Epigenetic regulation of transcription in Drosophila

Aishwarya Swaminathan¹,², Ambikai Gajan¹, Lori A Pile¹

¹Department of Biological Sciences, Wayne State University, Detroit, Michigan, 48202, USA, ²Current address: Department of Microbiology, University of Massachusetts, Amherst, Massachusetts, 01003, USA

TABLE OF CONTENTS

1. Abstract
2. Introduction
3. Where are the marks found along the genome?
4. Where are the marks relative to large chromatin domains?
   4.1. Regulation of the male X chromosome
   4.2. Heterochromatin
5. Locus specific function of histone marks
   5.1. Acetylation
   5.2. Methylation
   5.3. Arginine methylation
   5.4. Ubiquitination
   5.5. Phosphorylation
6. Conclusions and future perspectives
7. Acknowledgments
8. References

1. ABSTRACT

Post-translational modification of histones is a major mechanism of epigenetic regulation of eukaryotic transcription. Drosophila has proven to be an important model system for the study of histone modifying enzymes and the cross talk that occurs between the various modifications. Polytene chromosome analysis and genome-wide chromatin immunoprecipitation (ChIP) studies have provided much insight into the location of marks and many of the enzymes that perform the catalytic reactions. Gene specific effects have been determined through study of flies carrying mutations in histone modifying enzymes. This review will highlight classic studies and present recent progress on both the localization data and mutant analyses. This information has been used to assign function to the marks and to the enzymes that place or remove them, critical for the process of transcriptional regulation.

2. INTRODUCTION

Development is dependent on appropriate timing and cell specificity of gene expression. Gene expression is controlled in part by regulatory elements recognized by sequence specific activators and repressors and components of the general transcription machinery. In addition, the accessibility of these factors to their cognate DNA binding sites is defined in large part by chromatin structure. Transcription takes place on DNA packaged into chromatin, comprised of DNA wound around histone proteins. Histones are subject to post-translational modifications, or marks, that can affect chromatin structure either by directly affecting packaging or by affecting protein binding to the region (1). Individual modifications as well as combinations of histone marks generate a histone code that is associated with distinct transcriptional states (2). As such, histone modifications are considered
epigenetic regulators since they define gene expression activity. A number of enzymes have been identified that function to either add or remove these modifications. The importance of these enzymes is highlighted by the fact that many of the histone modifying enzymes are required for metazoan viability. Exactly how and when the marks are established and maintained by the various enzymes is an area of active research.

Drosophila is a well-established model system for the analysis of histone modifying enzymes and epigenetic studies. Drosophila is a classic model in which the organism has defined stages of development. Numerous mutants are available from stock centers and tissue specific knockdown via RNA interference (RNAi) can be achieved for essential genes. Polytene chromosomes allow direct visualization of chromatin structure, location of modified histones and binding of protein factors involved in the regulation of chromatin structure and transcription. Genome-wide studies have been used to determine gene expression profiles and localization studies have provided high resolution data indicating where many modifications are found.

In this review, we highlight both historic and emerging ideas of epigenetic regulation obtained using the Drosophila model system. We will review genome-wide localization data of epigenetic marks, from the original polytene chromosome binding analyses to high resolution ChIP-chip studies. Next, we will discuss the association of specific marks with large chromosomal domains. We will then highlight the enzymes that catalyze these reactions and provide information as to the developmental and signaling roles of these factors on the regulation of individual genes. Throughout, we will highlight instances where one mark is associated with another, indicative of cross talk between modifications. Transmission of these marks is still not fully understood and investigations using the Drosophila system are expected to shed light on this process.

3. WHERE ARE THE MARKS FOUND ALONG THE GENOME?

The location of acetylated histones was initially determined using polytene chromosome immunofluorescence analysis. Polytene chromosomes were probed with antibodies to individual acetylated histone residues, specifically lysines (K) 5, 8, 12 and 16 of histone H4 (3). Acetylation of K5 and K8 of histone H4 (H4K5Ac and H4K8Ac) is localized along euchromatin arms while the heterochromatic chromocenter is relatively depleted of these marks. Alternatively, the chromocenter is enriched in H4K12Ac. H4K16Ac exhibits the most distinct pattern of localization as it is found enriched specifically on the male X chromosome. The localization pattern of H4K16Ac strongly suggested that this mark plays a role in dosage compensation in Drosophila.

With the advent of new technologies, high resolution binding patterns have been determined for a number of different marks. An initial genome-wide localization study using ChIP-chip investigated binding patterns of acetylation and methylation of histones H3 and H4 and phosphorylation of histone H3 at a number of annotated genes and non-repetitive sequences of chromosome 2L (4). The authors found a strong correlation between binding of antibodies to acetylated histones H3 and H4 and methylated H3 (dimethylation of K4 of H3 (H3K4Me2), trimethylation of K4 of H3 (H3K4Me3) and dimethylation of K79 of H3 (H3K79Me2)). A comparison of the localization patterns to expression profiling analysis indicated that these marks are enriched at transcribed genes while absent or present at low levels at the majority of repressed genes. Phosphorylated serine 10 of histone H3 (H3S10P) has a distinct genome-wide localization pattern compared to all other marks in that there is no positive correlation of the location of this mark with the acetylated marks. Furthermore, unlike the acetylation and methylation marks, it is enriched to a similar level at most genes, regardless of activity.

Subsequent studies have investigated additional modifications, often including comparison to localization of proteins known to interact with a specific mark or an enzyme that carries out a particular modification. Shwartz et al. (5) investigated the genome-wide localization pattern of trimethylated K27 of H3 (H3K27Me3) in cultured cells along with binding of components of the repressive polycomb group (PcG) complex, including Polycomb (Pc), Posterior sex combs (Psc), and Enhancer of zeste (E(z)), which methylates K27 of histone H3 (6-7). The PcG proteins localize to known targets containing Polycomb response elements (PREs) that are associated with regulation of homeotic (Hox) genes (see Figure 1). H3K27Me3 is also found at those sites. Interestingly, the PcG proteins exhibit specific peaks of protein binding while the methyl mark encompasses this peak and also spreads out from the peak in a broad domain. PcG proteins and H3K27Me3 are predominantly associated with inactive genes, but are found localized to some sites that include moderately active transcribed genes.

To determine if PcG-associated histone modifications change with transcription, a second study analyzed the binding patterns of H3K27Me3 along with H3K27Ac and H3K4Me3 in three different cultured cell lines (8). The authors report similar patterns of localization as to those found in their initial 2006 study. H3K27Me3 is found at repressed genes containing PREs. In cells where the genes are active, H3K27Me3 levels are very low. Activation of PRE-containing genes is accompanied by acetylation of H3K27 (H3K27Ac) upstream of the transcription start site as well as in the gene body and H3K4Me3 at the promoter. The data indicate that H3K27Me3 is a repressive mark while H3K27Ac and H3K4Me3 are active marks (see Figure 1).

The distribution of H3K27Me3 along with components of PcG proteins in embryos was also recently determined (9). The authors of that study additionally investigated the localization of H3K4Me3. The results from embryos correlate very well with the earlier data produced in cultured cells. PcG proteins bind to previously described PREs at generally silenced gene regulatory regions while
Histone modifying enzymes and gene regulation

Figure 1. The Hox gene cluster is repressed by PcG proteins and activated by trxG proteins. Major enzymes and histone modifications are indicated. A. Silencing by PcG proteins is associated with deacetylation of H3K27 by RPD3, methylation of this same residue by E(z), demethylation of H3K4 and K36 by dKDM2, addition of ubiquitin to H2A by dRING and removal of ubiquitin from H2A by Calypso and H2B by USP7. B. Activation by trxG proteins is linked to demethylation of H3K27 by dUTX, acetylation of H3K27 by dCBP and methylation of H3K4 by Ash1 and Trx. The enzymes that carry out the modifications are often found in multi-subunit complexes. The additional proteins may direct or titrate the enzymatic activity. All methylation and acetylation marks occur on histone H3 lysine residues as indicated. Ubiquitination occurs on histone H2A K119 and H2B K120. Green arrows indicate the addition of the mark and red arrows with scissors indicate the removal of the specified modification. Dashed lines and margins indicate disappearance of the mark.

H3K27Me3 spreads out from those sites. There are very few regions of overlap between H3K27Me3 and H3K4Me3. Although there is strong overlap between the data sets generated in cultured cells and in embryos, cell specific binding sites were identified, indicating PcG cell type specific gene regulation.

Trimethylation of K36 of H3 (H3K36Me3) has also been analyzed in cultured cells. This mark is found to be associated with gene activation, but rather than being found at the promoter like many other marks, it is enriched over the body of the gene and peaks near the 3’ end (10-11). In contrast, dimethylation of K36 of H3 (H3K36Me2) peaks downstream of the promoter region of genes (10). Interestingly, mutations in the enzymes that are responsible for the methylation marks, dMes-4 for dimethylation and dSet2/dHypb for trimethylation, differentially affect bulk H4K16Ac levels. Reduction of dMes-4 by RNAi results in a decrease of H4K16Ac while knock down of dSet2 results in an increase in histones containing that mark. The data indicate that H3K36 methylation is generally localized to the body of genes. Additionally, these results provide one example of cross talk between modifications, in this case between methylation and acetylation.

The genome-wide distribution of H4K16Ac has been determined independently by two research groups (12-13). Consistent with the salivary gland polytene chromosome staining, ChIP-chip analysis of cultured cell and larval samples indicate that this mark is enriched on the male X chromosome. In addition to the male X chromosome enrichment, H4K16Ac is also localized to genes along autosomes and the female X chromosome. Interestingly, the localization pattern along a gene is different depending on the chromosome. On the male X, there is a bimodal distribution of this mark with a peak at
Histone modifying enzymes and gene regulation

Figure 2. Large chromosomal domains are associated with specific histone modifications. A polytene chromosome spread stained with DAPI indicates differential packaging along the genome. DAPI bright areas represent condensed regions while DAPI light areas are less compacted chromatin. The male X chromosome is enriched in H3S10P, H3K36Me3 and H4K16Ac. H3K36Me3 promotes H4K16Ac along the X. The heterochromatic chromocenter (area within dashed circle) is highly condensed and associated with high levels of H3K9Me2,3. H3S10P and K3K4Me act to prevent H3K9Me.

the promoter and then an increase toward the 3’ end, mirroring the H3K36Me3 mark. On the other chromosomes, the peak of H4K16Ac is found at the promoter regions of genes. The localization pattern of the H4K16Ac mark relative to the enzyme that adds this mark is similar to that of the H3K27Me3 mark and PcG protein distribution. There is a broad domain of histones containing the H4K16Ac mark. Found within the broad domain of H4K16Ac is a peak of binding for the protein complex containing males absent on the first (MOF), the enzyme responsible for acetylation of histone H4 at lysine 16.

4. WHERE ARE THE MARKS RELATIVE TO LARGE CHROMATIN DOMAINS?

Taken together, the results of genome-wide ChIP studies as well as polytene chromosome staining indicate that some sites along the genome have peaks of specific histone modifications that are associated with genes and gene activity. In other cases, large domains are associated with specific marks. Two examples of regulation at the chromosomal level include 1) dosage compensation, in which the male X chromosome is up-regulated two fold relative to the autosomes, and 2) heterochromatin versus euchromatin, in which the genome is differentially packaged into constitutively condensed regions or more open and accessible domains along the chromosome arms (see Figure 2). As discussed above, an additional example of a regulated domain is that which is associated with polycomb-mediated silencing of the Hox gene cluster (see Figure 1). Specific histone modifications have been functionally linked to regulation of genes located in these distinct broad domains of chromatin.

4.1. Regulation of the male X chromosome

Dosage compensation is the process by which male and female flies equalize the expression of genes located on the X chromosome. Males up-regulate the genes on their single X chromosome to match the level of gene expression from the pair of X chromosomes of the female (14). The dosage compensation process is mediated by the dosage compensation complex (DCC) that is comprised of two non-coding RNAs, roX1 and roX2, and proteins male specific lethal (MSL) 1, MSL2, MSL3, maleless (MLE) and MOF (15-16). Three histone modifications that appear to be associated with the male X are H4K16Ac, H3K36Me3 and H3S10P. H4K16Ac was found to be a major chromatin mark associated with the up-regulation of gene expression along the entire male X chromosome (see Figure 2) (17). In this paper the authors showed that the binding pattern of MSL1 and MLE are coincident with H4K16Ac on polytene chromosomes. Furthermore, mutations in genes of the DCC result in reduction of H4K16Ac. This suggests that the DCC is involved in the enrichment of H4K16Ac on the X chromosome and that H4K16Ac is an important mark for dosage compensation. MOF is an acetyltransferase belonging to the MYST family and acetylates histone H4 at lysine 16 (18). MOF can
relieve transcription repression if targeted to a gene promoter, suggesting that H4K16Ac is important to bring about up-regulation of the male X chromosome (19). Up-regulated genes show higher levels of H4K16Ac in the body or 3' end as compared to the 5' end of genes (19). The acetyltransferase activity of MOF is stimulated by its incorporation into the DCC and its physical contact with MSL1 and MSL3 (20). These results indicate that the recruitment and function of MOF, and therefore histone acetylation, play a vital role in dosage compensation.

Another mark shown to play an important role in dosage compensation is H3K36Me3 (see Figure 2). There is a strong correlation of H3K36Me3 and MSL binding, both on the X chromosome and autosomes, when MSL complex components are ectopically expressed (11). A mutation in dSet2, a histone methyltransferase (HMT) that catalyzes the addition of the trimethyl mark, reduces MSL binding to its targets (11). Reduction of H3K36Me3 was also found to lead to reduced recruitment of MSL proteins, including MSL1 and the acetyltransferase MOF, and loss of dosage compensation of male X genes (21). Reduction of H3K36Me3 by mutating dSet2 also leads to changes in H4K16Ac levels, indicating that the two marks are functionally linked (21). The results suggest that the DCC recognizes and binds to H3K36Me3 along the male X chromosome. This binding is followed by H4K16 acetylation by MOF. The recognition is likely by MSL3, which has been shown to bind H3K36Me3 in vitro (11). Interestingly, the male X and autosomes respond differently to reduced H3K36Me3. H4K16Ac levels decrease across the male X while they increase at genes located on autosomes. These results suggest that the cross talk between the two marks is context dependent. These data also imply that maintenance of acetylation and methylation levels across the Drosophila X chromosome are required to bring about proper dosage compensation and regulate gene expression on a chromosome wide level.

One additional mark that is associated with possible regulation of the male X is H3S10P (see Figure 2). Polytenic chromosome staining revealed that while this mark is found along autosomes, it is enriched on the male X (22). Subsequent studies investigating mutations of JIL-1 kinase, one enzyme that phosphorylates H3 at serine 10, indicate that the reduction of JIL-1 affects the overall chromosomal structure of the male X chromosome (23). The X chromosome is shorter than in wild type animals and the euchromatic banding pattern is severely disrupted. This finding suggests that the H3S10P mark might be important in establishing a chromatin structure to facilitate the two fold up-regulation in gene expression critical for dosage compensation. In sum, the data indicate that epigenetic regulation has a crucial role in regulation of structure and gene expression of the entire male X chromosome.

4.2. Heterochromatin

A second example of a region of the genome that is regulated at the domain level is heterochromatin (see Figure 2 and 3). Genes that are moved from their normal chromatin context by transposition or chromosome rearrangement to the boundary of heterochromatin are often expressed in some cells and silenced in others, a phenomenon known as position effect variegation (PEV). PEV of the white (w) gene is the best characterized example of this phenomenon. Drosophila that have w located near a boundary have patches of red and white eye pigmentation depending on whether or not w is expressed, reviewed in (24). Enhancement or suppression of PEV is often used as a read-out of heterochromatin formation. Factors that enhance variegation are referred to as enhancers of variegation, E(var)s, and those that suppress as suppressors of variegation, Su(var)s.

One characteristic feature of heterochromatin is the enrichment of methylated H3K9 (see Figures 2 and 3). The histone methyl transferase (HMT) Su(var)3-9 was the first Drosophila enzyme found to add this mark to histones (25). Mutations in Su(var)3-9 result in reduced global levels of methylated H3K9, as determined by Western blot (Schotta 2002). Additionally, the mutants show reduced H3K9 localization to pericentric heterochromatin along polytenic chromosomes. The enzymatic activity is regulated by dimerization of Su(var)3-9 via its N-terminal region (26). Subsequently, two additional factors, dG9a and dSETDB1, were found to have H3K9 methyltransferase activity (27-31). Research from Mis et al. (28) found that, similar to Su(var)3-9, a mutation in dG9a acts as a dominant suppressor of PEV, suggesting that both enzymes are important for the formation and/or maintenance of heterochromatin. More recent work on dG9a suggests that its main area of activity is along euchromatin (see section 5.2 below on methylation for more details) (27, 30). An analysis of the three enzymes was carried out by Brower-Toland et al. (32). These authors find that the concerted action of these enzymes helps establish and sustain methyl marks at heterochromatin. Yet, they play different roles in this process. It was determined that pericentric heterochromatin is established and maintained mainly by the actions of Su(var)3-9 and dSETDB1, while heterochromatin at the 4th chromosome requires essentially dG9a and dSETDB1. During early development, heterochromatin is maintained by the action of dSETDB1 (32). dSETDB1 also maintains heterochromatin during metamorphosis while Su(var)3-9 functions during late oogenesis and immediately after fertilization (33). Su(var)3-9 mutants are viable while dSETDB1 ubiquitous knockout flies exhibit decreased viability (32). Interestingly, loss of both Su(var)3-9 and dSETDB1 results in an increase in viability as compared to the single mutants. Taken together, the results indicate that multiple H3K9 methyltransferases have distinct functions in heterochromatin formation and maintenance.

A second defining feature of heterochromatin is the binding of heterochromatin protein 1 (HP1), perhaps the best characterized heterochromatin associated protein (see Figure 3). It binds to histone H3 methylated at K9 and acts to recruit Su(var)3-9 to heterochromatin (34). Methylation of histones in turn leads to recruitment of
Histone modifying enzymes and gene regulation

Figure 3. Heterochromatin is formed by the concerted action of distinct histone modifying enzymes. The major enzymes and their location of action are indicated. Green arrows indicate the addition of the mark and red arrows with scissors indicate the removal of the specified modification. Dashed arc indicates an indirect effect. Dashed lines and margins indicate disappearance of the mark.

Additional HP1, facilitating the spread of heterochromatin (34). HP1 was initially identified as a non-histone protein that is associated with chromatin in Drosophila embryos (35). It has a chromo and a chromo shadow domain separated by a hinge (36-37). The chromo domain allows for the recognition of methylated histones (38). The chromo shadow domain may be important to mediate self-association and/or interactions with other proteins (36).

Loss of Su(var)3-9 and HP1 leads to defective chromatin condensation of heterochromatic regions of the X chromosome, resulting in the formation of a pseudopuff (39). The HMT activity of Su(var)3-9 and binding of HP1 are thus important for heterochromatin structure.

Additional actions are important in the formation of heterochromatin, including the active removal of H3K4 methylation and incorporation of the histone variant H2Av (40-41). A mutation in Su(var)3-3, an H3K4 mono- and di-methylase, results in spreading of H3K4Me2 into the prospective heterochromatin compartment of blastoderm cells during embryogenesis (see Figure 3) (40). This in turn prevents H3K9 from being dimethylated, implying that removal of H3K4Me is a prerequisite for the addition of methyl groups to H3K9. Once established, heterochromatin is maintained by its interaction with HP1 via H3K9Me2 (40). A second mechanism implicated in the formation of heterochromatin is the incorporation of the histone variant H2Av into chromatin (41). H2Av mutants show a reduction in H4K12Ac and H3K9Me2 levels as well as a reduction in HP1 recruitment to the centromeric region. Furthermore, mutation in Su(var)3-9 does not affect deposition of H2Av. This suggests that the role of H2Av is upstream of the function of Su(var)3-9.

In contrast to the role of H2Av, H3S10P antagonizes H3K9Me2 and thus prevents the formation and subsequent spreading of heterochromatin (see Figure 3) (42). It was found that loss of JIL-1 kinase results in spreading of H3K9Me2 from the heterochromatic chromocenter to neighboring euchromatic loci that are not normally methylated in wild type chromatin (42). As mentioned above, H3S10P is enriched on the male X chromosome (22). Interestingly, male and female X chromosomes are particularly hypermethylated at H3K9 in the JIL-1 mutants (42). Additionally, loss of JIL-1 results in ectopic HP1 binding on polytene chromosomes in a pattern similar to that of H3K9Me2 in the mutants. The lethality due to loss of JIL-1 is counteracted by loss of function alleles of Su(var)3-9 HMT (43). Furthermore, the researchers found that, unlike loss of the HMT, loss of HP1 is unable to rescue lethality due to mutations in JIL-1. This suggests that the lethality caused by mutation of JIL-1 is not due to effects of loss or relocalization of HP1, but instead, is mainly due to perturbations in chromatin structure. These data indicate that while H3K9Me2, Su(var)3-9 and HP1 mark heterochromatin, other marks including H3S10P function to prevent spreading of heterochromatin into euchromatic regions. This combined action maintains a balance and ensures proper gene regulation via chromatin organization.
Histone modifying enzymes and gene regulation

In addition to the noted effects on PEV, loss of H3K9Me results in an increased frequency of DNA damage in heterochromatin (44). These observations suggest that the enzymes responsible for catalyzing the methylation of histone H3 at K9 also function in maintaining the genomic integrity of heterochromatin. The authors noted that similar increases in DNA damage and chromosomal defects in heterochromatin were observed in flies carrying mutations in the RNAi pathway. These findings suggest that, apart from the factors discussed above, the RNAi pathway is involved in maintaining heterochromatin. The role of RNA as an epigenetic regulator is outside the scope of this current work, interested readers may refer to (45).

The research on mechanisms of establishment and maintenance of histone modifications associated with dosage compensation, along with the heterochromatin structure-function studies, indicates that histone marks are critical for regulation of genes located within these regions (see Figure 2). H4K16Ac, H3K36Me3 and H3S10P are all important to allow for hyperactivation of genes along the male X chromosome. Methylation of K9 of histone H3 is critical for repression of genes located in heterochromatin. The generation of a specific chromatin landscape of marks that encompass large regions of DNA is thus important for epigenetic regulation.

5. LOCUS SPECIFIC ACTION OF HISTONE MARKS

So far we have focused on the role of histone modifications in the regulation of broad chromosomal domains, but what about the role of these marks in activation and repression of specific genes? Analysis of individual enzymes has provided much insight into the specificity of action and overall function of these important proteins (see Table 1).

5.1. Acetylation

The Drosophila genome-wide localization data supports classic studies indicating that acetylation is generally associated with transcription activation (46). Additional research using mutants in the enzymes that add or remove acetyl groups has provided much information regarding the role of histone acetylation in transcription and how this might then affect overall development. While histone acetylation at gene promoters is generally associated with transcription activation and deacetylation with transcription repression, there are examples of histone deacetylases (HDACs) involved in transcription activation and histone lysine acetyltransferases (KATs) in transcription repression. The examples that follow show that the combined action of HDACs and KATs create and maintain acetylation levels that facilitate chromatin packaging and/or recruitment of transcription factors essential for proper gene expression.

Six different HDACs are present in Drosophila: RPD3, HDAC3, HDAC4, HDAC6, HDACX and SIR2 (47-48). HDACs from yeast to mammals have been grouped into four distinct classes based on sequence similarity (49-50). Class I HDACs RPD3 and HDAC3 are similar to yeast Rpd3. Drosophila class II enzymes HDAC4 and HDAC6 resemble yeast Hda-1. SIR2 represents class III enzymes, which are distinct from the other classes in that they require nicotinamide adenine dinucleotide (NAD) as a cofactor. HDACX is most similar to mammalian HDAC11, the sole class IV enzyme identified to date. An early study indicated that RPD3, HDAC3, HDAC6 and SIR2 are able to deacetylate histone H4 as determined using an in vitro HDAC assay (47). Overexpression of RPD3, HDAC3, HDAC4, HDAC6 and SIR2 leads to misregulation of genes involved in a variety of processes (51). A systematic knockdown of individual HDACs in S2 cultured cells indicates that HDACs 1 and 3 are predominant in regulation of transcription (52). In the overexpression analysis, in addition to the predicted down-regulation of expression, genes were also observed to be up-regulated. Overexpression of RPD3 resulted in more genes being up-regulated than down-regulated, while overexpression of HDAC4, HDAC6 and SIR2 resulted in the majority of genes being down-regulated. Although it cannot be ruled out that the up-regulation observed upon overexpression of HDACs is an indirect effect, these results suggest that HDACs are not associated solely with transcription repression.

The Drosophila class I HDAC, RPD3, and an associated corepressor, SIN3, have been primarily linked to transcription repression. These factors are present throughout euchromatin on Drosophila polytene chromosomes but fail to overlap with acetylated histones or RNA polymerase II, marks of active transcription, suggesting that histone deacetylation results in transcription repression (53). Loss of RPD3 by RNAi results in increased global levels of acetylation of histone H3 and H4 (54). RPD3 knockdown leads to misregulation of genes involved in many processes, particularly nucleotide and lipid metabolism, DNA replication, cell cycle regulation and signal transduction (52). Loss of the corepressor SIN3 by RNAi in cultured cells leads to misregulation of about 3% of the genome (55). While most genes were up-regulated, some were also down-regulated upon loss of SIN3. Comparison of the studies conducted by the two groups indicates that there is a significant but not complete overlap in genes both up- and down-regulated by RPD3 and SIN3. These results imply that while RPD3 and SIN3 function together to regulate transcription, they may also have functions independent of each other. As to the question of which histone residues are targets of the SIN3/RPD3 HDAC complex, overexpression of one of three SIN3 isoforms (SIN3 187) in cultured cells results in decreased global levels of H3K9Ac and H3K14Ac, suggesting that these are the major target residues of the complex (56).

RPD3 has been implicated in regulating genes involved in development and is essential for viability (57). Inhibition of the deacetylase activity of class I and class II enzymes including RPD3 by Trichostatin A (TSA) results in lethality in flies (58). Additionally, RPD3 regulates genes involved in embryo segmentation through its association with the transcriptional corepressor groucho.
Histone modifying enzymes and gene regulation

Table 1. Histone modifying enzyme activity

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Activity</th>
<th>General Role</th>
<th>Cross Talk</th>
<th>Human Homolog</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deacetylases</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RPD3</td>
<td>H3K9, H3K14, H3K27, H4</td>
<td>Repression</td>
<td></td>
<td>HDAC1/HDAC2</td>
</tr>
<tr>
<td>HDAC3</td>
<td>H4</td>
<td>Repression</td>
<td></td>
<td>HDAC3</td>
</tr>
<tr>
<td>HDAC4</td>
<td></td>
<td>Repression</td>
<td></td>
<td>HDAC4</td>
</tr>
<tr>
<td>HDAC6</td>
<td>H4</td>
<td>Repression</td>
<td></td>
<td>HDAC6</td>
</tr>
<tr>
<td>HDACX</td>
<td></td>
<td>Repression</td>
<td></td>
<td>HDAC11</td>
</tr>
<tr>
<td>SIR2</td>
<td>H3K56, H4</td>
<td>Repression</td>
<td></td>
<td>SIR1</td>
</tr>
<tr>
<td>Acetyltransferases</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dGCN5</td>
<td>H3K9, H3K14, H4K12</td>
<td>Activation</td>
<td>H3S10P</td>
<td>KAT2A</td>
</tr>
<tr>
<td>Tip60</td>
<td></td>
<td>Repression</td>
<td></td>
<td>KAT5</td>
</tr>
<tr>
<td>Chm</td>
<td>H4</td>
<td>Activation and repression</td>
<td></td>
<td>MYST2</td>
</tr>
<tr>
<td>MOF</td>
<td>H4K16</td>
<td>Activation, dosage compensation</td>
<td></td>
<td>H3K36Me3</td>
</tr>
<tr>
<td>dCBP</td>
<td>H3K18, H3K27, H3K56</td>
<td>Activation</td>
<td></td>
<td>CREBBP</td>
</tr>
<tr>
<td>TAF1</td>
<td>H3 and H4</td>
<td>Activation</td>
<td></td>
<td>TAF1</td>
</tr>
<tr>
<td>Lysine Methylases</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trx</td>
<td>H3K4</td>
<td>Activation</td>
<td></td>
<td>MLL</td>
</tr>
<tr>
<td>Ash1</td>
<td>H3K4, H3K9, H3K36, H4K20</td>
<td>Activation</td>
<td></td>
<td>ASH1L</td>
</tr>
<tr>
<td>Ttr</td>
<td>H3K4Me3</td>
<td>Activation</td>
<td></td>
<td>MLL2/MLL3</td>
</tr>
<tr>
<td>dSet2/dHybp</td>
<td>H3K36Me3</td>
<td>Activation, dosage compensation</td>
<td>H4K16Ac</td>
<td>SETD2</td>
</tr>
<tr>
<td>dMes-4</td>
<td>H3K36Me2, Me3</td>
<td>Activation</td>
<td></td>
<td>NSD1</td>
</tr>
<tr>
<td>dMot1/Grappa</td>
<td>H3K36Me1, Me2</td>
<td>Activation</td>
<td></td>
<td>DOT1L</td>
</tr>
<tr>
<td>Sat(var)3-9</td>
<td>H3K9Me2</td>
<td>Heterochromatin</td>
<td></td>
<td>H3K4Me1, Me2</td>
</tr>
<tr>
<td>dSetDB1</td>
<td>H3K9</td>
<td>Heterochromatin</td>
<td></td>
<td>SETDB1/SETDB2</td>
</tr>
<tr>
<td>d9a</td>
<td>H3K9, H3K27</td>
<td>Repression</td>
<td></td>
<td>EHM1</td>
</tr>
<tr>
<td>E(z)</td>
<td>H3K27Me1, Me2, Me3</td>
<td>Repression</td>
<td></td>
<td>EZH1/EZH2</td>
</tr>
<tr>
<td>PR-Set7</td>
<td>H4K20Me1</td>
<td></td>
<td></td>
<td>SETD3</td>
</tr>
<tr>
<td>Sun4-20</td>
<td>H4K20Me2, Me3</td>
<td>Repression</td>
<td></td>
<td>SUV420H1/SUV420H2</td>
</tr>
<tr>
<td>Lysine Demethylases</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sat(var)3-3</td>
<td>H3K4Me1, Me2</td>
<td>Heterochromatin</td>
<td></td>
<td>LSD1</td>
</tr>
<tr>
<td>Ld</td>
<td>H3K4Me3</td>
<td>Activation and repression</td>
<td></td>
<td>KDM5A</td>
</tr>
<tr>
<td>dKDM4A</td>
<td>H3K9Me1, H3K36Me3</td>
<td>Repression</td>
<td></td>
<td>KDM4A</td>
</tr>
<tr>
<td>dKDM2</td>
<td>H3K4Me, H3K36Me2</td>
<td>Repression</td>
<td></td>
<td>KDM2B</td>
</tr>
<tr>
<td>dTX</td>
<td>H3K27Me2, Me3</td>
<td>Activation</td>
<td></td>
<td>KDM6A</td>
</tr>
<tr>
<td>Arginine Methylases</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DART1</td>
<td>H4R3Me2</td>
<td>Repression</td>
<td></td>
<td>PRMT1</td>
</tr>
<tr>
<td>DART4/CARMER</td>
<td>H3R17Me2</td>
<td>Activation</td>
<td></td>
<td>CARM1</td>
</tr>
<tr>
<td>Ubiquitinases</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dRING</td>
<td>uH2A</td>
<td>Repression</td>
<td></td>
<td>RING1</td>
</tr>
<tr>
<td>dbrec</td>
<td>uH2B</td>
<td>Activation</td>
<td>H3K4Me3</td>
<td>RNF40</td>
</tr>
<tr>
<td>TAF1</td>
<td>uH1</td>
<td>Activation</td>
<td></td>
<td>TAF1</td>
</tr>
<tr>
<td>Deubiquitinases</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calypso</td>
<td>uH2A</td>
<td>Repression</td>
<td></td>
<td>BAP1</td>
</tr>
<tr>
<td>Scrawny</td>
<td>uH2B</td>
<td>Repression</td>
<td>H3K4Me3</td>
<td>USP36</td>
</tr>
<tr>
<td>USP7</td>
<td>uH2B</td>
<td>Repression</td>
<td></td>
<td>USP7</td>
</tr>
<tr>
<td>Nonstop</td>
<td>uH2A, uH2B</td>
<td>Activation</td>
<td></td>
<td>USP22</td>
</tr>
<tr>
<td>Kinases</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JIL-1</td>
<td>H3S10</td>
<td>Activation</td>
<td></td>
<td>RPS6K5</td>
</tr>
<tr>
<td>TAF1</td>
<td>H2BS3</td>
<td>Activation</td>
<td></td>
<td>TAF1</td>
</tr>
</tbody>
</table>

1 Histone or residue if known 2 Interplay between marks 3 as given in databases FLIGHT and Flybase (217) (218)

(GRO) (59-60). Recruitment of RPD3 by GRO is also important for the repression of genes involved in the Wnt pathway in wing discs (61). In addition, genetic interactions with tramtrack69 suggest that RPD3 plays a role in wing development (62). In sum, these data suggest that RPD3 is involved in regulating many genes involved in development.

In addition to gene specific effects, RPD3 plays a role in regulating large domains, including the PcG group regulated region and heterochromatin. RPD3 has been implicated in long term silencing of Hox genes by PcG proteins (63). RPD3 is present in a PcG complex consisting of extra sexcombs (ESC) and E(z) among other proteins (see Figure 1). This complex directly binds to the PRE and recruits the polycomb repressive complex (PRC1) to bring about repression of Ultrabithorax (Ubx) during embryogenesis and imaginal disc development (63). RPD3 also genetically interacts with PcG proteins to affect Hox gene regulation (64). The authors found that a mutation in RPD3 enhances the formation of ectopic sex comb teeth on the second and third pair of legs in males. They also found that a deficiency in RPD3 is able to suppress homeotic transformation of the mesothorax commonly found in mutants of Trithorax group (trxG) proteins, which function in opposition to PcG proteins. These results indicate that RPD3 is important in bringing about transcription repression. Interestingly, the role of RPD3 in PcG silencing is deacetylation of histone H3 at K27, which must occur for methylation of this same residue by E(z) to take place (65). Loss of RPD3 also results in a suppression of silencing of genes that are subject to PEV (57, 66). In this context, RPD3 likely functions to deacetylate K9 of histone H3 prior to subsequent methylation by Su(var)3-9. In fact, RPD3 has been found to interact biochemically and genetically with the heterochromatin associated HMT
Histone modifying enzymes and gene regulation

Su(var)3-9 (67). These results suggest that RPD3 is also a regulator of chromatin structure.

HDAC3 is the other class I HDAC in *Drosophila* (68). It regulates genes that are both common to and distinct from RPD3, suggesting that it has unique functions in controlling transcription (52). Like RPD3, HDAC3 is also essential for *Drosophila* viability and acts as a suppressor of PEV, implying a function in heterochromatin regulation (69). Loss of HDAC3, however, does not result in global changes in histone acetylation (52). Studies in *Drosophila* cultured cells show that HDAC3 regulates genes involved in glycolysis, nucleic acid metabolism, cell growth and maintenance, cell-cell signaling, cell differentiation and metamorphosis (51). In imaginal discs, HDAC3 regulates growth through control of apoptosis whereas RPD3 controls cell proliferation (69). The HDAC activity of HDAC3 has specifically been linked to transcription repression (70). This study shows that Ebi, a member of the HDAC3/SMRTR corepressor complex, is recruited by the Snail transcriptional repressor to target genes, resulting in repression coincident with histone hypoacetylation. HDAC3 and Ebi have also been found to be important in repression of genes involved in the development of adult muscle through their interaction with SET and MYND domain protein 4 (SMYD4) (71). Thus, although HDAC3 and RPD3 belong to the same class of proteins and regulate many similar processes, they regulate unique sets of genes.

Fewer studies have been carried out on the class II and IV enzymes. Class II HDACs in *Drosophila* are represented by HDAC4 and HDAC6. Loss of HDAC4 by RNAi results in segmentation defects in embryos (72). Overexpression of HDAC4 results in misregulation of genes involved in protein biosynthesis, phosphorus metabolism, cell motility, signal transduction and development (51). Unlike class I HDACs, loss of HDAC4 or HDAC6 by RNAi does not result in significant changes in gene expression (52). Interestingly, loss of HDAC6 results in an increase in tubulin acetylation, a phenomenon observed upon the loss of its human ortholog as well (52, 73). In another study, loss of HDAC6 results in loss of dopaminergic neurons, retinal degeneration and locomotor dysfunction (74). All three phenotypes are associated with models of Parkinson’s disease (PD). Additionally, HDAC6 was found to associate with ectopically expressed alpha-synuclein, a key protein involved in the *Drosophila* model of PD. These studies suggest that HDAC6, unlike class I HDACs, is perhaps not involved in global gene regulation but rather regulates neuronal development mainly through its role in tubulin acetylation and protein:protein interactions. Knockdown of class IV HDACX by RNAi results in little to no change in global histone H3 acetylation, cell proliferation or gene expression as measured by microarray expression profiling (52). Additional work on these proteins is anticipated to illuminate discriminating functions of these multiple HDAC enzymes.

Finally, we look at the NAD+ dependent class III enzyme SIR2 (75). SIR2 has been implicated in repression of genes located in both euchromatic and heterochromatic chromosomal regions (76). SIR2 interacts genetically and biochemically with members of the Hairy/Deadpan/Enhancer of split (E(Spl)) family of basic helix-loop-helix euchromatic repressors (76). Mutations in SIR2 also affect PEV, implicating SIR2 in the regulation of heterochromatin structure (76-77). Furthermore, SIR2 facilitates the formation of a nuclease-resistant fast-sedimenting histone-DNA complex (78). This data lead the authors to propose a model in which SIR2 brings about transcription repression by introducing conformational changes in chromatin. SIR2 has been directly linked to deacetylation of H3K56 (79). This particular mark is principally associated with chromatin assembly during DNA replication and repair, but also likely has a role in nucleosome assembly and disassembly during transcription activation and repression (80-81). SIR2 has also been associated with caloric restriction and its effect on life span. Overexpression of SIR2 increases life span while a decrease in SIR2 counteracts the effect of caloric restriction on longevity (82-83). The data indicate that SIR2 can regulate transcription activity by modulating expression of genes located in euchromatin or by silencing genes in heterochromatin. Additionally, it is involved in regulating life span.

Just as multiple HDACs regulate deacetylation, histone acetylation is performed by multiple KAT proteins. KATs oppose HDACs and acetylation is generally associated with transcription activation. Four different categories of KATs are present in *Drosophila*: GNAT (GCN5-related N-terminal acetyltransferase), MYST domain, CREB binding protein (CBP)/p300 and TBP associated factor 1 (TAF1).

The GNAT homolog dGCN5 is an essential gene that has been implicated in various processes such as metamorphosis, neural, eye, leg and wing development (84-86). Initial investigations of dGCN5 mutants revealed strong reductions in global levels of H3K9Ac and H3K14Ac while H4K8Ac and H4K16Ac were essentially unchanged, indicating that dGCN5 is a major histone H3 KAT (85). dGCN5 is present in at least two multi-subunit complexes including *Drosophila* Spt-Ada-Gcn5-Acetyltransferase (dSAGA) complex and Ada two A containing (ATAC) complex, which contain common proteins dGCN5 and adenosine deaminase (ADA) 3 (87-88). Each complex also contains one of two distinct ADA2 proteins, ADA2a in ATAC and ADA2b in dSAGA. Two additional proteins found to be important components of dSAGA include AdaI and will decrease acetylation (WDA) (89).

Analysis of individual complex components has contributed to our understanding of the function of dGCN5, since its incorporation into these complexes appears to affect dGCN5 activity. Like dGCN5, ADA2a, ADA2b, ADA3 and WDA are essential for viability, implying that critical functions are performed by KAT complexes (85, 89-92). Using an *in vitro* KAT assay, Guelman et al. (89) demonstrated that dSAGA can acetylate either H3 or H4 as core histones or as nucleosomes, but noted that the activity...
Histone modifying enzymes and gene regulation

toward histone H3 is stronger. The authors also found that mutants in SAGA-specific WDA have reduced histone H3 acetylation as embryos with no change in histone H4 acetylation. Mutations in ADA2b, also part of dSAGA, lead to reduced H3K9Ac and H3K14Ac levels in embryos and along polytene chromosomes with no changes to histone H4 acetylation levels (90, 92). Taken together, the data suggests that dGCN5, as part of the dSAGA complex, functions to acetylate histone H3, specifically at K9 and K14. Acetylation by the dSAGA complex likely functions in multiple developmental pathways as mutants in ADA2b and Nipped-A, another dSAGA component, have been shown to be important for Notch signaling, eye pigmentation and p53-dependent apoptosis (90, 92-93).

Analysis of ADA2a has yielded information as to the activity of dGCN5 in the ATAC complex. Unlike ADA2b mutants, mutations in ADA2a do not result in altered levels of H3K9Ac and H3K14Ac on polytene chromosomes (92). The authors of that study also demonstrated that H4K8Ac and H4K12Ac are not affected by loss of either ADA2 protein. H4K12, however, has been found to be a target of ATAC (94). Global levels of H4K12Ac are reduced in ADA2a and dGCN5 mutants. In addition, Ciuciu et al (94) found that ATAC mutants have reduced levels of H3S10P, indicating that there is cross talk between acetylation and phosphorylation (see Figure 4). Interestingly, mutants in ADA3, a component of dSAGA and ATAC, have reduced levels of the dSAGA-associated marks H3K9Ac and H3K14Ac and the ATAC-associated marks H4K12Ac and H3S10P along polytene chromosomes, indicating that these residues are dGCN5-containing KAT complex targets (91).

ATAC has been linked to multiple gene regulatory functions. Loss of dGCN5, ADA2a and ADA3 has been shown to affect metamorphosis (85, 95). One strong possibility is that ATAC is responsible for regulating genes involved in metamorphosis by activating them in response to ecdysone. Interestingly, genes involved in the biosynthesis of ecdysone are regulated by the ATAC complex (95). Mutations in ADA2a and ADA3 result in decreased ecdysone levels and, in the ADA3 mutants, ecdysone receptor (EcR) binding along polytene chromosomes is reduced. ATAC has also recently been found to be an important factor in the response to mitogen-activated protein (MAP) kinase signaling and regulation of c-Jun N-terminal kinase (JNK) target genes (96). ATAC functions to regulate JNK activity, which in turns affects transcription of target genes. ATAC is thus critical for development and response to critical signaling pathways.

ATAC has also been linked to the activity of another chromatin modifying complex, nucleosome remodeling factor (NURF). X chromosome morphology is disrupted in dGCN5 and ADA2a mutants (97). The X chromosome takes on a shortened and bloated appearance, similar to what is seen in NURF complex mutants (98). The authors also identified genetic interactions between ATAC and NURF components and expression profiling analysis indicated that many genes are regulated by both complexes (97). Additionally, loss of the ISWI subunit of NURF results in reduced ADA2a binding to polytene chromosomes. Interestingly, imitation switch (ISWI) is a non-histone target of acetylation by dGCN5 (99). ATAC and NURF thus appear to be important co-regulators of gene transcription and the evidence of shared targets links histone modification and chromatin remodeling.

There are multiple MYST domain-containing proteins in *Drosophila* including Tip60, Chameau (Chm) and MOF. Tip60 is a *Drosophila* MYST domain KAT important for embryonic development (100). It is part of a multi-subunit complex (101). Mutations in components of the Tip60 complex including Enhancer of Polycomb (E(pc)), Domino, and MRG15 are able to suppress PEV (102). These results imply that the Tip60 complex has a role in the formation of heterochromatin. Tip60 knockdown in wing discs results in increased cell death and venation defects in the adult wings (103). Also as part of that study, genome-wide expression analysis found that Tip60 is involved in both activating and repressing transcription. Repression by the Tip60 complex is through a direct mechanism as Tip60 is found to ChIP on the promoters of genes repressed by the complex. Additionally, there is an overlap between the genes up-regulated upon loss of Tip60 and those up-regulated following treatment with the histone deacetylase inhibitor TSA. This finding suggests that Tip60 regulates gene expression by facilitating the recruitment of HDAC complexes, further supporting the idea that HDACs and KATs work in combination to tightly control acetylation levels for gene expression.

Another MYST domain KAT is Chm. Interestingly, rather than playing a role in activation, mutational analysis suggests that Chm is important for pericentric heterochromatin repression as well as Hox gene silencing by PcG proteins (104). Furthermore, the acetyltransferase activity of Chm is required for assembly of telomeric heterochromatin. Chm is also believed to have a role in transcription activation as it functions antagonistically to the HDAC RPD3 in regulating expression of genes involved in the JNK pathway during *Drosophila* metamorphosis (105). Chm promotes the tetra acetylation of histone H4 at the promoter of activator protein 1 (AP-1) responsive sequences upon transcription activation. This acetylation is reversed by RPD3 upon JNK signal inactivation. These results indicate that acetylation of histones is dynamic and that the combined role of KATs and HDACs is important for regulating gene expression critical for signaling.

As presented earlier, MOF is also a MYST family KAT. In addition to its role in the DCC, MOF is an important component of a second multi-subunit complex, non-specific lethal (NSL) (106-107). This complex contains NSL1, NSL2, NSL3, MCRS2, MBD-R2 and will die slowly (WDS) along with MOF. Complex components bind chromatin independent of MOF but MOF binding is affected by loss of MCRS2 (107). Genome-wide expression profiling revealed that this complex regulates over 4000 genes (107). Thus, apart from its role in dosage compensation, MOF is a major factor in transcriptional regulation of autosomal genes.
Histone modifying enzymes and gene regulation

The Drosophila homolog of the coactivator CBP is encoded by nejire (nej), representing a third category of KATs (108). It plays a role in early embryogenesis by facilitating the transcription of twist in response to Dorsal (109). Additionally, it is involved in signaling pathways like Hedgehog, Wnt/Wingless, Notch and decapentaplegic (Dpp) (110). dCBP is important for the expression of cubitus interruptus, a gene that activates hedgehog signaling, thus linking dCBP to pattern formation in embryos (110). It has also been shown to be involved in repression via its activity on non-histone proteins (111). dCBP binds TCF, a component of the Wnt signaling pathway, and acetylates a conserved Armadillo (Arm) binding domain, thus lowering the affinity of TCF for Arm (111). This reduced interaction results in down-regulation of the Wnt pathway (111). Furthermore, mutations in nej can suppress the dominant Ubx phenotype suggesting that dCBP is involved in Hox gene function (112). Embryos mutated in nej show defects similar to those seen in mutants that lack the extracellular signal Dpp or its effector Mothers against dpp (Mad) (113). dCBP acts as a coactivator during Dpp signaling through its recruitment to genes by Mad to affect the transcriptional activation of Dpp-responsive genes during development (113). Association of dCBP with Trithorax (Trx) results in activation of genes and antagonizes silencing by PcG proteins (see Figure 1) (65). In this study, the authors show that dCBP specifically acetylates H3K27, a site that is trimethylated by the PRC2 complex to bring about repression. Overexpression of dCBP leads to an increase in acetylation and a corresponding decrease in H3K27Me3. Interestingly, deacetylation of H3K27 is brought about by RPD3. These data suggest interplay between different transcription factors to affect gene expression.

In addition to H3K27Ac, dCBP is implicated in the regulation of acetylation of histone H3K56 (79). As stated above, while this mark does have a role in transcription, it is mainly associated with chromatin activities of DNA replication and repair (80-81). The H3K56 directed HAT activity is regulated through association with the histone chaperone Asf1 (Das et al. 2009). The specificity of dCBP enzymatic activity is thus affected by associated proteins, similar to the situation with dGCN5. These findings tell us that for a full understanding of epigenetic regulation, we need to understand not only the enzymes and their marks, but also the composition of protein complexes containing the enzymes.

TAF1, which represents the final category of KATs, is a component of the TFIIID complex important for transcription initiation (114). TAF1 is essential for ovary, eye, ocelli, wing, bristle and terminalia development as well as overall growth of the fly (115). Interestingly, the authors determined that mutations that affect these developmental phenotypes do not affect the KAT activity of TAF1. This was determined using an in vitro KAT activity assay testing recombinant TAF1 polypeptides, comparing the mutants to wild type. This finding suggests that the phenotypes of the mutants are due to some other

Figure 4. Examples of positive cross talk between histone modifications. A. Acetylation of H3K12 by dGCN5 in the ATAC KAT complex leads to phosphorylation of H3S10 by JIL-1 kinase. B. Monoubiquitination of H2B by dBre1 leads to methylation of H3K79 by dDot1. C. Monoubiquitination of H2B by dBre1 leads to increased methylation of H3K4 by an unidentified HMT. Deubiquitination of H2B by Scrawny leads to a decrease in methylation of H3K4.
component of TAF1 function and not directly related to the histone acetylation activity of TAF1. Although mainly associated with acetylation, TAF1 is thought to be necessary for the phosphorylation of serine 33 on histone H2B (116). This modification coincides with the activation of string and giant (116). TAF1 is also directly involved in monoubiquitinating histone H1 (117). Mutations in the histone specific ubiquitin-activating/conjugating domain of TAF1 result in the activation of genes that are targets of the maternal activator Dorsal. These data suggest that TAF1 can regulate transcription either via its role as a general transcription factor or as a histone modifier.

The early studies on histone modification show that RNA polymerase II activity is associated with histones enriched in acetylated lysine residues (46). Many years and experiments later, research from multiple laboratories has provided molecular data to support these early findings. Yet, there is a growing body of evidence linking histone acetylation to transcription repression and deacetylated histones to activation. In fact, in studies performed in mammalian and Drosophila cultured cells, HATs and HDACs have been found to colocalize at multiple chromosomal sites (118-119). How to explain this finding? One possibility is that the enzymatic activity of the factors is affected by the proteins with which they associate. In this case, the enzyme may be recruited, but not enzymatically active, or have activity to a non-histone substrate. A second possibility is that histone acetylation and deacetylation is very dynamic. Perhaps a promoter or regulatory region must be reset for each round of transcription activation by RNA polymerase II (119). Further mechanistic studies, supported by single cell analyses, should shed light on the dynamic role of acetylation in control of epigenetic regulation of gene expression.

5.2. Methylation

Unlike acetylation, which is most often associated with gene activation, methylation of histones is associated with both transcription activation and repression. Methylation may occur at lysine and arginine residues. The specific residue modified and the number of methyl groups attached determines the transcriptional state of the gene (120). The enzymes that catalyze these methylation events fall under three major categories: the protein arginine methyltransferases (PRMTs) family, the SET (Su(var)3-9, E(z) and Trx) domain-containing proteins and the non-SET domain-containing Dot1 protein, both of which catalyze lysine methylation (120).

Lysine methylation has been the most characterized. Five lysine residues in histone H3 and a single residue in histone H4 have been found to be methylated. Among these, methylation of H3K4, K36 and K79 is generally associated with active transcription whereas methylation of H3K9, K27 and H4K20 is associated with repression (120-121). All but one lysine HMT are known to utilize a SET domain for the transfer of one or more methyl groups to lysine residues from the cofactor S-adenosyl-L-methionine (122). The SET domain and its flanking regions both appear to play an important role in the catalytic function. Isolation and characterization of mammalian Lysine-specific demethylase 1 (LSD1) as a histone demethylase changed the long standing belief that methylation is a stable, irreversible mark (123). Subsequent research led to the identification of multiple histone lysine demethylases (KDMs), which fall into two major classes (124-125). One class is comprised of an amino-oxidase-domain-containing demethylase, which requires flavin adenine dinucleotide (FAD) as a cofactor. LSD1, the only member identified thus far, is capable of demethylating mono- or dimethyl marks. The second family is comprised of the Jumonji (Jmj) C-domain-containing demethylases, which utilize Fe²⁺, O₂, and alpha-ketoglutarate as cofactors. These enzymes can demethylate mono-, di- and trimethyl marks on histone H3K4, K9, K27 and K36 residues. In this section, we will focus on each of the lysine residues methylated and discuss the enzymes that add the methyl marks, followed by the demethylases that remove specific marks. First, we will discuss the methylation marks that are generally associated with active transcription.

Methylation of H3K4 in Drosophila is catalyzed by several enzymes. To date, Trx, absent, small and homeotic discs 1 (Ash1), and Trithorax-related (Trr) have been identified as H3K4 methyltransferases. Little imaginal discs (LID) and Su(var)3-3, the Drosophila homolog of LSD1, are H3K4 specific demethylases. Methylation of H3K4 is generally associated with active transcription. Trx and Ash1 are HMTs belonging to the trxG proteins. The trxG proteins are a well studied group of proteins involved in gene activation. trxG proteins were discovered as suppressors of the PcG mutant phenotypes and were found to be antagonists of Hox gene silencing by PcG proteins (see Figure 1) (126). Trx, for which the group is named, was first identified in mutants in which one segment was transformed to another (127). Ash1 was first identified in a screen for imaginal disc mutants and is classified as a trxG protein due to its functional interactions with other trxG proteins, for example, Trx and absent, small and homeotic discs 2 (Ash2) (128-129). In addition to the catalytic SET domain, both Trx and Ash1 contain a PHD (plant homeo domain) domain, which is a protein-protein interaction domain known to bind trimethylated lysine residues (130-131). The first evidence of the H3K4 methylation activity of Trx was shown by an in vitro histone methylation assay using lysine mutated derivatives of histone H3 (132). An assay using recombinant Ash1 protein showed that Ash1 can methylate histone H3 at K4 and K9 and H4 at K20 (133). Following up on this work, an in vivo approach using antibody staining of polytene chromosomes showed that Ash is required for nearly all H3K4 methylation and that the SET domain alone is sufficient to bring about H3K4 methylation in vitro (134). These authors further showed that only a small proportion of H3K9 methylation is affected in Ash1 mutants, while no effect is seen on H3K36 and H4K20 methylation levels. In contrast to these results, Tanaka et al. (135) performed a lysine scanning mutational assay and determined that the substitution of H3K36 with arginine leads to a loss of methyltransferase activity by Drosophila Ash1 and its mammalian homolog, while substitutions at other residues has no effect. This suggests that H3K36 may be the specific target of Ash1 while the effect on H3K4 methylation may be indirect. Additional
Histone modifying enzymes and gene regulation

work still needs to be done to dissect the role of Ash1 in methylating histones H3 and H4. Based on data generated thus far, the methylating activity of Ash1 and its interactions with other trxG proteins, including Trx and Ash2, are thought to bring about transcription activation.

Trx and Ash1 are components of distinct complexes, but biochemically interact with each other and with other trxG proteins to coordinate a repression of Hox gene expression (see Figure 1) (136-137). For instance, Ash1 has been shown to physically interact with Trx, and the two proteins bind within the same regulatory region of the Ubx gene to regulate its transcription (138). Ash1 has also been shown to interact with the transcriptional coactivator KAT dCBP, which is also found in a complex with Trx (139).

The authors showed that Ash1 and dCBP biochemically interact and colocalize along polytene spreads. It has also been shown that non-coding RNA transcripts of trithorax response elements (TREs) aid in the recruitment of Ash1 to the TREs to activate Ubx transcription, adding yet another layer of epigenetic regulation (140). Ash1 and Trx oppose the repressive effect mediated by PcG proteins, leading to the activation of Ubx (141). While Trx was found to bind Ubx in both transcriptional ON and OFF states, Ash1 was shown to bind only during the ON state, possibly regulating transcription activation by preventing the addition of repressive methylation marks at the promoter and 5' coding regions (142). Beisel et al. (133) showed through ChIP experiments that the trivalent methylation pattern (H3K4, K9 and H4K20) established by Ash1 activity coincides both with recruitment of Brahma (Brm), a trxG protein and a member of a chromatin remodeling complex, and with activation of Ubx. The recruitment of both Ash1 and Trx to chromatin is affected by another trxG protein, Kismet-L (143). In the absence of Kismet-L, there is a loss in recruitment of Ash1 and Trx and an associated increase in the repressive H3K27Me mark. In sum, these findings illustrate specific and overlapping roles for the trxG HMTs Trx and Ash1 in regulating Hox gene transcription.

In addition to regulating the Hox gene cluster, Trx and Ash1 enzymes have been implicated in transcriptional activation of a wide range of genes. For instance, the Trx and dCBP-containing trithorax acetyltransferase complex 1 (TAC1) complex has been shown to be recruited to coding regions of the heat shock induced gene hsp70, correlating with increases in H3K4Me and H3Ac levels at this region (132). Trx and dCBP mutants also show reduced expression of heat shock RNA relative to wild type flies. Another interaction of Trx with heat shock genes has been demonstrated in experiments showing that the phenotypic protein hsp90 colocalizes with Trx in heat shock puffs along polytene chromosomes (144). Hsp90 also affects both Trx binding to chromatin and Trx-dependent regulation of Hox genes. Furthermore, Ash1 has been implicated in activation of a subset of dMyc regulated genes (145). Apart from its role in activation, Ash1 is involved in repression or maintenance of low levels of a subset of dMyc regulated genes by acting alongside the PcG proteins Peiohomoetic (Pho) and Pc. Ash1 also interacts with Ash2, a trxG protein shown to play a role in H3K4Me3 (146). The authors performed transcriptome analysis of mRNA isolated from wing imaginal discs of Ash1 and Ash2 mutants and found a significant overlap in regulated genes. Ash2 further shows significant overlap with genes regulated by SIN3, the scaffold protein of the SIN3 HDAC complex, and shows biochemical interaction with SIN3 via Host Cell Factor (HCF). A genome-wide expression study of Trx and Ash1 targets was also performed and compared to the earlier analysis that included Ash2 (145-147). These combined studies identify a few hundred genes that are deregulated upon loss of each of these proteins, with some overlap amongst them. Mapping these regulated genes to the Drosophila genome, Blanco et al. (147) found that many of the functionally related genes form chromosomal clusters. This result suggests that Trx, Ash1 and Ash2 act on chromatin domains containing co-regulated genes.

In addition to these trxG HMTs, Trr also functions as a H3K4 methyltransferase. Trr was first identified in 1999 as a gene with structural similarity to Trx and found to contain both a SET and PHD domain like several trxG proteins (148). Despite the structural similarity of Trr to Trx, genetic interaction analysis did not show involvement of Trr in Hox gene regulation or PEV. Trr, however, has been shown to trimethylate histone H3K4 and to be involved in ecdysone-dependent regulation (149). Trr interacts with EcR and is involved in retinal development. ChIP experiments show enrichment of Trr, EcR and H3K4Me3 at the promoters of hedgehog and broad, two ecdysone responsive genes in S2 cells. These results demonstrate that Trr acts as a coactivator of EcR by modifying chromatin at ecdysone regulated genes.

Moving on to demethylases, Su(var)-3-3 and LID are two enzymes that have been found to remove the H3K4 methyl mark added by other trxG proteins. Su(var)-3-3, the homolog of LSD1, specifically demethylates H3K4 monomethylated and dimethylated marks (40). Interestingly, loss of this enzyme was shown to also affect H3K9 methylation levels. These authors showed that loss of Su(var)3-3 not only leads to an increase in H3K4Me2 in polytene chromosomes but also a reduction of H3K9Me2 in the heterochromatic chromocenter and to a wider distribution of H3K4Me3 along polytene chromosomes, providing additional evidence of cross talk, this time between distinct methylated residues (see Figure 3). Functional analysis of Su(var)-3-3 shows a role in heterochromatin formation during embryogenesis as discussed above (40).

Unlike Su(var)-3-3, LID specifically demethylates H3K4Me3, a mark of active transcription (150-153). Drosophila LID, a homolog of mammalian RBP2/JARID1A, was first identified in a genetic screen looking for new trxG genes (154). LID mutants enhance the phenotypes of Ash1, Brm and Trx mutations while suppressing Pc mutant phenotypes. Mutations in LID also lead to lethality in adults while exhibiting a small disc phenotype in larvae thus giving it the name little imaginal discs (154). LID contains an N-terminal and C-terminal JmjC domain, three PHD domains, an ARID (AT-rich interaction domain)/BRIGHT domain and a C5HC2 zinc finger domain (150, 152, 154). Analysis of global histones
Histone modifying enzymes and gene regulation

and polytene chromosomes in LID deficient flies showed a specific increase in H3K4Me3 levels while purified LID extracts were able to demethylate both H3K4Me2 and H3K4Me3 substrates in vitro (150, 152). Overexpression of LID in S2 cells leads to a decrease in H3K4Me3 (151-152). In sum, results from multiple laboratories indicate that LID is an important H3K4 demethylase.

As H3K4Me3 and the trxG proteins are associated with active transcription, it was surprising that trxG protein LID was shown to demethylate H3K4Me3. Unlike other trxG proteins, LID does not colocalize with RNA polymerase II on polytene chromosomes (151). Similar to trxG proteins, however, LID does play a role in gene activation. For example, loss of LID leads to reduced expression of the Hox gene Ubx in S2 cells and larval imaginal discs (151-152). LID mutants also show a strong enhancement of PEV, suggesting a role in active transcription (152). The increase in H3K4Me3 due to loss of LID is also accompanied by an altered distribution of the chromo helicase protein Chd1, a transcription activating protein that binds the H3K4Me3 mark (150). Additionally, LID is necessary for the maintenance of histone H3 acetylation of K9/K14, a mark of active transcription in polytene chromosomes (152). These findings suggest that the role of LID in active transcription may be indirect. LID has also been shown to genetically and physically interact with dMyc where LID is essential for dMyc function (153). In support of a role in activation, LID is essential for dMyc induced transcription of Nop60B. The authors further showed that dMyc binds the JmjC domain of LID and thus inhibits its demethylase activity. This suggests that the role of LID in activation maybe independent of the demethylase activity. It will be of interest to further dissect this factor that is important for activation and yet also has enzymatic activity to remove a modification typically associated with active gene expression.

Multiple LID-containing complexes have been isolated, suggesting that association of LID with other proteins may result in complex-dependent functional variations of LID. Lee et al. (155) purified a LID complex which includes the HDAC RPD3, Pfl, MRG15 and CG13367. The deacetylase activity of RPD3 is greatly diminished by its incorporation into this complex. LID was also shown to be a part of two other complexes, RLAF (RPD3 and LID associated factors) and LAF (LID associated factors), both of which also include SIN3, EMSY, Pfl and MRG15 (156). Only RLAF contains RPD3. As part of these complexes, unlike its role as a trxG protein, LID is involved in silencing of Notch target genes. The histone chaperones ASF1 and NAP1 respectively link the LAF and RLAF complexes to the Suppressor of Hairless (Su(H))/Hairless complex. The Su(H)/Hairless complex acts as the DNA binding factor that tethers these complexes to Notch target genes enabling silencing of these genes. LID was also recently purified as part of an isoform specific SIN3/RPD3 HDAC complex where LID specifically associates with the SIN3 220 isoform (56). Taken together, the data shows that the association of LID with other histone and chromatin modifiers affects transcriptional regulation at specific gene targets.

Similar to H3K4Me, H3K36Me is generally associated with active transcription. Drosophila Set2/Hypb is a SET domain-containing protein which specifically methylates histone H3K36 (157). RNAi knockdown of dSet2 leads to a lack of H3K36 methylation in larval salivary glands suggesting that it is solely responsible for H3K36Me in Drosophila. In another study, Bell et al. (10) show that dSet2 is responsible for H3K36 trimethylation while another methyltransferase, dMes-4, affects both di- and trimethylation of H3K36. Furthermore, these enzymes colocalize along polytene chromosomes and at specific gene targets. Additional support for the idea that multiple enzymes can methylate H3K36 comes from the study by Tanaka et al. (135) showing that the specific methylation target of Ash1 is H3K36. Further analysis will be required to unequivocally state which enzymes are responsible for methylation of H3K36. As discussed earlier, dSet2 is important for up-regulation of gene expression along the male X chromosome (11, 21). Additionally, dSet2 genetically interacts with EGR, suggesting a role of this histone modifier in ec dysone-dependent regulation (157). The authors further showed that dSet2 interacts with the elongating form of RNA polymerase II, which correlates with the genome-wide localization of the H3K36 methyl mark (10).

Drosophila KDM4A is a demethylase that has been associated with the removal of H3K36 methylation. dKDM4A was shown to demethylate H3K36Me3 in vitro and in vivo (152, 158). Overexpression of dKDM4A also leads to a reduction of H3K9Me3 in S2 cells and flies as determined by immunostaining experiments (152). The authors showed that dKDM4A localizes to euchromatic regions in polytene chromosomes corresponding with a decrease in the active mark H3K36Me3. Additionally, overexpression of the enzyme leads to spreading of the heterochromatin protein HP1 into euchromatic regions (152, 158). Lloret-Llinares et al. (152) speculate that the spreading of HP1 is primarily due to loss of H3K36Me at euchromatin rather than alternation in H3K9 methylation as they did not observe spreading of H3K9Me2,3 from the chromocenter. Further validating the association with HP1, Lin et al. (158) showed that the HP1 isoform, HP1α, physically interacts with dKDM4A, which enhances the demethylase activity of the enzyme. In sum, the data suggests that dKDM4A is likely involved in repression of transcription.

Another enzyme implicated in demethylation of H3K36Me is dKDM2. dKDM2 was identified as part of a novel PcG complex, dRAF, which also includes Psc and dRING, an E3 ubiquitin ligase (see Figure 1) (159). dRAF and PRC1 show an overlap in the transcriptomes that they regulate in S2 cells. Genetic analysis shows that dKDM2 is an enhancer of Psc and a suppressor of the trxG histone methyltransferases Ash1 and Trx. The authors further showed that a loss of dKDM2 in S2 cells leads to an increase in H3K36Me2 as well as decreased ubiquitination of H2A. This suggests that dKDM2 affects ubiquitination of H2A by dRING. Apart from the role in demethylating H3K36Me2, dKDM2 has also been shown to demethylate H3K4Me3 in flies (160). Loss of dKDM2 leads to an
Histone modifying enzymes and gene regulation

increase of H3K4Me3 specifically in polytene chromosomes and larval extracts analyzed by western blotting. The authors further showed that loss of dKDM2 also leads to defective nuclear organization. As dKDM2 is part of a PcG complex and involved in removing methylation marks associated with transcription activation, it is likely a factor important for silencing and gene repression.

H3K79Me is a third methylated lysine residue associated with active transcription. So far, all the HMTs discussed have been SET domain-containing lysine methyltransferases. Only a single HMT has been identified which does not contain a SET domain, dDot1/Grappa. This enzyme was first identified in a screen for novel PcG proteins and was seen to show phenotypes and genetic interactions with both PcG and trxG genes (161). Thus dDot1 is a member of the Enhancer of trithorax and Polycomb (ETP) class of genes. Immunostaining of wing discs and western blotting of imaginal discs and larval extracts showed that dDot1 is able to mono- or dimethylate H3K79 (161). Genome-wide ChIP-chip studies and polytene chromosome staining showed H3K79Me to be associated with chromatin domains containing actively transcribed genes (4, 161). In addition to its association with active transcription, dDot1 is able to play a role in telomeric silencing, but has no effect on centromeric silencing. dDot1 is also essential for viability and has been found to be important in regulation of development (161-162). dDot1 has been implicated in maintenance of normal lifespan, oxidative stress resistance and regulation of the Wnt signaling pathway (162-163). Reduction of dDot1 and homologs of the mammalian Dot complex (DotCom) by RNAi leads to reduced expression of senselss and several other wingless (wg) target genes (162). Mohan et al. (162) further showed that H3K79Me3 levels decrease in mutants of dBre1, a E3 ubiquitin ligase. Loss of dBre1 results in loss of monoubiquitination of histone H2B and affects the expression of some wg target genes. These findings suggest that dDot1 methylation of H3K79 is linked to monoubiquitination by dBre1 and that these marks may play a key role in the Wnt signaling pathway (see Figure 4). Together these data demonstrate that dDot1 predominantly plays a role in active transcription but also has some repressive functions.

Shifting focus from methylation marks that are typically linked to active transcription, we now concentrate on methyl marks that have been generally associated with repressive mechanisms. It must be noted, however, as discussed above, that several histone methyltransferases and demethylases, primarily associated with activation, also play roles in repression. This suggests that the functional roles of these enzymes are context dependent and are greatly influenced by associated proteins.

As a methyl mark associated with repression, H3K9Me has been implicated in heterochromatin formation, known above and reviewed in (164). In Drosophila, the enzymes Su(var)3-9, dSETDB1 and dG9a have been identified as H3K9 methyltransferases (see Figure 3). Su(var)3-9 is the major H3K9 methyltransferase associated with pericentric heterochromatin (34, 67). dSETDB1, a second enzyme that possesses H3K9 HMT activity, functions mainly on the 4th chromosome (29, 31). Further, the binding of HP1 and Painting of fourth (POF) to the 4th chromosome was shown to be dependent on methylation by dSETDB1 (31). dSETDB1 has also been shown to play a role in DNA methyltransferase 2 mediated transcriptional silencing (165) and to function along with Su(var)3-9 in methylating pericentric heterochromatin, as discussed above (32). Microarray analysis of gene expression in mutant flies of Su(var)3-9 and dSETDB1 shows a high proportion of overlap in misregulated genes suggesting coordinated function of the two enzymes (32). Several genes, however, are regulated by only one of the two enzymes, indicating that there are distinct functions for these proteins. In accord with the unique role of dSETDB1 in the 4th chromosome, dSETDB1 affects regulation of a higher percentage of 4th chromosome genes, while Su(var)3-9 affects a higher percentage of repeat elements.

A third enzyme, dG9a, was shown to have H3K9 methyltransferase activity both by in vitro and in vivo studies (27-28, 30). In vitro, dG9a is also capable of methylating H3K27 to a lesser extent and H4 at lysine residues 8, 12 or 16 (30). Polytenic staining showed localization of dG9a to euchromatin and telomeric regions (27, 30). Colocalization to polytenes showed no significant overlap with HP1 or phosphorylated RNA polymerase II (27). Further, dG9a mutants have a decrease in HP1 and an increase in phosphorylated RNA polymerase II at euchromatic regions. Genetic studies indicate that dG9a is required for normal development and plays a role in edysyne-mediated signaling pathways (30). Similar to Su(var)3-9, dG9a has also been shown to be a dominant suppressor of PEV (28). Additionally, this enzyme genetically interacts with Su(var)3-9, suggesting some overlap in function between the two methyltransferases. Taken together, the data implicate this enzyme in repression of genes located in euchromatic domains.

Methylated H3K27 is another mark associated with gene repression, and this modification has been specifically found to play an important role in polycomb-mediated silencing (see Figure 1). The key complexes that are important in PcG regulation are PRC1, PRC2 and PhoRC. Mono-, di- or trimethylation of H3K27 is linked to E(z), the SET domain-containing subunit of PRC2 (6-7). Methylation of H3K27 by E(z) increases the affinity of Pc, a PRC1 component, binding to histone H3 (6). Disruption of the HMT activity of E(z) affects the repression of the Hox gene Ubx (7). E(z) is present in a 600 kDa complex with ESC, p55 and Su(z)12 (7). E(z) has also been found in a one MDa complex that includes, in addition to the above proteins, polycomb like (Pcl) and the HDAC RPD3, suggesting that E(z) may exist in multiple complexes that have distinct functions (6, 166). Association of Pcl is essential for high levels of H3K27Me3 at Pc target genes, although it is not necessary for the genome-wide mono- and dimethylation of H3K27 by the PRC2 complex (167). ESC, which binds E(z) through C-terminal WD (tryptophan-aspartic acid) repeats, binds via its N-terminal tail to histone H3 and this interaction is essential for
Histone modifying enzymes and gene regulation

H3K27Me by E(z) (168). ESCL, which shows a similar protein interaction pattern to ESC, is thought to alternate with ESC during development, with ESC predominating in embryonic stages (169-170). Chen et al. (171) showed that a distinct subunit of the PRC2 complex, Su(z)12, is also required for the H3K27Me activity of the complex and that overexpression of Su(z)12 leads to increased H3K27Me. This data has lead to one model of polycomb-mediated silencing in which H3K27Me is necessary for the recruitment of PcG proteins to target genes. Alternatively, Kahn et al. (172) propose that the recruitment of PcG proteins to PREs is independent of histone methylation. Based on ChIP studies, they show that PcG proteins are mainly localized to a peak area at the PRE associated with the PcG target gene Ubx, the least methylated area within the target gene region, while H3K27Me3 is widespread along the gene. Based on these results the authors suggest a model in which looping occurs, allowing PcG proteins, bound to the PREs, to use their chromodomain to scan methylation status and to methylate nucleosomes along the length of the gene. Despite some differences in mechanism, both models implicate PcG proteins and H3K27Me3 in polycomb-mediated repression.

The *Drosophila* enzyme UTX is known to demethylate H3K27 and antagonize polycomb-mediated silencing (see Figure 1) (126). dUTX was shown to demethylate di- and trimethylated histone H3K27, both in vivo, by immunohistochemistry, and in vitro, using recombinant dUTX (173-174). This demethylase activity was shown to be JmjC domain dependent (174). Interestingly, dUTX also affects H3K4 monomethylation, independent of its JmjC domain (174). dUTX colocalizes with the elongating form of RNA polymerase II on polytenic chromosomes and is recruited to the hsp70 gene upon heat shock induction, suggesting a role in active transcription (173). Because dUTX is essential for viability and mutants show some phenotypic similarity with Trx mutants, dUTX can be potentially classified as a trxG gene (174). Additionally, loss of dUTX leads to increased proliferation due to higher Notch activity, indicating that dUTX antagonizes Notch signaling. Loss of dUTX also results in tumor-like growth in an Rbf (*Drosophila* retinoblastoma)-dependent manner. Thus, the demethylase activity of dUTX plays an important role in development.

H4K20 methylation is an additional methylation mark associated with gene repression. PR-Set7 and Suv4-20 are the two enzymes in *Drosophila* that have been identified to catalyze this reaction. Polytene chromosome staining and western blot analysis showed that loss of PR-Set7 leads to a loss in H4K20 methylation (175-176). Yang et al. (177) went on to show that PR-Set7 mainly contributes to H4K20 monomethylation while Suv4-20 is responsible for H4K20 di- and trimethylation. PR-Set7 was further shown to associate with dense chromatin regions in heterochromatin and euchromatin (175). The authors also found that a PR-Set7 mutant shows strong suppression of PEV, providing additional evidence for a role in silencing. PR-Set7 mediated H4K20 monomethylation has also been implicated in the recruitment of the HDAC RPD3, via the association of RPD3 with *Drosophila* lethal (3) malignant brain tumor (dl(3)MBT), which can bind monomethylated H4K20 (178). PR-Set7 has a role in cell cycle progression, as PR-Set7 mutants show a delay in progression through early mitotic stages and a reduction in cyclin B levels (179). The authors further showed that loss of PR-Set7 leads to defective chromosome condensation.

An RNAi screen for H4K20 methylation identified Suv4-20 as a major enzyme responsible for this mark in *Drosophila* S2 cells (177). The authors further showed that Suv4-20 is responsible for both H4K20 di- and trimethylation, where the dimethyl mark is predominant (~90%) relative to the trimethyl mark (~5%). Loss of Suv4-20 also leads to a loss in H3K4 di- and trimethylation in polytenic chromosomes (180-181). It was further shown that HP1 is required for Suv4-20 dependent H4K20Me3 enrichment at pericentric heterochromatin (177). This suggests that HP1 is required for the substrate specificity of Suv4-20 and implicates Suv4-20 in heterochromatic silencing. Furthermore, loss of Suv4-20 leads to a strong suppression of PEV (180). Suv4-20 has also been shown to associate with the cytosine 5-methyltransferase DNMT2, whereby DNMT2 initiates H4K20Me3 by Suv4-20 (182). The two enzymes function together in maintaining retrotransposon silencing. In sum, the data on PR-Set7 and Suv4-20 demonstrate that H4K20Me is important for establishment or maintenance of repressive chromatin.

5.3. Arginine Methylation

Similar to histone lysine methylation, arginine methylation has been implicated in epigenetic regulation of transcription. Arginine is methylated by the addition of either one or two methyl groups to its guanido nitrogen atoms (183). The reaction is catalyzed by two major classes of enzymes: Type I enzymes lead to the formation of asymmetrical dimethylated arginine residues whereas type II enzymes result in symmetrical dimethylated arginine residues (183-184). Arginine methylation is important for transcriptional regulation, nuclear trafficking, pre-mRNA splicing and protein-protein interactions (184).

One of the early studies on arginine methylation indicated that this mark has a role in the heat shock response. Heat shock and arsenite treatment induce a stress response in *Drosophila* leading to a decrease or silencing of the majority of transcription while increasing expression of stress response genes (185). These authors showed that during stress response there is novel dimethylation of arginine residues of histone H3. This is accompanied by a reduction in lysine methylation of histone H3 and an increase in methylation of histone H2B at proline residues. These finding show a role for arginine methylation in transcription activation of heat shock genes and also an interaction of multiple methylation marks in gene regulation.

Based on sequence homology studies with mammals, Boulanger et al. (186) first identified a family of nine PRMTs, which they termed DART1 to DART9 (*Drosophila* Arginine Methyltransferases 1-9). Of these, DART1 and DART4/CARMER were identified as homologs of mammalian PRMT1 and PRMT4/CARM1.
that are involved in arginine methylation of histones H4 and H3 respectively (186-187). Both enzymes belong to the type I category of arginine methyltransferases leading to asymmetrical dimethylation of arginine residues. Based on in vitro histone methylation assays, Kimura et al. (188) showed that, like mammalian PRMT1, the *Drosophila* homolog DART1 dimethylates histone H4 at arginine 3. They further showed that DART1 interacts with EcR in an ecdysone-dependent manner and acts as a corepressor of EcR, unlike the mammalian homolog PRMT1, which acts as a coactivator of nuclear hormone receptors (189). DART1 therefore plays an important role in maintaining normal pupal development in flies. DART4/CARMER, which dimethylates histone H3 at arginine 17, has also been associated with ecdysone-dependent regulation (187). DART4/CARMER associates with EcR and enhances ecdysone-mediated transcription and programmed cell death in *Drosophila* l(2)mbn cells. In flies, up-regulation of DART4/CARMER correlates with tissue hydrolysis in salivary glands. A later study by Cakouros et al. (190), showed that the *Drosophila* lysine ketoglutarate reductase/saccharopine dehydrogenase is a cofactor that binds histones H3 and H4, suppressing ecdysone-mediated programmed cell death by specifically inhibiting histone H3 arginine 12 dimethylation (H3R17Me2) by DART4/CARMER.

Though less well characterized than the lysine HMTs, the identification and initial study of PRMTs demonstrates an important role for these enzymes in transcription regulation. Although an arginine demethylase has been characterized in mammals, so far none have been identified in *Drosophila* (191). Further study of PRMTs and potential demethylases will clarify the role of this type of modification in important regulatory mechanisms.

5.4. Ubiquitination

Like acetylation and methylation marks, ubiquitination of histones has been implicated in transcription regulation. Ubiquitination refers to the covalent addition of ubiquitin, a 76 amino acid protein, to other proteins including nucleosomal histones. Although polyubiquitination of proteins is associated with proteolytic degradation of the protein, histones are generally monoubiquitinated, which serves as a functional epigenetic signal (192-193). Ubiquitination of chromatin mainly occurs on histone H2A and H2B and is the result of the covalent linkage of the C-terminal glycine of ubiquitin to the epsilon-amino group of the internal lysine 119 in H2A and lysine 120 of H2B (192-193).

An early study of the genomic localization of ubiquitin-H2A (uH2A) was done by Levinger and Varshavsky (194), where they performed a two dimensional hybridization mapping of nucleosomes in *Drosophila* cultured cells. In non-heat shocked cells, about one in two nucleosomes in the highly transcribed *copia* and *hsp70* genes have uH2A, while in poorly transcribed, tandemly repeated 1.668 satellite DNA, only one in 25 nucleosomes have uH2A. Staphylococcal nuclease digestion analysis further showed that while *hsp70* in non-shocked cells is enriched in uH2A nucleosomes, the nucleosomal organization is lost in heat shocked cells, rendering them highly susceptible to nuclease digestion. Because uH2A is prevalent in transcribed regions, these findings suggest that this modification may prevent higher order chromosomal structure formation, facilitating transcription. Additional evidence for the role of uH2A in chromosomal organization is seen during spermatogenesis in *Drosophila*. During the conversion from a nucleosome based structure to a protamine based structure during sperm development, de novo monoubiquitination of H2A as well as hyperacetylation of H4 occurs in early post meiotic stages, bringing about a relaxed chromatin conformation to allow access to chromatin remodelers (195).

Polycomb-mediated repression is another major pathway utilizing uH2A in *Drosophila* (see Figure 1). The PcG protein dRING/Sex combs extra (Sce) was found to have E3 ligase activity, ubiquitinating histone H2A (196). Using ChIP analysis of wing imaginal discs and S2 cultured cells, the authors found that both dRING and uH2A colocalize at the PRE and just downstream of the transcription start site of *Ubx*. Loss of dRING leads to derepression of *Ubx* in both S2 cells and wing imaginal discs, indicating that uH2A plays an important role in Polycomb-mediated silencing of *Ubx*. dRING has also been shown to associate with dKDM2, a demethylase, as a part of a novel PcG complex, dRING associated factors (dRAF), as discussed previously (159). Intriguingly, deubiquitination of H2A has additionally been implicated in polycomb-mediated repression. The polycomb repressive complex Polycomb repressive deubiquitinase (PR-DUB), which includes Calypso, a ubiquitin C-terminal hydrolase, and the PcG protein Additional sex combs, was identified as a deubiquitinase of H2A monoubiquitination (197). ChIP followed by whole genome tiling array analysis showed that PR-DUB complex components bind PREs at a large set of PcG target genes. Further, it was shown that the catalytic activity of PR-DUB is essential for Hox gene repression. While perhaps surprising that PcG silencing requires ubiquitination as well as deubiquitinase activities, the conclusion is supported by the finding that flies carrying mutations in members of both complexes have more severe changes in Hox gene expression and transformation phenotypes than single mutants alone. The authors suggest that either a dynamic balance between H2A ubiquitination and deubiquitination or specific chromosomal sites of uH2A are important for polycomb-mediated silencing.

Unlike the repressive role of H2A ubiquitination associated with polycomb-mediated silencing, H2B ubiquitination has been mainly linked to activation. One well known pathway regulated by H2B ubiquitination is the Notch signaling pathway. It was shown that dBre1 plays a role in activating Notch target genes by affecting histone modifications (198). In yeast, Bre1 monoubiquitinites H2B, and indirectly leads to increased methylation of histone H3K4, which suggests cross talk between the two marks (see Figure 4) (199). *Drosophila* Bre1 mutant clones have a loss in H3K4Me3, suggesting a similar molecular function to that found in yeast (198). In opposition to the role in activation of uH2B and consistent with the above results, H2B deubiquitination leads to repression of Notch.
target genes (200). These authors showed that mutations in the ubiquitin protease Scrawny, which deubiquinates H2B, leads to the suppression of PEV and activation of several differentiation genes including Notch targets. Scrawny mutant clones also have an increase in H3K4Me3. The data suggest that Scrawny functions in opposition to the H2B ubiquitinase dBre1, and also indirectly regulates H3K4Me levels. Scrawny is additionally required for maintenance of multiple Drosophila stem cells, potentially by repressing genes involved in differentiation (200).

Several other deubiquitinases have been characterized in Drosophila, including ubiquitin specific protease 7 (USP7) and Nonstop. USP7 tightly interacts with the metabolic enzyme GMPS, forming a complex that selectively deubiquinates H2B in vitro (201). USP7 was also found to bind many silenced regions in polytene chromosomes including Hox gene regions (see Figure 1). Loss of USP7 leads to enhanced homeotic transformations by the PcG proteins, providing additional evidence that deubiquitination of H2B is important for silencing. Another ubiquitin protease, Nonstop, deubiquinates H2A and H2B (202-203). Nonstop overexpression suppresses heterochromatic silencing by PEV, suggesting that deubiquitination is important in counteracting heterochromatin formation (203). Mutation of Nonstop leads to defects in axon targeting in flies, indicating that deubiquitination is also important for neuronal development (202). Nonstop is a component of the dSAGA complex, which as discussed earlier, also contains the HAT GCN5 (87-88). The dSAGA complex is thus comprised of two catalytic modules, one that acts as the deubiquitinae and the other as an acetyltransferase (202). The modules are functionally distinct but show some overlap in the genes they regulate. Interestingly, as discussed earlier, TAF1 also links acetylation and ubiquitination. TAF1 has acetylation and histone H1 monoubiquitination activity in the same protein (117). Taken together, data from multiple groups provide evidence of interplay between ubiquitinated and acetylated histones.

5.5. Phosphorylation
As mentioned above, phosphorylation of serine 10 of histone H3 is associated with transcription at autosomal genes as well as with regulation of chromatin structure involved in both dosage compensation and prevention of heterochromatin formation (see Figures 2 and 3). In addition to a function in transcription, phosphorylation of histone H3 is linked to chromatin condensation of mitotic chromosomes, reviewed in (204). Mutations in JIL-1 kinase, which is responsible for the phosphorylation of histone H3 at serine 10 outside of mitosis, lead to loss of viability (23). Additionally, loss of JIL-1 results in major changes in X chromosome structure, including disruption of the euchromatin banding pattern, suggesting loss of the more open, decondensed regions. In the previous sections, we have highlighted multiple examples of cross talk between other marks and H3S10P. H3K9Me2 binding spreads out from heterochromatin in JIL-1 mutants, suggesting that H3S10P prevents spreading of the methyl mark (42). A second example links H3S10P, H3K9Me2 and acetylation. Mutations in multiple components of the HAT complex ATAC lead not only to a reduction in H4K12Ac as expected, but also reduced H3S10P binding and spreading of H3K9Me2 into euchromatin (see Figure 4) (94).

In addition to the role of H3S10P in establishing chromatin domains important for male X chromosome gene expression and prevention of gene silencing, this mark has been proposed to accompany the activation process for the majority of RNA polymerase II transcribed genes (205). While two research groups have shown that transcription of heat shock genes is affected by mutations in JIL-1, the groups disagree as to the mechanism of action of this enzyme (205-206). Ivaldi et al. (205) propose that histone H3 phosphorylation by JIL-1 is important for the majority of RNA polymerase II transcription and that the activity is linked to release of promoter proximal pausing. Alternatively, rather than affecting chromatin structure to promote elongation, Cai et al. (206) propose that the major role of JIL-1 is to establish and maintain chromatin structure that will facilitate the transcription process. Further research focused on H3S10P including the identification and characterization of specific phosphatases as well as the cross talk with other histone modifications should shed light on the how this mark alone and in combination is read by other proteins resulting in transcription activation.

6. CONCLUSIONS AND FUTURE PERSPECTIVES
The accumulation of data demonstrates that a large number of histone modifying enzymes play a role in generating a chromatin template that will facilitate or silence transcription. Research identifying the location of the marks, as well as phenotypes resulting from altered levels of the modifying enzymes, indicates that there is cross talk between marks and that combinations of marks define transcription states, consistent with the histone code model of transcriptional regulation. Interplay between marks includes documented examples where one mark facilitates another or is even required for the second mark to occur. H4K12Ac by dGCN5 of the ATAC complex is important for H3S10P and is one example of this type of cross talk (94). Other examples are highlighted in Figure 4. An additional type of cross talk occurs when one modification prevents the modification of a second residue. An example of this type is seen where the loss of H3S10P by JIL-1 mutation allows H3K9Me to spread out from heterochromatin, indicating that one role of S10P here is to keep H3K9Me confined to heterochromatin compartments (see Figure 3) (42). A third type of cross talk occurs when a single residue is subject to different types of modification and these specific marks differentially affect functional read-out. These antagonistic marks include H3K27Ac, important for activation, in contrast to H3K27Me, linked to transcriptional silencing (see Figure 1) (8). An area of continued research is to identify signaling and developmental pathways that regulate enzymes involved in cross talk.

Adding additional complexity to the fact there are multiple enzymes that affect a single residue, either to add
Histone modifying enzymes and gene regulation

the same mark, or in some cases, catalyzing different marks, the specificity of individual enzymes can be altered by the proteins with which they associate. This is observed in the distinct KAT activity of dGCN5 in dSAGA and ATAC as discussed above. dCBP also has multiple substrates, likely regulated by interacting factors including Trx and Asf1 (65, 79). The HDAC RPD3 is found in multiple complexes, which exhibit differences in activity (56, 155-156, 207). Proteomic analysis, combined with the genomics and developmental studies, will help to sort out the specific roles of individual enzymes in establishing chromatin structure required for appropriate gene expression.

An emerging concept in the area of epigenetics and control of gene expression is the idea of bivalent domains of histone modification (208-209). Bivalent domains are chromosomal regions enriched for H3K4Me3 and H3K9Ac, generally considered active marks, and for H3K27Me3, associated with repression. Genes that reside in chromosomal regions with bivalent marks were first identified in genome-wide mapping studies of histone modifications found in mammalian embryonic stem cell populations (210-211). Many of the genes that contain these marks fall into the broad category of developmental regulators. Upon cell differentiation, genes that are expressed retain the active marks and lose the inactive ones (212). The opposite occurs to chromatin around genes that are inactive in the differentiated state. While Drosophila are currently being utilized for analysis of stem cell proliferation and differentiation (200), to date no direct studies of bivalency in pluripotent cells have been performed. Through the investigation of Drosophila stem cells and the comparison to mammalian counterparts, it will be interesting for researchers to uncover the evolutionary relationships of stem cell biology and histone modification.

As is evident from the research presented, studies using the Drosophila model system have provided a wealth of information regarding epigenetic regulation of transcription and its relationship to signaling and development. Major efforts are underway by members of the Drosophila research community to continue these lines of investigation. Work by the modENCODE Consortium is designed to map transcribed regions of the Drosophila genome as well as the location of a myriad of DNA binding transcription factors and histone modifications (213). In fact, two recent publications, one by the modENCODE Consortium members (214) and a second from Filion et al. (215), report findings of their initial comprehensive analyses. While a full discussion of this comprehensive data is beyond the scope of this review, briefly, the results are as follows. In one study, the authors analyzed the genome-wide binding of 53 distinct chromatin interacting factors in Kc167 cultured cells and compared these data to a subset of histone modification data (215). They define five major chromatin states including two types of heterochromatin, pericentric and polycomb-mediated, two types of euchromatin marked by the binding of distinct transcriptional activators and a large block of silent chromatin, quite distinct from classically defined heterochromatin. For further description of the findings of this work, the reader is referred to (216). The second report, from the modENCODE Consortium, analyzed the binding of multiple DNA-binding factors, histone modifiers and histone marks in multiple cultured cell lines and larval tissue samples (214). These authors report the presence of nine combinatorial chromatin states. The nine state model utilizes more information to classify chromatin states into additional categories in comparison to the five state model. The modENCODE Consortium authors present an initial comparison between the two data sets and noted the presence of both similarities and differences in the results. The nine state model includes multiple transcriptionally active states, separated into one with active marks at the promoter and transcription start site, two with distinct active marks of elongation, one associated with genes containing introns and one with the X chromosome. Repressive states include two types of heterochromatin, identified by different levels of repressive histone modifications, an additional state associated with polycomb repression and one “silent” domain, likely most similar to that described in the five state model. It will be of great interest to further analyze these very large recently published data sets. They provide a foundation for a detailed understanding of the role of epigenetic marks in the control of gene transcription.

The next major challenge for the field is to determine the molecular mechanisms that are involved in the transmission of the marks from one cell to another during cell division and from one generation to the next. The comprehensive data sets resulting from work of the modENCODE Consortium and others are anticipated to facilitate the required mechanistic studies that will illuminate epigenetic regulation.

7. ACKNOWLEDGMENTS

Authors Aishwarya Swaminathan and Ambikai Gajan contributed equally to this work. We thank members of the Pile laboratory for critical reading of the manuscript. We thank Marla Spain for literature research in the early stages of manuscript preparation. Research in the Pile laboratory is supported by a Research Scholar Grant awarded to LAP from the American Cancer Society.

8. REFERENCES


Histone modifying enzymes and gene regulation


Histone modifying enzymes and gene regulation


transcription of distinct genes. Genomics, 86(5), 606-17 (2005)


Histone modifying enzymes and gene regulation


Histone modifying enzymes and gene regulation


Histone modifying enzymes and gene regulation


Histone modifying enzymes and gene regulation


Histone modifying enzymes and gene regulation


Histone modifying enzymes and gene regulation


Histone modifying enzymes and gene regulation


**Key Words:** Chromatin, *Drosophila*, Epigenetics, Histone modification, Transcription, Review

**Send correspondence to:** Lori Pile, Department of Biological Sciences, 5047 Gullen Mall, Wayne State University, Detroit, Michigan 48202, Tel: 313-577-9104, Fax: 313-577-6891, E-mail: loripile@wayne.edu

http://www.bioscience.org/current/vol17.htm