DYNAMIC HISTONE MODIFICATIONS AT THE PROMOTERS OF HOX GENES IN EMBRYONIC STEM CELLS

By

HANNAH MARY STOWER

A thesis submitted to

The University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

Chromatin and Gene Expression Group,

Institute of Biomedical Research,

The University of Birmingham

June 2009

Abstract

Histone modifications have been closely associated with changing levels of gene expression, but their role in determining, or possibly predicting, patterns of expression is uncertain. Here, the link between histone modifications and *Hoxb* gene expression in mouse embryonic stem (ES) cells was explored.

Levels of the "active" modifications H3K9ac and H3K4me3 at *Hoxb* promoters varied widely from gene to gene, but were closely correlated in ES cells. Contrastingly, the repressive modification H3K27me3 was found at equivalent levels across the cluster. Treatment with the histone deacetylase inhibitor valproate induced a coordinate increase in the levels of H3K9ac and H3K4me3 at all *Hoxb* promoters, but not other genes, whilst H3K27me3 was unaffected. Such increases were not maintained upon removal of the inhibitor.

All *Hoxb* genes were silent in undifferentiated ES cells, but expression was activated at defined times of differentiation in the expected 3' to 5' sequence. The valproate-induced increase in active modifications did not induce *Hoxb* expression from the cluster in undifferentiated cells, nor was there any major shift in the timing of *Hoxb* expression in cells transiently exposed to valproate (ie. hyperacetylated) during the start of differentiation. Thus, active histone modifications at the *Hox* genes are uncoupled from transcription.

Acknowledgments

I would like to thank my supervisor Professor Bryan Turner for his stimulating ideas and enthusiasm in guiding this project, and supporting me throughout this thesis. My sincere thanks goes to Dr. Laura O'Neill for her determined encouragement, support and guidance throughout my PhD. In addition, the members of the Chromatin and Gene Expression group (both past and present) have provided me with much advice for which I am thankful. In particular I am grateful for the counsel of Dr. Karl Nightingale.

I would also not have been able to write this thesis without the homely surroundings provided by Michael Farrelly and Lucy Mayblin, nor without the support of Mark Lees.

Finally, I must thank my parents for their unquestioning support.

Table of Contents

Table of Figures

Table of tables

1 Introduction

1.1 Regulation of Transcription Initiation

1.1.1 Prokaryotic gene regulation

The basic principles of transcriptional regulation were first elucidated by Francois Jacob and Jacques Monod at the Lac operon in *E. coli* (Jacob & Monod, 1961). In prokaryotes the promoter, in the absence of any regulatory sequences, is in a permissive ground state. Transcription initiation is limited only by the rate at which RNA polymerase can gain access to, and clear the promoter (Struhl, 1999). However, even simple prokaryotes must be able to respond to external cues, particularly in response to nutritional status, since induction of metabolic pathways is important to allow adaptation to environmental changes. To this end, multiple genes required in metabolic pathways are transcribed as long polycistronic RNAs and hence are regulated in concert. At the lac operon this is achieved by repressive and activating systems that are able to sense the relative levels of lactose and glucose in the environment respectively (Figure 1.1). The protein repressor constitutively binds the cis regulatory sequence, the operator (Figure1a). The lac repressor is a tetramer Upon its metabolism into allolactose, lactose is able to bind the repressor protein causing a conformational change and its dissociation from the regulatory DNA (Figure1.1b). This clears the promoter for RNA polymerase binding and transcription of the genes required for lactose metabolism. Significant transcription, however, requires that the CAP (catabolite activator protein) binds at a sequence upstream of the promoter as the promoter is weak. CAP is only able to bind if it itself is bound to cAMP, a metabolite that is only abundant in the absence of glucose (Figure1.1c). Thus the bacterium is fine-tuned such that it only transcribes the genes for lactose metabolism in the presence of lactose, and when there is no alternative substrate.

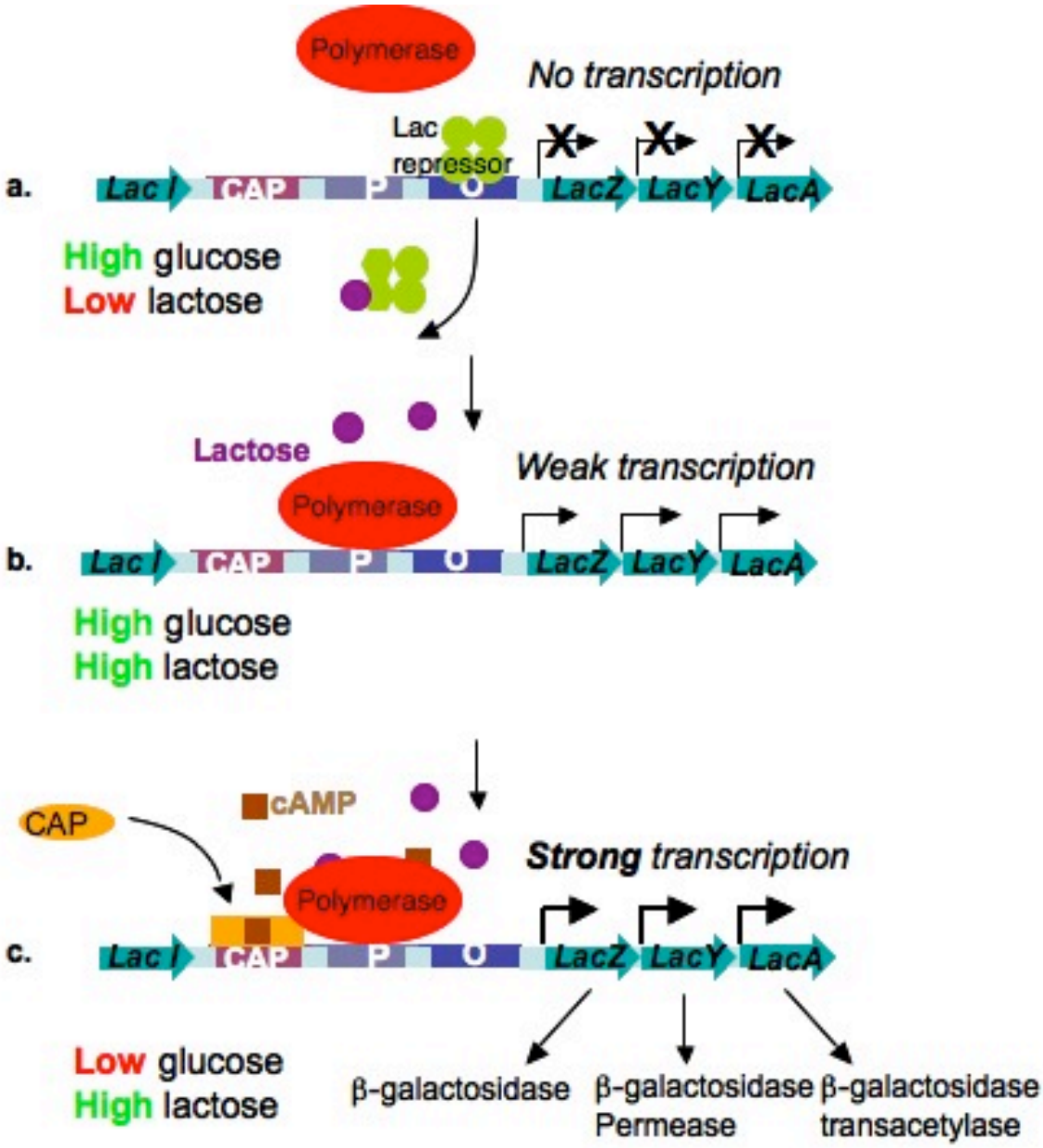


Figure 1.1 The control of the lac operon by external metabolites. a. In the presence of glucose, there is no transcription from the operon as the repressor protein binds and prevents the access of the polymerase. **b.** Lactose is able to bind the repressor protein resulting in a conformational change that causes its dissociation from the operator sequence. However, high transcription is only achieved in **c.** where low glucose levels causes an increase in cAMP concentration which binds to the CAP actively recruiting the polymerase and inducing the genes required for lactose metabolism

1.1.2 Eukaryotic Gene Regulation

The simple prokaryotic model of gene regulation is elegant and effective, however does not meet the complex needs of higher eukaryotes. In higher eukaryotes, genes are arranged adjacent to individual promoters and co-ordinately regulated genes are not transcribed as one. The genome of such organisms is vast when compared with prokaryotes. This size is compounded by the complexity that arises because the bulk of eukaryotic genomes are non-protein-coding. In addition, gene number is not correlated with complexity, but the number of non-coding sequences is, thus the *C. Elegans* genome is 30 times smaller than the human genome, implicating these additional non-coding sequences in gene regulation (Taft et al, 2007). As in prokaryotes, transcription initiation requires the recruitment of the RNA polymerase, but also general transcription factor proteins are required. In fact, across eukaryotes, the basic mechanism of transcription initiation and the proteins required are very well conserved. First, activators must bind upstream of the core promoter leading to recruitment of multi-protein adaptor complexes such as SAGA or mediator that are then able to bring in the general transcription factors to the site of transcription initiation. These general transcription factors, TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH are able to form the pre-initiation complex with RNA polymerase (Thomas & Chiang, 2006). TFIIH then “melts” the DNA at the transcription start site, opening up the two strands of DNA in order to poise the polymerase to commence transcription. The polymerase is then sequentially phosphorylated at its C terminal domain before it loses contact with the pre-initiation complex and moves on to the elongation stage. However, in contrast to the permissive ground state of a prokaryotic promoter, eukaryotic promoters are restricted, that is, trans-acting factors are a requirement for activation of these promoters (Struhl, 1999). This does however, present somewhat

of a challenge to the organism as these activating proteins must locate their cognate binding sites, as small as 6-8 base pairs long in the expanse of non-coding DNA. Furthermore, such binding must be strictly regulated as the small consensus binding sequences mean there is a high number of potential binding sites present in the genome (Thorne et al, 2009). This presents a challenge in our understanding of eukaryotic gene expression. Just what is it that produces the restrictive environment, and how are genes switched on, and indeed off?

1.2 Chromatin and transcriptional regulation

The increase in genome complexity through evolution is proposed to have been possible only through the coordinate introduction of a buffer to prevent inappropriate transcription (Bird & Tweedie, 1995). The central candidate for such a buffer is the structure of proteins and DNA termed chromatin into which DNA is packaged within the nucleus (Knezetic & Luse, 1986; Lorch et al, 1987).

Chromatin is made up of different levels of regular structure (Figure 1.2). At the lowest level, the fundamental repeating unit is the nucleosome (Luger et al, 1997). The nucleosome comprises, two histone H2A-H2B protein dimers flanking an histone H3-H4 tetramer to form an octameric globular protein core. The histone proteins are basic and so provide a good platform to package negatively charged DNA. 146 base pairs of DNA are wound around this octamer in 1 ¾ superhelical turns. Variable linker lengths are present between each nucleosome, producing a “beads on a string” structure. This structure can be further wound into the “30nm fibre”. The precise nature of the structure of this fibre is not known, and is still the topic of heated

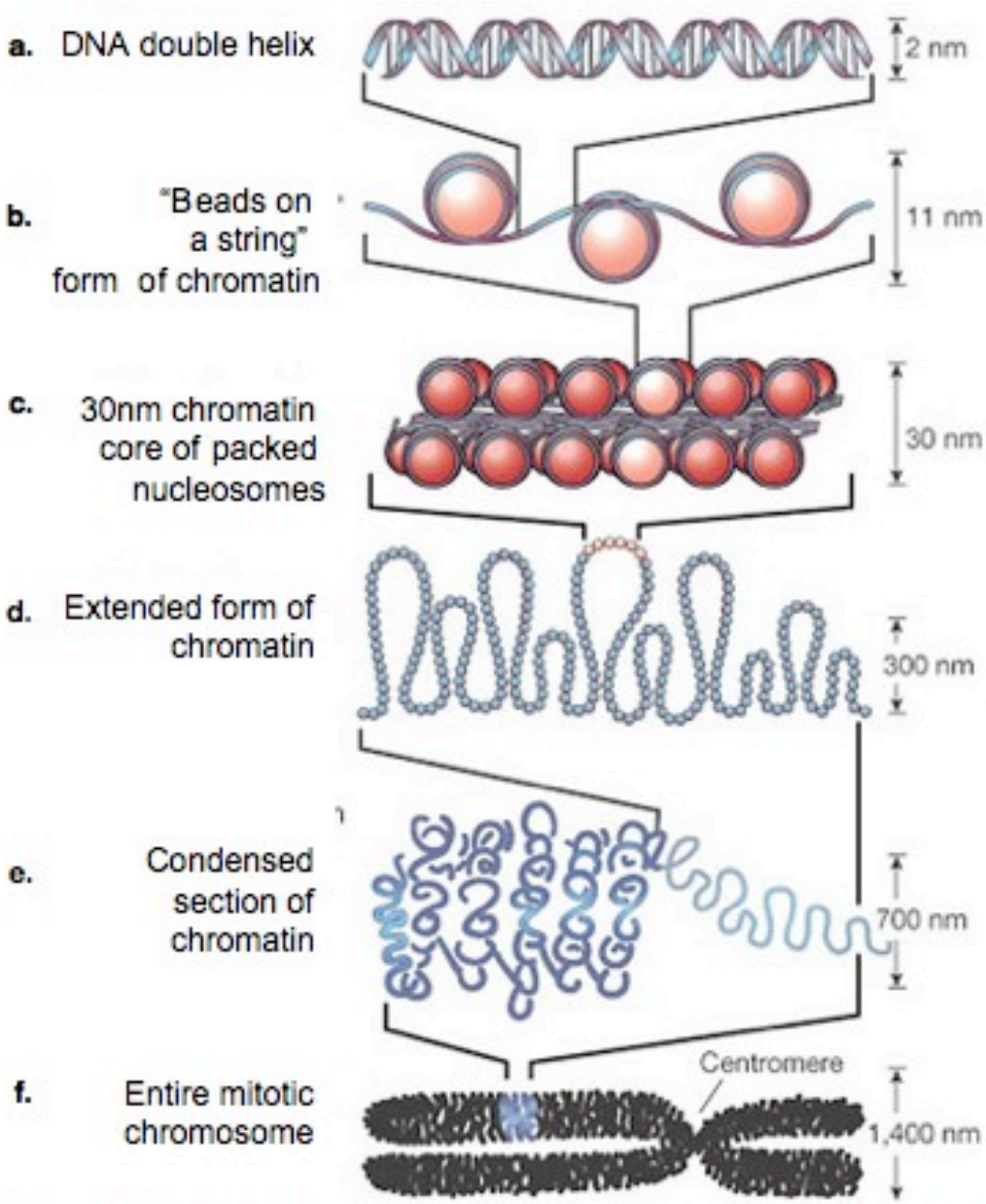


Figure 1.2 Chromatin packages DNA into the nucleus. DNA **a.** is wrapped around nuclosomes to form a "beads on a string" structure **b.** This, in turn, is further packaged into a 30nm, fibre **c.** Higher levels of packaging exist **d., e.,** with the mitotic form being the most compacted **f.** Adapted from Felsenfeld and Groudine, (2005).

debate, however it is known to involve a fifth “linker” histone protein, H1 (Satchwell & Travers, 1989). Further regular levels of structure exist that result in the compaction of DNA into the nucleus (Figure 1.2).

The existence of two forms of chromatin identifiable by electron microscopy; euchromatin, a transcriptionally more permissive gene-rich form and heterochromatin, a transcriptionally inactive gene-poor compacted form supports the role of chromatin in gene regulation. However, some essential genes are located within constitutive heterochromatin and are transcribed (Dimitri et al, 2005). Nonetheless one might speculate that chromatin does play a role in the regulation of gene transcription and that this might simply be to “block” the recruitment of both general and specific transcription factors to the DNA.

1.2.1 Regulating access to DNA

The discovery that in mammalian cells, the mouse mammary tumour virus promoter contains consistently ordered nucleosomes when integrated into mammalian cells indicated a possible role for these nucleosomes in transcriptional regulation at this promoter (Fragoso et al, 1998; Richard-Foy & Hager, 1987). This is highlighted by the fact that the second of these nucleosomes is positioned over the binding site for the glucocorticoid receptor. This protein, along with nuclear factor 1 (NF1) is involved in the transactivation of the virus. For transcription to commence, the glucocorticoid receptor must recruit ATP-dependent remodellers to the promoter (BRG or SWI/SNF) (Fryer & Archer, 1998; Muller et al, 2001). These open up the chromatin and allow the recruitment of NF1 and the assembly of the transcriptional initiation complex (Archer et al, 1992). Thus, here, the ordered positioning of nucleosomes before

activation, and their remodelling upon activation argues for a simple “blocking” role of chromatin.

The ATP-dependent remodelers were initially discovered in yeast. The SWI/SNF family were known to be transcriptional activators for a long time, though their mechanism of action remained elusive until it was discovered they could remodel chromatin upon hydrolysis of ATP (Clapier & Cairns, 2009). Thus, as is now evident, these proteins open local chromatin to allow, amongst others, transcriptional activators and the core components of the transcription complex to the DNA. However, there still remains the question of how these activators are targeted to specific genes. In addition, such transcriptional “blocking” seems a rather blunt tool with which to finely tune transcription. For this, we must look to the unstructured histone tails.

1.3 Histone tail modifications

The histone N-terminal tails that protrude from the nucleosome core particle are subject to an array of post-translational modifications. They include; phosphorylation at serine and threonine residues, acetylation at lysine residues and ubiquitination and methylation at lysine and arginine residues. These modifications are performed by specific enzymes and occur at specific amino acid residues.

1.3.1 The Dynamics of histone tail modifications

The enzymes regulating histone acetylation and deacetylation are well characterised. They consist of histone acetyl transferases (HATs) that catalyse the addition of acetyl groups onto the ε-amino group of lysine residues, and histone deactylases (HDACs)

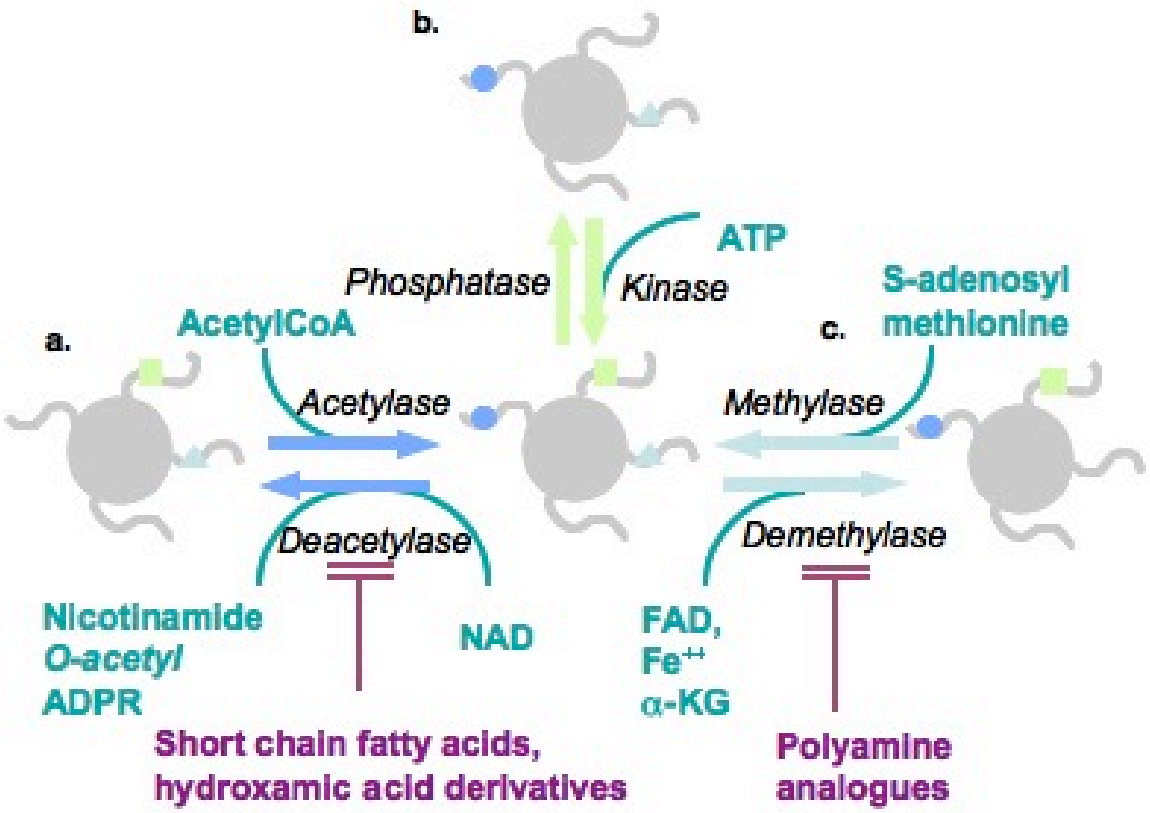


Figure 1.3 Histone tail modifications are dynamic and are regulated by metabolites. The level of post translational histone modifications on N-terminal histone tails is determined by a dynamic equilibrium of enzymes. These equilibria, in turn are balanced by the level of available metabolites. **a.** Lysine acetylation (blue circle) may be put in place by histone acetyl transferases and removed by histone deacetylases (HDACs). HDACs are sensitive to a variety of external factors. **B.** Serine phosphorylation (green square) may be put in place by kinases and removed by phosphatases and is sensitive to the levels of adenine phosphates.**c.** Lysine methylation is put in place by lysine methylases and removed by lysine demethylases. The demethylase may be repressed by polyamine analogues.

that catalyse their removal from the lysine group (Carmen et al, 1996; Johnson & Turner, 1999; Sterner & Berger, 2000) (Figure 1.3). The HATs are grouped into three main families; GNAT, MYST and CBP/p300. Generally histone acetyl transferases show specificity for more than one lysine. The histone deactylases are also classified into three groups; class I, class II and the class III NAD dependent family. In general, these enzymes are of broad specificity. Interestingly, the class I enzymes may be inhibited by naturally occurring short-chain fatty acids and class III by nicotinamide (Johnson & Turner, 1999).

The histone lysine methyl mark is deposited by the lysine methyltransferase enzymes. These transfer the methyl group of s-adenosyl methionine to the specific lysine group of the histone tail. They are highly specific when compared with the acetyl transferases (Zhang & Reinberg, 2001). The histone lysine demethylases, also show residue specific activity (Bannister & Kouzarides, 2005). In addition, methylation shows further complexity in that lysine molecules may be mono, di or tri methylated with the enzymes themselves showing specificity not only for the residue, but also for the level of methylation. Thus far, only the enzymes that methylate, not demethylate the arginines of histones have been elucidated. However, although no demethylating activity has been found, a deimination enzyme has and this activity is found to antagonise that of histone arginine methylation (Cuthbert et al, 2004). There are many more enzymes that have been characterised that may modify histone tails, further relatively-well characterised groups are the kinases and phopshatases involved in depositing and removing the phosphate mark (Nowak & Corces, 2004) and the enzymes involved in ubiquitination and deubiquitination (Shilatifard, 2006).

From this equilibrium of histone modifying enzymes, two important points have been made. Firstly, histone tail modifications are dynamic and are therefore a reflection of the action of the enzymes that deposit and remove them (Figure 1.3). They are thus likely to localise to different parts of the genome at different developmental stages, or in response to different external signals to maintain or alter histone modifications at different loci. Secondly, the dependence of many of the histone tail modifying enzymes upon both intracellular and extracellular cofactors means that the nucleosome must, at least to some extent, reflect the metabolic state of the cell and the environment in which the cell finds itself (Figure 1.3). Thus, the nucleosome could potentially be a signalling molecule at the end of a signal transduction cascade, regulating gene expression in response to external cues (Johnson & Turner, 1999; Thorne et al, 2009).

1.3.2 Charge-specific effects of histone modifications

At their initial discovery, it was proposed that histone modifications could induce their transcription-related outcomes via the modulation of the overall charge of the nucleosome. Thus, histone acetylation would result in a decrease in the positive charge of the nucleosome, decreasing the attraction between the negatively charged DNA and the nucleosome. Thus, transcriptional activation at a locally acetylated region of chromatin would merely be a result of an “opening” of chromatin due to neutralisation of the basic tails. Evidence for such charge-specific effects has been displayed in artificially reconstituted nucleosome arrays(Tse et al, 1998). The incorporation of acetylated histones prevented the formation of the 30-nanometre like fibres. This is consistent with a more “open” form of chromatin. However, charge effects alone could not explain the H4K16ac-specific effects seen in a different study. In this study, a series of yeast mutants were generated mutating each histone lysine

acetylation site at histone H4 to an arginine, both individually and combinatorially (Dion et al, 2005). This builds on previous work, that showed that mutation at these sites was not lethal but did confer phenotypes consistent with a role for these acetyl residues in gene expression, nuclear division and DNA replication (Megee et al, 1990). Another earlier study mutating the H4 tail also showed the H4 acetylation sites to be important in gene regulation (Durrin et al, 1991). The mutation of lysine to arginine mimics the charge effects of this residue in its unacetylated state and does not allow acetylation. Transcriptional effects on mutation at the sites were cumulative except for at histone H4K16. Thus, mutation at all other sites was consistent with charge effects alone (Dion et al, 2005). Such a result is not surprising if one considers that not all histone tail modifications result in the alteration of charge at the tail, for example methylation. Thus some other mechanism must exist by which histone tails may act to activate or repress transcription. The observation that differently acetylated H4 isoforms localise to specific regions in *Drosophila* polytene chromosomes led to the hypothesis that the nucleosome tails may act as specific docking modules for non-histone proteins by means of their post-translational modifications (Turner et al, 1992). Such non-histone proteins could then modulate chromatin structure and function themselves by bringing in further histone modifying activities or transcription factors to name but a couple of potential examples.

1.3.3 Chromatin readers

For such a model of protein recruitment to be feasible, there must be some means for the post-translational modifications to be interpreted; therefore proteins must exist that can firstly bind to the specifically modified histone, and secondly recruit gene repression or activation activities. Furthermore, it would be expected that some

combinations of modifications actively deter the binding of some transcriptional modulators.

The acetylated lysine motif is “translated” by the bromodomain. This domain was first described in *Drosophila* in the brahma transcriptional activator (Tamkun et al, 1992). In this study, it was noted that this motif was common to many transcriptional activators across eukaryotes. Since then, bromodomains have been found in many chromatin associated proteins, notably chromatin remodellers and most histone acetyl transferases (Horn & Peterson, 2001). A study in yeast went on to confirm the role of bromodomains in remodeller recruitment (Hassan et al, 2002). The study displayed the absolute dependence of the binding of the chromatin-remodelling factor SWI/SNF upon both its bromodomain and previous acetylation at the promoter by the yeast HAT complexes SAGA or Nu4; thus building up a sequential picture of gene activation. Since then human homologues have been elucidated that are also dependent upon their bromodomains (Chiba et al, 1994). A direct role for this domain in activating transcription is displayed in the TFIID constitutive transcription factor subunit TAF1. TAFIID is required for the recruitment of pol II to the TATA box. The structure of a fragment of the TAF1 protein containing the double bromodomain from humans revealed tandem domains suitable for binding to diacetylated histone H3K9 and K14 (Jacobson et al, 2000). Such a tandem arrangement of domains confers specificity of protein binding. Indeed, bromodomains, whilst conserved across proteins and species are variant, this variance no doubt influences binding specificity. A role for protein context in specificity has also been inferred (Hassan et al, 2002). In accordance with its recruitment role, the bromodomain is found widely distributed across histone modifying enzymes and chromatin remodellers (de la Cruz et al, 2005). Furthermore, its presence at histone acetyl transferases immediately suggests

a model of self-potentiation of the mark through binding, modifying and further recruitment.

The first discovered histone methyl mark translator was the chromodomain. It was first identified in the heterochromatin protein HP1 and *Drosophila* polycomb protein (Paro & Hogness, 1991). Both polycomb and HP1 are involved in stable gene silencing which will be further discussed later. HP1 binds specifically at H3K9me3 (Bannister et al, 2001). Gene repression may be induced by artificially targeting the H3K9 methylation enzyme activity to genes in human cells indicating that this modification is sufficient to initiate gene silencing (Snowden et al, 2002). In addition to its interaction with H3K9me3, HP1 interaction with the suv39 methyl transferase is also a requirement for the recruitment of this protein to heterochromatin in mammals (Stewart et al, 2005). The interaction between the two proteins, also suggests a self-propagating mechanism of heterochromatin whereby HP1 can recruit further methylase activity which, in turn, recruits further HP1. Thus, it can be envisaged that methylation spreads along the chromosome in concert with these two proteins.

Conversely, histone lysine methyl marks at H3K4me3 are known to be associated with active genes. This mark is catalysed by the Set 7/9 and MLL family of histone methyl transferases in humans (Milne et al, 2002; Nishioka et al, 2002; Wang et al, 2001) and in yeast by Set1 (Briggs et al, 2001). A translator of this mark is the binding protein Chd1. Chd1 has been demonstrated to function in chromatin remodelling, gene expression and transcriptional elongation in yeast (Simic et al, 2003; Tran et al, 2000). The CHD family of proteins have two chromodomains and an ATP dependent chromatin remodelling activity (Woodage et al, 1997). Chd1 was found to be a part of the yeast SAGA and SLK complexes; large histone acetyl

transferase containing gene-activating complexes (Pray-Grant et al, 2005). One of the chromodomains was shown to bind specifically to H3 trimethylated at lysine 4, thus coupling its role of transcriptional activation with binding to chromatin. Thus, reading the histone post-translational modifications gives a defined transcriptional activity. This role may be, however, organism specific since it was found in a different study that unlike human chd1, yeast Chd1 was unable to bind to H3K4me3 (Sims et al, 2005). Chromodomains are part of the tudor domain "royal family" that includes the tudor and MBT domains (Kim et al, 2006a). Tudor and MBT domains have also been shown on arrays to posses the ability to bind to methylated histone H3 and H4 tails (Kim et al, 2006a). There is also another domain, the PHD finger domain that has been shown to link H3K4me3 with active transcription (Wysocka et al, 2006). This domain is found in the nuclear remodelling complex NURF, an ISWI containing nuclear remodelling complex. Depletion of H3K4me3 causes a decrease in binding of the subunit containing this domain and the bromodomain preventing the recruitment of the ATPase subunit to chromatin. This causes the repression of a subset of genes including *Hoxc8*.

There still remain many modifications for which the details of the protein 'translators' are not known. An interesting mark, H3K36me3, currently has no clear 'readers'. In yeast, the H3K36me3 depositing enzyme set2 is known to associate with the elongating polymerase, providing a direct link for this modification in transcription, though the exact mechanism is not known (Kizer et al, 2005). It is known however, that acetylation at H4 is required for set2 binding, and hence the enzyme itself contains an H4 acetylation chromatin reader (Du et al, 2008). Furthermore, a common acute myeloid leukaemia (AML) gene nuclear receptor-binding SET domain protein 1 (NSD1) that is translocated in AML is able to methylate at H3K36me3. The

fused enzyme is able to induce the *HoxA7* and *HoxA9* proto-oncogenes by maintaining histone H3K36methylation by binding through its PHD domain(Wang et al, 2007). The direct mechanisms by which this modification is able to active these *HoxA* genes in mice is still unknown. Thus the translation of histone modifications is still very much an active part of research..

Interestingly, domains also exist that may only bind to unmodified tails. The *de novo* DNA methyltransferases DNMT3A is dependent upon the stimulation by DNMT3L, an associated protein. The N terminal domain of this protein is only able to bind H3 tails that are unmethylated H3K4, and hence those genes that are already marked for transcription by this modification are unable to be silenced by the methylation of their DNA(Ooi et al, 2007). Thus unmodified tails may 'mark' genes to be silenced and hence require their own translators.

1.3.4 Combinatorial possibilities and the histone code hypothesis

Whilst the exact significance of the chromodomain within the yeast SAGA and SLK complex remains controversial, it does introduce the concept that histone modifications may act in concert. Here, a methylated residue brings acetylation to adjacent residues. Since bromodomains are present in histone methylating enzymes, these may act to produce a defined combination of marks whereby a histone methytransferase binds chromatin via its bromodomain and then methylates, in cis, adjacent residues (Loyola & Almouzni, 2004). Thus, one specific modification may not result in a functional outcome alone, but when placed in the context of other modifications, it may result in active or inactive transcription by recruiting multiple non-histone proteins, or proteins with more than one domain that recognises modified histones. Indeed, it may be that the deposition of one mark predisposes that

histone tail for further enzyme modification establishing a mechanism for producing combinations of marks. Such a mechanism has been shown to exist in the cross-talk between histone H3K4 methylation and histone acetylation. The action of the H3K4-specific methyl transferase MLL is stimulated by histone substrates that are pre-acetylated (Nightingale et al, 2007). Given such combinatorial possibilities at the histone tails, a histone code hypothesis was proposed (Jenuwein & Allis, 2001; Strahl & Allis, 2000; Turner, 1993; Turner, 2000; Turner, 2002). This suggested that the combination of histone modifications result in a specific function being conveyed at a chromatin locality. Much of this can be explained by simple biochemistry. The modifications provide docking sites for (i) chromatin remodellers that may open or close the chromatin, (ii) transcription factors that may induce transcription, (iii) histone modifying enzymes that may add further modules to the code (iv) other proteins that convey distinct functional outcomes upon chromatin. Such recruitment or in other cases exclusion is the result of steric interactions between the proteins. To really take the code further the combination of modifications must be predictive of a certain functional outcome, i.e. they must be causative, not just permissive or consequential of such an outcome.

A good example of combinatorial outcomes is provided by the TFIID transcription factor. As discussed above, the TAF1 subunit contains tandem bromodomains able to bind at tandemly acetylated lysines. However, in addition to this, the TAF3 subunit of TFIID is able to bind to the actively associated H3K4me3 modification via its PHD domain and this interaction of TFIID with H3K4me3 is potentiated by the acetylation of H3K9Ac and H3K14Ac, presumably due to their interaction with TAF1 (Fig 1.4). Furthermore, the interaction of TAF3 with H3K4me3 is inhibited by the asymmetric methylation of H3R2 (Jacobson et al, 2000). Thus, here there is a combination of

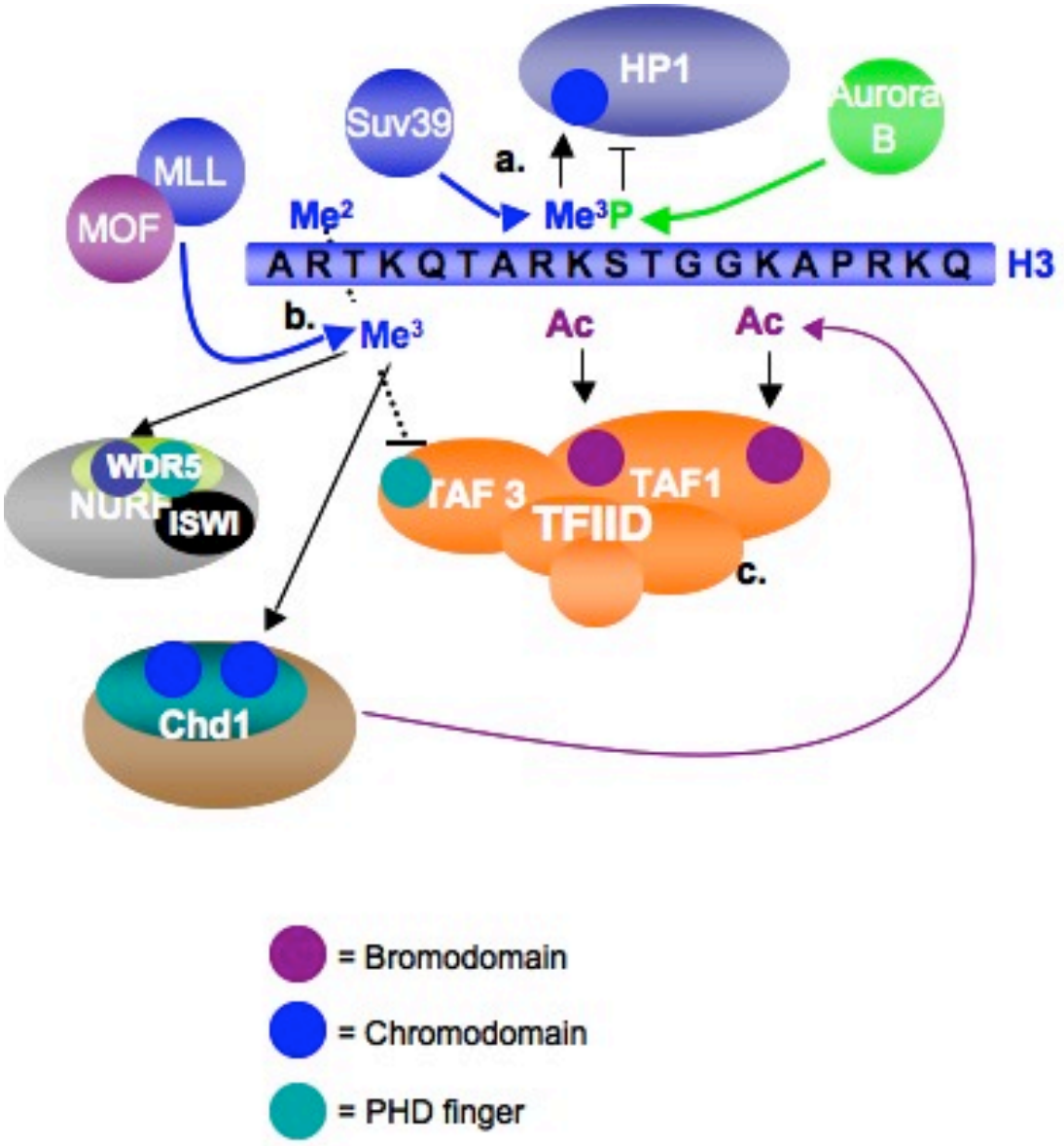


Figure 1.4 Histone tail modifications may act in combination and individually to recruit proteins to chromatin. Histone modifications are put in place by specific histone tail modifying enzymes. These modifications may modify the surface of the histone tail in a manner that favours or prevents protein-protein interactions. **a.** The trimethyl mark at lysine 9 on H3 is put in place by the Suv39 methyltransferase and favours HP1 interaction. Phosphorylation at the adjacent serine10 by Msk1 ejects HP1.**b.** Trimethylation at lysine 4 of H3 is performed by MLL. MLL may be found in a complex with MOF, an H4 acetyl transferase. Proteins may interact with this modification using a PHD finger as in the chromatin remodelling protein NURF, or using a chromodomain as in the SAGA/SLK complex. This interaction may lead to adjacent acetylation **c.** Proteins may interact with acetyl marks using bromodomains as in the TAF1 subunit of the TFIID core transcription factor which also contains a PHD finger in the TAF3 subunit. The interaction is inhibited by R2 asymmetric dimethylation.

modifications that may lead to the recruitment or 'blocking' of a constitutive transcription factor leading to transcription, or not, as appropriate.

Following on from this 'negative' outcome produced by H3R2me2 another 'negative' outcome is exemplified by HP1's interactions with chromatin. The HP1 protein, as previously discussed, is able to recognise the H3K9me3 mark, localising this protein to heterochromatin. Initial *in vitro* experiments showed that the binding of HP1 to H3K9me3 modified peptides was ablated by the adjacent modification phosphorylation of H3S10 resulting in the hypothesis of a "methyl-phosphoryl" switch (Fischle et al, 2003a) (Figure 1. 4). *In vivo,* however, the two modifications H3K9me3 and H3S10 phosphorylation were shown to coexist at the beginning of mitosis and HP1 ejection was a result of coordinated H3K14 acetylation and H3S10 phosphorylation (Mateescu et al, 2004). This is however, consistent with previous findings that the Aurora B kinase, the kinase responsible for depositing the H3S10 phosphorylation mark is required for the ejection of HP1 from heterochromatin during mitosis (Hirota et al, 2005).

The H3S10P mark is not only involved in the ejection of HP1 in mitosis. It has been shown to be a final phosphorylation mark in immediate early gene signal transduction upon exposure to mitogens or stress (Mahadevan et al, 1991). This is induced on the MAP kinase pathway by the Rsk-2 or Msk-1 kinases, or on the p38 stress activated by Msk-1 (Thomson et al, 1999). Again, as in mitotic heterochromatin, there is strong evidence for a coupling of this mark with histone acetylation upon mitogenic stimulation using antibodies to the doubly modified mark H3K9acH3S10P (Clayton et al, 2000). In addition, histone phosphorylation was shown to precede acetylation *in*

vivo and *in vitro.* H3 acetylation was stimulated by an H3S10 phosphorylated peptide (Cheung et al, 2000).

The presence of more than one histone modifying enzyme in a complex together further argues for a combinatorial code. For example, the histone acetyl transferase, MOF associates with the methyl transferase MLL1 (Figure 1.4). Both are responsible for the implementation of so-called "active" marks H4K16ac and H3K4me3 respectively. They were shown to coordinate methylation and acetylation both in vitro and in vivo, and to both be required for the activation of known MLL target genes and in an in vitro transcription assay (Dou et al, 2005). Furthermore, the mechanism of this complex recruitment is known to be reliant upon the WDR5 component of the MLL complex that is able to bind to histones di and tri-methylated at histone H3K4 using its WD40 repeat domain. Note that the specificity of WDR5 for di and tri-methyl H3K4 means that it is not involved in the initial recruitment of MLL to unmethylated H3K4, however, it is involved in a self-potentiation mechanism of this mark. A study in yeast further elucidated a mechanistic link between the active H3K4me3 and acetylation marks, though this time at H3K14ac. In this example, the histone "reader" is a PHD finger at the Yng1 protein in the NuA3 histone acetyl transferase complex. Disrupting the binding of the acetyltransferase complex at H3K4me3 resulted in a decrease in H3K14ac levels and altered transcription in a subset of NuA3 target genes (Taverna et al, 2006). This example thus further supports the hypothesis of a combinatorial code.

A final fascinating example of histone modifications talking to one another code is provided by the link of H2B ubiquitination in trans with subsequent H3K4 and H3K79 methylation, again in yeast (Dover et al, 2002; Ng et al, 2002; Sun & Allis, 2002). The

link between these reactions is mediated by the transcription elongation complex Paf1 which is proposed to recruited to the ubiquitination mark and remodel chromatin in order allow the recruitment of histone methyl transferases (Ezhkova & Tansey, 2004). The monoubiquitination H2B mark, this time at K120, has also shown to be important in transcriptional regulation of the *Hoxa* and *c* genes in humans. Importantly it appears that this mechanism of cross-talk between H2B ubiquitination and H3K4me3 and H3K79me3 is conserved between yeast and humans (Zhu et al, 2005).

The protein domains that translate combinations of marks must also occur in tandem if they are to read a combinatorial code. The tandem bromodomains of the TFIID transcription factor subunit TAF1 have already been discussed above, but equally proteins have been discovered that contain two different chromatin-binding domains that cooperate. For example, the PHD finger domain is often found adjacent to chromodomains and bromodomains as exemplified by TFIID and was shown to be important in conferring the specificity of binding of p300, a transcriptional activator (Ragvin et al, 2004).

1.3.5 Histone codes induced by cell signalling

A comprehensive study of the induction of the interferon promoter upon viral stimulation led to the elucidation of a specific combination of histone lysine modifications read by bromodomain containing proteins at this promoter. This study exemplified much of the major elements required for a histone code; modifications occurred sequentially and were accompanied by sequential recruitment of transcriptional activators followed by transcriptional activation (Agalioti et al, 2002). Through the use of recombinant mutant histone proteins, the relative importance of

sites of histone acetylation for protein recruitment was established. Acetylation of histone H4 at lysine 8 by GCN5 (a HAT) was required for the binding of the BRG1 bromodomain containing subunit of the SWI/SNF complex, and acetylation at lysines 9 and 14 by the same protein required for the recruitment of the transcription factor TFIID (TAF1 is the bromodomain containing subunit of this transcription factor as discussed above)(Jacobson et al, 2000). Surprisingly, despite its recruitment to the promoter, p300 was not required for these acetylation events. The conformational change induced at the promoter by the binding of TFIID causes the SWI/SNF protein to act by sliding the previously modified nucleosome downstream and allowing transcription to commence (Lomvardas & Thanos, 2001). Thus recruitment of the SWI/SNF by its bromodomain at specifically modified lysines results in chromatin remodelling and a specific transcriptional outcome.

Another example of cell signalling to the nucleus resulting in histone modifications changes is that of the immediate early signalling pathway as discussed above. Here, mitogenic stimulation of cells in G_0 results in the phosphorylation of a select group of immediately genes such as myc and fos at H3 serine10. This results in the recruitment of 14-3-3 proteins specifically to these regions of chromatin (Macdonald et al, 2005). Such an interaction is not adversely affected by the acetylation of H3K9Ac or H3K14Ac, and may even form a part of a combinatorial code. Such 14-3-3 proteins are linked to control of a wide range of biological processes such as cell cycle timing, signal transduction, cell death, and functions related to transcription and histone modification. Thus, this link of H3S10 with these proteins provides an insight into the function of the modification.

Finally, a more complex picture of sequential events was determined at the oestrogen receptor alpha's target promoter, pS2 (Metivier et al, 2004). Since this is a hormonally induced promoter, conclusions are most likely only relevant to such promoters but nevertheless show significant interest. The studies characterised the sequential recruitment of around thirty different proteins and four histone modifications to the promoters at the oestrogen receptor alpha target promoter pS2 in a cell culture system. Synchronicity in induction was achieved by treating the cells with a transcriptional inhibitor before introducing the oestrogen receptor ligand. The proteins studied included; histone modifying enzymes, transcription factors, central components of the transcriptional machinery and nucleosome remodelling enzymes. Changes in histone modifications and nucleosome remodelling directed transient recruitment of different protein complexes through three different cycles and were thus postulated as being the "clock" mediator. Interestingly, the first cycle does not result in transcription, but poising the promoter ready for transcription (Metivier et al, 2004). Also, protein complexes often associated with transcriptional repression were important for clearing active states between cycles. From this study, and that at the interferon promoter it is clear that a histone code is unique at each promoter. In addition, the code may be different dependent upon the stimulus since histone modifications induced at the Hsp70 promoter were shown to be unique dependent upon whether induction was caused by heat shock or by sodium arsenite, and therefore not simply a measure of transcriptional induction in general (Thomson et al, 2004).

1.3.6 Multiple roles for histone marks

Despite the link of the two marks H3K4me3, and H2B mono-ubiquitination discussed above by a transcriptional elongation complex, the readout of these marks does not

always result in gene activation. At some loci these two marks are involved in gene silencing, thus the manner in which the code is read is of utmost importance and most probably relies upon some other mark or mechanism to provide context for the mark reading (Fingerman et al, 2005). Similarly, in mammals, association of the ING family of PHD domain proteins at H3K4me3 confers a repressive read-out of the H3K4 trimethyl mark, contrary to active readout by Chd1 and Wdr5 95 (Santos-Rosa et al, 2003; Shi et al, 2006). This ING protein mediates its repressive role further by recruiting HDAC complexes to its target promoters. Thus this protein was speculated to have a particularly important role in rapid shut down of gene transcription during acute stress, for example during the DNA damage response. Furthermore, H3K9me3 and HP1 proteins are not always associated with gene silencing (Vakoc et al, 2005). H3K9me3 was shown to be enriched at induced erythroid specific genes in an erythroid cell line. This enrichment was shown to be coordinated with the binding of a form of the HP1 protein, HP1 gamma. Thus it appears that the result of a histone modification very much depends upon its protein translator and, as yet unidentified context.

Building on this context-dependence of histone modifications. The H3K36me3 mark, a mark often associated with transcriptional elongation, was recently shown to be preferentially associated with exons in *C. elegans*. (Kolasinska-Zwierz et al, 2009). This marking was present only in highly expressed genes and was not so at alternatively included exons, suggesting the modification marks the exon for inclusion. This pattern was shown to be conserved through to humans and mouse. A further study showed that, in fact the exons are enriched for nucleosomes, and other modifications including the H3K36me3 modification (Schwartz et al, 2009). In fact, the levels of these modifications varied in the exons from the 5' to the 3' end of the

gene, thus suggesting that the marking of exons by nucleosomes may have a role in defining the exon-intron structure of a gene. Thus, the histone 'code' need not apply only to transcription.

1.4 The location of histone marks within genes

If histone marks are truly involved in the coordination of transcription, be it initiation or elongation, then one would expect their distribution across active or inactive genes to follow a non-random distribution i.e. those that are involved in the initiation of transcription would be found at the promoter, and those involved in elongation in the body of the gene. Indeed, initial studies in yeast showed the exclusive association of H3K4me3 with transcriptionally active genes and H3K4me2 within euchromatic sites (Noma et al, 2001; Santos-Rosa et al, 2002). Further to this, the enzyme Set1, responsible for depositing the H3K4me3 mark was found to associate with a form of the polymerase that immediately precedes the elongation phase of transcription leaving H3K4me3 at the coding regions of genes (Ng et al, 2003). Thus H3K4me3 was postulated as having a direct role in initiation of transcription consistent with its interaction with Chd1 as discussed above. H3K36me2 has also been postulated as having a direct role in transcription from evidence in yeast. It is found in actively transcribed regions within yeast and its deposition enzyme, Set2, is found associated with the elongation-termination form of the polymerase (Krogan et al, 2003; Xiao et al, 2003).

1.4.1 Genome-wide technology and histone modifications

The advent of microarray and high-throughput sequencing technologies has meant that chromatin immunoprecipitation studies are no longer limited to the study of one

or a group of genes. These have allowed the further characterisation of histone modifications both within genes, from their promoters to their site of transcription termination, and across large expanses of chromosome. When examining these data, it is wise to keep in mind the limitations of such genome-wide experiments. Firstly, as with conventional ChIP, antibody specificity is of utmost importance. For example, antibodies against H3K9me3 and H3K27me3 must be tested thoroughly as their adjacent peptide sequences are identical. Second, again as with conventional ChIP, different methods for preparing chromatin may yield different results; the most fundamental difference being whether nucleosomes are generated by cross-linking and sonication, or by nuclease digestion (O'Neill & Turner, 1996). Finally, such experiments are performed on asynchronous populations of cells and provide a snapshot of this population at that given point in time. Thus, modification states are likely to be highly heterogenous. Also, when dealing with microarray data, the additional amplification step required to generate enough DNA label a slide may generate a bias toward specific sequences.

In yeast, the initial ChIP-chip studies confirmed the location of H3K4me3 at the promoters of active genes, with dimethylation and monomethylation of the same residue being sequentially more enriched toward the end of genes (Pokholok et al, 2005). Consistent with its role in elongation, H3K36me3 was found to localise to the body of active genes and acetylation of both H3 and H4 was found to localise to the promoters of genes (Krogan et al, 2003; Pokholok et al, 2005).

In mammals, ChIP-chip data exists for a variety of different cell types at different developmental stages including human and mouse embryonic stem cells, neural progenitor cells, fibroblasts, hepatocytes and T cells (Barski et al, 2007; Li et al,

2007; Mikkelsen et al, 2007; Roh et al, 2006). The localisation of marks within genes is shown to be highly consistent across the studies and so called “typical” patterns are highly predictive of transcriptional start sites (Birney et al, 2007). Active genes are enriched for H3K4me3 and H3 and H4 acetylation in their promoters (Pan et al, 2007) (Figure 1.5). They also show high levels of monomethylation at H3K27, H3K9, H4K20 and H3K79 (Figure 1.5). Inactive genes show high levels of trimethylation at H3K27 and H3K9. Thus, the consistent association of specific marks with active or inactive genes further supports a direct role for these marks in transcription by dictating a histone code. For the purpose of this thesis the marks that are found associated with active genes shall be termed active modifications and those associate with silent genes repressive marks. It must be remembered, however, that such studies are only correlative.

Intriguingly, many of these genome-wide studies show that there is a difference in the patterns at CpG-poor and CpG-rich promoters. At CpG poor promoters, there is a correlation between active transcription and H3K4 methylation (Schubeler et al, 2004). However, at CpG rich promoters, H3K4methylated forms are often found at the promoters independent of transcription (Barrera et al, 2008; Mohn et al, 2008; Weber et al, 2007). Interestingly, this is reliant upon the lack of methylation at the islands indicating a transcriptionally more “permissive” environment. Furthermore, the presence of methylated H3K4 prevents the binding of the de novo methylase Dnmt3a, and so methylation at H3K4 could also function to prevent de novo methylation at important regulatory elements.

A final point to note when considering the relationship of histone modifications and transcription is; what do we actually mean by “active” transcription? The tightly

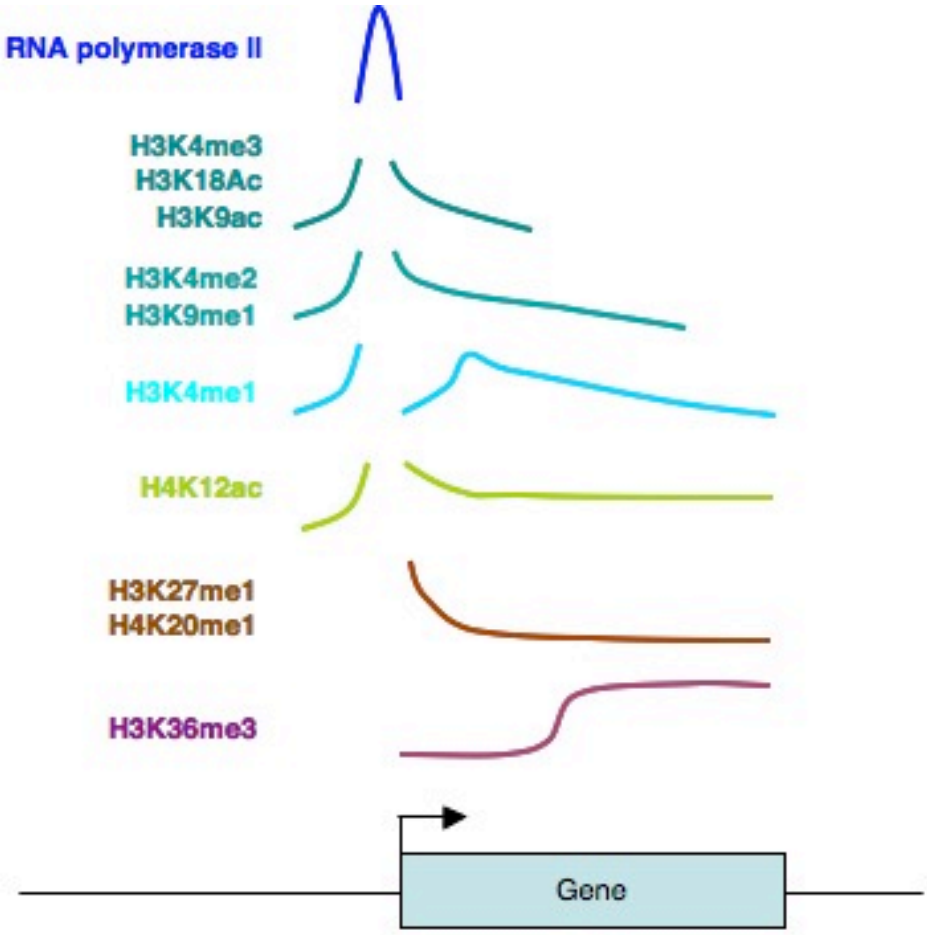


Figure 1.5. The typical pattern of histone modifications over an active human gene. The schematic shows typical distribution patterns as determined by ChIP:chip experiments. The gap over the transcriptional start site is a result of nucleosome depletion in that area. Adapted from Rando and Cheung 2009

regulated induction of gene expression has been called into question by the observed stochastic initiation of transcription. Indeed, such stochastic initiation has been shown to be of biological relevance. In erythroid cells, for example, stochastic gene expression is important in gene activation as initial stochastic gene expression patterns from alpha versus beta globin are clonally inherited (de Krom et al, 2002). More recently, transcription was shown to occur in bursts from a fluorescent reporter (Raj et al, 2006). Thus a gene that is considered active may only be periodically so. This “noise” can contribute significantly to the non-genetic heterogeneity of a cell population. Insertion of tandem copies of the transgene verses genome-wide insertion of multiple copies indicates that genes are under control of local, not global mechanisms. This is however, consistent with local chromatin structure playing a role in these transgene’s activation, keeping chromatin as a central candidate in gene control. In addition, measurements of stochastic wrapping and unwrapping of nucleosomal DNA could account for the access of proteins to the protein DNA-binding sequences on a biological timescale (Li et al, 2005).

1.5 Embryonic stem cells, pluripotency and chromatin

1.5.1 Embryonic stem cell properties

Embryonic stem cells (ES cells) are a population of pluripotent cells derived from epiblast cells of the inner cell mass of a pre-implantation stage embryo (Smith, 2001) and are a very interesting model in which to study transcriptional control (Figure 1.6). Their pluripotency confers the ability to divide to produce both a similarly pluripotent ES daughter cell and a more differentiated daughter cell (Figure 1.6). In culture, mouse ES cells and, to some extent, human ES cells, can be encouraged to differentiate down specific cell lineages providing a valuable tool in which to study differentiation and its associated changes in gene expression (Murry & Keller, 2008). Indeed, it is hoped that in the future such differentiation systems may be able to

provide a source of adult cells appropriate for regenerative medicine in the treatment of degenerative diseases such as Parkinson's.

The potential of ES cells to differentiate into all cell lineages requires that in the undifferentiated state, many genes are poised, ready for induction and therefore puts these cells in a "primed" state, whereby an external "push" in the right direction may result in commitment of the cells (Spivakov & Fisher, 2007). This has been proposed to result in ES cells existing in heterogenous populations, that puts them at an advantage as it continually presents opportunities for lineage specification (Silva & Smith, 2008). The pluripotency of stem cells is maintained by a transcription factor network, which includes Oct4, Nanog and Sox2 (Chambers et al, 2003; Ivanova et al, 2006; Mitsui et al, 2003; Niwa et al, 2000). They act to repress genes whose expression would result in terminal differentiation and to maintain their own expression in a self-regulatory network. As a result, expression of these factors is decreased upon differentiation. Proof of the absolute requirement for some of these pluripotent transcription factors can be found in the rapidly expanding field of induced pluripotent cells (Welstead et al, 2008). These induced poluripotent stem cells (iPS) were first created in 2006 by transfecting both mouse embryonic or adult fibroblasts with Oct3/4, Sox2, c-Myc, and Klf4 (Takahashi & Yamanaka, 2006). Such cells display nearly all of the characteristics of pluripotency and are the subject of high hopes for the generation of patient specific tissue and disease models. Since their original discovery, there has been a drive to decrease the number of factors required to induce pluripotency, and for a method that does not require the permanent modification of the host genome (Kaji et al, 2009; Okita et al, 2008; Stadtfeld et al, 2008; Woltjen et al, 2009).

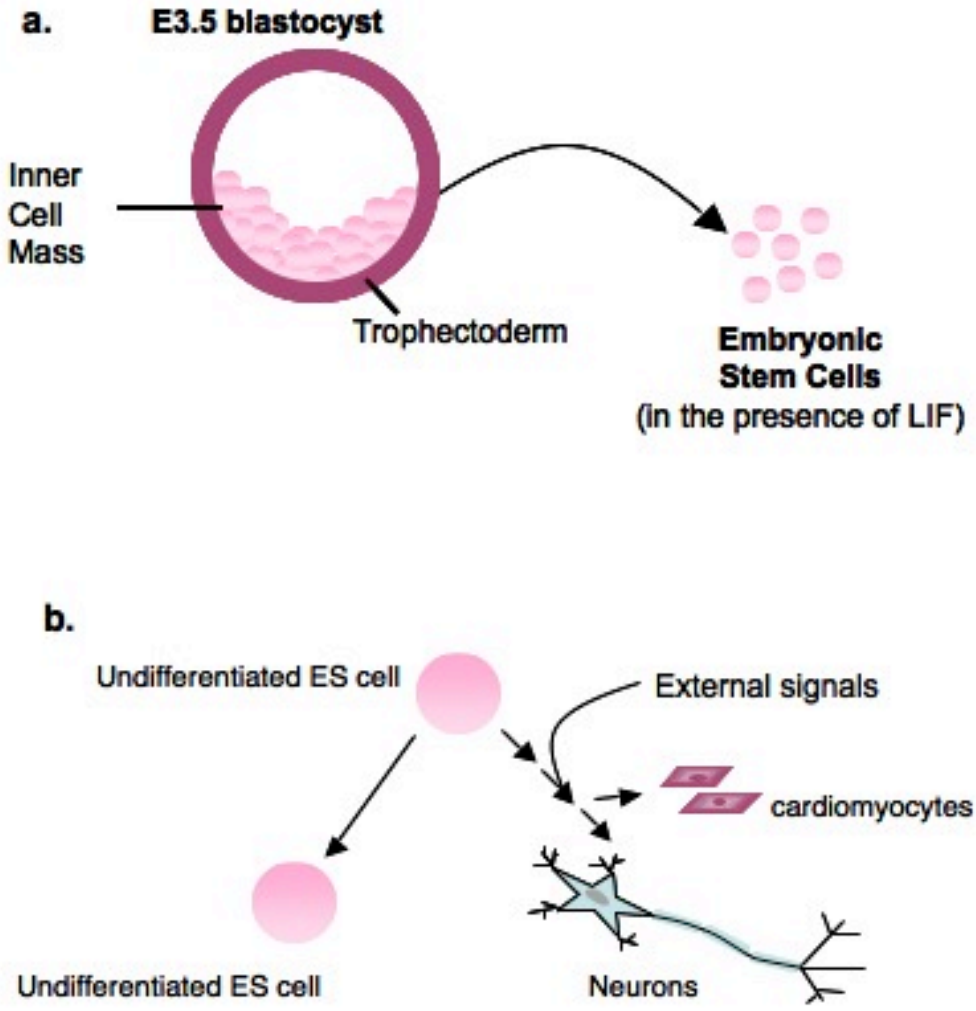


Figure 1.6. Embryonic stem cells are a population of pluripotent cells derived from the inner cell mass of the E3.5 blastocyst a. Embryonic stem cells are derived from a population of epiblast cells within the inner cell mass of a day 3.5 embryo.**b**. They may divide to self-renew or to develop into differentiated cells given the correct signals.

1.5.2 Embryonic stem cell chromatin

Embryonic stem cells are known to possess many unique chromatin features that are presumed to be a result of their being poised ready for differentiation (Spivakov & Fisher, 2007). The chromatin structure of embryonic stem cells is highly dynamic, with a high rate of exchange of the H1 histone protein and high levels of histone acetylation (Mikkelsen et al, 2007). In addition, early replication of non-expressed developmentally regulated transcription factors is a signature of ES cell chromatin (Azuara et al, 2006) such early replication timing was initially believed to correspond to actively transcribing loci but also corresponds well with histone acetylation (Vogelauer et al, 2002). There is also a marked difference in the higher order genome architecture in ES cell nuclei, inactive and active genes reside in the same compartment whereas in differentiated lymphocytes, inactive genes are located close to centromeric heterochromatin suggesting gene silencing must be maintained by different mechanisms in ES cell (Smale, 2003). However, detailed study of the neural-commitment gene *Mash1* showed that it located at the nuclear periphery in undifferentiated ES cells, and at the nuclear interior upon commitment into the neural lineage and concomitant initiation of expression. This suggests that location at the nuclear periphery is a potentially important mechanism of gene silencing in ES cells. The predominantly active chromatin features of ES cells have resulted in the proposal that the function of ES cell chromatin is to buffer against transcriptional "noise" rather than a more direct role in gene activation or repression (Chi & Bernstein, 2009). However, the increased efficiency of iPS cell generation by addition of valproic acid, (an HDAC inhibitor) or 5'-azacytidine (a DNA methyltransferase inhibitor) argues that chromatin alteration is required for the generation of pluripotent

cells and most likely plays an active role in the concomitant gene expression changes required for dedifferentiation (Huangfu et al, 2008).

In a landmark ChIP-chip study, Bernstein and colleagues confirmed a chromatin "signature" termed the bivalent domain at the promoters of a raft of genes in mouse ES cells (Bernstein et al, 2006). The study validated the signature already found at the promoters of early replicating genes in ES cells (Azuara et al, 2006). In the study, the histone modifications at the most highly conserved noncoding elements of mammalian cells were examined using antibodies to H3K4me3, H3K27me3 and H3K9ac. These highly conserved elements are located adjacent to transcription factors that are important in developmental regulation. The unique chromatin signature elucidated contained long stretches of H3K27me3, within which were contained peaks of H3K4me3, thus marks traditionally associated with active and repressive genes were found to be colocalised. This was confirmed by sequential ChIP. It was proposed that such domains serve to keep developmental genes silent in embryonic stem cells whilst keeping them poised for activation since the bivalent mark was associated with a low level of transcription in ES cells. Consistent with this is the observation that when the authors looked in differentiated cell types, many of these bivalent domains were resolved into either H3K4me3 or H3K27me3-only domains in a manner consistent with their subsequent transcription. That is, those genes that were actively transcribed resolved into a H3K4me3 containing domain, and those repressed resolved into a H3K27me3 containing domain. This signature has subsequently been confirmed in Human ES cells as well as mice, its presence has also been observed in differentiated cells at genes that require rapid activation in T cells (Roh et al, 2006).

Recently it has been argued that bivalent chromatin domains are merely a consequence of the CpG status of the promoters of developmental regulators. The many genome-wide studies have shown that CpG rich promoters in ES cells harbour the H3K4me3 mark constitutively and regardless of transcriptional status in the absence of DNA methylation, something that may actually protect such regulatory elements against DNA methylation (Guenther et al, 2007; Mikkelsen et al, 2007; Pan et al, 2007; Weber et al, 2007; Zhao et al, 2007). However, the repressive mark H3K27me3, is also shown to be preferentially targeted to CpG rich promoters, thus providing a sequence-based targeting approach of these two modifications that is not based upon biological function (Mohn & Schubeler, 2009).

The presence of H3K4me3 at non-transcribed genes is paradoxical regarding the evidence that points to a role of this modification in active transcription. However, such inactive bivalent genes were also shown to be enriched for another modification, ubiquitination at H2A, which was shown to be required for the binding of the “poised” ser 5 phosphorylated form of RNAPII. Furthermore, this poised form is surprisingly able to transcribe at low levels from these genes (Stock et al, 2007). Thus the H3K4me3 mark is again shown to localise with active genes. In a study in human ES cells transcription initiation was also shown to occur from the majority of “actively marked” genes, but that the polymerase does not proceed to elongation (Guenther et al, 2007). This would still be consistent with the role of the H3K4me3 mark in transcription initiation. This may represent another level at which transcriptional control can be achieved, with transcriptional processivity requiring additional factors. Note that many of these initiating transcripts were associated with genes that did not contain the elongation-associated mark H3K36me3 in the body of the gene. This abortive transcript production may represent a mechanism of poising

at these genes with transition into elongation being all that is required for active transcription. However, abortive transcription is not a phenomenon restricted to ES cells as a similar trait was observed in differentiated cells in the same study. Contrary, to this, in a different study in mouse ES cells, full transcripts were observed at a low level from the majority of ES cell genes (Efroni et al, 2008). Again, such a mechanism was proposed as being important to transcriptional “poising”. Therefore, hyper-transcription may play a role in maintenance of pluripotency.

1.6 The Homeotic genes

1.6.1 *Hox* genes as a model for transcriptional initiation

A group of genes proposed as being transcriptionally "poised" in embryonic stem cells are the *Hox* genes. The *Hox* genes are a group of much-studied genes required during development for defining the distinct morphological identities along the anterior-posterior axis in vertebrates. These genes were originally identified in *Drosophila* mutants that displayed segmental transformation known as homeotic transformations (Frischer et al, 1986). In these transformations, a correct structure was formed, but in the wrong place; for example, a leg may be formed in place of a wing. These mutations were shown to arise as a result of mutation of the homeotic, or *Hox* genes. The genes were first shown in Drosophila to be clustered on the genome in an orientation collinear to both their anterior to posterior segmental expression and their timing of expression (Lewis, 1978). Such a co-linear expression mechanism is present in all metameric animals (McGinnis & Krumlauf, 1992). In addition, these genes are known to contain a conserved motif; the homeobox domain, important for binding to DNA. In mammals, there are four *Hox* gene clusters located on four different chromosomes which, in mouse are termed A to D (Figure 1.7, (McGinnis & Krumlauf, 1992)). Between the cluster, paralogous genes have similar structures and equivalent expression patterns. That is, the gene termed "2" on all four clusters will have the same restricted expression as the gene termed "2" on the other clusters. However, between *Drosophila* and mammals, there must have been several rounds of duplication and deletion as equivalent genes are not always present (Figure 1.7). In mouse, Hox gene expression commences from the 3' end of the cluster at an early stage in development. The earliest expressed Hox genes are expressed in the mesoderm and more weakly in the epiblast of the posterior primitive streak in late streak stage embryos (E7.5)(Deschamps et al, 1999). Such an

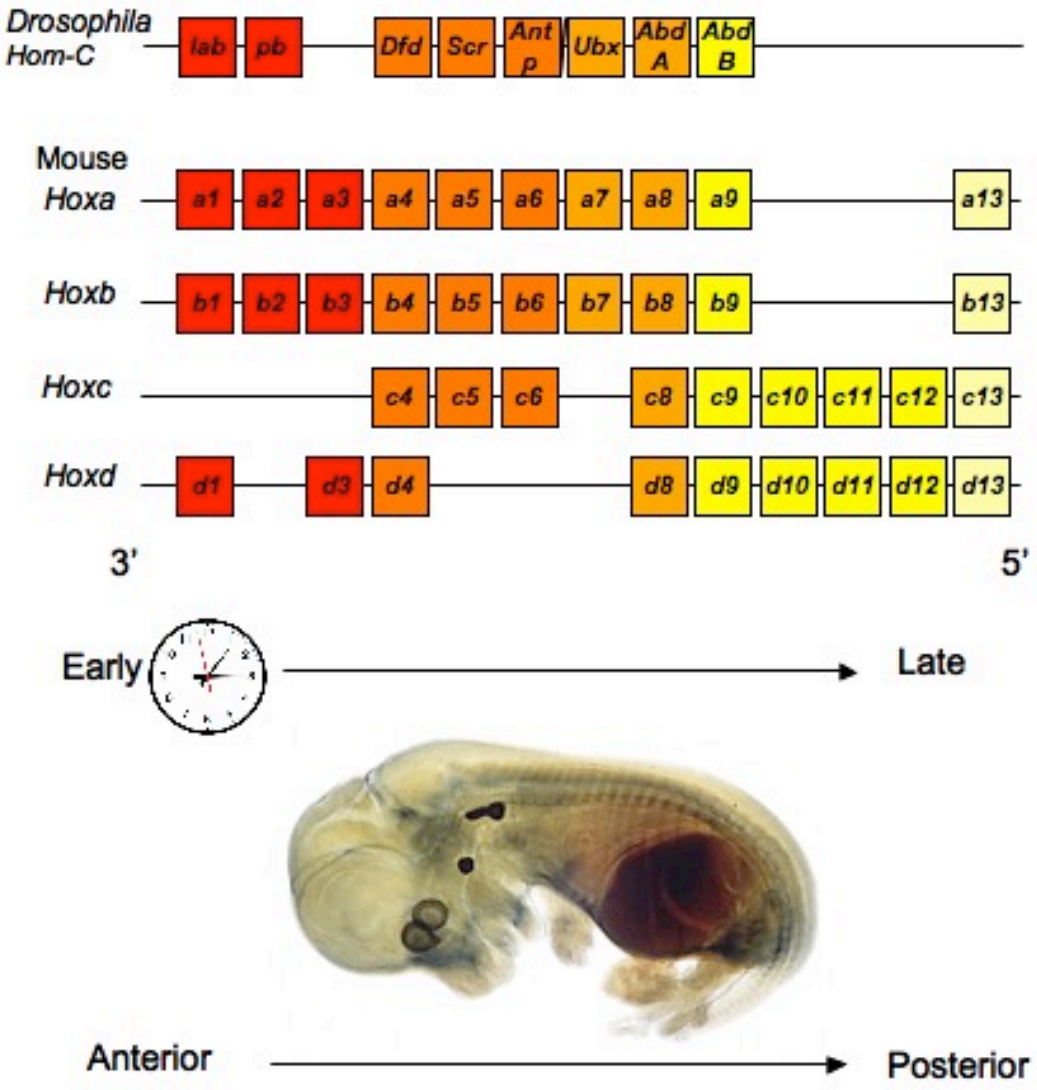


Figure 1.7. Hox genes are arranged in clusters and show co-linear timing of activation. Hox gene expression timing is sequential during development with 3' genes expressed first and 5' genes expressed last. 3' most genes are expressed in more anterior structures in the mouse and 5' most genes expressed in more posterior structures.

induction occurs concomitant with gastrulation and hence as the cells ingress, it results in widened field of induction of the Hox genes brought about by fibroblast growth factor and retinoic acid signalling, with the more 5' genes being expressed later and in the more posterior domain of the embryo. Through the use of the teratogen and homeotic mutation inducer retinoic acid in mice, it was argued that it is the combination of *Hox* genes within a segment that determines its identity resulting in the *Hox* code hypothesis (Kessel & Gruss, 1991). The need for tight transcriptional regulation of *Hox* genes is highly apparent. The *Hox* genes were thus proposed to undergo three phases of transcription; initiation, establishment and maintenance (Deschamps et al, 1999). Initially, it was proposed that the timing of *Hox* gene expression determined the ultimate pattern of gene expression in the "*Hox* clock" hypothesis since vertebrate development itself occurs in a rostral to caudal temporal progression (Kmita & Duboule, 2003).

1.6.2 A chromatin dependent mechanism of activation

Early genetic studies of *Hox* genes are consistent with a mechanism of activation that arises from progressive chromatin opening. Insertion of "anterior" transgenes *Hoxd11* and *d9* into a more posterior position in the *Hoxd* cluster resulted in the position-dependent timing of induction of these genes consistent with a repressive mechanism acting at the posterior end of the cluster (van der Hoeven et al, 1996). However, whilst early activation did result in the posteriorization of expression of the transgenes; the genes still maintained some of the spatial restriction of their endogenous copies. Thus, localisation of *Hox* gene expression does not correlate entirely with timing, though co-linearity does play some role in the initiation of gene expression patterns in early mice embryos. In addition to a higher (chromatin dependent) level of co-linear control, these experiments also displayed the

importance of *Hox* specific cis-acting elements since the transgene was able to induce spatially-dependent expression of the adjacent *Hox* gene (van der Hoeven et al, 1996). A further experiment, generating a fusion of the *Hox* d12 second exon and *Hox*d13 first exon resulted in a surprising *Hoxd*11-like expression of the fusion, further supporting a role for a progressive chromatin opening. This was proposed to be mediated by repressive protein binding (Kondo et al, 1998). Later experiments used a twofold approach to examine *Hox* gene regulation (Kondo & Duboule, 1999). First, sequential insertion of a *Hox*d9/lacZ transgene upstream of the cluster, revealed the location at which the repressive element no longer functions and second, deletion at this location confirmed its presence. Interestingly, deletion of the element resulted in aberrant gene expression only at the initiation phase of gene expression from the locus. Thus the higher-order control possibly mediated by chromatin only acts at this phase. Insertion of a *Hoxb1* transgene, (i.e. the most anterior gene) at *Hoxd13* the most posterior position of the *Hoxd* cluster did not result in complete “anteriorisation” of *hoxb1* (Kmita et al, 2000). It resulted in the surprise interplay between regulatory elements of the endogenous complex and integrated gene. The *Hoxb1* gene was activated in a manner consistent with both that of the endogenous gene and its newly inserted position and led to the activation of the adjacent *Hoxd13*. Thus the anterior transgene can recruit factors necessary for its own activation, resulting in a chromatin opening and active transcription of the adjacent gene. Therefore cluster control might be transcription-dependent, with co-linear activation relying upon an anterior entry point for activating factors that triggers the processing from a closed to open configuration across the cluster (Kmita et al, 2000).

Retinoic acid responsive elements (RAREs) are a control feature of many *Hox* genes. Notably, *Hoxb1* contains both 5' and 3' RAREs and a distal RARE (Marshall et al, 1994). *Hoxb1* and *Hoxb2*, unlike genes more 5' in the *HoxB* cluster are able to respond to exogenous retinoic acid addition by precocious activation at a stage termed pre-initiation (Roelen et al, 2002). This pre-initiation mimics *Hoxb1* patterns of gene expression in the developing embryo only 24 hours earlier. The 5'-most genes, do respond after their initial endogenous expression, indicating a different degree of accessibility of the genes at the 3' end. The RARE, therefore, may be able to open this 3' chromatin more easily through the recruitment of associated HATs (Bhattacharyya et al, 1997). However, a randomly integrated transgene containing all these RARE elements, was not able to respond to exogenous retinoic acid addition at the initiation phase, but only after the expression of the endogenous gene. This indicates a necessity for *Hoxb1* to be in its cluster for its complete regulation. Thus, the local chromatin opening at the 3' end of the cluster is a mechanism of gene control hierarchically higher than that conferred by the RAREs.

Finally, long-range enhancers have also been shown to have a role in *Hox* gene activation. These are distal sequences that occur outside of the cluster. For example, at *Hoxd*, a 5' "digit enhancer" has been shown to be important for a second round of co-linear expression of the 5' *Hoxd* genes in digits (Spitz et al, 2003). Such enhancers, by being positioned outside of the cluster bring an intrinsic asymmetry to control mechanisms. A strategy involving serial deletion and duplication analyses confirmed a two-wave expression mechanism from this cluster in forelimb and digit development. The first round relies on an early limb enhancer that is telomeric to the cluster and an element 5' to the cluster which spatially restricts the 5' genes. This enhancer acts against the repressive mechanisms seen at the 3' end of the cluster.

Direct evidence for a progressive chromatin opening mechanism came from studies at the *Hoxb* cluster (Chambeyron & Bickmore, 2004). In a landmark paper, the gradual induction of *Hox* genes upon addition of retinoic acid in culture systems was exploited (Simeone et al, 1990). Upon addition of retinoic acid, there was an increase in the activating modifications, H3K9ac and H3K4me2 at the promoters of *Hoxb1* and *Hoxb9*, despite the lack of induction of *Hoxb9* at this early stage. In addition, the authors observed a decondensation of the locus. This decondensation was not solely due to histone acetylation as induced increases in this mark by trichostain A, a histone deacetylase inhibitor, did not result in cluster opening. Further to this opening, *Hoxb1* was excluded from the local chromosome territory in a process parallel to the kinetics of its expression. *Hoxb9* also loops away at day 10, but never to the same extent as that of *Hoxb1*. Subsequently, it was confirmed that such a mechanism occurs *in vivo* in developing mice embryos (Chambeyron et al, 2005). Detailed analysis of *Hoxb* gene expression in distinct areas of the embryo suggested *in vivo* decondensation of the cluster and movement from out of the chromosome territory is differentially regulated at different developmental stages and no “priming” mechanism is visible.

Analysis of this mechanism at the *Hoxd* cluster, which makes complex use of different enhancers in its control was a key test of the role of chromatin opening in *Hox* gene activation. Here, it was displayed, however, that decondensation and relocation outside of the chromosome territory were not necessarily linked mechanisms of gene activation both *in vivo* and *in vitro* in ES cells (Morey et al, 2007). It was postulated that the difference in chromatin structure response to

activation might arise due to the different enhancers required for activation within different embryonic regions.

The link between induction of expression and chromatin decondensation was further dissected by the use of a transgene (Morey et al, 2008). As in earlier experiments, the *Hoxb1* transgene was placed 3' of the *Hoxd13* gene with this gene able to recapitulate some of the expression patterns of the endogenous gene (Kmita et al, 2000). It was able to ectopically reorganise the *Hoxd* cluster resulting in a large decondensation of the *Hox*d locus in the primitive streak (Morey et al, 2008). Furthermore, in rhombomere 4 where the *Hoxd* inserted transgene is silent, there was some movement of the transgene to the edge of the chromsome territory (CT) indicating a dominant effect of the transgene on the *Hoxd* cluster. Significantly, this provides evidence for movement to the edge of the CT being an event that occurs in advance of transcriptional activation.

Such cluster-wide mechanisms of co-linearity must not be ignored when considering the activation of individual *Hox* genes, however, these systems are good models in which to study the interplay of specific cis elements with domain-wide mechanisms. A detailed study of *Hoxd4* activation in neurally differentiating cells studied chromatin changes at this gene and at the 3' neural enhancer required for central nervous system (CNS) expression (Rastegar et al, 2004; Zhang et al, 1997). Other cis elements at this gene include a RARE and an autoregulatory element, both of which function in the model system used (Popperl & Featherstone, 1992; Popperl & Featherstone, 1993). Upon retinoic acid treatment, *Hoxd4* expression peaked between days 3 and 4 of treatment. This treatment induced changes in histone modifications; active H3 phospho-acetylation increases at the 3' enhancer preceded

those at the promoter and arose coordinate with the induction of *hoxd4*. However, active modifications at the enhancer persisted longer than at the promoter. In addition, retinoic acid induced the presence of H3 diacetylated at lysines 9 and 14 at the enhancer and a general increase of H4 acetylation across the locus concomitant with expression. Again, to further dissect the sequence of chromatin structure rearrangement and transcription, the histone modification changes were studied in vivo. Analysis of histone modifications in an embryonic tissue that precedes the activation of *Hoxd4* by a few hours showed active histone modifications had already been deposited at the neural enhancer, along with the activator CBP. Thus, here an important discovery has been made, that of the "priming" of the *Hox* locus before its transcription in the relevant tissue (Rastegar et al, 2004).

1.7 Heritability of gene expression patterns

1.7.1 Epigenetics

Thus far, in considering histone modifications, only those mechanisms directly involved in the initiation of *Hox* gene transcription have been considered. However, as an embryo divides and differentiates, the different lineages of the embryo must "remember" the gene expression changes that arise. Thus, the development of an organism and the execution of its genetic programme are intrinsically linked. For such a concept, the term epigenetics was coined by Conrad Waddington from the greek "epigenesis", a theory of development that proposed the embryo to be undifferentiated (Waddington, 1939). This idea was later revisited by Robin Holliday in his article "The inheritance of epigenetic defects" that discussed how some genetic defects and developmental changes could not be explained by Mendelian inheritance, but could be explained by DNA methylation (Holliday, 1987). This was a culmination of the idea gaining ground in the mid-1980s that not all inheritance was

based on DNA sequence. This resulted in Holliday's proposal of definitions for epigenetics in 1994 (Holliday, 1994). For the purpose of this thesis epigenetics will be defined as the nuclear inheritance of gene expression patterns that is not based on changes in DNA sequence. There is now little doubt that DNA methylation plays a strong part in epigenetics and epigenetic inheritance. Histone modifications too have been proposed to play a role in this epigenetic inheritance, however, their heritability has not been definitely proven, though their potential to maintain gene expression programmes in memory is highly attractive. Thus, the epigenetic code hypothesis was proposed (Turner, 2000; Turner, 2007). This hypothesis draws on the *predictive* role that histone modifications must play to any transcriptional outcome within a defined developmental stage. Thus the modifications are proposed to dictate not the actual gene expression programme, but the *potential* for that programme, such that a locus is predisposed to activation, or, indeed repression. In fact, the final example at *Hoxd* discussed in the previous section exemplifies this.

1.7.2 Epigenetic Maintenance of *Hox* gene expression, Polycomb and Trithorax proteins

The maintenance of *Hox* gene expression patterns in flies and mice has largely been attributed to the polycomb and trithorax proteins. Polycomb genes were originally identified in *Drosophila* as genetic repressors of homeobox proteins (Lewis, 1978). Trithorax genes were identified in the same study as mutations that suppressed those at the polycomb complexes. In mammals, a double knock-out confirmed the two complexes are antagonistic regulators of *Hox* gene expression (Hanson et al, 1999). Polycomb group proteins are known to exist in at least three separate complexes of which PRC1 and PRC2 are the best characterised, with a third complex, the Pho complex also being relatively well understood. The polycomb

Polycomb

Drosophila protein	Complex	Protein domains	Biochemical activity	Mouse protein homologues	Human protein Homologues
PC	PRC1	Chromodomain	H3K27me3 binding	NPCD, M33 (CBX2), CBX4, CBX6, CBX7, CBX8	HPC1, HPC2, HPC3
PH	PRC1	SAM	?	PHC1, PHC2, PHC3	HPH1, HPH2, HPH3
PSC	PRC1	RING	Cofactor for SCE	BMI1, MEL18	BMI1
SCE(RING)	PRC1?	RING	E3 ubiqutin ligase at H2AK119	RING1A, RING1B	RING1A
SCM	PRC2	SAM, MBT ZN finger	?	SCMH1, SCML2	SCMH1, SCMH2
E(Z)	PRC2	SET	methylation of H3K9, H3K27	EZH2, EZH1	EZH2
ESC	PRC2	WD40	Cofactor for E(Z)	EED	EED
ESCL	PRC2	WD40	Cofactor for E(Z)	EED	EED
Su(Z)12	PRC2	Zn-finher	?	SUZ12	SUZ12
PCL	PRC2	PHD, Tudor	?	PHF19, MTF (M96)	PHF1

Trithorax

Drosophila protein	Complex	Protein domains	Biochemical activity	Mouse protein homologues	Human protein Homologues
TRX	TAC1	PHD,SET	Methylation of H3K4	WBP7, MLL1	MLL1-3, WDR5, ASH2L, RbBP5, CFP1
ASH1	?	SET, PHD, BAH	Methylation of H3K4, H3K9, H4K20	ASH1L	
ASH2	?	PHD, SPRY	?	ASH2L	
BRM	SWI/SNF	SNF2, HELICc, Bromodomain	ATP-depependent nucleosome sliding	SMARCA4	BRM
MOR	SWI/SNF	SWIRM, SAINT	Cofactor for BRM	SMARCC1, SMARCC2	BAF170
OSA	SWI/SNF	BRIGHT	?	ARID1B	BAF250

Table 1.1 The polycomb and trithorax proteins in *Drosophila*, Mouse and human. Adapted from Schuettengruber *et al*

repressive complex (PRC2) complex in *Dropsophila* is made of a core of E(z) (Enhancer of Zeste), Esc (Extra sex combs), Su(z)12 (Suppressor of zeste 12) and Nurf (nucleosomal remodelling factor). The PRC1 complex has at its core Pc (Polycomb), Ph (polyhomeotic), Posterior Sex combs (Psc) and dRING. Mammalian homologues have been found for all of these proteins (Table 1.1). However, in mammals copies are often multiplied as a result of genome duplication events (Table 1.1). In flies, the PRC1 complex is found to associate with core components of the transcriptional machinery, though this is not the case in mammals (Saurin et al, 2001). It is believed that this may play a direct or indirect part in the silencing role of these proteins through recruitment of the PRC1 complex. Genetic studies in mice revealed components of the PRC1 complex to be important in anterior-posterior patterning and determining correct spatial expression patterns of different subsets of *Hox* genes (Akasaka et al, 1996; Akasaka et al, 2001; Core et al, 1997; del Mar Lorente et al, 2000; van der Lugt et al, 1996). Furthermore mouse knockouts of polycomb proteins *eed, suz12* and *Rnf2* show defects in anterior-posterior patterning during gastrulation, suggesting a role in Hox gene control (Faust et al, 1998; Pasini et al, 2004; Voncken et al, 2003). Interestingly, a double knock-out of the PRC1 components, *bmi-1* and *Rnf110* (both PSC equivalents) showed a window post 8.5 days where these proteins are required for maintenance of the spatial programme of *Hox* gene expression. That is the *Hox* gene expression patterns are established correctly in this double knock out, but deteriorate after day 8.5. This implicates a role for these proteins in the memory of gene expression (Akasaka et al, 2001). Furthermore, a conditional *Rnf2* knock-out indicates a role for this protein in the maintenance of *Hox* gene repression (Fujimura et al, 2006). Trithorax proteins are involved in the converse role of maintenance of active *Hox* gene expression programmes is shown by mouse *mll* knock-out studies (Yu et al, 1998).

1.7.3 The mechanism of polycomb silencing

Initial hypotheses in *Drosphila* suggested that polycomb proteins bind to Polycomb repressive elements (PREs) elements and then spread along the chromosome in a manner analogous to the repressor protein HP1 in heterochromatin, causing silencing by blanketing the chromosome (Paro, 1990). However, it has become apparent that such a smothering mechanism is unlikely given that transcription factors may bind to silenced promoters in *Drosophila* cells (Breiling et al, 2001). Also, genome-wide studies of the localisation of polycomb proteins at *Hox* promoters are inconsistent with the blanketing of repressed genes (Schwartz et al, 2006).

The elucidation that components of the PRC2 complex are able to methylate H3 at lysine 27 in a manner that is conserved across mammals and flies suggested a simplex "indexing" mechanism by which polycomb might achieve its silencing role (Cao et al, 2002; Czermin et al, 2002; Kuzmichev et al, 2002). *Suz12* depleted HeLa cells show derepression of a subset of *Hox* genes concomitant with the loss of H3K27me3 at these genes (Cao & Zhang, 2004). Furthermore, genome-wide studies consistently show across all species that the binding of polycomb group proteins correlates with H3K27me3 (Boyer et al, 2006; Bracken et al, 2006; Lee et al, 2006; Negre et al, 2006; Squazzo et al, 2006). Such studies indicate that despite the large conservation seen between mouse, human and *Drosophila* proteins, and the significant conservation of polycomb group targets, the binding pattern of these proteins varies significantly between mammals and *Drosophila*. Mouse and human proteins bind throughout the large domains of H3K27me3 at *Hox* genes and specifically at promoters of other genes, whereas *Drosophila* proteins bind at punctate sites (presumably the PREs). This seems to suggest a difference in the

mechanism by which polycomb-repressed genes are regulated in flies and mammals, and care must be taken when comparing between species.

A definitive mechanism for H3K27me3-maintained gene-silencing has not been elucidated, however, the PC subunit of the PRC1 complex is able to bind to H3K27me3 and to some extent H3K9me3 using its chromodomain, thus a sequential order of recruitment whereby PRC2 generates a mark promoting the binding of PRC1 has been proposed (Fischle et al, 2003b; Min et al, 2003). A further role for histone modifications in polycomb mediated silencing has been inferred from the capacity of the RING2/Ring1B subunit of PRC1 to monubiquitinate at H2A at lysine 119 both in *Drosophila* and mammals (Wang et al, 2004; Zhu et al, 2005). Ubiquitination at this residue was shown to be downstream of H3K27 methylation in a mouse *Bmi* (PRC1 RING subunit) knock-out (Cao et al, 2005). The loss of H2A ubiquitination correlated with an increase in expression of *Hoxc13* although the level of H3K27me3 at this gene did not change. Similarly, a *Suz12* knock-down, resulted in no H3K27me3 at the *Hoxc13* promoter preventing the binding of *bmi1* or *ring1* and resulting in the absence of H2A ubiquitination. This is consistent with PRC2 generating a binding site for PRC1 in the H3K27me3 mark (Figure 1.8). Recently, a direct role for H2A ubiquitination in arresting transcriptional elongation has been proposed (Reinberg & Sims, 2006). This study did not examine polycomb-related H2A ubiquitination, but H2A mono-ubiquitination catalysed by the N-Cor recruited ligase 2A-HUB/hRUL138 at a subset of genes in macrophages. This modification prevented the recruitment of FACT (facilitates chromatin transcription), a core transcription factor that would normally act to facilitate transcription by binding and displacing the H2A/H2B dimer from core nucleosomes (Reinberg & Sims, 2006) .

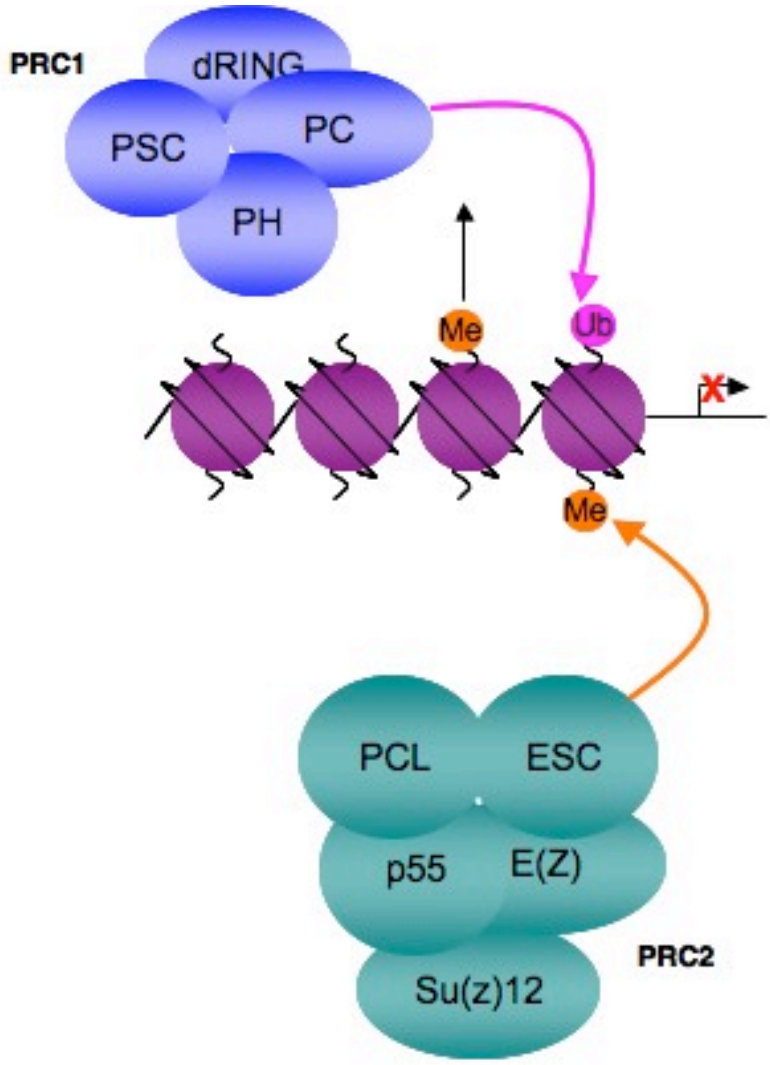


Figure 1.8. The biochemical activities of the *Drosophila* Polycomb group complexes PRC1 and PRC2. The PRC2 complex is shown in green. Its E(Z) protein is able to trimethylate lysine 27 of histone H3. This, in turn recruits PRC1, here shown in blue. PRC1 is recruited by the chromodomain of PC and is able to mono-ubitinylate histone H2A at lysine 119. This is believed to be the mark that prevents transcription from the surrounding loci. The *Drosophila* proteins are included as examples, though these activities are conserved across mouse and human.

To further delineate the relationship between polycomb-mediated *Hox* gene repression and histone modifications, a detailed study in the maintenance of *Hox* gene transcriptional programmes has been carried out at the *Hoxb8* locus (Fujimura et al, 2006). The PRC1 complex was previously shown to play a role in *Hoxb8* posterior restriction. The association of the full PRC1 complex with *Hoxb8* was confirmed in non-expressing tissues in the E12.5 embryo, however, a partial complex lacking the Rnf2 (Ring1B) component was also present in expressing regions. The role of this partial complex is not clear. H3K9ac and H3K4me3 levels were also correlated with active expression in the embryonic tissue whereas H3K27me3 was more abundant at the *Hoxb8* locus in the transcriptionally repressed tissue. Thus, active and repressive marks appear to play a role in the maintenance of expression of this gene. The role of polycomb group proteins in transcriptional repression has been challenged by evidence for PRC1 complexes playing a role in gene activation. A study in which a *Hox* reporter transgene was randomly inserted into the genome in the mouse polycomb null mutants, *Mel18* (PSC equivalent) and *Rae28* (PH equivalent), resulted in the lack of normal activation of this transgene (d Graaff et al, 2003). Intriguingly, in this study, the inactivated transgene displayed a high degree of DNA methylation at its promoter, therefore the role of polycomb proteins may be to protect from DNA methylation. Further to this, a drop in expression in *Hoxb8* occurs at around 9.5 dpc in Bmi/Rnf110 and phc1/phc2 double mutants (Akasaka et al, 2001; d Graaff et al, 2003; Isono et al, 2005). This coordinated with a drop in the level of H3K9 acetylation in those regions, whereas in wild type cells, *Hoxb8* expression and H3K9acetylation is maintained (Fujimura et al, 2006).

1.7.4 The mechanism of Trithorax regulated gene expression

Trithorax proteins maintain boundaries of active *Hox* gene expression in a manner that is also reliant upon the alteration of chromatin state. Several trithorax complexes have histone H3K4 methylation activity (Byrd & Shearn, 2003; Dou et al, 2005; Nakamura et al, 2002; Wysocka et al, 2005). The close link of this mark with transcription initiation has been discussed elsewhere. The *Hox* gene spatial maintenance defect of *mll* knock-out mice suggests a special relationship of *Hox* genes with trithorax related proteins (Yu et al, 1995). Further to this, the knock-down of components of the *mll* complex, rather than the enzyme itself also results in the lack of maintenance of expression of *Hoxc8* and *Hox*a9, coordinate with a loss in H3K4 methylation (Wysocka et al, 2005). ChIP-chip arrays across MEFs in differentiated cells i.e. cells that would display maintenance patterns, has revealed a peculiar pattern of H3K4me3. This mark appears to blanket across these domains (Bernstein et al, 2005). Consistent with this, the large scale mapping of MLL binding sites in a human lymphoma cell line reveals that whilst binding specifically at the promoters of most of its targets, at the *Hoxa* cluster, it displays a more global pattern of binding; binding both up and downstream of the transcriptional start site (Guenther et al, 2005). Consistent with observations in mouse knock-out studies showing MLL's preference for regulating 5' *Hox* genes, this unique binding pattern occurs at *Hox* genes A7 through to 10, though does also encompass *Hoxa1*.

Studies at individual genes have confirmed trithorax proteins are important in maintaining *Hox* gene active expression. For example, at *Hoxc8* MLL binds specifically at the promoter of this gene, not at 5' or 3' regulatory enhancer sequences (Milne et al, 2002). *mll* knock-out in MEFs resulted in a decrease of *Hoxc8* expression, but not interestingly, *hoxa1* since MLL primarily regulates 5' *Hox*

genes (Hanson et al, 1999). Activation of *Hoxc8* by *mll* knock-in in these knock-out cells was accompanied by histone acetylation at the 5' and 3' enhancer and H3K4 methylation at the promoter reliant upon the MLL set domain. Interestingly, *Hox*c8 experienced DNA hypermethylation at the promoter in mll-/- lines which could not be overcome in *Mll* knock-in lines, again showing the H3K4me3 may protect against DNA methylation. Another mammalian homologue of the trithorax proteins, ASH1L was shown to methylate H3K4me3 *in vitro* and was impaired by methylated H3K9, thus providing a conserved mechanism for trithorax-mediated *Hox* gene expression maintenance (Gregory et al, 2007).

It is interesting to note that in the above mouse genetic studies that in polycomb and trithorax protein knock-outs, it is only a subset of genes that are affected, revealing that different complexes are used at specific subsets of *Hox* genes, no doubt conferring tissue-specific regulation. In mammals, for example, the PRC2 complex is not always made of the same components and is believed to confer tissue-specific maintenance of *Hox* gene expression programs in such a manner. For example, the mammalian PRC2 complex may contain hPHF1, a homologue of *Drosophila* PHF (Cao et al, 2008). The knockdown of the mouse protein in cell culture resulted in *Hox* gene specific alterations in gene expression

1.7.5 A dynamic mechanism of gene regulation; the lysine demethylases and pluripotency

Contrary to what this gene expression *maintenance* role of polycomb and trithorax proteins suggests, the two modifications H3K4me3 and H3K27me3 are dynamic and may be removed by demethylases (Shi & Whetstine, 2007). H3K27me3/me2 may be demethylated by the mammalian proteins UTX and JMJD3, two demethylases that

associate with the promoters of *Hox* genes in a cell-type specific manner correlating with transcriptional activity (Agger et al, 2007; Lan et al, 2007). A functional role for these demethylases in the activation of *Hox* genes was demonstrated in the retinoic acid-induced transcription of *Hox* genes in embryonic carcinoma (EC) cells. Upon addition of retinoic acid, *Hoxb1* expression increased accompanied by an increased presence of UTX, and a decrease in H3K27me3 (Agger et al, 2007). This demethylase associated with the MLL2 complex and the further use of the retinoic acid differentiation system revealed an ordered recruitment in *Hox* gene activation whereby the MLL2 complex is first recruited to promoters followed by the UTX demethylase (Lee et al, 2007b). Also, knock-down studies displayed a role for UTX in regulating PRC1 binding and H2A ubiquitination at its regulated promoters. Note that the authors found this mechanism operating at a specific subset of *Hox* genes.

The expanding family of demethylases also includes those that demethlyate H3K4me3 (Shi et al, 2004). Specific to *Hox* gene regulation, RBP2 is a histone lysine demethylase that in retinoic acid-induced differentiating ES cells, was shown to dissociate from a subset of *Hoxa* genes, coordinate with their upregulation and increase in H3K4me3. (Christensen et al, 2007). In fact, the PRC2 complex is able to recruit this demethylase to polycomb targets, indicating a potentially rapid means of gene repression (Pasini et al, 2008).

Such rapid induction of gene expression may have a role in differentiation. Genome-wide profiling of a variety of different cell types has indicated that, in vertebrates significant proportion of polycomb group targets are transcription factors important in regulating developmental pathway. Their regulation must, by functional inference, be dynamic. Indeed, polycomb group proteins have been shown to play a role in

maintaining pluripotency, which has been proposed to be a direct result of the repression of transcription factors that determine differentiated cell states (Boyer et al, 2006; Lee et al, 2006). Thus, at least at these earlier stages in mammalian development, polycomb gene repression must be dynamic. Interestingly, a study that followed the association of polycomb proteins across genes in the differentiation of neuronal precursors showed that that genes that were activated in differentiation showed a loss of PcG binding, but those that were active in precursors still were associated with PcG proteins, and upon their repression, association only increased slightly (Bracken et al, 2006).

1.7.6 The role of active transcription in memory

Despite the majority of the genome being non-coding, much of it is transcribed (Carninci et al, 2005). In *Drosophila,* active transcription through PREs has been proposed as a mechanism whereby the expression of the related gene is maintained. In this model, silencing at PRE controlled genes occurs by default, and transcription from the element prevents silencing and maintains the gene in the “on” state (Schmitt et al, 2005). The molecular basis of this activation was elucidated with the discovery that the recruitment of the trithorax protein ASH1 to the UBX PRE/TRE relies on its interaction with non-coding transcripts from these PREs which directly recruit this protein by remaining associated with the PRE (Sanchez-Elsner et al, 2006). Memory of transcription was then proposed to persist through the cell cycle by the retention of these transcripts at the PRE in an epigenetic “bookmark” that is then able to re-recruit ASH1 and its deposition of H3K4me3. Alternatively, the H3K4me3 mark may persist though the cell cycle and result in the re-establishment of transcription of these non-coding RNAs, full recruitment of ASH1 and establishment of transcription patterns (Schmitt & Paro, 2006).

Mammals, too display significant transcription from non-coding domains. Transcription occurs, in fact, right across the *Hox* clusters in mouse and human (Sessa et al, 2007). This study focused on opposite strand synthesis of such noncoding RNAs. Interestingly, this transcription correlates with the co-linearity of transcription from the *HoxA* cluster upon retinoic acid induction in teratocarcinoma cells and appeared to be related to the activation of transcription from the genes themselves, particularly as such non-coding transcription persists. However, the role of these intergenic transcripts is unclear as they are not conserved and when transfected into teratocarcinmoa cells are unable to activate *Hoxa* genes in trans.

An impressive study in humans profiled the transcription from *Hox* clusters in fibroblasts derived from multiple differentiated tissues (Rinn et al, 2007). This study identified a long noncoding RNA transcribed antisense to the *Hoxc* locus termed HOTAIR, which interacts with, and is required for PRC2 occupancy and histone H3 lysine 27 trimethylation at *Hoxd*. Strikingly, its depletion resulted in the activation of the *Hoxd* cluster. Similarly, another noncoding RNA, the spliced long nc RNA, *Hox*b5/6 is expressed from a 15kb region in the *Hoxb* cluster that encompasses the *Hoxb5* and *b6* genes (Dinger et al, 2008). Its expression profile correlates with that of these genes, both in EB differentiation systems, and in developing mice embryos. It can also associate with MLL, leading to the speculation that it may help to recruit or maintain this protein at these genes. Thus, recruitment of protein complexes by RNAs appear to be important in regulating *Hox* gene expression.

1.8 Aims

In surveying histone modifications, it is clear that they are interlinked with transcription, be it as a consequence or as a cause. It is also clear that defined combinations of modifications are associated with defined outcomes of transcription at specific loci. Therefore this study aims to further delineate the relationship between histone modifications and gene expression at a defined set of inducible genes, the *Hoxb* cluster. Using this model, it is possible to ask the following questions.

- Are histone modifications combinatorial?
- Are relative levels of histone modifications predictive of transcriptional outcomes?
- Can a change in histone modification alone induce gene transcription? Or, are histone modification changes transcriptionally dependent.

Histone modifications have also been postulated as playing a role in the “memory” of gene activity therefore maintaining gene transcription programmes throughout development. The maintenance of expression at the *Hox* gene cluster is imposed by the trithorax and polycomb group proteins. Therefore, this model system is ideal for the study of memory associated histone marks as the histone modifications that are associated with this memory process within the cluster are clearly defined. Therefore, the following questions may be asked.

- Are externally induced histone modification changes “remembered” by the genes of the *Hoxb* cluster in a maintenance of elevated levels
- Are externally induced histone modification changes remembered by the genes of the *Hoxb* cluster by an alteration in gene expression programmes.

2 Materials and Methods

2.1 Cultured Cells

2.1.1 Mouse ES cells

Male CCE/R mouse embryonic stem cells were provided by G. Anderson. The cells were grown at 37°C in Dulbecco's modified Eagle's medium (DMEM)(Gibco) supplemented with 15 % Foetal calf serum (Gibco), 1% penicillin/streptomycin (Gibco), 1% L-glutamine (Gibco), non-essential amino acids (Gibco), 0.25% 2-mercaptoentanol (Gibco) supplemented with 1U/μl leukaemia inhibitory factor (LIF/ESGRO, Millipore). Cells were cultured in T25/T75 (Sarstedt) tissue culture flasks coated with 0.1% gelatin. Cells were harvested by trypsinisation (1 x Trypsin EDTA, Gibco) at 37°C and washed three times in ice-cold PBS containing 5mM sodium butyrate and centrifuged at 280g (MSE chilspin) for 5 minutes. For sodium valproate treatment, exponentially growing cells were treated with 1M sodium valproate (Sigma) resupended in water to a final concentration of 1 mM in tissue culture medium. For wash-out, the cells were trysinised, washed out of the culture medium, washed twice by centrifugation in valproate-free medium and replated as normal.

2.1.2 Differentiation of CCE/R cells

CCE/R cells were induced to differentiate by following the method developed by (Chambeyron & Bickmore, 2004). Briefly, cells were replated onto non-adherant plastic in the absence of LIF. Two days after removal of LIF, cells were differentiated along the neuroectoderm lineage by addition of 0.1μM retinoic acid. After 7 days of differentiation in this manner, cells were replated onto adherent plastic allowing the embryoid bodies to settle and continue differentiation.

2.1.3 Mouse embryonic fibroblasts

Primary mouse embryonic fibroblasts were used for ChIP experiments and as feeders for 3G4 cells. Mouse embryonic fibroblasts were derived from E12 BALB/c mice. Heads were first removed, along with any internal organs. A single cell suspension was created by first cutting the tissue with a scalpel and trypinising for 10 minutes. Any remaining large lumps of tissue were allowed to settle before removing the supernatant and plating the remaining single cells in a T25 tissue culture flask. Primary MEFs were then cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10 % Foetal calf serum (Gibco), 1 % penicillin/streptomycin (Gibco), 1% L-glutamine (Gibco).

MEFs for use as feeders were irradiated in order to arrest growth. Cells were first trysinised and resuspended at a concentration of 1 x 10^7 cells/ml in culture medium and then irradiated using a Caesium 137 source for 10 minutes. Irradiated MEFs were then plated at 1 x 10^6 / T25 or 3 x 10^6 / T75.

2.1.4 3G4 Cells

Wild type and knock-out 3G4 GCN5 -/- cells were a gift from Sharon Dent (Lin et al, 2007) grown at 37°C in Dulbecco's modified Eagle's medium (DMEM) supplemented with 15 % Foetal calf serum (Gibco), 1% penicillin/streptomycin (Gibco), 1% L-glutamine (Gibco), non-essential amino acids, 0.25% (Gibco) 2-mercaptoentanol supplemented (Gibco) with 1U/μl leukaemia inhibitory factor (LIF/ESGRO, Millipore). Cells were cultured in T25/T75 tissue culture flaks (Sarstedt) coated with 0.1% gelatin. Wild type cells were grown on irradiated MEFs as described above whereas knock-out cells were grown without feeders. Wild type cells were harvested by trysinisation and titrated from the MEF feeder layer by allowing the MEFs to settle first on tissue culture flasks for 20 minutes. Suspended ES cells were pipetted-off

and the process repeated three times. Cells were then washed 3 times in ice-cold PBS 5mM sodium butyrate. 3G4 GCN5 -/- cells were harvested as described for CCERs.

2.1.5 *Drosophila* SL2 cells

Drosophila SL2 cells were grown anaerobically at 26°C in Schneider's medium (Gibco) supplemented with 8% foetal calf serum (Gibco) and antibiotics (Penicillin/Streptomycin).

2.1.6 Alkaline Phosphatase activity assay

ES cells were rinsed once in ice cold PBS and fixed in ice cold NFB (3.7% paraformaldehyde, 112mM NaH_2PO_4, 30mM $NaH_2PO_4H_2O$) for 15 minutes. Cells were then incubated for 45 minutes using 1mg/ml Napthol-AS-MX-phosphate (Sigma) dissolved in 0.1M Tris HCl pH 8.3, 4% DMF as a substrate with 0.6mg/ml Fast Red Violet LB salt (Sigma) as a coupler. Pluripotent positive colonies appear red.

2.2 Flow cytometric analysis

MEFs were harvested by trypsinisation and resuspended in DMEM/10%FCS at a concentration of 10^7cells/ml. Cells were permeabilised in 0.1% NP40 DMEM and stained using 10mg/ml propidium iodide followed by immediate analysis. Flow cytometry analysis was carried out using a Beckman Coulter Epics XL bench-top flow cytometer (Beckman Coulter), data was analysed using the Windows Multiple Document Interface Flow Cytometry Application program (WinMDI Version 2.8).

2.2.1 Cell cycle fractionation by flow cytometry.

Vybrant Dyecycle stained MEFs were harvested by trypsinisation, and resuspended at a concentration of 10^6 cells/ ml in DMEM/10%FCS/Vybrant DyeCycle green in a volume of 1ml and labelled at 37^0C for 30 minutes. Cells were then passed through a nylon filter and transferred into polypropylene FACS tubes (Becton Dickinson) in the labelling medium. High speed sorting was carried out on a MoFlo high speed sorter (Dako cytomations, Gloustrop, Denmark). Cells were sorted into DMEM 10% FCS in polypropylene FACS tubes.

2.2.2 Chromatin Isolation from ES cells

CCE/R cells were grown to their exponential phase and labelled overnight with 3H-thymidine at 0.5μCi/ml (Amersham). After harvesting, cells were resuspended to a concentration of 2 x 10^7cells/ml in TBS (15mM NaCl, 0.15M NaCl , 10mM Tris-HCl (pH 7.5), 3mM $CaCl_2$, 2mM $MgCl_2$, 5mM Na butyrate) and an equal volume of 1%Tween40/TBS added. Cells were stirred on ice for 1 hour and homogenised using a Dounce all-glass homegnizer with a tight pestle in order to release nuclei (1 stroke/ 10^7 cells). This process was monitored by microscopy and stopped when 70-80% nuclei had been attained. Cell lysates were centrifuged at 800 x g for 20minutes at 4ºC (MSE Chilspin) and the pellet resupended in TBS/5% sucrose and spun at 800 x g for 20 minutes. The nuclei were resuspended in 5ml digestion buffer (0.32M Sucrose, 50mM Tris/HCl (pH 7.5), 4mM $MgCl_2$, 1mM $CaCl_2$, 0.1mM PMSF, 5mM Na butyrate) and the $A_{260/280}$ determined to get an estimate of the concentration. The concentration of the sample was readjusted to 0.5mg/ml. 1ml aliquots of the nuclei suspension were digested with 35-50U micrococcal nuclease (Pharmacia/Sigma) at 37ºC for 5 minutes. The reaction was stopped by adding 0.5mM EDTA to a final concentration of 15mM and chilled on ice for 5 minutes. After micrococcal nuclease

digestion, aliquots were centrifuged at 1,800g for 10 minutes (MSE microcentaur) and the first supernatant (S1) removed and kept at 4°C. The pellet was resuspended in 1ml lysis buffer (1mM Tris/HCl (pH7.4), 0.2mM Na_2EDTA, 0.2mMPMSF, 5mM Na butyrate) and dialysed overnight at 4°C against 2 litres of equivalent buffer. Solubilised chromatin (S2) was recovered by spinning dialysed S1 samples at 1,800g for 10 minutes (MSE Chilspin) and the remaining insoluble pellet (P) resuspended in 250μl of lysis buffer.

All three chromatin fractions, S1, S2 and P were analysed by agarose gel electrophoresis (Figure 2.1). S1 and S2 fractions were pooled and used for chromatin immunoprecipitation (Figure 2.1).

2.2.3 Chromatin Isolation from mixed SL2 / ES cells

For the carrier chromatin immunoprecipitation (CChIP) procedure, *Drosophila* SL2 cells were pelleted and washed 3 x in ice cold PBS, 5 mM sodium butyrate. Cells were resuspended to 5 x 10^7 cells / ml and 1 ml aliquots were mixed with a small number (10^4) of FACS cell cycle sorted populations of MEFs. Cell mixtures were washed twice in NB buffer (15 mM Tris-HCL pH 7.4, 60 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$, 0.1 mM EGTA, 0.5 mM 2-mercaptoethanol, 0.1 mM PMSF) supplemented with 5 mM sodium butyrate. The final pellet was resuspended in 2 ml NB buffer, and mixed with an equal volume 1% Tween 40 in NB buffer and stirred for one hour on ice. Nuclei were released by homogenisation in a Dounce all glass homogeniser with

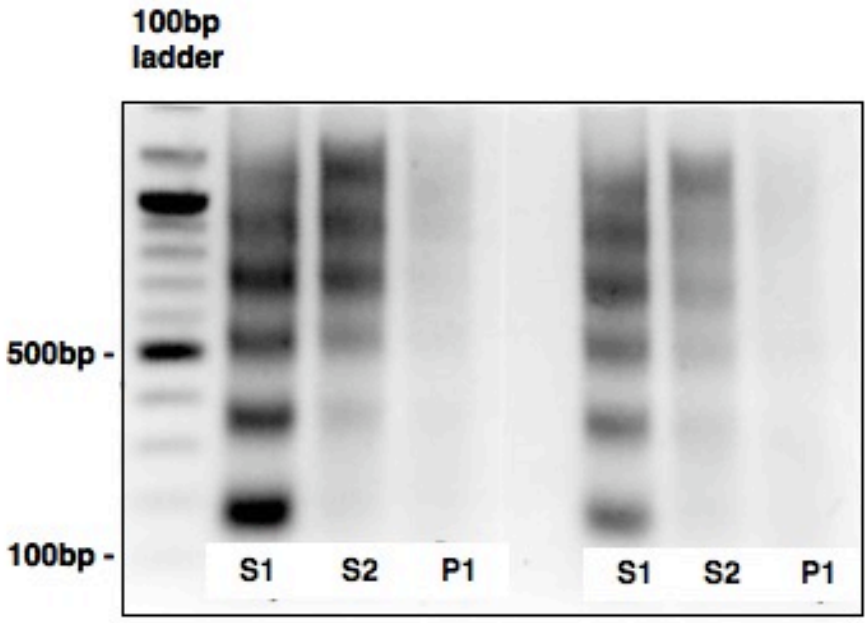


Figure 2.1 Chromatin agarose gel ectrophoresis. Chromatin was extracted from CCE/R cells and digested using micrococcal nuclease. S1 is the first supernatant obtained after digestion, S2 the supernatant obtained after overnight dialysis of the remaining pellet after digestion and P1 the pellet remaining after this overnight dialysis. 2.5 μg of DNA as measure by the A_{260} /A_{280} was electrophorised on a 1.2 % gel and stained using ethidium bromide.

a “tight” pestle using four cycles of 10 strokes with pauses of 10-15 min (on ice) between each cycle to prevent excessive foaming. This results in a 75-80% yield of intact nuclei as observed by counting under the microscope. Nuclei were pelleted (800 x g, MSE 3000, 15 min. 4oC), resuspended in 20 ml NB buffer, 5% (v/v) sucrose and pelleted (800 x g, MSE 3000, 25 min, 4oC). Nuclei were resuspended in 5 ml digestion buffer (50 mM Tris-HCl pH 7.4, 0.32 M sucrose, 4 mM MgCl2, 1 mM CaCl2, 0.1 mM PMSF). A260 was measured to give a rough estimate of DNA concentration and the sample adjusted to 250 μg/ml. 1 ml aliquots were mixed with 50 U micrococcal nuclease (Pharmacia) and incubated for 5 min at 28oC. The digested samples were spun at 217g, 5 min and the supernatant removed (fraction S1). The pellet was resuspended in 1 ml lysis buffer (2 mM Tris-HCl pH 7.4, 0.2 mM EDTA, 5 mM Na butyrate, 0.2 mM PMSF, 0.4 mM glycine) to give fraction S2. Both fractions were dialysed overnight at 4oC against lysis buffer, centrifuged (217g, 10 min MSE 3000) and the supernatants (fractions S1 and S2) analysed by agarose gel electrophoresis. If the chromatin was satisfactory (ie. predominantly but not exclusively mononucleosomes with minimal sub-nucleosomal fragments), S1 and S2 were combined and used for chromatin immunoprecipitation (ChIP).

2.2.4 N-ChIP

In order for maximum recovery of chromatin, siliconised pipettes, eppendorfs and 15ml Falcon tubes were used throughout. Affinity-purified antibody (50-125 μl, 50-100 mg/μl) was added to 50 μg unfixed chromatin and the final volume made up to 1 ml with incubation buffer (50 mM NaCl, 20 mM Tris-HCl, pH 7.5, 20 mM sodium butyrate, 5 mM Na_2EDTA, 0.1 mM PMSF). After overnight incubation (on a rotating platform) at 4°C, 200μl 50% w/v protein A-Sepharose (Pharmacia) was added and the incubation continued for a further 3hrs at room temperature. After centrifugation

(7500 x g, 10 min) the supernatant was removed and the protein A-Sepharose pellet resuspended in 1 ml wash buffer (50mM Tris-HCL (pH 7.5), 10mM EDTA, 5mM Na butyrate, 150mM NaCl) and layered onto 9ml of the same buffer. Following centrifugation (600 x g, 4°C, 10 min) the supernatant was removed by aspiration and the pellet washed twice more in 10ml wash buffer. Bound material was eluted from the protein A-Sepharose by addition of 125μl 1% SDS in incubation buffer and incubating for 15 min at room temperature on a rotating platform. After centrifugation (11 600 g, 10 min) the supernatant was removed and stored on ice. The protein A-Sepharose pellet was extracted as above with a further 125μl 1% SDS in incubation buffer. The two extracts were combined with an equal volume of incubation buffer to reduce the concentration of SDS to 0.5%. DNA was obtained from the input and bound fractions by two phenol/chloroform extractions and one chloroform extraction. DNA was ethanol precipitated using glycogen as a carrier and redissolved in 250μl water. All DNA samples were analysed initially by electrophoresis on a 1.2% agarose gels and stained with ethidium bromide. [^{3}H] Thymidine in each sample was determined by scintillation counting.

2.2.5 Preparation of cross-linked chromatin

Formaldehyde cross-linked chromatin was prepared essentially by the method outlined by (Orlando & Paro, 1993) Essentially, ES cells were harvested, washed 3 times in PBS / 5mM butyrate and resuspended at a concentration of 1 x 10^6 cells/ ml. Cells were cross-linked in 1% paraformaldehyde for 8 minutes at room temperature. The reaction was stopped by addition of glycine to a final concentration of 150mM. Cross-linked cells were washed twice with PBS 5mM butyrate and resuspended in 130 μl lysis buffer (50mM Tris-HCl, 10mM EDTA, 1%SDS, 5mM butyrate). Cells were sonicated using the Diagnode Biorupter for 10 minutes on medium at 4°C. An aliquot

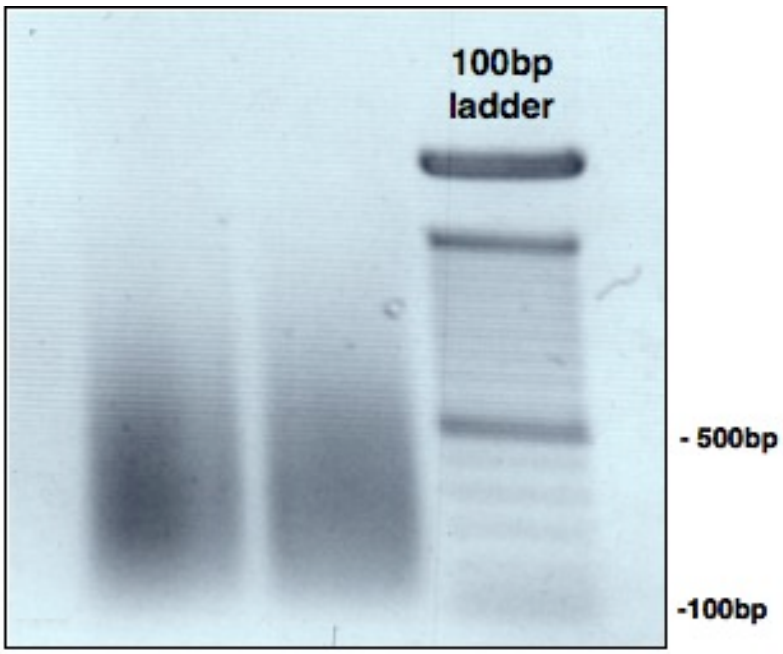


Figure 2.2 ES cell formaldehyde cross-linked chromatin. Cells were cross linked in 4% paraformaldehyde, lysed and sonicated. The cross links were reversed by proteinase K digestion and the resultant chromatin electrophorised on a 1% agarose gel. M=markers.

of cross-linked chromatin was taken for reversal of the cross links by proteinase K digestion at a concentration of 50 μg/ml at 68°C, 300rpm for 2 hours in an Eppendorf Thermomixer. DNA was extracted by two phenol/chloroform extractions and one phenol extraction. DNA was precipitated by centrifuging at 13,000rpm (MSE microcentaur) and resuspended in water. DNA was ran out on a 1% agarose gel to check the size of the fragments (Figure 2.2). Fragments were typically between 300 and 1000bp.

2.2.6 Precipitation of cross-linked chromatin

Antibody-bead complexes were formed by first washing Dynabeads protein A (Invitrogen) 4 times in RIPA buffer (10mM Tris-HCl pH 7.5, 1mM EDTA, 0.5mM EGTA, 1% Triton, 0.1% SDS, 0.1% Na deoxycholate, 150mM NaCl) and then incubating with 2.5μg antibody overnight at 4°C. 25μg cross-linked chromatin was added per antibody-bead complex and rotated at room temperature for 2 hours on a fast turntable. Beads were then washed with 200μl RIPA buffer five times before washing once with TE (1mM EDTA, 10mM Tris). Antibody-Bound DNA was eluted by addition of Elution buffer (20mM Tris-Hcl pH 7.5., 5mM EDTA, 5mM Na butyrate, 50mM NaCl) 1% SDS. Cross links were reversed by proteinase K digestion at a concentration of 50μg/ml at 68°C, 300 rpm for 2 hours in an Eppendorf Thermomixer. DNA was ethanol precipitated following 2 phenol/chloroform extractions and one phenol extraction. DNA was recovered by centrifuging at 13,000rpm (MSE microcentaur) and resuspended in 20 μl water.

2.3 Extraction of RNA and cDNA synthesis

RNA was extracted using the RNeasy mini kit (Qiagen) according to the manufacturers instructions. Concentration of RNA was analysed by nanodrop and 1 μg loaded on a 1% agarose gel to check RNA integrity (Figure 2.3). For cDNA

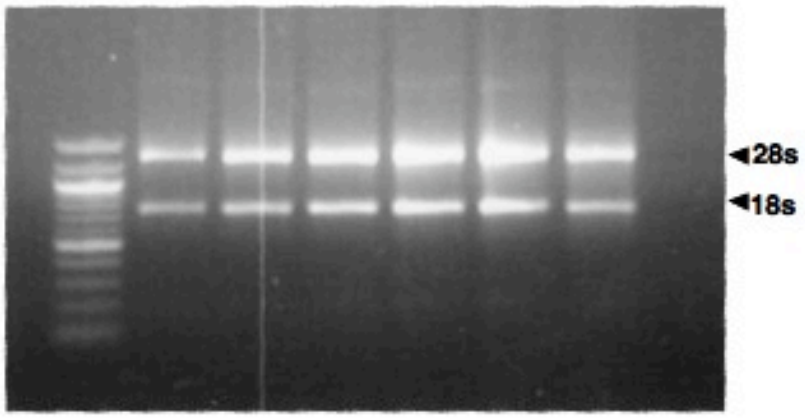


Figure 2.3 Extraction of total RNA from CCE/R cells. Total cell RNA was extracted using Qiagen RNAeasy kit according to the manufacturer's instructions. 1μg total RNA was loaded onto a 1% agarose gel to check the integrity of the RNA. The 28s and 18s RNA are labelled.

synthesis, 3μg RNA was first denatured by incubation at 65°C in 2.5μM oligo dT_{20} (Invitrogen) and 250 μM dNTP (Invitrogen). cDNA synthesis was performed by the addition of First Strand Buffer (final concentration 50mM tris-HCl, 75mNaCl, 3mM$MgCl_2$), 0.05U/μl SuperScript Reverse transcriptase (Invitrogen), 5μM DTT at 55°C and RNAase inhibitor and inactivated at 70°C. RNA was removed by addition of 0.1U/μl RNAaseH (NE Biolabs) and incubating at 37°C. The integrity of the cDNA was first checked by standard PCR using actin primers (Table 2.2) and running on an agarose to gel to ensure the production of a single amplicon.

2.4 Real-time SYBR green PCR

2.4.1 Native ChIP analysis

For native ChIP analysis input chromatin DNA was diluted to equal the concentration of the equivalent bound sample based upon the tritiated thymidine counts. All samples were then diluted x2 to achieve a concentration optimal for PCR. Real-time SYBR green PCR reactions were performed in a total volume of 10μl using SYBR green REDDY mix (Qiagen). Forward and reverse primers were added to a final concentration of 0.5μM each with 2μl DNA. Reactions were performed in an ABI 7900HT machine using the following conditions.

96°C 15mins

94 °C 15s
Tm 30s
72 °C 30s
} 40 cycles

95 °C 15s
60 °C 15s
95 °C 15s
} Dissociation curve

2.4.2 PCR primers for native ChIP analysis

All primers, (Table 2.1) except those to *Hoxb1, Hoxb9, Gapdh, Pou5f1* and *Nanog* were designed using sequences downloaded from the UCSC mouse genome assembly from February 2007. Primers were designed in the promoter region within 100bp of the first exon. Primers were designed using the web based programme primer 3 http://frodo.wi.mit.edu/ and ordered from Invitrogen. *Hoxb1, Hoxb9, Hoxb13* primers are detailed in Chamberyon *et al*, (2004). *Gapdh*, *Pou5fl* and *Nanog* primers were taken from O'neill *et al,* (2006). Sequences are in table 2.1. All primers were optimised such that only a single peak was produced on a dissociation curve.

2.4.3 Formaldehyde cross-linked ChIP analysis

For cross-linked ChIP analysis, input chromatin was serially diluted to generate a standard curve. Bound chromatin was added to the reaction without dilution. Real-time SYBR green PCR reactions were performed in a total volume of 10μl using SYBR green REDDY mix (Qiagen). Forward and reverse primers were added to a final concentration of 0.5μM each with 2μl DNA. Reactions were performed in an ABI 7900HT machine using the following programme.

96°C 15mins

94 °C 15s
Tm 30s
72 °C 30s
} 40 cycles

95 °C 15s
60 °C 15s
95 °C 15s
} Dissociation curve

Primers were the same as for native ChIP.

Gene	Forward	Reverse	Tm	Amplicon Length
Hoxb1	ttccatgtcgctctcagatg	taggaaggggctagggagtg	60	110
Hoxb3	gctttgtcttggttgtggtg	tccaaggagcagctttgact	60	139
Hoxb4	gctaaatgcgatttggaagg	gaggggtaggaggtgagtga	60	147
Hoxb5	acgactggtcaacaaaagca	gcgatgcactctacttcgtt	60	112
Hoxb6	atttccttctggccctcact	gccgggtttatgatttgttg	60	119
Hoxb7	gtaaatgcgagcaggaggag	agtcccagctggagcctact	60	110
Hoxb8	caacaacagactccggcttt	tatcgtgtggagggaattgg	60	100
Hoxb9	tccgaaagccctctgcac	gacattctcacacattatccgcg	60	110
hoxb13	gtgcgggagggattttatgag	gttttaaatcgctcccagctc	60	136
Hoxc8	aaagctggccttgtccatta	gggggacctggaggagag	60	101
Hoxc9	aggctgagttttccggttct	gtggcgaaaggagagacaat		
CCND1	cctccctcctagctgtcctc	gctcggactgcttctctcc	60	135
p57	gactgagagcaagcgaacag	tacctggctgattggtgatg	N/A	107
Pou5f1	ctgtaaggacaggccgagag	caggaggccttcattttcaa	60	140
Nanog	ctatcgccttgagccgttg	aactcagtgtctagaaggaaagatca	60	158
GAPDH	tgtgccaagcacttgtataac	tatgtctgaccagaggagagca	60	149

Table 2.1 Primers used for ChIP analysis in real-time SYBR Green PCR**.**

2.4.4 cDNA Analysis

For cDNA analysis, cDNA from reverse transcriptase reactions were first diluted 10 x to a concentration of 100ng/μl according to Nanodrop analysis. Real-time SYBR green PCR reactions were performed in a total volume of 10μl using SYBR green REDDY mix (Qiagen). Forward and reverse primers were added to a final concentration of 0.8μM each with 2μl cDNA. Reactions were performed in an ABI 7900HT machine using the following programme.

96°C 15mins

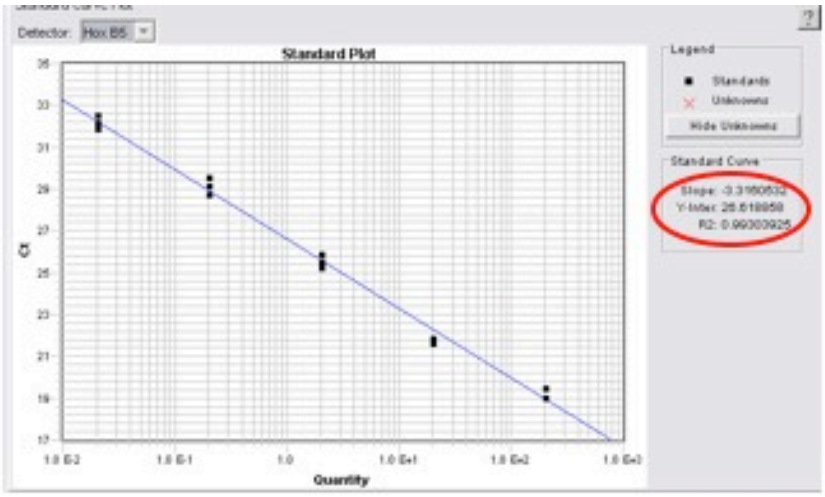
94 °C 15s
Tm 30s
72 °C 30s
} 40 cycles

95 °C 15s
60 °C 15s
95 °C 15s
} Dissociation curve

2.4.5 PCR primers for cDNA expression analysis

PCR primers were designed using primer3 such that they spanned an intron http://frodo.wi.mit.edu/. The sequences are in Table 2.2. Primers were optimised using serial dilutions of cDNA know to contain the target gene using different annealing temperatures until the standard curve was equal to -3.3 +/- 0.3 and the r^2 was no less than 0.95. An example is shown in Figure 2.4

a.



b.

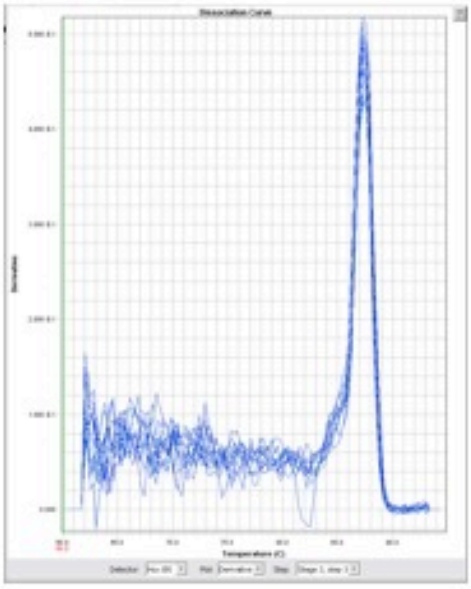


Figure 2.4 Standard curve and dissociation curve of *Hoxb5* expression primers. **a.** A standard curve was generated by amplifying using an annealing temperature of 62ºC from serial dilutions of cDNA derived from E12 embryos. The gradient and r^2 value are circled. **b.** The dissociation curve produced from the standard curve in **a.**

Gene	Forward 5' to 3'	Reverse 5' to 3'	Tm	Amplicon Length
Hoxb1	ccatatcctcgccgcag	cggactggtcagaggcatc	58	450
Hoxb2	cggcgcctccacccttcagagacc	ctttcggtgaggtccagcaaggc	60	373
Hoxb3	caactccaccctcaccaaa	gccaccaccacaaccttc	58	106
Hoxb4	ctggatgcgcaaagttcac	tccttctcaactccaggac	62	113
Hoxb5	cctctgagcccgaggaag	ccagggtctggtagcgagta	60	207
Hoxb6	gagaccgaggagcagaagtg	actgagctgagacgcactga	62	300
Hoxb7	aaccgagttccttcaacatg	cgagtcaggtagcgattgta	55	250
Hoxb8	ttctacggctacgaccctct	cgtgcgatacctcgatcctc	55	289
Hoxb9	cagggaggctgtcctgtctaatc	cttctctagctccagcgtctgg	58	176
hoxb13	ctggaacagccagatgtgtt	cctgctaaaggtgtcatctc	58	445
Hoxc8	cctccgaccaacactaacagt	caaggtctgataccggctgt	60	157
Hoxc9	cagcaagcacaaagaggaga	tcatcttcatcctccggtta	60	258
Nanog	ctcatcaatcaatgcctgcagtttttca	ctcctcagggcccttgtcagc	58	210
Pou5fl	ccaatcagctttgggctagag	cctgggaaacctgtcctgta	60	130

Table 2.2 Primers used for cDNA analysis in real time SYBR green PCR

2.4.6 CChIP PCR analysis

PCR reactions were performed in duplicate with mouse and *Drosophila* DNA controls run in parallel to monitor cross-hybridization. 45 μl of Reddy Mix PCR Master Mix (AB Gene, UK) was added to 3 μl input DNA (about 25 ng) and 2 μl primer mix. 0.1 μCi of dCTP radiolabelled with α-^{32}P (Amersham, UK) was added to each PCR reaction before cycling (41 cycles of 60 sec at 95°C, 60 sec at 60°C, and 90 sec at 72°C). In all reactions, unbound samples were diluted to equal the concentration of the equivalent bound sample using the [3H] Thymidine counts.

In the standard protocol, aliquots were removed after 38 and 41 cycles and loaded onto 5% polyacrylamide gels and electrophoresed at 400Volts and 30mA for 15 minutes. Gels were dried onto filter paper (SpeedGel System, Thermo Savant, UK) for a minimum of 2 hours. Filters were exposed to a phosphor screen overnight and scanned with a PhosphorImager (Typhoon 9200, Amersham, UK). Intensity values for each PCR product were analysed with 'Image Quant 5.2' software (Molecular Dynamics).

2.5 Antibodies

Rabbit polyclonal antisera to H3K9K9Ac (R607) and H3K4me3 (R612), H4K5ac (R401) H4K8ac (R403), H3K4me2 (R149) were raised by immunization with synthetic peptides conjugated to ovalbumin as previously described (Turner & Fellows, 1989; Turner et al, 1989; White et al, 1999). Specificity was assayed by inhibition ELISA for all in-house antisera used and checked by western blotting. For all antisera, cross-reaction with epitopes other than that against which the antiserum was raised was insignificant. Antibodies against H3K27me3 were from Millipore and

were re-tested in house to ensure specificity. Cross-reaction with H3K9me3, which shares the motif ARKS was minimal.

2.6 Histone Acid Extraction

Cells were harvested as detailed. Cells were suspended in Triton extraction buffer (1 xPBS, 5mM Na Butyrate, 0.5% Triton X-100, 2μM PMSF, 0.02%Na azide) at a concentration of 10^7 cells/ml. Cells were incubated on ice for 10 minutes and pelleted at 217 x g for 10 minutes (MSE chilspin). Cells were then resuspended in half the previous volume of TEB and centrifuged again. Histones were extracted in 0.2N HCl (50μl / 2 x 10^6 cells) at 4°C overnight. Histones were then isolated by centrifugation (217g for 10 minutes) and the histone containing supernatant removed.

The protein concentration of histone samples was calculated using the Pierce assay. Samples were diluted 50 x in Pierce reagent and left on ice for the colour change to develop. The concentration was analysed by absorption at 595nm against an equally treated sample of 0.2N HCl. The concentration of histones was determined using a BSA standard curve.

2.6.1 Sodium doecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were separated according to size using the method developed by (Laemmli, 1970). For histone separation, a resolving gel of 15% acrylamide, 0.4% NN'bisacrylamide, 375mM Tris-HCl pH 8.8%, 0.1% SDS polymerised with 300 μl 10% w/v ammonium persulphate and 30μl TEMED per 30ml gel solution. The gel was overlaid with iso-butanol during polymerisation to prevent evaporation which was then washed off before addition of the stacking gel. The stacking gel consisted of 3%

acrylamide, 0.16% N,N'bisacrylamide, 125mM Tris-HCl pH6.9 0.1% SDS, polymerised with 100 μl 10% w/v ammonium persulphate and 10μl TEMED per 10ml of gel solution.

Protein samples were prepared for SDS PAGE as follows:

x μl protein (5-10μg) + yμl H_20 (x + y = 60μl), 10μl standard dissociating buffer (1M Tris-HCl pH 7.2, 10mM Na_2EDTA, 10%SDS, 1.432M 2-mercaptoethanol) 25μl 50% glycerol, 5μl 0.02% bromophenol blue. Samples were incubated at 100°C for 10 minutes and cooled on ice for 5 minutes and loaded onto the gel though SDS reservoir buffer (50mM Tris, 0.384M glycine, 0.1% SDS). Gels were electrophoresed at 400V, 30mA, 20W until the bromophenol blue band had just ran off the end of the gel.

2.6.2 Acid-urea-Triton (AUT) polyacrylamide electrophoresis

Proteins were separated according to charge on AUT-PAGE essentially as described (Bonner et al, 1980). The resolving gel consisted of 12% acrylamide, 0.32% NN' bisacrylamide, 8M urea, 1M glacial acetic acid and 0.05M ammonia. Gel solutions were degassed before addition of 500 μl Triton X100, 150μl TEMED and 2ml 0.0004% riboflavin per 30ml gel solution and photopolymerised in front of a 2 x 15W light source. The resolving gel was overlaid with a few drops of water saturated isobutanol to prevent evaporation during polymerisation. Isobutanol was washed off the resolving gel with several changes of double distilled water before addition of the stacking gel. The stacking gel consited of 4% acrylamide, 0.21% bisacrylamide, 8M urea, 1m glacial acetic acid and 0.15M ammonia. It was degassed before addition of 100μl TEMED and 700μl 0.004% riboflavin per 10ml gel solution. Photopolymerisation was carried out as for the resolving gel.

Protein samples were prepared for AUT-PAGE as follows:

1 volume of protein sample (20-50μg) was mixed with 2 volumes of AUT-loading buffer (8M urea, 5% 2 – mercaptoethanol, 1M glacial acetic acid, plus a few drops of tracking dye pyroninY), vortexed and spun at 13,000rpm for 15minutes (MSE microcentaur) before loading through AUT reservoir buffer (1M glycine, 0.1M glacial acetic acid). The gels were electrophoresed for 15hours at 150V, 30mA, 10W at 10°C using reversed polarity.

2.6.3 Transfer of proteins onto nitrocellulose

Proteins separated by SDS-PAGE were transferred onto Hybond C nitrocellulose filters (Amersham) essentially as described (Towbin et al, 1979). Briefly, a piece of Hybond C was cut to the same size as the gel, pre-soaked in transfer buffer (25M Tris, 192mM Glycine, 20% Methanol) and overlaid on top of the gel. Gels were sandwiched between Whatman No1 filter paper and Scotch brite sponge pads. All air bubbles were removed and the gel plus pads placed in a cassette and slotted into a Biorad Transfer apparatus, ensuring the membrane was situated between the gel and the anode. Protein transfer was carried out at 300mA for 3 hours for SDS-PAGE.

2.6.4 Western Blotting

A non-radioactive method for detection of proteins on nitrocellulose was used, Enhanced Chemiluminescence (ECL) (Amesham). After transfer, filters were blocked in 5% powdered milk/TBST (TBS (20mM Tris pH7.5, 150mM NaCl), 0.1% Tween 20), for 1 hr at room temperature and rinsed in TBST before addition of the first antibody. After incubation at room temperature for 1hr, the filters were washed vigorously with

3 changes of TBST. The secondary antibody, goat anti-rabbit horse radish peroxidase was applied to each filter diluted at x20,000 in TBST and the incubation continued for a further 1 hour. All filters were washed vigorously with 3 changes of TBST before addition of substrate (Amersham). Filters were wrapped in Saran wrap and exposed to film several times in order to optimise signal to noise ratio.

3 Results

3.1 Histone modifications at the *Hoxb* cluster

3.1.1 What is the pattern of histone modifications at the promoters of the *Hoxb* cluster?

Specific histone modifications are associated with distinct transcriptional outcomes. H3K4me3 and histone acetylation, for example, are generally associated with transcriptionally active loci whereas H3K9me3 is generally associated with heterochromatin and transcriptionally repressed silent domains. These modifications are therefore often referred to as “active” and “repressive” respectively. However, these terms are broad generalisations and counter examples of active modifications being associated with inactive genes, and vice versa, do exist (Vakoc et al, 2005).

Genome wide studies in pluripotent cells have shown the *Hox* genes are located within chromatin domains that contain both active and repressive modifications, termed “bivalent domains” (Bernstein et al, 2006). It has been suggested that they may maintain gene repression in an undifferentiated state but also allow genes to be activated upon differentiation. The two modifications H3K4me3 and H3K27me3 are associated with these bivalent domains and are put in place by the trithorax and polycomb proteins respectively (Ringrose & Paro, 2004). These gene products are responsible for maintaining the spatial expression of *Hox* genes in the early embryo (Hanson et al, 1999). The patterns of H3K4me3 and H3K27me3 over the Hoxb genes as elucidated by ChIP-seq is shown in figure 3.1. Progressing from these genome-wide ChIP-seq studies here is described, a detailed study of the role of individual modifications at the *HoxB* gene promoters, performed using native ChIP coupled with qPCR. Such a study will a allow a high resolution picture to be generated at these *Hox* gene promoters.

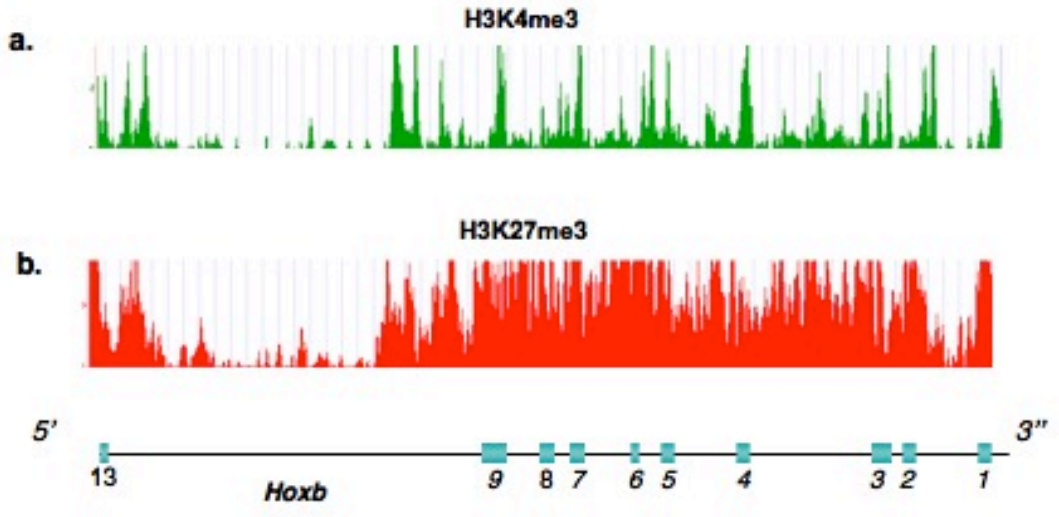


Figure 3.1 The distribution of histone modifications over *Hoxb* promoters identified by ChIP-seq. Chromatin was extracted from V6.5 murine ES cells, fixed with 1% formaldehyde and immunoprecipitated using antibodies directed at the histone modifications **a**. H3K4me3 **b**. H3K27me3 and their distribution across the *Hoxb* cluster examined by high-throughput sequencing. Adapted from Mikkelsen *et al, 2007.*

As appropriate for a study investigating bivalent domains, chromatin was extracted from the undifferentiated mouse embryonic stem cell line CCE/R and immunoprecipated using antibodies raised against the modifications H3K4me3, H3K27me3, and the activating modification H3K9ac. PCR primers were designed against the promoters of the genes *Hoxb1, b3, b4, b5, b6, b7, b8, b9* and *b13* within 100bp of the first exon and these primers were used to analyse the amount of these gene promoters in the input and bound chromatin by quantitative real time PCR (Figure 3.2). Primers designed against *Hoxb2* repeatedaly failed to amplify for unknown reasons. A ratio of bound DNA to input DNA greater than one indicates an enrichment for the modification at the location of that primer set, and a ratio of less than one indicates a depletion.

Analysis of the distribution of H3K9 acetylation shows a non-uniform distribution of this modification across *Hoxb* gene promoters (Figure 3.2a). *Hoxb9* shows much the highest enrichment for this modification with a bound:input ratio of 5; this gene therefore contains more than four times the levels of H3K9ac in its promoter than *Hoxb1*, and more than twice the level of enrichment seen across the rest of the cluster. *Hoxb3* is also conspicuous, but in its lack of H3K9ac compared with the other genes in the cluster. The majority of the genes display relatively low levels of the modification with bound:input ratios of between one and two.

Subsequent analysis of the distribution of H3K4me3, another activating mark, shows its levels are clearly linked with those of H3K9ac (Figure3.2b). Again, there is a striking enrichment for this active modification at the promoter of *Hoxb9.* In addition, *Hoxb3* displays the lowest level of enrichment for this modification as it did for H3K9ac. There are two other gene promoters that also contain elevated levels of H3K4me3; *Hoxb4* and *Hoxb5*. These two gene promoters did not stand out so prominently as having high levels of H3K9ac relative to the other gene promoters. It

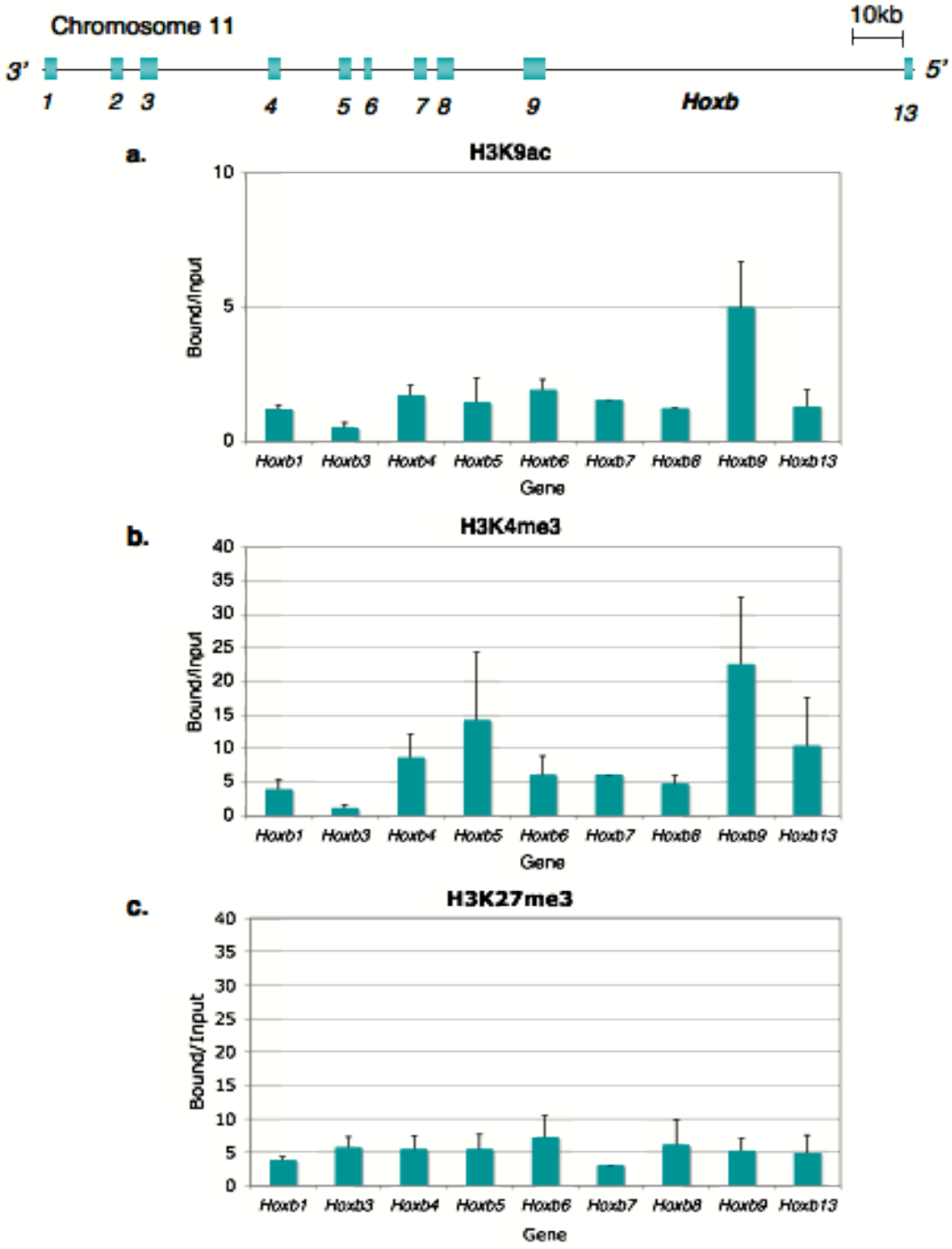


Figure 3.2 The distribution of histone modifications over *Hoxb* promoters. Chromatin was extracted from undifferentiated CCER cells and precipitated using antibodies directed at the histone modifications **a**. H3K9ac, **b**. H3K4me3 and **c**. H3K27me3. The levels in the bound and input fractions were determined by quantitative real time PCR in triplicate. Experiments were performed in duplicate. Error bars represent the standard error. A scale map of the cluster is provided at the top for reference.

is also important to note that despite the pattern of relative levels of H3K9ac and H3K4me3 being similar, H3K4me3 is present at higher absolute levels at *Hoxb* promoters than H3K9ac.

Finally, the abundance of the repressive mark H3K27me3 is generally more similar, in value, to that of H3K4me3 than H3K9ac (Figure 3.2 b and c). However, much less gene by gene variation is displayed compared with both previously discussed active modifications. This serves to emphasise the variation seen in the active modifications. For example, *Hoxb9* that displayed the highest levels of active modifications at its promoter does not show an equivalent high level of H3K27me3 and so has a high ratio of active to repressive modifications. Other genes showing a similar pattern are *Hoxb4, Hoxb5* and *Hoxb13.*

3.1.2 Are levels of histone modifications found in the *Hoxb* cluster promoters predictive of gene expression?

Histone modifications are known to play a role in transcriptional regulation, but this could reflect either ongoing transcription, or future activity. That is, in a gene not yet expressed, a high level of a specific combination of active modifications could indicate its imminent transcription, making these “predictive” marks. Since there is no gene transcription from *Hox* gene promoters in undifferentiated mouse embryonic stem cells, the active histone modifications seen at the *HoxB* promoters were postulated to predict future patterns of gene expression (Bernstein et al, 2006). Therefore, in an attempt to convey meaning to the varying level of activating modifications seen at the *Hoxb* gene cluster, gene expression patterns in differentiating CCE/R cells were determined to establish any correlation with histone

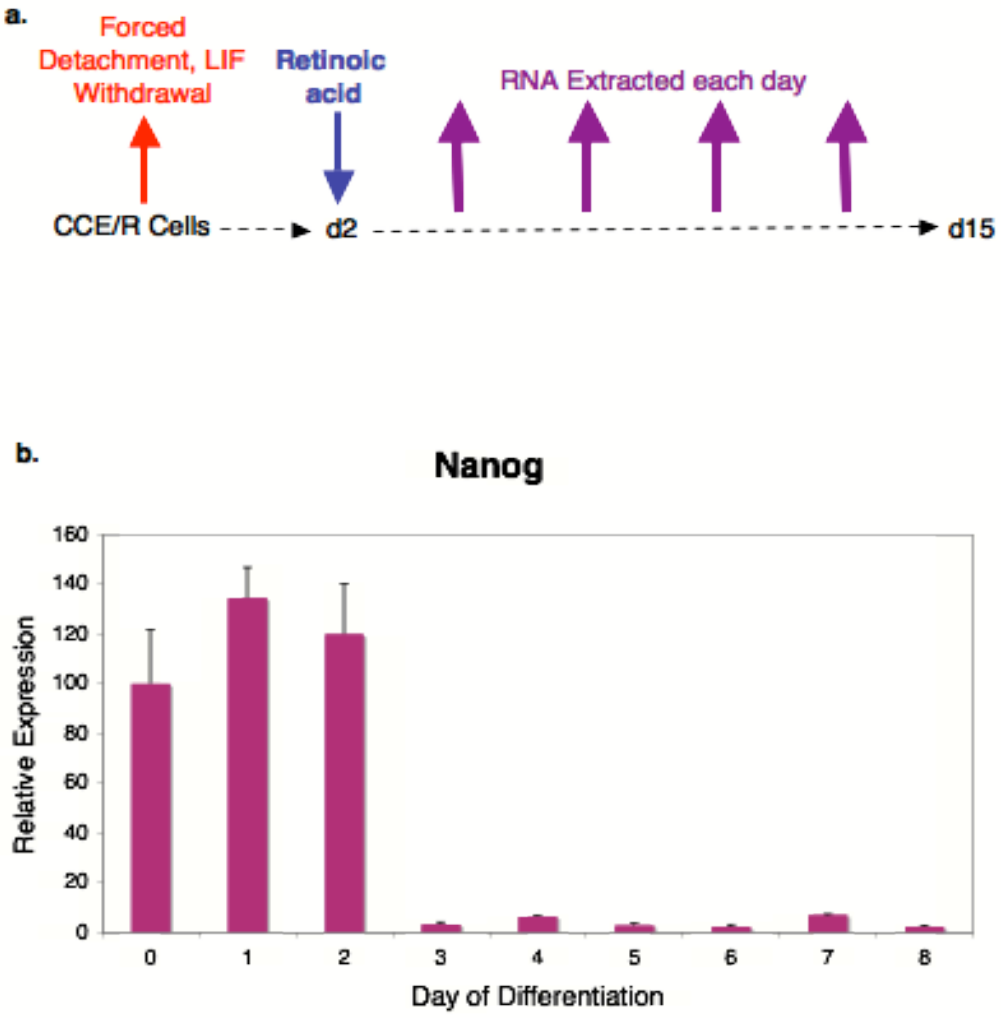


Figure 3.3 CCE/R cells show reduced expression of pluripotent markers upon addition of retinoic acid. Cells were differentiated as in **a.** CCE/R cells were induced to form embryoid bodies by replating and LIF withdrawal on day 0. On day 2, retinoic acid was added to a final concentration of 1 μM. RNA was extracted on each day and cDNA synthesised. The level of *Nanog* mRNA relative to day 0 was assessed by real time PCR and normalised to actin. Error bars represent the standard error of PCR triplicates.

modification levels at *Hoxb* gene promoters. In early embryos, the timing and spatial restriction of *Hox* gene expression is co-linear with their physical location within the cluster; the more 3' the gene, the earlier it is switched on, and the more anterior its expression (Kmita & Duboule, 2003). The sequential timing of induction of homeotic genes in response to retinoic acid in mouse embryonic carcinoma cells has been shown to reflect that of embryos (Papalopulu et al, 1991; Simeone et al, 1990).

In order to confirm that this sequential induction of *Hox* gene expression is conserved in CCE/R mouse embryonic stem cells, the cells were differentiated into embryoid bodies in the presence of retinoic acid (Figure 3.3a). RNA was extracted at each day of differentiation, cDNA synthesised, and the levels of *Hox* gene mRNA compared to that in undifferentiated cells and normalised to actin using real-time PCR. In addition, the level of the pluripotency marker *Nanog* was assayed. *Nanog* is a marker of pluripotency expressed by undifferentiated embryonic stem cells (Niwa, 2007). Upon differentiation, its expression is diminished. In this system as expected, its expression was found to decrease after day 2 of differentiation (Figure 3.3b).

The induction of genes from the *Hoxb* cluster occurred at defined times of differentiation with transcription generally commencing in a temporal sequential manner from the 3' to the 5' end of the cluster (Figure 3.4). In addition, all the genes seem to display a similar low level of transcription before the gene is fully induced. This may be a result of transcription from a minor lineage in the embryoid body before the major tissues that express the gene are fully formed. Alternatively, it may be due to the one drawback of this system. That is, when differentiated in this manner, embryonic stem cells differentiate asynchronously generating broad peaks of *Hox* gene expression.

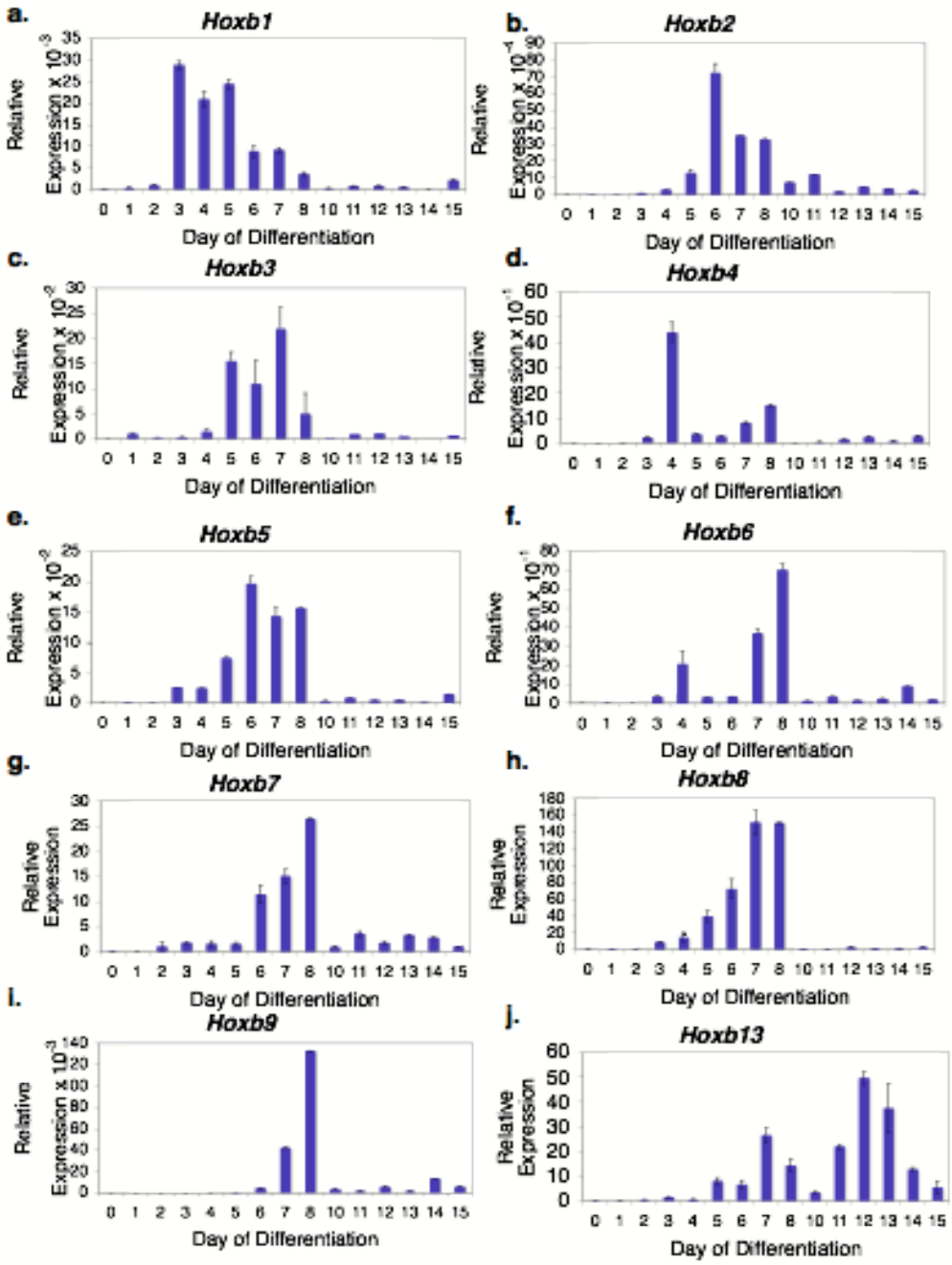


Figure 3.4 The *Hoxb* genes are sequentially induced upon CCE/R cell differentiation. CCE/R cells were differentiated by first inducing embryoid body formation by LIF withdrawal and replating. On day 2 retinoic acid was added to a final concentration of 1μM. RNA was extracted each day and cDNA synthesized. Expression relative to day 0 (with expression at day 0 being set at 100 or as powers of 100 as indicated) was determined by quantitative real time SYBR green PCR and normalised to actin. Error bars represent the standard error of the mean of PCR triplicates from a single experiment.

Hoxb1, the most 3' gene; and the first to be induced in differentiating embryos was also the first gene to be transcribed in this system (Figure 3.4a). There was a dramatic peak in the quantity of *Hoxb1* mRNA at day 3, which was maintained until day 5. *Hoxb2*, the adjacent gene reached maximal expression three days later (Figure 3.4b). An initial bout of expression from this locus was evident at day 4, and maximal expression was reached at day 6. At days 7 and 8 of differentiation, expression of the gene diminished. *Hoxb3* is switched on co-ordinately with *hoxb2* (Figure 3.4b and c) but reached its maximal expression previous to that of *Hoxb2,* at day 5. *Hoxb4* displays a very sharp peak in transcriptional activity at day 4 (Figure 3.4d) with expression again beginning to increase at day 8. *Hoxb5* is induced at a low level at day 3 (Figure 3.4e) and this low level is maintained until day 6 where maximal expression is reached. *Hoxb7* displays 3 phases of gene expression (Figure 3.4g); a minimal expression at days 2 to day 5, half maximal gene activity at days 6 and 7, and a final maximal period at day 8. *Hoxb8* also takes on a similar three-tiered pattern of gene expression; a very low level between day 3 and day 5, an intermediate level at day 6 and a final maximal level at days 7 and 8 (Figure 3.4h). Its expression does then, however enter a fourth phase until day 10. *Hoxb9* reaches a sharp peak of expression at day 8 which has disappeared by day 10. As would be expected, *Hoxb13* is the final gene to be fully induced, peaking in expression at day 12. However, there is significant leaky expression of this gene between days 5 and 8.

As regards histone modifications, levels of active marks are not a marker for early or late gene expression during differentiation. This is particularly evident at *Hoxb9* which displays the highest level of active modifications but is not induced until day 8.

3.2 Can a change in histone modifications induce transcription from the *Hoxb* cluster?

In order to further distinguish the direct relationship between histone modifications and control of gene expression at this locus, it was decided to directly alter the level of one of the active modifications and ascertain whether this induced any effects upon gene expression. Specifically, could an increase in an active mark override the presence of a repressive mark?

To alter the histone modifications, it was decided to treat with an HDAC inhibitor of the histone deacetylases. HDAC inhibitors fall into six structurally distinct classes; the small molecular weight carboxylates, the hydroxamic acids, benzmides, epoxyketones and the cyclic peptides. Several HDAC inhibitors have made it to the clinic as anti-cancer dugs owing to their anti-proliferative effects. Previous studies altering levels of histone acetylation at *Hox* genes have used the inhibitor TSA (Chambeyron & Bickmore, 2004). This inhibitor is a hydroxamic acid and is known to mediate its inhibitory effect by chelating the active zinc molecule of HDAC active sites, in combination with blocking the entrance to the active site and spanning the hydrophobic pocket of the active site (Drummond et al, 2005). In this study, however, valproate, a carboxylic acid was chosen. It too is able to chelate zinc ions, but is less effective than the hydroxamic acids at inhibiting HDACs and so has a much weaker HDAC inhibition constant. Valproate is a class II HDAC inhibitor, and also a teratogen, causing developmental defects in early embryos. It has been used widely in the clinic as an anti-epileptic drug but has been shown to cause birth defects if taken during pregnancy (Duncan, 2007). Studies in mice show it causes homeotic transformation and aberrant *Hox* gene expression in embryos exposed to the drug (Faiella et al, 2000). This may, or may not be a direct result of its histone deacetylase inhibitor activity. Valproate was, therefore was chosen above other available HDAC

inhibitors as a tool used to induce histone hyperacetylation across the *Hox* cluster. However, it must be noted that valproate is able to cause several potentially relevant effects within cells. Firstly it has been reported to be able to sequester acetyl CoA as it may be oxidised into its derivative valproylCoA, upsetting the metabolic balance in the cell (Becker & Harris, 1983). Secondly, as mentioned above, it chelates zinc ions, and finally it is a known inhibitor of the Wnt pathway, a pathway involved in control of stem cell differentiation, something particularly relevant to this study treating undiffentiated embryonic stem cells (Bug et al, 2005).

3.2.1 What are the global effects of valproate on CCE/R cells?

Initial experiments were performed in order to confirm that valproate is able to induce hyperacetylation in CCE/R cells through its HDAC inhibitor activity. Valproate was added to the culture medium of undifferentiated CCE/R mouse embryonic stem cells in the presence of LIF to a concentration of 1mM for a duration of 2 and 8 hours. Previous experiments had confirmed that this concentration was not toxic to these ES cells.

Upon treatment of CCE/R cells with valproate there was an alteration of the morphology of the cells (Figure 3.4a). Undifferentiated CCE/R cells grow in colonies and upon addition of valproate to the culture medium, after 2 hours, there is no visible change in the cells (Figure 3.5a). After 8 hours of treatment however, the colonies flatten out, the individual cells within these colonies become more distinct and those at the edge of the colony develop protruding spines (Figure 3.5a). It is possible that valproate causes the cells to differentiate. However, the flattening-out and development of spines on the cells could also be the result of hyperacetylation of structural proteins within the cell, causing an alteration in the cell cytoskeleton. Previous experiments with another HDAC inhibitor, trichostatin A (TSA) in HeLa cells

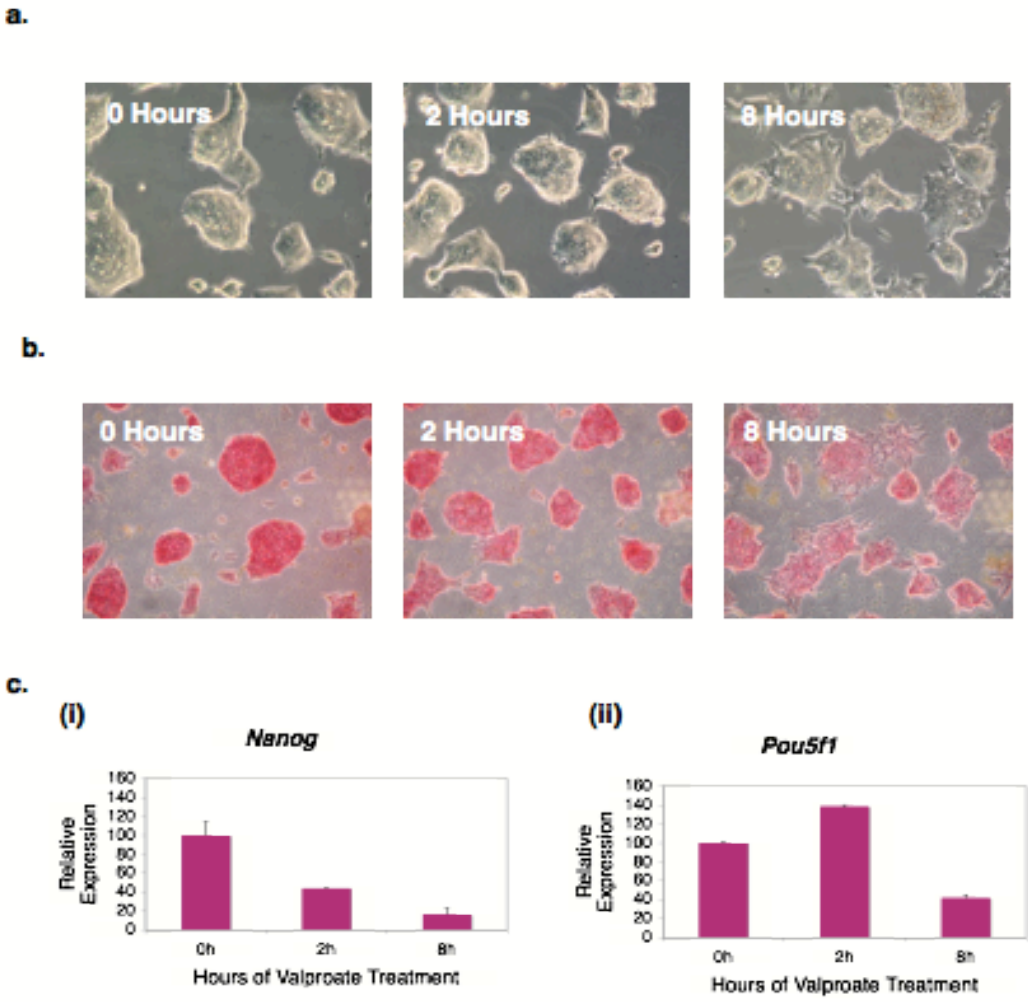


Figure 3.5. The global effects of sodium valproate on CCE/R cells. CCE/R cells were treated with 1mM sodium valproate for 2 and 8 hours as indicated. **a.** Morphology of cells was examined using a Zeiss Axiovert 25 and images captured using a Canon Digital EOS 400D **b.** Alkaline phosphatase assays were performed and images captured as for cell morphology **c.** RNA was extracted and the levels of (i) *Nanog* and (ii) *Pou5f1* RNA compared with that before treatment and normalized to actin by quantitative real time PCR.

have shown this type of hyperacetylation after a 24 hour treatment (Kim et al, 2006b).

To further investigate whether the morphological changes induced by valproate are due to differentiation of the embryonic stem cells, alkaline phosphatase assays were performed on the cells (Figure 3.5b) as this enzyme is seen as a marker for pluripotency in embryonic stem cells (Berstine et al, 1973). The assay results in the enzyme substrate being converted into a pink product and hence undifferentiated cells stain pink. The colonies of untreated CCE/R embryonic stem cells stain bright pink in the alkaline phosphatase assay. Similarly, colonies of the treated embryonic stem cells still stain pink after 2 and 8 hours (Figure 3.5b). However, after 8 hours, the staining is paler and closer inspection of the treated colonies reveals that there is, indeed, less staining for alkaline phosphatase at the edge of the flattened colonies where cells have spikey protrusions (Figure 3.5b).

Due to this decrease in alkaline phosphatase staining, the investigation into any possible differentiation upon valproate treatment was continued. *Pou5f1* and *Nanog* are markers of pluripotency expressed by undifferentiated embryonic stem cells (Niwa, 2007). Upon differentiation, their expression is diminished. RNA was extracted from valproate treated cells and the levels of *Pou5f1*and *Nanog* mRNA compared with that before treatment and normalised to actin (Figure 3.5c). If valproate is inducing differentiation then the levels of these markers will decrease with respect to actin upon treatment. The data clearly shows a marked reduction in the levels of Nanog RNA after 2 hours of treatment with valproate (Figure3.5c). Transcription from the *Pou5f1* locus is also decreased but not until 8 hours of treatment with valproate (Figure 3.5c). Thus, transcription from the *Nanog* locus is diminished before that from *Pou5f1*.

3.2.2 What are the effects of valproate upon global histone modifications?

To assess the global affect of valproate upon histone tail modifications, histones were extracted by acid extraction and western blots were performed using the extracted histones. Antibodies used to probe the blots were those raised against the modifications; H3K9ac, H4K5ac, H4K8ac, H3K4me3 and H3K27me3. Several antibodies raised against histone acetyl lysines were used in order to fully ascertain the affect of valproate on global histone acetylation (Figure 3.5). Anti-H3K4me3 and anti-H3K27me3 antibodies were used due to the presence of these marks on the *Hoxb* cluster and their possible importance in control of gene expression. As expected, there is an increase in the global level of all the histone acetyl marks in CCE/R cells upon treatment with 1mM valproate (Figure 3.5). H3K9ac and H4K8ac show a steady increase, whereas H4K5ac levels plateau after 2 hours. These acetyl marks are associated with active transcription, and there is also a concomitant rise in the active H3K4me3 mark. This is not unexpected since it has been shown previously in other cell lines, that histone acetylation and H3K4me3 are linked possibly due to MLL stimulation by the acetyl mark (Nightingale et al, 2007). The level of H3K27me3 stays constant throughout treatment. This is despite the increase in its partnering bivalent mark H3K4me3, thus these two marks do not behave globally in a concomitant manner.

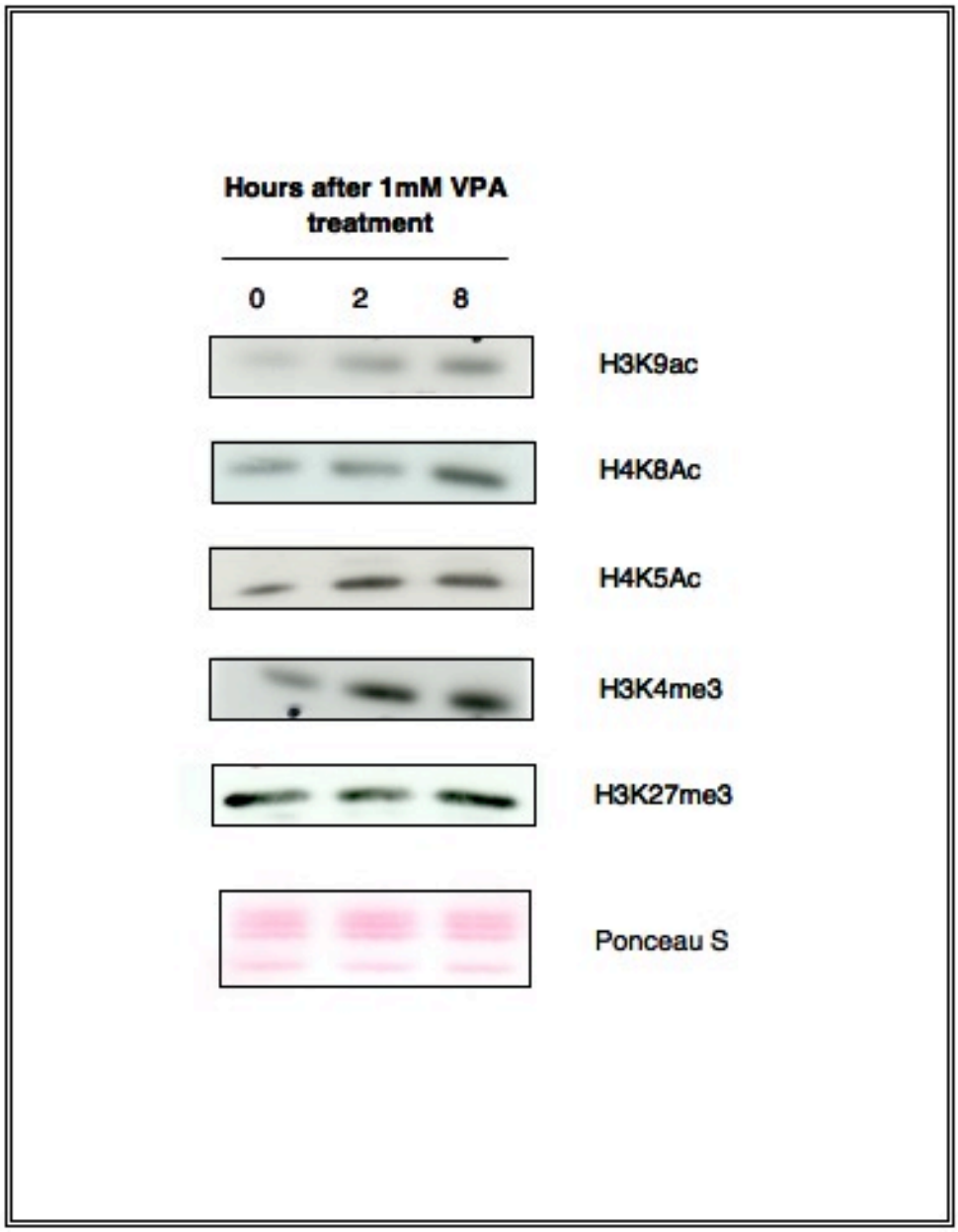


Figure 3.6 The effects of sodium valproate on global histone modifications in CCE/R cells. CCE/R cells were treated with 1mM valproate for 2 and 8 hours in the presence of LIF and histones acid extracted at each timepoint. Histones were separated by SDS PAGE and Western blotted using antibodies directed against H3K9ac, H4K8ac, H4K5ac, H3K4me3 and H3K27me3 (see materials and methods). This was repeated three times and representative blots are shown. Membranes were stained with Ponceau S to check for equal loading.

3.2.3 Can valproate induce hyperacetylation at individual genes?

The changes induced by valproate were next assessed at the level of histone modifications over *Hoxb* promoters. In order to assess the effect of valproate, undifferentiated CCE/R cells were cultured in 1mM valproate for 2 and 8 hours in the presence of LIF. After both these timepoints, chromatin was extracted and immunoprecipitated using antibodies directed against H3K9ac, H3K4me3 and H3K27me3. As before, a bound to input ratio was then calculated. As a control for the potentially unique nature in which *Hox* gene expression is controlled, primers were also designed against the *Gapdh* promoter, *Pou5f*1 and *Nanog*.

In CCE/R cells, the promoter of *Gapdh* is refractory to the effects of valproate (Figure 3.6a). Given its status as a housekeeping gene, *Gapdh* is expressed in CCE/R cells, and hence is likely to be enriched for active modifications. The active modification H3K9ac is, indeed present at the *Gapdh* promoter, albeit at low levels (Figure 3.6a) and its levels are not increased after treatment with valproate. Thus the H3K9ac modification at *Gapdh* responds to valproate in a manner that is contrary to the global response of the histone tails where a significant global hyperacetylation is seen upon valproate treatment (Figure 3.5). The active modification H3K4me3 is also found at the *Gapdh* promoter at similar levels to H3K9ac and once again, its levels, too are unaltered by treatment with valproate. This is again contradictory to the global response of this modification (Figure 3.5). H3K27me3 has a bound:input ratio

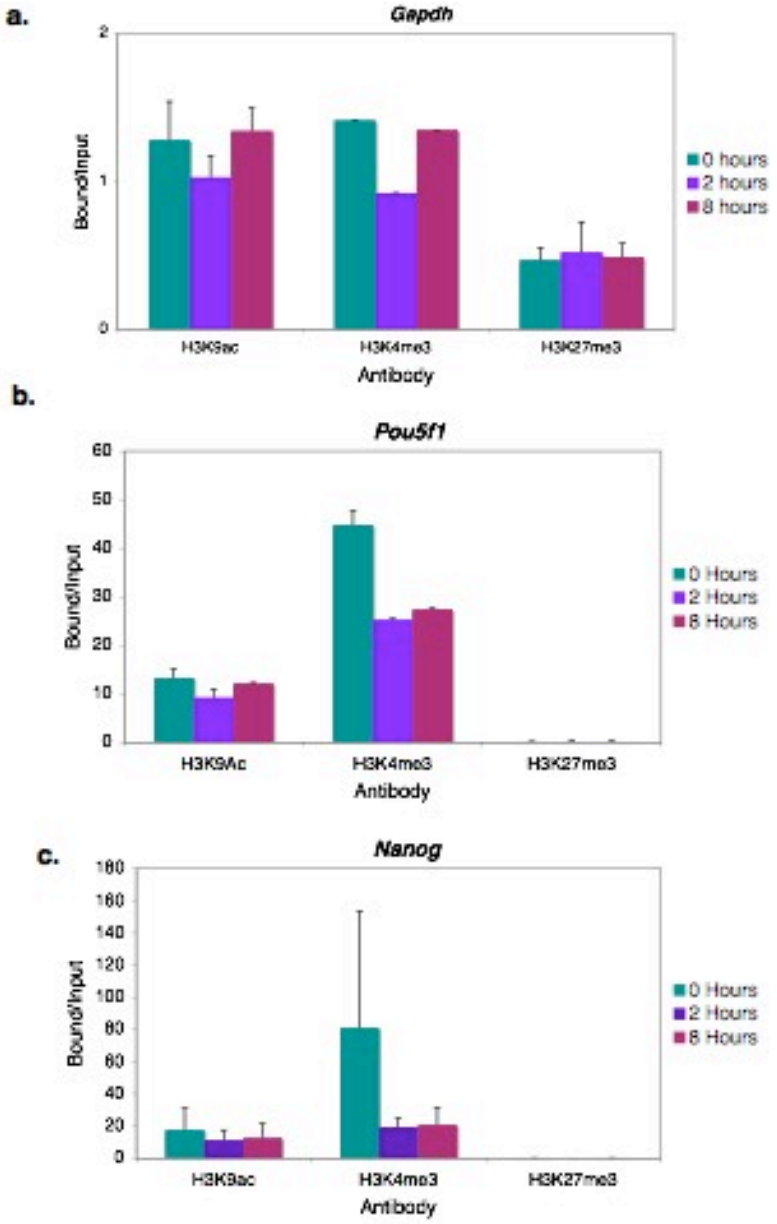


Figure 3.7 The effects of sodium valproate on histone modifications at *Gapdh*, *Nanog* and *Pou5f1* promoters. Undifferentiated CCE/R cells were treated with 1mM valproate for 2 and 8 hours, chromatin extracted at each timepoint and immunoprecipitated with antibodies against H3K9ac, H3K4me3 and H3K27me3. The amount of **a.** GAPDH **b.** *Pou5f1* and **c.** *Nanog* promoter DNA in the bound and input fractions was analyzed by quantitative real time PCR and a ratio generated. PCR was performed in triplicate from two ChIP experiments. Error bars represent the standard error of the mean.

of less than one at this promoter which indicates hypomethylation. This is unsurprising since it is a repressive modification and *Gapdh* is actively transcribed. Its level is not altered by treatment with valproate, coordinate with global levels (Figure 3.6, 3.7b).

In contrast to *Gapdh*, the *Pou5f1* promoter displays highly significant levels of both H3K9ac and H3K4me3 activating modifications (Figure 3.7b). The levels of H3K9ac do not change significantly upon treatment of valproate. Like *Gapdh,* this is contrary to the global increase seen in levels of this modification (Figure 3.6). The levels of H3K4me3 do change in response to valproate, however in a manner contrary to the global trend of increase, this modification decreases at the promoter upon treatment (Figure 3.7). The decrease is to around half the original untreated levels and is apparent after 2 hours of treatment with valproate. There is no enrichment for H3K27me3 at this promoter. Furthermore, the levels of this inactive modification do not change in response to valproate.

Like the *Pou5f1* promoter, the *Nanog* promoter contains highly significant levels of both H3K9ac and H3K4me3, the activating modifications (Figure 3.7c). The level of H3K9ac is refractory to valproate treatment. However, the level of H3K4me3, like that seen at the *Pou5f1* promoter plummets significantly after 2 hours of treatment (Figure 3.7c). In contrast, there is very little inactive H3K27me3 modification at this promoter, and it does not change upon valproate treatment.

In contrast to the observations seen at *Gapdh* and *Pou5f1*, valproate induces substantial changes to the histone modifications on the *Hoxb* cluster (Figure 3.8). This is most clear for H3K9 acetylation (Figure 3.8a). After the prolonged 8 hours treatment, there is a consistent increase in the level of H3K9ac of about 2-fold at all *Hoxb* gene promoters. This is preceded by a small increase in H3K9ac after 2 hours

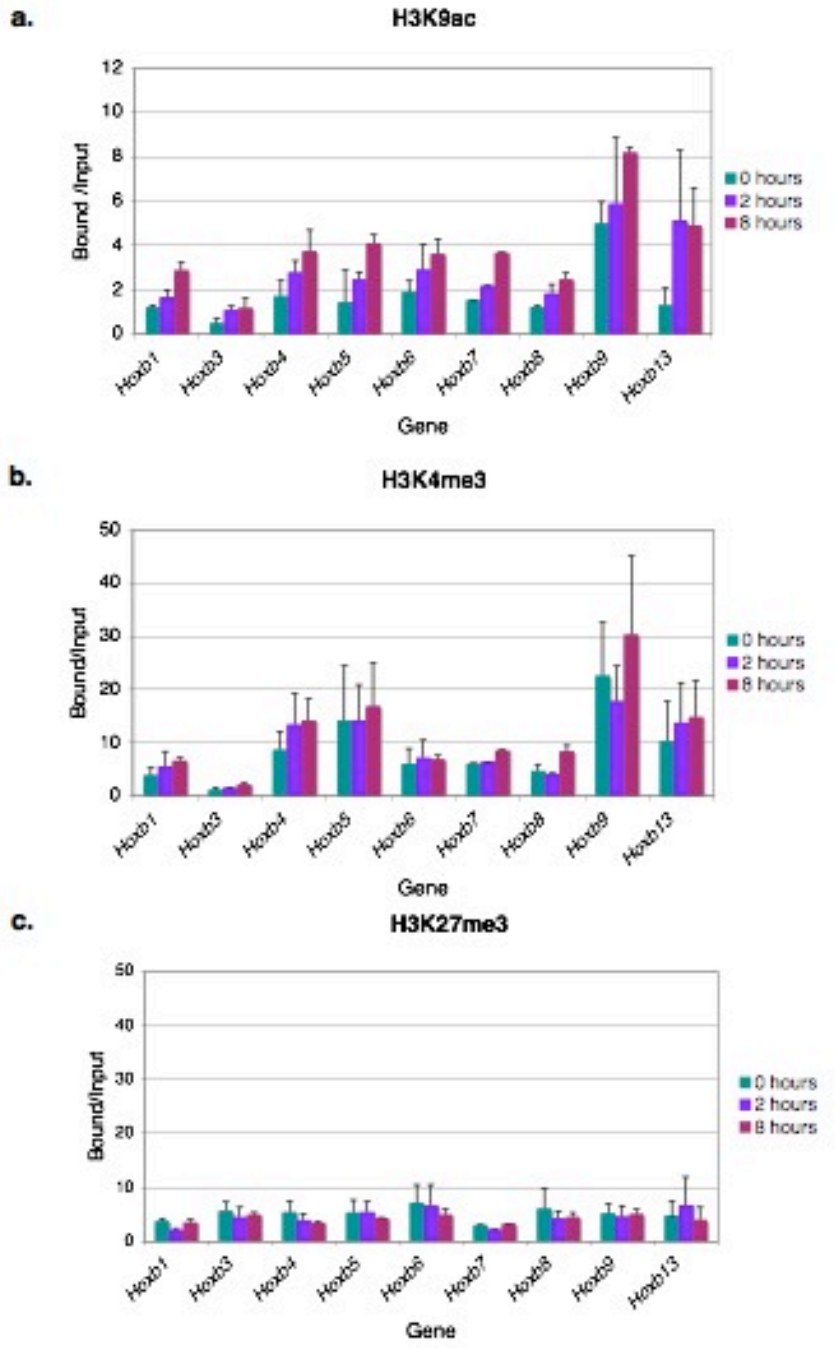


Figure 3.8 The effects of sodium valproate on histone modifications at *Hoxb* promoters. Undifferentiated CCE/R cells were treated with 1mM valproate for 2 and 8 hours. Chromatin was extracted at each timepoint and immunoprecipitated with antibodies against **a.** H3K9ac, **b.** H3K4me3 and **c.** H3K27me3. The amount of DNA in the bound and input fractions was analyzed by quantitative real time PCR and a ratio generated. PCR was performed in triplicate from two ChIP experiments. Error bars represent the standard error of the mean

of treatment, resulting in an overall steady increase in H3K9 acetylation at the promoters. Since this experiment was performed twice at each gene, it is not valid to perfom a statistical test of significance on these values, however the patterns of increase at each *HoxB* gene was highly repeatable, and error bars representing the standard error of the mean have been included to represent this. Separate experiments performed in chapter 4 also support this pattern of increase in histone acetylation after valproate treatment. Interestingly, since all the *Hoxb* genes respond to valproate in a similar manner with respect to this modification, the overall pattern of H3K9ac along the cluster is maintained. Therefore, upon treatment with valproate, *Hoxb9* remains the gene that is most enriched for H3K9ac at its promoter and *Hoxb3* the gene with the lowest levels.

The level of H3K4 trimethylation at the promoters of the *Hoxb* genes also increases in response to valproate and thus the two activating modifications H3K9ac and H3K4me3 are linked at the *HoxB* promoters (Figure 3.8b). The H3K4me3 hypermethylation seen upon treatment is less pronounced than the H3K9 hyperacetylation. In Figure 3.8b it is also apparent that valproate treatment leaves the relative pattern of H3K4me3 levels unaltered, with *Hoxb9* still displaying the highest level of H3K4me3 and *Hoxb3* much the lowest.

The repressive modification H3K27me3 displays a pattern of modifications quite different to the activating modifications in response to valproate as the levels of this modification do not change upon valproate treatment (Figure 3.8c). This is despite the increases in the active modifications H3K4me3 and H3K9ac. Thus, enrichments for active modifications in response to valproate are highlighted by this repressive modification's refractory behaviour.

3.2.4 Are changes in histone modifications at the promoters of *Hoxb* genes linked to changes in gene expression in embryonic stem cells?

The role that histone modifications play in direct control of the expression of *Hoxb* genes is unknown. Using the histone deacetylase inhibitor valproate, a system has been established in which the promoters of *Hoxb* genes may be enriched for the active modifications H3K9ac and H3K4me3 in the presence of the repressive H3K27me3 modification. Gene expression from the promoters of active modification enriched genes was examined in this system. This was an attempt to see if the active marks could override the repressive marks in relation to gene expression control.

In order to assess the direct relationship between histone modifications and transcription at the *Hoxb* gene cluster, RNA was extracted from the same populations of cells that were treated with 1mM valproate for 2 and 8 hours in the presence of LIF for the ChIP experiments. cDNA was generated from this RNA and the levels of the genes in the cDNA assessed by quantitative real time PCR, normalised to actin and compared with levels in untreated cells.

As above, it was difficult to amplify PCR products from the cDNA for these genes in these undifferentiated cells as they are not expressed. Products were often only being amplified at a cycle number greater than 35, if at all and triplicates occurred with little repeatability. This was taken as a marker that no significant expression from these loci occurred when compared with the consistent amplification and repeatability of actin amplification from the same samples. In fact, even after 8 hours of treatment with valproate, no significant gene expression was induced (data not shown) from any of the *Hoxb* gene promoters examined.

3.3 What is the role of GCN5, a histone acetyl transferase, in the response to valproate at *Hox* genes?

At the *Hoxc* cluster, evidence for a possible direct role for histone acetylation in the control of *Hox* gene transcription has been found. Mouse embryos homozygous for the histone acetyl transferase GCN5 flox (neo) mutation, display anterior homeotic mutations that coincide with a shift in the anterior expression boundary of *Hoxc8* and *hoxc9* (Lin et al, 2008). This may be the direct result of an alteration of histone acetylation levels at these gene promoters. The role that GCN5 related histone acetylation has at these gene promoters was therefore investigated in a GCN5 -/- mouse ES cell line, a gift from Sharon Dent (Lin et al, 2007). GCN5 knock-out cells were grown off feeders and grew faster than the equivalent wild type line. These wild-type cells were kept on feeders due to their tendency to spontaneously differentiate. This discrepancy is between cells growing on feeders is not ideal as it makes the 'control' different from the genetically alterered lines and may be the basis of any difference in response to valproate. However, due to the amount of cells necessary for the experiments, this compromise was made. Before any experiments were performed on these cells, an alkaline phosphatase assay was carried out in order to check for pluripotency. Both sets of cells stained positively for alkaline phosphatase (Figure 3.9).

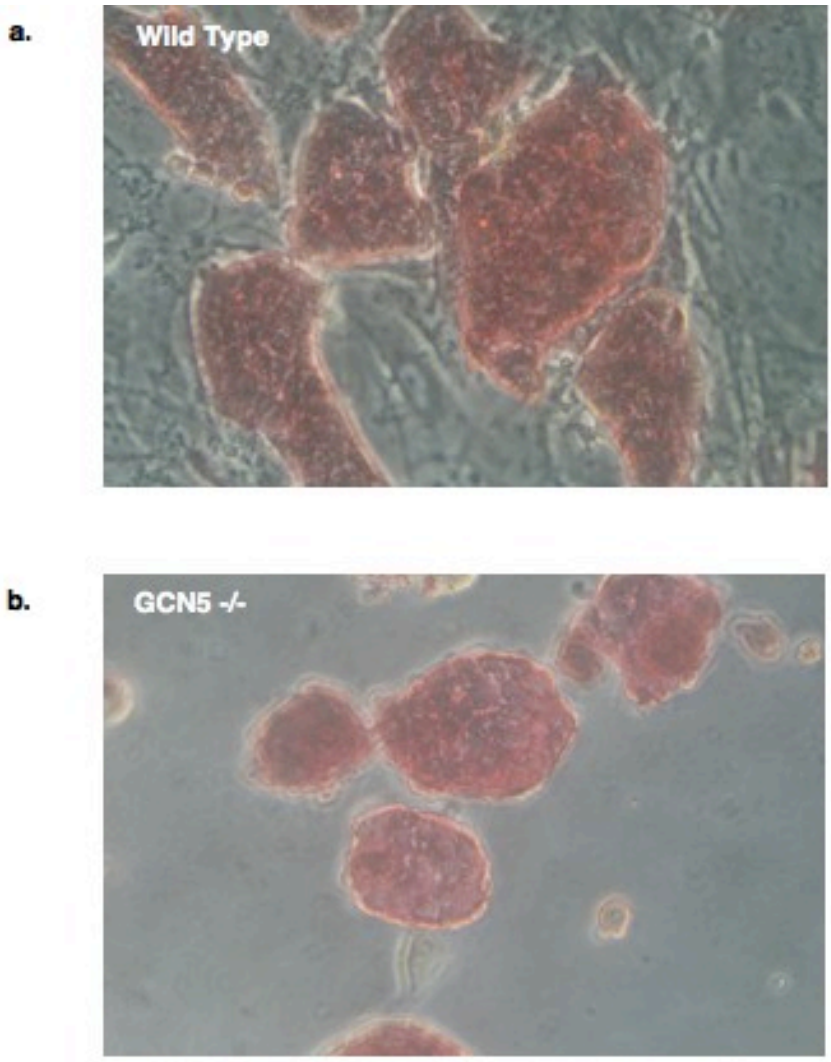


Figure 3.9 3G4 cells maintain an undifferentiated phenotype in the presence and absence of GCN5. a. Wild type and **b.** GCN5$^{-/-}$ cells were cultured in the presence of LIF and alkaline phosphatase assays were performed. The staining of cells was examined using a Zeiss Axiovert 25 and images captured using a Canon Digital EOS 400D

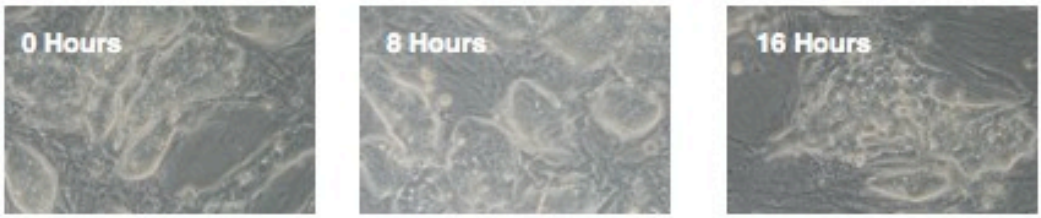
3.3.1 Are there global differences in histone acetylation in GCN5 -/- cells?

Before focusing on the *Hoxc* genes, global differences in histone modifications in wild type and GCN5 -/- ES cells were examined. These differences would only be apparent if histone acetyl transferases were unable to compensate for one another. In order to assess whether there are any global differences in histone acetylation, histones were extracted from these cells by acid extraction and the histones were ran on AUT gels which distinguish histones based upon their size and charge, allowing histones with different levels of acetylation to be distinguished. Since valproate is an HDAC inhibitor, treating with this drug will further highlight whether GCN5 is important for maintaining global levels of histone acetylation. If there is no enzyme that may compensate for GCN5 then the GCN5-/- cells will not display elevated levels of histone acetylation in response to treatment with valproate.

The morphological effect that valproate has on both the wild type and the knock-out lines is shown in Figure 3.10. It can be seen that as for the CCE/R cells, there is an induction of a spiny morphology in the knock out 3G4 line after 8 and 16 hours treatment (Figure 3.10b). However, such a change is not seen in the wild-type line (Figure 3.10a). This is probably because the wild type line is grown on feeders and so subtle changes in wild type colonies cannot be easily discerned.

The AUT gels show that there is no difference in histone acetylation between resting wild type and GCN5 -/- cells (Figure 3.11). Upon treatment with valproate for 8 hours, there is an increase in global hyperacetylation at both Histone H3 and H4 in wild type and GCN5-/- cells. Clearly, mono, di, tri and tetra acetylation of H4 is seen to increase upon treatment with valproate in both wild type and knock out cells. This increase is maintained in both cell types after 16 hours of treatment with 1mM

a.



b. GCN5 -/-

Figure 3.10 The morphology of a. wild type and b. knock out 3G4 cells upon valproate treatment. Cells were treated for 8 hours and 16 hours with 1mM valproate as marked. Morphology of cells was examined using a Zeiss Axiovert 25 and images captured using a Canon Digital EOS 400D

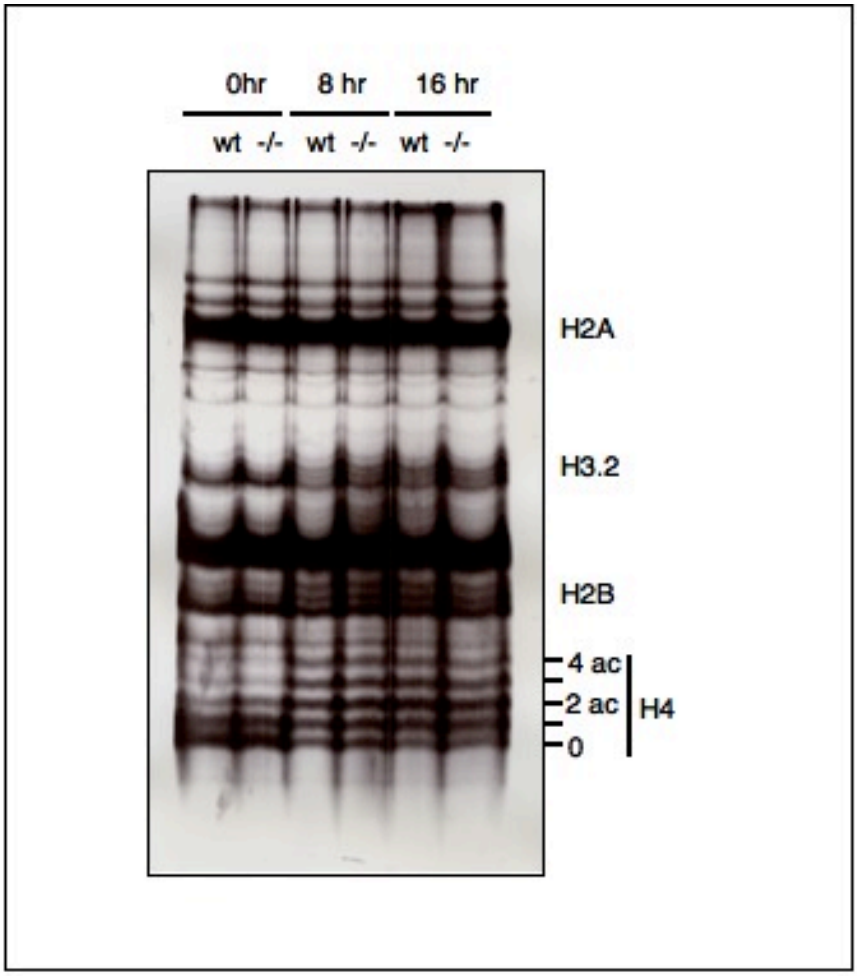


Figure 3.11. The effects of sodium valproate on global histone modifications in GCN5 knock out and wild type ES cells. 3G4 wild type and knock-out cells were treated with 1mM valproate for 8 and 16 hours and histones extracted by acid extraction. Equal amounts of protein were loaded onto AUT gels as marked. Gels were performed by K. P. Nightingale.

valproate. Thus GCN5-/- may be compensated for in order to maintain global levels of histone acetylation in GCN5 -/- cells, and this compensatory mechanism is still active in cells treated with 1mM valproate.

3.3.2 What is the role of GCN5 in histone hyperacetylation at the *Hoxc* cluster?

In order to discern the cause of *Hoxc* mis-expression in GCN5 -/- early embryos, ChIP assays were performed on GCN5 null and wild-type 3G4 ES cells. In addition ChIP assays were performed on valproate treated wild-type and knock-out cells. If GCN5 does, indeed play a significant role at these gene promoters, then there will be a less pronounced hyperacetylation in response to valproate. These experiments were performed once only as the cells were a technically difficult line to grow in large numbers.

To confirm that 3G4 cells are able to respond to valproate treatment, its effects upon the histone modifications at the *Hoxc8* and*Hoxc9* promoters were first examined in the wild-type line (Figure 3.12a). At these promoters, an increase in H3K9 acetylation was evident (Figure 3.12 a(i)). At both *Hoxc8* and *Hoxc9* the fold increase reached its maximum after 8 hours with an enrichment of roughly twice that seen in untreated cells. Immunoprecipitations with anti-H3K4me3 reveals that there was a rise in the level of this active modification at the *Hoxc9* but not the *Hoxc8* promoter (Figure 3.12 a(ii)). This increase is lower than that seen with H3K9ac at the *Hoxc9* promoter. Such an increase is, however, consistent with those at the *Hoxb* in CCE/R cells in response to valproate. In contrast with the *Hoxb* cluster in CCE/R cells, this increase is accompanied by an increase in H3K27me3 levels at *Hox*c9 (Figure 3.12 a (iii)). At *Hox*c9 after 16 hours of treatment with valproate the levels of H3K27me3 are twice that seen in untreated cells, though at *Hox*c8 the levels of this modification remain unchanged. Thus, at *Hox*c9, the modifications H3K9ac, H3K4me3, and H3K27me3

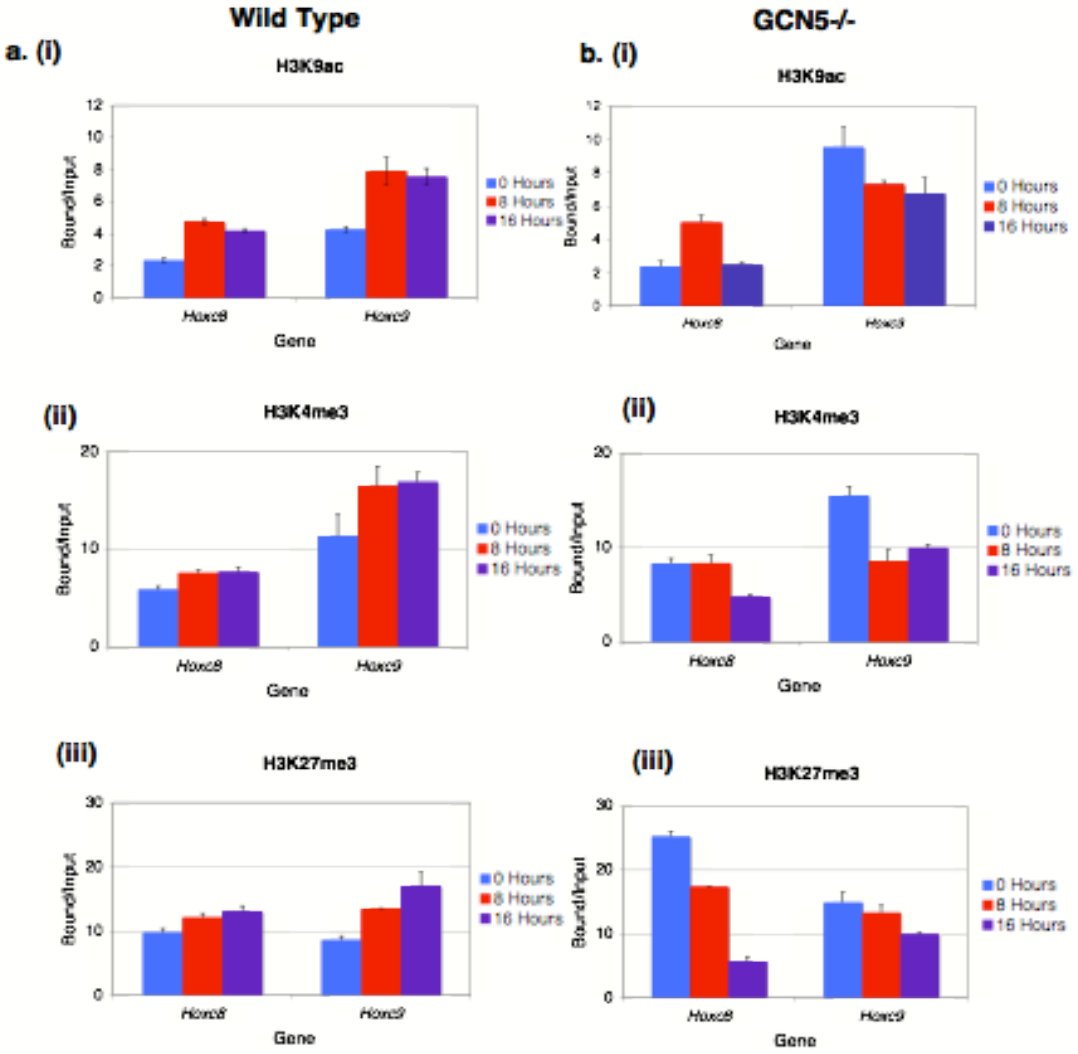


Figure 3.12 The distribution of histone modifications over valproate treated *Hoxc8* and *Hoxc9* promoters. a. Wild type, or **b.** GCN5 -/-. 3G4 cells were treated with 1mM valproate for 8 or 16 hours. Chromatin was extracted and immuno precipitated using antibodies directed against H3K9ac, H3K4me3 and H3K27me3 as marked. The levels in the bound and input fractions were determined by quantitative real time PCR in triplicate. Error bars represent the standard error of the mean of PCR triplicates.

are all linked and are enriched in valproate treated cells compared with resting cells. At *Hox*c8, however, only the H3K9ac modification responds to valproate. The response to valproate is very different in knock-out cells compared with wild- type 3G4 (Figure 3.12b). After treatment with valproate, the *Hox*c9 promoter shows a very slight decrease in the H3K9ac modification, contrasting with the increase seen in wild-type 3G4 (Figure 3.12 b(i)). The *Hox*c8 promoter displays an unexpected pattern of modifications; after 8 hours treatment, it shows an increase in its levels of H3K9ac. Subsequently, after 16hours of treatment, this increase is not maintained (Figure 3.12b(i)). The H3K4me3 mark at this gene does not follow a similar pattern, and, instead maintains similar levels of modifications in untreated cells, and cells treated for eight hours, followed by a decrease at 16 hours (Figure 3.12b(ii)). *Hox*c9 shows an earlier decrease in this modification at 8 hours, which is maintained at 16 hours. The repressive modification, H3K27me3 also shows a decrease at the *Hox*c8 and *Hox*c9 promoters upon treatment with valproate (Figure 3.12 b(iii). The *Hox*c8 decrease is most significant with a decrease of five times that of the initial resting level after 16 hours of treatment. The *Hox*c9 decrease is not as large but is still significant.

To ascertain any direct role that GCN5 plays in the acetylation differences seen in the wild-type and GCN5 knock-out 3G4 cells, X-ChIP was performed using an anti-GCN5 antibody and an anti-pol II antibody as a control. It was hypothesised that the GCN5 enzyme would be located at the promoters of both *Hox*c8 and *Hox*c9 in wild type cells but not in the GCN5 -/- cells. However, despite the use of anti-GCN5 antibodies from both Abcam and Santa Cruz, precipitated DNA could not be obtained using the anti-GCN5 antibody (data not shown).

3.3.3 Does GCN5 play an important role at the *Hoxb* cluster?

At the promoters of the two genes of the *Hoxc* cluster studied above, increases in histone acetylation in response to valproate were impaired in GCN5 null cells. The GCN5 -/- 3G4 cells were subsequently used to assess the importance of this enzyme at the *Hoxb* locus in order to determine whether GCN5 also plays a role in valproate response at this locus.

To assess the role that GCN5 plays at the promoters of the *Hoxb* cluster, quantitative real time PCR analysis with primers to the *Hoxb* gene promoters was performed on ChIP experiments using valproate treated wild-type and GCN5 -/- cells. It has been shown that treatment of the CCE/R mouse cell line with 1mM valproate resulted in H3K9 hyperacetylation at the *Hoxb* locus. If GCN5 is, indeed, important in depositing the histone H3K9acetyl mark at the promoters of the *Hoxb* gene cluster then in the GCN5 null cell line there will be a lack of histone hyperacetylation in response to valproate.

The levels of the H3K9ac modification do not appear to have been severely affected by the knock-out of the GCN5 enzyme (Figure 3.12). At the promoters of the *Hoxb* genes, alteration of the levels of H3K9ac in response to valproate follow a similar pattern in untreated wild type and GCN5 null 3G4 cells (Figure 3.12 a(i), b(i)). In addition, the numerical values of resting levels are not significantly altered by knock-out. The one exception is *Hoxb5,* which has five times the level of H3K9ac in its promoter in wild type cells compared with GCN5 null. The patterns of response to valproate at the promoters of *Hoxb3, b4, b6, b7, b8* and *b9* are near identical, and thus at these promoters, it is evident that the loss of GCN5 has not altered the acetylation/deacetylation equilibrium. At the *Hoxb1* promoter, the level of H3K9 acetylation is increased after 8 hours and further increased after 16 hours. In the knock-out cells, there is again an increase after eight hours, however, after a further

eight hours of treatment, this level has decreased back to that found at the resting level. It may be therefore that at *Hoxb1*, GCN5 is needed to maintain the level of hyperacetylation during longer treatment. At the promoter of *Hoxb5*, surprisingly,there is more H3K9 hyperacetylation in the GCN5 null cells in response to valproate than in the wild type cells. It is possible that this is because the histone acetylase that compensates for the loss of GCN5 at this gene has a higher turnover, and so when the deacetyalse is inhibited in the knock-out cells, there is a higher level of hyper-acetylation in these cells. Finally, at *Hoxb13* the GCN5 -/- cells are unable to respond as well to valproate as the wild type cells. Overall, at the *hoxb* cluster though, there is little change in acetylation and hyperacetylation levels in response to valproate.

As for the levels of H3K9 acetylation, the levels of H3K4me3 methylation are not altered significantly in resting GCN5 null ES cells in comparison to wild type cells (Figure 3.12 a(ii), b(ii)). This is with the exception of the *Hoxb6* promoter that shows an approximately 2 fold increase in the levels of H3K4me3 in the wild type than in the GCN5 null cells. The responses of the gene promoters to treatment with valproate also display a similar pattern in the wild type and null cells. Surprisingly, there is no increase in H3K4me3 at the promoters in either of these cells, unlike in CCE/R cells where treatment with valproate results in an enrichment for H3K4me3 at *Hoxb* gene promoters. However, it is does appear that there is a link in the levels of H3K9ac and H3K4me3 marks, at least in the wild type cells as there was in the CCE/R cells. Contrary to the expected increase, there is, however, a decrease in H3K4me3 at several gene promoters. The two notable exceptions are *Hoxb1* where the level of this modification increases in null cells upon 16 hours treatment with valproate but not in wild type cells, and *Hoxb3* where H3K4me3 levels are increased in wild type cells but not in knock out cells. Overall, any differences in H3K4me3 at the promoters

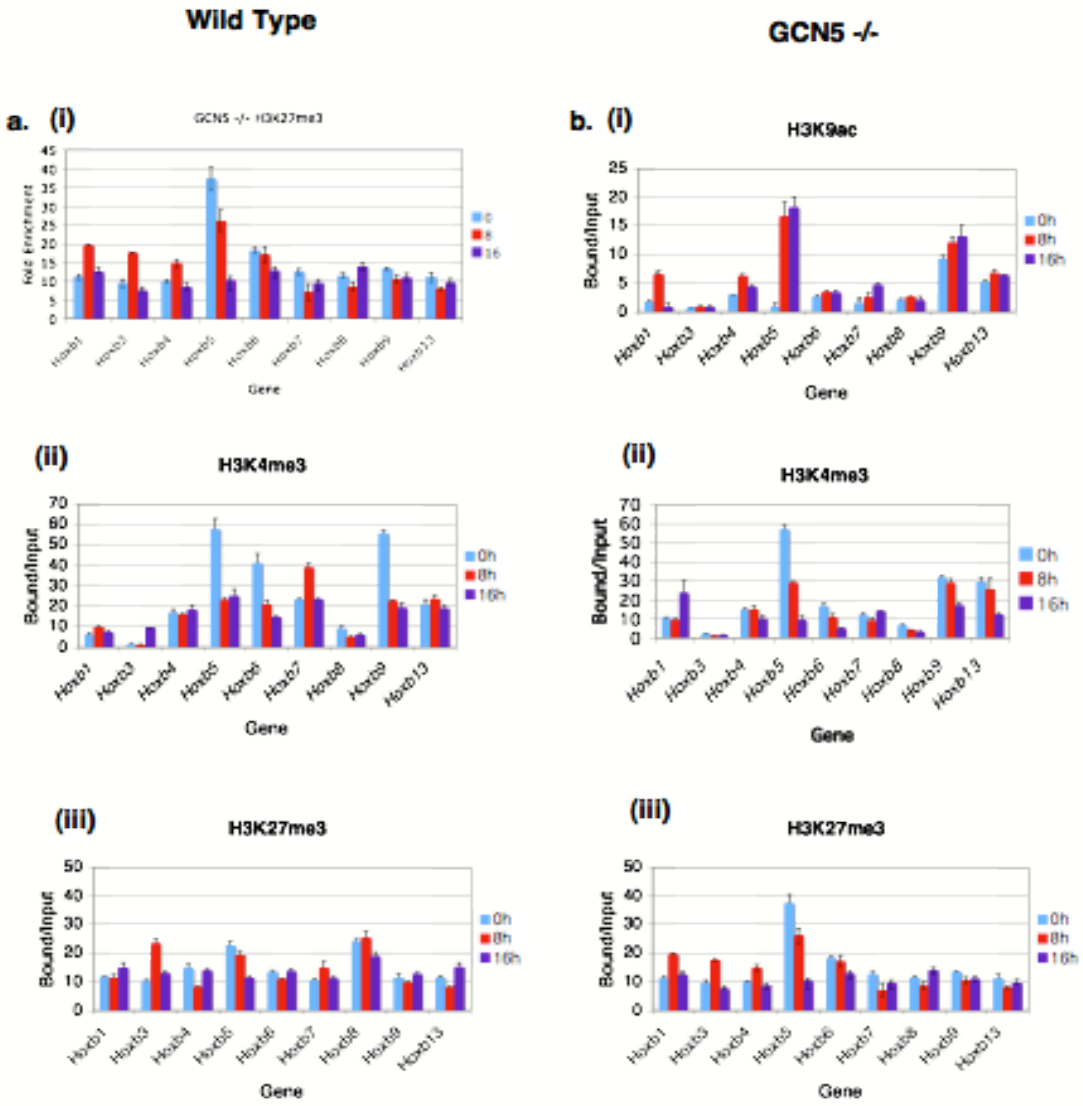


Figure 3.13 The distribution of histone modifications over valproate treated *Hoxb* promoters. a. Wild type, **or b.** GCN5 -/-. 3G4 cells were treated with 1mM valproate for 8 or 16 hours. Chromatin was extracted and immunoprecipitated using antibodies directed against H3K9ac, H3K4me3 and H3K27me3. The levels in the bound and input fractions were determined by quantitative real time PCR. Error bars represent the standard error of PCR triplicates.

of the *Hoxb* genes in the two cell types are slight, thus the mechanisms for deposition and maintenance of H3K4 tri-methylation are maintained in GCN5 null cells.

The levels of H3K27me3 at the promoters of the *Hoxb* genes are very similar between wild type and knock-out cells (Figure 3.12 a(iii), b(iii)). There is also a very similar pattern of change in levels in response to valproate with any differences being subtle between to the two types of cells.

Thus, it can be concluded that the presence or absence of the enzyme GCN5 makes little difference to the levels of H3K4me3, H3K27me3 and H3K9ac at the promoters of the genes of the *HoxB* cluster.

3.3.4 Do histone modification changes result in gene expression changes in 3G4 cells?

Although there were subtle differences in the changes in the histone modifications H3K4me3 and H3K9ac in 3G4 wild type and GCN5 -/- ES cells upon treatment with valproate, it was decided to extract RNA and see if there is any change in induction of these genes. cDNA was synthesised from the RNA extracted and changes in mRNA levels quantified as for CCE/R cells using quantitative real time PCR. As for the treatment of CCE/R cells, there was no induction of transcription in the 3G4 cells after valproate treatment for *Hoxb* and *Hoxc* genes studied (data not shown)

4 Memory of histone modifications at the *Hoxb* cluster

4.1 Are changes in histone modifications remembered at the *Hoxb* cluster?

In order to assess whether histone modifications at the promoters of the *Hoxb* locus are of a wider importance in memory at these promoters, cells were treated with valproate to induce the histone modification changes previously demonstrated (See Figure 3.8). However, in the following experiments, not only were the immediate effects of valproate upon *Hoxb* genes assessed, but also the potential persistence of induced histone modification changes was investigated by performing ChIP 48 hours after the initial valproate treatment (Figure 4.1a). This system was intended to investigate the epigenetic memory of histone modifications at the locus.

4.1.1 Are histone modifications inherited at the *Hoxb* cluster?

In order to investigate any possible epigenetic significance of histone modifications at the *Hoxb* cluster, CCE/R cells were treated for 8 hours with valproate. After this timepoint, the cells were harvested, chromatin extracted and immunoprecipitated using antibodies against H3K9ac, H3K4me3 and H3K27me3 (Figure 4.1a). In addition, another population of cells was valproate treated for 8 hours, the valproate washed-out of the tissue culture medium and the cells cultured for a further 48 hours whereupon ChIP was performed (Figure 4.1a). Untreated populations of cells were harvested concomitantly at 8 hours of valproate treatment and at washout. Their culture was exactly the same as the treated cells and so were considered a 'control' population. This allows a direct comparison of the levels of histone modifications in the treated cells with an equivalently cultured cell population.

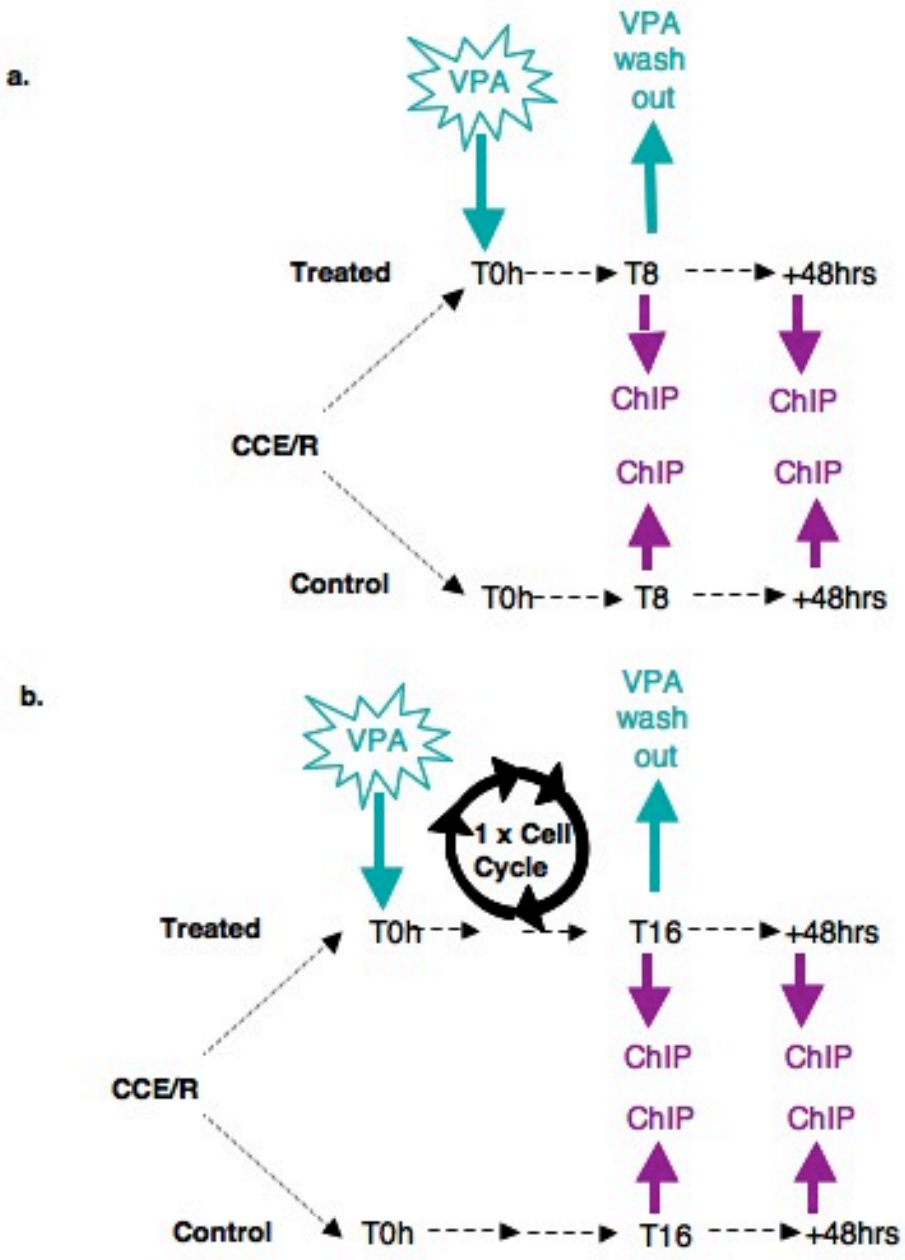


Figure 4.1. Schematic of valproate wash-out experiments.a. Undifferentiated ES cells were treated with 1mM valproate for 8 hours, the valproate removed from the culture medium and cells cultured for a further 48 hours. Control untreated cells were harvested at each timepoint. **b.**Undifferentiated ES cells were treated with 1mM valproate for 16 hours, the valproate removed from the culture medium and cells culture for a further 48 hours. Control untreated cells were harvested at each timepoint. Chromatin immunoprecipitation was performed at the time points indicated.

.

Thus any effects of cell culture on histone modification were controlled, and, in the analyses of this experiment, bound:input ratios were normalised to those of the co-ordinately harvested control cell population.

It is strikingly clear from the fold enrichments obtained for the ChIP experiments using the H3K9ac antibody that there is no memory of the induced increase in H3K9 acetylation at *Hoxb* promoters 48 hours after valproate treatment (Figure 4.2a). As displayed in previous experiments, there is a consistent increase in acetylation at all the promoters examined after treatment with valproate. This hyperacetylation then, for every gene, returns to levels equivalent to those in the respective untreated cells. Thus, no epigenetic memory is displayed at any of the *Hoxb* promoters, with all the genes behaving in the same manner.

The H3K4me3 modification shows a similar story to that seen for the H3K9ac modification; thus the two activating modifications echo one another in their epigenetic behaviour (Figure 4.2b). Upon treatment with valproate, the promoters of the genes of the *Hoxb* cluster are hypermethylated; this is exactly as that seen at these promoters previously. Then, after 48 hours in culture without valproate, the *Hoxb* promoters do not retain previous H3K4-hypermethylation. This pattern is displayed at every gene promoter of the *Hoxb* cluster studied and thus the treated cells show ratios of enrichment in H3K4me3 of one when compared with equivalent untreated cells 48 hours after treatment.

The H3K27me3 modification displays a different pattern of enrichment to the two active modifications (Figure 4.2c). Again, its levels at the *Hoxb* gene promoters are largely refractory to the effects of valproate. This is shown in the fold enrichment values for this modification, which do not stray far from one for all the genes at all

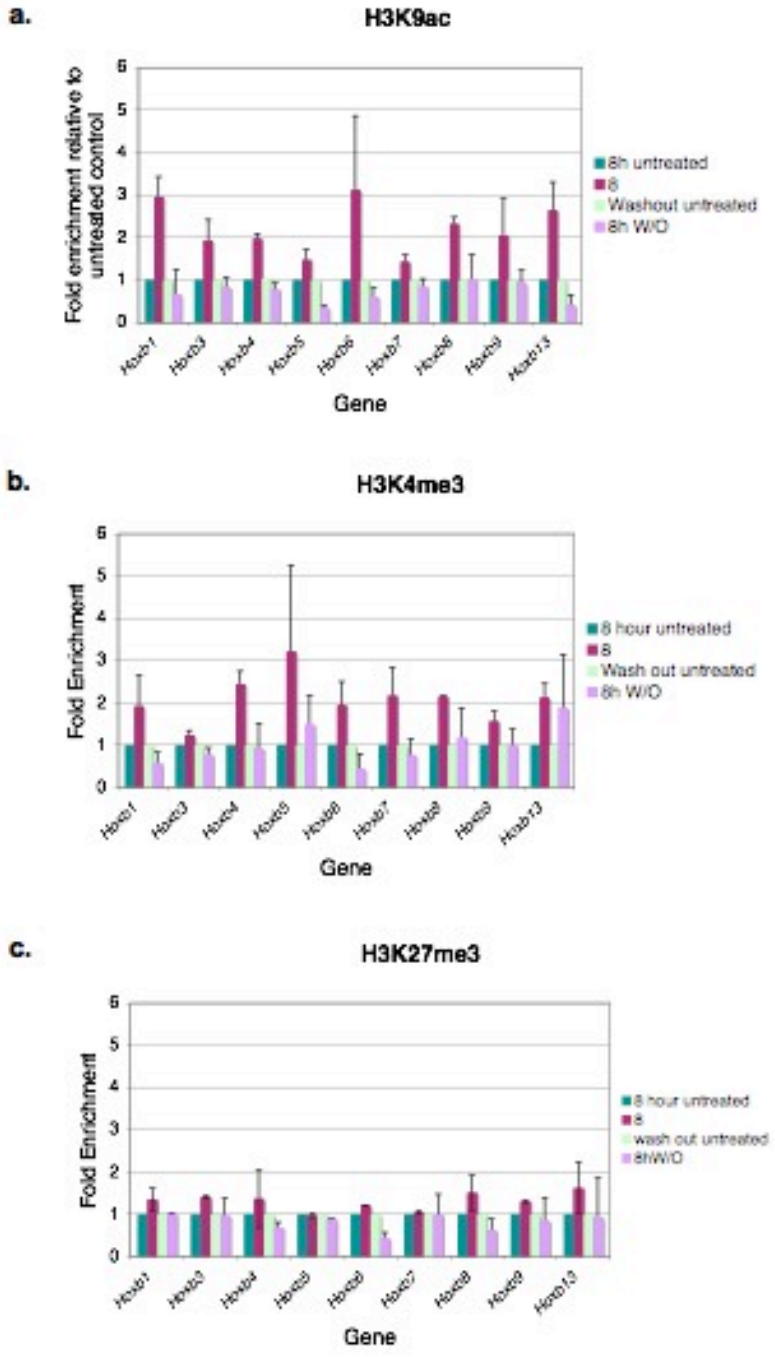


Figure 4.2 The effects of an 8 hour valproate treatment and wash out on histone modifications of the *Hoxb* cluster in undifferentiated cells. Undifferentiated CCE/R cells were treated with 1mM valproate for 8 hours and the valproate washed out. A population of equivalent untreated control cells was harvested at each timepoint. Chromatin was extracted at each timepoint and immunoprecipitated with antibodies against **a.** H3K9ac, **b.** H3K4me3 and **c.** H3K27me3. The amount of DNA in the bound and input fractions was analyzed by quantitative real time PCR, a ratio generated and normalised to the respective untreated. PCR was performed in triplicate from two ChIP experiments. Error bars represent the standard error of the mean

timepoints studied. Interestingly, after treatment and wash-out, levels of this modification in the treated cells at *Hoxb6* and *Hoxb8* are less than those in control cells.

4.1.2 Does valproate treatment result in delayed transcription induction from the *Hoxb* cluster?

In order to assess whether the previous induction in histone modifications caused expression changes from the *Hoxb* cluster after further cellular growth, RNA was extracted at timepoints equivalent to those in the ChIP experiments. No significant gene expression changes were seen from the *HoxB* cluster throughout the duration of the experiment as, again there was sporadic amplification from the cDNA, and a typical example is given on page 3. Similar plots were obtained for *Hoxb1,b5* and b9 (data not shown) thus it can be concluded that valproate treatment does not induce later gene expression changes from *Hoxb* genes.

4.1.3 Does increasing valproate treatment to one complete cell cycle result in result in fixed histone modification changes?

After 8 hours of valproate treatment increases in H3K9ac or H3K4me3 levels at *Hoxb* promoters are not retained. However, it is possible that for a memory of a histone modification to be fixed, the cell must go through either S phase or mitosis where mechanisms exist that are important for maintaining histone modification patterns throughout the cell cycle. Once the cell has passed through these phases, the histone mark may be consolidated and further remembered by the cell. To assess if this is indeed the case, CCE/R cells were treated with valproate for an extended period of 16 hours so that a full cell cycle was completed during the time of treatment (Figure 4.3b) The valproate was then washed out as for the 8 hours treatment and

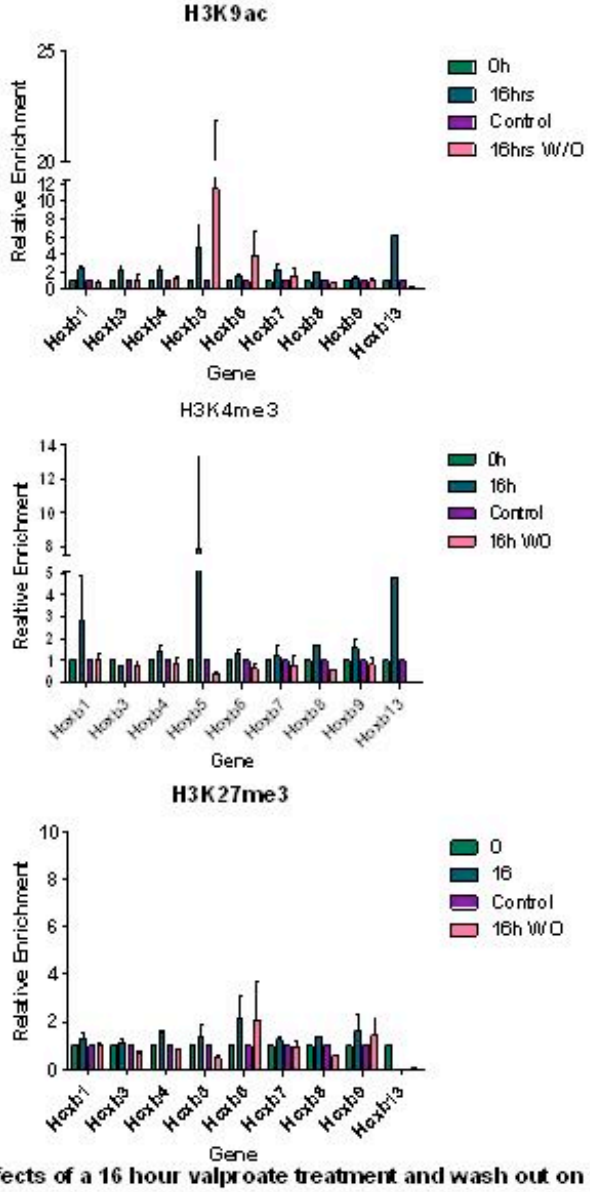


Figure 4.3 The effects of a 16 hour valproate treatment and wash out on histone modifications of the *Hoxb* cluster in undifferentiated cells. Undifferentiated CCE/R cells were treated with 1mM valproate for 16 hours and the valproate washed out. A population of equivalent untreated control cells was harvested at each timepoint. Chromatin was extracted at each timepoint and immunoprecipitated with antibodies against **a.** H3K9ac, **b.** H3K4me3 and **c.** H3K27me3. The amount of DNA in the bound and input fractions was analyzed by quantitative real time PCR, a ratio generated and normalised to the respective control. PCR was performed in triplicate from two ChIP experiments. Error bars represent the standard error of the mean

the levels of histone modifications at *Hoxb* gene promoters assessed. Despite the prolonged valproate treatment, the H3K9ac mark shows a similar pattern of hyperacetylation to that seen upon 8 hours treatment with valproate (Figure 4.3 a). That is, after 16 hours of exposure to valproate, there is a significant increase in H3K9 acetylation consistent across the two experiments performed, however, no cellular memory of this hyperacetylation remains after 48 hours culture without valproate. In one replicate, *Hoxb5* and *Hoxb6* did, however, show a maintenance of the increase in H3K9 acetylation. However, this was not repeatable and is reflected in the large error bar seen in the wash-out data. It could therefore be that something different occurred in this one experiment that resulted in the cementing of a memory. Generally, however, no memory of hyperacetylation was seen, despite completion of a full cell cycle with this hyperacetylation present.

The active modification H3K4me3 echoes the changes seen with H3K9ac (Figure 4.3b). However, it does not show as large an increase in its levels at the *Hoxb* promoters after 16 hours of treatment as does H3K9ac and does not appear to have maintained the levels seen after 8 hours treatment (Figure 4.2). There is, however, an increase in the levels of H3K4me3 at all gene promoters studied, except for *Hoxb3* whose levels of H3K4me3 have been consistently less responsive to valproate than other genes. After the increase at 16 hours, the levels of H3K4me3 at the promoters of the *HoxB* genes then return to levels equivalent to those in the respective control cells. Therefore, for this activating modification, there is no cellular memory displayed.

The response of H3K27me3 levels at *hoxb* promoters to valproate is contrary to that of the activating modifications (Figure 4.3c). No initial enrichment in response to valproate is displayed despite changes seen in activating modification levels. Thus, fold enrichments for this modification at these genes remained at one.

4.1.4 Does prolonged full cell cycle treatment with valproate result in altered transcription changes from the *Hoxb* cluster?

RNA was extracted concomitant with chromatin extraction. Again, amplification from the cDNA of the genes studied was sporadic and hence no real gene expression changes were seen from representative genes of the *Hoxb* gene cluster throughout the period of the experiment.

4.2 Are histone modification changes of epigenetic significance upon differentiation?

Induction of gene expression requires many signals, of which, histone modifications are just one layer. The induced increase in activating histone modifications at *Hoxb* gene promoters in undifferentiated ES cells were not sufficient to induce gene expression from the *Hoxb* cluster. However, *Hoxb* genes were shown to be induced in ES cells upon differentiation, therefore the onset of differentiation must provide signals in addition to the activating histone modifications associated with transcriptional activity. Thus, could induced increases in active histone modifications in undifferentiated ES cells cause premature transcription from *Hoxb* loci in differentiating cells if these modifications overlapped with the signal to differentiate? CCE/R cells were therefore incubated for 8 hours in valproate (a timepoint where it has been shown that these genes are hyperacetylate) the cells induced to differentiate and then treated for a further 8 hours in valproate before its removal (Figure 4.4). In this manner, the cells are hyperacetylated whilst they receive the differentiation signal, though hyperacetylation does not remain throughout the experiment. Any epigenetic significance of hyperacetylation at the *Hoxb* will reveal

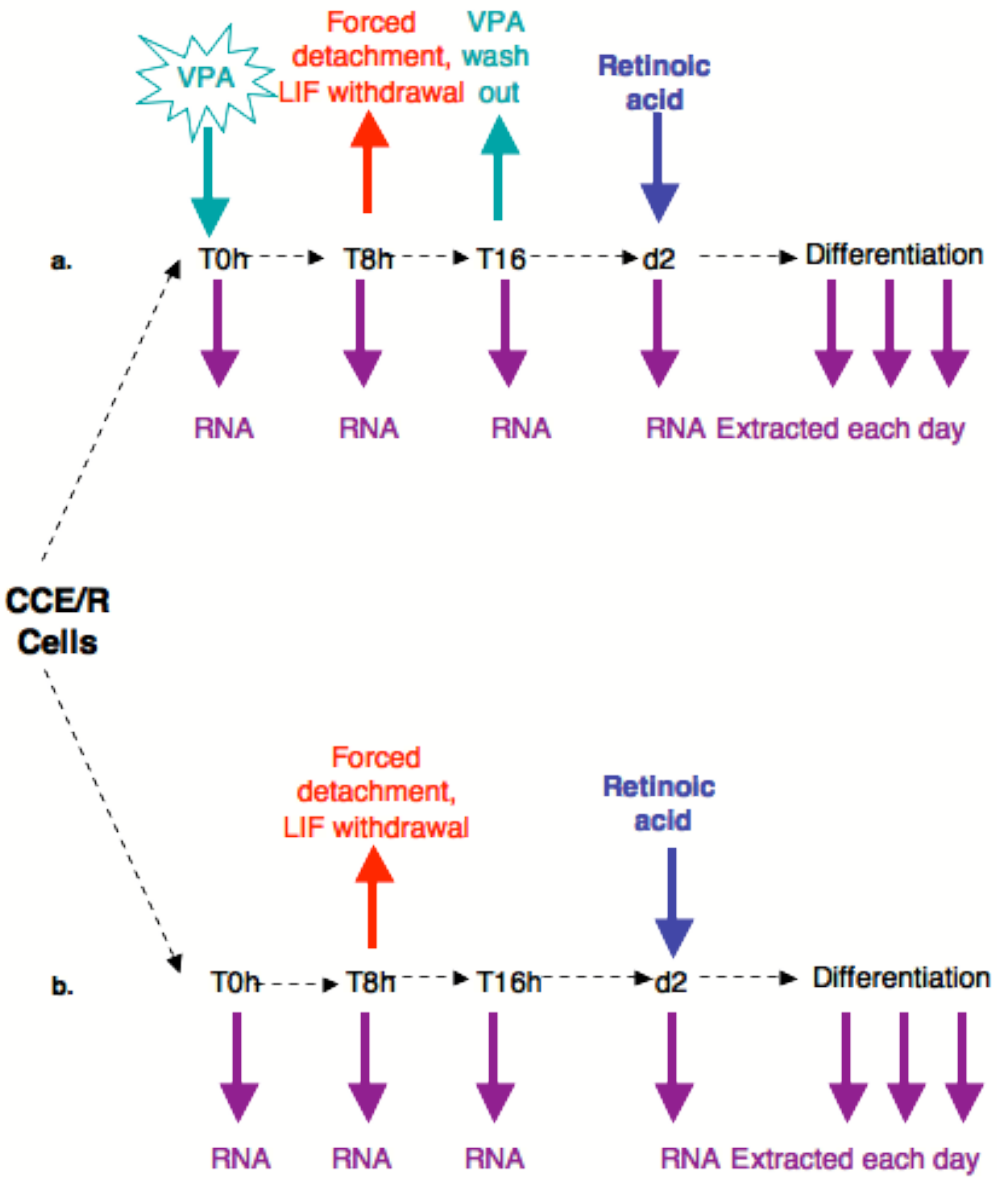


Figure 4.4 Schematic of valproate treatment and wash-out and differentiation of CCE/R cells.a. CCE/R cells were treated with valproate for 8 hours and then induced to differentiate by replating and removal of LIF. Cells remained in the presence of valproate for a further 8 hours until wash out. Retinoic acid was added to a final concentration of 1μM at day 2. RNA was extracted at the timepoints shown and on every day of differentiation. **B.** Control cells were treated exactly as in **a** but without valproate treatment.

itself during differentiation by an early or possibly late induction of transcription from that locus.

4.2.1 Do induced histone modification changes show epigenetic effects in differentiating cells?

The timeline of events are displayed in Figure 4.4. The control population is the same as in Figure 3.4, but is displayed again here in order to provide a direct comparison between treated and untreated differentiating cells. For each gene studied, the remarkable observation is that the overall pattern of gene expression throughout differentiation is unchanged, though the absolute levels of gene activity are altered by valproate treatment (Figure 4.5). For genes where this is more pronounced, as for *Hoxb2, Hoxb5* and *Hoxb9,* the change in absolute levels has resulted in a possible earlier induction of transcription from this gene (Figure 4.5b, e and i). There are similar effects at *Hoxb1, Hoxb5,* and *Hoxb7* though they are less pronounced (Figure 4.5 a,e,g). *Hoxb3* on the other hand, shows a remarkably reduced level of transcription in the valproate treated population though the transcription timing is similar to that in control cells (Figure 4.5 c). *Hoxb8* also displays a lower level of transcription from the treated cells (Figure 4.5h). Finally, *Hoxb4* and *b6* show transcription levels that are dependent on the day of extraction as to whether levels are higher or lower in treated cells (Figure 4.5 d,f). Thus, generally, there is no alteration in gene expression timing from the *Hoxb* cluster upon differentiation with retinoic acid after pre-treatment with valproate.

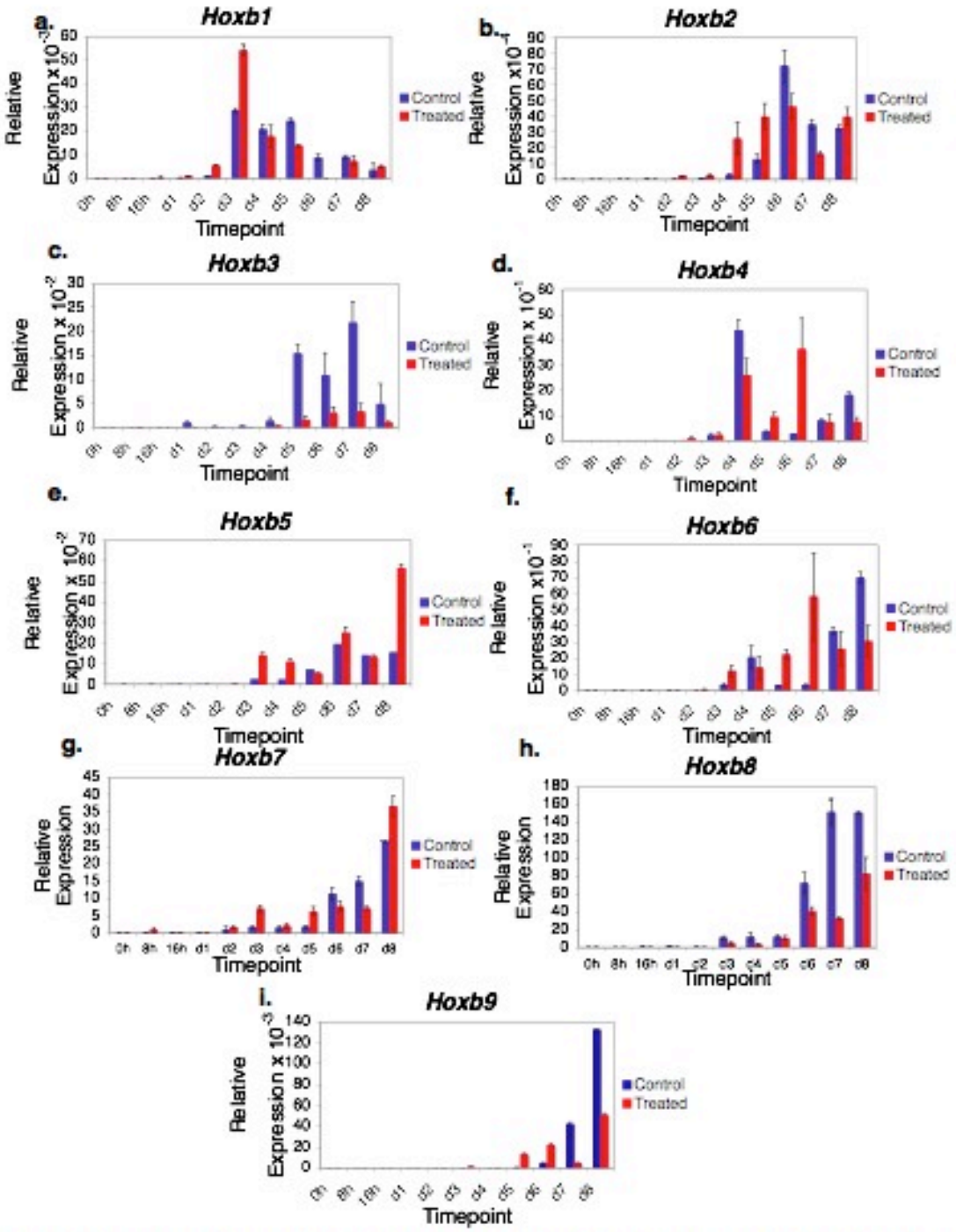


Figure 4.5 The effects of a 16 hour 1mM valproate treatment and wash out on timing of *Hoxb* gene expression during differentiation. CCE/R cells were treated as according to the scheme in Figure 4.4. RNA was extracted at each timepoint and cDNA synthesized. Expression relative to day 0 (with expression at day 0 being set at 100 or as powers of 100 as indicated) was determined by quantitative real time SYBR green PCR and normalised to actin. One experiment is shown. Error bars represent the standard error of the mean of PCR triplicates.

5 Histone modifications through the cell cycle: a means of looking at epigenetic inheritance and predictive modifications

For histone modifications to be truly epigenetic modifications instrumental in cell memory, they must be inherited through the cell cycle. Whilst no inheritance of externally induced histone modification changes could be seen at the *Hoxb* cluster, existing histone modification inheritance could result in the dictation of gene expression patterns required for differentiated cells to function within their niche. This inheritance includes their persistence or reimplementation during or after two potentially problematic stages of the cell cycle; mitosis and S phase (Figure 5.1).

At S phase, the DNA template is replicated, and the associated histones and their modifications too must be faithfully replicated and deposited at the relevant genome location. The mechanism by which this occurs is largely unknown, and is still hotly disputed. Some groups display evidence for a semi-conservative mechanism of nucleosome deposition, and others show that newly synthesized nucleosomes are deposited directly behind the replication fork (Probst et al, 2009). During mitosis, chromatin undergoes a dramatic compaction, furthermore, there is limited transcription from genes during this phase of the cell cycle. This is evidently problematic as regards inheritance of gene expression patterns, and, potentially histone modifications. The mechanisms that reintroduce transcription in G1 phase of the cell cycle must somehow know which genes to switch on and off. It is possible, and has been frequently suggested that this is the result of histone modification patterns maintained through this phase of the cell cycle. However, studies of histone modifications so far during this difficult phase of the cell cycle have been limited.

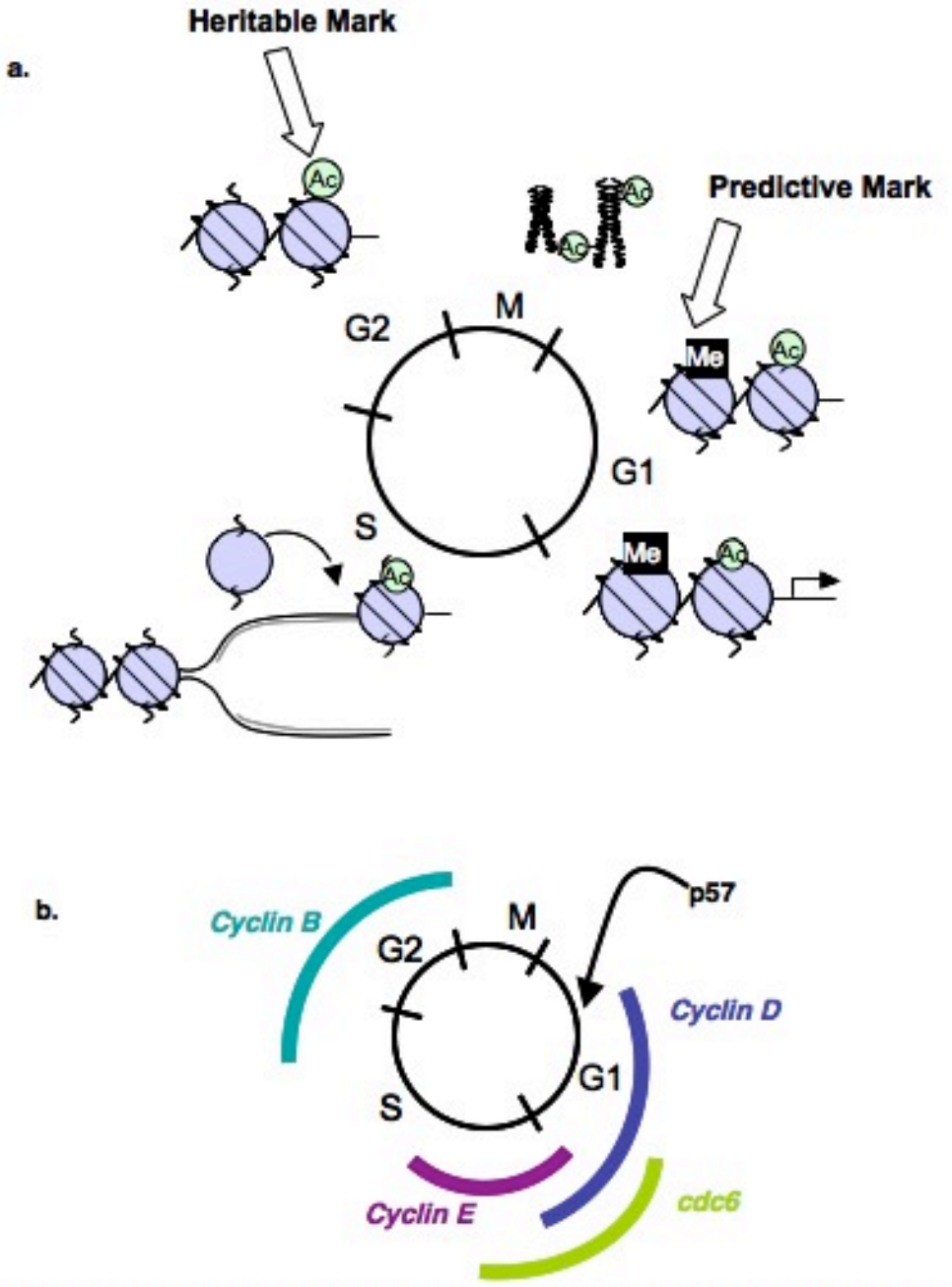


Figure 5.1 Studying histone modifications through the cell cycle. a. As an example, acetylation at a specific residue on a histone N- terminal tail represented by the green circle is shown as a memory mark and is maintained through the whole cell cycle at G1, S, G2 and M. Methylation at a specific mark is also shown as an example of a predictive mark represented by the black square, it is present just before (at early G1) and during transcription of the gene (late G1). **b.** Gives an indication of the cell cycle related expression restriction of the genes under study.

The methods used so far to study inheritance of histone modifications through the cell cycle have employed inhibitors to gain pure populations of metaphase cells. For example, one study treated with nocodazole for 16h in order to obtain a pure metaphase population, and another, treated with the same drug for 8 hours (Kouskouti & Talianidis, 2005; Valls et al, 2005). Both studies present evidence for inheritance of these marks through the cell cycle. Evidently, the addition of drugs might have additional effects upon the cells that could affect levels of histone modifications. Furthermore, both of these studies used transformed cell lines as their model. Transformed cells are known to have aberrantly regulated cell cycles, something that may be manifest in their gene regulation mechanisms and hence their inheritance patterns of histone modifications. Therefore, such studies could be improved by the use of primary cell lines.

The study of histone marks through the cell cycle may allow the distinction of not only memory marks, but marks that are predictive for genes that are about to be switched on. In this study, in experiments described earlier, the presence of active modifications on *Hox* genes that are silent in embryonic stem cells indicates a possible role for these modifications in predicting transcription. If genes could be studied that show periodic expression through the cell cycle, then it would be possible to examine which of the marks at the genes are predictive, and which are heritable. Thus, for the purpose of this pilot study, genes that were chosen with which to develop the technique were cyclin B1, cyclin D1, cyclin E1 and cdc6 as genes with periodic cell cycle expression, and GAPDH as a control (Berger et al, 1999; Kiyokawa et al, 1992; Smits & Medema, 2001; Williams et al, 1997). Their periodic expression with respect to the cell cycle is shown in Figure 5.1b. A cell cycle inhibitor,

p57, was also chosen, as its expression is inducible and may give a further insight into the roles of permissive modifications. In addition, there is evidence that the *Hox* genes *Hoxa5, Hoxa7* and *Hoxa10* show an expression pattern with cell cycle periodicity. This was shown in a study that separated populations of synchronised HeLa cells by double thymidine block (Mishra et al, 2009). At these genes, in the same study, the levels of H3K4me3 and the association of MLL was also shown to cycle periodically with the cell cycle. However, the time-points were not taken frequently enough to show whether the H3K4me3 mark occurred before the gene was expressed, and therefore whether the marks were predictive of expression, or consequential. Also, the use of a transformed cell lines is not necessarily an appropriate model in which to study the fundamental control of gene expression. It is, however, highly possible that the expression of genes from the *Hoxb* cluster also cycles within the cell cycle in differentiated cells. Therefore studying the dynamics of histone modifications and the expression of *Hoxb* genes in such a system may help to understand the significance of the histone modifications at *Hoxb* genes in undifferentiated stem cells.

Here, in this final chapter of results, a pilot study is described, aimed at delineating the relationship between heritable and predictive histone modifications through the cell cycle. The main aim was to develop a technique whereby the stages of the cell cycle could be resolved in live primary cells in a manner that would allow assay of histone modifications by chromatin immunoprecipitation.

5.1.1 MEFs may be labelled with a live dye to produce a cell cycle profile

Cells can be separated into the different stages of the cell cycle by fluorescent activated cell (FACs) sorting. Conventional labelling techniques to stain cells for DNA

content require either fixing or permeabilisation of the cell. Here, however, it was hoped to use a viable dye to label and sort the cells in order to preserve the true nature of the chromatin for immunoprecipitation. Both techniques come with their advantages and disadvantages. Fixing with ethanol, and permeabilisation with detergent will put the cells under significant stress, possibly causing a stress-response after prolonged exposure that alters the histone modifications. However, the turnover of histone modifications is such that if cells could be fixed immediately, and physiologically representative chromatin could be extracted from these fixed cells then this method might produce a clean snapshot of the chromatin modifications at a defined point in the cell cycle. On the other hand, with live cells, it is known that intact chromatin can be extracted which will represent the physiological state of the cells, although the effect of turnover of the enzymes through any prolonged cell sorting is unknown.

The dye chosen to label and sort the cells is a commercially available dye from Invitrogen; Vybrant DyeCycle green. This dye intercalates into the minor grove of the DNA, and so is able to permeabilise cells and label DNA without the need for fixing or addition of detergent. It is non-fluorescent until it binds DNA, at which point it has the same fluorescent properties as FITC, which enables its use in conventional FACs machines. Most importantly, this dye is a viable dye i.e. the cells may be labelled whilst live and cycling. The cells chosen to develop this technique are primary mouse embryonic fibroblasts (MEFs).

In initial experiments MEFs and HL60s, for whom the cell cycle profile was known, were labelled using the dye according to the manufacturers instructions (Figure 5.2). Both HL60s and MEFs were labelled in such a way that a typical cell cycle profile was obtained by FACs analysis with a large peak indicating G1 and a smaller peak

indicating G2/M (Figure 5.2a). The FACs profiles shown are deliberately un-gated in order to show a complete picture of the effects of the dye. Typically, the fluorescence of the G2/M peak (600 in HL60s and 500 in MEFs) was twice that of the G1 peak (300 in HL60s and 250 in MEFs), indicating double the content of DNA in G2/M than in G1. The final noticeable feature is that there is a larger sub-G1 population in the MEFs than in the HL60s. This is most likely due to their primary nature resulting in more cell death.

Using the Vybrant DyeCycle Dye, MEFs were sorted based upon DNA content into distinct G1, S and G2/M populations (Figure 5.2a). Cells from each population were then reanalysed to establish the purity of the population. It must be noted, however, that in order to create a stress-free environment for the cells, the cells were sorted into cell medium that did not contain the Vybrant DyeCycle dye and thus the dye may leach out of the cells after the sort causing the fluorescence profile to shift upon second analysis. However, in Figure 5.2a no shift in the fluorescence of the three populations is visible. Therefore, this is a viable technique with which to sort cells into the different phases of the cell cycle.

5.2 A technique for immunoprecipitating small numbers of cells is applicable to FACs sorted cells

5.2.1 Chromatin may be extracted from mixed populations of FACs sorted cells and *Drosophila* SL2 cells

One of the main priorities in the development of the technique was to conserve the integrity of the cells' chromatin. Therefore, we were reluctant to sort the cells any longer than necessary so as to prevent the induction of a stress response. 5-10,000 cells per cell cycle fraction only, therefore, were obtainable by cell sorting in a time of

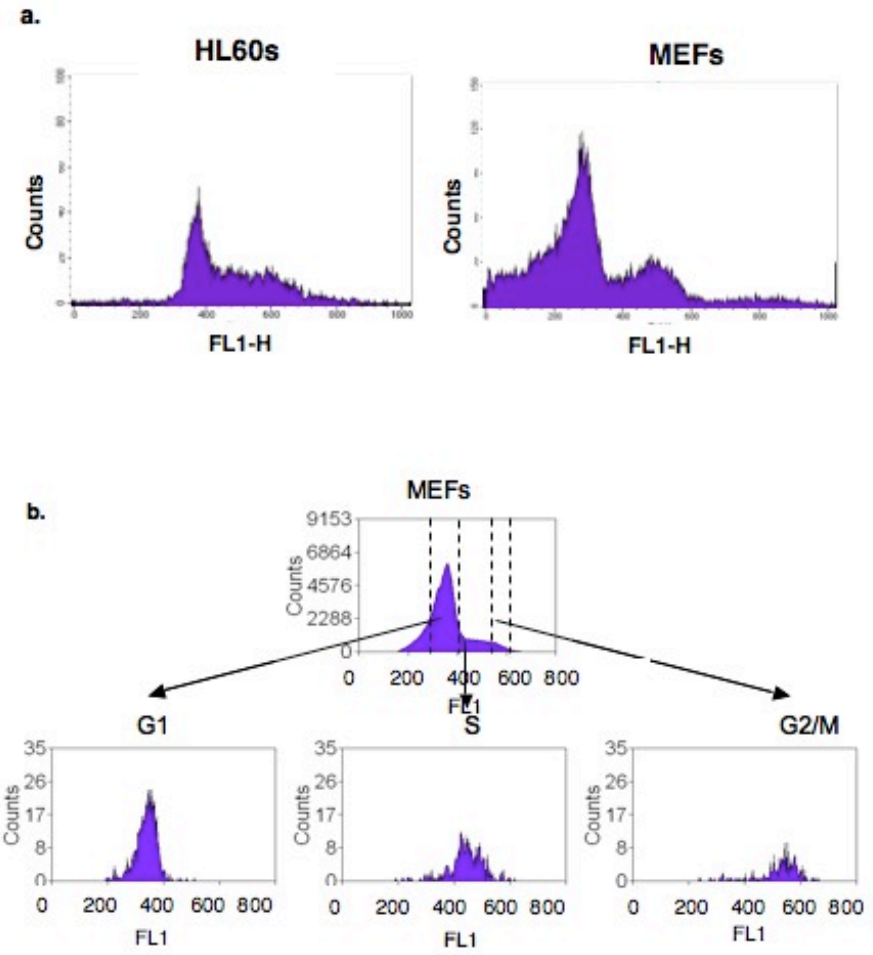


Figure 5.2 Vybrant DyeCycle dye produces a cell cycle profile with viable cells. **a.** HL60 cells were suspended at a concentration of 1 x 10^6/ml and incubated in 25mM Vybrant Dyecycle RPMI for 30 minutes before FACs analysis. MEFs were harvested by trypinisation and resuspended in 25mM Vybrant Dye Cycle DMEM at a concentration of 1 x 10^6/ml for 30 minutes before FACs analysis. **b.** Vybrant dye cycle stained MEFs were sorted into G1, S and G2/M populations based upon their DNA content.

15 minutes. This is too low to undertake the conventional ChIP technique used throughout the rest of thesis as this requires about 10^7 cells. Although the small number of sorted cells could be increased by sorting for longer, this would prolong the stress the cells would undergo, and hence further increase the possibility that the histone modifications would be altered. Therefore, the carrier CHIP (CChIP) procedure was used to immunoprecipitate the chromatin (Figure 5.3, (O'Neill et al, 2006). This is a technique developed to be able to immunoprecipitate native chromatin from low numbers of cells by "buffering" these low cell populations with *Drosophila* chromatin. It is reliant on the fact that mouse DNA may be distinguished from *Drosophila* DNA by using primers specific to mouse genes in PCR reactions during the final analysis.

5.2.2 Primers may be designed that identify specifically mouse DNA from a mixed population of mouse and *Drosophila DNA*

In order to detect mouse DNA in the bound and unbound fractions using the CChIP technique, mouse-specific primers must be used. Such primers must not cross-react with *Drosophila* DNA and so must be rigorously tested. Target genes used in the development of this technique were chosen for their cell cycle specific expression and so were frequently genes instrumental in the control of cell cycle progression and thus were highly conserved. As such, rigorous testing for cross-reaction was imperative. Initial testing was carried out using mouse and *Drosophila* DNA at a gradient of annealing temperatures from 56 to 64^{o}C. In such a manner, the optimum annealing temperature was determined whereby there is no cross-reaction of the primers with *Drosophila* genomic DNA. The PCR products were then analysed on polyacrylamide to ensure a single gene product had been amplified. Two examples of primer optimisation are given in Figure 5.4. It is clear from the gel that there is no

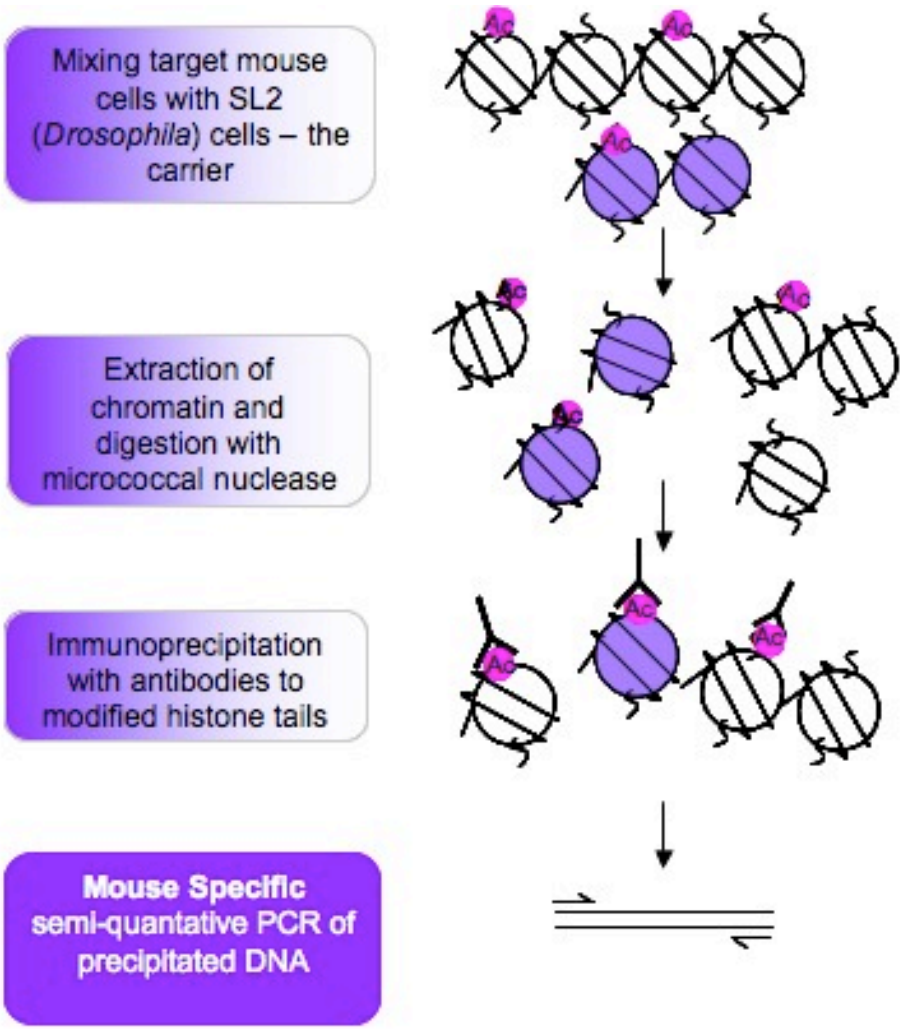


Figure 5.3 The CCHIP procedure. Chromatin is extracted from mixed populations of mouse and *Drosophila* cells. The chromatin is extracted and nucleosomes are precipitated using antibodies directed against histone modifications. The mouse DNA is then analysed using species specific PCR. The mouse nucleosomes are represented in purple. *Drosophila* nucleosomes are represented in white.

problem with cross-reaction with *Drosophila* DNA at any of the temperatures using primers designed against cyclin D1 (Figure 5.4a). However, for p57, there is multiple cross reactions at all temperatures (Figure 5.4b), and to continue with CChIP analysis, these primers had to be redesigned.

During the development of the FACs sort and CChIP technique, an acetyl mark, H4K8ac, and a histone H3 lysine 4 methylation mark were chosen for study as they were used in the previous study of the *Hox* genes. The sorted populations (around 5-10,000 cells) were first pooled with 5 x 10^7 *Drosophila* SL2 cells. The chromatin was then extracted from cell nuclei, and digested to yield mono and oligo-nucleosomes, and then analysed on an agarose gel (Figure 5.5b). This mixed *Drosophila* and mouse chromatin was then immunoprecipitated with antibodies raised to the histone modifications chosen above.

From Figure 5.5c, it is evident that there has been efficient precipitation of the SL2 DNA using the antibodies raised to H4K8Ac and H3K4me2. However, in the pre-immune control, there is no precipitated material and so there has been no non-specific binding of the SL2 chromatin.

The real test of the method was whether the specific primers were able to detect mouse DNA in mixed DNA populations obtained from the CChIP procedure. From Figure 5.5 d, it is evident that the radioactive PCR technique could be used to detect mouse DNA from equivalent concentrations of the unbound and bound DNA fractions from the S phase population of cells precipitated with the anti-H3K4me2 antibody. This is true at both 38 and 40 PCR cycles. This is significant since it is the

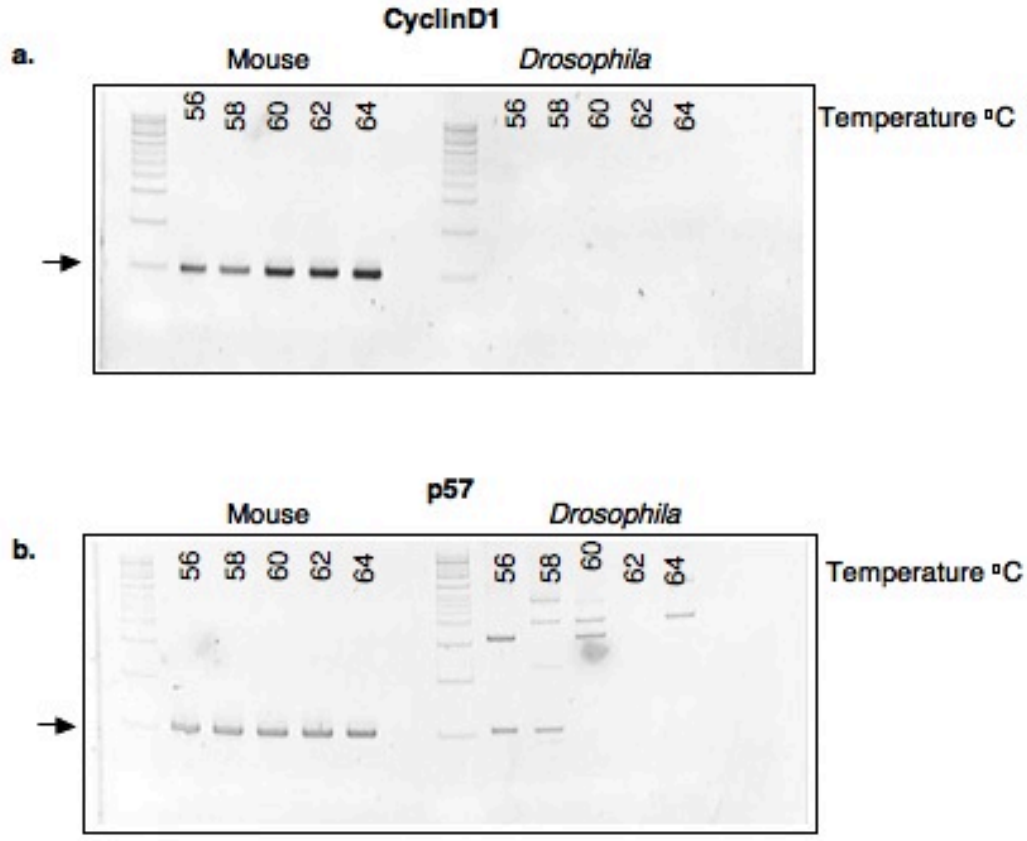
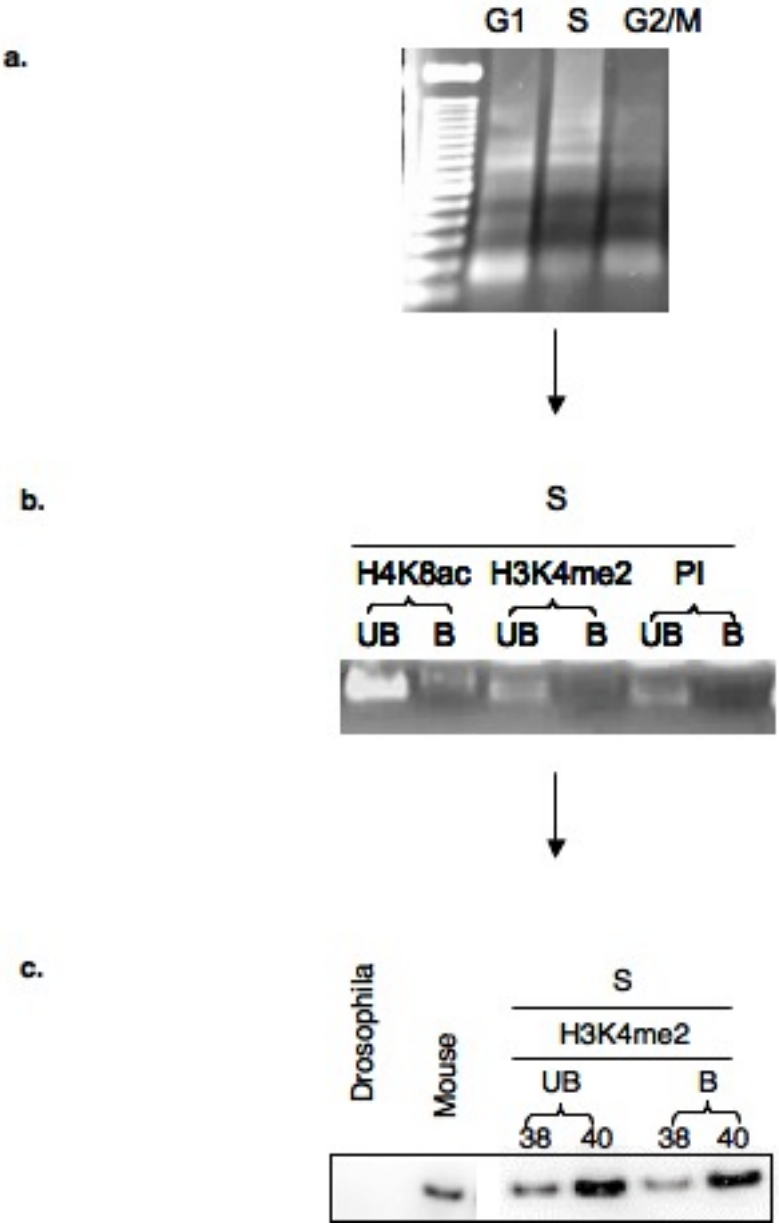


Figure 5.4 *Drosophila* and Mouse genes are identifiable based on their DNA sequence using PCR. a. PCR analysis was carried out using primers directed against CCND1 and **b.** p57. The correct amplicon is marked by an arrow for each.

Figure 5.5 Initial CChIP on FACs sorted cells. a Chromatin was prepared from these fractions by first combining each cell cycle fraction with SL2 cells. Chromatin was extracted from the nucleus and digested using micrococcal nuclease. The resultant chromatin ladder was ran on a 1.2% agarose gel. **b.** The chromatin was immunoprecipitated using rabbit anti-H4K8Ac, rabbit anti-H3K4me2 and a pre-immune control and an equal volume of DNA from the bound and unbound fractions ran on a 1.2% agarose gel. B=bound fraction, UB=unbound fraction **c.** Radioactive α-P32 PCR analysis was carried out using primers to GAPDH. An equal amount of DNA, determined by tritium counts was loaded into each PCR reaction and 0.1 uCI dCTP labelled with α-P32 added to the reaction. PCR reactions were sampled at 38 and 41 cycles and then analysed on 5 % polyacrylamide gels, dried onto filters, exposed to a phosphor screen and scanned using a Phosphorimager. D=*Drosophila*, negative control, M=mouse, positive control.


a.
G1
S
G2/M
b.
S
H4K8ac
H3K4me2
PI
UB
B
UB
B
UB
B
c.
Drosophila
Mouse
S
H3K4me2
UB
B
38
40
38
40

first time that the CChIP technology has been applied to FACs sorted cells, demonstrating the universal applicability of the technique.

There were a number of occasions where mouse DNA could be detected in the bound fractions but not the unbound fractions, most likely due to an inhibitory effect of the *Drosophila* DNA (data not shown). Here, both unbound and bound DNA had to be serially diluted in order to determine the optimal concentration for PCR amplification.

5.3 Analysis of housekeeping genes and cell cycle regulated genes

5.3.1 Housekeeping genes show a varied pattern of histone modifications through the cell cycle

As a control to genes whose expression alters through the cell cycle, a gene whose expression does not change through the cell cycle was also studied. This was important in order to check that any patterns of modifications seen were specific to these genes, and were not just an artefact of different efficiencies of precipitation at the different stages of the cell cycle (Fujii-Yamamoto et al, 2005). However, these effects are likely to be masked as the majority of the chromatin precipitated was asynchronous *Drosophila* chromatin. It could be envisioned that the condensed mitotic chromatin is difficult to digest at the micrococcal nuclease stage and hence might precipitate less efficiently which would go on to produce the artefactual result that all histone modifications are depleted at this stage in the cell cycle. If all genes including the housekeeping gene show similar results at one stage of the cell cycle then this result would have to be further investigated.

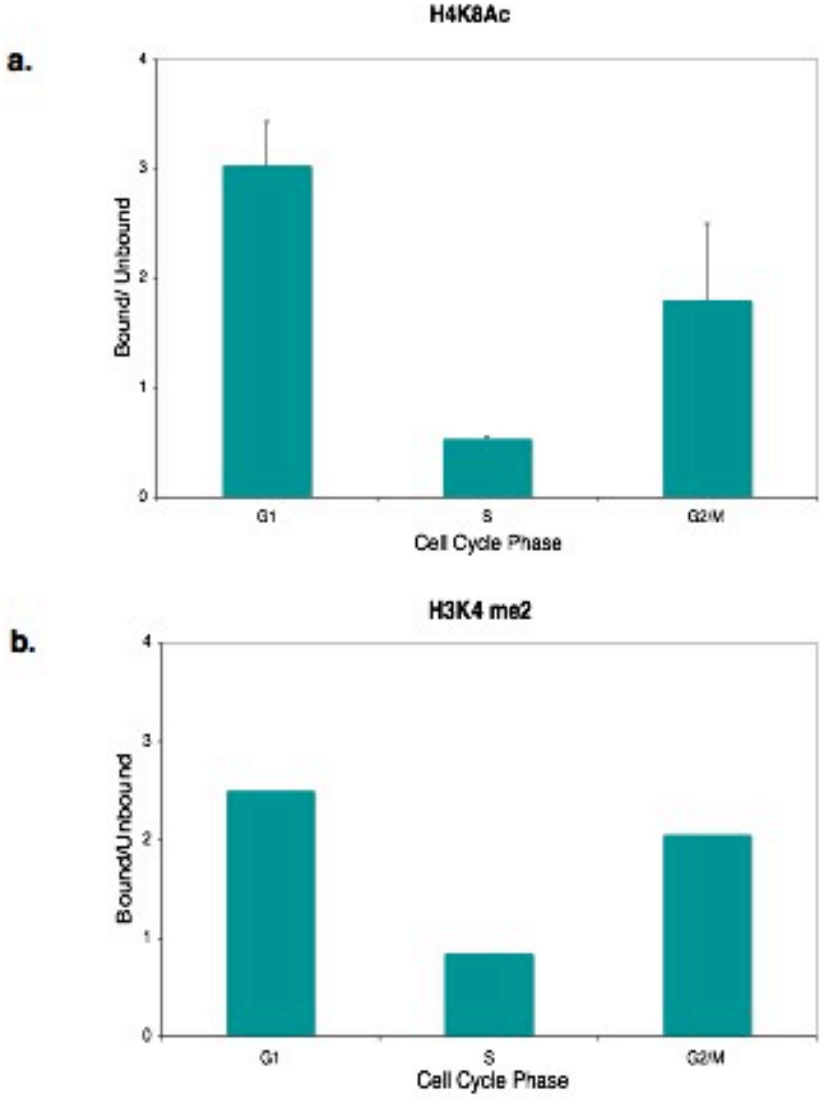


Figure 5.6 Histone modification changes over the GAPDH promoter during the cell cycle. Quantitation of H4K8 acetylation and H3K4 methylation over the GAPDH promoter following CChIP on FACs sorted cell cycle populations from MEFs **a.** H4K8Ac quatitation by radioactive PCR, data is from two biological replicates, the error bars represent the standard deviation. **b.** H3K4me2 data is from a single experiment

Gapdh is a housekeeping gene. From Figure 5.6, it is possible to see that during G1 *Gapdh* is enriched for acetylation at lysine 8 of histone H4. It is then depleted for this modification during S phase and again, slightly enriched during metaphase. Previously reported expression data shows that *Gapdh* is, in fact present throughout the cell cycle in MEFs, so if acetylation is associated with active transcription, then why its status should change during the cell cycle is unclear (Fujii-Yamamoto et al, 2005). It is possible that there is some previously unobserved cell cycle oscillation in *Gapdh* mRNA expression in MEFs which would require further investigation. From Figure 5.6b, it is very apparent that a second active mark H3K4me2 follows a similar pattern to H4K8Ac. H3K4me2 is also an active mark. Therefore, it could be that these marks are not directly involved in transcription at this *Gapd*h gene.

5.3.2 Cyclin D displays a possible consequential mark of transcription

The H4K8ac modification was examined at one location in the cyclin D1 promoter. The cyclin D1 gene is a cyclin expressed in early G1 (Kiyokawa et al, 1992). It is responsible for inducing cyclin E expression via its association with and activation of a kinase that phosphorylates and displaces a repressive E2F transcription factor complex from the cyclin E promoter. Given the timing of its transcription during the cell cycle, it might be expected that the active H4K8ac mark would be maximal during G1, if not predictively in metaphase. However, this is not the case. The results show that this active acetyl mark is almost depleted in G1 and is significantly enriched in S phase, with this enrichment tailing off during metaphase (Fig 5.7). Since this differs significantly from the pattern seen at the *Gapdh* promoter, from the same ChIP material, it is therefore, unlikely to be an artefact of precipitation at S phase. This

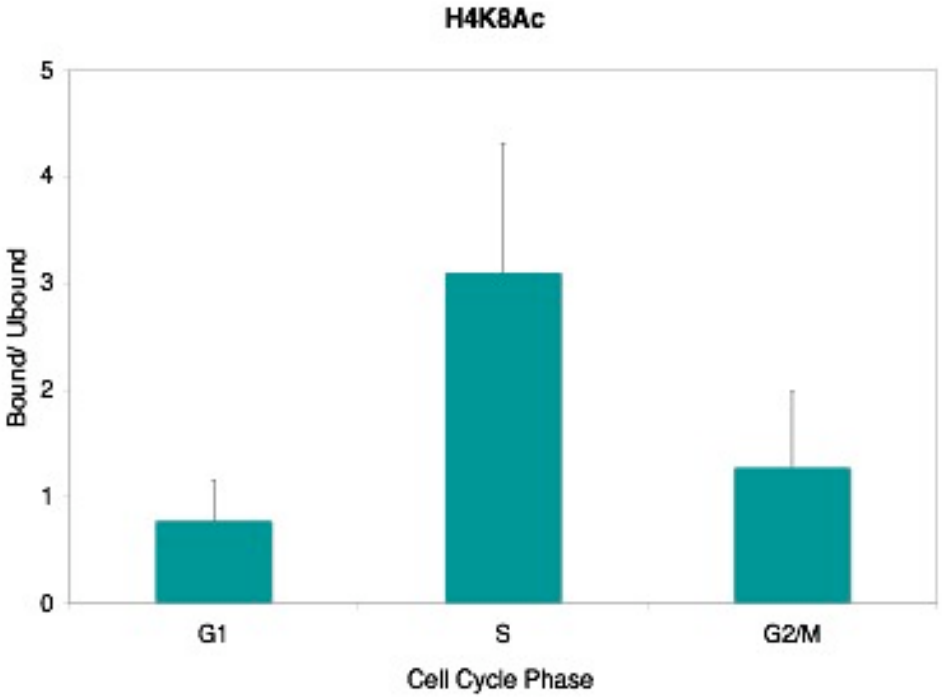


Figure 5.7 Histone modification changes over the *CyclinD1* promoter during the cell cycle. Quantitation of acetylation over the cyclin D1 promoter following CChIP on FACs sorted cell cycle populations from MEFs **a.** H4K8Ac quatitation by radioactive PCR, data is from two biological replicates, the error bars represent the standard deviation.

mark, therefore, might be a consequence of transcription, put in place as the polymerase passes, or be required for ongoing transcription.

5.4 Further investigation using the cell cycle as a model

The above study has shown that it is possible to use the CChIP technique upon cells separated into the different phases of the cell cycle using FACs analysis. This will help to resolve the roles of histone modifications in transcription. If applied to *Hox* gene analysis in embryonic stem cells, for example, this could provide fascinating insight into a study that showed that S phase must be crossed before *Hoxb* genes are switched on (Fisher & Mechali, 2003). There may, however, be one hurdle to studying ES cells using this technique. Multipotent adult stem cells are often characterised by what is known as the side population phenotype. This pheneotype is a reduced fluorescence when labelled with dyes such as Hoescht dye (Challen & Little, 2006). This results from the active efflux of the dye from the cell. Thus, in order to continue this study in ES cells, it is likely that a technique using fixed cells and a dye such as propidium iodide would have to be developed. However, the demonstration that CChIP may be applied to FACs sorted cells remains an exciting development in the investigation of the roles of histone modifications.

6 Discussion

6.1 The putative roles of histone modifications

Ever since the discovery that histone tails could be post-translationally modified, attempts have been made to correlate these modifications with a biological function. The modifications are placed by an array of enzymes at multiple locations upon histone tails. The conservation of these modifications, and their respective enzymes, through evolution suggests the marks are more than a mere “epi” phenomenon. One biological function that has been pinned on histone modifications is an active role in gene expression control (Allfrey et al, 1964). Initial evidence for this was demonstrated using radioactive acetate groups. The high turnover of these groups on histones was seen to correspond with high rates of RNA synthesis (Pogo et al, 1968). However, forty years later there is still a lack of solid evidence for a causative role for histone modifications in gene transcription. Linked with this role in active transcription, another function in which histone modifications are believed to play a part is that of the propagation of cellular transcription patterns, a so called “epigenetic” memory. It is these relationships that have been investigated here and will be further discussed.

6.2 The evidence for a causative link between histone modifications and transcription

The evidence for histone modifications playing a role in transcription is currently embroiled in genome-wide mapping experiments. A significant initial study using the ChIP:chip technique was a genome-wide study in yeast which found that the histone modification H3K4me3 was associated with the promoters of active genes in yeast (Santos-Rosa et al, 2002). This type of study has now been extended by mapping large areas of eukaryotic genomes in a variety of cell types using ChIP:chip and

ChIP-seq, a related technique in which “high throughput” DNA sequencing is used instead of microarrays (Barski et al, 2007; Li et al, 2007; Mikkelsen et al, 2007; Roh et al, 2006). Such studies are able to correlate certain marks with actively transcribing or silent genes, however, this kind of evidence merely associates specific histone marks with gene transcription status. It does not show that marks dictate gene expression patterns.

The potential for protein binding modules to recognise an altered binding surface on a nucleosome explains how an active role in transcription might be played by histone modifications. In this manner, histone modifications could serve in recruiting transcription factors or other proteins to chromatin. Evidence for such a mechanism may be found in the histone modification binding modules that do exist such as the bromodomains and chromodomains. However, the ability for these domains alone to be able to recruit proteins specifically to chromatin has been disputed (Ringrose & Paro, 2007). This is because interchanging the chromodomains of HP1 and polycomb is insufficient to target the heterologous proteins to targets enriched for the “targeting” modification *in vivo* (Platero et al, 1995; Ringrose & Paro, 2007). If the proteins were recruited to the DNA by their chromodomains alone, then swapping the domains would be enough to recruit them to the targets of the other protein. However, this is not the case, so other mechanisms must exist that recruit the proteins to the DNA resulting in gene repression.

6.2.1 Causative “Predictive” histone modifications

The finding that the normally active modification, acetylation may be present across the whole ß-globin locus at both the genes of the locus and the large non-transcribed region provided an “uncoupling” of histone modifications from the process of transcription itself (Hebbes et al, 1994). This led to the notion that histone

modifications are required for transcriptional “competence” of an entire locus with immediate control of transcription being mediated by additional factors. Such an idea is consistent with a “predictive” role for histone marks, that is, marks that are present in the absence of transcription that are later able to determine transcription from a locus.

In an extension of a predictive, causal role for histone modifications in transcription, the “histone code hypothesis” was proposed (Strahl & Allis, 2000; Turner, 2002). This hypothesis proposes that defined combinations of histone marks determine transcription patterns. In such a manner, multiple marks cooperate by recruiting proteins with multiple histone-modification recognition sites, or recruit multiple proteins, building up the environment for correct gene transcription. Thus, combinations of marks should predict specific transcriptional outcomes and could be examined in inducible or transcriptionally “poised” systems. The *Hox* genes provide us with such a system. Their transcription may be induced in culture by the addition of retinoic acid (Simeone et al, 1990). This gene induction programme follows the same temporal pattern of expression as occurs *in vivo*. Thus, it gives the opportunity to study the causal or consequential role of histone modifications, and the predictive aspect of the histone code. That is, those modifications present in undifferentiated cells may be indicative of future transcription.

6.3 Combinatorial predictive histone modifications at *Hox* genes

The histone modifications at the promoters of the *Hoxb* gene cluster were here characterised in undifferentiated embryonic stem cells, i.e. in the absence of detectable transcription from these genes. One of the major predictions of the histone

code, one of the hypotheses under examination, is that histone modifications function in a combinatorial manner. Indeed, without combinations working together, the code is no longer a code. Here, three modifications were investigated; H3K4me3, H3K9ac and H3K27me3. H3K4me3 and H3K9ac have been previously found to be associated with active genes and H3K27me3 with repressed genes. The presence of any combination of these marks together at the promoters of *Hoxb* in ES cells may therefore indicate cooperation between these modifications in determining later patterns of transcription as these genes are not expressed in pluripotent cells. That is, they may be "predictive" of transcription at a later stage of development.

At all the promoters of the *Hoxb* genes studied, the three modifications H3K4me3, H3K9ac and H3K27me3 were present. However, the pattern across the *Hoxb* cluster of the actively associated modifications, and the repressive associated modifications was very different. Both the actively associated modifications H3K4me3 and H3K9ac showed a large variation in their levels at the promoters of the different genes examined with their levels being highly correlated. Thus, there is a potential interaction between these two marks. There was, however, no general trend in these active modifications across the cluster. Both H3K9ac and H3K4me3 showed highest enrichment at the promoter of *Hoxb9,* and lowest enrichment at the promoter of *Hoxb3*. All genes were, nonetheless, consistently more enriched for H3K4me3 than they were H3K9ac. Care must be taken, however, when comparing absolute levels of different modifications due to the differing efficiencies of precipitation of different antibodies. Although this is controlled for by amplifying from the same amount of DNA in quantitative PCR reactions from the input and bound samples, this idea of different precipitation efficiency must be taken into account when comparing absolute levels of different modifications.

The polycomb-associated H3K27me3 mark and trithorax-associated H3K4me3 showed more equivalent absolute levels than the two activating modifications across the promoters of the *Hoxb* genes. Again, possible effects due to different antibody precipitation must be kept in mind. However, there is a more important observation related to the pattern of this modification across the *Hoxb* promoters. That is, the H3K27me3 polycomb-associated mark shows a more uniform enrichment across the cluster at the positions examined than the H3K4me3 mark. It could therefore be envisaged that it forms a "blanket" across the cluster. These data are consistent with previous data seen at so-called "bivalent" genes (Azuara et al, 2006; Bernstein et al, 2006). This "bivalent signature" describes the coordinate presence of both active and repressive marks originally described in pluripotent cells at the promoters of genes key for developmental regulation. At the *Hox* gene promoters, this was elucidated as being "large regions of H3 lysine 27 methylation harbouring smaller regions of H3 lysine 4 methylation". As discussed in the introduction to this thesis, the H3K27me3 mark is deposited by and co-associated with the polycomb repressive complex 2 (PRC2) and the H3K4me3 a mark deposited by the active transcription maintaining trithorax complex (Schuettengruber et al, 2007). The presence of these two marks at developmentally regulated gene promoters is believed to "poise" the genes ready for transcription whilst keeping them silent in the undifferentiated state. The data here generated using NChIP support the initial ChIP:chip characterisation in mouse embryonic stem cells showing the patterns of H3K4me3 at the *Hox* cluster to form peaks at the promoters, with H3K27me3 potentially covering the whole cluster.

The presence of *Hoxb* gene transcripts was analysed in these undifferentiated CCE/R cells by cDNA analysis. This revealed, as expected, that there were no detectable *Hoxb* gene transcripts in these undifferentiated ES cells. The lack of

transcription from these genes suggests that marks normally associated with active transcription are here not sufficient to be causative of gene transcription. Therefore, any role they play in transcriptional activation must occur later in differentiation, when the genes are induced. It is important to note, however, that in this study, only full transcripts were assessed (the *Hox* genes are small and primers crossed the major exon of all *Hox* genes). In a separate study, the different forms of Pol II were examined at bivalent genes, and showed the initiating form to be localised to these genes (Stock et al, 2007). Furthermore, it was shown that transcripts were produced, albeit at low levels. However, although RNA and cDNA was generated using similar methods as in this study, the *Hoxb* cluster was not examined. Therefore as shown here, the *Hoxb* genes may represent a subset of bivalent genes that are not expressed in ES cells, even at low levels.

6.3.1 The role of "predictive" modifications in differentiation

The possible predictive role of active marks at the *Hoxb* promoters was examined by differentiating these embryonic stem cells using retinoic acid and attempting to correlate levels of active marks at *Hox* genes with their induction of expression. For this, cells were first induced to differentiate by replating and withdrawal of LIF, and then pushed down the neuroectoderm lineage by the addition of retinoic acid (Rohwedel et al, 1999). This addition of retinoic acid has previously been shown to induce expression from the cluster in a co-linear fashion (Simeone et al, 1990). The genes do, indeed, show a co-linear timing of expression from the *Hoxb* cluster, co-linear with their location on the chromosome in this system. For example, *Hoxb1*, the most 3' gene is switched on first at day 2, and *Hoxb13,* the most 5' gene, is switched on last at day 9. This pattern of induction does not, however, correlate with the levels

of active modifications at these genes. For example, *Hoxb9* is the gene with the highest level of active modifications but is not the first gene to produce mRNA. Conversely, *Hoxb3* that shows the lowest level of actively associated histone modifications is not the last to be induced upon differentiation. Temporal co-linearity is achieved irrespective of the levels of active histone modifications before differentiation. This is still consistent with a predictive role for these modifications in permitting future transcription from the locus upon differentiation, however, absolute levels, appear to be of no importance in a "predictive" role.

These data are also consistent with another study in mouse ES cells studying the λ5-VpreB1 locus. The VpreB1 and λ5 genes are expressed in pre-B and pro-B cells to form a surrogate antibody light chain at these highly specific stages of B cell development (Szutorisz et al, 2005). They are, therefore, not expressed in embryonic stem cells. At this early undifferentiated stage, however, NChIP data showed an intergenic cis-acting region is marked by histone H3 acetylation and H3 lysine 4 methylation. These marks then spread towards the genes as the cells differentiate toward the B-cell lineage and are switched on. Importantly, if the ES cells were induced to differentiate down non-lymphoid lineages, then theses active marks disappear. Thus at this B cell-specific locus, active histone marks serve in setting up transcriptional competence in a predictive manner. The active marks elucidated at the *Hoxb* locus could be playing a similar role.

6.4 The effects of externally induced hyperacetylation upon CCE/R cells

In order to further elucidate the role that active histone modifications have at the *Hox* genes in transcription, embryonic stem cells were treated with the HDAC inhibitor

valproate. This inhibitor is know to cause global hyperacetylation in cultured cells, and its addition was intended to cause an increase in acetylation at the *Hoxb* locus. To further elucidate the role of these histone modifications at the *Hoxb* locus, any effects on transcription both in undifferentiated cells, and later on, in differentiation were also examined. However, the experiment revealed a valproate sensitivity that was surprisingly unique to the *Hox* genes and warrants further discussion.

6.4.1 High turnover of acetate groups at *Hox* genes

The addition of valproate to undifferentiated ES cells resulted in hyperacetylation at the promoters of the *Hoxb* cluster. However, its addition was unable to induce hyperacetylation at the promoters of any other genes examined, including the housekeeping gene *Gapdh* and the pluripotency markers, *Nanog* and *Pou5f1.* Such a refractory nature of gene promoter nucleosomes to valproate treatment has also been seen in HL60 cells (Vibhor Gupta, thesis 2008). *Hox* genes are, therefore, exceptional in their response to valproate. This sensitivity of *Hox* genes to valproate indicates that there is a high turnover of histone acetylation at these genes when compared with other genes. High turnover of histone acetylation has been used to define a class of genes that are highly responsive to mitogenic stimulation (Clayton et al, 2006; Hazzalin & Mahadevan, 2005). This was proposed to be a mechanism that "poised" these genes ready for transcription. *Hox* genes too are said to be "poised" and so could exhibit a similar high turnover of acetylation putting them in a similarly transcriptionally-ready state.

6.4.2 The enzymology of high modification turnover at *Hox* genes

The high turnover of histone acetylation seen at the promoters of the *Hox* genes, and their resultant sensitivity to valproate may be the result of the enzymes localised at

the promoters of these genes. An interesting early observation, indicating a possible role for histone deacteylation in polycomb silencing was that if both the *Drosophila* HDACI gene and polycomb genes were mutated in *Drosophila*, the two genetically interact, exaggerating the polycomb mutant phenotype (Chang et al, 2001). Subsequently The PRC2 complex component, *Eed* has been shown to interact with the class I HDAC, RPD3 and that the HDAC TSA is able to relieve transcriptional repression mediated by *Eed.* This suggests a role for histone deactylation in polycom-medited gene silencing (van der Vlag & Otte, 1999). Furthermore, the histone binding proteins RbAp46 and RbAp48 have been found in isolated PRC2 complexes from HeLa nuclear extracts. As these complexes are repressive complexes, it is easy to infer that the presence of these histone binding proteins within the complex is in order to target a 'repressive' chromatin environment within the gene locus (Kuzmichev et al, 2002). Thus as PRC2 targets, this may provide an explanation for the sensitivity of the *Hox* genes to hyperacetylation by valproate as *Hox* genes. This is particularly so as the HDAC isolated from, and active in PRC2 complexes is HDAC and valproate is a class I HADAC inhibitor, inhibiting HDACs I and II. This sensitivity due to the presence of polycomb complexes, would, however, have to be confirmed by proving localisation of polycomb and HDAC I at these genes. This could, however, explain the relative lack of sensitivity of *Gapdh*, *Pou5f1* and *Nanog* neither of which are polycomb group targets in ES cells.

In addition to the HDACs, the HATs at the *HoxB* cluster could also be responsible for this hypersensitivity to valproate. The possible role of the histone acetyl transferase GCN5, in depositing the H3K9ac mark was also investigated. GCN5 knock-out mice and ES cells have been generated in a previous study (Lin et al, 2008). In these knock-out mice, Homeotic transformations are seen coordinate with the mis-

localisation of expression of *Hoxc8* and *Hoxc9*. This HAT was therefore chosen as a likely candidate for *Hox* gene acetylation. The role of this enzyme in histone *Hoxc* and *Hoxb* histone acetylation was tested using valproate and any changes in histone modifications H3K9ac, H3K4me3 and H3K27me3 assessed. If GCN5 is, indeed, responsible for histone acetylation at these genes, then no response to valproate would be visible.

At the *Hoxc9* gene, in the wild-type cells, hyperacetylation in response to valproate was seen, while, in the knock-out cells, the gene was not able to respond to valproate. Therefore, GCN5 must play a role in acetylation in response to valproate at this gene. A similar effect was seen on H3K4 methylation which exhibits elevated levels in response to valproate in the wild type but not the knock-out cells. An analogous phenomenon was seen at the *Hoxc8* promoter, whereby there was hyperacetylation in the wild-type but not in the knock-out cells upon valproate treatment. However, there was no increase in the levels of H3K4me3 in either the wild type or the knock-out cells at these genes. Therefore at both *Hoxc8* and *Hoxc9*, GCN5 plays a role in acetylation of the genes in response to valproate, though it is worth noting that H3K9 acetylation is not abolished in the knock-out cells. In stark contrast with the changes seen at the *Hoxc8* and *Hoxc9* genes, at all the *Hoxb* genes examined, the changes in histone acetylation were very similar in wild-type and knock-out cells, indicating that it is not GCN5 that specifically is able to acetylate histones at the *Hoxb* promoters. Rather, either it plays no role in acetylation at these genes, or there are other enzymes that are able to compensate for GCN5. Thus, the mechanisms for histone acetylation, and most probably gene activation are not conserved at *Hox* gene clusters. The GCN5 enzyme most probably operates in two different complexes in mammals (Nagy & Tora, 2007). Firstly, in a complex

analogous to the yeast transcriptional activator SAGA complex, and secondly in a complex known as ATAC. The possible localisation of these complexes at the *Hoxc8* and *Hoxc9* genes would be an interesting further line of investigation.

6.4.3 Reversible effects of Valproate on pluripotent gene expression

Again, whilst treating with valproate, interesting, possibly linked, global phenomena occurred that also warrant further discussion. As mentioned above, the *pou5f1* and *Nanog* gene promoters show no change in histone acetylation in response to valproate, However, they do show a change in H3K4me3 levels. This change is a decrease that is contrary to the global increase seen in this modification. The decrease in H3K4me3 levels at the promoters of *Pou5f1* and *Nanog* are concomitant with a decrease in expression from these gene loci. In addition, a change in the cellular morphology upon treatment with valproate was observed, that hinted at differentiation of the cells. Since the changes in gene expression at *Pou5f1* and *Nanog* are coordinate with changes in histone modifications, it is difficult to say whether the histone modification change is a requirement or a consequence of transcription. The decrease in expression from this locus is, however, not linked with a rise in H3K27me3. This is consistent with results from a previous study, suggesting *Pou5f1* repression is under a mechanism of control independent from polycomb repression (Boyer et al, 2006).

The repression of the pluripotent markers *Nanog* and *Pou5f1* seen, and changes in morphology of these embryonic stem cells induced by valproate are consistent with previous studies using the histone deacetylase inhibitor TSA (Karantzal et al, 2008; McCool et al, 2007). One of these studies combined treating ES cells with 50nM TSA

for 12 hours and performing subsequent microarray analysis. Interestingly, in this study, similar changes in morphology of the cells were seen to those seen here upon addition of valproate (Karantzal et al, 2008). When the TSA treated cDNA populations were compared with cDNA generated from cells differentiated into embryoid bodies, genes were up or down regulated in the same direction under both conditions. In the second study, treatment with TSA was again shown to recapitulate the effects of differentiating by removal of LIF. Effects monitored include histone acetylation increases and induction of gene expression at *Hoxb1* and *Nestin* (a neuroectoderm marker), and a reduction of expression of *Nanog* and *Pou5f1*. Importantly in this study, markers of all three germ lineages were induced by both LIF withdrawal and TSA treatment, indicating a loss of pluripotency but not a specific differentiation (McCool et al, 2007). Furthermore these changes were reversible upon removal of TSA. The authors propose this represents a reversible phenotypic change in the loss of pluripotency, that requires an additional signal for the ES cells to commit to differentiation. In the studies here, the reduction of *Pou5f1* and *Nanog* expression are reversible upon wash out of the valproate (E. Boudadi, unpublished data) as are the changes in cellular morphology. Thus treatment of a HDAC inhibitor does not result in commitment to differentiation, merely a transient loss of pluripotency.

There is, of course, a second possibility, that all that is seen upon treatment with valproate is not a “higher” loss of pluripotency, but specific effects at specific valproate sensitive genes. It could, therefore be a mere coincidence that the genes displaying valproate-related effects on gene expression are genes related to pluripotency.

6.5 The effects of increasing active histone modifications at the *Hoxb* cluster

The addition of 1mM valproate for 2 and 8 hours in CCE/R cells was able to induce hyperacetylation at all of the *Hoxb* gene promoters examined as discussed above. Further to the valproate sensitivity of the acetylation marks at the *Hoxb* cluster, the H3K4me3 modification was also sensitive to valproate treatment. The H3K4me3 mark increases concomitant with the increase in acetylation. Therefore, there is an increase in two active histone modifications at the *Hoxb* cluster upon valproate treatment. These changes in histone modifications seen in response to valproate are independent of transcription from the *Hoxb* loci as assessed by changes in mRNA levels of the *Hoxb* genes. That is, despite an increase in active histone modifications at the loci, no transcription was induced. Again, this was assessed by RNA production only, and it is possible that there is a higher level of recruitment of RNA polymerase to the promoter.

6.5.1 Histone modification marks in determination of gene transcription programmes from the *Hoxb* locus

At the *Hoxb* genes studied, it is likely that the induction of transcription requires the extra signals received from retinoic acid signalling pathways. In undifferentiated ES cells, at the silent but actively marked *Hoxb* cluster, the retinoic acid signal may take its effect by removing the repressive H3K27me3 mark. The effect of valproate on *Hox* gene expression was therefore further explored in the differentiation system in an attempt to link increased levels of actively associated histone modifications with induction of gene transcription. It was postulated that, if actively associated histone modifications play an active role in "poising" *Hox* genes for expression then increases in their levels in undifferentiated cells may result in early expression upon the final

"push" from retinoic acid signalling. In addition, in a study using the HDAC inhibitor TSA, TSA treatment for 6 days resulted in inhibition of differentiation of these cells. However, upon removal of TSA, still in the absence of LIF, embryoid bodies were formed much more rapidly than had they not been pre-treated with TSA (Lee et al, 2004). Thus, HDAC treatment in the presence of the signal to differentiate may induce gene expression changes more rapidly upon subsequent differentiation, possibly due to hyperacetylation at these promoters. Cells were, therefore induced to differentiate in the presence of valproate.

In this experiment, the undifferentiated cells were first incubated in valproate for 8 hours, an incubation period shown in the previous experiment, to induce hyperacetylation at *Hoxb* gene promoters. Cells were then induced to differentiate by detachment and LIF withdrawal, and incubated in valproate for a further 8 hours. This was in order to achieve overlap between hyperacetylation, and the signal to differentiate (Figure 5.3). Retinoic acid was added two days after the initial signal to differentiate as previously. In this study no clear changes in gene expression timing from the *Hoxb* gene cluster upon treatment with valproate were observed. The cluster maintained its general feature of co-linear timing, with timing equivalent in untreated and treated cells. However, there were subtle differences between levels of gene transcripts at *Hoxb2*, *Hoxb5* and *Hoxb9* in treated and untreated cells. That is, these genes appeared to show earlier transcription in the treated cells. That this did not occur at the most 3' gene, *Hoxb1* disputes a mechanism of *Hox* gene activation whereby progressive chromatin opening induces sequential *Hox* gene transcription (Kmita & Duboule, 2003). However, it is difficult to tell whether these slight changes represent significant early induction, or just an increase in "leaky" expression that is seen at all genes before full induction. Such leaky expression may result from the

induction of the gene at different lineages of the embryoid body. Its increase may therefore result from valproate driving the cells down this different lineage. It may also be due to the low resolution of the differentiation system used. Overall, though, hyperacetylation and increased levels of H3K4me3 do not have a major effect upon the timing of *Hoxb* gene expression in differentiation when overlapped with the signal to differentiate.

6.6 Predictive, permissive marks are overridden by a repressive mechanism

The above data are consistent with evidence for predictive permissive histone modifications seen elsewhere in embryonic stem cells (Azuara et al, 2006; Bernstein et al, 2005). Such modifications are normally associated with active genes, though here are present in the absence of transcriptional activity. Even in their increase with the externally added agent, valproate, transcription cannot be induced in undifferentiated cells, or the timing altered in differentiating cells. However, in all of the scenarios examined, the H3K27me3 repressively modification was maintained, and it may be that this mark is intertwined in a dominant repressive mechanism over the activating modifications. That is, that the “blanket” of H3K27me3 seen over the *Hoxb* cluster is able to suppress any recruitment of transcriptionally activating proteins. This potential dominance of the H3K27me3 mark is confirmed in *Eed* knock-out ES cells (Boyer et al, 2006). *Eed* is a component of the PRC2 complex and works alongside *Ezh2,* the H3K27 methyl-transferase of the same complex. ES cells cannot be derived from blastocycts lacking the *Ezh2* enzyme, however ES cells deficient in the *Eed* cofactor are viable and do show a loss of the H3K27me3 mark at PRC2 targets (Montgomery et al, 2005). Importantly for this study, *Hoxb2, b4, b7* and *Hoxb13* show significant de-repression in these cells when compared with wild-type

ES cells (Boyer et al, 2006). Therefore, the loss of H3K27me3 results in the expression, in undifferentiated ES cells of a subset of the *Hox* genes examined in this study, showing a possible repressive dominance of this mark that reduces the activating marks to "predictive" marks. ES cells deficient in Suz12 are also available and this repressive dominance of the PRC2 related mark could be further investigated within these cells (Suz12 is another cofactor for the EZH2 enzyme activity) (Pasini et al, 2007). Furthermore, in another study, differentiating ES cells by concomitant retinoic acid addition and removal of LIF showed the loss of H3K27me3 at *Hoxb1* the retinoic acid response element coordinated with its gene expression upon differentiation (Lee et al, 2007a). Whilst this does not show removal of H3K27me3 to be the limiting factor in induction of expression, the coincident timing of the two events is significant with regards to a possible dominant repressive role for the H3K27me3 mark.

Contrary to data seen in this study with valproate, the HDAC inhibitor TSA was able to induce *Hoxb1* expression after 5 hours of treatment in mouse ES cells, suggesting that active histone modifications may sometimes override repressive modifications at this gene (McCool et al, 2007). The most likely cause of this difference is the different in HDAC inhibitor used. As mentioned above at the beginning of chapter 3, valproate subjects the cells to some non-specific effects to which TSA is not prone. Most, notably, the disruption of Wnt signalling (a pathway involved in stem cell differntiation) might here be most relevant as an explanation as to why *Hoxb1* may be induced by treatment with TSA and not valproate. The chelation of zinc and the sequestration of acetyl CoA may also provide some explanation, as might the need to use valproate at a much higher concentration than TSA. Such a higher concentration might exaggerate the non-specific effects of valproate with comparison

to TSA. It is possible also that their non HDAC deacteylase targets are slightly different, resulting the activation or repression of different signalling pathways. In addition, some differences between the two sets of experiments may be explained by the differences in ES cell line used. Whilst differences in gene expression patterns between mouse ES cell lines are subtle, they do exist, and so could result in slightly different responses to valproate (Sharova et al, 2007). In another study, treatment of undifferentiated ES cells with TSA for 14 hours upon removal of LIF, but without addition of retinoic acid resulted in the induction of expression of *Hoxa1* (Lee et al, 2007a). These conditions resulted in histone hyperacetylation but no removal of H3K27me3. In this same study, differentiation by removal of LIF and addition of retinoic acid results in the removal of this mark and histone hyperacetylation at *Hoxa1*. TSA treatment was also a requirement for *Hoxa1* expression upon LIF removal in the absence of retinoic acid treatment as removal of LIF alone does not induce a rise in H4 acetylation levels. Therefore, here, at *Hoxa1* histone acetylation caused by an external factor appears to be able to induce gene transcription that overrides any H3K27me3 present at the gene, so loss of H3K27me3 is not the final determining factor in transcription, though it does occur during "normal" differentiation. The experiments performed at the *HoxA* locus are most similar to the experiments performed here with valproate whereupon the genes were pre-treated and then induced to differentiate by the removal of LIF, overlapping with the valproate treatment for 8 hours. The most equivalent time point between this study and that at the *Hoxa* cluster is that after 8 hours of LIF removal. However, here, valproate is not able to induce expression of the *Hoxb* genes, despite inducing histone hyperacetylation. Indeed, *Hoxb* gene expression continues in a co-linear fashion much as in control cells. Since the main discrepancies between the two studies are the *Hox* genes examined and the HDAC inhibitor used, these must be the

main reason for any differences observed. Therefore, it is most likely that the *Hox* genes have differing control mechanisms, as seen in the different use of GCN5 at *Hoxb* and *Hoxc*, and utilise different tailored mechanisms for control of gene expression. The use of a different histone deacetylase inhibitor could, too be of significance. The two inhibitors may have subtly different specificities and this could even be applicable to non-histone targets such as transcription factors(Thorne et al, 2009).

There is further evidence that suggests that H3K27me3 is not a "master" regulator as indicated by the data in this study at *Hoxb.* Such evidence may be found in *Suz12* knock-out ES cells (Pasini et al, 2007). *Suz12* is a component of the PRC2 complex that is required for the H3K27me3 methylase activity of the complex These cells show a significant loss of the H3K27me3 mark and an increased expression of some but not all of the PRC2 targets that have lost this mark. Therefore, H3K27me3 is not always a "dominant repressive" mechanism and alternatives must be explored to explain the presence of active marks at the transcriptionally inactive *Hoxb* cluster in embryonic stem cells.

6.6.1 Alternative repressive control at the *Hoxb* cluster

The increase in acetylation at the *Hoxb* genes in response to valproate in undifferentiated ES cells here is consistent with data shown in embryonic stem cells at *Hoxb1* and *Hoxb9* using the histone deacetylase inhibitor TSA. In this study, TSA was able to induce hyperacetylation at these genes independent of a change in transcription (Chambeyron & Bickmore, 2004). Also, in this study, significant levels of activating modifications were seen at *Hoxb1* and *Hoxb9* before their transcription as in the experiments described in this thesis. However, an alternative explanation was proposed for the repression of transcription at this cluster not related to repressive

histone modifications, namely changes in the higher order genome architecture(Chambeyron & Bickmore, 2004). Upon activation, a large decondensation of the locus was seen, something that is not recreated by treatment with TSA. Regions of the cluster were also seen to loop away from their chromosome territory upon activation. Thus, for the cluster to become transcriptionally competent, it must overcome chromatin condensation. Note, however, chromatin decondensation could still result in the loss of H3K27me3 from the *Hoxb* genes, possibly in a complex interplay, that should not be thought of as hierarchical but cooperative. A cooperative mechanism could coordinate the multiple transcription factors that must be involved in this complex interplay. For example, retinoic acid signalling is known to be involved in the activation of the *Hox* genes, and therefore this too must play a role in activation of gene expression.

A discussion on gene silencing would be incomplete without the inclusion of DNA methylation. There is evidence that DNA methylation plays a role in *Hox* gene silencing. The polycomb component EZH2 may associate with the DNA methyl transferase for more permanent silencing of polycomb targets, such as the *Hox* genes (Vire et al, 2006). In addition, *Hox* genes were identified as CpG island containing tissue-specifically methylated genes, indicating a role for DNA methylation in their tissue-specific expression (Illingworth et al, 2008). *Hox* genes were also found to be derepressed in MEFs lacking *Lsh* a component involved in DNA methylation, though this study did not include a survey of the *Hoxb* cluster (Xi et al, 2007). Thus, although the evidence points towards DNA methylation being important in control of *Hox* genes, it also suggests that its role comes at a later stage in development than in the study here. Thus, it is unlikely that DNA methylation is the mechanism by which the activating modifications seen in undifferentiated cells are overridden.

6.6.2 Cross-talk implements the predictive code

The histone code hypothesis predicts that it is combinations of modifications that determine transcriptional outcomes. These combinations may serve to bring in multiple or individual proteins that may recognise more than one histone mark. They may also result in the ejection of proteins from chromatin. For such a hypothesis to be true, there must be mechanisms in place in cells for the deposition of such combinatorial marks. These may be manifest in cross-talk between histone modifications through their respective enzymes. For example, cross-talk may arise from different activities acting as complexes such as the MLL and MOF enzymes, histone methyl transferases and histone acetyl transferases that have been found in a complex together (Dou et al, 2005). Thus, the localisation of this complex to chromatin results in the deposition of these marks at similar locations on genes. Combinatorial modifications may also arise from one modification resulting in the recruitment of a second histone-modifying enzyme. For example, in yeast, the Yng1 PHD finger is able to bind to the H3K4me3 modification and recruit histone acetyl transferase activity to these loci. Here, there is evidence for cross-talk between H3K4me3 and histone acetylation as has already been described in HL60 cells (Nightingale et al, 2007). In this study in ES cells at the *Hoxb* cluster, the evidence for such a cross-talk is found firstly in the correlation between the patterns of H3K9ac and H3K4me3 modifications at the *Hoxb* cluster, and secondly in their similar responses to valproate. The correlation of the two marks means that the pattern of modifications across the promoters of the *Hoxb* genes is similar for the two active modifications (Figure 3.1). Indeed, if plotted against each other, a highly significant correlation can be calculated in both CCE/R cells and wild type 3G4 cells (Figure 6.1a and b). This cross talk also means that a significant increase in response to

valproate is observed at the promoters of *Hoxb* genes for both these modifications in CCE/R cells. Interestingly, in GCN5 knock-out 3G4 cells, there is no such significant correlation between the two marks (Figure 6.1c). However, there is one outlier, *Hoxb5* that is skewing the regression analysis, resulting in an insignificant correlation. Removal of this outlier results in an r^2 of 0.8. Thus, it cannot be concluded that GCN5 has a pivotal role to play in the cross talk of these marks. In these 3G4 cells, however, in neither wild type nor knock-out there is not an increase in H3K4me3 upon valproate treatment. This does indicates a difference in histone cross talk in these cells most likely due to the different ES cell lines used.

A self perpetuating mechanism has been proposed that may explain the link between the histone modifications H3K4me3 and H3K9ac (Figure 6.2) (Vermilyea *et al*, in press). The balanced action of histone acetyl transferases, and histone deactylases result in a dynamic level of histone acetylation at the H3 tail (Figure 6.2a). Acetylated peptides are shown to be better substrates for the lysine-specific methyl transferases (Figure 6.2b) (Nightingale et al, 2007). Therefore, any tail with high levels of histone acetylation, is likely to have high levels of H3K4methylation as shown here at the *Hox* genes. H3K4 methylation is then able to recruit the SAGA/Slk histone acetyl transferase via the H3K4me3-binding Chd1 thereby maintaining acetylation levels at these loci (Pray-Grant et al, 2005)(Figure 6.2c). Therefore, the levels of the two modifications would be maintained in close concert in a sort of “feed forward” loop. This would be still be true if the cells are treated with an HDAC inhibitor, and may explain why the levels of H3K4me3 increase upon treatment with the HDAC inhibitor

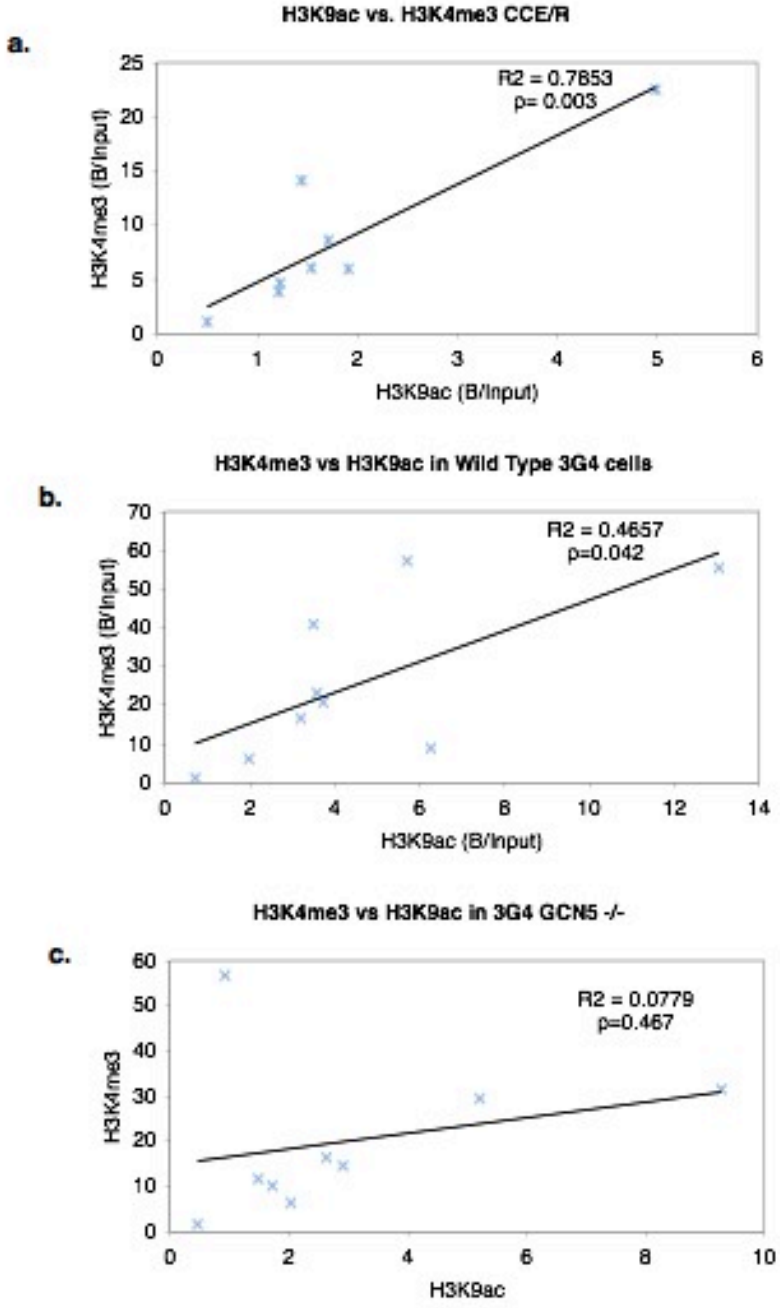


Figure 6.1 Correlation of H3K4me3 and H3K27me3. The levels of H3K9ac and H3K4me3 at each of the *Hoxb* genes in undifferentiated **a.** CCE/R cells **b** 3G4 wild type cells and **c.** 3G4 Gcn5 -/- at Hoxb9 genes were plotted against one another and an R^2 value calculated.

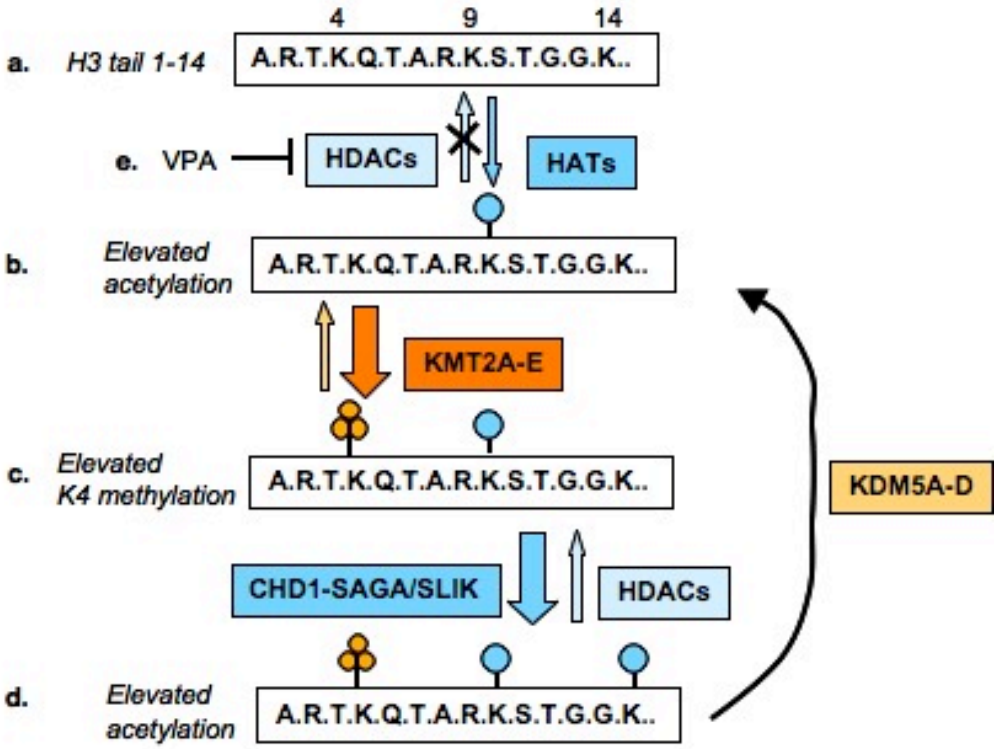


Figure 6.2 An epigenetic feedforward loop. a. Histone acetyl transferases and histone deactetlyases act on the H3 tail to result in a dynamic level of lysine acetylation (blue circle). **b.** elevated acetylation result in a better substrate for lysine methylases resulting in increased H3K4 trimethylation (orange circles) which, in turn results in the recruitment of Chd1 by its chromodomain. This brings in further acetylase activity **d.** The action of a lysine specific demethylase brings the histone tail to the stage in b. The addition of an HDAC inhibitor **e.**, results in increased acetylation in **b.**, and therefore an increase in H3K4 trimethylation. Adapted from Vermilyea *et al*, 2009

(Figure 6.2 b, d). Therefore, this model may explain cross talk between H3K9ac and H3K4me3 at the *Hoxb* cluster. Further study into histone cross-talk may provide further interactions that are involved in the placement of a "code". Here, at the *Hoxb* cluster this code serves in a predictive fashion as there is no transcription from the *Hoxb* genes despite the presence of active modifications.

6.7 Memory of histone modification changes

There is a final reason for choosing the *Hox* gene cluster as a model system in which to study histone modifications. Histone modifications are often defined as epigenetic modifications. The epigenetic information is the information necessary to interpret the underlying DNA sequence and such information is believed to be inherited as part of a cell "memory". That is, for a cell to retain its identity, it must retain its gene expression programme from one generation to the next via these epigenetic mechanisms. Due to their association with active or inactive patterns of gene expression, it has been postulated that histone modifications are inherited through the cell cycle as part of this epigenetic programme and are responsible for maintaining gene expression programmes through an active role in transcription.

The role of histone modifications in "epigenetic" maintenance of gene expression programmes is, however, widely disputed (Ptashne, 2007). For histone modifications to be truly epigenetic, they would need to be inherited through the cell cycle. This includes two critical phases where epigenetic modifications could be wiped clean. Firstly, in mitosis there is no gene expression and so expression patterns must be reintroduced in G1, and secondly in S-phase epigenetic mechanisms must be re-introduced to newly synthesised strands of DNA. The mechanistic explanation of how

histone modifications might be propagated through S-phase is still very much an ongoing work (Probst et al, 2009). For example, a recent study attempting to identify the mechanisms that transmit H3K27methylation through the cell cycle showed EZH2, the enzyme that deposits this mark, at the site of DNA replication, but was unable to coordinate this with the cell cycle inheritance of H3K27me3, EZH2 association and gene repression (Hansen et al, 2008).

6.7.1 The inheritance of histone modifications through mitosis

The inheritance of histone modifications through mitosis is potentially highly significant as they could then be responsible for re-introducing gene expression programmes in G1. Memory of this transcription is not maintained by transcription factors since most sequence-specific transcription factors are removed from the DNA during mitosis (Martinez-Balbas et al, 1995). Histone acetyl modifying enzymes, too are removed from the DNA during mitosis leaving histone modification patterns static (Kruhlak et al, 2001). There have been few studies that display inheritance of histone modification changes through mitosis. An initial indication that histone modifications could be remembered through mitosis was observed in the yeast *S. pombe*. Yeast were grown for several cell cycles in TSA, a treatment that was able to induce hyperacetylation at normally silent genes inserted into centric heterochromatin. Both the acetylation, and the active state of these genes was maintained through several cell cycles. However this concomitant inheritance of both functions (i.e. both marks and transcriptional activity) makes it difficult to resolve if it is the acetylation that is maintaining transcription, or vice versa. More recently, inappropriate transcriptional states were shown to be inherited in nuclear transfer studies in *Xenopus*. The first nuclear transfer took a muscle lineage nucleus and placed it in a recipient enucleated egg, the second round of transfer took non-muscle cells from the resulting blastula

and transferred this nucleus to a recipient enucleated egg. In this second embryo the muscle gene *MyoD* was expressed in non-muscle lineages even after these two successive nuclear transfer experiments. Therefore, even after two periods of 12 cell cycles of transcriptional silence after nuclear transfer, memory of *MyoD* expression remained (Ng & Gurdon, 2008). This inappropriate expression was reliant upon the histone H3.3 and its associated H3K4me3 modification, providing evidence for a role of chromatin in epigenetic memory. Finally, in an attempt to analyse memory though the cell cycle studies looking at active genes at defined points in the cell cycle have shown that histone modifications do remain in mitosis on genes that were active throughout G2 phase (Kouskouti & Talianidis, 2005; Valls et al, 2005).

The classical example for stable maintenance of gene expression programmes can be found in the polycomb and trithorax proteins. In the developing embryo, *Hox* gene expression programmes must be maintained into adulthood. The trithorax and polycomb group proteins are known to be largely responsible for maintenance at these genes, where this association has been proposed to be propagated by the maintenance of their associated histone modifications, H3K4me3 and H3K27me3. This is because the proteins themselves are displaced from chromatin during mitosis (Buchenau et al, 1998). Therefore the *Hox* genes were chosen in this study as a good model in which to study potential inheritance in externally-induced histone modifications. Indeed, a role for histone modifications in epigenetics is not possible if a change in their levels cannot be maintained.

Recent experiments in the early embryo have demonstrated that memory mechanisms for externally-induced histone modifications in the early embryo at *Hoxb* genes may exist (verMilyea *et al* in press). When embryos were cultured in the

presence of valproate from the 8-cell to the morula stage, the valproate washed out, and the embryos further cultured to blastocyst, higher levels of both histone acetylation and H3K4 trimethylation were observed in the treated versus the untreated blastocysts. Thus, changes in histone modifications earlier induced by valproate were remembered later in differentiation. Interestingly, the initial increase in "activating" histone modifications, and the maintenance of these modifications occurred in the absence of any transcription from the locus. Therefore, the memory mechanism was transcriptionally independent.

Given the results in the early embryo, it was decided to extend the valproate treatment studies here to analyse memory mechanisms present over the *Hox* clusters in undifferentiated CCE/R cells. Undifferentiated CCE/R cells were treated with valproate for 8 hours, the valproate washed-out, and any memory of valproate-induced histone modification changes analysed 48 hours later. It is known that such a wash-out results in rapid removal of any global histone hyperacetylation, however, the effect upon individual genes in this system is not known. An 8 hour treatment and wash-out did not result in any visible memory of histone modifications changes at the *Hoxb* genes. Both an initial 8 hour treatment and a 16 hour treatment resulted in hyper H3K9 acetylation and increases in H3K4 tri-methylation as seen previously, but these were consistently not maintained over the 48 hour period. Moreover, again the increases in 'active" modifications were transcription independent, both at their initial increase, and upon their wash-out 48 hours later. That is, at all timepoints examined, there was no detectable transcription from these genes. Thus there is a major difference in ES cells and early embryos, in that changes in histone modifications are not maintained in ES cells at *Hoxb* genes, but are maintained in early embryos. This is despite the similar treatment times (16hrs in ES cells and

18hrs in early embryos). However, the two studies do show commonality in their transcription-independence of histone modifications. That is, an increase in active histone modifications in both studies failed to result in an increase in transcription.

The discrepancies in histone modifications seen in the early embryo and ES cells are surprising given the mechanism discussed above that nicely models cross talk at the *Hoxb* locus (Figure 6.2). Given the likelihood that these enzymes operate in both the early embryo and in ES cells, it is surprising that no memory of histone modification changes was seen in ES cells. This could possibly be the result of different enzyme complexes being utilised at these different stages of development, therefore resulting in the cross-talk seen in ES cells not persisting into memory in this system.

The differences in memory of histone modification changes seen between memory in ES cells and in early embryos could, undoubtedly, result from the difference in developmental stage the embryonic stem cells represent compared with the early stages examined in early embryos. ES cells are derived from the epiblast of the inner cell mass of the mouse blastocyst. Thus, they represent a stage post that treated with valproate in the studies on the early embryo. Therefore, could it be that memory mechanisms are only displayed in response to valproate at this early stage? Possibly, this represents an early stage in which epigenetic mechanisms are put in place, and therefore any external alterations would be remembered along with the embryos normal epigenetic programme. Therefore, a later treatment with valproate at the blastocycst stage might recapitulate the findings found in this study, i.e. that external alterations in histone modifications are not remembered at the *Hoxb* cluster. Alternatively, the validity of ES cells as an *in vitro* model for epiblast cells has also been called into question (Silva & Smith, 2008). Therefore, the possible

discrepancies may be simply that ES cells are not a valid model for *in vivo* studies. However, the ability to differentiate ES cells into many different lineages contradicts this (Murry & Keller, 2008). That is, ES cells are able to respond to epigenetic alterations from external signals to differentiate. Thus, the discrepancies seen between the two studies are likely to arise due to the differences in developmental stage examined, with the situation in ES cells representing a true picture of the epigenetic mechanisms within epiblast cells of the blastocyst.

Thus the active histone modifications at the *Hoxb c*luster show a high turnover, but are not directly linked with transcription. This could be an effect of a dominant repressive H3K27me3 mark. Furthermore the modifications H3K9ac and H3K4me3 are highly correlated but this close relationship does not self perpetuate to leave a memory of externally-induced increases in histone modifications. Therefore transcription independent cross-talk is seen but the dynamic nature of the modifications at the cluster in ES cells means that the equilibrium soon finds its natural balance after external peturbation.

7 Bibliography

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

Agger K, Cloos PA, Christensen J, Pasini D, Rose S, Rappsilber J, Issaeva I, Canaani E, Salcini AE, Helin K (2007) UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* **449:** 731-734

Akasaka T, Kanno M, Balling R, Mieza MA, Taniguchi M, Koseki H (1996) A role for mel-18, a Polycomb group-related vertebrate gene, during theanteroposterior specification of the axial skeleton. *Development (Cambridge, England)* **122:** 1513-1522

Akasaka T, van Lohuizen M, van der Lugt N, Mizutani-Koseki Y, Kanno M, Taniguchi M, Vidal M, Alkema M, Berns A, Koseki H (2001) Mice doubly deficient for the Polycomb Group genes Mel18 and Bmi1 reveal synergy and requirement for maintenance but not initiation of Hox gene expression. *Development (Cambridge, England)* **128:** 1587-1597

Allfrey VG, Faulkner R, Mirsky AE (1964) Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **51:** 786-794

Archer TK, Lefebvre P, Wolford RG, Hager GL (1992) Transcription factor loading on the MMTV promoter: a bimodal mechanism for promoter activation. *Science (New York, NY* **255:** 1573-1576

Azuara V, Perry P, Sauer S, Spivakov M, Jorgensen HF, John RM, Gouti M, Casanova M, Warnes G, Merkenschlager M, Fisher AG (2006) Chromatin signatures of pluripotent cell lines. *Nature cell biology* **8:** 532-538

Bannister AJ, Kouzarides T (2005) Reversing histone methylation. *Nature* **436:** 1103-1106

Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, Kouzarides T (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410:** 120-124

Barrera LO, Li Z, Smith AD, Arden KC, Cavenee WK, Zhang MQ, Green RD, Ren B (2008) Genome-wide mapping and analysis of active promoters in mouse embryonic stem cells and adult organs. *Genome research* **18:** 46-59

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K (2007) High-resolution profiling of histone methylations in the human genome. *Cell* **129:** 823-837

Becker CM, Harris RA (1983) Influence of valproic acid on hepatic carbohydrate and lipid metabolism. *Archives of biochemistry and biophysics* **223:** 381-392

Berger C, Strub A, Staib C, Lepke M, Zisimopoulou P, Hoehn K, Nanda I, Schmid M, Grummt F (1999) Identification and characterization of a mouse homolog to yeast Cdc6p. *Cytogenetics and cell genetics* **86:** 307-316

Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ, 3rd, Gingeras TR, Schreiber SL, Lander ES (2005) Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120:** 169-181

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125:** 315-326

Berstine EG, Hooper ML, Grandchamp S, Ephrussi B (1973) Alkaline phosphatase activity in mouse teratoma. *Proceedings of the National Academy of Sciences of the United States of America* **70:** 3899-3903

Bhattacharyya N, Dey A, Minucci S, Zimmer A, John S, Hager G, Ozato K (1997) Retinoid-induced chromatin structure alterations in the retinoic acid receptor beta2 promoter. *Molecular and cellular biology* **17:** 6481-6490

Bird A, Tweedie S (1995) Transcriptional noise and the evolution of gene number. *Philosophical transactions of the Royal Society of London* **349:** 249-253

Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, Thurman RE, Kuehn MS, Taylor CM, Neph S, Koch CM, Asthana S, Malhotra A, Adzhubei I, Greenbaum JA, Andrews RM, Flicek P, Boyle PJ, Cao H, Carter NP, Clelland GK, Davis S, Day N, Dhami P, Dillon SC, Dorschner MO, Fiegler H, Giresi PG, Goldy J, Hawrylycz M, Haydock A, Humbert R, James KD, Johnson BE, Johnson EM, Frum TT, Rosenzweig ER, Karnani N, Lee K, Lefebvre GC, Navas PA, Neri F, Parker SC, Sabo PJ, Sandstrom R, Shafer A, Vetrie D, Weaver M, Wilcox S, Yu M, Collins FS, Dekker J, Lieb JD, Tullius TD, Crawford GE, Sunyaev S, Noble WS, Dunham I, Denoeud F, Reymond A, Kapranov P, Rozowsky J, Zheng D, Castelo R, Frankish A, Harrow J, Ghosh S, Sandelin A, Hofacker IL, Baertsch R, Keefe D, Dike S, Cheng J, Hirsch HA, Sekinger EA, Lagarde J, Abril JF, Shahab A, Flamm C, Fried C, Hackermuller J, Hertel J, Lindemeyer M, Missal K, Tanzer A, Washietl S, Korbel J, Emanuelsson O, Pedersen JS, Holroyd N, Taylor R, Swarbreck D, Matthews N, Dickson MC, Thomas DJ, Weirauch MT, Gilbert J, Drenkow J, Bell I, Zhao X, Srinivasan KG, Sung WK, Ooi HS, Chiu KP, Foissac S, Alioto T, Brent M, Pachter L, Tress ML, Valencia A, Choo SW, Choo CY, Ucla C, Manzano C, Wyss C, Cheung E, Clark TG, Brown JB, Ganesh M, Patel S, Tammana H, Chrast J, Henrichsen CN, Kai C, Kawai J, Nagalakshmi U, Wu J, Lian Z, Lian J, Newburger P, Zhang X, Bickel P, Mattick JS, Carninci P, Hayashizaki Y, Weissman S, Hubbard T, Myers RM, Rogers J, Stadler PF, Lowe TM, Wei CL, Ruan Y, Struhl K, Gerstein M, Antonarakis SE, Fu Y, Green ED, Karaoz U, Siepel A, Taylor J, Liefer LA, Wetterstrand KA, Good PJ, Feingold EA, Guyer MS, Cooper GM, Asimenos G, Dewey CN, Hou M, Nikolaev S, Montoya-Burgos JI, Loytynoja A, Whelan S, Pardi F, Massingham T, Huang H, Zhang NR, Holmes I, Mullikin JC, Ureta-Vidal A, Paten B, Seringhaus M, Church D, Rosenbloom K, Kent WJ, Stone EA, Batzoglou S, Goldman N, Hardison RC, Haussler D, Miller W,

Sidow A, Trinklein ND, Zhang ZD, Barrera L, Stuart R, King DC, Ameur A, Enroth S, Bieda MC, Kim J, Bhinge AA, Jiang N, Liu J, Yao F, Vega VB, Lee CW, Ng P, Shahab A, Yang A, Moqtaderi Z, Zhu Z, Xu X, Squazzo S, Oberley MJ, Inman D, Singer MA, Richmond TA, Munn KJ, Rada-Iglesias A, Wallerman O, Komorowski J, Fowler JC, Couttet P, Bruce AW, Dovey OM, Ellis PD, Langford CF, Nix DA, Euskirchen G, Hartman S, Urban AE, Kraus P, Van Calcar S, Heintzman N, Kim TH, Wang K, Qu C, Hon G, Luna R, Glass CK, Rosenfeld MG, Aldred SF, Cooper SJ, Halees A, Lin JM, Shulha HP, Zhang X, Xu M, Haidar JN, Yu Y, Ruan Y, Iyer VR, Green RD, Wadelius C, Farnham PJ, Ren B, Harte RA, Hinrichs AS, Trumbower H, Clawson H, Hillman-Jackson J, Zweig AS, Smith K, Thakkapallayil A, Barber G, Kuhn RM, Karolchik D, Armengol L, Bird CP, de Bakker PI, Kern AD, Lopez-Bigas N, Martin JD, Stranger BE, Woodroffe A, Davydov E, Dimas A, Eyras E, Hallgrimsdottir IB, Huppert J, Zody MC, Abecasis GR, Estivill X, Bouffard GG, Guan X, Hansen NF, Idol JR, Maduro VV, Maskeri B, McDowell JC, Park M, Thomas PJ, Young AC, Blakesley RW, Muzny DM, Sodergren E, Wheeler DA, Worley KC, Jiang H, Weinstock GM, Gibbs RA, Graves T, Fulton R, Mardis ER, Wilson RK, Clamp M, Cuff J, Gnerre S, Jaffe DB, Chang JL, Lindblad-Toh K, Lander ES, Koriabine M, Nefedov M, Osoegawa K, Yoshinaga Y, Zhu B, de Jong PJ (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447:** 799-816

Bonner WM, West MH, Stedman JD (1980) Two-dimensional gel analysis of histones in acid extracts of nuclei, cells, and tissues. *European journal of biochemistry / FEBS* **109:** 17-23

Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Wernig M, Tajonar A, Ray MK, Bell GW, Otte AP, Vidal M, Gifford DK, Young RA, Jaenisch R (2006) Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441:** 349-353

Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K (2006) Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes & development* **20:** 1123-1136

Breiling A, Turner BM, Bianchi ME, Orlando V (2001) General transcription factors bind promoters repressed by Polycomb group proteins. *Nature* **412:** 651-655

Briggs SD, Bryk M, Strahl BD, Cheung WL, Davie JK, Dent SY, Winston F, Allis CD (2001) Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in Saccharomyces cerevisiae. *Genes & development* **15:** 3286-3295

Buchenau P, Hodgson J, Strutt H, Arndt-Jovin DJ (1998) The distribution of polycomb-group proteins during cell division and development in Drosophila embryos: impact on models for silencing. *The Journal of cell biology* **141:** 469-481

Bug G, Gul H, Schwarz K, Pfeifer H, Kampfmann M, Zheng X, Beissert T, Boehrer S, Hoelzer D, Ottmann OG, Ruthardt M (2005) Valproic acid stimulates proliferation and self-renewal of hematopoietic stem cells. *Cancer research* **65:** 2537-2541

Byrd KN, Shearn A (2003) ASH1, a Drosophila trithorax group protein, is required for methylation of lysine 4 residues on histone H3. *Proceedings of the National Academy of Sciences of the United States of America* **100:** 11535-11540

Cao R, Tsukada Y, Zhang Y (2005) Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Molecular cell* **20:** 845-854

Cao R, Wang H, He J, Erdjument-Bromage H, Tempst P, Zhang Y (2008) Role of hPHF1 in H3K27 methylation and Hox gene silencing. *Molecular and cellular biology* **28:** 1862-1872

Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y (2002) Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science (New York, NY* **298:** 1039-1043

Cao R, Zhang Y (2004) SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. *Molecular cell* **15:** 57-67

Carmen AA, Rundlett SE, Grunstein M (1996) HDA1 and HDA3 are components of a yeast histone deacetylase (HDA) complex. *The Journal of biological chemistry* **271:** 15837-15844

Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, Oyama R, Ravasi T, Lenhard B, Wells C, Kodzius R, Shimokawa K, Bajic VB, Brenner SE, Batalov S, Forrest AR, Zavolan M, Davis MJ, Wilming LG, Aidinis V, Allen JE, Ambesi-Impiombato A, Apweiler R, Aturaliya RN, Bailey TL, Bansal M, Baxter L, Beisel KW, Bersano T, Bono H, Chalk AM, Chiu KP, Choudhary V, Christoffels A, Clutterbuck DR, Crowe ML, Dalla E, Dalrymple BP, de Bono B, Della Gatta G, di Bernardo D, Down T, Engstrom P, Fagiolini M, Faulkner G, Fletcher CF, Fukushima T, Furuno M, Futaki S, Gariboldi M, Georgii-Hemming P, Gingeras TR, Gojobori T, Green RE, Gustincich S, Harbers M, Hayashi Y, Hensch TK, Hirokawa N, Hill D, Huminiecki L, Iacono M, Ikeo K, Iwama A, Ishikawa T, Jakt M, Kanapin A, Katoh M, Kawasawa Y, Kelso J, Kitamura H, Kitano H, Kollias G, Krishnan SP, Kruger A, Kummerfeld SK, Kurochkin IV, Lareau LF, Lazarevic D, Lipovich L, Liu J, Liuni S, McWilliam S, Madan Babu M, Madera M, Marchionni L, Matsuda H, Matsuzawa S, Miki H, Mignone F, Miyake S, Morris K, Mottagui-Tabar S, Mulder N, Nakano N, Nakauchi H, Ng P, Nilsson R, Nishiguchi S, Nishikawa S, Nori F, Ohara O, Okazaki Y, Orlando V, Pang KC, Pavan WJ, Pavesi G, Pesole G, Petrovsky N, Piazza S, Reed J, Reid JF, Ring BZ, Ringwald M, Rost B, Ruan Y, Salzberg SL, Sandelin A, Schneider C, Schonbach C, Sekiguchi K, Semple CA, Seno S, Sessa L, Sheng Y, Shibata Y, Shimada H, Shimada K, Silva D, Sinclair B, Sperling S, Stupka E, Sugiura K, Sultana R, Takenaka Y, Taki K, Tammoja K, Tan SL, Tang S, Taylor MS, Tegner J, Teichmann SA, Ueda HR, van Nimwegen E, Verardo R, Wei CL, Yagi K, Yamanishi H, Zabarovsky E, Zhu S, Zimmer A, Hide W, Bult C, Grimmond SM, Teasdale RD, Liu ET, Brusic V, Quackenbush J, Wahlestedt C, Mattick JS, Hume DA, Kai C, Sasaki D, Tomaru Y, Fukuda S, Kanamori-Katayama M, Suzuki M, Aoki J, Arakawa T, Iida J, Imamura K, Itoh M, Kato T, Kawaji H, Kawagashira N, Kawashima T, Kojima M, Kondo S, Konno H, Nakano K, Ninomiya N, Nishio T, Okada M, Plessy C, Shibata K, Shiraki T, Suzuki S, Tagami M, Waki K, Watahiki A, Okamura-Oho Y, Suzuki H, Kawai J, Hayashizaki Y (2005) The transcriptional landscape of the mammalian genome. *Science (New York, NY* **309:** 1559-1563

Challen GA, Little MH (2006) A side order of stem cells: the SP phenotype. *Stem cells (Dayton, Ohio)* **24:** 3-12

Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A (2003) Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113:** 643-655

Chambeyron S, Bickmore WA (2004) Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes & development* **18:** 1119-1130

Chambeyron S, Da Silva NR, Lawson KA, Bickmore WA (2005) Nuclear re-organisation of the Hoxb complex during mouse embryonic development. *Development (Cambridge, England)* **132:** 2215-2223

Chang YL, Peng YH, Pan IC, Sun DS, King B, Huang DH (2001) Essential role of Drosophila Hdac1 in homeotic gene silencing. *Proceedings of the National Academy of Sciences of the United States of America* **98:** 9730-9735

Cheung P, Tanner KG, Cheung WL, Sassone-Corsi P, Denu JM, Allis CD (2000) Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation. *Molecular cell* **5:** 905-915

Chi AS, Bernstein BE (2009) Developmental biology. Pluripotent chromatin state. *Science (New York, NY* **323:** 220-221

Chiba H, Muramatsu M, Nomoto A, Kato H (1994) Two human homologues of Saccharomyces cerevisiae SWI2/SNF2 and Drosophila brahma are transcriptional coactivators cooperating with the estrogen receptor and the retinoic acid receptor. *Nucleic acids research* **22:** 1815-1820

Christensen J, Agger K, Cloos PA, Pasini D, Rose S, Sennels L, Rappsilber J, Hansen KH, Salcini AE, Helin K (2007) RBP2 belongs to a family of demethylases, specific for tri-and dimethylated lysine 4 on histone 3. *Cell* **128:** 1063-1076

Clapier CR, Cairns BR (2009) The Biology of Chromatin Remodeling Complexes. *Annual review of biochemistry*

Clayton AL, Hazzalin CA, Mahadevan LC (2006) Enhanced histone acetylation and transcription: a dynamic perspective. *Molecular cell* **23:** 289-296

Clayton AL, Rose S, Barratt MJ, Mahadevan LC (2000) Phosphoacetylation of histone H3 on c-fos- and c-jun-associated nucleosomes upon gene activation. *The EMBO journal* **19:** 3714-3726

Core N, Bel S, Gaunt SJ, Aurrand-Lions M, Pearce J, Fisher A, Djabali M (1997) Altered cellular proliferation and mesoderm patterning in Polycomb-M33-deficient mice. *Development (Cambridge, England)* **124:** 721-729

Cuthbert GL, Daujat S, Snowden AW, Erdjument-Bromage H, Hagiwara T, Yamada M, Schneider R, Gregory PD, Tempst P, Bannister AJ, Kouzarides T (2004) Histone deimination antagonizes arginine methylation. *Cell* **118:** 545-553

Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V (2002) Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* **111:** 185-196

d Graaff W, Tomotsune D, Oosterveen T, Takihara Y, Koseki H, Deschamps J (2003) Randomly inserted and targeted Hox/reporter fusions transcriptionally silenced in Polycomb mutants. *Proceedings of the National Academy of Sciences of the United States of America* **100:** 13362-13367

de Krom M, van de Corput M, von Lindern M, Grosveld F, Strouboulis J (2002) Stochastic patterns in globin gene expression are established prior to transcriptional activation and are clonally inherited. *Molecular cell* **9:** 1319-1326

de la Cruz X, Lois S, Sanchez-Molina S, Martinez-Balbas MA (2005) Do protein motifs read the histone code? *Bioessays* **27:** 164-175

del Mar Lorente M, Marcos-Gutierrez C, Perez C, Schoorlemmer J, Ramirez A, Magin T, Vidal M (2000) Loss- and gain-of-function mutations show a polycomb group function for Ring1A in mice. *Development (Cambridge, England)* **127:** 5093-5100

Deschamps J, van den Akker E, Forlani S, De Graaff W, Oosterveen T, Roelen B, Roelfsema J (1999) Initiation, establishment and maintenance of Hox gene expression patterns in the mouse. *The International journal of developmental biology* **43:** 635-650

Dimitri P, Corradini N, Rossi F, Verni F (2005) The paradox of functional heterochromatin. *Bioessays* **27:** 29-41

Dinger ME, Amaral PP, Mercer TR, Pang KC, Bruce SJ, Gardiner BB, Askarian-Amiri ME, Ru K, Solda G, Simons C, Sunkin SM, Crowe ML, Grimmond SM, Perkins AC, Mattick JS (2008) Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome research* **18:** 1433-1445

Dion MF, Altschuler SJ, Wu LF, Rando OJ (2005) Genomic characterization reveals a simple histone H4 acetylation code. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 5501-5506

Dou Y, Milne TA, Tackett AJ, Smith ER, Fukuda A, Wysocka J, Allis CD, Chait BT, Hess JL, Roeder RG (2005) Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell* **121:** 873-885

Dover J, Schneider J, Tawiah-Boateng MA, Wood A, Dean K, Johnston M, Shilatifard A (2002) Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. *The Journal of biological chemistry* **277:** 28368-28371

Drummond DC, Noble CO, Kirpotin DB, Guo Z, Scott GK, Benz CC (2005) Clinical development of histone deacetylase inhibitors as anticancer agents. *Annual review of pharmacology and toxicology* **45:** 495-528

Du HN, Fingerman IM, Briggs SD (2008) Histone H3 K36 methylation is mediated by a trans-histone methylation pathway involving an interaction between Set2 and histone H4. *Genes & development* **22:** 2786-2798

Duncan S (2007) Teratogenesis of sodium valproate. *Current opinion in neurology* **20:** 175-180

Durrin LK, Mann RK, Kayne PS, Grunstein M (1991) Yeast histone H4 N-terminal sequence is required for promoter activation in vivo. *Cell* **65:** 1023-1031

Efroni S, Duttagupta R, Cheng J, Dehghani H, Hoeppner DJ, Dash C, Bazett-Jones DP, Le Grice S, McKay RD, Buetow KH, Gingeras TR, Misteli T, Meshorer E (2008) Global transcription in pluripotent embryonic stem cells. *Cell stem cell* **2:** 437-447

Ezhkova E, Tansey WP (2004) Proteasomal ATPases link ubiquitylation of histone H2B to methylation of histone H3. *Molecular cell* **13:** 435-442

Faiella A, Wernig M, Consalez GG, Hostick U, Hofmann C, Hustert E, Boncinelli E, Balling R, Nadeau JH (2000) A mouse model for valproate teratogenicity: parental effects, homeotic transformations, and altered HOX expression. *Human molecular genetics* **9:** 227-236

Faust C, Lawson KA, Schork NJ, Thiel B, Magnuson T (1998) The Polycomb-group gene eed is required for normal morphogenetic movements during gastrulation in the mouse embryo. *Development (Cambridge, England)* **125:** 4495-4506

Fingerman IM, Wu CL, Wilson BD, Briggs SD (2005) Global loss of Set1-mediated H3 Lys4 trimethylation is associated with silencing defects in Saccharomyces cerevisiae. *The Journal of biological chemistry* **280:** 28761-28765

Fischle W, Wang Y, Allis CD (2003a) Histone and chromatin cross-talk. *Current opinion in cell biology* **15:** 172-183

Fischle W, Wang Y, Jacobs SA, Kim Y, Allis CD, Khorasanizadeh S (2003b) Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes & development* **17:** 1870-1881

Fisher D, Mechali M (2003) Vertebrate HoxB gene expression requires DNA replication. *The EMBO journal* **22:** 3737-3748

Fragoso G, Pennie WD, John S, Hager GL (1998) The position and length of the steroid-dependent hypersensitive region in the mouse mammary tumor virus long terminal repeat are invariant despite multiple nucleosome B frames. *Molecular and cellular biology* **18:** 3633-3644

Frischer LE, Hagen FS, Garber RL (1986) An inversion that disrupts the Antennapedia gene causes abnormal structure and localization of RNAs. *Cell* **47:** 1017-1023

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

Fujii-Yamamoto H, Kim JM, Arai K, Masai H (2005) Cell cycle and developmental regulations of replication factors in mouse embryonic stem cells. *J Biol Chem* **280:** 12976-12987

Fujimura Y, Isono K, Vidal M, Endoh M, Kajita H, Mizutani-Koseki Y, Takihara Y, van Lohuizen M, Otte A, Jenuwein T, Deschamps J, Koseki H (2006) Distinct roles of Polycomb group gene products in transcriptionally repressed and active domains of Hoxb8. *Development (Cambridge, England)* **133:** 2371-2381

Gregory GD, Vakoc CR, Rozovskaia T, Zheng X, Patel S, Nakamura T, Canaani E, Blobel GA (2007) Mammalian ASH1L is a histone methyltransferase that occupies the transcribed region of active genes. *Molecular and cellular biology* **27:** 8466-8479

Guenther MG, Jenner RG, Chevalier B, Nakamura T, Croce CM, Canaani E, Young RA (2005) Global and Hox-specific roles for the MLL1 methyltransferase. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 8603-8608

Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA (2007) A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130:** 77-88

Hansen KH, Bracken AP, Pasini D, Dietrich N, Gehani SS, Monrad A, Rappsilber J, Lerdrup M, Helin K (2008) A model for transmission of the H3K27me3 epigenetic mark. *Nature cell biology* **10:** 1291-1300

Hanson RD, Hess JL, Yu BD, Ernst P, van Lohuizen M, Berns A, van der Lugt NM, Shashikant CS, Ruddle FH, Seto M, Korsmeyer SJ (1999) Mammalian Trithorax and polycomb-group homologues are antagonistic regulators of homeotic development. *Proceedings of the National Academy of Sciences of the United States of America* **96:** 14372-14377

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

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

Hebbes TR, Clayton AL, Thorne AW, Crane-Robinson C (1994) Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain. *The EMBO journal* **13:** 1823-1830

Hirota T, Lipp JJ, Toh BH, Peters JM (2005) Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature* **438:** 1176-1180

Holliday R (1987) The inheritance of epigenetic defects. *Science (New York, NY* **238:** 163-170

Holliday R (1994) Epigenetics: an overview. *Developmental genetics* **15:** 453-457

Horn PJ, Peterson CL (2001) The bromodomain: a regulator of ATP-dependent chromatin remodeling? *Front Biosci* **6:** D1019-1023

Huangfu D, Maehr R, Guo W, Eijkelenboom A, Snitow M, Chen AE, Melton DA (2008) Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nature biotechnology* **26:** 795-797

Illingworth R, Kerr A, Desousa D, Jorgensen H, Ellis P, Stalker J, Jackson D, Clee C, Plumb R, Rogers J, Humphray S, Cox T, Langford C, Bird A (2008) A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS biology* **6:** e22

Isono K, Fujimura Y, Shinga J, Yamaki M, J OW, Takihara Y, Murahashi Y, Takada Y, Mizutani-Koseki Y, Koseki H (2005) Mammalian polyhomeotic homologues Phc2 and Phc1 act in synergy to mediate polycomb repression of Hox genes. *Molecular and cellular biology* **25:** 6694-6706

Ivanova N, Dobrin R, Lu R, Kotenko I, Levorse J, DeCoste C, Schafer X, Lun Y, Lemischka IR (2006) Dissecting self-renewal in stem cells with RNA interference. *Nature* **442:** 533-538

Jacob F, Monod J (1961) Genetic regulatory mechanisms in the synthesis of proteins. *Journal of molecular biology* **3:** 318-356

Jacobson RH, Ladurner AG, King DS, Tjian R (2000) Structure and function of a human TAFII250 double bromodomain module. *Science (New York, NY* **288:** 1422-1425

Jenuwein T, Allis CD (2001) Translating the histone code. *Science (New York, NY* **293:** 1074-1080

Johnson CA, Turner BM (1999) Histone deacetylases: complex transducers of nuclear signals. *Seminars in cell & developmental biology* **10:** 179-188

Kaji K, Norrby K, Paca A, Mileikovsky M, Mohseni P, Woltjen K (2009) Virus-free induction of pluripotency and subsequent excision of reprogramming factors. *Nature* **458:** 771-775

Karantzal E, Schulz H, Hummel O, Hubner N, Hatzopoulos A, Kretsovali A (2008) Histone deacetylase inhibition accelerates the early events of stem cell differentiation: transcriptomic and epigenetic analysis. *Genome biology* **9:** R65

Kessel M, Gruss P (1991) Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. *Cell* **67:** 89-104

Kim J, Daniel J, Espejo A, Lake A, Krishna M, Xia L, Zhang Y, Bedford MT (2006a) Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO reports* **7:** 397-403

Kim SC, Sprung R, Chen Y, Xu Y, Ball H, Pei J, Cheng T, Kho Y, Xiao H, Xiao L, Grishin NV, White M, Yang XJ, Zhao Y (2006b) Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Molecular cell* **23:** 607-618

Kiyokawa H, Busquets X, Powell CT, Ngo L, Rifkind RA, Marks PA (1992) Cloning of a D-type cyclin from murine erythroleukemia cells. *Proceedings of the National Academy of Sciences of the United States of America* **89:** 2444-2447

Kizer KO, Phatnani HP, Shibata Y, Hall H, Greenleaf AL, Strahl BD (2005) A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Molecular and cellular biology* **25:** 3305-3316

Kmita M, Duboule D (2003) Organizing axes in time and space; 25 years of colinear tinkering. *Science (New York, NY* **301:** 331-333

Kmita M, van Der Hoeven F, Zakany J, Krumlauf R, Duboule D (2000) Mechanisms of Hox gene colinearity: transposition of the anterior Hoxb1 gene into the posterior HoxD complex. *Genes & development* **14:** 198-211

Knezetic JA, Luse DS (1986) The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. *Cell* **45:** 95-104

Kolasinska-Zwierz P, Down T, Latorre I, Liu T, Liu XS, Ahringer J (2009) Differential chromatin marking of introns and expressed exons by H3K36me3. *Nature genetics* **41:** 376-381

Kondo T, Duboule D (1999) Breaking colinearity in the mouse HoxD complex. *Cell* **97:** 407-417

Kondo T, Zakany J, Duboule D (1998) Control of colinearity in AbdB genes of the mouse HoxD complex. *Molecular cell* **1:** 289-300

Kouskouti A, Talianidis I (2005) Histone modifications defining active genes persist after transcriptional and mitotic inactivation. *The EMBO journal* **24:** 347-357

Krogan NJ, Kim M, Tong A, Golshani A, Cagney G, Canadien V, Richards DP, Beattie BK, Emili A, Boone C, Shilatifard A, Buratowski S, Greenblatt J (2003) Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. *Molecular and cellular biology* **23:** 4207-4218

Kruhlak MJ, Hendzel MJ, Fischle W, Bertos NR, Hameed S, Yang XJ, Verdin E, Bazett-Jones DP (2001) Regulation of global acetylation in mitosis through loss of histone acetyltransferases and deacetylases from chromatin. *The Journal of biological chemistry* **276:** 38307-38319

Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D (2002) Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes & development* **16:** 2893-2905

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227:** 680-685

Lan F, Bayliss PE, Rinn JL, Whetstine JR, Wang JK, Chen S, Iwase S, Alpatov R, Issaeva I, Canaani E, Roberts TM, Chang HY, Shi Y (2007) A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* **449:** 689-694

Lee ER, Murdoch FE, Fritsch MK (2007a) High histone acetylation and decreased polycomb repressive complex 2 member levels regulate gene specific transcriptional changes during early embryonic stem cell differentiation induced by retinoic acid. *Stem cells (Dayton, Ohio)* **25:** 2191-2199

Lee JH, Hart SR, Skalnik DG (2004) Histone deacetylase activity is required for embryonic stem cell differentiation. *Genesis* **38:** 32-38

Lee MG, Villa R, Trojer P, Norman J, Yan KP, Reinberg D, Di Croce L, Shiekhattar R (2007b) Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science (New York, NY* **318:** 447-450

Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, Johnstone SE, Cole MF, Isono K, Koseki H, Fuchikami T, Abe K, Murray HL, Zucker JP, Yuan B, Bell GW, Herbolsheimer E, Hannett NM, Sun K, Odom DT, Otte AP, Volkert TL, Bartel DP, Melton DA, Gifford DK, Jaenisch R, Young RA (2006) Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **125:** 301-313

Lewis EB (1978) A gene complex controlling segmentation in Drosophila. *Nature* **276:** 565-570

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

Li G, Levitus M, Bustamante C, Widom J (2005) Rapid spontaneous accessibility of nucleosomal DNA. *Nature structural & molecular biology* **12:** 46-53

Lin W, Srajer G, Evrard YA, Phan HM, Furuta Y, Dent SY (2007) Developmental potential of Gcn5(-/-) embryonic stem cells in vivo and in vitro. *Dev Dyn* **236:** 1547-1557

Lin W, Zhang Z, Chen CH, Behringer RR, Dent SY (2008) Proper Gcn5 histone acetyltransferase expression is required for normal anteroposterior patterning of the mouse skeleton. *Development, growth & differentiation* **50:** 321-330

Lomvardas S, Thanos D (2001) Nucleosome sliding via TBP DNA binding in vivo. *Cell* **106:** 685-696

Lorch Y, LaPointe JW, Kornberg RD (1987) Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. *Cell* **49:** 203-210

Loyola A, Almouzni G (2004) Bromodomains in living cells participate in deciphering the histone code. *Trends in cell biology* **14:** 279-281

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

Macdonald N, Welburn JP, Noble ME, Nguyen A, Yaffe MB, Clynes D, Moggs JG, Orphanides G, Thomson S, Edmunds JW, Clayton AL, Endicott JA, Mahadevan LC (2005) Molecular basis for the recognition of phosphorylated and phosphoacetylated histone h3 by 14-3-3. *Molecular cell* **20:** 199-211

Mahadevan LC, Willis AC, Barratt MJ (1991) Rapid histone H3 phosphorylation in response to growth factors, phorbol esters, okadaic acid, and protein synthesis inhibitors. *Cell* **65:** 775-783

Marshall H, Studer M, Popperl H, Aparicio S, Kuroiwa A, Brenner S, Krumlauf R (1994) A conserved retinoic acid response element required for early expression of the homeobox gene Hoxb-1. *Nature* **370:** 567-571

Martinez-Balbas MA, Dey A, Rabindran SK, Ozato K, Wu C (1995) Displacement of sequence-specific transcription factors from mitotic chromatin. *Cell* **83:** 29-38

Mateescu B, England P, Halgand F, Yaniv M, Muchardt C (2004) Tethering of HP1 proteins to chromatin is relieved by phosphoacetylation of histone H3. *EMBO reports* **5:** 490-496

McCool KW, Xu X, Singer DB, Murdoch FE, Fritsch MK (2007) The role of histone acetylation in regulating early gene expression patterns during early embryonic stem cell differentiation. *The Journal of biological chemistry* **282:** 6696-6706

McGinnis W, Krumlauf R (1992) Homeobox genes and axial patterning. *Cell* **68:** 283-302

Megee PC, Morgan BA, Mittman BA, Smith MM (1990) Genetic analysis of histone H4: essential role of lysines subject to reversible acetylation. *Science (New York, NY* **247:** 841-845

Metivier R, Penot G, Carmouche RP, Hubner MR, Reid G, Denger S, Manu D, Brand H, Kos M, Benes V, Gannon F (2004) Transcriptional complexes engaged by apo-estrogen receptor-alpha isoforms have divergent outcomes. *The EMBO journal* **23:** 3653-3666

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, O'Donovan A, Presser A, Russ C, Xie X, Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES, Bernstein BE (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448:** 553-560

Milne TA, Briggs SD, Brock HW, Martin ME, Gibbs D, Allis CD, Hess JL (2002) MLL targets SET domain methyltransferase activity to Hox gene promoters. *Molecular cell* **10:** 1107-1117

Min J, Zhang Y, Xu RM (2003) Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genes & development* **17:** 1823-1828

Mishra BP, Ansari KI, Mandal SS (2009) Dynamic association of MLL1, H3K4 trimethylation with chromatin and Hox gene expression during the cell cycle. *The FEBS journal*

Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S (2003) The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113:** 631-642

Mohn F, Schubeler D (2009) Genetics and epigenetics: stability and plasticity during cellular differentiation. *Trends Genet*

Mohn F, Weber M, Rebhan M, Roloff TC, Richter J, Stadler MB, Bibel M, Schubeler D (2008) Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Molecular cell* **30:** 755-766

Montgomery ND, Yee D, Chen A, Kalantry S, Chamberlain SJ, Otte AP, Magnuson T (2005) The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. *Curr Biol* **15:** 942-947

Morey C, Da Silva NR, Kmita M, Duboule D, Bickmore WA (2008) Ectopic nuclear reorganisation driven by a Hoxb1 transgene transposed into Hoxd. *Journal of cell science* **121:** 571-577

Morey C, Da Silva NR, Perry P, Bickmore WA (2007) Nuclear reorganisation and chromatin decondensation are conserved, but distinct, mechanisms linked to Hox gene activation. *Development (Cambridge, England)* **134:** 909-919

Muller WG, Walker D, Hager GL, McNally JG (2001) Large-scale chromatin decondensation and recondensation regulated by transcription from a natural promoter. *The Journal of cell biology* **154:** 33-48

Murry CE, Keller G (2008) Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell* **132:** 661-680

Nagy Z, Tora L (2007) Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. *Oncogene* **26:** 5341-5357

Nakamura T, Mori T, Tada S, Krajewski W, Rozovskaia T, Wassell R, Dubois G, Mazo A, Croce CM, Canaani E (2002) ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. *Molecular cell* **10:** 1119-1128

Negre N, Hennetin J, Sun LV, Lavrov S, Bellis M, White KP, Cavalli G (2006) Chromosomal distribution of PcG proteins during Drosophila development. *PLoS biology* **4:** e170

Ng HH, Robert F, Young RA, Struhl K (2003) Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Molecular cell* **11:** 709-719

Ng HH, Xu RM, Zhang Y, Struhl K (2002) Ubiquitination of histone H2B by Rad6 is required for efficient Dot1-mediated methylation of histone H3 lysine 79. *The Journal of biological chemistry* **277:** 34655-34657

Ng RK, Gurdon JB (2008) Epigenetic memory of an active gene state depends on histone H3.3 incorporation into chromatin in the absence of transcription. *Nature cell biology* **10:** 102-109

Nightingale KP, Gendreizig S, White DA, Bradbury C, Hollfelder F, Turner BM (2007) Cross-talk between histone modifications in response to histone deacetylase inhibitors: MLL4 links histone H3 acetylation and histone H3K4 methylation. *The Journal of biological chemistry* **282:** 4408-4416

Nishioka K, Chuikov S, Sarma K, Erdjument-Bromage H, Allis CD, Tempst P, Reinberg D (2002) Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation. *Genes & development* **16:** 479-489

Niwa H (2007) How is pluripotency determined and maintained? *Development (Cambridge, England)* **134:** 635-646

Niwa H, Miyazaki J, Smith AG (2000) Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature genetics* **24:** 372-376

Noma K, Allis CD, Grewal SI (2001) Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science (New York, NY* **293:** 1150-1155

Nowak SJ, Corces VG (2004) Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends Genet* **20:** 214-220

O'Neill LP, Turner BM (1996) Immunoprecipitation of chromatin. *Methods in enzymology* **274:** 189-197

O'Neill LP, VerMilyea MD, Turner BM (2006) Epigenetic characterization of the early embryo with a chromatin immunoprecipitation protocol applicable to small cell populations. *Nature genetics* **38:** 835-841

Okita K, Nakagawa M, Hyenjong H, Ichisaka T, Yamanaka S (2008) Generation of mouse induced pluripotent stem cells without viral vectors. *Science (New York, NY* **322:** 949-953

Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD, Cheng X, Bestor TH (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448:** 714-717

Orlando V, Paro R (1993) Mapping Polycomb-repressed domains in the bithorax complex using in vivo formaldehyde cross-linked chromatin. *Cell* **75:** 1187-1198

Pan G, Tian S, Nie J, Yang C, Ruotti V, Wei H, Jonsdottir GA, Stewart R, Thomson JA (2007) Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. *Cell stem cell* **1:** 299-312

Papalopulu N, Lovell-Badge R, Krumlauf R (1991) The expression of murine Hox-2 genes is dependent on the differentiation pathway and displays a collinear sensitivity to retinoic acid in F9 cells and Xenopus embryos. *Nucleic acids research* **19:** 5497-5506

Paro R (1990) Imprinting a determined state into the chromatin of Drosophila. *Trends Genet* **6:** 416-421

Paro R, Hogness DS (1991) The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of Drosophila. *Proceedings of the National Academy of Sciences of the United States of America* **88:** 263-267

Pasini D, Bracken AP, Hansen JB, Capillo M, Helin K (2007) The polycomb group protein Suz12 is required for embryonic stem cell differentiation. *Molecular and cellular biology* **27:** 3769-3779

Pasini D, Bracken AP, Jensen MR, Lazzerini Denchi E, Helin K (2004) Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *The EMBO journal* **23:** 4061-4071

Pasini D, Hansen KH, Christensen J, Agger K, Cloos PA, Helin K (2008) Coordinated regulation of transcriptional repression by the RBP2 H3K4 demethylase and Polycomb-Repressive Complex 2. *Genes & development* **22:** 1345-1355

Platero JS, Hartnett T, Eissenberg JC (1995) Functional analysis of the chromo domain of HP1. *The EMBO journal* **14:** 3977-3986

Pogo BG, Pogo AO, Allfrey VG, Mirsky AE (1968) Changing patterns of histone acetylation and RNA synthesis in regeneration of the liver. *Proceedings of the National Academy of Sciences of the United States of America* **59:** 1337-1344

Pokholok DK, Harbison CT, Levine S, Cole M, Hannett NM, Lee TI, Bell GW, Walker K, Rolfe PA, Herbolsheimer E, Zeitlinger J, Lewitter F, Gifford DK, Young RA (2005) Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122:** 517-527

Popperl H, Featherstone MS (1992) An autoregulatory element of the murine Hox-4.2 gene. *The EMBO journal* **11:** 3673-3680

Popperl H, Featherstone MS (1993) Identification of a retinoic acid response element upstream of the murine Hox-4.2 gene. *Molecular and cellular biology* **13:** 257-265

Pray-Grant MG, Daniel JA, Schieltz D, Yates JR, 3rd, Grant PA (2005) Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature* **433:** 434-438

Probst AV, Dunleavy E, Almouzni G (2009) Epigenetic inheritance during the cell cycle. *Nature reviews* **10:** 192-206

Ptashne M (2007) On the use of the word 'epigenetic'. *Curr Biol* **17:** R233-236

Ragvin A, Valvatne H, Erdal S, Arskog V, Tufteland KR, Breen K, AM OY, Eberharter A, Gibson TJ, Becker PB, Aasland R (2004) Nucleosome binding by the bromodomain and PHD finger of the transcriptional cofactor p300. *Journal of molecular biology* **337:** 773-788

Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S (2006) Stochastic mRNA synthesis in mammalian cells. *PLoS biology* **4:** e309

Rastegar M, Kobrossy L, Kovacs EN, Rambaldi I, Featherstone M (2004) Sequential histone modifications at Hoxd4 regulatory regions distinguish anterior from posterior embryonic compartments. *Molecular and cellular biology* **24:** 8090-8103

Reinberg D, Sims RJ, 3rd (2006) de FACTo nucleosome dynamics. *The Journal of biological chemistry* **281:** 23297-23301

Richard-Foy H, Hager GL (1987) Sequence-specific positioning of nucleosomes over the steroid-inducible MMTV promoter. *The EMBO journal* **6:** 2321-2328

Ringrose L, Paro R (2004) Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annual review of genetics* **38:** 413-443

Ringrose L, Paro R (2007) Polycomb/Trithorax response elements and epigenetic memory of cell identity. *Development (Cambridge, England)* **134:** 223-232

Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, Chang HY (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129:** 1311-1323

Roelen BA, de Graaff W, Forlani S, Deschamps J (2002) Hox cluster polarity in early transcriptional availability: a high order regulatory level of clustered Hox genes in the mouse. *Mechanisms of development* **119:** 81-90

Roh TY, Cuddapah S, Cui K, Zhao K (2006) The genomic landscape of histone modifications in human T cells. *Proceedings of the National Academy of Sciences of the United States of America* **103:** 15782-15787

Rohwedel J, Guan K, Wobus AM (1999) Induction of cellular differentiation by retinoic acid in vitro. *Cells, tissues, organs* **165:** 190-202

Sanchez-Elsner T, Gou D, Kremmer E, Sauer F (2006) Noncoding RNAs of trithorax response elements recruit Drosophila Ash1 to Ultrabithorax. *Science (New York, NY* **311:** 1118-1123

Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL, Mellor J, Kouzarides T (2002) Active genes are tri-methylated at K4 of histone H3. *Nature* **419:** 407-411

Santos-Rosa H, Schneider R, Bernstein BE, Karabetsou N, Morillon A, Weise C, Schreiber SL, Mellor J, Kouzarides T (2003) Methylation of histone H3 K4 mediates association of the Isw1p ATPase with chromatin. *Molecular cell* **12:** 1325-1332

Satchwell SC, Travers AA (1989) Asymmetry and polarity of nucleosomes in chicken erythrocyte chromatin. *Embo J* **8:** 229-238

Saurin AJ, Shao Z, Erdjument-Bromage H, Tempst P, Kingston RE (2001) A Drosophila Polycomb group complex includes Zeste and dTAFII proteins. *Nature* **412:** 655-660

Schmitt S, Paro R (2006) RNA at the steering wheel. *Genome biology* **7:** 218

Schmitt S, Prestel M, Paro R (2005) Intergenic transcription through a polycomb group response element counteracts silencing. *Genes & development* **19:** 697-708

Schubeler D, MacAlpine DM, Scalzo D, Wirbelauer C, Kooperberg C, van Leeuwen F, Gottschling DE, O'Neill LP, Turner BM, Delrow J, Bell SP, Groudine M (2004) The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes & development* **18:** 1263-1271

Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G (2007) Genome regulation by polycomb and trithorax proteins. *Cell* **128:** 735-745

Schwartz S, Meshorer E, Ast G (2009) Chromatin organization marks exon-intron structure. *Nature structural & molecular biology* **16:** 990-995

Schwartz YB, Kahn TG, Nix DA, Li XY, Bourgon R, Biggin M, Pirrotta V (2006) Genome-wide analysis of Polycomb targets in Drosophila melanogaster. *Nature genetics* **38:** 700-705

Sessa L, Breiling A, Lavorgna G, Silvestri L, Casari G, Orlando V (2007) Noncoding RNA synthesis and loss of Polycomb group repression accompanies the colinear activation of the human HOXA cluster. *RNA (New York, NY* **13:** 223-239

Sharova LV, Sharov AA, Piao Y, Shaik N, Sullivan T, Stewart CL, Hogan BL, Ko MS (2007) Global gene expression profiling reveals similarities and differences among mouse pluripotent stem cells of different origins and strains. *Developmental biology* **307:** 446-459

Shi X, Hong T, Walter KL, Ewalt M, Michishita E, Hung T, Carney D, Pena P, Lan F, Kaadige MR, Lacoste N, Cayrou C, Davrazou F, Saha A, Cairns BR, Ayer DE, Kutateladze TG, Shi Y, Cote J, Chua KF, Gozani O (2006) ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* **442:** 96-99

Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA, Shi Y (2004) Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119:** 941-953

Shi Y, Whetstine JR (2007) Dynamic regulation of histone lysine methylation by demethylases. *Molecular cell* **25:** 1-14

Shilatifard A (2006) Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annual review of biochemistry* **75:** 243-269

Silva J, Smith A (2008) Capturing pluripotency. *Cell* **132:** 532-536

Simeone A, Acampora D, Arcioni L, Andrews PW, Boncinelli E, Mavilio F (1990) Sequential activation of HOX2 homeobox genes by retinoic acid in human embryonal carcinoma cells. *Nature* **346:** 763-766

Simic R, Lindstrom DL, Tran HG, Roinick KL, Costa PJ, Johnson AD, Hartzog GA, Arndt KM (2003) Chromatin remodeling protein Chd1 interacts with transcription elongation factors and localizes to transcribed genes. *The EMBO journal* **22:** 1846-1856

Sims RJ, 3rd, Chen CF, Santos-Rosa H, Kouzarides T, Patel SS, Reinberg D (2005) Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. *The Journal of biological chemistry* **280:** 41789-41792

Smale ST (2003) The establishment and maintenance of lymphocyte identity through gene silencing. *Nature immunology* **4:** 607-615

Smith AG (2001) Embryo-derived stem cells: of mice and men. *Annual review of cell and developmental biology* **17:** 435-462

Smits VA, Medema RH (2001) Checking out the G(2)/M transition. *Biochimica et biophysica acta* **1519:** 1-12

Snowden AW, Gregory PD, Case CC, Pabo CO (2002) Gene-specific targeting of H3K9 methylation is sufficient for initiating repression in vivo. *Curr Biol* **12:** 2159-2166

Spitz F, Gonzalez F, Duboule D (2003) A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. *Cell* **113:** 405-417

Spivakov M, Fisher AG (2007) Epigenetic signatures of stem-cell identity. *Nature reviews* **8:** 263-271

Squazzo SL, O'Geen H, Komashko VM, Krig SR, Jin VX, Jang SW, Margueron R, Reinberg D, Green R, Farnham PJ (2006) Suz12 binds to silenced regions of the genome in a cell-type-specific manner. *Genome research* **16:** 890-900

Stadtfeld M, Nagaya M, Utikal J, Weir G, Hochedlinger K (2008) Induced pluripotent stem cells generated without viral integration. *Science (New York, NY* **322:** 945-949

Sterner DE, Berger SL (2000) Acetylation of histones and transcription-related factors. *Microbiol Mol Biol Rev* **64:** 435-459

Stewart MD, Li J, Wong J (2005) Relationship between histone H3 lysine 9 methylation, transcription repression, and heterochromatin protein 1 recruitment. *Molecular and cellular biology* **25:** 2525-2538

Stock JK, Giadrossi S, Casanova M, Brookes E, Vidal M, Koseki H, Brockdorff N, Fisher AG, Pombo A (2007) Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. *Nature cell biology* **9:** 1428-1435

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

Struhl K (1999) Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. *Cell* **98:** 1-4

Sun ZW, Allis CD (2002) Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* **418:** 104-108

Szutorisz H, Canzonetta C, Georgiou A, Chow CM, Tora L, Dillon N (2005) Formation of an active tissue-specific chromatin domain initiated by epigenetic marking at the embryonic stem cell stage. *Molecular and cellular biology* **25:** 1804-1820

Taft RJ, Pheasant M, Mattick JS (2007) The relationship between non-protein-coding DNA and eukaryotic complexity. *Bioessays* **29:** 288-299

Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126:** 663-676

Tamkun JW, Deuring R, Scott MP, Kissinger M, Pattatucci AM, Kaufman TC, Kennison JA (1992) brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. *Cell* **68:** 561-572

Taverna SD, Ilin S, Rogers RS, Tanny JC, Lavender H, Li H, Baker L, Boyle J, Blair LP, Chait BT, Patel DJ, Aitchison JD, Tackett AJ, Allis CD (2006) Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. *Molecular cell* **24:** 785-796

Thomas MC, Chiang CM (2006) The general transcription machinery and general cofactors. *Critical reviews in biochemistry and molecular biology* **41:** 105-178

Thomson S, Hollis A, Hazzalin CA, Mahadevan LC (2004) Distinct stimulus-specific histone modifications at hsp70 chromatin targeted by the transcription factor heat shock factor-1. *Molecular cell* **15:** 585-594

Thomson S, Mahadevan LC, Clayton AL (1999) MAP kinase-mediated signalling to nucleosomes and immediate-early gene induction. *Seminars in cell & developmental biology* **10:** 205-214

Thorne JL, Campbell MJ, Turner BM (2009) Transcription factors, chromatin and cancer. *The international journal of biochemistry & cell biology* **41:** 164-175

Towbin H, Staehelin T, Gordon J (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences of the United States of America* **76:** 4350-4354

Tran HG, Steger DJ, Iyer VR, Johnson AD (2000) The chromo domain protein chd1p from budding yeast is an ATP-dependent chromatin-modifying factor. *The EMBO journal* **19:** 2323-2331

Tse C, Sera T, Wolffe AP, Hansen JC (1998) Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. *Molecular and cellular biology* **18:** 4629-4638

Turner BM (1993) Decoding the nucleosome. *Cell* **75:** 5-8

Turner BM (2000) Histone acetylation and an epigenetic code. *Bioessays* **22:** 836-845

Turner BM (2002) Cellular memory and the histone code. *Cell* **111:** 285-291

Turner BM (2007) Defining an epigenetic code. *Nature cell biology* **9:** 2-6

Turner BM, Birley AJ, Lavender J (1992) Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in Drosophila polytene nuclei. *Cell* **69:** 375-384

Turner BM, Fellows G (1989) Specific antibodies reveal ordered and cell-cycle-related use of histone-H4 acetylation sites in mammalian cells. *Eur J Biochem* **179:** 131-139

Turner BM, O'Neill LP, Allan IM (1989) Histone H4 acetylation in human cells. Frequency of acetylation at different sites defined by immunolabeling with site-specific antibodies. *FEBS Lett* **253:** 141-145

Vakoc CR, Mandat SA, Olenchock BA, Blobel GA (2005) Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Molecular cell* **19:** 381-391

Valls E, Sanchez-Molina S, Martinez-Balbas MA (2005) Role of histone modifications in marking and activating genes through mitosis. *J Biol Chem* **280:** 42592-42600

van der Hoeven F, Zakany J, Duboule D (1996) Gene transpositions in the HoxD complex reveal a hierarchy of regulatory controls. *Cell* **85:** 1025-1035

van der Lugt NM, Alkema M, Berns A, Deschamps J (1996) The Polycomb-group homolog Bmi-1 is a regulator of murine Hox gene expression. *Mechanisms of development* **58:** 153-164

van der Vlag J, Otte AP (1999) Transcriptional repression mediated by the human polycomb-group protein EED involves histone deacetylation. *Nature genetics* **23:** 474-478

Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, Morey L, Van Eynde A, Bernard D, Vanderwinden JM, Bollen M, Esteller M, Di Croce L, de Launoit Y, Fuks F (2006) The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439:** 871-874

Vogelauer M, Rubbi L, Lucas I, Brewer BJ, Grunstein M (2002) Histone acetylation regulates the time of replication origin firing. *Molecular cell* **10:** 1223-1233

Voncken JW, Roelen BA, Roefs M, de Vries S, Verhoeven E, Marino S, Deschamps J, van Lohuizen M (2003) Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. *Proceedings of the National Academy of Sciences of the United States of America* **100:** 2468-2473

Waddington CH (1939) *Introduction to modern genetics*, London: Allen and Unwin.

Wang GG, Cai L, Pasillas MP, Kamps MP (2007) NUP98-NSD1 links H3K36 methylation to Hox-A gene activation and leukaemogenesis. *Nature cell biology* **9:** 804-812

Wang H, Cao R, Xia L, Erdjument-Bromage H, Borchers C, Tempst P, Zhang Y (2001) Purification and functional characterization of a histone H3-lysine 4-specific methyltransferase. *Molecular cell* **8:** 1207-1217

Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS, Zhang Y (2004) Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431:** 873-878

Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler D (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature genetics* **39:** 457-466

Welstead GG, Schorderet P, Boyer LA (2008) The reprogramming language of pluripotency. *Current opinion in genetics & development* **18:** 123-129

White DA, Belyaev ND, Turner BM (1999) Preparation of site-specific antibodies to acetylated histones. *Methods* **19:** 417-424

Williams RS, Shohet RV, Stillman B (1997) A human protein related to yeast Cdc6p. *Proceedings of the National Academy of Sciences of the United States of America* **94:** 142-147

Woltjen K, Michael IP, Mohseni P, Desai R, Mileikovsky M, Hamalainen R, Cowling R, Wang W, Liu P, Gertsenstein M, Kaji K, Sung HK, Nagy A (2009) piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* **458:** 766-770

Woodage T, Basrai MA, Baxevanis AD, Hieter P, Collins FS (1997) Characterization of the CHD family of proteins. *Proceedings of the National Academy of Sciences of the United States of America* **94:** 11472-11477

Wysocka J, Swigut T, Milne TA, Dou Y, Zhang X, Burlingame AL, Roeder RG, Brivanlou AH, Allis CD (2005) WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* **121:** 859-872

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

Xi S, Zhu H, Xu H, Schmidtmann A, Geiman TM, Muegge K (2007) Lsh controls Hox gene silencing during development. *Proceedings of the National Academy of Sciences of the United States of America* **104:** 14366-14371

Xiao T, Hall H, Kizer KO, Shibata Y, Hall MC, Borchers CH, Strahl BD (2003) Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. *Genes & development* **17:** 654-663

Yu BD, Hanson RD, Hess JL, Horning SE, Korsmeyer SJ (1998) MLL, a mammalian trithorax-group gene, functions as a transcriptional maintenance factor in morphogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **95:** 10632-10636

Yu BD, Hess JL, Horning SE, Brown GA, Korsmeyer SJ (1995) Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* **378:** 505-508

Zhang F, Popperl H, Morrison A, Kovacs EN, Prideaux V, Schwarz L, Krumlauf R, Rossant J, Featherstone MS (1997) Elements both 5' and 3' to the murine Hoxd4 gene establish anterior borders of expression in mesoderm and neurectoderm. *Mechanisms of development* **67:** 49-58

Zhang Y, Reinberg D (2001) Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes & development* **15:** 2343-2360

Zhao XD, Han X, Chew JL, Liu J, Chiu KP, Choo A, Orlov YL, Sung WK, Shahab A, Kuznetsov VA, Bourque G, Oh S, Ruan Y, Ng HH, Wei CL (2007) Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. *Cell stem cell* **1:** 286-298

Zhu B, Zheng Y, Pham AD, Mandal SS, Erdjument-Bromage H, Tempst P, Reinberg D (2005) Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. *Molecular cell* **20:** 601-611