FOR GENES THAT ENCODE ONE COMPONENT OF A MULTIMERIC
PROTEIN COMPLEX, MEASURING ONLY ONE PHENOTYPE OFTEN
GIVES A BIASED VIEW OF FUNCTION: Su(var)3-9 AND CHROMATIN
ARCHITECTURE AS AN EXAMPLE.

by

Pamela Kalas

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Genetics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

December 2008

ABSTRACT

Eukaryotic genomes are organized into chromatin, a highly dynamic complex of DNA and proteins, which plays a critical role in the regulation of genes expression. This thesis focuses on the study of a non-histone chromatin protein, the SET domain-containing H3K9 methyltransferase (HMTase) SU(VAR)3-9, and its role in the packaging and regulation of a euchromatic locus, the *histone genes cluster* (HIS-C). SU(VAR)3-9 was discovered in *Drosophila melanogaster*, but it is highly conserved from yeast to mammals. It has two conserved domains, the chromo- and the SET domains, and both are required for its function in gene silencing. The SET domain is responsible for the catalytic activity of SU(VAR)3-9, while the exact function of the chromo domain is still unclear.

To gain an insight on the role(s) of SU(VAR)3-9 in the regulation of gene silencing, we first characterized a collection of *Su(var)3-9* EMS-induced mutants that had been isolated in a genetic screen for strong, dominant suppressors of position-effect variegation (PEV). These mutants were characterized at the molecular, enzymatic, and cellular level, and their effect on gene silencing was also examined. We found that all mutants have single amino acid substitutions in the conserved preSET/SET/postSET domain, and that they all display a dramatic or complete loss of HMTase activity, strongly suggesting that suppression of PEV is linked to SU(VAR)3-9's ability to methylate H3K9.

The HIS-C is a natural, euchromatic target of SU(VAR)3-9, and mutations in *Su(var)3-9* can alter its chromatin structure. To investigate the exact role(s) of SU(VAR)3-9 in the regulation of this locus, we analyzed the effects of a series of

Su(var)3-9 missense mutants on the chromatin architecture of the HIS-C and on the expression of the *histone genes*. We detected a drastic reduction in the levels of H3K9me2 and HP1 associated with the *his* genes in all *Su(var)3-9* missense mutants, although the mutant SU(VAR)3-9 still associate with the HIS-C. In addition, these mutants have elevated amounts of *histone H2A* and *histone H3* RNA, suggesting that the enzyme function of SU(VAR)3-9 is critical for the regulation of the *histone* genes.

TABLE OF CONTENTS

CHAPTER 1 *INTRODUCTION*

CHAPTER 2 *ALTERED HISTONE H3 METHYLTRANSFERASE ACTIVITY*

OF SU(VAR)3-9 IMPAIRS GENE SILENCING

CHAPTER 3 *THE ROLE OF SU(VAR)3-9 IN THE REGULATION OF*

DROSOPHILA'S HISTONE GENE *CLUSTER (HIS-C)*

CHAPTER 4 *GENERAL SUMMARY AND DISCUSSION*

APPENDICES

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ChIP   Chromatin immunoprecipitation

EGFP   Enhanced-green fluorescent protein

GST   Glutathione-S-transferase

H3K4   Lysine 4 of histone H3

H3K9   Lysine 9 of histone H3

H3K27   Lysine 27 of histone H3

H4K20   Lysine 20 of histone H4

HDAC1   Histone deacetylase 1

HIS-C   *Histone Gene* cluster

HP1   Heterochromatin Protein 1

MNase   Micrococcal nuclease

MOF   Males absent On the First

PEV   Position-effect variegation

PC   Polycomb

RTase   Reverse transcriptase

RT-PCR   Reverse transcription-PCR

S.E.M.   Standard error of the mean

Su(var)   Suppressor of position-effect variegation

CO-AUTHORSHIP STATEMENT

A version of chapters 2 and 3 of this thesis will be submitted for publication.

I (P. Kalas) performed all the experiments, and the data analysis, presented in

this thesis. Dr T.A. Grigliatti supervised the project and provided assistance with

experimental design and editing of this manuscript.

1. INTRODUCTION

1.1. Chromatin and gene regulation: an overview.

In metazoans, each somatic cell contains the same genetic information, but the fate and specificity of the cells comprising various tissue types is determined by and dependent upon the expression of different subsets of genes. While some genes go through "on" (expressed) and "off" (not expressed) states throughout the whole life of the organism, others remain permanently "off" in certain cell types, or after a certain stage of development. So, for example, nerve-specific genes are not expressed in muscle cells, and mitosis-specific genes are not expressed in cells that are not dividing anymore. Failure to express, or conversely, failure to silence a particular set of genes at the appropriate time in a specific cell type can have dramatic consequences such as congenital malformations, cancer, or death (reviewed for example by Ausio *et al.,* 2003; Oligny, 2003; Jaffe, 2003; Moss and Wallrath, 2007; Nelson *et al.,* 2007). Indeed, tight, accurate regulation of gene expression is absolutely crucial for the survival of any organism.

Gene expression is regulated at many levels. At the transcriptional level, protein complexes assembled on regulatory sequences such as enhancers interact with the transcriptional machinery assembled on a gene's promoter(s), stimulating gene transcription. However, in order for this to happen, the gene in question must first be "transcriptionally competent", or in other words, in a form that makes it accessible to transcription factors. In prokaryotes, the entire genome is transcriptionally competent, that is, the DNA is readily accessible to regulatory proteins. However, this is not the case in eukaryotes.

The eukaryotic genome is subdivided into chromosomes and organized and compacted into chromatin, a complex of DNA and proteins (see below, section 1.2.). Chromatin exists in different forms, depending on its level of compaction, the presence of specific proteins, and particular post-translational modifications of some of its components (recently reviewed by Ebert *et al.,* 2006; Razin *et al.,* 2007; Kouzarides, 2007). Some forms, often referred to as "open chromatin" or euchromatin, are more accessible to the transcriptional machinery, while "silent chromatin", or heterochromatin, in contrast, is typically refractory to transcription. Chromatin is extremely dynamic, having the ability to convert, under the appropriate circumstances, from an open to a silent conformation, and *vice-versa*). By modulating the chromatin architecture of certain regions, cells are able to make genes, and entire chromosome domains, transcriptionally competent or silenced (for a review see for example Struhl, 1999; Talbert and henikoff, 2006). Hence, in order to fully understand the biological process of gene regulation, it is necessary to uncover the mechanisms underlying chromatin biology and chromatin architecture. This thesis focuses on the study of SU(VAR)3-9, a key non-histone chromatin protein with a histone methyltransferase activity, and on its roles in modulating the packaging and thus the regulation of a euchromatic locus, the *histone gene* cluster.

1.2.   The components of chromatin.

Chromatin is a dynamic, highly organized complex of DNA and proteins. Histones are the most abundant chromatin proteins, but there are also numerous non-

histone chromatin proteins (NHCPs) that have crucial roles in regulating and maintaining chromatin architecture.

1.2.1. <u>Histones</u>

The basic unit of chromatin is the nucleosome, 146 bp of DNA wrapped around an octamer consisting of two copies of each of the core histones, H2A, H2B, H3 and H4 (Finch *et al.,* 1977; Klug *et al.,* 1980; Luger *et al.,* 1997), and chromatin at its simplest can be described as an array of nucleosomes (Kornberg, 1977). Although devoid of enzymatic activity, histones are far from being inert structural components of chromatin. Post-translational modifications of their N-terminal tails play an active role in recruiting and/or stabilizing the binding of non-histone chromatin proteins (NHCPs) to the chromatin fibre. To date, about a dozen histone modifications have been described, including methylation, acetylation, ubiquitylation and SUMOylation of lysines, phosphorylation of serines and threonines, methylation of arginines, and ADP-ribosylation (for a systematic review of the current nomenclature, see Turner, 2005). Table 1.1. summarizes the histone modifications thus far described in *Drosophila*, and the enzymes catalyzing them.

The availability of antibodies specific for most histone modifications, in combination with high throughput analyses in yeast, *Drosophila* and mammals have been providing high resolution, genome-wide maps of histone post-translational modifications (reviewed by Schones and Zhao, 2008). This has allowed the systematic study of the correlations between the chromatin structure and the transcriptional state of a locus (Crawford *et al*., 2006), as well as the

presence of particular histone modifications in that region. In all organisms analyzed, including yeast, flies and mammals, acetylated histone H3 (H3ac) and methylated lysines 4 and 36 of histone H3 (H3K4me, H3K36me) are enriched in regions corresponding to active promoters and transcribed genes (Roh *et al*., 2005; Schubeler *et al*., 2004; Liu *et al*., 2005; Bernsterin *et al*., 2005; Kim *et al*., 2005; Pokholok *et al*., 2005; Barski *et al*., 2007). In animals, the presence of phosphorylated histone H3 serine 10 (H3S10ph) is also characteristic of transcriptionally competent regions of the genome (Wang *et al*., 2001; Ebert *et al*., 2004; 2006). On the other hand, di- and tri-methylated histone H3 lysines 9 and 27, and methylated histone H4 lysine 20 (H3K9me2,3, H3K27me2,3, H4K20me3) are typically associated with large heterochromatic regions (Schotta *et al*., 2002; 2004; Peters *et al*., 2002; Ebert *et al*., 2004; 2006).

Two main mechanisms have been proposed to be responsible for the "translation" of a given set of histone modifications into a particular transcriptional state of the region in question. One of them postulates that the addition of a charged group (such as an acetyl or a phosphate group) to the histone tail can cause a localized decondensation of the chromatin fibre, making it accessible to the transcriptional machinery (Turner, 2000). The other, more widely applicable, proposes that specific histone modifications, or combinations thereof, serve as a binding platforms for specific chromatin proteins (for example Lachner and Jenuwein, 2002; de la Cruz *et al*., 2005). Evidence exists in support of both models, suggesting that both mechanisms probably play a role. For instance, it has been demonstrated that *in vitro* reconstituted chromatin arrays can form compact

fibres, but this compaction is prevented when H4K16 is acetylated, suggesting that the acetyl group could directly affect chromatin architecture (Shogren-Knaak *et al.*, 2006; Chodaparambil *et al*., 2007).

However, histone modifications have also been widely shown to affect chromatin structure and gene regulation via the action of NHCPs. The chromodomain of HP1 has been shown to recognize and bind to H3K9me2,3 (Lachner *et al*., 2001; Jacobs *et al*., 2001; Fischle *et al*., 2003), while the chromodomain of Polycomb specifically recognizes H3K27me2,3 (Cao *et al*., 2002) and those of CHD1 and CHD3 proteins bind to H3K4me and H3K36me, respectively (reviewed by Mellor, 2006). In contrast, bromodomains typically recognize acetylated histone residues (reviewed by Yang, 2004; Mujtaba *et al*., 2007).

The next section will focus on the roles of lysine methylation, since histone lysines are the target of SU(VAR)3-9's methyltransferase activity, and have important functions in the regulation of gene silencing as well as transcriptional competence.

Table 1.1. Post-translational modifications of canonical histones that have been formally detected in *Drosophila*.

MODIFICATION	PREDOMINANT LOCATIONS	ENZYME(S) RESPONSIBLE	SELECTED REFERENCES
H3K4me3	Euchromatin/interbands	TRX TRR ASH1, ASH2	Smith *et al.,* 2002; Sedkov *et al.,* 2003; Beisel *et al.,* 2002; Byrd and Shearn, 2003; Beltran *et al,* 2007.
H3K9me	IV chromosome	DmSetDB1	Seum *et al.,* 2007; Tzeng *et al.,* 2007.
H3K9me1,2	Heterochromatin	SU(VAR)3-9, dG9a	Schotta *et al.,* 2002; 2003; Eskeland *et al.,* 2004; Mis *et al.,* 2006; Stabell *et al.,* 2006.
H3K9me	Euchromatic sites	dG9a, SU(VAR)3-9	Ner *et al.,* 2002; Schotta *et al.,* 2002; Stabell *et al.,* 2006.
H3K9me3	Heterochromatin	SU(VAR)3-9	Schotta *et al.,* 2002; 2003; Eskeland *et al.,* 2004.
H3K9ac	Euchromatin, bands	dADA2b-containing complex	Pankotai *et al.,* 2005; Ebert *et al.,* 2006.
H3S10ph	Euchromatin, interbands	JIL-1	Jin *et al.,* 1999; Wang *et al.,* 2001; Ebert *et al.,* 2004; Zhang *et al.,* 2006.
H3K14ac	Euchromatin/bands	dGCN5/dADA2b	Cheung *et al.,* 2000; Pankotai *et al.,* 2005.
H3K27me	Heterochromatin	E(Z)/ESC/SU(Z)12 complex	Muller *et al.,* 2002; Ebert *et al.,* 2006.
H3K36me	Euchromatin/interbands	dSet2	Stabell *et al.,* 2007.
H3K79me	Euchromatin (puffs, interbands, some bands)	GRAPPA	Shanower *et al.,* 2005.
H4K5ac	Synticial blastoderm nuclei.	Undetermined	Ludlam *et al.,* 2002.
H4K8ac	Ubiquitos in embryonic nuclei.	Undetermined	Ludlam *et al.,* 2002.
H4K12ac	Heterochromatin, particular distribution in embryonic nuclei.	Undetermined	Turner *et al.*, 1992; Ludlam *et al.,* 2002; Swaminathan *et al.,* 2005.
H4K16ac	Euchromatin	CHAMEAU	Grienenberger *et al.,* 2002, Miotto *et al.,* 2006.
H4K16ac	Male X chromosome	MOF	Akhtar and Becker, 2000; Smith *et al.,* 2001.
H4K20me	Euchromatin, bands and interbands	PR-Set7, ASH1	Karachentsev *et al.*, 2005; Beisel *et al.,* 2002.
H4K20me3	Heterochromatin	SUV4-20	Schotta *et al.,* 2004.
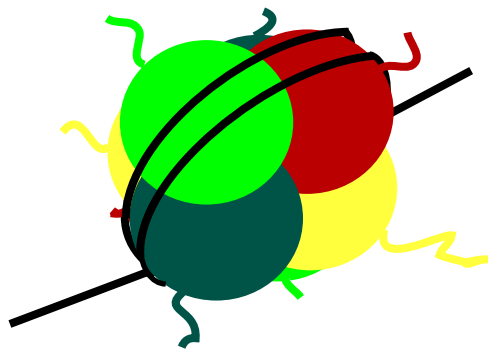
Histone H3

Histone H4

Histone H2B

Histone H2A

Figure 1.1. A schematic of the nucleosome core particle.

Figure 1.1. A schematic of the nucleosome core particle. The histone proteins are represented as solid circles. Their unstructured N-terminal tails, which comprise many of the residues targeted by histone-modifying enzymes, are shown as «squiggles» protruding from the core particle. Histones H2A and H2B are represented in light and dark green, respectively, H3 is represented in yellow and H4 in red (due to the perspective, only one H4 molecule is visible). The DNA double helix is represented as a plain black line.

1.2.1.1. DISTRIBUTION AND FUNCTIONS OF METHYLATED HISTONE LYSINES

Over a dozen lysine residues have been shown to be susceptible to methylation (reviewed by Kouzarides, 2007). These lysines reside mainly within the H3 and H4 histone proteins, but also on H2A and H2B. Lysines can be mono-, di-, or trimethylated (me1, me2, me3, respectively), giving rise to a large number of possible methylation states for each nucleosome. Distinct methylation states are observed not only in different regions of the genome (*i.e.* transcriptionally active *vs.* inactive), but also within different portions of a given gene. In addition, methylation of certain lysines seems to be incompatible with that of others, suggesting that the co-ordinated action of histone methyltransferases (HMTases) and demethylases (DMTases) is a critical factor in the regulation of gene expression.

Methylated H3K4 (H3K4me) is generally associated with transcriptionally active genes. It appears that H3K4me3 tends to be enriched within promoters and the 5'-most region of transcriptionally competent genes, while H3K4me2 is more abundant within the middle portion of active genes, and H3K4me1 is localized towards the 3'-most region (reviewed by Schones and Zhang, 2008). H3K36me3 is also associated with transcriptionally active chromatin (Barski *et al.*, 2007). However, unlike the H3K4me "marks", it seems to be less localized, and spread over the entire length of the transcribed regions and peaking at their 3' end (Bannister et al., 2005; Barski *et al.*, 2007). Also, while H3K4me3 is detected at the start of genes that are transcriptionally competent, but not necessarily transcribed, the presence of H3K36me3 seems to be restricted, at least in mammals, to genes that are actively transcribed (Mikkelsen *et al.*, 2007).

Trimethylated H3K9 and H4K20, on the other hand, are generally considered hallmarks of heterochromatin. In mammals, H3K9me3 and, to a lesser degree, H3K9me2, are associated with heterochromatin and also with silent regions of the genome, while H3K9me1 and H4K20me1 appear to be associated with active genes (Barski *et al*., 2007). In *Drosophila*, H3K9me1,2, H3K27me1,2,3 and H4K20me3 are all highly enriched in pericentric heterochromatin, but they are also present in other regions of the genome (Schotta *et al*., 2004; Ebert *et al.*, 2004; 2006). For instance, H3K9me1,2,3 are associated with a number of euchromatic sites and with telomeres and, in polytene chromosomes, H3K27me1,2 are associated with virtually all bands, which represent regions of the genome that are somewhat condensed (Ebert *et al*., 2006). H3K9me3, on the other hand, does not appear to be very abundant, and is only detected within the core of the chromocentre and at a few other sites (Ebert *et al*., 2004; 2006).

It is becoming increasingly apparent that these "methyl marks" may be involved in combinatorial, as well as in step-by-step mechanisms that regulate chromatin condensation/decondensation. In *Drosophila*, mutations in the H3K4 demethylase SU(VAR)3-3/dLSD1 are epistatic to the presence of additional copies of the H3K9 methyltransferase SU(VAR)3-9, indicating that demethylation of H3K4 must precede methylation of H3K9 (Rudolph *et al.*, 2007). However, the fact that no "methyl mark" is completely restricted to transcriptionally active or inactive chromatin, suggests that none of these modifications is sufficient, by itself, to determine the transcriptional state of a chromatin region.

The addition of methyl groups to lysine residues is catalyzed by histone methyltransferase enzymes (HMTases). The vast majority of HMTases characterized so far are non-histone chromatin proteins (NHCPs) containing the signature SET domain (see below). Differences within key residues of their catalytic region confer SET domain-containing HMTases their substrate specificity. A wealth of information is available about SET domain-containing proteins, and it will be discussed in section 1.2.2.1. Much less is known about histone demethylases, as their discovery is much more recent (Wang e*t al.,* 2004; Cuthberg *et al.,* 2004; Shi *et al.*, 2004; Tsukada *et al.,* 2006). So far, four histone lysine demethylases have been identified in *Drosophila*: SU(VAR)3-3/dLsd1, which demethylates H3K4me1,2, Lid, which appears to be specific for H3K4me3, and the JMJD2 homologs, dJMJD2(1) and (2) that demethylate H3K9me3 and H3K36me3, respectively (Rudolph *et al.,* 2007; Di Stefano *et al.,* 2007; Eissenberg *et al.,* 2007; Lee *et al.,* 2007; Secombe *et al.,* 2007; Lloret-Llinares *et al*., 2008).

Histone H3

Hs 1 artkqtarkstggkaprkqlatkaarksapatggvkkphryrpgtvalreirryqkstellirklpfqrlvreiaqd

Dm 1 artkqtarkstggkaprkqlatkaarksapatggvkkphryrpgtvalreirryqkstellirklpfqrlvreiaqd

Hs 78 fktdlrfqssavmalqeaseaylvglfedtnlcaihakrvtimpkdiqlarrirgera

Dm 78 fktdlrfqssavmalqeaseaylvglfedtnlcaihakritimpkdiqlarrirgera

Histone H4

Hs 1 sgrgkggkglgkggakrhrkvlrdniqgitkpairrlarrggvkrisgliyeetrgvlkvflenvirdavtytehak

Dm 1 tgrgkggkglgkggakrhrkvlrdniqgitkpairrlarrggvkrisgliyeetrgvlkvflenvirdavtytehak

Hs 78 rktvtamdvvyalkrqgrtlygfgg

Dm 78 rktvtamdvvyalkrqgrtlygfgg

12

Figure 1.2. Histone H3 and H4 post-translational modifications. Amino acid sequence of human (*H.s.*) and *Drosophila* (*D.m.*) histones H3 and H4. The sequences corresponding to H3 and H4's N-terminal tails are underlined in light blue, while yellow underlining denotes the sequences comprising the histone folds (according to Luger *et al.,* 1997). Formally identified post-translational modifications are represented as blue squares (methylation), green circles (phosphorylation) and red asterisks (acetylation) above the residues concerned. More details about the modifications identified in *Drosophila*, including all relevant references, are listed in Table 1.4. The diagrams relative to human H3 and H4 modifications are based on the Abcam/Millipore histone modification map (http://www.histone.com).

1.2.2. <u>Non-histone chromatin proteins</u>

Non-histone chromatin proteins (NHCPs) are generally defined as either "structural" components of chromatin or chromatin-modifying enzymes. The former include all the proteins that are physically associated with chromatin, but that do not appear to have catalytic functions, such as HP1, *Drosophila*'s SU(VAR)3-7 and *S. pombe*'s Rik1. Chromatin-modifying enzymes, in contrast, include histone acetyl-transferases and deacetylases (HATs and HDACs), HMTases and DMTases, histone kinases and phosphatases, and so forth. In many cases, these modifying enzymes also physically associate with chromatin and play a structural role in chromatin architecture. The next section will focus on HMTases and, in particular, SET domain-containing HMTases, since SU(VAR)3-9 belongs to this family of enzymes.

1.2.2.1.   SET DOMAIN-CONTAINING METHYLTRANSFERASES

The SET domain was originally identified as a ~140 amino acid region present and highly conserved in the gene products of <u>*Su(var)3-9*</u>, <u>*Enhancer of zeste*</u> and <u>*trithorax*</u> (Jones and Gelbart, 1993; Tschiersch et al., 1994). Dozens of SET-containing proteins have since been described in eukaryotes, and many of them appear to have HMTase or, more generally, protein lysine MTase activities (reviewed by Qian and Zhou, 2006). Table 1.2. summarizes the known *Drosophila* SET domain-containing HMTases and their substrate specificities.

SET domain-containing proteins can be classified into families based on phylogenetic analyses, on the presence of other protein domains (such as a

chromodomain, ankyrin repeats, zinc fingers, and so forth), and/or based on their substrate specificity. Most SET domain-containing HMTases can only catalyze the addition of methyl groups to one or two histone residues and, while some HMTases are strictly mono-methylases, others are able to catalyze the addition of multiple methyl groups (Eskeland *et al*., 2004; Xiao *et al*., 2005; Chin *et al*., 2006; Qian *et al*., 2006; Guo *et al*., 2007). The crystal structures of a dozen SET domains have been resolved (reviewed by Qian and Zhou, 2006), providing an explanation for their respective substrate specificities, and allowing investigators to rationalize the specificity of yet un-crystallized SET-containing HMTases.

In most cases, the SET domain is surrounded by a preSET and a postSET regions, which contain a number of conserved cysteines. These residues are not involved in the catalytic process, but they coordinate a set of zinc ions and are necessary to stabilize the structure of the SET domain (Min *et al*., 2002; reviewed in Qian and Zhou, 2006). The substrate (histone tail) and the methyl group/methyl donor complex (S-adenosyl-methyl-methionine) bind to two distinct clefts located at opposite sides of the SET domain. These two clefts are connected by a hydrophobic channel, and the substrate specificity of SET domains is likely determined by the side chains of the residues forming this channel (Qian and Zhou, 2006). The combination of side chains is thought to recognize particular residues that flank the target lysine in the substrate, which would explain why several HMTases also display lysine MTase activity on non-histone proteins (Kouskouti *et al*., 2004; Couture et al., 2006; Chin *et al.*, 2007; Sampath *et al*., 2007). In addition, the nature and size of the side chains in the hydrophobic

channel can determine the degree of methylation that can be achieved by a given SET domain. For example, the presence of a large, bulky side chain will not permit an already methylated histone tail access to the channel, thus precluding the transfer of multiple methyl groups on the target lysine.

SET-containing HMTases also have an additional layer of specificity, as it is becoming apparent that different HMTases act within different chromatin regions. In *Drosophila*, for example, SU(VAR)3-9 is responsible for H3K9me2,3 within centric and pericentric heterochromatin, as well as some euchromatic sites. In contrast, DmSetDB1 mono- and dimethylates H3K9 almost exclusively on the mainly heterochromatic fourth chromosome (Schotta *et al.,* 2002; Eskeland *et al.,* 2004; Mis *et al.,* 2006; Stabell *et al*., 2006; Seum *et al.,* 2007; Tzeng *et al.,* 2007; Ebert *et al.,* 2006, and references therein).

When this thesis was begun, the catalytic activity of one SET domain, that of SUV39H and CLR4 (the human and yeast homologs of SU(VAR)3-9, respectively) had just been discovered (Rea *et al*., 2000). Prior to that, SET domains were thought to have protein-protein interaction roles (for example Aagard *et al.,* 1999; Cui *et al.,* 1998), and no structure/function analyses had ever been performed.
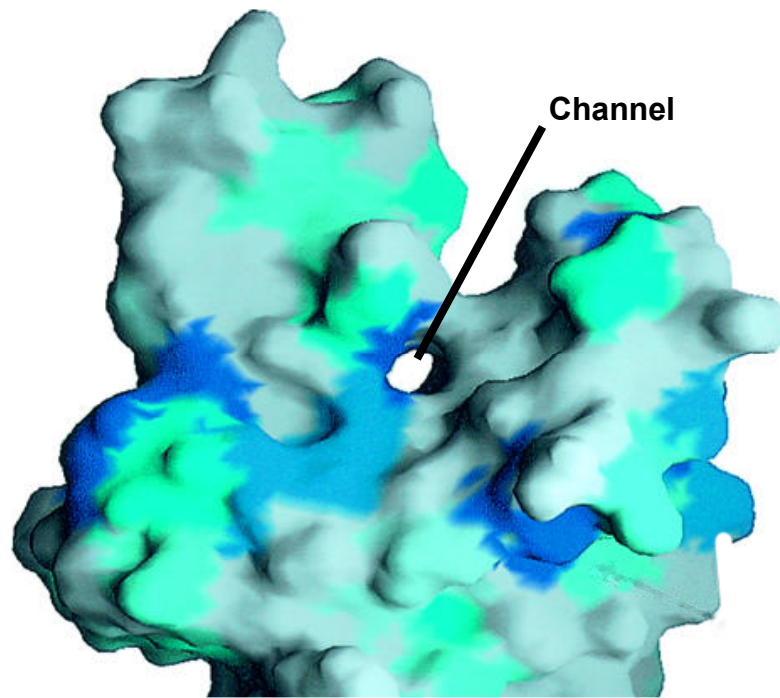
Channel

Figure 1.3.

Figure 1.3. Structure of a SET domain, in this case SET7/9. The substrate

(histone H3) binds to the shallow groove adjacent to the channel, on the side of the

protein that is facing the viewer in this figure. The methyl donor (AdoMet) binds to

a pocket on the opposite side of the protein (hidden in this figure). The narrow

channel indicated with an arrow allows the H3 tail to come in contact with the

AdoMet moiety. The blue shading indicates different levels of residue conservation

relative to other SET domain proteins. This figure was modified from Kwon *et al.*:

Mechanism of histone lysine methyl transfer revealed by the structure of SET7/9-

AdoMet. EMBO J *22*, 292-303 (2003).

Table 1.2. **A list of characterized *Drosophila* SET domain-containing proteins.** Note that

predicted/hypothetical proteins are not included.

Protein name	Activity and specificity reported	Selected references
SU(VAR)3-9	H3K9me (preferentially mono/dimethylation of H3K9me1)	Ebert *et al.,* 1994; Czermin *et al.,* 2001; Schotta *et al.*, 2002; Eskeland *et al.,* 2004.
E(Z)	H3K27me (as part of the E(Z)/ESC/SU(Z)12 complex)	Muller *et al.,* 2002; Ebert *et al.,* 2006.
TRITHORAX	H3K4me3	Smith *et al.,* 2004
TRR	H3K4me3	Sedkov *et al.,* 2003.
ASH1	H3K4me3, H3K9me3 H4K20me3	Beisel *et al.,* 2002; Byrd and Shearn, 2003.
ASH2	H3K4me3	Beltran *et al,* 2007.
SUV4-20	H4K20me3	Schotta *et al.,* 2004.
PR-Set7	H4K20me	Nishioka *et al.,* 2002; Karachentsev *et al.,* 2005.
dG9a	H3K9me1,2 H3K27me, H4K20me (*in vitro*)	Mis *et al.,* 2006; Stabell *et al.,* 2006.
DmSetDB1	H3K9me1,2	Seum *et al.,* 2007; Tzeng *et al.,* 2007.
dSet2	H3K36me2	Stabell *et al.,* 2007.

1.2.2.2.   SU(VAR)3-9

In *Drosophila*, SU(VAR)3-9, is one of the most prominent HMTases. It appears to be very conserved, with homologs in virtually every eukaryote from the fission yeast to mammals and plants (Krauss *et al*., 2006). SU(VAR)3-9, as well as most of its homologs, has two conserved domains: a SET domain (see above) and a chromodomain, which is also found in other chromatin proteins such as HP1 and Polycomb (Paro and Hogness, 1991). Chromodomains recognize and bind methylated histone residues, and each chromodomain seems to be very specific for one particular methyl-residue. For example, the chromodomain of HP1 binds to H3K9me (Bannister *et al.,* 2001; Lachner *et al.* 2001; Eskeland *et al.* 2007), that of PC binds to H3K27me, and it was recently shown that the chromodomain of CLR4, the yeast homolog of SU(VAR)3-9, binds to H3K9me (Zhang *et al.*, 2008). *Drosophila*'s SU(VAR)3-9 is 635 amino acids long, which is larger than its known homologs, and it is unique in that it contains an N-terminal domain responsible for dimerization (Eskeland *et al.,* 2004). The N-terminal moiety of SU(VAR)3-9 also comprises motifs that are required for interaction with two other chromatin proteins, HP1 and SU(VAR)3-7 (Schotta *et al.,* 2002).

In yeast, *clr4* mutants are viable, but show impaired silencing at centromeres and at the *mating type* locus (Ivanova *et al.,* 1998*)*. In mouse, homozygosity for knockout mutations in either *Suv39h1* or *Suv39h2* doesn't lead to any noticeable phenotypes (Peters *et al.,* 2001). However, a double knockout for both *Su(var)3-9* paralogs, *Suv39h1* and *Suv39h2*, is semi-lethal, and those few homozygous double mutant mice that survive display growth retardation , a high incidence of B cell lymphomas and chromosomal instability (Peters *et al.,* 2001).

In *Drosophila*, homozygous *Su(var)3-9* mutant adults are viable and fertile, but the chance of *Su(var)3-9* null embryos reaching adulthood is only about 50% of that of their wild-type counterparts (Mis *et al.,* 2006). These observations suggest that SU(VAR)3-9 is not absolutely essential for development, possibly because other HMTases are able to mimic some of SU(VAR)3-9's functions at least to some degree. On polytene chromosomes, SU(VAR)3-9 is detected at the chromocentre, as well as at some euchromatic sites on the chromosomal arms (Schotta *et al.,* 2002), including the HIS-C (Ner *et al.,* 2002). In addition to controlling methylation of H3K9 in heterochromatin, it interacts physically and functionally with a number of other chromatin proteins, including HP1 and the histone deacetylase HDAC1, which is responsible for deacetylating H3K9 (Schotta *et al.,* 2002; Ebert *et al.,* 2004; 2006; Czermin *et al.,* 2001). These and additional pieces of evidence obtained in yeast, have given rise to the current models for SU(VAR)3-9's role in chromatin-based gene silencing (see below, 1.3.1).

1.2.2.3. OTHER NON-HISTONE CHROMATIN PROTEINS IN <u>DROSOPHILA</u>

Lysine HMTases, and SU(VAR)3-9 in particular, are thought to work in conjunction with other NHCPs to form and maintain chromatin architecture. In *Drosophila*, the genes encoding many of these proteins were originally identified in genetic screens for dominant suppressors of position-effect variegation (PEV, reviewed by Spofford, 1976; Weiler and Wakimoto, 1995). Table 1.3 summarizes the genes originally identified through screens for dominant suppressors of PEV and the characteristics of their products, most of which are NHCPs. These include SU(VAR)3-7 and HDAC1, both of which are known to interact with SU(VAR)3-9 (Reuter *et al*., 1990; Cléard *et al*., 1997; Mottus *et al*., 2000; Czermin *et al*., 2001; Schotta *et al*., 2002).

SU(VAR)3-7 is a large, *Drosophila*-specific, zinc-finger protein that binds to repetitive DNA and physically interacts with both SU(VAR)3-9 and HP1 (Cléard and Spierer, 2001; Schotta *et al*., 2002; Jaquet *et al*., 2006). It is preferentially associated with heterochromatin and appears to be sufficient to induce SU(VAR)3-9-dependent heterochromatisation (Reuter *et al*., 1990; Cléard *et al*., 1995; Delattre *et al*., 2004). In contrast, localization of the histone deacetylase HDAC1, the homolog of *S. cerevisiae* Rpd3, does not appear to be heterochromatin-specific, as this protein is detected at hundreds of sites on polytene chromosomes (Chang *et al*, 2001; Tie *et al*., 2003). HDAC1 is an H3K9 deacetylase that has been shown to co-immunoprecipitate with SU(VAR)3-9, suggesting that the two proteins may be part of the same complex (Czermin *et al*., 2001; Rudolph *et al*., 2007). In addition, the deacetylation activity of HDAC1 is necessary for H3K9 methylation (Nakayama et al., 2001; Czermin et al., 2001; Vaute *et al*., 2002).

One NHCP, HP1, was originally identified as the specific target of a monoclonal antibody that associated preferentially with heterochromatin in polytene chromosomes (James and Elgin, 1986; James *et al.*, 1989) and *Su(var)2-5*, the gene encoding HP1, was identified as a strong dominant suppressor of PEV (Eissenberg *et al.*, 1990; 1992). HP1 has homologs in almost all eukaryotes from the fission yeast (Swi6) to mammals and plants; it is comprised of two conserved domains, the chromodomain and the chromo-shadow domain, linked by what is known as the "hinge" or "linker" region (reviewed by Lomberk *et al.*, 2006). The chromodomain is responsible for specifically recognizing and binding H3K9me, while the chromo-shadow domain is required for protein-protein interactions and for dimerisation (Bannister *et al.*, 2001; Lachner *et al.*, 2001; Jacobs *et al.*, 2001; Brasher *et al.*, 2000; Jones *et al.*, 2000; Nielsen *et al.*, 2001; Cowieson *et al.*, 2000). Both the chromo-shadow and the "hinge" domains are necessary for interaction with SU(VAR)3-9 and SU(VAR)3-7 (Schotta *et al.*, 2002; Jaquet *et al.*, 2002).

With the advent of genome-wide sequencing efforts, genomic databases and bioinformatics tools, a number of genes encoding NHCPs were cloned and characterized based on sequence homologies and *in silico* searches. These include two HMTase-encoding genes, *dG9a* and *Su(var)4-20*, as well as several DMTases (Schotta *et al.,* 2004; Mis et al., 2006; Stabell *et al.,* 2006; Schotta *et al.*, 2004; Secombe *et al.,* 2007; Lloret-Llinares *et al.*, 2008). Several NHCPs identified through an *in silico* approach have also been shown to suppress PEV, indicating that they are functionally involved in epigenetic silencing (Table 1.4).

Table 1.3. *Drosophila* **proteins identified through screens for strong, dominant suppressors of PEV.** «Location» refers to the cytological

location to which the gene in question has been mapped. Note that, under «homologs», hypothetical/predicted proteins are not included.

Gene name	Location	Gene product	Attributed functions	Homologs	Selected references
Su(var)2-5 (Sometimes referred to as *Su(var)205*)	28F2-F3	HP1 (Synonym: HP1a, C1A9 nuclear antigen)	Chromatin silencing and assembly, telomere maintenance, transcriptional regulation, chromosome segregation, binding to H3K9me, RNA-binding.	Orthologs: HP1a (mammals*), SWI6 (*S. pombe*), LHP1 (*Arabidopsis*), *M31 (mouse) Paralogs: HP1b, HP1c (*Drosophila*), HP1b, HP1g (mammals).	James and Elgin, 1986; Eissenberg *et al.*, 1992; Kellum and Alberts, 1995; Nielsen *et al.*, 2001; Greil *et al.*, 2003.
Su(var)2-10	45A8-A9	dPIAS* (Synonyms: Zimp, PIAS, SU(VAR)2-10) * multiple isomorphs	DNA binding, chromosome organization, structure and function, transcriptional regulation, regulation of the JAK-STAT cascade, haematopoiesis.	PIAS** (mammals) **several paralogs present in mammals.	Mohr and Boswell, 1999; Evans *et al.*, 2003; Hari *et al.*, 2001; Muller *at al.*, 2005;
Su(var)3-1 (Note that *Su(var)3-1* is a hypermorph allele of *JIL-1*)	68A5-A6	JIL-1	Histone kinase (specific for H3S10), regulation of H3K9methylation, chromatin architecture, chromosome segregation (meiosis), transcriptional regulation.	Undetermined	Jin *et al.*, 1999; Wang *et al.*, 2001; Ebert *et al.*, 2004; Bao *et al.*, 2005; Zhang *et al.*, 2006.
Su(var)3-3	77A3	SU(VAR)3-3 (Synonym: dLsd1)	Histone demethylase (specific for H3K4me1,2), gametogenesis, heterochromatin formation, transcriptional regulation.	LSD1 (mammals), SPR-5 (*C. elegans*), LSD1 (*S. pombe*).	Di Stefano *et al.*, 2007; Rudolph *et al.*, 2007.

Gene name	Location	Gene product	Attributed functions	Homologs	Selected references
Su(var)3-6	87B9-B10	PP1-87B (Synonyms: PP-1a, PP-1c, PP-1a, SU(VAR)3-6, CK19).	Ser/Thr phosphatase, cell cycle regulation/mitosis, glycogen metabolism, gametogenesis, wing, eye, nervous system and muscle development.	PP1 (vertebrates), TOPP2 (*Arabidopsis*), Ppz1p (*S. cerevisiae*).	Dombradi *et al.*, 1990; Baksa *et al.*, 1993; Bennett *et al.*, 2003; Babu *et al.*, 2005.
Su(var)3-7	87E3	SU(VAR)3-7	DNA-binding (preference for satellite sequences), interaction with HP1 and SU(VAR)3-9, chromatin compaction, architecture and silencing.	To date, no homologs identified outside the genus *Drosophila*.	Reuter *et al.*, 1990; Cléard and Spierer, 2001; Jaquet *et al.*, 2002; Delattre *et al.*, 2004; Demakova *et al.*, 2007.
Su(var)3-9 (Note: *pitkin*, a very strong *E(var)*, is a hypermorph allele of *Su(var)3-9*)	88E6-E8	SU(VAR)3-9	HMTase (specific for H3K9), gene regulation, chromatin architecture and silencing, interaction with HP1, SU(VAR)3-7, SU(VAR)3-3, HDAC1.	Orthologs: SUV39H1 (human), Suv39h1 (mouse), CLR4/Clr4p (*S. pombe*); DIM-5 (*N. crassa*). Paralogs: SUV39H2 (human), Suv39h2 (mouse). SUVH3,4,5 (*Arabidopsis*).	Tschiersch *et al.*, 1994; Ner *et al.*, 2002; Schotta *et al.*, 2002; Eskeland *et al.*, 2004; Swaminathan *et al.*, 2005.
Su(var)326/ Rpd3	64B12	HDAC1/RPD3	Histone deacetylase (preference for H3K9ac), interaction with SU(VAR)3-9, gene regulation, chromatin architecture, silencing, cell cycle regulation.	HDAC2/1 (mammals), Rpd3 (*S. cerevisiae*).	Maixner *et al.*, 1998; Mottus *et al.*, 2000; Huang and Kadonaga, 2001; Czermin *et al.*, 2001.

Table 1.4. Additional Su(var)s. Mutations in the following *Drosophila* genes were also found to act as dominant suppressors of PEV on at least two different variegating rearrangements. Note that the Su(var) phenotype associated *puc* and *dSas4* was determined in our lab (Toub, unpublished), and that, since *dG9a* and *Suv4-20* are located on the X chromosome, their Su(var) phenotypes have so far only been tested on one variegating rearrangement, Sb^V.

Gene name	Location	Gene product	Major attributed functions	Homologs	Selected references
dG9a	1A1	dG9a	Histone methyltransferase (preference for H3K9 and K27 in euchromatin), regulation of ecdysone pathways.	G9a (mammals).	Tachibana *et al.*, 2002; Mis *et al.*, 2006; Stabell *et al.*, 2006.
Suv4-20	1B13-14	SUV4-20	Histone methyltransferase (preference for H4K20), gene regulation, chromatin architecture and silencing.	Orthologs: Suv4-20h1 (mammals). Paralogs: Suv4-20h2 (mammals).	Schotta *et al.*, 2004.
chm	27F3-4	CHAMEAU (Synonym: HAT1)	HAT (preference for H4K16), chromatin architecture, silencing, regulation of the JNK cascade, transcriptional regulation.	HBO1 (human), Sas2 (S. cerevisiae).	Grienenberger *et al.*, 2002; Miotto *et al.*, 2006;
E(Pc)	47F13-14	E(PC)	Subunit of the Tip60 chromatin-remodeling complex.	Orthologs: EPC1 (human), Epc1 (mouse), Epl1 (S. cerevisiae). Paralogs: EPC2 (human), Epc2 (mouse).	Stankunas *et al.*, 1998; Sinclair *et al.*, 1998; Kusch *et al.*, 2004.
Su(UR)	68A4	SUUR	Chromatin architecture, regulation of genes arranged in clusters.	Undetermined	Belyaeva *et al.*, 1998; 2003; Belyakin *et al.*, 2005.

Gene name	Location	Gene product	Major attributed functions	Homologs	Selected references
dSas-4 (Previously known as S2214)	84C6-7	dSAS-4	Centriole replication, biogenesis of centrosomes, flagella and cilia.	SAS-4* (*C. elegans*). (*Spindle assembly 4)	Basto *et al.*, 2006.
puc	84E12-13	PUCKERED	Ser/Thr and Tyr phosphatase, regulation of JNK cascade, dorsal closure and metamorphosis.	Pyst1 (human); CEL-F081 (*C. elegans*).	Glise and Noselli, 1997; Martin-Blanco *et al.*, 1998; Agnès *et al.*, 1999.
mod	100E3	MODULO	Transcriptional regulation, gametogenesis, cell growth and proliferation, chromatin architecture.	Nucleolin (vertebrates), p67/NSR1 (*S. cerevisiae*).	Krejci *et al.*, 1989; Gazino *et al.*, 1992; Perrin *et al.*, 2003.

27

1.3. SU(VAR)3-9 and heterochromatin assembly

Heterochromatin is a transcriptionally inert form of chromatin that remains highly compacted throughout the cell cycle. Heterochromatin is characterized by the presence of repetitive DNA, transposable elements, highly ordered nucleosomal arrays (Wallrath and Elgin, 1995), specific subsets of histone modifications (*e.g.* H3K9me2,3, H4K20me3) and enrichment for particular NHCPs, such as SU(VAR)3-9, HP1, SU(VAR)3-7 (Eissenberg and Elgin, 2000; Schotta *et al.*, 2002; Reuter *et al.*, 1990; for a comprehensive review see Ebert *et al.,* 2006). Heterochromatic regions of the genome tend to be replicated late during S phase and appear underreplicated in polytene chromosomes. Chromosomes' centromeres and telomeres, the inactive X chromosome in female mammals and most of *Drosophila*'s fourth chromosome are heterochromatic.

However, chromatin is highly dynamic, and its packaging status, as well as the histone modifications and NHCPs present at a given locus, are not fixed; instead they change throughout the cell cycle, and during development, as the cell responds to cellular and extracellular signals (de Wit *et al.,* 2005; Ebert *et al.,* 2004; Dormann *et al.,* 2006). Results obtained from genetic and biochemical analyses in yeast, and from the study of polytene chromosomes in flies suggest that heterochromatin formation is a multi-step process. The model presented in section 1.3.1 focuses on the roles of SU(VAR)3-9, based on what is known from *Drosophila* and yeast.

1.3.1. <u>Postulated mechanism for heterochromatin formation</u>

Demethylation of H3K4me by SU(VAR)3-3/dLSD1 is thought to be one of the first steps in heterochromatisation of a region, since a loss of function mutation in *Su(var)3-3* is epistatic to the presence of extra copies of *Su(var)3-9*, and SU(VAR)3-9's activity is absolutely necessary for the formation of heterochromatin (Rudolph *et al*., 2007; Schotta *et al*., 2002; 2003). Deacetylation of H3K9ac by HDAC1 also needs to occur prior to SU(VAR)3-9-dependent methylation, as H3K9 can't be simultaneously acetylated and methylated. In addition, loss of function mutations in the gene encoding HDAC1 are also epistatic to the presence of extra copies of *Su(var)3-9* (Czermin *et al*., 2001).

Once H3K4 is demethylated and H3K9 is deacetylated, SU(VAR)3-9 can di- and trimethylate H3K9. It should however be noted that, since SU(VAR)3-9 seems to preferentially di- and trimethylate H3K9me1, its action is probably preceded by that of an H3K9 monomethylase, possibly dG9a (Mis *et al*., 2006). How SU(VAR)3-9 is recruited to its target sites is not completely understood, but we know that its recruitment is at least partially dependent on SU(VAR)3-7, and that the presence of HP1 is required to prevent the binding of SU(VAR)3-9 to ectopic sites (Delattre *et al*., 2004; Schotta *et al*., 2002; 2003). Since SU(VAR)3-9 is able to physically interact with HDAC1, SU(VAR)3-7 and HP1, it is possible that these chromatin proteins form a complex (Delattre *et al*., 2000; Czermin *et al*., 2001, Schotta *et al*., 2002; 2003).

The H3K9me2,3 probably acts as a binding platform for the chromodomain of HP1 (Bannister *et al*. 2001; Jacobs *et al*., 2001). There is also some evidence

showing that, at least in yeast, H3K9me2,3 helps recruit and stabilize the binding of Clr4, the yeast homolog of SU(VAR)3-9, to chromatin (Zhang *et al*., 2008). If the same is true in *Drosophila*, the presence of H3K9me2,3 may stabilize the association of SU(VAR)3-9 with chromatin, in addition to facilitating, with the help of auxiliary factors, the binding of HP1. This, in turn, is thought to recruit SU(VAR)4-20, an HMTase that trimethylates H4K20 (Ebert *et al*., 2006; Eskeland *et al*., 2007; Schotta *et al*., 2004).

SU(VAR)3-9 has also been associated with euchromatic gene regulation (Vandel *et al.,* 2001; Nielsen *et al.,* 2001; Ner *et al.,* 2002; Greil *et al.,* 2003). Because of its role in heterochromatin formation, the presence of SU(VAR)3-9 at euchromatic loci is hypothesized to be associated with silencing of the genes within these loci. Whether this is always the case, or not, has not been well documented. In addition, it is presently unclear whether SU(VAR)3-9 acts through similar or completely distinct mechanisms at heterochromatic and euchromatic sites. Investigating its role in the regulation of the *histone* genes, which are located in euchromatin should provide useful information in this regard, and is the subject of this thesis.
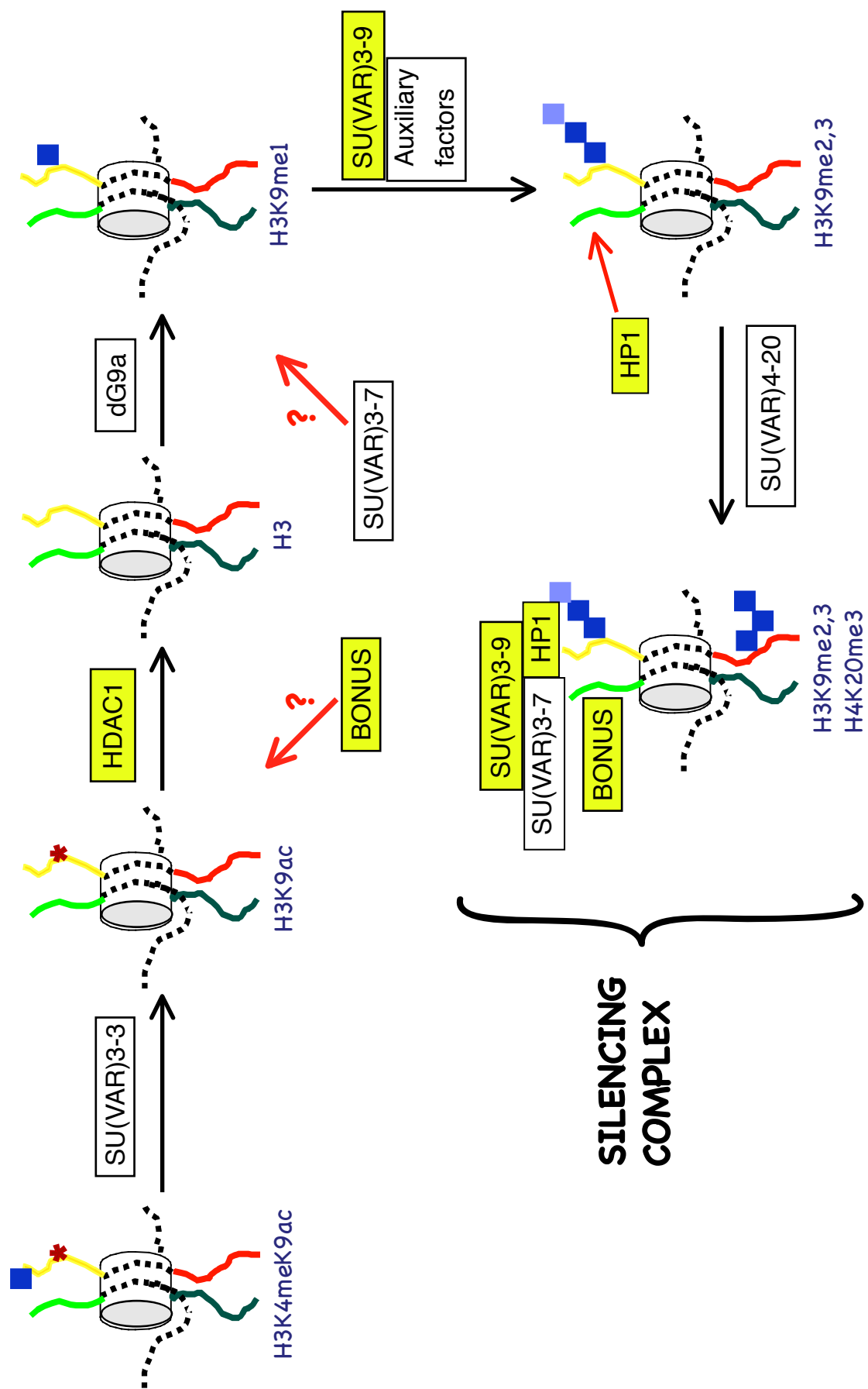
Figure 1.4.

Figure 1.4. A simplified diagram showing the assembly of heterochromatin.

The nucleosome core is represented as a cylinder, with the dotted line representing the DNA wrapped around it. For simplicity, only 4 N-terminal histone tails are shown instead of 8; the histone H3 tail in yellow, H4 in red, H2A in light green and H2B in dark green. The blue squares indicate methyl groups, the red asterisks represent acetyl groups. The enzymes known to catalyze the addition and removal of these groups are shown. Proteins known to be associated with the HIS-C are highlighted in yellow.

1.4. The *histone gene* cluster (HIS-C)

One of the targets of SU(VAR)3-9 in polytene chromosomes is the HIS-C, and some evidence exists that the regulation of the *histone* genes is, at least partially, dependent on SU(VAR)3-9 (Ner *et al.,* 2002; see below), thus making the HIS-C a very good model system to start dissecting the SU(VAR)3-9's mechanism of action at a euchromatic locus.

1.4.1. Structure of the HIS-C and characteristics of the *histone* genes

Drosophila's *histone* genes are organized into a "*histone unit*" consisting of one copy of each of the five *histone* genes (*h2a*, *h2b*, *h3*, *h4* and *h1*). About 110 tandemly repeated copies of this histone unit form the *Histone Gene cluster* (HIS-C). The order of the *histone* genes and their direction of transcription are shown in Figure 1.2. Located on the left arm of the second chromosome, *Drosophila*'s HIS-C spans over 500 kb (Saigo *et al.,* 1981) and has some peculiar characteristics. In spite of being a euchromatic locus, it displays some features that are usually associated with heterochromatin. It replicates slightly later in S-phase than other euchromatic loci, it is somewhat underreplicated in polytene chromososomes, nuclease sensitivity assays suggest that it is packaged as a higher order chromatin structure, and, at DNA level, it is a reiterated locus (Zhimulev and Belyaeva, 2003; Samal *et al.,* 1981).

In metazoans, *histone* genes differ from most genes in that they are intronless and, more importantly, their mRNAs are, for the most part, not polyadenylated in spite of being transcribed by RNA polymerase II. Instead, the

processing of *histone* pre-mRNAs depends on two elements located at their 3'

end: a highly conserved, 16 nucleotide stem-loop sequence, and a purine-rich

sequence, known as HDE (<u>h</u>istone <u>d</u>ownstream <u>e</u>lement). The processing also

requires several trans-acting factors: the stem-loop binding protein (SLBP), which

binds to the stem-loop structure, a U7 snRNP, which binds the HDE, plus some

additional (poorly characterized) factors (for a review, see for example Dominski

and Marzluff, 1999; 2007). Relatively little is known about the precise functions of

SLBP, and much of the data available comes from studies in vertebrates. We

know that SLBP remains associated with mature *histone* mRNAs as they are

transported into the cytoplasm, where it is thought to play a role in the stability and

translation of the transcripts (reviewed by Dominski and Marzluff, 1999; 2007).

More importantly, SLBP is only present in large amounts during S-phase, possibly

explaining why non-polyadenylated *histone* mRNAs only accumulate at this stage

of the cell cycle. One of its main roles may be to stabilize and protect the *histone*

mRNA from degradation (reviewed by Dominski and Marzluff, 1999; 2007).
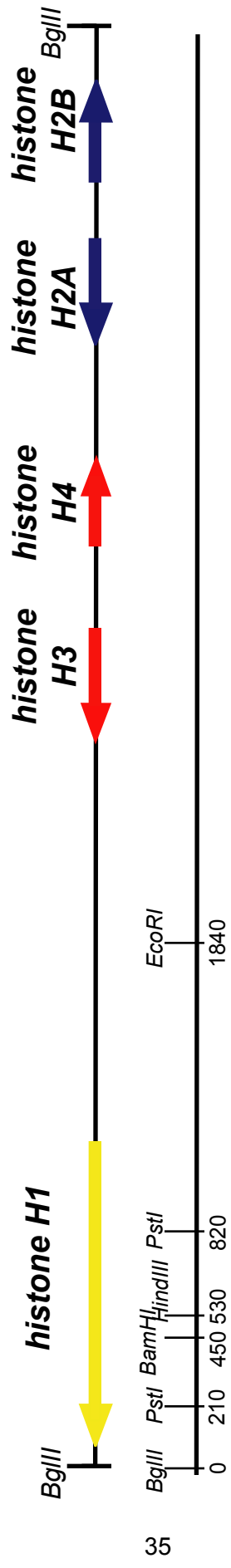
Figure 1.5.

35

Figure 1.5. A schematic of *Drosophila*'s *histone* unit. The *BglII* fragment

defining the *histone* unit is comprised of one copy of each *histone* gene. The

histone gene cluster (HIS-C) consists of a tandem array of about 110 copies of the

histone unit.The coding regions of the *histone* genes are represented in blue (H2A

and H2B), red (H3 and H4), and yellow (H1). The arrowheads indicate the direction

of transcription. Selected restriction sites, and their relative distances and positions

along the *histone* unit, are indicated.

1.4.2. H3K9me and SU(VAR)3-9 distribution across the HIS-C

Several pieces of evidence suggest that SU(VAR)3-9 must have a role in the regulation of the *histone gene cluster*. Based on MNase and DNaseI sensitivity assays, it appears that the HIS-C has an altered chromatin structure in *Su(var)3-9* mutants, and at least three *Su(var)3-9* mutants have increased levels of *histone h1* and *histone h4* transcripts (Ner *et al.*, 2002).

ChIP experiments have shown that SU(VAR)3-9 is associated with transcribed and non transcribed/ intergenic regions of the *histone unit* in staged embryos (Ner *et al*., 2002; Ner *et al*., in preparation), indicating that it probably plays a direct role in the regulation of the locus. The association of SU(VAR)3-9 with the HIS-C was confirmed in adults using a SU(VAR)3-9::DAM fusion protein (Ner *et al*., in preparation). In addition, localization of SU(VAR)3-9 at the HIS-C was also reported in *Drosophila* nurse cells (Koryakov *et al*., 2006). H3K9me2 follows the same distribution pattern as SU(VAR)3-9 along the histone unit, suggesting that SU(VAR)3-9 is not only present at the locus, but it is also enzymatically functional (Ner *et al*., 2002). Other NHCPs known to interact with SU(VAR)3-9, such as HP1 and HDAC1, have also been detected at the HIS-C (Greil *et al*., 2003; Koryakov *et al*., 2006; Ner *et al*., in preparation).

1.4.3. Possible function of SU(VAR)3-9 in the regulation of the *histone* genes

What could be the function of SU(VAR)3-9 in the regulation of *histone* gene expression? First of all, a role at the transcriptional level seems most compatible

with its nature as a chromatin-associated H3K9 HMTase. Assuming that this is the case, SU(VAR)3-9 could be involved in modulating the chromatin structure, and therefore transcriptional competency, of the HIS-C. It could do so in at least three ways. The first possibility is that it acts on individual *histone units*, determining which/how many of the ~110 of them are available for transcription. This possibility implies that the five *histone* genes comprising a given unit are all either accessible, or inaccessible to the transcriptional machinery. Alternatively, SU(VAR)3-9 could work as a transcriptional regulator that modulates the expression level of each *histone* gene independently. In this case, genes belonging to the same *histone* unit may be transcribed at different rates.

Finally, it is also possible that SU(VAR)3-9's function is somehow involved in the coupling of *histone* genes expression with the cell cycle. Chapter 3 of this thesis will address some of these issues.

1.5. Subject(s) of this study

Chromatin structure and dynamics play crucial roles in gene regulation. Thus, the importance of identifying chromatin components and understanding their functions and interaction is obvious. However, chromatin's size and complexity make it necessary to study its components one, or a few, at a time. The focus of this study is a non-histone chromatin protein, the H3K9-specific HMTase SU(VAR)3-9, and its role in the packaging and regulation of a euchromatic locus, the *histone genes cluster* .

Highly conserved from yeast to mammals (Krauss *et al.*, 2006), SU(VAR)3-9

is very well characterized. However, when this study was begun, the effects of

single amino acid substitutions on the protein's HMTase activity were not known.

Subsequently, the enzymatic activity of a dozen missense mutants has been

reported, but only with respect to a single synthetic substrate (a histone tail

peptide), which does not allow for a detailed characterization of the catalytic

characteristics of each mutant (Ebert *et al.*, 2004). Also, the mutants' phenotypes

at the enzymatic and cellular level, and their effect on chromatin-based gene

silencing, have never been systematically related to the position of the mutated

residue within the three dimensional structure of SU(VAR)3-9. Part of the reason

for this lack of detailed structure-function analyses of SU(VAR)3-9 may be that its

crystal structure has not yet been resolved, so one needs to "extrapolate" from

structural information relative to similar HMTases.

Finally, this thesis represents the first attempt to determine whether the effect(s) of

specific *Su(var)3-9* mutations on one phenotype (for example, catalytic activity on

a given substrate) are good predictors of how these mutations will affect other

phenotypes (for example, overexpression of the *histone* genes).

The first section of my thesis (chapter 2) is concerned with the

characterization of a series of *Su(var)3-9* missense alleles isolated in our

laboratory as strong, dominant suppressors of position-effect variegation (PEV,

see below, section 1.2.3.1.) at a number of different phenotypic levels, including:

the molecular, enzymatic, and cellular levels, as well as at their morphological

phenotype, suppression of PEV, the phenotype against which the mutants were

identified and isolated. Taking advantage of the crystal structure of CLR4 (Min *et al.,* 2002), the yeast homolog of SU(VAR)3-9, and of the high level of identity between the primary sequences of the two proteins, the results are then discussed in terms of structure-function. This analysis allows us to propose roles for the various mutated amino acids in the HMTase function of SU(VAR)3-9. In addition, it highlights the functional biases of the original genetic screen that was used to identify *Su(var)3-9* as well as the interpretive biases that result from characterizing mutants against a single molecular or cellular phenotype. Based on this molecular, biochemical and functional analyses we can postulate a possible mechanism by which SU(VAR)3-9 acts as a regulator of chromatin-based gene silencing (see "Discussion").

The second part of this thesis (chapter 3) deals with the role of SU(VAR)3-9 as a regulator of euchromatic gene expression. Although it was first identified as a chromatin protein, and found to be associated mainly with heterochromatic regions of the genome (Schotta *et al*., 2002; 2003; Ebert *et al*., 2004), there is reason to believe that SU(VAR)3-9 also plays an important role in the regulation of many euchromatic genes (Nielsen *et al.,* 2001; Greil *et al.,* 2003). Among the pieces of evidence supporting this idea is the relationship between *Su(var)3-9* function and regulation of the HIS-C transcription. The HIS-C appears to be one of the major euchromatic targets of SU(VAR)3-9, and at least two *Su(var)3-9* mutants show an altered chromatin structure at the HIS-C locus and elevated levels of *histone* transcripts (Ner *et al.,* 2002).

Assuming that SU(VAR)3-9 does indeed play a direct role in the regulation of *histone genes* expression, its mechanism of action is not known. In order to gain some insights on the function(s) of SU(VAR)3-9 at the HIS-C, its distribution along this locus is studied in wild-type and particular *Su(var)3-9* mutants. The distributions of dimethyl-H3K9 (H3K9me2), the histone modification elicited by SU(VAR)3-9, and that of HP1, another chromatin protein thought to bind to H3K9me2, are also analyzed. This study gives us a general picture of the HIS-C "landscape" in terms of SU(VAR)3-9, HP1 and H3K9me2, and allows us to determine the interdependence (or lack thereof) among these three factors. Finally, the relative amounts of *histone* transcripts are accurately quantified in the different *Su(var)3-9* mutants and, by relating this information to the distribution of SU(VAR)3-9, HP1 and H3K9me2 at the HIS-C in each mutant, we can propose a mechanism for SU(VAR)3-9's function at the HIS-C (chapter 3).

1.6. References

Aagaard, L., Laible, G., Selenko, P., Schmid, M., Dorn, R., Schotta, G., Kuhfittig, S., Wolf, A., Lebersorger, A., Singh, P.B., Reuter, G., Jenuwein, T.: Functional mammalian homologues of the *Drosophila* PEV-modifier *Su(var)3-9* encode centromere-associated proteins which complex with the heterochromatin component M31. EMBO J *18*, 1923-3 (1999)

Agnès, F., Suzanne, M., Noselli, S.: The *Drosophila* JNK pathway controls the morphogenesis of imaginal discs during metamorphosis. Development *126*, 5453-62 (1999)

Akhtar, A., Becker, P.B.: Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in *Drosophila*. Mol Cell *5*, 367-75 (2000)

Ausio, J., Levin, D.B., De Amorim, G.V., Bakker, S., Macleod. P.M.: Syndromes of disordered chromatin remodeling. Clin. Genet. *64*, 83-95 (2003)

Babu, K., Bahri, S., Alphey, L., Chia, W.: Bifocal and PP1 interaction regulates targeting of the R-cell growth cone in *Drosophila*. Dev Biol. *288*, 372-86 (2005)

Baksa, K., Morawietz, H., Dombradi, V., Axton, M., Taubert, H., Szabo, G., Torok, I., Udvardy, A., Gyurkovics, H., Szoor, B., *et al*.: Mutations in the protein phosphatase 1 gene at 87B can differentially affect suppression of position-effect variegation and mitosis in *Drosophila melanogaster.* Genetics *135*, 117-25 (1993)

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., Kouzarides, T.: Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120-4 (2001)

Bannister, A.J., Schneider, R., Myers, F.A., Thorne, A.W., Crane-Robinson, C., Kouzarides, T.: Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. J Biol Chem. *280*, 17732-6 (2005)

Bao, X., Zhang, W., Krencik, R., Deng, H., Wang, Y., Girton, J., Johansen, J., Johansen, K.M.: The JIL-1 kinase interacts with lamin Dm0 and regulates nuclear lamina morphology of *Drosophila* nurse cells. J Cell Sci. *118*, 5079-87 (2005)

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., Zhao, K.: High-resolution profiling of histone methylations in the human genome. Cell *129*, 823-37 (2007)

Basto, R., Lau, J., Vinogradova, T., Gardiol, A., Woods, C.G., Khodjakov, A., Raff, J.W.: Flies without centrioles. Cell *125*, 1375-86 (2006)

Beisel, C., Imhof, A., Greene, J., Kremmer, E., Sauer, F.: Histone methylation by the *Drosophila* epigenetic transcriptional regulator Ash1. Nature *419*, 857-62 (2002)

Beltran, S., Angulo, M., Pignatelli, M., Serras, F., Corominas, M.: Functional dissection of the ash2 and ash1 transcriptomes provides insights into the transcriptional basis of wing phenotypes and reveals conserved protein interactions. Genome Biol. 8, R67 (2007)

Belyaeva, E.S., Zhimulev, I.F., Volkova, E.I., Alekseyenko, A.A., Moshkin, Y.M., Koryakov, D.E.: *Su(UR)ES*: a gene suppressing DNA underreplication in intercalary and pericentric heterochromatin of *Drosophila melanogaster* polytene chromosomes. Proc Natl Acad Sci U S A *95*, 7532-7 (1998)

Belyaeva, E.S., Boldyreva, L.V., Volkova, E.I., Nanayev, R.A., Alekseyenko, A.A., Zhimulev, I.F.: Effect of the *Suppressor of Underreplication* (*SuUR*) gene on position-effect variegation silencing in *Drosophila melanogaster.* Genetics *165*, 1209-20 (2003)

Belyakin, S.N., Christophides, G.K., Alekseyenko, A.A., Kriventseva, E.V., Belyaeva, E.S., Nanayev, R.A., Makunin, I.V., Kafatos, F.C., Zhimulev, I.F.: Genomic analysis of *Drosophila* chromosome underreplication reveals a link between replication control and transcriptional territories. Proc Natl Acad Sci U S A *102*, 8269-74 (2005)

Bennett, D., Ször, B., Gross, S., Vereshchagina, N., Alphey, L.: Ectopic expression of inhibitors of Protein Phosphatase Type 1 (PP1) can be used to analyze roles of PP1 in *Drosophila* development. Genetics *164*, 235-45 (2003)

Bernstein, B.E., Humphrey, E.L., Erlich, R.L., Schneider, R., Bouman, P., Liu, J.S., Kouzarides, T., Schreiber, S.L.: Methylation of histone H3 Lys 4 in coding regions of active genes. Proc Natl Acad Sci U S A *99*, 8695-700 (2002)

Brasher, S.V., Smith, B.O., Fogh, R.H., Nietlispach, D., Thiru, A., Nielsen, P.R., Broadhurst, R.W., Ball, L.J., Murzina, N.V., Laue, E.D.: The structure of mouse HP1 suggests a unique mode of single peptide recognition by the shadow chromo domain dimer. EMBO J *19*, 1587-97 (2000)

Byrd, K.N., Shearn, A.: ASH1, a *Drosophila* trithorax group protein, is required for methylation of lysine 4 residues on histone H3. Proc Natl Acad Sci U S A *100*, 11535-40 (2003)

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., Zhang, Y.: Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science *298*, 1039-43 (2002)

Chang, Y.L., Peng, Y.H., Pan, I.C., Sun, D.S., King, B., Huang, D.H.: Essential role of *Drosophila* Hdac1 in homeotic gene silencing. Proc Natl Acad Sci U S A *98*, 9730-5 (2001)

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

Chin, H.G., Estève, P.O., Pradhan, M., Benner, J., Patnaik, D., Carey, M.F., Pradhan, S.: Automethylation of G9a and its implication in wider substrate specificity and HP1 binding. Nucleic Acids Res *35*, 7313-23 (2007)

Chodaparambil, J.V., Barbera, A.J., Lu, X., Kaye, K.M., Hansen, J.C., Luger, K.: A charged and contoured surface on the nucleosome regulates chromatin compaction. Nat Struct Mol Biol *14*, 1105-7 (2007)

Cléard, F., Matsarskaia, M., Spierer, P.: The modifier of position-effect variegation *Suvar(3)7* of *Drosophila*: there are two alternative transcripts and seven scattered zinc fingers, each preceded by a tryptophan box. Nucleic Acids Res *23*, 796-802 (1995). *Erratum in*: Nucleic Acids Res *23*, 3804 (1995)

Cléard, F., Delattre, M., Spierer, P.: SU(VAR)3-7, a *Drosophila* heterochromatin-associated protein and companion of HP1 in the genomic silencing of position-effect variegation. EMBO J. *16*, 5280-8 (1997)

Cléard, F., Spierer, P.: Position-effect variegation in *Drosophila*: the modifier *Su(var)3-7* is a modular DNA-binding protein. EMBO Rep. *2*, 1095-100 (2001)

Couture, J.F., Hauk, G., Thompson, M.J., Blackburn, G.M., Trievel, R.C.: Catalytic roles for carbon-oxygen hydrogen bonding in SET domain lysine methyltransferases. J Biol Chem *281*, 19280-7 (2006)

Cowieson, N.P., Partridge, J.F., Allshire, R.C., McLaughlin, P.J.: Dimerisation of a chromo shadow domain and distinctions from the chromodomain as revealed by structural analysis. Curr Biol *10*, 517-25 (2000)

Crawford, G.E., Davis, S., Scacheri, P.C., Renaud, G., Halawi, M.J., Erdos, M.R., Green, R., Meltzer, P.S., Wolfsberg, T.G., Collins, F.S.: DNase-chip: a high-resolution method to identify DNase I hypersensitive sites using tiled microarrays. Nat Methods *3*, 503-9 (2006)

Cui, X., De Vivo, I., Slany, R., Miyamoto, A., Firestein, R., Cleary, M.L.: Association of SET domain and myotubularin-related proteins modulates growth control. Nat Genet *18*, 331-7 (1998)

Cuthbert, G.L., Daujat, S., Snowden, A.W., Erdjument-Bromage, H., Hagiwara, T.,

Yamada, M., Schneider, R., Gregory, P.D., Tempst, P., Bannister, A.J., Kouzarides, T.: Histone deimination antagonizes arginine methylation. Cell *118*, 545-53 (2004)

Czermin, B., Schotta, G., Hulsmann, B.B., Brehm, A., Becker, P.B., Reuter, G., Imhof, A.: Physical and functional association of SU(VAR)3-9 and HDAC1 in *Drosophila*. EMBO Rep. *2*, 915-9 (2001)

de la Cruz, X., Lois, S., Sánchez-Molina, S., Martínez-Balbás, M.A.: Do protein motifs read the histone code? Bioessays *27*, 164-7 (2005)

Delattre, M., Spierer, A., Tonka, C-H., Spierer, P.: The genomic silencing of position-effect variegation in Drosophila melanogaster: interaction between the heterochromatin-associated proteins Su(var)3-7 and HP1. J Cell Sci *113*, 4253-61 (2000).

Delattre, M., Spierer, A,. Jaquet, Y., Spierer, P.: Increased expression of *Drosophila Su(var)3-7* triggers *Su(var)3-9*-dependent heterochromatin formation. J Cell Sci. *117*, 6239-47 (2004)

Demakova, O.V., Pokholkova, G.V., Kolesnikova, T.D., Demakov, S.A., Andreyeva, E.N., Belyaeva, E.S., Zhimulev, I.F.: The SU(VAR)3-9/HP1 complex differentially regulates the compaction state and degree of underreplication of X chromosome pericentric heterochromatin in *Drosophila melanogaster*. Genetics *175*, 609-20 (2007)

de Wit, E., Greil, F., van Steensel, B.: Genome-wide HP1 binding in *Drosophila*: developmental plasticity and genomic targeting signals. Genome Res. *15*, 1265-73 (2005)

Di Stefano, L., Ji, J.Y., Moon, N.S., Herr, A., Dyson, N.: Mutation of *Drosophila* Lsd1 disrupts H3-K4 methylation, resulting in tissue-specific defects during development. Curr Biol *17*, 808-12 (2007)

Dominski, Z., Marzluff, W.F.: Formation of the 3' end of histone mRNA. Gene *239*, 1-14 (1999)

Dominski, Z., Marzluff, W.F.: Formation of the 3' end of histone mRNA: getting closer to the end. Gene *396*, 373-390 (2007)

Dombradi, V., Axton, J.M., Brewis, N.D., da Cruz e Silva, E.F., Alphey, L., Cohen, P.T.: *Drosophila* contains three genes that encode distinct isoforms of protein phosphatase 1. Eur J Biochem. *194*, 739-45 (1990)

Dormann, H.L., Tseng, B.S., Allis, C.D., Funabiki, H., Fischle, W.: Dynamic regulation of effector protein binding to histone modifications: the biology of HP1

switching. Cell Cycle *5*, 2842-51 (2006)

Ebert, A., Schotta, G., Lein, S., Kubicek, S., Krauss, V., Jenuwein, T., Reuter, G.: *Su(var)* genes regulate the balance between euchromatin and heterochromatin in *Drosophila.* Genes Dev *18*, 2973-83 (2004)

Ebert, A., Lein, S., Schotta, G., Reuter, G.: Histone modification and the control of heterochromatic gene silencing in *Drosophila.* Chromosome Res. *14*, 377-92 (2006)

Eskeland, R., Czermin, B., Boeke, J., Bonaldi, T., Regula, J.T., Imhof, A.: The N-terminus of *Drosophila* SU(VAR)3-9 mediates dimerization and regulates its methyltransferase activity. Biochemistry *43*, 3740-9 (2004)

Eskeland, R., Eberharter, A., Imhof, A.: HP1 binding to chromatin methylated at H3K9 is enhanced by auxiliary factors. Mol Cell Biol. *27*, 453-65 (2007)

Eissenberg, J.C., James, T.C., Foster-Hartnett, D.M., Hartnett, T., Ngan, V., Elgin, S.C.: Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in *Drosophila melanogaster.* Proc Natl Acad Sci U S A *87*, 9923-7 (1990)

Eissenberg JC, Morris GD, Reuter G, Hartnett T.: The heterochromatin-associated protein HP-1 is an essential protein in *Drosophila* with dosage-dependent effects on position-effect variegation. Genetics *131*, 345-52 (1992)

Eissenberg, J.C., Elgin, S.C.: The HP1 protein family: getting a grip on chromatin. Curr Opin Genet Dev *10*, 204-10 (2000)

Eissenberg, J.C., Lee, M.G., Schneider, J., Ilvarsonn, A., Shiekhattar, R., Shilatifard, A.: The trithorax-group gene in *Drosophila* little imaginal discs encodes a trimethylated histone H3 Lys4 demethylase. Nat Struct Mol Biol. *14*, 344-6 (2007)

Evans, C.J., Hartenstein, V., Banerjee, U.: Thicker than blood: conserved mechanisms in *Drosophila* and vertebrate hematopoiesis. Dev Cell. *5*, 673-90 (2003)

Finch, J.T., Rhodes, D., Brown, R.S., Rushton, B., Levitt, M., Klug, A.: Structure of nucleosome core particles of chromatin. *Nature* 269, 29-36 (1977)

Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D., Khorasanizadeh, S.: Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes Dev. *17*, 1870-81 (2003)

Garzino, V., Pereira, A., Laurenti, P., Graba, Y., Levis, R.W., Le Parco, Y., Pradel, J.: Cell lineage-specific expression of modulo, a dose-dependent modifier of variegation in *Drosophila.* EMBO J. *11*, 4471-9 (1992)

Glise, B., Noselli, S.: Coupling of Jun amino-terminal kinase and Decapentaplegic signaling pathways in *Drosophila* morphogenesis. Genes Dev. *11*, 1738-47 (1997)

Greil, F., van der Kraan, I., Delrow, J., Smothers, J.F., de Wit, E., Bussemaker, H.J., van Driel, R., Henikoff, S., van Steensel, B.: Distinct HP1 and Su(var)3-9 complexes bind to sets of developmentally coexpressed genes depending on chromosomal location. Genes Dev. *17*, 2825-38 (2003)

Grienenberger, A., Miotto, B., Sagnier, T., Cavalli, G., Schramke, V., Geli, V., Mariol, M.C., Berenger, H., Graba, Y., Pradel, J.: The MYST domain acetyltransferase Chameau functions in epigenetic mechanisms of transcriptional repression. Curr Biol. *12*, 762-6 (2002)

Guo, H.B., Guo, H.: Mechanism of histone methylation catalyzed by protein lysine methyltransferase SET7/9 and origin of product specificity. Proc Natl Acad Sci U S A *104*, 8797-80 (2007)

Hari, K.L., Cook, K.R., Karpen, G.H.: The *Drosophila Su(var)2-10* locus regulates chromosome structure and function and encodes a member of the PIAS protein family. Genes Dev. *15,* 1334-48 (2001)

Huang, X., Kadonaga, J.T.: Biochemical analysis of transcriptional repression by Drosophila histone deacetylase 1. J Biol Chem. *276*, 12497-500 (2001)

Jacobs, S.A., Taverna, S.D., Zhang, Y., Briggs, S.D., Li, J., Eissenberg, J.C., Allis, C.D., Khorasanizadeh, S.: Specificity of the HP1 chromo domain for the methylated N-terminus of histone H3. EMBO J. *20*, 5232-41 (2001)

Jaffe, L.F.: Epigenetic theories of cancer initiation. Adv Cancer Res. *90*, 209-30 (2003)

James, T.C., Elgin, S.C.: Identification of a nonhistone chromosomal protein associated with heterochromatin in *Drosophila melanogaster* and its gene. Mol Cell Biol. *6*, 3862-72 (1986)

James, T.C., Eissenberg, J.C., Craig, C., Dietrich, V., Hobson, A., Elgin, S.C.: Distribution patterns of HP1, a heterochromatin-associated nonhistone chromosomal protein of *Drosophila*. Eur J Cell Biol *50*, 170-80 (1989)

Jaquet, Y., Delattre, M., Spierer, A., Spierer, P.: Functional dissection of the *Drosophila* modifier of variegation *Su(var)3-7*. Development *129*, 3975-82 (2002)

Jaquet, Y., Delattre, M., Montoya-Burgos, J., Spierer, A., Spierer, P.: Conserved domains control heterochromatin localization and silencing properties of SU(VAR)3-7. Chromosoma *115*, 139-5 (2006)

Jin, Y., Wang, Y., Walker, D.L., Dong, H., Conley, C., Johansen, J., Johansen, K.M.: JIL-1: a novel chromosomal tandem kinase implicated in transcriptional regulation in *Drosophila*. Molec. Cell *4*, 29-35 (1999)

Jones, D.O., Cowell, I.G., Singh, P.B.: Mammalian chromodomain proteins: their role in genome organisation and expression. Bioessays *22*, 124-37 (2000)

Jones, R.S., Gelbart, W.M.: The *Drosophila* Polycomb-group gene *Enhancer of zeste* contains a region with sequence similarity to *trithorax*. Mol Cell Biol *13*, 6357-66 (1993)

Karachentsev, D., Sarma, K., Reinberg, D., Steward, R.: PR-Set7-dependent methylation of histone H4 Lys 20 functions in repression of gene expression and is essential for mitosis. Genes Dev. *19*, 431-5 (2005)

Kellum, R., Alberts, B.M.: Heterochromatin protein 1 is required for correct chromosome segregation in *Drosophila* embryos. J Cell Sci. *108*, 1419-31 (1995)

Kim, T.H., Barrera, L.O., Zheng, M., Qu, C., Singer, M.A., Richmond, T.A., Wu, Y., Green, R.D., Ren, B.: A high-resolution map of active promoters in the human genome. Nature *436*, 876-80 (2005)

Klug, A., Rhodes, D., Smith, J., Finch, J.T., Thomas, J.O.: A low resolution structure for the histone core of the nucleosome. *Nature* 287, 509-516 (1980)

Kornberg, R.D.: Structure of chromatin. Annu Rev Biochem. *46*, 931-54 (1977)

Koryakov, D.E., Reuter, G., Dimitri, P., Zhimulev, I.F.: The SuUR gene influences the distribution of heterochromatic proteins HP1 and SU(VAR)3-9 on nurse cell polytene chromosomes of *Drosophila melanogaster*. Chromosoma *115*, 296-310 (2006)

Kouskouti A, Scheer E, Staub A, Tora L, Talianidis I.: Gene-specific modulation of TAF10 function by SET9-mediated methylation. Mol Cell *14*, 175-82 (2004)

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

Krauss, V., Fassl, A., Fiebig, P., Patties, I., Sass, H.: The evolution of the histone methyltransferase gene *Su(var)3-9* in metazoans includes a fusion with and a re-fission from a functionally unrelated gene. BMC Evol Biol *2*, 6:18 (2006)

Krejci, E., Garzino, V., Mary, C., Bennani, N., Pradel, J.: *Modulo*, a new maternally expressed *Drosophila* gene encodes a DNA-binding protein with distinct acidic and basic regions. Nucleic Acids Res. *17*, 8101-15 (1989)

Kusch, T., Florens, L., Macdonald, W.H., Swanson, S.K., Glaser, R.L., Yates, J.R. 3rd, Abmayr, S.M., Washburn, M.P., Workman, J.L.: Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions. Science *306*, 2084-7 (2004)

Kwon, T., Chang, J.H., Kwak, E., Lee ,C.W., Joachimiak, A., Kim, Y.C., Lee, J., Cho, Y.: Mechanism of histone lysine methyl transfer revealed by the structure of SET7/9-AdoMet. EMBO J *22*, 292-303 (2003)

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., Jenuwein, T.: Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature *410*, 116-20 (2001)

Lachner, M., Jenuwein, T.: The many faces of histone lysine methylation. Curr Opin Cell Biol *14*, 286-98 (2002)

Lee, N., Zhang, J., Klose, R.J., Erdjument-Bromage, H., Tempst, P., Jones, R.S., Zhang, Y.: The trithorax-group protein Lid is a histone H3 trimethyl-Lys4 demethylase. Nat Struct Mol Biol. *14*, 341-3 (2007)

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

Lloret-Llinares, M., Carré, C., Vaquero, A., de Olano, N., Azorín, F.: Characterization of *Drosophila melanogaster* JmjC+N histone demethylases. Nucleic Acids Res *36*, 2852-6 (2008)

Lomberk G, Bensi D, Fernandez-Zapico ME, Urrutia R.: Evidence for the existence of an HP1-mediated subcode within the histone code. Nat Cell Biol *8*, 407-15 (2006)

Ludlam, W.H., Taylor, M.H., Tanner, K.G., Denu, J.M., Goodman, R.H., Smolik, S.M.: The acetyltransferase activity of CBP is required for wingless activation and H4 acetylation in *Drosophila melanogaster*. Mol Cell Biol. *22*, 3832-41 (2002)

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., Richmond, T.J.: Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251-60 (1997)

Maixner, A., Hecker, T.P., Phan, Q.N., Wassarman, D.A.: A screen for mutations that prevent lethality caused by expression of activated Sevenless and Ras1 in the *Drosophila* embryo. Dev. Genet. *23*, 347-61 (1998)

Martin-Blanco, E., Gampel, A., Ring, J., Virdee, K., Kirov, N., Tolkovsky, A.M., Martinez-Arias, A.: *puckered* encodes a phosphatase that mediates a feedback loop regulating JNK activity during dorsal closure in *Drosophila.*
Genes Dev. *12*, 557-70 (1998)

Mellor, J.: It takes a PHD to read the histone code. Cell *126*, 22-4 (2006)

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., Lee, W., Mendenhall, E., O'Donovan, A., Presser, A., Russ, C., Xie, X., Meissner, A., Wernig, M., Jaenisch, R., Nusbaum, C., Lander, E.S., Bernstein, B.E.: Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553-60 (2007)

Min, J., Zhang, X., Cheng, X., Grewal, S.I., Xu, R.M.: Structure of the SET domain histone lysine methyltransferase Clr4. Nat Struct Biol. *9*, 828-32 (2002)

Min, J., Zhang, Y., Xu, R.M.: Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. Genes Dev. *17*, 1823-8 (2003)

Miotto, B., Sagnier, T., Berenger, H., Bohmann, D., Pradel, J., Graba, Y.: Chameau HAT and DRpd3 HDAC function as antagonistic cofactors of JNK/AP-1-dependent transcription during *Drosophila* metamorphosis.
Genes Dev. *20*, 101-12 (2006)

Mis, J., Ner, S.S., Grigliatti, T.A.: Identification of three histone methyltransferases in *Drosophila*: dG9a is a suppressor of PEV and is required for gene silencing. Mol Genet Genomics *275*, 513-26 (2006)

Mohr, S.E., Boswell, R.E.: Zimp encodes a homologue of mouse Miz1 and PIAS3 and is an essential gene in Drosophila melanogaster. Gene *229*, 109-16 (1999)

Moss, T.J., Wallrath, L.L.: Connections between epigenetic gene silencing and human disease. Mutat Res. *618*, 163-74 (2007)

Mottus, R., Sobel, R.E., Grigliatti, T.A.: Mutational analysis of a histone deacetylase in *Drosophila melanogaster*: missense mutations suppress gene silencing associated with position effect variegation. Genetics *154*, 657-68 (2000)

Mujtaba, S., Zeng, L., Zhou, M.M..: Structure and acetyl-lysine recognition of the bromodomain. Oncogene *26*, 5521-7 (2007)

Muller, P., Kuttenkeuler, D., Gesellchen, V., Zeidler, M.P., Boutros, M: Identification of JAK/STAT signalling components by genome-wide RNA interference. Nature *436,* 871-75 (2005)

Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D., Grewal, S.I.: Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science *292*, 110-3 (2001)

Nelson, W.G., Yegnasubramanian, S., Agoston, A.T., Bastian, P.J., Lee, B.H., Nakayama, M., De Marzo, A.M.: Abnormal DNA methylation, epigenetics, and prostate cancer. Front Biosci.*12*, 4254-66 (2007)

Ner, S.S., Harrington, M.J., Grigliatti, T.A.: A role for the *Drosophila* SU(VAR)3-9 protein in chromatin organization at the histone gene cluster and in suppression of position-effect variegation. Genetics *162*, 1763-74 (2002)

Nielsen, S.J., Schneider, R., Bauer, U.M., Bannister, A.J., Morrison, A., O'Carroll, D., Firestein, R., Cleary, M., Jenuwein, T., Herrera, R.E., Kouzarides, T.: Rb targets histone H3 methylation and HP1 to promoters. Nature *412*, 561-5 (2001)

Nishioka, K., Rice, J.C., Sarma, K., Erdjument-Bromage, H., Werner, J., Wang, Y., Chuikov, S., Valenzuela, P., Tempst, P., Steward, R., Lis, J.T., Allis, C.D., Reinberg, D.: PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin. Mol Cell *9*, 1201-13 (2002)

Oligny, L.L.: Cancer and epigenesis: a developmental perspective. Adv Pediatr. *50*, 59-80 (2003)

Pankotai, T., Komonyi, O., Bodai, L., Ujfaludi, Z., Muratoglu, S., Ciurciu, A., Tora, L., Szabad, J., Boros, I.: The homologous *Drosophila* transcriptional adaptors ADA2a and ADA2b are both required for normal development but have different functions. Mol Cell Biol. *25*, 8215-27 (2005)

Paro, R., Hogness, D.S.: The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila*. Proc Natl Acad Sci U S A *88*, 263-7 (1991)

Perrin, L., Benassayag, C., Morello, D., Pradel, J., Montagne, J.: Modulo is a target of Myc selectively required for growth of proliferative cells in *Drosophila*. Mech Dev. *120*, 645-55 (2003)

Peters, A.H., O'Carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schofer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A., Opravil, S., Doyle, M., Sibilia, M., Jenuwein, T.: Loss of the Suv39h histone methyltransferases

impairs mammalian heterochromatin and genome stability. Cell *107*, 323-37 (2001)

Peters, A.H., Mermoud, J.E., O'Carroll, D., Pagani, M., Schweizer, D., Brockdorff, N., Jenuwein, T.: Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin. Nat Genet. *30*, 77-80 (2002)

Pokholok, D.K., Harbison, C.T., Levine, S., Cole, M., Hannett, N.M., Lee, T.I., Bell, G.W., Walker, K., Rolfe, P.A., Herbolsheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D.K., Young, R.A.: Genome-wide map of nucleosome acetylation and methylation in yeast. Cell *122*, 517-27 (2005)

Qian, C., Wang, X., Manzur, K., Sachchidanand, Farooq, A., Zeng, L., Wang, R., Zhou, M.M.: Structural insights of the specificity and catalysis of a viral histone H3 lysine 27 methyltransferase. J Mol Biol *359*, 86-96 (2006)

Qian, C., Zhou, M.M.: SET domain protein lysine methyltransferases: Structure, specificity and catalysis. Cell Mol Life Sci *63*, 2755-63 (2006)

Razin, S.V., Iarovaia, O.V., Sjakste, N., Sjakste, T., Bagdoniene, L., Rynditch, A.V., Eivazova, E.R., Lipinski, M., Vassetzky, Y.S.: Chromatin domains and regulation of transcription. J Mol Biol. *369*, 597-607 (2007)

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., Jenuwein, T.: Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593-9 (2000)

Reuter G, Giarre M, Farah J, Gausz J, Spierer A, Spierer P.: Dependence of position-effect variegation in Drosophila on dose of a gene encoding an unusual zinc-finger protein. Nature *344*, 219-23 (1990)

Reuter, G., Spierer, P.: Position effect variegation and chromatin proteins. Bioessays *14*, 605-12 (1992)

Reuter, G., Wolff, I.: Isolation of dominant suppressor mutations for position-effect variegation in *Drosophila melanogaster*. Mol Gen Genet. *182*, 516-9 (1981)

Roh, T.Y., Cuddapah, S., Zhao, K.: Active chromatin domains are defined by acetylation islands revealed by genome-wide mapping. Genes Dev *19*, 542-52 (2005)

Rudolph, T., Yonezawa, M., Lein, S., Heidrich, K., Kubicek, S., Schafer, C., Phalke, S., Walther, M., Schmidt, A., Jenuwein, T., Reuter, G.: Heterochromatin formation in *Drosophila* is initiated through active removal of H3K4 methylation by the LSD1 homolog SU(VAR)3-3. Mol Cell. *26*, 103-15 (2007)

Saigo, K., Millstein, L., Thomas, C.A. Jr. The organization of *Drosophila*

melanogaster histone genes. Cold Spring Harb Symp Quant Biol. *45*, 815-27 (1981)

Samal, B., Worcel, A., Louis, C., Schedl, P.: Chromatin structure of the histone genes of *D. melanogaster*. Cell *23*, 401-9 (1981)

Sampath, S.C., Marazzi, I., Yap, K.L., Sampath, S.C., Krutchinsky, A.N., Mecklenbräuker, I., Viale, A., Rudensky, E., Zhou, M.M., Chait, B.T., Tarakhovsky, A.: Methylation of a histone mimic within the histone methyltransferase G9a regulates protein complex assembly. Mol Cell *27*, 596-608 (2007)

Schones, D.E., Zhao, K.: Genome-wide approaches to studying chromatin modifications. Nat Rev Genet. *9,* 179-91 (2008)

Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R., Reuter, G.: Central role of *Drosophila* SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. EMBO J. *21*, 1121-31 (2002)

Schotta, G., Ebert, A., Reuter, G.: SU(VAR)3-9 is a conserved key function in heterochromatic gene silencing. Genetica *117*, 149-58 (2003)

Schotta, G., Ebert, A., Dorn, R., Reuter, G.: Position-effect variegation and the genetic dissection of chromatin regulation in *Drosophila*. Semin Cell Dev Biol. *14*, 67-75 (2003)

Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D., Jenuwein, T.: A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes Dev. *18*, 1251-62 (2004)

Schübeler, D., MacAlpine, D.M., Scalzo, D., Wirbelauer, C., Kooperberg, C., van Leeuwen, F., Gottschling, D.E., O'Neill, L.P., Turner, B.M., Delrow, J., Bell, S.P., Groudine M.: The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. Genes Dev.*18*, 1263-71 (2004)

Secombe, J., Li, L., Carlos, L., Eisenman, R.N.: The Trithorax group protein Lid is a trimethyl histone H3K4 demethylase required for dMyc-induced cell growth. Genes Dev. *21*, 537-51 (2007)

Sedkov, Y., Cho, E., Petruk, S., Cherbas, L., Smith, S.T., Jones, R.S., Cherbas, P., Canaani, E., Jaynes, J.B., Mazo, A.: Methylation at lysine 4 of histone H3 in ecdysone-dependent development of *Drosophila*. Nature *426*, 78-83 (2003)

Seum, C., Reo, E., Peng, H., Rauscher, F.J. 3rd, Spierer, P., Bontron, S.: *Drosophila* SETDB1 is required for chromosome 4 silencing. PLoS Genet. *3*, e76 (2007)

Seum, C., Bontron, S., Reo, E., Delattre, M., Spierer, P.: *Drosophila* G9a is a nonessential gene. Genetics *177*, 1955-7 (2007)

Shanower, G.A., Muller, M., Blanton, J.L., Honti, V., Gyurkovics, H., Schedl, P.: Characterization of the grappa gene, the *Drosophila* histone H3 lysine 79 methyltransferase. Genetics *169*, 173-84 (2005)

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A., Shi, Y.: Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell *119*, 941-53 (2004)

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

Sinclair, D.A., Clegg, N.J., Antonchuk, J., Milne, T.A., Stankunas, K., Ruse, C., Grigliatti, T.A., Kassis, J.A., Brock, H.W.: *Enhancer of Polycomb* is a suppressor of position-effect variegation in *Drosophila melanogaster*. Genetics *148*, 211-20 (1998)

Smith, S.T., Petruk, S., Sedkov, Y., Cho, E., Tillib, S., Canaani, E., Mazo, A.: Modulation of heat shock gene expression by the TAC1 chromatin-modifying complex. Nat Cell Biol. *6*, 162-7 (2004)

Spofford, J.B.: Position effect variegation in *Drosophila*. In: The Genetics and Biology of Drosophila *Vol. 1c*, 955–1018. (1976) Edited by M. Ashburner & E. Novitski. Academic Press, New York.

Stabell, M., Eskeland, R., Bjorkmo, M., Larsson, J., Aalen, R.B., Imhof, A., Lambertsson, A.: The *Drosophila* G9a gene encodes a multi-catalytic histone methyltransferase required for normal development. Nucleic Acids Res. *34*, 4609-21 (2006)

Stankunas, K., Berger, J., Ruse, C., Sinclair, D.A., Randazzo, F., Brock, H.W.: The *enhancer of polycomb* gene of *Drosophila* encodes a chromatin protein conserved in yeast and mammals. Development *125*, 4055-66 (1998)

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

Swaminathan, J., Baxter, E.M., Corces, V.G.: The role of histone H2Av variant replacement and histone H4 acetylation in the establishment of *Drosophila* heterochromatin. Genes Dev. *19*, 65-76 (2005)

Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M.,

Takeda, N., Niida, H., Kato, H., Shinkai, Y.: G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes Dev. *16*, 1779-91 (2002)

Talbert, P.B., Henikoff, S.: Spreading of silent chromatin: inaction at a distance. Nat Rev Genet. *10,* 793-803 (2006)

Tie F, Prasad-Sinha J, Birve A, Rasmuson-Lestander A, Harte PJ.: A 1-megadalton ESC/E(Z) complex from *Drosophila* that contains polycomblike and RPD3. Mol Cell Biol *23*, 3352-62 (2003)

Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G., Reuter, G.: The protein encoded by the *Drosophila* position-effect variegation suppressor gene *Su(var)3-9* combines domains of antagonistic regulators of homeotic gene complexes. EMBO J. *13*, 3822-31 (1994)

Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P., Zhang, Y.: Histone demethylation by a family of JmjC domain-containing proteins. Nature *439*, 811-6 (2006)

Turner, B.M., Birley, A.J., Lavender, J.: Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in *Drosophila* polytene nuclei. Cell *69*, 375-84 (1992)

Turner, B.M.: Histone acetylation and an epigenetic code. Bioessays *22*, 836-45 (2000)

Turner, B.M.: Reading signals on the nucleosome with a new nomenclature for modified histones. Nat Struct Mol Biol. *12*, 110-2 (2005)

Tzeng, T.Y., Lee, C.H., Chan, L.W., Shen, C.K.: Epigenetic regulation of the *Drosophila* chromosome 4 by the histone H3K9 methyltransferase dSETDB1. Proc Natl Acad Sci U S A. *104*, 12691-6 (2007)

Vandel, L., Nicolas, E., Vaute, O., Ferreira, R., Ait-Si-Ali, S., Trouche, D.: Transcriptional repression by the retinoblastoma protein through the recruitment of a histone methyltransferase. Mol Cell Biol. *21*, 6484-94 (2001)

Vaute, O., Nicolas, E., Vandel, L., Trouche, D.: Functional and physical interaction between the histone methyl transferase Suv39H1 and histone deacetylases. Nucleic Acids Res *30*, 475-81 (2002)

Wallrath, L.L., Elgin, S.C.: Position effect variegation in *Drosophila* is associated with an altered chromatin structure. Genes Dev. *9*, 1263-77 (1995)

Wang, Y., Zhang, W., Jin, Y., Johansen, J., Johansen, K.M.: The JIL-1 tandem

kinase mediates histone H3 phosphorylation and is required for maintenance of chromatin structure in *Drosophila*. Cell *105*, 433-43 (2001)

Wang, Y., Wysocka, J., Sayegh, J., Lee, Y.H., Perlin, J.R., Leonelli, L., Sonbuchner, L.S., McDonald, C.H., Cook, R.G., Dou, Y., Roeder, R.G., Clarke, S., Stallcup, M.R., Allis, C.D., Coonrod, S.A.: Human PAD4 regulates histone arginine methylation levels via demethylimination. Science *306*, 279–83 (2004)

Weiler, K.S., Wakimoto, B.T.: Heterochromatin and gene expression in *Drosophila*. Annu Rev Genet *29*, 577-605 (1995)

Xiao, B., Jing, C., Kelly, G., Walker, P.A., Muskett, F.W., Frenkiel, T.A., Martin, S.R., Sarma, K., Reinberg, D., Gamblin, S.J., Wilson, J.R.: Specificity and mechanism of the histone methyltransferase Pr-Set7. Genes Dev *19*, 1444-54 (2005)

Yang, X.J.: Lysine acetylation and the bromodomain: a new partnership for signaling. Bioessays *26*, 1076-87 (2004)

Zhang, K., Mosch, K., Fischle, W., Grewal, S.I.: Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. Nat Struct Mol Biol *15*, 381-8 (2008)

Zhang, W., Deng, H., Bao, X., Lerach, S., Girton, J., Johansen, J., Johansen, K.M.: The JIL-1 histone H3S10 kinase regulates dimethyl H3K9 modifications and heterochromatic spreading in *Drosophila*. Development *133*, 229-35 (2006)

Zhimulev, I.F., Belyaeva, E.S.: Intercalary heterochromatin and genetic silencing. Bioessays *25*, 1040-51 (2003)

2. ALTERED HISTONE H3 METHYLTRANSFERASE ACTIVITY OF Su(var)3-9 IMPAIRS GENE SILENCING[1].

2.1. Introduction

The chromatin of eukaryotic organisms is a dynamic complex of DNA and proteins. The nucleosome, its basic structural unit, is comprised of 146 bp of DNA wrapped around an octamer of the core histones H3, H4, H2A and H2B (Finch *et al.* 1977; Klug *et al.*, 1980, Luger *et al.*, 1997). Higher order assemblies of nucleosomes package the genome into structured regions that typically have different functional properties. For example, regions of the genome surrounding the centromere have nucleosomes organized into highly condensed and tightly packaged chromatin, are transcriptionally inert, late replicating and are collectively termed heterochromatin (Gatti and Pimpinelli, 1992; Lohe and Hilliker, 1995). Regions that are less densely packaged, accessible to the transcriptional machinery, and which replicate earlier in S-phase, are collectively known as euchromatin (often also referred to as "open" chromatin). However, even euchromatin is a mosaic of silenced (repressed) and transcriptionally competent (open) domains. Numerous non-histone proteins and a range of covalent histone modifications influence nucleosome-nucleosome interactions and higher order packaging and thus control the transitions between transcriptionally competent (open) and transcriptionally repressed states of the euchromatic portion of the genome. The histone modifications, particularly at the N-terminal tails, include acetylation,

---

[1] A version of this chapter will be submitted for publication. Kalas, P. and Grigliatti, T.A. Altered histone H3 methyltransferase activity of SU(VAR)3-9 impairs gene silencing.

Specific patterns of histone modifications correlate with gene activity and chromatin structure - the "histone code" hypothesis (Jenuwein and Allis, 2001). For example, acetylation of lysine 9 of histone H3 (H3K9ac) is typically associated with gene expression, while methylation of H3K9 (H3K9me) and H4K20me are usually associated with gene repression or silenced regions of the genome (Litt et al., 2001; Schotta *et al.,* 2004; reviewed in Rice and Allis, 2001; Berger, 2002; Ebert *et al.,* 2006).

In *Drosophila*, methylation of H3K9 is catalyzed by SU(VAR)3-9, the archetypal SET domain-containing histone methyltransferase (HMTase) that is highly conserved in eukaryotes (Aagaard *et al.,* 1999; Rea *et al.,* 2000; Ivanova *et al.,* 1998). Mutations in *Su(var)3-9* were recovered in genetic screens (Reuter and Wolff, 1981; Sinclair *et al.,* 1983; Donaldson *et al.*, 2002) as strong dominant suppressors of the heterochromatin-associated gene silencing phenomenon, position-effect variegation (PEV) (Muller, 1930; Grigliatti, 1991; Lewis, 1950; Weiler and Wakimoto, 1995). All *Su(var)3-9* mutants known to date are strong, dominant suppressors of w^{m4}, Sb^V and bw^V variegation (Sinclair *et al.,* 1983; Reuter and Wolff, 1981). SU(VAR)3-9 has now been extensively characterized. It associates with heterochromatic regions (Schotta *et al.,* 2002; 2003) and numerous euchromatic loci (Greil *et al.,* 2003), including the tandemly reiterated *histone gene cluster* (HIS-C) (Ner *et al.,* 2002), where it epigenetically modulates the expression of the *histone* genes by altering the chromatin structure of the locus via methylation of H3K9 and recruitment of HP1 (Ner *et al.,* 2002; Ner *et al.*, in preparation).

The original EMS mutagenesis screen isolated over 2 dozen mutants that mapped near the *Su(var)3-9* locus (Sinclair *et al.* 1983). This chapter focuses on a subset of these mutants. We demonstrate that these mutations are indeed *Su(var)3-9*

alleles, and examine their phenotypes at four different levels: molecular, cellular, biochemical and morphological (eye colour/PEV). At the molecular level, we found that all alleles have missense mutations resulting in single amino acid substitutions in the catalytic region of the protein, which consists of the evolutionarily conserved preSET, SET and postSET domains. In contrast, no mutations were recovered in the other highly conserved region of SU(VAR)3-9, the chromodomain, or in the *Drosophila*-specific N-terminus, suggesting alterations in these regions don't influence the epigenetic silencing resulting from PEV. At the cellular level, the *Su(var)3-9* mutants had reduced levels of H3K9me2 and HP1 at both the chromocentre and a euchromatic locus, the HIS-C. Biochemically, we showed that all amino acid substitutions either abolish or dramatically reduce the HMTase activity of SU(VAR)3-9 *in vitro*. Substitutions of conserved cysteine residues in the preSET domain directly correlated with a complete loss-of-HMTase function *in vitro*, while changes in the SET or postSET regions resulted in partial loss-of-function. Finally, at the level of morphological phenotypes, eye colour/PEV, all *Su(var)3-9* mutants were able to suppress w^{m4} variegation in a strain that is hyperploid for *Su(var)3-9*, but we observed differences in their strength of suppression.

2.2. Results

2.2.1. <u>Recombination mapping and DNA Sequence analysis</u>

In the early 1980s our lab recovered about 50 third chromosome EMS-induced mutants that were strong, dominant suppressors of PEV (Sinclair *et al*., 1983). Over twenty of these mutants are homozygous viable, clustered on the right arm of the third chromosome and circumstantial evidence strongly indicated that a large subset of these

were *Su(var)3-9* alleles (Harrington, 2001). The mutations are homozygous viable, which obviated the use of complementation analyses to define allelism. Thus, to determine which of the mutations were *Su(var)3-9* alleles, we performed recombination analysis using a recessive lethal P-element insert, *Su(var)3-9^{P25}* (Harrington, 2001). For each of the 20 putative *Su(var)3-9* alleles, we generated heterozygous *Su(var)3-9^{P25}*/*Su(var)3-9^{putative}* females and crossed them to wild-type males, and examined the progeny. For 13 mutants we observed no wild-type recombinants (after screening > 2000 flies for each cross), indicating that the mutations they carry are very closely linked to *Su(var)3-9^{P25}* (<0.1 cM; Table 1). These 13 mutants were then characterized by DNA sequence analysis to determine the precise nature and position of the mutation. Nine out of the 12 were missense mutants in *Su(var)3-9* (Table 1). Accordingly, we renamed these mutants as *Su(var)3-9^{allele number}* (Table 1). Three mutants, *Su(var)3-9^{317}*, *Su(var)3-9^{327}*, and *Su(var)3-9^{329}*, did not contain mutations in the coding region of *Su(var)3-9*. Since the estimated distance between the mutations in each of these three mutants and *Su(var)3-9^{P25}* is <0.1 cM, which represents about 50kb or less in the region surrounding the *Su(var)3-9* locus (http://flybase.bio.indiana.edu), these mutants may still be alleles of *Su(var)3-9* and represent alterations in the regulatory region of *Su(var)3-9* or cause defects in mRNA processing. Alternatively, their Su(var) phenotype may be the result of mutations in one or more genes that are closely linked to *Su(var)3-9*, such as *Set* or *Oscp*. Indeed, Su(var) mutations are rather commonly found in clusters, closely mapping, but not contiguous (Sinclair *et al*., 1983).

The sequence analysis also revealed single base-pair changes leading to ALA304PHE and/or ILE375LEU in *Su(var)305*, *Su(var)3-9^{324}*, *Su(var)3-9^{329}* and *Su(var)331*

(Table 2.1). Since *Su(var)305*, which carries both substitutions, and *Su(var)331,* which carries the I375L substitution, are not allelic to *Su(var)3-9* (as determined by recombination mapping with *Su(var)3-9^{P25}*, see Materials and Methods), we concluded that these substitutions are not the cause of the Su(var) phenotype in *Su(var)3-9^{324}* or *Su(var)3-9^{329}*. In *Su(var)3-9^{324}*, the C428Y is most likely the cause of the Su(var) effect, while for *Su(var)3-9^{329}* the mutation responsible for this phenotype may be located in one of the *cis*-regulatory elements of the gene, which are not known and were therefore not sequenced.

Table 2.1. Putative *Su(var)3-9* alleles that were mapped and sequenced in this

study. Allelism to *Su(var)3-9* was determined by recombination mapping with the P-

element induced *Su(var)3-9^{P25}* mutant (Ner *et al.,* 2002). Mutants that failed to yield

wild-type recombinants were further analyzed by sequencing and, if allelism to

Su(var)3-9 was confirmed, they were renamed accordingly.

MUTANT	*Su(var)3-9* allele (genetically)?	Allelism confirmed by sequencing?	Mutation	New name
Su(var)301	NO	N/A (not tested)	N/A	
Su(var)305	NO	No	A304F[b], I375L[b]	
Su(var)306	NO [a]	N/A	N/A	
Su(var)309	YES	Yes	C462Y	*Su(var)3-9^{309}*
Su(var)311	YES	Yes	G521D	*Su(var)3-9^{311}*
Su(var)312	YES	Yes	C462Y	*Su(var)3-9^{312}*
Su(var)314	NO [a]	N/A	N/A	
Su(var)315	NO	N/A	N/A	
Su(var)317	YES	No	N/A	
Su(var)318	YES	Yes	S616L	*Su(var)3-9^{318}*
Su(var)319	YES	Yes	S616L	*Su(var)3-9^{319}*
Su(var)320	NO	N/A	N/A	
Su(var)324	YES	Yes	C428Y, I375L[b]	*Su(var)3-9^{324}*
Su(var)325	YES	Yes	P582Q	*Su(var)3-9^{325}*
Su(var)327	YES	No	N/A	
Su(var)329	YES	No	A304F[b]	
Su(var)C76	YES	Yes	C421S	*Su(var)3-9^{376}*
Su(var)330	YES	Yes	D536N	*Su(var)3-9^{330}*
Su(var)331	NO	N/A	I375L[b]	

[a] Harrington, 2001.
[b] These amino acid substitutions likely represent naturally occurring polymorphisms, and are not responsible for the Su(var) phenotype.

2.2.2. <u>Amino acid substitutions in Su(var)3-9 target the preSET/SET/postSET domain.</u>

Su(VAR)3-9 has two distinct, and highly conserved domains: a chromodomain, and a SET domain, which includes the flanking pre- and a postSET regions (Figure 2.1). The chromodomain of Su(VAR)3-9 is about 40 amino acids long and its function is not entirely clear, but it is speculated to have a role in protein-protein interactions that may control H3K9 methylation (Schotta *et al.,* 2003). The preSET/SET/postSET region is ~250 amino acids long, represents less than 40% of the entire length of the protein, and constitutes the catalytic region of the Su(VAR)3-9 HMTase activity (Tschiersch *et al.,* 1994; Rea *et al.,* 2000; Schotta *et al.,* 2002; Ner *et al.,* 2002). Both domains are also found in other chromatin proteins, where they play critical roles in silencing (Ivanova *et al.,* 1998; Nakayama *et al.,* 2001; Schotta *et al.,* 2002; Platero *et al.,* 1995: Messmer *et al.,* 1992; Ma *et al.,* 2001; Akhtar *et al.,* 2000; Jacobs *et al.*, 2001; Bannister *et al.*, 2001; Lachner *et al.,* 2001; Bouazoune *et al.,* 2002; Fischle *et al.,* 2003; Min *et al.,* 2003; Pray-Grant *et al.,* 2005). The *Drosophila* Su(VAR)3-9 has a longer N-terminus than its mammalian and yeast counterparts. This novel region of the protein appears to be required for interactions with chromatin proteins Su(VAR)3-7 and HP1, and for Su(VAR)3-9 dimerization (Schotta *et al.,* 2002; 2003; Eskeland *et al.,* 2004). Theoretically, mutations in the *Drosophila* specific N-terminus, the chromodomain, or the preSET/SET/postSET regions could lead to a loss of the protein's silencing function, and to a Su(var) phenotype. Mutations in the N-terminus may do so by preventing Su(VAR)3-9 dimerization and/or its interaction with other chromatin proteins, such as HP1 and Su(VAR)3-7. Mutations in the chromodomain could affect the targeting of Su(VAR)3-9 and/or its interaction with other proteins. Finally, alterations in the

63

preSET/SET/postSET may impair its catalytic activity, with or without affecting the protein's ability to participate in protein-protein interactions. In addition, mutations in any region could potentially affect the protein's folding, thus rendering it inactive.

All *Su(var)3-9* mutations were recovered in an EMS screen. EMS causes mainly base-pair substitutions, principally G:C to A:T transitions and thus, effectively causes random mutations. Accordingly, we reasoned that 1) if the *Su(var)3-9* mutations isolated in the original screen preferentially target certain regions of the protein, then these regions must be crucial for PEV-type gene silencing, and 2) if some of the EMS-induced lesions represent missense mutations, the affected residues are likely crucial for the silencing function of SU(VAR)3-9. Strikingly, we found that all 9 mutants are located solely in the preSET/SET/postSET domain. Not surprisingly, all are single base pair substitutions resulting in missense mutations. Since all the mutations cluster within the catalytic (preSET/SET/postSET) domain of SU(VAR)3-9, our results suggest the catalytic function of SU(VAR)3-9 is essential for the silencing associated with PEV, and perhaps, amino acid substitutions in the remainder of the protein do not influence the repression caused by PEV.

The nature of the amino acid substitutions in the various *Su(var)3-9* alleles was examined further. *Su(var)3-9*[309] and *Su(var)3-9*[312] arose independently, but have the same A-to-G transition, giving rise to a CYS to TYR substitution at position 462. CYS462 is part of the preSET domain, and is conserved in all homologues of SU(VAR)3-9, including SUV39H1, Suv39h1 and Clr4p. It corresponds to one of nine cysteine residues that coordinate three zinc ions (Min *et al.,* 2002). *Su(var)3-9*[324] (CYS428TYR) and *Su(var)3-9*[376] (CYS421SER) also cause substitutions in this highly conserved group

of cysteines (Fig. 2.1). Three alleles, *Su(var)3-9³²⁵*, *Su(var)3-9³³⁰* and *Su(var)3-9³¹¹*, cause amino acid substitutions in the SET domain. They correspond to PRO582GLU in *Su(var)3-9³²⁵*, GLY521ASP in *Su(var)3-9³¹¹*, and ASP536ASN in *Su(var)3-9³³⁰*. Both PRO582 and GLY521 are conserved in the mammalian homologues of SU(VAR)3-9, but not in the yeast homologue, Clr4p. ASP536 is conserved in all H3K9-specific HMTases and plays a key role in the interaction with the N-terminal tail of H3 (Zhang *et al.,* 2003). *Su(var)3-9³¹⁸* and *Su(var)3-9³¹⁹* represent a pair of independently induced mutations and each causes a SER616LEU substitution. This residue is part of the post-SET, but it is not conserved in SUV39H1, Suv39h1 or Clr4p. A similar screen performed by Reuter and colleagues (Reuter and Wolff, 1981; Ebert *et al.*, 2004) also recovered multiple alleles of *Su(var)3-9*, but none of them correspond to the nine mutations we recovered (see Discussion).

In summary, the screen yielded nine different *Su(var)3-9* mutants of which 4 carry mutations in the preSET domain, 3 in the SET domain, and 2 in the postSET domain. In two cases (*Su(var)3-9³⁰⁹* and *Su(var)3-9³¹²*, and *Su(var)3-9³¹⁸* and *Su(var)3-9³¹⁹*) the same residue was mutated independently, so the 9 different mutants represent 7 different, new alleles/mutations.

a.

SU(VAR)3-9

interactions with HP1,
SU(VAR)3-7, SU(VAR)3-9

catalytic region (HMTase)

318,319 | 325 | 311 330 | 312,309 | 324 | 376

| eIF2γ | N-terminal (fly specific) | chromo | preSET | SET | post SET |

b.

```
SU(VAR)3-9 361 FEKRMNHVEKPSPPIRVENNIDLDTIDS-NFMYIHDNIIGKDVPKPEAGI-VGKCTEDTEEC-TAST-KC-CARFAG--ELFAYERSTRRLRLRPGS 451
SUV39H1    145 -------------ITVENEVDLDGFFR-AFVYINEYRVGEGITLNQVA--VGEC----QDCLWAPTGGC-CPGASL--HKFAYNDQG-QVRLRAGL 214
CLR4       220 --------------VTLVNEVDDEPCPSLDFQFISQYRLTQGVIPPDPNFQSGCNCSSLGG-CDLNNPSRCECLDDLDEPTHFAYDAQG-RVRADTGA 302
```

SU(VAR)3-9376 — S
SU(VAR)3-9324 — Y

```
SU(VAR)3-9 452 AIYECNSRCSCDSSCSNRLVQHGRQVPLVLFKTANGSGWGVRAATALRKGEFVCEYIGEIITSDEANERGKAYDDNGRTYLFDLYNTAQDSEYTIDA 549
SUV39H1    215 PIYECNSRCRCGYDCPNRVVQKGIRYDLCIFRTDDGRGWGVRTLEKIRKNSFVMEYVGEIITSEEAERRGQTYDRQGATYLFDLY---VEDVYTVDA 312
CLR4       303 VIYECNSFCSCSMECPNRVVQRGRTLPLEIFKTKE-KGWGVRSLRFAPAGTFITCYLGEVITSAEAAKRDKNYDDDGITYLFDLDMFD-DASEYTVDA 398
```

SU(VAR)3-9309/312 — Y
SU(VAR)3-9311 — D
SU(VAR)3-9330 — N

```
SU(VAR)3-9 550 ANYGNISHFINHSCDPNLAVFPCWIEHLNWALPHLVFFTLRPIKAGEELSFDY-IRADNEDVPYENLST----------AVRVECRCGADNCRKVLF- 635
SUV39H1    313 AYYGNISHFVNHSCDPNLQVYNVFIDNLDERLPRIAFFATRTIRAGEELTFDYNMQVDPVDMESTRMDSNFGLAGLPGSPKKRVRIECKCGTESCRKYLF- 412
CLR4       399 QNYGDVSRFFNHSCSPNIAIYSAVRNHGFRTIYDLAFFAIKDIQPLEELTFDYAGAKDFSPVQSQKSQQN---------RISKLRRQCKCGSANCRGWLFG 490
```

SU(VAR)3-9325 — Q
SU(VAR)3-9318/319 — L

Figure 2.1.

Figure 2.1. *Su(var)3-9* **mutants are clustered in the preSET, SET and postSET domains. a)** A schematic of the domain structure of SU(VAR)3-9 (635 amino acids). The relative positions of the chromodomain (red box), the pre- and postSET domains (green boxes), the SET domain (blue box), and the region that SU(VAR)3-9 shares with eIF2γ (eIF2γ) are shown. The region N-terminal to the chromodomain is longer in the Drosophila protein than the mammalian or yeast homologues and is involved in protein-protein interactions with HP1, SU(VAR)3-7 and SU(VAR)3-9 (Schotta *et al.,* 2003, Eskeland *et al.,* 2004). The vertical bars mark the positions of amino acids substitutions in 9 *Su(var)3-9* mutants identified by sequence analysis. **b)** Amino acid alignment of the catalytic domain of SU(VAR)3-9, SUV39H1 (human) and CLR4P (fission yeast) showing major conserved residues (red), conserved preSET cysteines that coordinate 3 zinc ions in CLR4P (yellow) (Min *et al.* 2002) and the conserved aspartate that directly interacts with the substrate (green) (Zhang *et al.,* 2003). The amino acid changes observed in each EMS-induced mutant are shown below the alignment. Allele pairs *Su(var)3-9^309^* and *Su(var)3-9^312^*, and *Su(var)3-9^318^* and *Su(var)3-9^319^*, are independently generated mutants resulting in the same amino acid substitutions.

2.2.3. <u>The missense alleles express (mutant) Su(var)3-9 gene products.</u>

     At the morphological level (suppression of PEV in the w^{m4} strain), the phenotype

of the EMS-induced *Su(var)3-9* missense mutants is virtually indistinguishable from that

of the "protein null", *Su(var)3-9^{06}* (Reuter and Wolff, 1981; Sinclair *et al.,* 1983;

Tschiersch *et al.,* 1994; Schotta *et al.,* 2002). *Su(var)3-9^{06}* is an X ray-induced, null

allele that results in no RNA or protein products (Figures 2.2a and 2.2b, and Tschiersch

et al., 1994; Schotta *et al.,* 2002). Western blots were performed on each homozygous

mutant strain to ensure that the EMS-induced *Su(var)3-9* alleles are not effectively

"protein nulls", as a consequence of, for example, an instability of their mutant gene

products. All strains tested (*Su(var)3-9^{309}*, *Su(var)3-9^{330}*, *Su(var)3-9^{318}*, *Su(var)3-9^{319}*

and *Su(var)3-9^{311}*) showed approximately the same level of SU(VAR)3-9 as the wild-type,

confirming that the protein is still present in the mutant strains, in spite of single base-

pair substitutions. A representative blot is shown in Figure 2.2b.

Figure 2.2

a.

		RP49				Su(var)3-9		
cDNA amplified								
strain	wild-type		Su(var)3-9⁰⁶		wild-type		Su(var)3-9⁰⁶	
RTase	-	+	-	+	-	+	-	+

1　2　3　4　5　6　7　8

b.

α–SU(VAR)3-9

1　2　3　4　5
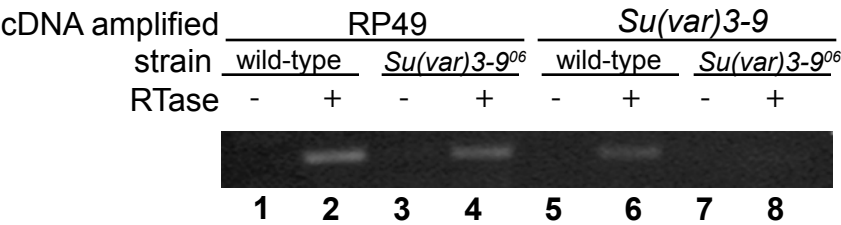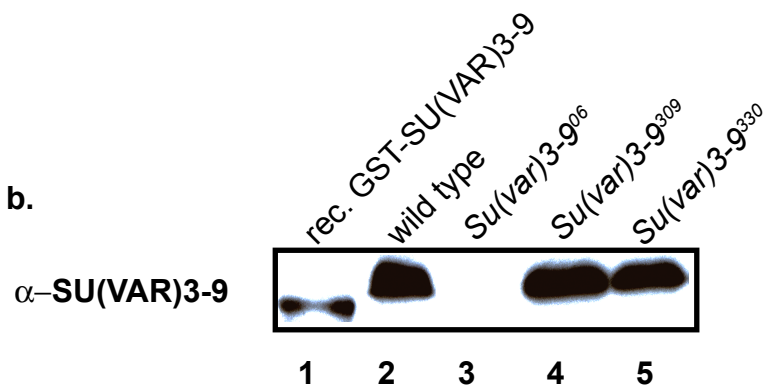
Figure 2.2. Detection of *Su(var)3-9* gene products in the EMS-induced *Su(var)3-9* missense mutants, and in the null allele *Su(var)3-9^06*.

a) RT-PCR reactions using total RNA extracted from wild-type (lanes 1, 2, 5 and 6) or *Su(var)3-9^06* female flies as starting material. RNA samples were reverse-transcribed with *Su(var)3-9*- and *RP49*-specific primers simultaneously, and the cDNA mixture obtained was PCR amplified separately using primers specific for *RP49* (lanes 1-4) or *Su(var)3-9* (lanes 5-8). The odd-numbered lanes represent PCR amplification of mock RT reactions (no reverse transcriptase). Identical results were obtained using total RNA from males. Note that primer 3906rt hybridizes to the first 17 nucleotides of *Su(var)3-9*'s third exon. **b)** Western blot analysis of wild-type and mutant embryo extracts. Lane 1: GST-SU(VAR)3-9[residues 310-635] recombinant protein (1μg), lane 2: wild-type embryo extract (150μg), lanes 3-5: extracts from homozygous *Su(var)3-9* mutant embryos (150μg).

2.2.4. _Su(var)3-9_ mutants display a reduction in the levels of H3K9me and HP1 associated with the chromocentre and with the HIS-C.

Since di- and trimethylated H3K9 (H3K9me2/H3K9me3) are the predominant products of SU(VAR)3-9 catalytic activity (Eskeland _et al.,_ 2004), we asked whether our EMS-induced _Su(var)3-9_ mutations had an effect on the level of H3K9me2 _in vivo_. First, a set of western blots was performed with an anti-H3K9me2 antibody, to compare the levels of H3K9me2 present in embryo extracts from the _Su(var)3-9_ homozygous mutant strains to the levels in wild-type and _Su(var)3-9[06]_ ("protein null") extracts. The western blots were also probed with anti-tubulin antibody to establish equal loading of protein extracts, and the total histone H3 present in each extract was detected on parallel blots to ensure that any difference we observed in H3K9me2 levels between the extracts were meaningful (Figure 3a). The _Su(var)3-9_ mutant extracts had significantly reduced levels of H3K9me2 relative to the wild-type extract (Figure 3a, compare lane 2 with lanes 3-6). However, all _Su(var)3-9_ missense mutant extracts showed a residual level of H3K9me2. We also detected residual H3K9me2 in the protein null strain, _Su(var)3-9[06]_. This was not unexpected since there are several other HMTases, including dG9a and DmSETDB1, that can methylate H3K9 (Ayyanathan _et al.,_ 2003; Mis _et al.,_ 2006; Seum _et al._, 2007).

H3K9me2 is generally considered a mark of heterochromatin, but is also detected at many euchromatic regions (Nielsen _et al.,_ 2001; Schotta _et al._, 2002; Ner _et al._, 2002; Greil _et al.,_ 2003; Ebert _et al._, 2006). Thus, we next asked whether the reduction in the level of methylated H3K9 observed in the _Su(var)3-9_ mutants affects heterochromatin exclusively, or whether it can also be observed in some euchromatic

regions. For this purpose, we examined the distribution of H3K9me2 at a the

chromocentre (heterochromatin) and at the *histone genes cluster* (HIS-C), a

euchromatic target of SU(VAR)3-9 (Ner *et al.,* 2002), in wild-type and *Su(var)3-9*

mutants. As expected, the chromocenter of wild-type nuclei was richly stained for

H3K9me2 (Figure 2.3c, panels A-C). In contrast, the amount of H3K9me2 detected at

the chromocentre was substantially reduced in the *Su(var)3-9* mutants (Figure 2.3c,

compare panel A with panels D, G, J and M). This reduction was particularly dramatic in

Su(var)3-9^{06}, confirming previous observations (Schotta *et al.,* 2002; Ebert *et al.,* 2006).

Note however that H3K9me2, despite its low level, was still detected at the

chromocenter. The presence of H3K9me2 at the HIS-C was examined by chromatin

immunoprecipitation (ChIP) of cross-linked extracts with α-H3K9me2. Extracts from 12-

16 hours old embryos of wild-type, *Su(var)3-9^{06}* and the various *Su(var)3-9* missense

mutant strains were tested, and HIS-C DNA sequences were detected in all extracts

(Figure 2.3b, lane 3). However, the proportion of HIS-C DNA pulled down by the

antibody was lower in the *Su(var)3-9* mutants than it was in the wild-type strain (Figure

2.3b, compare lanes 1 and 3 in the wild-type extracts (WT) with lanes 1 and 3 in the

Su(var)3-9 mutants). We conclude that the missense *Su(var)3-9* mutants, as well as the

null mutant *Su(var)3-9^{06}*, are associated with reduced levels of H3K9me2 at both the

highly compacted pericentric heterochromatin and at the largely silenced HIS-C, two

natural targets of SU(VAR)3-9.

Several pieces of evidence suggest that H3K9me2 serves as a substrate for HP1

in the formation of centric heterochromatin (Bannister *et al.* 2001; Lachner *et al.* 2001;

Nakayama *et al.* 2001; Ebert *et al.,* 2006), and HP1 has been shown to colocalize with

H3K9me2 in heterochromatin (Cryderman *et al.,* 2005; Ebert *et al.,* 2006). Co-localization of H3K9me2 and HP1 within euchromatic regions is much less frequent, but has been observed for a few loci, such as *cdc2*, the HIS-C and a few others (Cryderman *et al.,* 2005; Greil *et al.,* 2003; Ner *et al.,* in preparation). Accordingly, we asked whether the drastic reductions in the levels of H3K9me2 at the chromocentre and at the HIS-C observed in the *Su(var)3-9* mutants correlated with a disruption of HP1 targeting to these two regions. Consistent with previous results, we found that HP1 is still present at the chromocentre of *Su(var)3-9* mutants, but in much lower amounts (Figure 2.3b, panels E, H, K, N). Similarly, ChIP analyses demonstrated that although HP1 is present at the HIS-C locus in both wild-type and *Su(var)3-9* mutant embryos, it is considerably less abundant in the *Su(var)3-9* mutant extracts (Figure 3b, compare lanes 4 and 6 in the wild-type to lanes 4 and 6 in the mutants). We conclude that, like in the case of H3K9me, *Su(var)3-9* mutants display reduced levels of HP1 associated with pericentric heterochromatin and with the HIS-C.

a.

H3K9me2

H3

tubulin

core histones | wild type | Su(var)3-9^{06} | Su(var)3-9^{309} | Su(var)3-9^{330} | Su(var)3-9^{318}

1　2　3　4　5　6

b.

IN (1:10) | NO Ab | IP (H3K9me2) | IN (1:100) | NO Ab | IP (HP1)

wild-type

Su(var)3-9^{06}

Su(var)3-9^{309}

Su(var)3-9^{330}

Su(var)3-9^{318}

1　2　3　4　5　6

c.

H3 K9me2　HP1　merge

wild type | A | B | C

Su(var)3-9^{06} | D | E | F

Su(var)3-9^{309} | G | H | I

Su(var)3-9^{330} | J | K | L

Su(var)3-9^{318} | M | N | O
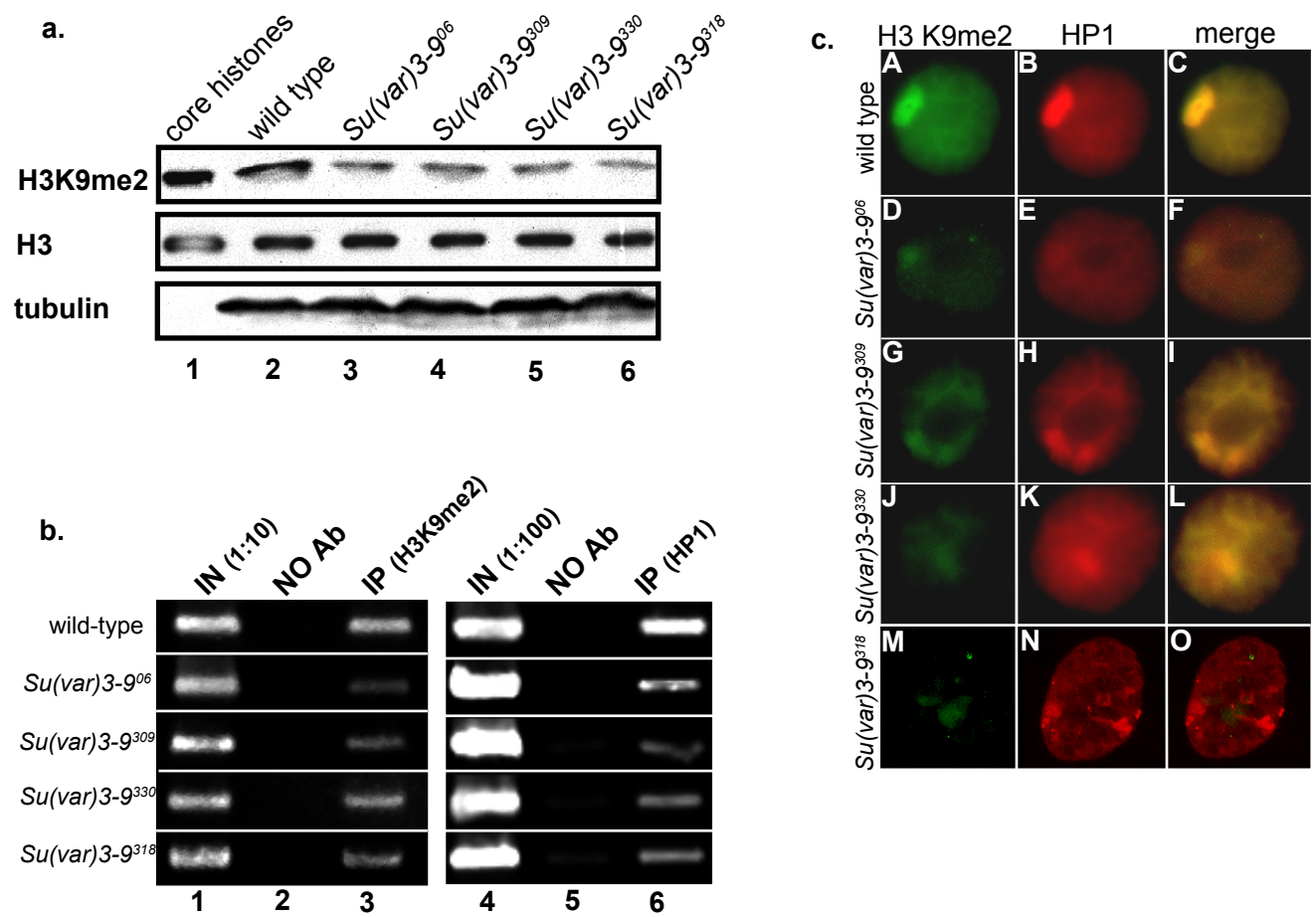
Figure 2.3.

Figure 2.3. Detection of H3K9me2 and HP1 *in Su(var)3-9* **mutants. a)** Western blot analysis of wild-type and mutant embryo extracts. Lane 1: bulk core histones (control), lanes 2: wild-type embryo extracts, lanes 3-6: extracts from homozygous *Su(var)3-9* mutant embryos. The allele numbers are indicated above the panel. The antibodies used in each case are indicated to the left of the panels. Each lane was loaded with the following: core histones (2 µg), extracts (15 µg for the H3 blot and 150 µg for the H3K9me2/tubulin blots). The H3K9me2 and tubulin panels represent the same western blot probed first with anti-H3K9me2 and subsequently with anti-tubulin. **b)** Chromatin immunoprecipitation of wild-type and *Su(var)3-9* mutant embryo extracts. In each case, the recovered DNA was PCR amplified using primers specific for the coding region of the *histone H3* gene (H3F and H3R, see Suppl.table 2). Lane 1: input DNA at a 1:10 dilution (IN (1:10)), lanes 2 and 5: mock IP, no antibody (NO Ab), lane 3: IP with anti-H3K9me2 (IP (K9me2)), lane 4: input DNA at a 1: 100 dilution (IN 1:100), lane 6: IP with anti-HP1 (IP (HP1)). **c)** Immunostaining of polytene nuclei from salivary glands of wild-type and *Su(var)3-9* mutant larvae with anti-H3K9me2 and anti-HP1 (HP1). The genotypes of the larvae are indicated to the left of each set of panels. The arrowheads point to the chromocenter region.

2.2.5. The HMTase activity of the Su(VAR)3-9 variants is impaired.

Since all *Su(var)3-9* missense alleles tested displayed a reduction in the levels of

H3K9me2, and carried mutations in the catalytic region of Su(VAR)3-9, we next asked

how much, if any, HMTase activity is retained by each of the mutant proteins, and

whether the levels of residual activity correlate with the positions of the different amino

acid substitutions. To address this, we tested the *in vitro* enzyme activity of five of our

seven Su(VAR)3-9 mutant proteins. We chose two mutants with substitutions in the

preSET (*Su(var)3-9^{376}* and *Su(var)3-9^{309}*), two in the SET domain (*Su(var)3-9^{311}* and

Su(var)3-9^{330}), and one in the postSET domain (*Su(var)3-9^{318}*). *Su(var)3-9^{330}* and

Su(var)3-9^{311} were selected because they represent substitutions of a highly conserved

(ASP536), and relatively un-conserved (GLY521) residues, respectively.

We produced various GST-Su(VAR)3-9 recombinant proteins (referred to as

Su(VAR)3-9^{allele} number for simplicity) and tested their HMTase activity. As an additional

control, we generated a GST-Su(VAR)3-9 fusion protein that carries the 2

polymorphisms (ALA304PHE, ILE375LEU) detected in some of our strains (Table 2.1). We

named this recombinant protein Su(VAR)3-9^{pol}, and expected it to retain wild-type levels

of enzyme activity since these polymorphisms don't result in suppression of PEV.

HMTase assays were first performed on bulk histones. Although the natural substrate

for Su(VAR)3-9 is histone H3, we used all the histones to rule out the possibility that the

mutants had acquired activity towards the other histones. As expected, Su(VAR)3-9^{WT}

(Fig. 2.4b, lane 2) and Su(VAR)3-9^{pol} (data not shown) efficiently catalyzed the transfer of

the methyl moiety onto H3. All mutant Su(VAR)3-9 proteins showed either no activity or a

dramatically reduced HMTase function compared to Su(VAR)3-9^{WT}. Su(VAR)3-9^{309} (Fig.

2.4b, lane 4) and $S_U(VAR)3-9^{376}$ (data not shown) had no detectable HMTase activity. $S_U(VAR)3-9^{330}$, $S_U(VAR)3-9^{318}$ and $S_U(VAR)3-9^{311}$ showed partial loss of function and they varied in the level of residual enzyme activity (Fig 2.4b, lanes 5-7). Finally, none of the mutants showed any obvious activity on histones other than H3.

Since bulk histones are prepared from nuclei, they contain extensively modified histones, which could have skewed our HMTase measurements. Therefore, we re-tested the mutants using unmodified recombinant histone H3. The results of this analysis were similar to those obtained for the bulk histones (Fig 2.4c and 2.4e). $S_U(VAR)3-9^{309}$ and $S_U(VAR)3-9^{376}$ showed no detectable HMTase function (Fig. 2.4d, lanes 11 and 15), while $S_U(VAR)3-9^{330}$, $S_U(VAR)3-9^{318}$ and $S_U(VAR)3-9^{311}$ had reduced levels of activity (Fig. 2.4d, compare lanes 12-14 with lane 9). The relative activities of the different mutants are shown in Figure 2.4e and Table 2.2. Finally, as expected, $S_U(VAR)3-9^{pol}$ methylated H3 as efficiently as $S_U(VAR)3-9^{WT}$ (99.0% +/-1.55%, data not shown).

In summary, the *in vitro* HMTase data indicate that all 5 variant $S_U(VAR)3-9$ proteins have dramatically reduced enzymatic activity. While preSET mutants displayed a complete loss-of-function, SET and postSET domain mutants were hypomorphs that retained a small fraction of the wild-type enzyme activity.

Figure 2.4
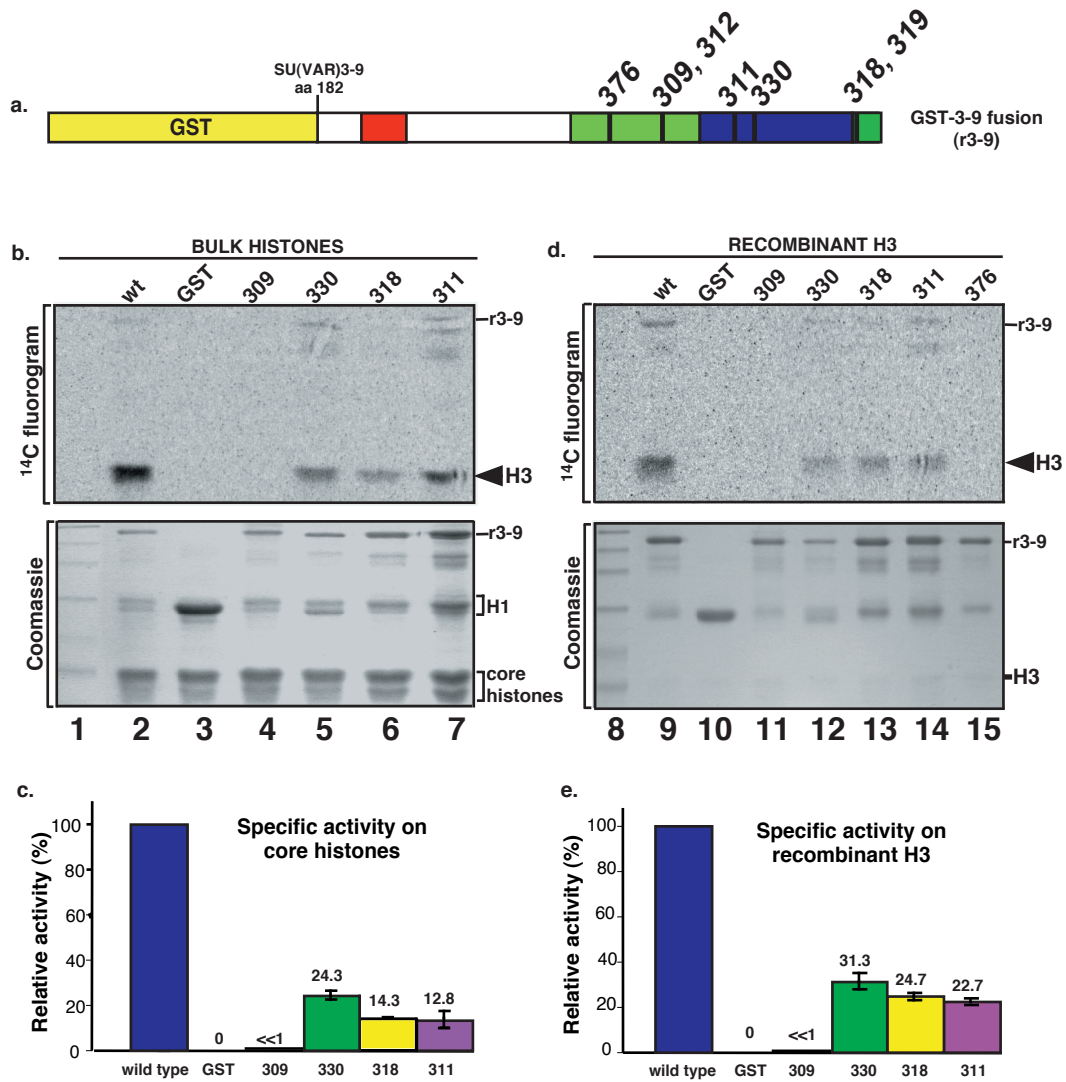
Figure 2.4. **The HMTase activity of variant SU(VAR)3-9 proteins. a)** Each

SU(VAR)3-9 variant protein was expressed and purified as a GST fusion polypeptide

(r3-9). The position of the chromodomain (red box), the pre- and postSET (green box)

and the SET domain (blue box) are indicated in the stick diagram. The vertical black

bars indicate the positions of the amino acid substitutions in 5 *Su(var)3-9* alleles that

were selected for HMTase analysis. The numbers above each bar indicate the

corresponding allele. **b)** and **d)** HMTase assay on bulk histones and recombinant H3

respectively. The enzyme activity of the variant proteins was tested using labeled Ado-

Met as the methyl donor. The fluorogram (top panel) and the corresponding Coomassie

stained gel (bottom panel) are shown. Lanes 1 and 8: molecular weight marker, lanes 2

and 9: wild-type GST-SU(VAR)3-9, lanes 3 and 10: recombinant GST (negative

control), lanes 4-7 and 11-15: GST-SU(VAR)3-9 mutant proteins. **c)** and **e)** Relative

HMTase activity of the various mutants on bulk histones and recombinant H3,

respectively. Presented as a bar graph. Following each assay, the methyl H3

radioactive signals and the Coomassie-stained bands corresponding to the appropriate

GST-SU(VAR)3-9 proteins were quantified using the ImageQuant and NIH Image

softwares respectively (see materials and methods). The bars represent the average of

3 independent assays and the error bars span two S.E.M. Both *Su(var)3-9^{309}* and

Su(var)3-9^{376} displayed no detectable activity, but for simplicity only *Su(var)3-9^{309}* is

shown on the graph.

2.2.6. <u>A mutation in the postSET prevents the addition of a third methyl group to</u>

<u>H3K9me2</u>**.**

There are a wide variety of SET domain-containing HMTases that differ in the

lysine residues that they target and/or the number of methyl groups they transfer

(Eskeland *et al.,* 2004; Mis *et al.*, 2006). Some methyltransferases can exclusively

mono-methylate H3K9 while others, like Su(VAR)3-9, can (mono-), di- or trimethylate

(Eskeland *et al.,* 2004). X-ray crystal structural analyses have revealed the key amino

acid residues in the preSET, SET and postSET domains that contribute to these

activities. For example, the presence of VAL569 in the active site and PHE602 in the

enzyme channel allows addition of multiple methyl groups to H3K9 by creating sufficient

space to accommodate tri-methylated H3K9 (Xiao *et al.,* 2003; Collins *et al.,* 2005;

Zhang *et al.,* 2003).

Since some of the mutations in our *Su(var)3-9* alleles affect amino acids located

in close proximity to residues in the active site, we asked if the variant Su(VAR)3-9

proteins, which show a partial loss of function, have an altered ability to mono/di- or

trimethylate. The bulk histones and recombinant H3 assays did not allow us to

distinguish between these three levels of K9 methylation. So, we addressed this issue

using H3 tail peptides. We took advantage of an unmodified peptide corresponding to

residues 1-20 of H3, and a peptide comprising amino acids 1-21 of H3 in which K9 is

dimethylated. Both H3 peptides are very efficient substrates for Su(VAR)3-9^{WT} and are

readily methylated (Fig 2.5a and 2.5c, lanes 1 and 7), clearly showing that Su(VAR)3-9 is

able to catalyze the addition of a third methyl group to H3K9me2.

The mutant protein Su(var)3-9[309] showed no enzyme function on either the unmodified or the dimethylated peptides (Fig. 2.5a, lane 3, and 2.5c, lane 9). Su(var)3-9[311] was a hypomorph, and it retained a fraction of the wild-type activity (Fig. 5a, lane 6, and 2.5c, lane 12, respectively) on both peptides. These results are similar to those obtained when full length (unmodified) H3 was used as a substrate. Thus the activity of the Su(var)3-9[309] and Su(var)3-9[311] mutant forms of the enzyme was consistent on all substrates tested.

In contrast, Su(var)3-9[318] behaved differently with the different substrates. While it had reduced activity towards the unmodified peptide and the full-length histone H3, it was completely inactive toward the H3K9me2 peptide (Fig. 2.5a and 2.5c, lanes 5 and 11). Thus, not only did Su(var)3-9[318] have a reduced overall catalytic activity, but it also appeared to be unable to tri-methylate K9. Su(var)3-9[318] has a SER616LEU substitution in the postSET domain.

Su(var)3-9[330] also showed a partial loss-of-enzyme function when either bulk histones or recombinant H3 were used as the substrate. However, rather surprisingly, this variant was unable to methylate the H3 peptides. It retained less than 1% and 4% activity towards the unmodified and H3K9me2 peptides, respectively (Fig. 2.5a and 2.5c, lanes 4 and 10).

In summary, the Su(var)3-9 mutants displayed an array of severely reduced enzyme activities, ranging from complete abolition to ~25% of the wild-type activity. In addition, some of the mutants showed distinct HMTase characteristics (summarized in Table 2.2) depending on the substrate used.

Figure 2.5
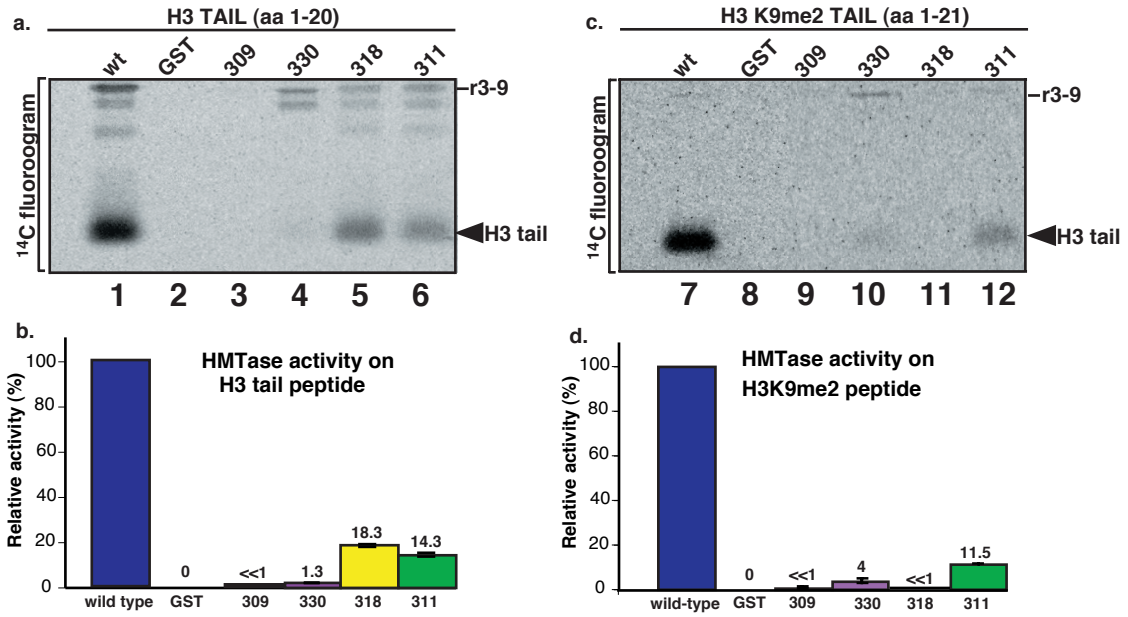
Figure 2.5. HMTase activity of *Su(var)3-9* alleles using unmodified and dimethylated H3 tail peptides. a) and **c)** Fluorograms of an HMTase assay on H3 peptide (H3 tail, aa 1-20) and dimethylK9H3 peptide (H3K9me2 tail, aa 1-21) respectively. Lanes 1 and 7: wild-type GST-SU(VAR)3-9, lanes 2 and 8: GST (negative control), lanes 3-6 and lanes 9-12: GST-SU(VAR)3-9 mutants. The allele numbers are indicated above the corresponding lanes. **b)** and **d)** Relative *in vitro* activity of the indicated *Su(var)3-9* alleles on unmethylated (b) and dimethylK9 (d), H3K9me2 H3 tail peptides. The numbers used in the bar graphs were obtained as described in the materials and methods.

2.2.7. <u>Effect of *Su(var)3-9* missense and null alleles on PEV in the presence of one mutant and 2 wild type copies of *Su(var)3-9*.</u>

All *Su(var)3-9* mutants were originally identified on the basis of a morphological phenotype, dominant suppression of PEV. Here, we examine the morphological phenotype of the *Su(var)3-9* mutants in more detail. We specifically asked whether the mutant alleles differ in their ability to suppress heterochromatic gene silencing that results from PEV and, if so, whether these are correlated to the differences observed at the biochemical level (*in vitro* enzyme function). Direct examination of the *Su(var)3-9* alleles with the "classical" variegating strain, w^{m4}, is inadequate for this study since all the *Su(var)3-9* mutants almost completely suppress w^+ gene silencing. So, we used a variation on the w^{m4} assay system and examined the suppression effect in the presence of an extra copy of wild-type *Su(var)3-9*, that is, in flies with w^{m4}; *Su(var)3-9^{mutation}/Su(var)3-9^+, Su(var)3-9^+* genotypes. We took advantage of a *Su(var)3-9::eGFP* transgene inserted in the third chromosome ($pP\{GS[ry^+,(10kb Su(var)3-9)EGFP]\}$, (Schotta and Reuter, 2000), here referred to as *P[3-9egfp]* for simplicity). Crosses were set up, in triplicate, between males homozygous for this transgene and for the wild-type, endogenous *Su(var)3-9* allele (w^{m4}/Y; *Su(var)3-9^+,P[3-9egfp]/Su(var)3-9^+, P[3-9egfp]*) and females homozygous for each of the *Su(var)3-9* alleles (see material and methods). The progeny of these crosses carry one maternally inherited mutant copy (or, in the case of the control cross, one wild-type copy) and two paternally derived wild-type copies of *Su(var)3-9* (the endogenous *Su(var)3-9^+* and the ectopically inserted *P[3-9egfp]*). For each cross, the eyes of the male progeny were visually scored for the amount of eye pigment and thus the level of suppression of PEV.

84

Over 98% of the flies derived from the control cross, which have 3 copies of the *Su(var)3-9⁺* allele (*w^{m4}/Y; Su(var)3-9⁺/Su(var)3-9⁺,P[3-9egfp]*), displayed a strong *E(var)* phenotype, confirming that the *Su(var)3-9::eGFP* transgene is functional and over-production of SU(VAR)3-9 enhances the gene silencing that results from PEV (Fig. 2.6). On the other hand, in the progeny bearing a mutant allele, a wild-type (endogenous) allele and an extra (transgenic) copy of *Su(var)3-9⁺*, we observed a broad range of variegating eye phenotypes. None of the individuals displayed the almost completely white eye phenotype of their control counterparts. However, we did note a significant proportion of flies with bilaterally unequal eye pigmentation, that is, one eye strongly suppressed and the other unsuppressed. For this reason, instead of performing standard pigment assays, we scored each eye visually and assigned it to one of three categories: strongly suppressed (~75-100% pigment), mildly suppressed (~20-75%), or unsuppressed (~5-20%) (Fig. 2.6b, panels A, B & C respectively). The proportion of strongly suppressed, mildly suppressed and unsuppressed eyes varied with the genotype (Fig. 2.6b). For the sake of simplicity, and since all these individuals have the identical wild type and transgenic alleles, and only differ in their mutant allele of *Su(var)3-9*, the description of their genotypes is limited to the identity of the mutant allele they carry.

Over half of the individuals carrying the *Su(var)3-9^{06}* allele had strongly suppressed eyes. Since *Su(var)3-9^{06}* is a null allele for which no mRNA or protein products are detected (Fig. 2.2 and Tschiersch *et al.,* 1994; Schotta *et al.,* 2002), these individuals should be phenotypically equivalent to *w^{m4}, Su(var)3-9⁺/Su(var)3-9⁺* flies, which typically have variegating, unsuppressed eyes (Tartof *et al.,* 1984; Grigliatti, 1991;

Reuter and Spierer, 1992, and references therein). Thus, the presence of a large proportion of strongly suppressed eyes in w^{m4}/Y; $Su(var)3$-9^{06}/$Su(var)3$-9^+, $P[3$-$9egfp]$ individuals suggests that the ectopically inserted $Su(var)3$-9 transgene, $P[3$-$9egfp]$, may not be functionally equivalent to the endogenous gene. The EGFP-tagged $Su(var)3$-9 may produce less protein product, and/or its product may function less efficiently than the wild-type SU(VAR)3-9.

The highest proportion of strongly suppressed eyes was observed in flies carrying the $Su(var)3$-9^{311} allele, which suggests $Su(var)3$-9^{311} is the strongest suppressor of PEV in this genetic test. Individuals bearing the $Su(var)3$-9^{318} and $Su(var)3$-9^{330} alleles displayed approximately the same fraction of strongly suppressed eyes, indicating that $Su(var)3$-9^{318} and $Su(var)3$-9^{330} suppress PEV roughly to the same degree. Curiously, in this assay, $Su(var)3$-9^{309} resulted in the lowest proportion of strongly suppressed, and the highest proportion of unsuppressed eyes, suggesting $Su(var)3$-9^{309} is a weaker suppressor of PEV than $Su(var)3$-9^{330}, $Su(var)3$-9^{318} or $Su(var)3$-9^{311}.

In summary, differences in the strength of PEV suppression were detected among the four $Su(var)3$-9 missense alleles. Using this particular genetic assay of function, $Su(var)3$-9^{311} was the strongest suppressor, while $Su(var)3$-9^{309} was the weakest, and also the only one to suppress PEV less efficiently than the «protein null» allele, $Su(var)3$-9^{06}. With the exception of $Su(var)3$-9^{309}, these results correlate well with the levels of remaining HMTase activity of each mutant (see Discussion).

a.

EGFP-tagged
Su(var)3-9

wild-type
Su(var)3-9

mutant
Su(var)3-9

b.

A | B | C | D

■ strongly sup ■ mildly sup ■ unsuppressed □ enhanced

$w^{m4}/Y; \dfrac{Su(var)3\text{-}9^{+}}{Su(var)3\text{-}9^{+},P[3\text{-}9egfp]}$

(n=3; tot=410)

A: 0.30 +/- 0.42
B: 0
C: 2.69 +/- 1.89
D: 97.01 +/- 2.31

$w^{m4}/Y; \dfrac{Su(var)3\text{-}9^{06}}{Su(var)3\text{-}9^{+},P[3\text{-}9egfp]}$

(n=3; tot=400)

A: 53.51 +/- 1.96
B: 29.29 +/- 2.38
C: 17.20 +/- 1.26
D: 0

$w^{m4}/Y; \dfrac{Su(var)3\text{-}9^{309}}{Su(var)3\text{-}9^{+},P[3\text{-}9egfp]}$

(n=3; tot=356)

A: 37.71 +/- 0.07
B: 21.79 +/- 0.01
C: 40.49 +/- 0.02
D: 0

$w^{m4}/Y; \dfrac{Su(var)3\text{-}9^{330}}{Su(var)3\text{-}9^{+},P[3\text{-}9egfp]}$

(n=3; tot=524)

A: 65.68 +/- 0.81
B: 19.88 +/- 1.03
C: 14.44 +/- 0.27
D: 0

$w^{m4}/Y; \dfrac{Su(var)3\text{-}9^{318}}{Su(var)3\text{-}9^{+},P[3\text{-}9egfp]}$

(n=3; tot=567)

A: 66.76 +/- 1.59
B: 24.28 +/- 2.30
C: 8.95 +/- 0.73
D: 0

$w^{m4}/Y; \dfrac{Su(var)3\text{-}9^{311}}{Su(var)3\text{-}9^{+},P[3\text{-}9egfp]}$

(n=3; tot=562)

A: 86.20 +/- 3.26
B: 10.73 +/- 2.53
C: 2.90 +/- 0.89
D: 0

Figure 2.6.

Figure 2.6. Effect of several *Su(var)3-9* alleles on PEV (here, *w^m4* variegation*)* in *w^m4/Y; Su(var)3-9^+, P[3-9egfp]/Su(var)3-9^mutant* individuals. a) A schematic of the genetic configuration of the flies examined in this experiment. The EGFP-tagged *Su(var)3-9* (also indicated as *P[3-9egfp]* for simplicity) represents the ectopic *Su(var)3-9* insertion pP{GS[ry^+,(10kb Su(var)3-9)EGFP} (Schotta and Reuter, 2000). This construct is inserted on the third chromosome, thus all flies harboring this insert also carry a wild-type copy of the endogenous *Su(var)3-9* gene.

b) Frequencies (%) of strongly suppressed, mildly suppressed, unsuppressed, and enhanced eyes in *w^m4/Y; Su(var)3-9^+, P[3-9egfp]/Su(var)3-9^mutant* individuals carrying different *Su(var)3-9* alleles. Pictures A-D show examples of strongly suppressed, mildly suppressed, unsuppressed and enhanced *w^m4* variegation, respectively. Genotypes are shown on the left, n=number of crosses analyzed, tot=total number of eyes scored. The bar graphs below each eye picture represent the fraction (%) of eyes displaying the depicted phenotype amongst the individuals of the genotype indicated (average of 3 independent crosses +/- S.E.M., rounded to the closest second decimal). Panels A and C show eyes of *w^m4/Y; Su(var)3-9^309/Su(var)3-9^+,P[3-9egfp]* flies. The eye in panel B is from a *w^m4/Y; Su(var)3-9^06/Su(var)3-9^+,P[3-9egfp]* fly and the eye in D is from a *w^m4/Y; Su(var)3-9^+/Su(var)3-9^+,P[3-9egfp]* individual.

2.3. Discussion

2.3.1. The *Su(var)3-9* mutations are single base-pair substitutions clustered in the catalytic region.

Su(VAR)3-9 has two highly conserved and functionally distinct regions: a chromodomain and a SET domain; the latter, together with the flanking pre- and postSET sequences, constitutes the HMTase activity of the protein. Both regions are present in several chromatin associated proteins and are highly conserved, and therefore it is reasonable to assume that the chromo and SET domains each have important roles in the function of Su(VAR)3-9. Other domains of the protein, as yet unidentified, may also be required for the function of Su(VAR)3-9. Indeed, we expected that the dominant suppressors of PEV would comprise mutations in the chromo and SET domains, and perhaps identify other regions of Su(VAR)3-9 that are required for the epigenetic gene silencing observed in PEV. Interestingly, the 9 confirmed EMS-induced *Su(var)3-9* mutations altered 7 residues, with 2 residues hit twice, all of which are located in the catalytic region of Su(VAR)3-9; none of the *Su(var)3-9* mutations occurred in or near the chromodomain of the protein. The concentration of missense mutations within the preSET/SET/postSET region of the protein suggests that the HMTase activity of Su(VAR)3-9 plays a crucial role in suppression of PEV. In contrast, single amino acid substitutions in the chromodomain are probably insufficient to cause a dominant suppression of PEV and thus affect the gene silencing function of Su(VAR)3-9. Alternatively, the absence of chromodomain mutants among the dominant Su(var)s recovered in the original screen may be due to the fact that the EMS mutagenesis failed to induce mutations in or around the chromodomain of *Su(var)3-9*. This latter hypothesis

seems unlikely, because Reuter and colleagues also failed to recover mutations in the chromodomain of SU(VAR)3-9 by screening for strong, dominant suppressors of PEV (Reuter and Wolff, 1981; Ebert *et al.*, 2004). Mutations in the N-terminal region of the protein were indeed isolated in several screens using P element transposition or gamma rays as the mutagenic agent, but they represent insertions and deletions that produce either truncated proteins, or no protein at all (Tschiersch et al., 1994; Harrington, 2001; Schotta *et al*., 2002; Ebert *et al*., 2004, 2006). These mutants act as dominant Su(var)s because the *Su(var)3-9* locus is dosage sensitive, that is hemizygosity for *Su(var)3-9* suppresses PEV (Grigliatti, 1991; Reuter and Spierer, 1992; and references therein). Nevertheless, the observation that mobile genetic elements can insert into or near the chromodomain suggests that this region should be amenable to mutagenesis via EMS.

Thus, the function of the chromodomain in SU(VAR)3-9 remains elusive, although it is more defined in other proteins, such as HP1, PC and MOF. In these proteins, the chromodomain appears to be required for chromatin binding (Platero *et al.,* 1995; Messmer *et al.,* 1992; Ma *et al.,* 2001; Akhtar *et al.,* 2000; Jacobs *et al*., 2001; Bannister *et al*., 2001; Lachner *et al.,* 2001; Bouazoune *et al.,* 2002; Fischle *et al.,* 2003; Min *et al.,* 2003; Pray-Grant *et al.,* 2005). By analogy, a similar role was suggested for the SU(VAR)3-9 chromodomain (Schotta *et al.,* 2002; 2003), but, to date, direct evidence is lacking. If the SU(VAR)3-9 chromodomain is involved in chromatin binding, then one would expect mutations within the chromodomain to cause mis-targeting, which should have a strong dominant Su(var) phenotype. The absence of chromodomain mutations among the respective collections of dominant *Su(var)3-9* mutants (this paper and Ebert

et al., 2004; Donaldson *et al.,* 2002), suggests that substitution of any of the ~40 amino acids that comprise the chromodomain is not enough to cause significant mis-targeting.

2.3.2. <u>Reduction of H3K9me2 and HP1 in *Su(var)3-9* mutants.</u>

At the cellular level, we found that all *Su(var)3-9* missense mutants display a dramatic reduction in the levels of both H3K9me2 and HP1 associated with centric heterochromatin and with the HIS-C. Firstly, these *in vivo* observations corroborate the results obtained from the *in vitro* enzyme assay, namely that the amino acid substitutions present in the mutant alleles of Su(var)3-9 impair the protein's HMTase function. Missense mutations in *Su(var)3-9* also altered the abundance of HP1 binding, as measured by *in situ* immunofluorescence and ChIP analysis, which is consistent with the hypothesis that H3K9me2 constitutes a binding platform for HP1 (Bannister *et al.,* 2001; Lachner *et al.,* 2001; Nakayama *et al.,* 2001; Ebert *et al.,* 2006). Secondly, the observation that *Su(var)3-9* mutants cause the same set of effects at the chromocentre and at the HIS-C, a euchromatic locus and natural target of Su(var)3-9, suggests that Su(var)3-9 may function as part of a silencing mechanism that affects numerous loci in the euchromatic as well as heterochromatic regions of the genome.

2.3.3. The SET mutants are hypomorphs

Based on the crystal structures of the SU(VAR)3-9 homologues Clr4p and DIM-5 (Min *et al.,* 2002; Zhang *et al.,* 2003), both the HMTase active site, and the substrate-binding cleft, are part of the SET domain. The mutant alleles *Su(var)3-9³¹¹, Su(var)3-9³²⁵* and *Su(var)3-9³³⁰* cause single amino acid substitutions in the SET domain. Each of the affected residues is predicted to lie within, or very near, the active site of the enzyme (Figure 2.7). *Su(var)3-9³³⁰*, which affects ASP536, is the most interesting of these SET mutants. ASP536 is conserved in all H3K9 and in some H3K4 methyltransferases (Aagard *et al.,* 1999; Zhang *et al.,* 2002; Mis *et al.*, 2006). It is located in the portion of the cleft that is involved in stabilizing the enzyme-substrate complex; its side chain forms a hydrogen bond with the hydroxyl oxygen of SER10 of H3 (H3S10) (Min *et al*., 2002; Zhang *et al.,* 2003). The HMTase activity of SU(VAR)3-9³³⁰, on full-length, unmodified histone H3, is reduced to about 31% of the wild-type. This partial loss of function is probably the result of an unstable enzyme-substrate interaction due to the replacement of ASP536 with ASN, which could disrupt the hydrogen bond formed between H3S10 and ASP536. Interestingly, the H3 peptides are much poorer substrates for SU(VAR)3-9³³⁰ compared with full length H3. The enzyme activity on unmodified, and H3K9me2 peptides is less than 5% of wild-type. Their smaller size, and the absence of backbone residues, which are present in full length H3, reduce the stability of binding of the peptides compared to the full length H3. Thus, the substitution of ASP536 probably has a more drastic effect on the peptides than it does on the full-length histones. Alternatively, the ASP536ASN substitution in SU(VAR)3-9³³⁰ may lead to an altered substrate specificity from H3K9 to another H3 lysine that is not present in the H3 tail

peptides, for example K27. Since K27 is absent in these peptides, the only activity detected would correspond to weak, residual methylation of K9. Since no appropriately modified peptides are available that include residues 1-27 of H3, this hypothesis cannot be tested without using mass spectrometry.

2.3.4. <u>The preSET mutants are enzymatically inactive.</u>

There are nine key cysteine residues in the preSET region of SU(VAR)3-9. These residues are highly conserved (Aagaard *et al.,* 1999; Min *et al.,* 2002) and crystal structures of the Clr4p and DIM-5 proteins reveal that these cysteines coordinate 3 zinc ions that form a "zinc cluster". This "cluster" has an important structural role as it holds together two random coils that form the bottom surface of the catalytic region of the protein (Min *et al.,* 2002; Zhang *et al.,* 2003). Although the preSET domain is not part of the enzyme active site, or the regions binding the substrate or cofactor *per se*, it is required for efficient H3 methylation (Rea *et al.*, 2000).

Four of our *Su(var)3-9* missense alleles have mutations in preSET cysteine residues. We tested two of them, SU(VAR)3-9[376] (CYS428TYR) and SU(VAR)3-9[309/312] (CYS462TYR), for their *in vitro* HMTase activity, and in both cases the variant proteins were completely inactive. Given the role of these residues in protein structure, the complete loss of enzyme function is likely due to misfolding of the protein, which dramatically alters many aspects of substrate binding and enzyme function.

2.3.4. <u>The postSET mutant lacks the ability to add a third methyl group to H3K9me2.</u>

Su(var)3-9[318] and *Su(var)3-9[319]* are two independently isolated mutations with the same SER616LEU mutation in the postSET region. Although this amino acid is not conserved in Clr4p, Suv39h1 and SUV39H1 (fig 1b), our results suggest that it is crucial for the addition of the third methyl group to H3K9. In addition, the fact that the corresponding residues are a GLN in Clr4p and an ASP in SUV39H1 and Suv39h1 (Fig.1b), and that the SER616LEU substitution in SU(VAR)3-9 causes a strong, dominant Su(var) phenotype, suggest that the presence of a polar residue at this position may be critical for SU(VAR)3-9 function.

In the Clr4p structure the flexible postSET region is positioned near the active site where it acts as a "lid" and creates a solvent-secluded space for optimal methyl transfer (Min *et al*., 2002). The postSET domain of DIM-5 works in a similar manner and in the presence of H3 substrate it interacts directly with the enzyme active site, with the substrate, and with the cofactor AdoMet (Zhang *et al.,* 2003). Hence, mutations in the postSET region could disrupt methyl transfer by altering the local architecture and exposing the active site to the solvent. Alternatively, the altered residue in the postSET may interfere with the normal positioning of the substrate or the cofactor in their respective binding pockets. As expected, SU(VAR)3-9[318] has reduced HMTase function on unmodified substrates such as bulk histones, recombinant H3 and H3 tail peptide (1-20). However, SU(VAR)3-9[318] fails to add a third methyl group to an H3 tail peptide already dimethylated at K9, suggesting that the mutation either interferes with the stability of the dimethylated peptide or the AdoMet in the catalytic cleft, or creates sufficient steric hindrance to impair methyl transfer.

a.

```
CLR4        220  --------------VTLVNEVDDEPCPSLDFQFISQYRLTQGVIPPDPNFQSGCNCSSLGG-CDLNNPSRCECLDDLDEPTHFAYDAQG-RVRADTGA  302
SU(VAR)3-9  361  FEKRMNHVEKPSPPIRVENNIDLDTIDS-NFMYIHDNIIGKDVPKPEAGI--VGCKCTEDTEEC-TAST-KC-CARFAG--ELFAYERSTRRLRLRPGS  451
```

B

```
CLR4        303  VTYECNSFCSCSMECPNRVVQRGRTLPLEIFKTKE-KGWGVRSLRFAPAGTFITCYLGEVITSAEAAKRDKNYDDGITYLFDLDMFD-DASEYTVDA  398
SU(VAR)3-9  452  ATYECNSRCSCDSSCSNRLVQHGRQVLVLFKTANGSGWVRAATALRKGEFVCEYIGEIITSDEANERGKAYDDNGRTYLFDLDYNTAQDSEYTIDA  549
```

C

```
CLR4        399  QNYGDVSRFFNHSCSPNIATYSAVRNHGFRTIYDLAFFAIKDIQPLEELTFDYAGAKDFSPVQSQKSQQNRISKLRQCKCGSANCRGWLFG  490
SU(VAR)3-9  550  ANYGNISHFIINHSCDPNLAVFPCWITEHLNWALPHLVFFTLRPIKAGEELSFDY-IRADNEDVPYENLST--A--VRVECRCGADNCRKVLF-  635
```

D

309 312   325   318 319   311   376   330   324

b.

N-terminus • C-terminus • zinc cluster • 311 • 330 • 309 312 • 318 319

Figure 2.7.

95

Figure 2.7. Probable tertiary structure of SU(VAR)3-9 and relative positions of the mutated residues. a) Amino acid sequence alignment of the catalytic domain of Clr4p and SU(VAR)3-9. Identical residues are highlighted in black, similar residues are highlighted in grey. Red dots indicate the residues that are mutated in the different *Su(var)3-9* alleles. Above the alignment is represented the secondary structure of Clr4p (aa 220-490), according to Min *et al.* (2002). The N-terminal region of the catalytic domain is shown in black, the pre- and postSET domains in green, and the SET domain in blue. The high level of homology between the two proteins at the primary sequence level suggests that their secondary and tertiary structures could also be very similar. **b)** A ribbon diagram representation of the Clr4 structure (residues 220-490). Under the assumption that the tertiary structure of SU(VAR)3-9 is very similar to Clr4's, the red dots show the approximate positions of the amino acids that are mutated in SU(VAR)3-9[309], SU(VAR)3-9[330], SU(VAR)3-9[318] and SU(VAR)3-9[311]. The diagram was generated partially using Raster3D (Kraulis, 1991) and MolScript (Merritt and Murphy, 1994).

2.3.5. <u>Effect of *Su(var)3-9* missense and null alleles on PEV in a strain that is hyperploid for *Su(var)3-9*.</u>

In our structure-function analysis of *Su(var)3-9* we wanted to examine the morphological phenotype of PEV and attempt to correlate the morphological end point (eye colour pattern) with the cytological and molecular effects of the different mutations in *Su(var)3-9*. However, all *Su(var)3-9* missense mutants suppress PEV very strongly. Therefore, we employed an assay for PEV, which examined the relative strength of the *Su(var)3-9* mutations in individuals that carried three copies of *Su(var)3-9*. The three alleles were one wild-type, endogenous allele (*Su(var)3-9*$^+$), one *Su(var)3-9*$^+$-*EGFP* transgene, and one mutant *Su(var)3-9* allele.

At the morphological level (suppression of PEV) we found that all *Su(var)3-9* mutant alleles were able to suppress PEV in at least a substantial fraction of w^{m4}/Y; *Su(var)3-9*^{mutant}/*Su(var)3-9*$^+$, *P[3-9egfp]* individuals, but they differed in their strength of suppression. These differences may be due to differences in the residual HMTase activity of the mutant SU(VAR)3-9 proteins. However, it should be kept in mind that, unlike the HMTase assay, which provides a direct measure of enzyme activity under given conditions, or the immunohistochemical analyses, which identify the distribution patterns and can delineate the relative abundance of given proteins, PEV suppression is a tertiary phenotype involving numerous unknown variables, and may therefore not be directly indicative of SU(VAR)3-9 HMTase function.

Flies carrying the *Su(var)3-9*^{06} mutant allele (w^{m4}/Y; *Su(var)3-9*^{06}/*Su(var)3-9*$^+$, *P[3-9egfp]*) served as "baseline", as *Su(var)3-9*^{06} does not produce any *Su(var)3-9* mRNA or protein, and the products of the endogenous *Su(var)3-9*$^+$ and the transgenic

Su(var)3-9⁺-EGFP must therefore account for the Su(VAR)3-9 function present in this strain. Individuals of this genotype showed strongly suppressed eyes at a frequency of about 53%. Curiously, the presence of missense *Su(var)3-9* allele (w^{m4}/Y; *Su(var)3-9^{missense}/Su(var)3-9⁺, P[3-9egfp]*) resulted in either stronger (*Su(var)3-9^{330}*, *Su(var)3-9^{318}*, *Su(var)3-9^{311}*)or weaker (*Su(var)3-9^{309}*) suppression of PEV than with the *Su(var)3-9^{06}* null allele (Fig. 6b). *Su(var)3-9* alleles resulting in stronger suppression of PEV than *Su(var)3-9^{06}* are likely antimorphs, while alleles that are weaker suppressors than *Su(var)3-9^{06}*, are probably hypomorphs. This logic suggests that *Su(var)3-9^{311}*, *Su(var)3-9^{318}* and *Su(var)3-9^{330}* are antimorphs, and *Su(var)3-9^{309}* is a hypomorph.

At a mechanistic level, the three antimorphic mutants possibly act as dominant negatives. The mutant Su(VAR)3-9 products may be incorporated into protein complexes like their wild-type counterparts, and these complexes would be correctly targeted, but would fail to efficiently methylate H3K9, interfering with the function of the wild-type Su(VAR)3-9. Indeed, the mutations present in *Su(var)3-9^{311}*, *Su(var)3-9^{318}* and *Su(var)3-9^{330}* are located in the catalytic region of the protein, and do not affect its N-terminal protein-protein interaction domains. However, the possibility that amino acid substitutions in the SET domain may play a role in the assembly, stability, or targeting of Su(VAR)3-9-containing complexes cannot be excluded.
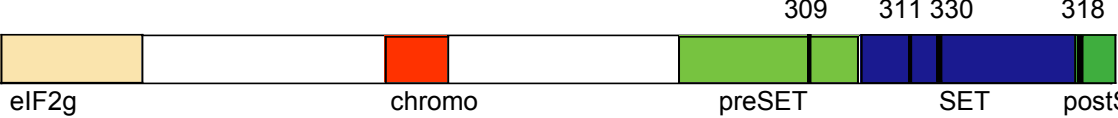
In this hyperploid genotype, *Su(var)3-9^{311}* is a stronger suppressor of PEV than *Su(var)3-9^{330}*. This may be due to the fact that Su(VAR)3-9^{311} retains significantly less HMTase activity than Su(VAR)3-9^{330}, as the *in vitro* enzyme assays indicate. In this genetic assay, *Su(var)3-9^{311}* is also a morphologically stronger PEV suppressor than *Su(var)3-9^{318}*, but the *in vitro* enzyme activities of Su(VAR)3-9^{311} and Su(VAR)3-9^{318} do

not differ significantly. This may be explained by the peculiar biochemical phenotype of SU(VAR)3-9^{318}. If, as suggested by the *in vitro* data (Figure 5), this mutant protein is an inefficient HMTase that is also completely unable to trimethylate histone H3, then its observed residual activity on unmodified substrates probably represents mono/dimethylation exclusively. Therefore, comparing the HMTase activity of SU(VAR)3-9^{318} and SU(VAR)3-9^{330} or SU(VAR)3-9^{311} based on the relative amount of radiolabeled methyl transferred onto unmodified H3 substrate, could be misleading. With respect to mono/dimethylation, SU(VAR)3-9^{318} may well retain as much activity as SU(VAR)3-9^{330} (*i.e.* significantly more than SU(VAR)3-9^{311}) and, since H3K9me2 is sufficient for heterochromatic silencing in *Drosophila* (Fischle *et al.,* 2003; Swaminathan *et al.,* 2005), it would not be surprising that *Su(var)3-9^{318}* and *Su(var)3-9^{330}* may be equivalent in their ability to suppress PEV, and that neither is as strong as *Su(var)3-9^{311}*. However, we cannot exclude the possibilities that 1) the *in vitro* HMTase activity of the mutant SU(VAR)3-9 proteins is not always a good indicator of their *in vivo* activity, and/or 2) the *Su(var)3-9* missense mutations affect more than just the protein's enzyme activity, and the enzyme activity of the different mutant SU(VAR)3-9 proteins is not the only factor determining the strength of the Su(var) phenotype.

Interestingly *Su(var)3-9^{309}*, which behaves as a hypomorph in the PEV assay employed here, results in a protein product with no *in vitro* HMTase activity at all. We propose that the amino acid substitution present in SU(VAR)3-9^{309} (CYS462TYR) causes the catalytic region of the protein to be misfolded since it affects a residue that likely plays an important structural role (Min *et al.,* 2002). Thus, *in vitro*, the complete loss of enzyme function associated with this mutation is due to severe misfolding of the

polypeptide. Furthermore, we suggest that the catalytic domain of SU(VAR)3-9[309] is

misfolded *in vivo*. This misfolding not only renders it inactive, but also prevents it from

being incorporated into SU(VAR)3-9-containing complexes, therefore not interfering with

the function of the wild-type protein. Hence, *Su(var)3-9[309]* is not an antimorph. We are

left with the observation that *Su(var)3-9[309]* is a weaker suppressor of PEV than the null

allele, *Su(var)3-9[06]*, indicating that *Su(var)3-9[309]* is a hypomorph. One possibility is that,

occasionally (*i.e.* at a low frequency, *e.g.* 10% of the time), SU(VAR)3-9[309] is still

incorporated into the SU(VAR)3-9-containing complex, and its incorporation into the

complex stabilizes its tertiary structure. In such cases, SU(VAR)3-9[309] is able to

methylate histone H3 like its wild-type counterpart, since its active site is intact.

Table 2.2. Summary of the molecular, biochemical, and Su(var) phenotypes of five

Su(var)3-9 mutant alleles examined in this study. The stick diagram of SU(VAR)3-9

shows the positions of the point mutations in *Su(var)3-9309*, *Su(var)3-9311*, *Su(var)3-9330*

and *Su(var)3-9318* (vertical black bars). The chromodomain (red box), pre- and postSET

domains (green boxes) and SET domain (blue box), and the N-terminal region that is in

common with eIF2g are also indicated. For each allele, the result of the mutation (amino

acid change), the relative amount of residual HMTase activity with 3 different

substrates, and the percentage of strongly suppressed eyes in the PEV assay are

shown.

Su(var)3-9 allele		H3K9me2 *in vivo*	HMTase *in vitro* (% of wild-type)			% strongly suppressed eyes
			H3	H3 tail	H3K9me2 tail	
06	No gene product	Reduced	N/A	N/A	N/A	53.51+/-1.96
309	C462Y	Reduced	0	0	0	37.71+/-0.07
311	G521D	Reduced	22.7+/-1.2	14.3+/-1.4	11.5+/-0.2	86.20+/-3.26
330	D536N	Reduced	31.3+/-3.8	1.2+/-0.2	4.0+/-1.1	65.68+/-0.81
318	S616L	Reduced	24.7+/-1.7	18.3+/-0.3	0	66.76+/-1.59

2.4. Materials and Methods

2.4.1. *Drosophila* strains

Unless otherwise specified, all fly strains were grown under standard conditions on glucose/yeast/cornmeal medium, with Tegosept (methyl-*p*-hydroxybenzoate) as a mold inhibitor.

2.4.2. Recombination mapping

Allelism to *Su(var)3-9* was determined by recombination analysis. Each putative EMS-induced *Su(var)3-9* mutant (w^{m4};*Su(var)X/Su(var)X*) was first crossed to w^{m4};*Su(var)3-9^{P25}*, which harbours a P-element insert in the first intron of the dual *Su(var)3-9*/eIF2g transcription units (Harrington, 2001; Ner *et al.*, 2002). F1 females (w^{m4}/w^{m4}; *Su(var)3-9^{P25}/Su(var)X*) were then crossed to w^{m4}/Y;+/+ males, and in each case >2000 offspring were scored with respect to PEV suppression. Mutants that did not yield any w^{m4};+/+ recombinants, indicating that the distance between *Su(var)3-9* and the *Su(var)X* mutation they carried was less than 0.1 cM, were further characterized by DNA sequence analysis. The maximum distance between *Su(var)X* and *Su(var)3-9* was calculated as if the next fly to eclose would have been a wild-type recombinant; max distance = 2 [since the reciprocal event yields a double mutant, indistinguishable from the parentals]/(1+total number of flies scored).

102

2.4.3. Suppression of PEV in the presence of 2 wild-type copies of *Su(var)3-9*

Homozygous w^{m4}; *Su(var)3-9* females were crossed to w^{m4}/Y; *pP{GS[ry+,(10kb Su(var)3-9)EGFP]}* homozygous males (Schotta and Reuter, 2000) and the eye colors of the offspring were scored. Only males were scored in order to avoid effects due to the presence of two copies of the *white* gene. All crosses were conducted in triplicate at 18 ºC. For each cross, a total of 356 to 580 eyes were examined. The phenotype of each eye was classified as "strongly suppressed", "mildly suppressed", "unsuppressed" (w^{m4}-like), or enhanced. The reciprocal crosses gave similar results, but the number of offspring was much lower due to the low fecundity of w^{m4}/ w^{m4}; *pP{GS[ry+,(10kb Su(var)3-9)EGFP]}* homozygous females at 18 ºC. Crosses were set up in triplicates, and strength of PEV suppression was assessed based on the percentage of strongly suppressed eyes in the male offspring of each cross (average +/- S.E.M.). Student's T-tests (p=0.05) were employed to determine whether differences between alleles were statistically significant.

2.4.4. Sequence analysis

Genomic DNA was extracted from the following stocks: Oregon-R, w^{m4}, w^{m4};*Su(var)3-9*^{309}/*Su(var)3-9*^{309}, w^{m4};*Su(var)3-9*^{312}/*Su(var)3-9*^{312}, w^{m4};*Su(var)3-9*^{311}/*Su(var)3-9*^{311}, w^{m4};*Su(var)3-9*^{317}/*TM3,Sb,Ser*; w^{m4};*Su(var)3-9*^{318}/*Su(var)3-9*^{318}, w^{m4};*Su(var)3-9*^{324}/*TM3,Sb,Ser*, w^{m4};*Su(var)3-9*^{325}/*TM3,Sb,Ser*, w^{m4};*Su(var)3-9*^{327}/*TM3,Sb,Ser*, *Su(var)3-9*^{319}/*TM3,Sb,Ser*, and w^{m4};*Su(var)3-9*^{330}/*Su(var)3-9*^{330}. Segments of the *Su(var)3-9* gene were amplified by PCR using the following primer

pairs: 39KYLE and 3SET, 3-95' and 39-1, 39-2 and 3-95', 5RI and 3RI, and 5SET and

3SET (all primer sequences are listed in supplementary table 2). The amplification

conditions were: 94ºC, 5 min; (94ºC 45 sec, 57-60ºC, 30 sec, 72ºC, 1.5 min) for 30

cycles, 72ºC, 10 min. Each PCR reaction was performed three times and both strands

of each product were sequenced twice.

2.4.5. RT-PCR

Total RNA was prepared from wild-type and *Su(var)3-9^{06}* homozygous flies by

TRIzol® extraction as recommended by the manufacturer (Invitrogen). 50 adults were

used for each extraction. 2 μg of each RNA sample were reverse-transcribed with

primers RP49rt (control) and 3906rt simultaneously, following standard procedures.

Mock reactions (no reverse transcriptase) were carried out to ensure that no

contaminating genomic DNA was present. cDNA (first strand) samples and mock

reactions were amplified using primer pairs RP495-RP493, and 3906rt-3906pcr,

separately. The amplification conditions were: 94°C, 5 min; (94°C, 30 sec; 52°C, 30 sec;

72°C, 30 sec) for 30 cycles; 72°C, 5 min.

2.4.6. Western blots

Western blot analyses were performed according to standard procedure (Lacey

et al., 1994). Embryo extracts were prepared from wild-type and mutant 12-16 hour old

embryos as described below under (ChIP analysis). About 150 μg of each extract were

used for the SU(VAR)3-9 and H3K9me analyses. The blots were probed with a polyclonal

anti- Su(VAR)3-9 antibody (a-3-9^{chr}) (Ner *et al.,* 2002) at 1:2000 dilution and subsequently the blots were reprobed with an anti-tubulin monoclonal antibody at 1:1000. To detect the methylation status of K9H3 we used commercial anti-H3K9me2 (Upstate Biotech #07-212) at 1:1000 and to detect total histone H3 the blots were probed with an anti-H3 monoclonal antibody at 1:30,000 dilution (Sauvé *et al.*, 1999).

2.4.7. Immunostaining of polytene nuclei

Salivary glands of wild-type and *Su(var)3-9* mutant 3rd instar larvae were dissected in PBS, fixed in PBS + 2% formaldehyde for 15 min at room temperature, washed 3 times in PBS^{2+}, and blocked in PBS^{2+} with 1% BSA for 60 min at room temperature (Cryderman *et al.,* 1999). Protease inhibitors were added as required. The anti-H3K9me2 (Upstate Biotech #07-212) and anti-HP1 (C1A9) (James *et al.,* 1989) were added at a final dilution of 1:250 each. The secondary antibodies were anti-rabbit Alexa488 and anti-mouse Alexa568 (Molecular Probes) at 1:1000 each.

2.4.8. ChIP analysis

12-16 hours old embryos were dechorionated in 50% bleach for 2 min and washed extensively with PBS+0.01% Triton-X 100. Cross-linking was achieved by incubation in 2% formaldehyde, 50mM Hepes pH 8.0, 100mM NaCl, 1mM EDTA, 0.5mM EGTA for 10 min at room temperature and then for 20 min at 4°C, and terminated by adding glycine at a final concentration of 250mM. The embryos were then washed twice with 10mM Tris-HCl, pH 8.0, 1mM EDTA, 0.5 mM EGTA and subjected to

sonication (9X15 sec at 30% output). Protease inhibitors were added as required. The soluble fraction of the lysate was adjusted to a final concentration of 3M urea and incubated on ice for 10 min. Nucleoprotein complexes were purified using a polyclonal anti-H3K9me2 (UPSTATE #02-441) or a polyclonal anti-HP1 (a-HP1 (Ner *et al.,* in preparation)) antibodies. Mock reactions (no antibody) were included in each set of experiments. Immunoprecipitated DNA was purified as described by Nelson *et al.,* (2006). Genomic sequences of interest (HIS-C) were detected by PCR using primer pair H3F/H3R. The amplification conditions were: 94°C, 5min; (94°C, 45 sec; 58°C, 30 sec; 72°C, 40 sec) for 26 cycles; 72°C, 5 min.

2.4.9. Expression and purification of active GST-SU(VAR)3-9 fusion proteins

DNA fragments encoding amino acids 182 to 635 of wild-type SU(VAR)3-9, SU(VAR)3-9[311], SU(VAR)3-9[309], SU(VAR)3-9[330], SU(VAR)3- SU(VAR)3-9[318] and SU(VAR)3-9[376], were amplified from genomic DNA of the corresponding strains using primers 39KYLE and 3SET. Each purified amplification product was cloned into the *EcoRV* site of pBluescript KS- and sequenced. An *EcoRI/NotI* fragment from the pBluescript *Su(var)3-9* constructs was then cloned into pGEX 4T-1 resulting in an in-frame construct that produces GST-SU(VAR)3-9 polypeptides. The junctions and specific *Su(var)3-9* mutation of each clone were verified by sequencing.

The pGEX-*Su(var)3-9* constructs were transformed into *BL21(DE3)pLysS*. For each construct, 500 ml of LB containing 100 μg/ml of ampicillin were inoculated with a single colony and incubated overnight in a 37° C shaker. The culture was induced by adding 150 ml of LB supplemented with 600 μg of ampicillin and 650 μl of 1M IPTG,

incubated for 5 hours at 37º C and then processed as described by Frangioni and Neel (1993) with the following modifications. The bacterial pellet was repeatedly frozen and thawed (four times) in liquid nitrogen and a 25º C water bath. The cells were resuspended in 36 ml of STE$^+$ (10mM Tris, pH 8.0, 300 mM NaCl, 1mM EDTA) and lysozyme added to a final concentration of 1 mg/ml. After 20 minutes at room temperature the cells were adjusted to 5 mM DTT and 1.4% N-lauryl-sarcosine. The lysate was placed on ice for 5 min and then sonicated (Sonic 300 dismembrator, 6x45 sec at 40% power). Triton X-100 was added to a final concentration of 3.4%. The sonicate was then incubated for 5 min. at room temperature and spun for 6 min at 12000 rpm (bench top centrifuge) at 4º C. The supernatant was collected and the GST-SU(VAR)3-9 fusions were bound to glutathione-coupled matrix (Pharmacia Biotech) as recommended by the manufacturer. After binding, the matrix was washed 4 times with PBS, 0.1% NP40, 6 times with PBS, 500mM NaCl, 0.1% NP40, and twice with PBS alone. The matrix-bound fusion protein was stored at –80°C in 50% glycerol in PBS.

2.4.10. <u>HMTase assays</u>

The recombinant variant GST- SU(VAR)3-9 polypeptides (1-10 μg) including the wild-type were incubated for 2 hours at room temperature in HMTase buffer (50mM Tris, pH 8.1, 20mM KCl, 10mM MgCl$_2$, 10mM 2-mercaptoethanol, 250mM sucrose (Rea *et al.,* 2000)) with 20 mg of bulk histones (Roche), or 2 μg of recombinant H3 (Upstate Biotech #14-411), or 2 μg of H3 tail peptide (Upstate Biotech #12-357, Upstate Biotech #12-430), and 0.125-0.25 μCi of S-adenosyl-methyl-methionine. The reactions were carried out in a final volume of 50 ml and stopped by adding 10 μl of 6X SDS loading

buffer. 30µl of each reaction was separated on a 13 or 15% polyacrylamide gel and stained with Coomassie. After drying, the gels were exposed and radioactive signal detected using a Phosphor Imager. The data were processed using the Image Quant software. The NIH Image software was used to quantify the amount of the various GST-Su(var)3-9 proteins in each Coomassie-stained band.

The specific activity of each mutant was calculated as the ratio of the radioactive signal corresponding to the (methylated) substrate and relative amount of recombinant GST- Su(var)3-9 used as determined by the intensity of the Coomassie stained band. The relative activity of each mutant, expressed as a percentage, is the ratio between its specific activity and the specific activity of the wild-type recombinant protein (run on the same gel). Each experiment was performed three times using recombinant proteins from independent preparations. The results are expressed as the average of three independent trials +/- the S.E.M. Student's T-tests (p=0.05) were used to determine whether differences between mutants were statistically significant.

Suppl. table 2.1. Enzymatic activity of selected SU(VAR)3-9 mutants on 4

different substrates (see Materials & Methods for details).

	Percent activity (average +/- SEM) on the indicated substrates			
	Bulk histones	Recombinant H3	H3 tail (1-20)	H3 K9me2 tail (1-21)
GST	0[1]	0[1]	0[1]	0[1]
SU(VAR)3-9[wt]	100	100	100	100
SU(VAR)3-9[309/312]	0[1]	0[1]	0[1]	0[1]
SU(VAR)3-9[311]	12.8 +/- 3.1	22.7 +/- 1.2	14.3 +/- 1.3	11.5 +/- 0.2
SU(VAR)3-9[318]	14.3 +/- 1.4	24.7 +/- 1.7	18.3 +/- 0.3	0[1]
SU(VAR)3-9[330]	24.3 +/- 1.9	31.3 +/- 3.8	1.2 +/- 0.2	4.0 +/- 1.1
SU(VAR)3-9[376]	0[1]	0[1]	0[1]	Not tested

[1] Values below 1% are listed as 0

Suppl. table 2.2. Sequences of the primers used in this study.

NAME	SEQUENCE (5' to 3')
3SET	TGTCTCAGGTGGGTAACGGCGTG
5SET	GCCAACGGCAGCGGATGGGGGG
3-95'	CGGGATCCCGAATTCATGGCCACGGCTGAAGCC
39-1	CTGCTGTCGCTGCTGCTTGGAGGT
39-2	CAATACGCTCCACAACGTACTCTC
39KYLE	TTCGCCAAACTGAAGCGTCG
5RI	CGATATCGAGATTTGATGCCG
3RI	TAGGGCACTACGGGGTTTAC
RP49rt	CGCGCTCGATAATCTCC
RP495	GCCCAAGATCGTGAAGAAGC
RP493	CTGTTGTCGATACCCTTGGG
3906rt	TTTTTCGTCAAGCGTTC
3906pcr	ATCCACGGTGGTCAAAG
H3F	GCTCGTACCAAGCAAACT
H3R	TGCCGTGTCAGCTTAAGCA

2.5. References

Aagaard, L., Laible, G., Selenko, P., Schmid, M., Dorn, R., Schotta, G., Kuhfittig, S., Wolf, A., Lebersorger, A., Singh, P.B. *et al*.: Functional mammalian homologues of the *Drosophila* PEV-modifier *Su(var)3-9* encode centromere-associated proteins which complex with the heterochromatin component M31. *Embo J* 18, 1923-1938 (1999)

Akhtar, A., Zink, D., Becker, P.B.: Chromodomains are protein-RNA interaction modules. *Nature* 407, 405-409 (2000)

Ayyanathan, K., Lechner, M.S., Bell, P., Maul, G.G., Schultz, D.C., Yamada, Y., Tanaka, K., Torigoe, K., Rauscher, F.J., 3rd: Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation. *Genes Dev 17*, 1855-1869 (2003)

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

Berger, S.L.: Histone modifications in transcriptional regulation. *Curr Opin Genet Dev* 12, 142-148 (2002)

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

Collins, R.E., Tachibana, M., Tamaru, H., Smith, K.M., Jia, D., Zhang, X., Selker, E.U., Shinkai, Y., Cheng, X.: *In vitro* and *in vivo* analyses of a Phe/Tyr switch controlling product specificity of histone lysine methyltransferases. *J Biol Chem* 280, 5563-5570 (2005)

Cryderman, D.E., Morris, E.J., Biessmann, H., Elgin SC, Wallrath LL.: Silencing at Drosophila telomeres: nuclear organization and chromatin structure play critical roles. *EMBO J* 18, 3724-3735 (1999)

Cryderman, D.E., Grade, S.K., Li**,**Y., Fanti, L., Pimpinelli, S., Wallrath, L.L.: Role of *Drosophila* HP1 in euchromatic gene expression. *Dev Dyn 232,* 767-74 (2005*)*

Czermin, B., Schotta, G., Hulsmann, B.B., Brehm, A., Becker, P.B., Reuter, G., Imhof, A.: Physical and functional association of SU(VAR)3-9 and HDAC1 in *Drosophila*. *EMBO Rep* 2, 915-919 (2001)

Donaldson, K.M., Lui, A., Karpen, G.H.: Modifiers of terminal deficiency-associated position effect variegation in *Drosophila*. *Genetics* 160, 995-1009 (2002)

Ebert, A., Schotta, G., Lein, S., Kubicek, S., Krauss, V., Jenuwein, T., Reuter, G.: *Su(var)* genes regulate the balance between euchromatin and heterochromatin in *Drosophila*. *Genes Dev* 18, 2973-2983 (2004)

Ebert, A., Lein, S., Schotta, G., Reuter, G.: Histone modification and the control of heterochromatin gene silencing in *Drosophila*. *Chromosome Research* 14, 377-392 (2006)

Eskeland, R., Czermin, B., Boeke, J., Bonaldi, T., Regula, J.T., Imhof, A.: The N-terminus of *Drosophila* SU(VAR)3-9 mediates dimerization and regulates its methyltransferase activity. *Biochemistry* 43, 3740-3749 (2004)

Finch, J.T., Rhodes, D., Brown, R.S., Rushton, B., Levitt, M., Klug, A.: Structure of nucleosome core particles of chromatin. *Nature* 269, 29-36 (1977)

Fischle, W., Wang, Y., Allis, C.D.:Histone and chromatin cross-talk. *Curr Opin Cell Biol* 15, 172-183 (2003)

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

Frangioni, J.V., Neel, B.G.: Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. *Anal Biochem* 210, 179-187 (1993)

Gatti, M., Pimpinelli, S.: Functional elements in *Drosophila melanogaster* heterochromatin. *Annu Rev Genet* 26, 239-275 (1992)

Greil, F., van der Kraan, I., Delrow, J., Smothers, J.F., de Wit, E., Bussemaker, H.J., van Driel, R., Henikoff, S., van Steensel, B.: Distinct HP1 and Su(var)3-9 complexes bind to sets of developmentally coexpressed genes depending on chromosomal location. *Genes Dev* 17, 2825-2838 (2003)

Grigliatti, T.: Position-effect variegation--an assay for nonhistone chromosomal proteins and chromatin assembly and modifying factors. *Methods Cell Biol* 35, 587-627 (1991)

Harrington, M.J.: Ph.D. Thesis *The University of British Columbia* (2001)

Ivanova, A.V., Bonaduce, M.J., Ivanov, S.V., Klar, A.J.: The chromo and SET domains of the Clr4 protein are essential for silencing in fission yeast. *Nat Genet* 19, 192-195 (1998)

Jacobs SA, Taverna SD, Zhang Y, Briggs SD, Li J, Eissenberg JC, Allis CD, Khorasanizadeh S (2001) Specificity of the HP1 chromo domain for the methylated N-terminus of histone H3. *Embo J* 20:5232-5241.

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

Kraulis, P.J.: MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *Appl Cryst* 24, 946-950 (1991)

Krauss, V., Reuter, G.: Two genes become one: the genes encoding heterochromatin protein Su(var)3-9 and translation initiation factor subunit eIF-2gamma are joined to a dicistronic unit in holometabolic insects. *Genetics* 156, 1157-67 (2000)

Klug, A., Rhodes, D., Smith, J., Finch, J.T., Thomas, J.O.: A low resolution structure for the histone core of the nucleosome. *Nature* 287, 509-516 (1980)

Krouwels, I.M., Wiesmeijer, K., Abraham, T.E., Molenaar, C., Verwoerd, N.P., Tanke, H.J., Dirks, R.W.: A glue for heterochromatin maintenance: stable SUV39H1 binding to heterochromatin is reinforced by the SET domain. *J Cell Biol* 170, 537-549 (2005)

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

Lewis, E.B.: The phenomenon of position effect. *Adv genet* 3, 73-115 (1950)

Litt, M.D., Simpson, M., Gaszner, M., Allis, C.D., Felsenfeld, G.: Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science* 293, 2453-2455 (2001)

Liu, L.P., Ni, J.Q., Shi, Y.D., Oakeley, E.J., Sun, F.L.: Sex-specific role of *Drosophila melanogaster* HP1 in regulating chromatin structure and gene transcription. *Nat Genet* 37, 1361-1366 (2005)

Lohe, A.R., Hilliker, A.J.: Return of the H-word (heterochromatin). *Curr Opin Genet Dev* 5, 746-755 (1995)

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

Ma, H., Baumann, C.T., Li, H., Strahl, B.D., Rice, R., Jelinek, M.A., Aswad, D.W., Allis, C.D., Hager, G.L., Stallcup, M.R.: Hormone-dependent, CARM1-directed, arginine-specific methylation of histone H3 on a steroid-regulated promoter. *Curr Biol* 11, 1981-1985 (2001)

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

Merritt, E.A., Murphy, M.E.: Raster3D Version 2.0.: A program for photorealistic molecular graphics. *Acta Crystallogr D Biol Crystallogr* 50:869-873 (1994)

Messmer, S., Franke, A., Paro, R.: Analysis of the functional role of the Polycomb chromo domain in *Drosophila melanogaster*. *Genes Dev* 6, 1241-1254 (1992)

Min, J., Zhang, X., Cheng, X., Grewal, S.I., Xu, R.M.: Structure of the SET domain histone lysine methyltransferase Clr4. *Nat Struct Biol* 9, 828-832 (2002)

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

Mis, J., Ner, S.S., Grigliatti, T.A.: Identification of three histone methyltransferases in *Drosophila*: dG9a is a suppressor of PEV and is required for gene silencing. *Mol Gen Genomics* 275, 513-526 (2006)

Muller, H.J.: Types of visible variations induced by X-rays in *Drosophila*. *J Genet* 22, 299-334 (1930)

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

Nelson, J.D., Denisenko, O., Sova, P., Bomsztyk, K.: Fast chromatin immunoprecipitation assay. *Nucleic Acids Res* 34, e2 (2006)

Ner, S.S., Harrington, M.J., Grigliatti, T.A.: A role for the *Drosophila* SU(VAR)3-9 protein in chromatin organization at the histone gene cluster and in suppression of position-effect variegation. *Genetics* 162, 1763-1774 (2002)

Nielsen, S.J., Schneider, R., Bauer, U.M., Bannister, A.J., Morrison, A., O'Carroll, D., Firestein, R., Cleary, M., Jenuwein, T., Herrera, R.E. *et al*.: Rb targets histone H3 methylation and HP1 to promoters. *Nature* 412, 561-565 (2001)

Peterson, C.L., Laniel, M.A.: Histones and histone modifications. *Curr Biol* 14, R546-551 (2004)

Platero, J.S., Hartnett, T., Eissenberg, J.C.: Functional analysis of the chromo domain of HP1. *Embo J* 14, 3977-3986 (1995)

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

Rea, S., Eisenhaber, F., O'Carrol, D., Strahl, B.D., Sun, Z-W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D. *et al.*: Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406, 593-599 (2000)

Reuter, G., Spierer, P.: Position effect variegation and chromatin proteins. *Bioessays* 14, 605-612 (1992)

Reuter, G., Wolff, I.: Isolation of dominant suppressor mutations for position-effect variegation in *Drosophila melanogaster*. *Mol Gen Genet* 182, 516-519 (1982)

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

Schneider, R., Bannister, A.J., Kouzarides, T.: Unsafe SETs: histone lysine methyltransferases and cancer. *Trends Biochem Sci* 27, 396-402 (2002)

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

Schotta, G., Ebert, A., Reuter, G.: SU(VAR)3-9 is a conserved key function in heterochromatic gene silencing. *Genetica* 117, 149-158 (2003)

Schotta, G., Reuter, G.: Controlled expression of tagged proteins in *Drosophila* using a new modular P-element vector system. *Mol Gen Genet* 262, 916-920 (2000)

Seum, C., Reo, E., Peng, H., Rauscher, F.J. III, Spierer, P., Bontron, S.: *Drosophila* SETDB1 Is Required for Chromosome 4 Silencing. *PLoS Genet* 3, *e76l* (2007)

Sinclair, D.A.R., Mottus, R.C., Grigliatti, T.A.: Genes Which Suppress Position-Effect Variegation in *Drosophila melanogaster* are Clustered. *Mol Gen Genet* 191, 326-333 (1983)

Strutt, H., Cavalli, G., Paro, R.: Co-localization of Polycomb protein and GAGA factor on regulatory elements responsible for the maintenance of homeotic gene expression. *Embo J* 16, 3621-3632 (1997)

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

Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G., Reuter, G.: The protein encoded by the *Drosophila* position-effect variegation suppressor *Su(var)3-9* combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J* 13, 3822-3831 (1994)

Weiler, K.S., Wakimoto, B.T.: Heterochromatin and gene expression in *Drosophila*. *Annu Rev Genet* 29, 577-605 (1995)

Xiao, B., Jing, C., Wilson, J.R., Walker, P.A., Vasisht, N., Kelly, G., Howell, S., Taylor, I.A., Blackburn, G.M., Gamblin, S.J.: Structure and catalytic mechanism of the human histone methyltransferase SET7/9. *Nature* 421, 652-656 (2003)

Zhang, X., Tamaru, H., Khan, S.I., Horton, J.R., Keefe, L.J., Selker, E.U., Cheng, X.: Structure of the *Neurospora* SET domain protein DIM-5, a histone H3 lysine methyltransferase. *Cell* 111, 117-127 (2002)

Zhang, X., Yang, Z., Khan, S.I., Horton, J.R., Tamaru, H., Selker, E.U., Cheng, X.: Structural basis for the product specificity of histone lysine methyltransferases. *Mol Cell* 12, 177-185 (2003)

3. THE ROLE OF SU(VAR)3-9 IN THE REGULATION OF *DROSOPHILA*'S *HISTONE GENE* CLUSTER (HIS-C)[2].

3.1. Introduction

In *Drosophila*, methylation of lysine 9 of histone H3 (H3K9me) is a hallmark of heterochromatin (Rea *et al.,* 2000; Nakayama *et al.,* 2001; Peters *et al.,* 2001; Schotta *et al.,* 2002; 2004; Ebert *et al.,* 2004; 2006), and SU(VAR)3-9 is one of the major methyltransferases responsible for this modification (Rea *et al.,* 2000; Schotta *et al.,* 2002; 2003). SU(VAR)3-9 itself is associated with heterochromatic regions of the genome, and particularly with centromeric and pericentric heterochromatin, which in polytene chromosomes form the chromocentre. The formation of heterochromatin is thought to involve several steps, and a relatively detailed model has emerged, which describes the sequence of events and the role of SU(VAR)3-9 in this process (Nakayama *et al.,* 2001; Czermin *et al.,* 2001; Schotta *et al.,* 2003; Swaminathan *et al*., 2005; Rudolph *et al.,* 2007). In this model, chromatin compaction is initiated by the demethylation of H3K4 by the demethylase SU(VAR)3-3/dLSD1, followed by the deacetylation of H3K9 by the histone deacetylase HDAC1/RPD3 (Czermin *et al.,* 2001; Rudolph *et al.,* 2007). HP1 and SU(VAR)3-7, two other NHCPs, are then responsible for targeting and restricting SU(VAR)3-9 to the chromocentre, where it methylates H3K9, thus creating a binding site for the chromodomain of HP1 (Jaquet *et al.,* 2002; Schotta *et al.,* 2002; Delattre *et al.,* 2004; Ebert *et al.,* 2006; Jacobs *et al.,* 2001;

---

[2] A version of this chapter will be submitted for publication. Kalas, P. and Grigliatti, T.A. The role of SU(VAR)3-9 in the regulation of *Drosophila*'s *Histone Gene* cluster.

help of auxiliary factors, HP1 then binds to H3K9me, allowing the recruitment of more SU(VAR)3-9 and other factors, such as the H4K20 HMTase SUV4-20 (Schotta *et al.,* 2004; Eskeland *et al.,* 2007).

However, SU(VAR)3-9 is also detected at a number of euchromatic loci, where it contributes to the regulation of gene expression (Nielsen *et al.,* 2001; Ner *et al.,* 2002; Greil *et al.,* 2003; Koryakov *et al.,* 2006). One of these euchromatic sites is the *histone genes cluster* (HIS-C), where SU(VAR)3-9 appears to affect gene expression by altering the chromatin structure of the locus (Ner *et al.*, 2002). Several pieces of evidence seem to connect the function of SU(VAR)3-9 in euchromatin with silencing or down regulation of the target genes, but its mechanism of action is largely unknown (Nielsen *et al.,* 2001; Ner *et al.*, 2002; Macaluso *et al.,* 2003). Moreover, it is becoming increasingly apparent that SU(VAR)3-9's role and function may vary in a context-dependent manner. For instance, genome-wide localization studies showed that SU(VAR)3-9 is present at a large number of euchromatic loci, and it colocalizes with HP1 only at a subset of these loci; unlike in heterochromatin, where SU(VAR)3-9 and HP1 appear to overlap very broadly (Greil *et al.,* 2003; Schotta *et al.,* 2002). This suggests that SU(VAR)3-9 may have slightly different functions at different loci, and/or that it may elicit its function(s) through different mechanisms (*e.g.* in collaboration with HP1 or with some other non-histone chromatin protein). Particular attention has been devoted to the study of SU(VAR)3-9's catalytic core, the preSET/SET/postSET domain (here referred to as "the SET domain" for simplicity). It has been demonstrated that the integrity of this domain is

necessary not only for the protein to carry out its enzymatic activity, which is believed to play a key role in heterochromatin formation, but also for its association with centric and pericentric heterochromatin (Schotta *et al.,* 2002). FRAP-based studies in mammalian cell lines have also shown that the SET domain of SUV39H1 (the human ortholog of SU(VAR)3-9) contributes to its stable association to (hetero)chromatin, and that this function seems independent from its catalytic activity (Krouwels *et al.,* 2005). Again, it is possible that the relative importance of each of SU(VAR)3-9's multiple functions are context-dependent.

Here, we use the HIS-C as a "model system" for investigating the function(s) of SU(VAR)3-9, and specifically those associated with its SET/preSET/postSET domain, in the regulation of a euchromatic locus. There are several reasons why the HIS-C was chosen. First, SU(VAR)3-9 has been shown to physically associate with the HIS-C, indicating that it probably plays a direct role in the chromatin architecture of this locus. Second, we know that the expression of at least two of the *histone* genes (H1 and H4) is altered in at least three *Su(var)3-9* mutants, suggesting that SU(VAR)3-9 must play a role in their regulation (Ner *et al.,* 2002). In addition, its mechanism of action seems to be chromatin-mediated, since the nucleosome spacing at the HIS-C appears altered in several *Su(var)3-9* mutants (Ner *et al.,* 2002).

The simplest initial working model is that SU(VAR)3-9 acts at the HIS-C through the same mechanism as it does in heterochromatin. In particular, we hypothesize that methylation of H3K9 across the HIS-C is mainly dependent on SU(VAR)3-9,

and that the presence of SU(VAR)3-9 and that of H3K9me2 are necessary for proper localization of HP1 at this locus. In turn, the presence of HP1 at the HIS-C would be necessary for proper regulation of the *histone* genes expression. In order to test this hypothesis, we first need to obtain a reasonably detailed picture of the HIS-C "landscape" in terms of the distribution of SU(VAR)3-9, H3K9me2 and HP1 across the locus in a wild-type strain. Then, we take advantage of a set of well-characterized *Su(var)3-9* mutants to investigate the functional relationships among SU(VAR)3-9, H3K9me2 and HP1 in the context of *histone* gene regulation. Specifically, we ask whether the distribution of SU(VAR)3-9, H3K9me2 and HP1 across the HIS-C, and the level of *histone* transcripts are altered in *Su(var)3-9* missense mutants. We show that, in the three missense mutants analyzed, the *Su(var)3-9* gene product is still present across the HIS-C. The levels of H3K9me2 and HP1 associated with the HIS-C are significantly reduced in all *Su(var)3-9* mutants tested. These mutants also display an increase in the amount of *histone H3* and *histone H2A* transcript levels, supporting the hypothesis that the enzymatic function of SU(VAR)3-9 is critical for regulation of the *histone* genes. Interestingly, in one *Su(var)3-9* missense allele (*Su(var)3-9^{330}*) the relative increase in the level of *H3* transcript is much more pronounced than those of *H2A* and *H2B*, suggesting that the stoechiometry of these core nucleosome proteins may be disrupted in this particular strain. Finally, we show that the elevated levels of *histone* transcripts detected in *Su(var)3-9* mutants are not due to an accumulation of abnormally high levels of *histone* mRNA synthesized outside S-phase. Hence, we conclude that the increased amount of

h2a and *h3* transcripts is most likely a consequence of an increase in the number

of templates transcribed/unit time, or an increase in the rate of transcription from

each of an invariant number of templates, or a combination of these two factors.
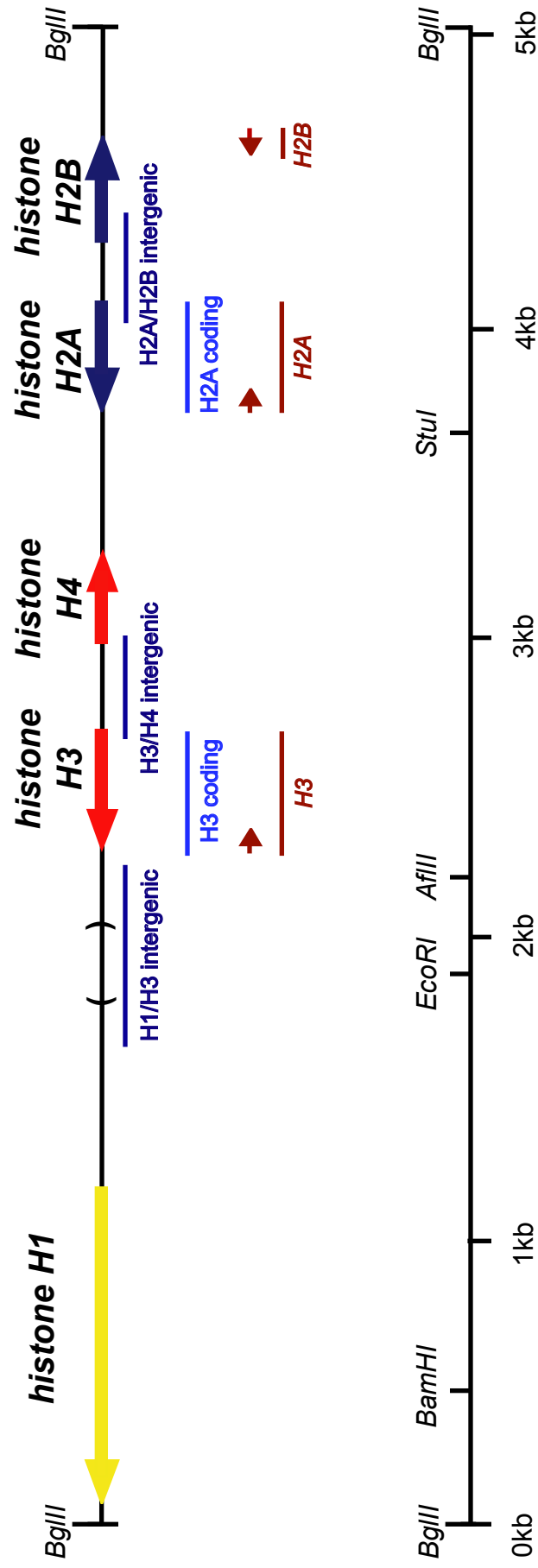
Figure 3.1.

Figure 3.1. A schematic of the *histone* unit. The *histone genes cluster* (HIS-C) is comprised of ~110 tandemly repeated histone units; the diagram shows the *BglII* genomic fragment representing the *histone* unit. The coding regions of histones H1 (yellow), H3 and H4 (red) and H2A and H2B (blue) are shown as plain boxes with an arrowhead pointing in the direction of transcription. The region delimited by the brackets represents the approximate extent and location of the deletion present in about 25% of the histone units. The blue lines labelled "H1/H3 intergenic", "H3/H4 intergenic", H2A/H2B intergenic", "H3 coding" and "H2A coding" indicate the sizes and positions of the fragments analyzed in the ChIP experiments. The dark red arrows show the positions and directions of the primers used to reverse-transcribe the histone RNAs, and the dark red lines labelled "H3", "H2A" and "H2B" indicate the extents of the amplified cDNAs produced. Below, the relative positions of four unique restriction sites are shown in relation to the two *BglII* sites defining the unit.

3.2. Results

3.2.1. SU(VAR)3-9 is associated with the HIS-C in wild-type and *Su(var)3-9* missense mutants

The *Histone Gene Cluster* (HIS-C) region of the second chromosome is comprised of about 100 tandemly reiterated copies of the *histone* unit (*his* unit). The *his* unit contains one copy of each of the five *histone* genes, with the core histone genes arranged into two gene pairs, the *h3/h4* and the *h2a/h2b* couplets (Figure 3.1). The two members of each couplet share a regulatory region. There are two versions of the *his* unit: one, representing about 75% of the *his* units present in the HIS-C, is 5kb, and the other, less represented, is about 4.75kb in length (Lifton *et al*., 1977). The 250 bp difference is due to an indel located in the H1/H3 intergenic region (see parentheses in Figure 3.1.).

SU(VAR)3-9 is associated with the HIS-C, where it is detected across the whole locus (Ner *et al.,* 2002; Koryakov *et al.,* 2006; Ner *et al.,* in preparation). In addition, several tested *Su(var)3-9* mutants, including the missense mutant *Su(var)3-9^{330}* display an alteration in the chromatin structure of the HIS-C and elevated levels of *histone* gene transcripts (Harrington, 2001; Ner *et al.,* 2002). However, whether the physical presence of SU(VAR)3-9 is necessary and/or sufficient for proper regulation of the *histone* genes, and whether the single amino acid substitutions in its catalytic region affect its localization at the HIS-C, is unclear.

In order to address these issues, we first used chromatin immunoprecipitation (ChIP) to determine the distribution of SU(VAR)3-9 at the HIS-C in wild-type and

Su(var)3-9 missense mutant embryos. Homozygous *Su(var)3-9^{06}* individuals, which represent complete nulls producing no *Su(var)3-9* mRNA or protein (Tschiersch *et al.,* 1994; Schotta *et al.,* 2002; this work, chapter 2) were also included in the analysis to serve as an internal standard/negative control. ChIP analysis of cross-linked extracts prepared from 12-16 hour old staged embryos was performed using an antibody raised against the N-terminal half of SU(VAR)3-9 (α-SU(VAR)3-9^{chr} (Ner *et al.,* 2002)), and the immunoprecipitated DNA was tested for the presence of 3 sequences belonging to the HIS-C: the *H3* coding region ("H3 coding"), a fragment including the intergenic/regulatory region between the *H2A* and *H2B* genes ("H2A/H2B intergenic") and a section of the intergenic region between the *H1* and *H3* genes ("H1/H3 intergenic") (see also Figure 3.1). This last fragment spans a region of the HIS-C that, in ~25% of the 110 or so copies of the *his* unit, contains a ~250bp indel (Lifton *et al.,* 1977). Thus, amplification with H1/H3 intergenic-specific primers gives rise to two different fragments: one representing the "longer" version of the *his* unit (5 kb) and one corresponding to the "shorter" one (4.75 kb). The sizes of these two fragments are roughly 600 bp and 350bp, respectively (Lifton *et al.,* 1977; Samal *et al.,* 1981; Worcel *et al.,* 1983).

As expected, the cross-linked material pulled down by the α-SU(VAR)3-9^{chr} antibody from *Su(var)3-9$^+$/ Su(var)3-9$^+$* extracts contains each of the three HIS-C sequences ("H3 coding", "H2 intergenic" and "H3/H1 intergenic"). In contrast, these fragments were not enriched in the material pulled down from *Su(var)3-9^{06}* extracts (Figure 3.2.b), allowing us to conclude that the enrichment observed in

the wild-type strain is indeed due to the association of SU(VAR)3-9 with the HIS-C. In all the *Su(var)3-9* missense mutants, fragments corresponding to the "H3 coding", "H2A/H2B intergenic" and "H1/H3 intergenic" regions were detected among the immunoprecipitated material (Figure 3.2). Accurate quantifications of one of the *his* unit fragments, "H3 coding", revealed that the relative enrichment obtained with α-SU(VAR)3-9^{chr} is relatively small, although significant (Figure 3.2.c and appendix 2). Still, the data obtained allow us to conclude that that wild-type SU(VAR)3-9 protein (SU(VAR)3-9^{WT}) is associated with the *his* unit. These data, together with those of Ner and colleagues suggest that SU(VAR)3-9^{WT} is distributed throughout the *his* unit (Ner *et al.*, 2002; Ner *et al.*, in preparation). A similar set of ChIP analyses, performed on *Su(var)3-9^{309}*, *Su(var)3-9^{330}* and *Su(var)3-9^{318}* 12-16 hours old embryos, demonstrated that the SU(VAR)3-9^{309}, SU(VAR)3-9^{330} and SU(VAR)3-9^{318} mutant proteins also associate with the *his* unit. The enrichment for the "H3 coding" fragment detected in the three missense mutants *Su(var)3-9^{309}*, *Su(var)3-9^{330}* and *Su(var)3-9^{318}* was not significantly different from that of the wild-type strain (Figure 3.2.c and appendix 2). Hence, our results also suggest that the single amino acid substitutions present in the SU(VAR)3-9^{309}, SU(VAR)3-9^{330} and SU(VAR)3-9^{318} variants do not prevent their association with the HIS-C.

a.

histone H1 — H1/H3 intergenic — H3 coding — histone H3 — histone H4 — histone H2A — histone H2B — H2A/H2B intergenic

b.

H3 coding | **H2A/H2B intergenic** | **H1/H3 intergenic**

IN (1:100), NO Ab, IP (3-9)

wild-type
Su(var)3-9^06
Su(var)3-9^309
Su(var)3-9^330
Su(var)3-9^318

1 2 3 4 5 6 7 8 9

c.

Relative amounts of "H3 coding" fragment in the material immunoprecipitated from 12-16 hrs old embryos.

% of the input — WT, 06, 309, 330, 318

IP (SU(VAR)3-9)
mock (no Ab)

d.

Average amounts of IPed target (relative to the wild-type)

ratio to WT — WT, 06, 309, 330, 318

Figure 3.2

127

Figure 3.2. Relative levels of SU(VAR)3-9 associated with three regions of the *histone unit* in wild type and *Su(var)3-9* mutant embryos. a) Schematic of the *BglII* fragment defining the *histone unit*. The five *histone* genes (H1, H3, H4, H2A and H2B) and the three regions analysed (H3 coding, H2A/H2B intergenic and H1/H3 intergenic) are shown. **b)** Chromatin immunoprecipitation of wild-type and *Su(var)3-9* mutant embryo extracts. The recovered DNA was PCR amplified using primers specific for the "H3 coding" region (lanes 1-3), the "H2A/H2B intergenic region" (lanes 4-6) or the "H1/H3 intergenic" region (lanes 7-9). The template used in each PCR reaction is indicated above the corresponding lane; IN (1:100): input material diluted 100X, NO Ab: mock (no antibody) reaction, IP (α-3-9^{chr}): immunoprecipitated material. The antibody used was raised against the N-terminal region (including the chromodomain) of SU(VAR)3-9 and has been previously described (Ner *et al.,* 2002). **c)** Relative amounts of "H3 coding" fragment in the precipitated material as determined by real-time PCR (see materials and methods for details). The bar graphs represent the average of 3 independent experiments and the error bars span two standard deviations. **d)** Average amounts of immunoprecipitated "H3 coding" fragment (minus the average for the respective "mock" reaction) expressed as fractions of the wild type.

3.2.2. The level of H3K9me2 associated with the HIS-C is significantly reduced in Su(var)3-9 missense mutants

H3K9me2 is enriched at the HIS-C (Figure 3.3 and Ner *et al.,* 2002) and, since *Su(var)3-9*[309], *Su(var)3-9*[330] and *Su(var)3-9*[318] have all been shown to produce SU(VAR)3-9 proteins with an impaired HMTase activity (chapter 2), we asked whether the levels and/or distribution of H3K9me2 across the HIS-C was altered in these *Su(var)3-9* mutants.

ChIP analyses of cross-linked embryo extracts from homozygous *Su(var)3-9*[+], *Su(var)3-9*[06], and the three *Su(var)3-9* missense mutant strains were performed using an antibody that specifically recognizes H3K9me2 (see material and methods for details). As shown previously (Ner *et al*., 2002), all three HIS-C fragments, "H3 coding", "H2A/H2B intergenic" and "H1/H3 intergenic" were detected in the material immunoprecipitated from the wild-type strain (Figure 3.3.a). Their enrichment was about 100-fold higher than that observed using non-specific IgG (Supplementary Figure 1 and appendices 1 and 3), allowing us to conclude that H3K9me2 is indeed present at the HIS-C in the wild-type strain. As expected, the three HIS-C fragments analyzed were significantly less abundant in the material immunoprecipitated from *Su(var)3-9*[06] mutant extracts (Figure 3.3). The *Su(var)3-9*[06] mutant completely lacks the SU(VAR)3-9 protein. However, it is not surprising that a small amount of H3K9me2 associated with its HIS-C in this strain since other HMTases, capable of methylating H3K9, are present in the nucleus (see discussion). The target *his* unit fragments were also detected in the material immunoprecipitated from each of the three *Su(var)3-9*

missense mutants, but, as expected, their relative abundance was significantly lower than in the wild-type strain (Figure 3.3).

The relative enrichment for one particular fragment, "H3 coding", was accurately quantified by real-time PCR. With regard to this fragment, *Su(var)3-9^{06}* and the missense mutant *Su(var)3-9^{330}* displayed the lowest level of enrichment, less than 5% of the wild-type (Figures 3.3.c and d). *Su(var)3-9^{309}* and *Su(var)3-9^{318}* showed a slightly higher enrichment for this fragment, corresponding to about 10% and 13% of the wild-type, respectively (Figure 3.3 and appendix 3). We conclude that the levels of H3K9me2 associated with the three regions of the HIS-C, "H3 coding", "H2A/H2B intergenic" and "H3/H1 intergenic" are significantly lower in all *Su(var)3-9* mutants tested than they are in the wild-type strain.

a.

histone H1 — histone H3 — histone H4 — histone H2A — histone H2B

H1/H3 intergenic   H3 coding   H2A/H2B intergenic

b.

H3 coding | H2A/H2B intergenic | H1/H3 intergenic

IN (1:10) | NO Ab | IP (K9me2)

wild-type
Su(var)3-9^06
Su(var)3-9^309
Su(var)3-9^330
Su(var)3-9^318

1   2   3   4   5   6   7   8   9

c.

Relative amounts of "H3 coding" fragment in the material immunoprecipitated from 12-16 hrs old embryos.

% of the input

IP (H3K9me2)
mock (no Ab)

WT   06   309   330   318

d.

Average amounts of IPed target (relative to the wild-type)

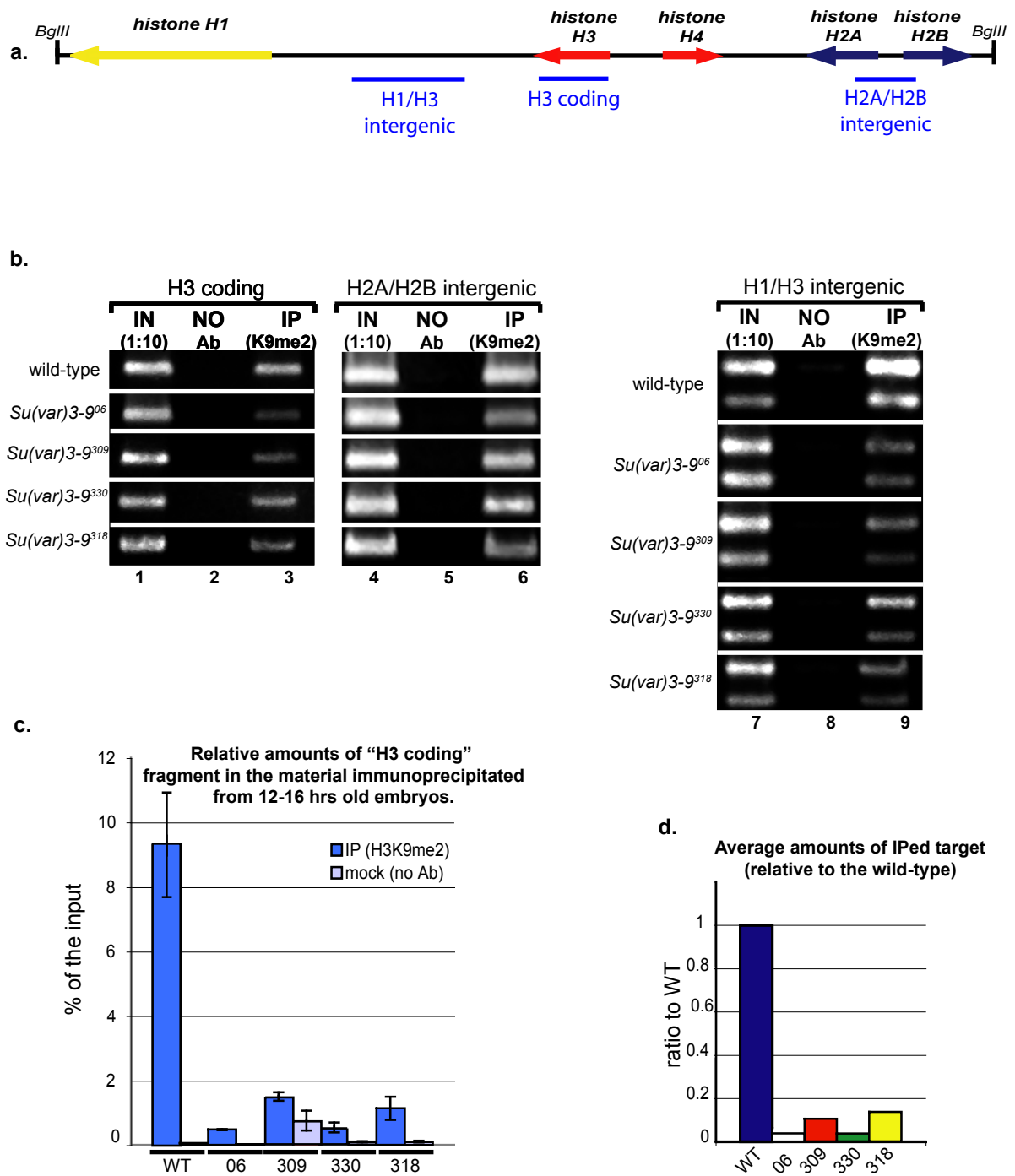ratio to WT

WT   06   309   330   318

Figure 3.3

131

Figure 3.3. Relative levels of H3K9me2 associated with three regions of the

histone unit **in wild type and** *Su(var)3-9* **mutant embryos. a)** Schematic of the

BglII fragment defining the *histone unit*. The five histone genes (H1, H3, H4, H2A

and H2B) and the three regions analysed (H3 coding, H2A/H2B intergenic and

H1/H3 intergenic) are shown. **b)** Chromatin immunoprecipitation of wild-type and

Su(var)3-9 mutant embryo extracts. The recovered DNA was PCR amplified

using primers specific for the "H3 coding" region (lanes 1-3), the "H2A/H2B

intergenic region" (lanes 4-6) or the "H1/H3 intergenic" region (lanes 7-9). The

template used in each PCR reaction is indicated above the corresponding lane;

IN (1:10): input material diluted 10X, NO Ab: mock (no antibody) reaction, IP

(K9me2): immunoprecipitated material. The antibody used was an anti-H3K9me2

from UPSTATE (#07-441). **c)** Relative amounts of "H3 coding" fragment in the

precipitated material as determined by real-time PCR (see materials and

methods for details). The bar graphs represent the average of 3 independent

experiments and the error bars span two standard deviations. **d)** Average

amounts of immunoprecipitated "H3 coding" fragment (minus the average for the

respective "mock" reaction) expressed as fractions of the wild type.

3.2.3. Association of HP1 with the HIS-C

In heterochromatin, H3K9me is thought to represent a binding platform for HP1 (Lachner *et al.,* 2001; Jacobs *et al.,* 2002; Bannister *et al.,* 2001; Nielsen *et al.,* 2002). In addition, SU(VAR)3-9 is known to interact with HP1 physically and genetically (Schotta *et al.,* 2002; 2003; our lab, unpublished data). Thus, not surprisingly, most models postulate an interaction among SU(VAR)3-9, HP1 and H3K9me as a central step in the formation of heterochromatin (Nakayama *et al.,* 2001; Schotta *et al.,* 2002; 2003; Ebert *et al.,* 2006; Rudolph *et al.,* 2007), and SU(VAR)3-9 and HP1 have been suggested to work together in the regulation of a subset of genes (Nielsen *et al.,* 2001; Greil *et al.,* 2003). Since HP1 has been detected at the HIS-C (Greil *et al.,* 2003; Koryakov *et al.,* 2006; chapter 2 of this work), we hypothesized that it may play a role in regulating the expression of the *histone* genes, as part of a SU(VAR)3-9-dependent mechanism. To test this hypothesis, we performed another set of ChIP analyses, in this case, with an anti-HP1 antibody (α–HP1 (Ner *et al.,* in preparation)).

The material immunoprecipitated with α–HP1 from wild-type embryo extracts contained significant, although not copious amounts, of *his* unit fragments (Figure 3.4 and appendix 4). This enrichment for *his* unit fragments was significantly higher than that obtained with pre-immune serum (Supplementary Figure 3.1 and Appendix 1), allowing us to conclude that HP1 is associated with the HIS-C in wild-type strains. In contrast, the material immunoprecipitated with α–HP1 from *Su(var)3-9* mutant extracts displayed very low levels of enrichment for all HIS-C fragments analyzed (Figure 3.4.b). For the "H3 coding" and "H2A coding"

133

fragments, relative quantifications were performed by real-time PCR, and the

Su(var)3-9 mutants showed a ~3 to 19-fold, and ~4 to 13-fold reduction in the

enrichment for HP1, respectively, relative to the wild-type.

We conclude that overall, in *Su(var)3-9* mutants, the levels of HP1 associated

with the HIS-C are significantly lower than in wild-type individuals.

a.

histone H1 · BglII · histone H3 · histone H4 · histone H2A · histone H2B · BglII · H1/H3 intergenic · H3 coding · H2A coding · H2A/H2B intergenic

b.

H3 coding — H2A/H2B intergenic — H1/H3 intergenic

IN (1:100) · NO Ab · IP (HP1)

wild-type, Su(var)3-9[06], Su(var)3-9[309], Su(var)3-9[330], Su(var)3-9[318]

1 2 3 4 5 6 7 8 9

c.

Relative amounts of "H2A coding" fragment in the material immunoprecipitated from 12-16 hrs old embryos.

% of the input — IP (HP1), mock (no Ab)

WT 06 309 330 318

d.

Relative amounts of "H3 coding" fragment in the material immunoprecipitated from 12-16 hrs old embryos.

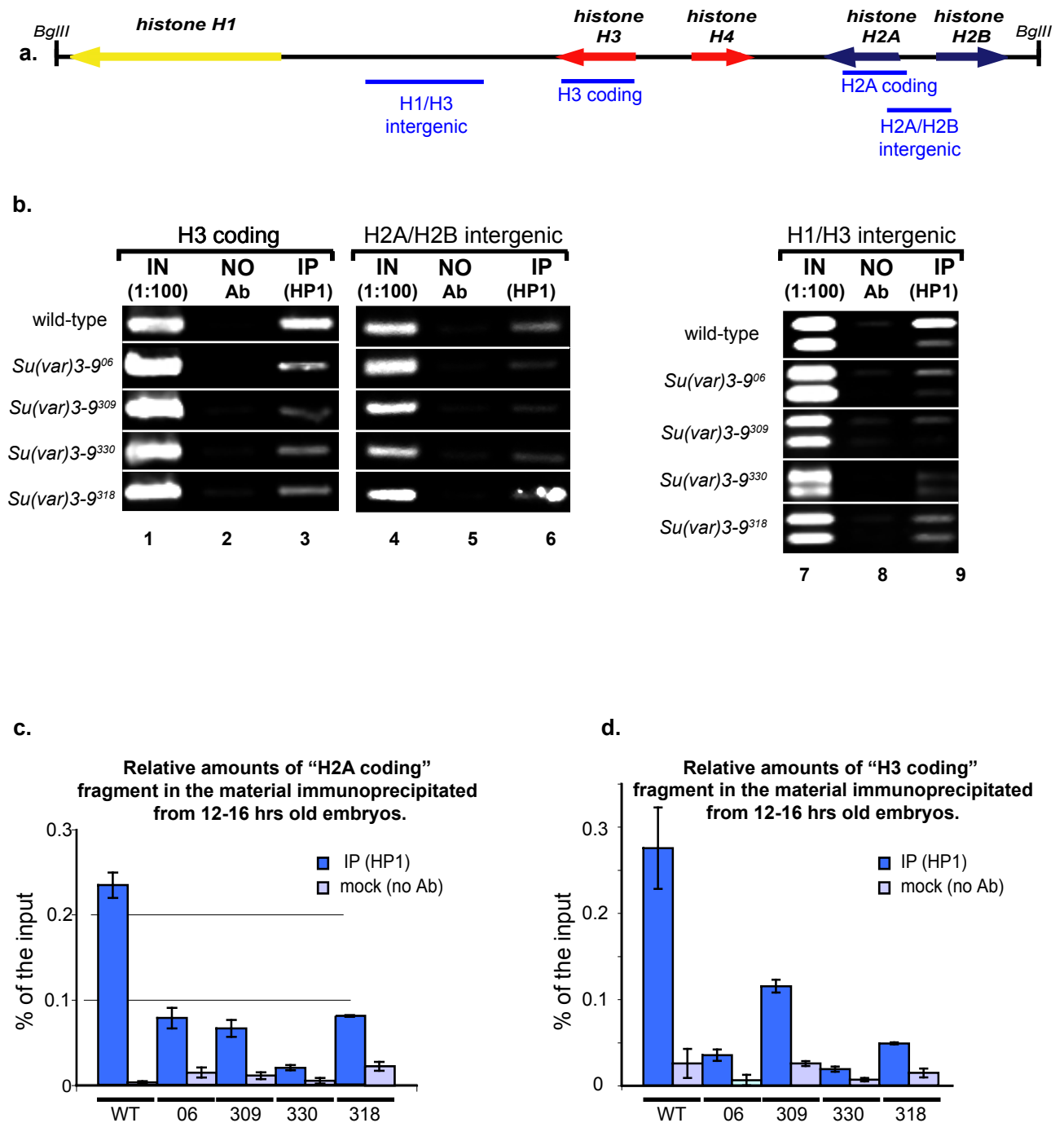% of the input — IP (HP1), mock (no Ab)

WT 06 309 330 318

Figure 3.4

135

Figure 3.4. Relative levels of HP1 associated with four regions of the

histone unit **in wild-type and** *Su(var)3-9* **mutant embryos. a)** Schematic of the

BglII fragment defining the *histone unit*. The five histone genes (H1, H3, H4, H2A

and H2B) and the four regions analysed (H3 coding, H2A/H2B intergenic, H2A

coding and H1/H3 intergenic) are shown. **b)** Chromatin immunoprecipitation of

wild-type and *Su(var)3-9* mutant embryo extracts. The recovered DNA was PCR

amplified using primers specific for the "H3 coding" region (lanes 1-3), the

"H2A/H2B intergenic region" (lanes 4-6) or the "H1/H3 intergenic" region (lanes

7-9). The material used as template is indicated above each lane; IN (1:100):

input material diluted 100X, NO Ab: mock (no antibody) reaction, IP (HP1):

immunoprecipitated material. See material and methods for details about the

antibody. **c)** and **d)** Relative amounts of "H2A coding" and "H3 coding"

fragments, respectively, in the precipitated material as determined by real-time

PCR (see materials and methods for details). The bar graphs represent the

average of 3 independent experiments and the error bars span two standard

deviations.

3.2.4. The null, as well as the missense *Su(var)3-9* mutants have elevated levels of *H2A* and *H3* transcripts

It was previously reported that the levels of *H1* and *H4* transcripts are elevated in three *Su(var)3-9* mutants, including the missense mutant *Su(var)3-9[330]* (Ner *et al.,* 2002). To determine whether this is a common feature of *Su(var)3-9* mutants, and its relationship to the abundance of H3K9me2, SU(VAR)3-9 and HP1 associated with the HIS-C, respectively, we quantified the relative amounts of two *histone* transcripts present in wild-type and several *Su(var)3-9* mutants. We used 12-16 hour old embryos as our source of RNA, for two reasons. Firstly, at this stage of embryogenesis a significant proportion of cells are still going through the cell cycle (reviewed by Lee and Orr-Weaver, 2003) and therefore synthesizing copious amounts of *histone* mRNAs. Thus, if SU(VAR)3-9 participates in the regulation of the *histone* genes, the effect of *Su(var)3-9* mutations on the levels of *histone* should be more pronounced at this stage of development than in adults. Secondly, we wanted to be able to relate the alterations (or lack thereof) in the relative levels of *histone* mRNA to the results of the ChIP analyses, which were performed on 12-16 hour old embryo extracts. For each strain (*Su(var)3-9[+]*, the various *Su(var)3-9* missense mutants and *Su(var)3-9[06]*), the relative levels of two *histone* genes transcripts, *H2A* and *H3*, were quantified by real time RT-PCR using *rp49* as an internal standard (see materials and methods). Within the histone unit, *H2A* and *H2B* are transcribed in opposite directions and they share a promoter region, and the same is true for *H3* and *H4* (Figure 3.1). *H2A* and *H3* were chosen for our analysis as
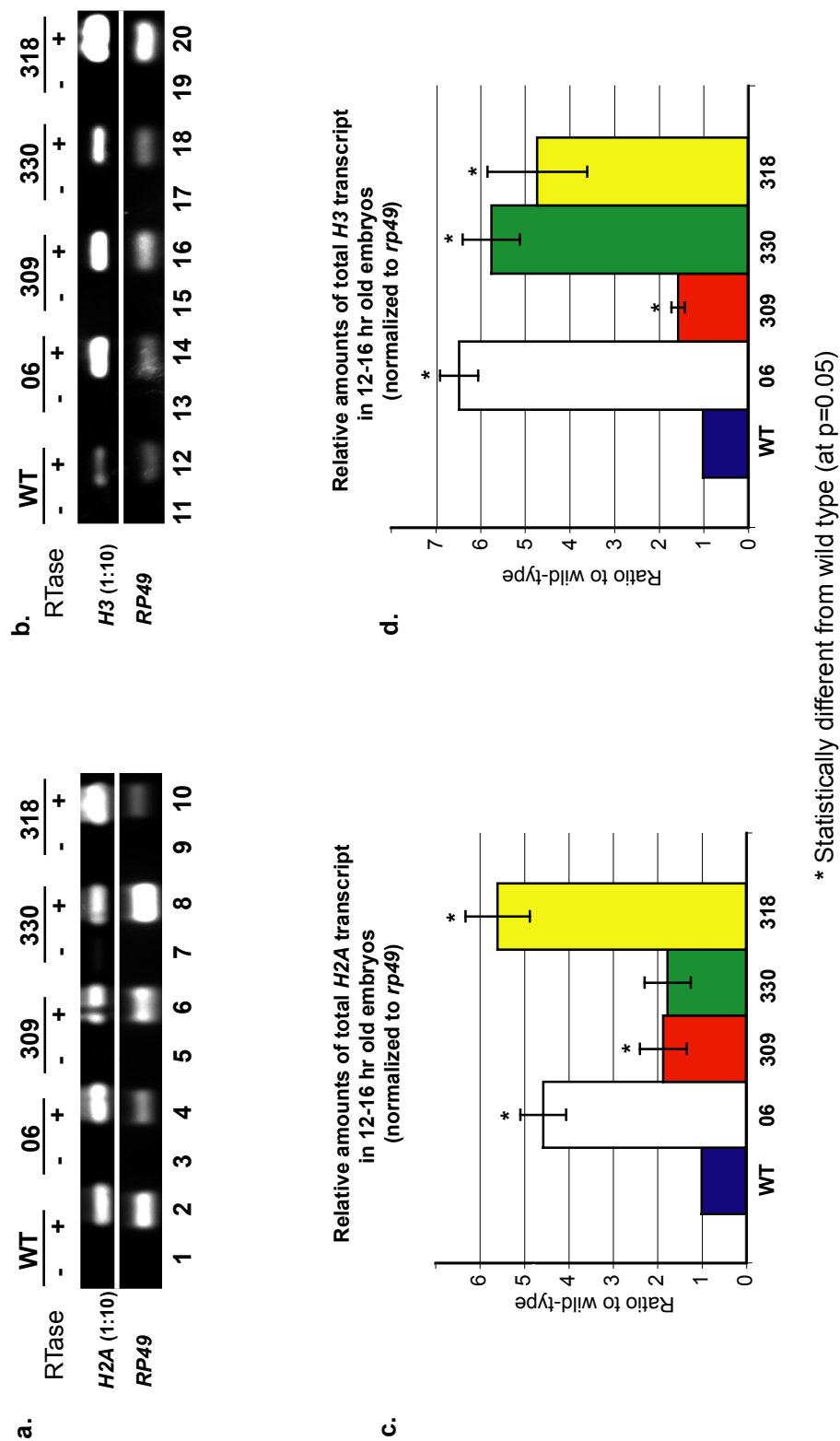
representative of each of the two "gene pairs", as we assumed that the two members of each pair would be co-regulated.

The relative amount of *H3* transcript detected in all *Su(var)3-9* mutants was significantly higher than in the *Su(var)3-9⁺* strain (Figure 3.5.b and 3.5.d, and Appendix 5). *Su(var)3-9⁰⁶*, *Su(var)3-9³³⁰*, and *Su(var)3-9³¹⁸* showed, on average a 6.46, 6.07 and 4.72 fold increase over the wild-type (*n*=3). For *Su(var)3-9³⁰⁹* the increase was less substantial (1.56 fold relative to the wild-type), but still statistically significant (*p*=0.05).

For *Su(var)3-9³⁰⁹* and *Su(var)3-9³¹⁸*, the relative increase in *H2A* was 1.86 and 5.60, respectively, which is comparable (not statistically different at p=0.05) to that observed for *H3*. In *Su(var)3-9⁰⁶*, the increase in *H2A* was only about 70% of that observed for *H3* (see Appendix 6), but still significantly higher than what was observed both in the wild-type and in *Su(var)3-9³⁰⁹* (Figure 3.5. and Appendix 6). *Su(var)3-9³³⁰* was an exception in that the relative increase (over the wild-type) in its level of *H2A* transcript was much lower (<30%) than that observed for its *H3* transcript. In fact, the relative abundance of *H2A* detected in *Su(var)3-9³³⁰* was so low as not to be statistically different from that observed in the wild-type (Appendix 5).

Despite the exceptional case of the *H2A* transcript in *Su(var)3-9³³⁰*, we conclude that elevated levels of *histone* transcripts are probably a common feature of those *Su(var)3-9* mutations that suppress PEV, although the magnitude of the increase seems to be, at least in part, allele-dependent (see section 3.2.5 and discussion). In addition, the data obtained with *Su(var)3-9⁰⁶*, and especially with

Su(var)3-9³³⁰, suggest that the regulation of the *H2A/H2B* and the *H3/H4* gene pairs may be uncoupled.

a.

b.

c. Relative amounts of total *H2A* transcript in 12-16 hr old embryos (normalized to *rp49*)

d. Relative amounts of total *H3* transcript in 12-16 hr old embryos (normalized to *rp49*)

* Statistically different from wild type (at p=0.05)

Figure 3.5

140

Figure 3.5. Relative quantifications of total *histone H3* and *histone H2A* transcript in wild-type and *Su(var)3-9* mutant embryos. a) and **b)** End point RT-PCR reactions on total RNA extracted from 12-16 hours old wild-type (lanes 1,2, 11 and 12) or *Su(var)3-9* mutant embryos (lanes 3-10 and 13-20). The transcripts amplified were *H2A* and *H3*, respectively, as well as *RP49* as an internal standard. In each case, the allele number is indicated above the corresponding lanes. "-" signs denote mock reactions (no reverse transcriptase). **c)** and **d)** Relative quantifications of *H2A* and *H3* transcripts by real-time PCR. For each reaction, the ratio of *H2A* or *H3* between mutants and wild-type (standardized for *RP49*) was reported. The histograms represent the average of three independent reactions and the error bars span two standard deviations. See material and methods for additional details.

3.2.5. <u>*Histone* genes expression in mutant *Su(var)3-9³³⁰*</u>

Since the four core histones are required in equal amounts, and SU(VAR)3-9 is associated with both the *H2A/H2B* and the *H3/H4* gene pairs, one would expect each *Su(var)3-9* mutation to affect the expression of the *core histone* genes to the same degree. This was indeed the case for *Su(var)3-9³⁰⁹* and *Su(var)3-9³¹⁸*, but it was not the case in the *Su(var)3-9⁰⁶* and *Su(var)3-9³³⁰* strains. In the *Su(var)3-9⁰⁶*, the "protein null" strain, there was, on average, a 4.65 and 6.46 fold increase in *H2A* and *H3*, respectively. Statistically ($p=0.05$), the difference between these two values is only marginally significant. The difference between the increase in *H2A* and *H3* was much more pronounced in the *Su(var)3-9³³⁰* strain (1.76 fold and 6.07 fold, respectively). Since in *Su(var)3-9³³⁰* there is such a dramatic difference in the increase of *H2A versus H3* mRNAs, we decided to focus on this strain, and to ask whether the regulation of the *histone* genes can be misregulated in such a way that the expression of each of the four core histones, or each of the two usually co-regulated pairs, *H2A*/*H2B* and H3/*H4* is decoupled. To examine this possibility, we measured the total *H2B* transcript present in embryo extracts from wild-type and *Su(var)3-9³³⁰*.

We found that, as it is the case for *H2A*, the relative levels of *H2B* detected in *Su(var)3-9³³⁰* embryo extracts are, on average, slightly less than 2 fold higher than in the wild-type (Figure 3.6). Similarly to what was observed for *H2A*, such difference is not statistically significant (at $p=0.05$), although it is probably significant biologically. We conclude that somehow, the amino acid substitution present in SU(VAR)3-9³³⁰ affects the relative abundance of *H3* transcript, but not

that of *H2A* or *H2B*. This suggests that the *H2A/H2B* and the *H3/H4* gene pairs may be independently regulated.

To examine this further, we next asked whether, in *Su(var)3-9^{330}*, the intergenic regions between the *H3*/*H4* and *H2A*/*H2B* pairs show any differences in the level of H3K9me2 versus the wild-type. If H3K9me2 is responsible for the regulation of the *histone* genes, we would expect to see a substantial difference between the wild-type and *Su(var)3-9^{330}*, with respect to such modification, within the *H3*/*H4* region, and a less dramatic difference within the *H2A/H2B* region. A ChIP analysis was performed, and relative quantifications of the fragments of interest were carried out by real-time PCR. The results showed that the material immunoprecipitated from both the wild-type and the mutant extracts is enriched for "H3/H4 intergenic" and "H2A/H2B intergenic" fragments (Figure 3.7). However, as for all other regions tested, this enrichment is significantly higher in the wild-type than in *Su(var)3-9^{330}* (Figure 3.7. and Appendix 3). Intriguingly, it was the "H2A/H2B intergenic" region that showed the most substantial difference between *Su(var)3-9^{330}* and the wild-type (Figure 3.7. and Appendix 3), suggesting that there is no direct correlation between the levels of *H2A*, *H2B* and *H3* transcripts and the amount of H3K9me2 associated with the genes' regulatory regions.
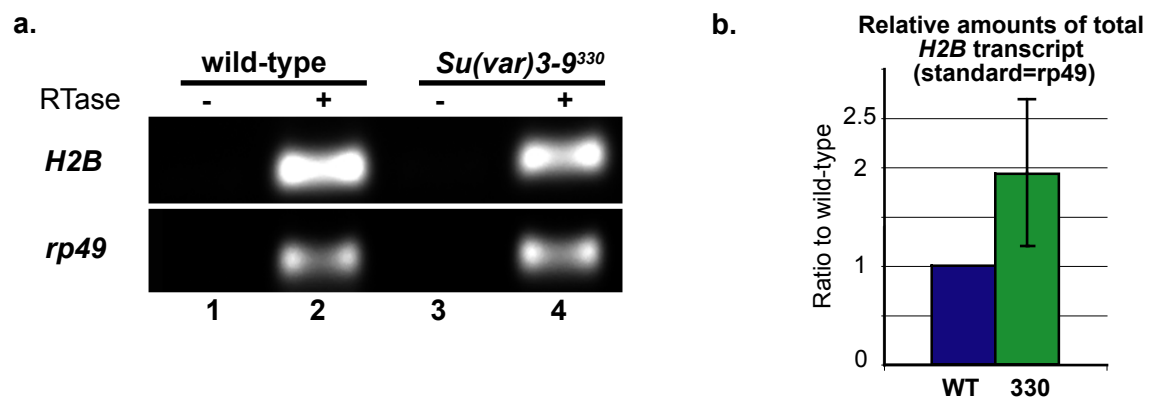
a.

wild-type   *Su(var)3-9³³⁰*

RTase   −   +   −   +

H2B

rp49

1   2   3   4

b.

Relative amounts of total
***H2B* transcript**
(standard=rp49)

Figure 3.6

Figure 3.6. Comparison of H2B transcript present in wild-type *vs. Su(var)3-9^{330}* mutant embryos. a) and **b)** Relative abundance of total *H2B* transcript in 12-16 hours embryos, quantified by real-time RT-PCR following the same procedure as in Figure 3.5. "-" signs denote mock reactions (no reverse transcriptase).
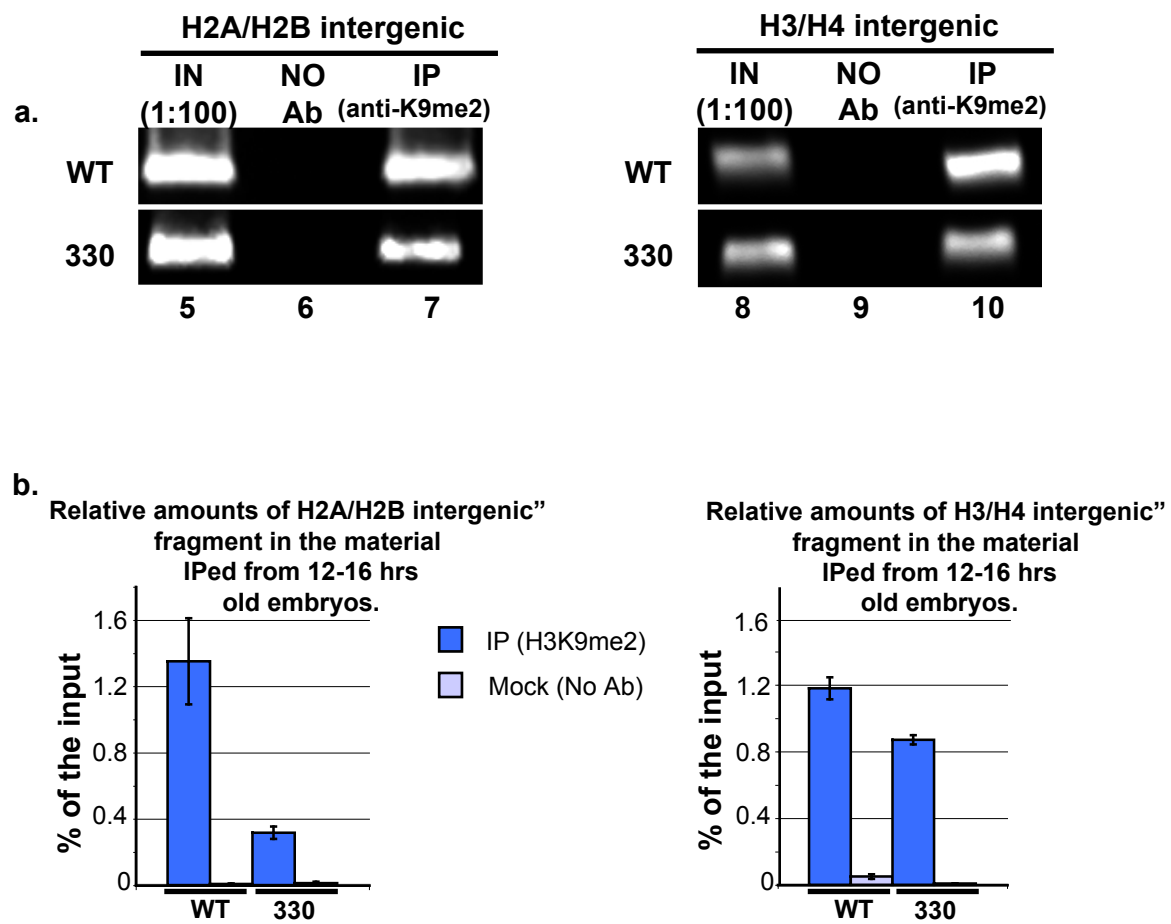
a.

H2A/H2B intergenic

	IN (1:100)	NO Ab	IP (anti-K9me2)
WT			
330			
	5	6	7

H3/H4 intergenic

	IN (1:100)	NO Ab	IP (anti-K9me2)
WT			
330			
	8	9	10

b.

Relative amounts of H2A/H2B intergenic" fragment in the material IPed from 12-16 hrs old embryos.

Relative amounts of H3/H4 intergenic" fragment in the material IPed from 12-16 hrs old embryos.

- IP (H3K9me2)
- Mock (No Ab)

% of the input

Figure 3.7

Figure 3.7. Relative levels of H3K9me2 associated with two intergenic regions of the *histone unit* in wild type and *Su(var)3-9[330]* mutant embryos.

a) Detection (end point PCR) and **b)** quantification (real-time PCR) of the "H2A/H2B intergenic" and "H3/H4 intergenic" fragments in the material immunoprecipitated with anti-H3K9me2 from wild-type and *Su(var)3-9[330]* cross-linked embryonic extracts.

In all cases the histograms represent the average of three independent experiments and the error bars span two standard deviations. See material and methods for additional details.

3.2.6. <u>The elevated levels of *histone* transcripts in *Su(var)3-9* mutants are not due to an increase in polyadenylated transcripts.</u>

The observed overproduction of *histone H2A* and *histone H3* transcripts is an interesting phenotype, and its mechanism is unknown. There are three simple possibilities. First, in the *Su(var)3-9* mutant strains there may be an increase in the number of *histone* templates transcribed at any point in time with respect to the wild-type, while the transcription rate remains constant. Second, *Su(var)3-9* may result in an increase in the transcription rate of the *histone* genes, while the number of *his* templates that are transcribed within the HIS-C region remains constant. Finally, it is possible that neither the number of transcribed templates, nor the transcription rate are affected, but, in *Su(var)3-9* mutants, expression of the *histone* genes may be uncoupled from the cell cycle (*i.e.* the *histone* genes may be transcribed outside, as well as during, S-phase).

The latter hypothesis is easily tested. *Histone* transcripts synthesized during S-phase are not polyadenylated, while those synthesized outside S-phase are (Akhmanova *et al.,* 1997). Therefore, *histone* transcripts produced outside S-phase (decoupled from DNA synthesis) can easily be detected and quantified using an appropriate primer (oligodT) for the reverse-transcription (RT) step. Taking this approach, we first measured the amount of polyadenylated *H2A* and *H3* mRNAs in *Su(var)3-9*[+]/ *Su(var)3-9*[+] embryos, in order to determine what proportion of the total *histone* transcripts are polyadenylated RNA. We found that the relative ratio of polyadenylated/total *H2A* and *H3* was roughly 1/1000 (Figures 3.8.a and 3.8.b).

Hence, if the increase in the level of total *H2A* and *H3* transcripts observed in the mutants is due, even only in part, to transcription outside S-phase, the *Su(var)3-9* mutants should show a very prominent increase in the amount of polyadenylated *H2A* and *H3* transcripts compared to the wild-type. This is not what we observed. Instead, we detected a small reduction in the levels of polyadenylated *H2A* and *H3* in *Su(var)3-9^{06}*, *Su(var)3-9^{309}* and *Su(var)3-9^{318}*, while *Su(var)3-9^{330}* showed a small reduction in *H3* and an increase in *H2A* (Figures 3.8.c and 3.8.d). The differences in the levels of polyadenylated *H2A* and *H3* between the wild-type and each of the *Su(var)3-9* mutants are statistically significant. However, since the polyadenylated *histone* RNAs represent less than 1% of the total, these differences certainly do not account for the increase in the levels of total *histone* transcripts. Thus, we conclude that increased expression outside S-phase is not the mechanism responsible for the observed elevation in total *histone* transcripts.
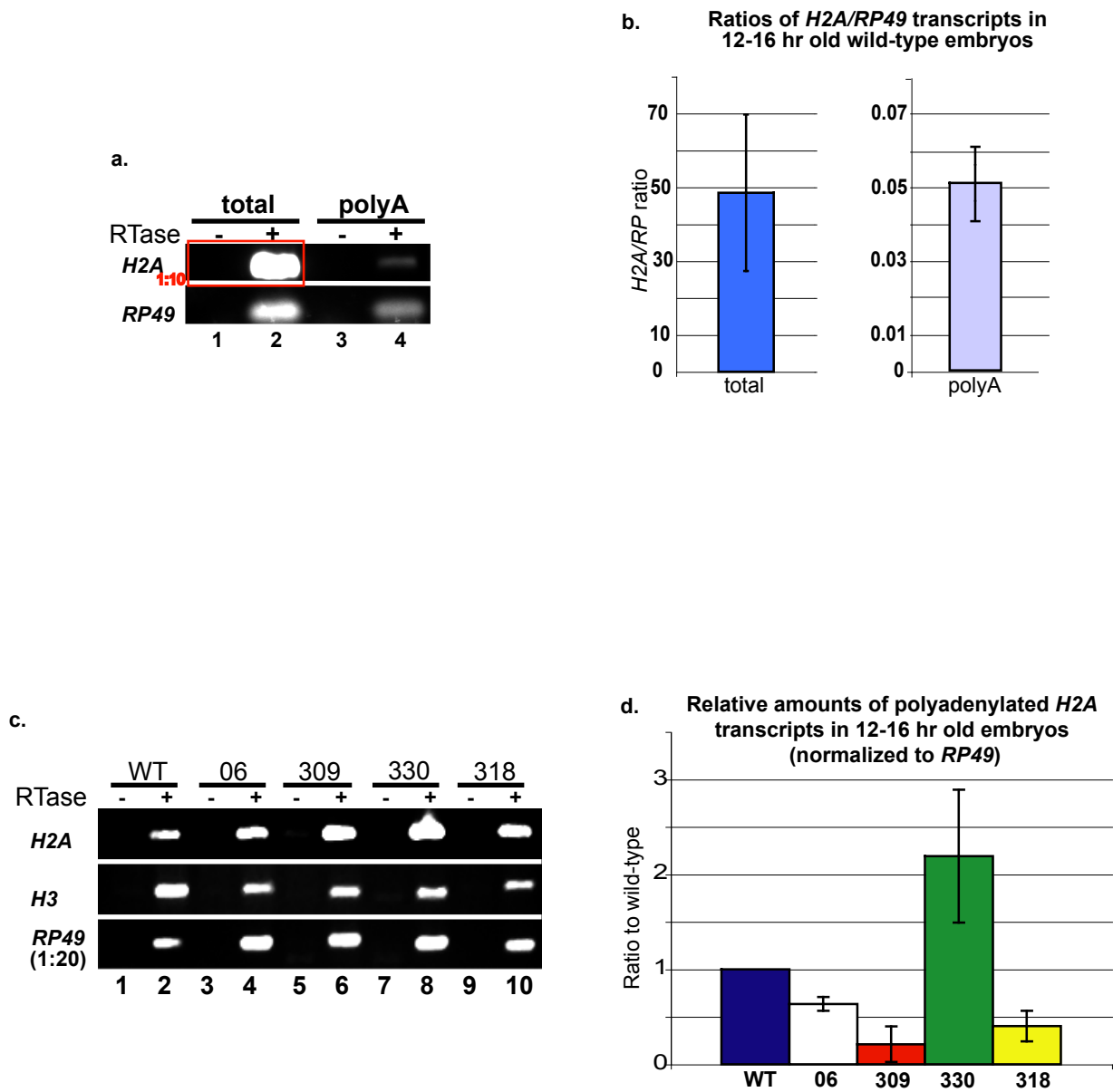
a.

	total		polyA	
RTase	−	+	−	+
H2A 1:10				
RP49				
	1	2	3	4

b. **Ratios of *H2A/RP49* transcripts in 12-16 hr old wild-type embryos**

c.

	WT		06		309		330		318	
RTase	−	+	−	+	−	+	−	+	−	+
H2A										
H3										
RP49 (1:20)										
	1	2	3	4	5	6	7	8	9	10

d. **Relative amounts of polyadenylated *H2A* transcripts in 12-16 hr old embryos (normalized to *RP49*)**

Figure 3.8.

Figure 3.8. Relative quantifications of polyadenylated (polyA) *histone H3* and *histone H2A* transcript in wild-type and *Su(var)3-9* mutant embryos.

a) Comparison between the relative amount of total and polyA *H2A* transcript in wild-type 12-16 hours old embryos. In lanes 1 and 2 the RNA was reverse-transcribed with primers H2AR and RP49rt, and in lanes 3 and 4 with an oligodT primer. In both cases the reverse-transcribed material was then PCR amplified (separately) with primer pairs H2AFc/H2AR and RP495/RP493. Lanes 1 and 3 represent "mock" (no RTase) reactions. For the total *H2A* transcript sample, only 1/10 of the amplified product was loaded on the gel. **b)** Ratios of total *H2A*/total *RP49*, and polyadenylated *H2A*/polyadenylated *RP49* transcripts, respectively, as determined by real-time PCR. The histograms represent the average of three independent experiments and the error bars span two standard deviations. **c)** RT-PCRs on total RNA extracted from 12-16 hours old wild-type and *Su(var)3-9* mutant embryos. The extracts were reverse-transcribed with an oligodT primer and the material obtained amplified with primer pairs H2AFc/H2AR, H3F/H3R and RP495/RP493, respectively. "Mock" (no RTase) reactions are indicated with a "-" sign above the corresponding lanes. **d)** Relative quantifications of polyadenylated *H2A* and *H3* transcripts by real-time PCR. For each reaction, the ratio of *H2A* or *H3* between mutants and wild-type (standardized for *RP49*) was reported. The histograms represent the average of three independent experiments and the error bars span two standard deviations.

3.3 Discussion

3.3.1. <u>Association of SU(VAR)3-9 with the HIS-C in 3-9 mutants</u>

The ChIP data demonstrate that SU(VAR)3-9 is present at the HIS-C in wild-type

and in *Su(var)3-9^{309}*, *Su(var)3-9^{330}* and *Su(var)3-9^{318}* embryos. This confirms

previous observations (Ner *et al*., 2002; Koryakov *et al.,* 2006) and suggests that

in all three cases, single amino acid substitutions in the catalytic region of

SU(VAR)3-9 do not substantially affect the protein's ability to be recruited to, or

associate with, the HIS-C. Since the three missense mutants tested carry distinct

amino acid substitutions, and all of them are still targeted to, and associated with

the HIS-C, we propose that their 3-dimensional structure, as well as their ability

to interact with the other chromatin components, are mainly unaffected. If this

were not the case, and one or more of the mutant gene products were misfolded

and/or unable to properly interact with the customary partners of SU(VAR)3-9,

we would expect them to fail to associate with their targets. The fact that no

SU(VAR)3-9 was detected in *Su(var)3-9^{06}* further validates this hypothesis.

All *Su(var)3-9* mutants tested, regardless of whether they were "protein nulls"

(*Su(var)3-9^{06}*) or missense alleles (*Su(var)3-9^{309}*, *Su(var)3-9^{330}* and *Su(var)3-*

9^{318}) display elevated levels of histone transcripts, suggesting that the physical

presence of SU(VAR)3-9 at the HIS-C is not sufficient for normal regulation of the

histone genes. It is certainly possible, at least in theory, that SU(VAR)3-9 has an

essential structural role at the HIS-C, and that each one of the single amino acid

substitutions present in the mutants analyzed impairs this function. We do not

favour this possibility because, as discussed above, the mutant proteins still

associate with their target (although in some cases less efficiently), which suggests that their ability to interact with other chromatin proteins does not differ significantly from that of the wild-type SU(VAR)3-9.

Although significant, the enrichment for HIS-C fragment detected in the material immunoprecipitated with α-SU(VAR)3-9^{chr} was very weak, which may reflect the fact that the association of SU(VAR)3-9 with this locus is very dynamic, and the amount of protein physically associated with the HIS-C chromatin at any given time is very small. This would fit with the observation that SUV39H1, the human homolog of SU(VAR)3-9, can be found stably associated with heterochromatin, but not with euchromatin (Krouwels *et al.,* 2005). It would also argue in favour of the hypothesis that SU(VAR)3-9 does not play a major structural role at the HIS-C. Alternatively, it is possible that the antibody was not very efficient in its recognition of SU(VAR)3-9 (possibly because the epitope against which it was raised is partially hidden by other chromatin proteins interacting with SU(VAR)3-9).

3.3.2. Reduced levels of H3K9me2 in *Su(var)3-9* mutants

Using ChIP analyses we were able to demonstrate conclusively that 1) at least three segments of the histone unit are enriched for H3K9me2, and 2) this enrichment is noticeably reduced in all four homozygous *Su(var)3-9* mutants tested (and, for at least one fragment, "H3 coding" the reduction is as high as ~7 to 25-fold). The product of *Su(var)3-9^{309}* (SU(VAR)3-9^{309}) is catalytically inactive *in vitro*, and those of *Su(var)3-9^{330}* (SU(VAR)3-9^{330}) and *Su(var)3-9^{318}*

(SU(VAR)3-9^{318}) have dramatically reduced enzymatic activity (chapter 2). Nevertheless, all strains, including *Su(var)3-9^{06}*, display a statistically significant enrichment for H3K9me2 within at least one region of the *his unit* ("H3 coding", see also Appendix 3). Although weak cross-reactivity of the antibody with H3K9me1 or H3K27me2 cannot be completely excluded, we favour the hypothesis that the residual signal observed in *Su(var)3-9^{06}* is due to the presence of other methyltransferases (MTases) that are able to methylate H3K9 at euchromatic loci, such as dG9a and DmSetDB1 (Mis *et al.,* 2006; Stabell *et al.,* 2006; Seum *et al.,* 2007; Tzeng *et al.,* 2007). This hypothesis could be tested by analysing cross-linked extracts from homozygous *dG9a$^-$*; *Su(var)3-9^{06}* and *DmSetdb1$^-$*; *Su(var)3-9^{06}* mutant embryos, respectively, and asking whether their level of H3K9me2 associated with the HIS-C is lower than that observed in their *Su(var)3-9^{06}* counterparts. Unfortunately, this approach may not be simple. In the case of *dSetdb1*, homozygous *dSetdb1$^-$* mutants survive until third instar larval stage (Seum *et al.,* 2007), so *dSetdb1$^-$*; *Su(var)3-9^{06}* homozygous embryos could in principle be produced, although the viability of the double mutant strain is not known. In addition, it is likely that traces of H3K9me2, generated by the dG9a protein, would still be present in homozygous *dSetdb$^-$*; *Su(var)3-9^{06}* embryos . Since the levels of H3K9me2 associated with the HIS-C in *Su(var)3-9^{06}* embryos are also very low, a comparison between *dG9a$^-$*; *Su(var)3-9^{06}* and *Su(var)3-9^{06}* would likely involve working with trace amounts of material and trying to detect a very small difference, which would require extremely accurate and reliable quantification methods, as well as exceedingly specific antibodies. In a way, the

task could be simpler for *dG9a*, since the *dG9a^{RG5}* mutant line survives, as a homozygote, in combination with *Su(var)3-9^{06}* (Seum *et al.,* 2007). However, this particular *dG9a* mutant line does not have any particular phenotype and fails to show reduced levels of H3K9me or H3K9me27 at larval stages, so there is no guarantee that a difference in the levels of such modifications could be detected in homozygous *dG9a^{RG5}*; *Su(var)3-9^{06}* vs. *dG9a^{+}*;*Su(var)3-9^{06}* embryos.

All *Su(var)3-9* missense mutants tested showed a significant reduction in the levels of H3K9me2 associated with the HIS-C. In principle, this could be due to either a reduction in the catalytic function of the mutant SU(VAR)3-9 products, or a reduction in the amount of SU(VAR)3-9 protein present at the HIS-C. Since we have shown that the level of (mutant) SU(VAR)3-9 associated with the HIS-C in the *Su(var)3-9* missense mutants are comparable to those observed in wild-type embryos (Figure 3.2), we conclude that the reduced levels of H3K9me2 are due to the altered enzymatic activity of the mutant products (SU(VAR)3-9^{309}, SU(VAR)3-9^{330} and SU(VAR)3-9^{318}, respectively). This was expected, as the catalytic activity of these mutant proteins is significantly impaired, at least *in vitro* (chapter 2).

Curiously, although SU(VAR)3-9^{309}, SU(VAR)3-9^{330} and SU(VAR)3-9^{318} differ in the strength of their catalytic phenotypes based on *in vitro* tests (chapter 2), we detected no significant differences in the levels of H3K9me2 that is associated with the HIS-C among the corresponding mutants. There are a number of possible reasons for this apparent discrepancy. The simplest explanation is that the conditions employed for the HMTase assay reported in chapter 2 do not

wholly simulate the nuclear environment in which SU(VAR)3-9 normally

functions. For instance, in the *in vitro* assay the reactions were allowed to

proceed uninterrupted for several hours, which is probably not the case in an *in

vivo* context. More importantly, the *in vitro* reactions were carried out on free

histones, in the absence of all the NHCPs and additional factors normally present

in a cell nucleus, and the results obtained may not be representative of the

enzyme's activity on a chromatin template in an *in vivo* context.

The dramatically reduced levels of HIS-C-associated H3K9me2 observed in the

Su(var)3-9 mutants correlate with an overall upregulation of the *histone* genes,

suggesting that dimethylation of H3K9 at the HIS-C is necessary to maintain

normal levels of *histone* transcripts in the cell. However, this conclusion may be

too simplistic since we don't know what other functions SU(VAR)3-9 may have,

and, if applicable, whether these other functions are affected in the *Su(var)3-9*

missense mutants. It has been suggested that the SET domain of SUV39H1, the

human homolog of SU(VAR)3-9, may play an important structural role in the

stable association of the protein with chromatin (Krouwels *et al.,* 2005). We know

that the residues that are affected in the *Su(var)3-9* mutants tested here are not

necessary for this process, since the association of these mutant forms of

SU(VAR)3-9 with the HIS-C is not significantly affected (Figure 3.2). However, as

discussed above, we can't exclude the possibility that such residues are required

for proper interaction with other chromatin components and/or regulators of the

histone genes, and that the elevated levels of *histone* transcripts are partially due

to the inability of the mutant SU(VAR)3-9 proteins to carry out such interactions.

3.3.3. Recruitment and role of HP1 at the HIS-C

The ChIP results obtained show that HP1 is present at the HIS-C in wild-type embryos, at least within the three regions of the histone unit that we analyzed (Figure 3.4). This is in agreement with previous reports (Greil *et al.*, 2003; Koryakov *et al.,* 2006). We also detected HP1 association with the histone units in the *Su(var)3-9* mutants, but its levels were drastically reduced in all mutants (Figure 3.4). We did not see a correlation between the magnitude of the reduction in the levels of H3K9me2 and the relative amount of HP1 associated with the different regions of the histone unit. This suggests that the presence of H3K9me2 is necessary for stable binding of HP1, but other factors, such as auxiliary proteins, are probably involved in the process, and these may influence the efficiency of HP1 binding to this chromatin domain.

Interestingly, the amount of HP1 associated with the histone units in the *Su(var)3-9* missense mutants was not significantly different from that found at the same locations in the "protein null" *Su(var)3-9[06]*. Thus the presence of a SU(VAR)3-9 protein with a single amino acid substitution in its catalytic region is not sufficient to recruit and stabilize the association of HP1 with the HIS-C. There are at least two possible explanations for this. The first one is that each of the SU(VAR)3-9 residues that are mutated in the missense alleles is critical for the binding of HP1. The second one is that the physical presence of SU(VAR)3-9, whether wild-type or mutant, is not sufficient to recruit and/or to stabilize HP1 at the HIS-C.

Although *a priori* it is difficult to decide which hypothesis is most plausible, we tend to favour the second one, for the following reasons. Firstly, it has been shown that, at least *in vitro*, the presence of auxiliary factors is required for stable binding of HP1 to a chromatin template, even in the presence of H3K9me2,3 (Eskeland *et al.,* 2007). Secondly, we showed that all the mutant SU(VAR)3-9 proteins, SU(VAR)3-9[309], SU(VAR)3-9[330] and SU(VAR)3-9[318], localize at the HIS-C like their wild-type counterpart (Figure 3.2), suggesting that *in vivo* their structure and folding are relatively unaffected. Thus, it is likely that their ability to interact with other chromatin proteins does not differ significantly from that of wild-type SU(VAR)3-9.

At this point, we propose that the most essential factor in the recruitment of HP1 at the HIS-C is the presence of a certain level of H3K9me2, and that this histone modification, in combination with the presence of NHCPs other than SU(VAR)3-9, are responsible for the stable association of HP1 with chromatin at this locus. Curiously, the relative enrichment for HP1 in *Su(var)3-9[330]* is significantly lower than that observed in *Su(var)3-9[06]*, suggesting that *Su(var)3-9[330]* may be acting as an antimorph, effectively hindering the stable association of HP1 with the HIS-C. In chapter 2 we showed that, *in vitro*, the recombinant SU(VAR)3-9[330] protein partially retains the ability to methylate histone H3, but can't methylate a peptide representing the histone H3 tail alone. One of the hypotheses proposed to explain this result was that the amino acid substitution present in SU(VAR)3-9[330] (D536N) might result in a change in specificity, causing SU(VAR)3-9[330] to methylate a histone H3 residue other than K9. Several studies have shown that

chromodomains can show very high specificity with respect to modified histone residues; for example, the chromodomain of HP1 specifically binds H3K9me2,3, but not H3K27me, while the opposite is true for the chromodomain of POLYCOMB (Bannister *et al.,* 2001; Nakayama *et al.,* 2001; Muller *et al.,* 2002; Jacobs and Khorasanizadeh, 2002; Nielsen *et al.,* 2002; Cao *et al.,* 2002; Min *et al.,* 2003; Fischle *et al.,* 2004; Pray-Grant *et al.,* 2005; reviewed by Daniel *et al.,* 2005). If the change in specificity hypothesized for SU(VAR)3-9[330] occurs *in vivo*, then the methylation of this other H3 residue may create a binding platform that recruits or stabilizes the binding of a different chromodomain protein, and the binding of this inappropriate NHCP may preclude binding of HP1. Alternatively, it is conceivable that SU(VAR)3-9[330] binds to histone H3, and may or may not methylate K9, but it remains tightly associated with the H3 tail, thus making it unavailable for the recruitment of HP1. We do not favour this possibility because, if SU(VAR)3-9[330] remained tightly bound to the H3 tail, and thus to chromatin, we would expect to see higher levels of (mutant) SU(VAR)3-9 associated with the HIS-C in the ChIP experiments on *Su(var)3-9[330]* extracts than in their wild-type counterparts. However, this is not the case (Figure 3.2).

3.3.4. Elevated levels of the *histone* transcripts in *Su(var)3-9* mutants

In a previous study, members of our lab (Ner *et al.,* 2002) showed that the levels of *histone H1* and *H4* mRNA present in two *Su(var)3-9* mutants (one missense allele, *Su(var)3-9[330]*, and one P element-induced allele, *Su(var)3-9[P25]*) is about two fold higher than in wild-type individuals. We confirmed and expanded this

observation. Our results show that the level of both the *H2A* and the *H3 histone* transcripts are elevated in three *Su(var)3-9* missense mutants and the complete null allele *Su(var)3-9^{06}*. We find the difference between the mutant and wild-type strains to range between ~1.8 and 6.5-fold.

The apparent discrepancies between these results and the data obtained in the 2002 study are likely due, at least in part, to differences in the experimental setup. In the former study the relative abundance of each *histone* transcript was quantified by northern blot analyses. In contrast, this study employed real-time/RT-PCR, a more accurate and reliable quantification system. Moreover, our mRNA samples are derived from staged embryos at a time when many of the cells are still undergoing mitosis (reviewed by Lee and Orr-Weaver, 2003). Since *histone* genes are expressed almost exclusively during S-phase, their misregulation would be more noticeable at this stage than in adult flies, which were used for the previous study (Ner *et al*., 2002).

In the present study, we were able to detect differences in the relative amounts of *histone h2a* and *h3* transcripts between some of the mutants. In particular, *Su(var)3-9^{06}* and *Su(var)3-9^{318}* have significantly higher levels of *h2a* and *h3* than *Su(var)3-9^{309}*. Since *Su(var)3-9^{06}* is a "protein null" (no RNA or protein product detected, see chapter 2) we suggest that *Su(var)3-9^{318}* acts as an amorph (at least functionally) and *Su(var)3-9^{309}* behaves as a hypomorph, with respect to the regulation of *h2a* and *h3* gene expression.

Once again, the *Su(var)3-9^{330}* mutant strain has a very curious phenotype. It shows dramatically elevated levels of *h3* (~6 fold over the wild-type), but only a ~1.8 to 1.9 fold increase in the *h2a* and *h2b* transcripts. Thus, the *Su(var)3-9^{330}* mutation is unique in that, unlike the other alleles studied, it seems to decouple the *h2a/h2b* from the *h3* expression levels. This mutant is therefore difficult to categorize, as it would be classified as a loss-of-function mutant (same phenotype as *Su(var)3-9^{06}* and *Su(var)3-9^{318}*) based on its *h3* expression level, but not with respect to *h2a* and *h2b*.

3.3.5. Possible mechanism for the functions of SU(VAR)3-9 at the HIS-C

In general, the results obtained in this study are consistent with the notion that the function of SU(VAR)3-9 at the HIS-C, a euchromatic locus, is very similar to its function in (pericentric) heterochromatin. In both cases SU(VAR)3-9 is required for downregulation or silencing, and its ability to methylate H3K9 is required to allow stable association of HP1 with chromatin.

Our results suggest that the physical presence of SU(VAR)3-9 probably does not play a major role in the regulation of the *histone* transcripts levels. *Su(var)3-9^{318}*, a mutant that produces a SU(VAR)3-9 protein with a single amino acid substitution, and which localizes at the HIS-C like its wild-type counterpart, shows the same *h2a* and *h3* hyperexpression phenotype as *Su(var)3-9^{06}*, a mutant that does not produce any SU(VAR)3-9 protein at all (no statistical difference in a t test at $p=0.05$). In addition, the missense mutant *Su(var)3-9^{330}* also displays the same hyperexpression phenotype as *Su(var)3-9^{06}* with respect

to the *histone h3* gene. If the physical presence of SU(VAR)3-9 played a major role in the regulation of the *histone* genes expression, one would expect most missense mutants to have a milder overexpression phenotype than the "protein null" *Su(var)3-9^{06}*.

All data presented point to the crucial role of SU(VAR)3-9's enzymatic activity in the regulation of the levels of *histone* transcripts; all mutants have an impaired HMTase activity (chapter 2), all of them display a drastic reduction in the abundance of H3K9me2 associated with the HIS-C, and all of them show elevated levels of *h2a* and *h3* transcripts. However, the relative levels of H3K9me2 associated with the various regions of the histone units are virtually indistinguishable in all *Su(var)3-9* mutants, while the abundance of *h3* and *h2a* transcripts may differ. The simplest explanation for this apparent disparity is that differences in the levels of H3K9me2 do exist among the mutants, but the resolution of the ChIP technique is not sensitive enough to detect them. It is also possible that the mutants do not differ significantly for H3K9me2, but they do with respect to another NHCP, or another histone modification, which could have a role in the fine-tuning of the regulation of *histone* genes expression.

Interestingly, the levels of HP1 associated with the "H3 coding" region in the *Su(var)3-9^{06}*, *Su(var)3-9^{318}* and *Su(var)3-9^{309}* mutants correlates with their respective *h3* transcript levels. For instance, the first two mutants show similar levels of HP1 associated with this region of the histone unit, and a similar increase in *h3* transcript. The *Su(var)3-9^{309}* has a higher level of HP1 associated with the "H3 coding" region and a lower increase in its *h3* transcript, which is

consistent with the view that the presence of HP1 is associated with silencing. As usual, *Su(var)3-9*[330] represents an exception. In this case, its levels of "H3 coding"-associated HP1 are significantly lower than those observed in the other mutants, yet the increase in its *h3* transcript level is similar to that of *Su(var)3-9*[06] and of *Su(var)3-9*[318].

In contrast, we did not detect a correlation between the level of HP1 associated with the "H2A coding" region and the relative increase in *h2a* transcript produced in the various mutants. The levels of HP1 associated with the "H2A coding" region did not differ significantly in *Su(var)3-9*[06], *Su(var)3-9*[318] and *Su(var)3-9*[309], but *Su(var)3-9*[309] has a much weaker *h2a* hyperexpression phenotype than the other two strains.

We have demonstrated that reduced levels of H3K9me2 and HP1 at the HIS-C correlate with an increase in the abundance of *h2a* and *h3* transcripts produced. We propose that there is a causal relationship between the two observations, and that the presence of appropriate amounts of H3K9me2 and HP1 at the HIS locus is absolutely necessary for the production of wild-type levels of *histone* transcripts. Our data also suggest that high levels of H3K9me2 and/or of HP1 may not be sufficient, by themselves, to ensure a normal regulation of the *histone* genes expression. We propose that additional factors are involved in the process of fine-tuning the production of *histone* transcripts.

In order to gain a thorough understanding of the mechanisms involved in the regulation of the HIS-C at the chromatin level, it will be necessary to identify the other factors associated with this *locus*, their respective functions and how they

interact with one another. So far, we know that the gene products of at least

three other suppressors of PEV (*abo*, *Bonus* and *Su(var)326/HDAC1*) localize to

the HIS-C (Berloco *et al.,* 2001; Beackstead *et al.,* 2005; Ner *et al.*, in

preparation). Investigations of their respective roles and functions in the

modulation of the *histone* gene expression will likely help our understanding of

this complex system.

3.3.6. <u>Regulation of the *histone* gene expression</u>

Our data show that the amounts of *h2a* and *h3* transcripts are higher in *Su(var)3-9* mutants than in *Su(var)3-9*[+] embryos and, as discussed previously, we think

that this is most likely the result of overexpression of the *histone* genes. We have

shown that the over production of *histone* transcripts is not due to decoupling of

histone genes expression from the cell cycle (*i.e*. synthesis of histones outside S

phase). Thus, we are left with two possibilities for the role of SU(VAR)3-9 in the

regulation of the *histone* genes expression.

In the first scenario, SU(VAR)3-9 determines the number of *histone* templates

that are actively transcribed in a nucleus. In this model, mutations in SU(VAR)3-9

would cause an increase in the number of transcribed templates without affecting

transcription rates. In a wild-type situation the HMTase activity of SU(VAR)3-9

may be mainly responsible for keeping a certain number of *histone units*

inaccessible to transcription factors. In *Su(var)3-9* mutant strains more units may

be accessible and thus more *histone* transcripts are produced during S-phase

while the transcription rate/template remains unchanged. This hypothesis is

supported by the fact that the chromatin structure of the HIS-C, as measured by the pattern of DNaseI and MNase hypersensitive sites, is altered in at least a subset of *Su(var)3-9* mutants (Ner *et al.,* 2002). The alternative model is that SU(VAR)3-9 modulates the rate of transcription of each *histone* template or unit. In this scenario, the same number of histone templates would be transcribed in wild-type and *Su(var)3-9* mutants, but the mutants would have higher transcription rates.

Of course, these two hypotheses are not mutually exclusive, and SU(VAR)3-9 could have several functions regulating template accessibility as well as modulating the transcription rate of each *histone* gene.

Finally, it should be noted that the peculiar phenotype of the *Su(var)3-9[330]* mutant, which shows a drastic increase in its levels of *h3*, but only a moderate elevation in *h2a* and *h2b* transcripts, suggests that the *h2a/h2b* and *h3/h4* gene pairs may be regulated somewhat independently. In a wild-type situation, the regulation of these two gene pairs is usually co-ordinated. In two *Su(var)3-9* mutants (*Su(var)3-9[309]* and *Su(var)3-9[318]*) this co-ordinated regulation appears to be maintained, as their increase in *h2a* transcript is virtually identical to their increase in *h3* transcript. *Su(var)3-9[06]* shows a slightly higher increase in *h3* than *h2a*, and for *Su(var)3-9[330]* this decoupling of *h2a/h2b* couplet from *h3* is very dramatic. We do not have an explanation for these results, but it is possible that SU(VAR)3-9 is involved, directly or indirectly in the co-regulation of the *h2a/h2b* and *h3/h4* pairs.

3.4. Materials and methods

3.4.1 *Drosophila* strains

All *Drosophila* strains used in this study are described in chapter 2, and were grown under standard conditions.
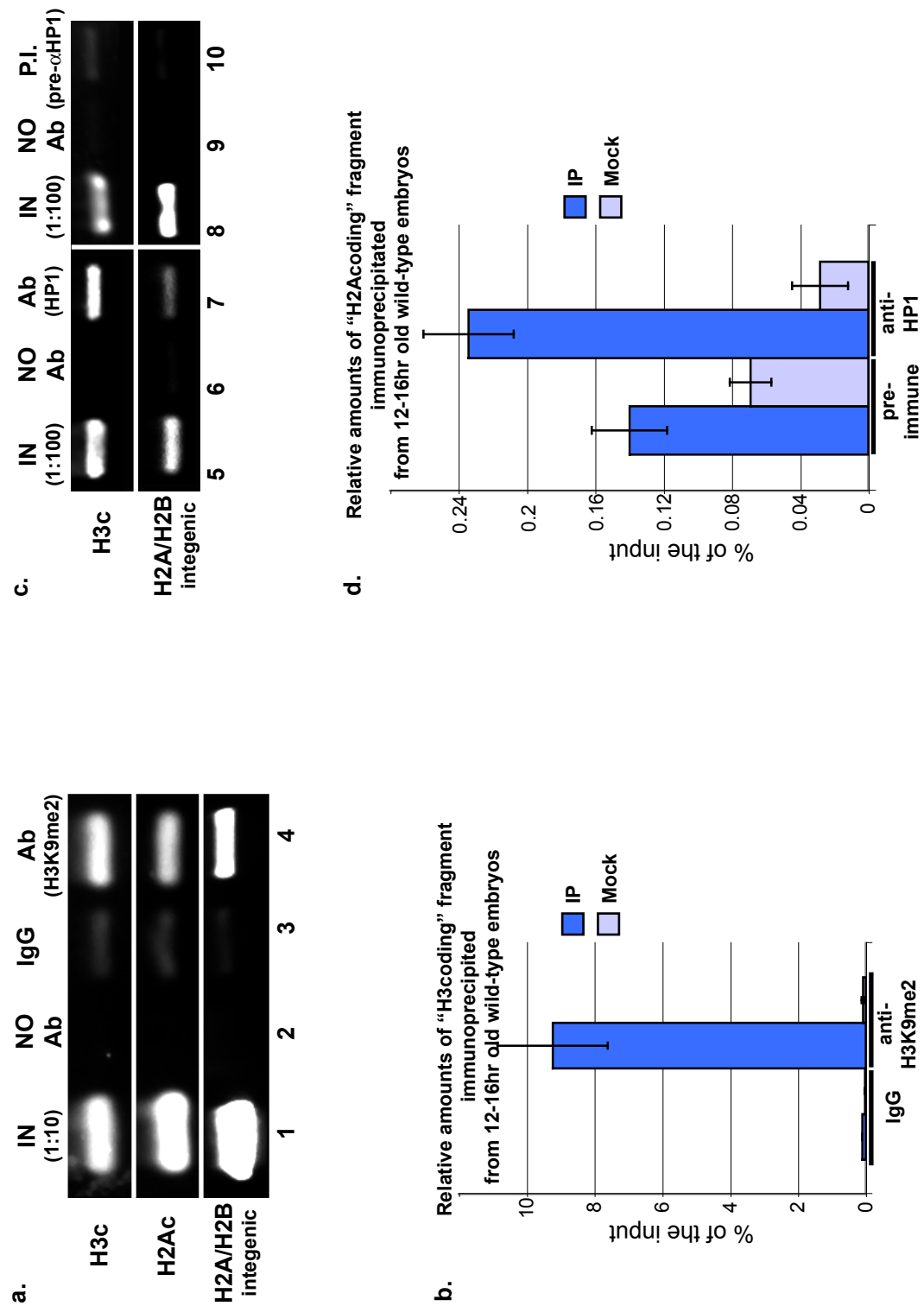
3.4.2. ChIP analysis

ChIP analysis of the HIS-C was performed as described in chapter 2. The antibody to H3K9me2 was purchased from Upstate (#07-441) and the anti-SU(VAR)3-9 antibody is described in chapter 2 and by Ner *et al.* (2002). The antibody to HP1 is a polyclonal and was raised in our lab (Ner *et al.,* in preparation). The primers pairs used for amplification were H3F/H3R ("H3 coding" region), iH2AF/iH2BR ("H2A/H2B intergenic" region), H2AFc/H2BR ("H2A coding" region), iH3F/iH4R ("H3/H4 intergenic" region) and iH1Bf/iH3r ("H1/H3 intergenic" region) (see supplementary table 3.1 and Ner *et al.,* 2002). The relative amounts of target DNA present in the immunoprecipitated material was quantified by real-time PCR. A dilution series (1:10-1:2000) of the input DNA was used to generate a standard curve, which was then used to estimate the amount of the fragment of interest present in the immunoprecipitated material (or in the mock reaction). In each case, the material pulled down in 3 independent experiments was individually analyzed for its enrichment in target (HIS-C) sequences, expressed as a % of the input for each IP (same for the mock IPs). The results obtained with each mutant strain were compared to those relative to the wild-type strain using Student's T test. Differences were considered

significant if the absolute t value was above 2.78 (critical value for p=0.05, 4 df, in a two-tailed T test). The same test was also employed to determine whether, in each strain, the enrichment observed was above the background.

3.4.3. Quantifications of *histone* mRNAs

Total RNA was isolated from 12-16 hours old embryos by TRIzol extraction as recommended by the manufacturer. Three independent extracts were prepared from each strain. For each extract, reverse-transcription reactions were performed as described in chapter 2. To reverse-transcribe total *histone H2A, H3* and *H2B* mRNA, primers H2AR, H3R and dH2BR were used, each one in combination with RP49rt (internal control). A 17 nucleotide oligodT primer was employed to reverse-transcribe polyadenylated mRNAs. The relative amounts of *histone* cDNAs obtained from each RT reaction were quantified by real-time PCR amplification. For each sample, three reactions were run and, since the resulting Ct values were almost identical (<1% difference) among the three measurements, they were averaged, and referred to as "one measurement". The ratio between the *histone* and the *rp49* measurements was calculated for three independent RNA extracts. In each case, the resulting value was then divided by the value obtained, by the same procedure, with the wild-type strain, giving rise to one "data point". For each strain, the results are presented as an average of three (independent) "data points" +/- the standard deviation. Differences among strains were considered significant if they yielded a |t value| >2.78 (critical value for p=0.05, 4df) in a two-tailed Student's t-test.

Supplementary figure 3.1

a.

	IN (1:10)	NO Ab	IgG	Ab (H3K9me2)
H3c				
H2Ac				
H2A/H2B integenic				
	1	2	3	4

b.

Relative amounts of "H3coding" fragment immunoprecipited from 12-16hr old wild-type embryos

% of the input

IP
Mock

anti-H3K9me2

IgG

c.

	IN (1:100)	NO Ab	Ab (HP1)	IN (1:100)	NO Ab	P.I. (pre-αHP1)
H3c						
H2A/H2B integenic						
	5	6	7	8	9	10

d.

Relative amounts of "H2Acoding" fragment immunoprecipitated from 12-16hr old wild-type embryos

% of the input

IP
Mock

anti-HP1

pre-immune

168

Supplementary figure 3.1. Control ChIP reactions from wild-type 12-16 hours old embryo extracts using commercial, non-specific IgG and pre-immune serum, respectively. a) Representative gels showing the PCR products obtained using the primers that amplify the "H3 coding" (H3c), "H2A coding" (H2Ac) and "H2A/H2B intergenic" fragments, respectively. The template was a 1:10 dilution of the ChIP input DNA (lane 1), the DNA obtained from a mock IP (no antibody, lane 2), or the DNA IPed with commercial non-specific IgG (lane 3) or with the anti-H3K9me2 antibody (lane 4). **b)** Relative abundance of the "H3 coding" fragment in the material IPed with IgG and with anti-H3K9me2 determined by real-time PCR (see materials and methods for details). **c)** Representative gels showing the PCR products obtained using the primers that amplify the "H2A coding" (H2Ac) and "H2A/H2B intergenic" fragments, respectively. The template was a 1:100 dilution of the ChIP input DNA (lanes 5 and 8), the DNA obtained from a mock IP (no antibody, lanes 5 and 9), or the DNA IPed with the anti-HP1 antibody (lane 7) or with pre-immune serum (lane 10). **d)** Relative abundance of the "H2A coding" fragment in the material IPed with IgG and with anti-H3K9me2 determined by real-time PCR (see materials and methods for details).

Supplementary table 3.1.
List of primers used in this study.

Primer name	Primer sequence (5'➔3')
H2AR	AACGTTTAGGCCTTCTTCT
H2AFc	TCTGGACGTGAAAAAGGTGG
H3R	TGCCGTGTCAGCTTAAGCA
H3F	GCTCGTACCAAGCAAACTG
RP495	GCCCAAGATCGTGAAGAAGC
RP493	CTGTTGTCGATACCCTTGGG
RP49rt	CGCGCTCGATAATCTCC
dH2BR	GTCCGCATTCGCAGGAG
dH2BF	CCTCCGAAAACTAGTGGA
iH2BR	ATGGCATAGCTCTCCTTCC
iH2AF	TTCCGGAGCAAACGGTGA
iH1Bf	TCCGCAACAAAATTAGCCAA
iH3R	AAGCGCTAGCGTACTCTATAA
iH3F	GCGTGGCGCCTTTCCACCAGTC
iH4R	CGCTTGGCGCCACCCTTT

3.5. References

Akhmanova, A., Miedema, K., Kremer, H., Hennig, W.: Two types of polyadenylated mRNAs and synthesized from *Drosophila* replication-dependent histone genes. Eur. J. Biochem. *244*, 294-300 (1997)

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., Kouzarides, T.: Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120-4 (2001)

Beckstead, R.B., Ner, S.S., Hales, K.G., Grigliatti, T.A., Baker, B.S., Bellen, H.J.: Bonus, a *Drosophila* TIF1 homolog, is a chromatin-associated protein that acts as a modifier of position-effect variegation. Genetics *169*, 783-94 (2005)

Berloco, M., Fanti, L., Breiling, A., Orlando, V., Pimpinelli, S.: The maternal effect gene, *abnormal oocyte* (*abo*), of *Drosophila melanogaster* encodes a specific negative regulator of histones. Proc Natl Acad Sci U S A. *98*, 12126-31 (2001)

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., Zhang, Y.: Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science *298*, 1039-43 (2002)

Czermin, B., Schotta, G., Hulsmann, B.B., Brehm, A., Becker, P.B., Reuter, G., Imhof, A.: Physical and functional association of SU(VAR)3-9 and HDAC1 in *Drosophila*. EMBO Rep. *2*, 915-9 (2001)

Daniel, J.A., Pray-Grant, M.G., Grant, P.A.: Effector proteins for methylated histones: an expanding family. Cell Cycle *4*, 919-26 (2005)

Danzer, J.R., Wallrath, L.L.: Mechanisms of HP1-mediated gene silencing in *Drosophila*. Development *131*, 3571-80 (2004)

de Wit, E., Greil, F., van Steensel, B.: High-resolution mapping reveals links of HP1 with active and inactive chromatin components. PLoS Genet. *3*, e38 (2007)

Delattre, M., Spierer, A,. Jaquet, Y., Spierer, P.: Increased expression of *Drosophila Su(var)3-7* triggers *Su(var)3-9*-dependent heterochromatin formation. J Cell Sci. *117*, 6239-47 (2004)

Dominski, Z., Marzluff, W.F.: Formation of the 3' end of histone mRNA: getting closer to the end. Gene *396*, 373-390 (2007)

Ebert, A., Schotta, G., Lein, S., Kubicek, S., Krauss, V., Jenuwein, T., Reuter, G.: *Su(var)* genes regulate the balance between euchromatin and heterochromatin in

Drosophila. Genes Dev. *18*, 2973-83 (2004)

Ebert, A., Lein, S., Schotta, G., Reuter, G.: Histone modification and the control of heterochromatic gene silencing in *Drosophila*. Chromosome Res. *14*, 377-92 (2006)

Eskeland, R., Czermin, B., Boeke, J., Bonaldi, T., Regula, J.T., Imhof, A.: The N-terminus of *Drosophila* SU(VAR)3-9 mediates dimerization and regulates its methyltransferase activity. Biochemistry *43*, 3740-9 (2004)

Eskeland, R., Eberharter, A., Imhof, A.: HP1 binding to chromatin methylated at H3K9 is enhanced by auxiliary factors. Mol Cell Biol. *27*, 453-65 (2007)

Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D., Khorasanizadeh, S.: Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes Dev. *17*, 1870-81 (2003)

Greil, F., van der Kraan, I., Delrow, J., Smothers, J.F., de Wit, E., Bussemaker, H.J., van Driel, R., Henikoff, S., van Steensel, B.: Distinct HP1 and Su(var)3-9 complexes bind to sets of developmentally coexpressed genes depending on chromosomal location. Genes Dev. *17*, 2825-38 (2003)

Harrington, M.J.: Ph.D. Thesis. The University of British Columbia (2001)

Jacobs, S.A., Taverna, S.D., Zhang, Y., Briggs, S.D., Li, J., Eissenberg, J.C., Allis, C.D., Khorasanizadeh, S.: Specificity of the HP1 chromo domain for the methylated N-terminus of histone H3. EMBO J. *20*, 5232-41 (2001)

Jacobs, S.A., Khorasanizadeh, S.: Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. Science *15*, 2080-3 (2003)

Jaquet, Y., Delattre, M., Spierer, A., Spierer, P.: Functional dissection of the *Drosophila* modifier of variegation *Su(var)3-7*. Development *129*, 3975-82 (2002)

Koryakov, D.E., Reuter, G., Dimitri, P., Zhimulev, I.F.: The *SuUR* gene influences the distribution of heterochromatic proteins HP1 and SU(VAR)3-9 on nurse cell polytene chromosomes of *Drosophila melanogaster*.
Chromosoma *115*, 296-31 (2006)

Krouwels, I.M., Wiesmeijer, K., Abraham, T.E., Molenaar, C., Verwoerd, N.P., Tanke, H.J., Dirks, R.W.: A glue for heterochromatin maintenance: stable SUV39H1 binding to heterochromatin is reinforced by the SET domain. J Cell Biol.*15*, 537-49 (2005)

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., Jenuwein, T.: Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature *410*, 116-20 (2001)

Lee, L.A., Orr-Weaver, T.L.: Regulation of cell cycles in *Drosophila* development: intrinsic and extrinsic cues. Annu. Rev. Genet. *37*, 545-78 (2003)

Lifton, R.P., Goldberg, M.L., Karp, R.W., Hogness, D.S.: The organization of the histone genes in Drosophila melanogaster: functional and evolutionary implications. Cold Spring Harb Symp Quant Biol. *42*, 1047-51 (1978)

Macaluso, M., Paggi, M.G., Giordano, A.: Genetic and epigenetic alterations as hallmarks of the intricate road to cancer. Oncogene *29*, 6472-8 (2003)

Min, J., Zhang, Y., Xu, R.M.: Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. Genes Dev. *17*, 1823-8 (2003)

Mis, J., Ner, S.S., Grigliatti, T.A.: Identification of three histone methyltransferases in *Drosophila*: dG9a is a suppressor of PEV and is required for gene silencing. Mol Genet Genomics *275*, 513-26 (2006)

Müller, J., Hart, C.M., Francis, N.J., Vargas, M.L., Sengupta, A., Wild, B., Miller, E.L., O'Connor, M.B., Kingston, R.E., Simon, J.A.: Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex.
Cell *111*, 197-208 (2002)

Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D., Grewal, S.I.: Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science *292*, 110-3 (2001)

Ner, S.S., Harrington, M.J., Grigliatti, T.A.: A role for the *Drosophila* SU(VAR)3-9 protein in chromatin organization at the histone gene cluster and in suppression of position-effect variegation. Genetics *162*, 1763-74 (2002)

Nielsen, S.J., Schneider, R., Bauer, U.M., Bannister, A.J., Morrison, A., O'Carroll, D., Firestein, R., Cleary, M., Jenuwein, T., Herrera, R.E., Kouzarides, T.: Rb targets histone H3 methylation and HP1 to promoters. Nature *412*, 561-5 (2001)

Nielsen, P.R., Nietlispach, D., Mott, H.R., Callaghan, J., Bannister, A., Kouzarides, T., Murzin, A.G., Murzina, N.V., Laue, E.D.: Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. Nature *416*, 03-7 (2002)

Peters, A.H., O'Carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schofer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A., Opravil, S., Doyle, M., Sibilia, M., Jenuwein, T.: Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell *107*, 323-37 (2001)

Peters, A.H., Mermoud, J.E., O'Carroll, D., Pagani, M., Schweizer, D., Brockdorff, N., Jenuwein, T.: Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin. Nat Genet. *30*, 77-80 (2002)

Pray-Grant, M.G., Daniel, J.A., Schieltz, D., Yates, J.R. 3rd, Grant, P.A.: Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. Nature *433*, 434-8 (2005)

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., Jenuwein, T.: Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593-9 (2000)

Rudolph, T., Yonezawa, M., Lein, S., Heidrich, K., Kubicek, S., Schafer, C., Phalke, S., Walther, M., Schmidt, A., Jenuwein, T., Reuter, G.: Heterochromatin formation in *Drosophila* is initiated through active removal of H3K4 methylation by the LSD1 homolog SU(VAR)3-3. Mol Cell. *26*, 103-15 (2007)

Samal, B., Worcel, A., Louis, C., Schedl, P.: Chromatin structure of the histone genes of *D. melanogaster*. Cell *23*, 401-9 (1981)

Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R., Reuter, G.: Central role of *Drosophila* SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. EMBO J. *21*, 1121-31 (2002)

Schotta, G., Ebert, A., Reuter, G.: SU(VAR)3-9 is a conserved key function in heterochromatic gene silencing. Genetica *117*, 149-58 (2003)

Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D., Jenuwein, T.: A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes Dev. *18*, 1251-62 (2004)

Seum, C., Reo, E., Peng, H., Rauscher, F.J. 3rd, Spierer, P., Bontron, S.: *Drosophila* SETDB1 is required for chromosome 4 silencing. PLoS Genet. *3*, e76 (2007)

Seum, C., Bontron, S., Reo, E., Delattre, M., Spierer, P.: *Drosophila G9a* is a nonessential gene. Genetics *177*, 1955-7 (2007)

Stabell, M., Eskeland, R., Bjorkmo, M., Larsson, J., Aalen, R.B., Imhof, A.,

Lambertsson, A.: The *Drosophila* G9a gene encodes a multi-catalytic histone methyltransferase required for normal development. Nucleic Acids Res. *34*, 4609-21 (2006)

Swaminathan, J., Baxter, E.M., Corces, V.G.: The role of histone H2Av variant replacement and histone H4 acetylation in the establishment of *Drosophila* heterochromatin. Genes Dev. *19*, 65-76 (2005)

Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G., Reuter, G.: The protein encoded by the *Drosophila* position-effect variegation suppressor gene *Su(var)3-9* combines domains of antagonistic regulators of homeotic gene complexes. EMBO J. *13*, 3822-31 (1994)

Tzeng, T.Y., Lee, C.H., Chan, L.W., Shen, C.K.: Epigenetic regulation of the *Drosophila* chromosome 4 by the histone H3K9 methyltransferase dSETDB1. Proc Natl Acad Sci U S A. *104*, 12691-6 (2007)

Worcel, A., Gargiulo, G., Jessee, B., Udvardy, A., Louis, C., Schedl, P.: Chromatin fine structure of the histone gene complex of *Drosophila melanogaster*. Nucleic Acids Res. *11*, 421-39 (1983)

4. GENERAL SUMMARY AND DISCUSSION

4.1. Summary of results

In chapter 2 we characterized a subset of EMS-induced *Su(var)3-9* missense alleles with respect to a variety of phenotypes, including the relative strength with which they suppress PEV, and the levels of residual HMTase activity of their respective gene products. Based on sequence information, and on the crystal structure of Clr4 (the *S. pombe* homolog of SU(VAR)3-9) (Min *et al.,* 2002) we were able to carry out a structure/function study of SU(VAR)3-9 and to rationalize the effects of each amino acid substitution on the various phenotypes analyzed. The nine mutants characterized were originally generated in a screen for strong, dominant modifiers of PEV (Sinclair *et al.*, 1983) but, remarkably, all of them have single amino acid substitutions only in the catalytic region of the protein and show significantly impaired HMTase activity. This result strongly suggests that the enzymatic activity of SU(VAR)3-9 is crucial for its silencing function and, perhaps more interestingly, that single amino acid substitutions elsewhere in the protein have no discernable effect on silencing, at least in this particular assay.

The data presented in chapter 3 shows the effects of one "protein null", and three of the *Su(var)3-9* missense mutants, on the regulation of a euchromatic locus, the HIS-C. Again, our results suggest that the HMTase activity of SU(VAR)3-9 is essential for the recruitment and/or binding of HP1 to the *his* units that comprise the HIS-C, and in the (down)regulation/silencing of the locus. Taken together, these two data sets strongly suggest that SU(VAR)3-9 likely acts in similar ways at the HIS-C and within the *w^{m4}* variegating

rearrangement and, by extension, heterochromatin. The results presented in chapter 3 strongly suggest that it is the catalytic activity of SU(VAR)3-9, rather than its physical association with the HIS-C locus, that plays a major role in the regulation of the *histone* gene transcription or in the association of HP1 with this locus.

Overall, we compared between 4 and 7 *Su(var)3-9* mutants with the wild-type strain (and with each other) in a number of biochemical, molecular, cytological and morphological assays. When trying to integrate the data from the different types of assays, in most cases it was possible to detect some general trends and to rationalize the results obtained. However, no two assays gave results that were completely consistent with each other for every mutant (Table 4.1 and Figure 4.1).

I suggest that these slight variations may derive mostly from the nature and limitations of the assays employed. However, slight differences in the role(s) of SU(VAR)3-9 in the regulation of the HIS-C and in PEV cannot be ruled out. The presence of these apparent "inconsistencies" between data sets from different assays highlights a third, very important point, namely, the advantage of testing mutants for a variety of parameters, rather than solely focusing on one specific aspect of their phenotypes. In this specific case, such a multi-assay approach has allowed us to uncover some subtle, and yet informative, differences among the *Su(var)3-9* mutants. Indeed, this thesis represents one of a few cases where the function of mutant alleles of a gene have been assayed in a wide variety of contexts. The results presented indicate that in many cases, the

effect(s) of a mutation on one particular phenotype do not accurately predict its effect(s) on a different phenotype. This may be true not only for *Su(var)3-9*, but also for other genes encoding chromatin and, most likely, other proteins. I suggest that this may be the case especially for those proteins that function as components of one or more multimeric complexes. Hence, it should be kept in mind that studies relying on only one or two phenotypic assays may give a biased, and possibly misleading idea of the function of the gene/gene product in question.

4.2. The function(s) of SU(VAR)3-9 in PEV and in the regulation of the HIS-C.

In this and the following sections I will try to briefly relate the role(s) and function(s) of SU(VAR)3-9 in the regulation of the HIS-C and in PEV, as well as some of the limitations of our experimental setup.

4.2.1. The *white* gene in the w^{m4} variegating rearrangement and the HIS-C.

In chapter 3 we set out to address the role of SU(VAR)3-9 in the regulation of one of its natural targets, the *histone* genes, and to compare it to what we know about its functions in heterochromatin, represented here by the w^{m4} variegating rearrangement. There is a valid rationale behind the choice of these two loci, but one should also keep in mind that the HIS-C and the *white* gene in the w^{m4} strain are only single examples of a euchromatic locus and a heterochromatin-induced variegating rearrangement, respectively, and as such may not be representative of what typically occurs in euchromatin and heterochromatin, respectively. The *white* locus in the w^{m4} variegating rearrangement has many features in common

with heterochromatin, for instance, it is enriched in heterochromatin marks such as H3K9me2, HP1 and HDAC1 (Rudolph *et al.,* 2007; Mottus, personal communication). However, the very fact that it shows variegated expression suggests that, at least in some nuclei, it is transcriptionally competent, and therefore presumably has an "open" chromatin structure. Moreover, in PEV the variegating gene does not reside in heterochromatin *per se*, but rather in *close proximity to* a disrupted heterochromatic region, and this distance varies depending on the *white*$^+$ rearrangement in question (Tartof *et al*., 1989). Hence, one should not automatically extrapolate all that applies to the role of SU(VAR)3-9 in the context of the *white* gene in the w^{m4} rearrangement to heterochromatin in general.

The case of the HIS-C is the opposite in that, while it is definitely a euchromatic locus, the HIS-C also has several characteristics that are perhaps more typical of heterochromatin: it is a reiterated locus, somewhat late-replicating, mildly underreplicated in polytene chromosomes, and it is organized into a higher order chromatin structure (Samal *et al.,* 1981; Ner *et al.,* 2002; Zhimulev and Belyaeva, 2003). For these reasons the HIS-C, albeit euchromatic, may not be representative of all euchromatic genes, especially single copy genes, and it would not be surprising if SU(VAR)3-9 acted through slightly different mechanisms and the recruitment of slightly distinct factors, in the regulation of other euchromatic loci.

4.2.2. <u>Considerations regarding the assays used to assess the effect of *Su(var)3-9* mutants on gene expression.</u>

For our purposes, the relative proportion of pigmented (red) *vs.* white pigment cells in flies carrying the w^{m4} variegating rearrangement can be considered as the readout of an assay for gene expression. In this sense, the PEV assay employed in chapter 2 is not substantially different from the RT-PCR approach used to determine the relative levels of *histone* transcripts in chapter 3. However, each of the two assays has its advantages and disadvantages. The RT-PCR approach is a direct assay of transcript accumulation, as it allows us to measure the relative amounts of transcripts of interest. Obviously, infer that an elevated transcript level reflects an increase in expression. However, we have not ruled out the possibility that the elevated transcript levels result from an increase in RNA stability, or a decrease in RNA turnover.

In contrast, the PEV assay is based on a tertiary phenotype, eye pigmentation, which is the result of the coordinated action of several factors, including events and proteins that are not related to the expression of the *white* gene *per se*. This assay has, however, one big advantage over the RT-PCR approach. Since the phenotype of each eye is individually recorded, it is possible to observe and document the phenotypic variability within each strain. This allows for a more accurate and, possibly, more relevant comparison among strains. For example, we were able to notice that the presence of the *Su(var)3-9^{06}* mutant allele in a strain that is hyperploid for *Su(var)3-9* (2 wild-type copies of *Su(var)3-9*, plus the mutant allele *Su(var)3-9^{06}*) results in a surprisingly high

frequency of "mildly suppressed" eyes (chapter 2, Figure 2.6.b), and that the presence of $Su(var)3-9^{318}$, instead of $Su(var)3-9^{06}$, in the same hyperploid strain often causes the eyes to be "sectored" (large, well-defined regions of red and large, well-defined regions of white). Such observations would be impossible with the approach employed to assess the levels of *histone* transcripts, since each RNA preparation used was derived from a population of embryos, which invariably contain cells that are in different stages of the cell cycle. Hence, it is formally possible, although probably highly unlikely, that embryos of a certain genotype could have a great variation in their levels of *histone* transcripts, with some showing strong overexpression and others being indistinguishable from the wild type, while those of a different genotype could all be overexpressing the *histone* genes. In both cases, the RT-PCR assay would show a moderate increase in the transcripts levels.

4.2.3. <u>Correlations between the effects of the *Su(var)3-9* mutants on PEV and on the levels of *histone* transcripts.</u>

Keeping in mind the respective limitations of the techniques used, if the function of SU(VAR)3-9 within the HIS-C is the same as in the w^{m4} variegating rearrangement, and if it acts through the same mechanism at both loci, then any given mutation in *Su(var)3-9* should affect the expression of the *white* gene in w^{m4}, and the *histone* genes transcription similarly. Thus, some general predictions can be made. Firstly, a screen for dominant mutants that result in overexpression of the *histone* genes should lead to the isolation of the same

181

Su(var)3-9 alleles that were recovered in the original screen for strong, dominant suppressors of PEV, as well as mutations in other loci. Secondly, all *Su(var)3-9* alleles that strongly suppress PEV should also show elevated levels of *histone* transcripts. Finally, mutations in *Su(var)3-9* that do not cause suppression of PEV should not result in elevated levels of *histone* transcripts.

An extension of this hypothesis predicts that, for each *Su(var)3-9* mutation, there is a correlation between overproduction of *histone* transcripts and the strength of Su(var) phenotype. That is, an allele that shows a very high increase in the level of *H2A* and *H3* transcripts should also be a very strong suppressor of PEV. Conversely, alleles that are weaker suppressors of PEV are expected to show a smaller elevation in the abundance of *H2A* and *H3* RNAs. In the present case, *Su(var)3-9^309^* was the weakest suppressor in our PEV assay (chapter 2), and it also showed the most limited increase in the levels of *H2A* and *H3* transcripts (only ~1.6-1.8 fold higher than the wild-type, see chapter 3).

The *H2A* and *H3* transcripts are significantly more abundant in *Su(var)3-9^06^* and *Su(var)3-9^318^* embryos (~4.5 to 6-fold higher than wild-type, see chapter 3), and both *Su(var)3-9^06^* and *Su(var)3-9^318^* displayed a stronger Su(var) phenotype than *Su(var)3-9^309^*. However, while no significant differences are detectable between *Su(var)3-9^06^* and *Su(var)3-9^318^* at the level of *histone* transcripts, *Su(var)3-9^318^* is a stronger Su(var) than *Su(var)3-9^06^* (chapters 2 and 3, respectively). This apparent discrepancy may be a result of the experimental systems used (discussed above), or may reflect a difference in the role of SU(VAR)3-9 in the regulation of the two loci, or in their respective sensitivity to

the presence of a malfunctioning SU(VAR)3-9 (as in *Su(var)3-9^{318}*) *versus* no SU(VAR)3-9 at all (as in *Su(var)3-9^{06}*).

4.2.4. The *Su(var)3-9^{330}* mutation appears to be a special case.

The *Su(var)3-9^{330}* mutant shows peculiar phenotypes in nearly every assay that we employed. In the *in vitro* enzyme assay this strain showed a unique characteristic. Its HMTase activity on full length histone H3 is higher than that of any other mutant SU(VAR)3-9 protein tested (>30% of wild-type), but it fails to function if a peptide representing the H3 N-terminal tail (20 amino acids) is used as the substrate. *Su(var)3-9^{330}* is also peculiar in that the levels of H3K9me2 associated with the *histone unit*, as determined by ChIP, are comparable to those observed in *Su(var)3-9^{06}*, yet the relative amount of HP1 associated with the same region is significantly lower than in *Su(var)3-9^{06}*. In fact, *Su(var)3-9^{330}* has the lowest level of *his* unit-associated HP1 among the mutants tested, suggesting that it may be acting as an antimorph. With respect to suppression of PEV, the strength of suppression of the *Su(var)3-9^{330}* mutant is similar to that of *Su(var)3-9^{318}* (although *Su(var)3-9^{330}* individuals do not show "sectored" eyes).

The most interesting feature of this mutant, however, is that the expression of its *h2a/h2b* and *h3 histone* genes appears to be dramatically decoupled, to the point where the relative abundance of the *h3* transcript is 6 fold higher than in wild-type (one of the strongest hyperexpression phenotypes observed), while the levels of *h2a* and *h2b* are only about 1.7-1.8 fold above the wild-type (Table 4.1). This is a very unique, and intriguing feature. It may be

worth pointing out that there are presently no data formally demonstrating that the *H2A*/*H2B*, and the *H3/H4* transcripts, are expressed in equal amounts in a wild type *Drosophila* (we are currently examining this). The data obtained with the *Su(var)3-9*[330] mutant strain suggest that there is some degree of independence between the regulation of *h2a/h2b* and *h3* expression, and that there may be a mechanism in place to ensure co-ordinated expression.
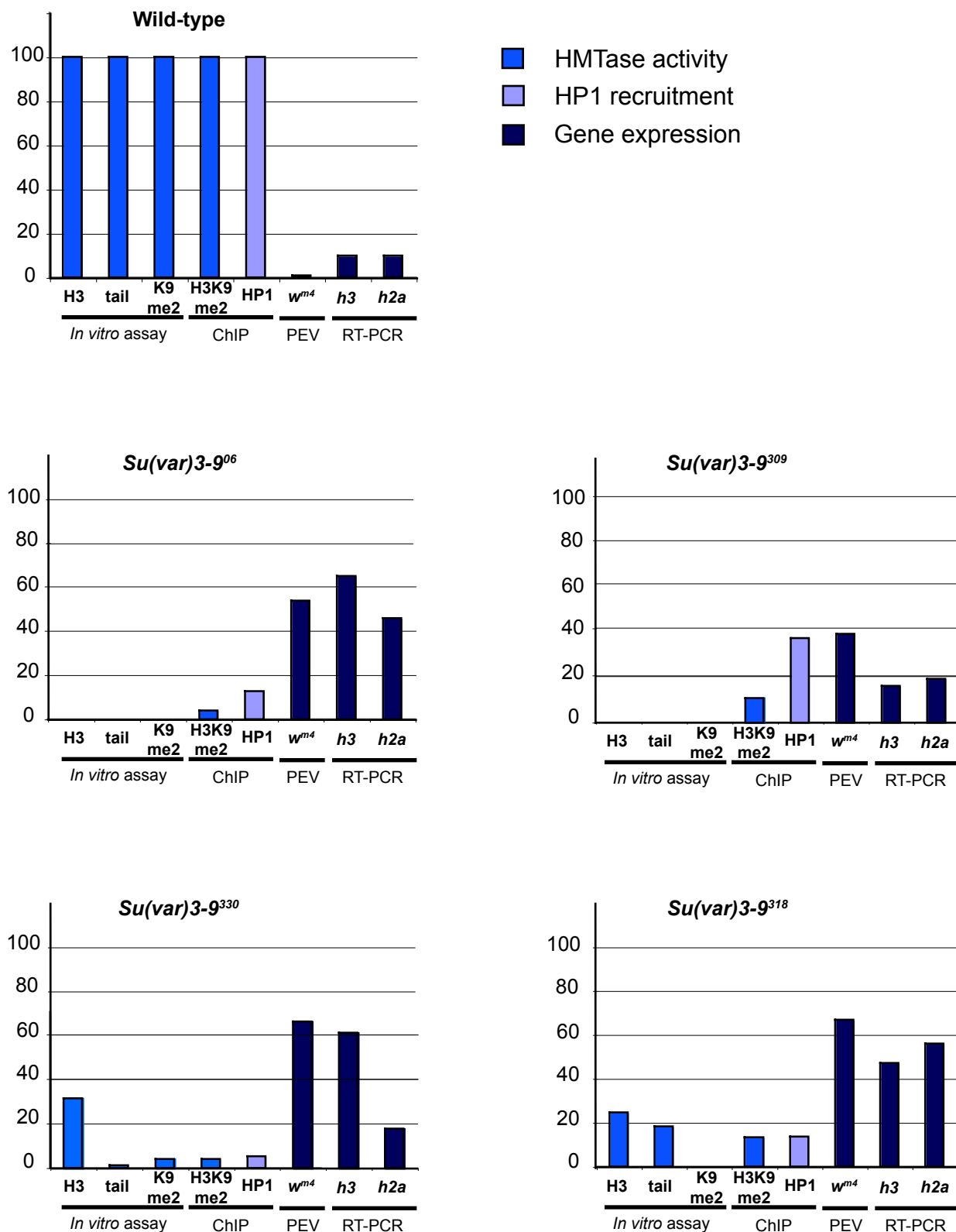
Figure 4.1. A semi-quantitative summary of the phenotypes of four *Su(var)3-9* mutants analyzed.

Figure 4.1. A semi-quantitative summary of the phenotypes of four

Su(var)3-9 **mutants analyzed.** The results presented in chapters 2 and 3

relative to *in vitro* HMTase activities of recombinant SU(VAR)3-9 proteins,

association of H3K9me2 and HP1 with the HIS-C, suppression of PEV in

Su(var)3-9 hyperploid strains and abundance of *h3* and *h2a* transcripts were

simplified and compiled in order to give a general "overview" of each mutant's

phenotype. For the *in vitro* enzyme assay, only the averages of 3 independent

trials are indicated. The same is true for the ChIP data reported. Such data refer

to the enrichment for "H3 coding" fragment in the material IPed with anti-

H3K9me2 and anti-HP1, respectively. Note that no data regarding the relative

enrichment for H3K9me2 or HP1 within the chromocentre is reported here.

For the abundance of *h3* and *h2a* transcripts, the wild-type strain was arbitrarily

assigned the value of 10, and the mutants are expressed as proportions of the

wild type. For the *in vitro* assays, the activity of the wild-type strain was always

arbitrarily assigned a value of 100. The ChIP data is expressed in each case as

% of the input, and the PEV value represent the % of fly eyes showing a strong

Su(var) phenotype in the assay described in chapter 2.

4.2.5. <u>A possible mode of action for SU(VAR)3-9</u>

The genetic screens for dominant Su(var)s were originally designed as a way to identify non-histone chromatin proteins (NHCPs). The term NHCP generally describes proteins that are either relatively stable components of chromatin, or that act as chromatin-modifying enzymes. There are of course proteins that fit both descriptions: SU(VAR)3-9, for example, is both a chromatin-modifier and a chromatin-associated protein (Rea *et al.,* 2000; Nakayama *et al.,* 2001; Schotta *et al.,* 2002; Ebert *et al.,* 2004; 2006). Given these two functions, it is easy to hypothesize how SU(VAR)3-9 is involved in the variegated, epigenetic silencing of the *white* gene in the w^{m4} rearrangement: it is recruited to centromeric and pericentric chromatin, where it methylates H3K9, thus allowing the recruitment and binding of HP1, a crucial step in the "heterochromatinisation" of the region. Therefore, the absence of functional SU(VAR)3-9 will result in a drastic reduction on H3K9me2 within pericentric chromatin; consequently low levels of HP1 are recruited and/or bound to the region (Rudolph *et al*., 2007; Mottus, personal communication). This will prevent or reduce the rate at which heterochromatin is assembled in most cells, leading to the Su(var) phenotype.

However, we now know that *Su(var)3-9* mutations that suppress PEV cause increased levels of *histone* transcripts (chapter 3 and Ner *et al*., 2002). We also know that strains that are hemizygous for the HIS-C have a Su(var) phenotype (Moore *et al*., 1979; 1983) and surprisingly these strains also produce increased levels of *histone* transcripts (Ner *et al*., 2002). Therefore, it is possible that the Su(var) phenotype observed in w^{m4}*; Su(var)3-9* mutant individuals is due

to the combination of a lack of methylation of H3K9 within pericentric chromatin, where the variegating *white* gene is located, and an increase in the amount of *histone* transcripts produced.

It has previously been shown that individuals with elevated levels of *histone* gene transcripts also have an increased level of nucleus-associated histone proteins (Ner *et al.,* 2002). We hypothesize that, since the euchromatic portion of the genome generally replicates earlier than the heterochromatic regions, the increased availability of histone proteins may give the euchromatic portion of the genome a competitive advantage over the later replicating heterochromatin. Therefore, in *Su(var)3-9* mutants the formation of heterochromatin would be impeded in two ways. Firstly, by the failure to methylate H3K9 and secondly, by an excess of histone proteins that may allow the euchromatin to replicate slightly more rapidly or more completely, thus interfering with heterochromatin assembly.

It will be interesting to investigate whether other *Su(var)*s also have elevated levels of *histone* transcripts, and how their levels of *histone* transcripts correlate with the strength of their Su(var) phenotype, and with that of individuals that are hemizygous for the HIS-C. Our hypothesis would predict that 1) all *Su(var)* mutations that result in increased levels of *histone* transcripts should be stronger suppressors of PEV than the lack of one copy of the HIS-C, and 2) hemizygosity for the HIS-C and *Su(var)* mutations should have an additive effect on the Su(var) phenotype, particularly if the *Su(var)* mutation in question does not result in increased levels of *histone* gene products.

Table 4.1. Summary of the characteristics of the *Su(var)3-9* mutants analyzed. All numerical data reported represent the average of 3 independent experiments (for details, see corresponding sections in chapters 2 and 3). The detection of H3K9me2 and HP1 associated with the chromocentre was carried out by immunofluorescence on polytene nuclei; which only provided qualitative data (expressed as "strong" vs. "reduced"). No HMTase assays were run for mutant *Su(var)3-9^06* because such allele results in no gene product (see chapter 2). *Su(var)3-9^311* was not chosen for the HIS-C and *histone* transcripts analysis (NT=not tested).

Allele	Result of mutation	HMTase function					HP1 recruitment/ association		Loss of silencing/ overexpression		
		In vitro (enzyme assays)			*In vivo* (H3K9me2)				*white* in *w^m4*	*histone RNAs*	
		rH3	H3 tail	K9me tail	Chromo-centre	HIS-C (H3 coding)	Chromo-centre	HIS-C (H3 coding)		h3	h2a
+	N/A	100%	100%	100%	strong	100%	strong	100%	<1%	1	1
06	No product	N/A	N/A	N/A	reduced	3.9%	reduced	12.7%	53.5%	6.46	4.56
309	C462Y	0%	0%	0%	reduced	10.4%	reduced	35.9%	37.7%	1.56	1.86
330	D536N	31.3%	1.2%	4.0%	reduced	4.0%	reduced	5.3%	65.68%	6.07	1.76
318	S616L	24.7%	18.3%	0%	reduced	13.3%	reduced	13.7%	66.76%	4.72	5.60
311	G521D	22.7%	14.3%	11.5%	reduced	NT	reduced	NT	86.20%	NT	NT

4.3. *Drosophila* SU(VAR)3-9 has a unique N-terminal region

Like many *Su(var)* genes, *Su(var)3-9* is highly conserved from yeast to mammals. Therefore, one might infer that the basic function of the protein and our findings in *Drosophila* are transferable to other organisms. However, *Drosophila*'s SU(VAR)3-9 is larger than its homologs, and it contains a "unique" N-terminal region. This region is comprised of the 82 amino acids that are in common with (eIF2γ common domain) and a region, sometimes called "region 2" (residues 83-219), that is required for the dimerization and full enzymatic function of SU(VAR)3-9, as well as for the interactions with SU(VAR)3-7 and HP1 (Krauss and Reuter, 2000; Schotta *et a*l., 2002; Eskeland *et al.*, 2004; Krauss *et al.*, 2006).

The lack of these N-terminal domains in the yeast and vertebrate homologs of SU(VAR)3-9 suggest that they may function slightly differently from the *Drosophila* protein. Neither the vertebrate nor the yeast homologs have the regions required for dimerisation, implying that they may not need to dimerize in order to be fully functional. The fact that they also lack the region of interaction with SU(VAR)3-7 should not be surprising, since SU(VAR)3-7 does not exist outside the *Drosophila* genus (Jaquet *et al.*, 2006), where it is required for proper localization of SU(VAR)3-9 (Delattre *et al.*, 2004). In the fission yeast this function is carried out, at least in part, by an RNAi-based mechanism (Bühler *et al.*, 2006; Zhang *et al.*, 2008). In spite of its lack of region 2, the mammalian homolog of SU(VAR)3-9 can interact with HP1 (Aagaard *et al.*, 1999; Melcher *et*

al., 2000). This interaction occurs through the first 44 amino acids of Suv39h1 (Melcher *et al*., 2000), which correspond to the last 44 residues of SU(VAR)3-9's region 2. In addition, SUV39H1 (the human homolog) can partially rescue the dominant Su(var) phenotype of *Su(var)3-9* mutants (Schotta *et al.*, 2002). A physical interaction between *S. pombe*'s Clr4 and Swi6 (the homolog of HP1) has not been formally demonstrated, and in any case it would be unlikely to occur through the N-terminal region of Clr4, since this portion of the protein is only 7 amino acids long.

Curiously, although the N-terminal region of SU(VAR)3-9 has such important functions, no single amino acid substitutions within this portion of the protein have been isolated in the screens for dominant suppressors of PEV (Sinclair *et al*., 1983; Tschiersch *et al*., 1994; Ebert *et al*., 2004). As in the case of the chromodomain (see discussion in chapter 2), this suggests that mutating any single amino acid may not be sufficient to cause a dominant Su(var) phenotype. In order to dissect the function of region 2, it would be interesting to create transgenic *Drosophila* lines that express tagged SU(VAR)3-9 proteins with single and multiple amino acid substitutions within this region, and investigate the targeting and chromatin association of the mutant proteins, as well as their effects on the targeting of the endogenous (wild-type) SU(VAR)3-9 and some of its known partners, HP1 and SU(VAR)3-7. We would not expect the single amino acid substitutions to have any effects on the behaviour of the endogenous protein. The same experiment, carried out in a *Su(var)3-9* null background, would allow us to determine whether any missense mutations in region 2 result in a

recessive suppression of PEV phenotype, and what their effect on *in vivo*

H3K9me2,3 is.

4.4. References

Delattre, M., Spierer, A., Jaquet, Y., Spierer, P.: Increased expression of *Drosophila Su(var)3-7* triggers Su(var)3-9-dependent heterochromatin formation. *J Cell Sci* 117, 6239-47 (2004)

Ebert, A., Schotta, G., Lein, S., Kubicek, S., Krauss, V., Jenuwein, T., Reuter, G.: *Su(var)* genes regulate the balance between euchromatin and heterochromatin in *Drosophila.* Genes Dev. *18*, 2973-83 (2004)

Ebert, A., Lein, S., Schotta, G., Reuter, G.: Histone modification and the control of heterochromatic gene silencing in *Drosophila*. Chromosome Res. *14*, 377-92 (2006)

Greil, F., van der Kraan, I., Delrow, J., Smothers, J.F., de Wit, E., Bussemaker, H.J., van Driel, R., Henikoff, S., van Steensel, B.: Distinct HP1 and Su(var)3-9 complexes bind to sets of developmentally coexpressed genes depending on chromosomal location. Genes Dev. *17*, 2825-38 (2003)

Jaquet, Y., Delattre, M., Montoya-Burgos, J., Spierer, A., Spierer, P.: Conserved domains control heterochromatin localization and silencing properties of SU(VAR)3-7. *Chromosoma* 115, :139-50 (2006)

Krauss, V., Reuter, G.: Two genes become one: the genes encoding heterochromatin protein Su(var)3-9 and translation initiation factor subunit eIF-2gamma are joined to a dicistronic unit in holometabolic insects. *Genetics* 156,1157-67 (2000)

Krauss V, Fassl A, Fiebig P, Patties I, Sass H.: The evolution of the histone methyltransferase gene Su(var)3-9 in metazoans includes a fusion with and a re-fission from a functionally unrelated gene. *BMC Evol Biol* 6/18 (2006)

Min, J., Zhang, X., Cheng, X., Grewal, S.I., Xu, R.M.: Structure of the SET domain histone lysine methyltransferase Clr4. Nat Struct Biol. *9*, 828-32 (2002)

Moore, G.D., Procunier, J.D., Cross, D.P., Grigliatti, T.A.: Histone gene deficiencies and position-effect variegation in *Drosophila.* Nature *282*, 312-4 (1979)

Moore, G.D., Sinclair, D.A., Grigliatti, T.A.: Histone gene multiplicity and position effect variegation in *Drosophila melanogaster.* Genetics *105*, 327-44 (1983)

Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D., Grewal, S.I.: Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science *292*, 110-3 (2001)

Ner, S.S., Harrington, M.J., Grigliatti, T.A.: A role for the *Drosophila* SU(VAR)3-9 protein in chromatin organization at the histone gene cluster and in suppression of position-effect variegation. Genetics *162*, 1763-74 (2002)

Nielsen, S.J., Schneider, R., Bauer, U.M., Bannister, A.J., Morrison, A., O'Carroll, D., Firestein, R., Cleary, M., Jenuwein, T., Herrera, R.E., Kouzarides, T.: Rb targets histone H3 methylation and HP1 to promoters. Nature *412,* 561-5 (2001)

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., Jenuwein, T.: Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593-9 (2000)

Rudolph, T., Yonezawa, M., Lein, S., Heidrich, K., Kubicek, S., Schafer, C., Phalke, S., Walther, M., Schmidt, A., Jenuwein, T., Reuter, G.: Heterochromatin formation in *Drosophila* is initiated through active removal of H3K4 methylation by the LSD1 homolog SU(VAR)3-3. Mol Cell. *26*, 103-15 (2007)

Samal, B., Worcel, A., Louis, C., Schedl, P.: Chromatin structure of the histone genes of *D. melanogaster*. Cell *23*, 401-9 (1981)

Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R., Reuter, G.: Central role of *Drosophila* SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. EMBO J. *21*, 1121-31 (2002)

Sinclair, D.A.R., Mottus, R.C., Grigliatti, T.A.: Genes which suppress position effect variegation in *Drosophila melanogaster* are clustered. Molec. gen. Genet. *191***,** 326-33 (1983)

Tartof, K.D., Bishop C, Jones M, Hobbs CA, Locke J.: Towards an understanding of position effect variegation. Dev. Genet. *10*, 162-76 (1989)

Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G., Reuter, G.: The protein encoded by the Drosophila position-effect variegation suppressor gene *Su(var)3-9* combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J*. 13, 3822-31 (1994)

Zhimulev, I.F., Belyaeva, E.S.: Intercalary heterochromatin and genetic silencing. Bioessays *25*, :1040-51 (2003)

Appendix 1: ChIP with commercial non-specific IgG, and "SN1 pre-immune" (Ner *et al.*, in preparation) serum.

(|Critical value| for a two-sided, unpaired T-test at p=0.05: 2.78)

Strain: WT
Region: H3c
Antibody: non-specific IgG (CALBIOCHEM NI01, control for H3K9me2 IPs)

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.102	0.101	0.004	0.0367	0.026	0.0136	9.20
2	0.105			0.0306			
3	0.097			0.0107			

Strain: WT
Region: H2Ac
Antibody: "SN1 pre-immune serum" (control for HP1 IPs)

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.1642	0.139	0.022	0.0779	0.069	0.012	4.85
2	0.1306			0.0742			
3	0.1343			0.0550			

Comparison between "control IPs" and corresponding "specific IPs"

	Control Ab-NO Ab (average)	Specific Ab-NO Ab (average, in % input)	Significantly different?
H3K9me2	0.097	9.154	t=9.67; YES at p=0.05
HP1	0.142	0.205	t=4.72; YES at p=0.05

Appendix 2: ChIP with anti-SU(VAR)3-9 (α-SU(VAR)3-9[chr]; Ner *et al.*, 2002).

(|Critical value| for a two-sided, unpaired T-test at p=0.05: 2.78)

Strain: WT
Region: H3c

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.0652			0.0104			
2	0.0413	0.0504	0.0128	0.0079	0.0103	0.0023	5.31
3	0.0449			0.0126			

Strain: 06
Region: H3c

IP#	Ab(%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.0174			0.0232			
2	0.0160	0.0177	0.0019	0.0138	0.0173	0.0051	0.117
3	0.0198			0.0151			

Strain: 309
Region: H3c

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.0311			0.0128			
2	0.0405	0.0345	0.0051	0.0118	0.0113	0.0017	7.35
3	0.0320			0.0095			

Strain: 330
Region: H3c

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.0383			0.0127			
2	0.0254	0.0358	0.0094	0.0100	0.0121	0.0018	4.26
3	0.0439			0.0136			

Strain: 318
Region: H3c

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.0874			0.0244			
2	0.0530	0.0653	0.0191	0.0251	0.0247	0.0004	3.67
3	0.0556			0.0246			

Comparison of mutants to WT

Mutant	Avg-avg NO	Ratio to WT	Significantly different from WT?
06	0.0004 (%)	0.007	t= 4.99; YES at p=0.05
309	0.0232 (%)	0.578	t= 2.15; NO at p=0.05
330	0.0237 (%)	0.591	t= 1.92; NO at p=0.05
318	0.0406 (%)	1.012	t= -0.035; NO at p=0.05

Appendix 3: ChIP with anti-H3K9me2 (UPSTATE#07-441)

(|Critical value| for a two-sided, unpaired T-test at p=0.05: 2.78)

Strain: WT
Region: H3c

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	7.474	9.232	1.63	0.058	0.078	0.045	9.67
2	10.71			0.130			
3	9.512			0.046			

Strain: 06
Region: H3c

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.323	0.378	0.079	0.030	0.013	0.015	7.81
2	0.343			0.002			
3	0.470			0.007			

Strain: 309
Region: H3c

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	1.707	1.707	0.047	1.052	0.754	0.375	4.37
2	1.660			0.878			
3	1.755			0.333			

Strain: 330
Region: H3c

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.630	0.480	0.178	0.114	0.106	0.027	3.60
2	0.283			0.128			
3	0.526			0.075			

Strain: 318
Region: H3c

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	1.615	1.21	0.410	0.049	0.088	0.076	4.67
2	1.224			0.039			
3	0.796			0.177			

Comparison of mutants to WT

Mutant	Avg-avg NO	Compared to WT	Significantly different from WT?
06	0.365 (%)	0.039	t= 9.36; YES at p=0.05
309	0.953 (%)	0.104	t= 7.96; YES at p=0.05
330	0.374 (%)	0.040	t= 9.21; YES at p=0.05
318	1.122 (%)	0.133	t= 8.24; YES at p=0.05

Strain: WT
Region: H3/H4 intergenic

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	1.208			0.038			
2	1.226	1.179	0.065	0.048	0.050	0.014	29.0
3	1.104			0.066			

Strain: 330
Region: H3/H4 intergenic

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.922			0.008			
2	0.836	0.871	0.045	0.010	0.009	0.001	32.9
3	0.854			0.011			

Strain: WT
Region: H2A/H2B intergenic

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	1.640			0.012			
2	1.134	1.350	0.261	0.006	0.006	0.005	8.91
3	1.276			0.002			

Strain: 330
Region: H2A/H2B intergenic

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.314			0.024			
2	0.354	0.315	0.038	0.006	0.013	0.009	13.4
3	0.278			0.011			

Comparison of 330 to WT for H3/H4 intergenic region

Strain	Avg-avg NO	Compared to WT	Significantly different from WT?
WT	1.13 (%)	N/A	N/A
330	0.862 (%)	0.762	t= 5.07; YES at p=0.05

Comparison of 330 to WT for H2A/H2B intergenic region

Strain	Avg-avg NO	Compared to WT	Significantly different from WT?
WT	1.344 (%)	N/A	N/A
330	0.304 (%)	0.226	t= 6.93; YES at p=0.05

Appendix 4: ChIP with anti-HP1 (α-HP1; Ner *et al.,* in preparation).
(|Critical value| for a two-sided, unpaired T-test at p=0.05: 2.78)

Strain: WT
Region: H2Ac

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.2179	0.234	0.0548	0.0461	0.0282	0.0164	11.4
2	0.2194			0.0137			
3	0.2646			0.0252			

Strain: 06
Region: H2Ac

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.0942	0.078	0.0146	0.0207	0.0162	0.0040	7.06
2	0.0652			0.0136			
3	0.0754			0.0149			

Strain: 309
Region: H2Ac

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.0553	0.066	0.0174	0.0082	0.011	0.0040	5.36
2	0.0568			0.0088			
3	0.0862			0.0154			

Strain: 330
Region: H2Ac

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.0230	0.020	0.0040	0.0084	0.0048	0.0032	5.10
2	0.0156			0.0018			
3	0.0224			0.0040			

Strain: 318
Region: H2Ac

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.0781	0.0804	0.0024	0.0342	0.0242	0.0086	10.8
2	0.0824			0.0207			
3	0.0808			0.0178			

Comparison of mutants to WT

Mutant	Avg-avg NO	Ratio to WT	Significantly different from WT?
06	0.0618	0.3002	t= 8.89; YES at p=0.05
309	0.0550	0.2682	t= 9.16; YES at p=0.05
330	0.0152	0.0758	t= 13.8; YES at p=0.05
318	0.0562	0.2730	t= 9.99; YES at p=0.05

Strain: WT
Region: H3c

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.3238	0.2750	0.0472	0.0061	0.0255	0.0168	8.60
2	0.2705			0.0352			
3	0.2295			0.0352			

Strain: 06
Region: H3c

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.0426	0.0351	0.0066	0.0089	0.0061	0.0025	7.03
2	0.0339			0.0039			
3	0.0292			0.0055			

Strain: 309
Region: H3c

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.1098	0.1150	0.0074	0.0230	0.0253	0.0029	19.4
2	0.1238			0.0244			
3	0.1123			0.0286			

Strain: 330
Region: H3c

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.0208	0.0187	0.0030	0.0084	0.0066	0.0020	5.86
2	0.0153			0.0070			
3	0.0200			0.0045			

Strain: 318
Region: H3c

IP#	Ab (%IN)	Avg	SD	NO (%IN)	Avg	SD	t
1	0.0473	0.0487	0.0013	0.0204	0.0144	0.0052	10.9
2	0.0500			0.0123			
3	0.0489			0.0105			

Comparison of mutants to WT

Mutant	Avg-avg NO	Ratio to WT	Significantly different from WT?
06	0.0290	0.1273	t= 6.04; YES at p=0.05
309	0.0897	0.3595	t= 4.35; YES at p=0.05
330	0.0121	0.0531	t= 6.52; YES at p=0.05
318	0.0343	0.137	t= 18.0; YES at p=0.05

Appendix 5: relative quantifications of *histone* transcripts by RT-PCR.
Comparison between the relative abundance of total *h2a* and *h3* transcripts in
wild-type *vs. Su(var)3-9* mutant embryo extracts.
(|Critical value| for a two-sided, unpaired T-test at p=0.05: 2.78)

Exp #	Strain	*h2a/rp49* in mut. / *h2a/rp49* in WT	Avg	St. dev.	t* (mutant/WT)
1		(1)			
2	WT	(1)	(1)	N/A	N/A
3		(1)			
1		5.01			
2	06	4.00	4.56	0.51	-12.0 √
3		4.67			
1		1.53			
2	309	1.58	1.86	0.53	-2.82 √
3		2.47			
1		2.21			
2	330	1.87	1.76	0.52	-2.56 X
3		1.20			
1		5.17			
2	318	5.22	5.60	0.72	-11.0 √
3		6.43			

Exp #	Strain	*h3/rp49* in mut. / *h3/rp49* in WT	Avg	St. dev.	t*
1		(1)			
2	WT	(1)	(1)	N/A	N/A
3		(1)			
1		6.70			
2	06	5.85	6.46	0.53	-17.9 √
3		6.82			
1		1.71			
2	309	1.40	1.56	0.15	-6.25 √
3		1.57			
1		6.03			
2	330	6.58	6.07	0.50	-17.7 √
3		5.59			
1		6.24			
2	318	3.56	4.72	1.38	-4.67 √
3		4.35			

Appendix 6.

Comparison between the relative increase in *h2a* and *h3* transcripts detected in the *Su(var)3-9* mutants. The data presented in appendix 5 were compared using a Student's T test.

(|Critical value| for a two-sided, unpaired T-test at p=0.05: 2.78)

Strain	Relative increase *h2a*	Relative increase *h3*	Statistically different?
06	4.56 +/- 0.51	6.46 +/- 0.53	YES at p=0.05; t=-4.46
309	1.86 +/- 0.53	1.56 +/- 0.15	NO at p=0.05; t=0.943
330	1.76 +/- 0.52	6.07 +/- 0.50	YES at p=0.05; t=-10.5
318	5.60 +/- 0.72	4.72 +/- 1.38	NO at p=0.05; t=0.994

Appendix 7.

Comparison between the relative abundance of total *h2b* transcript in wild-type *vs. Su(var)3-9*[330] embryo extracts. The indicated t* refers to the result of an unpaired T-test in which the *h2b/rp49* ratio of each mutant was compared to that of the wild-type strain.

(|Critical value| for a two-sided, unpaired T-test at p=0.05: 2.78)

Exp #	Strain	$\dfrac{\textit{h2b/rp49 in mut.}}{\textit{h2b/rp49 in WT}}$	Avg	St. dev.	t*
1	WT	(1)	(1)	N/A	N/A
2		(1)			
3		(1)			
1	330	2.77	1.95	0.74	-2.67 X
2		1.32			
3		1.76			