transformation

2.2.6.1 Bacterial strains

Bacterial strains used in our study are listed in Table 2-3 below.

Table 2-3. Bacterial strains.

Name	Genotype	Transformation method used	Description	Supplier
DH5α	*F- φ80lacZΔM15 Δ(lacZYA-argF)U169 deoR recA1 endA1 hsdR17(rk-, mk+) phoA supE44 thi-1 gyrA96 relA1 λ-*	Heat shock	Plasmid DNA amplification and purification	Invitrogen
BL21 CodonPlus RIL	*E. coli B F– ompT hsdS(rB– mB–) dcm+ Tetr gal endA The [argU ileY leuW Camr]*	Heat shock	GST fusion protein expression	Stratagene
DH10	*F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80dlacZΔM15 ΔlacX74 endA1 recA1 deoR Δ(ara,leu)7697 araD139 galU galK nupG rpsL λ-*	Electroporation	Strain delivering BAC plasmid DNA	P[acman] Resources
EPI300	*F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80dlacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ- rpsL (StrR) nupG trfA dhfr*	Electroporation	BAC plasmid DNA constructs amplification	EPICENTRE
SW102	*F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80dlacZ M15 ΔlacX74 deoR recA1 endA1 araD139 Δ(ara, leu) 7649 galU galK rspL nupG [λcI857 (cro-bioA) tet]*	Electroporation	BAC recombineering strain	National Cancer Institute Frederick

2.2.6.2 Preparation of SW102 competent cells

SW102 *E.coli* strain (Warming et al., 2005) was streaked on to a LB agar plate containing tetracycline (12.5 µg/ml) and incubated overnight at 30℃. A single colony was inoculated and cultured as a seed culture in 3 ml of LB medium containing tetracycline (12.5 µg/ml) overnight at 30℃ in a shaking incubator (INFORS HT) at 250 rpm. The culture was transferred into 100 ml LB medium without antibiotics and grown for three hours at 30℃ until the OD_{600} reached 0.6. The cultured cells were cooled on ice slurries for five minutes. The cells were harvested by centrifugation at 3,200 *g* for 10 minutes at 4℃ and as much as possible of the LB medium was removed. 1 ml of sterilized ice-cold MiliQ water was added to the collected cells to wash, and the cells were resuspended by gentle swirling on ice. We added 1 ml more of sterilized ice-cold MiliQ water before the cells were centrifuged at 3,200 *g* for 10 minutes at 4℃. The cells were washed one more time with 1 ml of sterile ice-cold 10% glycerol and pelleted. We removed all of the supernatant and added 125 µl of sterile ice-cold 10% glycerol to resuspend the cells, so that the final volume was equal to 1/800 volume of original cell culture volume.

2.2.6.3 Heat shock transformation

1-50 ng of DNA was incubated with 50 µl of host bacteria competent cells (see Table 2-3) for 30 minutes on ice. A heat-pulse was conducted for each transformation reaction in a 42℃ water bath for 45 seconds. The reaction samples were incubated on ice for two minutes. SOC medium (see 2.1.1.2) was added to each transformation reaction for recovery and incubated for one hour at 37℃ in water bath. Each reaction sample was spread on to LB plates containing ampicillin (100 µg/ml) and incubated overnight at 37℃.

2.2.6.4 Electroporation

1-50 ng of DNA was added to 40 µl host bacteria electrocompetent cells (see Table 2-3) in a microcentrifuge tube and mixed by gentle pipetting. After 30-60 seconds to allow DNA adsorption to the cells, the mixture was placed in a 0.1 cm electroporation cuvette (Bio Rad) chilled on ice. Electroporation was performed using GenePulser Xcell System (Bio Rad) with a pulse of 1.8 kV, capacitance at 25 µF and resistance at 200 Ω. For the cell recovery, 1 ml of SOC medium was added and incubated for one hour at 30℃ in a shaking incubator at 250 rpm. The cells were pelleted in 1.5 ml microcentrifuge tube at 21,000 *g* for 15 seconds, resuspended with 100 µl of LB medium and then spread on the LB medium plate containing antibiotics. The plates were incubated for two days at 30℃.

2.2.7 DNA and Protein quantification

The Quant-iT Broad-Range DNA assay kit (Invitrogen) was used for DNA quantification after DNA extraction or PCR. The Quant-iT Protein assay kit (Invitrogen) was used for protein quantification after protein purification. The samples were mixed with fluorophores in the kit that become fluorescent upon binding to DNA or proteins, and then the fluorescence intensity was quantified by a Qubit Fluorometer (Invitrogen) giving sample concentrations. These assays were performed according to the supplier's protocol.

2.2.8 DNA sequencing

2.2.8.1 Template preparation

We followed the sequencing method supplied from Functional Genomics Laboratory in the University of Birmingham. Template DNA was prepared as genomic DNA, plasmid DNA or PCR product purified by method of choice (see 2.2.3). Different DNA template quantities

were used for cycle sequencing reaction as shown in Table 2-4 below.

Table 2-4. Quantities of template DNA required for cycle sequencing reaction.

Template type	**Quantity**
PCR product:	
100 – 200 bp	1 – 3 ng
200 – 500 bp	3 – 10 ng
500 – 1000 bp	5 – 20 ng
1000 – 2000 bp	10 – 40 ng
>2000 bp	40 – 100 ng
Plasmid Single Stranded	50 – 100 ng
Plasmid Double Stranded	200 – 500 ng
Cosmid BAC	0.5 – 1.0 µg
Bacterial Genomic DNA	2 – 3 µg

2.2.8.2 PCR reaction

We used the Big Dye Cycle Sequencing Kit (Applied Bioscience) for sequencing. The reaction sample was prepared with 8 µl of terminator ready reaction mix previously diluted 1:3 in dilution buffer (200 mM tris-Cl at pH 9.0, 5 mM $MgCl_2$), 3.2 pmol of primer, the required amount of template DNA and the volume adjusted to 20 µl with distilled water. The PCR programme used for Big Dye Cycle Sequencing reaction is described below (Table 2-5).

Table 2-5. Big Dye Cycle Sequencing programme.

Step	Temperature	Time
Initial denaturation	96℃	1 min
25 cycles	96℃	10 secs
	50℃	5 secs
	60℃	4 mins
Hold	4℃	∞

2.2.8.3 Purification of DNA for sequencing

After the PCR reaction, DNA was purified. The reaction sample was entirely transferred into a 1.5 ml microcentrifuge tube. Subsequently, 2 µl of 250 mM EDTA and 64 µl of non-denatured 95% ethanol were added. The sample was vortexed briefly and then left for 15 minutes at room temperature for precipitation. The tube was centrifuged at 17,400 *g* for 20 minutes at 4℃ in a Hawk 15/05 centrifuge (MSE). The orientation was marked on the tube for the next step. Immediately, we carefully aspirated and discarded the supernatant. 250 µl of 70% ethanol was then added and mixed briefly by vortexing. This step was repeated once and the sample was centrifuged in the same orientation at 17,400 *g* for 10 minutes at 4℃. The supernatant was removed carefully and the pellet was air dried for 20 minutes. After drying, the purified DNA was dissolved with 20 µl of Hi Di formamide and denatured in a heating block at 95℃ for four minutes. Finally, DNA samples were sent for sequencing by a high-throughput automated capillary electrophoresis machine (Applied Biosystems) which has fluorescence-based detection.

2.2.9 GST fusion protein overexpression and purification.

2.2.9.1 GST fusion cloning

A DNA fragment encoding the NURF301 PHD domain (Figure 2-1) was PCR amplified using the PHD F and PHD R primers described in Table 2-6 using *Drosophila* embryonic cDNA as a template. A 205 bp PCR product was purified and then digested using the restriction enzymes BamHI and EcoRI. The digested DNA fragment was cloned into BamHI and EcoRI digested pGEX-2T vector (Millipore) to create an in-frame fusion with GST. Three plasmids encoding GST fusions with the NURF301 Bromodomain and, PHD1 and PHD2 fingers that have previously been generated (Wysocka et al. 2006) were also utilized. All GST fusion constructs were amplified in *E.coli* strain DH5α (see Table 2-3) and purified using the plasmid midiprep procedure as described above (see 2.2.3.2). These constructs were confirmed by sequencing (see 2.2.8) using the PHD F and PHD R primers, and this sequencing data was analyzed by BLAST (Basic local alignment search tool, Altschul et al., 1990) searches.

Table 2-6. PCR primers used to amplify and sequence *Nurf301* PHD domain gene.

All primers were obtained from Eurofins MWG Operon.

Primer sequence (5' to 3')	Name	Description
CGCGGATCCATGCTGCAGGAGGGACCTAT	PHD F	Amplify *Nurf301*-PHD forward
CCGGAATTCCCACTACGCCACTTACCTTAT	PHD R	Amplify *Nurf301*-PHD reverse

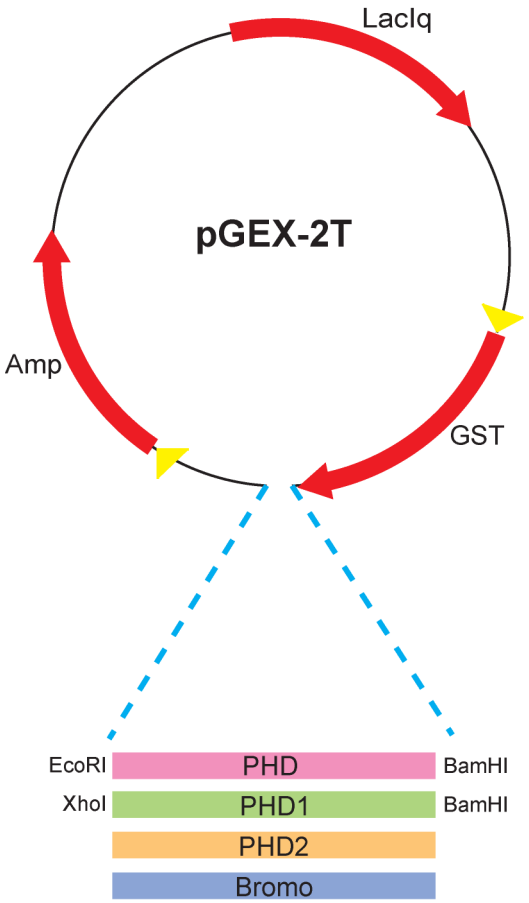


Figure 2-1. GST fusion NURF301 domain cloning structure in the pGEX-2T plasmid vector. GST conjugated NURF301 PHD was generated by cloning, and other domain plasmids encoding GST fusions with the NURF301 PHD1, PHD2 and Bromodomain had previously been generated (Wysocka et al. 2006).

2.2.9.2 Protein overexpression

E.coli strain BL21 CodonPlus RIL (see Table 2-3) was used as the host for GST-fusion protein over-expression. 5 ml overnight cultures of each strain were grown at 30℃ and diluted into 1L LB medium containing ampicillin (25 µg/ml) in a 2 L conical flask. For each construct, 3 L of cells were grown at 32℃ in a shaking incubator at 230 rpm, until the cell culture had reached an OD_{660} of 0.5~0.6. Protein expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG, Promega) to 0.5 mM, and the cells were incubated for two and a half hours at 32℃ in a shaking incubator at 230 rpm.

2.2.9.3 Purification

Cells were harvested by centrifugation at 7,900 *g* for five minutes in an Avanti J20XP centrifuge (JLA10.5 rotor). Harvested cells were washed once with 1x PBS, and the cell pellet was frozen at -80℃ until needed. Cell pellets were resuspended in resuspension buffer (12 mM hepes at pH 7.9, 12.5 mM $MgCl_2$, 2 mM EDTA, 5 mM DTT, 20% glycerol, 0.1% NP-40). Cells were lysed by sonication on ice (30% power, for one minute, 10 repeats) using a Vibra cell VC130 sonicator (Sonics & Materials). The lysate was cleared by centrifugation twice at 27,000 *g* for 30 minutes at 4℃ in an Avanti J20XP centrifuge (JA25.5 rotor). The cleared lysate was applied to 0.5 mL of Glutathione Sepharose 4B beads (Amersham) in a Poly-Prep Chromatography Column (Bio Rad). The column was washed with wash buffer (12 mM hepes pH 7.9, 12.5 mM $MgCl_2$, 2 mM EDTA, 5 mM DTT, 0.15 M KCl, 20% glycerol, and 0.1% NP-40). Bound proteins were then eluted in elution buffer (12 mM hepes pH 7.9, 12.5 mM $MgCl_2$, 2 mM EDTA, 5 mM DTT, 20% glycerol, 0.1% NP-40, 0.15 M KCl, 10 mM glutathione). The purity and yield of protein was analysed by SDS PAGE.

2.2.10 SDS PAGE and western blot

2.2.10.1 10% Acrylamide gel preparation

A Gel was prepared in a Invitrogen gel cassette (Mini, 1.0 mm). The lower resolution gel was prepared by mixing 4.96 ml distilled water, 2.5 ml 4x lower buffer (2 M tris-Cl at pH 8.8), 2.5 ml 40% acrylamide:bis-acrylamide solution (Fisher Scientific), 100 µl freshly made 10% ammonium persulfate and 10 µl TEMED, which was poured into a cassette. The resolving gel was allowed to polymerize for 20 minutes. Then the stacking gel was made with 2.47 ml distilled water, 1 ml 4x upper buffer (1.3 M tris-Cl at pH 6.8), 450 µl 40% acrylamide:bis-acrylamide solution, 68 µl freshly made 10% ammonium persulfate and 4 µl TEMED and

layered on the top of the resolving gel. A 10 well comb was inserted into the stacking gel, and the gel was allowed to polymerize for about 45 minutes.

2.2.10.2 Sample preparation

Protein samples were prepared for polyacrylamide gel electrophoresis by mixing 20 µl of each elution sample and an equal volume of 2x tris-glycine SDS sample buffer (126 mM tris-Cl pH 6.8, 20% glycerol, 4% SDS, 0.005% bromophenol blue, Invitrogen) with 10% β-mercaptoethanol and denatured at 100°C for five minutes. After cooling on ice, 30 µl of each sample was loaded per well. Precision Plus Protein All Blue Standards (Bio Rad) were used to determine the molecular weight.

2.2.10.3 SDS PAGE electrophoresis

The gel was run in a XCell SureLock Mini-Cell Electrophoresis System (Invitrogen) with tris-glycine running buffer (25 mM tris, 192 mM glycine, 0.1% SDS) at a constant voltage of 120 for two hours or until the 10 kDa prestained marker band reached the bottom of the resolving gel. Then the gel was removed from the cassette and stained with Brilliant Blue R Concentrate (Sigma) for two minutes. Subsequently, the gel was put into destaining solution (20% methanol, 10% acetic acid) to reveal protein bands.

2.2.10.4 Western blot

After SDS PAGE, PVDF membrane (Bio Rad) activated in 99% methanol, six fibre pads (Invitrogen) and two whatman papers (pre-cut, Whatman) were prepared and thoroughly soaked in cold tris-glycine transfer buffer (12 mM tris, 96 mM glycine) with 20% methanol for few minutes in a shallow tray. PAGE gel was disassembled from the gel cassette, and the

stacking gel was removed with a clean razor blade. The gel was then transferred and soaked in the tris-glycine transfer buffer with 20% methanol for few minutes. The transfer sandwich was prepared as following orders: three fibre pads, one whatman filter paper, polyacrylamide gel, PVDF membrane, one whatman paper, three fibre pads, stacked from the bottom of XCell II blot module (Invitrogen). The module with the transfer sandwich was placed into the electrode module and transfer was conducted at 30 Volt for two hours. Subsequently, the membrane was removed from the module, and blocked in 20 ml of TBST (50 mM tris-Cl pH 7.4, 150 mM NaCl, 0.1% tween 20) containing 5% dried skimmed milk powder overnight at 4℃ on a shaker. Antibody incubation was carried out with HRP conjugated rabbit anti-GST antibody (1:5000, Novus Biologicals) in 10 ml of TBST buffer containing 5% dried skimmed milk powder for three hours at room temperature. The membrane was then washed with 20 ml TBST buffer four times for 10 minutes at room temperature. Supersignal West Pico Chemiluminescent Substrate (ECL substrate, Thermo Fisher Scientific) was applied to membrane and allowed to react for five minutes at 25℃ on a shaker. Finally, protein bound spots were visualised by direct exposure to Hyperfilm ECL (Amersham).

2.3 Histone modified peptide library array analysis

2.3.1 Histone H3 peptide library array and Histone H3, H4 N-terminal peptide library array analysis

Histone H3 peptide library array (Alta Bioscience) and Histone H3, H4 N-terminal peptide library array (Alta Bioscience) assays were conducted with purified GST-conjugated Bromo, PHD, PHD1 and PHD2 domain proteins respectively. The proteins were dialysed in Slide-A-Lyzer Dialysis Cassettes (10K MWCO, Thermo Fisher Scientific) in binding buffer (50 mM tris-Cl pH 7.5, 150 mM NaCl, 0.1% NP-40) to remove the glutathione. The protein

concentrations were determined before they were applied to the peptide arrays. The array was blocked with blocking buffer (50 mM tris-Cl pH 7.5, 150 mM NaCl, 0.1% NP-40, 20% FBS, 5 mg/ml biotin) containing protease inhibitor cocktail (one Tablet per 50 ml, Roche) for five minutes at room temperature. The array was washed twice with binding buffer containing 20% FBS for five minutes at room temperature. Each domain protein (concentration range: 1 mg/ml to 2 mg/ml) was then introduced to the array, and the binding reaction was performed with binding buffer in a hybridization gasket slide kit overnight at 4℃ on a rotator in a microarray Hybridization chamber (Agilent Technologies). Subsequently the chip was washed with wash buffer (50 mM tris-Cl pH 7.5, 150 mM NaCl, 0.1% NP-40) three times for five minutes. The array was incubated with rabbit anti-GST HRP conjugated antibody (1:5000 dilution in binding buffer containing 20% FBS) for one hour at 25℃. Subsequently, the array was washed with wash buffer three times for five minutes. Supersignal West Pico Chemiluminescent Substrate was applied to the array and allowed to react for five minutes at 25℃ in the chamber on a rotator. Finally, protein bound peptide spots were detected by direct exposure to Hyperfilm ECL film.

2.3.2 MODified histone H3, H4, H2A and H2B peptide array analysis

MODified histone H3, H4, H2A and H2B peptide array (Active Motif) assays were conducted as above (see 2.3.1), except for that the domain protein and array binding reaction was performed in a small shallow tray. Moreover, the array slide was washed in a large volume wash chamber we designed for this experiment. To visualize protein bound peptide spots, the primary antibody reaction was performed with HRP conjugated rabbit anti-GST antibody (1:5000 dilution) followed by a secondary antibody reaction with IRDye 800 CW rabbit anti-HRP antibody diluted in Chemi IR diluent (1 μg/ml, LI-COR Biosciences) for one hour at

room temperature in the dark. The array was washed three times as described above, rinsed with PBS and spun down briefly in a centrifuge to remove residual buffer on the array. The protein bound peptide spots were visualized by direct exposure to 800 nm channel of Odyssey scanner (with Odyssey V3.0 programme, LI-COR Biosciences).

2.4 Peptide pull down assay

2.4.1 Peptide pull down

Peptide pull down was carried out to confirm the peptide array data as described in Wysocka 2006. Standard reactions contained 1 µg of biotinylated peptide (see Table 2-7) and 5 µg purified GST-conjugated protein (Bromo, PHD, PHD1 and PHD2 domain) with 300 µl binding buffer (50 mM tris-Cl at pH 7.5, 150 mM NaCl, 0.1% NP-40). However, the amount of peptide in the reaction was varied from 0.1 µg to 50 µg depending on the binding affinity of the domain protein to the peptide. Binding reactions were incubated overnight at 4℃ on a rotator. Streptavidin Plus Ultralink resin (Thermo Fisher Scientific) was equilibrated by washing with binding buffer four times, and 30 µl of 50% bead slurry was added to each binding reaction. Beads were allowed to bind to biotinylated peptides for two hours at 4℃ on a rotator. Unbound protein was removed by washing with binding buffer four times for five minutes at 4℃. In the last wash, buffer was completely removed, and 60 µl of 1x tris-glycine SDS sample buffer with 10% β-mercaptoethanol was added to the sample for SDS PAGE and western blot (see 2.2.10).

Table 2-7. Histone peptide samples.

Histone peptide samples used in this study are listed below

Name	Peptide sequence (N-terminal to C-terminal)	Suppiler
Histone H3(1-21)	ARTKQTARKSTGGKAPRKQLAGGK(Biotin)	Millipore
Histone H3(1-21) K4me3	ARTK(me3)QTARKSTGGKAPRKQLA	Anaspec
Histone H3(1-21) K4me3	ARTK(me3)QTARKSTGGKAPRKQLAGGK(Biotin)	Millipore
Histone H3(1-21) T3p/K4me3	ART(pho)K(me3)QTARKSTGGKAPRKQLAGGK(Biotin)	Cambridge Peptides
Histone H3(1-21) K4me3/K9ac	ARTK(me3)QTARK(ac)STGGKAPRKQLAGGK(Biotin)	Anaspec
Histone H3(1-21) K4me3/K9ac/S10p	ARTK(me3)QTARK(ac)S(pho)TGGKAPRKQLAGGK(Biotin)	Anaspec
Histone H3(21-44)	ATKAARKSAPATGGVKKPHRYRPGGGK(Biotin)	Anaspec
Histone H3(18-36)	KQLATKAARKSAPATGGVK(Biotin)	Johns Hopkins School of Medicine, USA
Histone H3(18-36) K23me3	KQLATK(me3)AARKSAPATGGVK(Biotin)	Johns Hopkins School of Medicine, USA
Histone H3(14-34) K23me3	(Biotin)KAPRKQLATK(me3)AARKSAPATGG	Alta Bioscience
Histone H3(21-44) K23me3	ATK(me3)AARKSAPATGGVKKPHRYRPGGGK(Biotin)	Anaspec
Histone H3(21-44) K36me2	ATKAARKSAPATGGVK(me2)KPHRYRPGGGK(Biotin)	Anaspec
Histone H3(21-44) K36me3	ATKAARKSAPATGGVK(me3)KPHRYRPGGGK(Biotin)	Anaspec
Histone H3(21-44) K36ac	ATKAARKSAPATGGVK(ac)KPHRYRPGGGK(Biotin)	Anaspec
Histone H4(1-21)	SGRGKGGKGLGKGGAKRHRKVGSGSK(Biotin)	Millipore
Histone H4(1-21) K16ac	SGRGKGGKGLGKGGAK(ac)RHRKVGSGSK(Biotin)	USBiological
Histone H4(1-25) K16ac	SGRGKGGKGLGKGGAK(ac)RHRKVLRDNGSGSK(Biotin)	Anaspec

Histone H4(1-18) K5/8/12/16ac	SGRGK(ac)GGK(ac)GLGK(ac)GGAK(ac)RHGSGSK(Biotin)	Millipore
Histone H4(1-25) K5/8/12/16ac	SGRGK(ac)GGK(ac)GLGK(ac)GGAK(ac)RHRKVLRDNGSGSK(Biotin)	Anaspec

2.4.2 Peptide competition pull down assay

Peptide competition pull down assay was conducted to determine the relative affinities of the NURF301 PHD2 domain to different histone modifications on the peptides. Each sample contained 0.1 µg of biotinylated histone modified peptide (H3(1-21) K4me3, H3(1-21) K4me3/K9ac, H3(1-21) K4me3/K9ac/S10p, see Table 2-7) and 5 µg of purified GST conjugated PHD2 domain protein with 300 µl of binding buffer and was incubated overnight at 4℃ on a rotator. Streptavidin plus ultralink resins were equilibrated by washing with binding buffer four times, and 30 µl of 50% bead slurry was added to each sample. The streptavidin resins and protein-biotinylated peptides were allowed to bind to each other for two hours at 4℃ on a rotator. Unbound proteins were removed by washing with binding buffer four times for five minutes at 4℃. Competitor peptides, H3(1-21) K4me3 peptide un-biotinylated (see Table 2-7), at a 0, 100, 250, 500 times concentration (0 µg, 10 µg, 25 µg, 50 µg) were then added to each sample and incubated for one hour at room temperature. Unbound peptides were washed with binding buffer four times for five minutes at room temperature, followed by SDS PAGE and western blot were performed as above (see 2.2.10).

2.5 Biacore

2.5.1 Sample preparation

GST conjugated PHD2 domain protein, which contains a cleavage site for the prescission protease (Amersham) between GST and the PHD2 domain, was dialysed in protease digestion buffer (50 mM tris-Cl pH 7.0, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.01% NP-40) overnight at 4℃. Dialysed protein concentration was measured by Quant-iT™ protein assay (see 2.2.7). For our studies, 0.7-1 mg of protein was digested with prescission protease (2 units per 100 µg protein) for four hours at 5℃. After digestion, the sample was applied to a mini-column (Bio Rad) containing 200 - 250 µl of glutathione sepharose (Amersham) equilibrated with Biacore running buffer (10 mM HEPES pH 7.9, 150 mM NaCl, 0.005% NP-40) to separate the PHD2 domain protein from uncleaved protein, cleaved GST and prescission protease.

2.5.2 Chip preparation and peptide binding

A Biacore 3000 (Amersham) was used to confirm whether binding of PHD2 to H3K4me3 could be affected by a flanking histone post-translational modification like H3T3p. Sensor chip CM5 (Amersham) was used, and the surface of the chip was activated by injection of EDC (0.4 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide, Amersham)/NHS (0.1 M N-hydroxysuccinimide, Amersham) (1:1) mixture at 10 µl/min for six minutes. Next, 10 mg/ml streptavidin in 10 mM sodium acetate (pH 5.0) was injected at a flow-rate of 10 µl/min for seven minutes to immobilize the streptavidin. The remaining activated sites were then blocked by injection of 1 M ethanolamine-HCl pH 8.5 (Amersham) followed by a low pH wash with 0.1 M glycine (pH 2.5). Finally, 0.1 ng/µl of biotinylated H3(1-21), H3(1-21) K4me3 and H3(1-21) T3p/K3me3 peptides were injected to flow cells Fc2, Fc3 and Fc4 on the

streptavidin coated biosensor chip respectively. Fc1 was a control served as reference for F2, F3 and F4. The immobilization response units of H3(1-21), H3(1-21) K4me3 and H3(1-21) T3p/K4me3 peptides were 167, 191 and 173, respectively.

2.5.3 Loading programme

The previously published Kd for the PHD2 domain binding to H3K4me3 is 2.7 µM (Li et al. 2006). We injected a serial dilution of the PHD2 domain protein over all four flow cells that spanned this concentration, from 78 nM (-32x) to 40 µM (16x). A 10 minute-sample injection was performed under the flow rate of 10 µl/min, followed by a 10 minute-wash step was carried out with running buffer to allow determination of the on and off rates.

2.6 Drosophila genetics

The table below shows the *Drosophila* strains that we used for our studies.

Table 2-8. *Drosophila* strains.

The following strains were provided from Bloomington *Drosophila* Stock Centre (BDSC)

Genotype	**Description**
w^{1118}	Control
$y^1\ w^{67}c^{23}$; *In(2LR)Gla, wgGla-1/CyO, P{hsH\T-2}CyO-1*	Delivering heat shock transposase
$y^1\ w^{67}c^{23}$; $P\{wHy\}CG^{16971}DG^{23208}$	Delivering hobo transposon
$y^1\ w^{67}c^{23}$; *In(3L)D,* D^1*/TM3,* Sb^1, Ser^1	Chromosome 3 balancer
y^1 *M{vas-int.Dm}ZH-2A* w^*; *M{3xP3-RFP.attP}ZH-86Fb*	Delivering *attP* docking site/ φC31 integrase expression in germline cell
w^{1118}; $Gcn5^{E333st}$ $P\{FRT(w^{hs})\}2A\ e^1$*/TM3,* *P{ActGFP}JMR2,* Ser^1	*Gcn5* mutant stop codon generated at E333 (E333st)
$y^1\ w^{1118}$; $Gcn5^{Q186st}$ $P\{FRT(w^{hs})\}2A\ e^1$*/TM3,* *P{ActGFP}JMR2,* Ser^1	*Gcn5* mutant stop codon generated at Q186 (Q186st)
w^{1118}; $Gcn5^{C137T}$ $P\{FRT(w^{hs})\}2A\ e^1$*/TM3,* *P{ActGFP}JMR2,* Ser^1	*Gcn5* mutant delivering substitution of cysteine to trosine (C137T)
$y^1\ w^{1118}$; $Gcn5^{\Delta T280\text{-}F285}$ $P\{FRT(w^{hs})\}2A\ e^1$*/TM3,* *P{ActGFP}JMR2,* Ser^1	*Gcn5* small deletion mutant from tyrosine 280 to phenylalanine 285 (ΔT280-F285)
$y^1\ w^*$; *Df(3L)sex204/TM6C,* $Sb^1\ Tb^1$	Large gene deletion with *sex204 gene (contains Gcn5 exon)* gene deficiency
$y^1\ sc^*\ v^1$; *P{TRiP.HMS01468}attP2*	Haspin RNAi
UAS-Dcr 2; hml-Gal 4, UAS-GFP	Dicer 2 expression in blood cells for RNA interference

2.6.1 Transposon-induced NURF301 mutant fly generation

To generate *Nurf301* C-terminal excision lines, we used the P{wHy} transposon system (Mohr and Gelbart, 2002; Huet et al., 2002). As the first step, heat shock transposase *y^1 $w^{67}c^{23}$; In(2LR)Gla, wgGla-1/CyO, P{hsH\T-2}CyO-1* male flies were crossed with *y^1 $w^{67}c^{23}$; P{wHy}$CG^{16971}DG^{23208}$* virgin females that contain a hobo transposon 15 kb downstream of *Nurf301* (Figure 2-2). Transposon excision was induced by heat shocking the flies three times every two days for 30 minutes at 38℃ in a water bath. When the F_2 progeny emerged, white eye deletion males were selected. These potentially contain an excision that removed parts of *Nurf301*. Individual males were crossed with virgin *y^1 $w^{67}c^{23}$; In(3L)D, D^1/TM3, Sb^1, Ser^1* females delivering a TM3 balancer to stabilize the excision events on chromosome 3. Finally, we made F_3 brother-sister crosses to stabilize the mutant lines.

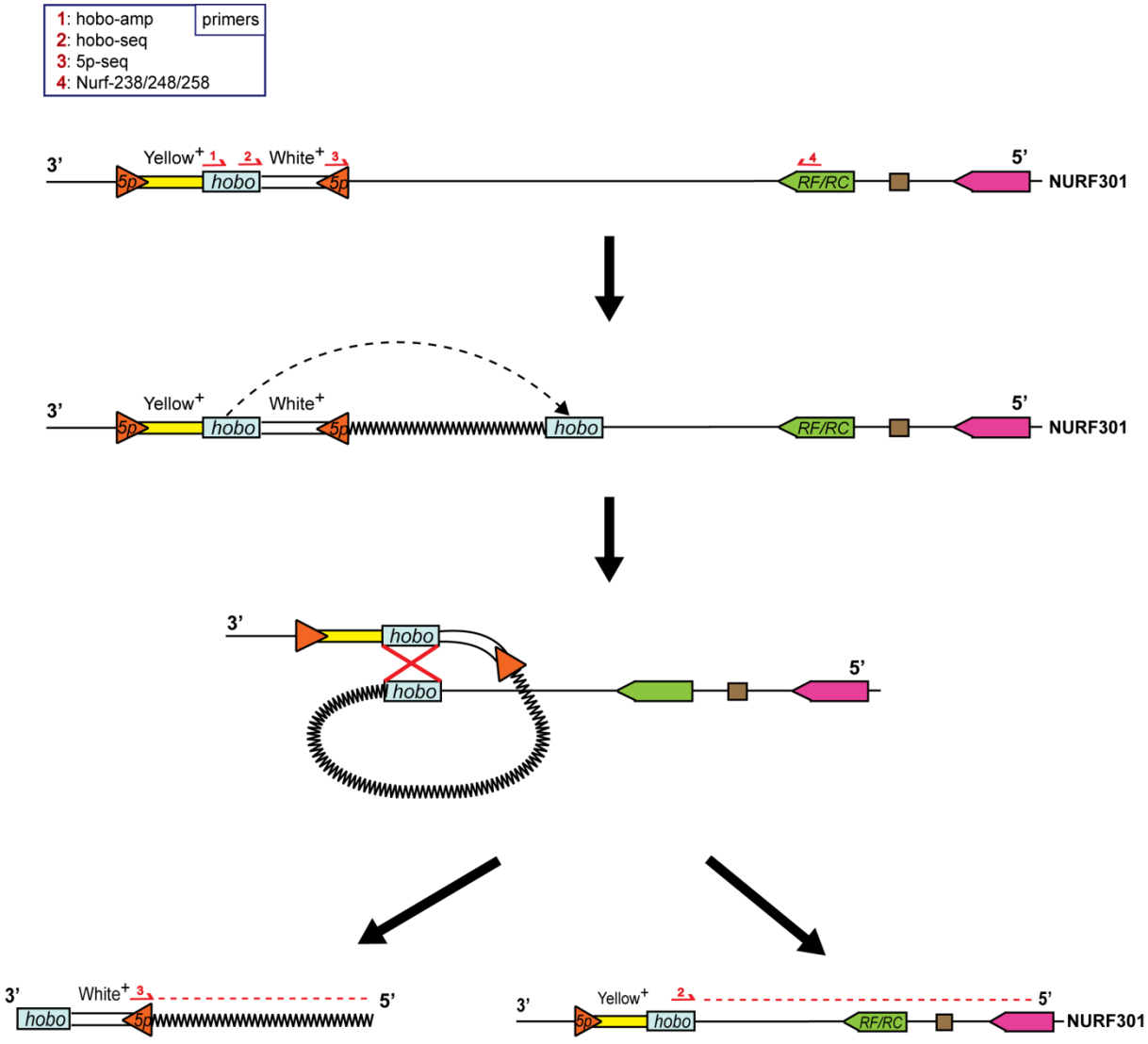


Figure 2-2. The p{wHy} transposon system to generate NURF301 c-terminal excised fly lines. Transposase was activated by heat shock, so transposon excision was induced. Excised regions were then confirmed by sequencing.

2.6.2 *Drosophila* genomic DNA preparation

Twenty frozen adult flies were put into a 1.5 ml eppendorf tube on dry ice, and 500 µl of HB buffer (7 M urea, 2% SDS, 50 mM tris-Cl pH 7.5, 10 mM EDTA, 0.35 mM NaCl) was added. Flies were immediately homogenized using a pellet pestle, and then 500 µl phenol/chloroform was added and mixed by gentle inversion. The tube was placed on a rotator for 30 minutes at room temperature. The reaction sample was centrifuged at 21,000 *g* for 10 minutes at room

temperature. The upper aqueous phase was transferred to a clean 1.5 ml tube. Subsequently, 1 ml 100% ethanol was added and centrifuged at 21,000 *g* for one minute at room temperature. The supernatant was removed, and pellet was washed with 1 ml of 70% ethanol. The tube was centrifuged again at 21,000 *g* for one minute at room temperature. The supernatant was removed and the pellet was briefly air dried. The dried genomic DNA was resuspended with 500 µl of TE buffer (10 mM tris-Cl pH 8.0, 1 mM EDTA). Subsequently, 50 µl 3 M sodium acetate solution and 1 ml 100% ethanol were added to the sample, and the DNA was precipitated overnight at -80°C. The sample was centrifuged at 21,000 *g* for 10 minutes at room temperature. The supernatant was removed and 1 ml of 70% ethanol was added for washing. The sample was centrifuged at 21,000 *g* for one minute at room temperature. The supernatant was eliminated and the pellet was air dried. Finally, the genomic DNA was resuspended with 50 µl TE buffer and kept at -20°C until needed. The *Nurf301* C-terminal excision regions were analysed by sequencing (see 2.2.8) with primers (Figure 2-2, Table 2-9) and BLAST (Altschul et al., 1990).

Table 2-9. PCR primers used to amplify and sequence *Nurf301* C-terminal deleted regions.

All primers were obtained from Eurofins MWG Operon.

Primer sequence (5' to 3')	**Name**	**Description**
GCGACGCAAAACACCGTATTGATT	hob-amp	amplify from hobo element
CCACTCGACTCACACCCTACA	hob-seq	sequence out from hobo element
ACCTTTCCTCTCAACAAGCAA	5P-seq	sequence out from 5P-element
		amplify from *Nurf*-RF/RC region
GAACAACTTTGCTCCGATTAGCTT	Nurf-238	[238439-238416]
GAGCCCGTTGTCGCACATCGAATC	Nurf-248	[248514-248491]
CATCGTCCGAGAAATCGCCCAACT	Nurf-258	[258558-258535]

2.7 *galK* recombineering and gap repair system

2.7.1 Transferring BAC plasmid DNA into SW102

2.7.1.1 BAC plasmids

We ordered CH321(85D11) and CH322(169D05) BAC plasmid DNA for recombination (P*[acman]* Resources). The clones deliver *Nurf301* genomic DNA (*E(bx)*) with some flanking DNA on both 3' and 5' sides in the BAC vectors as shown in figure 2-3. To ensure that the necessary flanking regulatory elements were present, both long and short *Nurf301* genomic constructs were tested.

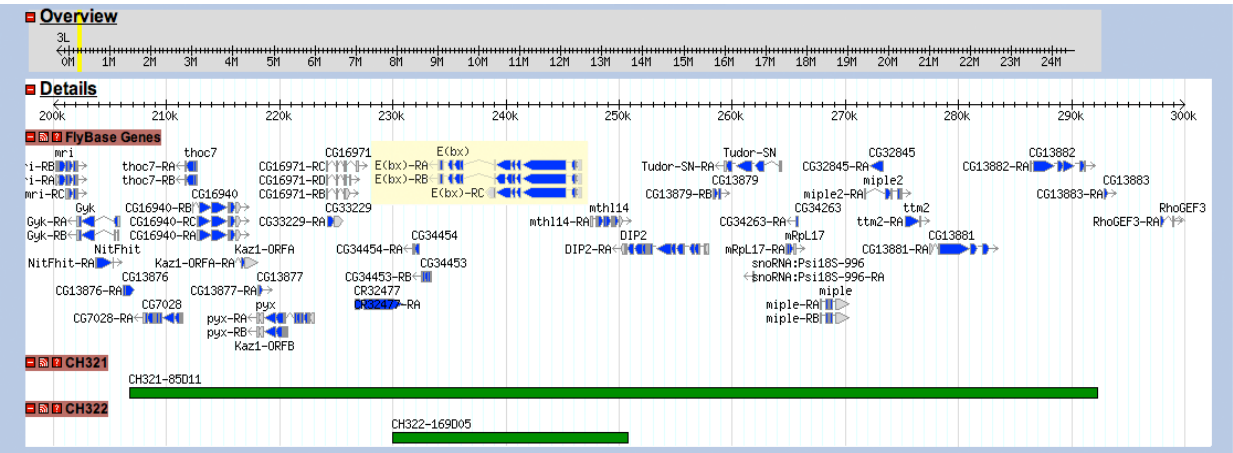


Figure 2-3. Map of CH321(85D11) and CH322(169D05) BAC plasmid DNA. CH321 and CH322 BAC plasmid DNA clones contain *Nurf301* genomic DNA (*E(bx)*) with some flanking DNA on both 3' and 5' sides in the BAC vectors as indicated.

2.7.1.2 BAC plasmid DNA miniprep

CH321 and CH322 BAC clones in *E.coli* strain DH10 were streaked onto LB agar plate containing chloramphenicol (25 µg/ml) respectively and incubated for two days at 30℃. A single colony was picked from each plate and cells were grown in 7.5 ml of LB medium with chloramphenicol (25 µg/ml) in a shaking incubator at 250 rpm for two days at 30℃. Then the cells were pelleted at 4,500 *g* for five minutes at 4℃, and the remaining medium was

removed. The BAC plasmid DNA was prepared by plasmid miniprep (see 2.2.3.1). Air dried CH321 and CH322 BAC DNA pellets were resuspended with 40 µl double distilled water. We stored the DNA samples at 4℃ until needed.

2.7.1.3 Transfer of BAC DNA into the recombinogenic *E.coli* strain SW102

Isolated BAC DNA was then transferred into the SW102 strain which contains the lambda prophage based recombineering system. SW102 competent cells were prepared (see 2.2.6.2), and BAC DNA constructs were delivered into the host strain by performing electroporation as described above (see 2.2.6.4). Cells were plated on LB agar plates containing tetracycline (12.5 µg/ml) and chloramphenicol (12.5 µg/ml) as selective markers.

Table 2-10. Primers for *galK*/CTAP containing *Nurf301* 50 bp/500 bp homology arms.

All primers were supplied from Eurofins MWG Operon.

Primer sequence (5' to 3')	**Number/Name**		**Description**
TTGTGCAAAAAATTAAAAATTTTCGCGAAAATGTTTTTGACCAAAGAACACCTGTTGACAATTAATCATCGGCA	**1**	*galK*-full-5p	*galK*/CTAP recombineering DNA templates of *Nurf301*variant A/B
CTTTATAAACTATAATCATTTGTTCTAAAGTTTTCAAGGGCTATTATTTCTCAGCACTGTCCTGCTCCTT	**2**	*galK*-full-3p	
TTGTGCAAAAAATTAAAAATTTTCGCGAAAATGTTTTTGACCAAAGAACAGAGCAGAAGCTTATCTCCGAG	**3**	GSTAP-full-5p	
CTTTATAAACTATAATCATTTGTTCTAAAGTTTTCAAGGGCTATTATTTCCTATTCAGTGACATGAAAGT	**4**	GSTAP-full-3p	
GTGAATCAAATATCCAATTCGTTC	**5**	Lomgarm 5-3	
GTTCTTTGGTCAAAAACATTTTCG	**6**	Longarm 5-5	
TAGCCCTTGAAAACTTTAGAACAA	**7**	Longarm 3-3	
AGTCCAATGAGCAGTAAACAAACC	**8**	Longarm 3-5	
GTTCACCAATCGGCTGCGCATCAGCAAGTGAGTAGTTTTCCTGTTGACAATTAATCATCGGCA	**1'**	*galK*-NurfC-5p	*galK*/CTAP recombineering DNA templates of *Nurf301*variant C
CTGTTCCTATTCCTCATTTATACTTTTACATTATAATTATAATCAGCACTGTCCTGCTCCTT	**2'**	*galK*-NurfC-3p	
GTTCACCAATCGGCTGCGCATCAGCAAGTGAGTAGTTTTGAGCAGAAGCTTATCTCCGAG	**3'**	GSTAP-NurfC-5p	
CTGTTCCTATTCCTCATTTATACTTTTACATTATAATTATAACTATTCAGTGACAGTGAAAGT	**4'**	GSTAP-NurfC-3p	
TCCCACCACAAAGTCCACAACAAG	**5'**	CLongarm 5-3	
AAAACTACTCACTTGCTGATGC	**6'**	CLongarm 5-5	
ATAATTATAATGTAAAAGTATAAATGAGG	**7'**	CLongarm 3-3	
GCTTAATGTGCTTCCTACTCACT	**8'**	CLongarm 3-5	

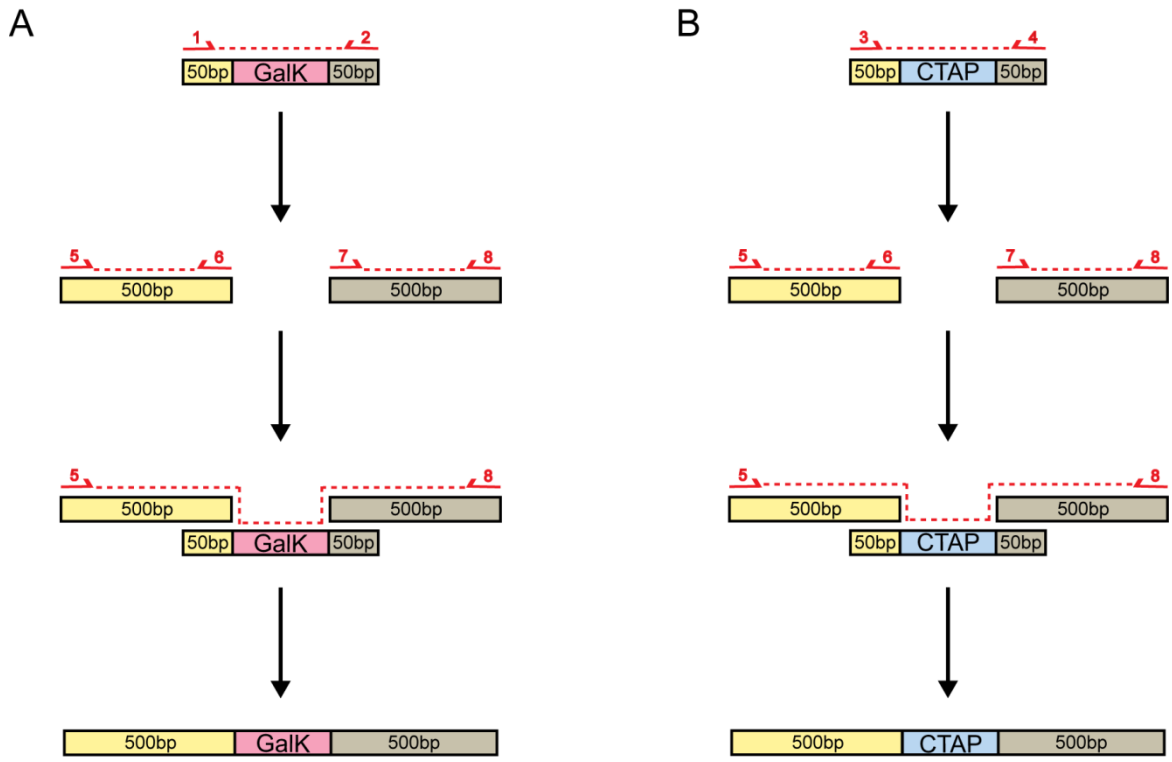


Figure 2-4. Generation of DNA templates for *galK* recombineering. (A) The *galK* +/- selection templates and (B) C-terminally TAP tagging of NURF301 templates used 50 bp or 500 bp homology arms on both 3' and 5', so they can be incorporated into the C-terminal of *Nurf301* genomic region during recombination. The primers utilized are listed and numbered as in Table 2-10 above (both NURF full A/B and NURF C isoforms).

2.7.2 Preparation of the recombination DNA templates

2.7.2.1 *galK* positive and negative selection templates

To generate a galactokinase (*galK*) recombination template DNA with the *Nurf301* full A/B isoform that contains 50 bp length homology arms on both 3' and 5', we conducted PCR (see 2.2.1) with the *galK*-full-5p and *galK*-full-3p primers described in Table 2-10 and p*galK* plasmid (Addgene) as a template. After the reaction, the 1.2 kb of *galK* PCR product delivering 5'/3' 50 bp homology arms was run on a 1.0% agarose gel electrophoresis and gel purified (see 2.2.2). 5' and 3' 500 bp homology arm DNA fragments were also prepared by PCR amplification with Longarm 5-5/Longarm 5-3 for 5' end 500 bp homology arm and Longarm 3-5/Longarm 3-3 for 3' end 500 bp homology arm using CH321 BAC plasmid DNA

as a template (see Figure 2-4, Table 2-10). Subsequently, we performed another PCR to generate a *galK* recombination template DNA with *Nurf301* full A/B isoform 500 bp homology arms on both the 3’ and 5’ ends. Three different templates were required namely: 1) 5’ end 500 bp arm DNA fragment, 2) 3’ end 500 bp arm DNA fragment and 3) the *galK* recombination DNA delivering 50 bp homology arms generated above. The three templates were mixed in equimolar ratios. The final *galK* recombination template DNA with *Nurf301* full A/B isoform 500 bp length homology arms on both the 3’ and 5’ ends was generated by PCR with the Longarm 5-3/3-5 primers as shown in figure 2-4 A. The resultant 2 kb of PCR product was purified by DNA purification (see 2.2.2). We performed the same procedures for generation of *Nurf301 C* isoform *galK* recombination templates delivering 50 bp or 500 bp homology arms on both 3' and 5' ends (see Table 2-10).

2.7.2.2 Templates for C-terminally TAP-tagged NURF301

To generate templates for C-terminally TAP tagging recombination of *Nurf301* full A/B isoform with 50 bp length homology arms on both 3’ and 5’ ends, we conducted PCR with the GSTAP-full-5p and GSTAP-full-3p primers described in Table 2-10. The pCeMM CTAP plasmid (Research Centre for Molecular Medicine of the Austrian Academy of Sciences) was used as a template for this reaction. The CTAP PCR product delivering 5'/3' 50 bp arms (673 bp) was run on a 1.0% agarose gel and gel purified (see 2.2.2). Next, we performed another PCR to make CTAP recombination template DNA of *Nurf301* full A/B isoform with 500 bp length homology arms on both 3’ and 5’ ends. The same 5’ and 3’ 500 bp arm DNA fragments prepared above (see 2.7.2.1) were used to this reaction. Three different templates were required namely: 1) 5’ 500 bp arm DNA fragment, 2) 3’ 500 bp arm DNA fragment and 3)' the CTAP recombination DNA delivering 50 bp homology arms generated above. The three

templates were then mixed in equimolar ratios, and a PCR reaction was carried out with the Longarm 5-3/3-5 primers as shown in figure 2-4 B. The final PCR product, approximate size 1.5 kb, was purified (See 2.2.2). We performed the same procedures for generation of *Nurf301 C* isoform CTAP recombination templates delivering 50 bp or 500 bp homology arms on both 3' and 5' ends.

2.7.3 BAC recombineering using *galK* selection

2.7.3.1 Induction of recombineering function in SW102

SW102 single colonies delivering the CH321/CH322 BAC construct were selected from the plates respectively (see 2.7.1.3) and cultured in 3 ml LB containing chloramphenicol (12.5 µg/ml) for two days at 30℃. The culture was transferred into 100 ml fresh LB without antibiotics and grown for three hours at 250 rpm at 30℃ in a shaking incubator until the OD_{600} reached 0.6 - 0.7. 50 ml of culture was transferred into a sterile falcon tube and reserved at 32℃ as an uninduced control. The rest of the culture (50 ml) was incubated for 15 minutes exactly at 42℃ in a shaking incubator to induce the expression of the recombineering function. Subsequently, we prepared uninduced and induced SW102 electrocompetent cells as described in session 2.2.6.2.

2.7.3.2 Introducing the *galK* containing cassette

Uninduced and induced competent cells were transformed with 300 ng of *galK* template DNA (*Nurf301* full A/B and C isoforms) with either 50 bp or 500 bp arms by electroporation (see 2.2.6.4). After recovery in SOC for one hour at 37℃, 1 ml of cells was transferred into a 1.5 ml eppendorf tube and centrifuged at 21,000 *g* for 15 seconds at room temperature. All the medium was removed, and cells were resuspended with 1 ml M9 salts (see 2.1.1.3). The

sample was then pelleted and the supernatant was discarded followed by another two washes with 1 ml M9 salts. This M9 salt wash was necessary to remove any rich medium from the bacterial culture. Finally, 100 µl of the uninduced control and 100 µl of the induced sample with serial dilutions (1/1, 1/10 and 1/100) were plated on *galK*+ selection M63 minimal medium (see 2.1.1.4). The plates were incubated for 2-4 days at 30℃. Subsequently, the eight colonies of *galK*+ from each plate were selected and streaked on McConkey agar plate (see 2.1.1.5), containing galactose as the sole source of carbon, in order to select against *galK*- contaminating "hitchhikers".

2.7.3.3 Removal of the *galK* containing cassette and TAP tagging

One of the *galK*+ SW102 red colonies from the McConkey agar plate, containing *galK* tagged CH321/CH322 BAC construct, was selected for another round of recombineering to replace the *galK* cassette with CTAP at either the C-terminus of full length *Nurf301* or at the C-terminus of the *Nurf301 C* variant. The colony cell was cultured, and uninduced/induced electrocompetent cells were prepared as described in section 2.2.6.2. This time, uninduced and induced competent cells were transformed with 300 ng of CTAP template DNA (*Nurf301* full A/B and C isoforms) with either 50 bp or 500 bp arms by electroporation (see 2.2.6.4). Finally, 100 µl of the uninduced control and 100 µl of the induced sample with serial dilutions (1/1, 1/10 and 1/100) were plated on *galK*- selection M63-DOG minimal medium (see 2.1.1.4). The plates were incubated for 2-4 days at 30℃. Twelve DOG-resistant colonies were analysed by Spe I restriction enzyme digestion and also by PCR to verify the presence of the CTAP tag with *Nurf301* genes.

2.7.4 C-terminally TAP tagged BAC DNA amplification

Each of the CTAP-tagged CH321 full length NURF301, CH322 full length NURF301, CH321 NURF301-C variant and CH322 NURF301-C variant DNA in host SW102 was purified by plasmid DNA miniprep (see 2.2.3.1). The DNA constructs were then transformed into the amplification host *E.coli* strain EPI300 (Epicentre) by electroporation (see 2.2.6.4). Cells were plated on LB agar plates medium containing chloramphenicol (12.5 µg/ml) as a selective marker. Single colonies were selected from each plate, and seed cell cultures were performed in 3 ml of LB medium containing chloramphenicol (12.5 µg/ml) overnight at 30 ℃ in a shaking incubator. Cells were then transferred to 200 ml of LB medium containing chloramphenicol (12.5 µg/ml) and incubated at 30℃ until the cell culture reached an OD_{600} of 0.4~0.6. The DNA was amplified by addition of copycontrol induction solution (Epicentre) to a 1x final concentration. The cells were incubated for two hours at 30℃ in a shaking incubator. Finally, all CTAP-tagged NURF301 BAC DNA constructs were purified by Nucleobond Xtra midi Endotoxin Free kit (see 2.2.3.2). At the final step, DNA was dissolved in 100 µl double distilled water by gentle pipetting. The purified DNA samples were visualized by 0.7% agarose gel electrophoresis and quantified using a UV spectrometer (Ultrospec 2100 pro, Amersham).

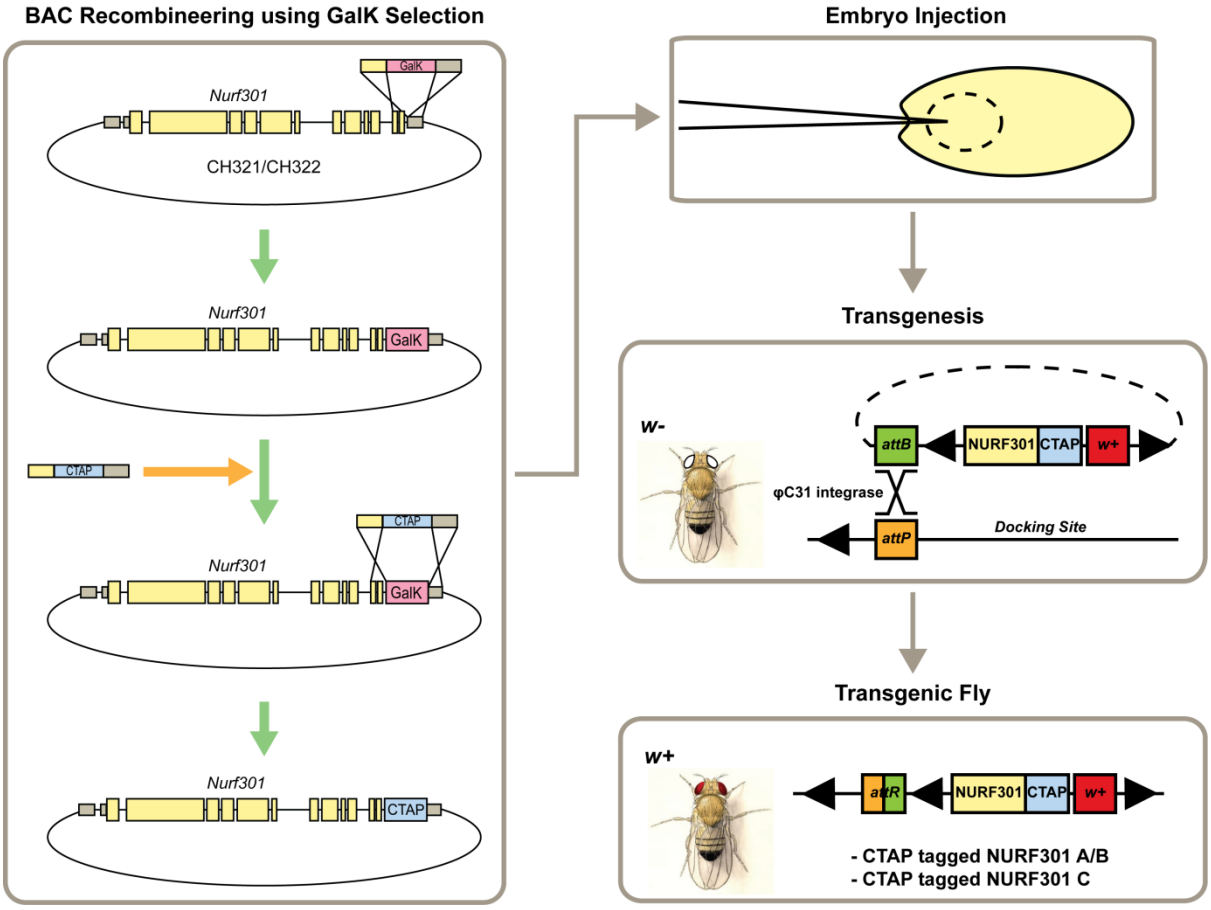


Figure 2-5. Overview of the generation of CTAP tagged *Nurf301* constructs by using BAC recombineering system and transgenic fly expressing CTAP tagged NURF301. CH321 and CH322 BAC plasmids delivering *Nurf301* genomic DNA (*E(bx)*) were used as templates for recombination, and y^1 *M{vas-int.Dm}ZH-2A* w^**; M{3xP3-RFP.attP}ZH-86Fb Drosophila* strain was prepared for embryo injection.

2.7.5 Injection of CTAP tagged NURF301 BAC DNA constructs

To generate CTAP-tagged NURF301 transgenic lines, the φC31 integrase system was utilized (Venken et al., 2006; Bischof et al., 2007). DNA was injected into embryos of the strain y^1 *M{vas-int.Dm}ZH-2A* w^**; M{3xP3-RFP.attP}ZH-86Fb* (see Table 2-8), which carries an *attP* docking site, at 86F on chromosome 3R and a construct that expresses φC31 integrase in the germline cells. To prepare embryos for injection, approximately 500 young flies (2-3 days) were transferred to an egg-laying cage and fed with yeast paste on apple agar collection plates overnight. The next day, embryos were collected every 20-30 minutes, hand dechorionated

and lined up on a glass slide coated with double-sided tape (3M). Embryos were injected with 0.8 µg/µl purified endonuclease free BAC plasmid DNA using an Eppendorf FemtoJet Microinjector and Leitz Labovert FS Inverted Microscope with injection stand and needle holder. Injected embryos were incubated at 18℃ for recovery. Larvae were collected after 2-3 days and transferred to fresh vials with food at 25℃. Viable adults were singly mated to w^{1118} flies of the corresponding sex, and then the progenies were examined for red-eyed transgenic founders (Figure 2-5).

2.8 Immunostaining of *Drosophila* embryos, polytene chromosomes, testes and ovaries

2.8.1 Fixation and permeabilization of *Drosophila* embryos

Drosophila embryos were collected on the apple agar collection plates and rinsed with water onto a nitex mesh sieve to remove residual yeast. Then they were then dechorionated in 5% of sodium hypochlorite solution (Sigma) for three minutes. The embryos were transferred to a 1.5ml eppendorf tube containing 50% of n-heptane (Sigma) and 50% of PEM-formaldehyde solution (0.1 M PIPES pH 6.95, 2 mM EGTA, 1 mM MgSO4, 4% formaldehyde). Embryos were fixed for 25 minutes at room temperature on a rotator. The aqueous (bottom) layer was removed, 1 volume of 99% methanol was added, and the tube was vortexed for one minute to remove the vitelline membrane. The heptane/methanol phases were allowed to separate and the heptane (top layer), all of the interface and the methanol above the devitellinised embryos, were removed. Two washes were performed with 1 ml methanol and then with 1:1 dilution of 99% methanol and PBTw (0.1% Tween 20 in PBS). Embryos were washed with 1ml PBTw three times for five minutes on a rotator, and single or double immunostaining was performed (see 2.8.4 and 2.8.5)

2.8.2 Immunostaining of *Drosophila* polytene chromosomes

2.8.2.1 Culturing of third instar larvae

About fifteen female flies were allowed to lay eggs in a vial containing normal fly medium (see 2.1.2.1) and transferred to a new vial every two days. In order to get optimal polytene chromosome morphology, larvae were raised at 22℃ and uncrowded culturing of larvae was essential. The wandering stage third instar larvae were used for all polytene chromosome squash techniques.

2.8.2.2 Conventional squash

The third instar larvae were washed in water to remove residual culture medium. The salivary glands of larvae were dissected in solution 1 (PBS containing 0.1% Triton X-100 (pH 7.5)). Glands were transferred to a 20 µl drop of solution 2 (PBS containing 3.7% formaldehyde, 1% triton X-100 (pH 7.5)) on a 22 mm x 22 mm Sigmacote-treated coverslip and incubated for two and a half minutes. The solution was replaced with solution 3 (3.7% formaldehyde, 50% acetic acid). The glands were crushed into small pieces with a tungsten needle during three minutes of incubation. Then a poly-L-lysine-coated microscope slide was overlaid to the sample and inverted. The chromosomes were spread with gentle tapping on the coverslip followed by drawing a trace over the coverslip. The slide was put into liquid nitrogen to freeze and the coverslip was removed using a razor blade. The slide was stored in PBS until ready to perform immunostaining.

2.8.2.3 Acid free squash

Wandering third instar larvae were washed in water and used for the preparation of polytene chromosome squashes. The larvae were transferred to Brower's fixation buffer (Blair, 2000) (0.15 M Pipes, 3 mM MgSO4, 1.5 mM EGTA, 1.5% NP40 at pH 6.9) containing 2% formaldehyde, and salivary glands were directly dissected in this solution. The glands were fixed for 2-3 minutes and transferred to PBT (PBS containing 0.1% triton X-100) for another 2-3 minutes. They were moved into 50% glycerol and soaked for five minutes. The glands were transferred to 13 µl 50% glycerol on a 22 mm x 22 mm Sigmacote-treated coverslip and broken down to small pieces using a tungsten needle. A poly-L-lysine-coated microscope slide was gently placed onto the coverslip. Then the microscope slide was inverted, lightly tapped on the coverslip followed by drawing a trace over the coverslip and squashed to spread

the chromosomes. The slide was directly submerged in liquid nitrogen for a few seconds to freeze, and the coverslip was removed using a razor blade. The frozen slide was quickly submerged in PBS until needed for immunostaining.

2.8.3 *Drosophila* testis/ovary immunostaining

Young healthy two to three day old adult flies were prepared to isolate male testes or female ovaries. They were dissected in 0.7% NaCl, and a small number of testes or ovaries were pooled in one tube. The sample was fixed in PBT containing 4% formaldehyde for 30 minutes at room temperature on a rotator. The fixed sample was washed with PBT three times for 10 minutes on a rotator. To increase the antibody penetration of tissue, the tissues were treated with higher concentration of detergent (0.3% Triton X-100 in PBS) for one hour at room temperature on rotator. Primary antibody binding reaction was performed in blocking solution (PBT containing 10% FBS) at 4℃ overnight on a rotator. The sample was washed with PBT three times for 10 minutes. Secondary antibody binding was conducted using Cy3 conjugated donkey anti-rabbit IgG (Jackson ImmunoResearch Laboratories) at a 1:400 dilution in PBT for two hours at room temperature on a rotator in the dark. The sample was washed with PBT three times for 10 minutes and mounted in Vectashield mounting medium with DAPI (Vector Laboratories). Microscope slides were visualized by confocal microscopy (Zeiss Axiovert 100M Confocal Laser Scanning Microscope (LSM), Zeiss).

Table 2-11. Antibodies used for either single or double immunostaining.

Primary antibodies used			
Antibody	**Host**	**Source**	**Dilution**
Anti- Phospho-Histone H3 (Thr3), ChIPAb+	Rabbit	Millipore	1:1500
Anti-Histone H3 (phospho T3, tri metyl K4)	Rabbit	Abcam	1:500
Anti-Histone H3 (tri methyl K4), ChIP Grade	Rabbit	Abcam	1:2000
Anti- Acetyl-Histone H3 (Lys9)	Rabbit	Prof. Bryan Turner (University of Birmingham, UK)	1:250
Anti- Acetyl-Histone H3 (Lys9), ChIPAb+	Rabbit	Millipore	1:250
Anti-Histone H3 (acetyl K9, phospho S10), ChIP Grade	Rabbit	Abcam	1:250
Anti- Phospho-Histone H3 (Ser10), ChIPAb+	Rabbit	Millipore	1:1000
Anti-Histone H3 (tri methyl K23)	Rabbit	Dr. Sean Taverna (Johns Hopkins School of Medicine, USA)	1:100
Anti-NURF301 (Kwon5)	Rabbit	Dr. So Yeon Kwon (University of Birmingham, UK)	1:25
Anti-Protein G [HRP]	Rabbit	Novus Biologicals	1:500
Secondary antibodies used			
Antibody	**Host**	**Source**	**Dilution**
Anti-rabbit IgG, Cy3-conjugated	Donkey	Jackson ImmunoResearch Laboratories	1:500
AffiniPure Fab Fragment Anti-Rabbit IgG, Cy3-conjugated	Goat	Jackson ImmunoResearch Laboratories	1:400
AffiniPure Fab Fragment Anti-Rabbit IgG	Goat	Jackson ImmunoResearch Laboratories	1:20
AffiniPure Fab Fragment Anti-Rabbit IgG, FITC-conjugated	Goat	Jackson ImmunoResearch Laboratories	1:400

2.8.4 Single immunofluorescence staining

Samples/slides were incubated with the primary antibody (see Table 2-11) in PBTw containing 10% BSA at 4℃ overnight. Samples then were washed with PBTw four times for 5-10 minutes at room temperature. Secondary antibody incubation was performed using Cy3 conjugated donkey anti-rabbit IgG (see Table 2-11) in PBTw for three hours at room temperature in the dark. The samples were washed with PBTw four times for 5-10 minutes at room temperature. The samples were mounted in Vectashield mounting medium with DAPI. Confocal microscopy was performed using a Zeiss Axiovert 100M Confocal Laser Scanning Microscope (LSM).

2.8.5 Double immunofluorescence staining

This protocol was used for immunostaining using two primary antibodies from the same host species. All the primary and secondary antibodies with dilutions used in this study are described in Table 2-11. Stored samples/slides were incubated with the first primary antibody (host rabbit) in PBTw containing 10% BSA overnight at 4℃. The samples were washed with PBTw three times for 10 minutes. Then the first secondary antibody reaction was conducted using Cy3-conjugated AffiniPure Fab Fragment Goat Anti-Rabbit IgG in PBTw for two hours at room temperature in the dark. The samples were washed with PBTw three times for 10 minutes then fixed in 1% formaldehyde for 10 minutes in the dark. Samples were washed three times with PBTw for 10 minutes in the dark. Samples were incubated with unconjugated AffiniPure Fab Fragment Goat Anti-Rabbit IgG in PBTw for two hours at room temperature in the dark. The first rabbit IgG is blocked in this step, so that the second secondary antibody would not be able to bind to it. The samples were washed with PBTw three times for 10 minutes, and the second primary antibody (host rabbit) reaction was performed in PBTw

containing 10% BSA overnight at 4℃. The samples were washed with PBTw three times for 10 minutes on the following day. The second secondary antibody reaction was performed using FITC-conjugated AffiniPure Fab Fragment Goat Anti-Rabbit IgG in PBTw for two hours at room temperature in the dark. Finally, the samples were washed with PBTw three times for 10 minutes, mounted in Vectashield mounting medium with DAPI and examined by confocal microscopy.

2.9 Chromatin immunoprecipitation (ChIP) sequencing

2.9.1 *Drosophila* primary blood cell preparation

Fresh primary blood cells were collected manually by ripping 50 third instar larvae in a tube containing HyQ-CCM3 insect medium (Thermo Fisher Scientific) and protease inhibitors (Roche). The cells were fixed with 1% formaldehyde solution in PBS for 15 minutes at 25℃ and pelleted at 410 *g* for five minutes in a benchtop centrifuge. The supernatant was removed. Then the cells were washed three times with cold PBS containing protease inhibitors and spun down. After the final wash, all buffer was removed, and the tube was kept at -80℃ until required. Fifteen to twenty tubes were prepared in this way for each male and female sample.

2.9.2 Micrococcal nuclease (MNase) digestion

The stored blood samples were thawed on ice and transferred to one tube. The cells were resuspended with buffer A (15 mM tris-Cl pH 7.4, 15 mM NaCl, 60 mM KCl, 0.34 M sucrose mixed with freshly made 1 mM DTT, 25 mM sodium metabisulfite, 0.5 mM spermidine, 0.15 mM spermine). Then the cells were homogenized with pellet pestle. $CaCl_2$ and MNase (Worthington) were added for final concentrations of 1 mM and 100 unit/µl respectively, and the sample was incubated for 12 minutes at 16℃ in water bath. As soon as the reaction was completed, an equal volume of STOP buffer (0.1 M tris-Cl at pH 8.5, 0.1 M NaCl, 50 mM EDTA, 1% SDS) was added to stop the reaction. The sample was centrifuged at 17,400 *g* for three minutes at 4℃. Subsequently, the supernatant was transferred to a new tube and diluted with ChIP dilution buffer (0.01% SDS, 1.1% triton X-100, 1.2 mM EDTA, 16.7 mM tris-Cl pH 8.1, 167 mM NaCl). This process produced mononucleosomal DNA fragments about 146 base pairs in size.

2.9.3 Chromatin immunoprecipitation (ChIP)

Antibody coated beads were prepared the day before ChIP. 25 µl of the Dynabeads® M-280 Sheep anti-Rabbit IgG antibody (Invitrogen) was prepared for each sample in a 1.5 ml LoBind tube (Eppendorf). The beads were washed five times with freshly made PBS containing 5 mg/ml BSA and protease inhibitors. After the final wash, the beads were resuspended with the same buffer, and 4 µg rabbit anti-protein G antibody (Novus Biologicals) was added per 25 µl of beads. The antibody and beads were incubated overnight at 4℃ on a rotator. Antibody coated beads were then washed five times with PBS containing 5 mg/ml BSA and protease inhibitors in a magnetic rack. Pre-cleared beads, that lack antibody, were also simultaneously prepared as above. The MNase digested chromatin prepared as described above (see 2.9.2) was incubated with pre-cleared beads for 15 minutes, and then the supernatant was incubated with antibody coated beads for two and a half hours at room temperature on a rotator. The beads were washed once with Low salt buffer (0.1% SDS, 1% triton X-100, 2 mM EDTA, 20 mM tris-Cl pH 8.1, 150 mM NaCl), once with High salt buffer (0.1% SDS, 1% triton X-100, 2 mM EDTA, 20 mM tris-Cl pH 8.1, 500 mM NaCl), once with LiCl immune complex wash buffer (0.25 M LiCl, 1% IGEPAL CA-630, 1% deoxycholic acid, 1 mM EDTA, 10 mM tris-Cl pH 8.1) and twice with TE buffer for five minutes each at room temperature. The precipitated DNA was eluted by incubation of the beads with Elution buffer (1% SDS, 0.1 M $NaHCO_3$) for 15 minutes at room temperature on a rotator. The residual proteins in the sample were removed by treating with proteinase K (Invitrogen) and NaCl for final concentrations of 100 µg per sample and 80 mM respectively, and the sample was incubated overnight at 65℃ in a heat block. The immunoprecipitated DNA was purified using a volume of Agencourt AMPure XP beads (Beckman Coulter) equal to 1.8x the volume of the sample following the supplier's protocol.

2.9.4 End repair and P1 adaptor/barcode ligation

Our ChIP library DNA preparation was based on the barcoded fragment library preparation method of the SOLiD 4 sequencing system (Applied Biosystems). The ChIP DNA fragments were end repaired using SOLiD™ Fragment Library Construction Kit (Life Technologies) following the supplier's protocol. For the reaction, purified ChIP DNA fragments, 1x End-polishing buffer, 0.4 mM dNTP mix, 5 units End polishing enzyme 1 and 40 units of End polishing enzyme 2 were used in 50 µl reaction in a 1.5 ml LoBind tube, and the sample was incubated for 30 minutes at 25℃. After the reaction, the end repaired DNA was purified using Agencourt AMPure XP beads equal to 1.8x the volume of the sample according to the supplier's protocol. The purified and end-repaired DNA was then ligated with multiplex P1 and P2 adaptors making a reaction with end-repaired DNA, 25 pmol of Multiplex library P1 adaptor, 0.5 µM of Barcode, 1x T4 ligase buffer and 25 units of T4 ligase in 100 µl of ligation reaction. The sample was incubated for 10 minutes at 25℃, and then another DNA purification was performed using Agencourt AMPure XP beads equal to 1.3x the volume of the sample following the supplier's protocol.

2.9.5 Sequencing library preparation with barcodes

The barcoded and purified library DNA was then nick-translated and amplified by PCR reaction using 1 µM of Multiplex library PCR primer-1, 1 µM of Multiplex library PCR primer-2, barcoded library DNA and 200 µl of Platinum PCR amplification mix in a total reaction volume of 250 µl. The reaction mix was split into two PCR tubes, and the PCR was performed with the thermocycler conditions as suggested in the supplier's protocol (see Table 2-12). After PCR amplification, the amplified barcoded library DNA was purified twice using Agencourt AMPure XP beads and visualized quickly using a flash gel (Lonza) in a flash gel dock (Lonza).

Table 2-12. PCR conditions for nick-translation and amplification of the barcoded library DNA

Step	Temperature	Time
Nick translation	72℃	20 mins
Denaturation	95℃	5 mins
15-17 cycles	95℃	15 secs
	62℃	15 secs
	70℃	1 min
Final extension	70℃	5 mins
Hold	4℃	∞

2.9.6 SOLiD sequencing

Clonal amplification of library fragments on the sequencing beads surface was done by emulsion PCR (ePCR) according to the manufacturer's instructions (Applied Biosystems). 500 pM of double stranded library DNA was added to 5.6 ml PCR mix (1x PCR buffer, 3.5 mM of each deoxynucleotide, 25 mM $MgCl_2$, 40 nM ePCR primer 1, 3 µM ePCR primer 2, 3000 U AmpliTaq Gold DNA Polymerase, 1.6 billion SOLiD sequencing P1 DNA beads). The PCR mix was added to SOLiD ePCR Tube containing 9 ml of oil phase while swirling on ULTRA-TURRAX Tube Drive. The emulsion was then transferred to a 96-well PCR plate and run for 30 cycles on the 96-Well Thermal cycler. The emulsion was broken with 2-butanol and washed with 1x Bead wash buffer. Beads were enriched in 60% glycerol for template positive beads selection, and 3' end modified to apply onto sequencing slides. Work Flow Analysis (WFA) was performed before the actual full sequencing run to confirm the beads quality. Finally, a full sequencing run was conducted using cycled ligation sequencing on a SOLiD 4 general analyzer.

CHAPTER 3. IDENTIFICATION OF THE MODIFIED HISTONE BINDING SPECIFICITIES OF THE NURF301 PHD DOMAINS AND BROMODOMAIN *in vitro*

3.1 Introduction

Previous studies showed that *Drosophila* NURF301 can bind to the histone H3K4me3 and histone H4K16ac post-translational modifications (Wysocka et al., 2006, Kwon et al., 2009). Such binding enables functional cooperation between two of the principal mechanisms of encoding epigenetic information in the genome, histone post-translational modifications and ATP-dependent chromatin remodelling. In this way histone modifications could provide a mark to recruit the NURF ATP dependent chromatin remodelling enzyme to specific sites in the genome. Significantly, *Drosophila* NURF301 contains three PHD domains and a bromodomain, all of which have potential to bind additional modified histone marks.

Previous research had shown that *Drosophila* NURF301 and human BPTF bind to the H3K4me3 post-translational modification via the PHD2 domain. NMR and crystal structure analysis of the interaction between the BPTF PHD2 domain and H3K4me3 defined a cage composed of four aromatic amino acids (Y10, Y17, Y23 and W32) that interacts with trimethylated Lys 4 residues. (Li et al., 2006), but which in principle could bind to other trimethylated lysine residues. In the initial reports, the binding specificity of the PHD2 domain to a limited number of trimethylated lysine containing peptides was analysed and showed that the PHD2 interaction was specific to H3K4me3. However, the full-binding specificity had not been determined. We first proposed to characterise the full modified histone-binding specificity of the PHD2 domain to determine if there were additional histone

post-translational modifications that could recruit NURF.

In addition, NURF contains three other domains that have the potential to bind to modified histone residues. These include a single bromodomain and two additional PHD fingers. The bromodomain is an evolutionally conserved motif that has been shown to interact with acetylated lysine residues in N-terminal tails of histones H3 and H4 *in vivo* and *in vitro* (Ornaghi et al. 1999; Jacobson et al. 2000; Ladunder et al. 2003). Previously in our lab, it was also shown that purified NURF301 bromodomain binds to histone H4 peptide acetylated at Lys 16 (Kwon et al. 2009). However, as with the PHD2 domain the full-binding specifity of this domain had not be exhaustively characterised. It was still possible that the bromodomain could interact with other modifications on histone tails.

Drosophila NURF301 contains an additional PHD domain at the C-terminus of the protein that has been designated PHD1 (Wysocka et al., 2006). The PHD1 domain is not present in the annotated BPTF protein isoforms, but is only identified in annotated *Drosophila* NURF301 protein isoforms. The PHD1 domain showed amino acid sequence similarity with PHD2 and also contains the highly conserved aromatic amino acids (Y10, Y17, Y23 and W32), which have been proposed to form a cage that potentially binds to metylated lysine residues as described in figure 1-8. However, the modified histone binding specificity of the PHD1 domain has not been determined.

Finally, both *Drosophila* NURF301 and human BPTF contain an additional PHD domain at the N-terminal region of the protein, designated PHD (Wysocka et al., 2006). Unlike the PHD1 and PHD2 domains, this PHD domain does not contain the aromatic cage required for

binding to trimethylated histone residues. Rather the sequence of the PHD domain resembles more closely that of the PHD domain of BHC80 which has been shown to bind to unmodified histone tails (Lan et al. 2007).

To elucidate the complete modified histone binding specificities of these domains of NURF301, we separately purified the PHD, PHD1, PHD2 and Bromodomain of NURF301 as GST fusion proteins. These domains were then used to screen peptide arrays that contained spots decorated with the principal reported histone modifications. These peptide arrays deliver histone modifications both singly and in conjunction with other flanking histone modifications. By using peptide arrays containing combinations of histone tail modifications, we were able to investigate the effect of combinations of histone modifications on the binding ability of the domain.

An emerging theme in research on the binding of histone post-translational modifications has been the concept that combinations of histone modifications can provide "rheostats" that can modify binding to a core histone post-translational modification. In the case of the PHD2-H3K4me3 interaction for example, flanking modification of H3K4me3-modified histone tails could mask PHD2 binding, and act as a rheostat to control NURF recruitment without the loss of the H3K4me3 mark. Thus, four states could be envisaged, in which i) the unmodified H3 tail is not bound by PHD2 (off), ii) the H3K4me3 modified tail is bound by PHD2 (on), iii) further flanking modification of the H3K4me3 tail blocks binding (off) or iv) enhances binding.

Binding-specificities determined using peptide arrays were then confirmed by peptide pull

down, peptide competition pull down and Biacore assays. The goal of this work was to define specific modifications and combinations of modifications that would allow the NURF complex to be recruited to sites of action in the genome.

3.2 Results

3.2.1 Expression and purification of GST fusions with NURF301 domains

All GST fusion proteins were expressed at high levels in the *E.coli* BL21 strain and purified using GST-sepharose columns. As shown in the figure 3-1, purified proteins migrated at the predicted molecular masses of the GST fusion: bromodomain (37 kDa), PHD (33 kDa), PHD1 (31 kDa) and PHD2 (31 kDa). Eluted proteins were of high yield and high purity with one single band or minimal background protein bands. In the case of the PHD domains, additional bands of approximated 26 kDa were detected below the expected protein bands. These corresponded to the size of GST protein alone and probably came from either protease cleavage of the link between the PHD domain and GST during lysis or ribosome stalling at this location during protein synthesis. Nevertheless, the purified GST domain protein does not pose a problem to the peptide array screening as we have shown that GST domain alone does not react with the peptide arrays (Figure 3-3). Of all the domains purified, the GST-Bromodomain fusion exhibited the best yield and purity. After dialysis to remove the glutathione from the samples, the purified GST fusion domain proteins were applied to modified histone peptide library arrays.

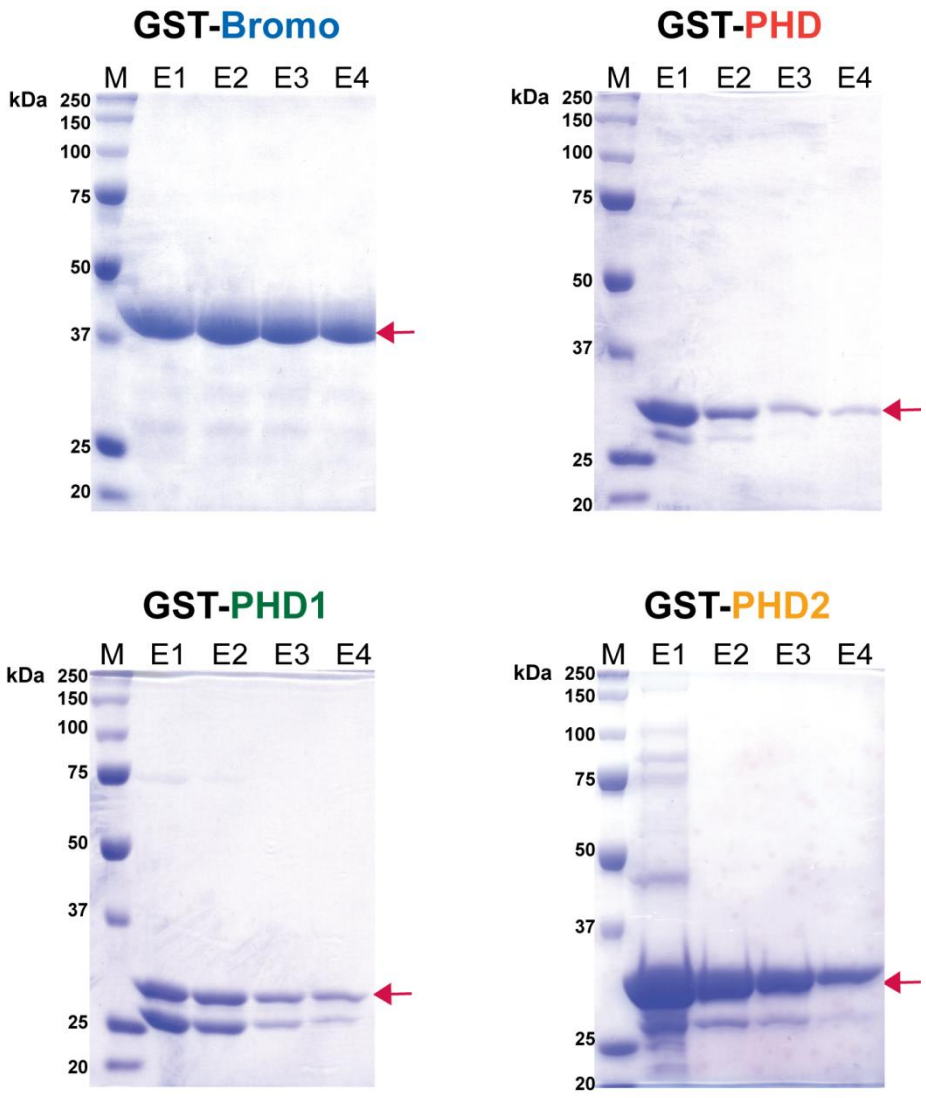


Figure 3-1. SDS PAGE gel image of GST column elution fractions of GST fusion proteins. The predicted molecular size of GST fusion domains are indicated in red arrows: Bromodomain (37 kDa), PHD (33 kDa), PHD1 (31 kDa) and PHD2 (31 kDa). The elution fractions are indicated (E1, E2, E3, E4), and additional bands detected below the expected GST fusion domains are GST proteins (26 kDa).

3.2.2 Characterisation of the histone modification targets of the PHD domains and Bromodomain by modified histone peptide library array analysis

Modified histone peptide array analysis was performed using the purified GST fusion domain proteins. In a first series of experiments, we used custom-generated modified histone peptide arrays produced by a University of Birmingham spin-off company, Alta Bioscience, as there were no existing commercially available arrays available at the start of this work. Two series of arrays were used: H3 peptide library arrays and histone H3, H4 N-terminal peptide library arrays. During the course of this work, commercially available histone peptide arrays were generated by Active Motif (Modified histone H3, H4, H2A and H2B peptide array), and were also used. The binding of GST-domain fusion proteins to peptide spots was visualized in two different ways, either using anti-GST HRP-conjugated antibody followed by ECL substrate reaction and ECL film exposure, or using anti-GST HRP-conjugated antibody as a primary antibody, followed by IRDye 800CW-conjugated anti-HRP antibody as a secondary antibody and visualized by 800 nm channel scanning using an Odyssey scanner (Figure 3-2).

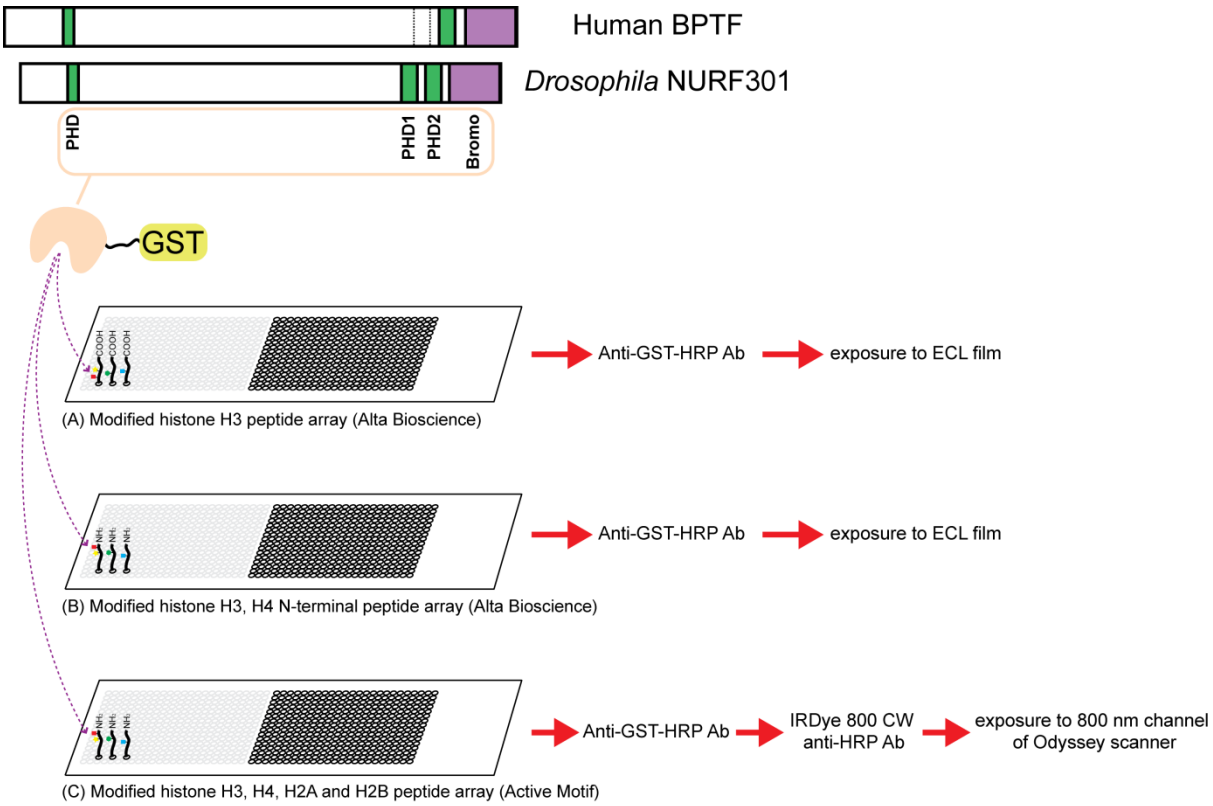


Figure 3-2. Overview of histone peptide library array assays. Purified GST conjugated fusion domain proteins (PHD, PHD1, PHD2 and Bromodomain) and the three different histone peptide arrays were used in these assays, (A) Modified histone H3 peptide array (Alta Bioscience), (B) Modified histone H3, H4 N-terminal peptide array (Alta Bioscience), (C) Modified histone H3, H4, H2A and H2B peptide array (Active Motif). GST fusion domain protein bound peptide spots were visualized as described above.

3.2.2.1 GST fusion PHD2 domain and bromodomain binding to the histone H3 peptide library array

We first tested the binding of GST-PHD2 domain protein to the histone H3 peptide library array as described above (Figure 3-2 A). GST-PHD2 was used as we already know that PHD2 can bind at least to the H3K4me3 mark. This provides a positive control to validate the array screening protocol. Both undialized purified GST-PHD2 and dialysed GST-PHD2 proteins were applied to arrays at a concentration of 0.1 mg/ml in the binding reaction to determine whether reduced glutathione affected the interaction of PHD2 domain with histone modifications. Purified GST domain protein alone was also used as a negative control. We

expected to see signals at peptide spots containing the H3K4me3 modification in both panels of the array, as the array is duplicated on each slide (see Appendix 1). However, we were unable to detect binding to any spots on the arrays (Figure 3-3). Moreover, the array background was increased when reduced glutathione was present, so all fusion proteins were dialysed before peptide binding reaction with these peptide library arrays. To validate that modified histone peptides were present on the arrays, we washed the slides to remove proteins and performed control immunostainings using an antibody that recognizes histone H3 phosphorylated at Serine 10 (anti-H3S10p antibody). As shown in figure 3-4, strong signals corresponding to H3S10p-containing peptide spots on the modified histone H3 peptide library array slides (see Appendix 2) were detected.

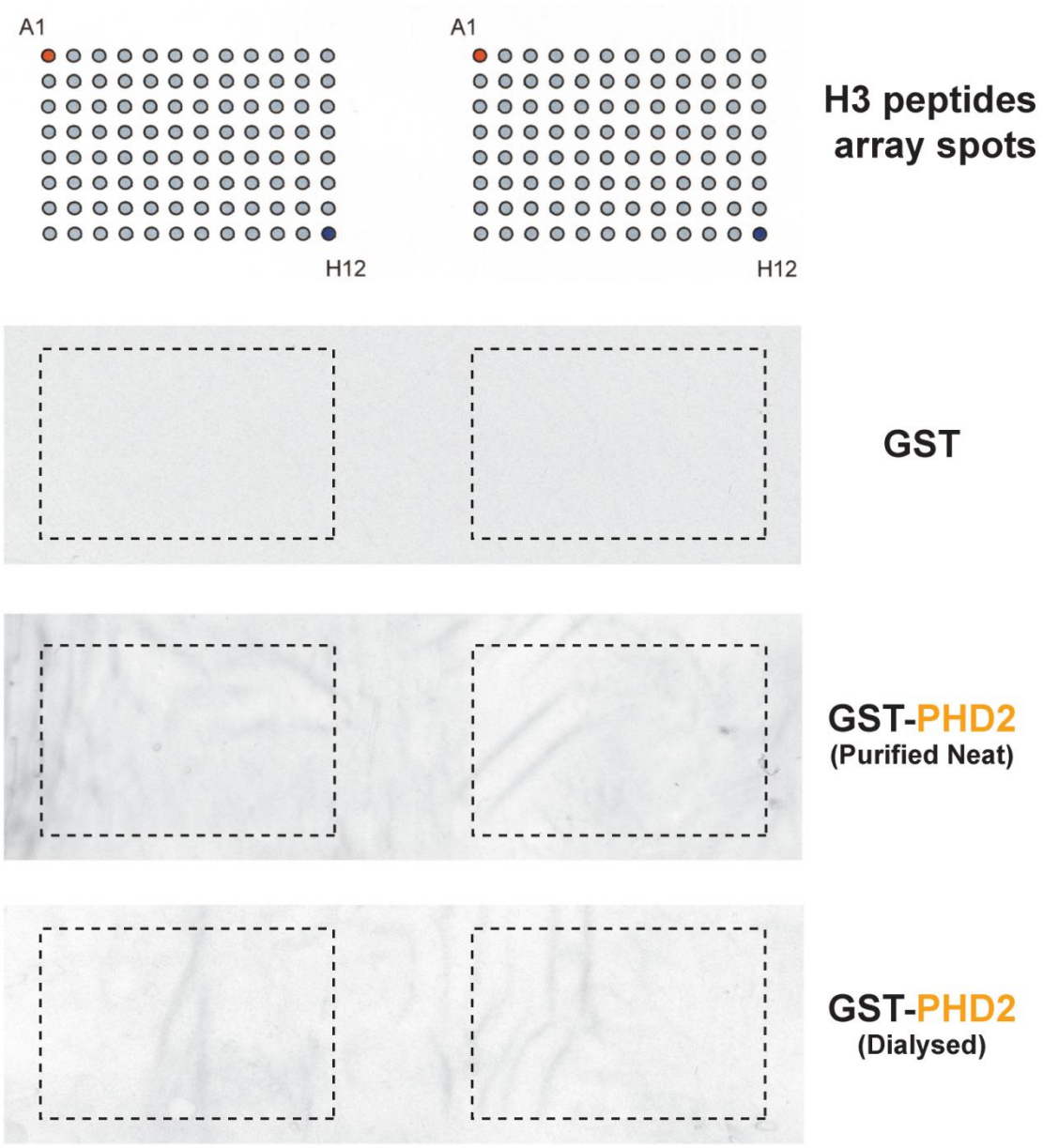


Figure 3-3. Modified histone H3 peptide library array assays of GST and GST-PHD2 domain proteins. The histone H3 peptide array grid is shown on the top. Binding of GST (control), GST-PHD2 (purified neat) and GST-PHD2 (dialysed) to this array was tested respectively and did not show any binding.

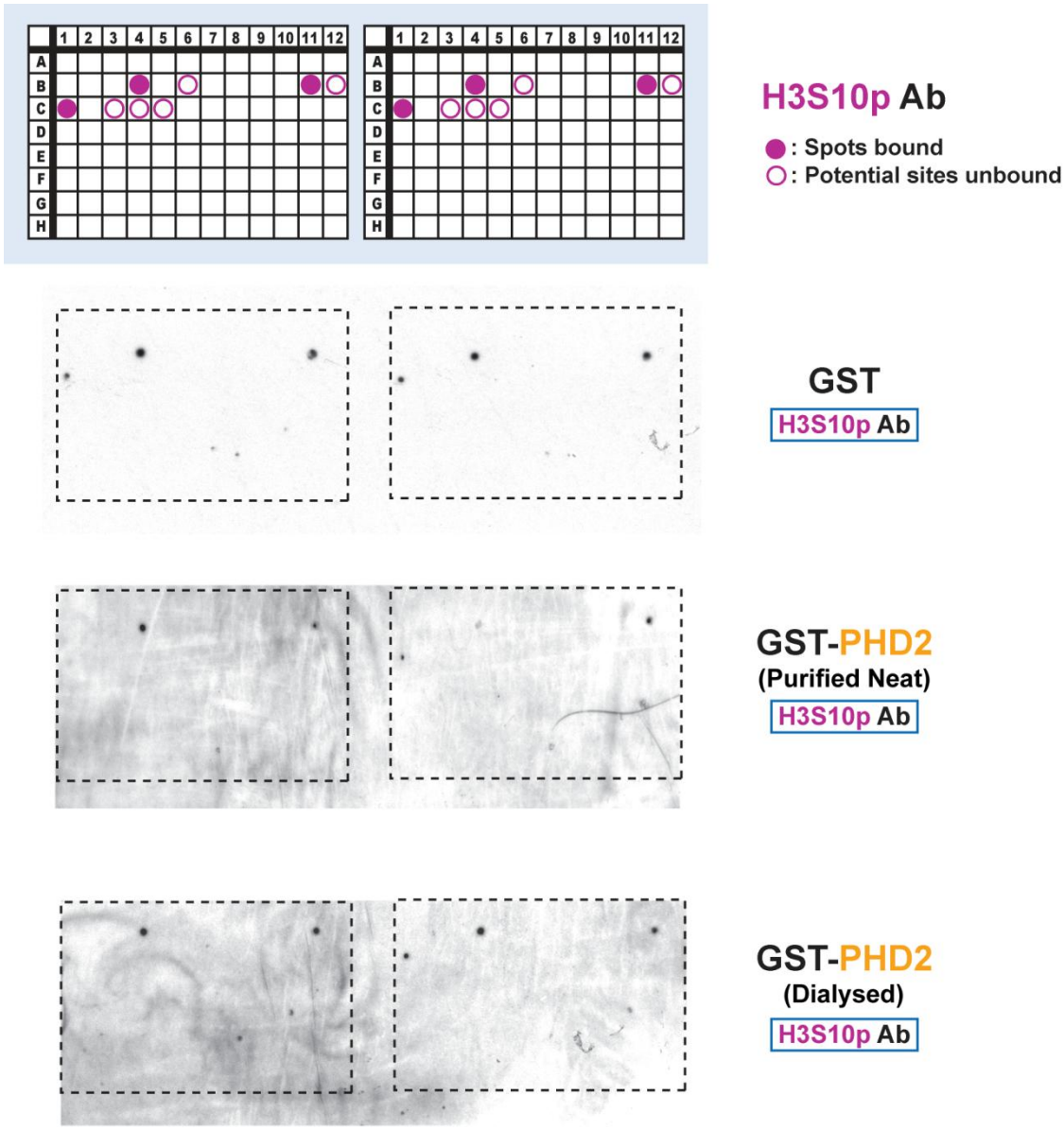


Figure 3-4. Modified histone H3 peptide library array validation using anti-H3S10p antibody. The array slides from figure 3-3 were immunostained with anti-H3S10p antibody followed by HRP conjugated anti-rabbit IgG secondary antibody staining to validate that modified histone peptides were present on the peptide arrays.

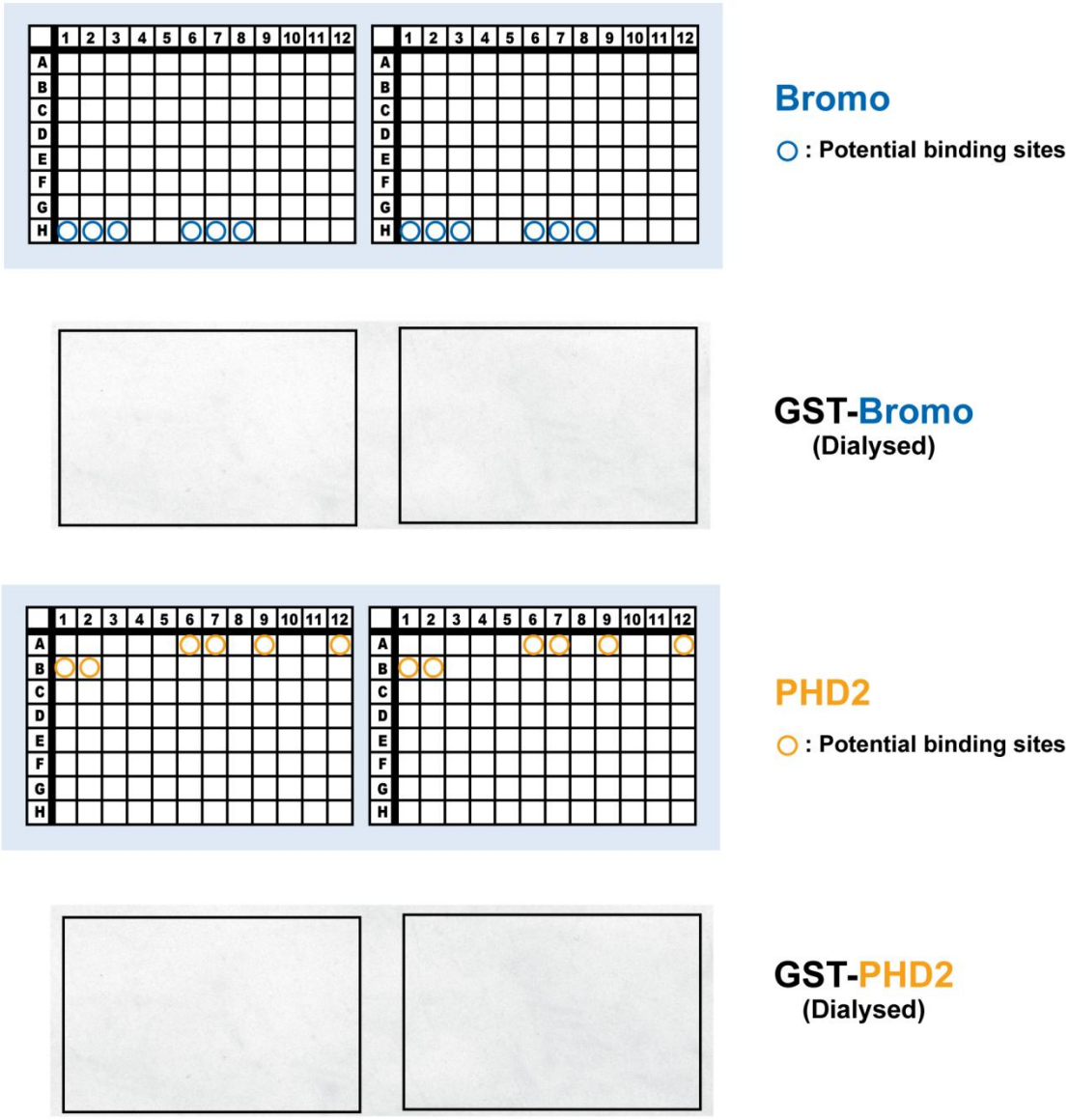


Figure 3-5. Modified histone H3 peptide library array assays of GST-Bromodomain and GST-PHD2 domain proteins using higher protein concentration. No binding was detected with anti-GST HRP-conjugated antibody.

To exclude that the concentration of GST-PHD2 protein used in the first experiments was not high enough to stabilize and detect GST-PHD2 binding, we repeated the hybridizations but increased the protein concentration to 0.2 mg/ml in the binding reaction. Assays were performed with dialyzed GST-Bromo and GST-PHD2 with new modified histone H3 peptide library array slides. However, no binding was detected by either GST-Bromo or GST-PHD2

peptide array assays (Figure 3-5). We suspected that although modified histone peptides are present on the arrays, they were not accessible to be bound by GST-PHD2 and GST-Bromodomain. Inspection of the array design (Appendix 2) revealed that peptides on the slide were immobilized to the slide surface via N-terminal biotin modifications. We hypothesized that the histone modifications on the peptides may not be recognized by our domain proteins because of steric hindrance, with the modifications too close to the slide surface. Therefore, we generated new histone peptide library arrays in which the peptides were bound to the slide surface via C-terminal biotin modifications, leaving the N-terminal sequences freely exposed.

3.2.2.2 Different interaction profiles of NURF301 domains with histone H3 and H4 modifications on histone H3, H4 N-terminal peptide library arrays

As described in Appendix 3, modified histone H3 peptides and H4 peptides were arrayed via the C-terminus (Modified Histone H3, H4 N-terminal peptide library array, Figure 3-2 B). These slides also contained modified histone H4 peptides allowing us to analyze the binding of domains to modifications on the histone H4 N-terminal tail. Agilent slide gaskets, chambers and a rotating hybridization chamber were also used for these binding reactions to maximize the protein concentration and hybridization efficiency. First modified histone binding assay was performed using purified GST-PHD2 protein. Significant binding of the PHD2 domain protein to H3K4me3 were observed. However, the slides contained significant background (data not shown). To improve signal to noise ratios, we asked the array supplier to generate arrays with larger peptide spots that would have greater potential for binding of domain fusion proteins. We also improved wash conditions to minimize backgrounds by using a slide stand and a wash bucket with 500-600ml of wash buffer stirred using a magnetic stirrer bar.

PHD2 domain peptide array binding assay was conducted again using these optimized conditions and with a higher concentration of protein sample in binding reaction (1-2 mg/ml). As expected, the PHD2 domain showed strong binding to peptide spots containing H3K4me3. No binding was detected to peptides that contained methylation of other residues or the mono- and di-methylation states of H3K4 indicating that PHD2 domain specifically binds H3K4me3 (Figure 3-6). Interestingly, however, the binding of PHD2 domain to H3K4me3 was prevented by additional modifications at the Thr 3 position. Thus the PHD2 domain failed to bind a peptide that was both phosphorylated at H3 Thr 3 and trimethylated at H3 Lys 4 (H3T3p/K4me3). In contrast, methylation at Arg 2 and acetylation or methylation of Lys 9 did not prevent PHD2 binding to H3K4me3. This indicated that PHD2 binding to H3K4me3 was hindered by the phosphorylation at Thr 3, but not by other flanking modifications such as Arg 2 di-methylation or Lys 9 trimethylation/acetylation. This data suggested that phosphorylated histone H3 threonine 3 could act as a rheostat function to control PHD2 binding to H3K4me3.

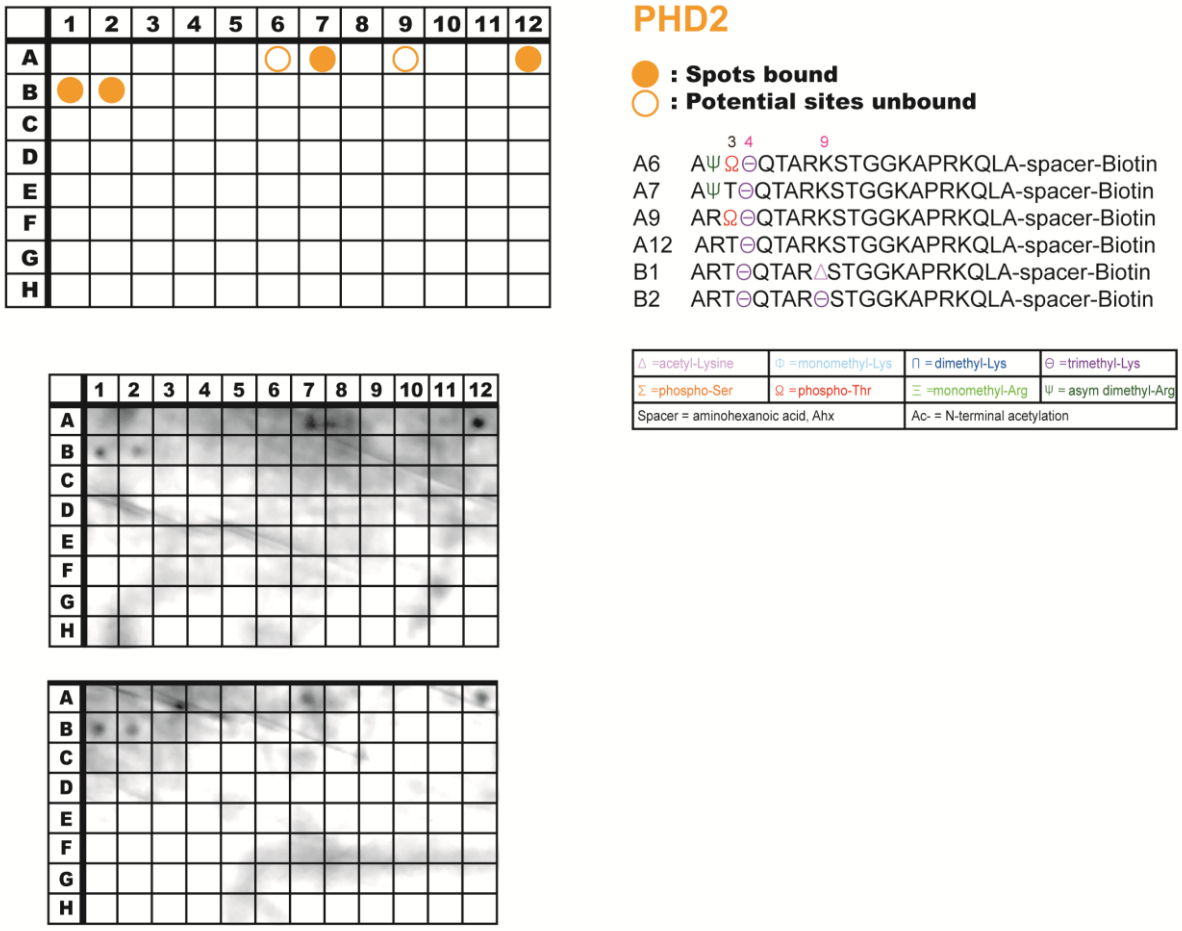


Figure 3-6. Binding of PHD2-GST fusion protein to histone H3, H4 N-terminal peptide library arrays. H3, H4 N-terminal peptide library arrays containing N-terminal free peptides were probed using PHD2-GST fusion proteins and showed specific binding to H3K4me3. PHD2-GST fusion protein bound spots were detected using anti-GST HRP-conjugated antibody.

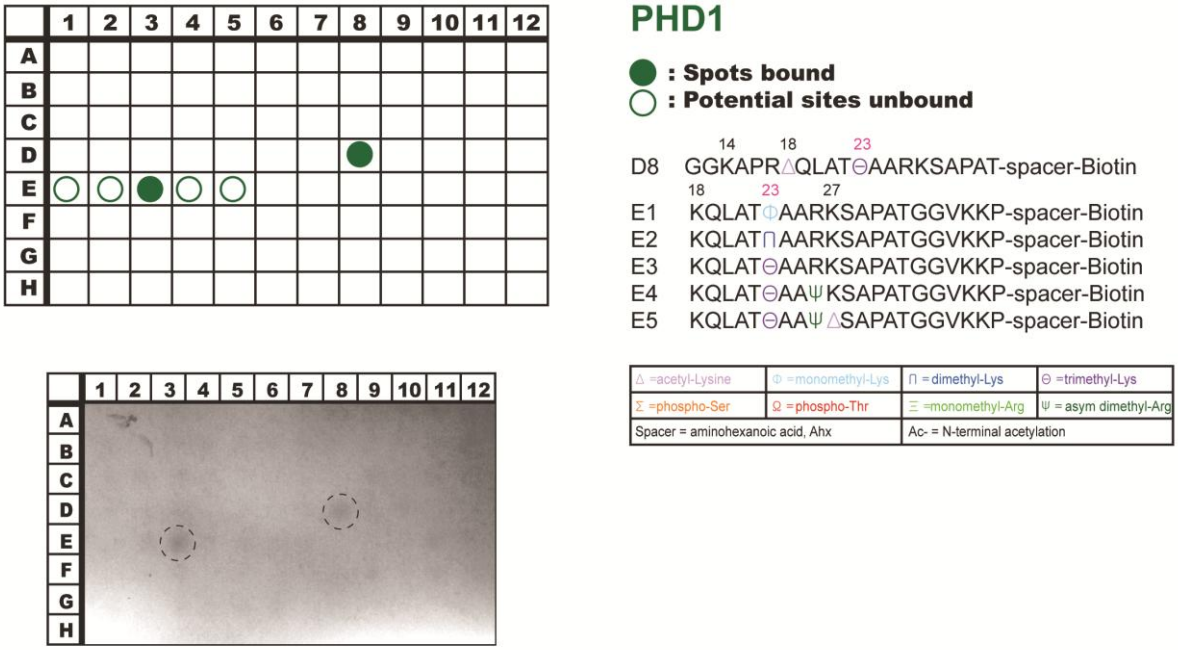


Figure 3-7. Binding of PHD1-GST fusion protein to histone H3, H4 N-terminal peptide library arrays. H3, H4 N-terminal peptide library arrays were probed using PHD1-GST fusion protein and showed binding to H3K23me3. PHD1-GST fusion protein bound spots were detected using anti-GST HRP-conjugated antibody.

We next examined the binding specificity of the PHD1 domain. Like the PHD2 domain, the PHD1 domain contains the core aromatic cage structure including the key residue Trp 32 which is critical for methylated lysine recognition (Li et al. 2006). Based on this, we anticipated that the PHD1 domain would also be able to bind to tri-methylated lysine residue. Indeed, the PHD1 domain specifically interacted with peptide trimethylated at histone H3 Lys 23 (H3K23me3) but did not recognize the mono- or di-methyl states of H3K23 (Figure 3-7). Binding to H3K23me3 was not influenced by acetylation at Lys 18 but was prevented by asymmetric dimethylation at Arg 26. As observed during our PHD2 domain analysis, these marks represent additional di-modification states that can act as a rheostat regulating recruitment of the PHD1 domain to H3K23me3.

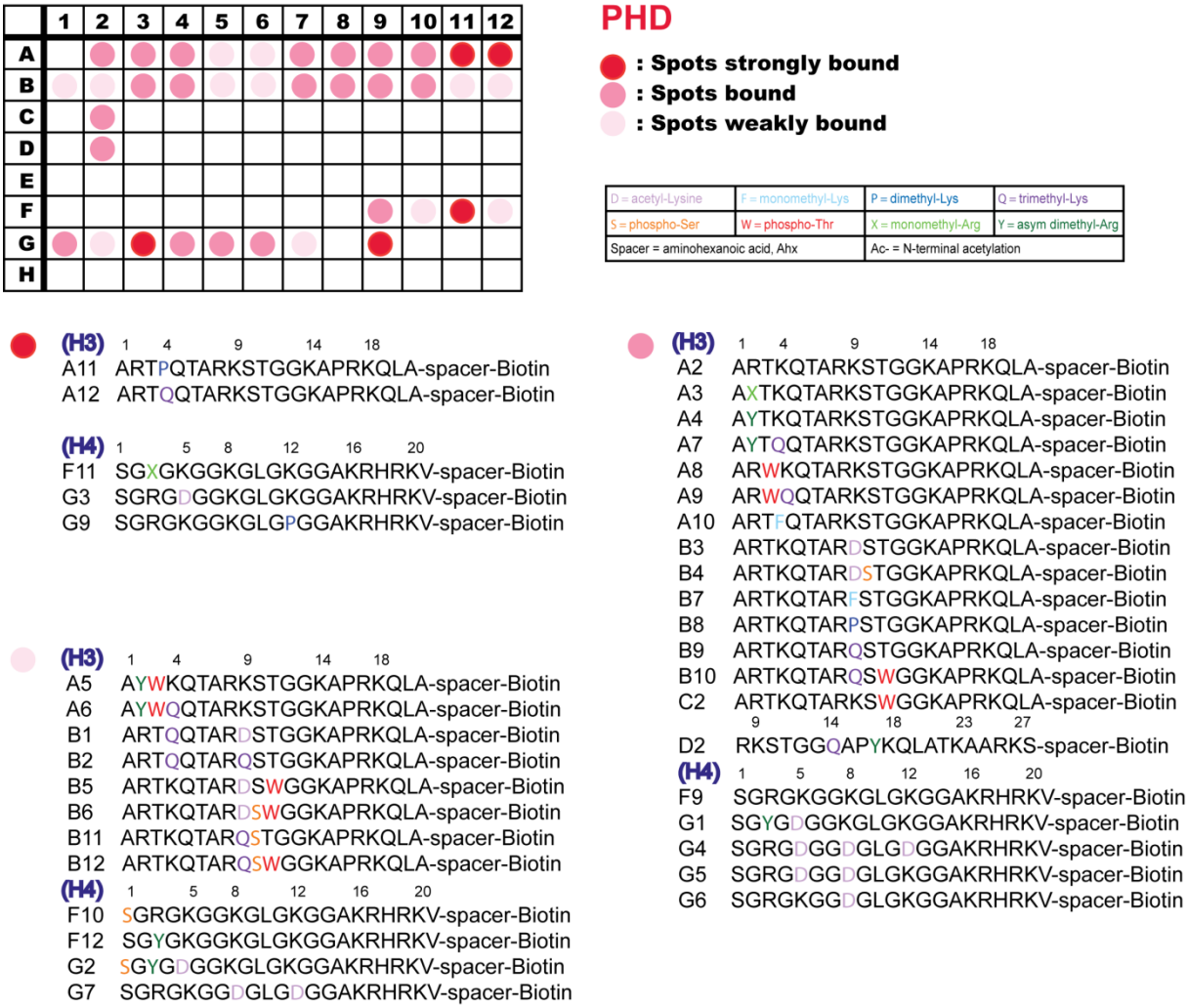


Figure 3-8. Schematic of binding of PHD-GST fusion proteins to histone H3, H4 N-terminal peptide library array. No clear modified histone binding specificity was observed.

Unlike PHD1 and PHD2 domains, the NURF301 PHD domain does not contain the aromatic cage required for binding to trimethylated lysine residues. Rather the sequence resembles more closely that of the BHC80 PHD domain which binds to unmodified histone tails (Lan et al. 2007). To identify residues to which the PHD domain could bind or that could potentially inhibit the PHD domain binding, the same peptide array assay was carried out with the GST-PHD domain. As expected, we did not observe specific binding of PHD domain. However interestingly, the PHD domain appeared to interact with the histone H3 and H4 N-terminal tails generally. As shown in the figure 3-8, a broad range of peptide spots were bound by the

PHD domain. Although no clear modification binding specificity was observed, there was a clear sequence preference of the PHD domain binding as the binding was only detected to peptides containing the N-terminal tails of histone H3(1-21) and H4(1-21).

3.2.2.3 Identification of the modified histone-binding specificities of NURF301 PHD domains and Bromodomain using MODified histone H3, H4, H2A and H2B peptide arrays

Histone peptide array analysis using arrays from Alta Bioscience identified some potential histone post translational modifications that could be recognized by the NURF301 PHD domains. However, the peptide arrays showed a high level of background and variability between arrays. It remained possible that weak binding would not be detected using these arrays and that some potential recruiting modifications could be overlooked. To clarify this, we used commercially available histone peptide arrays from another supplier (MODified histone H3, H4, H2A and H2B peptide array, Active Motif) for binding analysis with the PHD domains and bromodomain (Figure 3-2 C). These peptide arrays contain 59 histone modifications, including acetylation, methylation, phosphorylation and citrullination modifications on the H2A, H2B, H3 and H4 N-terminal tails. The arrays deliver 384 different combinations of those markers from single to quadruple modifications on the peptide (Appendix 4. and 5.) and are duplicated on the slide. Moreover, peptides are bound to the slide surface via C-terminal biotin modifications allowing domain fusion proteins to access more easily free N-terminal sequences. The existence of proprietary software also allowed us to analyse signal intensities from peptide spots and binding specificities to be calculated automatically.

The new arrays displayed significantly better binding efficiency and a lower background level.

Using these arrays, the PHD2 domain was screened again with 1-2 mg/ml of protein sample in the binding reaction. As expected, the PHD2 domain showed significant binding to peptide spots containing H3K4me3. No binding was detected to peptides that contained mono- and di-methylation states of H3 Lys 4 indicating that PHD2 specifically binds to tri-methylation of H3 Lys 4 (Figure 3-9 A and B). No significant binding was observed to any other singly methylated or acetylated histone marks (Figure 3-9 B and C). Moreover, as we observed from our previous peptide arrays, binding of PHD2 to H3K4me3 was prevented by additional modification at the Thr 3 position (H3T3p). Thus the PHD2 domain failed to bind to the peptides containing phosphorylated H3T3 and trimethylated H3K4 (H3T3pK4me3). Dimethylation of Arg 2 (H3R2me2s, H3R2me2a) also showed a minor inhibitory effect on PHD2 binding to H3K4me3. Relative binding intensity analysis revealed that compared to H3K4me3 alone, PHD2 binding was decreased to 12% when Thr 3 was also phosphorylated. Histone H3 Arg 2 dimethylation (H3R2me2s, H3R2me2a) also decreased the binding to 30% and 60% (Figure 3-9 D). Therefore, our data suggest that PHD2 domain has strict binding specificity for H3K4me3, but this binding can be inhibited by phosphorylation of Thr 3.

Interestingly, the new peptide array data also showed that flanking trimethylation or acetylation of the histone H3 Lys 9 significantly increased the binding of the PHD2 domain to H3K4me3. Although the PHD2 domain did not bind to H3K9me3 or H3K9ac alone, these modifications increased binding when combined with H3K4me3. Compared with the binding signal intensity of H3K4me3 alone, peptides combining H3K4me3 with methylations or acetylation of H3K9 showed 300-400 % stronger binding (Figure 3-9 E). We speculate that methylation or acetylation at H3 Lys 9 can increase the binding capacity of PHD2 domain to H3K4me3. In combination, these results identified flanking histone modifications that could

inhibit or enhance the binding of NURF301 PHD2 domain to H3K4me3 and provide possible epigenetic rheostats that could control NURF binding to chromatin.

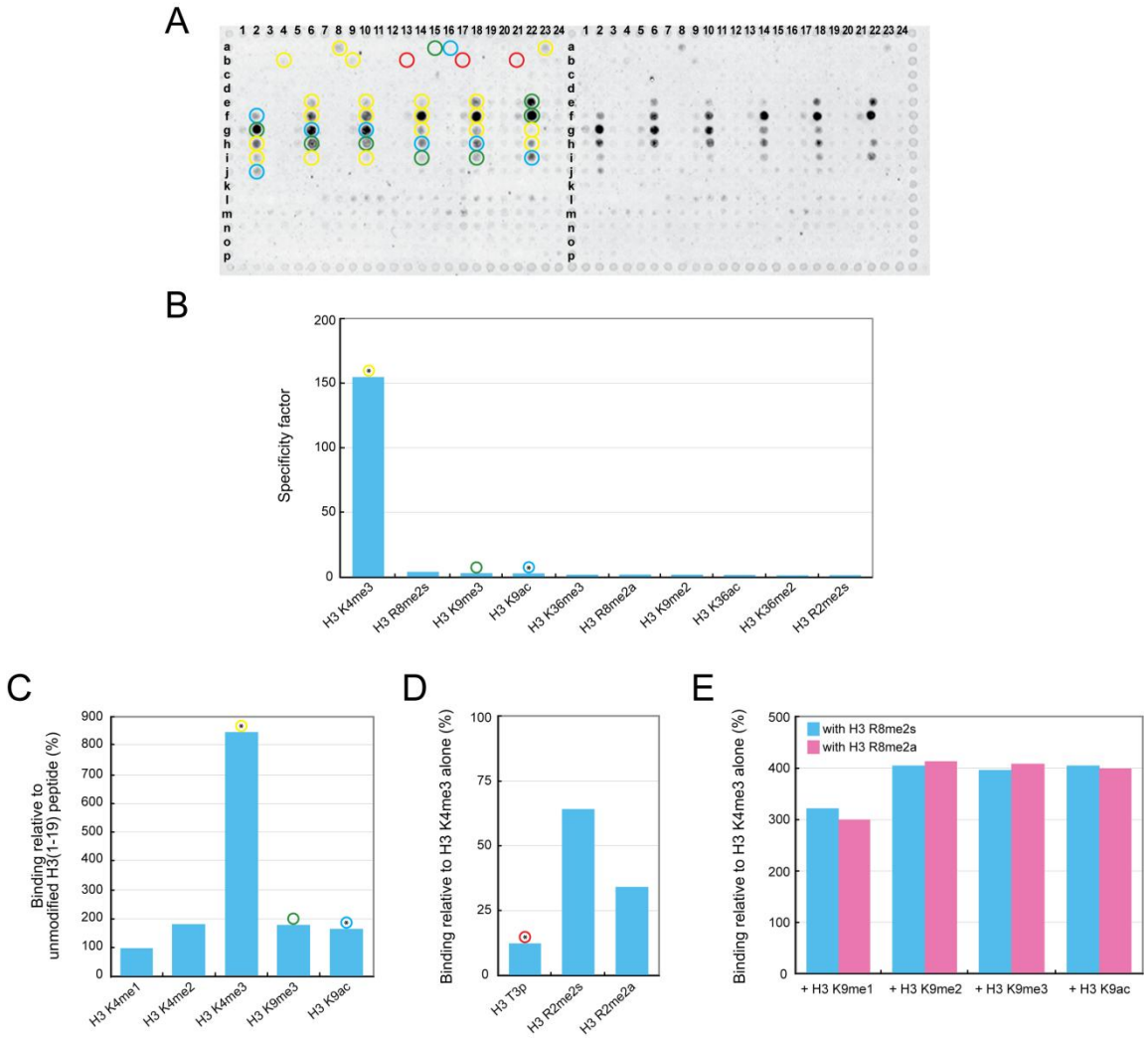


Figure 3-9. Binding of PHD2-GST fusion protein to MODified histone H3, H4, H2A and H2B peptide arrays. (A) Modified histone peptide arrays containing different modification combinations were probed with PHD2-GST fusion protein. PHD2-GST domain bound spots were detected using anti-GST HRP-conjugated primary antibody and IRDye 800CW-conjugated rabbit anti-HRP secondary antibody. (B) The domain bound peptide spot signal intensities were analysed by using the Array Analyse software program (Active Motif). The top 10 modifications often detected from the domain bound peptides are listed on the graph. The PHD2 binding specificity factor list shows that PHD2 binds most specifically to H3K4me3 (yellow circle). H3K9me3 (green circle) or H3K9ac (blue circle) modifications along with H3 K4me3 also appeared to be weaker binding sites. (C) Relative binding intensity of singly modified histone H3 peptides in comparison to unmodified H3 peptide. (D) Relative binding intensity analysis showed that H3T3p (red circle), H3R2me2s and H3R2me2a marks inhibit PHD2 binding to H3K4me3. (E) H3K9me1, H3K9me2, H3K9me3 and H3K9ac marks enhance binding of PHD2 to H3K4me3. (D, E) The PHD2 binding to these combinatorial marks was calculated relative to the binding to the singly modified H3K4me3 peptide.

We next examined the binding specificity of the PHD1 domain to the modified histone peptide arrays. Like the PHD2 domain, the PHD1 domain also contains the core aromatic cage structure including the key Trp 32 residue that is critical for binding to trimethylated histone tails (Li et al. 2006). We observed that PHD1 was able to bind trimethylated histone tails but also to acetylated histone tails. The PHD1 domain bound modifications were distinct from our previous PHD1 domain peptide array data (Figure 3-7). It is important to note that the H3K23me3 mark is not found on these arrays and as such it was not possible to confirm our previous PHD1 domain binding data. Thus, array analysis showed that the PHD1 domain mainly interacted with histone H3 acetylated H3K36 (H3K36ac) and also independently bound to di- and tri-methylated Lys 36 (H3K36me2 and H3K36me3) (Figure 3-10 A and B). The binding of the PHD1 domain to H3K36ac was two times stronger than that of unmodified H3(26-45) peptide, followed by H3K36me2 and H3K36me3 as 1.75 times and 1.25 times stronger respectively (Figure 3-10 C). By comparison binding of the PHD2 domain to H3K4me3 peptide was 8 times stronger than the unmodified reference peptide.

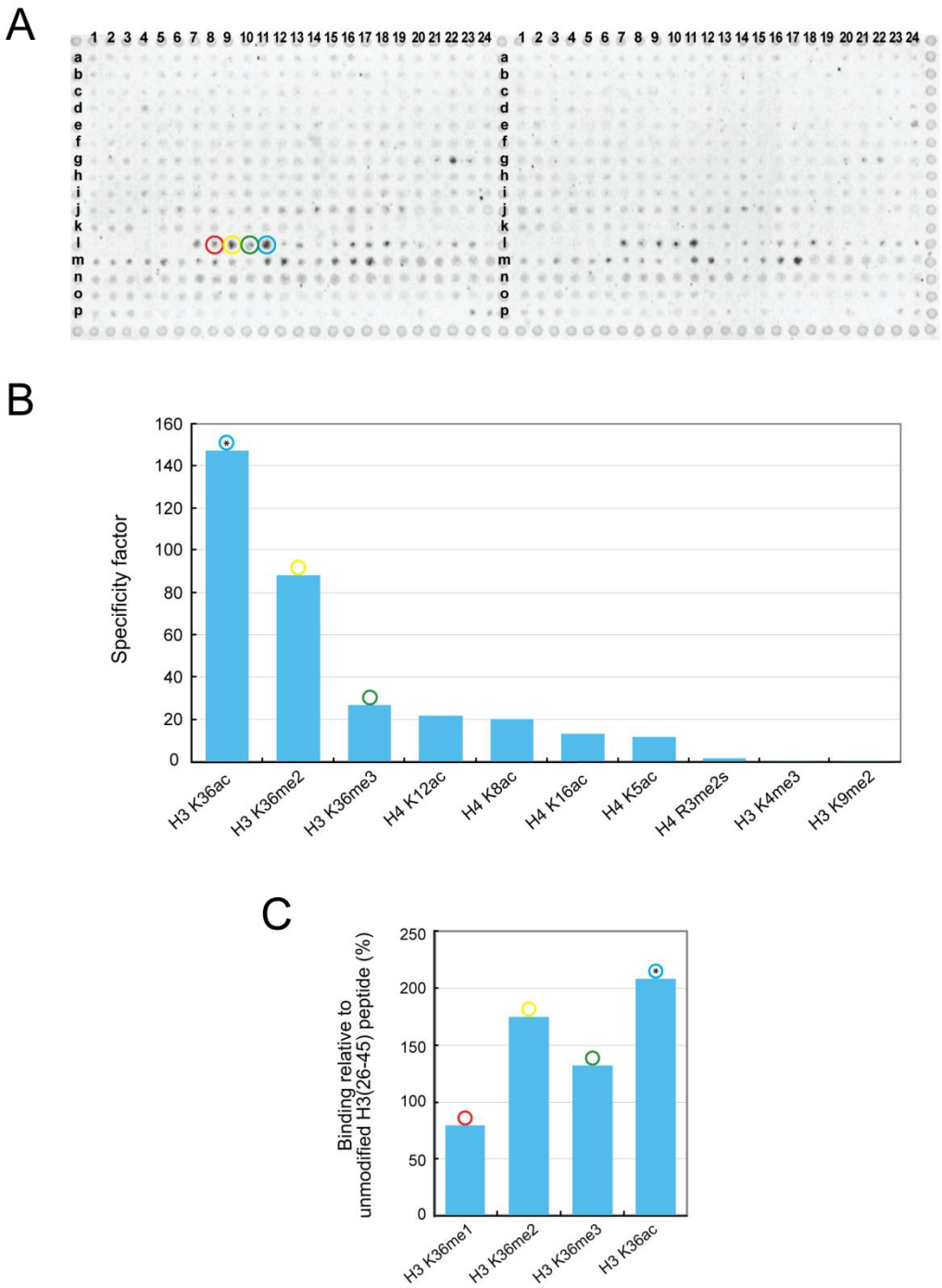


Figure 3-10. Binding of PHD1-GST fusion protein to MODified histone H3, H4, H2A and H2B peptide arrays. (A) Peptide arrays were probed with PHD1-GST fusion proteins and protein bound spots were detected as described previously. (B) PHD1 binding specificity factor analysis shows that the PHD1 domain potentially binds to H3K36ac (blue circle), followed by H3K36me2 (yellow circle) and H3K36me3 (green circle). (C) Relative binding intensity analysis shows that of peptides present on the array, H3K36ac is the preferred binding site relative to the unmodified H3 (26-45) peptide.

To identify residues that the PHD domain could potentially bind, we performed peptide array binding analysis with the PHD-GST fusion protein. While we did not observe highly specific binding of PHD domain, some binding was detected to H3K36me2, H3K36me3 and H3K36ac which is similar to PHD1 domain binding specificity. In addition, the PHD domain also interacted with histone H4 N-terminal sequences that contained acetylated H4K16 residue (Figure 3-11 A and B). Though PHD domain does not have strong binding specificity, it still appears to be some sequence preference of histone H4 N-terminal free sequences, as PHD domain binding was not observed to peptide spots having modifications beyond H4K20 peptide sequences.

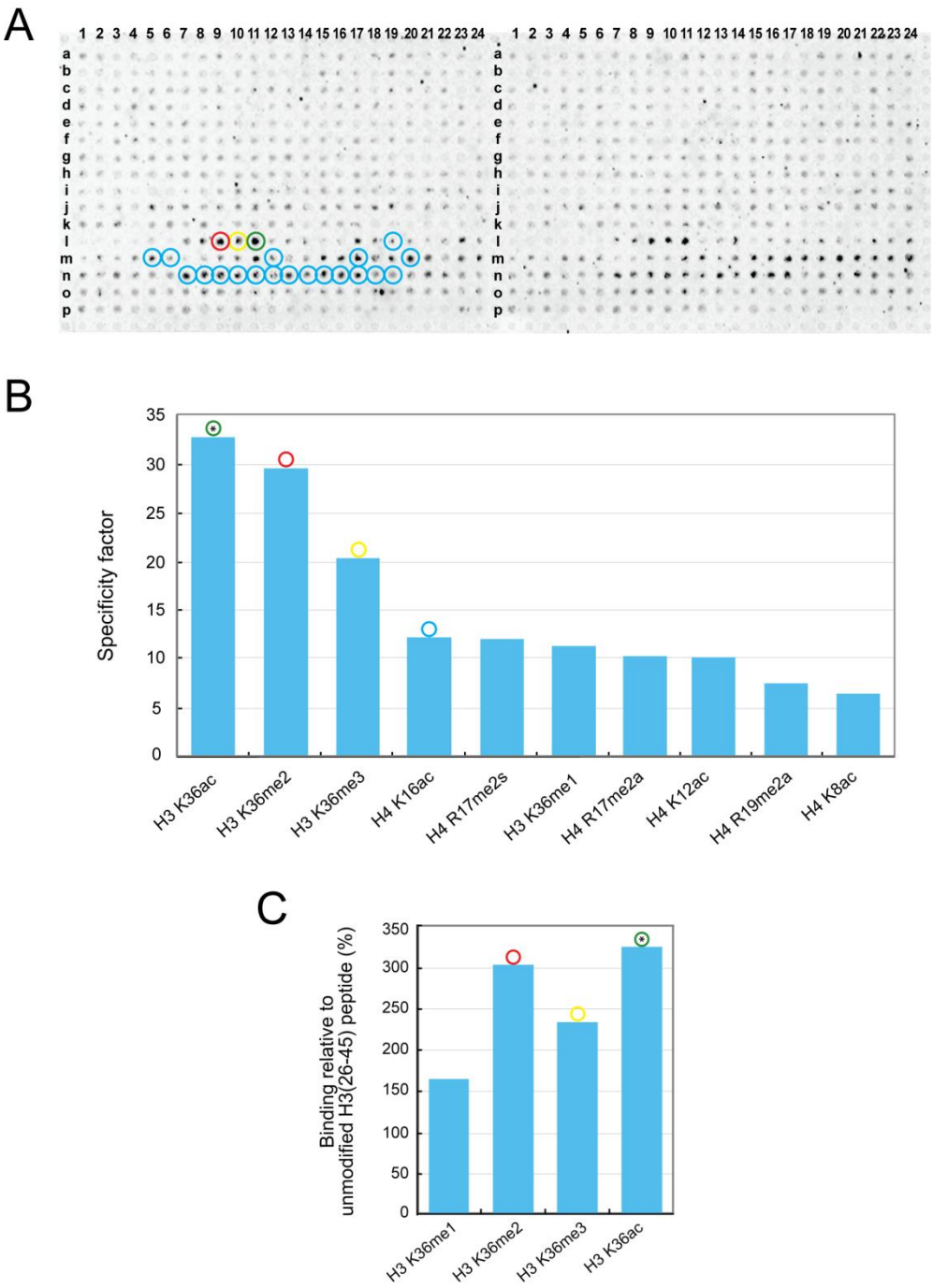


Figure 3-11. Binding of PHD-GST fusion protein to MODified histone H3, H4, H2A and H2B peptide arrays. (A) Peptide arrays were probed with PHD-GST fusion proteins and protein bound spots were detected as described above. (B) PHD binding specificity factor list shows that PHD domain binds to H3K36ac (green circle), H3K36me2 (red circle), H3K36me3 (yellow circle) and H4K16ac (blue circle), but more likely binds generally to histone H4 N-terminal free sequences. (C)) Relative binding intensity of modified histone H3 peptides in comparison to unmodified H3 peptide.

Finally, to investigate the interaction of the bromodomain with modified histones, MODified histone H3, H4, H2A and H2B peptide arrays were incubated with purified GST-bromodomain protein. The array analysis showed that bromodomian binds to the acetylated H4K16 as expected, but other singly acetylated lysine residues were also detected to bind including H4K5ac, H4K8ac and H4K12ac (Figure 3-12 A and B), H4K12ac actually was shown as more specific binding site (Figure 3-12 C). However, interestingly, our peptide array and relative binding analysis data indicated enhanced binding of the bromodomain to di-, tri- or tetra- acetylation combinations of those modifications (Figure 3-12 D). Among peptides that bound, tetra-acetylated at H4K5, K8, K12 and K16 (H4K5, 8, 12, 16ac) showed the strongest binding and appeared to be the most significant binding site of the NURF301 bromodomain. This data was consistent with a recent study showing that bromodomain binding pocket is able to bind to two acetylated lysines (Filippakopoulos et al. 2012) raising the possibility that the binding strength of NURF301 bromodomain was also increased by multiple acetylations on histone H4 N-terminal tail.

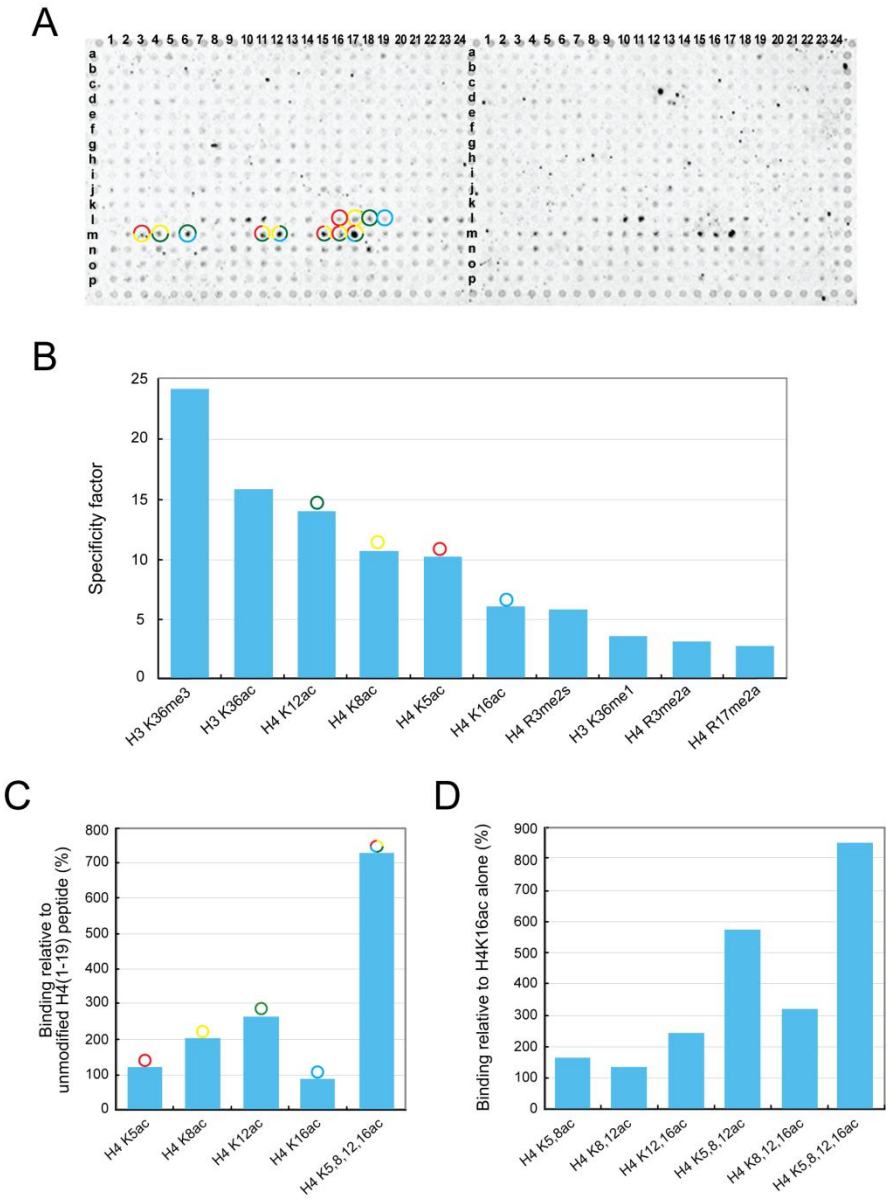


Figure 3-12. Binding of Bromo-GST fusion protein to MODified histone H3, H4, H2A and H2B peptide arrays. (A) Peptide arrays were probed with Bromo-GST fusion proteins and protein bound spots were detected as described previously. (B) Bromodomain binding analysis shows that the bromodomain binds to residues including H4K5ac (red circle), H4K8ac (yellow circle), H4K12ac (green circle) and H4K16ac (blue circle) with a distinct preference for tetra-acetylated H4 tails (H4K5/K8/K12/K16ac). (C) Relative binding intensity of modified histone H4 peptides in comparison to unmodified H4 peptide. (D) Relative binding intensity analysis showed that multiple acetylations enhance binding of the bromodomain to H4 N-terminal peptides in comparison to the binding to the singly modified H4K16ac peptides.

3.2.3 Validation of modified histone interactions by peptide pull down assays

3.2.3.1 PHD2 domain

To confirm the interaction of PHD2 domain with binding markers identified in our peptide library array assays, peptide pull down assay was performed using both the *Drosophila* NURF301 PHD2 domain and the human BPTF PHD2 domain. This was prepared to confirm whether binding was functionally conserved. The GST--conjugated PHD2 domain proteins were allowed to bind to modified histone peptides (biotinylated). Peptide bound PHD2 proteins were recovered on streptavidin-coated agarose beads, and the amount of copurified PHD2 domain proteins were determined by western blotting. In addition to modifications detected on arrays, we also examined the effect of flanking phosphorylation at histone H3 Ser 10, located next to H3K9ac, which was found as an enhancing marker on PHD2 binding to H3K4me3 in our study. H3S10p is well known histone marker involved in mitotic/meiotic chromosome condensation (Prigent and Dimitrov, 2003). We initially postulated that phosphorylation of H3S10 could antagonize H3K9ac like H3T3p does to H3K4me3, and mask the enhancing effect on PHD2 domain binding to the H3K4me3/K9ac mark.

As we expected, both NURF301 (*Drosophila*) and BPTF (Human) PHD2 domains bound to H3K4me3, but the flanking phosphorylation of H3T3 significantly inhibited the binding of PHD2 domains to H3K4me3 (H3T3p/K4me3) showing no detectable binding to H3T3pK4me3 peptides. In contrast, the flanking acetylation of H3K9 resulted in enhanced binding of PHD2 domains to H3K4me3 (H3K4me3/K9ac). Interestingly additional phosphorylation of H3S10 resulted in a greatly enhanced binding (H3K4me3/K9ac/S10p). Differences in peptide-PHD2 domain binding affinities were only seen with 0.1 µg of peptide pull down reactions, while these were not seen when 1µg of peptides were supplied in our

standard reactions suggesting that higher levels of peptide can override any differences in the affinity of the PHD2 domain (Figure 3-13 A).

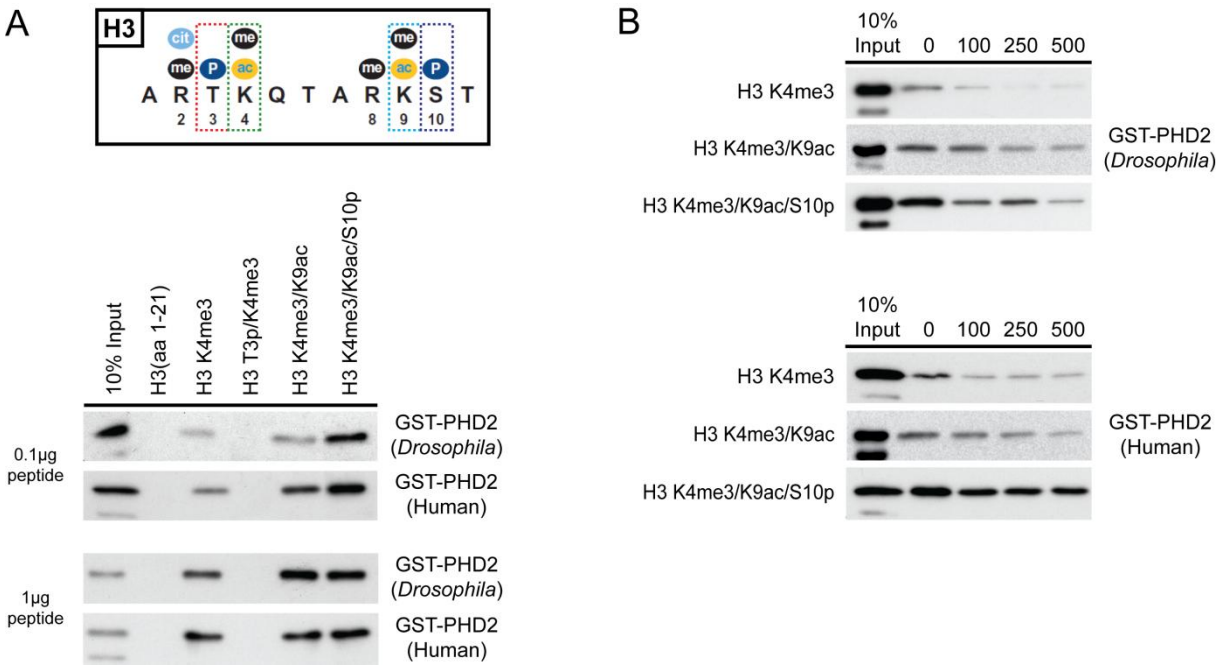


Figure 3-13. The PHD2 domain peptide pull down and peptide competition pull down assays. The binding specificities of the *Drosophila* NURF301 PHD2 domain and the human BPTF PHD2 domain (GST conjugated) to histone markers identified from peptide library array assays were determined by peptide pull down and peptide competition pull down assays. (A) Interactions between GST-PHD2 domains and either 0.1 µg or 1 µg of biotinylated histone peptides were analysed by peptide pull down and visualized by western blot using rabbit anti-GST (HRP) antibody. (B) Relative affinities of PHD2 domains to different histone modifications were analysed by peptide competition pull down, in which constant 0.1µg of biotinylated peptides (H3K4me3, H3K4me3/K9ac and H3K4me3/K9ac/S10p) and 0 to 500 molar excess of competitor H3K4me3 peptide mixed in the GST-PHD2 fusion protein binding reactions. Bound proteins were then visualized by western blot using rabbit anti-GST (HRP) antibody.

To further determinate differences in binding affinities for these peptides, we also performed peptide competition pull down assays to determine the relative affinities of both NURF301 PHD2 domain and BPTF PHD2 domain to different histone modifications on the peptides. In particular, relative PHD2 binding affinities between H3K4me3 and H3K4me3/K9ac, between H3K4me3 and H3K4me3/K9ac/S10p were examined. In each PHD2 domain peptide pull

down reaction, a 0 fold, 100 fold, 250 fold and 500 fold molar excess of H3K4me3 peptide as competitor over biotinylated H3K4me3/K9ac or H3K4me3/K9ac/S10p peptide was added. The same reactions between H3K4me3 and biotinylated H3K4me3 peptide were also conducted as controls. Even in the 250 and 500 fold excess of H3K4me3 peptides, binding of PHD2 domain to biotinylated H3K4me3/K9ac and H3K4me3/K9ac/S10p peptides was still observed, while binding to the control biotinylated H3K4me3 peptide was lost at 100 fold excess of the competitor (Figure 3-13 B). Titration of competitor H3K4me3 peptide indicates that the binding specificity of PHD2 to either H3K4me3/K9ac or H3K4me3/K9ac/S10p are about 100 times and 250 times respectively stronger than the binding to the singly modified H3K4me3 histone H3 peptides. In addition, the binding affinity of PHD2 to H3K4me3/K9ac/S10p is stronger in comparison to H3K4me3/K9ac. This result contradicted our initial hypothesis that phosphorylation of H3S10 would block PHD2 binding. Instead, this suggests that phosphorylation itself does not antagonize the PHD2 binding to H3K4me3 as phosphorylation of H3T3 inhibits binding, whereas phosphorylation of H3S10 enhances binding. It is the location of the phosphorylated residue that determines the affect on binding.

The X-ray crystal and NMR based structures of the human BPTF PHD2 domain in complex with H3(1-15)K4me3 modified histone tail peptide (Li et al., 2006) exist. We sought to examine if there were features of the H3K4me3 binding pocket that could account for the antagonistic effect of the H3T3p and H3K9ac/S10p residues. The structure was visualized, modified and analysed using Accelrys Discovery Studio Visualizer 2.0 (Accelys). In the original structure, the trimethyl group of K4me3 is stably positioned within a cage composed of four aromatic amino acids, Tyr 10, Tyr 17, Tyr 23 and Trp 32 (Y10, Y17, Y23 and W32), as described in the previous study (Li et al., 2006). We then modified the structure of the H3 tail

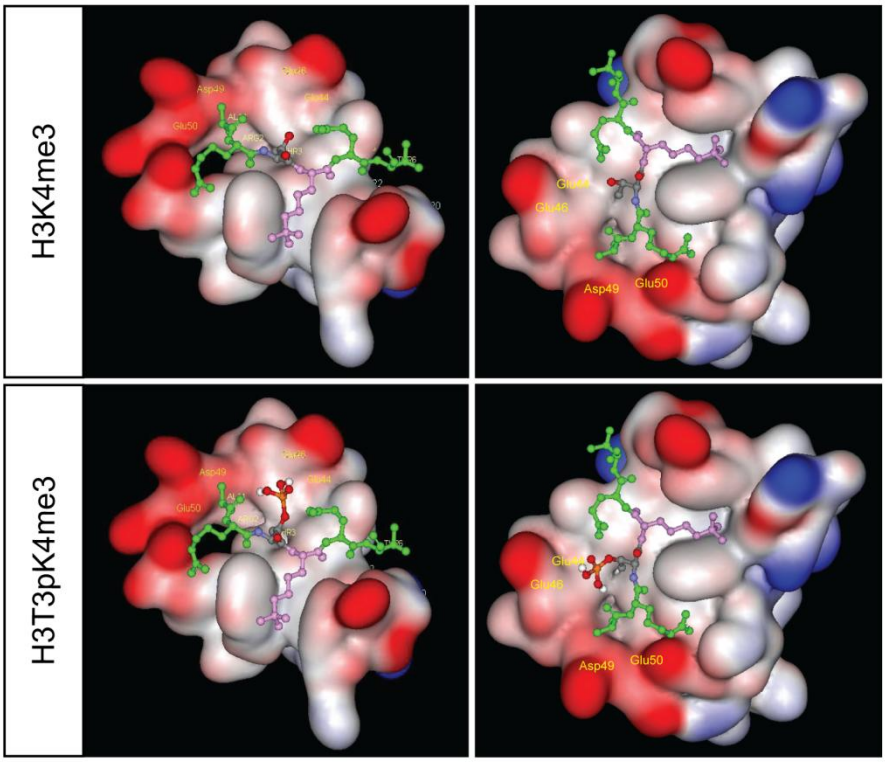
peptide to model the addition of a negatively charged phosphate group at Thr 3. Analysis of electrostatic charge distribution in the PHD2 binding pocket shows that negatively charged phosphate at Thr 3 is possibly repulsed by negatively charged amino acids, Glu 44, Glu 46, Asp 49 and Glu 50 (E44, E46, D49 and E50), in the PHD2 binding pocket. This provides a structural basis of how the binding of PHD2 to H3K4me3 can be inhibited by flanking phosphorylation of H3T3 (Figure 3-14 A).

We also attempted to modify the H3 tail peptide from the NMR structure to model the addition of an acetyl group at Lys 9 and negatively charged phosphate group at Ser 10. Due to the flexibility of this region of the H3 peptide, these residues are not localised in the X-ray crystal structure precluding similar analysis of the X-ray structure. However, this region of the H3 peptide can be visualized in the related NMR structures. When we looked at the potential binding region of BPTF PHD2 binding pocket and unmodified or modified H3K9/S10 peptides, we observed that H3K9ac and S10p, which are neutralized by acetylation and negatively charged by phosphorylation respectively, were surrounded by positively charged amino acids, Lys 21 and Arg 36 (K21 and R36) (Figure 3-14 B). Electrostatic interaction between the residues could enhance the binding capacity of PHD2 domain to H3K4me3, as an additional negatively charged phosphate group at Ser 10 would favourably interact with positively charged regions in the PHD2 domain when it binds to H3K4me3 on the peptide. Altough we can not localise this region of the H3 peptide in the X-ray crystal structure, analysis of the region where the peptide would be predicted to exit (based on the NMR sructures) shows clear representation of the favourable electrostatic surface.

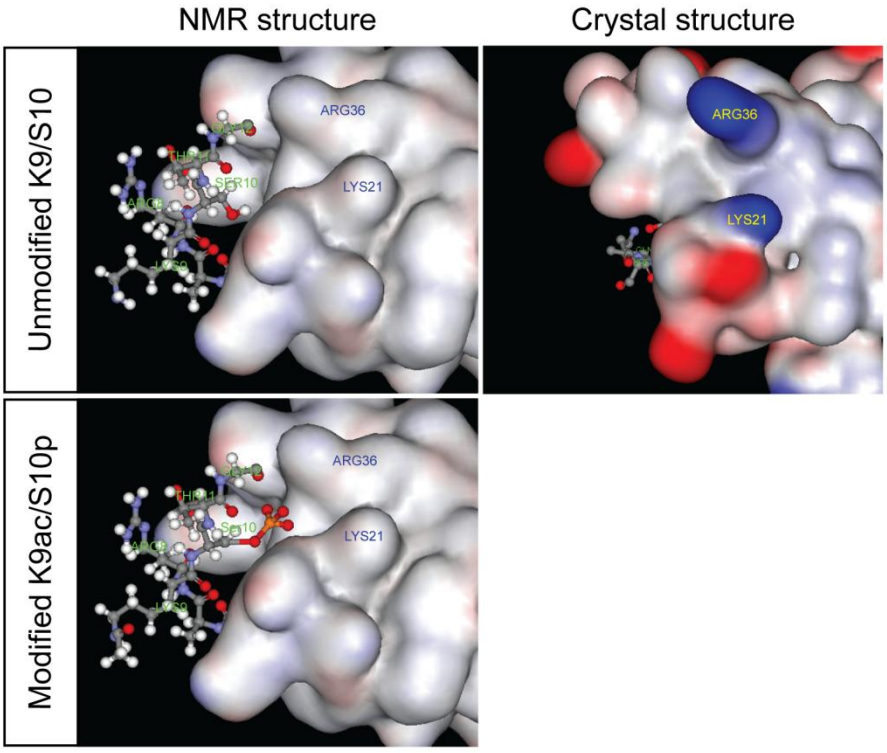
Finally, the protein sequence alignments of the PHD2 domains in *Drosophila* NURF301,

human and mouse BPTF showed high conservation of these amino acid residues that we postulate are essential for binding of the rheostat histone modifications (Figure 3-14 C). One exception is the positively charged Lys 21 residue that may be a critical residue for interaction with negatively charged Ser 10. This residue is not present in *Drosophila*, which may explain the differences in our peptide competition pull down analysis data (Figure 3-13 B), which showed weaker binding of *Drosophila* NURF301 PHD2 to H3K4me3/K9ac/S10p peptide in comparison to human BPTF PHD2.

A



B



C

```
                10        17   21 23        32   36          44 46  49 50
                *         *    +  *         *    +           -  -   -- 
Drosophila NURF301 PHD2  YCSCRQPYDESQFYICCDKCQDWFHGRCVGILQSEAEFIDEYVCPECQ
Human BPTF PHD2          YCICKTPYDESKFYIGCDRCQNWYHGRCVGILQSEAELIDEYVCPQCQ
Mouse BPTF PHD2          YCICKTPYDESKFYIGCDRCQNWYHGRCVGILQSEADLIDEYVCPQCQ
```

Figure 3-14. Structure modelling of the interaction between rheostat modifications and the human BPTF PHD2 domain. (A) Crystal structure of the BPTF PHD2 binding pocket with H3K4me3 and H3T3p/K4me3 peptides. Thr 3 residue is coloured by element of the atom, and trimethylated Lys 4 residue is in purple on the tail peptide backbone. (B) NMR and crystal structure modelling of the BPTF PHD2 binding pocket with H3K4me3 and H3K4me3/K9ac/S10p peptides. In (A) and (B), the PHD2 domain binding pocket is displayed in electrostatic surface representation coloured in red as negatively charged and blue as positively charged. (C) Protein sequence alignments of the NURF PHD2 domains in *Drosophila*, human and mouse show highly conserved amino acid residues that are critical for binding of these histone modifications. Amino acid residues important for H3K4me3 binding are indicated in green(*), negatively charged residues involved in repulsion of H3T3p in red(-) and positively charged residues involved in H3K9ac/S10p binding in blue(+).

3.2.3.2 PHD1 domain

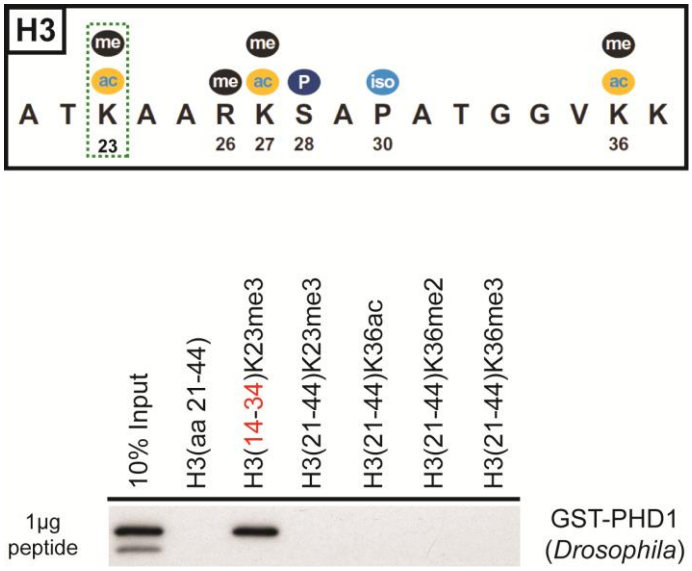


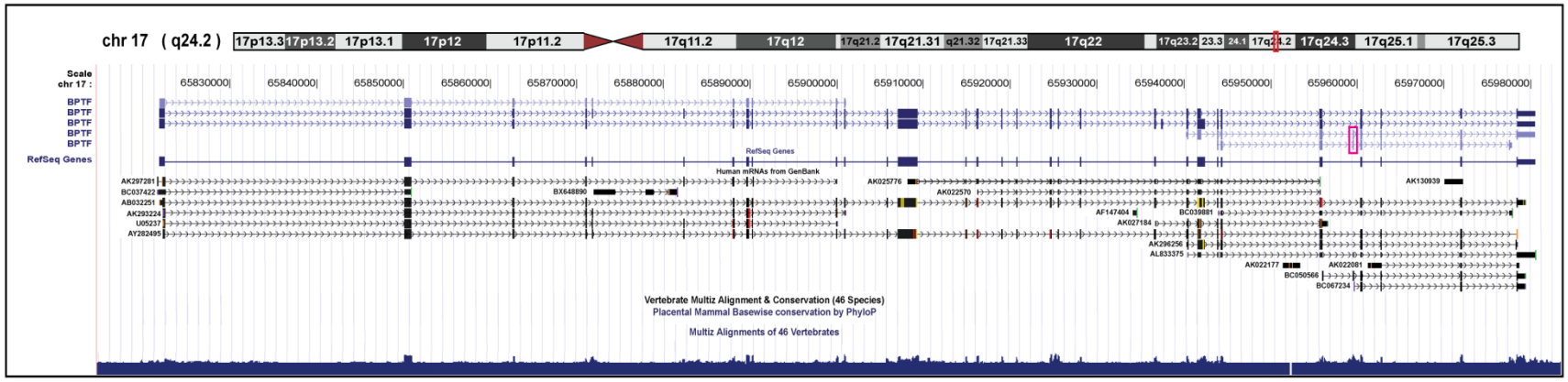
Figure 3-15. The PHD1 domain peptide pull down assays. The binding specificities of the *Drosophila* NURF301 PHD1 domain (GST conjugated) to histone modifications identified from peptide library arrays were validated by peptide pull down and western blot analysis using rabbit anti-GST (HRP) antibody.

To verify that the observed interaction between the NURF301 PHD1 domain and histone modifications in the peptide array data, we performed a peptide pull down assay using biotinylated histone H3K23me3 peptides or H3K36 peptides delivering acetyl-, di-methyl or tri-methyl groups at Lys 36. Unmodified histone H3(21-44) peptide was used as a control. No binding of PHD1 was observed to H3K36ac, H3K36me3 and H3K36me3 peptides (Figure 3-15), even when the bindings were repeated with excess amount of peptide to 25 µg (data not shown). As shall be shown later, increasing peptide amounts can reveal lower affinity interaction. This contrast with what we observed from peptide arrays. However, analysis of peptide array data of the PHD domains and the bromodomain suggests that binding to H3K36 peptides may result from non-specific binding. In contrast, we could detect robust binding of PHD1 domain to H3K23me3. Two H3K23me3-containing peptides were used. Strong binding

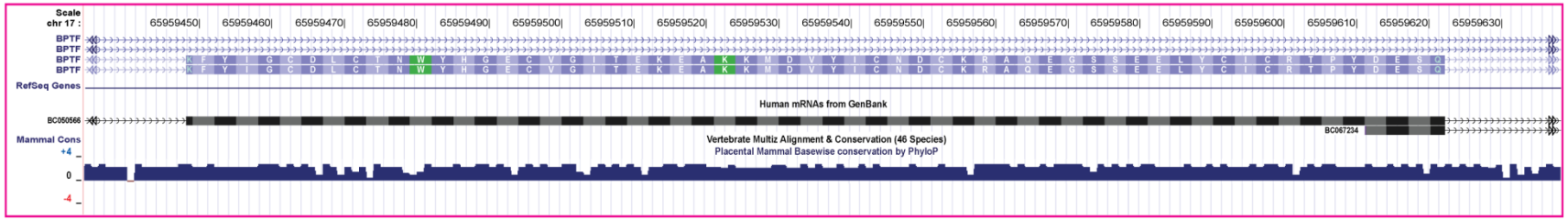
was observed when the H3(14-34)K23me3 peptide was used. In this peptide the modification is located in the middle of the peptide. Binding was also observed to a peptide, H3(21-44)K23me3, in which the modification is positioned near to N-terminal end of the peptide, but binding required higher peptide amounts of at least 25 μg.

In the previous studies, it has been known that human BPTF contains only two PHD finger domains: the N-terminal PHD domain and a single (PHD2) domain at the C-terminus. The PHD1 domain that is present in *Drosophila* NURF301 is not present in annotated mammalian BPTF transcripts. However, we have shown that PHD1 shows a unique modified histone-binding specificity and we wanted to determine if human BPTF may in fact contain a PHD1 domain. We examined the genomic region that includes BPTF using the UCSC genome browser (http://genome.ucsc.edu). This browser provides genome annotation data by graphically displaying assembly contigs and gaps, mRNA and expressed sequence tag (EST) alignments, multiple gene predictions, cross-species homologies, single nucleotide polymorphisms, sequence-tagged sites, radiation hybrid data and transposon repeats (Kent et al., 2002). Although a BPTF PHD1 domain was not annotated in the reference sequence genes (RefSeq), we identified the PHD1 domain in two N-terminal truncated short isoforms highlighted in pink (Figure 3-16 A) and in some mRNAs and ESTs. The amino acid and DNA sequence of the putative BPTF PHD1 were highly conserved in different species and resembled the *Drosophila* PHD1 sequence. The aromatic cage residues, Tyr 10, Tyr 17, Tyr 23 and Trp 32 (Y10, Y17, Y23 and W32) that are required for methylated lysine binding were present in the human BPTF PHD1 domain (Figure 3-16 B and C).

A


chr 17 (q24.2)
17p13.3 17p13.2 17p13.1 17p12 17p11.2 17q11.2 17q12 17q21.2 17q21.31 q21.32 17q21.33 17q22 17q23.2 23.3 24.1 17q24.2 17q24.3 17q25.1 17q25.3
Scale
chr 17 :
65830000 65840000 65850000 65860000 65870000 65880000 65890000 65900000 65910000 65920000 65930000 65940000 65950000 65960000 65970000 65980000
BPTF
BPTF
BPTF
BPTF
BPTF
RefSeq Genes
RefSeq Genes
Human mRNAs from GenBank
AK297281
BC037422
AB032251
AK293224
U05237
AY282495
BX648890
AK025776
AK022570
AF147404
AK027184
AK296256
AL833375
AK022177
AK022081
BC050566
BC067234
AK130939
BC039881
Vertebrate Multiz Alignment & Conservation (46 Species)
Placental Mammal Basewise conservation by PhyloP
Multiz Alignments of 46 Vertebrates

B


Scale
chr 17 :
65959450 65959460 65959470 65959480 65959490 65959500 65959510 65959520 65959530 65959540 65959550 65959560 65959570 65959580 65959590 65959600 65959610 65959620 65959630
BPTF
BPTF
BPTF
BPTF
RefSeq Genes
Human mRNAs from GenBank
BC050566
BC067234
Mammal Cons
+4
0
-4
Vertebrate Multiz Alignment & Conservation (46 Species)
Placental Mammal Basewise conservation by PhyloP

C


Human
Drosophila
Rhesus
Mouse
Dog
Elephant
Opossum
Chicken
X_tropicalis
Zebrafish

Figure 3-16. Identification of a human BPTF PHD1 domain. (A) Genomic view of human BPTF using the UCSC genome browser.. The genomic location of *Bptf* gene is indicated in a red box on the chromosome 17 map. The BPTF PHD1 domain was identified in two of the predicted BPTF isoforms (in pink box) and in some mRNAs (BC050566 and BC067234). (B) Enlarged image of BPTF PHD1 domain (in pink) and its amino acid sequence. (C) Protein and DNA sequence alignments of conserved PHD1 domains in most vertebrates. The amino acid residues in green boxes indicate critical residues forming aromatic cage of PHD1 domain, and the DNA bases in gray indicate wobble positions.

3.2.3.3 Bromodomain

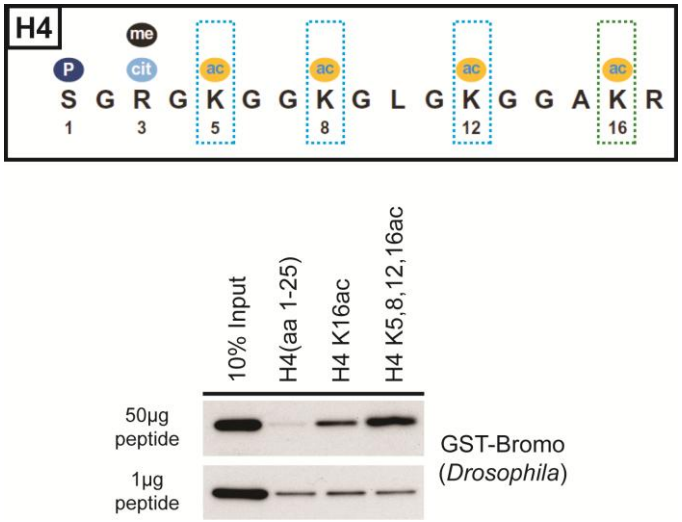


Figure 3-17. The bromodomain peptide pull down assays. The binding specificities of the *Drosophila* NURF301 bromodomain (GST conjugated) to histone modifications identified from peptide library arrays were validated by peptide pull down and western blot using rabbit anti-GST (HRP) antibody.

The interaction of the NURF301 bromodomain with histone modifications from the peptide array data was also confirmed by peptide pull down assay using biotinylated histone H4 peptides. Affinities of singly acetylated H4K16 and multiply acetylated H4K5, 8, 12 16 peptides were validated and unmodified H4(1-25) peptide used as a control. As observed in peptide arrays, peptide pull down showed interaction between NURF301 bromodomain and H4K16ac. However, binding to multiply-acetylated peptide (H4K5,8,12,16ac) was stronger (Figure 3-17). This was more apparent when excess peptide was used (50 µg). The result suggests that binding of the NURF301 bromodomain is higher to peptide that contains the multiple acetylations at histone H4 N-terminal tail than to singly acetylated H4K16.

3.2.4 Biacore

From our peptide array, peptide pull down and structure modelling analysis, we found that the binding of NURF301 PHD2 domain to H3K4me3 could be inhibited by phosphorylation of Thr 3 and enhanced by acetylation of Lys 9 and phosphorylation of Ser 10. To analyse the kinetics of these interactions, we performed biacore analysis. The injection PHD2 concentration was from 78 nM (-32x) to 40 μM (16x) for each peptide. The binding of PHD2 was monitored using surface plasmon resonance and the binding affinity of PHD2 to H3T3p/K4me3 decreased dramatically in comparison to H3K4me3 alone at different PHD2 concentrations (Figure 3-18). We explored models to calculate the dissociation constants (Kd) for binding of H3K4me3 and H3T3pK4me3 from the biacore plots, but the analysis indicated that the binding does not follow simple first order kinetics. Biacore analysis with H3K4me3/K9ac peptide was not successful (data not shown).

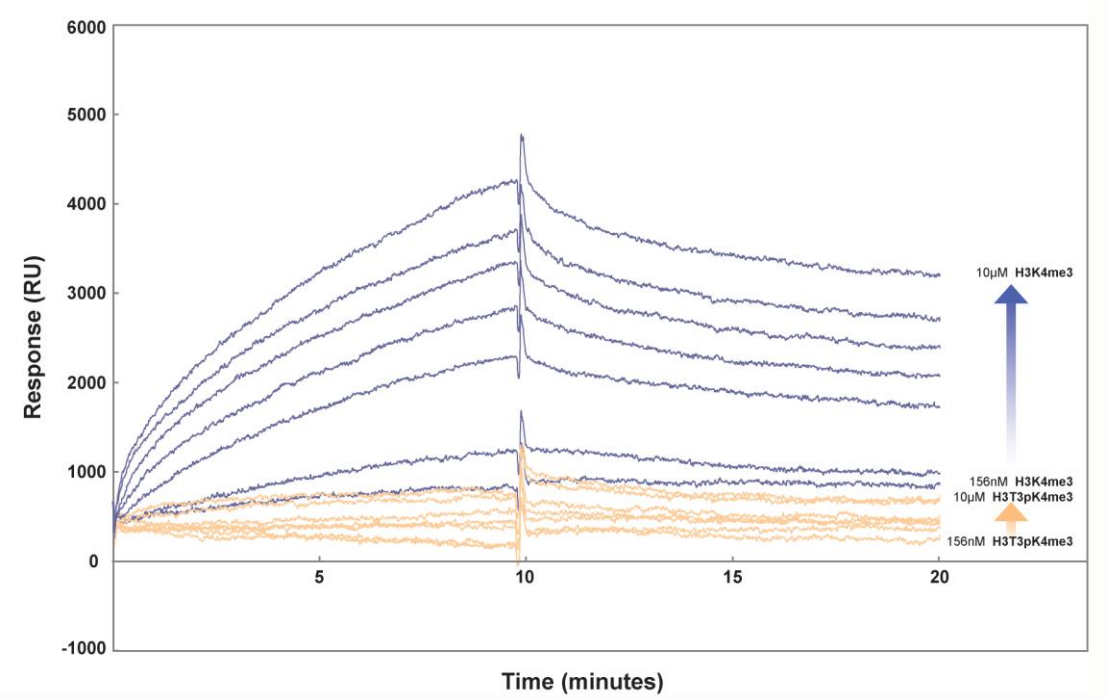


Figure 3-18. Surface plasmon resonance-based studies of the binding of the NURF301 PHD2 domain to H3(1-21) K4me3 and H3(1-21) T3pK4me3 peptides. Surface plasmon resonance-based binding and elution curves for the NURF301 PHD2 domain at different concentrations with biotin labelled H3(1-21) peptides as a function of K4 methylation and T3 phosphorylation states.

3.3 Conclusion

Taken together these data imply that PHD domains and bromodomain of NURF301 can make a broad range of interactions with histone modifications on the histone H3 and H4 N-terminal tails. Here we found the NURF301 PHD2 domain specifically binds to H3K4me3. However, the binding to this primary recognition site can also be influenced by interaction with flanking modifications. The flanking phosphorylation of H3T3 inhibits the PHD2 domain binding to H3K4me3 while acetylation of H3K9 enhances the binding. Moreover, the additional phosphorylation of H3S10 with H3K9ac can increase affinity of the PHD2 domain binding to H3K4me3. Together, these modifications provide rheostats to control NURF binding. Structural modelling analysis also suggested the electrostatic interactions between the PHD2 binding pocket and these modifications. The negatively charged surface of the PHD2 repulses the negatively charged phosphate group of H3T3, so PHD2 binding to H3K4me3 is inhibited. This was confirmed by Biacore assay which showed dramatic decrease of binding affinity of PHD2 domain to H3T3p/K4me3 in comparison to H3K4me3. In contrast, the neutralization of H3K9 by acetylation and the addition of a negative charge on H3S10 by phosphorylation (H3K9ac/S10p) can interact with positively charged surfaces on the PHD2 binding pocket, which enhances the binding capacity of the PHD2 domain to H3K4me3. We also revealed that the NURF301 PHD1 domain interacts with H3K23me3 specifically and identified that this PHD1 domain also exists in human BPTF and is highly conserved. The NURF301 PHD domain does not have any modified-histone binding specificity, but still showed binding preference to histone H4 N-terminal free sequences. Finally, we observed that NURF301 bromodomain binds to H4K16ac and this binding affinity can be increased by multiple acetylations at Lysine 5,8,12,16 (H4K5,8,12,16ac) of the histone H4 N-terminal tail. These data allow us to derive a model of the PHD domains and bromodomain with their specific

interactions of modifications on histone H3, H4 N-terminal tails of nucleosome.

CHAPTER 4. IMMUNOFLUORESCENCE MICROSCOPY TO LOCALIZE THE HISTONE MODIFICATIONS: H3K4me3, H3T3p, H3K9ac, H3S10p, H3K23me3 and NURF301 in *Drosophila in vivo* AND ChIP SEQUENCING OF NURF301 ISOFORMS

4.1 Introduction

4.1.1 *In vivo* localization of NURF and histone modifications

In work in the previous chapter we have used modified histone peptide arrays and *in vitro* histone peptide pull-down assays to define the binding specificities of each of the PHD domains and bromodomain of NURF301. However, it is important to verify whether these histone modifications are involved in targeting of NURF *in vivo*. In this chapter we examine the tissue distribution of NURF301 and histone modifications as well as co-localization of NURF and histone modification on chromatin.

Previous research has shown that *Drosophila* NURF, as a chromatin remodelling enzyme, generally localizes to the nucleus in cells. For instance, immunofluorescence on testes showed NURF301 is expressed within the primary spermatocytes although it is enriched on two of the three bivalents in mature primary spermatocytes (Kwon et al., 2009). Human BPTF has also been shown to localize to the nucleus although some expression is also detected in the cytoplasm (Jones et al., 2000). While BPTF appears to be expressed ubiquitously, the highest levels of expression have been observed in the testis (Jones et al., 2000).

The distribution of the histone marks we have shown to be bound by NURF301 (or influence

its binding) have been characterized to varying degrees. One of the best-studied is the histone H3K4me3 modification, which is found to be enriched at many transcriptionally active genes in yeast, *Drosophila* and mammals (Pokholok et al., 2005, Schuettengruber et al., 2009, Bernstein et al., 2005). On *Drosophila* polytene chromosomes, the majority of H3K4me3 staining colocalizes with Pol II and dSet1, which is known as a component of a conserved H3K4 trimethyltransferase (Ardehali et al., 2011).

Histone H3 phosphorylation is a well-known mark involved in regulation of gene expression and chromatin reorganization during mitosis. High levels of phosphorylation of histone H3 Thr 3 (H3T3p) are detected on mitotic chromosomes. Strong H3T3p staining commences at early prophase, spreads to pericentromeric chromatin during prometaphase but is fully reversed by late anaphase in mitotic cells (Polioudaki et al., 2004). Haspin is the major histone H3 Thr 3 kinase in mammalian cells, required for phosphorylation of Thr 3 at inner centromeres and normal metaphase chromosome alignment (Dai et al., 2005). Phosphorylation of histone H3 Ser 10 (H3S10p) is also detected at high levels during mitosis and meiosis (Prigent and Dimitrov, 2003), and structural alterations of chromatin are found to be related to dephosphorylation of H3S10p in *Drosophila* (Cai et al., 2008). However, H3S10p is also detected in interphase nuclei, albeit at lower levels, where it is associated with induction of gene expression (Mahadevan et al., 1991; Barratt et al., 1994; Sassone et al., 1999; Anest et al., 2003; Soloaga et al., 2003; Yamamoto et al., 2003). In *Drosophila*, higher levels of H3S10p are also detected on the male X-chromosome (Johansen and Johansen, 2006), where it is associated with structural changes and increased transcription from the male X-chromosome.

Acetylation of histone residues also correlates with transcription activation. Genome-wide profiling indicates that acetylation of histone H3 Lys 9 (H3K9ac) is corresponds with transcriptional activity and occurs predominantly at the 5' end of genes (Pokholok et al., 2005, Bernstein et al., 2005). This overlaps with the reported distributions of H3K4me3 and H3S10p (Schübeler et al., 2004; Roh et al., 2006). In particular the double modification H3K9acS10p mark occurs at promoters and is correlated with gene expression (Clayton et al., 2000; Kellner et al., 2012). Localisation of the histone H3K9ac mark *in vivo* in *Drosophila* has also been characterized using transgenic flies expressing GFP conjugated to an anti-H3K9ac antibody. These results reveal the distribution of H3K9ac in the nucleus during late embryogenesis and in euchromatic bands on polytene chromosome (Sato et al., 2013). The H3K9ac mark is established through the action of the acetyltransferase Gcn5, which when part of multisubunit complexes such as the Spt-Ada-Gcn5-acetyltransferase (SAGA) and ADA complexes (Grant et al., 1999), particularly acetylates histone H3 lysine 9 and lysine 14 (Qi et al., 2004). Loss of Gcn5 or other the Ada2b component of the SAGA complex results in the loss of H3K9ac on polytene chromosomes and in embryos, and cell cycle defects and apoptosis (Carré et al., 2005).

Of the modifications that were identified during our binding studies, the least characterized is the H3K23me3 modification. To a large extent studies of this modification have been limited by the absence of available antibodies. However, recently anti-H3K23me3 antibodies have been generated and used to show H3K23me3 on heterochromatic H3 from *Tetrahymena*, and that levels of H3K23me3 increase during meiosis in *Tetrahymena* micronuclei, *C. elegans*, and mice (Papazvan et al., 2013). To visualize the distribution of H3K23me3 in *Drosophila*, we obtained the anti-H3K23me3 antibody in a collaboration with the Taverna group in Johns Hopkins School of Medicine that generated this antibody.

4.1.2 Chromatin immunoprecipitation (ChIP) sequencing

The development of new techniques such as chromatin immunoprecipitation (ChIP) have been important in understanding how the dynamic properties of chromatin can be altered by modification of the DNA and histones. In particular, ChIP followed by hybridization to microarrays (ChIP-chip) or by high-throughput sequencing (ChIP-seq) has been widely used to characterize histone modifications as well as chromatin binding regulatory proteins. For example, most of the key modifications, such as H3K4me3, H3K9me3, H3K9ac and H3K27me3, have been mapped using S2 cells and adult flies in *Drosophila* (Kharchenko et al., 2011, Nègre et al., 2011, Yin et al., 2011).

In previous work in the laboratory, NURF was immunoprecipitated from testis and hemocytes, and the distribution of NURF was determined by semi-quantitative PCR analysis. This work identified that NURF binds to the transcriptional start site of the *fuzzy onions (fzo)* gene and that this was correlated with the distribution of the H3K4me3 and H4K16ac modifications (Kwon et al., 2009). In hemocytes, NURF301 was found to interact with the *Drosophila* Bcl6 homologue Ken and NURF-binding was shown to overlap Ken binding sites in hemocytes (Kwon et al., 2008). However, beyond these two studies, little is known about genome wide localization of NURF isoforms as it has been difficult to generate ChIP grade antibodies that target NURF301 and the only anti-NURF301 antibody available recognizes both NURF isoforms (Kwon and Badenhorst, unpublished data).

To identify differences in localization between the NURF301-A/B and NURF301-C isoforms, here we performed ChIP-Seq using CTAP tagged NURF301-A/B and NURF301-C transgenic flies. Hemocytes were isolated from 3rd instar larvae for ChIP as this provides a pure

population of cells and also allow us to collect male and female samples separately which would be useful to look at the sex-dependent differences in recruitment.

4.2 Results

4.2.1 H3K4me3 and NURF301 localisation *in vivo*

4.2.1.1 Detection of Histone H3K4me3 and NURF301 in embryos by immunostaining

Data from the PHD2 domain histone peptide arrays and peptide pull down assays suggested that the NURF301 PHD2 domain binds primarily to the H3K4me3 modification on histone tails, but this binding can be inhibited by flanking H3T3p or enhanced by flanking H3K9ac and H3S10p modifications. To understand the localisation of NURF and H3K4me3 and its interaction on chromatin *in vivo*, immunostaining of *Drosophila* embryo was performed using anti-H3K4me3, anti-NURF301(Kwon5), anti-H3T3p and anti-H3T3pK4me3 antibodies. The embryos are nearly transparent and can be easily visualized by microscopy and deliver representatives of each cell type at different developmental stages. These characters allow us to analyse the spatial and temporal distribution of histone modifications and NURF301 during development.

First, we examined the distribution of the single H3K4me3 modification across the time-course of embryonic development. Overnight collections of embryos (which includes stages from 0-18 hr) were analysed. Among the different stages of embryos, minimal antibody staining is detected during early embryonic stages, prior to cellularization. After cellularization (Fig 4-1, St 4), weak H3K4me3 could be observed, after which expression rose progressively during development with intense staining of tri-methylation at H3K4 detected in cells of late embryonic stages (Fig 4-1, St 7-9). The appearance of high levels of H3K4me3 at cellularization were striking as this corresponded with the large scale activation of zygotic transcription that occurs at this stage (Edgar and Schubiger, 1986). Significantly, anti-H3K4me3 staining was not only observed in interphase nuclei but also detected on the chromosomes in dividing cells (Figure 4-1, stage 7 dividing).

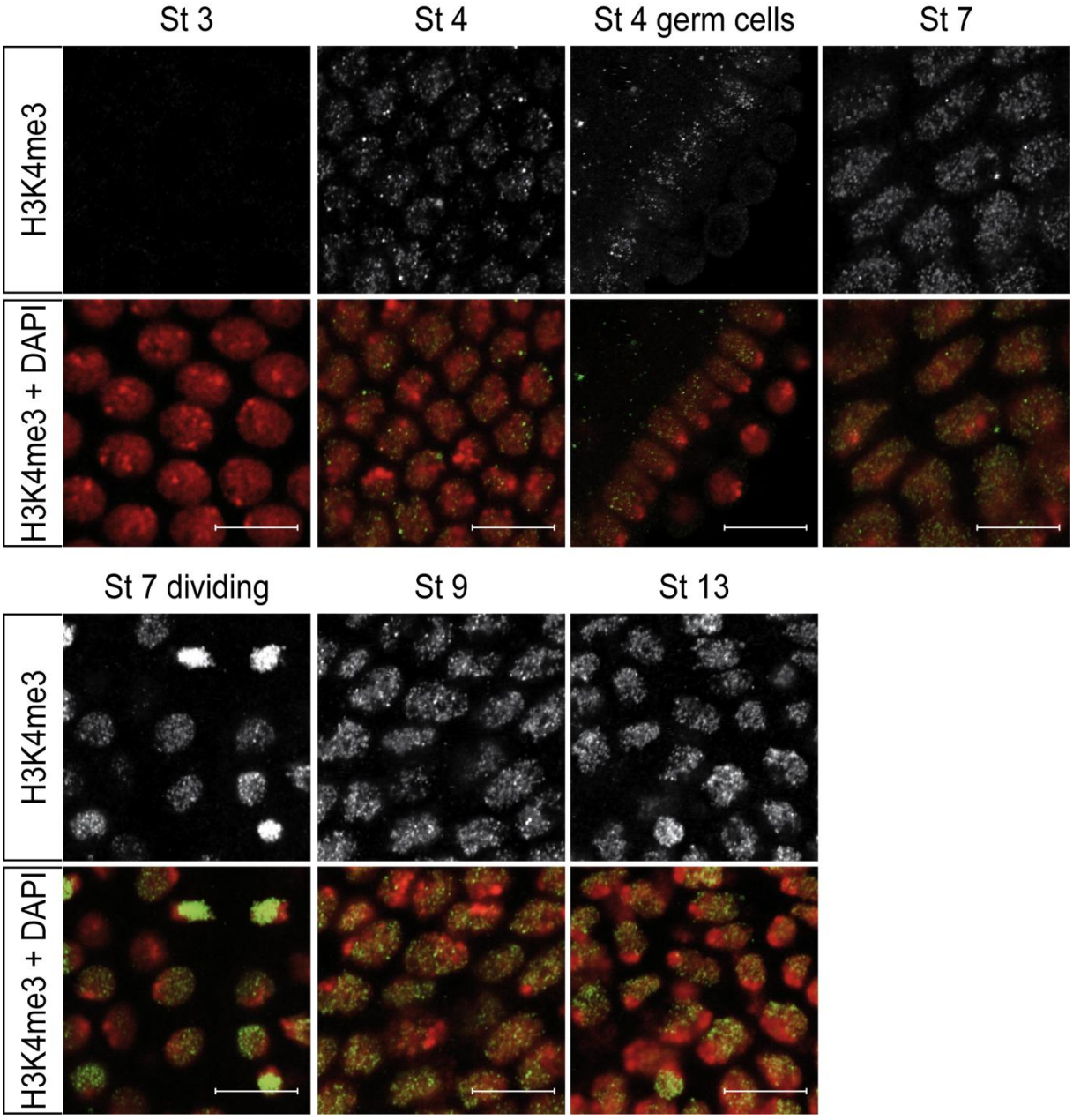


Figure 4-1. The histone H3K4me3 modification was detected after cellularizaton. *w*^{1118} embryos were fixed and immunostained using rabbit anti-H3K4me3 antibodies (green in merge), and DNA was visualized by DAPI staining (red in merge). The embryos at the different developmental stages (from stage 3 to stage 13) were observed. Scalebar represents 10μm.

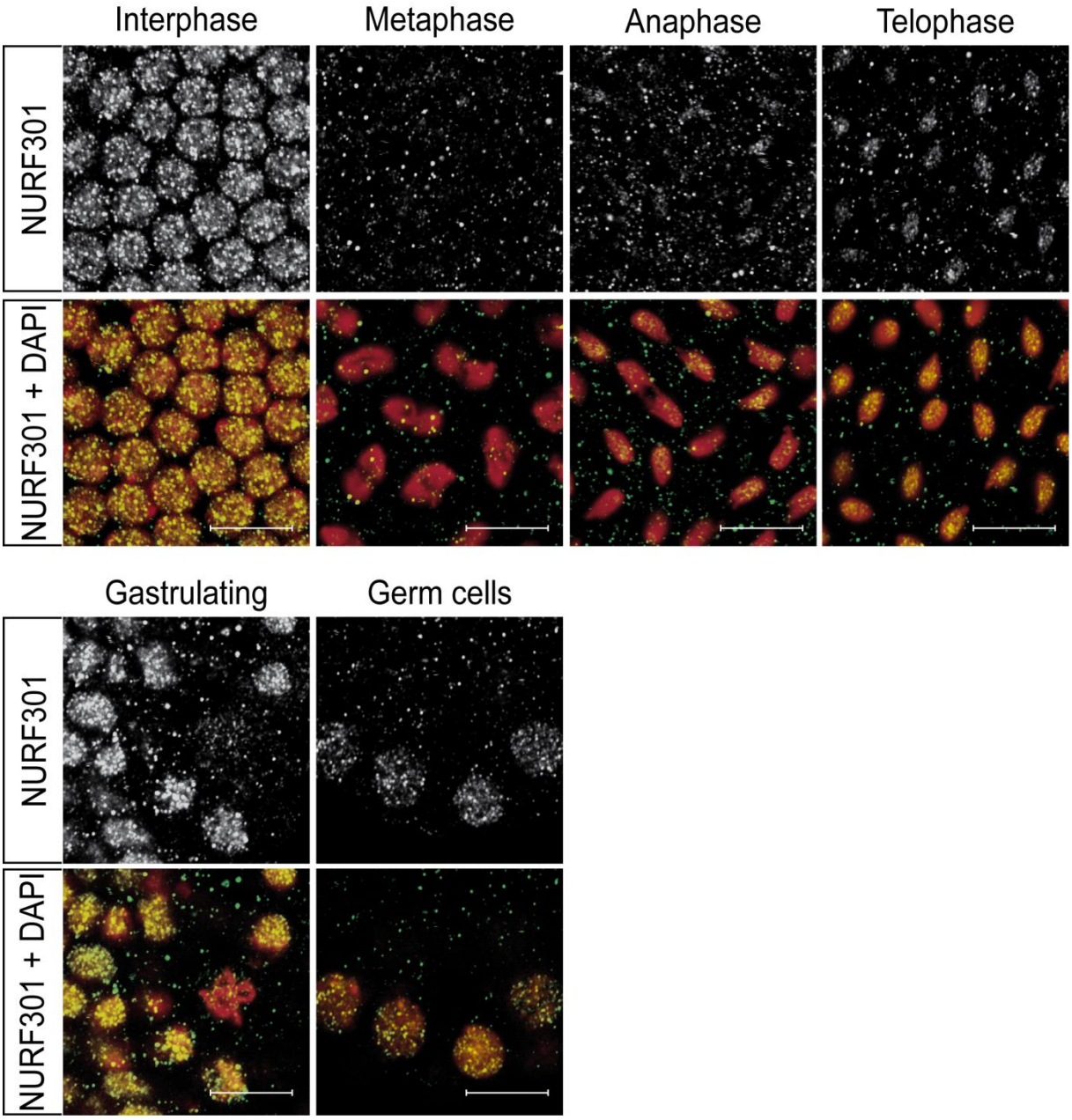


Figure 4-2. *Drosophila* NURF301 is delocalized from chromatin during mitosis. w^{1118} embryos were fixed and stained with rabbit anti-NURF301 antibodies (green in merge), and DNA was visualized by DAPI staining (red in merge). NURF301 is localized to chromatin during interphase, is not chromatin-bound during metaphase and anaphase, but starts to re-localize to chromatin during late telophase. Scalebar represents 10µm.

We next examined NURF301 distribution in embryos collected in the same manner to compare with the localisation of H3K4me3. Rabbit anti-NURF301 antibodies that had been generated in the laboratory (Kwon, unpublished data) were used. NURF301 staining was

detected in nuclei in embryos throughout development (data not shown). This corresponds with the high maternal contribution of *Nurf301* transcripts that has been observed previously (Badenhorst et al., 2002). However, NURF301 was only detected on chromatin during interphase (Figure 4-2, Interphase). During metaphase and anaphase, NURF301 was delocalized from chromosomes. NURF301 begins to associate with chromatin again during late telophase. NURF301 staining was detected in all cell types, however expression is higher in gastrulating cells, and the lowest levels were detected in primordial germ cells (Figure 4-2).

Consistent with the NURF301 PHD2 binding specificity data, we observed colocalization of NURF301 with H3K4me3 at least in interphase cells, but not in the dividing cells. The H3T3p mark, which we identified as an inhibitory marker of NURF PHD2 domain binding, has been reported to be elevated on mitotic chromatin (Polioudake et al. 2004; Dai et al. 2005). We therefore tested whether the delocalisation of NURF301 during mitosis was linked to H3T3p levels.

4.2.2 Identification of H3T3p as an inhibitor of the *Drosophila* NURF301 binding to H3K4me3 *in vivo*

4.2.2.1 Embryo immunostaining of Histone H3T3p, H3T3p/K4me3 and NURF301 in dividing cells

The distribution of the H3T3p mark in *Drosophila* embryos was examined using rabbit anti-H3T3p antibodies. Previous studies in human Hela cells had indicated high levels of H3T3p occur on metaphase chromosomes, declining substantially during anaphase and telophase, with H3T3p comparatively absent during interphase (Polioudake et al., 2004; Dai et al., 2005). We observed similar results using *Drosophila* embryos. The phosphorylation of H3T3 is barely detectable in interphase cells (Figure 4-3, Interphase). However, significant staining of H3T3p is detected on metaphase chromosomes, which then declines during anaphase and telophase as shown in figure 4-3.

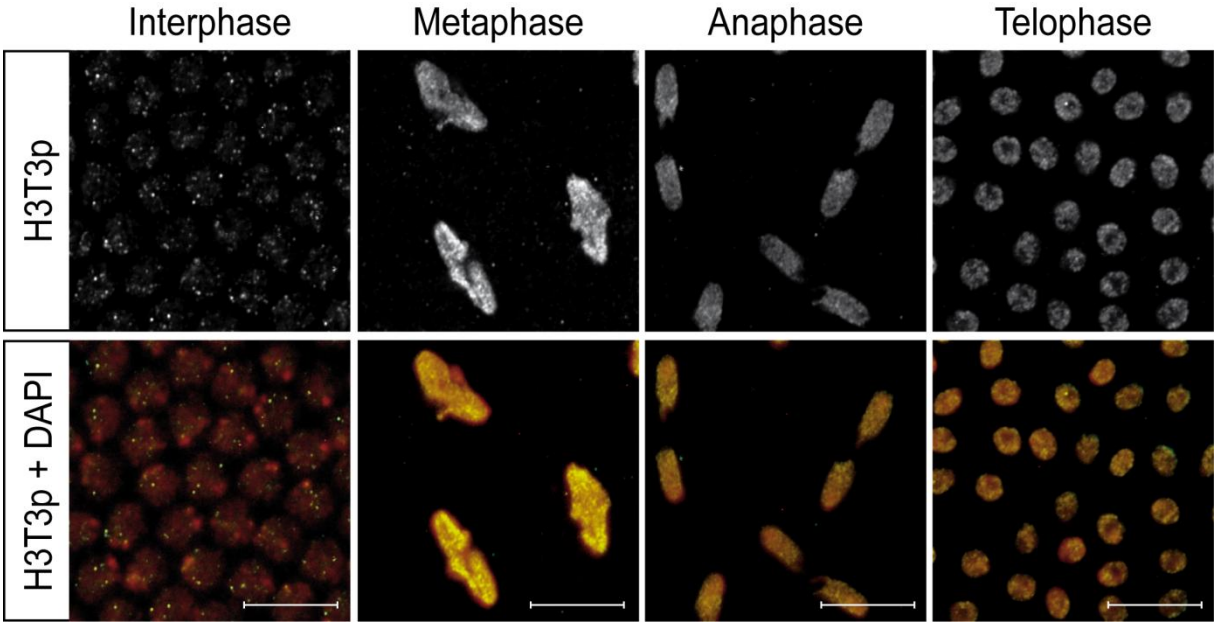


Figure 4-3. High levels of anti-H3T3p staining were detected on mitotic chromatin. w^{1118} embryos were fixed and stained with rabbit anti-H3T3p antibodies (green in merge), and DNA was visualized by DAPI staining (red in merge). Scalebar represents 10µm.

To investigate whether phosphorylation of H3T3 could affect the binding of NURF301 to H3K4me3 *in vivo*, we double-stained embryos using antibodies specific for *Drosophila* NURF301 and histone H3T3p/K4me3 respectively, to determine the extent of overlap. As noted previously, NURF301 is detected in interphase nuclei but is delocalized during metaphase and anaphase. NURF301 then starts to come back onto chromatin in late telophase (Figure 4-4 A). In contrast, H3T3p/K4me3 staining is detected significantly during metaphase and anaphase, but is absent during interphase (Figure 4-4 B). Double immunostaining with antibodies against H3T3p/K4me3 and NURF301 shows no colocalization during any stage of the cell cycle (Figure 4-4 C).

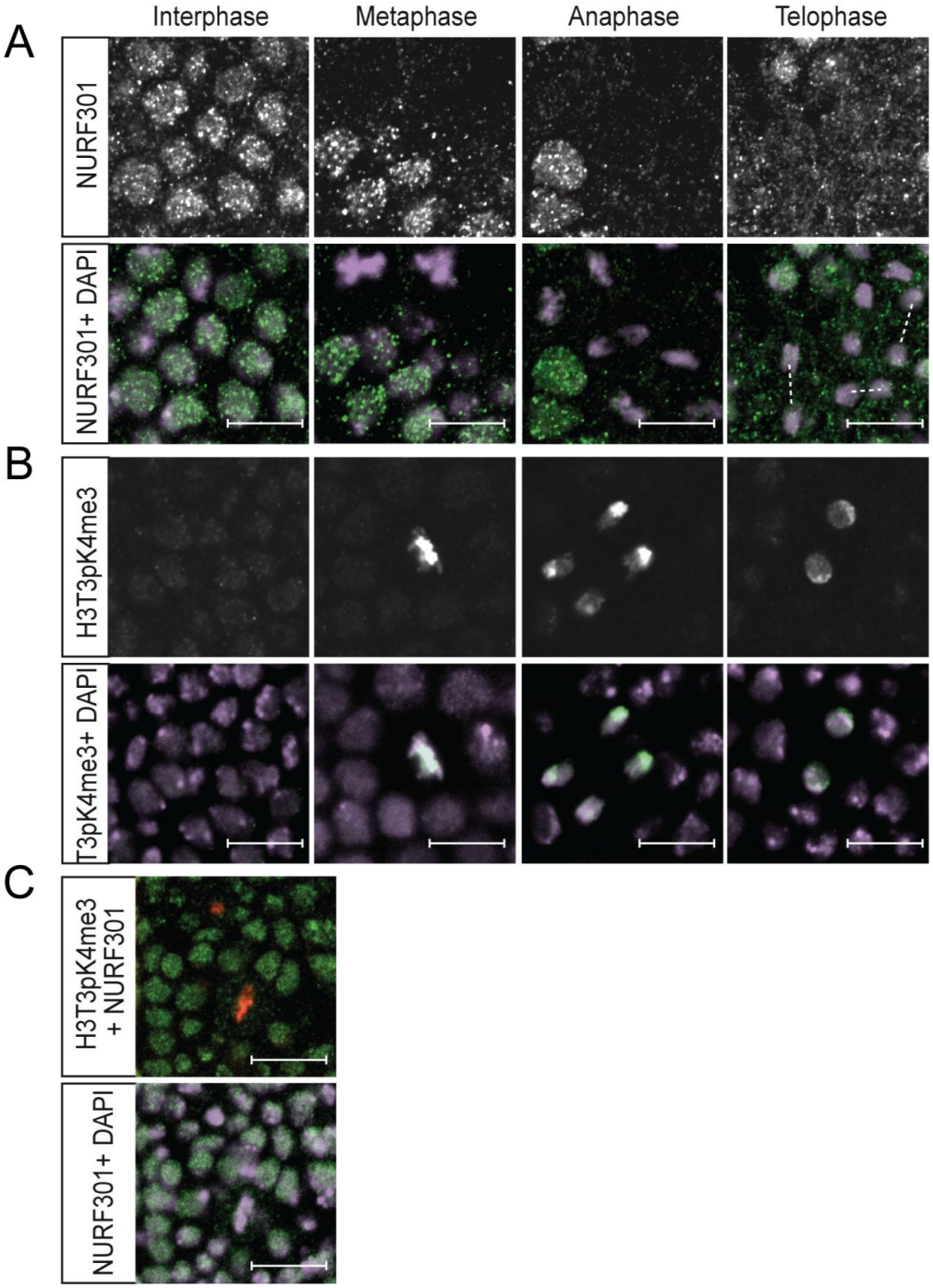


Figure 4-4. Distribution of NURF301 relative to H3T3p/K4me3 in *Drosophila* embryos at different stages of mitosis. (A) w^{1118} embryos were fixed and stained with rabbit anti-NURF301 antibodies (green in merge) and DAPI (purple in merge), (B) with rabbit anti-phospho-methyl-histone H3 (H3T3p/K4me3) antibodies (green in merge) and DAPI (purple in merge). (C) Cells sequentially double-stained using anti-NURF301 antibodies (green in merge) and anti-phospho-methyl-histone H3 antibodies (H3T3p/K4me3, red in merge). DNA was visualized by DAPI staining (purple). Scalebar represents 10 μm.

4.2.2.2 Visualization of Histone H3T3p, H3T3p/K4me3 and NURF301 in neuroblast cells by immunostaining

Immunostaining of *Drosophila* embryonic cells with anti-NURF301, anti-H3K4me3, anti-H3T3p and anti-H3T3p/K4me3 antibodies at different mitotic stages showed that binding of NURF301 could potentially be regulated by the histone H3T3p modification in chromatin *in vivo*. To confirm whether this also occurred at other stages of development and to profile localisation on chromatin at higher resolution, *Drosophila* third instar larval neuroblasts were immunostained using the same antibodies. *Drosophila* larval neuroblasts are well-established model for visualizing mitotic chromasomes (Gatti et al., 1994). They divide asymmetrically and following mitosis generates two cells, a larger daughter cell, which retains neuroblast identity, and smaller ganglion mother cell (GMC) that generates neurons (Doe and Bowerman, 2001). The larger size of neuroblasts allows mitotic figures to be observed easily at high resolution.

We examined the distributions of NURF301 and the H3T3p/K4me3 modification on double-immunostained interphase nuclei and mitotic chromosomes from larval neuroblasts. As expected, NURF301 was detected during interphase nuclei but was delocalized during metaphase, while H3T3p/K4me3 was detected at high levels on metaphase choromosomes (Figure 4-5 A and B). In interphase nuclei, lower levels of H3T3p/K4me3 were observed, distributed in a punctate manner. Double labelling with anti-NURF301 antibodies suggested no overlap with H3T3p/K4me3 foci in interphase neuroblasts. These data suggested that two populations of H3T3p exist. The first is the broadly expressed modification that occurs during mitosis. The second is a more restricted modification during interphase. This suggested that the H3T3p mark could also play a role on interphase chromatin in ejecting NURF at specific loci.

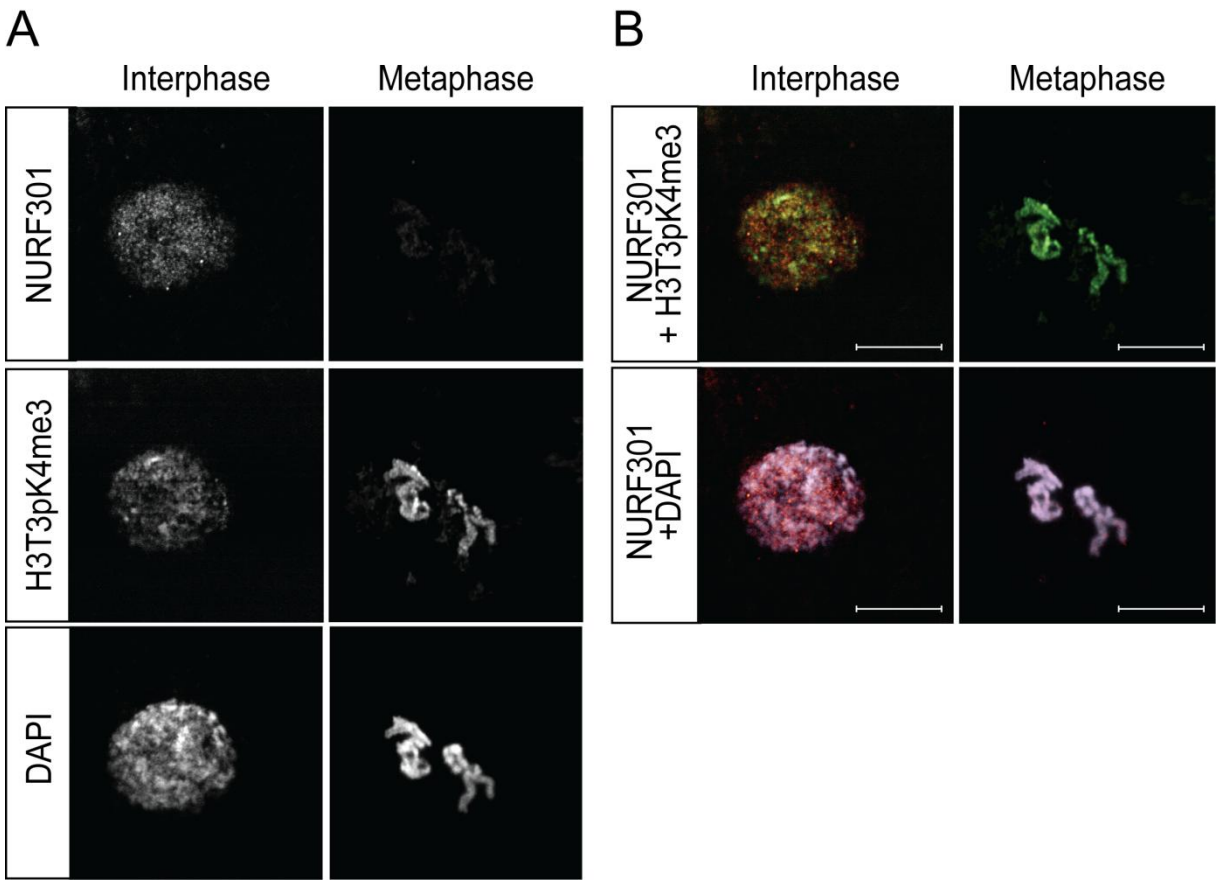


Figure 4-5. Distributions of NURF301 and H3T3p/K4me3 on interphase nuclei and mitotic chromosomes of *Drosophila* larval neuroblasts. (A) Double immunostaining images of NURF301, H3T3p/K4me3 and DAPI in interphase and metaphase neuroblasts from w^{1118} shown in separate panels in black and white. (B) Merged double immunostaining images of NURF301 (red)/H3T3p/K4me3 (green) and NURF301 (red)/DAPI(purple). Scalebar represents 10 μm.

4.2.2.3 Visualization of Histone H3T3p, H3T3p/K4me3 and NURF301 in polytene chromosomes by immunostaining

To resolve loci at which H3T3p antagonizes NURF-binding in interphase chromatin, *Drosophila* third instar larval polytene chromosomes were immunostained using the same antibodies. As the polytene chromosomes are non-mitotic interphase chromosomes consisting of thousands of copies of the DNA strands and display specific banding patterns, we expected to understand and visualise the more detailed distributions of NURF and histone modifications on chromosomes *in vivo*.

We first used the conventional polytene chromosome squash technique in which highly acidic fixation conditions are employed (Badenhorst et al., 2002). However, the resultant antibody staining was poor and we were unable to detect good single antibody staining with either anti-NURF301, anti-H3T3p or anti-H3T3pK4me3 antibodies (data not shown). We reasoned that the highly acidic (45% acetic acid) step that is used to spread polytene chromosomes in the conventional squash procedure strips chromatin of NURF301 and the histone modifications. To address this we adopted an alternative squash procedure that has been shown to result in reliable antibody staining of the H3S10p histone phosphorylation mark (Cai et al., 2008). This acid-free procedure uses glycerol instead of acetic acid to spread chromosomes after fixation (DiMario et al., 2006) and has been used for all subsequent polytene chromosome immunostaining.

The immunostainings of NURF301, H3T3p and H3T3p/K4me3 were performed respectively with the acid-free protocol. We detected significant staining of NURF301 on the non-mitotic interphase polytene chromosomes. NURF301 was predominantly localized to inter-bands

which are in the decondensed state on chromosomes, or to the interface between bands and interbands (Figure 4-6 A, green bands in the merge). However, there was no significant staining of H3T3p detected on polytene chromosomes (Figure 4-6 B). Low levels of background staining were detected using anti-H3T3p/K4me3 antibodies (Figure 4-6 C). This data was consistent with the embryo immunostaining results above, where significant chromatin localisation of NURF301 was observed during interphase when H3T3p and H3T3p/K4me3 staining was absent. However, it did not allow us to address the interphase function(s) of the H3T3p mark as it is absent from polytenized chromosomes.

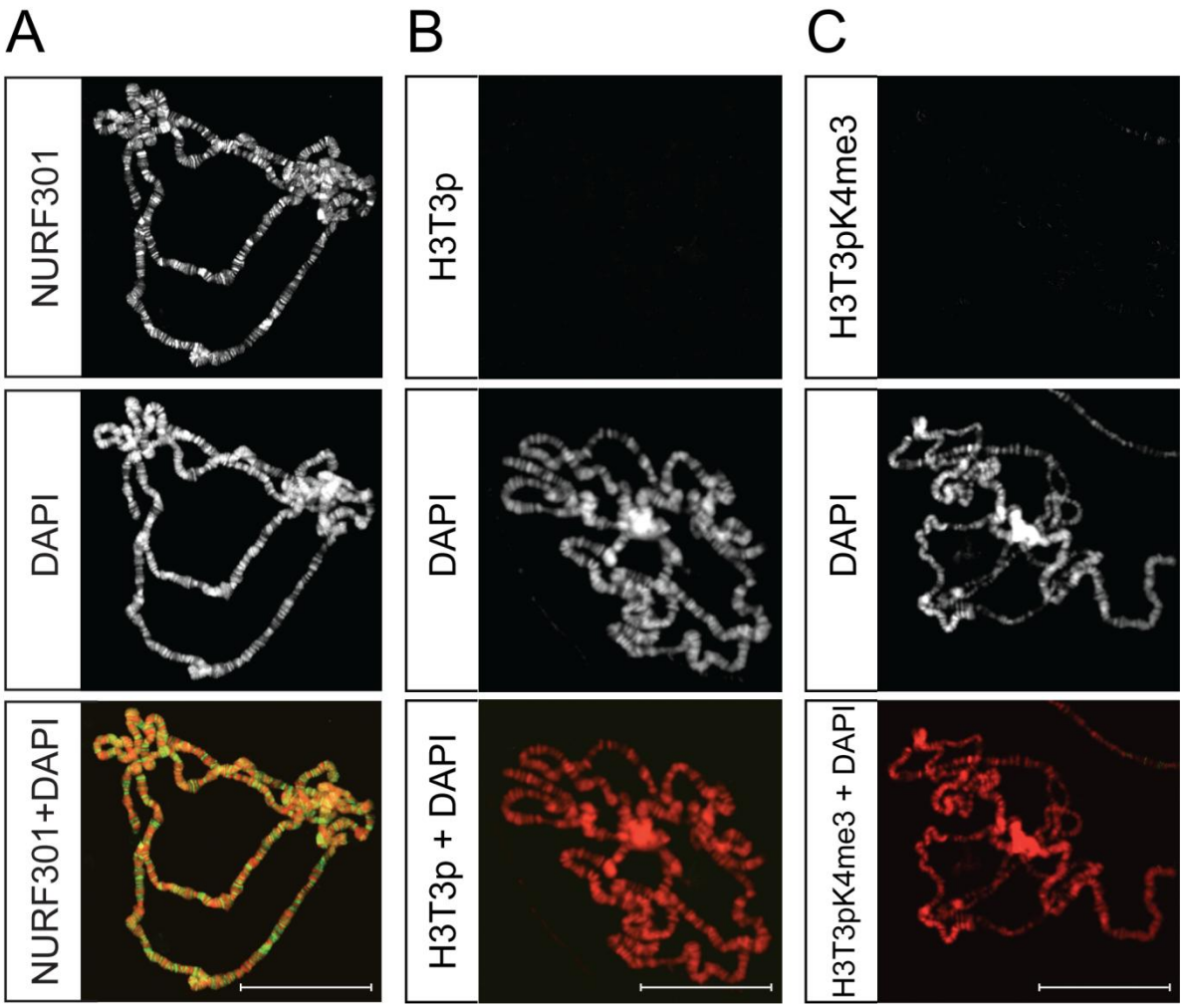


Figure 4-6. Distribution of NURF301 relative to H3T3p and H3T3p/K4me3 in *Drosophila* non-mitotic interphase polytene chromosomes. For immunostaining, w^{1118} third instar larvae were used. The polytene chromosomes were immunostained using (A) rabbit anti-NURF301, (B) anti-H3T3p and (C) anti-phospho-methyl-histone H3 (H3T3p/K4me3) antibodies (green). DNA was visualized by DAPI staining (red). Scalebar represents 50 μm.

Overall, these immunostaining data support the idea that the flanking phosphorylation of H3T3 has an inhibitory effect on the binding of NURF301 to H3K4me3 during mitosis. The H3T3p mark is established by the atypical kinase Haspin (Dai et al., 2006). In a next set of experiments we attempted to generate mutants in the *Drosophila Haspin* gene and examine NURF localization to chromatin during mitosis in *Haspin* mutants. We obtained fly lines containing transposon insertions in *Haspin* and attempted to generate imprecise excision mutants from these. However, this was not successful. In addition, we also tried to generate

small deletions in the *Haspin* locus by X-ray mutagenesis. These experiments also failed to generate a *Haspin* mutant line. *Haspin* is an example of a rare type of gene in *Drosophila* that is expressed even though it is embedded in pericentric heterochromatin (location 2RHet:3,199,376..3,234,607). Given its location near the centromere of the second chromosome we postulate that deletion of *Haspin* may disrupt the centromere making it difficult to recover mutations through conventional excision techniques.

To overcome this we also attempted to knock-down *Haspin* expression using inducible RNAi (siRNA). The *Drosophila* Transgenic RNAi Project at Harvard Medical School (TRiP) has generated transgenic fly strains that express short hairpin interfering RNAs (siRNAs) under the control of the GAL4 UAS sequence. By crossing these strains to GAL4 driver fly strains that express the GAL4 transcription factor in specific tissues, tissue specific knockdown can be achieved. The TRiP project has generated TRiP lines that target most *Drosophila* genes (Ni et al., 2008). We used a strain that is designed to target *Haspin* (y^1 sc^* v^1; *P{TRiP.HMS01468}attP2*) to attempt to knock-down *Haspin* levels. As a first test we used the hemocyte-specific GAL4 driver *Hemolectin-GAL4* to drive expression of *Haspin* siRNAs in blood cells to test whether this siRNA strain could knockdown *Haspin* expression. No antibodies against *Drosophila* Haspin exist but we could assay knockdown by looking for effects on H3T3p staining of hemocytes. However, as shown in figure 4-7, the *Haspin* siRNA strain was unable to significantly affect H3T3p levels, precluding the use of this strain for functional analysis of the effect of H3T3p on NURF301 localization.

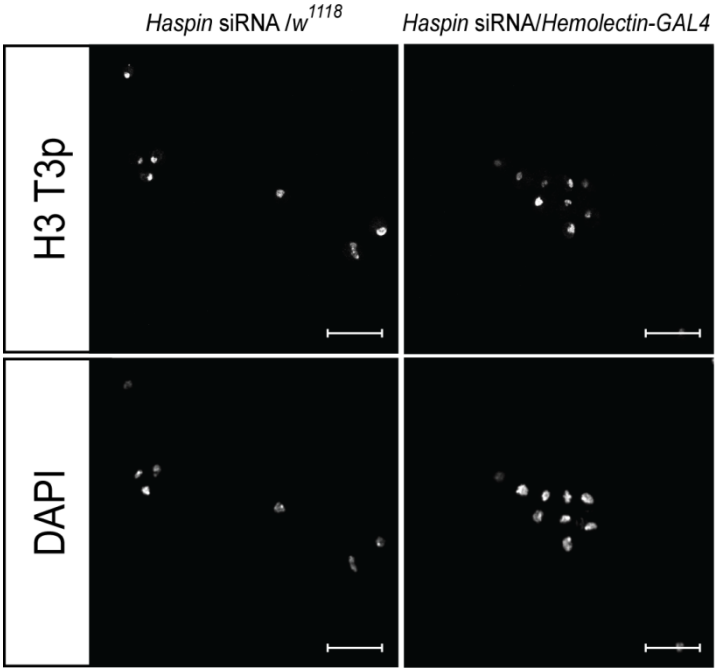


Figure 4-7. Expression of *Haspin* siRNAs in blood cells. A *Haspin* siRNA strain was crossed to w^{1118} as control, or *Haspin* siRNA expression was driven in blood cells by crossed it to *Hemolectin-GAL4* strain. Staining using anti-H3T3p antibodies showed that the Haspin siRNA strain was unable to block H3T3p. Scalebar represents 20 μm.

4.2.3 Identification of H3K9ac and H3K9ac/S10p as enhancing markers of *Drosophila* NURF binding to H3K4me3 *in vivo*

4.2.3.1 Visualization of Histone H3K9ac, H3K9ac/S10p and NURF in polytene chromosomes by immunostaining

Results obtained using peptide arrays and peptide pull down assays also identified two marks that enhanced NURF301 PHD2 domain binding to H3K4me3, the histone marks acetylated H3 Lys 9 (H3K9ac) and phosphorylated H3 Ser 10 (H3S10p). In a similar manner to that described for the H3T3p mark, we examined whether these modifications affect NURF localisation to chromatin using immunofluorescence microscopy of third instar larval polytene chromosomes. We have already demonstrated that the optimized acid-free polytene squash procedure preserves NURF301 staining of polytene chromosomes (Section 4.2.2.2). We first confirmed if this squash procedure was effective for H3K9ac and H3S10p. We examined the distribution of these histone modifications either alone or in combination on polytene chromosome using antibodies against H3K9ac, H3S10p and H3K9ac/S10p, before we analysed their correlation with NURF. The H3K9ac mark was detected at numerous loci on wild-type polytene chromosomes, and localized to both interbands and DAPI stained bands (Figure 4-8 A). Anti-H3S10p antibody staining displayed clearer banding patterns, mostly localised to interbands (Figure 4-8 B). Antibodies that detect the combined phospho-acetyl H3K9ac/S10p modification showed restricted staining compared to H3K9ac or H3S10p, but which was localised to interbands (Figure 4-8 C).

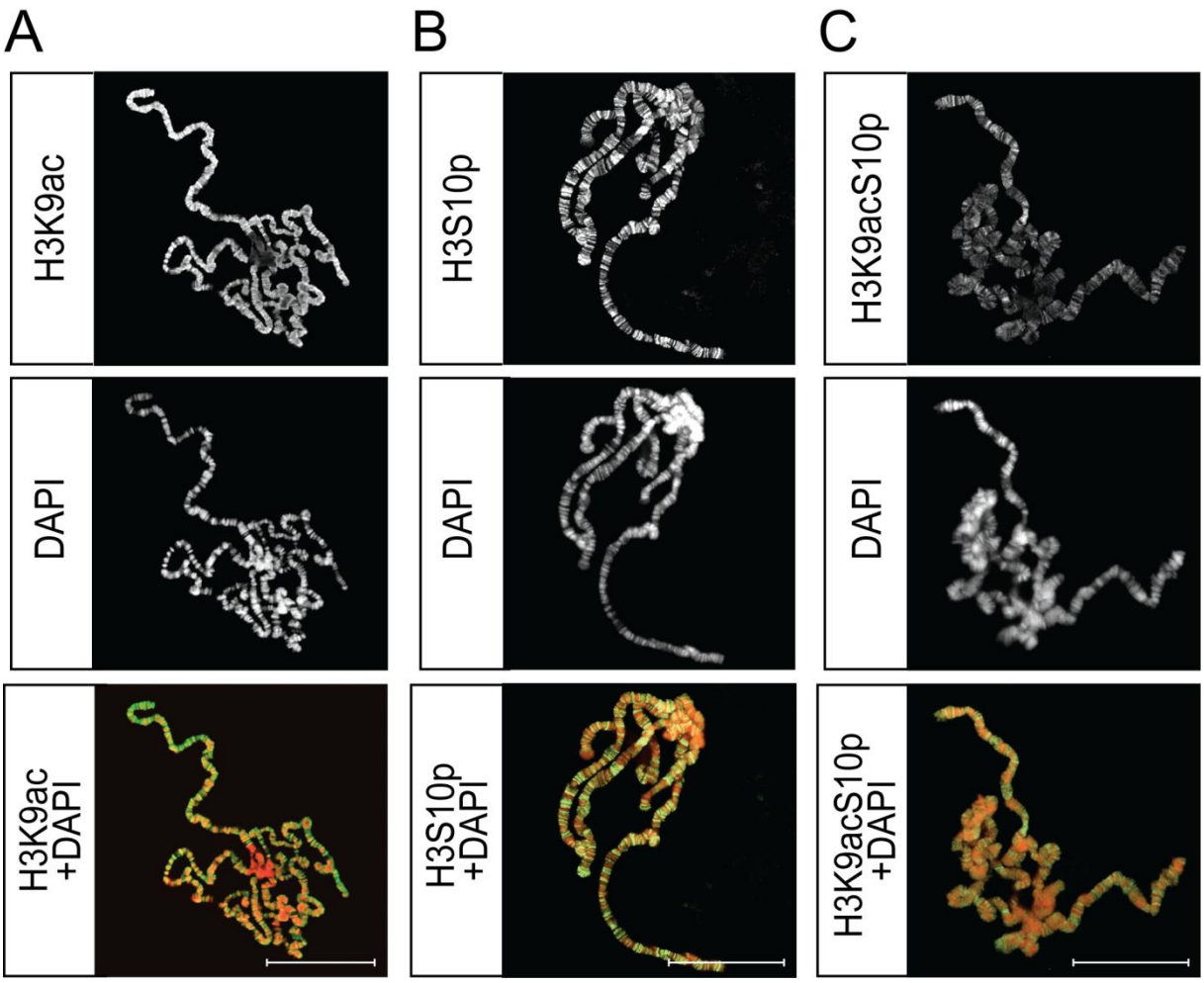


Figure 4-8. The distribution of the histone H3K9ac, H3S10p and the combined phospho-acetyl H3K9ac/S10p mark on wild-type *w*^{1118} *Drosophila* interphase polytene chromosomes. Polytene chromosomes were immunostained with (A) rabbit anti-H3K9ac, (B) anti-H3S10p and (C) anti-H3K9ac/S10p antibodies (green in merge). DNA was visualized by DAPI staining (red). Scalebar represents 50 μm.

4.2.3.2 Visualization of NURF301 in histone acetyl transferase Gcn5 knock-out condition by polytene chromosome immunostaining

If histone H3K9ac modification is required for NURF binding, we would expect to see that NURF301 and H3K9ac colocalize on chromatin. Conversely, we would also predict that mutation of the histone acetyl transferase Gcn5, which acetylates histone H3 Lys 9, would alter NURF301 localization. To test these predictions we looked at the relative distribution of NURF and acetylated histone H3 K9 on polytene chromosomes of wild-type (w^{1118}) and Gcn5 delete strains. We generated animals lacking Gcn5 by crossing *Df(3L)sex204,* a deficiency that removes *Gcn5* and the flanking *tra1* locus, to the EMS point mutant $Gcn5^{E333st}$ which has previously been shown to be a *Gcn5* null allele (Carre et al., 2005). Polytene chromosomes from mutant and wild-type third instar larvae were double stained using antibodies against NURF301 and H3K9ac. The same chromosome arm was then selected and observed at higher magnification to map the relative distribution of NURF301 and H3K9ac.

We observed that NURF was mainly localized to interbands or the interface between bands (Figure 4-9A). H3K9ac staining was more generally distributed, but we were able to detect bands at which NURF301 and H3K9ac were co-localized on wild-type polytene chromosome (Figure 4-9A, merge). Clearly H3K9ac is not the sole determinant of NURF301 recruitment as bands were observed were H3K9ac positive but which showed no NURF301 binding (green bands in the merge in Figure 4-9A). This is not surprising given that we postulate that a combination of H3K4me3 and the H3K9ac mark determines recruitment. Significantly, we observed a more restricted number of bands that showed NURF301 staining but did not display high levels of H3K9ac (green bands in the merge in Figure 4-9A). However, the antibody we used to detect NURF301 recognizes both the full-length NURF301-A/B isoform

and the C-terminally truncated NURF301-C isoform, which does not contain PHD1, PHD2 and the bromodomain. Bands that stain with NURF301 antibodies but do not show H3K9ac may reflect staining of the NURF301-C isoform.

As predicted, H3K9ac was barely detectable on *dGcn5* mutant polytene chromosomes compared to wild-type chromosomes (Figure 4-9 B). Staining of NURF301 on *dGcn5* mutant polytene chromosomes was not prevented completely but did show reduced binding to a restricted set of bands(Figure 4-9 B, merge). We postulate that the sites of residual NURF301 staining on *dGcn5* mutant polytene chromosomes correspond to the NURF301-C isoform, which does not contain the C-terminal PHD2 domain. We conclude that NURF interacts with histone H3K9ac on polytene chromosomes and that binding is modified by loss of H3K9ac.

As the antibody we used against NURF301 recognizes both the full-length NURF301-A/B isoform and the C-terminally truncated NURF301-C isoform which does not contain PHD1, PHD2 and bromodomain, it was possible that the residual NURF staining we observed in *dGcn5* mutants may have come from staining of the NURF301-C isoform. To discriminate between staining of the NURF301-A/B and NURF301-C isoforms *in vivo*, we generated epitope-tagged NURF301-A/B and NURF301-C isoforms using BAC recombineering. These experiments will be discussed later in chapter 4.2.8.1.1.

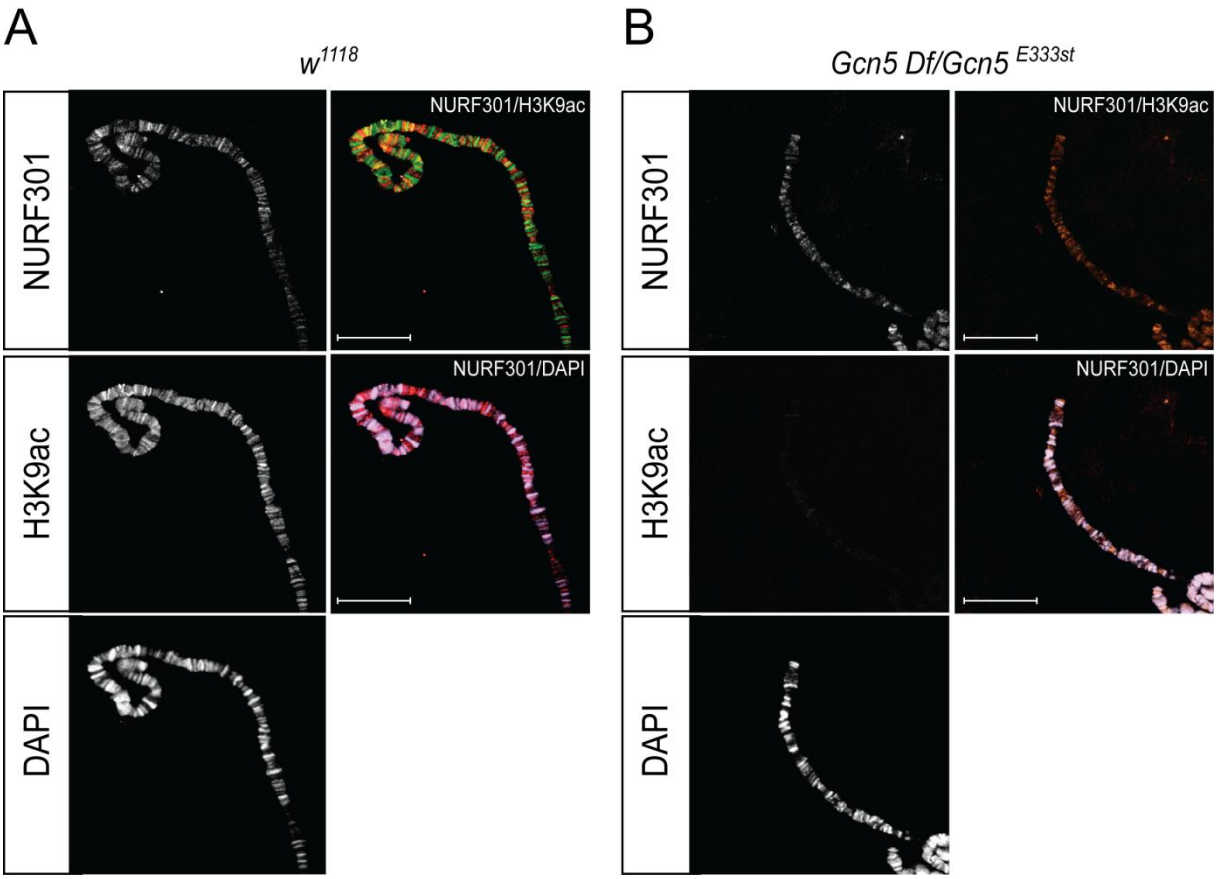


Figure 4-9. Loss of Gcn5 reduces the binding of NURF to some chromatin sites. Double immunostaining using antibodies against NURF301 and H3K9ac of (A) wild-type (w^{1118}) and (B) *dGcn5* (*Gcn5 Df* crossed to $Gcn5^{E333st}$) mutant polytene chromosome arms. In merged double immunostaining images, NURF301 (red), H3K9ac (green) and DAPI (purple). Scalebar represents 20 μm.

4.2.4 Identification of H3K23me3 *in vivo*

4.2.4.1 Immunostaining of embryos, testes and ovaries.

Data from the NURF301 PHD1 domain histone peptide arrays and peptide pull down assays suggested that the PHD1 domain interacts primarily with the H3K23me3 modification. The H3K23me3 modification is a relatively novel modification having been first described in Papazvan et al. (2013). Before characterizing the interaction between NURF301 and H3K23me3 *in vivo*, we needed to test whether the histone H3K23me3 modification occurs in *Drosophila*. We tested anti- H3K23me3 staining using a range of cell types and developmental stages. We first performed immunostaining of wild-type *Drosophila* embryos collected as described in section 4.2.1.1. However, no anti-H3K23me3 antibody staining could be detected in embryos at any stages of development or the cell cycle including dividing cells (Figure 4-10).

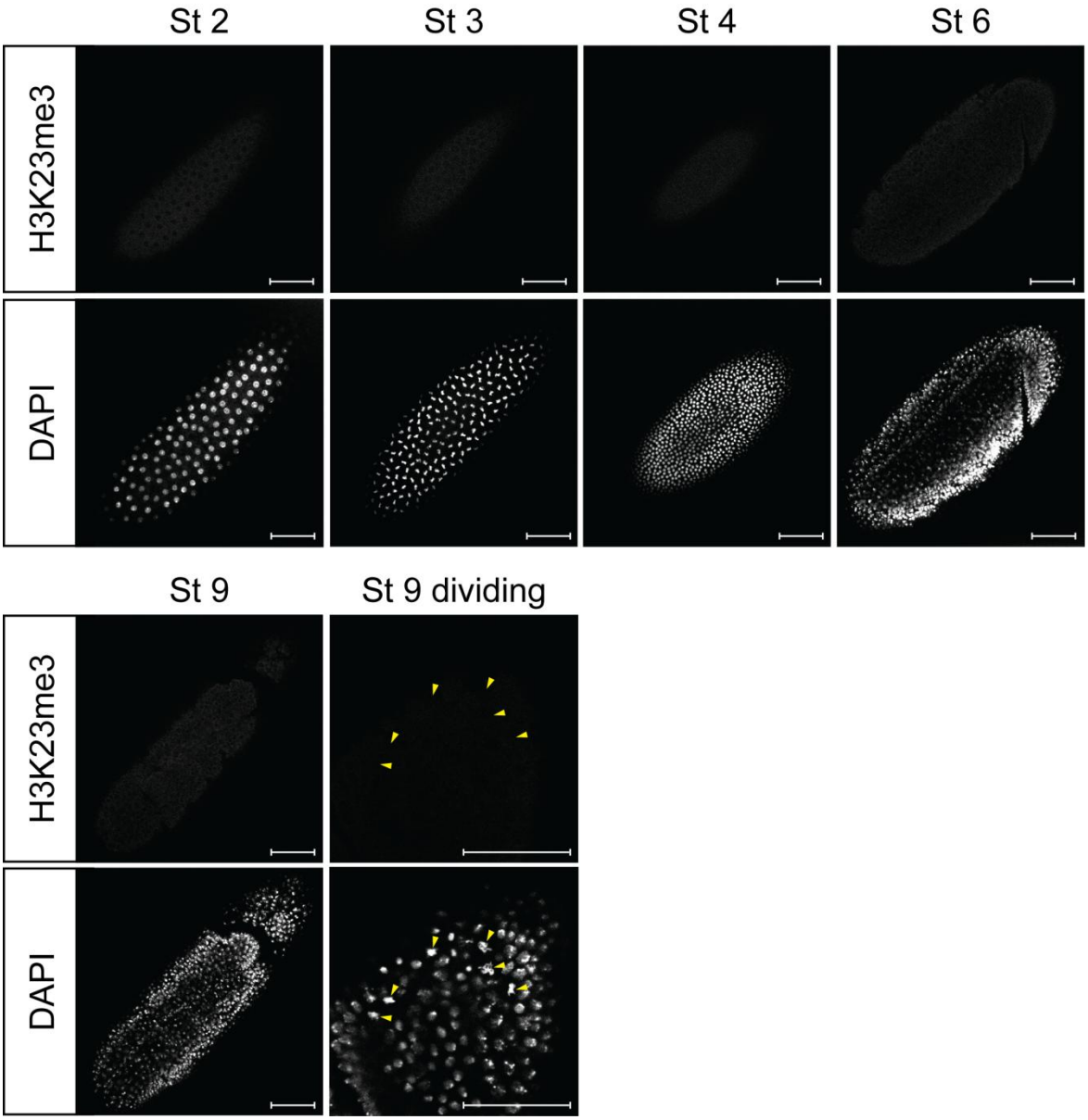


Figure 4-10. Immunostaining of wild-type(w^{1118}) *Drosophila* embryos using rabbit anti-H3K23me3 antibodies. No H3K23me3 staining is detected in staged embryos (from stage 2 to stage 9) and dividing cells of embryo in stage 9 indicated by yellow arrowheads. DNA was visualized by DAPI staining. Scalebar represents 50 μm.

As the H3K23me3 modification has been postulated to play a role in meiosis (Papazvan et al., 2013), we also tested whether we could detect H3K23me3 in the adult testis or ovary. However, immunostaining of *Drosophila* testes and ovaries using anti-H3K23me3 antibodies failed to exhibit significant staining in both testes and ovaries (Figure 4-11 A and B). Use of the antibody at higher concentrations or with modifications to improve permeabilization of

the testes and ovaries (using higher concentrations of detergent treatment or squashing) did not alter these results (data not shown).

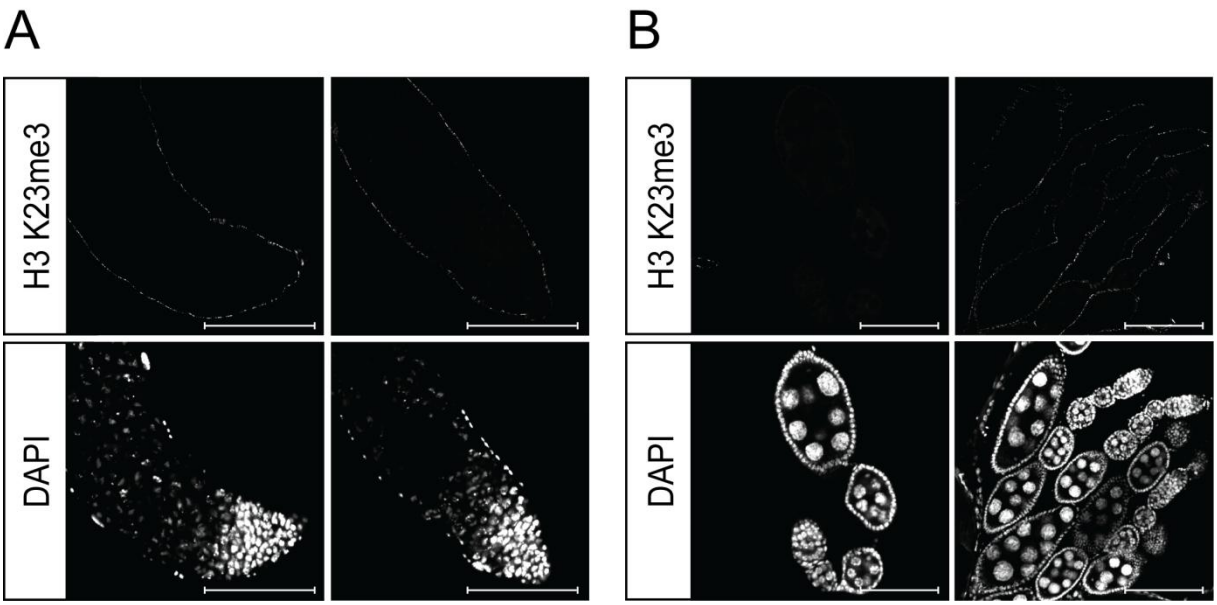


Figure 4-11. Immunostaining of wild-type(w^{1118}) *Drosophila* male testes and female ovaries using rabbit anti-H3K23me3 antibodies. No anti- H3K23me3 staining was detected in (A) male testes or (B) female ovaries. DNA was visualized by DAPI staining. Scalebar represents 100 μm.

4.2.4.2 Validation of the anti-H3K23me3 antibody by peptide dot blot analysis

These data suggest that the H3K23me3 modification does not exist in flies. However, before being able to conclude this, it was important to confirm whether the antibody used was effective at detecting H3K23me3. To validate whether the anti-H3K23me3 antibody we obtained indeed recognizes the histone H3K23me3 modification, we performed peptide dot blot using biotinylated modified H3K23me3 peptides. These peptides all contain the H3K23me3 mark, but vary in peptide length and the position of the H3K23me3 mark relative to the N-terminus of the peptide. An unmodified histone H3 peptide was used as a control. Significantly, although the anti-H3K23me3 antibody was able to detect some peptides that contain H3K23me3, the antibodies only reacted against peptides that commenced after H3

residue 18 (Figure 4-12 B and C). No reaction was detected against a peptide that contained H3K23me3 but which initiated after H3 residue 14 (Figure 4-12 D). This suggests that the anti-H3K23me3 antibodies may not bind in the context of the full-length H3 tail. It has been shown previously that proteolytic cleavage of the H3 tail occurs in a number of systems, with cleavage (called tail-clipping) occurring after Ala 21 (Duncan et al., 2008; Santos-Rosa et al., 2009). It is possible that anti-H3K23me3 staining that has been observed in other systems reflects H3K23me3 in the context of a clipped H3. In the absence of tail-clipping, the available anti-H3K23me3 antibody may not be able to detect genuine H3K23me3 marks. This may account for why we could not detect the H3K23me3 modification in our immunostaining.

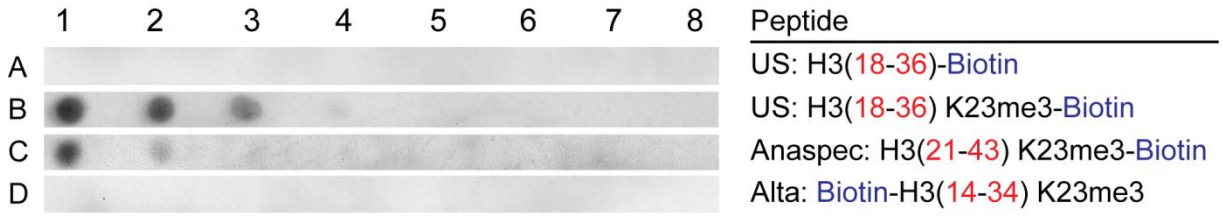


Figure 4-12. Peptide dot blot analysis was used to validate the binding specificity of anti-H3K23me3 antibody (1:100 dilution). Peptides containing the H3K23me3 modification were obtained from different suppliers as indicated above, and also vary in length and biotin position. Antibody-bound peptides were probed with anti-rabbit HRP-conjugated antibody (1:10,000 dilution). The amount of peptides spotted was as follows: (lane 1: 1 μg, lane 2: 500 ng, lane 3: 200 ng, lane 4: 40 ng, lane 5: 8 ng, lane 6: 1.6 ng, lane 7: 0.32 ng, lane 8: 0 ng)

4.2.5 Generation of transposon-induced *Nurf301* mutant lines

Research in our laboratory has previously shown that the *Nurf301* locus encodes three NURF301 isoforms by alternative splicing (Kwon et al., 2009) and two of these, NURF301-A and NURF301-B, are essentially full-length isoforms. However, the third isoform NURF301-C encodes a truncated variant that lacks the C-terminal PHD1 and PHD2 domains and the bromodomain. To establish the function of the NURF301-C isoform and to determine the function of the C-terminal PHD1, PHD2 domains and bromodomain, we attempted to generate mutations that specifically affect the C-terminus of NURF301 and thus NURF301-A and NURF301-B. To generate these we relied on the hobo deletion generator system (Mohr and Gelbart, 2002; Huet et al., 2002) to attempt to generate small deletions from a *P{wHy}* insertion (*P{wHy}CG*^{16971}*DG*^{23208}) inserted 15 kb downstream of the full-length *Nurf301* 3' UTR.

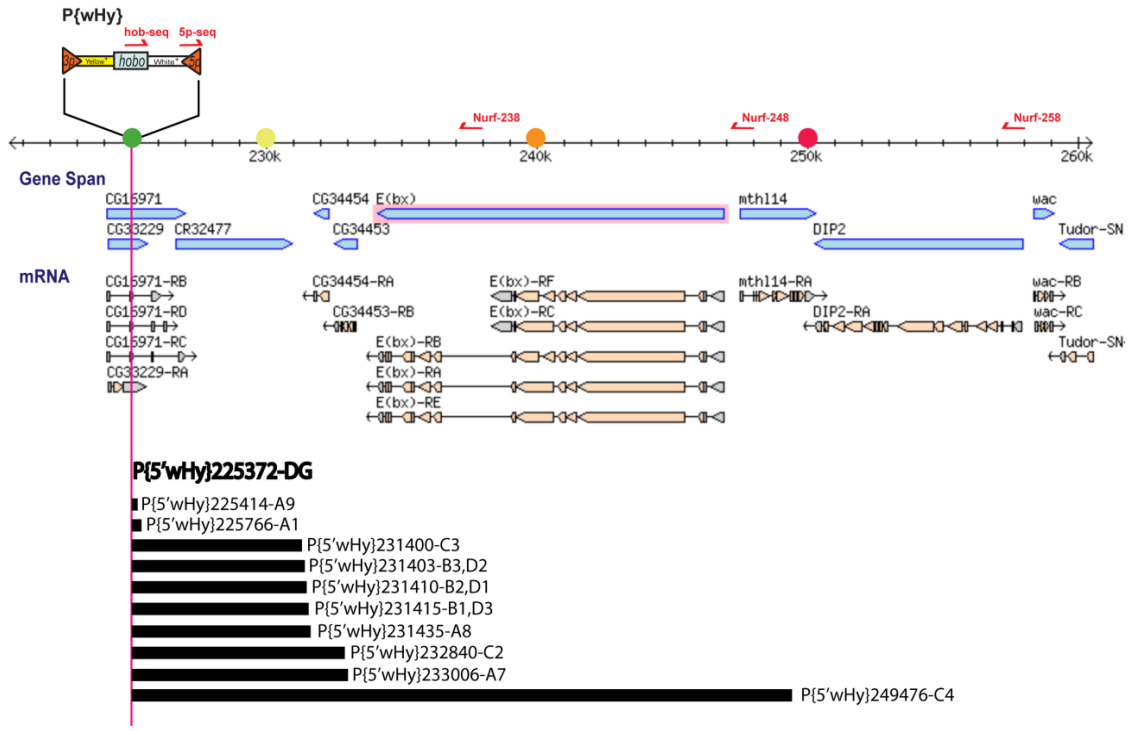


Figure 4-13. *P{wHy}* transposon induced *Nurf301* mutant lines. Black bars are showing different C-terminal deletion series of *Nurf301* we generated.

As the first step, heat shock transposase $y^1\ w^{67}c^{23}$*; In(2LR)Gla, wgGla-1/CyO, P{hsH\T-2}CyO-1*male flies were crossed with $y^1\ w^{67}c^{23}$*; P{wHy}CG*^{16971}*DG*^{23208} virgin females that contain a hobo transposon 15 kb downstream of *Nurf301*. The transposon excision was induced by heat shocking flies three times every two days in a 38℃ water bath for 30 minutes. We obtained 25 potential *Nurf301* C-terminal deleted lines and divided these into groups of 10 to verify the extent of the deletions induced by hobo transposase by sequencing. (Figure 4-13). Initially, mutant fly lines were sequenced by PCR amplifying the *Nurf301* 3' region using the hob-amp primer and either primers Nurf-238, Nurf-248 or Nurf-258. We anticipated obtaining serial deletions running into the C-terminus of *Nurf301* that would generate a sequential deletion series of different NURF301 domains. However, sequencing of the deletion lines revealed that we did not obtain a uniform series of deletions. Rather, it appeared that there were a small number of hotspots near the *Nurf301* C-terminal region into which

hobo element preferentially inserted. These did not generate a smooth series of deletions but rather a stepped series of deletions. None selectively affected the NURF301 C-terminus and only NURF301-A or NURF301-B. Most of the deleted fly lines were homozygous fertile and viable, with the exception of one in which the whole *Nurf301* locus was deleted, which was homozygous lethal. This confirmed former studies of *Nurf301* deletion mutants which were homozygous lethal (Kwon et al., 2009). As we were unable to obtain the targeted deletions required, we switched our strategy to use a targeted approach using recombineering (Vencken et al., 2006) to tag each NURF301 isoform and analyse their distributions on chromatin.

4.2.6 Recombineering to generate CTAP-tagged NURF301 isoforms

Previous work in the laboratory had indicated that traditional methods of generating tagged proteins (cDNA constructs in P-element transformation vectors) are not successful for *Nurf301*. Firstly, *Nurf301* cDNAs are toxic to bacteria and undergo high levels of recombination, presumably to inactivate the protein expressed at low levels from these cDNAs. Secondly, the large size of P-element transformation vectors after integration of the large *Nurf301* cDNA (8.6 kb) results in plasmids that are poorly maintained in *E.coli*. To circumvent these problems, we decided to use BACs that contain the *Nurf301* locus (and which are not toxic due to the presence of intron sequences) as a way of epitope-tagging the *Nurf301* locus.

These BACs were engineered using recombineering (Lee et al., 2001; Vencken et al., 2006) to introduce a C-terminal TAP (CTAP) tag to *Nurf301*. The CTAP tag contains two tandomly expressed copies of protein G and one streptavidin domain that allow purification and immunostaining of the resultant NURF301 protein. Two different BACs were selected for

initial tagging, CH321(85D11) and CH322(169D05). 85D11 contains a large insert including a number of flanking genes. 169D05 contains a smaller insert that spans the *Nurf301* cDNA and approximately 3kb upstream and downstream. We used recombineering to introduce a CTAP tag at the C-terminus of either NURF301-A/B (the full-length isoform) or NURF301-C (the C-terminally truncated isoforms). Recombination was a two-step process in which first the selectable/counter-selectable marker *galK* was introduced at the C-terminus. Then the *galK* marker was replaced with the CTAP cassette and CTAP-containing BACs selected by counter-selection against *galK* by growing bacteria on the GalK substrate DOG that generates a toxic product if *galK* is present. We successfully generated CTAP tagged NURF301-A/B and NURF301-C constructs for both the large CH321 BAC and small CH322 BAC rescue constructs.

4.2.7 Generation of CTAP tagged *Drosophila* NURF301 transgenic flies

As an additional tool to examine NURF distribution in cells and on chromatin, we generated CTAP-tagged NURF301-A/B and CTAP-tagged NURF301-C BAC constructs for injection into flies as described above. Using the φC31 integrase system (Venken et al., 2006; Bischof et al., 2007), embryo injection was performed with purified BAC constructs into embryos of *y^1 M{vas-int.Dm}ZH-2A w^*; M{3xP3-RFP.attP}ZH-86Fb* mothers expressing the φC31 integrase in germline cells. However, the survival rate of injected embryos was too low and no transgenic flies were recovered. We suspected that the BAC DNA samples purified using standard plasmid DNA midi-prep kit contained low level of endotoxin, which could affect the survival of embryos after injection. To resolve this, we utilized endotoxin-free plasmid DNA midi-prep reagents (Macherey-Nagel) to purify BAC DNA and performed injections again. This time, survival rate of larvae was increased up to 20-25 %, and injection of CH322

construct delivering smaller size of *Nurf301* genomic DNA in the BAC had a better survival rate than that of the CH321 construct. Red-eyed transgenic founder lines were obtained. We generated CTAP tagged NURF301-A/B: three different lines of CH321 and seven lines of CH322 from injection of 500 embryos for each construct. We also generated CTAP tagged NURF301-C: one line of CH321 and three lines of CH322 from injection of 500 embryos for each construct.

4.2.8 Functional discrimination of NURF301-A/B and NURF301-C isoforms

4.2.8.1 Visualization of CTAP tagged NURF301-A/B and NURF301-C distributions on polytene chromosomes

The CTAP-tagged NURF301-A/B and C transgenic flies we generated above (see 4.2.7) were used to confirm distributions of NURF301-A/B and C isoforms in cells. We first performed immunostainings of salivary gland polytene chromosomes using anti-protein G antibodies that recognise the CTAP tag to show the distributions of the NURF301-A/B and NURF301-C isoforms on polytene chromosomes. As shown in figure 4-12, both NURF301-A/B and C were localized to interbands or the interface between bands. However, comparisons of anti-protein G antibody staining indicated that the relative expression of the NURF301-A/B isoform was lower than that of the NURF301-C isoform (Figure 4-14 A and B). This was more apparent when double immunostaining of either NURF301-A/B or NURF301-C strains was performed using anti-protein G antibodies (to detect tagged proteins) and anti-NURF301 antibodies (to detect all NURF301 isoforms). Merged images of the NURFA/B or NURF301-C tag staining together with the anti-NURF301 staining (Figure 4-14 C and D) suggests that the NURF301-C isoform is the predominant variant on polytene chromosomes. More bands are shared between NURF301-C-CTAP and anti-NURF301 (visible as yellow merged bands) than with NURF301-A/B-CTAP and anti-NURF301.

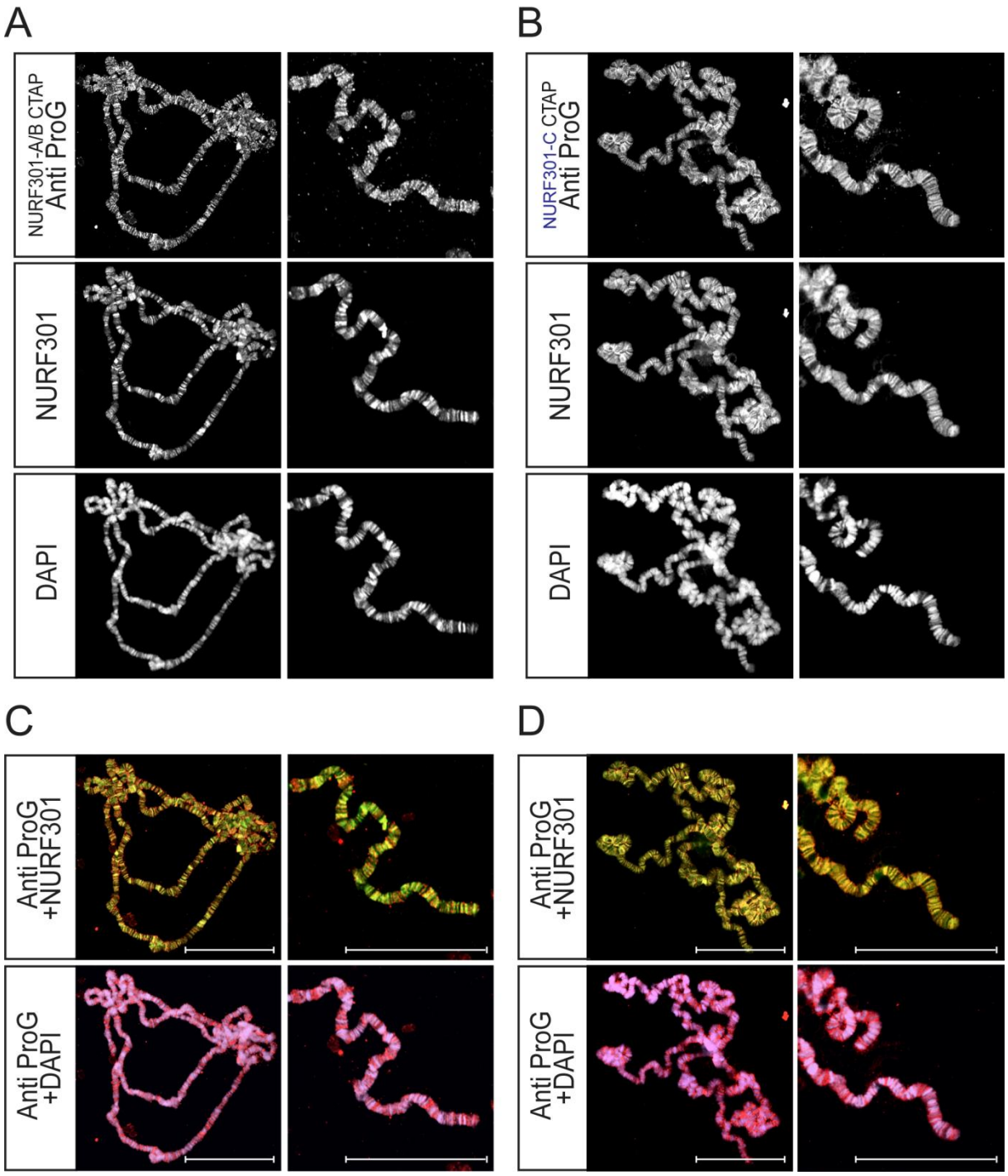

A
NURF301-A/B CTAP
Anti ProG
NURF301
DAPI
B
NURF301-C CTAP
Anti ProG
NURF301
DAPI
C
Anti ProG
+NURF301
Anti ProG
+DAPI
D
Anti ProG
+NURF301
Anti ProG
+DAPI

Figure 4-14. Immunostaining of polytene chromosome to reveal the distributions of the NURF301-A/B and NURF301-C isoforms. A) Double immunostaining of polytene chromosomes from the CH322 NURF301-A/B CTAP transgenic line that expresses CTAP-tagged NURF301-A/B. Anti-protein G antibody reveals the CTAP tag, anti-NURF301 antibodies (show all NURF301 isoforms). DAPI staining reveals DNA. (B) Double immunostaining of polytene chromosomes from the CH322 NURF301-C CTAP transgenic line that expresses CTAP-tagged NURF301-C. Anti-protein G antibody reveals the CTAP tag, anti-NURF301 antibodies. (show all NURF301 isoforms. DAPI staining reveals DNA. (C,D) Merged images of the corresponding panels in A and B to show tag and anti-NURF301 overlap and overlap between DAPI and tags. Higher magnifications of chromosome arms are shown on the right to allow better discrimination of chromosome banding patterns. Scalebar represents 50 μm.

To determine whether there are any localization sites specific for either NURF301-A/B or NURF301-C isoform, magnifications of double immunostained polytene chromosome arms described above (Figure 4-14 C and D, right) were shown as “split” chromosome images. Displaying multiple staining patterns in a ‘‘split’’ format helps avoiding visual artefacts when high levels of one protein can mask low levels of another (Corona et al., 2004; Wysocka et al., 2006). We identified some NURF301-A/B or NURF301-C isoform specific binding sites by looking for bands that are detected by anti-NURF301 staining but were not present in anti-protein G staining that targets the CTAP tagged NURF301-A/B or NURF301-C isoforms. As expected, more NURF301-C isoform specific sites were identified on chromosome arms (Figure 4-15 A). Relatively few NURF301-A/B isoform specific sites were found on chromosome arms when using the NURF301-C CTAP transgenic line(Figure 4-15 B) which were double immunostained using anti-protein G and anti-NURF301 antibodies. Most of those sites were non-overlapping but specific to the NURF301-A/B or NURF301-C isoforms. This could mean that NURF301-A/B and NURF301-C isoforms work independently on chromatin as NURF301-C isoforms do not contain PHD1, PHD2 and the bromodomain, which are critical to interact with histone modifications we identified above. However, it was difficult to conclude this given the level of resolution provided by immunostaining. To confirm this, these results need to be validated by more precise and high resolution mapping techniques such as ChIP sequencing, which will be discussed in chapter 4.2.9.

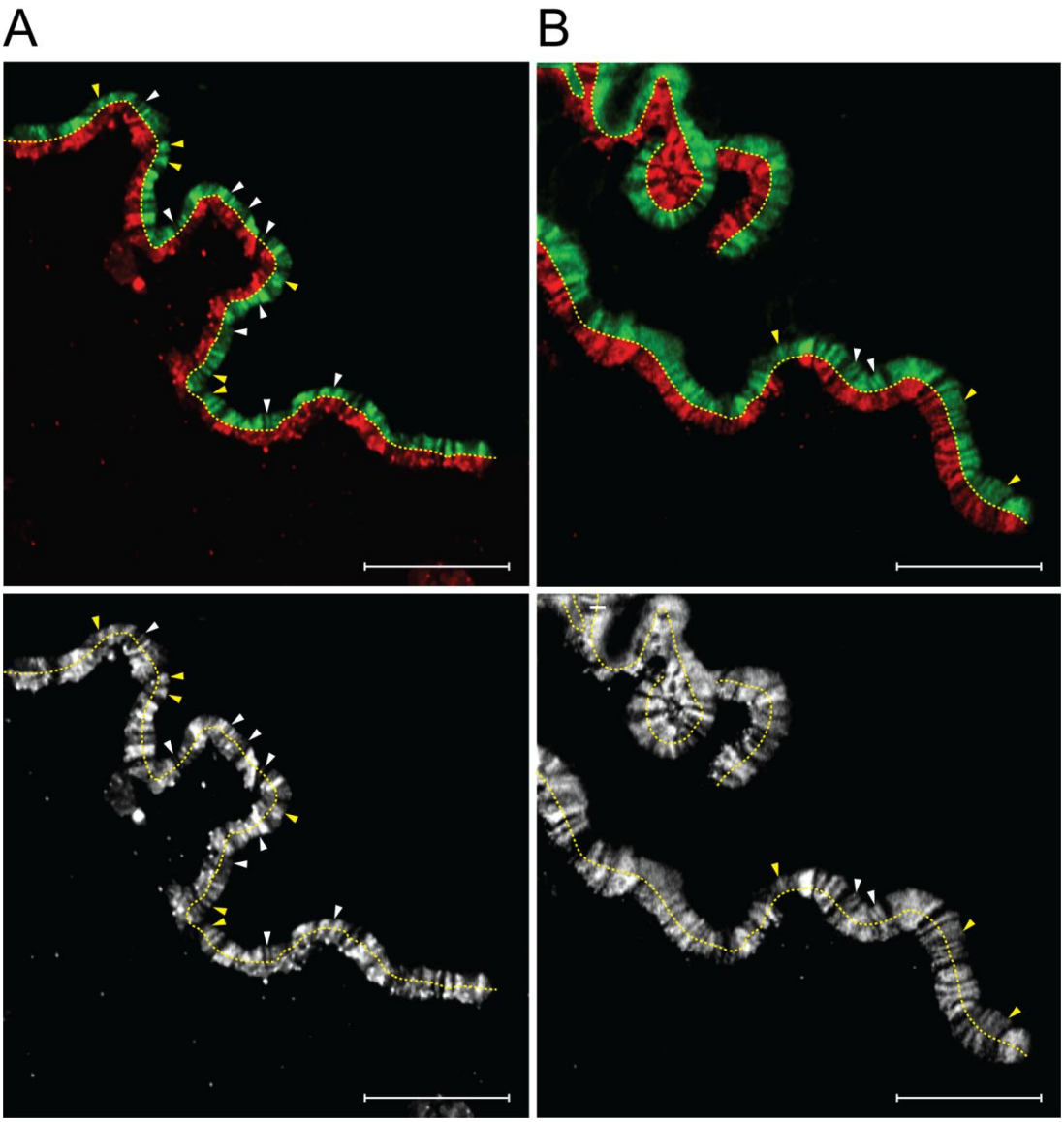


Figure 4-15. Double immunostained chromosome arms are shown as "split" chromosome images. Distribution of (A) the NURF301-A/B isoform (red) and anti-NURF301 antibody staining (green), (B) NURF301-C isoform (red) and anti-NURF301 antibody staining (green) on salivary gland polytene chromosome arms from CH322 NURF301-A/B CTAP and CH322 NURF301-C CTAP transgenic lines respectively. The "spilt" images are also shown in black and white (below) for better discrimination of banding patterns, which are not distinguished in the color image. The white and yellow arrowheads indicate the bands missing in anti-protein G antibody staining, but yellow shows more significant bands. Scalebar represents 20 μm.

4.2.8.1.1 Visualization of histone H3K9ac, H3K9acS10p and NURF301-A/B isoform distributions on polytene chromosome by immunostaining

As shown in figure 4-9, some of bands on polytene chromosomes recognized by anti-NURF301 staining colocalized with histone H3K9ac sites. However, it was unclear whether the NURF301-A/B isoform was responsible for this colocalization with H3K9ac, as the antibody we used recognized both NURF301-A/B isoform and the NURF301-C isoform which lacks the PHD2 domain. Using the NURF301-A/B CTAP transgenic line we generated, we could resolve this as we could specifically visualize the NURF301-A/B isoform with respect to the histone H3K9ac and phospho-acetyl H3K9ac/S10p modifications on polytene chromosomes.

Double immunostainings of polytene chromosomes from the NURF301-A/B CTAP transgenic line were performed. As we identified previously, the NURF301-A/B isoform was localized mainly at interbands or the interface between bands. Histone H3K9ac and H3K9ac/S10p modifications were detected at numerous loci over the polytene chromosomes, but H3K9ac/S10p displayed a more restricted distribution (Figure 4-16 A and B). We observed clear overlap between NURF301-A/B isoform binding sites and the distribution of the H3K9ac or H3K9ac/S10p modifications (Figure 4-16 C and D), in comparison to the previous anti-NURF301 staining performed previously (Figure 4-9 B, left). Unlike the previous anti-NURF301 staining, no bands that were only stained by anti-protein G antibody staining were detected. These data confirmed that recruitment of the full-length NURF301-A/B isoform to chromatin could be correlated with the presence of histone H3K9ac or H3K9ac/S10p modifications.

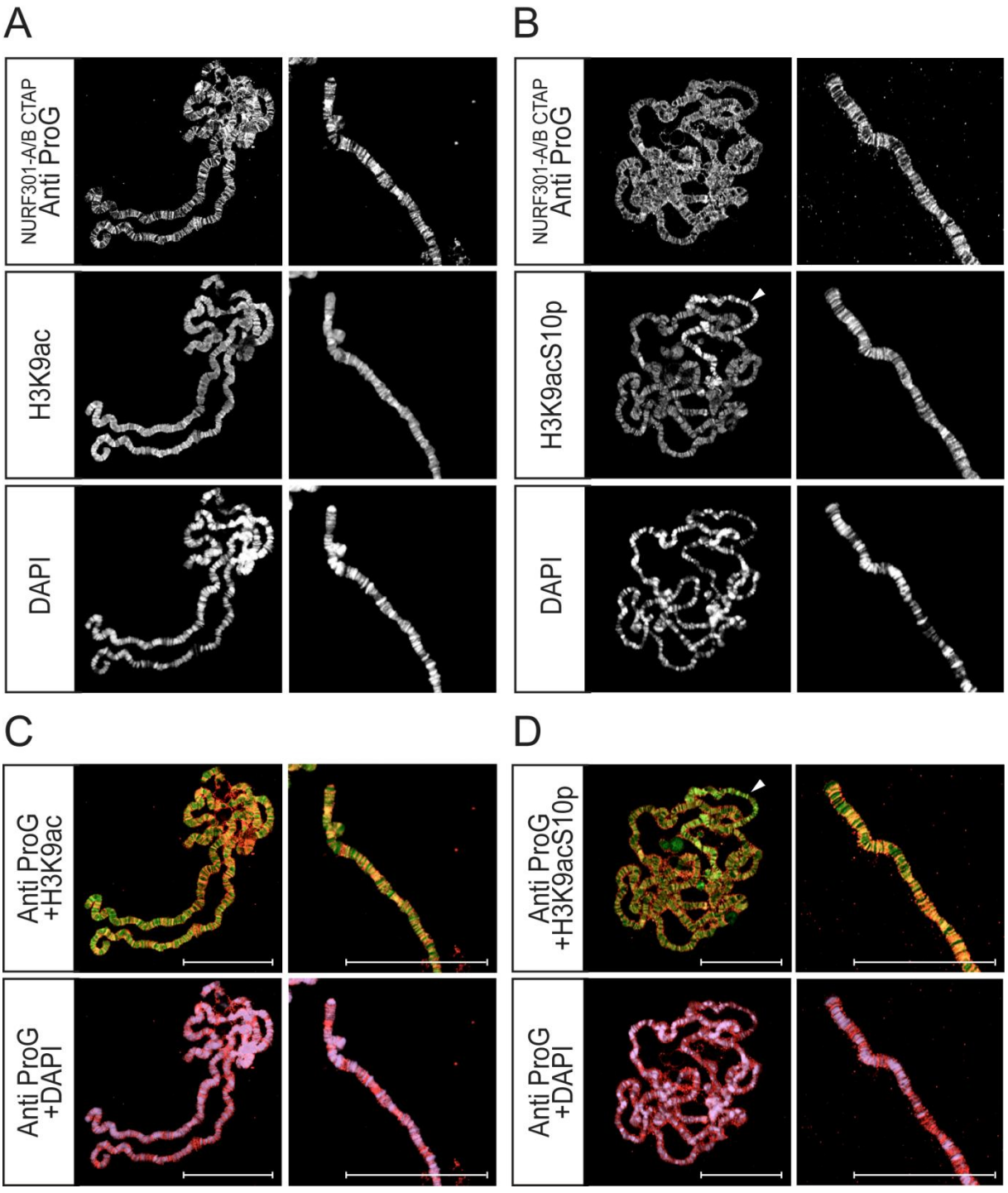

A
NURF301-A/B CTAP Anti ProG
H3K9ac
DAPI
B
NURF301-A/B CTAP Anti ProG
H3K9acS10p
DAPI
C
Anti ProG +H3K9ac
Anti ProG +DAPI
D
Anti ProG +H3K9acS10p
Anti ProG +DAPI

Figure 4-16. The full-length NURF301-A/B isoform colocalizes with the histone H3K9ac and H3K9ac/S10p modifications. The double immunostained chromosome images show the distributions of histone H3K9ac, H3K9ac/S10p modifications and the CTAP-tagged NURF301-A/B isoform (revealed by anti-protein G antibody). (A) Double immunostaining images of H3K9ac, NURF301-A/B and DAPI on polytene chromosome and (B) H3K9ac/S10p, NURF301-A/B and DAPI on polytene chromosomes from CH322 NURF301-A/B CTAP transgenic line shown in separate panels in black and white. (C) Merged double immunostaining images of the NURF301-A/B isoform (red)/H3K9ac (green) and NURF301-A/B isoform (red)/DAPI(purple), (D) NURF301-A/B isoform (red)/H3K9ac/S10p (green) and NURF301-A/B isoform (red)/DAPI(purple). Arrowheads indicate the male X-chromosome. Higher magnification images of the chromosome arms are shown on the right to better discriminate chromosome banding patterns. Scalebar represents 50 μm.

4.2.8.2 Identification of localization of NURF301-A/B and NURF301-C isoforms in male testes and ovaries by immunostaining

Previous research in the laboratory and by other groups has shown that NURF301 is required for testis and ovary development (Kwon et al., 2009; Cherry and Matunis, 2010; Ables and Drummon-Barbosa, 2010). In the male testis, spermatogenesis begins with germline stem cell division to produce a new stem cell and a gonialblast. The gonialblast then undergo four rounds of mitotic amplification divisions to generate 16 primary spermatocytes. This is followed by two meiotic divisions after which, spermatid differentiation occurs and mature sperm are produced (reviewed in Fuller, 1998). Previous data indicated differential requirements for the two NURF301 isoforms during testes development, with NURF301-C required for germline stem cell development and NURF301-A/B required for maturation of spermatocytes and oocytes (Kwon et al., 2009). One explanation for differential requirements of the two NURF301 isoforms at different stages of testis development is that these isoforms may be differentially expressed in testes tissues. To resolve this, we examined the distribution our CTAP tagged NURF301-A/B and NURF301-C isoforms respectively by immunostaining of testes.

As shown in figure 4-17 A, the majority of NURF301-A/B staining was observed at later stages in primary spermatocytes and their descendants suggesting that NURF301-A/B isoforms more likely to be involved in later stage of spermatogenesis. In contrast, NURF301-C isoforms were highly expressed at all stages including in germline stem cells and gonialblasts as well as primary spermatocytes (Figure 4-17 B). These data suggest that differential expression of the NURF301 isoforms accounts for the previously observed isoform tissue dependencies.

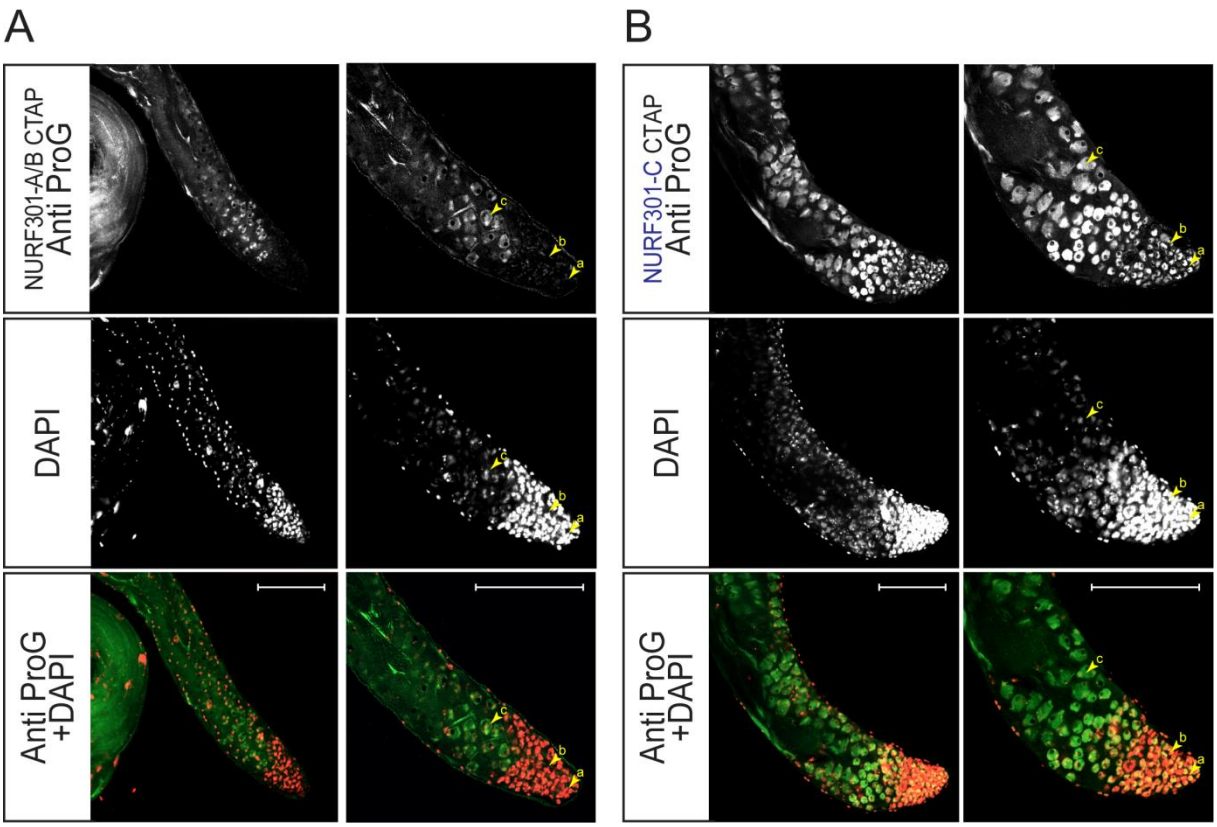


Figure 4-17. Different localizations of NURF301-A/B and NURF301-C isoforms observed in *Drosophila* male testes. CTAP tagged NURF301-A/B and NURF301-C transgenic fly male testes were dissected and immunostainined using anti-protein G antibody. DNA was visualized by DAPI staining. (A) NURF301-A/B isoform (green) and DAPI (red), (B) NURF301-C isoform (green) and DAPI (red). Schematic of spermatogenesis is indicated by yellow arrowheads: (a) germline stem cells, (b) gonialblasts, (c) primary spermatocytes. Scalebar represents 100 μm.

We next tested whether the same differential expression was observed during ovary development. During ovary development, germline stem cells at anterior tip begin asymmetric divisions that give rise to one daughter stem cell and one cytoblast. The cytoblast then undergoes four rounds of cell division generating a cyst of 16 cells, one of which differentiates as oocyte while the other 15 cells become nurse cells (reviewed in Bastock and St Johnston (2008)). We tested whether the distributions of the NURF301-A/B and NURF301-C isoforms were distinct by immunostaining CTAP tagged NURF301-A/B and NURF301-C transgenic female ovaries. Unlike testes, the distribution of NURF301-A/B and NURF301-C isoforms in ovaries were similar. In each ovariole, we observed high levels of expression anteriorly in stem cells. For both isoforms expression was reduced in maturing egg chambers although expression of the NURF301-C isoform appeared to be higher (Figure 4-18 A and B, left and middle panels). At vitellogenic stages staining of both NURF301-A/B and NURF301-C were lost from nurse cell nuclei but high level expression was retained in the nucleus of the single cell that has been singled out to become the oocyte (Figure 4-18 A and B, right panels). These results suggest that the both NURF301-A/B and NURF301-C isoforms could be involved in stem cell differentiation as well as oocyte differentiation and maintenance.

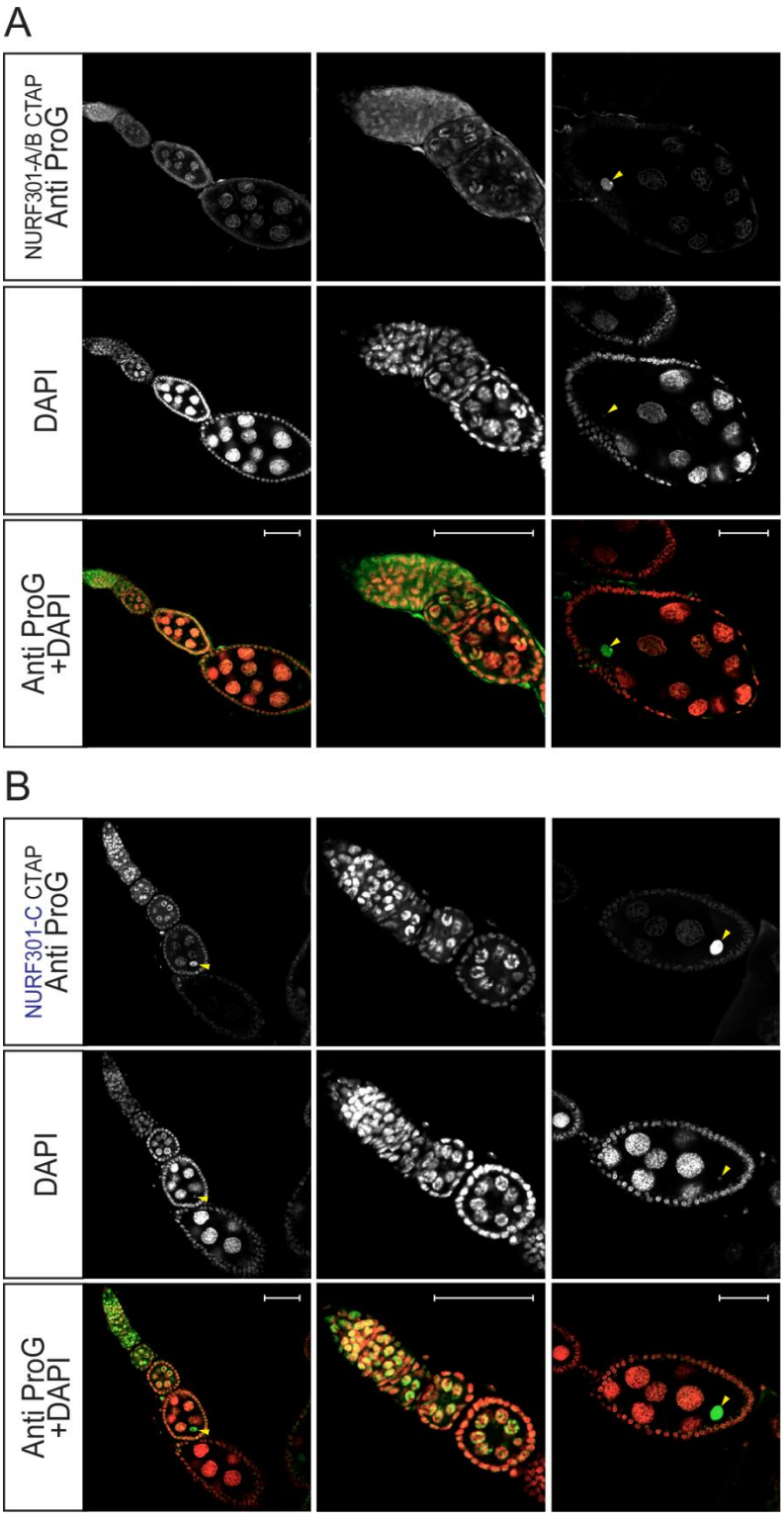


Figure 4-18. Similar localization of NURF301-A/B and NURF301-C isoforms in *Drosophila* female ovaries. CTAP tagged NURF301-A/B and NURF301-C transgenic fly female ovarioles were dissected and immunostainined using anti-protein G antibody. DNA was visualized by DAPI staining. (A) NURF301-A/B isoform (green) and DAPI (red), (B) NURF301-C isoform (green) and DAPI (red). The yellow arrowheads indicate the oocyte nuclei. Scalebar represents 50 μm.

4.2.9 ChIP sequencing

4.2.9.1 ChIP-seq DNA library preparation of NURF301-A/B and NURF301-C isoforms and work flow analysis (WFA)

To map and accurately characterize the NURF isoforms-DNA interactions of an entire genome, we performed ChIP sequencing using CTAP tagged NURF301-A/B and NURF301-C hemocytes. The samples were also prepared male and female separately to determine whether NURF301 isoforms have any characteristics of sex dependent regulation of gene expression as we observed such properties in our testes/ovaries immunostaining (see Figure 4-17 and Figure 4-18).

DNA for sequencing on the SOLiD sequencer needs to be fragmented in the range 100-300 bp for optimal sequencing. We used two approaches to fragment ChIP DNA for library preparation: i) sonication of chromatin using a Covaris S2 Adaptive Focused Acoustics ultrasonicator followed by DNA size selection or ii) micrococcal nuclease (MNase) digestion of chromatin to generate mononucleosome sized 146 bp DNA fragments.

CTAP tagged NURF301-A/B male/female and NURF301-C male/female hemocyte samples were used for ChIP-seq DNA library preparations, and cross-linked chromatin was first sheared by sonication using Covaris. However, we were unable to obtain enough immunoprecipitated DNA after purification for library preparation (data not shown). This may be due to fragmentation of NURF301 as well as DNA during Covaris ultrasonication and the separation of antibody epitopes and DNA-crosslinking sites. To circumvent this, we next performed MNase digestion of cross-linked chromatin and were able to obtain reasonable amount of immunoprecipitated DNA. This was converted into P1 adaptor/barcode ligated

sequencing libraries, PCR amplified and purified. FlashGel electrophoresis of the resultant ChIP DNA libraries indicated that they were sized in the region of 250 bp, the expected size of a library that includes 146 bp MNase digested mononucleosomes and ~100 bp of adaptor/barcode fragments (Figure 4-19).

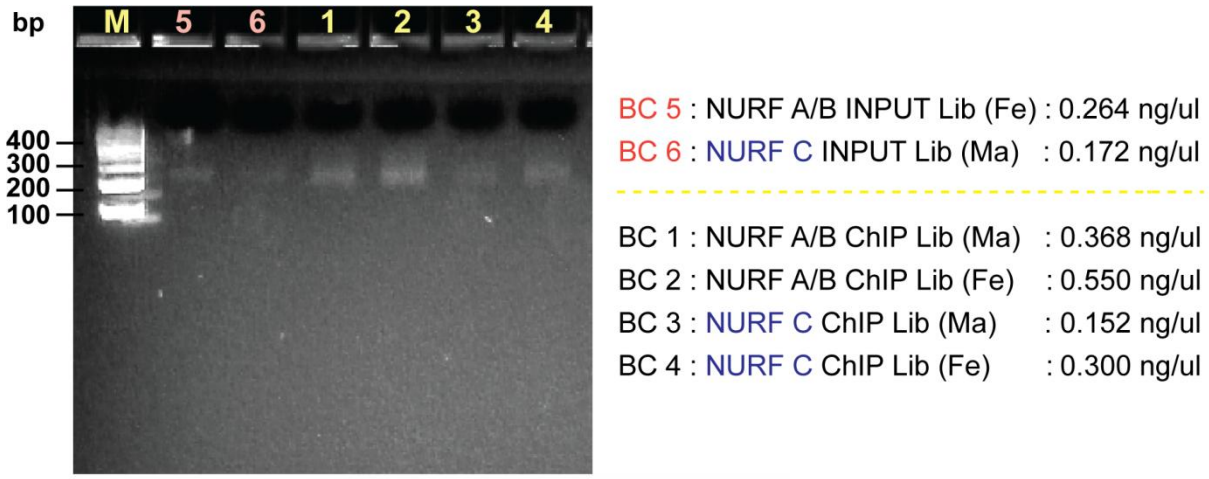
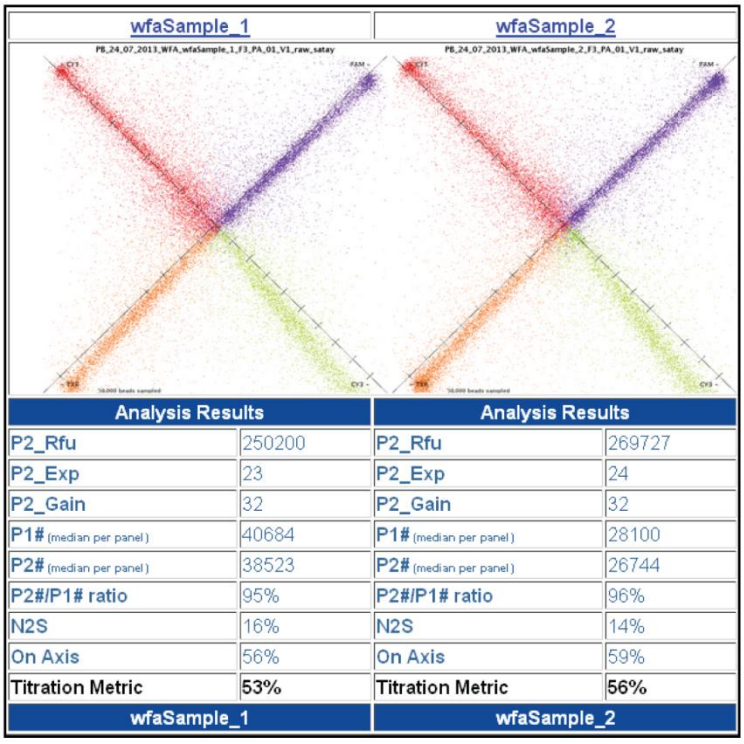


Figure 4-19. FlashGel image showing purified barcoded ChIP DNA libraries. Each lane is labeled with the barcode (BC) number of the library sample (BC 1: NURF301-A/B male, BC 2: NURF301-A/B female, BC 3: NURF301-C male, BC 4: NURF301-C female, BC 5: NURF301-A/B input female, BC 6: NURF301-C input male). Purified library DNA was quantified as indicated.

Templated beads were prepared using equal amounts of barcoded DNA fragment libraries were pooled into a single sample, and quality of templated beads was determined by work flow analysis (WFA) prior to full sequencing. We prepared two different sets of beads and so ran two WFA analyses. The first set of WFA data is shown in figure 4-20A and the second set is shown in figure 4-20 B. The quality of beads could be graphically visualized on a Satay plot (Figure 4-20 A and B above). These plots show the relative intensities of the signals for each bead and how well separated the signal is across the four fluorecent dye channels in four axes in a X, Y two dimension (red: Cy5, yellow: Texas Red, green: Cy3 and blue: 6-FAM). The P2:P1 ratio of our templated beads was in optimal level of 95-97%, which assess the

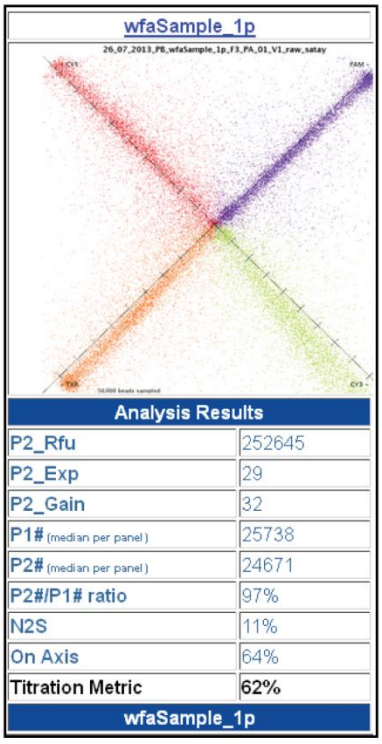
efficiency of emulsion and enrichment of template positive beads. The noise to signal ratio (N2S) was 11-16%, and beads on-axis were 56-64% which could have been higher for better sequencing. Titration meric value, which is percentage of P2 positive beads times percentage of beads on axis, was 53-62% lower than we expected. The Satay plot displayed these bead spectral purity and signal intensity graphically (Figure 4-20 A and B). In total, templated beads were qualified and applied to run full sequencing based on this data.

A



wfaSample_1		wfaSample_2	
Analysis Results		Analysis Results	
P2_Rfu	250200	P2_Rfu	269727
P2_Exp	23	P2_Exp	24
P2_Gain	32	P2_Gain	32
P1# (median per panel)	40684	P1# (median per panel)	28100
P2# (median per panel)	38523	P2# (median per panel)	26744
P2#/P1# ratio	95%	P2#/P1# ratio	96%
N2S	16%	N2S	14%
On Axis	56%	On Axis	59%
Titration Metric	53%	Titration Metric	56%
wfaSample_1		wfaSample_2	

B



wfaSample_1p	
Analysis Results	
P2_Rfu	252645
P2_Exp	29
P2_Gain	32
P1# (median per panel)	25738
P2# (median per panel)	24671
P2#/P1# ratio	97%
N2S	11%
On Axis	64%
Titration Metric	62%
wfaSample_1p	

Figure 4-20. WFA run report shows the quality of templated beads. Two different sets of beads were prepared (A) and (B), and the data displays P2_Rfu, P2_Exp, P2_Gain, P1 positive beads, P2 positive beads, P2/P1 ratio, noise to signal ratio (N2S), beads on axis and titration metric value.

4.2.9.2 ChIP sequencing and peak calling

ChIP DNA libraries were sequenced on a SOLiD 4 genome analyser. Libraries were sequenced using a SOLiD ToP Fragment Barcoded Sequencing Kit MM50/5 (Life

Technologies). Beads in the pool are first assigned to a ChIP library after 5bp barcode sequencing, based on the unique barcode adaptor used to generate each library. Thereafter a 50bp read is produced from every DNA fragment and assigned to the correct ChIP library based on the initial barcode sequence. Quality control of the sequencing run was analyzed using the SETS (Solid experiment trafficking system) software of the SOLiD 4 analyzer. The total number of reads generated for each library are shown in Table 4-1 (The 1st column). Subsequently the 50 bp DNA sequence reads were mapped onto the reference *Drosophila* genome (release dm3) using either local mapping parameters in the Bioscope software (Life Tecnologies). The total number of reads mapped for each library are shown in Table 4-1 (The 2nd column). It is noteworthy that with the exception of Input DNA samples, mapping percentages were low (Table 4-1, the 3rd column shaded yellow), suggesting that ChIP-DNA libraries may have been contaminated by adaptor-dimers, which are prevalent when small amounts of DNA are used for library preparation. Mapped reads were also filtered for high quality reads using Samtools (Li et al., 2009) which resulted in even greater loss of mapping reads as shown in Table 4-1 (The 4th and 5th column shaded green).

Table 4-1. ChIP sequencing map reads

Library	Total reads	Mapped reads	% Mapped	Filtered reads	% Filtered
NURF301-A/B Input (Female)	36,543,183	9,488,788	25.97	3,272,750	8.96
NURF301-A/B Male	92,241,527	2,404,941	2.61	993,940	1.08
NURF301-A/B Female	77,317,411	1,921,861	2.49	760,217	0.98
NURF301-C Input (Male)	44,791,213	8,294,403	18.52	3,148,436	7.03
NURF301-C Male	76,033,813	4,297,962	5.65	3,148,436	4.14
NURF301-C Female	86,147,157	4,931,614	5.72	2,771,168	3.22

Even though mapped read numbers were low, especially for NURF301-A/B ChIP samples, we attempted to call genome-wide peaks of NURF301-A/B and NURF301-C. Peaks were identified using the MACS programme (Zhang et al., 2008). MACS ChIP wiggle tracks were then used to generate averaged genome profiles of the distribution of the NURF301-A/B and NURF301-C peaks using the CEAS (Cis-regulatory Element Annotation System) script of the Cistrome package (Liu et al., 2011).

Due to the low number of mapped and filtered reads that were obtained, ChIP profiles did not allow NURF recruitment at individual genes to be analyzed in any detail due to the low sequence tag coverage. However, we did attempt to extrapolate information on NURF recruitment by generating gene-averaged profiles. Using CEAS we were able to average NURF301-A/B and NURF301-C signals over all genes with respect to either transcription start sites (TSSs), transcription termination sites (TTSs) or by performing meta-gene analysis were all genes are artificially expanded or compressed to fit a hypothetical gene structure consisting of 1 kb of upstream and downstream regulatory regions and 3 kb of coding sequence. The averaged profiles of NURF301-A/B suggested that it was enriched at the TSS and TTS in both male and female samples (Figure 4-21, A and B). A peak of high occupancy of NURF301-C was seen around the transcription termination site (TTS) and flanking the transcription start site (TSS) in both male and female (Figure 4-22, A and B). Meta-gene analysis appears to indicate that there is a trend of increased enrichment of both NURF301-A/B and NURF301-C towards the 3' end (TTS), indicating that they might exhibit more remodelling functions at transcription termination regions (Figure 4-21 and Figure 4-22, average gene profiles). However, it is important to restate that sequence coverage is low and ChIP efficiency would need to be increased to generate more sequence reads to allow NURF

ChIP profiles to be determined with any confidence. The low sequence coverage obtained makes profiles generated susceptible to mapping artefacts.

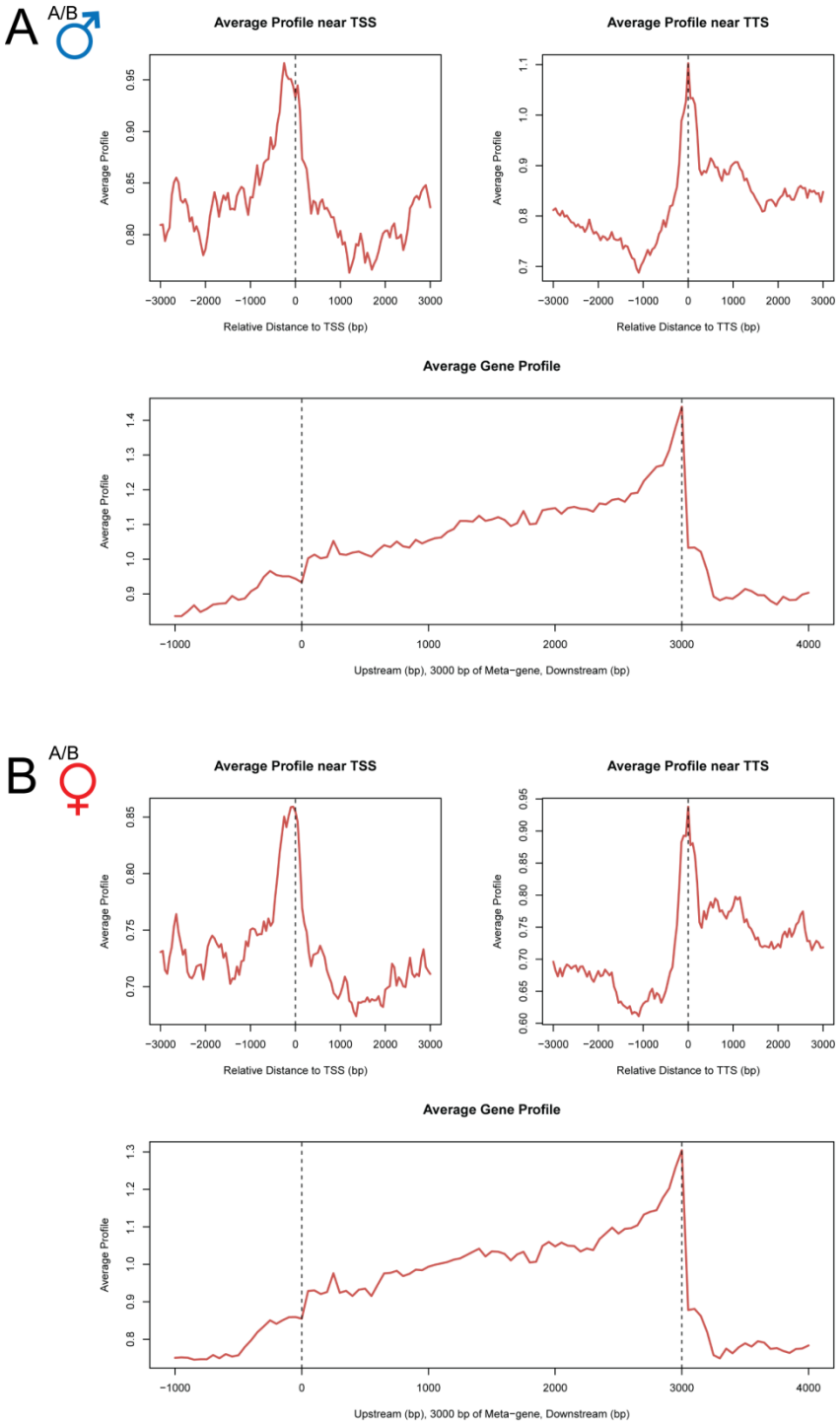


Figure 4-21. Average ChIP-seq read density profiles of the NURF301-A/B isoform. NURF301-A/B distribution is plotted relative to transcription start sites (TSS) and transcription termination sites (TTS). Both (A) male and (B) female show similar patterns.

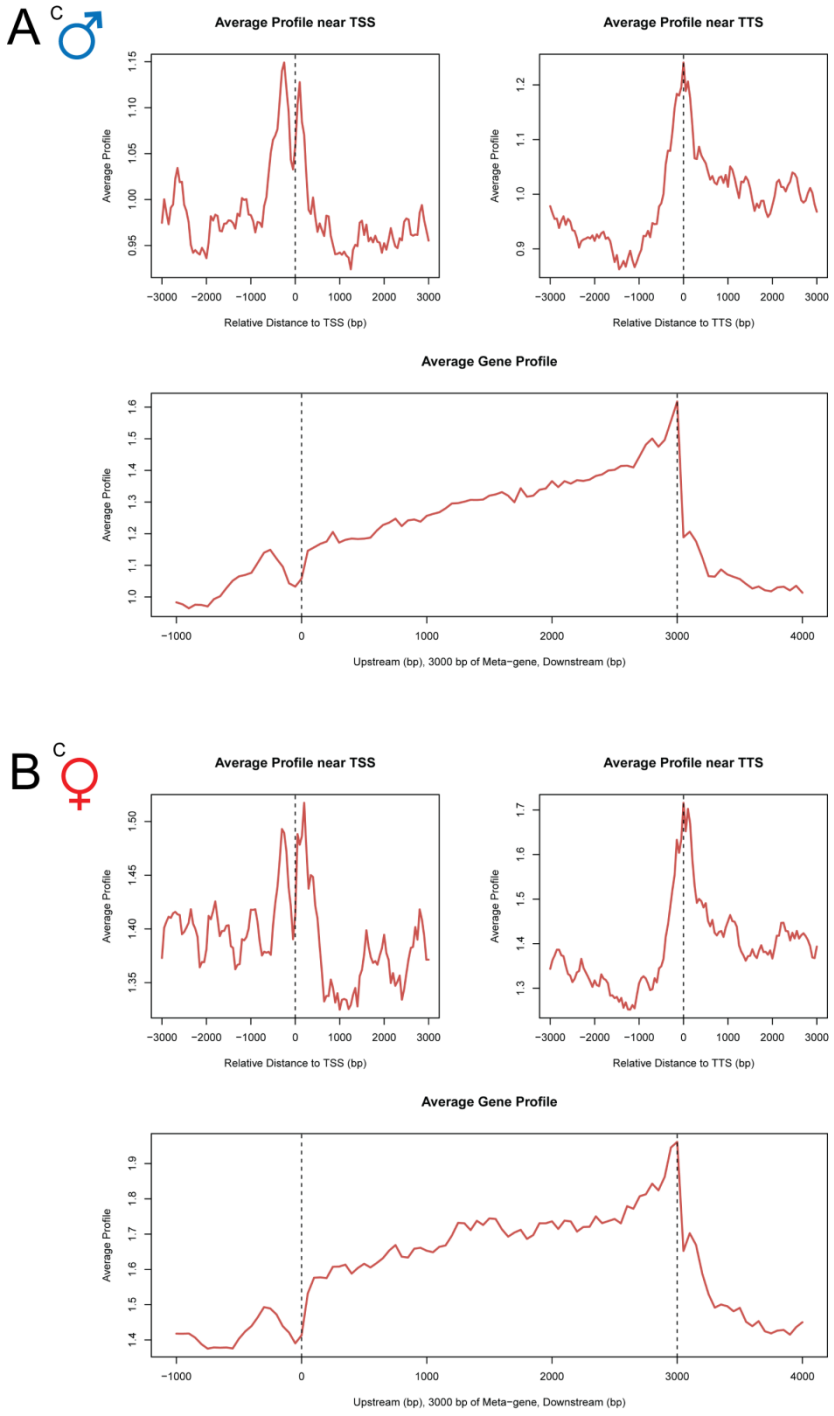


Figure 4-22. Average ChIP-seq read density profiles of NURF301-C isoform. NURF301-C distribution is plotted relative to transcription start sites (TSS) and transcription terminate sites (TTS). Both (A) male and (B) female show similar patterns.

4.3 Conclusion

Immunostaining data of different cell types is consistent with our *in vitro* data of interactions between the NURF and histone modifications. We first examined the distributions of NURF301, its primary recognition mark H3K4me3, and H3T3p, a postulated flanking inhibitory mark, by immunostaining of wild-type (w^{1118}) embryos. We observed that H3K4me3 was present in both interphase and mitotic cells in embryos after cellularization. NURF301 colocalized with H3K4me3 during interphase, but NURF301 was ejected from chromatin during mitosis. This corresponded with an increase in the H3T3p modification, which was highly expressed during metaphase and anaphase when NURF was absent from chromatin. These data were validated by double immunostaining of NURF and H3T3p/K4me3 showing no colocalization of NURF with H3T3p/K4me in dividing cells of embryos and neuroblasts, suggesting NURF301 could be delocalized by phosphorylation of H3T3 during mitosis.

Immunostaining of non-mitotic interphase chromosomes from wild-type (w^{1118}) larvae were also performed to analyse *in vivo* localization of NURF301, and the H3K9ac and H3K9ac/S10p modifications, which are postulated to be flanking enhancers of NURF binding to H3K4me3. NURF was mainly detected at interbands or the interface between bands. H3K9ac and H3K9ac/S10p were observed at numerous loci not only interbands but also DAPI stained bands. By double immunostaining of polytene chromosome, we observed that the NURF301 and H3K9ac colocalized at some interband sites, but NURF301 localization was not completely affected by loss of the H3K9ac modification in *dGcn5* mutants. However, we later showed that the NURF301-A/B isoform colocalized well with H3K9ac and H3K9ac/S10p. This suggested that localization of full-length NURF301-A/B isoform

containing PHD2 domain, as we found interacting with those markers *in vitro,* could be affected by the presence of H3K9ac or H3K9ac/S10p modifications.

We also attempted to verify that NURF301 binds the histone H3K23me3 modification, which was found as a primary recognition mark of the PHD1 domain. In a first series of experiments, the *in vivo* distribution of H3K23me3 was analysed by immunostaining of embryos, testes and ovaries. However no staining was detected using the available anti-H3K23me3 antibody. However, characterization of the binding specificity of this antibody using histone peptides suggested that this antibody is only capable of recognizing the H3K23me3 modification in the context of N-terminally clipped histone H3. Without alternative antibodies, the absence of staining prevents us from concluding whether this modification is truly absent or not detectable. .

For more functional analysis of these domains, we tried to generate serial deletions of the C-terminus of *Nurf301* by transposon excision. This would generate a sequential deletion series of *Nurf301* deletion mutants in which chromatin-targetting domains were progressively deleted. However, these experiments failed to generate a progressive series of mutants as transposon insertion hotspots near to the *Nurf301* C-terminal region resulted in a stepped series of deletions. However, we have established the BAC recombineering system which will allow us to engineer *Nurf301* deletions more precisely and efficiently. We have used this system to generate C-terminally CTAP epitope-tagged variants of the NURF301-A/B and NURF301-C isoforms. We have used transgenic flies that express these to further analyse *in vivo* function and localization of NURF301. We expect this optimized system will also allow us to generate BAC constructs in which the PHD, PHD1, PHD2 and bromodomain can be

selectively deleted to examine function and targeting requirements.

By using the CTAP tagged NURF301-A/B and NURF301-C transgenic lines we could discriminate the function of the two NURF301 isoforms in different cell types. Immunostainings of polytene chromosome NURF301-A/B and NURF301-C revealed that each isoform had specific and independent binding sites on chromatin. Moreover, in testes immunostaining, NURF301-A/B was observed to be expressed only at late stages of spermatogenesis in contrast to NURF301-C which was broadly expressed in germline stem cells and primary spermatocytes. These data indicate that alternative splicing of *Nurf301* can generate distinct chromatin remodelling complexes. We suggest that these complexes have unique functions as a result of differential chromatin targeting as well as differential tissue expression.

To extend our understanding of genome wide localizations of NURF301-A/B and C accurately, ChIP sequencing was performed using CTAP tagged NURF301-A/B and NURF301-C hemocytes. Although it was possible to generate gene averaged NURF profiles, the low sequence coverage precluded detailed analysis of NURF301-A/B or NURF301-C isoforms and comparisons with histone modifications.

CHAPTER 5. DISCUSSION

5.1 The recruitment and function of NURF is influenced by histone modifications

In this study, we analysed the full histone modification-binding specificities of the NURF301 PHD domains and bromodomain by using modified histone peptide library arrays and peptide pull down assays containing individual modifications and combinations of histone modifications.

5.1.1 Modified histone binding specificity of the PHD2 domain

In the case of PHD2 domain, we observed clear binding specificity for histone H3 tri-methylated Lys 4 as previously identified (Li et al., 2006; Wysocka et al., 2006, Kwon et al., 2009). No binding was observed to any other singly-modified histone tails on arrays. This is significant as it was previously speculated that the NURF301/BPTF PHD2 domain may also recognize other methylation states of the H3K4 residue, for example H3K4me2 (Sims and Reinberg, 2006). However, our binding studies using two different sets of peptide arrays are consistent with specific recognition of H3K4me3 over H3K4me2.

In addition, we found that the binding of the PHD2 domain to H3K4me3 was able to be inhibited by phosphorylation of flanking Thr 3, and enhanced by acetylation of Lys 9 and phosphorylation of Ser 10. The enhancement of binding by the H3K9ac modification is consistent with recent MS (mass spectrometry) based proteomic analysis revealed that BPTF binding to H3K4me3 was increased by H3K9ac (Vermeulen et al., 2010). These studies did not discriminate the mechanism of this enhancement. It was possible that binding of the double marks was mediated by distinct domains such as the PHD2 and bromodomains. However, our data conclusively shows that this enhancement is mediated through the single

PHD2 domain and that H3K9ac modifies the binding of the PHD2 domain to H3K4me3. Our structural modelling indicates the existence of favourable charge interactions that stabilise binding of the PHD2 domain H3K4me3/K9ac/S10p peptides. It is important to note that binding to H3K9ac and H3S10p is not observed unless the H3K4me3 mark is present. H3K4me3 is the primary target of the PHD2 domain with other modifications acting as rheostats to increase or decrease binding. These data are consistent with studies of other remodelling complexes, such as the RSC complex, which identified that recruitment of RSC can be increased by specific acetylation at histone H3 Lys 14 (H3K14ac) (Ferreira et al., 2007).

In contrast, we have clearly established that phosphorylation of the Thr3 residues is able to inhibit the binding of the PHD2 domain to H3K4me3. Our structural modelling suggests that this is the result of unfavourable charge interactions between H3T3p and negatively charged residues in the H3K4me3 binding pocket. Phosphorylation of H3T3p provides a convenient mechanism by which binding of NURF to H3K4me3 could be masked in a regulated manner, without requiring the removal of H3K4me3. Our data suggest that this may be responsible for controlling chromatin targeting of NURF during mitosis. While we have not been able to test this model by removing the kinase the deposits H3T3p (Haspin), analysis of NURF and H3T3p distribution indicates an inverse relationship. During mitosis high levels of H3T3p and H3T3pK4me3 are detected. At these stages NURF is ejected from chromatin.

The NURF301 PHD2 finger provides another example of an emerging theme in methylated histone tail recognition, namely the existence of so called “phospho-methyl switches” that can regulate methylated histone tail recognition, blocking tail binding without requiring the

removal of the methylated lysine residue. This crosstalk has been suggested in several studies including regulation of HP1 (heterochromatin protein 1) (Fischle et al., 2005; Hirota et al., 2005), PRC2 (polycomb repressive complex 2) (Gehani et al., 2010; Lau and Cheung, 2011), the WDR5 subunit of MLL1 (Couture et al., 2006; Ruthenburg et al., 2006) and the TAF3 subunit of TFIID (Varier et al. 2010).

5.1.2 Modified histone binding specificity of the PHD1 domain

The PHD1 domain, which shows high sequence similarity to PHD2 domain (Figure 1-8), had also clear binding specificity for tri-methylated Lys 23 residue of histone H3. Initial reports suggested that this domain was only present in *Drosophila* NURF301 and not in human BPTF (Wysocka et al., 2006), but here we have identified an exon in human BPTF that encodes PHD1. This raises the possibility that human BPTF may exist in multiple splice variants and that alternative splicing of the PHD1 domain may generate human NURF complexes that have distinct modified histone recognition profiles. Research in our laboratory has shown that *Drosophila* NURF301 is controlled by alternative splicing (Kwon et al., 2009) and multiple splice variants of *C. elegans* NURF301 have been shown to occur (Andersen et al., 2006). Although alternative splicing of human BPTF has not previously been reported, splice variants of the SNF2L catalytic subunit have been reported (Barak et al., 2004). This data contributes to an emerging concept in chromatin remodelling that suggest that functionally diverse tissue- or stage-specific variants of chromatin remodelling complexes exist. These can be generated either by alternative splicing of core subunits as we have shown for NURF, or by the flexible incorporation of subunits in the case of SWI/SNF complexes (Olave et al., 2007; Lessard et al., 2007; Ho et al., 2009; Vogel-Ciernia et al., 2013).

The biological function of H3K23me3 has not been fully characterized, although there are suggestions that H3K23me3 is present on heterochromatic H3 from *Tetrahymena*, and that levels are increased during meiosis in *Tetrahymena* micronuclei (Papazvan et al., 2013). We have used the only available anti-H3K23me3 antibody to determine the distribution of H3K23me3 in *Drosophila* but are unable to detect a response. It is possible that the modification is absent from *Drosophila*, however our peptide dot blot analysis suggests that the antibody is affected by the presence of residues N-terminal to position K18 and thus may not recognise unclipped histone tails. This histone H3 tail clipping has been found in several organisms such as yeast, chicken, and humans and is suggested to regulate gene expression and histone modifications by removing the histone H3 N-terminal region (Santos-Rosa et al., 2009; Mandal et al., 2013). In *Drosophila*, however, it is possible that histone H3 tail clipping activity is not prevalent, so the anti-H3K23me3 antibody was not able to recognize the H3K23me3 modification.

Inspection of the histone H3 tail sequence reveals interesting parallels between the H3K4 and H3K23 residues. As shown in figure 5-1A, the H3K23 residue is flanked by an identical set of residues that we have shown can influence binding of PHD2 to H3K4me3. We speculate that the PHD1 binding to H3K23me3 could be able to be inhibited by phosphorylation of flanking Thr 22, and enhanced by acetylation of Lys 27 and phosphorylation of Ser 28. The absence of reagents (antibodies) that would allow us to detect these residues precludes further *in vivo* analysis of these modifications, however, it would be tempting to speculate the binding of PHD1 to multiply modified H3K23me3 peptides would be regulated by flanking T22 phosphorylation and phosphoacetylation of H3K27 and H3S28. Inspection of the amino acid sequence of the PHD1 domain compared with the PHD2 domain shows that core tri-

methylated lysine binding pocket residues (Y10, Y17, Y23, W32) are conserved as expected. However, charged residues, that we previously speculated hinder or interact with negatively charged phosphate groups of H3T3p and H3S10p, are divergent (Figure 5-1 B). In the PHD1 domain, some of the negatively charged residues (E44, E46, D49 and E50 in PHD2) which we showed could repulse negatively charged H3T3p and inhibit PHD2 binding to H3K4me3, are conserved, but two positively charged lysine residues (K46, K47 in PHD1) also exist that could favour binding with a phosphate group on H3T22 and thus enhance the binding of PHD1 to H3K23me3. Moreover, one of the positively charged residues (R36 in PHD2) which we speculated interacted with negatively charged H3S10p, was changed to a negatively charged residue in the PHD1 domain (D36 in NURF301 PHD1 and E36 in BPTF PHD1). This suggests that flanking phosphorylation at Ser 28 may not enhance the binding of the PHD1 domain to H3K23me3, but reduce binding. Thus it is possible that the PHD1 domain binding to H3K23me3 may be regulated by phospho-methyl and phospho-acetyl switches but in an opposite manner to PHD2 binding to H3K4me3.

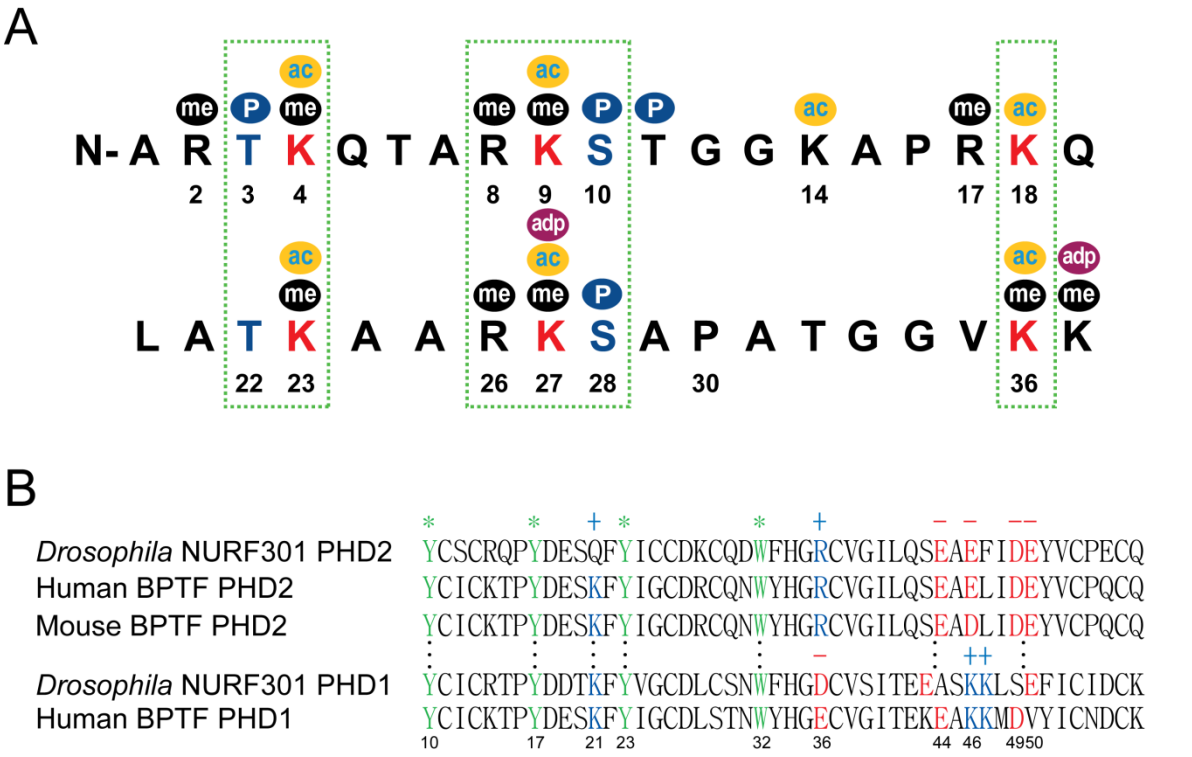


Figure 5-1. Protein sequence alignment of histone H3 N-terminal sequence and PHD2/PHD1 domains. (A) N-terminal sequence alignment of H3 (1-19) and (20-37) shows conservation of modification target residues that flank H3K4 and H3K23. (B) PHD2 and PHD1 domain sequence alignment shows conserved trimethyl lysine-binding residues (green), while positively charged residues (blue) and negatively charged residues (red) which could interact with flanking phosphorylation vary.

5.1.3 Modified histone binding specificity of the Bromodomain

The C-terminal bromodomain showed more enhanced binding to multiply acetylated H4 (K5, K8, K12, K16) compared to the single H4 K16 acetylation. This property of bromodomains was also found in recent study showing high-affinity binding of the BRDT bromodomain to multiple acetylated histone tails (Morinière et al., 2009; Filippakopoulos et al., 2012). Analysis of the crystal structures of bromodomains in a complex with multiply acetylated histone tails suggests that the bromodomain is only able to bind to two acetylated residues at any one time (Morinière et al., 2009; Filippakopoulos et al., 2012). The bromodomain does not bind simultaneously to all the acetylated residues on tetra-acetylated histone tails, rather it

is more likely that the bromodomain is in rapid exchange between flanking di-acetylated residues. The apparent increase in binding to the tetra-acetylated histone tails can be explained as an effect on avidity rather than affinity. Effects on avidity are frequently used to describe antibody-antigen interactions in immunology. "Affinity" is the strength of binding of one molecule to a substrate, while "Avidity" is the sum of total strength of binding of more than one molecule to substrates. They are not separate concepts but correlated, as the strength of "Avidity" depends on the "Affinity" and the valency of interactions (Murphy et al., 2008). Therefore, binding preference of NURF301 bromodomain to multiply acetylated histone H4 tails might be correlated to the high "Avidity" of the bromodomain binding pocket that has "Affinity" for two acetylated lysines on histone H4 tail.

5.1.4 Modified histone binding specificity of the N-terminal PHD domain

Our data suggests that the PHD domain does not have clear modified-histone binding specificity, but binds generally the free N-terminal tails of H4. This is consistent with BLAST comparison of the NURF301 PHD domain that indicated that the PHD domain is most similar to the PHD domains BHC80 (also known as PHF21A, a plant homeodomain (PHD) finger-containing protein, Lan et al., 2007) and AIRE (Human autoimmune regulator, Koh et al., 2010) both of which have been shown to bind to unmodified histone tails. However, the observed binding interactions of the PHD domain were weak and it is equally possible that the PHD domain may serve no strong role in nucleosome recognition. The PHD may act as a protein-protein interaction surface that serves a structural role in the organisation of the NURF complex. Alternatively, it is possible that the PHD domain acts as the reader of as yet unknown ligands. Early reports suggested that the PHD domain of ING2 could bind to nuclear phosphoinositides (Gozani et al., 2003), providing an attractive mechanism by which chromatin remodelling activity could be regulated by signal transduction pathways.

5.1.5 Multivalent histone recognition by NURF

An important question is whether the modified histone binding specificities observed using the isolated PHD and bromodomains are the same as those of the full-length NURF301 or the assembled NURF complex. It has been shown that the modified histone recognition of the tandem PHD domains of the muscle regulator DPF3 is not shared by the individual PHD fingers (Lange et al., 2008). We have not exhaustively addressed this issue by expressing all combinations of PHD fingers and bromodomains, however, in the case of PHD1 and PHD2 we have been able to show that a tandem PHD1+PHD2 GST fusion protein shows binding to modified histone peptide arrays that is a composite of the binding specificities of the individual domains (data not shown). New recognition specificities were not observed.

We believe that the modified histone recognition of full-length NURF301 and the NURF complex will be a composite of the modified histone binding of the individual domains. In combination histone tail recognition by the PHD2 and PHD1 fingers and the bromodomain will allow the recognition of multiply-modified nucleosomes and stabililise binding through increased valency of interaction. As shown in figure 5-2, this allows the construction of detailed models of recruitment and interaction of NURF301 with nucleosomes via the PHD domains and bromodomain (Figure 5-2). Some support for this is provided by the analysis of interactions between H3K4me3 and H4K16ac where our laboratory has shown overlap between these marks and NURF *in vivo* (Kwon et al., 2009) and *in vitro* studies using multiply-modified nucleosomes shows enhanced recognition of H3K4me3, H4K16ac modified nucleosomes by BPTF (Ruthenburg et al., 2011). We speculate that the presence of multiple protein-interaction modules could generate high target selectivity and increased binding affinity with substrates due to avidity through simultaneous binding of several interaction domains (Wu et al., 2009; Muller et al., 2011).

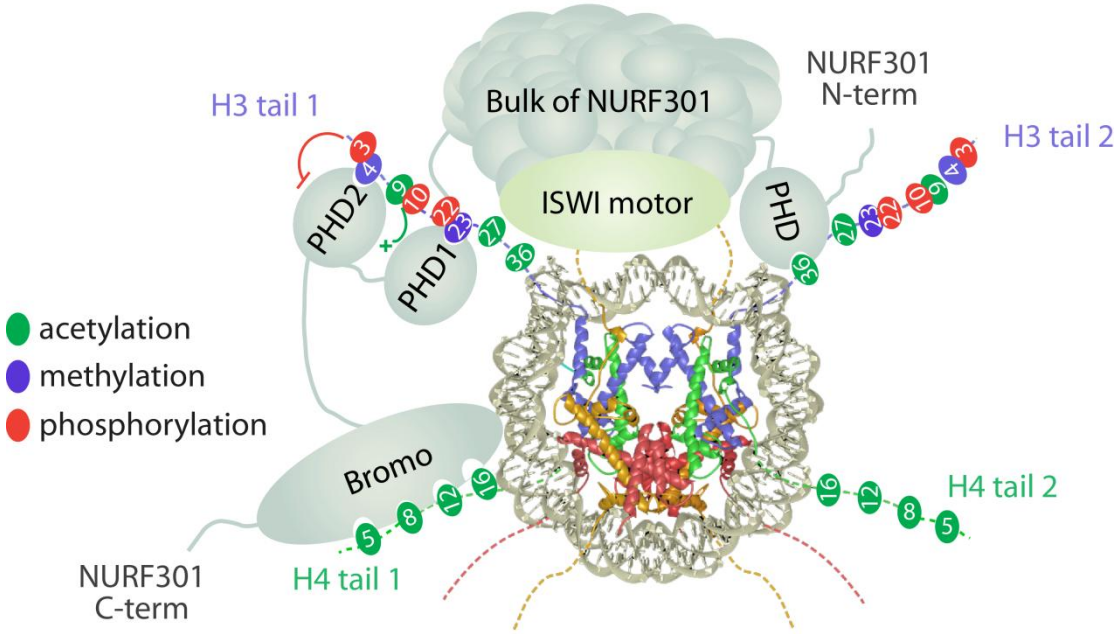


Figure 5-2. Predicted binding model of *Drosophila* NURF and histone modifications based on our data (P. Badenhorst (unpublished data)). Modified histone binding specificities of domains (PHD, PHD1, PHD2 and Bromodomain) may lead the NURF recruitment to a nucleosome as indicated.

5.2 NURF may interact with other epigenetic regulators: Haspin and Gcn5

The deposition and effect of histone post-translational modifications requires many proteins which can be classified in three simple groups: epigenetic writers, erasers and readers (Figure 5-3). The epigenetic "writers" attach modifications to DNA or histones (acetyl-transferases, methyl-transferases, kinases), "erasers" remove these marks (deacetylases, demethylases, phosphatases), while "readers" are proteins that are able to bind to the modifications through binding domains such as the PHD domain, bromodomain and chromodomain. The “readers” recognize specifically modified sites producing changes in gene expression (Gardner et al., 2011). We have found that there is a close relationship between the “reader” NURF and two histone modification "writers" – the H3T3-specific kinase Haspin, and the H3K9 and K14 specific acetyltransferase Gcn5.

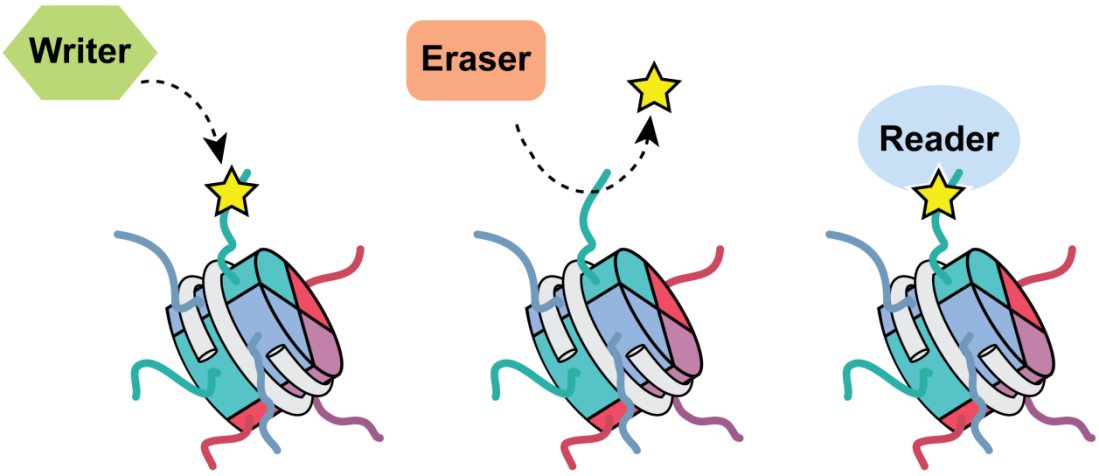


Figure 5-3. Schematic illustrating the concept "writer", "eraser" and "reader". The "writer" places PTMs on histone proteins (left), an "eraser" removes such modifications from histone proteins (middle), and a "reader" interacts with these covalent modifications (right) to mediate diverse downstream processes. Figure was adapted from Gardner et al. (2011).

Here we have shown that H3T3 phosphorylation blocks binding of the PHD2 domain to H3K4me3 by immunostaining assay *in vivo*. Double immunostaining of H3T3 phosphorylation and NURF301 showed no overlap during any stage of the cell cycle indicating that H3T3p can act as a mark to eject NURF from chromatin. Haspin is known to be required to phosphorylate histone H3T3 and this modification is critical for chromosome condensation during metaphase (Dai et al., 2005). In previous studies, it was shown that *Drosophila Nurf301* mutants cause aberrant structure of interphase chromosomes (Badenhorst et al. 2002). Surprisingly, however, metaphase chromosome structure is unaffected in *Nurf301* mutants (Badenhorst and Kwon unpublished data). As the Haspin kinase mainly affects metaphase chromosome structure by phosphorylation of H3T3 (Dai et al., 2005; Higgins, 2009), it is possible that H3T3 phosphorylation by Haspin allows a switch that removes NURF from chromosomes by masking H3K4me3 sites that might normally recruit NURF. This ejection of NURF may be important in moving the cell cycle from interphase into metaphase.

However, we have not been able to conclusively show that H3T3p is required to eject NURF from chromatin by either ectopically providing or removing this mark. To functionally dissect the interaction between NURF and Haspin we tried a number of genetic approaches to ablate Haspin by generating *Haspin* mutant flies. But neither approach including transposon excision mutation or X-ray mutagenesis were successful. We attempted to use RNAi-knockdown to ablate *Haspin* but existing *Haspin* siRNA strains were unable to reduce H3T3p levels indicating that Haspin levels were not affected by *Haspin* siRNAs when expressed in blood cells. One approach that may be attempted is to use inhibitors of Haspin. Recent studies have identified small-molecule inhibitors of Haspin that could be used for this study (De Antoni et

al., 2012; Wang et al., 2012). One potential concern with inhibitors is the possibility of off-target effects when using an inhibitor. Moreover, the use of Haspin inhibitors require that cells are synchronized at the G2/M boundary prior to use of the inhibitor as effects are also observed of these inhibitors on interphase cells (De Antoni et al., 2012; Wang et al., 2012). It is challenging to synchronize *Drosophila* cultured cells which would require that these experiments be performed using mammalian cultured cells. Using synchronized HeLa cells and Haspin inhibitors the interplay of NURF and Haspin in chromosomal condensation and mitosis would be able to be determined.

Unlike the mark/writer pair of H3T3p/Haspin, which inhibit NURF binding, we speculate that writers Gcn5 (the acetylase for H3K9) and Jil-1 (the H3S10 kinase, Wang et al., 2001) enhance binding of the PHD2 domain to H3K4me3. Our immunostaining of polytene chromosomes is consistent with Gcn5 and Jil-1 stabilizing NURF binding to chromatin as we see overlap between NURF and H3K9ac/S10p modifications. The availability of mutants Gcn5 has allowed use to functionally dissect whether the phospho-acetyl H3K9ac/S10p mark is required for NURF recruitment. We observe in our study that loss of Gcn5 influences localization of the NURF301 *in vivo* by double immunostaining of polytene chromosomes. This data is consistent with previously published interactions of NURF301 with the Gcn5 containing complexes ATAC (known to acetylate K5/K12 of histone H4, Ciurciu et al., 2006) and SAGA (known to acetylate K9/K14 of histone H3, Carre et al., 2005) that were reported by Carre and co-workers (Carré et al., 2008). These authors suggested that NURF301 interacts with the ATAC complex, which may impact on H4 tail recognition by the bromodomain. ATAC is also required for H4K16 acetylation, which has been shown to influence nucleosome sliding by ISWI which is the catalytic subunit of the NURF, ACF and

CHRAC complexes in *Drosophila* (Carre et al., 2005; Suganuma et al., 2008). However, it is also possible that the SAGA Gcn5-containing complexes is responsible for the affect on NURF301 recruitment and nucleosome remodelling. SAGA is required for H3K9ac and has been shown to synergise with NURF nucleosome remodelling to facilitate *in vitro* transcription of chromatin templates (Mizuguchi et al., 2001).

This suggests that both Gcn5 complexes regulate the level of histone acetylation which might affect NURF301 binding to chromatin and remodelling activity. Altough our data suggests a positive correlation between histone acetylation and the recruitment and function of NURF, some other reports have suggested an antagonistic relationship. For example, acetylation at lysine 12 and 16 of the histone H4 tail have been reported to inhibit substrate recognition by ISWI and reduce its ATPase activity (Clapier et al., 2002; Corona et al., 2002). Although some workers have shown that acetylation enhances sliding by ISWI-containing remodelling complexes (Suganuma et al., 2008), other reports have suggested that modified mononucleosomes containing histone H4K16ac inhibit sliding activity by ACF (Shogren-Knaak et al., 2006). This indicates that relationship between acetylation and chromatin remodelling activity is still controversial. Thus, the mechanism of how these modifications affect NURF remodelling activity and chromatin structural changes remains to be determined. For this reason, it will be of interesting to define the recruitment as well as remodelling activity of NURF for nucleosomes with/without these modifications.

5.3 Generation of epitope tagged NURF301-A/B and -C isoforms would extend our understanding of NURF

Previous research in our laboratory has shown that NURF301 expresses three distinct isoforms: the full-length NURF301 A and B isoforms, and the C-terminally truncated NURF301 C isoform (Kwon et al., 2009). The existence of similar multiple splice variants of mammalian BPTF remains to be conclusively established but our data here indicates the existence of transcript variants that contain or lack reader domains such as the PHD1 domain. This has the potential to encode variant NURF complexes that have distinct targeting specificity and activity. Epitope-tagged variants of these isoforms provide important reagents for determining these distinct functions. Previously in our laboratory, we attempted to generate epitope tagged variants of the NURF301-A and NURF301-C isoforms using *Nurf301* cDNAs cloned into plasmid vectors in bacteria. However, the large size of the resultant plasmids and toxicity of the cDNAs meant that the bacteria grew poorly and inactivated the cDNA during culture (Badenhorst and Kwon unpublished data). To circumvent this, here we have used a more controllable expression system and genomic constructs containing introns. Using BAC plasmids harbouring *Nurf301* genomic DNA and the GalK recombineering system, we successfully generated CTAP tagged NURF301-A and NURF301-C constructs. Those constructs have been injected to flies containing φ C31 integrase system (Venken et al., 2006; Bischof et al., 2007) and we generated flies that express CTAP tagged full-length NURF301-A/B isoform and C-terminally truncated NURF301-C isoform transgenic flies.

Polytene chromosome immunostaining data using these epitope tagged transgenic lines, showed that NURF301-A/B and NURF301-C colocalize at some regions, but genes where NURF301-A/B or NURF301-C did not overlap were also detected. It is possible that lack of

overlap at these regions may be due to technical reasons – a difference in the relative intensity of isoforms rather than absence of localization – but it is most likely that this reflects the, loss of important histone modification binding modules (C-terminal PHD1, PHD2 and bromodomain) in the NURF301-C isoform that restricts localization and function. This provides a clear example how alternative splicing could regulate chromatin modifying and remodelling factors by generating functionally distinct variants in which domains critical for function are substituted (Lois et al., 2007). These results are consistent with reports that the function and localization of the human NURF can be regulated by differential splicing of SNF2L, the catalytic subunit of human NURF (Barak et al., 2004; Lazzaro et al., 2008), and studies showing that mutation or loss of histone binding modules change target gene expression by altering histone binding specificity or interactions with other transcription factors (Matangkasombut and Buratowski, 2003; Chen et al., 2008).

Alternative splicing of NURF301 can affect the diversity of chromatin remodelling complexes in two ways: either generating functional diversity within the same cell, or generating functional diversity between distinct cell types. If two splice variants are expressed in the same cells, as we have shown occurs in salivary glands (polytene chromosomes), alternative splicing can generate chromatin remodelling complexes that are possess different functions due to the chromatin territories to which they are targeted. Thus the NURF301-A containing complexes would be expected to be targeted to regions that contain the combinations of histone modifications we have identified here. NURF301-C complexes, in contrast, which do not contain the histone-modification binding PHD1, PHD2 and bromodomains would not be anticipated to be enriched in these regions. The resolution of polytene chromosome bands precludes us from mapping the extent of overlap between NURF301-A, NURF301-C and

histone modifications and we have made a major effort to map these two isoforms on chromatin by chromatin immunoprecipitation coupled sequencing (ChIP-Seq). However, as shown in results section 4.2.9 (Table 4-1), whether due to the nature of the NURF-chromatin interaction or the low efficiency of NURF ChIP and the low cell numbers that are used in our hemocyte ChIP procedures, we have not been able to generate ChIP-Seq profiles of NURF301-A/B or NURF301-C isoforms.

Alternative splicing of NURF301 could also be used to generate tissue and/or stage specific NURF variants. Previously, the whole genome expression profiling of *NURF* null mutant and *Nurf301ΔC* mutant revealed that a subset of genes require full-length NURF301-A/B. Loss of the full-length NURF A/B isoform appeared to selectively affect genes involved in primary spermatocyte development, for example the *fuzzy onions (fzo)* gene (Hales and Fuller, 1997; Kwon et al., 2009). Our male testes immunostaining data was also consistent with these findings as we showed that NURF301-A/B was only expressed in the later stage of spermatogenesis, while NURF301-C was observed in all stages. This suggest the NURF301-A/B and NURF301-C isoforms can also have unique functions as a result of differential chromatin targeting as well as differential tissue expression. This latter result is consistent with an emerging body of data where functional diversity of SWI/SNF complexes is generated by alternative incorporation of subunits at different stages and in different tissues. During neural development switching occurs in SWI/SNF complexes (Olave et al., 2002; Lessard et al., 2007; Yoo et al., 2009; Vogel-Ciernia et al., 2012) with the exchange of BAF53a and BAF45a (PHF10) for the related BAF53b and BAF45b subunits. Equally a stem cell-specific SWI/SNF complex with a distinct subunit composition has been identified (Ho et al., 2009).

5.4 Final conclusion and future work

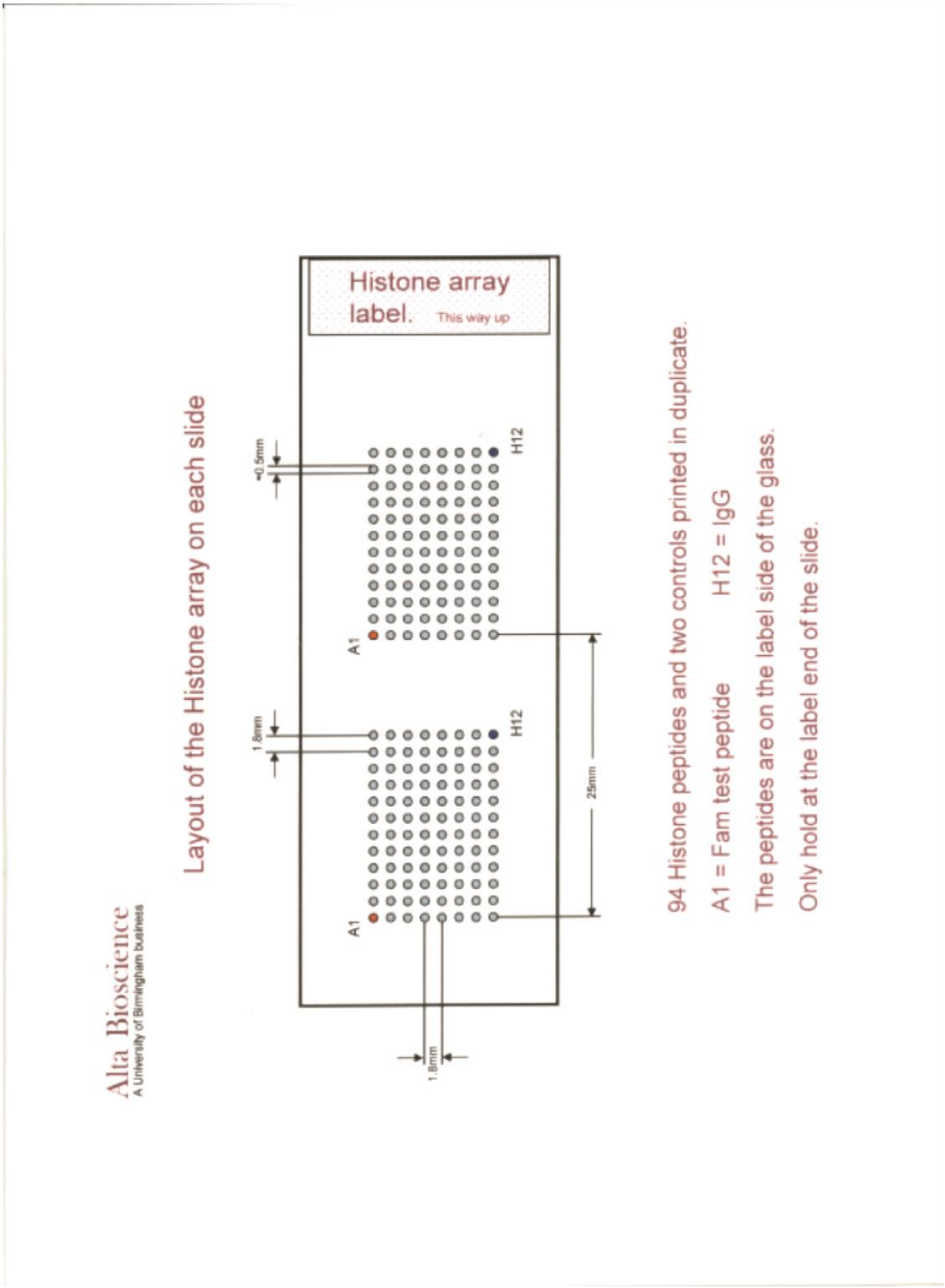
In summary, each chromatin targeting module, the PHD, PHD1 and PHD2 domains and bromodomain in NURF301 exhibits different binding specificity. The PHD2 domain binds specifically to H3K4me3, but binding is inhibited by phosphorylation of flanking Thr 3, and enhanced by acetylation of Lys 9 and phosphorylation of Ser 10. The PHD1 specifically interacts with H3K23me3, while PHD does not have binding specificity but generally binds to the free N-terminal tails of H4. The Bromodomain binds preferentially to a multiply acetylated H4 tail. These results allow us to draw models of recruitment and interaction of NURF301 with nucleosomes via the PHD domains and bromodomain. These data also suggest interactions between NURF301 with histone modifying enzymes, the H3T3-specific kinase Haspin, and the H3K9 and K14 specific acetyltransferase Gcn5. Finally, we have successfully generated CTAP tagged *Nurf301-A/B* and *C* BAC constructs and transgenic flies that provide important reagents for determining distinct functions of the NURF301-A/B and C isoforms by looking at immunostaining of different cell types and ChIP sequencing.

In future work it will be important to demonstrate functional requirements of the PHD domains and bromodomain for recruitment of NURF *in vivo*. One way in which this can be achieved is by using our established recombineering method to generate CTAP- or NTAP-tagged NURF301-A/B constructs containing deletions of the PHD, PHD1 and PHD2 domains and bromodomain respectively. Using these constructs, we would be able to generate transgenic flies expressing specific domain deleted forms of NURF301-A/B and analyze biological functions of the each domain precisely by comparing with wild type and NURF-A/B and C full deletion mutant.

To extend our finding that the PHD1 domain is present in human BPTF and that alternative splice variants of BPTF may exist in which the PHD fingers and Bromodomains are alternatively incorporated, we will determine the tissue- and disease-specific splicing patterns of BPTF. The Birmingham Human Tissue Biorepository collects and stores human tissue for use in biomedical research. We will obtain tissue samples from the biorepository, isolate mRNA, generated cDNA and profile BPTF isoform variants using real time PCR. A particular focus will be to determine whether PHD1 is alternatively incorporated in ovary and testes samples.

Lastly, to generate ChIP-Seq profiles of NURF, we will use *Drosophila* S2 cells for our next ChIP experiment using anti-NURF301 antibody (recognizes all NURF301-A.B and C isoforms). We can generate large numbers of S2 cells which will allow enough ChIP DNA to be prepared even if ChIP efficiency is low. Although we will not able to discriminate NURF301-A/B or NURF301-C isoforms directly it will provide a first NURF profile. By comparing S2 cells and Kc cells, we may also be able to explore sex dependent NURF distributions.

CHAPTER 6. APPENDIX



Appendix 1. Layout of the histone array on the slide.

Alta Bioscience
A University of Birmingham business

Histone H3 library

A1		Control 1	E1	*Biotin-spacer-*	KAPRKQLATΦAARKSAPATGG
A2	*Biotin-spacer-*	ARTKQTARKSTGGKAPRKQLA	E2	*Biotin-spacer-*	KAPRKQLATΠAARKSAPATGG
A3	*Biotin-spacer-*	AΞTKQTARKSTGGKAPRKQLA	E3	*Biotin-spacer-*	KAPRKQLATΘAARKSAPATGG
A4	*Biotin-spacer-*	AΨTKQTARKSTGGKAPRKQLA	E4	*Biotin-spacer-*	KAPRKQLATΘAAΨKSAPATGG
A5	*Biotin-spacer-*	AΨΩKQTARKSTGGKAPRKQLA	E5	*Biotin-spacer-*	KAPRKQLATΘAAΨΔSAPATGG
A6	*Biotin-spacer-*	AΨΩΘQTARKSTGGKAPRKQLA	E6	*Biotin-spacer-*	KAPRKQLATKAAΞKSAPATGG
A7	*Biotin-spacer-*	AΨTΘQTARKSTGGKAPRKQLA	E7	*Biotin-spacer-*	KAPRKQLATKAAΨKSAPATGG
A8	*Biotin-spacer-*	ARΩKQTARKSTGGKAPRKQLA	E8	*Biotin-spacer-*	KAPRKQLATKAAΨΔSAPATGG
A9	*Biotin-spacer-*	ARΩΘQTARKSTGGKAPRKQLA	E9	*Biotin-spacer-*	KAPRKQLATKAAΨΔΣAPATGG
A10	*Biotin-spacer-*	ARTΦQTARKSTGGKAPRKQLA	E10	*Biotin-spacer-*	KAPRKQLATKAAΨKΣAPATGG
A11	*Biotin-spacer-*	ARTΠQTARKSTGGKAPRKQLA	E11	*Biotin-spacer-*	KAPRKQLATKAARΔSAPATGG
A12	*Biotin-spacer-*	ARTΘQTARKSTGGKAPRKQLA	E12	*Biotin-spacer-*	KAPRKQLATKAARΔΣAPATGG
B1	*Biotin-spacer-*	ARTΘQTARΔSTGGKAPRKQLA	F1	*Biotin-spacer-*	KAPRKQLATKAARΦSAPATGG
B2	*Biotin-spacer-*	ARTΘQTARΘSTGGKAPRKQLA	F2	*Biotin-spacer-*	KAPRKQLATKAARΠSAPATGG
B3	*Biotin-spacer-*	ARTKQTARΔSTGGKAPRKQLA	F3	*Biotin-spacer-*	KAPRKQLATKAARΘSAPATGG
B4	*Biotin-spacer-*	ARTKQTARΔΣTGGKAPRKQLA	F4	*Biotin-spacer-*	KAPRKQLATKAAΨΘSAPATGG
B5	*Biotin-spacer-*	ARTKQTARΔSΩGGKAPRKQLA	F5	*Biotin-spacer-*	KAPRKQLATKAARΘΣAPATGG
B6	*Biotin-spacer-*	ARTKQTARΔΣΩGGKAPRKQLA	F6	*Biotin-spacer-*	KAPRKQLATKAAΨΘΣAPATGG
B7	*Biotin-spacer-*	ARTKQTARΦSTGGKAPRKQLA	F7	*Biotin-spacer-*	KAPRKQLATKAARKΣAPATGG
B8	*Biotin-spacer-*	ARTKQTARΠSTGGKAPRKQLA	F8	*Biotin-spacer-*	KQLATKAARKΣAPAΩGGVKKP
B9	*Biotin-spacer-*	ARTKQTARΘSTGGKAPRKQLA	F9	*Biotin-spacer-*	SAPATGGVKKPHRYRPGTVAL
B10	*Biotin-spacer-*	ARTKQTARΘSΩGGKAPRKQLA	F10	*Biotin-spacer-*	SAPAΩGGVKKPHRYRPGTVAL
B11	*Biotin-spacer-*	ARTKQTARΘΣTGGKAPRKQLA	F11	*Biotin-spacer-*	SAPATGGVΦKPHRYRPGTVAL
B12	*Biotin-spacer-*	ARTKQTARΘΣΩGGKAPRKQLA	F12	*Biotin-spacer-*	SAPATGGVΠKPHRYRPGTVAL
C1	*Biotin-spacer-*	ARTKQTARKΣTGGKAPRKQLA	G1	*Biotin-spacer-*	SAPATGGVΘKPHRYRPGTVAL
C2	*Biotin-spacer-*	ARTKQTARKSΩGGKAPRKQLA	G2	*Biotin-spacer-*	SAPAΩGGVΘKPHRYRPGTVAL
C3	*Biotin-spacer-*	ARTKQTARKΣΩGGKAPRKQLA	G3	*Biotin-spacer-*	SAPATGGVKΦPHRYRPGTVAL
C4	*Biotin-spacer-*	ARTKQTARKΣΩGGΔAPRKQLA	G4	*Biotin-spacer-*	SAPATGGVKΠPHRYRPGTVAL
C5	*Biotin-spacer-*	ARTKQTARKΣΩGGΘAPRKQLA	G5	*Biotin-spacer-*	SAPATGGVKΘPHRYRPGTVAL
C6	*Biotin-spacer-*	ARTKQTARKSΩGGΔAPRKQLA	G6	*Biotin-spacer-*	SAPAΩGGVKΘPHRYRPGTVAL
C7	*Biotin-spacer-*	ARTKQTARKSΩGGΘAPRKQLA	G7	*Biotin-spacer-*	SAPATGGVΘΘPHRYRPGTVAL
C8	*Biotin-spacer-*	ARTKQTARKSTGGΔAPRKQLA	G8	*Biotin-spacer-*	SAPAΩGGVΘΘPHRYRPGTVAL
C9	*Biotin-spacer-*	ARTKQTARKSTGGΔAPΨKQLA	G9	*Biotin-spacer-*	VREIAQDFKTDLRFQSSAVMA
C10	*Biotin-spacer-*	ARTKQTARKSTGGΦAPRKQLA	G10	*Biotin-spacer-*	VREIAQDFΦTDLRFQSSAVMA
C11	*Biotin-spacer-*	ARTKQTARKSTGGΠAPRKQLA	G11	*Biotin-spacer-*	VREIAQDFΠTDLRFQSSAVMA
C12	*Biotin-spacer-*	ARTKQTARKSTGGΘAPRKQLA	G12	*Biotin-spacer-*	VREIAQDFΘTDLRFQSSAVMA
D1	*Biotin-spacer-*	ARTKQTARKSTGGΘAPΨKQLA	H1	*Biotin-spacer-*	CAIHAKRVTIMPKDIQLARRI
D2	*Biotin-spacer-*	ARTKQTARKSTGGKAPΞKQLA	H2	*Biotin-spacer-*	CAIHAΔRVTIMPKDIQLARRI
D3	*Biotin-spacer-*	ARTKQTARKSTGGKAPΨKQLA	H3	*Biotin-spacer-*	CAIHAΔRVΩIMPKDIQLARRI
D4	*Biotin-spacer-*	KAPRKQLATKAARKSAPATGG	H4	*Biotin-spacer-*	CAIHAKRVΩIMPKDIQLARRI
D5	*Biotin-spacer-*	KAPΨΔQLATKAARKSAPATGG	H5	*Biotin-spacer-*	CAIHAKRVTIMPΔDIQLARRI
D6	*Biotin-spacer-*	KAPRΔQLATKAARKSAPATGG	H6	*Biotin-spacer-*	CAIHAKRVΩIMPΔDIQLARRI
D7	*Biotin-spacer-*	KAPRΔQLATΘAARKSAPATGG	H7	*Biotin-spacer-*	CAIHAΔRVΩIMPΔDIQLARRI
D8	*Biotin-spacer-*	KAPRΔQLATΔAARKSAPATGG	H8	*Biotin-spacer-*	KRVTIMPKDIQLARRIRGERA-acid
D9	*Biotin-spacer-*	KAPRΔQLATΔAARKSAPATGG	H9	*Biotin-spacer-*	KRVTIMPKDIQLAΞRIRGERA-acid
D10	*Biotin-spacer-*	KAPRKQLATΔAARKSAPATGG	H10	*Biotin-spacer-*	KRVTIMPKDIQLAΨRIRGERA-acid
D11	*Biotin-spacer-*	KAPRKQLATΔAAΨKSAPATGG	H11	*Biotin-spacer-*	KRVTIMPΔDIQLAΨRIRGERA-acid
D12	*Biotin-spacer-*	KAPRKQLATΔAAΨΔSAPATGG	H12		Control 2

Δ = acetyl-Lysine	Φ = monomethyl-Lys	Π = dimethyl-Lys	Θ = trimethyl-Lys
Σ = phospho-Ser	Ω = phospho-Thr	Ξ = monomethyl-Arg	Ψ = asym dimethyl-Arg
Spacer = aminohexanoic acid, Ahx		All the peptides have C-terminal amide groups unless specified.	

19 March 2010

www.altabioscience.bham.ac.uk/pdfs/Histone H3 library.pdf

Appendix 2. Histone H3 peptide library array.

Alta Bioscience
A University of Birmingham business

Histone H3, H4 N-terminal library

A1	Control 1	E1	Ac-KQLATΦAARKSAPATGGVKKP -spacer-Biotin
A2	ARTKQTARKSTGGKAPRKQLA -spacer-Biotin	E2	Ac-KQLATΠAARKSAPATGGVKKP -spacer-Biotin
A3	AΞTKQTARKSTGGKAPRKQLA -spacer-Biotin	E3	Ac-KQLATΘAARKSAPATGGVKKP -spacer-Biotin
A4	AΨTKQTARKSTGGKAPRKQLA -spacer-Biotin	E4	Ac-KQLATΘAAΨKSAPATGGVKKP -spacer-Biotin
A5	AΨΩKQTARKSTGGKAPRKQLA -spacer-Biotin	E5	Ac-KQLATΘAAΨΔSAPATGGVKKP -spacer-Biotin
A6	AΨΩΘQTARKSTGGKAPRKQLA -spacer-Biotin	E6	Ac-LATKAARKSAPATGGVKKPHR -spacer-Biotin
A7	AΨTΘQTARKSTGGKAPRKQLA -spacer-Biotin	E7	Ac-LATKAAΞKSAPATGGVKKPHR -spacer-Biotin
A8	ARΩKQTARKSTGGKAPRKQLA -spacer-Biotin	E8	Ac-LATKAAΨKSAPATGGVKKPHR -spacer-Biotin
A9	ARΩΘQTARKSTGGKAPRKQLA -spacer-Biotin	E9	Ac-LATKAAΨΔSAPATGGVKKPHR -spacer-Biotin
A10	ARTΦQTARKSTGGKAPRKQLA -spacer-Biotin	E10	Ac-LATKAAΨΔΣAPATGGVKKPHR -spacer-Biotin
A11	ARTΠQTARKSTGGKAPRKQLA -spacer-Biotin	E11	Ac-LATKAAΨKΣAPATGGVKKPHR -spacer-Biotin
A12	ARTΘQTARKSTGGKAPRKQLA -spacer-Biotin	E12	Ac-LATKAARΔSAPATGGVKKPHR -spacer-Biotin
B1	ARTΘQTARΔSTGGKAPRKQLA -spacer-Biotin	F1	Ac-LATKAARΔΣAPATGGVKKPHR -spacer-Biotin
B2	ARTΘQTARΘSTGGKAPRKQLA -spacer-Biotin	F2	Ac-LATKAARΦSAPATGGVKKPHR -spacer-Biotin
B3	ARTKQTARΔSTGGKAPRKQLA -spacer-Biotin	F3	Ac-LATKAARΠSAPATGGVKKPHR -spacer-Biotin
B4	ARTKQTARΔΣTGGKAPRKQLA -spacer-Biotin	F4	Ac-LATKAARΘSAPATGGVKKPHR -spacer-Biotin
B5	ARTKQTARΔSΩGGKAPRKQLA -spacer-Biotin	F5	Ac-LATKAAΨΘSAPATGGVKKPHR -spacer-Biotin
B6	ARTKQTARΔΣΩGGKAPRKQLA -spacer-Biotin	F6	Ac-LATKAARΘΣAPATGGVKKPHR -spacer-Biotin
B7	ARTKQTARΦSTGGKAPRKQLA -spacer-Biotin	F7	Ac-LATKAAΨΘΣAPATGGVKKPHR -spacer-Biotin
B8	ARTKQTARΠSTGGKAPRKQLA -spacer-Biotin	F8	Ac-LATKAARKΣAPATGGVKKPHR -spacer-Biotin
B9	ARTKQTARΘSTGGKAPRKQLA -spacer-Biotin	F9	SGRGKGGKGLGKGGAKRHRKV -spacer-Biotin
B10	ARTKQTARΘSΩGGKAPRKQLA -spacer-Biotin	F10	ΣGRGKGGKGLGKGGAKRHRKV -spacer-Biotin
B11	ARTKQTARΘΣTGGKAPRKQLA -spacer-Biotin	F11	SGΞGKGGKGLGKGGAKRHRKV -spacer-Biotin
B12	ARTKQTARΘΣΩGGKAPRKQLA -spacer-Biotin	F12	SGΨGKGGKGLGKGGAKRHRKV -spacer-Biotin
C1	ARTKQTARKΣTGGKAPRKQLA -spacer-Biotin	G1	SGΨGΔGGKGLGKGGAKRHRKV -spacer-Biotin
C2	ARTKQTARKSΩGGKAPRKQLA -spacer-Biotin	G2	ΣGΨGΔGGKGLGKGGAKRHRKV -spacer-Biotin
C3	ARTKQTARKΣΩGGKAPRKQLA -spacer-Biotin	G3	SGRGΔGGKGLGKGGAKRHRKV -spacer-Biotin
C4	Ac-KQTARKΣΩGGΔAPRKQLATKA -spacer-Biotin	G4	SGRGΔGGΔGLGΔGGAKRHRKV -spacer-Biotin
C5	Ac-KQTARKΣΩGGΘAPRKQLATKA -spacer-Biotin	G5	SGRGΔGGΔGLGKGGAKRHRKV -spacer-Biotin
C6	Ac-KQTARKSΩGGΔAPRKQLATKA -spacer-Biotin	G6	SGRGKGGΔGLGKGGAKRHRKV -spacer-Biotin
C7	Ac-KQTARKSΩGGΘAPRKQLATKA -spacer-Biotin	G7	SGRGKGGΔGLGΔGGAKRHRKV -spacer-Biotin
C8	Ac-KQTARΔSTGGΔAPRKQLATKA -spacer-Biotin	G8	SGRGKGGKGLGΦGGAKRHRKV -spacer-Biotin
C9	Ac-RKSTGGΔAPRKQLATKAARKS -spacer-Biotin	G9	SGRGKGGKGLGΠGGAKRHRKV -spacer-Biotin
C10	Ac-RKSTGGΔAPΨKQLATKAARKS -spacer-Biotin	G10	SGRGKGGKGLGΘGGAKRHRKV -spacer-Biotin
C11	Ac-RKSTGGΦAPRKQLATKAARKS -spacer-Biotin	G11	SGRGKGGΔGLGΦGGAKRHRKV -spacer-Biotin
C12	Ac-RKSTGGΠAPRKQLATKAARKS -spacer-Biotin	G12	SGRGKGGΔGLGΘGGAKRHRKV -spacer-Biotin
D1	Ac-RKSTGGΘAPRKQLATKAARKS -spacer-Biotin	H1	Ac-GGKGLGΦGGAΔRHRKVLRDNI -spacer-Biotin
D2	Ac-RKSTGGΘAPΨKQLATKAARKS -spacer-Biotin	H2	Ac-GGKGLGΘGGAΔRHRKVLRDNI -spacer-Biotin
D3	Ac-RKSTGGKAPΞKQLATKAARKS -spacer-Biotin	H3	Ac-GGKGLGKGGAΔRHRKVLRDNI -spacer-Biotin
D4	Ac-RKSTGGKAPΨKQLATKAARKS -spacer-Biotin	H4	Ac-GKGGAKRHRKVLRDNIQGITK -spacer-Biotin
D5	Ac-RKSTGGΔAPRΔQLATKAARKS -spacer-Biotin	H5	Ac-GKGGAKRHRΔVLRDNIQGITK -spacer-Biotin
D6	Ac-GGKAPΨΔQLATKAARKSAPAT -spacer-Biotin	H6	Ac-GKGGAΔRHRΔVLRDNIQGITK -spacer-Biotin
D7	Ac-GGKAPRΔQLATKAARKSAPAT -spacer-Biotin	H7	Ac-GKGGAΔRHRΦVLRDNIQGITK -spacer-Biotin
D8	Ac-GGKAPRΔQLATΘAARKSAPAT -spacer-Biotin	H8	Ac-GKGGAΔRHRΘVLRDNIQGITK -spacer-Biotin
D9	Ac-GGKAPRΔQLATΔAARKSAPAT -spacer-Biotin	H9	Ac-GKGGAKRHRΦVLRDNIQGITK -spacer-Biotin
D10	Ac-KQLATΔAARKSAPATGGVKKP -spacer-Biotin	H10	Ac-GKGGAKRHRΠVLRDNIQGITK -spacer-Biotin
D11	Ac-KQLATΔAAΨKSAPATGGVKKP -spacer-Biotin	H11	Ac-GKGGAKRHRΘVLRDNIQGITK -spacer-Biotin
D12	Ac-KQLATΔAAΨΔSAPATGGVKKP -spacer-Biotin	H12	Control 2

Δ = acetyl-Lysine	Φ = monomethyl-Lys	Π = dimethyl-Lys	Θ = trimethyl-Lys
Σ = phospho-Ser	Ω = phospho-Thr	Ξ = monomethyl-Arg	Ψ = asym dimethyl-Arg
Spacer = aminohexanoic acid, Ahx		Ac- = N-terminal acetylation	

15 October 2009

www.altabioscience.bham.ac.uk/pdfs/Histone H3_H4 N-term library.pdf

Appendix 3. Histone H3, H4 N-terminal peptide library array.

Appendix 4. Layout of the Active Motif MODified histone peptide array

Active Motif

MODified™ Histone Peptide Array*

Catalog Nos. 13001 & 13005

Modification number	Peptide location	Peptide sequence	Name	Mod1	Mod2	Mod3	Mod4	N-terminus
1	A1	ARTKQTARKSTGGKAPRKQ	H3 1-19	unmod				Free
2	A2	ARme2sTKQTARKSTGGKAPRKQ	H3 1-19	R2me2s				Free
3	A3	ARme2aTKQTARKSTGGKAPRKQ	H3 1-19	R2me2a				Free
4	A4	ACitTKQTARKSTGGKAPRKQ	H3 1-19	R2Citr				Free
5	A5	ARpTKQTARKSTGGKAPRKQ	H3 1-19	T3P				Free
6	A6	ARTKme1QTARKSTGGKAPRKQ	H3 1-19	K4me1				Free
7	A7	ARTKme2QTARKSTGGKAPRKQ	H3 1-19	K4me2				Free
8	A8	ARTKme3QTARKSTGGKAPRKQ	H3 1-19	K4me3				Free
9	A9	ARTKacQTARKSTGGKAPRKQ	H3 1-19	K4ac				Free
10	A10	ARTKQTARme2sKSTGGKAPRKQ	H3 1-19	R8me2s				Free
11	A11	ARTKQTARme2aKSTGGKAPRKQ	H3 1-19	R8me2a				Free
12	A12	ARTKQTACitKSTGGKAPRKQ	H3 1-19	R8Citr				Free
13	A13	ARTKQTARKme1STGGKAPRKQ	H3 1-19	K9me1				Free
14	A14	ARTKQTARKme2STGGKAPRKQ	H3 1-19	K9me2				Free
15	A15	ARTKQTARKme3STGGKAPRKQ	H3 1-19	K9me3				Free
16	A16	ARTKQTARKacSTGGKAPRKQ	H3 1-19	K9ac				Free
17	A17	ARTKQTARKpSTGGKAPRKQ	H3 1-19	S10P				Free
18	A18	ARTKQTARKSpTGGKAPRKQ	H3 1-19	T11P				Free
19	A19	ARTKQTARKSTGGKacAPRKQ	H3 1-19	K14ac				Free
20	A20	ARme2spTKQTARKSTGGKAPRKQ	H3 1-19	R2me2s	T3P			Free
21	A21	ARme2sTKme1QTARKSTGGKAPRKQ	H3 1-19	R2me2s	K4me1			Free
22	A22	ARme2sTKme2QTARKSTGGKAPRKQ	H3 1-19	R2me2s	K4me2			Free
23	A23	ARme2sTKme3QTARKSTGGKAPRKQ	H3 1-19	R2me2s	K4me3			Free
24	A24	ARme2sTKacQTARKSTGGKAPRKQ	H3 1-19	R2me2s	K4ac			Free
25	B1	ARme2apTKQTARKSTGGKAPRKQ	H3 1-19	R2me2a	T3P			Free
26	B2	ARme2aTKme1QTARKSTGGKAPRKQ	H3 1-19	R2me2a	K4me1			Free
27	B3	ARme2aTKme2QTARKSTGGKAPRKQ	H3 1-19	R2me2a	K4me2			Free
28	B4	ARme2aTKme3QTARKSTGGKAPRKQ	H3 1-19	R2me2a	K4me3			Free
29	B5	ARme2aTKacQTARKSTGGKAPRKQ	H3 1-19	R2me2a	K4ac			Free
30	B6	ACitpTKQTARKSTGGKAPRKQ	H3 1-19	R2Citr	T3P			Free
31	B7	ACitTKme1QTARKSTGGKAPRKQ	H3 1-19	R2Citr	K4me1			Free
32	B8	ACitTKme2QTARKSTGGKAPRKQ	H3 1-19	R2Citr	K4me2			Free

33	B9	ACitTKme3QTARKSTGGKAPRKQ	H3 1-19	R2Citr	K4me3			Free
34	B10	ACitTKacQTARKSTGGKAPRKQ	H3 1-19	R2Citr	K4ac			Free
35	B11	ARpTKme1QTARKSTGGKAPRKQ	H3 1-19	T3P	K4me1			Free
36	B12	ARpTKme2QTARKSTGGKAPRKQ	H3 1-19	T3P	K4me2			Free
37	B13	ARpTKme3QTARKSTGGKAPRKQ	H3 1-19	T3P	K4me3			Free
38	B14	ARpTKacQTARKSTGGKAPRKQ	H3 1-19	T3P	K4ac			Free
39	B15	ARme2spTKme1QTARKSTGGKAPRKQ	H3 1-19	R2me2s	T3P	K4me1		Free
40	B16	ARme2spTKme2QTARKSTGGKAPRKQ	H3 1-19	R2me2s	T3P	K4me2		Free
41	B17	ARme2spTKme3QTARKSTGGKAPRKQ	H3 1-19	R2me2s	T3P	K4me3		Free
42	B18	ARme2spTKacQTARKSTGGKAPRKQ	H3 1-19	R2me2s	T3P	K4ac		Free
43	B19	ARme2apTKme1QTARKSTGGKAPRKQ	H3 1-19	R2me2a	T3P	K4me1		Free
44	B20	ARme2apTKme2QTARKSTGGKAPRKQ	H3 1-19	R2me2a	T3P	K4me2		Free
45	B21	ARme2apTKme3QTARKSTGGKAPRKQ	H3 1-19	R2me2a	T3P	K4me3		Free
46	B22	ARme2apTKacQTARKSTGGKAPRKQ	H3 1-19	R2me2a	T3P	K4ac		Free
47	B23	ARTKQTARme2aKme1STGGKAPRKQ	H3 1-19	R8me2s	K9me1			Free
48	B24	ARTKQTARme2aKme2STGGKAPRKQ	H3 1-19	R8me2s	K9me2			Free
49	C1	ARTKQTARme2aKme3STGGKAPRKQ	H3 1-19	R8me2s	K9me3			Free
50	C2	ARTKQTARme2aKacSTGGKAPRKQ	H3 1-19	R8me2s	K9ac			Free
51	C3	ARTKQTARme2aKpSTGGKAPRKQ	H3 1-19	R8me2s	S10P			Free
52	C4	ARTKQTARme2aKSpTGGKAPRKQ	H3 1-19	R8me2s	T11P			Free
53	C5	ARTKQTARme2aKme1STGGKAPRKQ	H3 1-19	R8me2a	K9me1			Free
54	C6	ARTKQTARme2aKme2STGGKAPRKQ	H3 1-19	R8me2a	K9me2			Free
55	C7	ARTKQTARme2aKme3STGGKAPRKQ	H3 1-19	R8me2a	K9me3			Free
56	C8	ARTKQTARme2aKacSTGGKAPRKQ	H3 1-19	R8me2a	K9ac			Free
57	C9	ARTKQTARme2aKpSTGGKAPRKQ	H3 1-19	R8me2a	S10P			Free
58	C10	ARTKQTARme2aKSpTGGKAPRKQ	H3 1-19	R8me2a	T11P			Free
59	C11	ARTKQTACitKme1STGGKAPRKQ	H3 1-19	R8Citr	K9me1			Free
60	C12	ARTKQTACitKme2STGGKAPRKQ	H3 1-19	R8Citr	K9me2			Free
61	C13	ARTKQTACitKme3STGGKAPRKQ	H3 1-19	R8Citr	K9me3			Free
62	C14	ARTKQTACitKacSTGGKAPRKQ	H3 1-19	R8Citr	K9ac			Free
63	C15	ARTKQTACitKpSTGGKAPRKQ	H3 1-19	R8Citr	S10P			Free
64	C16	ARTKQTACitKSpTGGKAPRKQ	H3 1-19	R8Citr	T11P			Free
65	C17	ARTKQTARKme1pSTGGKAPRKQ	H3 1-19	K9me1	S10P			Free
66	C18	ARTKQTARKme1SpTGGKAPRKQ	H3 1-19	K9me1	T11P			Free
67	C19	ARTKQTARKme1STGGKacAPRKQ	H3 1-19	K9me1	K14ac			Free
68	C20	ARTKQTARKme2pSTGGKAPRKQ	H3 1-19	K9me2	S10P			Free
69	C21	ARTKQTARKme2SpTGGKAPRKQ	H3 1-19	K9me2	T11P			Free
70	C22	ARTKQTARKme2STGGKacAPRKQ	H3 1-19	K9me2	K14ac			Free
71	C23	ARTKQTARKme3pSTGGKAPRKQ	H3 1-19	K9me3	S10P			Free
72	C24	ARTKQTARKme3SpTGGKAPRKQ	H3 1-19	K9me3	T11P			Free
73	D1	ARTKQTARKme3STGGKacAPRKQ	H3 1-19	K9me3	K14ac			Free
74	D2	ARTKQTARKacpSTGGKAPRKQ	H3 1-19	K9ac	S10P			Free
75	D3	ARTKQTARKacSpTGGKAPRKQ	H3 1-19	K9ac	T11P			Free
76	D4	ARTKQTARKacSTGGKacAPRKQ	H3 1-19	K9ac	K14ac			Free
77	D5	ARTKQTARKpSpTGGKAPRKQ	H3 1-19	S10P	T11P			Free

78	D6	ARTKQTARKpSTGGKacAPRKQ	H3 1-19	S10P	K14ac			Free
79	D7	ARTKQTARKSpTGGKacAPRKQ	H3 1-19	T11P	K14ac			Free
80	D8	ARTKQTARme2sKme1pSTGGKAPRKQ	H3 1-19	R8me2s	K9me1	S10P		Free
81	D9	ARTKQTARme2sKme2pSTGGKAPRKQ	H3 1-19	R8me2s	K9me2	S10P		Free
82	D10	ARTKQTARme2sKme3pSTGGKAPRKQ	H3 1-19	R8me2s	K9me3	S10P		Free
83	D11	ARTKQTARme2sKacpSTGGKAPRKQ	H3 1-19	R8me2s	K9ac	S10P		Free
84	D12	ARTKQTARme2sKme1SpTGGKAPRKQ	H3 1-19	R8me2s	K9me1	T11P		Free
85	D13	ARTKQTARme2sKme2SpTGGKAPRKQ	H3 1-19	R8me2s	K9me2	T11P		Free
86	D14	ARTKQTARme2sKme3SpTGGKAPRKQ	H3 1-19	R8me2s	K9me3	T11P		Free
87	D15	ARTKQTARme2sKacSpTGGKAPRKQ	H3 1-19	R8me2s	K9ac	T11P		Free
88	D16	ARTKQTARme2aKme1pSTGGKAPRKQ	H3 1-19	R8me2a	K9me1	S10P		Free
89	D17	ARTKQTARme2aKme2pSTGGKAPRKQ	H3 1-19	R8me2a	K9me2	S10P		Free
90	D18	ARTKQTARme2aKme3pSTGGKAPRKQ	H3 1-19	R8me2a	K9me3	S10P		Free
91	D19	ARTKQTARme2aKacpSTGGKAPRKQ	H3 1-19	R8me2a	K9ac	S10P		Free
92	D20	ARTKQTARme2aKme1SpTGGKAPRKQ	H3 1-19	R8me2a	K9me1	T11P		Free
93	D21	ARTKQTARme2aKme2SpTGGKAPRKQ	H3 1-19	R8me2a	K9me2	T11P		Free
94	D22	ARTKQTARme2aKme3SpTGGKAPRKQ	H3 1-19	R8me2a	K9me3	T11P		Free
95	D23	ARTKQTARme2aKacSpTGGKAPRKQ	H3 1-19	R8me2a	K9ac	T11P		Free
96	D24	ARTKQTARme2aKme1pSpTGGKAPRKQ	H3 1-19	R8me2a	K9me1	S10P	T11P	Free
97	E1	ARTKQTARme2aKme2pSpTGGKAPRKQ	H3 1-19	R8me2a	K9me2	S10P	T11P	Free
98	E2	ARTKQTARme2aKme3pSpTGGKAPRKQ	H3 1-19	R8me2a	K9me3	S10P	T11P	Free
99	E3	ARTKQTARme2aKacpSpTGGKAPRKQ	H3 1-19	R8me2a	K9ac	S10P	T11P	Free
100	E4	ARme2sTKme1QTARme2sKSTGGKAPRKQ	H3 1-19	R2me2s	K4me1	R8me2s		Free
101	E5	ARme2sTKme2QTARme2sKSTGGKAPRKQ	H3 1-19	R2me2s	K4me2	R8me2s		Free
102	E6	ARme2sTKme3QTARme2sKSTGGKAPRKQ	H3 1-19	R2me2s	K4me3	R8me2s		Free
103	E7	ARme2sTKacQTARme2sKSTGGKAPRKQ	H3 1-19	R2me2s	K4ac	R8me2s		Free
104	E8	ARme2aTKme1QTARme2aKSTGGKAPRKQ	H3 1-19	R2me2a	K4me1	R8me2a		Free
105	E9	ARme2aTKme2QTARme2aKSTGGKAPRKQ	H3 1-19	R2me2a	K4me2	R8me2a		Free
106	E10	ARme2aTKme3QTARme2aKSTGGKAPRKQ	H3 1-19	R2me2a	K4me3	R8me2a		Free
107	E11	ARme2aTKacQTARme2aKSTGGKAPRKQ	H3 1-19	R2me2a	K4ac	R8me2a		Free
108	E12	ARme2sTKme1QTARKme1STGGKAPRKQ	H3 1-19	R2me2s	K4me1	K9me1		Free
109	E13	ARme2sTKme2QTARKme1STGGKAPRKQ	H3 1-19	R2me2s	K4me2	K9me1		Free
110	E14	ARme2sTKme3QTARKme1STGGKAPRKQ	H3 1-19	R2me2s	K4me3	K9me1		Free
111	E15	ARme2sTKacQTARKme1STGGKAPRKQ	H3 1-19	R2me2s	K4ac	K9me1		Free
112	E16	ARme2aTKme1QTARKme2STGGKAPRKQ	H3 1-19	R2me2a	K4me1	K9me2		Free
113	E17	ARme2aTKme2QTARKme2STGGKAPRKQ	H3 1-19	R2me2a	K4me2	K9me2		Free
114	E18	ARme2aTKme3QTARKme2STGGKAPRKQ	H3 1-19	R2me2a	K4me3	K9me2		Free
115	E19	ARme2aTKacQTARKme2STGGKAPRKQ	H3 1-19	R2me2a	K4ac	K9me2		Free
116	E20	ARme2sTKme1QTARKme3STGGKAPRKQ	H3 1-19	R2me2s	K4me1	K9me3		Free
117	E21	ARme2sTKme2QTARKme3STGGKAPRKQ	H3 1-19	R2me2s	K4me2	K9me3		Free
118	E22	ARme2sTKme3QTARKme3STGGKAPRKQ	H3 1-19	R2me2s	K4me3	K9me3		Free
119	E23	ARme2sTKacQTARKme3STGGKAPRKQ	H3 1-19	R2me2s	K4ac	K9me3		Free
120	E24	ARme2aTKme1QTARKacSTGGKAPRKQ	H3 1-19	R2me2a	K4me1	K9ac		Free
121	F1	ARme2aTKme2QTARKacSTGGKAPRKQ	H3 1-19	R2me2a	K4me2	K9ac		Free
122	F2	ARme2aTKme3QTARKacSTGGKAPRKQ	H3 1-19	R2me2a	K4me3	K9ac		Free

123	F3	ARme2aTKacQTARKacSTGGKAPRKQ	H3 1-19	R2me2a	K4ac	K9ac		Free
124	F4	ARTKme1QTARme2sKme1STGGKAPRKQ	H3 1-19	K4me1	R8me2s	K9me1		Free
125	F5	ARTKme2QTARme2sKme1STGGKAPRKQ	H3 1-19	K4me2	R8me2s	K9me1		Free
126	F6	ARTKme3QTARme2sKme1STGGKAPRKQ	H3 1-19	K4me3	R8me2s	K9me1		Free
127	F7	ARTKacQTARme2sKme1STGGKAPRKQ	H3 1-19	K4ac	R8me2s	K9me1		Free
128	F8	ARTKme1QTARme2aKme1STGGKAPRKQ	H3 1-19	K4me1	R8me2a	K9me1		Free
129	F9	ARTKme2QTARme2aKme1STGGKAPRKQ	H3 1-19	K4me2	R8me2a	K9me1		Free
130	F10	ARTKme3QTARme2aKme1STGGKAPRKQ	H3 1-19	K4me3	R8me2a	K9me1		Free
131	F11	ARTKacQTARme2aKme1STGGKAPRKQ	H3 1-19	K4ac	R8me2a	K9me1		Free
132	F12	ARTKme1QTARme2sKme2STGGKAPRKQ	H3 1-19	K4me1	R8me2s	K9me2		Free
133	F13	ARTKme2QTARme2sKme2STGGKAPRKQ	H3 1-19	K4me2	R8me2s	K9me2		Free
134	F14	ARTKme3QTARme2sKme2STGGKAPRKQ	H3 1-19	K4me3	R8me2s	K9me2		Free
135	F15	ARTKacQTARme2sKme2STGGKAPRKQ	H3 1-19	K4ac	R8me2s	K9me2		Free
136	F16	ARTKme1QTARme2aKme2STGGKAPRKQ	H3 1-19	K4me1	R8me2a	K9me2		Free
137	F17	ARTKme2QTARme2aKme2STGGKAPRKQ	H3 1-19	K4me2	R8me2a	K9me2		Free
138	F18	ARTKme3QTARme2aKme2STGGKAPRKQ	H3 1-19	K4me3	R8me2a	K9me2		Free
139	F19	ARTKacQTARme2aKme2STGGKAPRKQ	H3 1-19	K4ac	R8me2a	K9me2		Free
140	F20	ARTKme1QTARme2sKme3STGGKAPRKQ	H3 1-19	K4me1	R8me2s	K9me3		Free
141	F21	ARTKme2QTARme2sKme3STGGKAPRKQ	H3 1-19	K4me2	R8me2s	K9me3		Free
142	F22	ARTKme3QTARme2sKme3STGGKAPRKQ	H3 1-19	K4me3	R8me2s	K9me3		Free
143	F23	ARTKacQTARme2sKme3STGGKAPRKQ	H3 1-19	K4ac	R8me2s	K9me3		Free
144	F24	ARTKme1QTARme2aKme3STGGKAPRKQ	H3 1-19	K4me1	R8me2a	K9me3		Free
145	G1	ARTKme2QTARme2aKme3STGGKAPRKQ	H3 1-19	K4me2	R8me2a	K9me3		Free
146	G2	ARTKme3QTARme2aKme3STGGKAPRKQ	H3 1-19	K4me3	R8me2a	K9me3		Free
147	G3	ARTKacQTARme2aKme3STGGKAPRKQ	H3 1-19	K4ac	R8me2a	K9me3		Free
148	G4	ARTKme1QTARme2sKacSTGGKAPRKQ	H3 1-19	K4me1	R8me2s	K9ac		Free
149	G5	ARTKme2QTARme2sKacSTGGKAPRKQ	H3 1-19	K4me2	R8me2s	K9ac		Free
150	G6	ARTKme3QTARme2sKacSTGGKAPRKQ	H3 1-19	K4me3	R8me2s	K9ac		Free
151	G7	ARTKacQTARme2sKacSTGGKAPRKQ	H3 1-19	K4ac	R8me2s	K9ac		Free
152	G8	ARTKme1QTARme2aKacSTGGKAPRKQ	H3 1-19	K4me1	R8me2a	K9ac		Free
153	G9	ARTKme2QTARme2aKacSTGGKAPRKQ	H3 1-19	K4me2	R8me2a	K9ac		Free
154	G10	ARTKme3QTARme2aKacSTGGKAPRKQ	H3 1-19	K4me3	R8me2a	K9ac		Free
155	G11	ARTKacQTARme2aKacSTGGKAPRKQ	H3 1-19	K4ac	R8me2a	K9ac		Free
156	G12	ARme2sTKme1QTARme2sKme1STGGKAPRKQ	H3 1-19	R2me2s	K4me1	R8me2s	K9me1	Free
157	G13	ARme2sTKme2QTARme2sKme1STGGKAPRKQ	H3 1-19	R2me2s	K4me2	R8me2s	K9me1	Free
158	G14	ARme2sTKme3QTARme2sKme1STGGKAPRKQ	H3 1-19	R2me2s	K4me3	R8me2s	K9me1	Free
159	G15	ARme2sTKacQTARme2sKme1STGGKAPRKQ	H3 1-19	R2me2s	K4ac	R8me2s	K9me1	Free
160	G16	ARme2aTKme1QTARme2sKme1STGGKAPRKQ	H3 1-19	R2me2a	K4me1	R8me2s	K9me1	Free
161	G17	ARme2aTKme2QTARme2sKme1STGGKAPRKQ	H3 1-19	R2me2a	K4me2	R8me2s	K9me1	Free
162	G18	ARme2aTKme3QTARme2sKme1STGGKAPRKQ	H3 1-19	R2me2a	K4me3	R8me2s	K9me1	Free
163	G19	ARme2aTKacQTARme2sKme1STGGKAPRKQ	H3 1-19	R2me2a	K4ac	R8me2s	K9me1	Free
164	G20	ARme2sTKme1QTARme2sKme2STGGKAPRKQ	H3 1-19	R2me2s	K4me1	R8me2s	K9me2	Free
165	G21	ARme2sTKme2QTARme2sKme2STGGKAPRKQ	H3 1-19	R2me2s	K4me2	R8me2s	K9me2	Free
166	G22	ARme2sTKme3QTARme2sKme2STGGKAPRKQ	H3 1-19	R2me2s	K4me3	R8me2s	K9me2	Free
167	G23	ARme2sTKacQTARme2sKme2STGGKAPRKQ	H3 1-19	R2me2s	K4ac	R8me2s	K9me2	Free

168	G24	ARme2aTKme1QTARme2sKme2STGGKAPRKQ	H3 1-19	R2me2a	K4me1	R8me2s	K9me2	Free
169	H1	ARme2aTKme2QTARme2sKme2STGGKAPRKQ	H3 1-19	R2me2a	K4me2	R8me2s	K9me2	Free
170	H2	ARme2aTKme3QTARme2sKme2STGGKAPRKQ	H3 1-19	R2me2a	K4me3	R8me2s	K9me2	Free
171	H3	ARme2aTKacQTARme2sKme2STGGKAPRKQ	H3 1-19	R2me2a	K4ac	R8me2s	K9me2	Free
172	H4	ARme2sTKme1QTARme2sKme3STGGKAPRKQ	H3 1-19	R2me2s	K4me1	R8me2s	K9me3	Free
173	H5	ARme2sTKme2QTARme2sKme3STGGKAPRKQ	H3 1-19	R2me2s	K4me2	R8me2s	K9me3	Free
174	H6	ARme2sTKme3QTARme2sKme3STGGKAPRKQ	H3 1-19	R2me2s	K4me3	R8me2s	K9me3	Free
175	H7	ARme2sTKacQTARme2sKme3STGGKAPRKQ	H3 1-19	R2me2s	K4ac	R8me2s	K9me3	Free
176	H8	ARme2aTKme1QTARme2sKme3STGGKAPRKQ	H3 1-19	R2me2a	K4me1	R8me2s	K9me3	Free
177	H9	ARme2aTKme2QTARme2sKme3STGGKAPRKQ	H3 1-19	R2me2a	K4me2	R8me2s	K9me3	Free
178	H10	ARme2aTKme3QTARme2sKme3STGGKAPRKQ	H3 1-19	R2me2a	K4me3	R8me2s	K9me3	Free
179	H11	ARme2aTKacQTARme2sKme3STGGKAPRKQ	H3 1-19	R2me2a	K4ac	R8me2s	K9me3	Free
180	H12	ARme2sTKme1QTARme2sKacSTGGKAPRKQ	H3 1-19	R2me2s	K4me1	R8me2s	K9ac	Free
181	H13	ARme2sTKme2QTARme2sKacSTGGKAPRKQ	H3 1-19	R2me2s	K4me2	R8me2s	K9ac	Free
182	H14	ARme2sTKme3QTARme2sKacSTGGKAPRKQ	H3 1-19	R2me2s	K4me3	R8me2s	K9ac	Free
183	H15	ARme2sTKacQTARme2sKacSTGGKAPRKQ	H3 1-19	R2me2s	K4ac	R8me2s	K9ac	Free
184	H16	ARme2aTKme1QTARme2sKacSTGGKAPRKQ	H3 1-19	R2me2a	K4me1	R8me2s	K9ac	Free
185	H17	ARme2aTKme2QTARme2sKacSTGGKAPRKQ	H3 1-19	R2me2a	K4me2	R8me2s	K9ac	Free
186	H18	ARme2aTKme3QTARme2sKacSTGGKAPRKQ	H3 1-19	R2me2a	K4me3	R8me2s	K9ac	Free
187	H19	ARme2aTKacQTARme2sKacSTGGKAPRKQ	H3 1-19	R2me2a	K4ac	R8me2s	K9ac	Free
188	H20	ARme2sTKme1QTARme2aKme1STGGKAPRKQ	H3 1-19	R2me2s	K4me1	R8me2a	K9me1	Free
189	H21	ARme2sTKme2QTARme2aKme1STGGKAPRKQ	H3 1-19	R2me2s	K4me2	R8me2a	K9me1	Free
190	H22	ARme2sTKme3QTARme2aKme1STGGKAPRKQ	H3 1-19	R2me2s	K4me3	R8me2a	K9me1	Free
191	H23	ARme2sTKacQTARme2aKme1STGGKAPRKQ	H3 1-19	R2me2s	K4ac	R8me2a	K9me1	Free
192	H24	ARme2aTKme1QTARme2aKme1STGGKAPRKQ	H3 1-19	R2me2a	K4me1	R8me2a	K9me1	Free
193	I1	ARme2aTKme2QTARme2aKme1STGGKAPRKQ	H3 1-19	R2me2a	K4me2	R8me2a	K9me1	Free
194	I2	ARme2aTKme3QTARme2aKme1STGGKAPRKQ	H3 1-19	R2me2a	K4me3	R8me2a	K9me1	Free
195	I3	ARme2aTKacQTARme2aKme1STGGKAPRKQ	H3 1-19	R2me2a	K4ac	R8me2a	K9me1	Free
196	I4	ARme2sTKme1QTARme2aKme2STGGKAPRKQ	H3 1-19	R2me2s	K4me1	R8me2a	K9me2	Free
197	I5	ARme2sTKme2QTARme2aKme2STGGKAPRKQ	H3 1-19	R2me2s	K4me2	R8me2a	K9me2	Free
198	I6	ARme2sTKme3QTARme2aKme2STGGKAPRKQ	H3 1-19	R2me2s	K4me3	R8me2a	K9me2	Free
199	I7	ARme2sTKacQTARme2aKme2STGGKAPRKQ	H3 1-19	R2me2s	K4ac	R8me2a	K9me2	Free
200	I8	ARme2aTKme1QTARme2aKme2STGGKAPRKQ	H3 1-19	R2me2a	K4me1	R8me2a	K9me2	Free
201	I9	ARme2aTKme2QTARme2aKme2STGGKAPRKQ	H3 1-19	R2me2a	K4me2	R8me2a	K9me2	Free
202	I10	ARme2aTKme3QTARme2aKme2STGGKAPRKQ	H3 1-19	R2me2a	K4me3	R8me2a	K9me2	Free
203	I11	ARme2aTKacQTARme2aKme2STGGKAPRKQ	H3 1-19	R2me2a	K4ac	R8me2a	K9me2	Free
204	I12	ARme2sTKme1QTARme2aKme3STGGKAPRKQ	H3 1-19	R2me2s	K4me1	R8me2a	K9me3	Free
205	I13	ARme2sTKme2QTARme2aKme3STGGKAPRKQ	H3 1-19	R2me2s	K4me2	R8me2a	K9me3	Free
206	I14	ARme2sTKme3QTARme2aKme3STGGKAPRKQ	H3 1-19	R2me2s	K4me3	R8me2a	K9me3	Free
207	I15	ARme2sTKacQTARme2aKme3STGGKAPRKQ	H3 1-19	R2me2s	K4ac	R8me2a	K9me3	Free
208	I16	ARme2aTKme1QTARme2aKme3STGGKAPRKQ	H3 1-19	R2me2a	K4me1	R8me2a	K9me3	Free
209	I17	ARme2aTKme2QTARme2aKme3STGGKAPRKQ	H3 1-19	R2me2a	K4me2	R8me2a	K9me3	Free
210	I18	ARme2aTKme3QTARme2aKme3STGGKAPRKQ	H3 1-19	R2me2a	K4me3	R8me2a	K9me3	Free
211	I19	ARme2aTKacQTARme2aKme3STGGKAPRKQ	H3 1-19	R2me2a	K4ac	R8me2a	K9me3	Free
212	I20	ARme2sTKme1QTARme2aKacSTGGKAPRKQ	H3 1-19	R2me2s	K4me1	R8me2a	K9ac	Free

213	I21	ARme2sTKme2QTARme2aKacSTGGKAPRKQ	H3 1-19	R2me2s	K4me2	R8me2a	K9ac	Free
214	I22	ARme2sTKme3QTARme2aKacSTGGKAPRKQ	H3 1-19	R2me2s	K4me3	R8me2a	K9ac	Free
215	I23	ARme2sTKacQTARme2aKacSTGGKAPRKQ	H3 1-19	R2me2s	K4ac	R8me2a	K9ac	Free
216	I24	ARme2aTKme1QTARme2aKacSTGGKAPRKQ	H3 1-19	R2me2a	K4me1	R8me2a	K9ac	Free
217	J1	ARme2aTKme2QTARme2aKacSTGGKAPRKQ	H3 1-19	R2me2a	K4me2	R8me2a	K9ac	Free
218	J2	ARme2aTKme3QTARme2aKacSTGGKAPRKQ	H3 1-19	R2me2a	K4me3	R8me2a	K9ac	Free
219	J3	ARme2aTKacQTARme2aKacSTGGKAPRKQ	H3 1-19	R2me2a	K4ac	R8me2a	K9ac	Free
220	J4	ARKSTGGKAPRKQLATKAAR	H3 7-26	unmod				Acetylated
221	J5	ARKSTGGKacAPRKQLATKAAR	H3 7-26	K14ac				Acetylated
222	J6	ARKpSTGGKacAPRKQLATKAAR	H3 7-26	K14ac	S10P			Acetylated
223	J7	ARKSpTGGKacAPRKQLATKAAR	H3 7-26	K14ac	T11P			Acetylated
224	J8	ARKSTGGKAPRme2sKQLATKAAR	H3 7-26	R17me2s				Acetylated
225	J9	ARKSTGGKAPRme2aKQLATKAAR	H3 7-26	R17me2a				Acetylated
226	J10	ARKSTGGKAPCitKQLATKAAR	H3 7-26	R17Citr				Acetylated
227	J11	ARKSTGGKAPRKacQLATKAAR	H3 7-26	K18ac				Acetylated
228	J12	ARKSTGGKacAPRme2sKQLATKAAR	H3 7-26	K14ac	R17me2s			Acetylated
229	J13	ARKSTGGKacAPRme2aKQLATKAAR	H3 7-26	K14ac	R17me2a			Acetylated
230	J14	ARKSTGGKacAPRKacQLATKAAR	H3 7-26	K14ac	K18ac			Acetylated
231	J15	ARKSTGGKAPRme2sKacQLATKAAR	H3 7-26	R17me2s	K18ac			Acetylated
232	J16	ARKSTGGKAPRme2aKacQLATKAAR	H3 7-26	R17me2a	K18ac			Acetylated
233	J17	ARKSTGGKAPCitKacQLATKAAR	H3 7-26	R17Citr	K18ac			Acetylated
234	J18	ARKSTGGKacAPRme2sKacQLATKAAR	H3 7-26	K14ac	R17me2s	K18ac		Acetylated
235	J19	ARKSTGGKacAPRme2aKacQLATKAAR	H3 7-26	K14ac	R17me2a	K18ac		Acetylated
236	J20	PRKQLATKAARKSAPATGG	H3 16-35	unmod				Acetylated
237	J21	PRKQLATKAARme2sKSAPATGG	H3 16-35	R26me2s				Acetylated
238	J22	PRKQLATKAARme2aKSAPATGG	H3 16-35	R26me2a				Acetylated
239	J23	PRKQLATKAACitKSAPATGG	H3 16-35	R26Citr				Acetylated
240	J24	PRKQLATKAARKme1SAPATGG	H3 16-35	K27me1				Acetylated
241	K1	PRKQLATKAARKme2SAPATGG	H3 16-35	K27me2				Acetylated
242	K2	PRKQLATKAARKme3SAPATGG	H3 16-35	K27me3				Acetylated
243	K3	PRKQLATKAARKacSAPATGG	H3 16-35	K27ac				Acetylated
244	K4	PRKQLATKAARKpSAPATGG	H3 16-35	S28P				Acetylated
245	K5	PRKQLATKAARme2sKme1SAPATGG	H3 16-35	R26me2s	K27me1			Acetylated
246	K6	PRKQLATKAARme2sKme2SAPATGG	H3 16-35	R26me2s	K27me2			Acetylated
247	K7	PRKQLATKAARme2sKme3SAPATGG	H3 16-35	R26me2s	K27me3			Acetylated
248	K8	PRKQLATKAARme2sKacSAPATGG	H3 16-35	R26me2s	K27ac			Acetylated
249	K9	PRKQLATKAARme2sKpSAPATGG	H3 16-35	R26me2s	S28P			Acetylated
250	K10	PRKQLATKAARme2aKme1SAPATGG	H3 16-35	R26me2a	K27me1			Acetylated
251	K11	PRKQLATKAARme2aKme2SAPATGG	H3 16-35	R26me2a	K27me2			Acetylated
252	K12	PRKQLATKAARme2aKme3SAPATGG	H3 16-35	R26me2a	K27me3			Acetylated
253	K13	PRKQLATKAARme2aKacSAPATGG	H3 16-35	R26me2a	K27ac			Acetylated
254	K14	PRKQLATKAARme2aKpSAPATGG	H3 16-35	R26me2a	S28P			Acetylated
255	K15	PRKQLATKAACitKme1SAPATGG	H3 16-35	R26Citr	K27me1			Acetylated
256	K16	PRKQLATKAACitKme2SAPATGG	H3 16-35	R26Citr	K27me2			Acetylated
257	K17	PRKQLATKAACitKme3SAPATGG	H3 16-35	R26Citr	K27me3			Acetylated

258	K18	PRKQLATKAACitKpSAPATGG	H3 16-35	R26Citr	S28P			Acetylated
259	K19	PRKQLATKAARKme1pSAPATGG	H3 16-35	K27me1	S28P			Acetylated
260	K20	PRKQLATKAARKme2pSAPATGG	H3 16-35	K27me2	S28P			Acetylated
261	K21	PRKQLATKAARKme3pSAPATGG	H3 16-35	K27me3	S28P			Acetylated
262	K22	PRKQLATKAARKacpSAPATGG	H3 16-35	K27ac	S28P			Acetylated
263	K23	PRKQLATKAARme2sKme1pSAPATGG	H3 16-35	R26me2s	K27me1	S28P		Acetylated
264	K24	PRKQLATKAARme2sKme2pSAPATGG	H3 16-35	R26me2s	K27me2	S28P		Acetylated
265	L1	PRKQLATKAARme2sKme3pSAPATGG	H3 16-35	R26me2s	K27me3	S28P		Acetylated
266	L2	PRKQLATKAARme2sKacpSAPATGG	H3 16-35	R26me2s	K27ac	S28P		Acetylated
267	L3	PRKQLATKAARme2aKme1pSAPATGG	H3 16-35	R26me2a	K27me1	S28P		Acetylated
268	L4	PRKQLATKAARme2aKme2pSAPATGG	H3 16-35	R26me2a	K27me2	S28P		Acetylated
269	L5	PRKQLATKAARme2aKme3pSAPATGG	H3 16-35	R26me2a	K27me3	S28P		Acetylated
270	L6	PRKQLATKAARme2aKacpSAPATGG	H3 16-35	R26me2a	K27ac	S28P		Acetylated
271	L7	RKSAPATGGVKKPHRYRPG	H3 26-45	unmod				Acetylated
272	L8	RKSAPATGGVKme1KPHRYRPG	H3 26-45	K36me1				Acetylated
273	L9	RKSAPATGGVKme2KPHRYRPG	H3 26-45	K36me2				Acetylated
274	L10	RKSAPATGGVKme3KPHRYRPG	H3 26-45	K36me3				Acetylated
275	L11	RKSAPATGGVKacKPHRYRPG	H3 26-45	K36ac				Acetylated
276	L12	SGRGKGGKGLGKGGAKRHR	H4 1-19	unmod				Free
277	L13	pSGRGKGGKGLGKGGAKRHR	H4 1-19	S1P				Free
278	L14	SGRme2sGKGGKGLGKGGAKRHR	H4 1-19	R3me2s				Free
279	L15	SGRme2aGKGGKGLGKGGAKRHR	H4 1-19	R3me2a				Free
280	L16	SGRGKacGGKGLGKGGAKRHR	H4 1-19	K5ac				Free
281	L17	SGRGKGGKacGLGKGGAKRHR	H4 1-19	K8ac				Free
282	L18	SGRGKGGKGLGKacGGAKRHR	H4 1-19	K12ac				Free
283	L19	SGRGKGGKGLGKGGAKacRHR	H4 1-19	K16ac				Free
284	L20	pSGRme2sGKGGKGLGKGGAKRHR	H4 1-19	S1P	R3me2s			Free
285	L21	pSGRme2aGKGGKGLGKGGAKRHR	H4 1-19	S1P	R3me2a			Free
286	L22	pSGRGKacGGKGLGKGGAKRHR	H4 1-19	S1P	K5ac			Free
287	L23	SGRme2sGKacGGKGLGKGGAKRHR	H4 1-19	R3me2s	K5ac			Free
288	L24	SGRme2sGKGGKacGLGKGGAKRHR	H4 1-19	R3me2s	K8ac			Free
289	M1	SGRme2aGKacGGKGLGKGGAKRHR	H4 1-19	R3me2a	K5ac			Free
290	M2	SGRme2aGKGGKacGLGKGGAKRHR	H4 1-19	R3me2a	K8ac			Free
291	M3	SGRGKacGGKacGLGKGGAKRHR	H4 1-19	K5ac	K8ac			Free
292	M4	SGRGKGGKacGLGKacGGAKRHR	H4 1-19	K8ac	K12ac			Free
293	M5	SGRGKGGKacGLGKGGAKacRHR	H4 1-19	K8ac	K16ac			Free
294	M6	SGRGKGGKGLGKacGGAKacRHR	H4 1-19	K12ac	K16ac			Free
295	M7	pSGRme2sGKacGGKGLGKGGAKRHR	H4 1-19	S1P	R3me2s	K5ac		Free
296	M8	pSGRme2aGKacGGKGLGKGGAKRHR	H4 1-19	S1P	R3me2a	K5ac		Free
297	M9	SGRme2sGKacGGKacGLGKGGAKRHR	H4 1-19	R3me2s	K5ac	K8ac		Free
298	M10	SGRme2aGKacGGKacGLGKGGAKRHR	H4 1-19	R3me2a	K5ac	K8ac		Free
299	M11	SGRGKacGGKacGLGKacGGAKRHR	H4 1-19	K5ac	K8ac	K12ac		Free
300	M12	SGRGKGGKacGLGKacGGAKacRHR	H4 1-19	K8ac	K12ac	K16ac		Free
301	M13	pSGRme2sGKacGGKacGLGKGGAKRHR	H4 1-19	S1P	R3me2s	K5ac	K8ac	Free
302	M14	pSGRme2aGKacGGKacGLGKGGAKRHR	H4 1-19	S1P	R3me2a	K5ac	K8ac	Free

303	M15	SGRme2sGKacGGKacGLGKacGGAKRHR	H41-19	R3me2s	K5ac	K8ac	K12ac	Free
304	M16	SGRme2aGKacGGKacGLGKacGGAKRHR	H41-19	R3me2a	K5ac	K8ac	K12ac	Free
305	M17	SGRGKacGGKacGLGKacGGAKacRHR	H41-19	K5ac	K8ac	K12ac	K16ac	Free
306	M18	GKGGAKRHRKVLRDNIQGIT	H411-30	unmod				Acetylated
307	M19	GKacGGAKRHRKVLRDNIQGIT	H411-30	K12ac				Acetylated
308	M20	GKGGAKacRHRKVLRDNIQGIT	H411-30	K16ac				Acetylated
309	M21	GKGGAKRme2sHRKVLRDNIQGIT	H411-30	R17me2s				Acetylated
310	M22	GKGGAKRme2aHRKVLRDNIQGIT	H411-30	R17me2a				Acetylated
311	M23	GKGGAKRHRme2sKVLRDNIQGIT	H411-30	R19me2s				Acetylated
312	M24	GKGGAKRHRme2aKVLRDNIQGIT	H411-30	R19me2a				Acetylated
313	N1	GKGGAKRHRKme1VLRDNIQGIT	H411-30	K20me1				Acetylated
314	N2	GKGGAKRHRKme2VLRDNIQGIT	H411-30	K20me2				Acetylated
315	N3	GKGGAKRHRKme3VLRDNIQGIT	H411-30	K20me3				Acetylated
316	N4	GKGGAKRHRKacVLRDNIQGIT	H411-30	K20ac				Acetylated
317	N5	GKGGAKRHRKVLRme2aDNIQGIT	H411-30	R24me2a				Acetylated
318	N6	GKGGAKRHRKVLRme2sDNIQGIT	H411-30	R24me2s				Acetylated
319	N7	GKacGGAKacRHRKVLRDNIQGIT	H411-30	K12ac	K16ac			Acetylated
320	N8	GKGGAKacRme2sHRKVLRDNIQGIT	H411-30	K16ac	R17me2s			Acetylated
321	N9	GKGGAKacRme2aHRKVLRDNIQGIT	H411-30	K16ac	R17me2a			Acetylated
322	N10	GKGGAKacRHRme2sKVLRDNIQGIT	H411-30	K16ac	R19me2s			Acetylated
323	N11	GKGGAKacRHRme2aKVLRDNIQGIT	H411-30	K16ac	R19me2a			Acetylated
324	N12	GKGGAKacRHRKme1VLRDNIQGIT	H411-30	K16ac	K20me1			Acetylated
325	N13	GKGGAKacRHRKme2VLRDNIQGIT	H411-30	K16ac	K20me2			Acetylated
326	N14	GKGGAKacRHRKme3VLRDNIQGIT	H411-30	K16ac	K20me3			Acetylated
327	N15	GKGGAKacRHRKacVLRDNIQGIT	H411-30	K16ac	K20ac			Acetylated
328	N16	GKacGGAKacRHRKme1VLRDNIQGIT	H411-30	K12ac	K16ac	K20me1		Acetylated
329	N17	GKacGGAKacRHRKme2VLRDNIQGIT	H411-30	K12ac	K16ac	K20me2		Acetylated
330	N18	GKacGGAKacRHRKme3VLRDNIQGIT	H411-30	K12ac	K16ac	K20me3		Acetylated
331	N19	GKacGGAKacRHRKacVLRDNIQGIT	H411-30	K12ac	K16ac	K20ac		Acetylated
332	N20	GKGGAKRHRme2aKme1VLRDNIQGIT	H411-30	R19me2a	K20me1			Acetylated
333	N21	GKGGAKRHRme2aKme2VLRDNIQGIT	H411-30	R19me2a	K20me2			Acetylated
334	N22	GKGGAKRHRme2aKme3VLRDNIQGIT	H411-30	R19me2a	K20me3			Acetylated
335	N23	GKGGAKRHRme2aKacVLRDNIQGIT	H411-30	R19me2a	K20ac			Acetylated
336	N24	GKGGAKRHRme2sKme1VLRDNIQGIT	H411-30	R19me2s	K20me1			Acetylated
337	O1	GKGGAKRHRme2sKme2VLRDNIQGIT	H411-30	R19me2s	K20me2			Acetylated
338	O2	GKGGAKRHRme2sKme3VLRDNIQGIT	H411-30	R19me2s	K20me3			Acetylated
339	O3	GKGGAKRHRme2sKacVLRDNIQGIT	H411-30	R19me2s	K20ac			Acetylated
340	O4	GKGGAKRHRKme1VLRme2aDNIQGIT	H411-30	R24me2a	K20me1			Acetylated
341	O5	GKGGAKRHRKme2VLRme2aDNIQGIT	H411-30	R24me2a	K20me2			Acetylated
342	O6	GKGGAKRHRKme3VLRme2aDNIQGIT	H411-30	R24me2a	K20me3			Acetylated
343	O7	GKGGAKRHRKacVLRme2aDNIQGIT	H411-30	R24me2a	K20ac			Acetylated
344	O8	GKGGAKRHRKme1VLRme2sDNIQGIT	H411-30	R24me2s	K20me1			Acetylated
345	O9	GKGGAKRHRKme2VLRme2sDNIQGIT	H411-30	R24me2s	K20me2			Acetylated
346	O10	GKGGAKRHRKme3VLRme2sDNIQGIT	H411-30	R24me2s	K20me3			Acetylated
347	O11	GKGGAKRHRKacVLRme2sDNIQGIT	H411-30	R24me2s	K20ac			Acetylated

348	O12	SGRGKQGGKARAKAKSRSS	H2a 1-19	unmod				Free
349	O13	pSGRGKQGGKARAKAKSRSS	H2a 1-19	S1P				Free
350	O14	SGRGKacQGGKARAKAKSRSS	H2a 1-19	K5ac				Free
351	O15	SGRGKQGGKacARAKAKSRSS	H2a 1-19	K9ac				Free
352	O16	SGRGKQGGKARAKacAKSRSS	H2a 1-19	K13ac				Free
353	O17	pSGRGKacQGGKARAKAKSRSS	H2a 1-19	S1P	K5ac			Free
354	O18	pSGRGKQGGKacARAKAKSRSS	H2a 1-19	S1P	K9ac			Free
355	O19	pSGRGKQGGKARAKacAKSRSS	H2a 1-19	S1P	K13ac			Free
356	O20	SGRGKacQGGKacARAKAKSRSS	H2a 1-19	K5ac	K9ac			Free
357	O21	SGRGKacQGGKARAKacAKSRSS	H2a 1-19	K5ac	K13ac			Free
358	O22	SGRGKQGGKacARAKacAKSRSS	H2a 1-19	K9ac	K13ac			Free
359	O23	pSGRGKacQGGKacARAKAKSRSS	H2a 1-19	S1P	K5ac	K9ac		Free
360	O24	pSGRGKacQGGKARAKacAKSRSS	H2a 1-19	S1P	K5ac	K13ac		Free
361	P1	pSGRGKQGGKacARAKacAKSRSS	H2a 1-19	S1P	K9ac	K13ac		Free
362	P2	SGRGKacQGGKacARAKacAKSRSS	H2a 1-19	K5ac	K9ac	K13ac		Free
363	P3	pSGRGKacQGGKacARAKacAKSRSS	H2a 1-19	S1P	K5ac	K9ac	K13ac	Free
364	P4	PDPAKSAPAPKKGSKKAVT	H2B 1-19	unmod				Free
365	P5	PDPAKacSAPAPKKGSKKAVT	H2B 1-19	K5ac				Free
366	P6	PDPAKSAPAPKKacGSKKAVT	H2B 1-19	K12ac				Free
367	P7	PDPAKSAPAPKKGpSKKAVT	H2B 1-19	S14P				Free
368	P8	PDPAKSAPAPKKGSKacKAVT	H2B 1-19	K15ac				Free
369	P9	PDPAKacSAPAPKKacGSKKAVT	H2B 1-19	K5ac	K12ac			Free
370	P10	PDPAKacSAPAPKKGpSKKAVT	H2B 1-19	K5ac	S14P			Free
371	P11	PDPAKacSAPAPKKGSKacKAVT	H2B 1-19	K5ac	K15ac			Free
372	P12	PDPAKSAPAPKKacGpSKKAVT	H2B 1-19	K12ac	S14P			Free
373	P13	PDPAKSAPAPKKacGSKacKAVT	H2B 1-19	K12ac	K15ac			Free
374	P14	PDPAKSAPAPKKGpSKacKAVT	H2B 1-19	S14P	K15ac			Free
375	P15	PDPAKacSAPAPKKacGpSKKAVT	H2B 1-19	K5ac	K12ac	S14P		Free
376	P16	PDPAKacSAPAPKKacGSKacKAVT	H2B 1-19	K5ac	K12ac	K15ac		Free
377	P17	PDPAKacSAPAPKKGpSKacKAVT	H2B 1-19	K5ac	S14P	K15ac		Free
378	P18	PDPAKSAPAPKKacGpSKacKAVT	H2B 1-19	K12ac	S14P	K15ac		Free
379	P19	PDPAKacSAPAPKKacGpSKacKAVT	H2B 1-19	K5ac	K12ac	S14P	K15ac	Free
380	P20	BioAANWSHPQFEKAA	Biotin, control peptide					Biotinylated
381	P21	EQKLISEEDLA	c-myc tag					Free
382	P22	HAc	neg. contol					Acetylated
383	P23	KKme1Kme2Kme3KacRRme2sRRme2aRCitKKme1KacKme3RK	background 01					Acetylated
384	P24	RRme2sKKme1KacRRme2aKme2KKme3RKme1Rme2sKKacRK	background 02					Acetylated

Appendix 5. Active Motif Histone H3/H4/H2A/H2B N-terminal peptide library

CHAPTER 7. REFERENCES

Ables E.T., Drummond-Barbosa D. (2010) The steroid hormone ecdysone functions with intrinsic chromatin remodeling factors to control female germline stem cells in *Drosophila*. *Cell Stem Cell.* 7(5):581-92.

Ahmad K., Henikoff S. (2002) The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell.* 9(6):1191-200.

Alkhatib S.G., Landry J.W. (2011) The nucleosome remodeling factor. *FEBS Lett.* 585(20):3197-207.

Allfrey V.G., Faulkner R., Mirsky A.E. (1964) Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc Natl Acad Sci U S A.* 51:786-94.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. (1990) Basic local alignment search tool. *J Mol Biol.* 215(3):403-10.

Alvarez F, Muñoz F, Schilcher P, Imhof A, Almouzni G, Loyola A. (2011) Sequential establishment of marks on soluble histones H3 and H4. *J Biol Chem.* 286(20):17714-21.

Andersen E.C., Lu X., Horvitz H.R. (2006) *C. elegans* ISWI and NURF301 antagonize an Rb-like pathway in the determination of multiple cell fates. *Development.*133(14):2695-704.

Anest V., Hanson J.L., Cogswell P.C., Steinbrecher K.A., Strahl B.D., Baldwin A.S. (2003) A nucleosomal function for IkappaB kinase-alpha in NF-kappaB-dependent gene expression. *Nature.* 423(6940):659-63.

Annunziato A. T. (2008) DNA packaging: nucleosome and chromatin. *Nature Education.* 1(1).

Ardehali M.B., Mei A., Zobeck K.L., Caron M., Lis J.T., Kusch T. (2011) *Drosophila* Set1 is the major histone H3 lysine 4 trimethyltransferase with role in transcription. *EMBO J.* 21;30(14):2817-28.

Awad S., Hassan A.H. (2008) The Swi2/Snf2 bromodomain is important for the full binding and remodeling activity of the SWI/SNF complex on H3- and H4-acetylated nucleosomes. *Ann N Y Acad Sci.* 1138:366-75.

Badenhorst P., Voas M., Rebay I., Wu C. (2002) Biological functions of the ISWI chromatin remodelling complex NURF. *Genes. Dev.* 16: 3186-3198.

Badenhorst P., Xiao H., Cherbas L., Kwon S. Y., Voas M., Rebay I., Cherbas P., Wu C. (2005) The *Drosophila* nucleosome remodelling factor NURF is required for Ecdysteroid signalling and metamorphosis. *Genes. Dev.* 19: 2540-2545.

Badenhorst P. (2013) What can we learn from flies: Epigenetic mechanisms regulating blood cell development in *Drosophila*. In Epigenetics and Human Health (series) Vol. on Transcriptional and Epigenetic Mechanisms Regulating Normal and Aberrant Blood Cell Development. Springer-Verlag. In press.

Bao Y., Shen X. (2007) INO80 subfamily of chromatin remodeling complexes. *Mutat Res.* 618(1-2):18-29.

Barak O., Lazzaro M. A., Lane W. S., Speicher D. W., Picketts D. J., Shiekhattar R. (2003) Isolation of NURF: a regulator Engrailed gene expression. *EMBO J.* 22: 6089-6100.

Barak O., Lazzaro M.A., Cooch N.S., Picketts D.J., Shiekhattar R. (2004) A tissue-specific, naturally occurring human SNF2L variant inactivates chromatin remodeling. *J Biol Chem.* 279(43):45130-8.

Barlev N. A., Poltoratsky V., Owen-Hughes T., Ying C., Liu L., Workman J. L., Berger S. L. (1998) Repression of GCN5 histone acetyltransferase activity via Bromodomain-mediated binding and phosphorylation by the Ku-DNA-dependent protein kinase complex. *Mol. Cell. Biol.* 18(3): 1349-1358.

Barratt M.J., Hazzalin C.A., Cano E., Mahadevan L.C. (1994) Mitogen-stimulated

phosphorylation of histone H3 is targeted to a small hyperacetylation-sensitive fraction. *Proc Natl Acad Sci U S A.* 91(11):4781-5.

Bastock R, St Johnston D. (2008) *Drosophila* oogenesis. *Curr Biol.* 9;18(23):R1082-7.

Becker P. B., Hörz W (2002) ATP-dependent nucleosome remodelling. *Annu. Rev. Biochem.* 71: 247-273.

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., Schreiber S.L., Lander E.S. (2005) Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell.* 28;120(2):169-81.

Bernstein B.E., Meissner A., Lander E.S. (2007) The mammalian epigenome. *Cell.* 128(4):669-81.

Bischof J., Maeda R. K., Hediger M., Karch F., Basler K. (2007) An optimized transgenesis system for *Drosophila* using germ-line-specific ΦC31 integrases. *Proc. Natl. Acad. Sci.* 104(9): 3312-3317.

Blair, S. (2000). Imaginal discs. In: *Drosophila* Protocols, ed. W. Sullivan, M. Ashburner, and R. S. Hawley, Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 159–173.

Bouazoune K., Mitterweger A., Langst G., Imhof A., Akhtar A., Becker P.B., Brehm A. (2002) The dMi–2 chromodomains are DNA binding modules important for ATP-dependent nucleosome mobilization. *EMBO J.* 21:2430–2440.

Boyer L. A., Latek R. R., Peterson C. L. (2004) The SANT domain: a unique histone-tail-binding module? *Nat Rev. Mol. Cell. Biol.* 5: 158-163.

Bruno M., Flaus A., Stockdale C., Rencurel C., Ferreira H., Owen-Hughes T. (2003) HistoneH2A/H2B dimmer exchange by ATP-dependent chromatin remodelling activities. *Mol.*

Cell. 12(6): 1599-1606.

Cairns B. R., Kim Y. J., Sayre M. H., Laurent B. C., Kornberg R. D. (1994) A multisubunit complex containing the SWI1/ADR6, SWI2/SNF2, SWI3, SNF5 and SNF6 gene products isolated from yeast. *Proc. Natl. Acad. Sci.* 91: 1950-1954.

Cai W., Bao X., Deng H., Jin Y., Girton J., Hohansen J., Johansen K. M. (2008) Pol II mediated transcription at active loci does not require histone H3S10 phosphorylation in *Drosophila*. *Development.* 135(17): 2917-2915.

Carré C., Szymczak D., Pidoux J., Antoniewski C. (2005) The histone H3 acetylase dGcn5 is a key player in *Drosophila melanogaster* metamorphosis. *Mol Cell Biol.* 25(18):8228-38.

Carré C., Ciurciu A., Komonyi O., Jacquier C., Fagegaltier D., Pidoux J., Tricoire H., Tora L., Boros I.M., Antoniewski C. (2008) The *Drosophila* NURF remodelling and the ATAC histone acetylase complexes functionally interact and are required for global chromosome organization. *EMBO Rep.* 9(2):187-92.

Chadwick B.P., Willard H.F. (2001) A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome. *J Cell Biol.* 152(2):375-84.

Chakravarty S., Zeng L., Zhou M.M. (2009) Structure and site-specific recognition of histone H3 by the PHD finger of human autoimmune regulator. *Structure.* 17(5):670-9.

Chatterjee N., Sinha D., Lemma-Dechassa M., Tan S., Shogren-Knaak M.A., Bartholomew B. (2011) Histone H3 tail acetylation modulates ATP-dependent remodeling through multiple mechanisms. *Nucleic Acids Res.* 39(19):8378-91.

Chen J., Santillan D.A., Koonce M., Wei W., Luo R., Thirman M.J., Zeleznik-Le N.J., Diaz M.O. (2008) Loss of MLL PHD finger 3 is necessary for MLL-ENL-induced hematopoietic stem cell immortalization. *Cancer Res.* 68(15):6199-207.

Cherry C.M., Matunis E.L. (2010) Epigenetic regulation of stem cell maintenance in the *Drosophila* testis via the nucleosome-remodeling factor NURF. *Cell Stem Cell.* 6(6):557-67.

Churikov D., Siino J., Svetlova M., Zhang K., Gineitis A., Morton Bradbury E., Zalensky A. (2004) Novel human testis-specific histone H2B encoded by the interrupted gene on the X chromosome. *Genomics.* 84(4):745-56.

Ciurciu A., Komonyi O., Pankotai T., Boros I.M. (2006) The *Drosophila* histone acetyltransferase Gcn5 and transcriptional adaptor Ada2a are involved in nucleosomal histone H4 acetylation. *Mol Cell Biol.* 26(24):9413-23.

Clapier C.R., Nightingale K.P., Becker P.B. (2002) A critical epitope for substrate recognition by the nucleosome remodeling ATPase ISWI. *Nucleic Acids Res.* 30(3):649-55.

Clapier C. R., Cairns B. R. (2009) The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* 78: 273-304.

Clayton A.L., Rose S., Barratt M.J., Mahadevan L.C. (2000) Phosphoacetylation of histone H3 on c-fos- and c-jun-associated nucleosomes upon gene activation. *EMBO J.* 19(14):3714-26.

Corona D. F., Lāngst G., Clapier C. R., Bonte E. J., Ferrari S., Tamkun J. W., Becker P. B. (1999) ISWI is an ATP-dependent nucleosome remodelling factor. *Mol. Cell.* 3: 239-245.

Corona D.F., Clapier C.R., Becker P.B., Tamkun J.W. (2002) Modulation of ISWI function by site-specific histone acetylation. *EMBO Rep.* 3(3):242-7.

Corona D.F., Armstrong J.A., Tamkun J.W. (2004a) Genetic and cytological analysis of *Drosophila* chromatin-remodeling factors. *Methods Enzymol.* 377:70-85.

Corona D.F., Tamkun J.W. (2004b) Multiple roles for ISWI in transcription, chromosome organization and DNA replication. *Biochim Biophys Acta.* 1677(1-3):113-9.

Couture J.F., Collazo E., Trievel R.C. (2006) Molecular recognition of histone H3 by the WD40 protein WDR5. *Nat Struct Mol Biol.* 13(8):698-703.

Côté J., Quinn J., Workman J. L., Peterson C. L. (1994) Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. *Science.* 265: 53-60.

Dai J., Sultan S., Taylor S. S., Higgins J. M. G. (2005) The kinase haspin is required for mitotic histone H3 Thr 3 phosphorylation and normal metaphase chromosome alignment. *Genes & Dev.* 19: 472-488.

Dai J., Sullivan B. A., Higgins J. M. (2006) Regulation of mitotic chromosome cohesion by Haspin and Aurora B. *Developmental Cell.* 11: 741-750.

De Antoni A., Maffini S., Knapp S., Musacchio A., Santaguida S. (2012) A small-molecule inhibitor of Haspin alters the kinetochore functions of Aurora B. *J Cell Biol.* 199(2):269-84.

Denslow S. A., Wade P. A., (2007) The human Mi-2/NURD complex and gene regulation. *Oncogene.* 26(37): 5433-5438.

Deuring R., Fanti L., Armstrong J. A., Sarte M., Papoulas O., Prestel M., Daubress G. Verardo M., Moseley S. L., Berloco M. (2000) The ISWI chromatin-remodeling protein is required for gene expression and the maintenance of higher order chromatin structure *in vivo*. *Mol. Cell.* 5: 355-365.

Dey A., Chitsaz F., Abbasi A., Misteli T., Ozato K. (2003) The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. *Proc Natl Acad Sci U S A.* 100(15):8758-63.

Dhalluin C., Carlson J.E., Zeng L., He C. Aggarwal A.K., Zhou M.M. (1999) Structure and ligand of a histone acetyltransferase bromodomain. *Nature.* 399: 491-496.

DiMario P., Rosby R., Cui Z. (2006) Direct visualization of GFP-fusion proteins on polytene chromosomes. *Dros Inf Serv.* 89: 115-118.

Dingwall A.K., Beek S.J., McCallum C.M., Tamkun J.W., Kalpana G.V., Goff S.P., Scott M.P. (1995) The *Drosophila* snr1 and brm proteins are related to yeast SWI/SNF proteins and are components of a large protein complex. *Mol Biol Cell.* 6(7):777-91.

Doe C.Q., Bowerman B. (2001) Asymmetric cell division: fly neuroblast meets worm zygote. *Curr Opin Cell Biol.* 13(1):68-75.

Doyen C. M., An W., Angelov D., Bondarenko V., Mietton F., Studitsky V. M., Hamiche A., Roeder R. G., Bouvet P., Dimitrov S. (2006a) Mechanism of polymerase II transcription repression by the histone variant macroH2A. *Mol. Cel.l Biol.* 26: 1156-1164.

Doyen C. M., Montel F., Gautier T., Menoni H., Claudet C., Delacour-Larose M., Angelov D., Hamiche A., Bednar J., Faivre-Moskalenko C., et al. (2006b) Dissection of the unusual structural and functional properties of the variant H2A.Bbd nucleosome. *EMBO J.* 18: 4234-4444.

Duncan E.M., Muratore-Schroeder T.L., Cook R.G., Garcia B.A., Shabanowitz J., Hunt D.F., Allis C.D. (2008) Cathepsin L proteolytically processes histone H3 during mouse embryonic stem cell differentiation. *Cell.* 135(2):284-94.

Eberharter A., Vetter I., Ferreira R., Becker P. B. (2004) ACF1 improves the effectiveness of nucleosome mobilization by ISWI through PHD-histone contacts. *EMBO J.* 23: 4029-4039.

Edgar B.A., Schubiger G. (1986) Parameters controlling transcriptional activation during early *Drosophila* development. *Cell.* 44(6):871-7.

Elfring L.K., Deuring R., McCallum C.M., Peterson C.L., Tamkun J.W. (1994) Identification and characterization of *Drosophila* relatives of the yeast transcriptional activator SNF2/SWI2. *Mol Cell Biol.*14(4):2225-34.

Euskirchen G.M., Auerbach R.K., Davidov E., Gianoulis T.A., Zhong G., Rozowsky J., Bhardwaj N., Gerstein M.B., Snyder M. (2011) Diverse roles and interactions of the SWI/SNF chromatin remodeling complex revealed using global approaches. *PLoS Genet.* 7(3):e1002008.

Fasulo B., Deuring R., Murawska M., Gause M., Dorighi K.M., Schaaf C.A., Dorsett D., Brehm A., Tamkun J.W. (2012) The *Drosophila* MI-2 chromatin-remodeling factor regulates higher-order chromatin structure and cohesin dynamics *in vivo*. *PLoS Genet.* 8(8):e1002878.

Felsenfeld G., Groudine M. (2003) Controlling the double helix. *Nature.* 421(6921):448-53.

Ferreira H., Flaus A., Owen-Hughes T. (2007) Histone modifications influence the action of Snf2 family remodeling enzymes by different mechanisms. *J. Mol. Biol.* 374: 563-579.

Filippakopoulos P., Picaud S., Mangos M., Keates T., Lambert J.P., Barsyte-Lovejoy D., Felletar I., Volkmer R., Müller S., Pawson T., Gingras A.C., Arrowsmith C.H., Knapp S. (2012) Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell.* 149(1):214-31.

Fischle W., Tseng B.S., Dormann H.L., Ueberheide B.M., Garcia B.A., Shabanowitz J., Hunt D.F., Funabiki H., Allis C.D. (2005) Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature.* 438(7071):1116-22.

Flaus A., Martin D. M. Barton G. J., Owen-Hughes T. (2006) Identification of multiple distinct Snf2 subfamilies with conserved structural motif. *Nucleic Acids Res.* 34: 2887-2905.

Fuller MT. (1998) Genetic control of cell proliferation and differentiation in *Drosophila* spermatogenesis. *Semin Cell Dev Biol.* 9(4):433-44.

Fyodorov D.V., Kadonaga J.T. (2002) Binding of Acf1 to DNA involves a WAC motif and is important for ACF-mediated chromatin assembly. *Mol Cell Biol.* 22(18):6344-53.

Gardner K.E., Allis C.D., Strahl B.D. (2011) Operating on chromatin, a colorful language where context matters. *J Mol Biol.* 409(1):36-46.

Gatti M., Bonaccorsi S., Pimpinelli S. (1994) Looking at *Drosophila* mitotic chromosomes. *Methods Cell Biol.* 44:371-91.

Gdula D. A., Sandaltzopoulos R., Tsukiyama T., Ossipow V., Wu C. (1998) Inorganic pyrophosphatase is a component of the *Drosophila* nucleosome remodelling factor complex. *Genes. Dev.* 12: 3206-3216.

Gehani S.S., Agrawal-Singh S., Dietrich N., Christophersen N.S., Helin K., Hansen K. (2010) Polycomb group protein displacement and gene activation through MSK-dependent H3K27me3S28 phosphorylation. *Mol Cell.* 39(6):886-900.

Georgel P. T., Tsukiyama T., Wu C. (1997) Role of histone tails in nucleosome remodelling by *Drosophila* NURF. *EMBO J.* 16: 4717-4726.

Gozani O., Karuman P., Jones D.R., Ivanov D., Cha J., Lugovskoy A.A., Baird C.L., Zhu H., Field S.J., Lessnick S.L., Villasenor J., Mehrotra B., Chen J., Rao V.R., Brugge J.S., Ferguson C.G., Payrastre B., Myszka D.G., Cantley L.C., Wagner G., Divecha N., Prestwich G.D., Yuan J. (2003) The PHD finger of the chromatin-associated protein ING2 functions as a nuclear phosphoinositide receptor. *Cell.* 114(1):99-111.

Grant P.A., Eberharter A., John S., Cook R.G., Turner B.M., Workman J.L. (1999) Expanded lysine acetylation specificity of Gcn5 in native complexes. *J Biol Chem.* 274(9):5895-900.

Grigoryev S.A., Woodcock C.L.. (2012) Chromatin organization - the 30 nm fiber. *Exp Cell Res.* 318(12):1448-55.

Guillemette B., Bataille A.R., Gévry N., Adam M., Blanchette M., Robert F., Gaudreau L. (2005) Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes

and regulates nucleosome positioning. *PLoS Biol.* 3(12):e384.
Hamiche A., Sandaltzopoulos R. (1999) ATP dependent histone octamer sliding mediated by the chromatin remodelling complex NURF. *Cell.* 97: 833-842.
Hake S.B., Garcia B.A., Duncan E.M., Kauer M., Dellaire G., Shabanowitz J., Bazett-Jones D.P., Allis C.D., Hunt D.F. (2006) Expression patterns and post-translational modifications associated with mammalian histone H3 variants. *J Biol Chem.* 281(1):559-68.

Hales K.G., Fuller M.T. (1997) Developmentally regulated mitochondrial fusion mediated by a conserved, novel, predicted GTPase. *Cell.* 11;90(1):121-9.

Hargreaves D.C., Horng T., Medzhitov R. (2009) Control of inducible gene expression by signal-dependent transcriptional elongation. *Cell.* 138(1):129-45.

Hassa P.O., Haenni S.S., Elser M., Hottiger M.O. (2006) Nuclear ADP-ribosylation reactions in mammalian cells: where are we today and where are we going? *Microbiol Mol Biol Rev.* 70(3):789-829.

Hassan A.H., Prochasson P., Neely K.E., Galasinski S.C., Chandy M., Carrozza M.J., Workman J.L. (2002) Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell.* 111: 369-379.

Hassan A.H., Awad S., Al-Natour Z., Othman S., Mustafa F., Rizvi T.A. (2007) Selective recognition of acetylated histones by bromodomains in transcriptional co-activators. *Biochem J.* 402(1):125-33.

Hendzel M.J., Kruhlak M.J., Bazett-Jones D.P. (1998) Organization of highly acetylated chromatin around sites of heterogeneous nuclear RNA accumulation. *Mol Biol Cell.* 9(9):2491-507.

Henikoff S., Ahmad K. (2005) Assembly of variant histones into chromatin. *Cell Deb Biol.* 32: 133-153.

Hennig L., Bouveret R., Gruissem W. (2005) MSI1-like proteins: an escort service for chromatin assembly and remodeling complexes. *Trends Cell Biol.* 15(6):295-302.

Hietz E. (1928) Das Heterochromatin der Moose. *Jahrb Wiss Bot.* 69:762-818.

Henry K. W., Wyce A., Lo W. S., Duggan L. J., Emre N. C., Kao C. F., Pillus L., Shilatifard A., Osley M. A., Berger S. L. (2003) Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes. Dev.* 17(21): 2648-2663.

Higgins J.M.G. (2009) Haspin: a newly discovered regulator of mitotic chromosome behaviour. *Chromosoma.* 119(2): 137-147.

Hirota T., Lipp J.J., Toh B.H., Peters J.M. (2005) Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature.* 438(7071):1176-80.

Ho L., Ronan J.L., Wu J., Staahl B.T., Chen L., Kuo A., Lessard J., Nesvizhskii A.I., Ranish J., Crabtree G.R. (2009) An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *Proc Natl Acad Sci U S A.* 106(13):5181-6.

Hong L., Schroth G.P., Matthews H.R., Yau P., Bradbury E.M. (1993) Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. *J Biol Chem.* 268(1):305-14.

Hudson B.P. Martinez-Yamout M.A. Dyson H.J. Wright P.E. (2000) Solution structure and acetyl-lysine binding activity of the GCN5 bromodomain1. *J Mol Biol.* 304(3):355-70.

Huet F, Lu J. T., Myrick K. V., Baugh L. R. Crosby M. A., Gelbart W. M. (2002) A deletion-generator compound element allows deletion saturation analysis for genomewide phenotypic annotation. *Proc Natl Acad Sci U S A.* 99(15): 9948-9953.

Imbalzano A. N., Kwon H., Green M. R., Kingston R. E. (1994) Facilitated binding of TATA-binding protein to nucleosomal DNA. *Nature.* 370: 481-485.
Ito T., Tyler J. K., Kadonaga J. T. (1997a) Chromatin assembly factors: a dual function in nucleosome formation and mobilization? *Genes Cells.* 2: 593-600.

Ito T., Bulger M., Pazin M. J., Kobay ashi R., Kadonaga J. T. (1997b) ACF, an ISWI containing and ATP-utilizing chromatin assembly and remodeling factor. *Cell.* 90: 145-155.

Ito T., Chiba T., Ozawa R., Yoshida M., Hattori M., Sakaki Y. (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci U S A.* 98(8):4569-74.

Jacobson R. H., Ladurner A. G., King D. S., Tjian R. (2000) Structure and function of a Human TAFII250 double bromodomain module. *Science.* 288(5470): 1422-1425.

Jiang C., Pugh B. F. (2009) Nucleosome positioning and gene regulation: advances through genomics. *Nature Reviews Genetics.* 10: 161-172.

Jin J., Cai Y., Li B., Conaway R.C., Workman J. L., Conaway J. W., Kusch T. (2005) In and out: histone variant exchange in chromatin. *Trends Biochem Sci.* 30(12): 680-687.

Johansen K.M., Johansen J. (2006) Regulation of chromatin structure by histone H3S10 phosphorylation. *Chromosome Res.* 14(4):393-404.

Jones M.H., Hamana N., Shimane M. (2000) Identification and characterization of BPTF, a novel bromodomain transcription factor. *Genomics.* 1;63(1):35-9.

Jonsson Z.O., Dhar S.K., Narlikar G.J., Auty R., Wagle N., Pellman D., Pratt R.E., Kingston R., Dutta A. (2001) Rvb1p and Rvb2p are essential components of a chromatin remodeling complex that regulates transcription of over 5% of yeast genes. *J Biol Chem.* 276:16279–16288.

Jordan-Sciutto K.L., Dragich J.M., Caltagarone J., Hall D.J., Bowser R. (2000) Fetal Alz-50 clone 1 (FAC1) protein interacts with the Myc-associated zinc finger protein (ZF87/MAZ) and alters its transcriptional activity. *Biochemistry.* 39(12):3206-15.

Josling G.A., Selvarajah S.A., Petter M., Duffy M.F. (2012) The role of bromodomain proteins in regulating gene expression. *Genes.* 3(2):320-343.

Kal A.J., Mahmoudi T., Zak N.B., Verrijzer C.P. (2000) The *Drosophila* brahma complex is an essential coactivator for the trithorax group protein zeste. *Genes Dev.* 14(9):1058-71.

Kamakaka R. T., Biggins S. (2005) Histone variants: deviants? *Genes Dev.* 19(3): 295-310.

Kellner W.A., Ramos E., Van Bortle K., Takenaka N., Corces V.G. (2012) Genome-wide phosphoacetylation of histone H3 at *Drosophila* enhancers and promoters. *Genome Res.* 22(6):1081-8.

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. (2002) The human genome browser at UCSC. *Genome Res.* 12(6): 996-1006.

Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, Sabo PJ, Larschan E, Gorchakov AA, Gu T, Linder-Basso D, Plachetka A, Shanower G, Tolstorukov MY, Luquette LJ, Xi R, Jung YL, Park RW, Bishop EP, Canfield TK, Sandstrom R, Thurman RE, MacAlpine DM, Stamatoyannopoulos JA, Kellis M, Elgin SC, Kuroda MI, Pirrotta V, Karpen GH, Park PJ. (2011) Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature.* 471(7339):480-5.

Kim J., Hake S. B., Roeder R. G. (2005) The human homolog of yeast BRE1 functions as a transcriptional coactivator through direct activator interactions. *Mol. Cell.* 20(5): 759-770.

Kimura H, Cook PR. (2001) Kinetics of core histones in living human cells: little exchange of H3 and H4 and some rapid exchange of H2B. *J Cell Biol.* 153(7):1341-53.

Kobor M.S., Venkatasubrahmanyam S., Meneghini M.D., Gin J.W., Jennings J.L., Link A.J., Madhani H.D., Rine J. (2004) A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. *PLoS Biol.* 2(5):E131.

Koh A.S., Kingston R.E., Benoist C., Mathis D. (2010) Global relevance of Aire binding to hypomethylated lysine-4 of histone-3. *Proc Natl Acad Sci U S A.* 107(29):13016-21.

Kugler S.J., Nagel A.C. (2007) Putzig is required for cell proliferation and regulates Notch activity via an epigenetic mechanism in *Drosophila*. *Mol. Biol. Cell.* 18: 3733–3740.

Kugler S.J., Nagel A.C. (2010) A novel Pzg-NURF complex regulates Notch target gene activity. *Mol. Biol. Cell.* 21: 3443–3448.

Kugler S.J., Gehring E.M., Wallkamm V., Krüger V., Nagel A.C. (2011) The Putzig-NURF nucleosome remodeling complex is required for ecdysone receptor signaling and innate immunity in *Drosophila melanogaster*. *Genetics.* 188(1):127-39.

Kwon H., Imbalzano A. N., Khavari P. A., Kingston R. E., Green M. R. (1994) Nucleosome disruption and enhancement of activator binding by a human SWI/SNF complex. *Nature.* 370: 477-481.

Kwon S. Y., Xiao H., Glover B. P., Tjian R., Wu C., Badenhorst P. (2008) The nucleosome remodeling factor (NURF) regulates genes involved in *Drosophila* innate immunity. *Dev. Biol.* 316: 538-547.

Kwon S. Y., Xiao H., Wu C., Badenhorst P. (2009) Alternative splicing of NURF301 generate distinct NURF chromatin remodelling complexes with altered modified histone binding specificities. *PLOS Genetics.* 5(7): e1000574.

Kornberg R. D., Lorch Y. (1999) Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell.* 98(3): 285-294.

Ladurner A. G., Inouye C., Jain R., Tjian R. (2003) Bromodomains mediate an acetyl-histone encoded antisilencing function at heterochromatin boundaries. *Molecular Cell.* 11(2): 365-376.
Lamonica J.M., Deng W., Kadauke S., Campbell A.E., Gamsjaeger R., Wang H., Cheng Y., Billin A.N., Hardison R.C., Mackay J.P., Blobel G.A. (2011) Bromodomain protein Brd3 associates with acetylated GATA1 to promote its chromatin occupancy at erythroid target genes. *Proc Natl Acad Sci U S A.* 108(22):E159-68.

Lan F., Collins R. E., Cegli R. D., Alpatov R., Horton J. R., Shi X., Gozani O., Cheng Z., Shi Y. (2007) Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. *Nature.* 448: 718-722.

Lange M., Kaynak B., Forster U.B., Tönjes M., Fischer J.J., Grimm C., Schlesinger J., Just S., Dunkel I., Krueger T., Mebus S., Lehrach H., Lurz R., Gobom J., Rottbauer W., Abdelilah-Seyfried S., Sperling S. (2008) Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. *Genes Dev.* 22(17):2370-84.

Längst G., Becker P.B. (2001) Nucleosome mobilization and positioning by ISWI-containing chromatin-remodeling factors. *J Cell Sci.* 114(Pt 14):2561-8.

Lau P.N., Cheung P. (2011) Histone code pathway involving H3 S28 phosphorylation and K27 acetylation activates transcription and antagonizes polycomb silencing. *Proc Natl Acad Sci U S A.* 108(7):2801-6.

Lazzaro M.A., Pépin D., Pescador N., Murphy B.D., Vanderhyden B.C., Picketts D.J. (2006) The imitation switch protein SNF2L regulates steroidogenic acute regulatory protein expression during terminal differentiation of ovarian granulosa cells. *Mol Endocrinol.* 20(10):2406-17.

Lazzaro M.A., Todd M.A., Lavigne P., Vallee D., De Maria A., Picketts D.J. (2008) Characterization of novel isoforms and evaluation of SNF2L/SMARCA1 as a candidate gene

for X-linked mental retardation in 12 families linked to Xq25-26. *BMC Med Genet.* 9:11.
Lee E. C., Yu D., Martinez de Velasco J., Tessarollo L., Swing D. A., Court D. L., Jenkins N. A., Copeland N. G. (2001) A highly efficient Escherichia Coli-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics.* 73: 56-65.

Lessard J., Wu J.I., Ranish J.A., Wan M., Winslow M.M., Staahl B.T., Wu H., Aebersold R., Graef I.A., Crabtree G.R. (2007) An essential switch in subunit composition of a chromatin remodeling complex during neural development. *Neuron.* 55(2):201-15.

Lewis Carl S. A., Gillete-Ferguson I., Ferguson D. G. (1993) An indirect immunofluorescence procedure for staining the same Cryosection with two mouse monoclonal primary antibodies. *J Histochem Cytochem.* 41: 1273-1278.

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

Li G., Widom J.(2004) Nucleosomes facilitate their own invasion. *Nat Struct Mol Biol.* 11(8):763-9.

Li H., Ilin S., Wang W., Ducan E.M., Wysocka J., Allis C.D., Patel D.J. (2006) Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature.* 442: 91-95.

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R.; 1000 Genome Project Data Processing Subgroup. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 25(16):2078-9.

Liu N., Balliano A., Hayes J.J. (2011) Mechanism(s) of SWI/SNF-induced nucleosome mobilization. *Chembiochem.* 12(2):196-204.

Liu T, Ortiz JA, Taing L, Meyer CA, Lee B, Zhang Y, Shin H, Wong SS, Ma J, Lei Y, Pape UJ,

Poidinger M, Chen Y, Yeung K, Brown M, Turpaz Y, Liu XS. (2011) Cistrome: an integrative platform for transcriptional regulation studies. *Genome Biol.* 12(8):R83.
Lois S., Blanco N., Martínez-Balbás M., de la Cruz X. (2007) The functional modulation of epigenetic regulators by alternative splicing. *BMC Genomics.* 8:252.

Lo W.S., Duggan L., Emre N.C., Belotserkovskya R., Lane W.S., Shiekhattar R., Berger S.L. (2001) Snf1--a histone kinase that works in concert with the histone acetyltransferase Gcn5 to regulate transcription. *Science.* 10;293(5532):1142-6.

Luger K., Maeder A. W., Richmond R. K., Sargent D. F., Richmond T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature.* 389: 251-259.

Luger K., Dechassa M.L., Tremethick D.J. (2012) New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nat Rev Mol Cell Biol.* 13(7):436-47.

Mahadevan L.C., Willis A.C., Barratt M.J. (1991) Rapid histone H3 phosphorylation in response to growth factors, phorbol esters, okadaic acid, and protein synthesis inhibitors. *Cell.* 65(5):775-83.

Malik H. S., Henikoff S. (2003) Phylogenomics of the nucleosome. *Nature Structural Biology.* 10: 882-891.

Mandal P., Verma N., Chauhan S., Tomar R.S. (2013) Unexpected histone H3 tail-clipping activity of glutamate dehydrogenase. *J Biol Chem.* 288(26):18743-57.

Mansfield R.E., Musselman C.A., Kwan A.H., Oliver S.S., Garske A.L., Davrazou F., Denu J.M., Kutateladze T.G., Mackay J.P. (2011) Plant homeodomain (PHD) fingers of CHD4 are histone H3-binding modules with preference for unmodified H3K4 and methylated H3K9. *J Biol Chem.* 286(13):11779-91.

Marino-Ramirez L., Kann M. G., Shoemaker B. A., Landsman D. (2005) Histone structure

and nucleosome stability. *Expert Rev. Proteomics.* 2(5): 719-729.
Matangkasombut O., Buratowski S. (2003) Different sensitivities of bromodomain factors 1 and 2 to histone H4 acetylation. *Mol Cell.* 11(2):353-63.

Martin C. and Zhang Y. (2005) The diverse function of histone lysine methylation. *Nat. Rev. Mol. Cell. Bio.* 6: 838-849.

Martinez-Balbas M. A., Tsukiyama T., Gdula D., Wu C. (1998) *Drosophila* NURF-55, a WD repeat protein involved in histone metabolism. *Proc Natl Acad Sci U S A.* 95: 132-137.

Marzluff W.F., Gongidi P., Woods K.R., Jin J., Maltais L.J. (2002) The human and mouse replication-dependent histone genes. *Genomics.* 80(5):487-98.

McKittrick E., Gafken P.R., Ahmad K., Henikoff S. (2004) Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proc Natl Acad Sci U S A.* 101(6):1525-30.

Messner S., Hottiger M.O. (2011) Histone ADP-ribosylation in DNA repair, replication and transcription. *Trends Cell Biol.* 21(9):534-42.

Mizuguchi G., Vassilev A., Tsukiyama T., Nakatani Y., Wu C. (2001) ATP-dependent nucleosome remodeling and histone hyperacetylation synergistically facilitate transcription of chromatin. *J Biol Chem.* 276(18):14773-83.

Mizuguchi G., Shen X., Landry J., Wu W.H., Sen S., Wu C. (2004) ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science.* 303(5656):343-8.

Mohr S. E., Gelbart W. M. (2002) Using P{wHy} hybrid transposable element to disrupt genes in region 54D-55B in *Drosophila melanogaster*. *Genetics.* 162: 165-176.

Morinière J., Rousseaux S., Steuerwald U., Soler-López M., Curtet S., Vitte A.L., Govin J.,

Gaucher J., Sadoul K., Hart D.J., Krijgsveld J., Khochbin S., Müller C.W., Petosa C. (2009) Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature.* 461(7264):664-8.

Mujtaba S. He Y. Zeng L. Farooq A. Carlson J.E. Ott M. Verdin E. Zhou M.M. (2002) Structural basis of lysine-acetylated HIV-1 Tat recognition by PCAF bromodomain. *Mol Cell.* 9(3):575-86.

Muller S., Filippakopoulos P., Knapp S. (2011) Bromodomains as therapeutic targets. *Expert Rev Mol Med.* 13:e29.

Murphy K., Travers P., Walport M. (2008) Janeway's Immunobiology (7th edition), Garland Science.

Musselman C.A., Kutateladze T.G. (2009) PHD fingers: epigenetic effectors and potential drug targets. *Mol Interv.* 9(6):314-23.

Musselman C.A., Kutateladze T.G. (2011) Handpicking epigenetic marks with PHD fingers. *Nucleic Acids Res.* 39(21):9061-71.

Mu X., Springer J.E., Bowser R. (1997) FAC1 expression and localization in motor neurons of developing, adult, and amyotrophic lateral sclerosis spinal cord. *Exp Neurol.* 146(1):17-24.

Nakamura Y., Umehara T., Nakano K., Jang M.K., Shirouzu M., Morita S., Uda-Tochio H., Hamana H., Terada T., Adachi N., Matsumoto T., Tanaka A., Horikoshi M., Ozato K., Padmanabhan B., Yokoyama S. (2007) Crystal structure of the human BRD2 bromodomain: insights into dimerization and recognition of acetylated histone H4. *J Biol Chem.* 282(6):4193-201.

Nathan D., Ingvarsdottir K., Sterner D. E., Bylebyl G. R., Dokmanovic M., Dorsey J. A., Whelan K. A., Krsmanovic M., Lane W. S., Meluh P. B., Johnson E. S., Berger S. L. (2006) Histone sumoylation is a negative regulator in Saccharomyces cerevisiae and shows dynamic

interplay with positive-acting histone modifications. *Genes & Development.* 20: 966-976.
Nègre N, Brown CD, Ma L, Bristow CA, Miller SW, Wagner U, Kheradpour P, Eaton ML, Loriaux P, Sealfon R, Li Z, Ishii H, Spokony RF, Chen J, Hwang L, Cheng C, Auburn RP, Davis MB, Domanus M, Shah PK, Morrison CA, Zieba J, Suchy S, Senderowicz L, Victorsen A, Bild NA, Grundstad AJ, Hanley D, MacAlpine DM, Mannervik M, Venken K, Bellen H, White R, Gerstein M, Russell S, Grossman RL, Ren B, Posakony JW, Kellis M, White KP. (2011) A cis-regulatory map of the *Drosophila* genome. *Nature.* 471(7339):527-31.

Ni J.Q., Markstein M., Binari R., Pfeiffer B., Liu L.P., Villalta C., Booker M., Perkins L., Perrimon N. (2008) Vector and parameters for targeted transgenic RNA interference in *Drosophila melanogaster*. *Nat Methods.* 5(1):49-51.

Norton V.G., Imai B.S., Yau P., Bradbury E.M. (1989) Histone acetylation reduces nucleosome core particle linking number change. *Cell.* 57(3):449-57.

Norton V.G., Marvin K.W., Yau P., Bradbury E.M. (1990) Nucleosome linking number change controlled by acetylation of histones H3 and H4. J Biol Chem. 265(32):19848-52.

Nowak S.J., Corces V.G. (2004) Phosphorylation of histone H3: A balancing act between chromosome condensation and transcriptional activation. *Trends Genet.* 20: 214-220.

Olave I., Wang W., Xue Y., Kuo A., Crabtree G.R. (2002) Identification of a polymorphic, neuron-specific chromatin remodeling complex. *Genes Dev.* 16(19):2509-17.

Ornaghi P., Ballario P., Lena A. M., Gonzalez A., Filetici P. (1999) The bromodomain of Gcn5p interacts *in vitro* with specific residues in the N terminus of histone H4. *J Mol Bio.* 287(1): 1-7.

Owen D.J., Ornaghi P., Yang J.C., Lowe N., Evans P.R., Ballario P., Neuhaus D., Filetici P., Travers A.A. (2000) The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p. *EMBO J.* 19(22):6141-9.

Papazvan R., Voronina E., Chapman J.R., Gilbert T.M., Meier E., Shabanowitz J., Hunt D.F., Liu Y.,Taverna S.D. (2013) Methylation of histone H3 at lysine 23 in meiotic heterochromatin. *Epigenetics Chromatin.* 6(Suppl 1): O13.

Pehrson J.R., Fuji R.N. (1998) Evolutionary conservation of histone macroH2A subtypes and domains. *Nucleic Acids Res.* 26(12):2837-42.

Peña P.V., Davrazou F., Shi X., Walter K.L., Verkhusha V.V., Gozani O., Zhao R., Kutateladze T.G. (2006) Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature.* 442(7098):100-3.

Peng J., Dong W., Chen L., Zou T., Qi Y., Liu Y. (2007) Brd2 is a TBP-associated protein and recruits TBP into E2F-1 transcriptional complex in response to serum stimulation. *Mol Cell Biochem.* 294(1-2):45-54.

Peterson C. L., Dingwall A., Scott M. P. (1994) Five SWI/SNF gene products are components of a large multisubunit complex required for transcriptional enhancement. *Proc Natl Acad Sci U S A.* 91: 2905-2908.

Peterson C. L., Ramkun J. W. (1995) The SWI-SNF complex: a chromatin remodelling machine? *Trends Biochem.* Sci. 20: 143-146.

Phelan M.L., Sif S., Narlikar G.J., Kingston R.E. (1999) Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. *Mol Cell.* 3(2):247-53.

Phillips D.M. (1963) The presence of acetyl groups of histones. *Biochem J.* 87:258-63.

Pogo B.G., Allfrey V.G., Mirsky A.E. (1966) RNA synthesis and histone acetylation during the course of gene activation in lymphocytes. *Proc Natl Acad Sci U S A.* 55(4):805-12.

Pokholok D.K., Harbison C.T., Levine S., Cole M., Hannett N.M., Lee T.I., Bell G.W., Walker K., Rolfe P.A., Herbolsheimer E., Zeitlinger J., Lewitter F., Gifford D.K., Young R.A. (2005) Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell.* 26;122(4):517-

27.
Polesskaya A., Naguibneva I., Duquet A., Bengal E., Robin P., Harel-Bellan A. (2001) Interaction between acetylated MyoD and the Bromodomain of CBP and/or p300. *Mol. Cell. Biol.* 21(16): 5312-5320.

Polioudaki H., Markaki Y., Kourmouli N., Dialynas G., Theodoropoulos P. A., Singh P. B., Georgatos S. D. (2004) Mitotic phosphorylation of histone H3 at threonine 3. *FEBS Letters.* 560(2004): 39-44.

Portela A., Esteller M. (2010) Epigenetic modifications and human disease. *Nat Biotechnol.* 28(10):1057-68.

Prigent C, Dimitrov S. (2003) Phosphorylation of serine 10 in histone H3, what for? *J Cell Sci.* 116(Pt 18):3677-85.

Qi D., Larsson J., Mannervik M. (2004) *Drosophila* Ada2b is required for viability and normal histone H3 acetylation. *Mol Cell Biol.* 24(18):8080-9.

Ramón-Maiques S., Kuo A.J., Carney D., Matthews A.G., Oettinger M.A., Gozani O., Yang W. (2007) The plant homeodomain finger of RAG2 recognizes histone H3 methylated at both lysine-4 and arginine-2. *Proc Natl Acad Sci U S A.* 104(48):18993-8.

Redon C., Pilch D., Rogakou E., Sedelinikova O., Newrock K., Bonner W. (2002) Histone H2A variants H2AX and H2AZ. *Curr Opin Genet Dev.* 12(2): 162-169.

Reeves R., 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(15):8573-82.

Reeves R., Beckerbauer L. (2001) HMGI/Y proteins: flexible regulators of transcription and chromatin structure. *Biochim Biophys Acta.* 1519(1-2):13-29.

Roh T.Y., Cuddapah S., Cui K., Zhao K. (2006) The genomic landscape of histone modifications in human T cells. *Proc Natl Acad Sci U S A.* 103(43):15782-7.

Ruthenburg A.J., Wang W., Graybosch D.M., Li H., Allis C.D., Patel D.J., Verdine G.L. (2006) Histone H3 recognition and presentation by the WDR5 module of the MLL1 complex. *Nat Struct Mol Biol.* 13(8):704-12.

Ruthenburg A.J., Li H., Milne T.A., Dewell S., McGinty R.K., Yuen M., Ueberheide B., Dou Y., Muir T.W., Patel D.J., Allis C.D. (2011) Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. *Cell.* 145(5):692-706.

Sakabe K., Wang Z., Hart G.W. (2010) Beta-N-acetylglucosamine (O-GlcNAc) is part of the histone code. *Proc Natl Acad Sci U S A.* 107(46):19915-20.

Sala A., La Rocca G., Burgio G., Kotova E., Di Gesù D., Collesano M., Ingrassia A.M., Tulin A.V., Corona D.F. (2008) The nucleosome-remodeling ATPase ISWI is regulated by poly-ADP-ribosylation. *PLoS Biol.* 6(10):e252.

Santos-Rosa H., Kirmizis A., Nelson C., Bartke T., Saksouk N., Cote J., Kouzarides T. (2009) Histone H3 tail clipping regulates gene expression. *Nat Struct Mol Biol.* 16(1):17-22.

Sarma K., Reinberg D. (2005) Histone variants meet their match. *Nat Rev Mol Cell Biol.* 6(2):139-49.

Sassone-Corsi P., Mizzen C.A., Cheung P., Crosio C., Monaco L., Jacquot S., Hanauer A., Allis C.D. (1999) Requirement of Rsk-2 for epidermal growth factor-activated phosphorylation of histone H3. *Science.* 285(5429):886-91.

Sato Y., Mukai M., Ueda J., Muraki M., Stasevich T.J., Horikoshi N., Kujirai T., Kita H., Kimura T., Hira S., Okada Y., Hayashi-Takanaka Y., Obuse C., Kurumizaka H., Kawahara A., Yamagata K., Nozaki N., Kimura H. (2013) Genetically encoded system to track histone modification *in vivo*. *Sci Rep.* 14;3:2436.

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., Bell S.P., Groudine M. (2004) The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev.* 18(11):1263-71.
Schuettengruber B., Ganapathi M., Leblanc B., Portoso M., Jaschek R., Tolhuis B., van Lohuizen M., Tanay A., Cavalli G. (2009) Functional anatomy of polycomb and trithorax chromatin landscapes in *Drosophila* embryos. *PLoS Biol.*13;7(1):e13.

Shen X., Mizuguchi G., Hamiche A., Wu C. (2000) A chromatin remodelling complex involved in transcription and DNA processing. *Nature.* 406(6795):541-4.

Shogren-Knaak M., Ishii H., Sun J.M., Pazin M.J., Davie J.R., Peterson C.L. (2006) Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science.* 311(5762):844-7.

Sims R.J. 3rd., Reinberg D. (2006) Histone H3 Lys 4 methylation: caught in a bind? *Genes Dev.* 20(20):2779-86.

Sohn D.H., Lee K.Y., Lee C., Oh J., Chung H., Jeon S.H., Seong R.H. (2007) SRG3 interacts directly with the major components of the SWI/SNF chromatin remodeling complex and protects them from proteasomal degradation. *J Biol Chem.* 282(14):10614-24.

Soloaga A., Thomson S., Wiggin G.R., Rampersaud N., Dyson M.H., Hazzalin C.A., Mahadevan L.C., Arthur J.S. MSK2 and MSK1 mediate the mitogen- and stress-induced phosphorylation of histone H3 and HMG-14. *EMBO J.* 22(11):2788-97.

Staynov D.Z. (2008) The controversial 30 nm chromatin fibre. *Bioessays.* 30(10):1003-9.

Stewart M.D., Li J., Wong J. (2005) Relationship between histone H3 lysine 9 methylation, transcription repression, and heterochromatin protein 1 recruitment. *Mol Cell Biol.* 25(7):2525-38.

Strachan G.D., Morgan K.L., Otis L.L., Caltagarone J., Gittis A., Bowser R., Jordan-Sciutto K.L. (2004) Fetal Alz-50 clone 1 interacts with the human orthologue of the Kelch-like Ech-associated protein. *Biochemistry.* 43(38):12113-22.

Strohner R., Nemeth A., Jansa P., Hofmann-Rohrer U., Santoro R., Langst G., Grummt I. (2001) NoRC- a novel member of mammalian ISWI containing chromatin remodeling machines. *EMBO J.* 20: 4892-4900.

Struhl K. (1998) Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev.* 12(5):599-606.

Suganuma T., Gutiérrez J.L., Li B., Florens L., Swanson S.K., Washburn M.P., Abmayr S.M., Workman J.L. (2008) ATAC is a double histone acetyltransferase complex that stimulates nucleosome sliding. *Nat Struct Mol Biol.* 15(4):364-72.

Sullivan K.F., Hechenberger M., Masri K. (1994) Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. *J Cell Biol.* 127(3):581-92.

Tamkun J.W., Deuring R., Scott M.P., Kissinger M., Pattatucci A.M., Kaufman T.C., Kennison J.A. (1992) Brahma: a regulator of *Drosophila* homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. *Cell.* 68(3):561-72.

Tsukiyama T., Wu C. (1995a) Purification and properties of an ATP-dependent nucleosome remodelling factor. *Cell.* 83: 1011-1020.

Tsukiyama T., Daniel C., Tamkun J., Wu C. (1995b) ISWI, a member of the SWI2/SNF2 ATPase family, encodes the 140 kDa subunit of the nucleosome remodelling factor. *Cell.* 83: 1021-1026.

Tsukiyama T., Palmer J., Landel C.C., Shiloach J., Wu C. (1999) Characterization of the

imitation switch subfamily of ATP-dependent chromatin-remodeling factors in Saccharomyces cerevisiae. *Genes Dev.* 13(6):686-97.

Van Ingen H., van Schaik F.M., Wienk H., Ballering J., Rehmann H., Dechesne A.C., Kruijzer J.A., Liskamp R.M., Timmers H.T., Boelens R. (2008) Structural insight into the recognition of the H3K4me3 mark by the TFIID subunit TAF3. *Structure.* 16(8):1245-56.

Varga-Weisz P. D., Wilm M., Bonte E., Dumas K., Mann M., Becker P. B. (1997) Chromatin-remodeling factor CHRAC contains the ATPases ISWI and topoisomeraseII. *Nature.* 388: 598-602.

Varga-Weisz P. D. (2001) ATP-dependent chromain remodelling factors: Nucleosome shufflers with many missions. *Oncogene.* 20:3076-3085.

Varier R. A., Outchkourov N. S., de Graaf P., van Schaik F. M., Ensing H. J., Wang F., Higgins J. M., Kops G. J., Timmers H. J. (2010) A phosphor/methyl switch at histone H3 regulates TFIID association with mitotic chromosomes. *EMBO.* 29(23): 3967-3978.

Venken K. J. T., He Y., Hoskins R. A., Bellen H. J. (2006) P[acman]: A BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. *Science.* 314(5806): 1747-1751.

Vermeulen M., Eberl H.C., Matarese F., Marks H., Denissov S., Butter F., Lee K.K., Olsen J.V., Hyman A.A., Stunnenberg H.G., Mann M. (2010) Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell.* 142(6):967-80.

Vogel-Ciernia A., Matheos D.P., Barrett R.M., Kramár E.A., Azzawi S., Chen Y., Magnan C.N., Zeller M., Sylvain A., Haettig J., Jia Y., Tran A., Dang R., Post R.J., Chabrier M., Babayan A.H., Wu J.I., Crabtree G.R., Baldi P., Baram T.Z., Lynch G., Wood M.A. (2013) The neuron-specific chromatin regulatory subunit BAF53b is necessary for synaptic plasticity and memory. *Nat Neurosci.* 16(5):552-61.

Vollmuth F., Blankenfeldt W., Geyer M. (2009) Structures of the dual bromodomains of the P-TEFb-activating protein Brd4 at atomic resolution. *J Biol Chem.* 284(52):36547-56

Wang F., Ulyanova N.P., Daum J.R., Patnaik D., Kateneva A.V., Gorbsky G.J., Higgins J.M. (2012) Haspin inhibitors reveal centromeric functions of Aurora B in chromosome segregation. *J Cell Biol.* 199(2):251-68.

Wang J. (2011) Computational study of associations between histone modification and protein-DNA binding in yeast genome by integrating diverse information. *BMC Genomics.* 12:172.

Wang Y., Zhang W., Jin Y., Johansen J., Johansen K.M. (2001) The JIL-1 tandem kinase mediates histone H3 phosphorylation and is required for maintenance of chromatin structure in *Drosophila*. *Cell.* 105(4):433-43.

Warming S., Costantino N., Court D. L., Jekins N. A., Copeland N. G. (2005) Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acid Res.* 33(4): e36.

Wei J., Zhai L., Xu J., Wang H. (2006) Role of Bmi1 in H2A ubiquitylation and Hox gene silencing. *J Biol Chem.* 281(32): 22537-22544.

White C.L., Luger K. J. Defined structural changes occur in a nucleosome upon Amt1 transcription factor binding. *Mol Biol.* 342(5):1391-402.

Widom J. (1998) Chromatin structure : Linking structure to function with histone H1. *Currunt Biology.* 8: R788-R791.

Wilson B.G., Roberts C.W. (2011) SWI/SNF nucleosome remodellers and cancer. *Nat Rev Cancer.* 11(7):481-92.

Winston F., Carlson M. (1992) Yeast SNF/SWI transcriptional activators and the SPT/SIN chromatin connection. *Trend Genet.* 8: 387-391.

Wood A., Krogan N. J., Dover J., Schneider J., Heidt J., Boateng M. A., Dean K., Golshani A., Zhang Y., Greeblatt J. F., Johnston M., Shilatifard A. (2003). Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. *Mol Cell.* 11(1): 267-274.

Woodcock CL, Skoultchi AI, Fan Y. (2006) Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. *Chromosome Res.* 14(1):17-25.

Wu J.I., Lessard J., Crabtree G.R. (2009) Understanding the words of chromatin regulation. *Cell.* 136(2):200-6.

Wysocka J. (2006) Identifying novel proteins recognizing histone modifications using peptide pull-down assay. *Elsevier Methods.* 40(4): 339-343.

Wysocka J., Swigut T., Xiao H., Milne T. A., Kwon S. Y., Landry J., Kauer M., Tackett A. J., Chait BT., Badenhorst P., Wu C., Allis CD. (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature.* 6;442(7098): 86-90.

Xiao H., Sandaltzopoulos R., Wang H. M., Hamiche A., Ranallo R., Lee K. M., Fu D., Wu C. (2001) Dual functions of largest NURF subunit NURF301 in nucleosome sliding and transcription factor interactions. *Molecular Cell.* 8: 531-543.

Yamamoto Y., Verma U.N., Prajapati S., Kwak Y.T., Gaynor R.B. (2003) Histone H3 phosphorylation by IKK-alpha is critical for cytokine-induced gene expression. *Nature.* 423(6940):655-9.

Yang Z., Yik J.H., Chen R., He N., Jang M.K., Ozato K., Zhou Q. (2005) Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol Cell.* 19(4):535-45.

Yin H., Sweeney S., Raha D., Snyder M., Lin H. (2011) A high-resolution whole-genome map of key chromatin modifications in the adult *Drosophila melanogaster*. *PLoS Genet.*

7(12):e1002380.
Yoo A.S., Staahl B.T., Chen L., Crabtree G.R. (2009) MicroRNA-mediated switching of chromatin-remodelling complexes in neural development. *Nature.* 460(7255):642-6.

Zahn L.M., Travis J. (2013) Cancer genomics. A medical renaissance? *Science.* 339(6127):1539.

Zalensky A.O., Siino J.S., Gineitis A.A., Zalenskaya I.A., Tomilin N.V., Yau P., Bradbury E.M. (2002) Human testis/sperm-specific histone H2B (hTSH2B). Molecular cloning and characterization. *J Biol Chem.* 277(45):43474-80.

Zeng L., Zhang Q., Gerona-Navarro G., Moshkina N., Zhou M.M. (2008) Structural basis of site-specific histone recognition by the bromodomains of human coactivators PCAF and CBP/p300. *Structure.* 16(4):643-52.

Zeng L., Zhang Q., Li S., Plotnikov A.N., Walsh M.J., Zhou M.M. (2010) Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. *Nature.* 466(7303):258-62.

Zentner G.E., Tesar P.J., Scacheri P.C. (2011) Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res.* 21(8):1273-83.

Zentner G.E., Henikoff S. (2013) Regulation of nucleosome dynamics by histone modifications. *Nat Struct Mol Biol.* 20(3):259-66.

Zhang S., Roche K., Nasheuer H.P., Lowndes N.F. (2011) Modification of histones by sugar β-N-acetylglucosamine (GlcNAc) occurs on multiple residues, including histone H3 serine 10, and is cell cycle-regulated. *J Biol Chem.* 286(43):37483-95.

Zhang Y., Liu T., Meyer C.A., Eeckhoute J., Johnson D.S., Bernstein B.E., Nusbaum C., Myers R.M., Brown M., Li W., Liu X.S. (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9(9):R137.