2006

Studies on the Histone Methyltransferase G9a

Srihari Cidambi Sampath

Studies on the histone methyltransferase G9a

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Srihari Cidambi Sampath

June 2006

ABSTRACT

The size and complexity of eukaryotic genomes require that specific mechanisms exist for ensuring both the stability as well as the accessibility of DNA. One such mechanism is the association of DNA with histones to form chromatin, the physiological substrate of gene expression. An important means by which histones impact transcriptional activity is through site-specific enzymatic modification of the amino terminal histone "tails", which can alter the spectrum of chromatin-associated proteins and hence transcriptional states. Among the known modifications of histones, lysine methylation has been proposed to represent a relatively stable mark which might mediate stable activation or repression, depending upon the site modified.

The immune system provides an ideal system in which to test the physiological functions of particular chromatin-modifying activities, since proper lymphocyte development and function requires integration of multiple signals, both cell autonomous and receptor-mediated, with complex DNA recombination reactions which are unique to lymphocytes. We have exploited these features to explore the possible functions of histone 3, lysine 3 (H3K9) methylation in the immune system through conditional inactivation of the H3K9-specific methyltransferases G9a and GLP.

These studies demonstrate that G9a is essential for B cell development in the mouse, but is dispensable for T cell development. The defect in B lymphopoiesis in the absence of G9a is caused by a block in development at the pro-B cell stage, corresponding to the onset of immunoglobulin heavy chain recombination. The overall normal behavior of G9a-deficient peripheral B cells argues in favor of the specificity of this effect.

Furthermore, through analysis of G9a and GLP protein sequences, we have identified a novel mechanism by which chromatin-modifying complexes can be regulated. We find that both G9a and GLP contain conserved H3K9-like motifs, on which the main biochemical features of H3K9 itself are recapitulated. Considering both the sequence and functional conservation between these sites and histones, we term these motifs in non-histone proteins "histone mimics". Our initial analysis indicates that many chromatin-associated proteins potentially contain H3K9-type histone mimics, and that this phenomenon is therefore likely to be a general one.

If you seek Truth, you will not seek to gain a victory by every possible means;
and when you have found Truth, you need not fear being defeated.

Epictetus

ACKNOWLEDGEMENTS

This work was made possible by the continuous support and guidance of many individuals. I owe the deepest debt of gratitude to my supervisor, Sasha Tarakhovsky, who has been—in every sense—a true mentor to me. By his generosity, understanding, creativity and rigor, he has set a standard which I will find it hard to live up to. Beyond all this, he served as a strong support when times were rough, and saved me from myself on more than one occasion. All this was done without ever dampening my sometimes overly-enthusiastic attempts to go in my own direction, a balance which I will always remember gratefully. In a similar vein, I would like to acknowledge Drs. Sidney Orlov and Peter Bruns, without whose help and encouragement I would never have even begun graduate school, much less finished it.

I thank all of my labmates in the Tarakhovsky lab, past and present, for their help and encouragement. Special thanks go to Christian Schmedt, who served as an outstanding role model and single-handedly taught me much of what I learned in graduate school, and Dónal O'Carroll, without whom the targeting projects described here would have failed. My baymates, Sukhvinder Sahota and Laura Donlin, are also deeply appreciated for patiently listening to my music and unsolicited opinions.

I was fortunate to have outstanding collaborators in the course of this work. These include Andrew Krutchinsky and Brian Chait for mass spectrometry analysis, Agnes Viale and Eric Miska for microarray experiments, and Miheala Zavolan and Terry Gaasterland for bioinformatic analyses. Critical contributions of reagents were made by numerous individuals and labs. I am particularly indebted to Thomas Jenuwein and colleagues for generously providing the α4xdiMeK9 antibody as well as numerous plasmids, and to C. David Allis for providing H3 peptides. Similarly, I would like to acknowledge the members of my Thesis Committee, Drs. C. David Allis (Chairman), Charles Rice and Carl Nathan for their indispensable contributions to this work, in particular their many insightful questions and comments.

Finally, my heartfelt thanks and appreciation to all of my family and friends for keeping me sane and happy these last few years—and especially to Amma and Srinath, who know how I feel much better than I can describe.

TABLE OF CONTENTS

FIGURES AND TABLES

CHAPTER 1: INTRODUCTION

1.1 Histone modifications and gene expression

A fundamental problem facing all organisms is the question of how to package, access and transmit genetic information. These three aims would appear to be mutually exclusive, since packaging and transmission of the genome clearly require compaction of DNA, whereas accessing DNA to allow transcription requires decondensation. All three of these problems are addressed in eukaryotes through the association of DNA with histones. Histones are small, highly basic proteins which interact directly and sequence non-specifically with DNA, forming the nucleoprotein complex known as chromatin. Packaging of DNA into chromatin allows compaction of up to 10,000-fold, allowing stable storage as well as propagation of the genome through mitosis.

The fundamental repeating unit of chromatin is the nucleosome core particle, consisting of 147 bp of DNA wrapped in 1.75 turns around a core histone octamer. This octamer consists of a histone 3/histone 4 (H3/H4) tetramer (itself made of two H3/H4 dimers) and two H2A/H2B dimers (Khorasanizadeh, 2004; Luger et al., 1997). The compaction of DNA caused by its packaging into chromatin can be increased by the further association with the linker histone, H1. Incorporation of H1 stabilizes a more condensed packing of nucleosomes *in vitro*, and promotes compaction into a 30 nm diameter fiber (Hayes and Hansen, 2001).

Histones were originally thought to be relatively inert "scaffolding" molecules, which themselves played little part in the regulation of gene expression. This view has given ground, though, with recent rapid advances made in understanding the many active processes which manipulate histones. These include ATP-dependent chromatin

remodeling (Lusser and Kadonaga, 2003), histone replacement with variants having distinct functional properties (Ahmad and Henikoff, 2002), and direct post-translational modification of core histones (Strahl and Allis, 2000). The study of histone modification in particular has seen explosive growth as it has become clear that a vast array of modifications can and do occur on all of the core histones. These modifications include serine and threonine phosphorylation (Nowak and Corces, 2004), lysine acetylation (Kouzarides, 2000; Mizzen and Allis, 1998; Roth et al., 2001), lysine and arginine methylation (Kouzarides, 2002; Lachner and Jenuwein, 2002), lysine ubiquitination (Briggs et al., 2002; Sun and Allis, 2002), as well as other less studied modifications (e.g. sumoylation, ADP-ribosylation, glycosylation and biotinylation). A list of the most well-studied modifications and the enzymes carry them out is given in Figure 1.1.

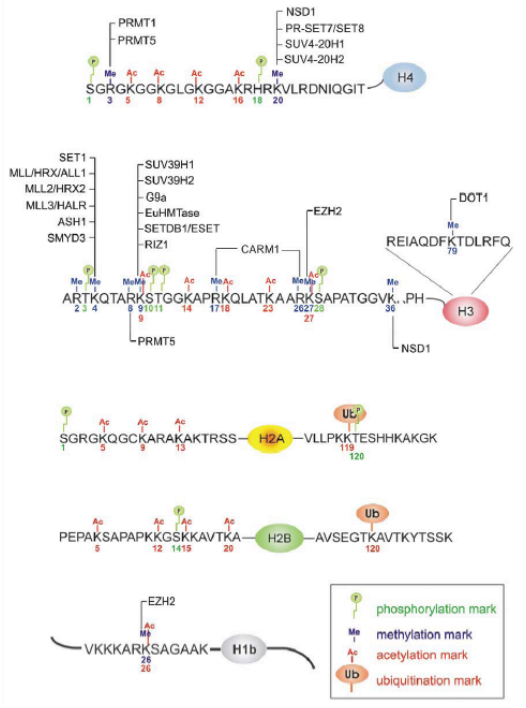

Figure 1.1 Partial description of histone modifications and modifying enzymes. Reproduced from Margueron et al., 2005.

Histone modifications can be broadly divided into those associated with active transcription (euchromatin-associated), and those associated with transcriptional repression (heterochromatin-associated). A subset of these are shown schematically in Figure 1.2. One prominent mark seen on active loci is histone phosphorylation, especially on H3 serine 10 (H3S10), which is associated both with chromosome condensation during mitosis as well as transcriptional activation (Nowak and Corces, 2000; Wei et al., 1999). Histone acetylation is also strongly correlated with transcriptional activity, and is perhaps the most well-studied mark of transcriptional activation (Allfrey et al., 1964; Kouzarides, 2000; Kuo et al., 1996). While it is still not completely clear how or indeed whether these modifications directly promote transcription, the existing evidence strongly indicates a causal relationship (Anest et al., 2003; Brownell et al., 1996; Chan and La Thangue, 2001; Yamamoto et al., 2003).

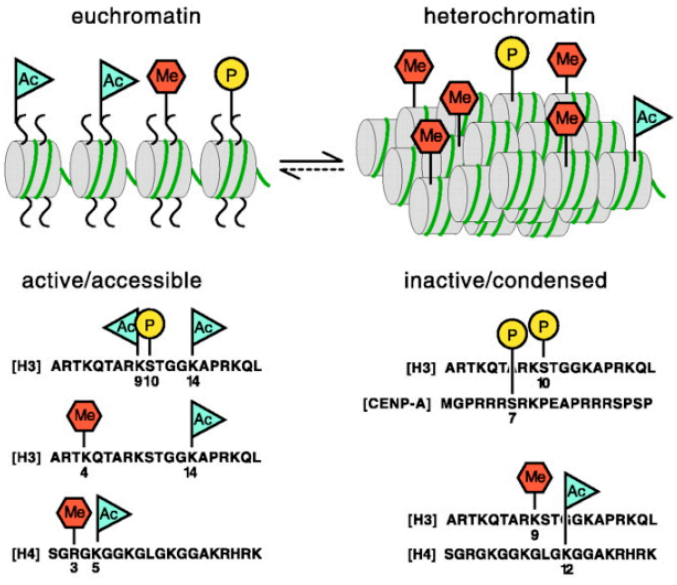

Figure 1.2 Histone modifications associated with euchromatin and heterochromatin. Abbreviations: Ac, acetylation; Me, methylation; P, phosphorylation. Adapted from Jenuwein and Allis (2001)

In contrast to phosphorylation and acetylation, the effects of histone lysine methylation on transcriptional activity seem to depend critically on the site of modification. Thus, while methylation of lysine 4 in the amino-terminal "tail" of histone 3 (H3K4) correlates with activation of transcription (Santos-Rosa et al., 2002), methylation of lysines 9 or 27 of the H3 tail (H3K9 and H3K27) is strongly correlated with transcriptional repression (Cao et al., 2002; Rea et al., 2000; Snowden et al., 2002). Lysine methylation is further complicated by the possibility of mutually exclusive modification on the same residue (e.g. acetylation versus methylation of H3K9), as well as by antagonism of modifications on different residues. One example of this antagonism is the described inhibition of H3K9 methylation by preexisting H3S10 phosphorylation (Rea et al., 2000). Furthermore, differences in modification density (e.g. mono-, di- and tri-methylation) can significantly effect the distribution of these marks: while H3K4 trimethylation in yeast is found mainly at the 5' end of actively transcribed genes, dimethylation in generally located in the middle of the open reading frame, and monomethylation at the 3' end (Pokholok et al., 2005).

1.2 Writing and reading histone modifications

To date, all except one of the known histone methyltransferases (HMTases) contains the evolutionarily conserved SET domain (Dillon et al., 2005). The SET domain was named for the three founding members of the family, **S**uVar(3-9), **E**nhancer of Zeste [E(z)], and **T**rithorax. All three of these proteins were known from their respective mutant phenotypes in *Drosophila* to be involved in epigenetic regulation (Alvarez-Venegas and Avramova, 2002; Schotta et al., 2002; Schumacher and Magnuson, 1997).

It has since become clear that these proteins all have intrinsic SET domain-dependent HMTase activity: SuVar(3-9), E(z) and Trithorax (as well as their mammalian homologs, Suv39H1/2, Ezh1/2 and MLL), catalyze H3K9, H3K27 and H3K4 methylation respectively.

For the most part, multiple HMTases seem to contribute to methylation of any given site in histones (Figure 1.1). However, knockout studies in mice have demonstrated that a subset of these HMTases are necessary for the vast majority of the methylation mark which they place. For instance, disruption of the murine E(z) homolog Ezh2 or its obligate binding partner Eed leads to a dramatic decrease in total H3K27 trimethylation, demonstrating that little redundancy exists for methylation of this site (Montgomery et al., 2005; Su et al., 2003; Su et al., 2005). Similarly, double mutation of Suv39H1/2 leads to complete loss of heterochromatic H3K9 trimethylation (Peters et al., 2001; Rea et al., 2000). Remarkably, although triMeH3K9 is lost, H3K9 dimethylation (diMeH3K9), a transcriptionally repressive mark normally present in euchromatic chromosomal territories, is unaffected (Peters et al., 2003; Rice et al., 2003). This is logical, as placement of the diMeH3K9 mark is catalyzed not by Suv39H1/2, but by the HMTases G9a and G9a-like protein (GLP, also known as EuHMTase-1). Loss of either G9a or GLP severely decreases euchromatic diMeH3K9, while leaving triMeH3K9 levels unchanged (Peters et al., 2003; Rice et al., 2003; Tachibana et al., 2002; Tachibana et al., 2005).

A major breakthrough in our understanding of the physiological functions of these varying histone modifications came with the discovery that particular modifications produce binding sites for “effector modules”, which recognize modified residues with

high specificity. This is true, for instance, of acetyllysine, which is recognized and bound by proteins containing the bromodomain (Dhalluin et al., 1999; Jacobson et al., 2000; Owen et al., 2000), which is found in nearly all histone acetyltransferases (HATs). Similarly, methyllysine can be specifically recognized by the chromodomain (Bannister et al., 2001; Jacobs et al., 2001; Lachner et al., 2001), and has more recently been shown to interact with Tudor domains as well (Huyen et al., 2004).

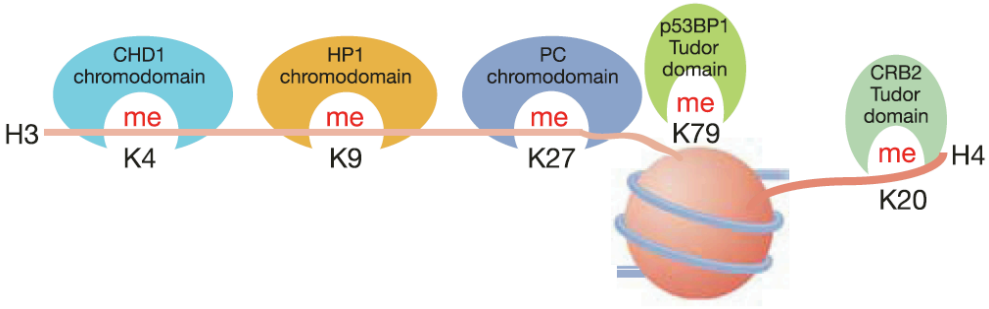

Figure 1.3 Summary of methylated histone binding proteins. Histones shown in red, DNA in blue. Abbreviations: CHD1, chromodomain helicase DNA-binding protein 1; HP1, heterochromatin protein 1; PC, Polycomb protein; p53BP1, p53-binding protein 1; CRB2, Cut5-repeat-binding protein 2. Adapted from Bannister and Kouzarides (2005).

Methylation of lysine 9 in H3 (H3K9) is recognized by chromodomain-containing proteins of the Heterochromatin Protein 1 (HP1) family, whereas methylation of H3K27 is recognized by chromodomain-containing Polycomb group proteins (Fischle et al., 2003b). In at least one protein, the *Arabidopsis* chromodomain-containing DNA methyltransferase CMT3, recognition of methylated H3 seems to require simultaneous methylation of both K9 and K27 (Lindroth et al., 2004). In addition, a recent report demonstrated that the chromodomain helicase CHD1 specifically binds methylated H3K4 (Pray-Grant et al., 2005). These interactions are summarized in Figure 1.3. Binding of

effector modules to methylated lysine is thought to recruit protein complexes which enforce either active or inactive transcriptional states, and thus perpetuate transcriptional memory (Kouskouti and Talianidis, 2005; Lorentz et al., 1994; Nakayama et al., 2001; Ogawa et al., 2002; Peters and Schubeler, 2005; Stewart et al., 2005).

The complexity of histone modifications and their recognition by specific effector modules led to the proposal that these modification/recognition systems constitute a "histone code", defined as a system in which "multiple histone modifications, acting in a combinatorial or sequential fashion on one or multiple histone tails, specify unique downstream functions" (Strahl and Allis, 2000). The existence of such a code is highly controversial, since it is difficult to formally prove that multiple modifications occurring after histone deposition on a single histone tail are distinguished combinatorially (Henikoff, 2005), as predicted by the model. What is clear, however, is that histone modifications can have profound effects on gene expression, and that some cross-talk between modifications is possible, as with the previously discussed example of recognition of dual-modified H3 by CMT3.

1.3 Methylation of non-histone targets

Recent reports have begun to expand the "language" of covalent histone modifications to non-histone proteins. It has been known for some time, for instance, that proteins other than histones can be acetylated by canonical HATs (Chan and La Thangue, 2001; Imhof et al., 1997). Most of the other known modifications of histones (phosphorylation, ubiquitination, ADP-ribosylation etc.) have similarly been well characterized to occur on non-histone proteins. Similarly, arginine methylation is known to exist on numerous

proteins, many of them transcriptional regulators or proteins involved in translation (Blanchet et al., 2005; Lee et al., 2005a; Miranda et al., 2005; Strahl et al., 2001; Teyssier et al., 2005; Tidwell et al., 1968). Until recently, the only exception was lysine methylation, which had been poorly studied outside of histones. Although reports of protein lysine methylation have existed for over thirty years (Allfrey et al., 1964; Paik and Kim, 1971), the identification of both the targets of methylation as well as the enzymes that modify them remained elusive.

Recent years have seen significant progress, beginning with the observation that the basal transcription factor TAF10 is lysine methylated, and that this methylation is carried out by the SET9 methyltransferase (Kouskouti et al., 2004). SET9 was previously described as a H3K4-specific enzyme, which was proposed to facilitate transcription both through methylation of H3K4, as well as through an inhibitory effect of H3K4 methylation on methylation of H3K9 (Chuikov et al., 2004; Nishioka et al., 2002). TAF10 monomethylation was shown to increase interaction of TAF10 with RNA Polymerase II, thereby potentiating transcription of target genes. This finding was consistent with initial studies indicating that SET9 could stimulate transcription, despite the fact that nucleosomal histones were poor *in vitro* substrates for SET9-mediated methylation (Nishioka et al., 2002).

Subsequent experiments have demonstrated that TAF10 is not the only non-histone target of SET9 methyltransferase activity. It was recently shown that the critical tumor-suppressor protein p53 also undergoes lysine monomethylation, and that this methylation is mainly carried out by SET9 (Chuikov et al., 2004). Methylated p53 was strictly nuclear, and was stabilized relative to unmethylated p53, suggesting a possible increase

in transactivation activity. Indeed, methylated p53 could be detected by chromatin immunoprecipiation (ChIP) at the p53-responsive p21 promoter after treatment of cells with adriamycin, although the relative affinity of methylated versus unmethylated p53 for promoter binding was not examined. As p53 is known to be subject to myriad post-translational modifications (Yang, 2005), it is possible that methylation may influence the presence and/or absence of these other marks as well.

The most recently reported non-histone target of lysine methylation is the yeast kinetochore protein Dam1 (Zhang et al., 2005). Dam1 was found to be methylated on multiple lysines, and this methylation was greatly reduced, but not eliminated, by mutation of the H3K4 methyltransferase Set1. Set1 itself is best known for its role in transcriptional activation through interaction with the RNA Pol II holoenzyme, via the Paf1 complex (Briggs et al., 2001; Roguev et al., 2001; Tenney and Shilatifard, 2005). Zhang et al. also demonstrated that lysine methylation of Dam1 reduces the ability of Ipl1 (the endogenous H3S10 kinase) to phosphorylate an adjacent serine, possibly suggesting a functional crosstalk between the two sites. However, the physiological relevance of these findings are questionable, since while mutation of a single methylation site in Dam1 is lethal, deletion of Set1, the enzyme claimed to be responsible for carrying out the methylation, has no effect on cell viability (Zhang et al., 2005).

Overall, three important points stand out from the existing data on non-histone lysine methylation. To begin with, all three of the proteins discussed above (TAF10, p53 and Dam1) are involved directly in transcriptional control or chromosome function. While this could be due to chance (considering the small number of non-histone targets identified to date), it is also possible that the pool of lysine methylated proteins is

inherently enriched in chromatin-associated proteins. Such enrichment in particular cellular compartments is plausible, considering that a large fraction of the arginine methylated proteins identified previously in an unbiased mass-spectrometry based analysis were unexpectedly found to be cytoplasmic proteins involved in translation (Ong et al., 2004).

The second important point is that the enzymes that methylate non-histone proteins are the same enzymes that methylate histones. The mouse genome encodes for approximately 55 SET-domain containing proteins (SCS, unpublished observation), only a small number of which have been characterized to date. Thus, it was formally possible that histone and non-histone methylation were carried out by separate groups of enzymes; however, this does not seem to be the case. An important implication of this is that mutant phenotypes for histone methyltransferases studied to date must be interpreted with caution, as they may reflect loss of both histone and more general protein methylation functions.

Finally, it is striking that none of the methylation sites identified to date bear any similarity to known target sites in histones (Chuikov et al., 2004; Kouskouti et al., 2004; Zhang et al., 2005). This is certainly surprising, given the remarkable target specificity of histone methyltransferases in general. The fact that targets sites in histones and non-histones are so dissimilar raises important questions about whether the fundamental tenets of the histone code will be generally applicable to all proteins, or whether they are unique to histones.

1.4 Lymphocyte development

In adult mammals, the development of B and T lymphocytes begins in the bone marrow, where common lymphoid progenitors (CLPs) reside. B cell development continues further in the bone marrow, whereas T cell precursors migrate to the thymus, where their development is completed. B cell development is a complex program, during which clonal antigen specificity is produced through the process of V(D)J recombination. The earliest steps of this process are carried out in progenitor (pro) B cells, in which recombinational coupling of diversity (D) to joining (J) segments occurs biallelically at the immunoglobulin heavy chain (IgH) locus. Successful completion of D-J joining is followed by monoallelic V-DJ joining, ultimately producing functional heavy chain of the μ isotype (IgH μ). The expression of membrane-bound μ chain in combination with surrogate light chain (SLC, comprised of λ5 and v-pre-B subunits) and various signaling components constitutes a functional pre-B cell receptor (pre-BCR), which signals clonal expansion, allelic exclusion at IgH (preventing further V-DJ recombination), and V-J recombination at the Ig light chain (IgL). Synthesis of complete surface IgM (IgH + IgL) signals the cessation of V(D)J recombination and the testing of immature B cells for both signaling capacity (positive selection) and autoreactivity (negative selection). Only newly produced B cells which have correctly rearranged Ig loci and display little reactivity towards self-antigens are stably selected into the peripheral population of mature, naïve B cells.

Whereas naïve B cells are marked by expression of surface IgM and IgD, antigen-activated B cells produce Ig of other functional classes, known as isotypes. Isotype switching is accomplished through the process of class switch recombination (CSR),

which, similar to V(D)J recombination, is a lymphocyte-specific, developmentally programmed DNA elimination event. In contrast to V(D)J recombination, however, the purpose of DNA rearrangement in this context is not to generate antigen specificities, but to alter the functional class of immunoglobulin produced by activated cells (IgA, IgG, IgE, etc.). This is accomplished by altering constant chain (C_H) usage via a deletional mechanism. CSR is initiated when naïve peripheral B cells are activated in the appropriate cytokine and cellular context, which together act to initiate transcription from cytokine-responsive promoters (I regions) upstream of the target switch (S) regions (Snapper et al., 1997). Transcription of the switch region corresponding to the target C_H is an invariant feature of CSR, and is required for efficient switching (Gu et al., 1993; Jung et al., 1993). After transcription has begun, the excision of Cμ/Cδ is accomplished by recombination between Sμ and the target switch region, releasing the intervening DNA as an episome (Iwasato et al., 1990). S regions have a highly unusual sequence composition, consisting mainly of repetitive and palindromic elements (Davis et al., 1980; Shimizu et al., 1982). This finding has fueled speculation that higher order structures produced during S region transcription might contribute to the specificity of DNA breakage at these sites during CSR (Kinoshita and Honjo, 2001; Ramiro et al., 2003; Shinkura et al., 2003).

While V(D)J and class switch recombination share many features in common, including transcription of target sequences, dependence on cytokine signaling for specificity, and restriction to particular developmental stages, they also differ in several significant respects. Most importantly, while CSR and the process of somatic hypermutation (SHM), a targeted mutagenesis program which results in increased

antibody affinity, are strictly dependent on the activity of the recently-discovered Activation-Induced cytidine Deaminase (AID), V(D)J recombination is not (Muramatsu et al., 2000). While the exact role of AID in control of CSR and SHM remains highly controversial, the dispensability of AID for V(D)J recombination strongly suggests basic mechanistic differences between this process and CSR. Conversely, while V(D)J recombination is completely dependent on activity of the Recombinase Activating Genes (RAG), CSR is not (Rolink et al., 1996). Finally, the sequence of Recombination Signal Sequences (RSSs), the recognition sites for the RAG recombinase during V(D)J recombination, bears no resemblance to that of S regions, suggesting that while RAGs may act as sequence specific endonucleases, S regions may rely on more general structural features to accomplish targeting during CSR.

1.5 Histone modification in the immune system

Several data point to a role for histone modification in the control of B cell development. For instance, while it has been recognized for some time that transcription of target loci precedes DNA rearrangement during V(D)J recombination (Sleckman et al., 1996), it was recently demonstrated that this developmentally programmed shift in transcriptional activity and recombinase accessibility is correlated with increased local levels of histone acetylation (Chowdhury and Sen, 2001; McMurry and Krangel, 2000). Moreover, D and J(H) segments were found to be acetylated earlier than V(H) segments, mirroring the order of recombination events at IgH. These findings suggest that at the molecular level, the established concept of "locus accessibility" might be defined by differential patterns of histone modification.

This possibility is supported by data suggesting that accessible D and J loci in both T and B cells are associated with diMeH3K4 and a lack of phosphorylation of H4 serine 1 (Morshead et al., 2003). Furthermore, recombinationally active gene segments are also enriched in H3K79 methylation, a mark placed by the single known non-SET domain HMTase, Dot1 (Feng et al., 2002; Ng et al., 2003). In contrast, recombinationally inactive loci are enriched in H3K9 methylation (Morshead et al., 2003), which in B cells is specifically removed in a process requiring the critical B cell lineage transcription factor Pax5 (Johnson et al., 2004). These findings are summarized in Figure 1.4.

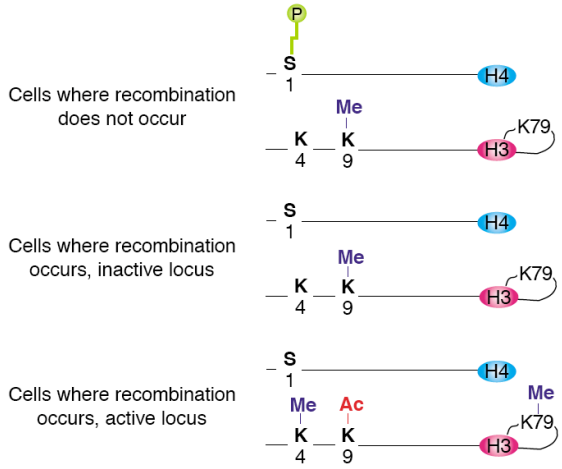

Figure 1.4 Patterns of histone modification on active and inactive antigen receptor loci. Abbreviations: P, phosphorylation; Me, methylation; Ac, acetylation; K, lysine. Adapted from Margueron et al. (2005).

A possible link between locus accessibility and particular histone modification patterns was further supported by the observation that 3' V(H) genes, which are preferentially utilized early in B cell development, are acetylated prior to 5' V(H) segments (Chowdhury and Sen, 2001; Yancopoulos et al., 1984). In addition, 5' V(H) genes were found to be hyperacetylated specifically in response to interleukin-7 (IL-7),

correlating well with the observation that mice lacking IL-7Rα have defects specifically in usage of 5' V(H) genes, and that these defects are associated with a lack of germline transcripts corresponding to these segments (Chowdhury and Sen, 2001; Corcoran et al., 1998).

Strikingly, it was recently found that mice lacking the H3K27-specific histone methyltransferase Ezh2 harbor very similar defects in variable chain usage to those observed in IL-7Rα mice, despite any discernible effect of Ezh2 deficiency on IL-7 signaling per se (Su et al., 2003). The defect in 5' V(H) usage in Ezh2-null mice was, in fact, even more specific than in IL-7Rα-null mice, since germline transcription of these sequences was not altered by loss of Ezh2 function. This finding suggests the possibility that while histone acetylation may regulate locus accessibility broadly defined, histone methylation may be responsible for recombinase targeting itself (Su et al., 2003). Together, the existing studies on histone methylation and acetylation strongly imply that histone modification may be a key regulatory pathway in lymphocyte development, and that ordered recruitment of various histone-modifying activities may underlie the exquisite specificity of DNA recombination reactions in lymphocytes.

The hypothesis that histone methylation may "mark" target sequences during DNA rearrangement events in the immune system is supported by a similar function of histone methylation in the unicellular eukaryote *Tetrahymena thermophila*. This organism is unique in that it contains two nuclei: the germinal micronucleus is transcriptionally silent and responsible only for transmission of genetic information during mating (called conjugation), while the somatic macronucleus is transcriptionally active and responsible for the phenotype of the cell, but is lost during mating. The process of conjugation

between cells of compatible mating type involves serial meiotic and mitotic division of each micronucleus, reciprocal transfer of haploid pronuclei, pronuclear fusion to form a zygotic micronucleus, and production of a new macronucleus from the zygotic micronucleus. During development of the new macronucleus, thousands of interspersed genetic elements, termed Internal Eliminated Sequences (IESs), are removed from the genome through a deletional process (Coyne et al., 1996). Recently, it was found that this DNA elimination program is tightly correlated with H3K9 methylation of histones associated with DNA destined for elimination, and that "tethering" to ectopic sites of a chromodomain protein which recognizes the methyl-lysine mark (Pdd1p) was sufficient to promote deletion (Taverna et al., 2002). Likewise, disruption of parental Pdd1p was found to interfere with DNA deletion (Coyne et al., 1999; Madireddi et al., 1996).

The similarities between programmed DNA elimination in *Tetrahymena* and processes in the immune system, such as V(D)J recombination and CSR, are numerous. Both nuclear downsizing and CSR, for instance, are developmentally programmed, depend upon transcription of target sequences, and occur at sites marked by irregular repetitive elements. Similarly, DNA elimination in Tetrahymena has been reported to coincide with bidirectional transcription of the sequences to be eliminated, and processing of these transcripts to form small RNAs capable of guiding DNA elimination (Chalker et al., 2005; Chalker and Yao, 2001; Mochizuki et al., 2002; Mochizuki and Gorovsky, 2004; Yao et al., 2003). Remarkably, a recent report indicated that V(D)J recombination at the Ig heavy chain locus is also associated with production of antisense intergenic transcripts, which might be capable of being further processed via an RNAi-like mechanism (Bolland et al., 2004). Considering that H3K9 methylation is required for DNA elimination in

Tetrahymena (Liu et al., 2004), it is possible that DNA modification programs in the immune system, such as V(D)J recombination and CSR, might also be guided by H3K9 methylation.

In addition to its possible role in providing specificity for particular processes in B cell development, H3K9 methylation may hold more general meaning for the function of peripheral lymphocytes. In their normal unactivated state, naïve B and T cells are non-proliferative and have limited lifespan. Conversely, spontaneous activation and proliferation of these cells in the absence of appropriate stimuli is pathogenic, and thus associated with autoimmune disease. Recently, it was demonstrated that a transcriptional complex capable of repressing E2F, Myc and Brachyury-responsive promoters is specifically active in quiescent cells (Ogawa et al., 2002). Strikingly, this complex was found to contain the human homologs of the related H3K9-specific methyltransferases G9a and Eu-HMTase (also known as G9a-like protein, GLP). These proteins have now been shown to associate with multiple transcriptionally repressive complexes, and are thought to be the dominant HMTases mediating MeH3K9-dependent repression in euchromatic chromosomal domains (Ogawa et al., 2002; Shi et al., 2003; Tachibana et al., 2002). This conclusion is consistent with the embryonic lethality caused by complete loss of G9a or GLP function (Tachibana et al., 2002; Tachibana et al., 2005).

1.6 Hypothesis and goals of this study

Based on the observations mentioned above, we propose that G9a and GLP may be involved in the transcriptional repression of cell-cycle associated genes in quiescent, naïve lymphocytes. Moreover, multiple lines of evidence point to the possibility that

H3K9 methylation might play a role in guiding DNA recombination reactions in developing lymphocytes. Should this be the case, the proven importance of G9a and GLP for proper euchromatic H3K9 dimethylation make them excellent candidates for carrying out this function, as well. Taken together, it is clear that significant insight into the functions of histone methylation in the immune system could be gained through production of mutants lacking G9a or GLP function. Since complete deletion of either of these genes in the mouse is lethal (Tachibana et al., 2002; Tachibana et al., 2005), we chose to investigate these issues further by producing conditional knockout alleles of G9a and GLP.

CHAPTER 2: MATERIALS AND METHODS

2.1 Cloning of targeting constructs

For isolation of the G9a genomic clone, a BAC (RPCI24-165L4; BL/6; CHORI, Oakland CA, USA) containing the murine G9a locus was obtained. This was digested with a cocktail of enzymes to release a ~10kb fragment containing the entire region containing the long and short arms of homology, which was subcloned into pBlueScriptIIKS+. This was partially digested with SapI, and an annealed double-stranded oligonucleotide containing the upstream loxP site and a HindIII site was inserted. The vector was then digested with BsaBI, and a NsiI fragment from pZero-loxP-FRT-neo-FRT(-) (D. O'Carroll, Rockefeller University) containing the downstream loxP site and FRT-flanked *neo* gene was inserted and screened for proper orientation. The resulting plasmid was digested with NheI and AflII to release the targeted locus, which was inserted into XhoI-digested pDTA-TK (D. O'Carroll, Rockefeller University) to produce the final targeting construct.

The GLP genomic locus was recombinogenically subcloned from a BAC clone (RPCI24-156K12; BL/6; CHORI, Oakland CA, USA) essentially as described (Yu et al., 2000; Lee et al., 2001) into pBlueScriptIIKS+ using primers which inserted a 5' AscI site and 3' FseI site. Briefly, EL350 cells were grown overnight in 5 ml of LB at 32C. 0.5-1 ml of the overnight culture was innoculated into 50ml of fresh LB in a 500ml flask and grown at 32C with shaking to OD600 0.5-0.8. The cells were then shaken in an ice bath slurry by hand to cool the sample down quickly. 10 ml of the culture was pelleted for 8 minutes at 7000 rpm (5500g) at 4C. The cell pellet was washed 3 times with 1 ml of cold sterile water and resuspended in 100ul of ice-cold water for electroporation. 50 ul of

competent cells was transferred to pre-chilled electroporation cuvettes (0.1 cm Bio-Rad Cat# 165-2089), 1 ug of freshly prepared BAC DNA was added and electroporated at 1.8kV on a Bio-Rad Gene Pulser set at 25uF and 200 ohms. 1 ml of LB medium was added to the cuvette and transferred to an eppendorf tube, incubated at 32C for 1-1.5 hours and plate on chloramphenicol plates.

To excise the genomic fragment by recombineering, the BAC-containing EL350 strain was grown overnight in 5 ml of LB at 32^{o}C. 0.5-1 ml of the overnight culture was innoculated in 50ml of fresh LB in a 500ml flask and grown at 32C with shaking to OD600 0.5-0.8. RED induction was performed by transferring 10 ml of the growing culture into a 125ml conical flask in a water bath at 42C with shaking (200 revs/min) for 15 min. Immediately after induction, the flask was placed into an ice bath slurry and shaken by hand to cool the sample. 10 ml of the culture was spun for 8 min at 7000 rpm (5500g) at 4C, and the cell pellet washed 3 times with 1ml of cold sterile water. After the final wash the cells were resuspended in 100 ul of ice-cold water, and 50 ul of cells was transferred to a pre-chilled electroporation cuvette (0.1 cm). 100ng of pBS PCR amplicon was added and electroporated at 1.8kV on a Bio-Rad Gene Pulser set at 25 uF and 200 ohms. 1 ml of LB medium was immediately added to the cuvette, transferred to an eppendorf tube, and shaken at 32C for 1-1.5 hours before plating on ampicillin plates.

The properly recombined plasmid was digested with NheI, and an annealed double-stranded oligonucleotide containing the upstream loxP site and a BsrGI site was inserted. The vector was then digested with BamHI, and a NsiI fragment from pZeroloxP-FRT-neo-FRT(-) (D. O'Carroll, Rockefeller University) containing the downstream loxP site and FRT-flanked *neo* gene was inserted and screened for proper orientation. The

resulting plasmid was digested with AscI and FseI to release the targeted locus, which was inserted into XhoI-digested pDTA-TK (D. O'Carroll, Rockefeller University) to produce the final targeting construct.

2.2 ES cell methods and production of mice

Embryonic stem cell culture was performed as described (Torres and Kuhn, 1997) using the ES cell line E14.1 (129/Ola). ES cells were cultured on embryonic feeder (EF) cells in ES cell medium containing leukemia inhibitory factor (LIF). Both LIF conditioned medium and EF cells harboring a *neo*^{r}-gene were prepared as described (Torres and Kuhn, 1997). The batch of FCS used for ES cell culture was tested to yield optimal plating efficiency without differentiation of ES cells. EF cells were passaged at most four times and treated with Mitomycin C (Sigma) (10 μg/ml in EF cell medium) for 2 hours at 37C immediately before use for ES cell culture. To maintain their undifferentiated state, ES cells were always passaged at a subconfluent stage. For passaging, ES cells were washed twice with PBS and trypsinised at 37C for 1-3 minutes. Trypsinization was stopped by adding 3-5 volumes of ES cell medium, followed by pipetting to yield a single cell suspension. ES cells were passed at 1:10-1:20 dilutions. Cells were frozen at –80C in ES cell medium containing 10% dimethylsulfoxide (DMSO). After freezing, ES cells were transferred to liquid nitrogen for long term storage.

All procedures concerning the manipulation of ES cells *in vitro* were performed as described (Torres and Kuhn, 1997). Briefly, $1x10^7$ ES cells were mixed with 25 μg of linearised targeting vector DNA (0.5 μg/ml) in 0.8 ml PBS. Following transfection by

electroporation (500 μF, 230V), ES cells were plated at $2x10^6$ per 10 cm tissue culture dish. G418 selection (350 μg/ml) for the presence of the *neo*^{r}-gene was started at 24 hrs and gangcyclovir selection (2 μM; to select for absence of the TK gene) was started 5 days after transfection. The selection medium was changed daily. Between day 7-10 after transfection, G418/gangcyclovir resistant ES cell colonies were picked and trypsinized. Cells were split in two, and half the sample used for PCR screening (see below, "PCR"). Positive clones were serially expanded in EF coated 48, 12 and 6-well plates subconfluently, and also plated at time of passage into 6 wells onto 6-well dishes pre-coated with gelatin as described (Torres and Kuhn, 1997). ES cells grown in the absence of feeders were used for Southern typing, while the cells grown on feeders in 6-well dishes were frozen as described above.

Preparation of ES cells for blastocyst injection was done as described (Torres and Kuhn, 1997). Injections were performed by the Rockfeller University Transgenics Facility. Chimeric mice were crossed to BL/6 and assessed for germline transmission by coat color. Deletion of the FRT-flanked *neo* gene was accomplished by crossing to eFLP-transgenic mice (Rodriguez et al., 2000). Routine typing was performed by tail biopsy and PCR.

2.3 Mice

Mice were maintained in the Rockefeller University Laboratory Animal Resource Center under Specific Pathogen Free (SPF) conditions. All procedures were approved by the Institutional Animal Care and Use Committee (IACUC). G9a and GLP mice were maintained on a mixed BL/6-129 genetic background.

2.4 Genomic DNA isolation

For analysis of targeted mice, genomic DNA was isolated from single cell suspensions of total splenocytes. ~1X10^8 cells were lysed in 0.5 ml tail lysis buffer [10 mM Tris-HCl pH8.5, 5mM EDTA, 0.2% sodium dodecyl sulfate (SDS), 200 mM NaCl, 100 µg/ml proteinase K] at 55°C over night. After centrifugation at 15,000g for 10 min, the supernatant was transferred to a fresh 1.5 ml reaction tube. 1 ml 100% ethanol was added and shook vigorously. The precipitate was spun for 5 min at 15,000g, washed once in 70% ethanol and briefly air dried. Samples were resuspended in 100-500 µl 10 mM Tris-HCl pH 8.0. For analysis of ES cells, pelleted cells were lysed directly in tail lysis buffer and processed as described above.

2.5 Genomic DNA digestion

For standard Southern blot analysis, 15 ug of genomic DNA prepared as described above was digested overnight in a 50 ul reaction with 100 units of the appropriate enzyme in 1X reaction buffer (all enzymes and buffers from NEB). After overnight digestion, 20 units of enzymes were added and the reaction continued for a further 2 hours. Reactions were stopped with 1X Orange G loading buffer for electrophoresis.

For analytic restriction digests, 5-10 fold excess of enzyme was used to digest DNA for 1 hour at 37C per manufacturers protocol (all enzymes purchased from NEB) before stopping with loading buffer.

2.6 Southern blotting

DNA was electrophoresed on either 2% (for RT-PCR products) or 0.8% (for genomic DNA) agarose gels and rotated in transfer buffer (0.6 M NaCl, 0.4 M NaOH) for 45 min. DNA was transferred by upward capillary transfer onto Hybond-N membrane (Amersham) overnight. Blots were rotated in neutralization buffer [0.5 M Tris-HCl (pH 7.0), 1 M NaCl] for 15 min and crosslinked in a Stratalinker (Stratagene) on "auto" mode. Blots were incubated briefly in 2XSSC, and transferred into pre-warmed hybridization buffer [50 mM Tris-HCl (pH 7.5), 1 M NaCl, 1% SDS, 10% dextran sulfate, 300 ug/ml sonicated salmon sperm DNA] and pre-hybridized for 3 hours at 65C.

Probes were PCR amplified from 0.01-0.1 ng BAC DNA and gel purified (Qiagen). DNA (50 ng) was labeled with 5 ul ^{32}P-α-dCTP using Ready-to-Go labeling beads (Amersham) and purified over ProbeQuant G50 spin columns per manufacturers protocol. Eluted DNA was boiled for 5 min and iced for 5 min prior to addition to pre-hyrbidized blot. Hybridization was carried out overnight at 65C. Blots were washed twice for 10 min each in 2XSSC/0.1% SDS, 1XSSC/0.1% SDS and 0.5XSSC/0.1% SDS, wrapped in SaranWrap and exposed to film (Kodak XAR).

2.7 RT-PCR and oligonucleotide Southern blot

A maximum of 5X10^6 cells were lysed in 1 ml Trizol (Gibco) and total RNA was prepared per manufacturers protocol. GlycoBlue coprecipitant (Ambion) was added to 50 ug/ml before isopropanol precipitation. 1 ug total RNA was reverse transcribed using oligo(dT) primer (Invitrogen) and SuperScript III reverse transcriptase (Invitrogen) per manufacturers protocol. 1 ul first strand cDNA was used for PCR amplification in a 20

ul reaction volume with the appropriate primers as detailed under "General PCR". Southern blots were performed as described under "Southern blotting" with the following modifications. Oligonucleotide probes (20 pmol) were labeled with 5 ul ^{32}P-γ-ATP in a 20 ul reaction using T4 Polynucleotide Kinase (NEB) and purified over ProbeQuant G25 spin columns (Amersham) per manufacturers protocols. Hybridization was performed at 50C overnight and washed/exposed as described under "Southern blotting". When needed, blots were exposed on a PhosphorImager screen (Amersham) and quantitated using ImageQuant software.

2.8 PolyI:C treatment

For induction of deletion in mice carrying the MX-cre transgene, mice were injected twice (separated by 72 hr) intraperitoneally with 200 ug polyI:C (Amersham). polyI:C was prepared by dissolving solid powder in sterile DPBS, heating at 56C for 3 hours and cooling slowly to room temperature. The dissolved polyI:C was sterile filtered and stored in aliquots at –20C.

2.9 FACS analysis

For routine FACS analysis of lymphocytes, mice were sacrificed by CO_2 euthanasia and organs collected sterilely. For splenocytes and thymocytes, single cell suspensions were prepared by scraping through a sterile plastic mesh in Hanks Balanced Salt Solution (BSS) with phenol red. Erythrocytes were lysed in 1 ml per spleen erythrocyte lysis buffer (0.75% NH4Cl, 100 mM Tris/HCl pH 7.65) for 3 min at room temperature, and the lysis stopped by addition of 10 ml cold BSS. $5X10^5$ cells were used per staining,

which was performed in U-bottom 96 well plates (Corning). Cells were stained in 15 ul BSS with antibody at the concentration listed below for 20 min at 4C , washed once with BSS and either stained with secondary antibody or resuspended in BSS with 2 nM ToPro-3 (Molecular Probes) for FACS.

2.10 Antibodies

FACS antibodies (all purchased from BD Pharmingen unless otherwise noted)

α-mouse CD19 (clone 1D3) FITC 1:100; α-mouse CD4 (clone GK1.5) PE 1:200; α-mouse CD43 (clone S7) FITC 1:100; α-mouse CD8 (clone 53-6.7) FITC 1:100; α-mouse CD90.2 (clone 53-2.1) PE 1:400; α-mouse IgD (clone 11-26C.2a) FITC 1:100; α-mouse IgM (clone 711116152) PE 1:300 (Jackson lab); α-B220 Cychrome (1:75).

Western blot/immunoprecipitation antibodies

αG9a (RU1061), rabbit, 1 ug/ml; α4xdiMeK9 (kind gift of T. Jenuwein, IMP, Vienna), rabbit, 6 ug/ml; αTubulin (source unknown), mouse, 1:5,000; αFLAG M2 (Sigma), mouse, 10 ug/ml; αdiMeG9aMS2 (RU1218), rabbit, 2 ug/ml; αLamin B (Santa Cruz Biotech), goat, 1:500; αHP1γ (Upstate), rabbit, 1:500; αHP1α (Upstate), rabbit, 1:500; αGLP (RU1160), rabbit, 1 ug/ml; αdiMeH3K9 (Upstate, #07-212), rabbit, 1:500.

2.11 PCR

For routine PCR (typing, etc) 1 ul sample was used in a 20 ul reaction containing 1X KlenTherm buffer [67 mM Tris-HCl (pH 9.1), 16 mM ammonium sulfate, 2.5 mM $MgCl_2$, 150 ug/ml BSA], 2.5 mM each dNTP, 0.4 uM each primer and 1 unit Taq polymerase (Sigma). Typing PCR conditions: 94C, 2 min; 94C, 1 min; 65C, 1 min, 72C,

1 min; 35 cycles. For ES cell screening PCR, PlatinumTaq (Amersham) was substituted in the reaction, and 42 cycles of amplification were performed.

2.12 PCR primers

Primer name	Sequence	Purpose
G9a type 1	TGCAGACTCTGCACCTTGCTCTTCTG	G9a typing
G9a type 2	GTGTGAGCCTGTGTTCTGGGGATTA	G9a typing
G9a type 3	CCGGAGATGAGGAAGAGGAGAACAG	G9a typing
GLP type 1	GGGTTGTGCTCAGAGTTTCTACCTC	GLP typing
GLP type 2	TCCCTCATCGCCCACATTTCTG	GLP typing
GLP type 3	CCGGAGATGAGGAAGAGGAGAACAG	GLP typing
G9a Probe G FWD	GGGCTCCTGGGCTCTATGAG	G9a probe G
G9a Probe G REV	ATCAGCTGGCAGAGGCCCAAC	G9a probe G
Probe E FWD	CGGGATCCGGACGTAGCCCGAGGCTATGAG	G9a probe E
Probe E REV	CGGGATCCCTTGTCATACCAGCATCGGATAC	G9a probe E
Probe 18-19 FWD	CCTCTTTCTTTCTCGGGATTCAG	GLP probe A
Probe 18-19 REV	CTGCAAATGAGTGATGTTCCTG	GLP probe A
LAH extl probe	CCTGTTAAACATGGCTGCTTG	GLP probe B

FWD (should be SAH)		
LAH extl probe REV (should be SAH)	TTCTCTTTTTCCTTGTCTGCCC	GLP probe B
CD19C	AACCAGTCAACACCCTTCC	CD19cre typing
CD19D	CCAGACTAGATACAGACCAG	CD19cre typing
CRE7	TCAGCTACACCAGAGACGG	CD19cre typing
MX CRE F	CATGTGTCTTGGTGGGCTGAG	MXcre typing
MX CRE R	CGCATAACCAGTGAAACAGCAT	MXcre typing
SD24	CTAATGTTGTGGGAAATTGGA	FLPer typing
SD25	CTCGAGGATAACTTGTTTATTG	FLPer typing

Table 2.1 PCR Primers used in this study. All sequences are given 5'-3'.

2.13 B cell purification

Splenic B cells were purified using α-CD43 microbeads (Miltenyi Biotec) per manufacturers protocol. Purity as judged by B220 staining was consistently >90%.

2.14 CFSE labeling and B cell stimulation

Purified splenic B cells were washed twice in plain RPMI and resuspended at $1X10^7$ cells/ml in plain RPMI with 2.5 μM carboxyfluoroscein diacetate succinimidyl ester (CFSE; Molecular Probes). After incubation at 37C for 10 min, 10 ml cold serum was added and cells were washed twice with complete medium [RPMI 1640, 10% Fetal Bovine Serum, 1X Penicillin/Streptomycin (Gibco), 1X L-Glutamine (Gibco), 50 uM betamercaptoethanol] before resuspension in complete medium for cell culture. $2x10^5$ cells were cultured per well of a 96-well flat bottom plate for 3-4 days in complete medium with varying stimuli at the following concentrations: α-F(ab)2 of IgM (1-5 μg/ml), rIL-4 (25 U/ml), LPS (5 μg/ml).

2.15 Protein lysate preparation

For routine protein lysate preparation, cells or purified nuclei were resuspended in an appropriate volume M2 lysis buffer [50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 1 mM EDTA, 1% Triton X-100) with 1X protease inhibitor cocktail (Sigma) and rotated at 4C for 30 min. After centrifugation at 14000 rpm for 15 min at 4C, the supernatant was moved to a new tube and quantitated by Bradford assay (Biorad). 50 ug of lysate were typically loaded per lane.

2.16 Western blotting

Gels were blotted in a submerged transfer chamber (Biorad) in 1X transfer buffer [48 mM Tris-HCl, 39 mM Glycine, 0.037% (w/v) SDS, 20% MeOH] overnight at 15 volts, 4C, onto 0.2 micron pore nitrocellulose (Schleicher and Schuell). Membranes were

stained with Ponceau S (Sigma) and destained with dH_2O before cutting individual strips for probing. Strips were blocked 1 hr at room temperature with 4% autoclaved milk in PBS. Antibodies (see separate list for concentrations) were diluted in block and added for 1 hr at room temperature. For all antibodies except FLAG M2, blots were washed 4 times 5 min in PBS/0.1% Tween 20 (PBST). For FLAG M2 blots, washing was done twice for 5 min in PBS. All secondary anibodies were used at 1:10,000 dilution in block, and all blots were washed 4 times 5 min in PBST. Detection was performed with ECL (Amersham) or Visualizer (Upstate) as required.

2.17 Peptides

All G9a peptides were produced by the Rockefeller University Proteomics Resource Center and were at least 85% pure. G9a MS2 peptides all included a C terminal cysteine and contained the base sequence KVHRARKTMSKPGC. H3 peptides were kindly provided by C.D. Allis (Rockefeller University).

2.18 Calcium phosphate transfection

For transfection of 1 10 cm dish, $5X10^6$ 293 cells were plated 1 day prior to transfection, and the medium was changed the next day prior to transfection. 10 ug plasmid DNA in 480 ul was mixed with 120 ul 1.25 M $CaCl_2$ and added dropwise to 600 ul 2X HBS (pH 7.0) with gentle vortexing. After 30 min incubation at room temperature, the precipitate was added dropwise to cells. Medium was replaced after 24 hr, and all cells were harvested 48 hr after transfection.

2.19 Nuclear fractionation

For fractionation of nuclei, $1X10^7$ cells were collected by trypsinization, washed once with cold DPBS, and resuspended in Buffer A [10 mM HEPES (pH 7.9), 5 mM $MgCl_2$, 0.25 M sucrose]. NP-40 was added to 0.1% final concentration, the mixture pipetted through an 18 guage needle 4 times, incubated at 4C 10 min, and spun 10 min at 8000 rpm, 4C. The supernatant was taken as the cytosolic fraction, and the pellet as the nuclear fraction.

2.20 Histone methyltransferase assays

HMTase assays were performed essentially as described (Rea et al., 2000) with the following modifications: reactions were scaled down to 20 ul volumes, using 2.5-5 ug peptide per reaction. For assay on beads, immunoprecipitations from ~0.5-$1X10^7$ transiently transfected 293 cells were used per reaction. Samples were stopped with 4X Laemmli buffer (to 1X final concentration) and separated by SDS-PAGE on 15% acrylamide gels. After Coomassie staining and destaining in 30% MeOH/10% acetic acid, gels were dried 1hr at 80C under vacuum before exposure to a PhosphorImager screen (described above).

2.21 Immunoprecipitation

For small scale immunoprecipiation (IP) of FLAG-tagged protein, 20 ul packed volume M2-agarose beads (Sigma) was washed extensively with M2 lysis buffer and added to lysed protein samples (see above). Samples were rotated overnight at 4C, washed 3 times with 0.5 ml lysis buffer and boiled in 1X Laemmli buffer.

For large scale FLAG M2 IP and small scale IP using αG9a antibody, M-270 Epoxy Dynabeads (FLAG) or Protein A Dynabeads (G9a; both from Dynal Biotech) were coated per manufacturers instructions. 100 ul (FLAG) or 10 ul (G9a) beads were used per IP, and washing was done as above. To elute FLAG protein for mass spectrometry, Dynabeads were resuspended after washing in 250 ug/ml 3xFLAG peptide (Sigma) and rotated for 1 hr at 4C. The supernatant was used for SDS-PAGE.

For coimmunoprecipitation of G9a with endogenous HP1γ, transfected 293 cells were arrested for 12 hr with 2 mM thymidine (Sigma) in medium 24 hr after transfection. 18 hr later, the fresh medium containing 20 uM MG-132 (Peptides International) was added, and this was left on the cells until harvesting.

2.22 Antibody generation

For generation of αG9a antibody, residues 1-281 of murine G9a were fused to GST in pGEX-6P-1. Protein production in transformed bacteria (*E. coli* BL21) was induced with 1 mM IPTG (Sigma) for 3 hr at 37C. Cells were lysed by sonication in lysis buffer (PBS, 1 mM EDTA, 1 mM EGTA, 1X protease inhibitor cocktail) and clarified by centrifugation. The supernatant was passed over a Glutathione-Sepharose column (Amersham Pharmacia Biotech), the column washed extensively with PBS and eluted with excess glutathione per manufacturers protocol.

GST-G9a antigen was coupled to KLH and used to immunize rabbits by Cocalico Biologicals. Sera were screened by Western blot against the immunizing antigen and a highly reactive serum (RU1061) was affinity purified against immobilized antigen on AffiGel 10 (Biorad) per manufacturers protocol. Peak fractions were pooled, dialyzed

against PBS/50% glycerol and stored at –80C. For peptide antibodies, affinity purification was performed by immobilization of peptides via a C-terminal cysteine on SulfoLink coupling gel (Pierce) per manufacturers protocol. Peak fractions were dialyzed and stored as above.

2.23 Peptide dot blots

For dot blots, 2 ul 0.5 mM peptide solution was spotted on pre-wetted moist nitrocellulose (0.2 micron pore, Scheicher and Schuell). Blots were dried overnight at room temperature before use, and were processed as described under "Western blotting".

2.24 Peptide pulldown assays

G9a peptides were coupled to SulfoLink beads (Pierce) per manufacturers protocol. 25 ul (bed volume) beads were added to nuclear extracts prepared in M2 lysis buffer and rotated overnight at 4C. Beads were washed 3 times in 0.5 ml lysis buffer and boiled in 1X Laemmli sample buffer.

Binding to *in vitro* translated proteins was performed essentially as described (Lachner et al., 2001). Where necessary, peptides were dephosphorylated with 25 units PP1 (NEB) for 30 min per manufacturers protocol. Samples were electrophoresed on 7.5-15% linear gradient acrylamide gels, and stained with Coomassie blue, destained, dried and exposed to PhosphorImager as described above.

2.25 Adenoviral infection and subcloning

Cre-expressing adenovirus was a kind gift of I. Su (Rockefeller University). Infections were carried out as described (Su et al., 2005). 2-3 passages after deletion, cells were plated at 0.2 cells/well in 96 well plates and allowed to expand to confluence. Wells showing growth were further expanded and tested for deletion. Two clones (#3 and #16) showing 100% deletion by PCR were selected for retroviral reconstitution. All experiments were performed with derivatives of clone #3.

2.26 Retroviral infection

SV40-large T retrovirus was kindly provided by S. Buonomo (Rockefeller University). FLAG-hG9a and mutations thereof were subcloned into pMSCV-Puro (Invitrogen). For production of retrovirus, $7.5X10^5$ BOSC packaging cells were plated in 1 well of 6 well dish and transfected 24 hr later with 1.4 ug retroviral vector and 0.6 ug pCL helper plasmid using Fugene (Roche) per manufacturers protocol. 48 hr later the supernatant was collected, passed through a 0.2 micron filter, and added directly to subconfluent target cells in the presence of 20 mM HEPES and 10 ug/ml polybrene (Sigma). 48 hr after infection, cells were selected by addition of medium containing 4 ug/ml puromycin (Sigma). Lines were subsequently maintained as bulk populations to avoid integration site effects.

2.27 Indirect Immunofluorescence on Mouse Embryonic Fibroblasts

Cells were seeded into 6 well dishes containing acid washed, poly D-lysine (Sigma) coated coverslips at $1x10^5$ cells per well. After 12-18 hr, the medium was removed, the

coverslips were washed twice with PBS at room temperature, and were fixed 10 min at room temperature in freshly prepared 4% parafomaldehyde/PBS. Coverslips were then washed twice in PBS, permeabilized with 0.5% Triton-X 100/PBS for 10 min at room temperature, washed twice with PBS, and blocked overnight at 4°C in PBS/5% skim milk/0.2% Tween 20. For detection, the coverslips were washed four times with AbDil (TBS/0.1% Triton-X 100/2% BSA) and incubated for 1 hr at room temperature with affinity purified αG9a (RU1061, 2 μg/ml), αHP1γ (2355 μg/ml; AbCam #ab10480), αdiMeH3K9 (1:200; Upstate #07-212), or control rabbit IgG (2 μg/ml) diluted in AbDil. Coverslips were then washed 4 times with AbDil and incubated with Alexa 488- (Molecular Probes), FITC- (Jackson), or X-Rhodamine-conjugated (Jackson) secondary antibodies diluted in AbDil for 1 hr at room temperature, washed 4 times with AbDil, incubated briefly with 0.25 μg/ml Hoechst 33258 in AbDil to counterstain DNA, washed once with AbDil, and mounted in 90% glycerol/PBS.

Imaging was performed using a Carl Zeiss Axioplan 2 microscope equipped with a Photometrics CoolSnap HQ cooled CCD camera, and controlled by MetaMorph software (Universal Imaging). Images were processed with MetaMorph and Adobe Photoshop.

2.28 Kinase Assay

Preparation of *Xenopus* egg extracts, immunoprecipitation using αIncenp antibody and Aurora B kinase assays were performed essentially as described (Sampath et al., 2004; Ohi et al., 2004). 2.5 ug peptide were used per reaction and samples were separated by SDS-PAGE, Coomassie stained, destained, dried and imaged as discussed above.

CHAPTER 3: GENERATION AND ANALYSIS OF G9A AND GLP MUTANTS

3.1 Domain structure and expression of G9a and GLP

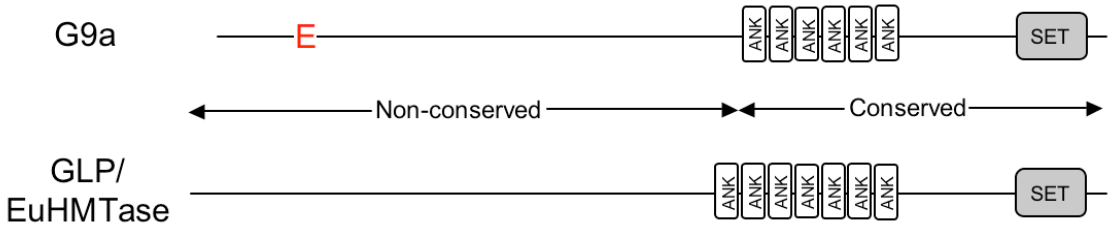

Figure 3.1 Domain structure of G9 and GLP. Schematic domain structure of murine G9a and GLP/EuHMTase-1. ANK, ankyrin repeat; SET, Su(var)3-9, E(z), Trithorax domain; E, gluatmic/aspartic acid repeats.

G9a and GLP are SET-domain containing proteins of the Suv39 family, all of which to date have been shown to act as H3K9-specific histone methyltransferases *in vitro* and *in vivo* (Dillon et al., 2005). Both proteins contain a C-terminal SET domain, which is responsible for enzymatic activity, as well as pre-SET and post-SET domains, which contribute to catalysis (Figure 3.1;(Dillon et al., 2005; Tachibana et al., 2001; Tachibana et al., 2002; Tachibana et al., 2005). Both G9a and GLP also contain 5-6 centrally located ankyrin domains, whose functions are unknown (Figure 3.1). In contrast to the high degree of sequence similarity between the C-termini of G9a and GLP (>65% sequence identity), the amino termini of the two proteins are relatively poorly conserved with one another (<35% sequence identity; data not shown). The lack of conservation in this region is consistent with the absence of identifiable protein domains, the only exception being a stretch of 27 consecutive glutamic and aspartic acid residues in G9a (denoted "E" in Figure 3.1). Despite this poor conservation, small "islands" of sequence

identity can still be identified in the amino termini of G9a and GLP (discussed further in Chapter 4).

Homologs of both G9a and GLP can be found in most vertebrates, while G9a homologs are also present in invertebrates, including *Drosophila*. Interestingly, in addition to the domains found in vertebrates, *Drosophila* G9a also contains an amino-terminal AT-hook, a protein domain thought to be involved in direct binding to AT-rich DNA sequences (data not shown). Strikingly, this AT-hook is embedded within one of the microdomains of sequence conservation between vertebrate G9a and GLP homologs (see Chapter 4).

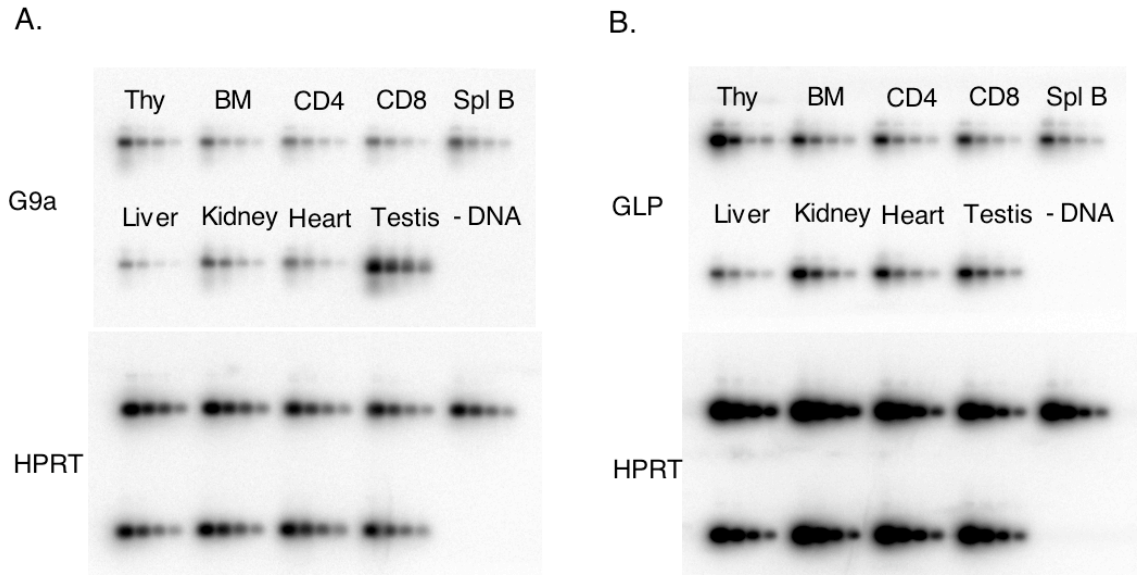

Figure 3.2 RT-PCR analysis of G9a and GLP expression patterns. Total RNA was extracted from the indicated tissues and amplified by RT-PCR with primers specific for either G9a (A) or GLP (B), followed by Southern blotting and hyrbidization with a transcript-specific internal probe. –DNA indicates no RT samples was added for PCR.

In order to determine whether G9a and GLP might have functional roles in the immune system, we first sought to examine the expression pattern of these genes in various immune and non-immune cell types. Total RNA was extracted from five different immune cell populations (total thymocytes, total bone marrow, $CD4^+$ T cells, $CD8^+$ T cells and splenic B cells) as well as four control tissues (liver, kidney, heart and testis). These samples were reverse-transcribed and PCR-amplified using primers specific for either G9a, GLP or HPRT (loading control). The PCR products were separated by agarose gel electrophoresis, Southern blotted, and hybridized with internal probes specific for each transcript. As shown in Figure 3.2A and B, expression of G9a and GLP was readily detectable in all cell types tested, and at approximately equal levels. Expression of G9a was moderately increased in testis, but neither G9a nor GLP showed any significant overexpression in immune cell types.

3.2 Generation of conditional G9a and GLP mutant mice

In order to produce genetic systems for understanding the physiological functions of G9a and GLP, it was necessary to produce targeted deletions in these genes. Both G9a and GLP have previously been disrupted by conventional gene targeting, resulting in total loss of function and embryonic lethality beginning at approximately gestational day E9.5 (Tachibana et al., 2002; Tachibana et al., 2005). For this reason, it was necessary to produce conditional mutations, which allow specific deletion of the target gene in particular tissues, thereby bypassing the complication of premature lethality (Lewandoski, 2001). In order to conditionally inactivate G9a, a targeting construct was generated which contained exons 23 and 24 flanked by loxP sites ("floxed"; Figure 3.3).

These two exons encode the last third of the pre-SET domain and first quarter of the SET domain, respectively, both of which are necessary for catalytic activity (Tachibana et al., 2001). Thus, deletion of these exons *in vivo* by Cre recombinase expression should cause production of a catalytically inactive protein. In addition, Cre-mediated deletion is predicted to cause out-of-frame splicing of exons 22 and 25, leading to nonsense-mediated decay (NMD) of the mutant transcript. The targeting vector also contained a neomycin-resistance gene (*neo*) for positive selection, as well as *Thymidine Kinase* (TK) and *Diphtheria Toxin* (DTA) cassettes outside the arms of homology to allow negative selection. The *neo* gene was flanked by FRT sequences (recognition sites for the *FLP* recombinase), in order to allow removal of the selection cassette *in vivo*.

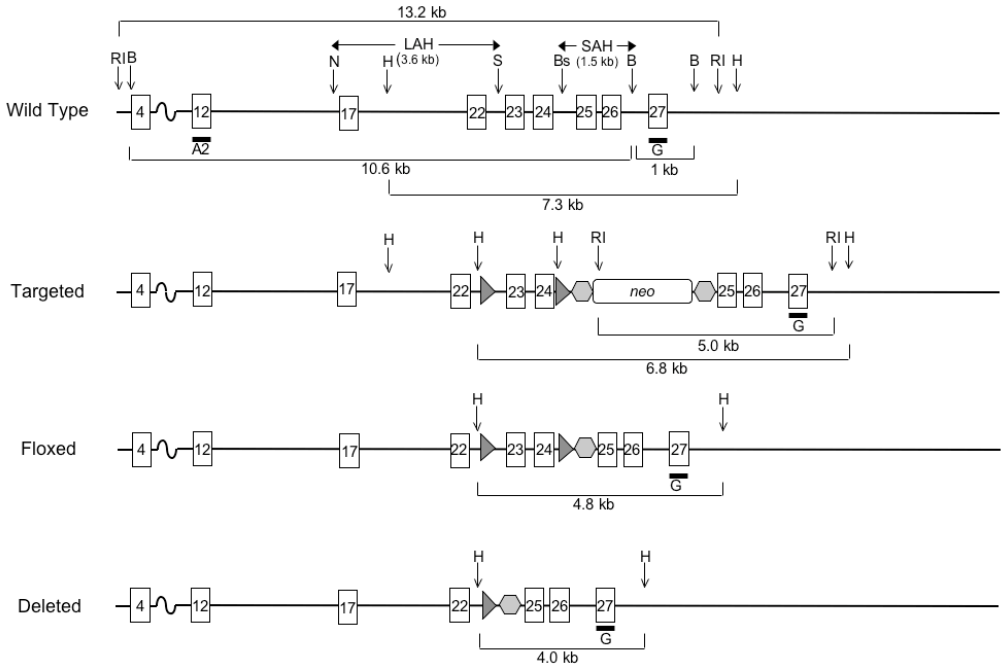

Figure 3.3 G9a conditional targeting strategy. Partial diagram of the G9a genomic locus. Exons are indicated by unfilled rectangles, loxP sites by filled triangles, FRT sites by shaded octagons, and the location of Southern probes by dark rectangles. Restriction sites are abbreviated as follows: RI, EcoRI; B, BamHI; H, HindIII.

E14.1 ES cells (129/Ola genetic background) were electroporated with the G9a targeting construct, and cells were selected for resistance to G418. Resistant colonies were expanded and screened by Southern blot for proper targeting using probes G (external to the short arm of homology) and E (internal to the long arm of homology). Proper targeting was achieved in multiple clones (Figure 3.4A), and these were injected by the Rockefeller University Transgenics facility into blastocysts from pseudopregnant BL/6 female mice. Chimeric progeny were bred to BL/6 partners, and germline transmission was obtained from a single male founder (derived from ES cell clones 1A4 and 1F10). Germline progeny were crossed to mice expressing the enhanced *FLP* transgene (Rodriguez et al., 2000) in order to delete the *neo* cassette, leaving behind a single scarring FLP site and the conditionally targeted locus (Figure 3B.1). Proper deletion of the *neo* cassette was confirmed by Southern blotting of splenocyte DNA from floxed (fl) mice (Figure 3.4B).

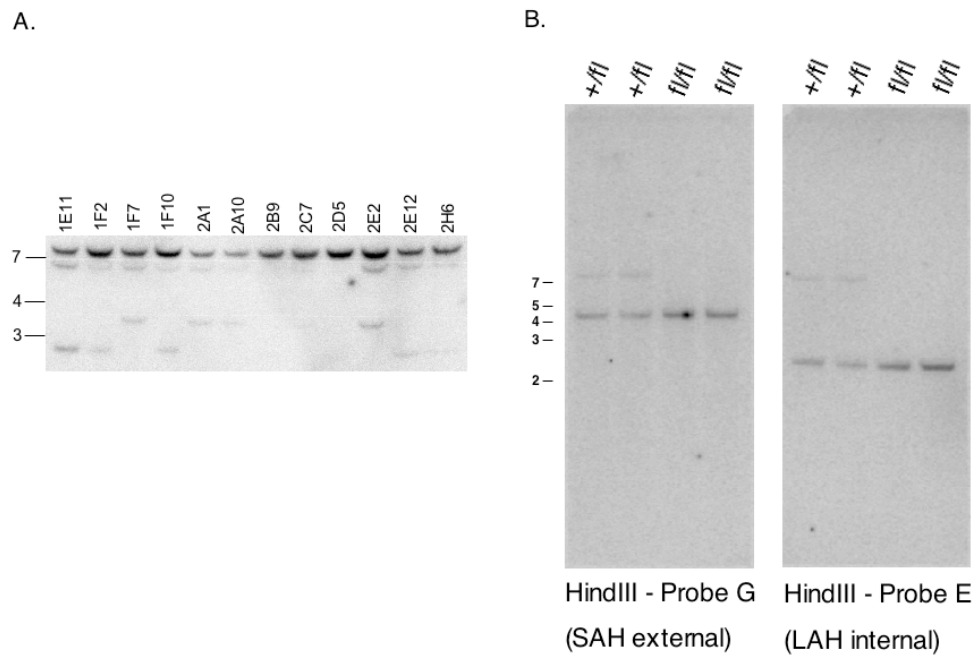

Figure 3.4 Evaluation of G9a targeting by Southern blot. A. Genomic DNA samples from the indicated ES cell clones were digested with HindIII and blotted sequentially with probes E and G. B. Splenocyte DNA from progeny of the indicated genotypes after *neo* cassette removal was digested with HindIII and probed with probe G (left) or probe E (right). Probe E fragment sizes: WT, 7.3 kb; Targeted (with upstream loxP site), 2.5 kb; Targeted (without upstream loxP site), 3.4 kb; Floxed, 2.5 kb. Probe G fragment sizes: WT, 7.3 kb; Targeted (with or without upstream loxP site), 6.1 kb; Floxed, 4.8 kb.

Prior to conditional targeting of G9a with the loxP-FRT construct, we also attempted to produce G9a mutants using the more traditional "three loxP" system, in which the *neo* gene is flanked by loxP sites, and removed in ES cells after targeting by transient transfection with a Cre-expressing plasmid. While proper targeting was achieved with this construct, we were never able to derive heterozygously floxed ES cells, since transient expression of Cre invariably drove complete recombination to produce the deleted allele (data not shown). This was most likely due to the presence of an extremely small floxed region in our targeting strategy (957 bp), making complete deletion highly

efficient. Nonetheless, this targeting did produce heterozygously deleted ES cells, which we were able to use to examine whether the mutant transcript underwent NMD.

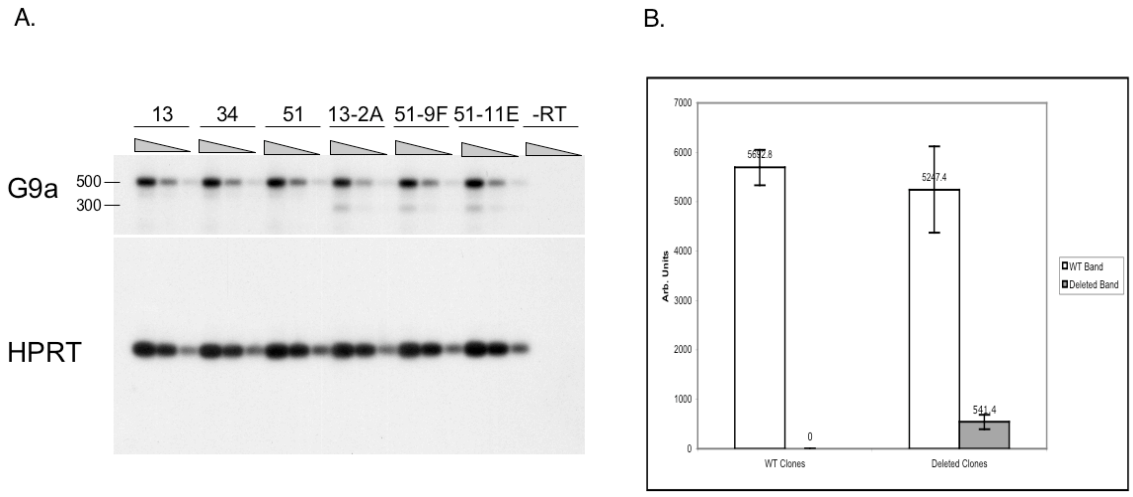

Figure 3.5 Nonsense-mediated decay of G9a$^{\Delta}$ transcript. A. RT-PCR and Southern blotting was performed on three-fold serial dilutions (1, 1:3, 1:9) of first-strand cDNA from three G9a$^{Targ/+}$ ES cell clones (13, 34, 51) and three G9a$^{\Delta/+}$ clones (13-2A, 51-9F, 51-11E). Expected band sizes: WT transcript, 509 bp; deleted transcript, 306 bp. B. The results from A. were quantitated by PhosphorImager analysis and normalized to HPRT values. Error bars represent the standard deviation for three clones.

Total RNA was extracted from three parental and three deleted ES cell clones, and used for RT-PCR with primers spanning the deleted exons. After electrophoresis and Southern blotting, the bands were quantitated by PhosphorImager analysis and normalized to the HPRT control. While a smaller band corresponding to the deleted transcript was not seen in the parental clones, such a band was readily seen with all three heterozygously deleted clones (Figure 3.5A). Quantitation revealed that the deleted band was present at only 10% the level of the wild type band in the deleted cells (Figure 3.5B). Assuming biallelic transcription of G9a, this reduced abundance indicates that the mutant transcript efficiently undergoes NMD.

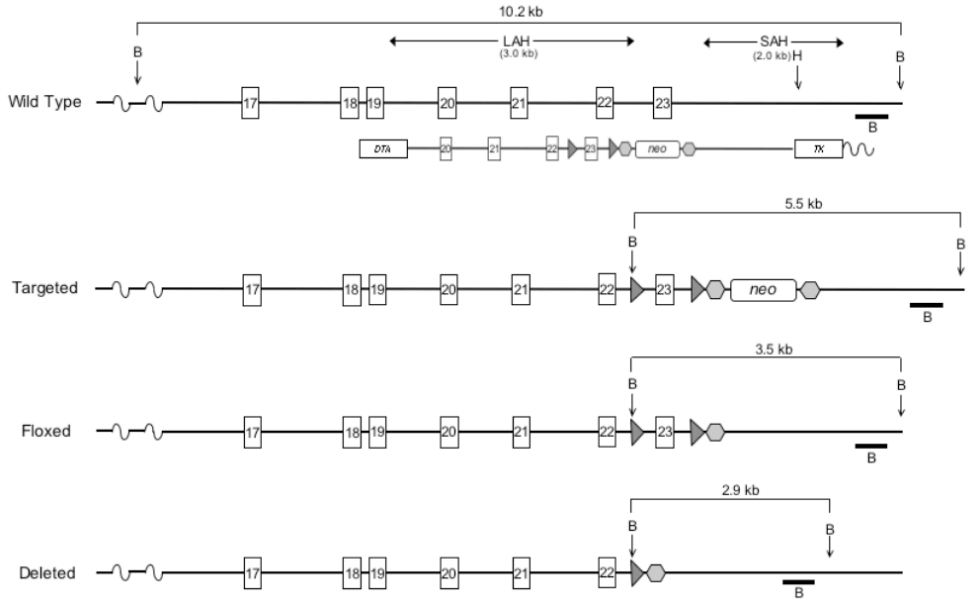

Figure 3.6 GLP conditional targeting strategy. Symbols are as described in Figure 3.3. Abbreviations: DTA, Diphtheria Toxin gene; neo, Neomycin resistance cassette; TK, Thymidine Kinase gene.

The targeting strategy for GLP was very similar to that for G9a. The overall exon-intron structure of the two proteins is well conserved (Dillon et al., 2005), making a similar strategy feasible. The SET domain of GLP is encoded by exons 22, 23 and 24, but deletion of exon 22 (corresponding to exon 23 of G9a) was not possible due to the likelihood of subsequent in-frame splicing. Instead we chose to flox exon 23 of GLP (corresponding to exon 25 of G9a), which encodes the second quarter of the SET domain, and deletion of which should lead to out-of-frame splicing of exons 22 and 24 (Figure 3.6).

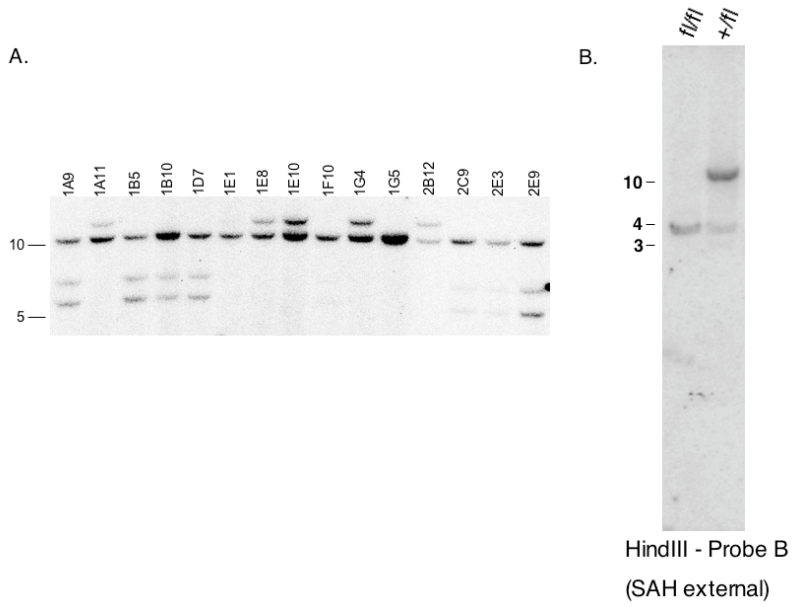

Figure 3.7 Evaluation of GLP targeting by Southern blot. A. Genomic DNA samples from the indicated ES cell clones were digested with BsrGI and blotted sequentially with probes A and B. B. Splenocyte DNA from progeny of the indicated genotypes after *neo* cassette removal was digested with HindIII and probed with probe B. Probe A fragment sizes: WT, 10.2 kb; Targeted (with upstream loxP site), 6.7 kb; Targeted (without upstream loxP site), 12.3 kb. Probe B fragment sizes: WT, 10.2 kb; Targeted (with upstream loxP site), 5.5 kb; Targeted (without upstream loxP site), 12.3 kb Floxed, 3.5 kb.

A loxP-FRT targeting construct was prepared for GLP in essentially the same manner as for G9a. Proper targeting in ES cells was assessed by hybridization of Southern blots of BsrGI-digested ES cell DNA with probes A (external to the long arm of homology) and B (external to the short arm of homology). Correctly targeted ES cells (Figure 3.7A) were injected into BL/6 blastocysts by the Rockefeller University Trangenics facility. Germline transmission was obtained from several male chimeras, all derived from clone 1D7. These were bred to eFLP-transgenic mice to delete the *neo* cassette, leaving the properly targeted floxed allele (Figure 3.7B).

3.3 Analysis of G9a deletion in central lymphoid organs

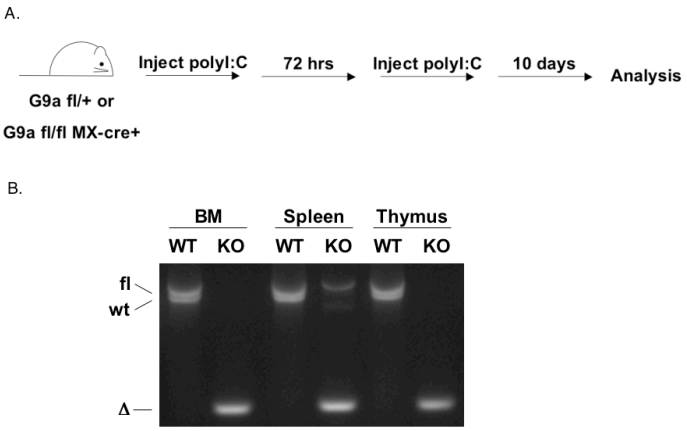

Figure 3.8 Deletion of G9a *in vivo* with MX-cre. A. Strategy for deletion using MX-cre. B. PCR analysis of genomic DNA harvested from the indicated tissues 10 days after the final polyI:C injection.

To begin to address the issue of the contribution of G9a to the development of lymphocytes, we crossed the $G9a^{flox}$ allele to the MX-cre driver strain (Kuhn et al., 1995). This strain carries Cre under the control of the interferon-responsive *MX* promoter, which can be induced by intraperitoneal injection of mice with the double-stranded RNA mimic, polyI:C. $G9a^{fl/+}$ or $G9a^{fl/fl}$ MX-cre^{+} mice were injected with 200 μg polyI:C twice over 72 hours, and sacrificed ten days later for analysis (Figure 3.8A). Deletion of the G9a locus was first assessed by PCR typing of deleted tissues. As has been observed with other conditional mutants using the MX-cre driver, deletion was extremely efficient in both bone marrow and thymus, but incomplete in peripheral lymphoid organs such as the spleen (Su et al., 2003).

We first examined B cell development in the bone marrow by flow cytometry. Pro- and pre-B cells can be distinguished from more mature B cell lineages in the bone

marrow by expression of the CD45 glycosylation isoform B220 and their lack of surface μ chain (IgM) expression ($B220^{+}IgM^{-}$). As cells complete V(D)J recombination to produce a functional B cell receptor, they become surface IgM^{+} immature B cells. These eventually emigrate to the peripheral lymphoid organs (spleen and lymph nodes), and recirculate through the bone marrow as $B220^{hi}IgM^{+}$ cells. As shown in Figure 3.9 (left upper and lower panels), the overall fraction of $B220^{+}IgM^{-}$ pro- and pre-B cells was unchanged in G9a-deleted mice ten days after the final injection of interferon.

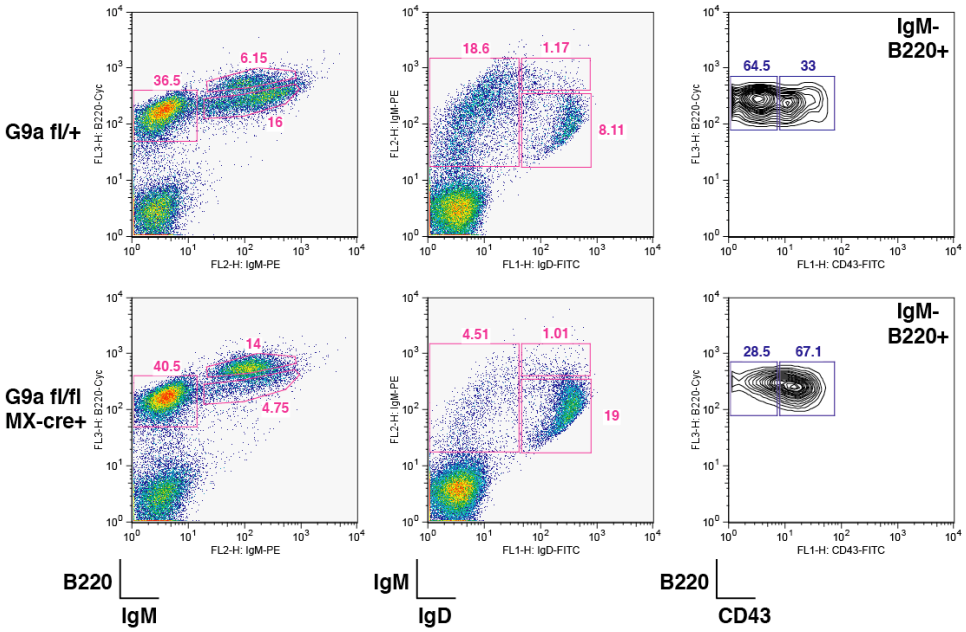

Figure 3.9 Loss of G9a function causes a block in B cell development. Mice of the indicated genotypes were sacrificed 10 days after the final polyI:C injection and total bone marrow was stained as indicated and analysed by FACS. B220/CD43 staining in the right-hand panels was gated on IgM-B220+ cells as indicated in the left-hand panes.

In contrast, a clear reduction in the frequency of $B220^{+}IgM^{+}$ immature B cells was noted, with a slight increase in the frequency of recirculating B cells. As B cells mature,

they lose IgM expression and begin to express heavy chain of the IgD isotype due to biased alternative splicing of the heavy chain transcript. The depletion of IgM^+ immature B cells was also apparent in stainings for IgM and IgD in bone marrow (Figure 3.9, middle panels), where a corresponding increase in IgD^+ was also noted. These cells represent mature recirculating cells, most of which are likely to be undeleted (Figure 3.8B).

The block in development of B cells to the $sIgM^+$ stage could reflect a defect either in progression of pro-B cells to the pre-B cell stage, which depends upon successful heavy chain recombination, or failure of pre-B cells to complete light chain recombination. The precise location of this block can be determined by staining of the pro/pre-B cell population with αCD43 antibody, which preferentially stains pro-B cells. This staining revealed that the $B220^+IgM^-$ population in G9a-deleted mice was dramatically enriched in $CD43^+$ cells, indicative of a block in pro-to-pre B cell development. This phenotype is qualitatively very similar to the effect of deletion of the H3K27-specific histone methyltransferase Ezh2 in bone marrow (Su et al., 2002), and also resembles the phenotype of mice lacking a functional receptor for Interleukin 7 (IL-7), a key trophic factor for pro-B cells (Corcoran et al., 1998).

A major concern in interpretation of lymphocyte developmental phenotypes is the possibility that removal of key molecules in cells that are undergoing rapid rounds of cell division (such as occur during development to the immature B cell stage) can have non-specific toxic effects which may superficially appear quite specific. One way to address this possibility is to examine the effect of loss of function on the development of other

cell lineages, such as T cells. We thus examined whether T cell development in the thymus was affected by loss of G9a function.

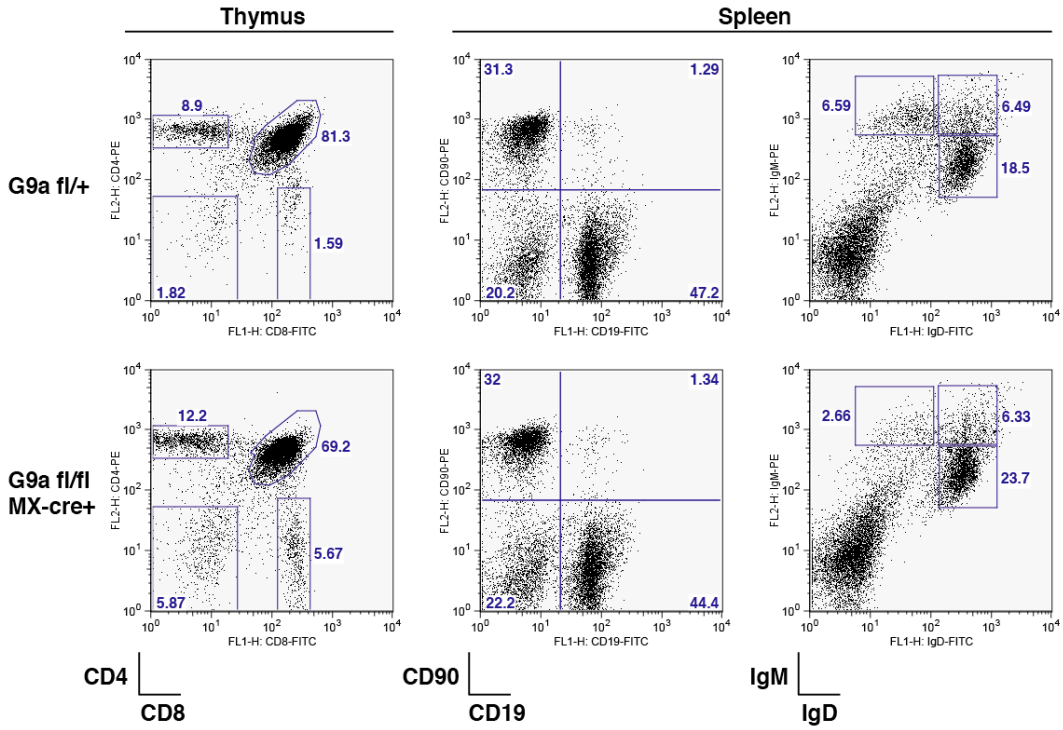

Figure 3.10 Loss of G9a function does not grossly perturb T cell development. Total thymocytes or splenocytes from the mice described in Figure 3.9 were stained as indicated and analysed by FACS.

The earliest stages of thymic α:β T cell development are characterized by the lack of expression of both the CD4 and CD8 co-receptors (double negative stage). Thymocytes that successfully undergo recombination of the TCRβ chain expand rapidly and mature to the CD4-CD8 double positive (DP) stage, during which TCRα recombination occurs. The small percentage of DP cells which survive positive and negative selection silence either CD4 or CD8 expression to become CD8 single positive (SP) or CD4 SP cells, respectively. Although deletion of G9a in thymocytes with MX-cre was very efficient

(Figure 3.8B), we saw no effect of G9a deficiency on T cell development by FACS (Figure 3.10, left panels). Consistent with this finding and the inefficient deletion observed in total splenocytes, the ratio of T cells ($CD90^+$) to B cells ($CD19^+$) in the spleen was also unchanged (Figure 3.10, middle panels). However, the number of immature B cells (IgM^+IgD^{lo}) in the periphery was reduced (Figure 3.10, right panels), as would be expected on the basis of decreased IgM^+ B cell production in the bone marrow. Thus while G9a function is essential for B cell development, it is dispensable for T cell development.

3.4 Analysis of G9a function in mature B cells

G9a and GLP were previously identified as components of a E2F6-containing transcriptional repressor complex linked to maintenance of cells in the G_0 (quiescent) stage of the cell cycle (Ogawa et al., 2002). This stage corresponds to the resting state of naive mature lymphocytes, thus raising the possibility that G9a and/or GLP might be involved in maintenance of cell quiescence. It was therefore of significant interest to us to investigate the effect of G9a deletion in peripheral lymphocytes. In order to begin to address this issue, we first examined expression of G9a and GLP in resting and activated peripheral B cells. Purified wild type splenic B cells were activated with either LPS alone or LPS and Interleukin 4 (IL-4), and G9a and GLP expression were measured by RT-PCR and Southern blotting as previously. Since triggering of naive B cells causes dramatic upregulation of "housekeeping" genes such as HPRT, expression of IgM germline transcripts was used as a loading control. As shown in Figure 3.11A, neither G9a nor GLP was significantly upregulated in response to B cell activation.

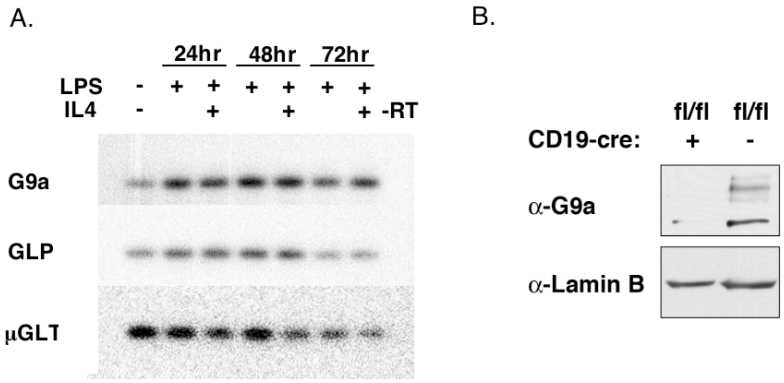

Figure 3.11 G9a/GLP expression during B cell activation and deletion of G9a with CD19-cre. A. Purified wild type splenic B cells were activated with the indicated stimuli and total RNA was collected on a timecourse. RT-PCR and Southern blotting were performed as in Figure 3.2. B. Splenic B cells were purified from $G9a^{fl/fl}$ mice with or without the CD19-cre knock-in allele, and lysates were immunoblotted with the indicated antibodies. The lower band in the αG9a blot represents a cross-reacting cytoplasmic protein (discussed in Chapter 4).

Since G9a deletion in bone marrow causes a block in B cell development, it was necessary to use a different driver strain in order to investigate possible functions of G9a in mature B cells. As mentioned previously, deletion of Ezh2 in hematopoietic stem cells with MX-cre also causes a severe block in B cell development at the pro-to-pre B cell transition. However, deletion starting at the pro-B cell stage with the CD19-cre driver allows production of peripheral B cells (Rickert et al., 1997; Su et al., 2003). We therefore crossed $G9a^{fl/fl}$ mice to the CD19-cre driver strain. G9a deletion in peripheral B cells, as measured by Western blotting with a αG9a polyclonal rabbit antiserum produced in the lab (described in Chapter 4), was complete (Figure 3.11B). As observed for Ezh2, conditional deletion of G9a with the CD19-cre driver strain did not significantly perturb B cell development in the bone marrow (Figure 3.12). Likewise, B cells were present in spleen and the ratio of IgM^+/IgD^+ cells was unaltered (data not shown).

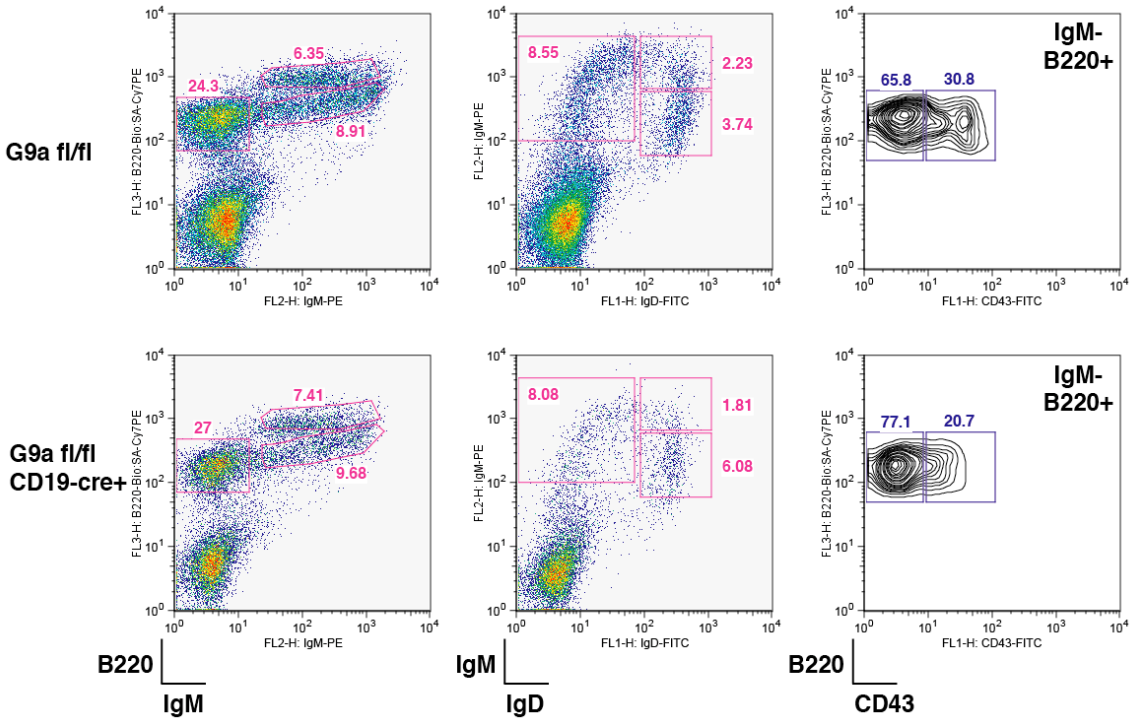

Figure 3.12 Unaltered B cell development in G9a$^{fl/fl}$ CD19-cre mice. Total bone marrow cells from mice of the indicated genotypes were stained as in Figure 3.9 and analysed by FACS.

In order to assess the contribution of G9a to maintenance of the G_0 state, we examined the response of wild type and G9a-deficient cells to mitogenic stimuli *in vitro*. Purified splenic B cells loaded with the cell division tracker dye carboxyfluorescein succinimidyl ester (CFSE) were stimulated with sub-optimal (1 μg/ml) or saturating (5 μg/ml) concentrations of αIgM antibody or LPS in the presence or absence of IL-4. Cell division is accompanied by linear dilution of the CFSE dye, allowing quantitation of cell division by FACS.

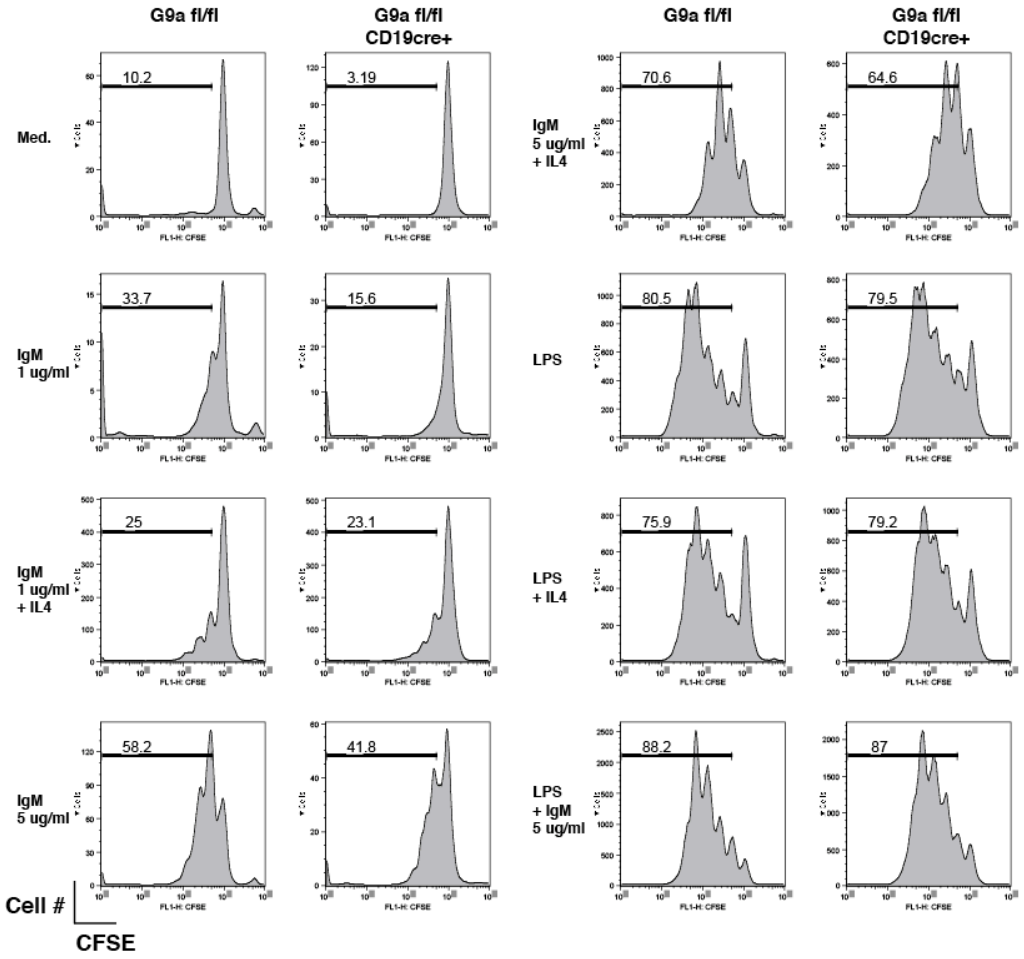

Figure 3.13 ***In vitro*** **activation profile of G9a-deficient B cells.** Purified splenic B cells from G9a$^{fl/fl}$ and G9a$^{fl/fl}$ CD19-cre mice were labeled with CFSE and activated with the indicated stimuli *in vitro*. Cells were analysed by FACS 72 hours after stimulation.

As shown in Figure 3.13, both wild type and G9a-deficient B cells proliferated robustly in response to triggering by αIgM+IL-4, LPS ± IL-4, or LPS+αIgM. The only difference in proliferation was seen under conditions of limited triggering with αIgM in the absence of IL-4, in which samples a modest decrease in division of the mutant cells could be measured. Importantly, no spontaneous division of G9a-deficient cells was observed in the absence of stimulation, arguing against a fundamental role for G9a activity in maintenance of quiescence.

Recently, it was reported that G9a physically interacts with the SET-domain containing protein Blimp-1, a known "master regulator" of B cell differentiation into plasma cells (Gyory et al., 2004; Shapiro-Shelef et al., 2003). Although Blimp has not been shown to possess histone methyltransferase activity itself, it is thought to promote plasma cell differentiation by inactivation of the B cell gene expression program through interaction with multiple families of transcriptional repressors (Lin et al., 2003; Piskurich et al., 2000; Shaffer et al., 2002; Yu et al., 2000). The interaction between G9a and Blimp was shown to be essential for a separate Blimp-dependent activity, namely repression of Interferon β (IFN-β) after induction *in vitro*. This raised the possibility that G9a might contribute to Blimp-mediated plasma cell development as well.

We addressed this possibility by measuring induction of CD138/Syndecan-1 expression, an established plasmablast marker (Shapiro-Shelef et al., 2003), in the B cell activation experiment described previously (Figure 3.13). Four days after stimulation, strong plasmablast differentiation was seen in both wild type and mutant cells stimulated with LPS only, and this was accompanied by extensive division of the cells as measured by CFSE dilution (Figure 3.14). Similar results were obtained for wild type cells stimulated with LPS+IL-4, but under this condition approximately half as many plasmablasts were seen in the mutant sample. Stimulation of B cells with both LPS and αIgM has been described to deliever an inhibitory signal which inhibits plasma cell differentiation, even in the presence of strong division. This inhibitory effect was maintained in mutant B cells (Figure 3.14), demonstrating that this effect in unlikely to be mediated by induction of specific H3K9 dimethylation.

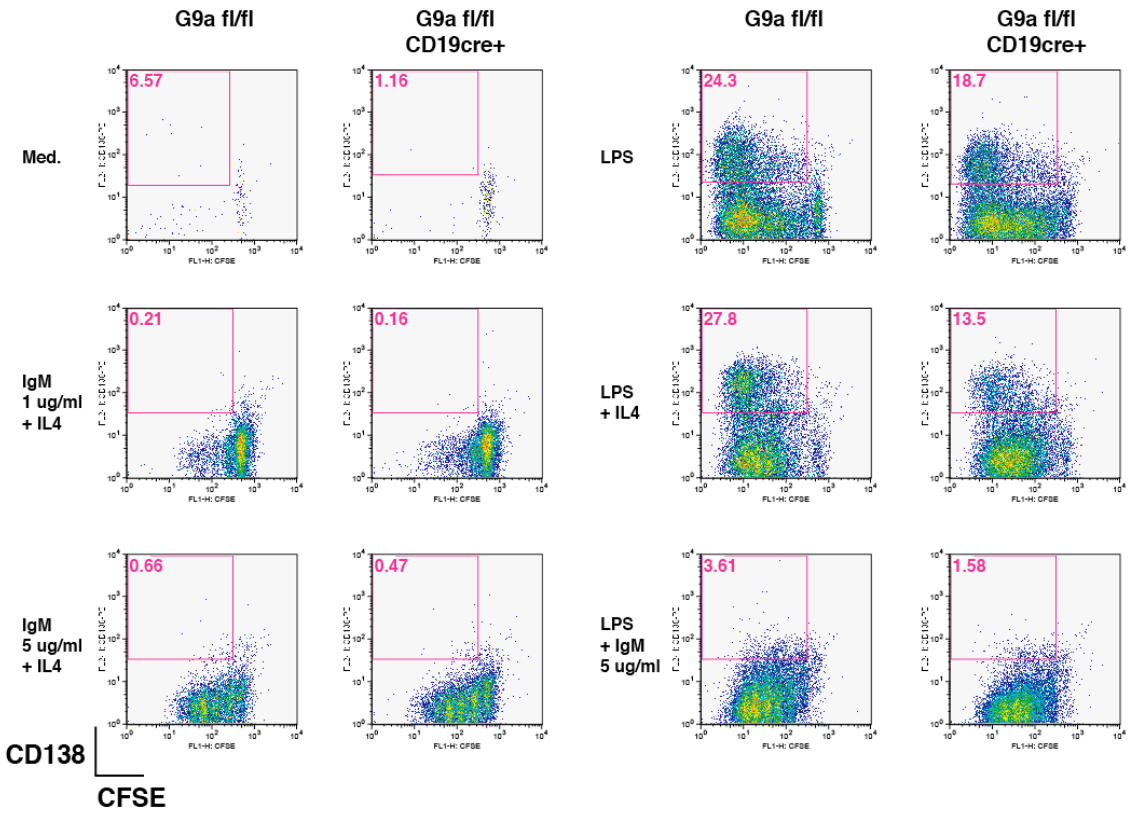

Figure 3.14 Effect of G9a deficiency on plasma cell differentiation *in vitro*. CFSE-labeled wild type or mutant B cells were activated *in vitro* with the indicated stimuli and analysed by FACS 4 days after stimulation for plasmablast differentiation, as measured by expression of CD138/Syndecan-1.

3.5 Summary

Overall, our studies of G9a conditional knockout mice indicate that G9a activity is critically required for B cell development, but is not required for T cell development. Within the B lineage, we find that G9a contributes to the pro-B to pre-B cell transition during development, but is largely dispensable for activation and proliferation in the periphery. We found no defect in upregulation of activation markers after stimulation *in vitro* (data not shown), and cells proliferated at wild type levels in response to most stimuli. In addition, plasmablast differentiation was completely unaffected in response to

LPS, ruling out an essential role for G9a in Blimp-mediated repression in differentiating B cells. Small differences were seen in the proliferative response of G9a-deficient B cells to purely BCR-driven activation in the absence of IL-4, as well as in plasma cell differentiation in response to LPS+IL-4. However, since the CD19-cre driver used in these experiments is a knock-in into the CD19 locus, G9a-deficient cells are also heterozygously deficient for the CD19 co-receptor. Complete loss of CD19 function is known to cause defective T-cell dependent immune responses *in vivo* (Rickert et al., 1995), and thus quantitatively different results on a heterozygous background must be interpreted with caution.

Chapter 4: Analysis of G9a methylation

4.1 G9a is lysine methylated *in vitro* and *in vivo*

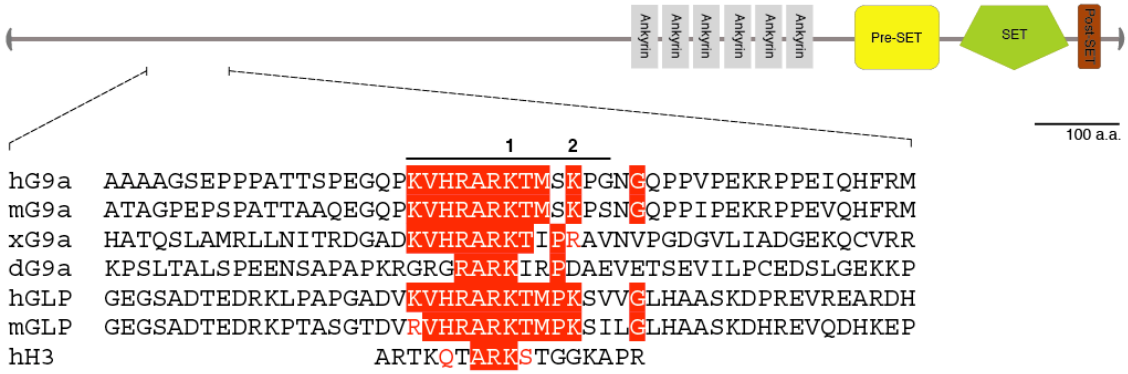

Figure 4.1 G9a and GLP carry a conserved H3K9-like motif in their N terimini. Protein sequences of G9a and GLP homologs from several species were aligned using ClustalW software. Identical residues in at least four of six sequences are shown in red, and the amino terminus of H3 is shown for comparison. The sequence listed corresponds to residues 140-190 of human G9a. The positions marked "1" and "2" above the sequences denote the two lysine residues conserved between human and mouse G9a/GLP homologs. Abbreviations: h, human; m, mouse; x, *Xenopus*; d, *Drosophila*.

While performing sequence alignments between the amino-termini of murine G9a and GLP, we identified a stretch of eight amino acids which showed perfect identity between the two (Figure 4.1). Extending the alignment to include human G9a and GLP, as well as G9a homologs in *Xenopus* and *Drosophila*, we noted a high degree of conservation within this motif, which contrasts with the relatively poor conservation of the surrounding sequence. Most intriguing was the degree to which the core conserved motif resembled the canonical H3K9 target sequence of G9a (Figure 4.1). This observation led us to ask whether this H3-like sequence in G9a might be a target for lysine methylation.

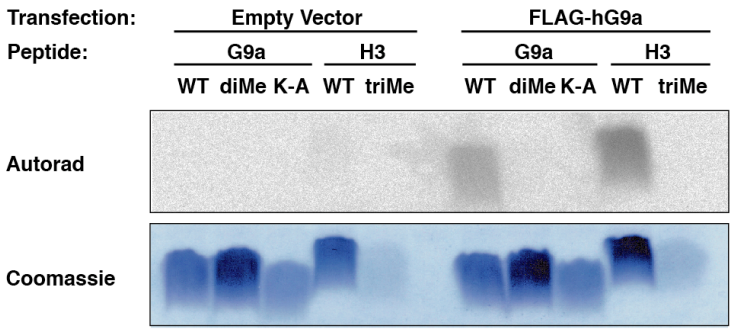

Figure 4.2 G9a K194 can be methylated *in vitro*. αFLAG immunoprecipitates from 293 cells transiently transfected with the indicated constructs were used in an *in vitro* HMTase assay against the peptide substrates listed. WT, unmodified peptide; diMe, dimethyl K194 peptide; K-A, K165A mutant peptide; triMe, trimethyl H3K9 peptide.

To address this possibility, we first assayed the ability of a peptide containing the putative methylation site of human G9a (K165) to act as a substrate in an *in vitro* methylation reaction. As a source of methyltransferase activity, we initially chose to use human G9a itself, since this enzyme has been shown to possess high intrinsic HMTase activity towards H3K9 (Tachibana et al., 2001). Methyltransferase assays using immunoprecipitated FLAG-tagged hG9a demonstrated robust methylation of peptides containing either the wild type H3K9 sequence or the wild type G9a K165 sequence (Figure 4.2). Since the K165 peptide contains an additional conserved lysine, K169, we also attempted to methylate a peptide containing the lysine of interest mutated to an alanine (K165A). This mutant peptide failed to be methylated by hG9a *in vitro*, demonstrating the specificity of the methylation reaction for the single lysine which most closely resembles H3K9 (Figure 4.1). In addition, peptides carrying chemically dimethylated K165 were completely refractive to further methylation (Figure 4.2), a

finding consistent with previous reports demonstrating that G9a is specifically a di-, but not tri-methylating enzyme (Peters et al., 2003).

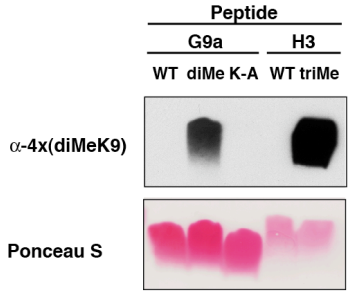

Figure 4.3 Methylated G9a K194 peptide is recognized by a "multi-methyl" lysine-specific antibody. The indicated peptides were probed by Western blot with an antibody (α4xdiMeK9) which recognized methyllysine in multiple sequence contexts (see text). WT, unmodified peptide; diMe, dimethyl K194 peptide; K-A, K165A mutant peptide; triMe, trimethyl H3K9 peptide.

Having established that the H3K9-like K165 site in G9a can be methylated *in vitro*, we next sought to determine whether this site is actually methylated *in vivo*. Studies of histone methylation have been greatly aided by the production of antisera capable of specifically recognizing modified residues. Of particular interest is an antiserum that was raised against a branched-peptide antigen bearing the dimethyl H3K9 sequence (α4xdiMeK9); this antibody has been shown to recognize various degrees of methylation of H3K9, as well as lysine methylation of other sites in various histones, and has therefore been termed a "multi-methyl" lysine specific antibody (Perez-Burgos et al., 2004). We investigated whether this antibody might also react with methylated G9a K165 peptide. As shown in Figure 4.3, the α4xdiMeK9 antibody was able to recognize the dimethylated but not unmethylated or alanine mutant K165 peptide by Western blot,

demonstrating specific recognition of the methyl-modified form of G9a. We therefore further used this antibody to determine whether G9a exists in a methylated state *in vivo*.

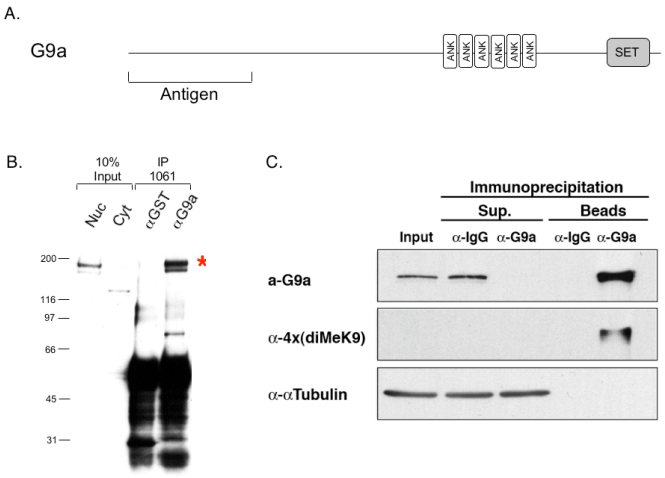

Figure 4.4 G9a exists in a methylated state *in vivo*. A. Schematic representation of the region used to raise αG9a antibody. B. Western blot of immunoprecipitations from nuclear or cytoplasmic 293 extracts using αG9a antibody. Red asterisk marks the position of G9a; the strong signal below ~50 kD corresponds to cross reaction of the secondary antibody wih that used for IP. C. Immunoprecipitation of endogenous hG9a from 293 cells, followed by Western blot with αG9a or α4x(diMeK9) antibody as indicated.

In order to purify endogenous G9a, we raised an antiserum against an amino-terminal fragment of murine G9a. The affinity-purified antibody recognizes both human and mouse G9a in nuclear extracts, and immunoprecipitates H3K9- and G9a K165-methylating activity (Figure 4.4A and B, and data not shown). Endogenous hG9a immunoprecipitated from HEK293 nuclear extracts using αG9a antibody was probed by Western blot with the α4xdiMeK9 antibody (Figure 4.4C). Reactivity with the multi-methyl lysine antibody was readily observed, indicating that G9a is in fact methylated *in vivo*.

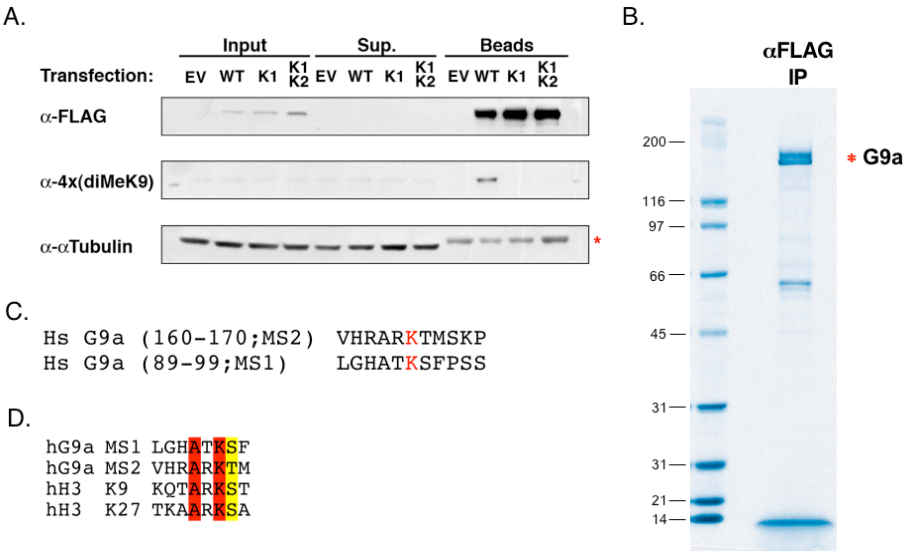

Figure 4.5 Identification of G9a methylation sites. A. Wild type or point mutant G9a was immunoprecipiated from transiently transfected 293 cells and analysed by Western blot with the indicated antibodies. Asterisk indicates cross-reaction with light chain of the immunoprecipitating antibody. B. Immunoprecipitated FLAG-hG9a was separated by SDS-PAGE followed by staining with Coomassie Blue. C. Sites of methylation identified by mass spectrometry (indicated in red); both sites were purely dimethylated. D. Alignment of MS1 and MS2 with H3K9 and H3K27. Identical residues are shown in red, and conservative substitutions in yellow. Abbreviations: EV, empty vector; WT, wild type FLAG-hG9a; K1, FLAG-hG9a K165A; K1K2, FLAG-hG9a K94A K165A.

To determine the site of methylation recognized by the α4xdiMeK9 antibody, we began by constructing point mutations of human G9a, mutating either one or both conserved lysines in the originally-identified motif to alanine (Figure 1A, K165A or K165,169A). To our surprise, mutation of a single lysine to alanine (K165A, corresponding to the site of methylation *in vitro*) completely abolished recognition of FLAG-hG9a by the α4xdiMeK9 antibody (Figure 4.5A).

In order to directly confirm the finding of G9a methylation on K165, we investigated methylation of this site by mass- and tandem-mass spectrometry (MS and MS-MS). Trypsin digests of immunoprecipitated FLAG-hG9a (Figure 4.5B) produced an MS peak

consistent with the mass of a peptide (KTMSKPGNGQPPVPEK) bearing an oxidized methionine and two methyl groups (Figure 4.5C). Further fragmentation of this peptide and analysis by MS-MS yielded a spectrum consistent with the sequence of the predicted peptide, including dimethylation of K165 (data not shown). Intriguingly, no peptides bearing mono- or trimethylated lysines were observed, and unmethylated peptides were similarly not detected, suggesting that the vast majority of the expressed G9a exists in the dimethylated state. In addition to K165, we also identified an additional site of lysine methylation, K94 (Figure 4.5C). Like K165, K94 was also dimethylated, and also existed in a sequence context which closely resembled H3K9 (Figure 4.5D). For clarity, the upstream K94 site is hereafter referred to as MS1 (**M**ethylation **S**ite **1**) and the downstream K165 site as MS2 (**M**ethylation **S**ite **2**).

In order to further confirm our finding of endogenous G9a methylation, we raised an antiserum against the dimethylated G9a MS2 peptide. Dot blot analysis of the affinity-purified antibody demonstrated strict requirement of methylation for recognition of the immunizing peptide (Figure 4.6A). Surprisingly, peptides bearing dimethylation on MS2 as well as phosphorylation of the adjacent threonine were not recognized by this serum, suggesting that recognition is specific for the methylated, non-phosphorylated version of this site (data not shown). We further found that the immune, but not the preimmune serum specifically recognized bands of the correct size for G9a in 293 nuclear extracts (Figure 4.6B, indicated by red asterisk). Despite the sequence similarity between MS1 and MS2, it is clear that both the α4xdiMeK9 and αdiMeMS2 antisera specifically recognize MS2 methylation, as MS1 point mutation had no effect on recognition by antibody (Figure 4.6C). Since the MS2 site bears the strongest resemblance to H3K9 and

is uniquely recognized by the α4xdiMeK9 antibody, we chose to further investigate the biochemical consequences of methylation on this lysine.

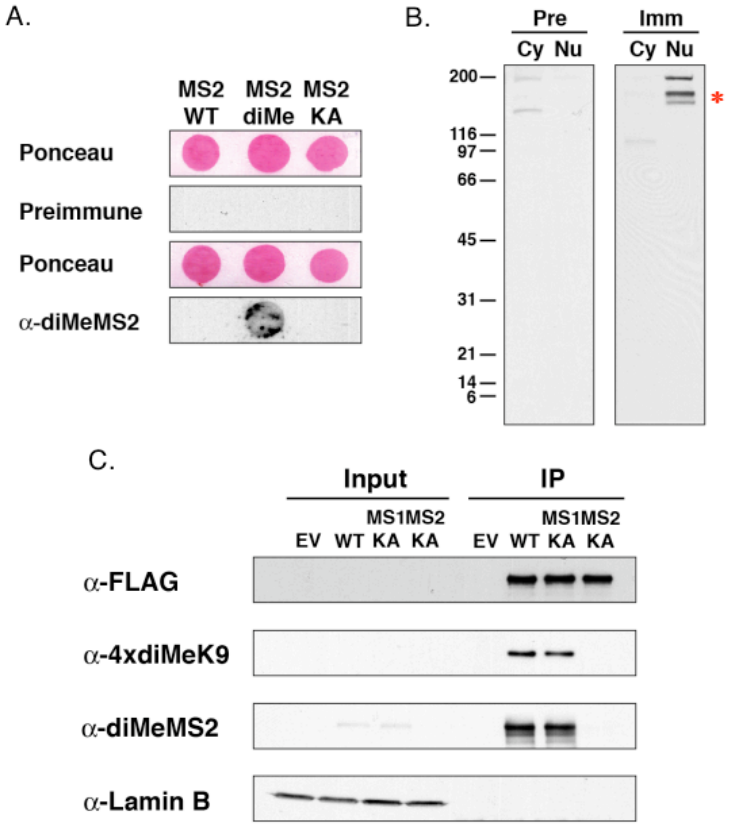

Figure 4.6 Two different methyl-specific antibodies recognize methylation of G9a MS2. A. Dot blot analysis using the indicated peptides and either the affinity purified αG9a diMeMS2 antibody or pre-immune serum from the same rabbit. B. Western blot analysis of cytoplasmic (Cy) or nuclear (Nu) 293 cells extracts, using the sera as in (A). Red asterisk marks the position of endogenous G9a. C. Immunoprecipitates of wild type or point mutant FLAG-hG9a probed with the indicated antibodies. Signal from input samples probed with αFLAG antibody was visible on longer exposure (data not shown).

4.2 G9a methylation controls interaction with HP1

Groundbreaking studies previously demonstrated that methylation of H3K9 is sufficient to create a specific binding site for chromodomain-containing proteins of the Heterochromatin Protein 1 (HP1) family (Lachner et al., 2001). Intriguingly, G9a has

been shown to exist in a complex containing the predominantly euchromatic HP1 isoform, HP1γ, and to interact directly with HP1γ *in vitro* (Ogawa et al., 2002; Roopra et al., 2004). These observations, together with the extensive sequence similarity between MS2 and H3K9, as well as the finding that G9a exists in a methylated state *in vivo*, led us to speculate that methylation of G9a on K165 could create a binding site for HP1γ. To address this possibility, we asked whether interaction of G9a with HP1γ depended on the presence of a methylatable lysine at the MS2 site. Wild type or MS2 KA point mutant hG9a were expressed in 293 cells, and assayed for their ability to associate with HP1γ. As shown in Figure 2D, wild type G9a was able to co-immunoprecipitate HP1γ, whereas the MS2 K165A mutation completely eliminated this interaction, despite the presence of an unmutated MS1 site. By contrast, the previously reported interaction between G9a and GLP was not affected by the MS2 KA mutation (data not shown).

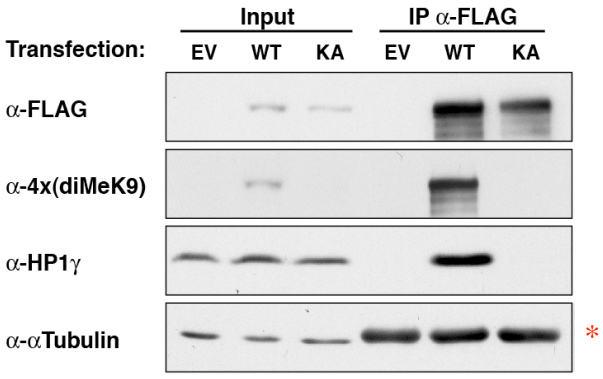

Figure 4.7 Interaction of G9a with HP1γ requires an intact MS2 methylation site. Immunoprecipitates of transiently transfected wild type (WT) or K165A (KA) FLAG-hG9a were analysed by Western blot for co-immunoprecipitation of endogenous HP1γ. Asterisk indicates cross-reaction with the heavy chain of the immunoprecipitating antibody.

These results suggest, but do not directly prove that interaction of HP1γ with G9a depends on methylation of MS2. We therefore examined whether the presence of methyl groups *per se* was required for HP1 binding. MS2 peptide matrices containing either unmodified lysine, dimethylated lysine or a lysine to alanine mutation were incubated with unmanipulated nuclear extracts in order to determine whether HP1γ could directly interact with the methylated peptide. We found that recovery of HP1γ as well as HP1α on the beads was significantly enhanced by pre-existing methylation of the MS2 peptide, and was completely blocked by mutation of the methylated lysine to alanine (Figure 4.8). In contrast, the nuclear protein Lamin B was not recovered with any peptide.

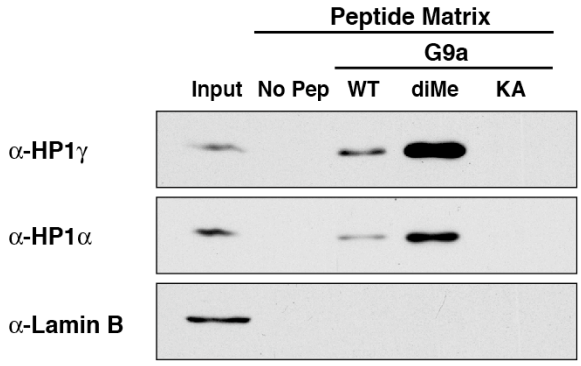

Figure 4.8 Methylation of MS2 strongly promotes binding of HP1 isoforms. The indicated peptides were covalently coupled to beads and incubated in unmanipulated 293 extracts. After thorough washing, associated proteins were examined by Western blot. WT, unmodified peptide; diMe, dimethyl MS2; KA, MS2 K165A peptide.

The interaction of HP1 isoforms with unmethylated peptide, while substantially less efficient than in the presence of methylation, was still detectable. This could represent either low-affinity interaction between HP1 and unmethylated lysine, or *de novo* methylation of the MS2 peptide by methyltransferases present in the extract. These possibilities can be distinguished through the use of *in vitro* translated HP1 protein, as contaminating HMTases should be less abundant or absent in the reticulocyte lysates

used for translation. We therefore incubated wild type, dimethylated or point mutant MS2 peptides with *in vitro*-translated HP1 proteins, followed by SDS-PAGE and autoradiography (Figure 4.9). Only the methylated MS2 peptide was able to interact with HP1 under these conditions, demonstrating that the interaction of methylated G9a with HP1 requires lysine methylation.

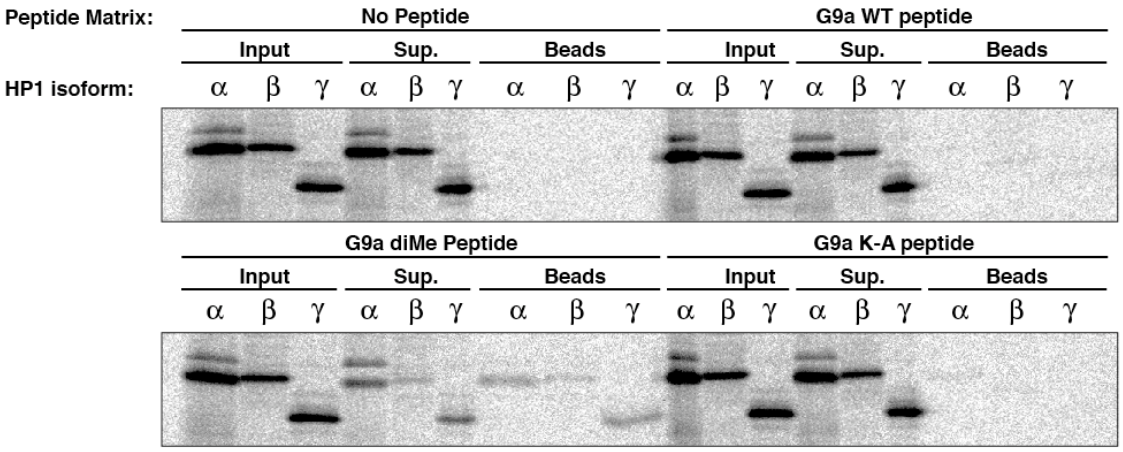

Figure 4.9 Methylation of G9a MS2 is required for HP1 binding. Beads conjugated with the indicated peptides were incubated with *in vitro*-translated HP1 isoforms, and binding assessed by SDS-PAGE and autoradiography.

4.3 Studies of G9a methylation *in vivo*

Further investigation of the mechanism and possible consequences of G9a methylation required a cell system lacking endogenous G9a expression. To create such a system, we utilized the conditional G9a mouse mutant described earlier (Chapter 3). Since G9a-deficient Mouse Embryonic Fibroblasts (MEFs) fail to proliferate in culture (Tachibana et al., 2002), we chose to use a multistep strategy to derive deficient MEFs. First, immortalized $G9a^{fl/fl}$ ("floxed") cell lines would be derived by infection of primary MEFs

with a retrovirus expressing SV40 large-T antigen. These would be subsequently infected in bulk with a Cre-expressing adenovirus to produce G9a deficiency ($G9a^{\Delta/\Delta}$), and subcloned by limiting dilution to produce clonal populations. Finally, the subcloned cells would be infected with retroviruses carrying either an empty vector (KO), wild type FLAG-hG9a (WT), MS2 K165A FLAG-hG9a (KA) or FLAG-hG9a carrying a single amino acid point mutation in a residue known to be critical for enzymatic activity in other SET-domain contain HMTases (H1093K, abbreviated HK). The overall strategy is summarized in Figure 4.10.

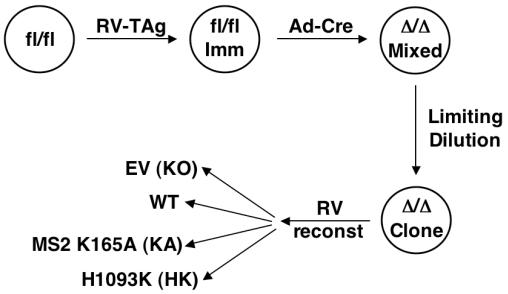

Figure 4.10 Scheme for creation of G9a-deficient reconstituted MEFs. Abbreviations: RV, retrovirus; TAg, SV40 Large T antigen.

To examine the efficiency of adenoviral Cre expression, we first performed a timecourse of Adeno-Cre deletion in primary MEFs. Two independent $G9a^{fl/+}$ or $G9a^{fl/fl}$ embryos were used to establish primary cultures, which were infected at high multiplicity of infection with Adeno-Cre. Samples were taken at 24 hour interval, and deletion was monitored by Western blot with αG9a antibody. As shown in Figure 4.11, complete deletion at the protein level was seen within 24 hours. Interestingly, a band cross-reacting with the αG9a antibody was seen to increase in expression level in parallel with G9a deletion.

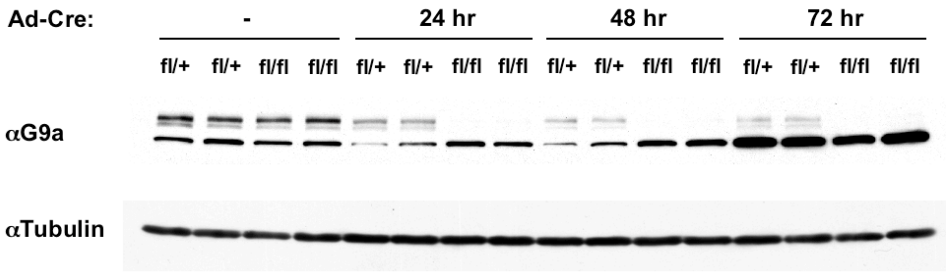

Figure 4.11 Deletion of G9a *in vitro* with Adeno-Cre. Independent primary MEF littermate cultures of the indicated genotypes were infected with Cre-expressing adenovirus and followed by Western blot to monitor deletion.

The presence of a cross-reacting band which increased upon G9a deletion raised the possibility that G9a deletion might actually produce an in-frame splice product. To understand the properties of this cross-reacting band further, we immunoprecipitated G9a from floxed cells using our αG9a antibody either with or without deletion using Adeno-cre treatment, and then performed Western blot with the same antibody (Figure 4.12A). Surprisingly, the cross-reacting band was immunoprecipitated even in deleted cells, while co-IP of endogenous GLP (examined using our own αGLP antibody) was abolished. We sought to further understand the nature of the cross-reacting protein by performing nuclear fractionation (Figure 4.12B). This demonstrated that the cross-reacting protein is purely cytoplasmic, and thus most likely not significant for our experiments relating to nuclear functions of G9a and G9a methylation, although this observation may be very significant in other contexts (discussed below).

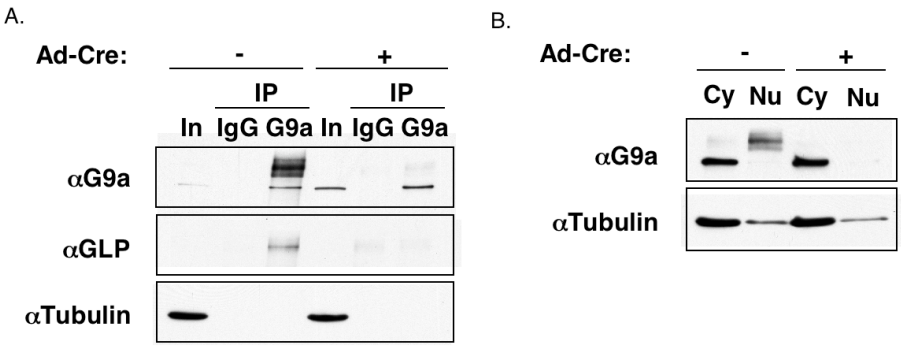

Figure 4.12 G9a deletion with Adeno-cre disrupts interaction with GLP and does not produce a nuclear truncation product. Primary MEF cultures were mock-infected or infected with Adeno-cre. A. Whole cell extracts were immunoprecipitated with αG9a antibody and examined for co-IP of GLP by Western blot. B. Nuclear and cytoplasmic fractions were prepared to examine the distribution of the cross-reacting band seen with αG9a antibody.

Deleted immortalized cells were subcloned to produce homogeneous populations, and then reconstituted with retrovirus as indicated in Figure 4.10. Genotyping of the retrovirally infected cell lines by PCR demonstrated that all lines remained completely deleted for G9a at the genomic DNA level after extensive passage in culture (Figure 4.11A). Western blot analysis of nuclear lysates from the reconstituted cell lines demonstrated the expected overexpression of G9a in the retrovirally transduced cell lines (Figure 4.11B). Complete loss of G9a expression has been shown to result in a severe decrease in euchromatic H3K9 dimethylation, with partial relocalization of the dimethyl H3K9 mark to pericentric heterochromatin (Peters et al., 2003; Rice et al., 2003). Consistent with these observations, both fully mutant MEFs and those reconstituted with catalytically inactive G9a showed a dramatic reduction in the overall level of H3K9 dimethylation, as detected by immunoblotting with a modification-specific antibody (Figure 3B). In contrast, the overall levels of HP1γ remained unchanged (Figure 3B).

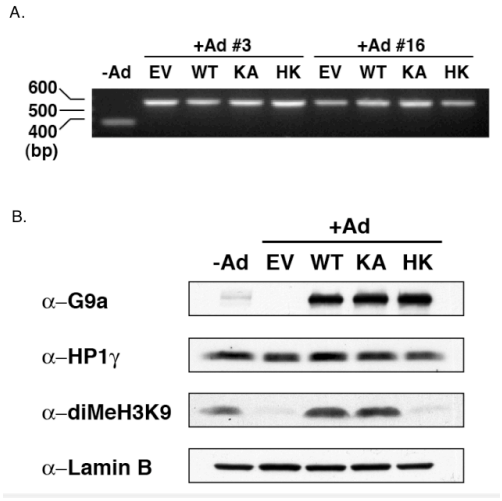

Figure 4.13 Analysis of reconstitution in G9a$^{\Delta/\Delta}$ MEFs. A. G9a genotyping PCR on genomic DNA from undeleted G9a$^{fl/fl}$ MEFs (-Ad) or two independent subclones of G9a$^{\Delta/\Delta}$ MEFs (+Ad) reconstituted with empty retrovirus (EV), wild type (WT), MS2 KA (KA) or catalytically inactive (HK) FLAG-hG9a. Expected product sizes: floxed, 371 bp; deleted, 574 bp. B. Reconstituted lines from clone #3 were analysed by Western blot with the indicated antibodies.

In order to determine whether the euchromatic distribution of HP1γ was altered in the presence of the MS2 point mutation, we performed indirect immunofluorescence (IF) using either HP1γ, G9a or diMeH3K9-specific antibodies. G9a-deficient cells showed the expected loss of euchromatic diMeH3K9, with weak accumulation of dimethyl H3K9 at pericentric heterochromatin (Figure 4.12 and data not shown). G9a deficiency also induced increased localisation of HP1γ to pericentric heterochromatin, which, given the unchanged overall levels of HP1γ, implies impaired euchromatic localisation. Reconstitution of deficient cells with wild type G9a completely rescued the effects on both diMeH3K9 and HP1γ distribution (Figure 4.12).

Cells lacking endogenous G9a and expressing only the non-methylatable point mutant showed an HP1γ localisation pattern that was grossly identical to cells expressing wild type G9a, indicating that methylation of G9a is not required for bulk euchromatic

localisation of HP1γ. Likewise, both G9a and diMeH3K9 were distributed as in wild type cells, indicating that MS2 methylation is dispensable for G9a localization and histone methylation activity. Interestingly, cells reconstituted with catalytically inactive G9a continued to demonstrate mislocalisation of HP1γ, suggesting that, as with targeting of HP1α to pericentric heterochromatin (Lachner et al., 2001), proper euchromatic targeting of HP1γ depends mainly on histone methylation *per se* (Figure 3C).

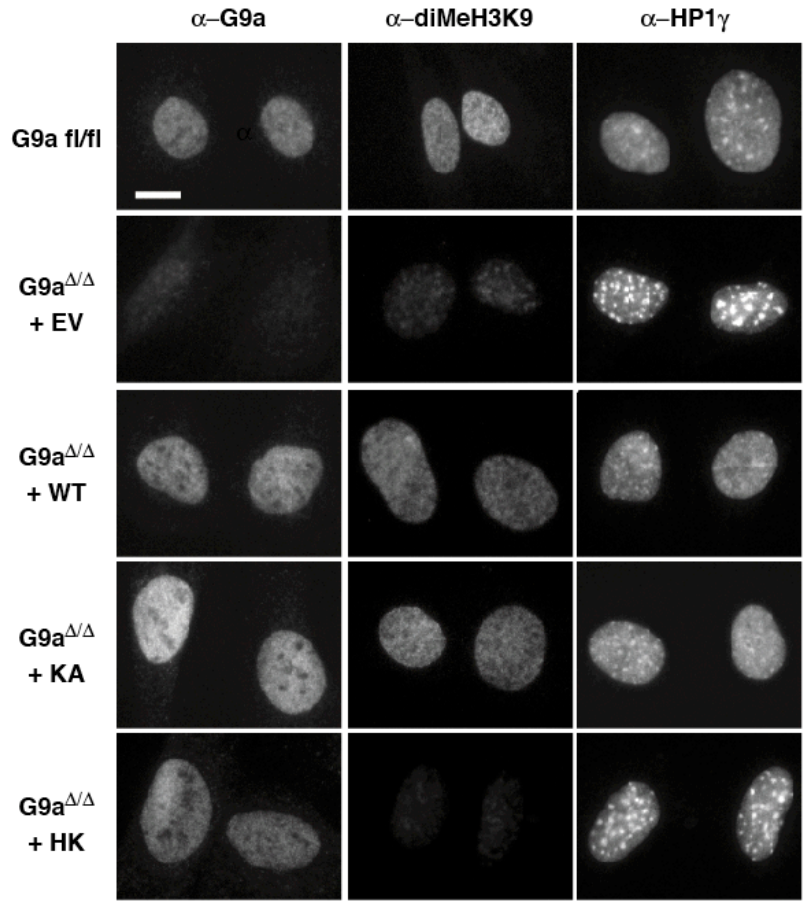

Figure 4.14 G9a methylation on MS2 is not necessary for bulk G9a, HP1γ or diMeH3K9 localization *in vivo*. MEF lines of the indicated genotypes were examined by indirect immunofluorescence using the listed antibodies and couterstained with DAPI. Only the specific stain is shown (DAPI channel excluded).

The creation of cell lines which do not express endogenous G9a further allowed us to address the question of which HMTase is responsible for G9a methylation on MS2. Initial experiments indicated that while G9a can robustly methylate G9a peptides *in vitro*, the related H3K9 methyltransferase Suv39H1 cannot (data not shown). This finding is consistent with the observation that G9a catalytic activity is ~20-fold higher than that of Suv39H1, most likely due to a histidine to arginine substitution within the G9a SET domain which hyperactivates the enzyme (Tachibana et al., 2001). However, the two non-histone targets of lysine methylation reported to date, TAF10 (Kouskouti et al., 2004) and p53 (Chuikov et al., 2004), were both shown to be targets of the SET9 methyltransferase, which has not yet been demonstrated to possess any *in vivo* HMTase function, thus raising the possibility that the enzymes targeting histones and non-histone proteins may be mutually exclusive.

We first chose to test whether G9a might be responsible for its own methylation. FLAG-hG9a was immunoprecipitated from reconstituted MEF clones expressing either G9a WT, MS2 K165A or the H1093K catalytic mutant, followed by immunoblotting with the α4xdiMeK9 antibody, which uniquely recognizes MS2 methylation. Whereas wild type G9a was readily recognized by this antibody, both the MS2 and catalytic point mutants were not (Figure 4.13). The lack of G9a methylation in the presence of a catalytic point mutation demonstrates that G9a is required for its own methylation. Whether this methylation occurs in *cis* or in *trans* remains an outstanding question.

Figure 4.15 Methylation on MS2 requires G9a catalytic activity. Wild type (WT), MS2 KA (KA) or catalytically inactive (HK) G9a was immunoprecipiated from reconstituted G9a-deficient MEFs and examined by Westwen blot with the indicated antibodies.

The presence of a methylatable target site within the G9a methyltransferase clearly raised the possibility that this site could behave in a manner similar to that seen in kinases, where pseudo-substrate sites located in *cis* are able to modulate kinase activity (Kemp et al., 1994). Since G9a is known to homodimerize (Tachibana et al., 2005), we therefore took advantage of the reconstituted MEFs to measure the catalytic activity of MS2 KA mutant G9a in a background devoid of endogenous G9a methyltransferase activity. FLAG-tagged wild type or MS2 KA G9a were immunoprecipitated from MEFs and assayed *in vitro* for methyltransferase (MTase) activity against either H3 or G9a MS2 peptides. Equal amounts of G9a were recovered from all cell lines, and MTase activity of wild type and MS2 KA G9a were equal on both substrates (Supp. Figure 4). Since MS2 KA G9a retains its ability to heterodimerize with endogenous wild type GLP, it was possible that the activity measured with the methylation site point mutation was due to co-IP of GLP. We therefore also assayed the MTase activity of the G9a H1093K catalytic dead mutant, which is unable to reconstitute G9a activity *in vivo* but still interacts with GLP (Figure 3B, C, and data not shown). This mutant still exhibited an

~65% reduction in catalytic activity *in vitro*, demonstrating that the activity measured for the MS2 point mutation is not simply due to interaction with GLP (Supp. Figure 4).

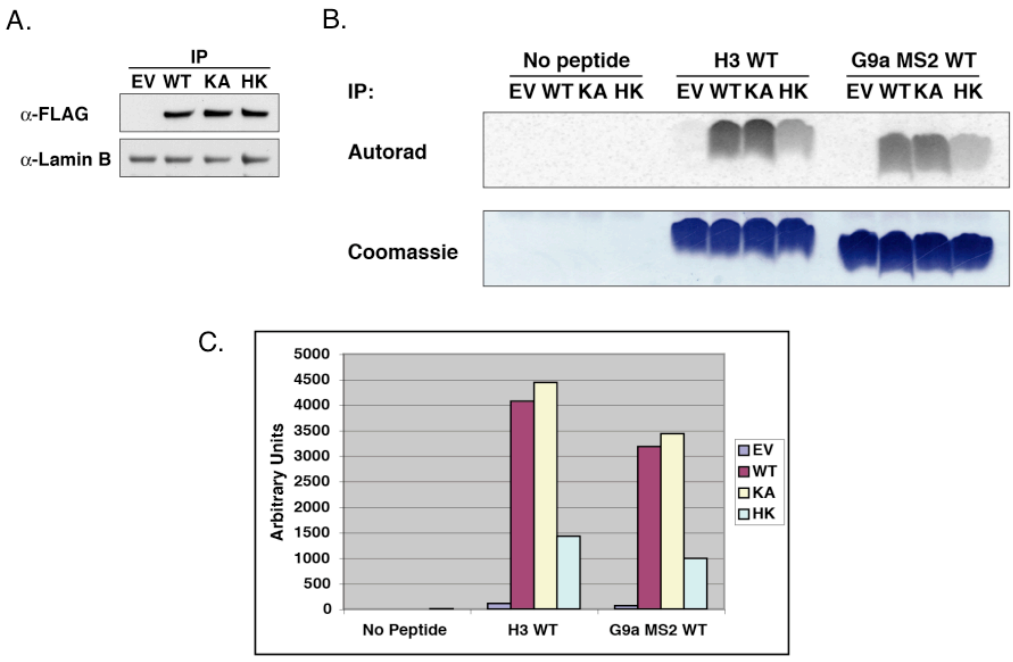

Figure 4.16 G9a MS2 methylation is not required for catalytic activity. Wild type (WT), MS2 KA (KA) or catalytically inactive (HK) FLAG-hG9a were immunoprecipitated from reconstituted MEF cell lines. A. Western blot analysis of G9a recovery from each cell line. B. HMTase assay on unmodified H3(1-20) or G9a MS2 peptide using immunoprecipitated protein from (A). C. Quantitation of the results shown in (B).

4.4 Potential control of HP1 binding by phosphorylation

A major issue in the study of lysine methylation has been the possibility of reversal of this mark. Significant progress has been made recently through identification of an enzymatic activity capable of active demethylation of lysines 4 and 9 in H3 (Lee et al., 2005b; Metzger et al., 2005; Shi et al., 2004). It has also been proposed that ejection of HP1 bound to methyllysine could be achieved through phosphorylation of serine residues found adjacent to H3K9 and H3K27 (Fischle et al., 2003a). Such "methyl-phos

switches" have been proposed to comprise a reversible mechanism for localized regulation of HP1 binding to histones. The presence of phosphorylatable residues (serine and threonine, respectively) at the positions corresponding to H3 serine 10 in the MS1 and MS2 sites (Figure 4.15A) raised the possibility that phosphorylation at these sites could have a similar regulatory effect on HP1 binding to G9a.

To investigate this possibility, we first sought to determine whether the threonine at position 166 in MS2 could be phosphorylated by the dominant mitotic H3S10 kinase, Aurora B (Hsu et al., 2000). As a source of endogenous Aurora B, we used extracts prepared from unfertilized *Xenopus* oocytes, which are naturally arrested at meiotic metaphase II and contain high Aurora B kinase activity (Bolton et al., 2002; Murray, 1991). The Aurora B-containing Chromosomal Passenger Complex (CPC) was immunoprecipitated from *Xenopus* egg extracts using an antibody directed against the CPC member Incenp (Sampath et al., 2004), and the immunoprecipitated material was used in an *in vitro* kinase assay on various peptide substrates. As shown in Figure 4.15B, the Aurora B complex readily phosphorylated a peptide containing the H3K9 sequence. This phosphorylation was completely blocked by mutation of serine 10 to alanine, consistent with the known specificity of Aurora B (Bolton et al., 2002; Crosio et al., 2002; Giet and Glover, 2001). A G9a peptide containing the MS2 methylation site was also robustly phosphorylated, and this phosphorylation was unaffected by the presence of dimethylation on MS2. Since the MS2 peptide contains multiple phosphorylatable residues, we also produced a peptide carrying both MS2 dimethylation as well as phosphorylation on T166; this peptide was completely resistant to further

phosphorylation, demonstrating that phosphorylation by Aurora B is specific for T166, the single residue in the MS2 peptide which aligns with H3 serine 10.

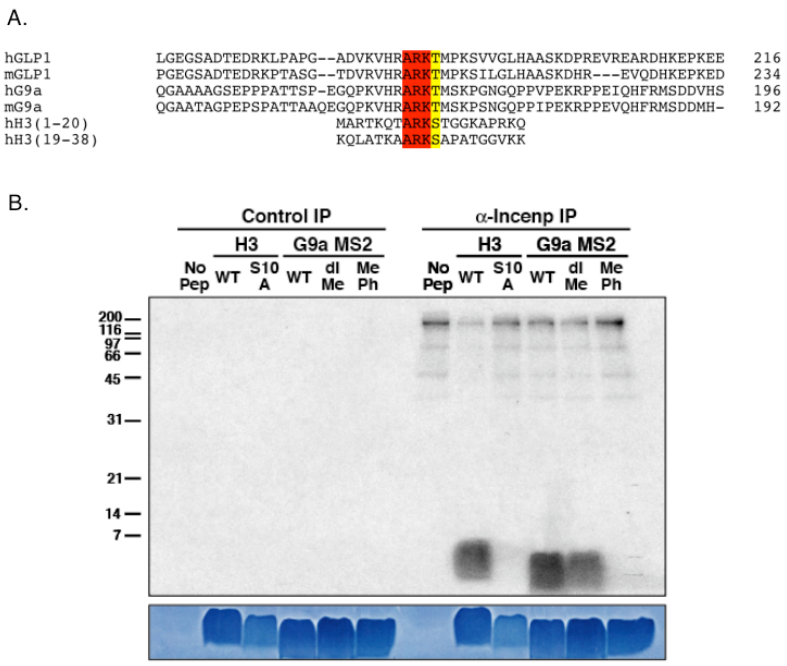

Figure 4.17 G9a MS2 peptide can be phosphorylated by Aurora B *in vitro*. A. Alignment G9a/GLP MS2 sites with H3K9 and H3K27 demonstrates conservation of a phosphorylatable residue (shown in yellow) adjacent to the methylation site (shown in red). B. The Aurora B-containing *Xenopus* chromosomal passenger complex (CPC) was immunoprecipitated and used in an *in vitro* kinase assay with the indicated peptides as substrates. Coomassie staining of the same gel is shown below. The labeled high molecular weight species correspond to known Aurora B targets within the CPC. WT, unmodified peptide; S10A, H3 Ser10Ala mutant peptide; diMe, dimethyl MS2 peptide; MePh, dimethyl MS2 T166 phosphorylated double-modified peptide.

We next investigated whether phosphorylation of T166 could modulate HP1 binding to G9a *in vitro*. Bead matrices coupled with either no peptide, wild type, dimethyl or dimethyl-phospho G9a MS2 peptide were incubated with *in vitro*-translated HP1γ as previously. As expected, neither the empty beads nor the unmodified peptides bound significant amounts of HP1γ, while the dimethyl MS2 beads bound HP1γ strongly

(Figure 4.16A). Significantly, this binding was completely eliminated by the presence of phosphorylation on the adjacent threonine (T166). The lack of HP1γ binding to the dimethyl-phospho beads was not due to reduced coupling efficiency of the peptide, since binding could be quantitatively restored by pre-treatment of the phosphorylated peptide matrix with Protein Phosphatase 1 (PP1; Figure 4.16A, quantitation shown in 4.16B). These results demonstrate that, as with H3, phosphorylation adjacent to a site of methylation in G9a has the potential to reverse the biochemical readout of this mark. Thus the "tail" of G9a mimics the biochemical properties of the H3 tail itself in the following ways: the G9a amino terminus is multiply methylated; this methylation is at least in part carried out by an H3K9-specific methyltransferase; methylation creates a binding site for HP1; a G9a tail peptide can be phosphorylated by a mitotic kinase; and finally, methyl-dependent HP1 binding is blocked by adjacent phosphorylation.

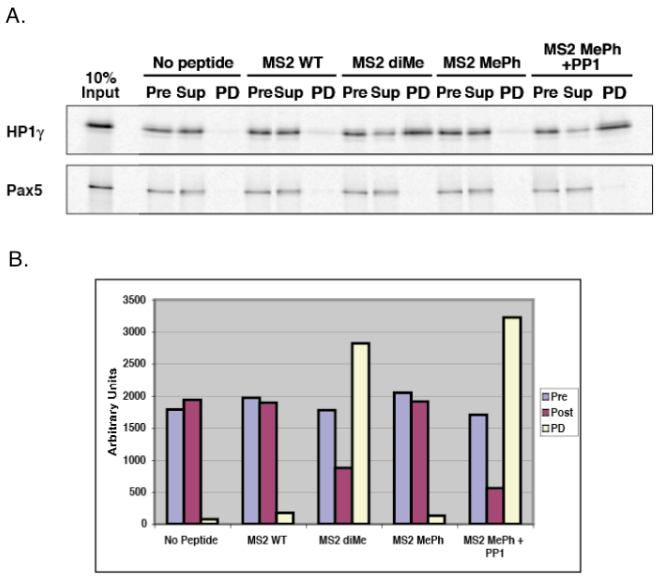

Figure 4.18 Phosphorylation of T166 blocks binding of HP1γ to methylated MS2. A. Beads conjugated with the indicated peptides were incubated with *in vitro* translated ^{35}S-labeled HP1γ or Pax5, and binding examined by SDS-PAGE and autoradiography. The dimethyl-phospho MS2 peptides were also pre-treated with Protein Phosphatase 1 (PP1) to control for peptide coupling efficiency. Abbreviations: Pre, pre-pull down sample; Sup, post-pull down supernatant; PD, bead fraction after pull down. B. Quantitation of the results in (A).

CHAPTER 5: DISCUSSION

5.1 Studying histone modification in the immune system

The study of chromatin structure and function has been challenging for many of the same reasons that make the subject compelling: the extreme conservation of histones and the proteins that bind and modify them make genetic approaches in higher eukaryotes difficult. It is therefore not surprising that significant insights into the physiological functions of histone modification in vertebrates have been made through studies of the immune system, where loss of function phenotypes do not necessarily cause lethality (Corcoran et al., 1998; Johnson et al., 2004; Su et al., 2003; Su et al., 2005).

Within the immune system, the processes that have been examined the most intensively from a chromatin perspective have unsurprisingly been those involving DNA recombination reactions: V(D)J recombination and class switch recombination (CSR). While the potential contribution of chromatin-based mechanisms for control of CSR are only beginning to be addressed (see Chapter 1), "locus accessibility" has been an issue at the forefront of studies on V(D)J recombination for quite some time (Corcoran et al., 1998; McMurry and Krangel, 2000; Sleckman et al., 1996). V(D)J recombination in B cells poses a particular challenge, as the immunoglobulin heavy chain locus spans several megabases, and must be activated in a stepwise manner for recombination in a process that may require physical contraction of the locus (Chowdhury and Sen, 2001; Fuxa et al., 2004).

Intriguingly, many of the molecules and pathways demonstrated by knockout studies to be important for immunoglobulin V(D)J recombination demonstrate similar loss of function phenotypes: disruption of the IL-7 receptor α chain, the transcription factor Pax5

or the histone methyltransferase Ezh2 all cause a block in pro- to pre-B cell development, with progressively less efficient use of variable genes located at a distance from the D and J loci (Corcoran et al., 1998; Hesslein et al., 2003; Su et al., 2002). However, an important point to consider in analysing the results of such gene targeting studies is that the marked similarity in phenotype may reflect either a common function for these disparate molecules, or a non-specific effect of perturbing lymphocyte development.

The possibility that failure of B cell development might reflect a more general cell biological problem is significant, since many genes whose deletion has a phenotype in lymphocyte development demonstrate developmental failure at precisely the stages during which cells undergo their most dramatic bursts of cell division (Chi et al., 2003; Cobb et al., 2005; Corcoran et al., 1998; Muljo et al., 2005; Su et al., 2002). It is therefore plausible that deletion of crucial chromatin regulators with cell-type independent functions causes a non-specific stress leading to developmental arrest. This would imply that many of the "specific" phenotypes generated in developing lymphocytes are specific mainly due to the experimental system used. This issue is of particular concern in the setting of conditional mutagenesis, which by its nature produces an artifactual degree of specificity.

However the argument in favor of a specific effect is supported by many lines of evidence, including the fact that while the overall effect of these knockouts may be similar, there are many specific differences as well. For instance, while loss of either Pax5 or IL-7Rα blocks development at the pro-B stage with reduced usage of 5' V genes, Pax5 deficiency is not associated with a lack of accessibility of these gene segments (as measured by histone acetylation and germline transcription), while IL-7Rα deficiency is

(Corcoran et al., 1996; Hesslein et al., 2003). Similarly, while the pro-B cell block in Ezh2 deficiency can be rescued by providing a pre-rearranged Ig heavy chain, Pax5 deficiency cannot (Su et al., 2002; Thevenin et al., 1998), arguing for a specific function for Ezh2 in V(D)J recombination *per sé*.

Also arguing for the specificity of these developmental blocks are the observations that some proteins, such as Ezh2 and G9a, appear critical for B cell development, but largely dispensable for activation and proliferation of peripheral B cells (Figure 3.13 and 3.14;(Su et al., 2002). Similarly, while G9a deletion severely blocks B cell development, its loss appears to have no obvious effect on α:β T cell development, which is also accompanied by rapid cell division during the double negative to double positive transition (Figure 3.10). These instances of undisturbed cell division in the absence of proteins such as G9a and Ezh2 lend credence to the suggestion that the observed developmental phenotypes are specific. However, it is difficult to overcome the formal argument that developing B cells are intrinsically more sensitive to the overall "health" of the cell, and therefore more non-specifically sensitive to manipulation. It will therefore be necessary to continue to interpret developmental phenotypes with caution.

5.2 Possible functions for G9a and GLP in developing lymphocytes

Our studies and those of other groups have shown that G9a and GLP are essential for the vast majority of euchromatic H3K9 dimethylation (Tachibana et al., 2002; Tachibana et al., 2005). Several groups have attempted to address the possible roles of euchromatic H3K9 methylation in the control of V(D)J recombination in T and B cells. This has mainly taken the form of chromatin immunoprecipitation (ChIP) analysis of the Ig and

TCR loci with modification-specific antibodies (Johnson et al., 2004; Morshead et al., 2003), which collectively suggested that diMeH3K9 might be enriched on inactive gene segments. This hypothesis was strengthened by the finding that direct recruitment of the G9a SET domain to the VEGF promoter was sufficient to initiate silencing (Snowden et al., 2002), and that artifically "tethering" a catalytic fragment of G9a to a TCR locus minigene was sufficient to inhibit germline transcription and subsequent V(D)J recombination (Osipovich et al., 2004). Such experiments imply that localized recruitment of G9a and/or GLP could represent a powerful mechanism for diminishing locus accessibility. However, such experiments are limited by their artificiality and inability to distinguish between simple correlation and a more causal role for diMeH3K9 in gene silencing.

The main alternate hypothesis, that localized histone methylation might "mark" target loci for recombination, has to date not been addressed in the immune system. However, there is a precedent for such a marking system in the unicellular eukaryote *Tetrahymena*, where localized domains of H3K9 methylation may function to tag sites of recombination during macronuclear development (Taverna et al., 2002). The unimpaired germline trancription observed in Ezh2-deficient B cells is also compatible with a model in which histone methylation either directly or indirectly targets recombination activity to particular variable region genes. If this were indeed the case, it is possible that such an effect might be overridden in the "tethering" experiments discussed earlier by the use of truncated G9a, which is unlikely to partner correctly with its endogenous binding proteins [Figures 4.7 and 4.12; (Osipovich et al., 2004; Roopra et al., 2004)].

The results reported here may allow discrimination between these two models for the function of histone methyltransferases in lymphocyte development since, in contrast to Ezh2, we only observe defective development of B cells in the absence of G9a (Figure 3.9 and 3.10; (Su et al., 2005). The grossly normal development of G9a-deficient T cells would seem to exclude a non-redundant role for H3K9 dimethylation in "marking" of V_H genes for recombination, although cell-type specific effects cannot be excluded. Overall, the existing data on G9a are most consistent with a role for diMeH3K9 in promoting immunoglobulin heavy chain recombination specifically, possibly through heterochromatin-induced contraction of the Ig locus.

While our data do not directly address whether G9a-mediated methylation might also restrict recombination to particular TCR loci as proposed previously (Osipovich et al., 2004), this question is addressable using the genetic systems reported here. One prediction of models in which diMeH3K9 limits recombination is that loss of this mark should allow promiscuous recombination. It will therefore be of interest to determine whether disruption of G9a expression allows, for instance, productive TCRα/β recombination in non-T cells, or co-expression of κ and λ light chains in individual B cells. Similarly, a classical experiment in molecular immunology was the demonstration that RAG-1 and RAG-2 were capable of potently inducing V(D)J recombination when co-transfected into NIH-3T3 cells (Oettinger et al., 1990; Schatz et al., 1989). It would be interesting to repeat such an experiment in G9a-deficient fibroblasts, which would be expected to undergo recombination at a much higher rate if diMeH3K9 functioned endogenously to decrease recombinase access to its substrates.

5.3 G9a methylation as an extension of the histone code

Over four decades ago, Allfrey, Mirsky and colleagues made the initial observation that histones could exist in a lysine methylated state *in vivo* (Allfrey et al., 1964). The first evidence for protein lysine methylation, however, came from studies a decade earlier of a non-histone target, the flagellar protein of *Salmonella* (Ambler and Rees, 1959). A possible connection between histone and non-histone methylation systems was proposed by Jenuwein and colleagues in their description of the first cloned histone methyltransferase, Suv39H1, where it was speculated that the same HMTases that modified histones might also act as more general "MTases" on non-histone targets (Rea et al., 2000). Credibility was lent to this hypothesis by the recent description in yeast of Set1-mediated methylation of a kinetochore component, Dam1 (Zhang et al., 2005). However, as with the previously described targets TAF10 and p53, lysine methylation of Dam1 was observed on a site lacking similarity to histones, leaving possible functional similarities between histone and non-histone methylation obscure (Chuikov et al., 2004; Kouskouti et al., 2004; Zhang et al., 2005). Thus the possibility that "modification cassettes" might exist in non-histone proteins has until now rested on the basis of sequence similarity alone (Fischle et al., 2003a).

The results reported here on G9a methylation directly demonstrate that the modification systems operating on histones are indeed conserved to non-histone proteins, and therefore likely represent a universal feature of protein regulation in eukaryotes. We find that highly conserved histone-like sequence motifs, which we term histone mimics, exist in the amino-terminal "tail" of the histone methyltransferase G9a, and that at least one of these motifs behaves in a manner almost identical to H3K9 with respect to

methylation, binding of methyllysine by HP1, and the ability of this interaction to be relieved by phosphorylation of an adjacent residue (Figures 4.4, 4.7, 4.15 and 4.16).

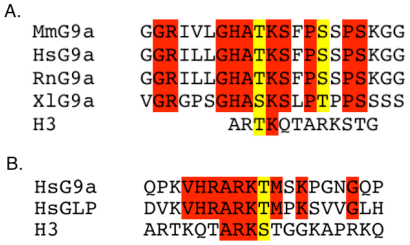

Figure 5.1 Alignments of G9a MS1 and MS2 A. Alignment of MS1 from various species with H3K4. B. Alignment of MS2 with H3K9. Residues identical in G9a homologs (A) or between G9a and GLP (B) are indicated in red; conservative substitutions are indicated in yellow. Abbreviations: Mm, *Mus musculus*; Hs, *Homo sapiens*; Rn, *Rattus norvegicus*; Xl, *Xenopus laevis*.

While our studies to date have been confined to the MS2 methylation site, it is clear that an important area of investigation for the future will be to define the degree to which the MS1 site behaves analogously to MS2 and H3K9. Recently the *S. cerevisiae* chromodomain-containing protein Chd1 and the mammalian WD40-repeat protein WDR5 were described to specifically interact with methylated H3K4 (Pray-Grant et al., 2005; Wysocka et al., 2005). The MS1 site demonstrates little sequence similarity to H3K4 (Figure 5.1A), but also differs from H3K9 and MS2 in that MS1 does not contain a consensus Aurora phosphorylation site [R/K-X-S/T, Figure 5.1A; (Ohi et al., 2004)]. This raises the possibility that, similar to H3K4, lysine methylation of MS1 may define interaction with a distinct class of effector modules.

In addition to the potential interaction of unique binding proteins with methylated MS1, the scope for combinatorial modification of G9a may also be significantly extended by additional modifications. For instance, we note that the MS2 site contains a lysine

(K169) that is conserved between G9a and GLP homologs in human and mouse, and which aligns within one amino acid of H3K14, a major site of histone acetylation in H3 [Figure 5.1b; (Strahl and Allis, 2000)]. We believe that G9a K169 will also prove to exist in an acetylated state *in vivo*, and indeed that the entire range of post-translational modifications on histones may also be present on G9a.

The complexity of G9a modification may itself represent only one mode of regulation of G9a activity. We have consistently observed that our G9a antibody cross-reacts with a cytoplasmic band of ~120 kD (Figure 4.4). This band is reproducibly induced by stimuli which cause cell stress, including adenoviral infection and freeze-thawing (Figure 4.11 and data not shown). Importantly, deletion of G9a with CD19-cre, which should not induce a stress response, causes a dramatic reduction in the amount of this cytoplasmic cross-reacting band, implying that it is produced from the G9a locus (Figure 3.11B). We believe that this band may therefore represent a stress-inducible cytoplasmic G9a splice variant. This possibility is supported by the fact that the G9a genomic locus is contained within the MHC class III region, which is known to encode many proteins upregulated in response to stress and immune activation (Brown et al., 2001; Hauptmann and Bahram, 2004; Spike and Lamont, 1995). The recent finding that the predominantly nuclear histone methyltransferase Ezh2 may have unexpected cytoplasmic functions (Su et al., 2005) suggests that cytoplasmic G9a might have important roles in the cellular response to "innate" activation.

5.4 Possible functions of G9a methylation

What are the physiological corollaries of G9a automethylation and HP1 binding? One obvious possibility is that methylation of G9a regulates its ability to control transcription. However, we did not detect any significant difference in steady-state transcriptional activity by microarray analysis of G9a-deficient MEFs reconstituted with either wild type or MS2 point mutant G9a (data not shown). This may reflect inherent redundancy in this system, since the related protein GLP carries an identical set of methylation sites, and has been described to heterodimerize with G9a [Figure 4.1; (Tachibana et al., 2005)]. We find this explanation unsatisfactory however, given that HP1γ was not recovered following immunoprecipitation of $G9a^{MS2KA}$, despite the presence of co-precipitated GLP (Figures 4.7, 4.12 and data not shown). Alternatively, although methylation of MS2 uniquely defines the interaction between G9a and HP1γ, it is possible that the MS1 site is physiologically redundant to MS2 *in vivo*. However the evolutionary conservation of both sites would not be expected if they were in fact redundant to one another.

Rather, we believe that our findings reflect the inherent plasticity of chromatin modification systems, where several activities are typically brought to bear simultaneously to regulate gene expression. For example, G9a has been found to exist in large repressor complexes containing HMTases (G9a and GLP), repressive transcription factors (CtBP, E2F6 and CDP/*cut*), histone deacteylases (HDACs 1 and 2) and an H3K4 demethylase (LSD1/NPAO), the activity of each of which is capable of transcriptional repression (Lee et al., 2005b; Nishio and Walsh, 2004; Oberley et al., 2003; Ogawa et al., 2002; Shi et al., 2004; Shi et al., 2003; Trimarchi et al., 2001). Indeed, it has been reported that H3K9-methylation by the G9a SET domain alone is sufficient to induce

HDAC-dependent gene silencing in the absence of HP1 recruitment (Stewart et al., 2005). These findings are consistent with our observation that reintroduction of the catalytically inactive G9a^{H1093K} allele into G9a$^{\Delta/\Delta}$ cells itself only deregulates the transcription of a relatively small number of genes (data not shown), despite the near-total absence of euchromatic H3K9 dimethylation (Fig. 3C). This may reflect a relative stability of the existing transcriptional program in these cells, which might be maintained through mechanisms requiring neither histone nor G9a methylation (e.g. histone deacetylation and demethylation).

A clearer assessment of the role of G9a methylation in G9a-mediated repression will ultimately require the analysis of animals harbouring both conditional deletions and single and compound knock-in alleles at the methylation sites of interest. Such tools will allow the analysis of methylation function at endogenous expression levels, and in a developmental context in which gene expression must be dynamically controlled in order to establish and perpetuate transcriptional memory (Peters and Schubeler, 2005). Considering the impossibility of performing genetics in mammals on the multicopy histone genes, as well as the potential for histone-like complexity of G9a modification, it is possible that the G9a conditional knockout model will also provide the first genetic platform in mammals for studying key predictions of the histode code hypothesis.

5.5 Potential histone mimics in chromatin-associated proteins

In addition to their relevance for our understanding of G9a function, the present studies strongly suggest that histone mimicry is a widespread and conserved phenomenon. An informative example of this conservation is seen in the G9a homologue

of *Drosophila* (dG9a, Figure 4.1); dG9a clearly contains a potential methylation site, but this exists in a simplified sequence context lacking both an upstream (MS1-type) methylation site and an adjacent phosphorylatable residue. One intriguing possibility is that the H3K9 mimic in dG9a represents an ancestral methylation site, which has subsequently been expanded and elaborated upon in vertebrate G9a homologues. Such a scenario is reminiscent of the C-terminal domain (CTD) of the RNA polymerase II RPB1 subunit, in which the number of heptapeptide repeats has increased through evolutionary time (26 repeats in yeast, 44 in *Drosophila* and 52 in humans), in step with increasing developmental and organismal complexity (Meinhart et al., 2005; Stiller and Hall, 2002).

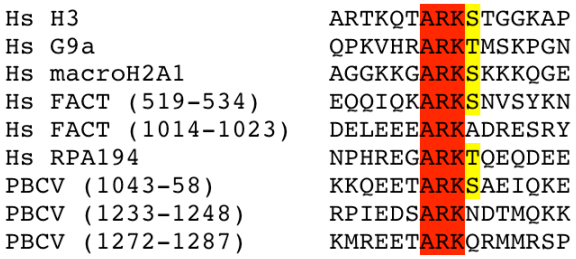

Figure 5.2 Potential H3K9 histone mimics in chromatin-associated proteins. H3K9-like sites are highlighted in red, and potential H3S10-like sites in yellow. Abbreviations: FACT, facilitates chromatin transcription; RPA194, RNA Polymerase I 194 kD subunit; PBCV, *Paramecium bursaria* chlorella virus.

In contrast to the simplified methylation site of dG9a, we find complete potential H3K9-type methylation/phosphorylation mimics in many vertebrate proteins involved in transcription and chromatin regulation. These include the heterochromatic histone variant macroH2A1, the H2A/H2B-interacting histone chaperone FACT (Facilitates Chromatin Transcription) and the RNA polymerase I subunit RPA194 (Figure 5.2). The occurence of H3 mimics in so many proteins involved in transcriptional regulation

suggests a basic role in the coordination of transcription with histone dynamics (e.g. modification, deposition and/or removal). This again raises an analogy with the RNA Pol II CTD, which is known to function in coordination of transcription with co-transcriptional events such as capping, splicing and polyadenylation (Proudfoot et al., 2002). Indeed, it was recently demonstrated that both diMeH3K9 and HP1γ are present specifically in the coding regions of actively transcribed genes, and that their presence requires Pol II-mediated elongation through chromatin (Vakoc et al., 2005). The possibility that H3K9 HMTases might be recruited to the Pol II holoenzyme is supported by the known association of the H3K4-specific Set1 and H3K36-specific HYPB HMTases with the elongating Pol II complex (Hampsey and Reinberg, 2003; Sun et al., 2005).

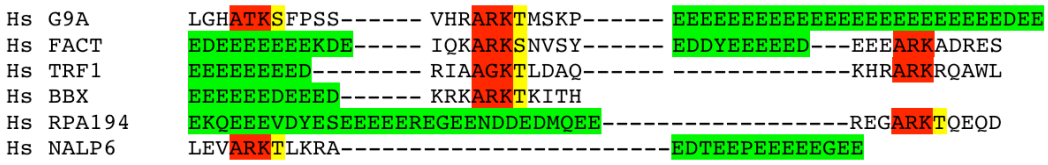

Figure 5.3 Expanded alignment of potential histone mimics. H3K9-like motifs are shown in red with potential phosphorylation sites in yellow. Acidic domains indicated in green. Abbreviations: FACT, facilitates chromatin transcription; TRF1, Telomeric Repeat-binding Factor 1; RPA194, RNA Polymerase I 194 kD subunit; NALP6, NACHT/LRR/Pyrin-domain containing protein 6.

Significantly, many proteins containing potential histone mimics further resemble G9a in that they contain multiple mimic sites and/or an extremely acidic region (in hG9a, a stretch of 28 consecutive glutamic or aspartic acid residues). Examples include FACT, the telomere repeat binding factor TRF1, RNA Pol I β subunit, and the cytosolic pro-

inflammatory signal transducer NALP6 (Figure 5.3). Highly acidic domains are in general characteristic of molecules which possess histone chaperone activity (e.g. nucleoplasmin, CAF-1), the ability to promote histone deposition onto naked DNA templates (Philpott et al., 2000). FACT in particular is known to behave as a histone chaperone, and is required both for transcription through chromatinized templates, as well as for redeposition of histones after Pol II-mediated transcription (Belotserkovskaya and Reinberg, 2004; Schwabish and Struhl, 2004). The acidic C-terminus of FACT has been studied in detail, and is critically required for histone chaperone activity (Belotserkovskaya et al., 2003).

Taken together, these finding suggest a model in which G9a and/or GLP could become associated with the elongating RNA Pol II complex. This association would allow co-recruitment of HP1γ to actively transcribed genes via G9a MS2 automethylation, in addition to recruiting other (as yet unidentified) proteins via MS1 methylation. The acidic domain of G9a might then function as a histone chaperone, similar to FACT, in order to evict histones ahead of the Pol II complex or redeposit replacement histones behind the complex. Such association between G9/GLP and Pol II could allow redeposition of modified histones co-transcriptionally, thus perpetuating the pre-exising pattern of histone modification, as has been observed to occur endogenously (Kouskouti and Talianidis, 2005).

5.6 Possible extension of histone mimicry to viruses

While the precise functions of histone mimics currently remain elusive, the cumulative evidence indicates that histone mimicry may represent a fundamental protein

regulatory mechanism. An interesting test of this hypothesis is whether histone mimicry as a mechanism is conserved in viruses, which naturally evolve to exploit key cellular processes. Abundant evidence exists that viruses actively utilize host chromatin functions to promote their replication. An example of this is the adenoviral E1A protein, which functions to force quiescent G_0 cells into the cell cycle, thus promoting viral replication. This is accomplished through derepression of target genes normally silenced by the pRb-related protein p130, and this reactivation is associated with dramatic loss of H3K9 methylation on target promoters and replacement of this mark with H3K9 acetylation (Ghosh and Harter, 2003). An even clearer example of viral use of host chromatin factors is the double bromodomain-containing protein Brd4, which is directly used by the bovine papillomavirus E2 protein to tether the viral genome to mitotic chromosomes (McBride et al., 2004; You et al., 2004). Finally, it was recently found that both Flock House Virus (FHV, an animal virus normally infecting insects), as well as Human Immunodeficiency Virus (HIV) encode proteins which act as suppressors of cellular RNAi-mediated immunity (Bennasser et al., 2005; Li et al., 2002; Lu et al., 2005).

To date, only one viral SET-domain containing histone methyltransferase has been described, that encoded by the double-stranded DNA virus *Paramecium bursaria* chlorella virus, PBCV-1 (Manzur et al., 2003). PBCV is a large (180 nm diameter, >330 kb genome) virus which infects zoochlorellae, eukaryotic algae which normally live symbiotically within hosts such as *Paramecium bursaria* (Van Etten and Meints, 1999). Chlorella viruses are unusual in that they encode enzymes similar to eukaryotic DNA methyltransferases, as well as prokaryote-like restriction-modification systems (Van

Etten and Meints, 1999), suggesting a complex interplay between virally-encoded proteins and the host genome.

The PBCV HMTase, referred to as vSET, was found to specifically methylate H3K27 (Manzur et al., 2003). Searches of the viral sequence databases with the H3K9 sequence demonstrated that a protein from PBCV contains a very similar site (Figure 5.2). Remarkably, manual inspection revealed that this protein (accession number NP_048536) contains three different H3K9-type sequences. Moreover, this protein is annotated as having similarity to the mammalian SWI/SNF chromatin remodeling complex subunit OSA2. Thus, the genome of PBCV simultaneously encodes both a functional histone methyltransferase, as well as a protein potentially capable of being multiply modified by this enzyme. The similarity of this putative methylation target to a known chromatin remodeler suggests that histone mimicry in this context may allow viral control of host chromatin remodeling machines, potentially contributing to the competition between viral and host genomes.

5.7 Conclusion

Studies of histone biology have entered an exciting phase, in which it has become possible to directly test the contribution of particular histone modifications to cellular and organismal physiology. Using conditional mutagenesis, we have begun to address the role of the H3K9-specific histone methyltransferase G9a in lymphocyte development and function. We find that while G9a is largely dispensable for T cell development, it is critically important for developing B cells. While the mechanism by which G9a, and H3K9 dimethylation more generally, contributes to B cell development remains to be

determined, the grossly normal behavior of G9a-deficient peripheral B cells argues for possible specificity of the developmental defect.

In the course of studying G9a and GLP, we found that histone methyltransferases are themselves capable of recapitulating many of the main biochemical features of histones. Through automethylation on an H3K9-like "histone mimic" site, G9a regulates its association with the methyllysine binding protein, HP1. The ability of this association to be further modified by phosphorylation of an adjacent residue speaks to the generality of the histone code, as well as the complexity of the regulatory mechanisms to which histone-modifying activities are subject. The finding that many other chromatin-associated proteins have the potential to be similarly modified, and that this mode of regulation may even extend to viral proteins, should guide the way for further studies to elucidate the varied functions of protein lysine methylation.

BIBLIOGRAPHY

Ahmad, K., and Henikoff, S. (2002). Histone H3 variants specify modes of chromatin assembly. Proc Natl Acad Sci U S A *99 Suppl 4*, 16477-16484.

Allfrey, V. G., Faulkner, R., and Mirsky, A. E. (1964). Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. Proc Natl Acad Sci U S A *51*, 786-794.

Alvarez-Venegas, R., and Avramova, Z. (2002). SET-domain proteins of the Su(var)3-9, E(z) and trithorax families. Gene *285*, 25-37.

Ambler, R. P., and Rees, M. W. (1959). Epsilon-N-Methyl-lysine in bacterial flagellar protein. Nature *184*, 56-57.

Anest, V., Hanson, J. L., Cogswell, P. C., Steinbrecher, K. A., Strahl, B. D., and Baldwin, A. S. (2003). A nucleosomal function for IkappaB kinase-alpha in NF-kappaB-dependent gene expression. Nature *423*, 659-663.

Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120-124.

Belotserkovskaya, R., Oh, S., Bondarenko, V. A., Orphanides, G., Studitsky, V. M., and Reinberg, D. (2003). FACT facilitates transcription-dependent nucleosome alteration. Science *301*, 1090-1093.

Belotserkovskaya, R., and Reinberg, D. (2004). Facts about FACT and transcript elongation through chromatin. Curr Opin Genet Dev *14*, 139-146.

Bennasser, Y., Le, S. Y., Benkirane, M., and Jeang, K. T. (2005). Evidence that HIV-1 encodes an siRNA and a suppressor of RNA silencing. Immunity *22*, 607-619.

Blanchet, F., Cardona, A., Letimier, F. A., Hershfield, M. S., and Acuto, O. (2005). CD28 costimulatory signal induces protein arginine methylation in T cells. J Exp Med *202*, 371-377.

Bolland, D. J., Wood, A. L., Johnston, C. M., Bunting, S. F., Morgan, G., Chakalova, L., Fraser, P. J., and Corcoran, A. E. (2004). Antisense intergenic transcription in V(D)J recombination. Nat Immunol *5*, 630-637.

Bolton, M. A., Lan, W., Powers, S. E., McCleland, M. L., Kuang, J., and Stukenberg, P. T. (2002). Aurora B kinase exists in a complex with survivin and INCENP and its kinase activity is stimulated by survivin binding and phosphorylation. Mol Biol Cell *13*, 3064-3077.

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

Briggs, S. D., Xiao, T., Sun, Z. W., Caldwell, J. A., Shabanowitz, J., Hunt, D. F., Allis, C. D., and Strahl, B. D. (2002). Gene silencing: trans-histone regulatory pathway in chromatin. Nature *418*, 498.

Brown, S. E., Campbell, R. D., and Sanderson, C. M. (2001). Novel NG36/G9a gene products encoded within the human and mouse MHC class III regions. Mamm Genome *12*, 916-924.

Brownell, J. E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S. Y., and Allis, C. D. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. Cell *84*, 843-851.

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

Chalker, D. L., Fuller, P., and Yao, M. C. (2005). Communication between parental and developing genomes during tetrahymena nuclear differentiation is likely mediated by homologous RNAs. Genetics *169*, 149-160.

Chalker, D. L., and Yao, M. C. (2001). Nongenic, bidirectional transcription precedes and may promote developmental DNA deletion in Tetrahymena thermophila. Genes Dev *15*, 1287-1298.

Chan, H. M., and La Thangue, N. B. (2001). p300/CBP proteins: HATs for transcriptional bridges and scaffolds. J Cell Sci *114*, 2363-2373.

Chi, T. H., Wan, M., Lee, P. P., Akashi, K., Metzger, D., Chambon, P., Wilson, C. B., and Crabtree, G. R. (2003). Sequential roles of Brg, the ATPase subunit of BAF chromatin remodeling complexes, in thymocyte development. Immunity *19*, 169-182.

Chowdhury, D., and Sen, R. (2001). Stepwise activation of the immunoglobulin mu heavy chain gene locus. Embo J *20*, 6394-6403.

Chuikov, S., Kurash, J. K., Wilson, J. R., Xiao, B., Justin, N., Ivanov, G. S., McKinney, K., Tempst, P., Prives, C., Gamblin, S. J.*, et al.* (2004). Regulation of p53 activity through lysine methylation. Nature *432*, 353-360.

Cobb, B. S., Nesterova, T. B., Thompson, E., Hertweck, A., O'Connor, E., Godwin, J., Wilson, C. B., Brockdorff, N., Fisher, A. G., Smale, S. T., and Merkenschlager, M. (2005). T cell lineage choice and differentiation in the absence of the RNase III enzyme Dicer. J Exp Med *201*, 1367-1373.

Corcoran, A. E., Riddell, A., Krooshoop, D., and Venkitaraman, A. R. (1998). Impaired immunoglobulin gene rearrangement in mice lacking the IL-7 receptor. Nature *391*, 904-907.

Corcoran, A. E., Smart, F. M., Cowling, R. J., Crompton, T., Owen, M. J., and Venkitaraman, A. R. (1996). The interleukin-7 receptor alpha chain transmits distinct signals for proliferation and differentiation during B lymphopoiesis. Embo J *15*, 1924-1932.

Coyne, R. S., Chalker, D. L., and Yao, M. C. (1996). Genome downsizing during ciliate development: nuclear division of labor through chromosome restructuring. Annu Rev Genet *30*, 557-578.

Coyne, R. S., Nikiforov, M. A., Smothers, J. F., Allis, C. D., and Yao, M. C. (1999). Parental expression of the chromodomain protein Pdd1p is required for completion of programmed DNA elimination and nuclear differentiation. Mol Cell *4*, 865-872.

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

Davis, M. M., Kim, S. K., and Hood, L. E. (1980). DNA sequences mediating class switching in alpha-immunoglobulins. Science *209*, 1360-1365.

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

Dillon, S. C., Zhang, X., Trievel, R. C., and Cheng, X. (2005). The SET-domain protein superfamily: protein lysine methyltransferases. Genome Biol *6*, 227.

Feng, Q., Wang, H., Ng, H. H., Erdjument-Bromage, H., Tempst, P., Struhl, K., and Zhang, Y. (2002). Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. Curr Biol *12*, 1052-1058.

Fischle, W., Wang, Y., and Allis, C. D. (2003a). Binary switches and modification cassettes in histone biology and beyond. Nature *425*, 475-479.

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

Fuxa, M., Skok, J., Souabni, A., Salvagiotto, G., Roldan, E., and Busslinger, M. (2004). Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. Genes Dev *18*, 411-422.

Ghosh, M. K., and Harter, M. L. (2003). A viral mechanism for remodeling chromatin structure in G0 cells. Mol Cell *12*, 255-260.

Giet, R., and Glover, D. M. (2001). Drosophila aurora B kinase is required for histone H3 phosphorylation and condensin recruitment during chromosome condensation and to organize the central spindle during cytokinesis. J Cell Biol *152*, 669-682.

Gu, H., Zou, Y. R., and Rajewsky, K. (1993). Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. Cell *73*, 1155-1164.

Gyory, I., Wu, J., Fejer, G., Seto, E., and Wright, K. L. (2004). PRDI-BF1 recruits the histone H3 methyltransferase G9a in transcriptional silencing. Nat Immunol *5*, 299-308.

Hampsey, M., and Reinberg, D. (2003). Tails of intrigue: phosphorylation of RNA polymerase II mediates histone methylation. Cell *113*, 429-432.

Hauptmann, G., and Bahram, S. (2004). Genetics of the central MHC. Curr Opin Immunol *16*, 668-672.

Hayes, J. J., and Hansen, J. C. (2001). Nucleosomes and the chromatin fiber. Curr Opin Genet Dev *11*, 124-129.

Henikoff, S. (2005). Histone modifications: combinatorial complexity or cumulative simplicity? Proc Natl Acad Sci U S A *102*, 5308-5309.

Hesslein, D. G., Pflugh, D. L., Chowdhury, D., Bothwell, A. L., Sen, R., and Schatz, D. G. (2003). Pax5 is required for recombination of transcribed, acetylated, 5' IgH V gene segments. Genes Dev *17*, 37-42.

Hsu, J. Y., Sun, Z. W., Li, X., Reuben, M., Tatchell, K., Bishop, D. K., Grushcow, J. M., Brame, C. J., Caldwell, J. A., Hunt, D. F.*, et al.* (2000). Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. Cell *102*, 279-291.

Huyen, Y., Zgheib, O., Ditullio, R. A., Jr., Gorgoulis, V. G., Zacharatos, P., Petty, T. J., Sheston, E. A., Mellert, H. S., Stavridi, E. S., and Halazonetis, T. D. (2004). Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. Nature *432*, 406-411.

Imhof, A., Yang, X. J., Ogryzko, V. V., Nakatani, Y., Wolffe, A. P., and Ge, H. (1997). Acetylation of general transcription factors by histone acetyltransferases. Curr Biol *7*, 689-692.

Iwasato, T., Shimizu, A., Honjo, T., and Yamagishi, H. (1990). Circular DNA is excised by immunoglobulin class switch recombination. Cell *62*, 143-149.

Jacobs, S. A., Taverna, S. D., Zhang, Y., Briggs, S. D., Li, J., Eissenberg, J. C., Allis, C. D., and Khorasanizadeh, S. (2001). Specificity of the HP1 chromo domain for the methylated N-terminus of histone H3. Embo J *20*, 5232-5241.

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

Johnson, K., Pflugh, D. L., Yu, D., Hesslein, D. G., Lin, K. I., Bothwell, A. L., Thomas-Tikhonenko, A., Schatz, D. G., and Calame, K. (2004). B cell-specific loss of histone 3 lysine 9 methylation in the V(H) locus depends on Pax5. Nat Immunol *5*, 853-861.

Jung, S., Rajewsky, K., and Radbruch, A. (1993). Shutdown of class switch recombination by deletion of a switch region control element. Science *259*, 984-987.

Kemp, B. E., Parker, M. W., Hu, S., Tiganis, T., and House, C. (1994). Substrate and pseudosubstrate interactions with protein kinases: determinants of specificity. Trends Biochem Sci *19*, 440-444.

Khorasanizadeh, S. (2004). The nucleosome: from genomic organization to genomic regulation. Cell *116*, 259-272.

Kinoshita, K., and Honjo, T. (2001). Linking class-switch recombination with somatic hypermutation. Nat Rev Mol Cell Biol *2*, 493-503.

Kouskouti, A., Scheer, E., Staub, A., Tora, L., and Talianidis, I. (2004). Gene-specific modulation of TAF10 function by SET9-mediated methylation. Mol Cell *14*, 175-182.

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

Kouzarides, T. (2000). Acetylation: a regulatory modification to rival phosphorylation? Embo J *19*, 1176-1179.

Kouzarides, T. (2002). Histone methylation in transcriptional control. Curr Opin Genet Dev *12*, 198-209.

Kuhn, R., Schwenk, F., Aguet, M., and Rajewsky, K. (1995). Inducible gene targeting in mice. Science *269*, 1427-1429.

Kuo, M. H., Brownell, J. E., Sobel, R. E., Ranalli, T. A., Cook, R. G., Edmondson, D. G., Roth, S. Y., and Allis, C. D. (1996). Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. Nature *383*, 269-272.

Lachner, M., and Jenuwein, T. (2002). The many faces of histone lysine methylation. Curr Opin Cell Biol *14*, 286-298.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature *410*, 116-120.

Lee, D. Y., Teyssier, C., Strahl, B. D., and Stallcup, M. R. (2005a). Role of protein methylation in regulation of transcription. Endocr Rev *26*, 147-170.

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

Lewandoski, M. (2001). Conditional control of gene expression in the mouse. Nat Rev Genet *2*, 743-755.

Li, H., Li, W. X., and Ding, S. W. (2002). Induction and suppression of RNA silencing by an animal virus. Science *296*, 1319-1321.

Lin, K. I., Tunyaplin, C., and Calame, K. (2003). Transcriptional regulatory cascades controlling plasma cell differentiation. Immunol Rev *194*, 19-28.

Lindroth, A. M., Shultis, D., Jasencakova, Z., Fuchs, J., Johnson, L., Schubert, D., Patnaik, D., Pradhan, S., Goodrich, J., Schubert, I.*, et al.* (2004). Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with CHROMOMETHYLASE3. Embo J *23*, 4286-4296.

Liu, Y., Mochizuki, K., and Gorovsky, M. A. (2004). Histone H3 lysine 9 methylation is required for DNA elimination in developing macronuclei in Tetrahymena. Proc Natl Acad Sci U S A *101*, 1679-1684.

Lorentz, A., Ostermann, K., Fleck, O., and Schmidt, H. (1994). Switching gene swi6, involved in repression of silent mating-type loci in fission yeast, encodes a homologue of chromatin-associated proteins from Drosophila and mammals. Gene *143*, 139-143.

Lu, R., Maduro, M., Li, F., Li, H. W., Broitman-Maduro, G., Li, W. X., and Ding, S. W. (2005). Animal virus replication and RNAi-mediated antiviral silencing in Caenorhabditis elegans. Nature *436*, 1040-1043.

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution [see comments]. Nature *389*, 251-260.

Lusser, A., and Kadonaga, J. T. (2003). Chromatin remodeling by ATP-dependent molecular machines. Bioessays *25*, 1192-1200.

Madireddi, M. T., Coyne, R. S., Smothers, J. F., Mickey, K. M., Yao, M. C., and Allis, C. D. (1996). Pdd1p, a novel chromodomain-containing protein, links heterochromatin assembly and DNA elimination in Tetrahymena. Cell *87*, 75-84.

Manzur, K. L., Farooq, A., Zeng, L., Plotnikova, O., Koch, A. W., Sachchidanand, and Zhou, M. M. (2003). A dimeric viral SET domain methyltransferase specific to Lys27 of histone H3. Nat Struct Biol *10*, 187-196.

McBride, A. A., McPhillips, M. G., and Oliveira, J. G. (2004). Brd4: tethering, segregation and beyond. Trends Microbiol *12*, 527-529.

McMurry, M. T., and Krangel, M. S. (2000). A role for histone acetylation in the developmental regulation of VDJ recombination. Science *287*, 495-498.

Meinhart, A., Kamenski, T., Hoeppner, S., Baumli, S., and Cramer, P. (2005). A structural perspective of CTD function. Genes Dev *19*, 1401-1415.

Metzger, E., Wissmann, M., Yin, N., Muller, J. M., Schneider, R., Peters, A. H., Gunther, T., Buettner, R., and Schule, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature.

Miranda, T. B., Webb, K. J., Edberg, D. D., Reeves, R., and Clarke, S. (2005). Protein arginine methyltransferase 6 specifically methylates the nonhistone chromatin protein HMGA1a. Biochem Biophys Res Commun *336*, 831-835.

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

Mochizuki, K., Fine, N. A., Fujisawa, T., and Gorovsky, M. A. (2002). Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in tetrahymena. Cell *110*, 689-699.

Mochizuki, K., and Gorovsky, M. A. (2004). Conjugation-specific small RNAs in Tetrahymena have predicted properties of scan (scn) RNAs involved in genome rearrangement. Genes Dev *18*, 2068-2073.

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

Morshead, K. B., Ciccone, D. N., Taverna, S. D., Allis, C. D., and Oettinger, M. A. (2003). Antigen receptor loci poised for V(D)J rearrangement are broadly associated with BRG1 and flanked by peaks of histone H3 dimethylated at lysine 4. Proc Natl Acad Sci U S A *100*, 11577-11582.

Muljo, S. A., Ansel, K. M., Kanellopoulou, C., Livingston, D. M., Rao, A., and Rajewsky, K. (2005). Aberrant T cell differentiation in the absence of Dicer. J Exp Med *202*, 261-269.

Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., and Honjo, T. (2000). Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. Cell *102*, 553-563.

Murray, A. W. (1991). Cell cycle extracts. Methods Cell Biol *36*, 581-605.

Nakayama, J., Rice, J. C., Strahl, B. D., Allis, C. D., and Grewal, S. I. (2001). Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science *292*, 110-113.

Ng, H. H., Ciccone, D. N., Morshead, K. B., Oettinger, M. A., and Struhl, K. (2003). Lysine-79 of histone H3 is hypomethylated at silenced loci in yeast and mammalian cells: a potential mechanism for position-effect variegation. Proc Natl Acad Sci U S A *100*, 1820-1825.

Nishio, H., and Walsh, M. J. (2004). CCAAT displacement protein/cut homolog recruits G9a histone lysine methyltransferase to repress transcription. Proc Natl Acad Sci U S A *101*, 11257-11262.

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

Nowak, S. J., and Corces, V. G. (2000). Phosphorylation of histone H3 correlates with transcriptionally active loci. Genes Dev *14*, 3003-3013.

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

Oberley, M. J., Inman, D. R., and Farnham, P. J. (2003). E2F6 negatively regulates BRCA1 in human cancer cells without methylation of histone H3 on lysine 9. J Biol Chem *278*, 42466-42476.

Oettinger, M. A., Schatz, D. G., Gorka, C., and Baltimore, D. (1990). RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. Science *248*, 1517-1523.

Ogawa, H., Ishiguro, K., Gaubatz, S., Livingston, D. M., and Nakatani, Y. (2002). A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells. Science *296*, 1132-1136.

Ohi, R., Sapra, T., Howard, J., and Mitchison, T. J. (2004). Differentiation of cytoplasmic and meiotic spindle assembly MCAK functions by Aurora B-dependent phosphorylation. Mol Biol Cell *15*, 2895-2906.

Ong, S. E., Mittler, G., and Mann, M. (2004). Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. Nat Methods *1*, 119-126.

Osipovich, O., Milley, R., Meade, A., Tachibana, M., Shinkai, Y., Krangel, M. S., and Oltz, E. M. (2004). Targeted inhibition of V(D)J recombination by a histone methyltransferase. Nat Immunol *5*, 309-316.

Owen, D. J., Ornaghi, P., Yang, J. C., Lowe, N., Evans, P. R., Ballario, P., Neuhaus, D., Filetici, P., and Travers, A. A. (2000). The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p. Embo J *19*, 6141-6149.

Paik, W. K., and Kim, S. (1971). Protein methylation. Science *174*, 114-119.

Perez-Burgos, L., Peters, A. H., Opravil, S., Kauer, M., Mechtler, K., and Jenuwein, T. (2004). Generation and characterization of methyl-lysine histone antibodies. Methods Enzymol *376*, 234-254.

Peters, A. H., Kubicek, S., Mechtler, K., O'Sullivan, R. J., Derijck, A. A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y*., et al.* (2003). Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. Mol Cell *12*, 1577-1589.

Peters, A. H., O'Carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schofer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A*., et al.* (2001). Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell *107*, 323-337.

Peters, A. H., and Schubeler, D. (2005). Methylation of histones: playing memory with DNA. Curr Opin Cell Biol *17*, 230-238.

Philpott, A., Krude, T., and Laskey, R. A. (2000). Nuclear chaperones. Semin Cell Dev Biol *11*, 7-14.

Piskurich, J. F., Lin, K. I., Lin, Y., Wang, Y., Ting, J. P., and Calame, K. (2000). BLIMP-I mediates extinction of major histocompatibility class II transactivator expression in plasma cells.PG - 526-32. Nat Immunol *1*.

Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Lee, T. I., Bell, G. W., Walker, K., Rolfe, P. A., Herbolsheimer, E*., et al.* (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. Cell *122*, 517-527.

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

Proudfoot, N. J., Furger, A., and Dye, M. J. (2002). Integrating mRNA processing with transcription. Cell *108*, 501-512.

Ramiro, A. R., Stavropoulos, P., Jankovic, M., and Nussenzweig, M. C. (2003). Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. Nat Immunol *4*, 452-456.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z. W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C. P., Allis, C. D., and Jenuwein, T. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593-599.

Rice, J. C., Briggs, S. D., Ueberheide, B., Barber, C. M., Shabanowitz, J., Hunt, D. F., Shinkai, Y., and Allis, C. D. (2003). Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. Mol Cell *12*, 1591-1598.

Rickert, R. C., Rajewsky, K., and Roes, J. (1995). Impairment of T-cell-dependent B-cell responses and B-1 cell development in CD19-deficient mice. Nature *376*, 352-355.

Rickert, R. C., Roes, J., and Rajewsky, K. (1997). B lymphocyte-specific, Cre-mediated mutagenesis in mice. Nucleic Acids Res *25*, 1317-1318.

Rodriguez, C. I., Buchholz, F., Galloway, J., Sequerra, R., Kasper, J., Ayala, R., Stewart, A. F., and Dymecki, S. M. (2000). High-efficiency deleter mice show that FLPe is an alternative to Cre-loxP. Nat Genet *25*, 139-140.

Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W. W., Wilm, M., Aasland, R., and Stewart, A. F. (2001). The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. Embo J *20*, 7137-7148.

Rolink, A., Melchers, F., and Andersson, J. (1996). The SCID but not the RAG-2 gene product is required for S mu-S epsilon heavy chain class switching. Immunity *5*, 319-330.

Roopra, A., Qazi, R., Schoenike, B., Daley, T. J., and Morrison, J. F. (2004). Localized domains of G9a-mediated histone methylation are required for silencing of neuronal genes. Mol Cell *14*, 727-738.

Roth, S. Y., Denu, J. M., and Allis, C. D. (2001). Histone acetyltransferases. Annu Rev Biochem *70*, 81-120.

Sampath, S. C., Ohi, R., Leismann, O., Salic, A., Pozniakovski, A., and Funabiki, H. (2004). The chromosomal passenger complex is required for chromatin-induced microtubule stabilization and spindle assembly. Cell *118*, 187-202.

Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C., Schreiber, S. L., Mellor, J., and Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. Nature *419*, 407-411.

Schatz, D. G., Oettinger, M. A., and Baltimore, D. (1989). The V(D)J recombination activating gene, RAG-1. Cell *59*, 1035-1048.

Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R., and Reuter, G. (2002). Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. Embo J *21*, 1121-1131.

Schumacher, A., and Magnuson, T. (1997). Murine Polycomb- and trithorax-group genes regulate homeotic pathways and beyond. Trends Genet *13*, 167-170.

Schwabish, M. A., and Struhl, K. (2004). Evidence for eviction and rapid deposition of histones upon transcriptional elongation by RNA polymerase II. Mol Cell Biol *24*, 10111-10117.

Shaffer, A. L., Lin, K. I., Kuo, T. C., Yu, X., Hurt, E. M., Rosenwald, A., Giltnane, J. M., Yang, L., Zhao, H., Calame, K., and Staudt, L. M. (2002). Blimp-1 orchestrates plasma cell differentiation by extinguishing the mature B cell gene expression program. Immunity *17*, 51-62.

Shapiro-Shelef, M., Lin, K. I., McHeyzer-Williams, L. J., Liao, J., McHeyzer-Williams, M. G., and Calame, K. (2003). Blimp-1 is required for the formation of immunoglobulin secreting plasma cells and pre-plasma memory B cells. Immunity *19*, 607-620.

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

Shi, Y., Sawada, J., Sui, G., Affar el, B., Whetstine, J. R., Lan, F., Ogawa, H., Luke, M. P., and Nakatani, Y. (2003). Coordinated histone modifications mediated by a CtBP co-repressor complex. Nature *422*, 735-738.

Shimizu, A., Takahashi, N., Yaoita, Y., and Honjo, T. (1982). Organization of the constant-region gene family of the mouse immunoglobulin heavy chain. Cell *28*, 499-506.

Shinkura, R., Tian, M., Smith, M., Chua, K., Fujiwara, Y., and Alt, F. W. (2003). The influence of transcriptional orientation on endogenous switch region function. Nat Immunol *4*, 435-441.

Sleckman, B. P., Gorman, J. R., and Alt, F. W. (1996). Accessibility control of antigen-receptor variable-region gene assembly: role of cis-acting elements. Annu Rev Immunol *14*, 459-481.

Snapper, C. M., Marcu, K. B., and Zelazowski, P. (1997). The immunoglobulin class switch: beyond "accessibility". Immunity *6*, 217-223.

Snowden, A. W., Gregory, P. D., Case, C. C., and Pabo, C. O. (2002). Gene-Specific Targeting of H3K9 Methylation Is Sufficient for Initiating Repression In Vivo. Curr Biol *12*, 2159-2166.

Spike, C. A., and Lamont, S. J. (1995). Genetic analysis of three loci homologous to human G9a: evidence for linkage of a class III gene with the chicken MHC. Anim Genet *26*, 185-187.

Stewart, M. D., Li, J., and Wong, J. (2005). Relationship between histone H3 lysine 9 methylation, transcription repression, and heterochromatin protein 1 recruitment. Mol Cell Biol *25*, 2525-2538.

Stiller, J. W., and Hall, B. D. (2002). Evolution of the RNA polymerase II C-terminal domain. Proc Natl Acad Sci U S A *99*, 6091-6096.

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

Strahl, B. D., Briggs, S. D., Brame, C. J., Caldwell, J. A., Koh, S. S., Ma, H., Cook, R. G., Shabanowitz, J., Hunt, D. F., Stallcup, M. R., and Allis, C. D. (2001). Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. Curr Biol *11*, 996-1000.

Su, I. H., Basavaraj, A., Krutchinsky, A. N., Hobert, O., Ullrich, A., Chait, B. T., and Tarakhovsky, A. (2002). Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. Nat Immunol.

Su, I. H., Basavaraj, A., Krutchinsky, A. N., Hobert, O., Ullrich, A., Chait, B. T., and Tarakhovsky, A. (2003). Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. Nat Immunol *4*, 124-131.

Su, I. H., Dobenecker, M. W., Dickinson, E., Oser, M., Basavaraj, A., Marqueron, R., Viale, A., Reinberg, D., Wulfing, C., and Tarakhovsky, A. (2005). Polycomb group protein ezh2 controls actin polymerization and cell signaling. Cell *121*, 425-436.

Sun, X. J., Wei, J., Wu, X. Y., Hu, M., Wang, L., Wang, H. H., Zhang, Q. H., Chen, S. J., Huang, Q. H., and Chen, Z. (2005). Identification and characterization of a novel human histone H3 lysine 36 specific methyltransferase. J Biol Chem.

Sun, Z. W., and Allis, C. D. (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature *418*, 104-108.

Tachibana, M., Sugimoto, K., Fukushima, T., and Shinkai, Y. (2001). Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. J Biol Chem *276*, 25309-25317.

Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H., and Shinkai, Y. (2002). G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes Dev *16*, 1779-1791.

Tachibana, M., Ueda, J., Fukuda, M., Takeda, N., Ohta, T., Iwanari, H., Sakihama, T., Kodama, T., Hamakubo, T., and Shinkai, Y. (2005). Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. Genes Dev *19*, 815-826.

Taverna, S. D., Coyne, R. S., and Allis, C. D. (2002). Methylation of histone h3 at lysine 9 targets programmed DNA elimination in tetrahymena. Cell *110*, 701-711.

Tenney, K., and Shilatifard, A. (2005). A COMPASS in the voyage of defining the role of trithorax/MLL-containing complexes: linking leukemogensis to covalent modifications of chromatin. J Cell Biochem *95*, 429-436.

Teyssier, C., Ma, H., Emter, R., Kralli, A., and Stallcup, M. R. (2005). Activation of nuclear receptor coactivator PGC-1alpha by arginine methylation. Genes Dev *19*, 1466-1473.

Thevenin, C., Nutt, S. L., and Busslinger, M. (1998). Early function of Pax5 (BSAP) before the pre-B cell receptor stage of B lymphopoiesis. J Exp Med *188*, 735-744.

Tidwell, T., Allfrey, V. G., and Mirsky, A. E. (1968). The methylation of histones during regeneration of the liver. J Biol Chem *243*, 707-715.

Trimarchi, J. M., Fairchild, B., Wen, J., and Lees, J. A. (2001). The E2F6 transcription factor is a component of the mammalian Bmi1-containing polycomb complex. Proc Natl Acad Sci U S A *98*, 1519-1524.

Vakoc, C. R., Mandat, S. A., Olenchock, B. A., and Blobel, G. A. (2005). Histone H3 Lysine 9 Methylation and HP1gamma Are Associated with Transcription Elongation through Mammalian Chromatin. Mol Cell *19*, 381-391.

Van Etten, J. L., and Meints, R. H. (1999). Giant viruses infecting algae. Annu Rev Microbiol *53*, 447-494.

Wei, Y., Yu, L., Bowen, J., Gorovsky, M. A., and Allis, C. D. (1999). Phosphorylation of histone H3 is required for proper chromosome condensation and segregation. Cell *97*, 99-109.

Wysocka, J., Swigut, T., Milne, T. A., Dou, Y., Zhang, X., Burlingame, A. L., Roeder, R. G., Brivanlou, A. H., and Allis, C. D. (2005). WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. Cell *121*, 859-872.

Yamamoto, Y., Verma, U. N., Prajapati, S., Kwak, Y. T., and Gaynor, R. B. (2003). Histone H3 phosphorylation by IKK-alpha is critical for cytokine-induced gene expression. Nature *423*, 655-659.

Yancopoulos, G. D., Desiderio, S. V., Paskind, M., Kearney, J. F., Baltimore, D., and Alt, F. W. (1984). Preferential utilization of the most JH-proximal VH gene segments in pre-B-cell lines. Nature *311*, 727-733.

Yang, X. J. (2005). Multisite protein modification and intramolecular signaling. Oncogene *24*, 1653-1662.

Yao, M. C., Fuller, P., and Xi, X. (2003). Programmed DNA deletion as an RNA-guided system of genome defense. Science *300*, 1581-1584.

You, J., Croyle, J. L., Nishimura, A., Ozato, K., and Howley, P. M. (2004). Interaction of the bovine papillomavirus E2 protein with Brd4 tethers the viral DNA to host mitotic chromosomes. Cell *117*, 349-360.

Yu, J., Angelin-Duclos, C., Greenwood, J., Liao, J., and Calame, K. (2000). Transcriptional repression by blimp-1 (PRDI-BF1) involves recruitment of histone deacetylase. Mol Cell Biol *20*, 2592-2603.

Zhang, K., Lin, W., Latham, J. A., Riefler, G. M., Schumacher, J. M., Chan, C., Tatchell, K., Hawke, D. H., Kobayashi, R., and Dent, S. Y. (2005). The Set1 Methyltransferase Opposes Ipl1 Aurora Kinase Functions in Chromosome Segregation. Cell *122*, 723-734.