universität wien

DISSERTATION

Titel der Dissertation

Quantitative Analysis and Live Imaging of Polycomb Group Proteins in Drosophila melanogaster

angestrebter akademischer Grad

Doktor/in der Naturwissenschaften (Dr. rer.nat.)

Verfasserin / Verfasser:	Cornelia Gänger
Matrikel-Nummer:	0003240
Dissertationsgebiet (lt. Studienblatt):	Dr.-Studium der Genetik - Mikrobiologie
Betreuerin / Betreuer:	Dr. Leonie Ringrose

Wien, im März 2010

Table of contents

1 Abstract

1.1 Abstract (English)

The Polycomb(PcG) and Trithorax (TrxG) proteins ensure epigenetic memory of transcription for several hundred developmentally important genes by binding to cis regulatory elements called Polycomb/Trihorax response elements (PRE/TREs). Paradoxically, although the PcG and TrxG proteins pass the transcriptional state very stably from one cell generation to the next, this is a dynamic system in constant flux between chromatin bound and free protein pools.

For a complete understanding of this mechanism it is crucial to be able to study functional PcG repression complexes, TrxG activation complexes and complexes of DNA binding factors, which recruit other PcG/TrxG proteins, in living organisms. Advanced microscopy in combination with fluorscent protein tagging with GFP or its variants have been established as a minimally invasive tool to visualize, track and quantify proteins of interest *in vivo*. In this thesis work I have selected important members of different complexes involved in the PRE-mediated regulatory mechanism (E(Z), PHO, DSP1), and tagged them fluorescently with GFP. I quantitatively analyzed the expression level and functionality of the fusion proteins in the respective transgenic flies and their distribution in different tissues at different developmental stages.

This resulted in the identification of expression strategies that gave full rescue of mutant phenotypes for PHO and E(Z), and also sufficient levels of protein to be visualised by live imaging, thus providing a set of tools for future photobleaching studies of these proteins to investigate the behavior, stability and lifetime of functional complexes *in vivo*.

In addition, I have generated transgenic flies carrying a visually tagged transgene PRE to study GFP fusion protein kinetics at a single locus. I have established a system to follow PcG/TrxG protein binding and locus-specific behavior in the living organism.

1.2 Zusammenfassung (Deutsch)

Proteine der Polycomb- (PcG) und Trithorax- (TxG) Gruppen bewahren das epigenetische Transkriptionsgedächtis von hunderten Genen, die für die Entwicklung wichtig sind. Sie binden an cis-regulatorische DNA-elemente die man Polycomb/Trithorax Response Elements (PRE/TREs) nennt. Obwohl der Transkriptionsstatus eines durch Proteine der PcG und TrxG regulierten Gens stabil von einer Zellgeneration zur nächsten weitervererbt wird, ist der Mechanismus dynamisch und es herrscht ein ständiger Austausch zwischen freiem Protein und Protein, welches an Chromatin gebunden ist.

Um dieses System verstehen zu können ist es wichtig, funktionstüchtige Proteinkomplexe, welche an die DNA binden, die Gentranskription unterdrücken (PcG) oder aktivieren (TxG), in lebenden Organismen studieren zu können. Proteine, die mit kleinen fluoreszierenden Proteinen verbunden sind, können mittels mikroskopischer *in vivo* Methoden sichtbar, verfolgbar und quantifizierbar gemacht werden. Ich habe wichtige und representative Proteine verschiedender Komplexe ausgewählt (E(Z), PHO, DSP1), um sie mit einem grün fluoreszierenden Protein zu fusionieren. Ich habe den Expressionslevel dieser Proteine und deren Funktionalität in den jeweiligen transgenen Fliegen quantitativ analysiert und ihre Verteilung in verschiedenen Geweben observiert. Das hat zur Identifikation der Expressionsstrategien geführt, die mutante Phenotypen von E(Z) und PHO ersetzen können und ausreichend Fusionsprotein erzeugen, um mit Live Imaging Methoden studiert werden zu können. Das ermöglicht zukünftige kinetische Studien dieser Proteine, um die Stabilität und Lebensdauer von funktionstüchtigen Proteinen zu erforschen.

Zudem habe ich auch transgene Fliegen geschaffen, die ein sichtbar markiertes PRE/TRE in ihrem Genom tragen, um kinetische Studien der grün fluoreszierenden Fusionsproteine an einem konkreten DNA-element vergleichend studieren zu können. Somit habe ich ein System etabliert, das die spezifische Bindung der Proteine der PcG und TrxG im lebenden Gewebe verfolgen lässt.

2 Introduction

2.1 Chromatin and Epigenetics

Chromatin is a dynamic polymer composed of DNA, which contains the genetic information, specialized proteins, and RNA and forms a multi-layered organized structure for information storage and programmed read-out of genes.
Mutations in the DNA sequence can cause phenotypic changes that are heritable through the germline. However, also changes in covalent and non-covalent modifications of chromatin, which allow a stable but flexible and reversible propagation of specific gene expression patterns that form cellular identity, can lead to heritable traits from cell its daughter cells and also transgenerationally through the germline. Therefore, the term "epigenetics" describes changes in a phenotype (or in the state of the expression of a gene) that do not involve a mutation in the DNA, but that are nevertheless inherited in the absence of the signal that initiated the changes [1]. A variety of epigenetic phenomena have been discovered, such as the PcG and TrxG mediated transcriptional memory, which maintains correct cell identity in multicellular organisms. Hence, not only genetic but also epigenetic information is important for determining fundamental biological processes. [2]

2.2 PcG and TrxG protein complexes

Polycomb group (PcG) and Trithorax group (TrxG) proteins are essential factors in sustaining cellular identity by taking effect on chromatin. They were first identified in *Drosophila melanogaster* as genes that are required for maintaining body segment identity by controlling *hox* gene activity (reviewed in [3]). The *Hox* genes encode a highly conserved class of developmental regulators, which are important for the positioning of structures along the anterior-posterior axis, and which are expressed in a particular pattern defined by the transient segmentation gene products. [4] The first *PcG* mutations leading to sex comb anomalies in males were identified in the 1940s. These mutants showed homeotic transformations not because of mutations that affect the *Hox* gene products themselves, but the way *Hox* genes are segmentally expressed [5]. Furthermore,

they are not required for the initiation of correct *hox* gene expression, but rather to maintain their expression state after the disappearence of the initial transcription factors [6-8]. Detailed genetic and molecular studies identified two antagonizing groups of genes involved in maintaining the pattern of *Hox* gene expression: the *Polycomb* group and *Trithorax* group of genes. PcG and TrxG proteins act in large multimeric complexes to maintain repression and activation of their target genes, respectively (reviewed in [3]).

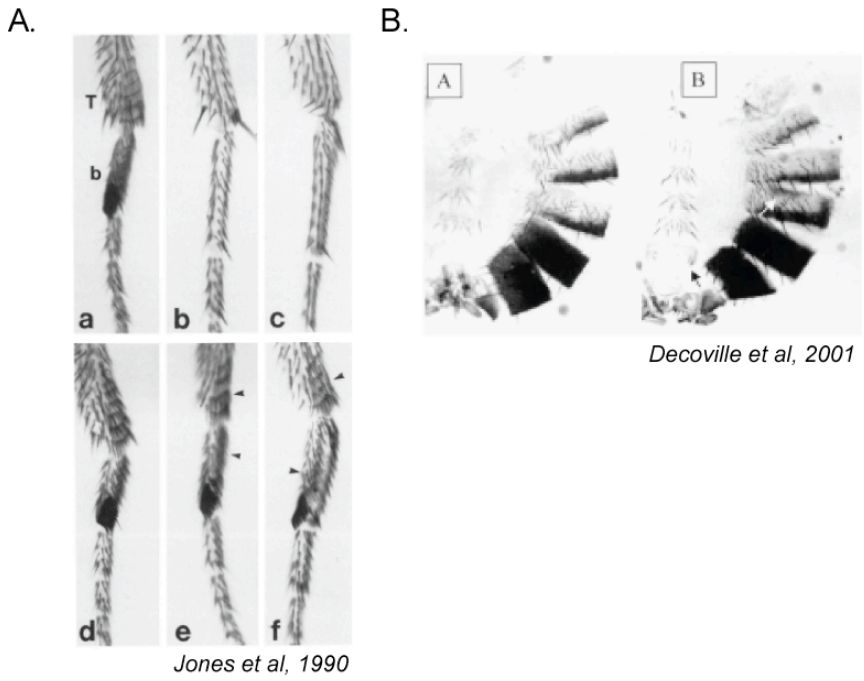

Figure 1: Typical phenotypes of PcG mutations

A. Prothoracic (a,d), mesothoracic (b,e) and metathoracic (c,f) legs of wildtype (a-c) and hemizygous mutant for *E(z)* (d-f) adult males. Meso- and metahoracic legs show anteriorly directed transformations towards the prothoracic state, recognizable by additional sex combs (e,f). This phenotype is due to ectopic expression of the *Hox* gene from the *Antennapedia* complex *Scr* (*Sex combs reduced*) in larval leg discs. [9]
B. Posteriorly directed homeotic transformation of abdominal segment A4 (white arrow) in wildtype (A) and *dsp1* mutant (B) males due to activation of *iab-5* and expression of the *Hox* gene *Ubx* from the *Bithorax* complex. [10]

2.2.1 PcG repressive complexes

Different PcG protein complexes have been described with alternative forms depending on method and material used. Biochemical purifications of *Drosophila* embryos and cells have shown that the core PcG complexes Polycomb Repressive Complex 1 (PRC1) and Polycomb Repressive Complex 2 (PRC2) act in conjunction (reviewed in [11, 12]). More

recently, the Pho Repressive Complex (PhoRC) has been described [13]. Generally, the identification of common components in PcG complex composition and their evolutionary conservation, have determined a minimal set of stably interacting proteins [14]. However, the varying components of these complexes may reflect modulated complex-composition in different tissues and at different target genes, or different states of repression or activation (reviewed in[15]).

The *Drosophila* PRC1 core complex consists of stoichiometric amounts of four proteins: PC (*Polycomb*), PH (*Polyhomeotic*), PSC (*Posterior sex combs*) and dRING/SCE (*Sex combs extra*). The core PRC1 has been shown to block chromatin remodeling *in vitro* [16]. This complex has also been purified from embryos with additional accessory proteins, notably the DNA binding protein Zeste, SCM (Sex combs on midleg), and several other components, including TBP-associated factors, implying more than one type of PRC1 [17]. Reconstitution of the repressive complex including Zeste was shown to increase the binding of the complex to DNA containing Zeste binding sites [18, 19]. PC has an N-terminal chromodomain that has been shown to bind to methylated H3K27 and H3K9, which are histone modifications linked to transcriptional repression [20, 21]. PH has two isoforms (PH-P and PH-D) and has been shown to be glycosylated by SXC/OGT (super sex combs/ O-linked N-acetylglucosamine transferase) [22]. The mammalian dRing homologue has been identified as a H2A ubiquitin ligase and ubiquitination has also been linked to PcG mediated gene silencing in *Drosophila*. [23]. SCM has been shown to have methyl-lysine binding activity and interacts with dSfmbt, which is part of the PcG protein complex PhoRC [24, 25].
The complex variant CRASCH (Chromatin associated silencing complex for homeotics) was purified from Drosophila Schneider cells and consists of PC, PH, a histonedeacetylase (HDAC1) and PSQ (*pipsqueak*), which is a PcG protein that specifically binds to specific DNA motifs [26, 27].

The core complex PRC2 is composed of the PcG proteins ESC (*Extra sex combs*), SU(Z)12 (*Suppressor (12) of Zeste*), E(Z) (*enhancer of zeste*), and p55 (a histone-binding protein) and is a complex with histone methyl lysine transferase (HMKT) activity mediated by the SET domain of E(Z), which methylates H3K27 and/or K9 [28, 29]. PRC2 protein orthologues are present in all multicellular model species. The histone deacetylase RPD3 has been reported to co-purify with PRC2 [30]. ESC serves as a platform for protein-protein interactions and needed for HMKT activity [29]. Mutation of *E(z)* in the fly

dominantly enhanced the effect of the *zeste*^{1} allele on *white* repression, produced homeotic transformations [9, 31], and was therefore defined as a PcG protein. *E(z)* encodes a 760-amino acid protein (87kDa) and contains a SET domain [32], which mediates the HMKT activity, methylating H3K27 and H3K9. The N-terminal region of E(Z) binds to ESC, which is indispensable for PcG mediated silencing [33, 34].
Additional PcG proteins like Pcl (Polycomb-like) and Scm (Sex combs on midleg) have been shown to be important for PcG silencing in association with the PRC2 and PRC1, respectively (reviewed in [15]).

The third described PcG protein complex contains PHO. PHO and its paralogue PHOL (*Pho-like*) are sequence-specific DNA binding proteins and members of the PcG. *Pho* in *Drosophila* encodes a 520 amino acid protein (58kDa) and contains a zinc-finger domain [35]. PhoRC is a complex of PHO with dSfmbt, which binds specifically to mono- and dimethylated H3K9 and H4K20 via its MBT repeats and combines DNA-targeting with histone binding [25, 36]. PHO has been purified in a second complex with dINO80, a nucleosome-remodelling complex [13]. PHO has also been shown to contact the BRM (*Brahma*) remodelling complex and interacts with the PRC2 via E(Z) and the PRC1 via PC and PH [37-39].

2.2.2 TrxG activating complexes

The TrxG of genes is a large group that encodes functionally diverse regulatory proteins, which form a variety of complexes that are important for maintaining the transcriptional active state of target genes. Many of them have been identified in screens for suppressor of *Pc* mutations [40] or because they mimic loss-of-function *Hox* mutations [41, 42]. Some proteins were classified as TrxG proteins due to sequence homology or physical association with known members of that group. Many TrxG members do not exclusively function in epigenetic maintenance, but are involved in general transcriptional processes like chromatin remodelling or being part of the transcriptional machinery.
Two of the purified TrxG protein complexes contain the ATPase BRM (*Brahma*), MOR (*moira*) and alternatively OSA or a polybromodomain protein as well as some other accessory factors [43-46]. This complex is highly related to the yeast SWI/SNF nucleosome-remodeling complex that uses the energy of ATP hydrolysis to move nucleosomes and counteracts repression. The TrxG protein KIS (*kismet*) also contains an

ATPase domain, belongs to the CHD family of chromatin-remodelling factors [41, 47] and is thought to form a distinct TrxG complex [48].

Other TrxG proteins with multiple roles are the SKD (*skuld*) and KTO (*kohtalo*) proteins, which are part of the mediator complex and link gene-specific activator proteins with the pre-initiation complex containing PolII [49]. *fsh* (*female sterile homeotic*) [50], *Tna* (*Tonalli*) [51] and *sls* (*sallimus* encoding Titin) [52] are also genes of the TrxG encoding proteins with distinct biochemical functions.

Proteins that have specialized roles for transcriptional memory are ASH1, TRX, and ASH2. Two complexes of 2-MDa and 1-MDa contain the ASH1 (*Absent small or homeotic discs 1*) and TRX (*Trithorax*) proteins respectively. Both possess a SET domain with histone methyltransferase activity and are primarily methylating H3K4, which is typically associated with active chromatin [40, 53, 54]. The SET domain of Ash1 has also been shown to bind non-coding transcripts from the *bxd* PRE [55]. The complex that contains TRX is called TAC1 and additionally includes the histone acetyltransferase dCBP and the antiphosphatase Sbf1 [56]. Another identified TrxG complex contains ASH2 [45], but homologues of TRX and ASH2 have been found together in one complex in humans [57, 58].

In summary, the TrxG complexes thus far isolated and characterized all contain enzymatic activities that help to activate transcription by modifying chromatin properties or are involved in any of the numerous steps of transcriptional activation. However, not for all components of TrxG complexes do mutations in their genes also show a TrxG phenotype.

2.2.3 DNA binding proteins recruit PcG/TrxG protein complexes

PcG and TrxG protein complexes act via Polycomb response elements (PREs), which are specific regulatory sites consisting of characteristic DNA motifs. Besides the PcG complex PhoRC and the PRC1 variant CHRASCH containing the DNA-binding proteins PHO and PSQ respectively, PcG and TrxG protein complexes that have been described do not bind the regulatory DNA elements called Polycomb Response Elements (PREs) directly. Therefore, they are recruited via more or less specialized proteins involved in PcG/TrxG-mediated “cellular memory”, which specifically bind DNA sequence motifs.

As already mentioned, PSQ [26, 59, 60] and PHO as well as PHOL (Pho-like), which has partially redundant functions to PHO [39, 61], are two of six different proteins that have been characterized as DNA-binding proteins with a role in PcG/TrxG regulation. The others are GAF (GAGA binding factor) [59, 60, 62, 63], Zeste [19] and DSP1 [64], but

these may not be the only proteins involved in PcG and TrxG protein complex targeting. For example the SP1 [65] and Grainyhead [66] DNA binding proteins have been shown to play a role at specific PREs. As there is conflicting evidence for the function of some DNA binding factors, they may bind the PRE in both repressive and active mechanisms, have different roles at different PREs or they may act redundantly. [3, 67]

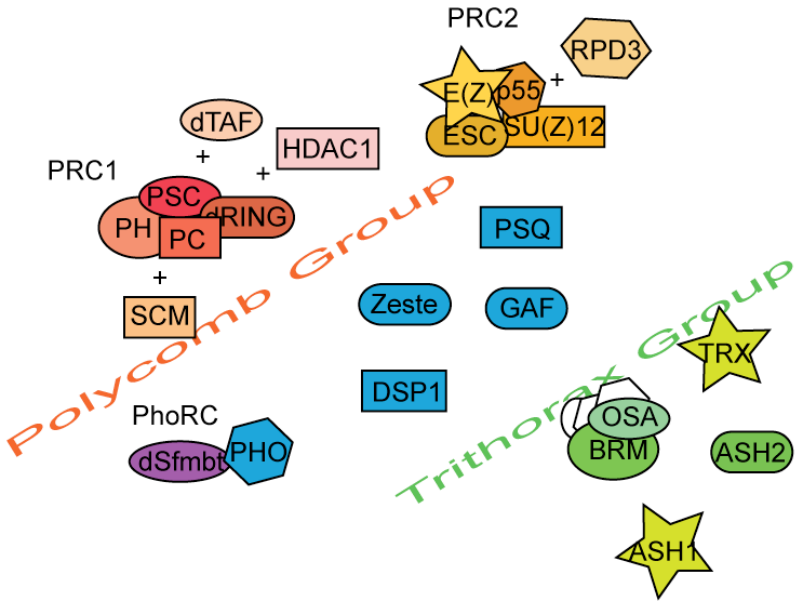

Figure 2: PcG, TrxG and DNA binding proteins in Drosophila

Overview of Polycomb group (red/orange) and Trithorax group (green) proteins. DNA binding proteins are coloured in blue. Histone methyl transferases are drawn as stars. Not all PcG and TrxG proteins known are presented.
Abbreviations are listed in (7).

In summary, although much progress in the field has been made by the purification and biochemical characterization of multimeric protein complexes *in vitro*, we have very little insight into the behavior of these complexes *in vivo*.

2.3 Epigenetic memory mechanisms by the PcG and TrxG in Drosophila

Members of the PcG and TrxG act for the most part antagonistically. PcG protein mediated mechanisms are required for the transcriptional silent state of the target gene, whereas TrxG proteins are generally important for target gene transcription. The mechanisms by which PcG and TrxG protein complexes are targeted to chromatin and

maintain long-term transcriptional memory through many rounds of replication and mitosis and therefore allow stability as well as flexibility during developmental changes are not yet fully understood.

2.3.1 PcG proteins act via Polycomb Response Elements (PREs)

PREs are epigenetic switchable cis-regulatory elements that enable PcG and TrxG proteins to maintain the transcriptional status of the target gene over many cell generations [6, 68]. The first PREs have been identified in the fly as DNA stretches needed for *Hox* gene silencing dependent on PcG gene function, which has been referred to as PRE-mediated silencing [69, 70]. The *bxd* and *Fab7* PREs regulating the *Drosophila Hox* complex genes *Ubx* and *AbdB* respectively, have been shown to silence reporter genes in a PcG dependent manner in transgenic assays, which was also true for many other functionally tested PREs. Recent evidence in our lab reveals different behavior of different PREs in a genome position dependent manner (*H. Okulski, L. Ringrose*, unpublished observation). Potential and functional PREs have been identified at many different loci in *Drosophila*, containing various target genes [3, 71-75].

Studies that have addressed the question of what makes a functional PRE came to the conclusion that at least one type of the DSP1, PHO, Zeste and GAF/PSQ (GA)n binding motif can be found in a PRE. Although their numbers and order vary, they share particular spacings of pairs of binding motifs. This may be important for cooperative or competitive binding. [63, 64, 71, 76]. However, other yet unidentified sequences/DNA binding factors could be needed. In a genome wide PRE prediction another motif out of three unknown was found, the GTGT repeats, but an associated motif-binding protein has not been found so far [71]. Therefore, the sequence definition will likely be further refined in the future.

PREs can be located upstream, in introns, or downstream and several 10kb away from their target gene. Silencing of the target gene over a distance could be achieved as suggested by the PRE/TRE looping model [77]. Additionally, each homeotic gene is controlled by two or more PRE elements, accompanied by enhancers and boundary elements [69, 78-80]. This suggests that PREs operate in groups, either cooperatively or individually and dependent on other regulatory sequences. Furthermore, many PREs are used differently at different times in development, underlining the dynamic nature of regulation [73].

In mammals, specific target genes for PcG/TrxG mediated transcriptional control, e.g. the homeotic genes [81] and the tumor suppressor locus *Ink4a* [82] have been identified, but no PREs were known for a long time. Two studies in human and murine ES cells have

analyzed the genome-wide distribution of PcG protein binding locations [83, 84] and two recent reports identified DNA sequence regions that target PcG mediated silencing [85, 86].

Although recent effort heads towards the identification and characterization of mammalian PREs, Drosophila is not only the historical and best known, but also the model organism with most tools to study the field of PcG/TrxG-mediated cellular memory.

2.3.2 Mechanisms of epigenetic regulation by the PcG and TrxG proteins in Drosophila

Before the transcription factors that initiate silent or active states of their target genes disappear, the PcG and TrxG mediated memory mechanisms have to perceive these transcriptional states and take instructions to form a repressive or an active PRE configuration. It has been shown that the recruitment of PcG proteins is independent of the decision to silence or activate [87, 88]. ChIP analysis has identified PREs as binding platforms for proteins of the PcG, TrxG and DNA binding factors [72-75]. Interestingly, PREs are nucleosome-depleted [88], which could allow room for DNA-binding factors. In other words, DNA binding proteins are thought to recognize specific DNA motifs at the PRE, which then tether the PRC2, the PRC1 and/or TrxG protein complexes to these sites. Complexes establishing covalent modifications on chromatin could facilitate complex binding and transmit the formation of a silent or active chromatin structure.

The PcG complex PRC2 has H3K27me3 specific histone methyl transferase activity [89]. There are also other chromatin-modifying activities present in PRC2 variants like deacetylation by the HDAC RPD3 [30] and it has also been shown that mammalian PRC2 complexes can methylate H1K26 [90], but the functional relevance for PcG mediated silencing remains unclear. Genome-wide analysis has revealed that PcG protein binding correlates with H3K27me3, which distributes in broad domains rather than discrete binding peaks like investigated members of the PRCs [74] (reviewed in [91]).

PRC1 components have been shown to interact with PHO [37, 38] and the complex is thought to bind via the chromodomain of PC specifically to H3K27me3 [21] to stabilize target gene silencing. However, there is strong suggestion that histone modification is a downstream event of PcG mediated silencing and not the installing mechanism [3, 39, 92]. The recent discovery of demethylases and histone methylation being reversible modifications [93-96] and the fact that the interaction of chromodomains to methylated lysines is weak [21], this epigenetic mark offers opportunities for kinetic modulation and differential regulation [97]. How this is finally achieved still remains to be solved. Several

models have been proposed for the PcG-mediated repressive system, such as closed chromatin formation that does not allow access to promoters [98]; inhibition of the assembly of the preinitiation complex [39]; and interference with transcription of initiation and/ or elongation [99], therefore establishing structures that are unfavourable for transcription.

TrxG proteins act antagonistically to the PcG mediated silencing and can be activators and anti-repressors. Activation involves fundamental changes from exchange of bound protein complexes [53, 100], transcription of RNA through the PRE [55, 101], chromatin remodeling and modification [7, 40, 53, 56]. Experimental data imply that TrxG proteins are not necessarily needed for *de novo* activation of target genes, but rather that they act as antirepressors that fight a constant battle against the default state of silencing by the PcG [102, 103].

2.3.3 Implications for replication and mitosis

The PcG/TrxG-mediated memory has to pass through the cell cycle. DNA replication demands that both strands inherit the same transcriptional configuration. Recent studies have investigated the propagation of trimethylated H3K27 by the PRC2 complex through the cell cycle [104, 105]. During mitosis, chromosomes are highly condensed [106], transcription is globally silenced and the bulk of PcG proteins dissociates [107, 108]. Nevertheless, it has been shown that inactive chromatin as in mitotic chromosomes and heterochromatin, is still accessible to transcription factors and chromatin structural proteins and that binding is not static [109-111]. Few studies have examined the association of PcG proteins to chromatin during mitosis. In Drosophila, an antibody staining approach revealed the PcG proteins PC, PH and PSC to dissociate from chromatin during the onset of mitosis and reassociation between anaphase and next interphase, depending on the protein [107]. The examination of the distribution of PC::GFP during blastoderm cell cycles demonstrated that a minor fraction resided on mitotic chromosomes [112]. There is much insight to be gained by looking at both PcG and TrxG members during the cell cycle in living cells of an organism, which also presents a technical challenge. However, the well-characterized development of the *Drosophila* embryo exhibits many features like synchronous cell divisions to enable the study of mitotic processes in living animals.

2.4 Dynamics

Functional PcG complexes are assembled sequentially, with a particular hierarchy and directly on chromatin [39, 107]. Examining these protein complexes has been very informative, but most data have come from *in vitro* studies but still it is not clear exactly how the transcriptional memory mechanism works at the molecular level *in vivo*. There is accumulating evidence that interactions of PcG and TrxG proteins with their targets are different at different target loci in vivo [3, 113-115]. At each gene, the PRE mediated "cellular memory" system has to preserve the correct gene expression state. To get a full picture of this epigenetic memory system, it will be vital to take dynamic aspects into account.

Most studies to date on PcG-mediated repression have been heavily based on chromatin IP (ChIP), but kinetic data [115, 116] suggest a model in which the transcriptional state of chromatin domains is determined in a stochastic manner by the dynamic competition of activating and repressing factors. In addition, there is evidence that specific protein modifications of PcG and TrxG proteins with respect to the expression state may also contribute stability to the memory mechanism [117-119].

In 2005, Ficz et al. used photobleaching methods to investigate fluorescently tagged PC and PH fusion protein kinetics. Their experiments revealed a fast exchange of complexes faster than the cell cycle within few minutes and they measured rate constants suggesting heterogeneity of PRC1 complexes. Therefore, the authors ruled out models of repression in which access of transcription activators to chromatin is limited, but proposed the idea that long-term repression primarily reflects mass-action chemical equilibria [115].

The amount and distribution of repressive chromatin marks may affect the flexibility by influencing these equilibria. Similar conclusions have been drawn for HP1, a repressor protein that is targeted to heterochromatin [109, 110]. Since silencing is a dynamic process, it is reasonable to expect that activation, which counteracts silencing, will be similarily dynamic [3]. However, there are no reports in the literature addressing the *in vivo* binding kinetics of Drosophila TrxG proteins. In particular, the timing of associated functional protein complexes during development is of major interest, as well as the study of not only the repressive, but also the activating mechanism.

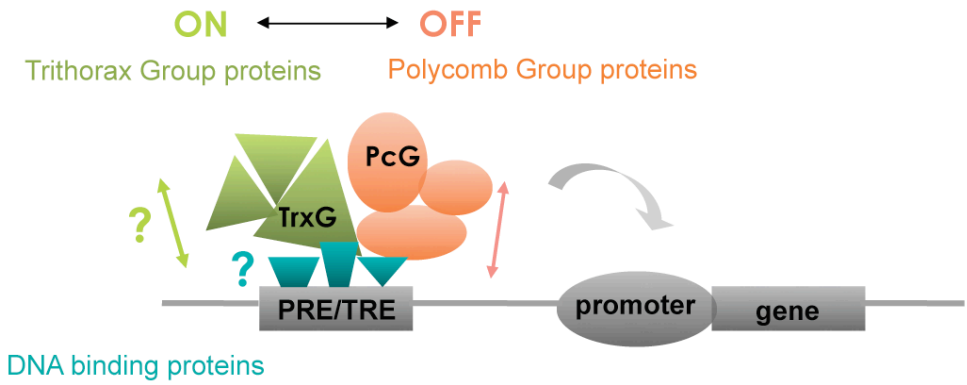

Figure 3: Dynamic PcG and TrxG protein mediated cellular memory

The cellular memory mechanism by the PcG and TrxG implicates the maintainance of the transcriptional state of the target gene through replication and mitosis to the next cell generation. This cellular memory mechanism is dynamic. Although the transcriptional memory is stably passed on to the next cell generation, PcG protein complexes bind dynamically to chromatin and complex stabilities vary, as shown by *Ringrose et al, 2004* and *Ficz et al, 2005*. This led to the idea that activation and repression is controlled by equilibrium between activating and repressing complexes. The protein complexes do not bind directly to Polycomb/Trithorax response elements (PRE/TRE) but via dann binding proteins, adding another possible layer of regulation for this system.

A comparison of the kinetics of protein binding in the activated and silenced states is required to better understand the dynamic dialogue between the TrxG and PcG proteins at their PRE targets. Likewise, our understanding of the transmission of this flexible but stable state through mitosis will be greatly enhanced by kinetic studies of complex behavior during the cell cycle.

2.5 Live Imaging

To understand a biological system it is essential to dissect the functions of proteins and other molecular components of this system in vivo. Fluorescent protein tags like GFP variants have been developed as a minimally invasive tool to visualize, track and quantify proteins of interest with live imaging methods. Advanced Confocal Microscopy and fluorescence imaging methods like FRAP (fluorescence recovery after photobleaching) have been essential to study the localization and kinetic behaviour of GFP-tagged fusion proteins (reviewed in [120]).

The first quantitative data of PcG complex formation and kinetic behavior in living Drosophila organisms and tissues using photobleaching fluorescence microscopy have been obtained in the lab of Donna Arndt-Jovin: Ficz and coworkers [115] have demonstrated that PC and PH in the repressive complex are exchangeable at all developmental stages studied. These data are in accordance with the mathematical predictions based on experimental results obtained by Ringrose and coworkers [116] and argue against models for repression that invoke blocking chromatin access. Regarding PcG/TrxG mediated regulatory mechanisms, the two PcG proteins investigated by Ficz et al. represent the PRC1 complex, but no PRC2 proteins have yet been examined. Also, association of TrxG proteins has been examined with immunofluorescence staining on polytene chromosomes and by chromatin immunoprecipitation analysis [75, 121, 122], but no quantitative data exist on the stability and time-dependent distribution of TrxG activating complexes. Finally, none of the sequence-specific DNA-binding proteins, which are important for PRE targeting, have been investigated with real-time kinetics.

In summary, dynamic aspects of complex formation and binding to chromatin and at mitosis remain elusive. For these studies, tools have to be established like a variety of GFP-tagged PcG and TrxG representatives. Technical challenges for quantitative studies and live imaging have to be taken to gain biological relevant data. Furthermore, the published microscopy data of Ficz and coworkers that show interactions of PcG and TrxG complexes with PREs have been obtained by examining a population of several hundred targets. It has not yet been possible to follow one of these PREs in the nucleus and correlate kinetic measurements. Therefore, a strategy to study single locus kinetics *in vivo* is required.

2.6 Aims of this thesis

The aim of this PhD project was to establish transgenic fly lines, which express fluorescently tagged members of the PcG and TrxG protein complexes to set up the tools for studying their function with live imaging methods in whole organisms and tissues. In order to gain insight into this dynamic epigenetic memory mechanism, which is able to ensure stability as well as flexibility, and to quantify its kinetic behavior, it was necessary to thoroughly characterize the established transgenic fly lines in terms of expression levels, ability of the fusion protein to rescue lethal mutations in the endogenous protein, and suitablility for live imaging. Expression of fluorescent fusion proteins in larval salivary glands identifies single binding loci of all tagged proteins in vivo. This thesis also provides the first data of protein binding of a DNA binding factor and a histone methyl transferase of the PcG to chromatin during mitosis.

In addition, I designed a visually tagged transgenic PRE locus that represses a reporter gene and will be switched to an activated state by an inducible signal. This setup will enable the study of PcG and TrxG protein kinetics at a single locus during DNA replication and mitosis and to characterize kinetic differences between the “on” and “off” states and will thus give fundamental insight into this mechanism in vivo.

3 Material and Methods

3.1 Molecular Biology Material and Methods

3.1.1 Material

3.1.1.1 Chemicals, enzymes and kits

- CIP Calf Intestine Phophatase[1U/μl] (Fermentas)
- T4 DNA Ligase [1U/μl] (Roche)
- Proteinase K [10μg/μl] (Invitrogen)
- RNaseA [10μg/μl] (Roche)
- Quick Ligation kit (New England Biolabs)
- Expand High Fidelity PCR System (Roche)
- T4 Polynucleotide Kinase (Roche)
- TURBO DNA-free™ (Ambion)

All restriction enzymes were purchased from Fermentas, New England Biolabs or Roche.

- Antibiotics: Sigma
- Complete Protease Inhibitor Mix, EDTA-free (Roche)
- Milk powder (Fixmilch Instant)
- DanKlorix (Colgate-Palmolive)
- BSA [10mg/ml] (New England Bio Labs)
- GeneRuler™DNA Ladder mix, ready to use (Fermentas)
- SYBR Safe™ DNA gel stain (Invitrogen)
- MassRuler™ 6x Loading Dye Solution (Fermentas)

All other chemicals were obtained from Merck, Fluka or Sigma-Aldrich or otherwise indicated in the text.

- QIAquick® Gel Extraction Kit (Qiagen)
- QIAquick® PCR Purification Kit (Qiagen)
    - DNeasy® Blood&Tissue Kit (Qiagen)
    - QIAprep® Spin Miniprep Kit (Qiagen)

- PureYield™Plasmid Midiprep System (Promega)
- High Pure RNA isolation Kit (Roche)
- Amersham ECL Plus™ Western Blotting Detection Reagents (GE Health care)
- SYBR®-Green Jump Start™ Taq Ready Mix™ for Quantitative PCR (Sigma)

3.1.1.2 Buffers, Media

- 4x NuPage LDS sample buffer (Invitrogen)

- LB-Medium pH 7.5 (prepared by the IMBA service department): 5g NaCl, 10g Tryptone, 5g Yeast extract, ad 1L ddH_2O pH adjusted with NaOH
- S.O.C. medium (Invitrogen)
- Complete Mini, Protease Inhibitor Cocktail Tablets (Roche), stock solution (7x): one tablet/1.5ml H_2O
- 3 M NaAc pH 5.2 (prepared by the IMBA service department): 408.1g NaAc3H_2O [3M], ad 1 L ddH_2O, pH adjusted with acetic acid
- 5 M NaCl (prepared by the IMBA service department): 292.2g NaCl [5M], ad 1l ddH_2O
- 1xPBS (prepared by the IMBA service department): 8g NaCl [137mM], 200mg KCl [2.7mM], 1.44g Na_2HPO_4O [10mM], 240mg KH_2PO_4O [2mM], ad 1L ddH_2O
- 10 x TAE buffer: 48.4g TRIS [0.4M], 3.7g EDTA [0.01M], ad 1 L ddH_2O, pH adjusted with acetic acid (prepared by the IMBA service department)
- 10x Tris-EDTA (TE) pH 8: 10mM Tris-HCl (pH 8.0), 1mM EDTA, H_2O ad to 1l
- 1 M TRIS ph 7.5 (prepared by the IMBA service department): 121.1g TRIS [1M], 29.2g EDTA [0.1M] ad 1 l ddH_2O, pH adjusted with acetic acid

Specific buffers are listed with the protocols (Methods, 3.2.1).

3.1.1.3 Bacterial strains

- DH5α (F^-_80*lacZ*ΔM15Δ*pho*AgyrA96, IMBA Service Department)
- Library Efficiency®DH5α^{TM} competent cells (F^- ϕ80*lacZ*ΔM15 Δ(*lac*ZYA-*arg*F)U169 *rec*A1 *end*A1 *hsd*R17(r_k-,m_k+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ^-, Invitrogen)

- MAX Efficiency®Stbl2™ competent cells (F- *mcr*A Δ(*mcr*BC-hsdRMS-*mrr*) *rec*A1 *end*A1 *lon gyr*A96 *thi sup*E44 *rel*A1 λ Δ(*lac-pro*AB), Invitrogen)

3.1.1.4 Primers and oligos

Oligo nucleotides for cloning multiple cloning sites (MCSs) (3.1.2.1 Fig A) and primers in pKC27 adjacent to the MCSs:

MCSforLacI GATCCGGCGCGCCTCCTAGGGCGGCCGCTGGTACCCTCGAGCAGGGCCCA
MCSforLacIrc CTAGTGGGCCCTGCTCGAGGGTACCAGCGGCCGCCCTAGGAGGCGCGCCG
MCSforprot GATCAGGCGCGCCTTGCGGCCGCAAGGTACCGCCGGCACCGGTGCTAGCCTCGAGGGATCCACGCGTGGGCCCTCT AGACCTAGGT
MCSforprotrc GATCACCTAGGTCTAGAGGGCCCACGCGTGGATCCCTCGAGGCTAGCACCGGTGCCGGCGGTACCTTGCGGCCGCA AGGCGCGCCT
MCSforPRE GATCACACGTCGACGGAGGATCCTTTAGATCTGAAGCGGCCGCTGAATTCATGGGCCCATCCTAGGTCT
MCSforPRErc GATCAGACCTAGGATGGGCCCATGAATTCAGCGGCCGCTTCAGATCTAAAGGATCCTCCGTCGACGTGT
primer5'ofMCS AGCTCACTCATTAGGCAC
primer3'ofMCS TTCCCAGTCACGACGTTG

Primer pair for amplifying SV40 with attached restriction sites for cloning:

3'SV40SpeI ACAACCACTAGTGATCCAGACATGATAAG
5'SV40XbaI ACCAATCTAGAGATCTTTGTGAAGGAAC

Primer pairs for amplification of coding sequences:

Dsp1primer5' AGGGCCTCGAGATGGAACACTTTCATC
Dsp1primer3' ATTATCCCGGGCTATTGGTTCTCGTCATC
Dsp1primer3'noStop ATTATCCCGGGTTGGTTCTCGTCATC
phoprimer5' TGGAGGAAGATCTATGGCCGGATCCGAATTC
phoprimer3' GTTACCCCGGGTCAGTCTGCATATACCACAAAC
phoprimer3'noStop GTTATCCCGGGGTCTGCATATACCACAAAC
phoprimer3'UTR TTTATACCCGGGCTCGAGCATCGTCTGATC **E(z)primer5'** TTGGACTCGAGATGAATAGCACTAAAGTG
E(z)primer3' TAATGGGGCCCTCAAACAATTTCCATTTC
E(z)primer3'noStop TAATGGGGCCCAACAATTTCCATTTC
mCherry5' ACTTGCTCGAGATGGTGAGCAAGGGCGAG
mCherry3' ACTTGGAATTCCTTGTACAGCTCGTCCATC

Primer pairs for amplification of promoter sequences:

pDsp1primer5' AAGGGGCGGCCGCTTAACATAATAAAACTAGCAG
pDsp1primer3' CGGGCGGTACCAACTGTGAGATTTCGGATC
pphoprimer5' TATATGCGGCCGCTGTCAGATTGTGTGCATGTTG
pphoprimer3' GCGGGGGTACCATTTACAATACTATAATCCACATAC
pE(z)primer5' AATTAGCGGCCGCACTGATAACGGGCTGCGTTTC
pE(z)primer3' AAGGGGGTACCAATGCCTTCGAGGGACTTTATATTTG
5'Ubi AGTTGAATGCGGCCGCGACTGGAAATTTAAAATGGAG
3'Ubi ACGGTTGGTACCTTGTCGCCGGAAC

Primer pairs with attached restriction sites for cloning transgenic PRE constructs:

5'pcrUAS TTCATTACAGATCTAGGCCACTAGTAAG
3'pcrUAS ACATTATGGGCCCTATTCAGAGTTC
Fab75' AGTTGGACGCGGCCGCAAGCTTGATGCTATCGCG
Fab73' AATGGTACTAGTGCCATAATGCCCCTTG

Primer used in single fly PCR:

5'PCRmCherryflies ATGGTGAGCAAGGGCGAG
3'PCRmCherryflies TCTTGACCTCAGCGTCGTAG

Real-time PCR primer:

Tbp5' CATCGTGTCCACGGTTAATCT
Tbp3' GAAACCGAGCTTTTGGATGAT
Gfpa5' CTTCAAGGAGGACGGCAACATC
Gfpa3' GAACTCCAGCAGGACCATGTG
Pho35' CATCAATGCCAAATTCTTCCT
Pho33' AATTCAAACGGCAAGTCAAGA
Ez45' GCAGCTTGTCCTTGACTTCTT
Ez43' TGGATGTAGACGCAGATTGTG

Primer for sequencing were selected every 400-500bp in the sequence of interest, 17-19bp in length and according to general rules for primer design.

3.1.1.5 Vectors and plasmids

- pGEX-E(z) (R. Jones)
- pMalC2 (D. Locker)
- pT7link (J. Müller)
- FRED11 9.9 (L.Ringrose)
- pSR301a1tubulin (S.Rotkopf, B.Dickson)
- pGD269actin5c (G.Dietzel, B.Dickson)
- pWRPUbi (J.Mummery-Widmer, J.Knoblich)
- pAFS-UASP (J.Vazquez)
- pLacO153 (2.5kb LacO, 64-mer) (J.Vazquez)
- pLacO150 (5kb LacO, 128-mer) (J.Vazquez)
- pV (10kb lacO, D. Arndt-Jovin)
- pSV-dhfr 8.32 (A.Belmont)
- pEGFP-N2,-C1,C3 (Clontech)
- pmCherryIII (R.Tsien)
- pUZ (L.Ringrose)

3.1.1.6 Antibodies

Primary antibodies:

- Polyclonal rabbit Anti-PC (Santa Cruz)
- Polyclonal rabbit Anti-DSP1 (D.Locker)
- Polyclonal rabbit Anti-PHO (J.Kassis)
- Polyclonal rabbit Anti-E(Z) (R.Jones)
- Monoclonal mouse Anti-β-Actin (AC-15) (GeneTex)
- Polyclonal rabbit Anti-histone H3 (Abcam)

Secondary antibodies:

- Alexa Fluor® 488 goat anti mouse (Invitrogen)
- Alexa Fluor® 488 goat anti rabbit (Invitrogen)
- Alexa Fluor® 647 goat anti rabbit (Invitrogen)
- Anti rabbit IgG-Horse Raddish Peroxidase (Sigma)

3.1.1.7 Equipment, Machines and Software

- Standard laboratory equipment

- PCR machine, Mastercycler ep gradient S (Eppendorf)
- Real time PCR machine, Mastercycler ep gradient S realplex^2 (Eppendorf)
- Fly cages (IMBA Workshop)
- Agarose gel electrophoresis apparatus (Mupid-ExU, Submarine Electrophoresis)
- Non-UV transilluminator, Safe-Imager™ (Invitrogen)
- UV DNA Imager (Biorad)
- Pellet Pestle (Kontes)
- Spectrophotometer, Nanodrop ND-1000 (peqlab)
- Double-sided adhesive tape (Scotch)
- Western blot developer (Colenta® MP900F)
- ECL film, Amersham Hyperfilm™ High Performance chemiluminescence film (GE Health care)
- Transmitted light scanner (Canon Scan 4200F)
- XCell SureLock Mini-Cell Electrophoresis System (Invitrogen)
- Forceps
- Kodak Molecular Imager (Carestream Health, Inc.)

- Gene Constructing Kit, GCK (Textco)
- MacVector 7.2.3 (Accelrys, Inc.)
- IDT SciTools OligoAnalyzer 3.1 (IDT, http://eu.idtdna.com/analyzer/applications/oligoanalyzer/default.aspx)
- Primer3 0.4.0 (http://frodo.wi.mit.edu/primer3/)
- CLC Combined Workbench 3.5.1 (CLCbio)
- SPOT Digital camera systems 4.5.9.8. (DIAGNOSTIC instruments, Inc.)
- Kodak Molecular Imaging Software 4.5.1. (CARESTREAM HEALTH, Inc.)
- Zeiss LSM510 software AIM4.2
- Huygens Pro 3.5 (SVI)
- Deltavision images – SoftWorx 3.7.0 (Applied Precision)
- MetaMorph 7.6 (Molecular Devices)

3.1.2 Methods

3.1.2.1 Cloning Strategies

In order to clone GFP-tagged proteins of interest, mCherry-tagged LacI and the transgenic PRE, two basic vectors were constructed from the pKC27 (Su, K and Dickson, B unpublished; described in detail in Springer Protocols, Transgenesis Techniques ©Humana Press *Transgenesis in Drosophila melanogaster*, Ringrose, L) plasmid.

All GFP-fusion protein constructs (4.1) were cloned into pKC27p, which was generated by cutting pKC27 with BamHI and inserting a multiple cloning site (MCSp) fragment (Fig. 4A).

All constructs of the transgenic PRE (4.6.1) were cloned into the MCSpre inserted into the BamHI cut vector backbone of pKC27, generating pKC27pre (Fig. 4A).

The correct orientation of the MCSs in the pKC27vector backbone in clones, which were further used in the cloning procedure, was tested by PCR.

The MCSl for mCherryLacI fusion protein constructs (Fig 4A) was directionally cloned into BamHI/SpeI double digested vector pKC27, making up pKC27l.

In order to sequence the MCS and inserts or their direction, primers left and right of the MCS in the pKC27 vector (*primer5'ofMCS* and *primer3'ofMCS*) were designed that were used at various steps during the cloning procedure.

Clones of all generated vectors were sequenced by the in house sequencing department and correct ones were selected for further cloning.

A

MCSl

```
                                        KpnI     XhoI
BamHI  AscI        AvrII    NotI        Asp718I          PspOMI  SpeI
ggatccGGCGCGCCtCCTAGGgcggccgctggtaccctcgagcagggcccactagt
cctaggCCGCGCGGaGGATCCcgccggcgaccatgggagctcgtcccgggtgatca
```

MCSp

```
                                KpnI    NgoMIV
BclI    AscI          NotI      Asp718I       AgeI    NheI    XhoI    BamHI  MluI     PspOMI XbaI    AvrII    BclI
tgatcaGGCGCGCCttgcggccgcaaGGTACCgccggcACCGGTgctagcCTCGAGggatccACGCGTgggcccTCTAGAcctaggTGATCA
actagtCCGCGCGGaacgccggcgttCCATGGcggccgTGGCCAcgatcgGAGCTCcctaggTGCGCAcccgggAGATCTggatccACTAGT
```

MCSpre

```
BclI        SalI        BamHI      BglII       NotI        EcoRI     PspOMI    AvrII      BclI
tgatcacacgtcgacggaggatcctttagatctgaagcggccgctgaattcatgggcccatcctaggtctgatca
actagtgtgcagctgcctcctaggaaatctagacttcgccggcgacttaagtacccgggtaggatccagactagt
```

B

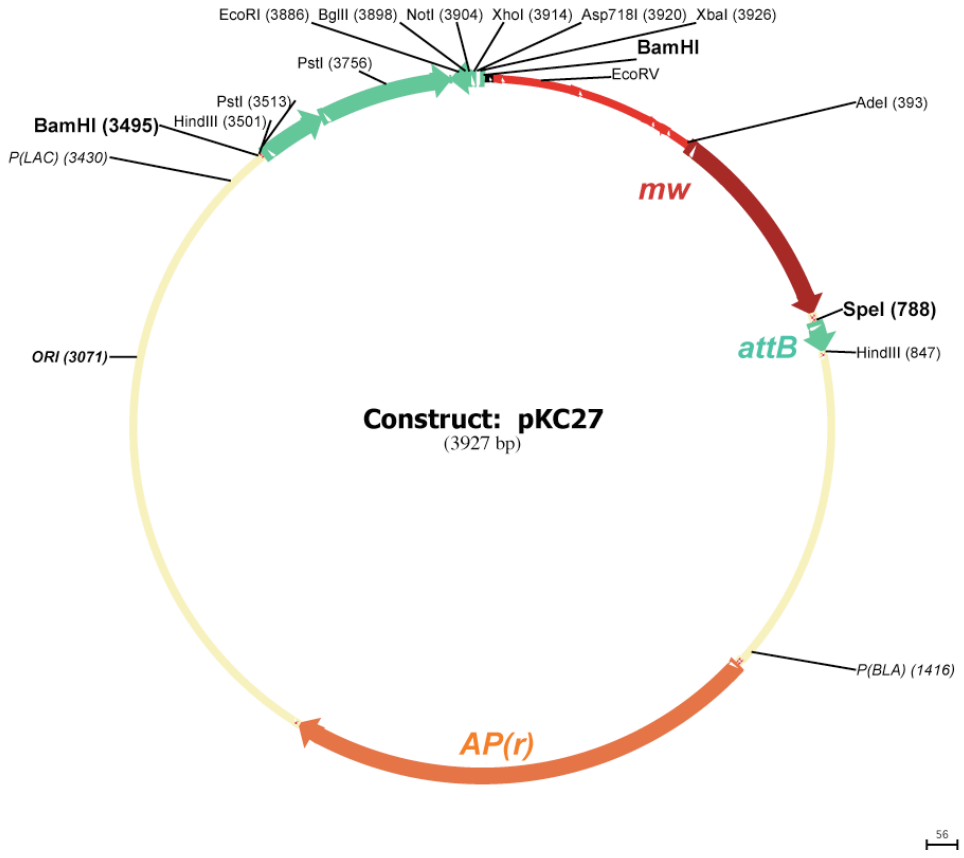

Figure 4: Multiple cloning sites and original vector for pKC27 based plasmids

The multiple cloning site MCSl (A) for cloning the mCherry tagged LacI repressor constructs was inserted into the BamHI and SpeI (restriction sites marked in bold) cut vector backbone containing the origine of replication, the *attB* site for site specific recombination by the phiC31 integrase into an *attP* landing site and the ampicilline restistance gene (*AP(r)*).

For generating the Gfp fusion protein constructs and the vectors containing the transgenic Fab7 PRE, the *miniwhite* cassette with the minimal promoter and the 5' coding sequence for complementation of the *white* gene was needed. Therefore, the MCSp and MCSpre were cloned into BamHI digested vector backbone, removing a smaller fragment from the original vector than for *mCherryLacI* cloning. *ORI*…origine of replication, *P(LAC)*… *lac* promoter, *P(BLA)*… *bla* gene promoter, *mw*…5' *miniwhite* cassette.

Cloning strategy for *Gfp* fusion protein constructs

BclI/BamHI ligation of the MCSp into BamHI digested pKC27 destroyed the recognition sequences at the junction site for both enzymes. A clone with the correct orientation of the MCSp was selected to proceed with further cloning steps.

Both N- and C-terminally tagged GFP fusion proteins were cloned.

For the C-terminally tagged constructs, an SV40 polyA sequence, which was PCR amplified out of pKC27 with primers *3'SV40SpeI* and *5'SV40XbaI* was inserted into the MCSp by cutting the vector with XbaI and the SV40 polyA fragment with XbaI and SpeI in order to keep an intact upstream XbaI site.

The cDNA sequences for *Dsp1*, *pho* and *E(z)* were PCR amplified out of templates listed in (3.1.1.5) and directionally cloned into pEGFP vectors (Clontech). The coding sequences were proofed against mutations with sequencing and correct clones were chosen for the next cloning step.

Fusion protein	Rsites(primers)	pEGFP vector
Dsp1Gfp	XhoI (Dsp1primer5')/ XmaI (Dsp1primer3'noStop)	pEGFP-N2
phoGfp	BglII (phoprimer5')/ XmaI (phoprimer3'noStop)	pEGFP-N2
E(z)Gfp	XhoI (E(z)primer5')/ ApaI (E(z)primer3'noStop)	pEGFP-N2
GfpDsp1	XhoI (Dsp1primer5')/ XmaI (Dsp1primer3')	pEGFP-C3
Gfppho	BglII (phoprimer5')/ XmaI (phoprimer3') or (phoprimer3'UTR)	pEGFP-C1
GfpE(z)	XhoI (E(z)primer5')/ ApaI (E(z)primer3')	pEGFP-C3

Table 1: Combination of restriction sites and target pEGFP vectors for cloning Gfp fusion proteins

(Table 1)
To fuse *Gfp* with the PcG proteins of interest, an intermediate cloning step of inserting the PCR amplified coding sequences of the proteins of interest generated with listed primers and attached restriction sites into pEGFP vectors (Clontech) was performed. Subsequently, the fusion constructs were excised and inserted into pKC27 based vectors (Fig. 4).

The C-terminally tagged proteins were cut out at BglII and XbaI (*pho*) or XhoI and XbaI restriction sites and were cloned upstream of the SV40 *polyA* signal in the pKC27p vector. For the *phoGfp* insert, the vector was cut with BamHI and ligated with the BglII end.

The N-terminally tagged proteins were cut out of pEGFP-C1/3 with the downstream SV40 *polyA* sequence using NheI and MluI restrictions sites and were directionally cloned into the plain pKC27p vector.

As the last cloning step, promoter sequences were added for transgenic fusion protein expression using NotI and KpnI restriction sites. The promoter sequences of *Actin5c*, *α1Tubulin* were cut out of vectors *pGD269Actin5c* and *pSR301α1Tubulin*, respectively, and the presumed endogenous promoters were PCR ampified using genomic wildtype DNA from OregonR flies and primers listed in (3.1.1.4).

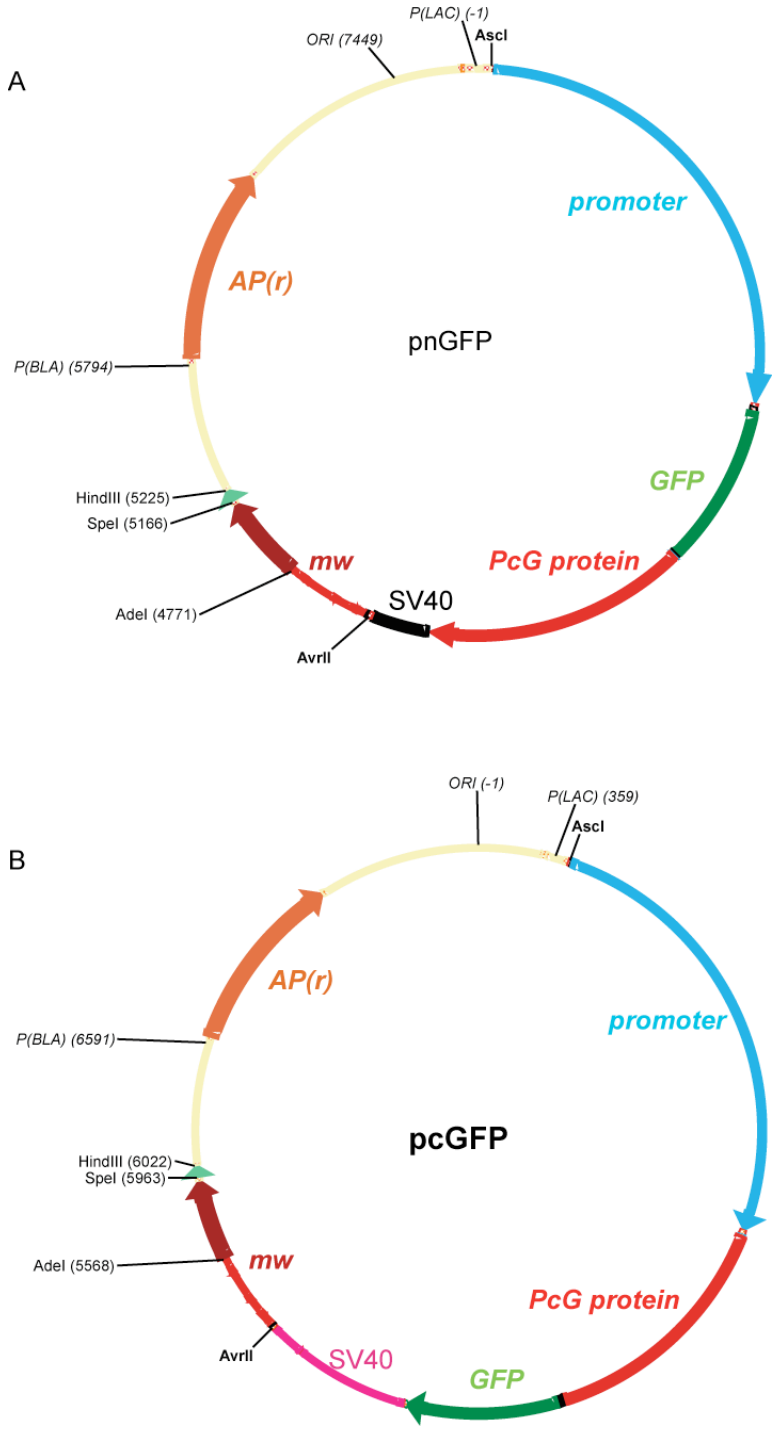

Figure 5:Schematic presentation of final Gfp fusion protein vectors for site-specific integration by the phiC31 integrase.

pnGFP vectors contained n-terminally tagged proteins of interest (*Dsp1, pho, E(z)*) and pcGFPvectors were established with C-terminally tagged versions. The promoter-fusionprotein cassette was cloned between AscI and AvrII restriction sites (marked in bold letters) for recloning into *pmCherryLacI* (Fig. 6 and 5.6). *ORI*…origine of replication, *P(LAC)*…*lac* promoter, *P(BLA)*…*bla* gene promoter, *mw*…5' *miniwhite* cassette, SV40…termination sequence, *polyA*.

Cloning strategy for mCherry-LacI constructs

In order to distinguish the correct orientation of the MCSI, which was ligated into the vector at BamHI and SpeI compatible ends, a short PCR product as for the other vectors/+MCS was sequenced.

GfpLacI was cut out of purified pAFS-UASP using XhoI and NotI and ligated into XhoI and PspOMI cut vector pKC27I. *mCherry* was PCR amplified from an *mCherry* containing plasmid (R. Tsien) using *5'mCherry* and *3'mCherry* primers and exchanged *Gfp* in the plasmid using XhoI and EcoRI. A selected number of clones was sequenced to control for mutations. A PCR amplified SV40 *polyA* sequence was added at XbaI and SpeI restriction sites downstream of *LacI*. The last step was generating three different constructs by adding three different ubiquitous promoter sequences: *pActin5c*, *pα1Tubulin* (as for Gfp fusion protein constructs) and *pUbiquitin63E*. The *Ubiquitin63E* promoter was PCR amplified using primers *5'Ubi* and *3'Ubi* adding NotI and KpnI restriction sites.

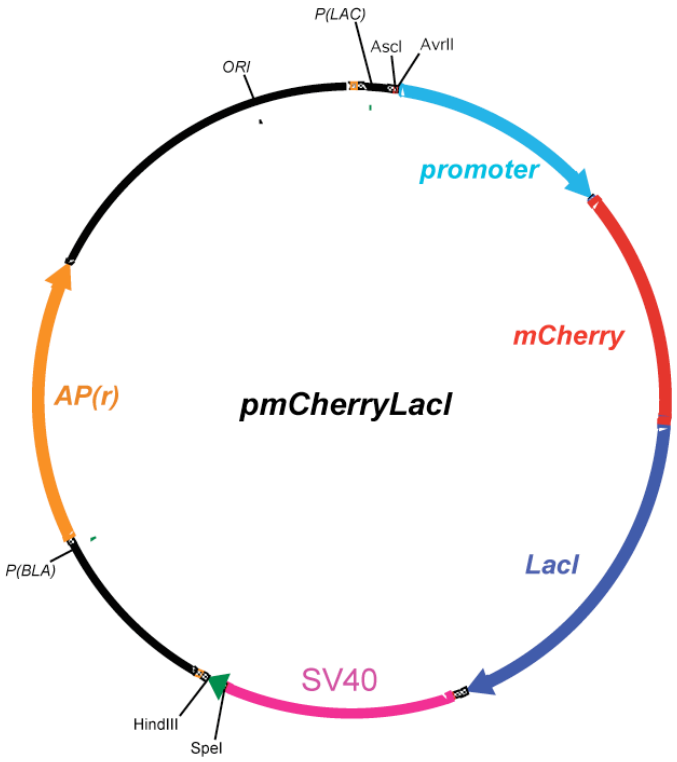

Figure 6: Schematic representation of mCherryLacI containing vectors

(Figure 6)
mCherryLacI fusion protein vectors were constructed in *miniwhite*-less plasmids so that transgenic flies could be recombined with transgenic PRE-*miniwhite* containing flies without interference with the function of *miniwhite* as a reporter gene in the PRE locus (4.6.1).
AscI and AvrII sites are upstream of the promoter sequence for Gfp fusion protein cassette insertion (Fig 4A and 5.6). *attB*…homologous recombination sequence for the phiC31 integrase, *AP(r)*…ampicilline resistance gene, *ORI*…origine of replication, *P(LAC)*…*lac* promoter, *P(BLA)*…*bla* gene promoter, *mw*…5' *miniwhite* cassette, SV40…termination sequence, *polyA*.

Cloning strategy for the transgenic PRE

In order to clone an inducibly regulated transgenic PRE that can be visually tagged into pKC27pre. The first step was adding a PCR amplified Gal4 inducible *Upstream Activating Sequence* (*UAS*) from pUZ plasmid using primers *5'pcrUAS* and *3'pcrUAS* and BglII and PspOMI directed insertion. This plasmid was cut with PspOMI and AvrII to add the 3.2kb *Fab7* PRE sequence, which was amplified out of genomic wildtype DNA using the primer pair *Fab75'* and *Fab73'* and added compatible NotI and SpeI restriction sites. This plasmid was used to generate the *Fab7* control fly line (4.6.1).

To visually tag the transgenic PRE multiple *LacO* repeats, which can be specifically recognized by LacI and a GFP::LacI fusion protein in flies [123, 124], were cloned upstream of the *UAS*. For subsequent cloning steps, MAXEfficiency®Stbl2^{TM} competent cells, which are suitable for unstable inserts such as large repeats, were employed. 10kb, 5kb and 2.5kb *LacO* repeats were cut out of purified plasmids (3.1.1.5) and were then inserted at SalI and BamHI restriction sites.

The yeast DNA fragment in the FRED11 9.9 plasmid of approximately 10kb in size was tested for PRE motifs using the PREdictor [71]. The fragment did not contain any PRE similarity, and was used as a spacer to separate the *LacO* array from the transgenic PRE and therefore obviate potential influence of the large repetitive sequence and LacI binding on the *Fab7* element. The spacer fragment from digested pFRED was inserted at the BglII site of the MCSI. A control plasmid without LacO inserted, but with the spacer fragment in the BglII site was also generated for establishing the *sp-Fab7* transgenic fly line (4.6.1).

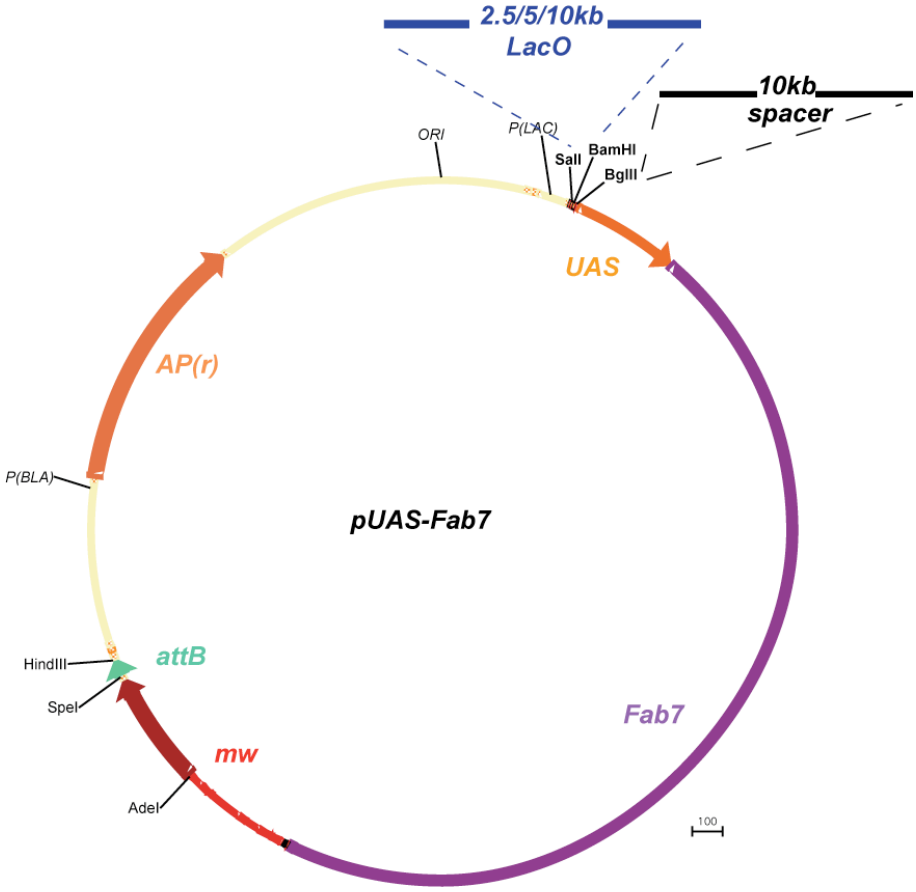

Figure 7: Scheme for UAS-Fab7 containing plasmids

UAS-Fab7, sp-UAS-Fab7 and *LacO-sp-UAS-Fab7* transgenics (4.6.1) were generated in pKC27 vector backbone with the 5'miniwhite cassette for transformation marker complementation in *attP* landing sites. In these construct, the full *miniwhite* cassette is designed to serve as a reporter gene for Fab7 PRE activity in transgenic flies. As indicated, inserting a 10kb spacer DNA into the BglII site generated the vector for *sp-UAS-Fab7*. 2.5kb, 5kb and 10kb *LacO* repeat containing constructs as part of the mCherryLacI tagging strategy were constructed by first adding the *LacO* repeat fragments into *UAS-Fab7* constructs at SalI and BamHI sites and then separating the repeat from the *UAS-Fab7* sequence by adding the 10kb spacer DNA into the BglII site. *attB*…homologous recombination sequence for the phiC31 integrase, *AP(r)*…ampicilline resistance gene, *ORI*…origine of replication, *P(LAC)…lac* promoter, *P(BLA)…bla* gene promoter, *mw*…5' *miniwhite* cassette.

3.1.2.2 Polymerase Chain Reaction (PCR)

Polymerase chain reaction was used to amplify cloning fragments with appropriate restriction sites and as an analytical tool.

For cloning purposes the Expand High Fidelity PCR system (Roche) was used. 300nM of each primer and 0.5ng (plasmids) – 100ng (gDNA) template DNA were used. Usually 25µl or 50µl reactions were pipetted.

Generally, “touchdown” PCR was programmed:

2’94°C

5x: 30’’94°C
30’’45-65°C (-1°C/cycle)
45’’68°C or 72°C

31x: 30’’94°C
30’’56°C
45’’68°C or 72°C

7’68°C or 72°C

4°C

For analytical purposes Taq polymerase (Eppendorf, 5prime) was used, except for single fly PCR (3.1.2.18).

3.1.2.3 Oligonucleotide cloning

Complementary oligonucleotides were resuspended with ddH_2O to a concentration of 100pmol/μl and stored at -20°C. 50 pmol of each oligonucleotide, 10U Polynucleotide Kinase, 5μl 10x buffer (Roche) and 1mM ATP were used in 50μl reactions. The reaction was incubated for 1h at 37°C to catalyze the transfer of the terminal phosphate group of ATP to the 5'-hydroxylated terminus of the oligonucleotides and stopped with 10mM EDTA.

To anneal the oligonucleotides, the volume was heated to 70°C for 15minutes for denaturation and slowly cooled down. Subsequently, the oligonucleotide kination was used like any other insert in 3:1 molar excess over cut vector for ligation.

3.1.2.4 Agarose Gel Electrophoresis

For 0.6%-1.5% w/v agarose gels the agarose was dissolved in 100 ml 0.5x TAE buffer and boiled. After cooling down 10μl SYBR Safe™ DNA gel stain were added (0.1μl/ml gel), the gel was poured into an electrophoresis tray and the comb was placed. After solidification of the gel the comb was removed and the gel was put into an electrophoresis apparatus filled with 0.5x TAE buffer. 6x Loading Dye (Fermentas) was added to DNA. In addition to DNA samples a marker (GeneRuler™DNA Ladder mix, ready to use (Fermentas)) was loaded into one well. Gel electrophoresis was performed for 30 minutes at 100V.

Using an imager (Biorad) DNA bands were visualized under UV light and an image taken. Alternatively, in order to avoid UV damage to DNA the non-UV transilluminator (Invitrogen) was used in case of subsequent use of separated DNA for cloning.

3.1.2.5 DNA extraction from gels and purification of PCR products

PCR product purification

PCR products were purified using the PCR purification kit (Qiagen) or alternatively Montage® PCR Centrifugal Filter Devices.

DNA gelextraction

DNA fragments after restriction nuclease digestion were separated by agarose gel electrophoresis, the gel volume containing the fragment of appropriate size was cut out and the DNA was gel extracted using the Gel extraction Kit (Qiagen).

3.1.2.6 Restriction Enzyme Digestion and Ligation

The restriction enzymes were used as recommended by the supplier to generate linear DNA fragments with compatible ends for directed cloning. In case of double digestions of with different optimal digestion conditions two subsequent single enzyme digests were performed.

Digested vector DNA, oligonucleotides and PCR products were gel extracted using the Gel extraction Kit (Qiagen) or purified with the PCR purification Kit (Qiagen).

Calf intestine phosphatase treatment

In order to remove 5' phosphates from double-digested vector DNA, 0.5U/μg (vector DNA) CIP Calf intestine phosphatase (Roche) were used at conditions recommended by the supplier. The reaction was incubated for 1h at 37°C and subsequently loaded onto an agarose gel for separation and extraction of the vector fragment.

Ligation

In order to ligate vector DNA with insert DNA the T4 DNA Ligase (Roche) was used. 50 ng vector DNA and a 3 fold excess of insert DNA were added to 2μl 10xLigation Buffer and 1μl T4 DNA Ligase (1U/μl). H_2O was added to a final volume of 20μl. The reaction was incubated at RT for 4hrs or o/n at 16°C.

Alternatively, the Quick Ligation Kit and was used for the same amount of material as for the T4 DNA ligation as recommended by the supplier.
Control experiments for efficiencies of restriction enzyme digest, CIP treatment and the occurrence of religation of the vector were routinely performed.

3.1.2.7 Transformation of chemical competent bacteria

CaCl-competent E.coli bacteria (DH5α - IMBA service department) were transferred from -80°C to ice. After thawing a 100µl aliquot *E.coli* was transferred into an Eppendorf tube. Approximately 10ng DNA were added and incubated on ice for 20 minutes. Then the bacteria were heat-shocked at 42°C for 90 seconds and the mixture was immediately put on ice afterwards. 650µl pre-warmed S.O.C. medium were added and the mixture was shaken for 45 minutes (1200 rpm) at 37°C. According to antibiotic resistance of the transformed vector, *E.coli* bacteria were plated onto LB agarose containing the selective antibiotic and incubated over night at 37°C.
For increased transformation efficiency MAX Efficiency®DH5α™ competent cells (Invitrogen) were used. Transformation was adapted as recommended by the supplier.
For cloning large repetitive sequences (*LacO* containing plasmids), MAX Efficiency®Stbl2™ (Invitrogen) competent cells were used. The protocol provided by the supplier was used for transformation.

3.1.2.8 Plasmid DNA purification (Mini/Midiprep)

2ml o/n culture of transformed *E.coli* bacteria were directly used for miniprep plasmid purification or inoculated 1:500 in 50ml LB medium with the selective antibiotic and incubated for 8 hours for midiprep plasmid purification. For minipreps, plasmid preparations were performed with the QIAprep Spin Miniprep Kit (Qiagen) and midipreps with the PureYield™Plasmid Midiprep System (Promega). Midiprep purifications for injection into embryos were eluted in ddH_2O and further precipitated.

3.1.2.9 Genomic DNA extraction

Genomic DNA preparation of Drosophila melanogaster was performed using the DNeasy®Blood&Tissue Kit (Qiagen). 10 anesthetized flies were transferred into a sterile 1.5ml Eppendorf tube and frozen in liquid nitrogen. Then the flies were ground with a

tissue grinder in 100 µl ATL buffer. The following gDNA preparation was performed according to the kit description.
Genomic DNA was eluted with 100µl AE buffer. This elution step was repeated in order to obtain an increased yield and the DNA concentration was determined.

3.1.2.10 Phenol-chloroform extraction and DNA precipitation

Phenol-chloroform extraction and DNA precipitation were performed to concentrate DNA samples and remove protein remains. An equal volume of phenol, chloroform and isoamyl alcohol (25:24:1) was added to the DNA suspension and centrifuged 3 min. at 13000 rpm in a table-top centrifuge. Afterwards, the DNA containing aqueous phase was transferred to another Eppendorf tube and an additional extraction with an equal volume of chloroform and isoamly alcohol (24:1) was performed. Again the mixture was centrifuged for 3 min. at 13000 rpm. The DNA containing aqueous phase was transferred to a sterile tube and subjected to DNA precipitation.
2.5x of the DNA suspension volume was calculated and this volume of absolute ethanol (-20°C) was added. Then 1/10 of the DNA suspension volume was calculated and this volume of 3M NaAc pH 5.2 was added. The solution was mixed and incubated at -20°C over night or at -80°C for 45 minutes. After incubation the solution was centrifuged at 4°C for 30 minutes, full speed. The supernatant was discarded and the pellet was washed with 100µl 70% ethanol (4°C). Again the solution was centrifuged at 4°C for 10 minutes, full speed, followed by the removal of the supernatant. The pelleted precipitated DNA was dried (open Eppendorf tube transferred to heating block 37°C). Finally the DNA was resuspended in H_2O and the concentration was measured.

3.1.2.11 Determination of DNA concentration

The DNA concentration was measured by loading 1,5µl of DNA suspension onto a Spectrophotometer (NanoDrop), measurement of 260nm absorbance and calculation of ng/µl based on Beer's Law.

3.1.2.12 Total Protein Extraction from Drosophila

The protocol used was adapted from Ficz, G. (2005) *Protein dynamics in the nucleus: Implications for gene expression*, Georg August University Göttingen.

20 Drosophila embryos from overnight collections (0-12h) were dechorionated by rolling them on double-sided adhesive tape and collected into a 1.5ml Eppendorf tube containing 50µl DLB. For protein extract from larval tissue, 15 pairs of salivary glands and 25-30 wing discs were collected in 40µl DLB. Rotating and moving a clean pestle up and down for 30 seconds squashed the embryos. Afterwards the tube was frozen in liquid nitrogen. The freezing-thawing-squashing procedure was repeated 3 times. The pestle was rinsed with 20µl DLB and the solution was mixed by resuspension. 10µl of the extract were transferred to a 1.5ml Eppendorf tube for subsequent protein concentration measurement. 4x NuPage LDS Sample buffer (Invitrogen) and DTT to a final concentration of 50mM were added to the extract. Then the extract was boiled for 10 minutes at 70°C followed by centrifugation at full speed for 10 minutes (at RT). The water-soluble fraction was separated from the pelleted lipid layer and the aliquoted extracts (30µl/tube) were frozen or immediately used to perform a Western Blot analysis.

- Drosophila lysis buffer (DLB): Hepes-KOH pH 7.5 [20mM], KCl [100mM], EDTA [2mM], Triton X-100 [0.5%], Aprotinin 0.3U/ml, Complete cocktail stock solution 14% v/v, DTT [5mM], $MgAc_2$ [1mM]

3.1.2.13 Determination of protein concentration by Bradford

Determination of the protein concentration was performed according to the method of Bradford (Stoscheck, 1990), 795 µl H_2O were mixed with 5 µl of protein solution and 200 µl of Bradford-solution (BioRad, Germany) in a plastic- cuvette. The absorption of this mixture was measured in a photometer at a wavelength of 595 nm against a blank solution containing 800 µl H2O and 200 µl of Bradford-solution. The corresponding protein concentration was calculated from a calibration curve using bovine serum albumin (BSA) as a standard.

3.1.2.14 Western blotting (SDS-Polyacrylamide gel electrophoresis (SDS-PAGE) and Immunoblotting)

SDS-PAGE:

Depending on the protein quality of the extract and the antibody for the investigated protein, 7-12µg protein were loaded into precast gels from Invitrogen.

In order to achieve optimal separation of investigated proteins NuPAGE® Novex 10% Bis-Tris gels with MOPS running buffer (E(Z)) or MES running buffer (DSP1) and 4-12% Bis Tris gels with MOPS running buffer (PHO) were used in SDS-PAGE. Gel electrophoresis was performed according to the NuPAGE® electrophoresis system (Invitrogen).

Blotting:

After gel electrophoresis the gel was incubated in transfer buffer (Invitrogen) with gentle shaking for 20 minutes.

The membrane was placed between three upper and three lower layers of Whatmann paper and below the gel onto the blotting apparatus, with all layers presoaked in transfer buffer. In order to have optimal transfer, air bubbles have been removed. Finally the proteins were blotted onto the membrane at 1.2mA/cm^2 for 1 hour.

Immunodetection:

After blotting the membrane was incubated in blocking buffer and slowly shaking for 40 minutes. The membrane was rinsed with TBST and the first antibody diluted in antibody buffer at appropriate volume was added, followed by incubation at 4°C over night and slowly shaking. On the following day the membrane was washed three times for 5 minutes with antibody buffer and rinsed with TBST. Afterwards the second antibody (Horse radish peroxidase conjugated anti rabbit IgG, 1:5000) diluted in 5 ml antibody buffer was added to the membrane, followed by a 1-hour incubation at room temperature with slowly shaking.

After incubation the membrane was washed three times for five minutes in TBST and one time for five minutes in TBS. Finally ECL detection was performed as described by the supplier. The membrane was placed between plastic sheets and transferred to a cassette. In the darkroom an ECL film was placed onto the membrane covered by plastic sheets and the cassette was closed. Depending on signal intensity different exposure times were applied. The ECL films were developed using the Colenta® MP900F Western Blot developer.

Alternatively to quantify DSP1, the CCD camera in the Kodak Molecular Imager directly captured the photons that were emitted from the membrane.

Buffers:

- TBS: Tris pH 8.0 [10mM], NaCl [150mM]

- TBST: TBS, 0.2% Tween 20

- Block buffer: TBS, 5% non-fat milk powder

- Antibody buffer: TBST, 2% BSA

Western Blot data analysis:
Scanned in (Canon Scan 4200F) developed ECL films or captured images from the Kodak Molecular Imager were used for subsequent signal intensity analysis with the Molecular Imaging Software (Kodak, Molecular Imaging Software). Manual regions of interest (ROIs) were selected and net intensities of different protein bands within these ROIs were compared. The background was subtracted by the median of the perimeter of each ROI. In order to improve the significance of the calculated ratios blots were averaged.

3.1.2.15 RNA isolation from embryos

Embryo collection:
0-12 hours o/n embryo collections were subjected to dechorionation. Embryos were washed off the plate by egg-wash and using a brush. The embryos were transferred into a net-tube and dechorionated with 1.5ml bleach solution for 2 minutes and 30 seconds. Immediately afterwards the embryos were extensively washed with egg wash solution. Finally the dechorionated embryos were washed onto a net with 1xPBS. Using the brush they were transferred into a 1.5ml Eppendorf tube containing 200µl 1xPBS. Afterwards the embryos were subjected to RNA extraction.

RNA isolation:
In order to isolate RNA from Drosophila embryos, the embryos were squashed by rotating a clean pestle up and down and RNA isolation was performed with the High Pure RNA isolation Kit (Roche) as recommended by the supplier. The RNA was eluted in 50µl elution buffer. Subsequent to RNA isolation the RNA concentration was measured by loading 1.5µl eluate onto a spectrophotometer (NanoDrop) and RNA integrity analysis for EukaryoteTotal RNA (Bioanalyzer (Agilent)) was performed, as described in the RNA 6000 Nano Assay Protocol (Edition Nov. 2003) in order to investigate the quality of all RNA samples.

To remove DNA contaminations, the RNA was treated with TURBO DNAse according to the TURBO DNA-free protocol.

Buffers:

- Bleach (used for embryo dechorionation): 2.8% sodium hypochlorite (50ml DanKlorix bleach, 50ml egg wash)

- Egg wash: 14g NaCl [0.7%], 0.6ml Triton X-100 [0.03%], ad 2L ddH_2O

3.1.2.16 cDNA synthesis

For each RNA sample two cDNA preparations and a negative control without reverse transcriptase (RT) were performed ("cDNA1", "cDNA2" and "no RT"). 9µl RNA (110ng/µl) were transferred into a sterile 1.5ml Eppendorf tube for each reaction and 1µl Oligo dT (500µg/ml), 1µl random 10er and 1µl dNTP (10mM) were added. Then the reactions were heated at 65°C for 5 minutes and quickly chilled on ice. For primer annealing 4µl 5xfirst strand buffer, 2µl 0.1M DTT and 1µl RNAseOUT (40U/µl) were added to the reactions, which were then placed into a 42°C water bath for two minutes. Finally 1µl SuperscriptII RT (200U/µl) was added to the reactions cDNA1 and cDNA2, whereas 1µl H_2O was added to the control "no RT". For cDNA synthesis all reactions were incubated at 42°C (water bath) for 50 minutes. Afterwards the reverse transcriptase was heat inactivated at 70°C for 15 minutes. The cDNA and control preparations were spinned down and stored at -20°C.

3.1.2.17 Real-Time PCR

Real time PCR was performed using the SYBR®Green JumpStart™ Taq ReadyMix™ (Sigma) with a primer concentration of 150nM. Primers are listed in (3.1.1.4). Dilution series of 50ng- 10ng and 2ng (final concentration) genomic template and 1µl, 0.2µl and 0.04µl of each cDNA template were used. In addition the negative control "no RT" and a non-template control with H_2O instead of DNA were included in the experiment.

Real time PCR program:

Denaturation 50°C, 2min
95°C, 2min
94°C, 15sec

40x:

Annealing 60°C, 30sec

Amplification 72°C, 30sec

Melting curve 95°C, 15sec

60-95°C, 20min

95°C, 15sec

Cooling 4°C

3.1.2.18 Single fly PCR

Single fly genomic DNA preparation:

In order to identify transgenic *mCherryLacI* and *mCherryLacI*-recombined flies by PCR, the male progeny of the flies that were injected as embryos were crossed separately to 5 *pin/CyO* balancer line virgins for few days. The anesthesized males were removed from the vials with forceps and transferred into one of the 96-plate wella with 50μl of squishing buffer with freshly added proteinase K on ice. The flies were carefully mashed for 5-10 seconds with pipette tips on a multichannel pipette. The plates were sealed with a pierce-it-lite foil and incubated at 37°C for 30 minutes and then heated to 95°C for 2 minutes to inactivate the proteinase K. The plate was shortly spinned down and stored at 4°C if necessary.

PCR detection:

For single fly PCR, PeqGOLD Taq-DNA Polymerase, Y-mix (Peqlab) was used.

25μl reactions were pipetted into 96 well PCR plates:

12,5μl 2x master mix

5μl primer mix (*5'PCRmCherryflies* and *3'PCRmCherryflies*, for a final concentration of 300nM)

7μl ddH_2O

Large-scale mixes were prepared and 24.5μl were pipetted in each well of the PCR plate. 0.5μl single fly genomic DNA template were added and negative and positive controls with respective ddH_2O and *mCherryLacI* vector DNA as templates were included on each plate. PCR plates were sealed with pierce-it-lite foil.

PCR program:

2’94°C
5x: 30’’94°C
30’’54°C (-1°C/cycle)
45’’68°C
31x: 30’’94°C
30’’56°C
45’’68°C
7’68°C
4°C

Positive transgenics were identified on a 1%agarose gel showing a band for the 550bp PCR product in the *mCherry* coding sequence. Positive progeny was selected and subjected to further analysis.

3.1.2.19 Immunostaining of larval tissue

The protocol used to stain larval tissue against PHO has been taken from Methods in Molecular Biology #420, Drosophila methods and protocols ©Humana Press, *Immunolabeling of Imaginal Discs* by Thomas Klein.
Wildtype and transgenic *GfpPho* third instar larvae were dissected to investigate PHO and GFP::PHO protein distribution by immunolabeling the tissue with a rabbit anti-PHO (1:500) and goat anti-rabbit alexa488 antibody (1:100) combination. DNA was stained with 5μg/ml Hoechst. The preparations were mounted in vectashield (Vector Laboratories, Inc.) and mounted for confocal microscopy.

3.1.2.20 Larval salivary gland squashes and immunostaining of polytene chromsosomes

The protocol used was adapted from Drosophila protocols ©2000 by CSHL Press, *Mapping protein distributions on polytene chromosomes by immunostaining*, R.Paro.

Keeping fly stocks for larvae:
Fly stocks were kept in bottles at 18°C. Every 2-3 days the adult flies were flipped into new bottles. To raise fat larvae, which have bigger salivary glands than under nromal groth conditions, fresh yeast was added to the empty vials containing eggs and first instar larvae. 11-12 day-old larvae were used for polytene chromosome stainings.

Polytene chromosome staining:

Salivary glands of female larvae were dissected in 0.7% NaCl and transferred to a 45µl drop of fix solution on a labeled Poly-L-lysine slide, with each 2 pairs of salivary gland per slide. After 15-20 minutes of incubation a coverslip was placed onto the fixed glands and by moving the coverslip slightly and tapping a used pencil against it the salivary glands were spread and squashed. Then the slide was transferred into liquid nitrogen and the coverslip was flipped off immediately after freezing. The fixed polytene chromosomes were either stored in methanol at 4°C for up to one week or immediately subjected to immunostaining.

Immunostaining:

The slides were first rinsed in 1XPBS before they were washed with rapid shaking two times 10 minutes in 1XPBS and additional 10 minutes in PBST. After washing the slides were incubated in blocking buffer for 40 minutes, followed by rinsing the slides in PBS in a beaker. The slides were dried without touching the chromosome spreads with paper towels. Finally the slides were transferred to a humid box. 50µl of antibody dilution in antibody buffer were placed onto the chromosomes. Carefully a cover slip (24x40mm) was lowered onto the slide. The humid box was closed and the polytene chromosomes were incubated with the antibody at 4°C (cold room) over night.

Primary antibody	conc.	**Secondary antibody**	conc.
anti-PC (rabbit)	1.100	Alexa Fluor® 647 goat anti rabbit	1:250
anti-DSP1 (rabbit)	1:150		
anti-PHO (rabbit)	1:250		
anti-E(z) (rabbit)	1:100		
anti-GFP (mouse)	1:50	Alexa Fluor® 488 goat anti mouse	1:250

Table 2: Antibodies used for immunostaining of polytene chromosomes

Primary and secondary antibody combinations at indicated working concentrations are listed. Information on the origin of the antibodies can be found in chapter*antibodies*.

On next day the slides were rinsed in PBS in order to remove the cover slips. Afterwards the fixed chromosomes were washed three times for 5 minutes in block solution with rapid shaking, followed by rinsing the slides in PBS. After drying each slide on paper towel (without touching the chromosomes) the slides were transferred to a humid box and 50µl

of secondary antibody diluted in secondary antibody buffer were placed onto the chromosomes. Carefully a cover slip was lowered onto each slide and the chromosomes were incubated with the 2nd antibody for one hour at room temperature in darkness.
After incubation the slides were rinsed in PBS in order to remove the cover slips, whereas foil covered containers were used to keep the slides in darkness. Then the stainings were washed for 10 minutes with the washing solution "300" and additional 10 minutes with the washing solution "400".
Finally the slides were rinsed in PBS and placed into a dry box containing paper towels. The slides were shaken dry and left in the box for 10 minutes. A drop of ProLong®Gold antifade reagent with Dapi (Invitrogen) was placed onto the staining and a coverslip was carefully lowered onto the slide. After 24 hours of hardening of the mounting medium at room temperature the stainings were sealed with nail polish, collected in a slide folder and stored at 4°C in the cold room.

Imaging of polytene chromosme stainings:
The polytene chromosome immunostainings were imaged using the microscope Axioplan2 imaging/Coolsnap fx /2.48 (Zeiss) with 25x and 40x objectives. In order to visualize the immunostainings a dedicated set of filters was used. Dapi: Excitation 406/15nm, Emission 457/50nm, exposure time 15msec.; Alexa fluor 488 goat anti-mouse: Excitation 740/40nm, Emission 525/50nm, exposure time 3000msec.; Alexa fluor 647 goat anti rabbit: Excitation 620/60, Emission 700/70, exposure time 500 msec. (Monochromatic light source HBO 103 W Selfadjusting Halogen Lamp 50W.)

Buffers:

- PBST (500mL): 1x PBS, 1% Triton
- 0.7% NaCl: 0.7g NaCl, 100ml ddH_2O
- 37% Paraformaldehyde: 1.85g Paraformaldehyde, 5ml ddH_2O, 70µl 1N KOH heated at 80°C for 30 minutes, subjectede to sterile filtration, 100 aliquots frozen at -20°C
- Block solution (250mL): 1x PBS, 5% non-fat dried milk powder (keep for 2 days at 4°C)
- 1st antibody buffer: (50µL per slide; 10mL prepared, 1mL aliquots frozen at -20°C) 1x PBS, BSA [1%]
- 2nd antibody buffer: (50µL per slide) 1x PBS, 1% BSA, 2% goat serum (made fresh from 1st antibody buffer)

- Wash buffer “300” (500ml): 1x PBS, 5M NaCl (300mM), NP40 (0.2%), Tween 20 (0.2%)
- Wash buffer “400” (500ml): 1x PBS, 5M NaCl (400mM), NP40 (0.2%), Tween 20 (0.2%)

3.1.2.21 Live Imaging

Mounting of Drosophila embryos for live imaging:

0-20min embryos were collected on apple juice plates from fly cages and dechorinated by rolling embryos on double-sided sticky tape. The embryos were immediately transferred into a drop of mounting oil according to the “hanging drop protocol” [125] on a #1.5 cover slip and fixed to a deep-well slide. This slide was mounted upside down onto a inverted laser-scanning microscope, therefore applied as “sitting drop” method.

Mounting of third instar larval tissue

Larval tissue was dissected in PBS, transferred into a Tyrodes buffer drop in a small well on a glass slide, and directly mounted onto the microscope.

Drosophila embryos and tissue:

Live Imaging was performed with the confocal microscope LSM 510/Axiovert (Zeiss) using 25x/0.8 LD LCI plan apochromat oil objective, 40x/1.3 plan-neofluor oil objective, C-Apochromat 63x/1.2 W corr UV-VIS-IR and Plan-Apochromat 63x/1.4 oil DIC objectives. GFP was excited with the 488nm line of an Argon laser and mCherry was excited with 594nm of the HeNe laser. Appropriate filter, laser intensity, pinhole size and detector gain were set.

- Immersion media: immersol 518N (oil), ddH_2O (water)
- Tyrodes Buffer: 135mM NaCl, 10mM KCl, 0.4mM $MgCl_2$, 5.6mM glucose, 10mM HEPES pH7.2

Image Processing:

Data for deconvolution were sampled at Nyquist rate in x, y and z and collected z-stacks were deconvolved using Huygen’s Pro software. Huygens Pro was also used for converting files.

Movies were cut using Deltavision and Metamorph.

3.2 Fly Work and Genetics

3.2.1 Generation of Transgenic Flies using site-specific integration

DNA preparation

Plasmid DNA for injection was prepared with the PureYield^{TM}Plasmid Midiprep System (Promega) (See 3.2.2.1) and eluted with ddH_2O.

Embryo injection

Based on the phiC31 site-specific integration tool the designed constructs were injected into 43.4 landing site line embryos carrying one P-element established defined transgenic locus with an *attP* recombination site (*Sheetal Bhalerao*, IMP/IMBA unpublished) as described by Ringrose,L. in [126].

For each injection 250ng/μl donor plasmid containing the *attB* site and 600ng/μl "helper plasmid pKC40" expressing the phiC31 integrase, which performes *attP-attB* specific recombination, were used. Injections were performed by the in house fly facility.

3.2.2 Fly Strains and Culture

- Landing site line 43.4: $y^- w^-/y^- w^-$; p[43.4],y^+/*CyO*; +/+
  The generated integration site for site-specific integration is at position 38E3 (2L).
- Landing site line 43.16a: $y^- w^-/y^- w^-$; *P*{43.16a},y^+/*CyO*; +/+
  The generated integration site for site-specific integration is at position 46E1 (2R).
- PcGFP: w^{1118}; *P{pPc-PcGFP,*w^+*}*
  Kindly obtained from Arndt-Jovin, D.
- *sc* $z^1 w^{is}$;$E(z)^{63} e^{11}$/*TM3,Sb,Ser*
- *Df(3L)lxd15/TM3,Sb,Ser*
  Kindly obtained from Rick Jones
- $E(z)^{731}$/*TM6C, Tb,Sb*
  Kindly obtained from Jürg Müller
- *C(l)Dx yw/ywDsp1*^{1}*f*
- *C(l)Dx yw/wDsp1*^{1}*f*
- *Dsp1*^{1}*f*

- *Binsn/Dsp1^{1}f*
  From Daniel Locker
- *pho^{1}/ci^{D}*
  from Judith Kassis
- Balancer fly line: *Binscy, w^{1}/C(1)DX, y^{1}, f^{1}*
  (Bloomington stock# 4354)
- Balancer fly line: yw/yw; pin/CyO; +/+
  - Balancer fly line: yw/yw; CyO/sp; +/+
  - Balancer fly line: *Eip74EF^{DL-1}st^{1}p^{p}e^{11}/TM6C, Sb^{1} Tb^{1}*

  (Bloomington, stock# 4435)
- Balancer fly line: *Dr/TM3, Sb, Ser*
- Chromosome IV *ci^{D}/ey^{D}* marked fly line:
  y^{1}*w^{1118};;; In(4)ci^{D},ci^{D}pan^{ciD}/Dp(2;4)ey^{D},Alp^{eyD}:eyD*
  (Bloomington stock# 6645)

Buffers and media for fly work were supplied from the Media Kitchen and Fly Food Kitchen in house.
Fly food:
7.5g Agar, 80g Corn meal, 18g dried yeast, 80g Malzym, 0.5ml O-phosphoric acid, 8.4ml Propionic acid, 10g Soya meal, 22g Sugar beets syrup, ad 1L ddH_2O
Apple juice plates (used for embryo collection):
17.5g Agar, 250ml Apple juice, 10ml Nipagin, 25g sugar, ad 1L ddH_2O

3.2.3 Rescue Experiments

3.2.3.1 Attempt to rescue the *Dsp1f* null mutant by GFP::DSP1

For the purpose of the rescue experiment described in (4.2.1), the following genetic crosses were performed, considering *Dsp1* encoded on the X chromosome:

Binscy, w^{1}/C(1)DX, y^{1}, f^{1} balancer fly lines were crossed to *GfpDsp1* transgenic flies. For each transgenic line, two crosses were set up:

f *y^{-}w^{-}/y^{-}w^{-}; pin/ CyO*

x

m *Binscy, w^{1}/Y; +/+*

This cross was performed to obtain *Binscy,w^{1}/Y; pin/CyO* males.

f *C(1)DX, y^1, f^1/Y; +/+*

x

m *y^-w^-/Y; GfpDsp1/ GfpDsp1*

This cross was performed to obtain *C(1)DX, y^1, f^1/Y; GfpDsp1* females.

As *C(1)DX, y^1, f^1* is not *w^-*, homo- and heterozygosity of the trangene could not be determined.Therefore, the following cross was performed:

M *Binscy,w^1/Y; pin/CyO*

X

C(1)DX, y^1, f^1/Y; GfpDsp1

And subsequently:

C(1)DX, y^1, f^1/Y; GfpDsp1/CyO

x

M *Binscy,w^1/Y; GfpDsp1/CyO*

To get a stable stocks

Binscy, w^1/C(1)DX, y^1, f^1; pα1Tubulin-GfpDsp1/pα1Tubulin-GfpDsp1 and

Binscy, w^1/C(1)DX, y^1, f^1; pα1Tubulin-GfpDsp1/CyO.

A line with first chromosome balancers and the transgene under *pActin5c* promoter expression could not be established.

For the penultimate cross, the stocks were expanded and the following crosses performed (see also Results, 4.2.1, Fig. 10).

Fem Binsn/Dsp1^1f; +/+

x

Binscy,w^1/Y; GfpDsp1/GfpDsp1 (/CyO)

From the progeny the following cross was set up:

Binscy,w^1/ Dsp1^1f; GfpDsp1/+

x

Binsn/Y; GfpDsp1/+

Homo- and heterozygously transgenic flies were screened for by the absence of the *Bar (B,* mutant eye shape) phenotype in males and females. The last crosses were performed at 18°C and 25°C.

3.2.3.2 *pho* rescue experiment

For the purpose of the rescue experiment described in (4.3.1), the following genetic crosses were performed:

y⁻w⁻/y⁻w⁻;;; ci^D/ey^D **x** ***y⁻w⁻/y⁻w⁻; Gfppho/Gfppho***

This cross was performed to create a stable stock carrying the transgene on the second chromosome and marked fourth chromosomes:

y⁻w⁻/y⁻w⁻; Gfppho/Gfppho; ci^D/ey^D

This was arranged for all three stable transgenic fly lines, however only

y⁻w⁻/y⁻w⁻; ppho-Gfppho/ppho-Gfppho;; ci^D/ey^D and

y⁻w⁻/y⁻w⁻; pα1Tubulin-Gfppho/pα1Tubulin-Gfppho;; ci^D/ey^D

were realized. In the case of *pActin5c-Gfppho*, a stable stock of *y⁻w⁻/y⁻w⁻; pActin5c-Gfppho/pActin5c-Gfppho; ci^D/ey^D* could not be established.

For the penultimate cross, the stocks were expanded and the following crosses performed (see also Results, 4.3.1, Fig. 16).

y⁻w⁻/y⁻w⁻; Gfppho/ Gfppho;; ci^D/ey^D

x

y⁻w⁻/y⁻w⁻;;; pho^1/ci^D

Gfppho implicates both *ppho-Gfppho* and *pα1Tubulin-Gfppho.*

The ultimate cross was set up from the progeny of the penultimate cross:

y⁻w⁻/y⁻w⁻; Gfppho/ +;; pho^1/ci^D

x

y⁻w⁻/y⁻w⁻; Gfppho/ +;; pho^1/ci^D

All possible genotypes are listed in (Results, 4.3.1, Fig. 17) and the observed occurrence of each genotype was evaluated as percent of the total number of flies counted in the respective experiment and presented in ratio to the theoretical percentage of these genotypes if GFP::PHO can fully replace endogenous PHO under *α1Tubulin* and the presumed *pho* promoter expression. Rescued homo- and heterozygously transgenic flies were screened for by the absence of the *ciD* phenotype (disrupted wing vein L4).

3.2.3.3 *E(z)* rescue experiment

For the purpose of the rescue experiment described in (4.4.1), the following genetic crosses were performed:

Dr/TM3, Sb, Ser **x** ***Eip74EF^{DL-1}st^{1}p^{p}e^{11}/TM6C, Sb^{1} Tb^{1}***

This cross was performed to generate the *yw/yw; +/+; TM6c,Tb,Sb/Dr* balancer line, in order to have a larval marker in the final rescue experiment.

y^{-}w^{-}/y^{-}w^{-};; TM6c,Tb,Sb/Dr **x** ***sc z^{1}w^{is} ;; E(z)^{63}e^{11}/TM3,Sb,Ser***

From this cross a y^{-}w^{-}/y^{-}w^{-};; *E(z)^{63}e^{11}/TM6c,Tb,Sb* stable stock was generated.

y^{-}w^{-}/y^{-}w^{-};; TM6c,Tb,Sb/Dr **x** ***Df(3L)lxd15/TM3,Sb,Ser***

From this cross a y^{-}w^{-}/y^{-}w^{-};; *Df(3L)lxd15/TM6c,Tb,Sb* stable stock was generated.

y^{-}w^{-}/y^{-}w^{-};; TM6c,Tb,Sb/Dr **x** ***y^{-}w^{-}/y^{-}w^{-}; GfpE(z)/GfpE(z)***

This cross was performed to create a stable stock carrying the transgene on the second chromosome and the third chromosome balancer:

y^{-}w^{-}/y^{-}w^{-}; GfpE(z)/GfpE(z); TM6c,Tb,Sb/Dr

This was arranged for all three stable transgenic fly lines, however only *y^{-}w^{-}/y^{-}w^{-}; pE(z)-GfpE(z)/ pE(z)-GfpE(z); TM6c,Tb,Sb/Dr* and *y^{-}w^{-}/y^{-}w^{-}; pα1Tubulin-GfpE(z)/ pα1Tubulin-GfpE(z); TM6c,Tb,Sb/Dr* were realized. In the case of *pActin5c-GfpE(z)/CyO*, a stable stock of *y^{-}w^{-}/y^{-}w^{-}; pActin5c-GfpE(z)/CyO; TM6c,Tb,Sb/Dr* could not be established.

For the penultimate cross, the previously described stocks were expanded and the following crosses performed (see also Results, 4.4.1, Fig. 22).

y^{-}w^{-}/y^{-}w^{-}; GfpE(z)/GfpE(z); TM6c,Tb,Sb/Dr

x

y^{-}w^{-}/y^{-}w^{-};; *E(z)^{63}e^{11}/TM6c,Tb,Sb*

y^{-}w^{-}/y^{-}w^{-}; GfpE(z)/GfpE(z); TM6c,Tb,Sb/Dr

x

y^{-}w^{-}/y^{-}w^{-};; *E(z)^{731}/TM6c,Tb,Sb*

For simplicity, only one mutant combination is shown, which was the cross that rescued the trans-heterozygous *E(z)^{63}/E(z)^{731}* lethality. Analogously, the crosses were previously attempted to rescue hemizygous *E(z)^{63}/Df(3L)lxd^{15}* and homozygous *E(z)^{63}*, but were not

successful as described in (4.4.1). GfpE(z) implicates both *pE(z)-GfpE(z)* and *pα1Tubulin-GfpE(z).*

y⁻w⁻/y⁻w⁻; GfpE(z)/+; E(z)^{63}e^{11}/TM6c,Tb,Sb

x

y⁻w⁻/y⁻w⁻; GfpE(z)/+; E(z)^{731}/TM6c,Tb,Sb

From the progeny of the penultimate cross, the ultimate cross was set up as indicated. The ultimate cross gives the result of the rescue in the progeny.

All possible genotypes are listed in (Results, 4.4.1, Fig. 23) and the observed occurrence of each genotype was evaluated as percent of the total number of flies counted in the respective experiment and presented in ratio to the theoretical percentage of these genotypes if GFP::E(Z) can fully replace endogenous E(Z) under *α1Tubulin* and the presumed *E(z)* promoter expression. Rescued homo- and heterozygously transgenic flies were screened for by the absence of the *Stubble* (*Sb*, short bristles) and *Tubby* (*Tb*, longitudinally shortened larvae and pupae) phenotypes that mark the balancer chromosome.

4 Results

4.1 Stable transgenic fly lines for live imaging of PcG proteins

Fluorescent tagging and quantitative live imaging are established as powerful tools to study the dynamics of molecular processes *in vivo*. The autofluorescent *green fluorescent protein* (*Gfp*) from the jellyfish *Aequorea victoria* (reviewed in [127]) is broadly used to study proteins in intact cells and organisms as it can easily be fused to a protein of interest by molecular cloning techniques.

To gain understanding of the PcG and TrxG mediated memory mechanism and its regulation, the dynamic behaviour functional PcG repression complexes, TrxG activation complexes and complexes of DNA binding factors, which recruit other PcG/TrxG proteins, have to be studied. PcG proteins bind dynamically to chromatin and complex stabilities vary at different loci on polytene chromosomes of larval salivary glands [115, 116]. To enable the quantification of the kinetic behavior of this system, I selected important members of different complexes involved in the PRE/TRE-mediated regulatory mechanism and tagged them fluorescently with *Gfp*.

Four proteins of different complexes were selected. *E(z)* belongs to the PcG and is part of the Polycomb Repressive Complex 2 (PRC2). *ash1* is a member of the TrxG and like *E(z),* codes for an enzyme with histone methyl transferase activity. *pho* and *Dsp1* are DNA-binding proteins and important factors for the recruitment of PcG complexes to Polycomb Response Elements (PREs). Neither PRC2 members, DNA-binding proteins nor TrxG proteins have been studied in terms of their kinetic properties.

The project on *ash1* was perfomed by Eva Dworschak under my supervision has been published in her diploma thesis (*Generation and Characterization of Ash1-EGFP Transgenic Fly Lines for Live Imaging of Ash1*, at Fachhochschule Campus Wien, 2009).

In order to test for the best expression level of functional GFP-fusion proteins I designed cloning strategies to obtain all four proteins either N- or C-terminally tagged with GFP, and each protein under control of 3 different promoters (*Actin5c*, *α1Tubulin* and the presumed endogenous promoter). The cloned *Actin5c* and *α1Tubulin* promoters come from constitutative genes and are ubiquitously expressed and were selected for strong transgene expression in all tissues and developmental stages. To simulate the expression

of the endogenous proteins of interest with the transgenes, the presumed endogenous promoters were cloned. As the endogenous promoters have not been described, the upstream sequence of the coding region of the respective protein of interest, until the gene region of the adjacent gene as defined in flybase was cloned as endogenous promoter sequence.

The coding sequence of the *Gfp* mutant *enhanced Gfp (Egfp)* [128] was used for tagging of all proteins, although *Gfp* in this work refers for simplicity to *Egfp*.

All constructed plasmids were based on the pKC27 vector (*Kuang-chung Su* and *Barry Dickson*, unpublished; described in detail in [126]) carrying the *attB* recognition site for the phage ΦC31 integrase [129] and containing the 5' coding sequence of the transformation marker *miniwhite*. The promoter-fusion protein cassette was cloned between AscI and AvrII restriction sites to be able to combine a functional and characterized tagged protein of interest with the mCherry tagged LacI for studying single-locus kinetics at a transgenic PRE. (For single locus-tagging see chapter 4.6 and for details on the cloning strategy see Materials and Methods.)

Transgenic flies were generated by microinjection of 43.4 landing site line embryos (y^-,w^-; p{pKC43.4}; +), which carry a single *attP* recognition site for the ΦC31 integrase on chromosome 2L at genomic position 38E3, and the 3' coding sequence of *miniwhite*, established by P-element transgenesis (*Sheetal Bhalerao*, unpublished). Supply of the integrase from a co-injected helper plasmid was used to achieve site-specific integration of the vector into the landing site. The complemented *miniwhite* transformation marker gene facilitated screening for transgenics by eye colour phenotype. (See 3. Materials and Methods.)

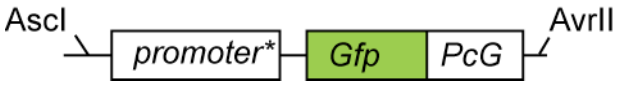

Figure 8: Schematic representation of GFP::PcG fusion protein transgenes

Selected proteins of interest (*PcG*) were fused either N-terminally (as indicated in Figure) or C-terminally to the *Egfp* DNA sequence (*Gfp*) and downstream of one out of three different promoters to test for best fusion protein expression (*… the presumed endogenous, the *Actin5c*, or the α*1Tubulin* promoter). An SV40 termination signal was placed downstream of the PcG protein sequences (not indicated in Figure). The plasmids were constructed on pKC27 vector backbones and two unique restriction sites were inserted adjacent to the cassette for recombining the construct with the mCherry::LacI vector for the locus-tagging project (See chapter 4.6).

protein	promoter	designation	stable transgenic line
GFP::E(Z)	*Actin5c* *α1Tubulin* *E(z)*	*pActin5c-GfpE(z)* *pα1Tubulin-GfpE(z)* *pE(z)-GfpE(z)*	✓ (het) ✓ ✓
GFP::DSP1	*Actin5c* *α1Tubulin* *Dsp1*	*pActin5c-GfpDsp1* *pα1Tubulin-GfpDsp1* *pDsp1-GfpDsp1*	x ✓ ✓ (het)
GFP::PHO	*Actin5c* *α1Tubulin* *pho*	*pActin5c-Gfppho* *pα1Tubulin-Gfppho* *ppho-Gfppho*	✓ ✓ ✓
GFP::ASH1*	*Actin5c* *α1Tubulin* *ash1*	*pActin5c-Gfpash1* *pα1Tubulin-Gfpash1* *pash1-Gfpash1*	x ✓ ✓

Figure 9: Fluorescently tagged PcG proteins expressed in stable transgenic fly lines

Constructs for N-terminally tagged fusion proteins under the *Actin5c*, the *α1Tubulin* and the presumed endogenous promoters, containing the genomic DNA sequence upstream of the coding sequence of the respective gene until the adjacent gene region according to flybase Gbrowse, were injected into embryos of the 43.4 landing site line and transgenes selected by the *miniwhite* transformation marker. Stable stocks were established by crossing to the *pin*/CyO balancer line and a subsequent cross to generate homozygous transgenes. "p" promoter, "X" no transgenes obtained, *established and analyzed by Eva Dworschak

All N-terminally tagged constructs were injected, which were regarded as less likely to disturb protein functions than the C-terminally tagged constructs based on the published information on tagging these proteins for purification and antibody production [121, 130-132]. Transgenes were found in several rounds of injection for almost all constructs (Fig.9), however a *pActin5c-GfpDsp1* transgenic fly was not obtained, possibly due to strong overexpression under the *Actin5c* promoter. Consistent with this interpretation, the presumed *Dsp1* promoter, which strongly expresses the transgene (4.2.2), was only obtained in a heterozygous transgenic fly line, indicating homozygous lethality. Interestingly, *pActin5c-GfpE(z)* was not stable in a homozygous stock, again suggesting an overdose in protein expression.

Summarizing this chapter, transgenic fly lines expressing GFP-tagged proteins of interest under different promoters have been established. They serve the basis for establishing a set of selected members of the PcG, TrxG and DNA binding proteins with the best

expression level for live imaging experiments. Detailed characterization of the transgenic lines is described in (4.2–4.4) and (*Generation and Characterization of Ash1-EGFP Transgenic Fly Lines for Live Imaging of Ash1* Diploma thesis, Eva Dworschak, 2009).

4.2 Quantitative analysis and live imaging of GFP::DSP1

4.2.1 Transgenic GFP::DSP1 fusion protein binds to correct target sites on larval polytene chromosomes

The GFP tagged DSP1 has to be a functional protein to yield conclusive results in live imaging experiments. The most conclusive proof for the functionality of the GFP::DSP1 fusion protein is a rescue experiment, testing whether the fusion protein as the only expressed copy of *Dsp1* can rescue the homozygous lethality of *Dsp1* mutants. Furthermore, to detect if the GFP::DSP1 fusion protein binds the same target loci as endogenous DSP1 and therefore participates in correct complex formation, colocalization studies on polytene chromosomes from larval salivary glands were performed by immunostaining with antibodies against GFP and DSP1.

***GfpDsp1* transgenes do not rescue the *Dsp1*^{1} chromosome**

In order to test whether GFP::DSP1 can fulfill all essential functions of DSP1, the transgenes were crossed to a null-mutant background. (See also 3. Materials and Methods.)

Dsp1^{1} [130, 133] is a loss-of function allele that has been characterized in detail [10] and the only described lethal mutation of *Dsp1*. Homozygous or hemizygous *Dsp1*^{1} flies can be maintained at temperatures below 22°C and show severe morphological and physiological defects, die prematurely and show very low fertility. At higher temperatures, only balanced *Dsp1*^{1} mutants were observed. *Dsp1*^{1} adult males show the phenotype of a reduced number of sex combs, with an average of 6 instead of the 11 wildtype teeth.

I used the *Dsp1*^{1} allele to test whether *GfpDsp1* under the control of the presumed *Dsp1* promoter or α*1Tubulin* promoter can rescue the loss-of-function phenotype. The mutant allele was kept over the chromosome I balancer Binsn. Binsn/*Dsp1*^{1} females were crossed to males of stable stocks that were C(I)Dx,*y,f*/Binscy,*w* on the first chromosome and *pDsp1-GfpDsp1/CyO* or *p*α*1Tubulin-GfpDsp1/p*α*1Tubulin-GfpDsp1* on the second chromosome (Fig. 10 cross (1)). Both Binsn and Binscy,*w* chromosomes are *In(1)sc*^{S1L}*sc*^{8R}*+dl49* balancers, which are viable and fertile in both males and homozyous females and effective in suppressing crossing over with normal X-chromosome

sequences (Lindsley and Zimm, 1992). They are marked with the viable dominant mutation B^1, visible in an eye phenotype.

In the progeny of these crosses, no hemizygous $Dsp1^1$/Y males with B^+ phenotype were detected. Instead all males were B^1, indicating the Binsn/Y genotype. From the progeny of the first cross (Fig. 10 cross (1)), Binscy,*w*/$Dsp1^1$;*pDsp1-GfpDsp1*/+ and Binscy,w/$Dsp1^1$;*pα1Tubulin-GfpDsp1*/+ females were crossed to Binsn/Y;*pDsp1-GfpDsp1*/+ and Binsn/Y;*pα1Tubulin-GfpDsp1*/+ males, respectively (Fig. 10 cross (2)).

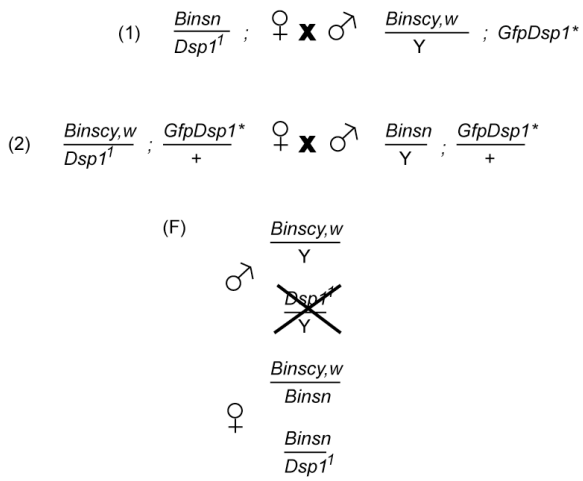

Figure 10: Attempt to rescue the *Dsp1^1* null allele by transgenic *GfpDsp1* expression

Binsn/Dsp1^1 females were crossed to *Binscy,w/Y* males of the *C(I)Dxyf/Binscy,w* balancer stock (1). Progeny of cross (1) with indicated genotypes (2) were selected and recrossed. Possible first chromosome genotypes in the next generation are listed in (F), however no $Dsp1^1$/Y males were obtained. **GfpDsp1* implicates crosses for both the *pDsp1-GfpDsp1*/CyO and *pα1Tubulin-GfpDsp1* rescue experiments.

As mentioned above, the $Dsp1^1$ mutation is temperature sensitive. Therefore, two crosses at different temperatures (18°C and 25°C) were performed for each genotype.

All flies hatching from crosses of both genotypes and at different temperatures carried balancer chromosomes as identified by the B^1 phenotype and males did not have a reduced number of sex combs (Fig. 10 (F)). Conclusively, none of the progeny of this cross gave hemizygous $Dsp1^1$/Y males, independently of the temperature at which the crosses were kept. Therefore, no further cross to get homozygous $Dsp1^1$ females could be performed. Although $Dsp1^1$/Y males were expected at temperatures below 22°C, these flies could not be obtained. Thus, I conclude that the $Dsp1^1$ chromosome was recessively lethal at all temperatures, which could be due to an additionally gained lethal recessive

mutation or *Dsp1*^{1}-mediated hypersensitivity to environmental impact. Remarkably, a large number of pupae did not hatch in all crosses.

Binscy,*w*/Binsn and Binscy,*w*/*Dsp1*^{1} or Binsn/*Dsp1*^{1} females (Fig. 10 (F)) could be distinguished by the strength of the *B*^{1} phenotype. Binscy,*w*/Binsn females showed a stronger deformed eye similar to Binscy,*w*/Y or Binsn/Y males. This phenotypic difference revealed that Binscy,*w*/*Dsp1*^{1} or Binsn/*Dsp1*^{1} comprised 2/3 of all female progeny instead of 50%, probably due to a negative effect by two *In(1)sc*^{S1L}*sc*^{8R}*+dl49* balancers on viability. In the rescue experiment performed, *GfpDsp1* does not compensate for lethality by the *Dsp1*^{1} chromosome, which is temperature independent. This is in conflict with the description of homozygous and hemizygous *Dsp1*^{1} flies obtained at temperatures lower than 22°C [130] and could be due to an additionally gained mutation on the *Dsp1*^{1} chromosome. Therefore, this *Dsp1*^{1} chromosome is not useful for a rescue experiment with *GfpDsp1*.

GFP::DSP1 and DSP1 bind the same target loci on salivary gland polytene chromosomes from transgenic *pα1Tubulin-GfpDsp1* larvae

The expression of *GfpDsp1* transgenes could not rescue the null allele lethality of *Dsp1*^{1}. As this could have other reasons besides GFP::DSP1 not being a functional protein as discussed above, I wished to ascertain whether GFP::DSP1 can bind to DSP1 target sites and therefore is able to form correctly targeted complexes. To test this, I performed coimmunostainings of larval polytene chromosomes from both stable transgenic fly lines. Salivary gland squashes and polytene chromosome immunostaining of *pDsp1-GfpDsp1/CyO* and *pα1Tubulin-GfpDsp1* transgenic larvae were performed as described in detail in Materials and Methods. The monoclonal mouse anti-GFP and polyclonal rabbit anti-DSP1 primary antibodies were tested and showed specific banding patterns on polytene chromosomes from transgenic *PcGfp* or wildtype larvae, respectively (data not shown). Secondary antibodies were anti-rabbit alexa647 and anti-mouse alexa488 conjugated antibodies giving red and green fluorescent signals (see Fig. 11) when excited with the appropriate wavelength.

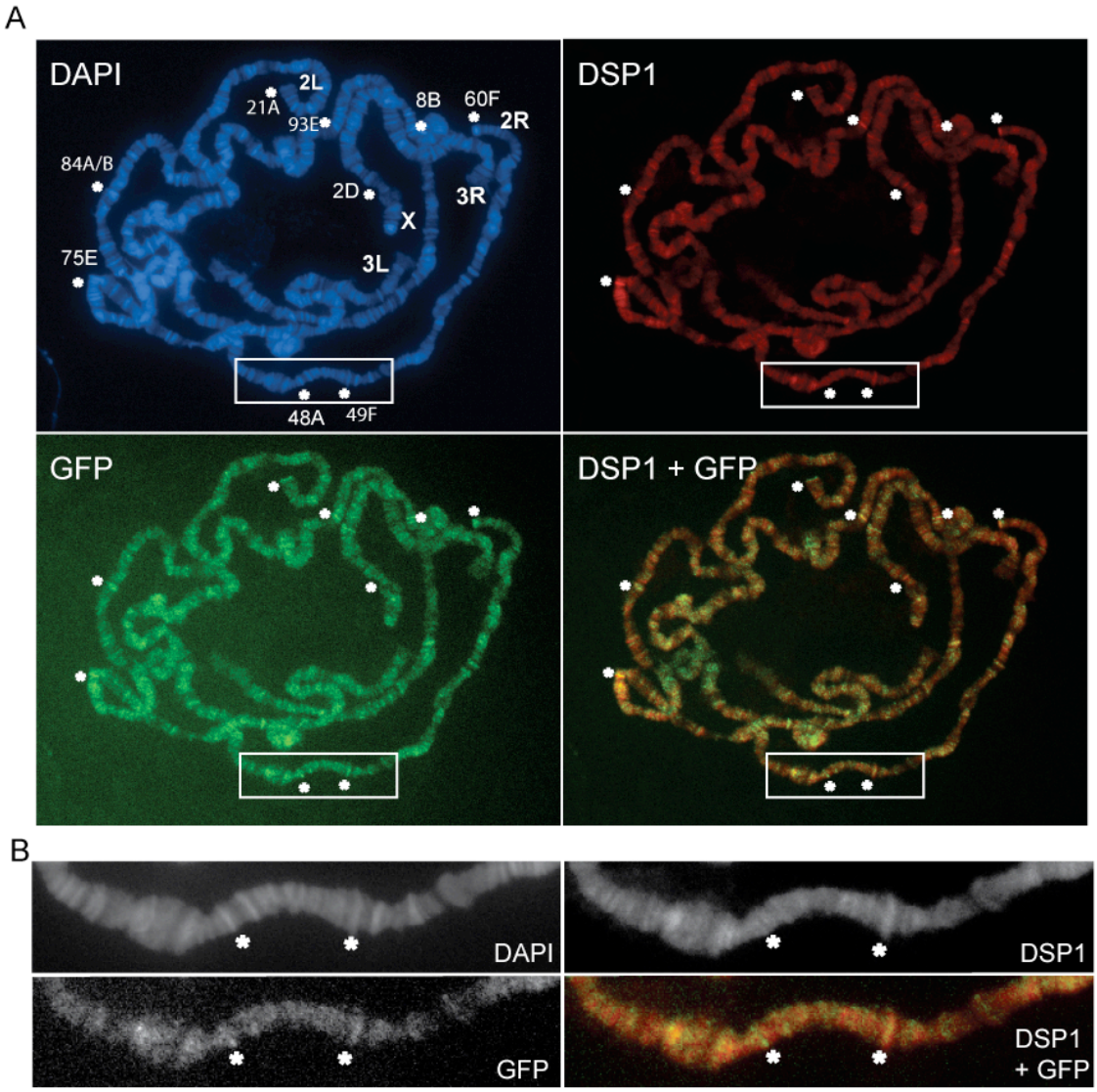

Figure 11: Immunostaining of polytene chromosomes from transgene *pα1Tubulin-GfpDsp1* larvae

Polytene chromosomes were costained with anti-mouse alexa488/mouse anti-GFP and anti-rabbit alexa647/rabbit anti-DSP1 antibody combinations. DNA was stained with DAPI. Chromosome arms were identified and prominent colocalizing bands were mapped (A) according to Lindsley and Zimm (1992). Selected sites that correspond to known PC sites were marked with asterisks. (B) shows a magnification on chromosome arm 2R and colocalization on the *engrailed* (48A) and *Psc* (49F) PRE/TRE loci is marked as in (A).

Due to a weak specific signal and therefore low signal-to-noise ratio from the anti-GFP antibody and a low abundance of fusion protein level (see 4.2.2), GFP from *pDsp1-GfpDsp1/+* larval salivary glands could not be detected for sufficient image quality. Coimmunostained polytene chromosomes from *pα1Tubulin-GfpDsp1* transgenic larvae

revealed specific signals from GFP and DSP1 binding. Anti-DSP1 detection revealed a better signal-to noise ratio than anti-GFP staining and bound to many loci as described previously [64].

Analysis of GFP and DSP1 binding patterns illustrate almost complete overlap of the signals detected, minor differences appear as differences in staining efficiencies of the antibodies used. (Fig. 11 A and B)

Strong colocalizing bands were mapped on depicted chromosome arms and compared to previously published PC binding sites [134]. Selected sites are marked in Fig A and B. Prominantly, the homeotic *antennapedia complex (ANT-C)* at 84A/B on chromosome arm 3R, the *engrailed* locus and the nearby *Psc* locus on chromosome arm 2R (magnified in Fig. 11 B) are examples of PC-bound and PRE/TRE containing loci [71].

Summarizing these results, GFP and DSP1 binding colocalize, which demonstrates that GFP::DSP1 and DSP1 protein can bind the same target sites.

4.2.2 Different transgenic fly lines reveal different levels of GFP::DSP1 fusion protein in the embryo and salivary glands

The PcG gene *polyhomeotic* (*ph*) is controlled by feedback inhibition of the *ph* promoter [135] and the transgenic expression of a GFP tagged fusion protein has been shown to affect the endogenous relative expression of the proximal to the distal *ph* genes. The proximal gene of *ph* was used for PH::GFP*,* and its expression caused an increase in endogenous proximal PH, and a decrease in distal PH protein levels [115].

To define the levels of fusion protein from the two different expression strategies and to investigate whether there is an effect of transgenic protein expression on the endogenous protein level I quantified protein levels on Western Blots of total protein extracts from embryos and larval salivary glands of wildtype and stable transgenic fly lines.

The fluorescent tagging with EGFP gave rise to a fusion protein of 70kDa predicted MW. Western blot analysis with an anti-DSP1 antibody showed that both the fusion protein and the 44kDa endogenous protein ran at higher molecular weight relative to the marker proteins in SDS-PAGE (Fig A). Similar results have been described for *Dsp1* in [130].

For detecting the signals on the Western Blots I used the conventional method of densitometry of developed a chemiluminescence films, and in addition I captured the emitted photons directly in the Kodak Molecular Imager (See 3. Material and Methods). Despite being less sensitive, use of the second applied method was possible due to clean and strong signals. Both techniques measured similar ratios of the amount of GFP::DSP1

to the amount of endogenous DSP1 as can be seen by size of the standard deviation in Fig. 12 (B).

Transgenic expression driven by the presumed *Dsp1* promoter did not result in viable homozygous flies. However, a stable heterozygous fly line was established which revealed the fusion protein being expressed at approximately 0.1-fold of level of the endogenous protein in embryos and larval salivary glands. The stable homozygous fly line constitutatively expressing GFP::DSP1 from the *α1Tubulin* promoter showed a similar expression level of the fusion and the endogenous proteins in the embryo but approximately two fold more fusion protein in salivary glands. (See Fig. 12 A and B).

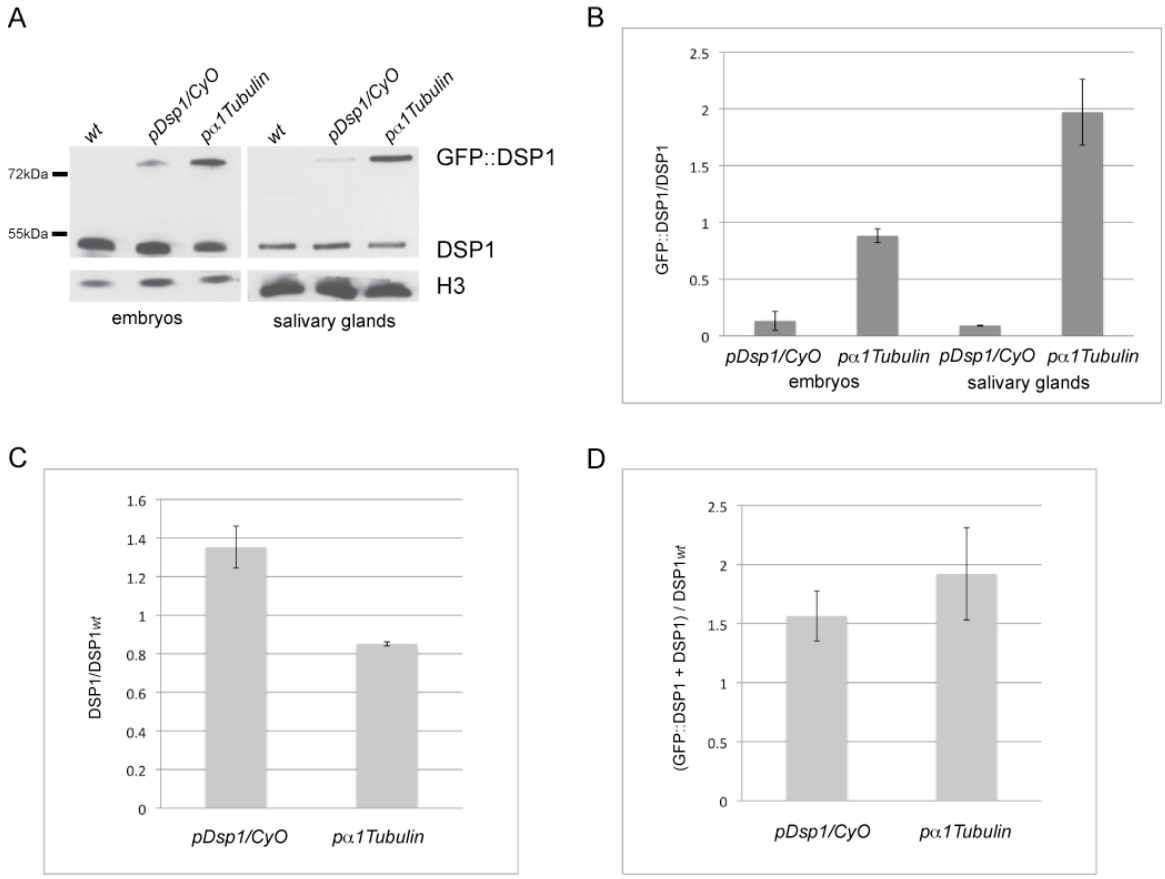

Figure 12: Western blot and quantification of DSP1 and GFP::DSP1 proteins in transgenic flies.

GFP::DSP1 fusion protein expression in transgenic fly lines was induced using the presumed endogenous *Dsp1* promoter (*pDsp1*) or the *α1Tubulin* promoter (*pα1Tubulin*) and expression levels in embryos and salivary glands were quantified on a film (A) or on the Kodak Molecular Imager. The ratio of the GFP fusion protein compared to the level of endogenous DSP1 in both experiments is shown in (B). Error bars in (B) reflect the standard deviation obtained by averaging the results of one experiment of each applied method. Endogenous and total DSP1 protein levels in

transgenic extracts of embryos and salivary glands on the western blot developed on film were normalized to the amount of protein loaded and compared to respective wildtype (wt) levels. Results from embryos and salivary glands of each transgenic line were combined. (C and D)

In order to evaluate whether the transgenic GFP::DSP1 expression affects the endogenous protein level, I normalized the signal intensities of endogenous DSP1 to the wildtype DSP1 level by the amount of protein loaded according to the loading control detected by an anti-histone H3 antibody in the experiment developed on a chemiluminescence film.

In both embryos and salivary glands, the endogenous DSP1 levels in the transgenic fly lines increased to approximately 1.3 fold when the presumed *Dsp1* promoter controlled the transgene, and decreased to 0.8 fold in the constitutatively expressing *pα1Tubulin* line. (Fig. 12 C) This result suggests that the heterozygous expression of GFP::DSP1 protein under the control of the endogenous promoter has an upregulating effect on endogenous DSP1 levels and this might be increased to a level by homozygous expression that is lethal.

To ascertain the increase of total DSP1 protein in transgenic fly lines, the sum of GFP::DSP1 normalized to wildtype DSP1 levels and normalized values of endogenous DSP1 levels from embryos and salivary glands of each transgenic fly line were plotted in ratio to the levels of wildtype DSP1.

This showed that in both lines the total DSP1 level is less than two fold above the DSP1 level in wildtype flies. Although the difference in the levels of GFP::DSP1 in the transgenic lines were significant, the total DSP1 protein levels were similar in both lines relative to the DSP1 level in wildtype flies (Fig. 12 D).

These results demonstrate that the different expression strategies for GFP::DSP1 have intrinsic differences in terms of both fusion protein level and effect on the amount of endogenous DSP1.

4.2.3 GFP::DSP1 localizes to specific bands on polytene chromosomes in larval salivary gland nuclei *in vivo*

Immunostaining against GFP on transgenic polytene chromosomes from fixed larval salivary glands revealed binding to descrete sites (4.2.1). To investigate whether the GFP::DSP1 fusion protein is expressed sufficiently for live imaging experiments on

specific bands, I dissected salivary glands from transgenic third instar larvae and tested imaging conditions for confocal microscopy.

As expected from quantification of protein levels (4.2.2), there was more GFP signal detected from transgenic larvae that were homozygous for *pα1Tubulin-GfpDsp1* (Fig. 13 B) than from the *pDsp1-GfpDsp1/CyO* (Fig. 13 A) line. The fluorescence intensity was sufficient to reduce the laser output to 1-2% (see Fig.13). Under these imaging conditions discrete binding loci on the polytene chromosomes were detected without bleaching.

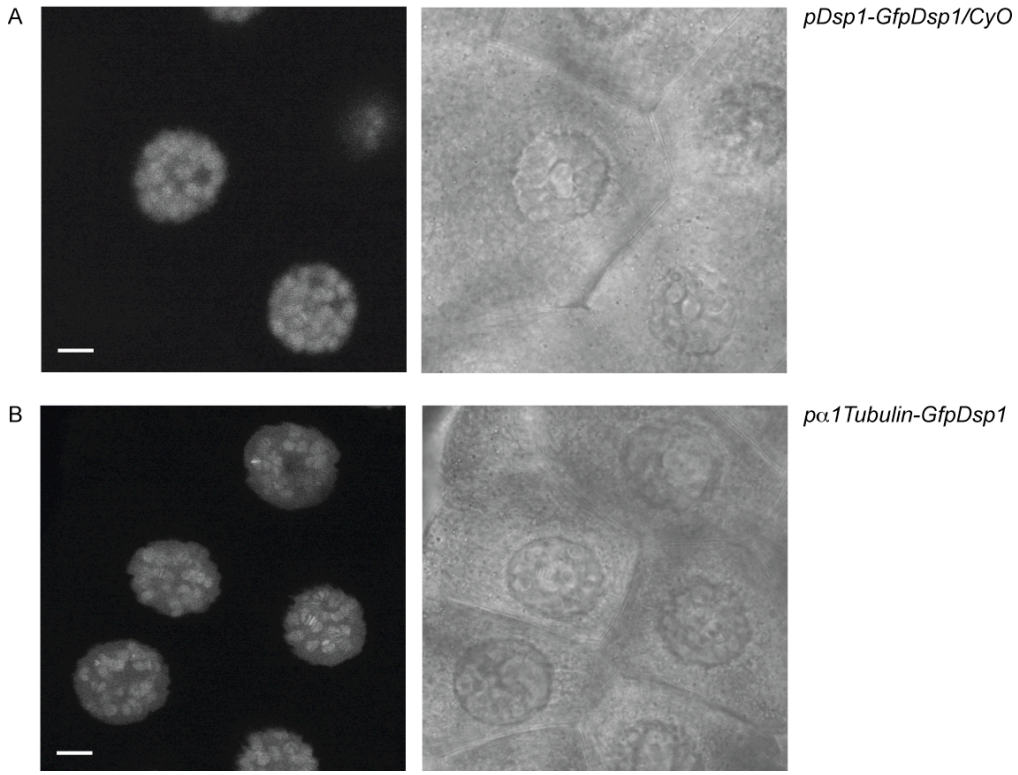

Figure 13: Live Imaging of GFP::DSP1 in larval salivary gland nuclei

GFP::DSP1 distribution in living salivary gland nuclei is shown of *pDsp1-GfpDsp1/CyO* (A) and *pα1Tubulin-GfpDsp1* (B) transgenic larvae. The images were taken under conditions for live imaging with a 40x 1.3NA oil immersion objective. 488nm excitation with 2.1% laser output and a pinhole of 260μm in (A) and 1.1% laser output and a pinhole of 66μm (1 Airy Unit) in (B) was used. The corresponding DIC images are shown on the right. Scaling was x: 0.11μm, y: 0.11μm, z: 0.6μm. The white bars represent 100x10 pixels.

Taken together these results identify the GFP::DSP1 expression levels in both stable transgenic lines as sufficient for live imaging experiments in larval salivary gland nuclei. The GFP fusion protein binds specific sites in vivo, which can be further investigated.

4.2.4 GFP::DSP1 is detected on mitotic chromosomes during synchronous nuclear divisions in early embryo development

Dsp1 codes for a high mobility group (HMG) 1-like DNA-binding protein [136]. It has been shown to be an essential factor for PcG protein recruitment and silencing in a transgenic assay *in vivo* [64] and therefore might be one of the first factors establishing silencing after mitosis. To address whether this protein stays on chromatin during restructuring changes at mitosis I investigated GFP::DSP1 disposition at nuclear divisions in early embryo development.

In zygotes of *Drosophila*, the first thirteen nuclear divisions (stage 1-4 of embryo development) occur syncytially. After division seven most of the nuclei move outwards as they continue to divide. After nine divisions they are mostly out near the egg surface and metasynchronous divisions 10-13 proceed in waves. After the 13th division in cycle 14, cellularization takes place. A monolayer of cells is formed by plasma membrane extension centripedally into the egg and forming the blastoderm embryo (stage 5 of embryo development).

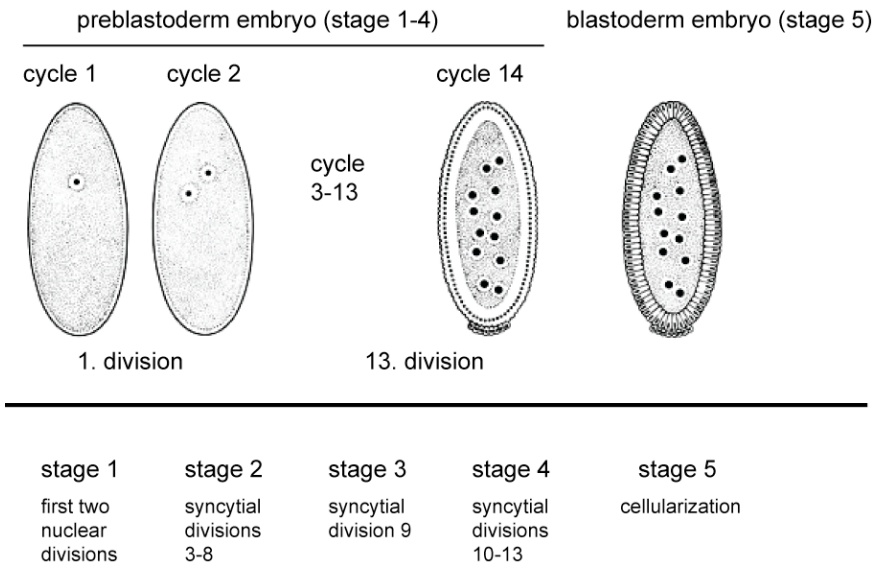

Figure 14: Early embryo development and nuclear divisions imaged in transgenic embryos

Volker Hartenstein and José Campos-Ortega have subdivided the complete embryonic development of Drosophila melanogaster into 17 stages (Flymove, [137]). This figure shows the first 5 stages, which were imaged in embryos of transgenic GFP fusion protein expressing fly lines. The preblastoderm embryo (stages 1-4) is characterized by 13 syncytial nuclear divisions and lasts

approximately two hours. At cycle 14 cellularization forms a peripheral monolayer of cells, the blastoderm. Images were adapted from Foe and Alberts, 1983 [138].

Dechorionated preblastoderm embryos were imaged with a laser-scanning confocal microscope as described in Materials and Methods.

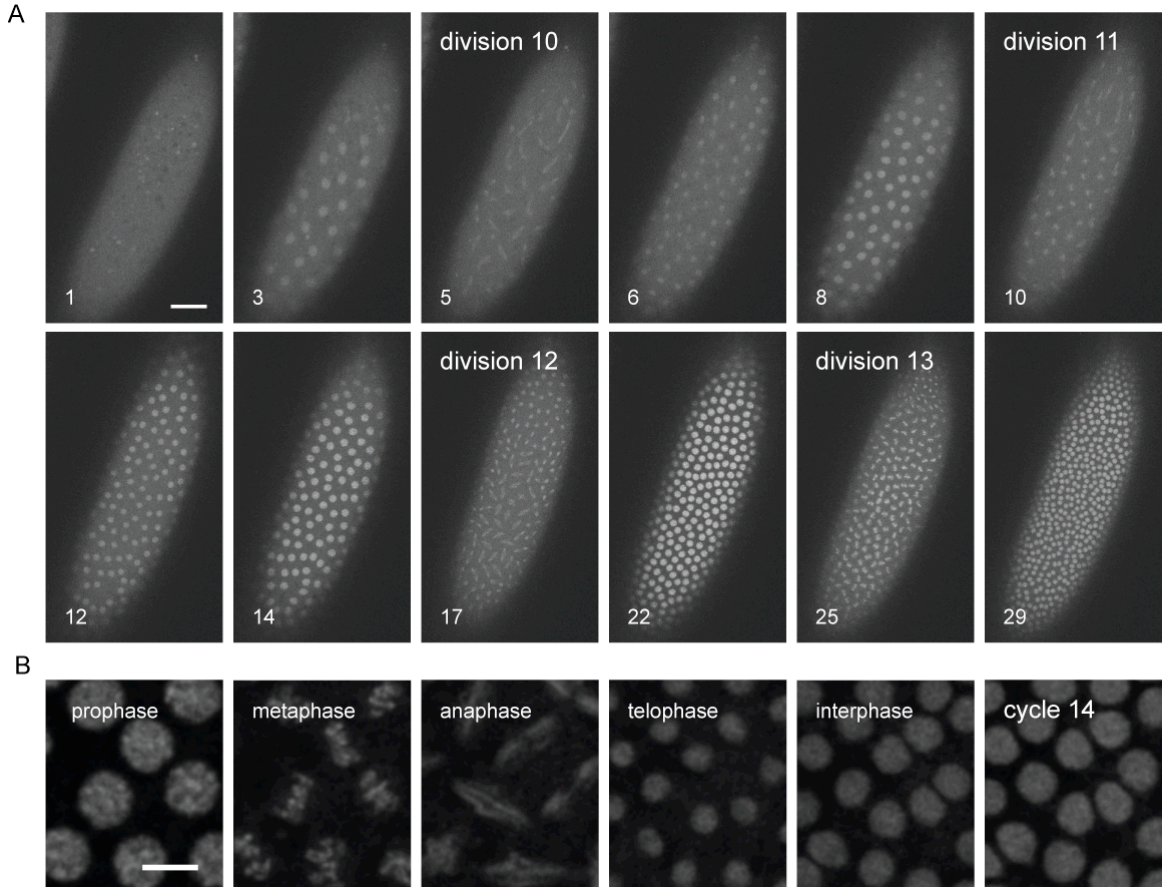

Figure 15: Live Imaging of GFP::DSP1 during metasynchronous nuclear divisions (cycles 10-14) in transgenic preblastoderm embryos

Nuclear GFP::DSP1 in *pα1Tubulin-GfpDsp1* transgenic embryos was detected from stage 4 (cycle 10) on and imaged during subsequent cell divisions. The 13th division between cycle 13 and 14 is shown in (B). Movies were generated taking a new z-stack every 91sec (A) or 163sec and deconvolved (B). Numbers in (A) indicate the number of the z stack acquired. Pictures are presented as maximum intensity projections of a selected number of slices. A 25x/0.8 NA oil immersion corrected objective, 5.1% laser output at 488nm and a pinhole of 174µm were used for imaging. Scaling was in (A) x: 0.35µm, y: 0.35µm, z: 0.5µm and zoomed in (B) x: 0.07µm, y: 0.07µm, z: 0.5µm. The white bars represent 100x10 pixels.

Before stage 4, fluorescence was observed only in the egg cytoplasm and yolk granules, but at cycle 10 the synchronously dividing nuclei, which had moved peripherally, showed clear GFP fusion protein expression. During subsequent metasynchronous nuclear divisions, which were observed to occur in waves, GFP::DSP1 was located on the chromosomes during all steps of mitosis (Fig. 15). At interphase the fluorescence in the nuclei increased until next prophase, metaphase and anaphase, where the condensed chromatin showed strong fluorescence. The number and size of nuclei changed as has been described [112, 138] and the duration of cleavage divisions 10-13 increased progressively. (Fig. 15)

Taken together these findings demonstrate that transgenic GFP::DSP1 expression in the early embryo fulfills the requirements for live imaging of DSP1 during the cell cycle. As DSP1 has many functions beyond being a component of the PcG silencing mechanism, the correlation of DSP1 remaining on mitotic chromosomes and PcG protein mediated memory still have to be ascertained.

4.3 Quantitative analysis and live imaging of GFP::PHO

4.3.1 GFP::PHO is a functional fusion protein

To verify the functionality of a fusion protein it must be able to rescue a homozygously lethal mutant phenotype and accordingly fully replace the endogenous protein. *pho*^{1} is a presumed null mutant with a retrotransposon insertion upstream of the DNA binding domain [35], and homozygous flies die at the pupal stage [87].

To evaluate whether GFP:PHO fusion proteins can rescue the lethality of *pho*^{1} homozygotes, the cross shown in Fig. 16 was performed. Briefly, *pα1Tubulin-Gfppho* and *ppho-Gfppho* transgenic flies were crossed to a *pho*^{1}/*ci*^{D} mutant background and homozygous *pho*^{1} mutants in the progeny were identified by the lack of the fourth chromosome marked by the dominant mutation of *ci*^{D} (*cubitus interruptus*), which creates wings with a disrupted vein L4. The progeny was analyzed with respect to the copynumber of the transgenic *Gfppho* on the second chromosome. (For details see 3. Materials and Methods.) A *pActin5c-Gfppho* fly line with the marked fourth chromosomes *ci*^{D}/*ey*^{D}, which is necessary for the genetic crosses of the rescue experiment, could not be established. The rescue experiment was also performed with *pα1Tubulin-Gfppho* transgenic flies and a fourth chromosome marked by the dominant *eyeless* mutation *ey*^{D}, exhibiting an eye phenotype. However, the *ey*^{D} phenotype was partially suppressed by the *Gfppho* transgene. PcG proteins have been identified to regulate the eye development

[139] and PHO has been mapped on *eyeless* in chromatin immunoprecipitation [75]. Therefore, this experiment was not evaluated.

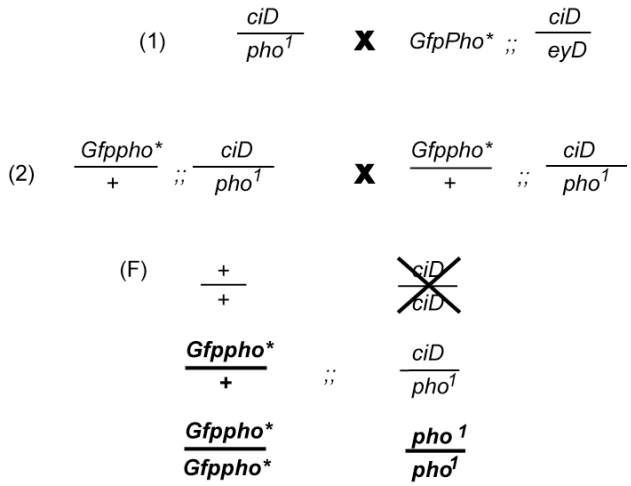

Figure 16: Rescue of the *pho[1]* null allele by transgenic *Gfppho* expression

ci^D/pho^1 flies were crossed to homozygous *Gfppho* flies that carried marked 4th chromosomes (ci^D/ey^D) (1). Progeny of cross (1) with indicated genotypes (2) were selected and recrossed. Possible second and fourth chromosome genotypes in the next generation are listed in (F). The observed combinations of second and fourth chromosome genotypes that are presented in bold letters constitute the rescue. **Gfppho* implicates crosses for both the *ppho-Gfppho* and *pα1Tubulin-Gfppho* rescue experiments.

The experiment with *Gfppho* regulated by the *α1Tubulin* promoter resulted in a rescue, however the expression with the cloned presumed *pho* promoter did not yield in statistically interpretable data, as only 3 out of 416 counted males and no female fly out of 398 counted were homozygous for the *pho[1]* mutation. The observed expression levels in 4.3.2 indicate that the presumed *pho* promoter is not sufficient to drive transgenic *Gfppho* expression to an adequate level such that flies survive without an endogenous copy of *pho*.

genotype	phenotype	**no rescue** (expected % of total number of flies)	**full rescue** (expected % of total number of flies)	**flies counted (observed % of total flies counted)**
pα1Tubulin-Gfppho / pα1Tubulin-Gfppho ;; pho1 / ci D	dark orange *ciD* wing	25%	18,2%	78(18,6%) 144(20,7%)
pα1Tubulin-Gfppho / pα1Tubulin-Gfppho ;; pho1 / pho1	dark orange *wt* wing	0%	9,1%	35(8,4%) 73(10.5%)
pα1Tubulin-Gfppho / + ;; pho1 / ci D	light orange *ciD* wing	50%	36,4%	215(51,3%) 322(46,3%)
pα1Tubulin-Gfppho / + ;; pho1 / pho1	light orange *wt* wing	0%	18,2%	22(5,3%) 37(5,3%)
+ / + ;; pho1 / ci D	white *ciD* wing	25%	18,2%	69(16,5%) 120(17,2%)
+ / + ;; pho1 / pho1	lethal	0%	0%	0(0%) 0(0%)

Figure 17: *pho*^{1} null mutant rescue statistics with *pα1Tubulin-Gfppho*

The genotypes obtained from the ultimate genetic cross of the *pho*^{1} null mutant rescue experiment and their respective expected distribution, if *pα1Tubulin-Gfppho* is not able to substitute or fully replaces endogenous *pho,* are listed. The indicated phenotypes were used for genotype identification and are presented in red for the eye colour and in black for the wing vein L4. The number of flies counted for each genotype and their percentage of the total number of flies counted is presented in green (n_1=419) and blue (n_2=696) for two independent experiments.

Generally, males and females showed a different expression level of the *miniwhite* transformation marker cloned downstream of the *Gfp* fusion protein coding sequence. In order to evaluate the efficiency of the rescue by homo- and heterozygous transgenes separately, the hatched flies of the *pα1Tubulin-Gfppho* rescue cross were separated into males and females. They were then further divided into groups of homozygous transgenics, which have the strongest *miniwhite* expression due to the presence of two copies of the *miniwhite* gene, heterozygous transgenics, which have only one copy of *miniwhite* and therefore a lighter eye colour than homozygous transgenics, or flies without the transgene, which have white eyes. There was no significant difference between counted numbers of genotypes in males and females and therefore the data in Fig 17 represent the sum of the respective genotypes in males and females.

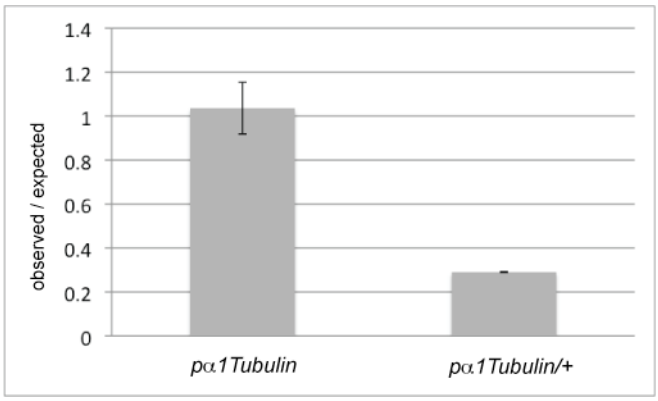

Figure 18: Ratio of the observed percentage of rescued flies to the theoretically expected percentage of flies in a full rescue situation with *pα1Tubulin-Gfppho*

Numbers of *pα1Tubulin-Gfppho* homozygous (*pα1Tubulin*) and heterozygous (*pα1Tubulin/+*) adult flies that are pho^{1}/pho^{1} (wildtype wing vein L4) were calculated as percent of the total number of flies counted in the respective experiment and presented in ratio to the theoretical percentage of these genotypes if GFP::PHO can fully replace endogenous PHO under *α1Tubulin* promoter expression.

The distribution of the respective genotypes with the assumption that all have the same viability with exception of the homozygous lethal mutant without a copy of the transgene was defined as a full rescue. The observed occurrence of each genotype was calculated as the ratio over the expected occurrence and plotted in Fig 18.

Clearly, the N-terminally GFP-tagged PHO is a functional protein. Under the expression of the *α1Tubulin* promoter two copies of *Gfppho* are sufficient to reach a full rescue. A heterozygous level revealed approximately a quarter of a fully rescued distribution.

These results show that the rescue of the functional fusion protein GFP::PHO depends on its expression level. The *ppho* sequence cloned to drive transgene expression is not sufficient for replacing the endogenous protein, but constitutative *pα1Tubulin* control can restore full viability to fertile adults.

4.3.2 Constitutative *Actin5c* and α1*Tubulin* promoters yield detectable transgenic *pho* transcript- and protein levels

In order to analyze the previously generated stable transgenic fly lines expressing the *Gfppho* transgene homozygously and controlled by the *Actin5c*, the presumed endogenous, or the *α1Tubulin* promoters, I used quantitative PCR and Western Blot techniques.

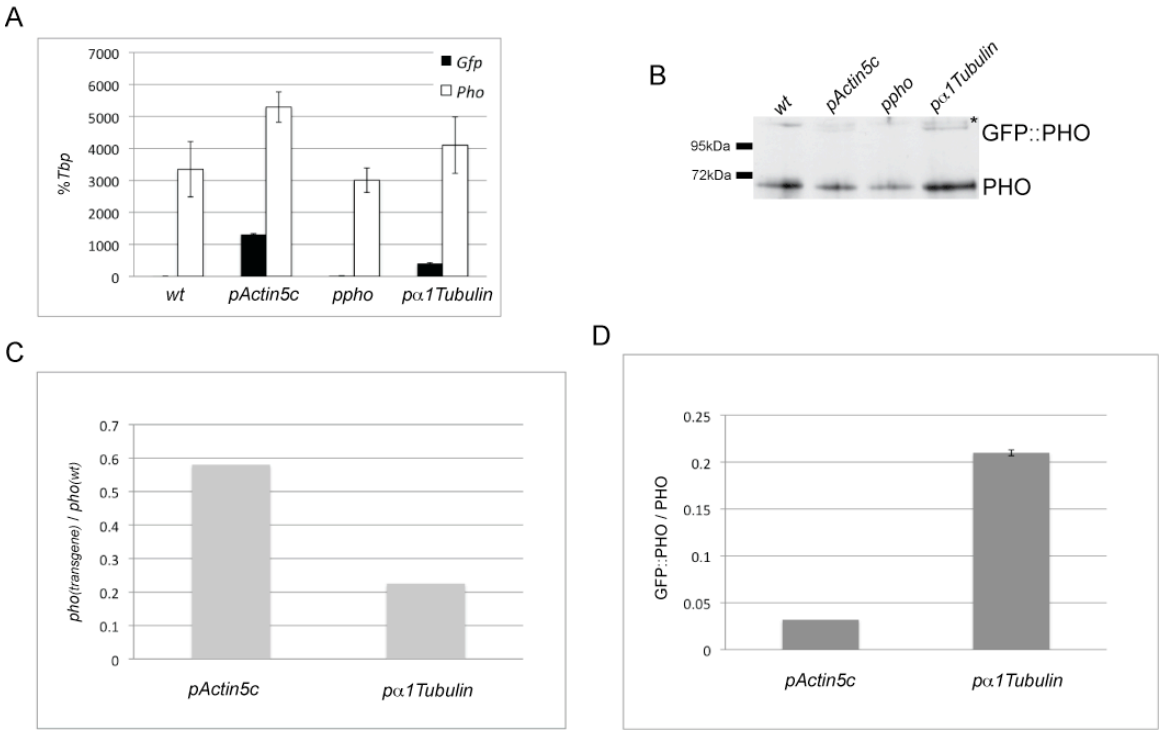

Figure 19: Realtime PCR, western blot and respective quantification of transcript and protein levels of *Gfppho* in embryos.

Gfp and *pho* transcripts of 0-12h collections of wildtype and homozygous transgenic embryos were analyzed by real-time PCR (A and C). RNA was isolated from each genotype and PCR reactions were performed on dilution series of two technical cDNA replicates. Ct values were calculated as %*Tbp*, averaged for each cDNA replicate and plotted in (A), however the mean values of *Gfp* on transgenic cDNA were obtained only from those Ct values that were in linear correlation to eachother. The transgenic *pho* transcript level was calculated by subtracting the average of both means (shown in A) from wildtype technical replicates from the average of both means from transgenic technical replicates and plotted in (C) in ratio to the wildtype level of *pho* transcript.
GFP::DSP1 protein levels in total protein extracts of wildtype and transgenic 0-12h embryo collections were quantified from Western Blots (B) and shown as fold of endogenous PHO levels (D). The asterisk in (B) marks a background band running above GFP::PHO.

Both constitutative promoters tested drive verifiable transgenic transcript levels

Due to the fact that GFP::PHO protein levels could not be reliably quantified in all transgenic lines by Western Blot (described below) I investigated transcript levels by quantitative real time PCR to examine expression by independent means.

Primer pairs in the coding sequence of *Egfp* (referred to for simplicity as *Gfp)*, within the third exon of the *pho* gene and for *Tbp* as control showed similar efficiencies and were used for real time PCR. RNA was isolated from 0-12h embryos of wildtype and transgenic

Gfppho fly lines and two technical cDNA replicates were prepared of which dilution series were made for independent PCR reactions.

Significant levels of the transcript of *Gfp* and an increase in *pho* transcript over wildtype level were only detected in the transgenic fly lines that have constitutative promoters (*pActin5c* and *pα1Tubulin*) to drive the *Gfppho* expression. *Ppho* did not show detectable transcription of the transgene in embryos. (Fig. 19 A)

To quantify *Gfppho* transcript in transgenic fly lines, the mean value of wildtype *pho* transcript was subtracted from transcript levels of transgenic *pho* (shown in Fig. 19 A) and plotted as fold of the wildtype level.

In the *pActin5c* line this analysis detected approximately half the amount of endogenous *pho* for *Gfppho* and a quarter of endogenous *pho* for *Gfppho* under *pα1Tubulin* controlled expression. As indicated this calculation did not show a positive result for the *ppho* line. (Fig. 19 C)

Taken together these results demonstrate that the *Actin5c* and the *α1Tubulin* promoters drive the transgene to a level detectable in real-time PCR in embryos.

Constitutatively expressed transgenic fusion protein is detected in reverse relation to the respective transcript levels in the embryo

The quantitatification of Western Blots reveals the impact of transgenic expression strategies on GFP fusion protein level and regulatory feedback effects on the endogenous protein level as already described above (See 4.2.2). In order to test this, GFP::PHO expression in 0-12h total embryo extracts of wildtype and all three transgenic lines homozygous for *Gfppho* controlled by *Actin5c*, the presumed endogenous *pho* or *α1Tubulin* promoters was analyzed.

Western blot analysis with an anti-PHO antibody showed as observed for the other proteins studied that both the fusion protein (predicted MW is 102kDa) and the approximately 58kDa endogenous protein ran at higher molecular weight relative to the marker proteins in SDS-PAGE. Due to a suboptimal combination of the quality of anti-PHO antibody, low amount of protein of interest in the sample and a background band running in close distance above GFP::PHO, the fusion protein in embryos was hardly detectable and quantifiable as a separate band on Western Blot. (Fig. 19 B) The correct identity was confirmed by anti-GFP detection. The protocols and methods applied did not give a usable result with salivary gland preparations (data not shown).

No GFP::PHO band was detectable in samples of flies that carry the presumed endogenous *pho* promoter to drive transgene expression, hence I could not determine the

ratio of GFP::PHO to endogenous PHO in the embryo (Fig. 19 D), which is consistent with the real-time PCR result. The other lines having constitutative promoters for transgene expression show low levels of fusion protein compared to endogenous protein (0.03 fold for *pActin5c* and 0.2 fold for *pα1Tubulin*).

In summary, the presumed endogenous promoter that was cloned does not drive expression in the embryo to an amount of GFP::PHO fusion protein detectable in Western Blot analysis. In the *pActin5c* and *pα1Tubulin* driven transgenic fly lines GFP::PHO is detected and the levels are significantly below those of endogenous PHO.

These results show that the cloned presumed *pho* promoter is not effective in transgene fusion protein expression. Real-time PCR analysis demonstrates that the *Actin5c* promoter features transgene expression at approximately half the level of the endogenous transcript, which is 3 times more transcript than in the *α1Tubulin* promoter embryos. However, the fusion protein level with *pActin5c* is low compared to the level under *α1Tubulin* promoter control, which makes a fifth of endogenous PHO protein.

4.3.3 GFP::PHO under constitutative promoter control is detected as descrete bands on polytene chromosomes in larval salivary gland nuclei *in vivo*

In order to investigate if binding of functional GFP::PHO to specific sites on polytene chromosomes in salivary gland nuclei is visible under live imaging conditions, I dissected transgenic third instar larvae and imaged GFP with confocal microscopy in the living tissue as described in Materials and Methods.

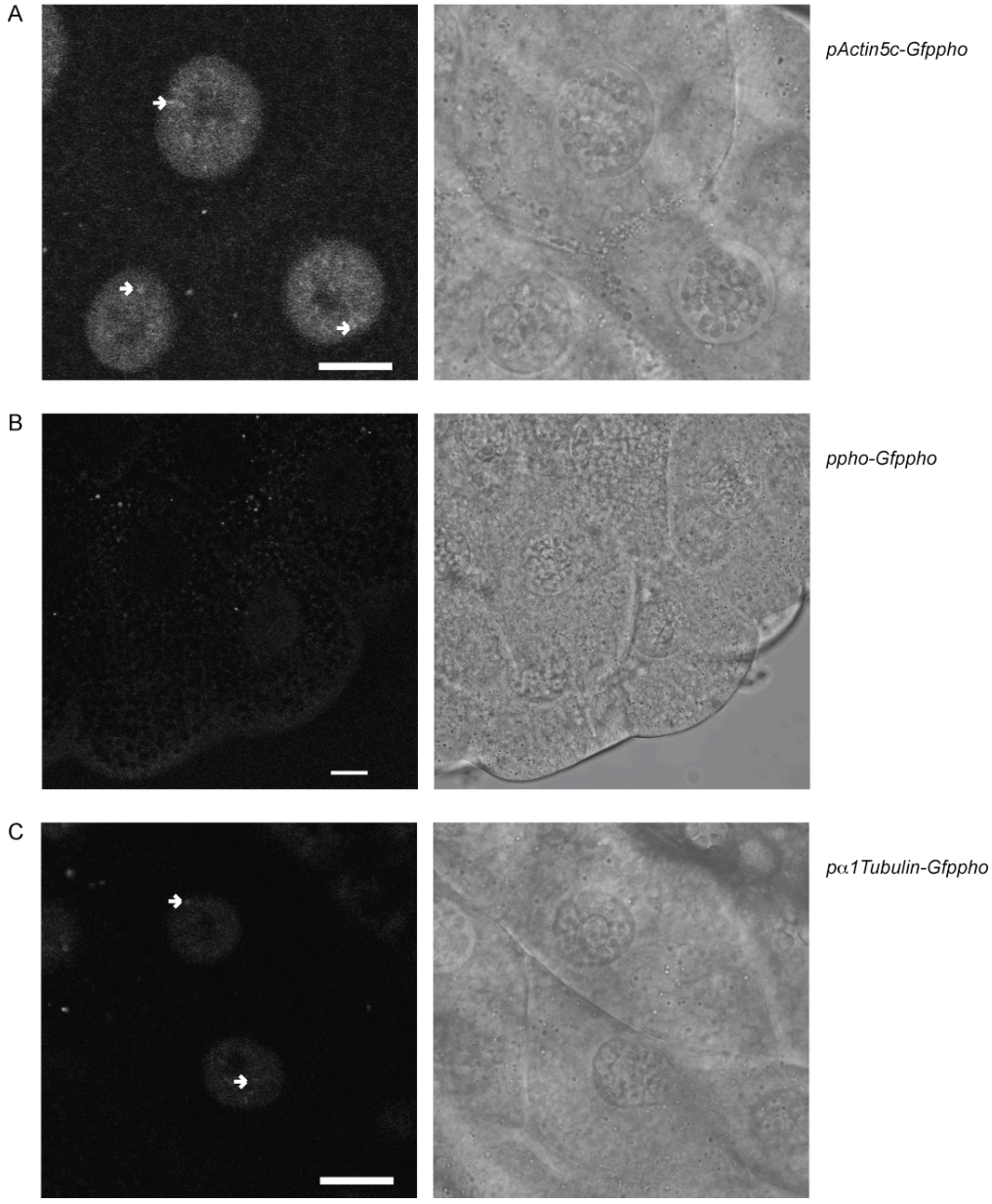

Figure 20 Live Imaging of GFP::PHO in larval salivary gland nuclei

GFP::PHO distribution in living salivary gland nuclei is shown of *pActin5c-Gfppho* (A), *ppho-Gfppho* (B), and *pα1Tubulin-Gfppho* (C) transgenic larvae. The images were taken with a 40x/1.3NA oil immersion objective. 488nm excitation with 2.1% laser output (A and C) or 8.1% (B) and a pinhole of 260µm were used. The corresponding DIC images are shown on the right. Scaling was x: 0.22µm, y: 0.22µm, z: 0.6µm. The white bars represent 100x10 pixels. Arrowheads point to selected bands.

In salivary gland nuclei from larvae expressing the *Gfp* fusion transgene under constitutative promoter control (*pActin5c* and *pα1Tubulin*) the amount of fluorescent

protein was sufficient to detect single bands. Compared to GFP::DSP1 (4.2.3) there was less GFP signal in the nuclei, the contrast between chromosomal and free nuclear GFP was not as strong and there was a smaller number of binding sites clearly descernable. Under the same imaging conditions, there was more GFP::PHO detected in the *pActin5c* line than in the *pα1Tubulin* line, in contrast to the situation in the embryo as can be seen on western blot analysis (4.3.2).

The amount of GFP::PHO expressed under the presumed endogenous promoter sequence was too low for live imaging to measure relevant fluorescence intensities. Therefore, this transgenic line could not be used for live imaging experiments on specific binding sites in salivay gland nuclei.

In salivary glands of all three transgenic lines structured cytoplasmic fluorescence was detected (Fig. 20). In order to investigate the observed signal in the cytoplasm, I performed immunostaining against PHO in third instar larval tissue of all three stable transgenic fly lines (data not shown). There was no significant difference between the cytoplasmic signals from wildtype or transgenic *pα1Tubulin-Gfppho* tissue observed in salivary glands, imaginal discs and brain. Conclusively, the observed fluorescence was not an overexpression artefact of GFP::PHO in the transgenic lines.

In summary, both transgenic lines with constitutative expression strategies are useful for live imaging experiments on GFP::PHO binding to polytene chromosomes in larval salivary glands.

4.3.4 Live Imaging of GFP::PHO in early embryo development

PHO, like DSP1, is a DNA binding protein in the PcG mediated epigenetic memory mechanism and recognizes a specific DNA sequence motif. *pho* has been suggested to a play a role in early embryo development [35, 131, 140, 141]. It has been shown [13] that PHO binding to PREs is essential for maintaining repressive *Hox* gene chromatin and that recruitment of PcG proteins is hierarchical, with PHO recruiting E(Z)-containing complexes [39]. As GFP::DSP1 has been shown in 4.2.3 to reside on mitotic chromatin during synchronous nuclear divisions in the early embryo (Fig. 15), I used confocal microscopy to ascertain the distribution of GFP::PHO.

Embryos of the transgenic Drosophila line expressing *Gfppho* under control of the *α1Tubulin* promoter were used for this study, because this expression strategy showed highest GFP::PHO protein levels and was able to rescue *pho^1* lethality. During the time investigated the large part of fluorescence was mainly autofluorescence in the yolk

cytoplasm and granules (Fig. 21 A). After deconvolution (Fig. 21 B) of data from cycle 13-14 embryos there was higher intensity in nuclear structure observed. Nevertheless, the amount of GFP fusion protein in the embryo was not sufficiently high to answer the question if it persists on chromatin during the cell cycle with this method.

With the expression strategies tested and the methodology used, it could not be ascertained whether PHO remains on chromatin during the upheavals of mitosis. However, the presence of readily detectable GFP::PHO in salivary gland nuclei suggests that further investigation of other larval tissues that are mitotic may yield an answer to this question.

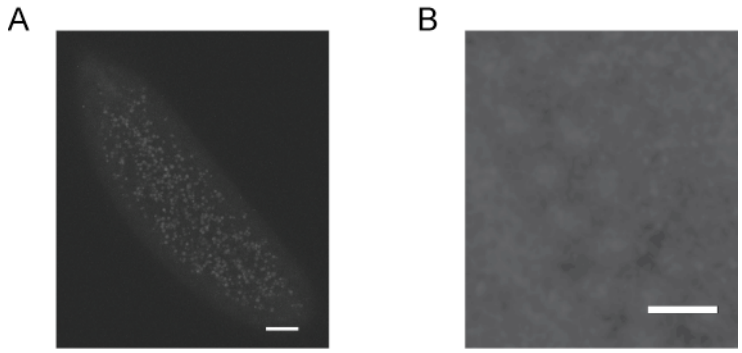

Figure 21: GFP::PHO under in *pα1Tubulin-Gfppho* preblastoderm embryos

GFP::PHO under *pα1Tubulin* control in transgenic embryos was faintly detectable. Movies were generated taking a new z stack every 133sec (A) or 121sec and deconvolved (B). Pictures are presented as maximum intensity projections of a selected number of slices. A 25x/0.8 NA oil immersion corrected objective, 6.1% laser output at 488nm and a pinhole of 178µm were used for imaging. Scaling was in (A) x: 0.35µm, y: 0.35µm, z: 0.5µm and zoomed in (B) x: 0.07µm, y: 0.07µm, z: 0.5µm. The white bars represent 100x10 pixels. The gamma in (B) was changed to 1.5.

4.4 Quantitative analysis and live Imaging of GFP::E(Z)

4.4.1 GFP::E(Z) is a functional fusion protein

In order to assess the functionality of the GFP::E(Z) fusion protein and to evaluate the optimal expression level strategy, I performed rescue experiments. As described in detail in Materials and Methods, *GfpE(z)* transgenes were brought into a homozygously recessive lethal mutant background and the ability of the transgene to support *E(z)* function was determined. Three different strategies were used, with third chromosomal states of hemizygous $E(z)^{63}/Df(3L)lxd^{15}$, homozygous $E(z)^{63}$ or trans-heterozygous

E(z)^{63}*/E(z)*^{731}. *E(z)*^{63} and *E(z)*^{731} have both been described to be null mutations [28, 132]. Out of the three different stable transgenic lines obtained (*pActin5c-GfpE(z)/CyO*, *pE(z)-GfpE(z)*, and *p*α*1Tubulin-GfpE(z)*) the lines with the *pE(z)* and *p*α*1Tubulin* controlled transgenic expression could be established with the third chromosome balancer TM6c,*Tb,Sb/Dr,* which were used for rescue crosses. However *E(z)*^{63}*/Df(3L)lxd*^{15} was viable, probably due to a recombination event in the *Df(3L)lxd*^{15}/TM6c,*Tb,Sb* stock, and homozygous *E(z)*^{63} could not be rescued because of a likely additional lethal recessive mutation on the *E(z)*^{63} chromosome (Rick Jones, personal communication, and mentioned in [28]). Therefore, the rescue was performed with the trans-heterozygous combination of *E(z)*^{63}*/E(z)*^{731}. Fig. 22 shows the final cross of this experiment.

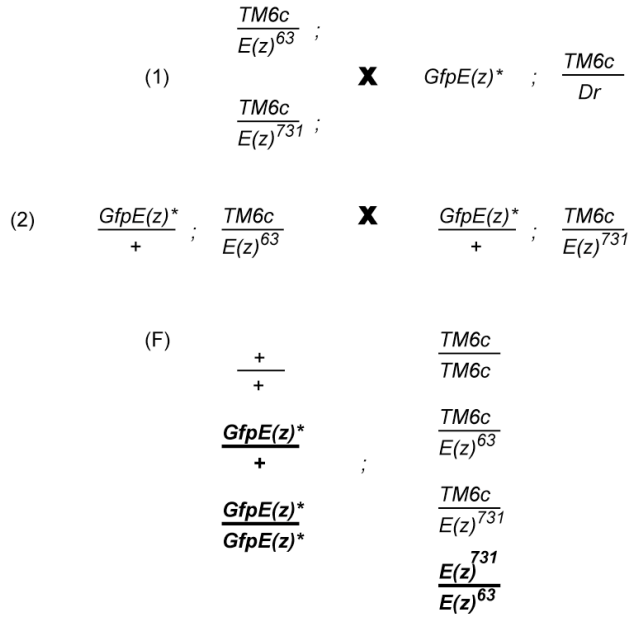

Figure 22: Rescue of *E(z)* null alleles by transgenic *GfpE(z)* expression

TM6c/E(z)^{63} and *TM6c/E(z)*^{731} mutant flies were crossed to homozygous *GfpE(z)* flies that carried TM6c,*Tb,Sb*/*Dr* on the third chromosomes (1). Progeny of cross (1) with indicated genotypes (2) were selected and recrossed. Possible second and third chromosome genotypes in the next generation are listed in (F). The combinations of second and third chromosome genotypes that are presented in bold letters constitute the rescue. **GfpE(z)* implicates crosses for both the *pE(z)-GfpE(z)* and *p*α*1Tubulin-GfpE(z)* rescue experiments.

pE(z)-GfpE(z) and *p*α*1Tubulin-GfpE(z)* transgenic flies were crossed to a mutant background and in the retried experiment heterozygous *E(z)*^{63}*/E(z)*^{731}mutants in the progeny were identified by *Tb*+ pupal shape and *Sb*+ bristle phenotypes. Counted flies were analyzed with respect to the copynumber of the transgenic *GfpE(z)* on the second

chromosome. Heterozygous $E(z)^{63}/E(z)^{731}$ was lethal in the final progeny, and this lethality was rescued by the expression of *GfpE(z)* transgenes.

genotype	phenotype	**no rescue** (expected % of total number of flies)	**full rescue** (expected % of total number of flies)	**flies counted (observed % of total flies counted)** *pE(z)*	*pα1Tubulin*
GfpE(Z) / *GfpE(Z)* ; *E(z)* mutant* / TM6c, *Tb, Sb*	dark orange *Tb* pupae *Sb* bristle	25%	18,2%	70(12,9%) 89(15,5%)	53(7,8%) 97(14,6%)
GfpE(Z) / *GfpE(Z)* ; *E(z)731* / *E(z)63e11*	dark orange *wt* pupae *wt* bristle	0%	9,1%	70(12,9%) 48(8,4%)	85(12,4%) 75(11,3%)
GfpE(Z) / + ; *E(z)* mutant* / TM6c, *Tb, Sb*	light orange *Tb* pupae *Sb* bristle	50%	36,4%	273(50,4%) 295(51,5%)	224(32,8%) 236(35,6%)
GfpE(Z) / + ; *E(z)731* / *E(z)63e11*	light orange *wt* pupae *wt* bristle	0%	18,2%	0(0%) 0(0%)	207(30,3%) 158(23,8%)
+ / + ; *E(z)* mutant* / TM6c, *Tb, Sb*	white *Tb* pupae *Sb* bristle	25%	18,2%	129(23,8%) 141(24,6%)	114(16,7%) 97(14,6%)
+ / + ; *E(z)731* / *E(z)63e11*	lethal	0%	0%	0(0%) 0(0%)	0(0%) 0(0%)

Figure 23: *E(z)* null mutant rescue statistics with *pE(z)-GfpE(z)* and *pα1Tubulin-GfpE(z)*

The genotypes obtained from the ultimate genetic cross that brings together two *E(z)* null mutant alleles $E(z)^{731}$ and $E(z)^{63}$ and their respective expected distribution, if *pE(z)-GfpE(z)* and *pα1Tubulin-GfpE(z)* are not able to substitute or can fully replace endogenous *E(z),* are listed. The indicated phenotypes were used for genotype identification and are presented in red for the eye colour and in black for the shapes of pupae and adult bristles. The number of flies counted for each genotype and their percentage of the total number of flies counted is presented in green ($n_{pE(z)\text{-}1}$= 542, $n_{p\alpha 1tubulin\text{-}1}$= 683) and blue ($n_{pE(z)\text{-}2}$= 573, $n_{p\alpha 1tubulin\text{-}2}$= 663) for two independent experiments in each transgenic fly line.

Interestingly, two copies of *pE(z)-GfpE(z)* were able to fully rescue the trans-heterozygous $E(z)^{63}/E(z)^{731}$ mutant background, whereas no single rescued adult fly was observed with only one copy of transgenic *GfpE(z)* under the control of the presumed *E(z)* promoter.

The GFP::E(Z) expression level under control of the *α1Tubulin* promoter resulted in a full rescue in both homo- and heterozygous states of the transgenic *GfpE(z)*. Counted numbers of respective genotypes with a trans-heterozygous $E(z)^{63}/E(z)^{731}$ mutant background were higher than the expected occurrence and comparison of expected versus counted numbers of all possible genotypes revealed that the presence of the

balancer chromosome TM6c,*Tb,Sb* in context with *E(z)* mutations had an adverse effect on proper development until adulthood.

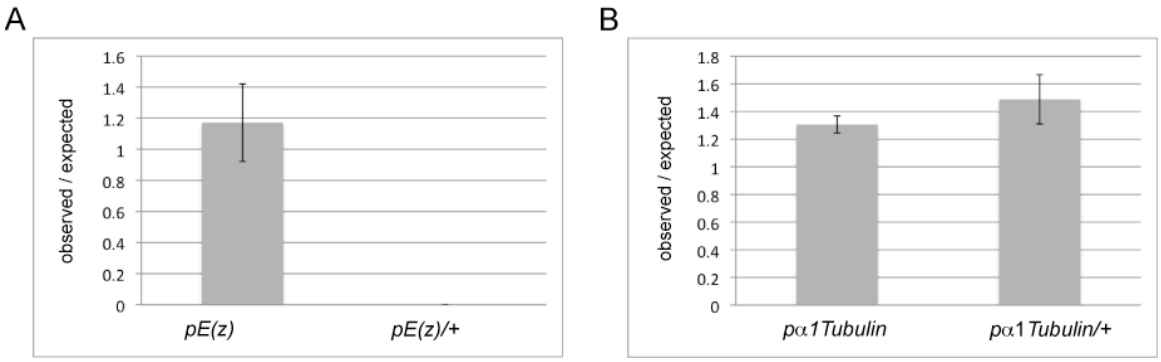

Figure 24: Ratio of the observed percentage of rescued flies to the theoretically expected percentage of flies in a full rescue situation with *GfpE(z)*

Numbers of *pE(z)-GfpE(z)* (A) and *pα1Tubulin-GfpE(z)* (B) homozygous and trans-heterozygous transgenic adult flies that are $E(z)^{731}/E(z)^{63}$ (wildtype bristles) were calculated as percent of the total number of flies counted in the respective experiment and presented in ratio to the theoretical percentage of these genotypes if GFP::E(Z) can fully replace endogenous E(Z) under *α1Tubulin* and the presumed *E(z)* promoter expression.

Taken together these results demonstrate that GFP::E(Z) is a fusion protein that can fulfill all of the functions of endogenous E(Z). Remarkably, the cloned presumed *E(z)* promoter drives transgene expression to a vital level only in homozygotes.

4.4.2 Different promoters controlling *GfpE(z)* show fundamental differences in transcript and protein levels

Homozygous stable fly lines were obtained with transgenes driven by the presumed endogenous promoter and the *α1Tubulin* promoter, however the *Actin5c* promoter controlled transgene construct did not give rise to a stable homozygous fly line. The homozygous genotype was not lethal but not healthy and hardly gave rise to progeny. Therefore, the *pActin5c/CyO* transgenic fly line reflects a mixture of heterozygous and to a minor but undefined amount homozygous embryo collections used for quantification of the transgenic expression level.

The expression strategies for transgenic GFP-tagged *Dsp1* and *pho* quantified in this work (chapters 4.2.2 and 4.3.2) showed overexpression of total protein levels. The same

methods as for these two proteins were applied to the PRC2 member *E(z)* to evaluate the effect of the different expression strategies for *GfpE(z)* transcripts and protein levels.

The *GfpE(z)* transgene is weakly transcribed by the cloned *E(z)* promoter and strongly transcribed by constitutative promoters in the embryo

To investigate the transcription levels under the different promoters cloned for *GfpE(z)*, which makes a functional fluorescently tagged E(Z) fusion protein (See 4.4.1), I used primer pairs in the coding sequence of Egfp (for simplicity referred to as *Gfp)*, within the second exon of the *E(z)* gene and for *Tbp* as control that showed similar efficiencies in real time PCR.

RNA was isolated from 0-12h embryos of wildtype and transgenic *GfpE(z)* fly lines and dilution series of two technical cDNA replicates were made for independent PCR reactions (Fig. 25 A).

All three *GfpE(z)* transgenic lines reveal an increase in *E(z)* transcript. In the transgenic lines with constitutative promoters the detected increase is larger than under the presumed endogenous promoter of *E(z)* (Fig. 25 A). The corresponding *Gfp* transcript levels did not correlate to the rise of *E(z)* product level, although the control experiments for *Gfp* on genomic DNA and experiments on *Gfppho* lines did not suggest inaccurate primer function. (See 3. Material and Methods.)

The increase of *E(z)* transcript present in transgenic fly lines was plotted as ratio of transcript levels of transgenic *E(z)* lines to the mean value of wildtype *E(z)* transcript (Fig. 25 C). The level of *E(z)* transcript in the transgenic *pActin5c* and *pα1Tubulin* lines were approximately respectively 11 fold and 6 fold higher than the endogenous transcript level in wildtype samples. However, if the increase in *E(z)* product originated from the transgenic *GfpE(z)* transcription, *Gfp* transcript levels would be accordingly detectable, namely to levels of 10000-13000% *Tbp*. In fact, *Gfp* levels were significantly lower (Fig 25 A). Therefore, high *E(z)* transcript level in transgenic samples were not due to *GfpE(z)* transcription, but increased endogenous transcription.

Taken together, these results demonstrate that in the fly lines with *GfpE(z)* transgenes controlled by *Actin5c* and *α1Tubulin* promoters, *E(z)* is expressed at remarkably higher levels than in the line with the transgene under control of the presumed endogenous promoter, which shows no overtranscription measured by real-time PCR in embryos. This suggests an upregulation of the endogenous gene caused by constitutative promoter control of the transgene.

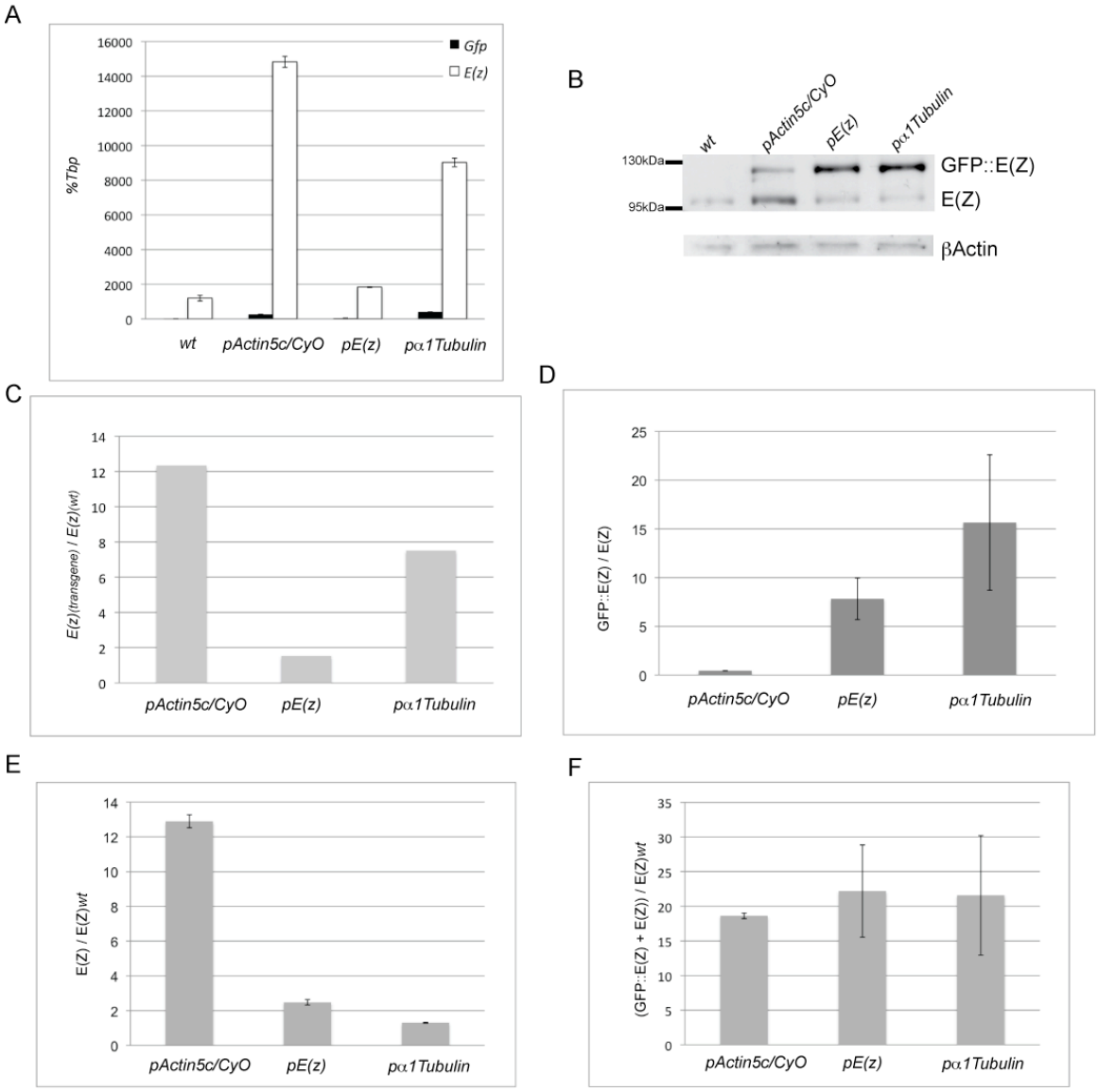

Figure 25: Realtime PCR, western blot and respective quantification of transgenic and endogenous transcripts and proteins in embryos of stable *GfpE(z)* fly lines

Transgene expression in respective fly lines was induced using the *actin5c* promoter (*pactin5c*) presumed endogenous *E(z)* promoter (*pE(z)*), or the *α1Tubulin* promoter (*pα1Tubulin*) and transcript and protein levels in embryos were quantified. Real-time PCR was employed to analyze *Gfp* and *E(z)* transcripts of 0-12h collections of wildtype and transgenic embryos of obtained stable fly lines (A and C). RNA was isolated from each genotype and PCR reactions were performed on dilution series of two technical cDNA replicates. Ct values were calculated as %*Tbp*, averaged for each cDNA preparation and plotted in (A) However, the mean values of *Gfp* on transgenic cDNA were obtained only from those Ct values that were in linear correlation to eachother. The transgenic *E(z)* transcript level was calculated by deviding the average of both means from transgenic technical cDNA replicates (heights of white bars from transgenic samples in A) and shown in (C) by the wildtype level of *E(z)* transcript (average of both means from wildtype technical cDNA replicates).

The ratio of the GFP fusion protein compared to the level of endogenous E(Z) is shown in (B) and (D). Error bars in (D) reflect the standard deviation obtained by averaging the signals from two developed chemiluminescence films of one experiment that had different exposure times. Endogenous and total E(Z) levels were normalized to the amount of protein loaded (B) and compared to respective wildtype levels. Results from both developed films were averaged. (E and F).

Differential fusion protein expression and effects on the endogenous protein level by *pActin5c*, *pE(z)* and *pα1Tubulin*

To define fusion protein levels from different expression strategies and to investigate if there is an effect of transgenic protein expression on the endogenous protein level as described previously (4.2.2 and 4.3.2), I quantified Western Blots of total protein extracts from embryos of wildtype and stable transgenic fly lines using an anti-E(Z) antibody.

Quantification of western blots showed the 87kDa endogenous E(Z) protein and the fusion protein whose predicted MW is 112kDa ran at higher molecular weight relative to the marker proteins in SDS-PAGE (Fig. 25 B) as also described for DSP1 and PHO and each of their fusion proteins.

GFP::E(Z) was strongly expressed under control of *pE(z)* and *pα1Tubulin*, with levels of 7 and 15 fold that of endogenous E(Z), whereas the *pActin5c* revealed much less fusion protein expression (0.45 times endogenous E(Z)). (Fig. 25 B and D)

In order to quantify the effect of transgenic GFP::E(Z) expression on the endogenous protein level, I normalized the signal intensities of endogenous E(Z) to the wildtype E(Z) level by the amount of protein loaded according to the loading control detected by an anti-βACTIN antibody.

Interestingly, the amount of endogenous E(Z) levels in the transgenic fly lines increased approximately 13 fold for the *pActin5c/+* line but increased only slightly (1-2 fold) when the transgene was controlled by the presumed *E(z)* or the *α1Tubulin* promoters which showed much higher GFP::E(Z) protein levels. (Compare Fig. 25 B, D and E) This upregulation of endogenous *E(z)* in the *pActin5c* transgenic embryos may explain why the *Gfp* signal in the real-time PCR experiment was so low compared to the increase in *E(z)* transcript levels measured in *pE(z)-GfpE(z)* and *pα1Tubulin-GfpE(z)* embryos(C).

The differences in transcript and protein levels of the transgene in embryos of the *pActin5c-GfpE(z)* and *pα1Tubulin-GfpE(z)* lines suggest also translational control.

To determine the increase of total E(Z) protein in transgenic fly lines, the sum of GFP::E(Z) normalized to wildtype E(Z) levels and normalized values of endogenous E(Z)

levels from embryos of each transgenic fly line were plotted in ratio to the levels of wildtype E(Z). (Fig. 25 F)
This showed that in all three transgenic fly lines there was a similar and high overexpression of total E(Z) protein compared to the wildtype condition, although the difference in the levels of GFP::E(Z) were large. (Compare Fig. 25 D and F).
Thus, the three different tested expression strategies for GFP::DSP1 lead to significant differences on protein level and affect the regulation of endogenous *E(z)*.

The results obtained from quantifying transcript and protein levels in fly lines carrying *GfpE(z)* show different effects of different promoters on endogenous and transgenic products. In contrast to *pE(z)* and *pα1Tubulin* the constitutative promoter of *actin5c* is not stably tolerated homozygously. This promoter shows the highest *E(z)* transcript level, has only a small amount GFP::E(Z) in embryos but a large upregulation of endogenous E(Z). *pE(z)* does not drive a significant increase in *E(z)* transcript but GFP::E(Z) is overexpressed compared to endogenous E(Z) levels whereas for *pα1Tubulin,* both transcript and protein levels are increased to a considerable extent.

4.4.3 GFP::E(Z) in transgenic larval salivary glands binds to descrete loci on polytene chromosomes *in vivo*

To investigate fluorescence intensities in transgenic GFP::E(Z) *Drosophila* lines for live imaging experiments on single binding loci on polytene chromosomes of third instar larval salivary glands, I dissected and mounted living tissue from all stable transgenic *GfpE(z)* lines as described in previous chapters. GFP::E(Z) in salivary gland nuclei was imaged with confocal microscopy.
All transgenic lines had sufficient GFP::E(Z) expressed to distinguish single bands from free nuclear fluorescence in imaged nuclei. In salivary gland nuclei from *pActin5c-GfpE(z)* larvae, signals of descrete binding loci were weak and chromosomal regions were visible that were nearly free of GFP::E(Z) (Fig. 26 A). This might be due to a lower level of fluorescent fusion protein, and similar effects may have been masked by stronger nucleoplasmic fluorescence in nuclei from *pE(z)-GfpE(z)* and *pα1Tubulin-GfpE(z)* transgenic larvae.
The signals were stronger in the *pE(z)* and *pα1Tubulin* controlled lines than in the mostly heterozygous *pActin5c* line (Fig 26 B and C). Specific bands on the polytene chromosomes were bright and easily detectable with conditions for live imaging. Under

protein expression with *pα1Tubulin* there was a higher intensity of nucleoplasmic GFP signal than with *pE(z)*.

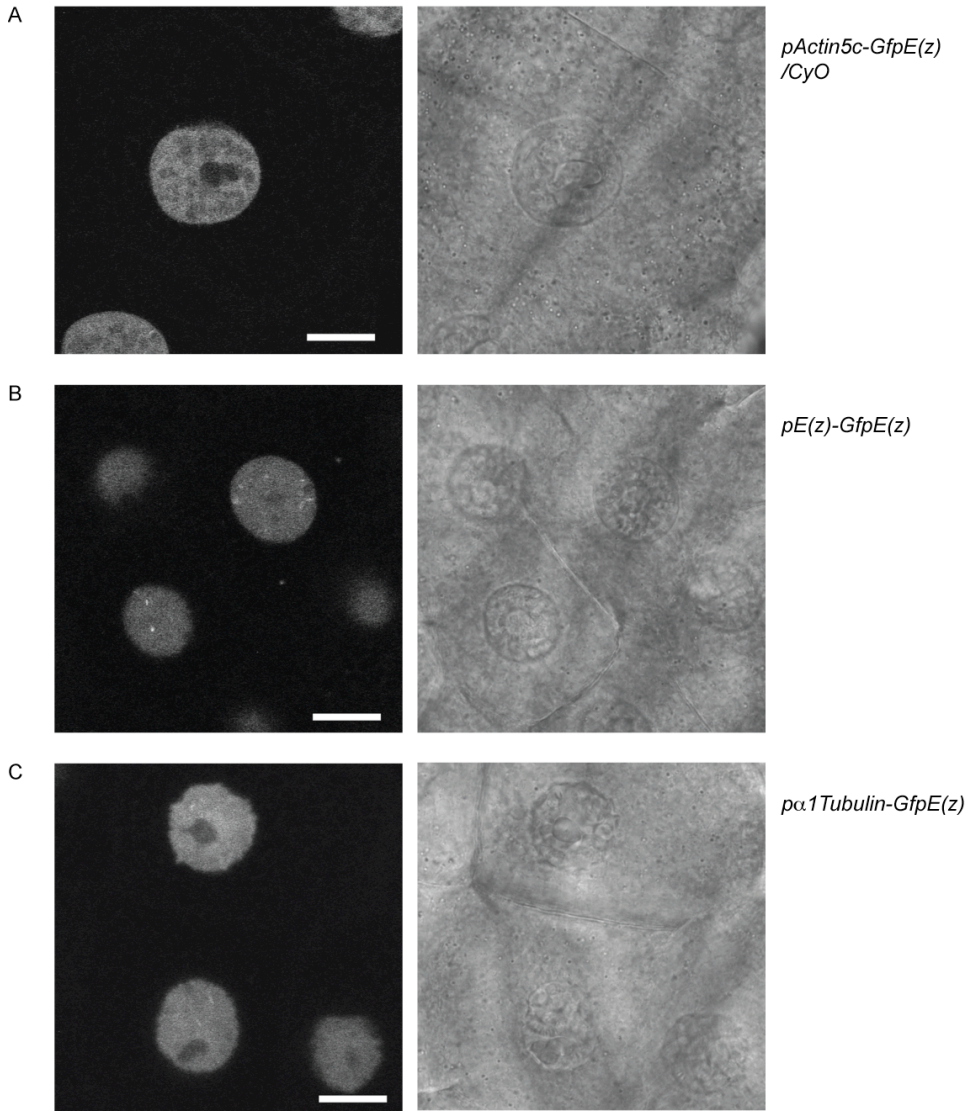

Figure 26: Live Imaging of GFP::E(Z) in larval salivary gland nuclei

(Fig. 26)
GFP::E(Z) distribution in living salivary gland nuclei is shown of *pActin5c-GfpE(z)* (A), *pE(z)-GfpE(z)* (B), and *pα1Tubulin-GfpE(z)* (C) transgenic larvae. The images were taken with a 40x/1.3NA oil immersion objective. 488nm excitation with 2.1% laser output and a pinhole of 66µm (A) or 260µm (B and C) were used. The corresponding DIC images are shown on the right. Scaling was x: 0.22µm, y: 0.22µm, z: 0.6µm. The white bars represent 100x10 pixels.

Although more GFP is detectable in *GfpE(z)* transgenic salivary gland nuclei than with GFP::PHO, there were less clear discrete binding loci of GFP::E(Z) than observed with the other GFP fusion proteins.

In conclusion, this tagged PcG protein can be studied in live imaging experiments on specific binding loci on polytene chromosomes of larval salivary glands.

4.4.4 Residual GFP::E(Z) remains on chromatin during metasynchronous nuclear divisions in early embryo development

Despite drastic chromosomal reorganization during mitosis, selective dissociation of non-histone proteins [142], my results for DSP1 have shown that this DNA-binding protein stays attached to mitotic chromosomes. The bulk of the PRC1 PcG protein PC c-terminally fused to GFP has been shown to be depleted from the chromosomes during mitosis in preblastoderm embryos by time-lapse confocal microscopy, however a detectable fraction remains associated [112]. In contrast, immunostaining of later embryos have revealed no detectable localization of the PRC1 PcG proteins PC, PH and PSC on metaphase chromosomes, with different time points of dissociation and reassociation for each protein [107].

To investigate the distribution of the PRC2 protein E(Z) during mitosis, I employed laser-scanning confocal microscopy as described above for GFP::DSP1 and GFP::PHO. Transgenic *pE(z)-GfpE(z)* and *pα1Tubulin-GfpE(z)* embryos at early developmental stages were used for live imaging. GFP::E(Z) in stage 4 embryos featuring metasynchronous nuclear divisions (10^{th}-13^{th}) are documented in Fig. 27.

Like DSP1 (4.2.4) and PC [112] the first distinctive GFP signals were detected at cycle 10 in peripheral interphase nuclei. During stage 4 yolk granules and autofluorescence diminished as also observed for GFP::DSP1 expressing embryos (Fig 15). GFP derived signal in the nuclei decreased during mitosis, however dissociation occurred only partially, as could be ascertained at higher magnification and after deconvolution (Fig. 27 B). A detectable fraction of GFP::E(Z) on chromatin was observed througout cycle 10-14.

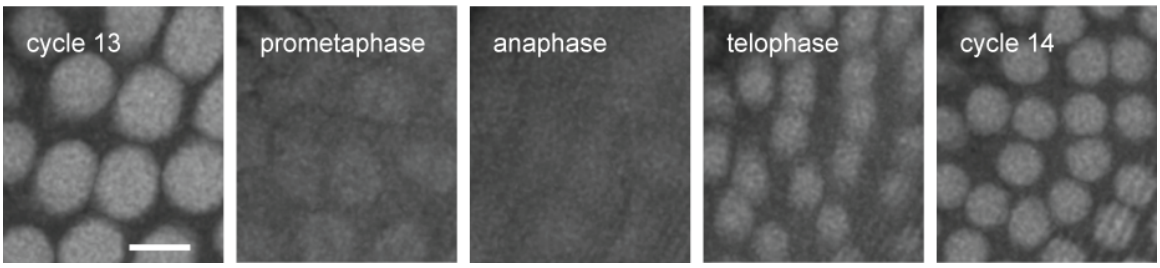

Figure 27: Live Imaging of GFP::E(Z) during metasynchronous nuclear divisions (cycle 10-14) in transgenic *pα1Tubulin-GfpE(z)* preblastoderm embryos

Nuclear GFP::E(Z) in *pα1Tubulin-GfpE(z)* transgenic embryos was detected from stage 4 (cycle 10) on and imaged during subsequent nuclear divisions. The 13th mitosis is presented. Movies were generated taking a new z stack every 125sec and deconvolved. Pictures are presented as maximum intensity projections of a selected number of slices. A 25x/0.8 NA oil immersion corrected objective, 5.1% laser output at 488nm and a pinhole of 178µm were used for imaging. Scaling was x: 0.07µm, y: 0.07µm, z: 0.5µm. The white bars represent 100x10 pixels.

To summarize, GFP::E(Z) remains, although to a residual amount, associated with chromatin during mitosis. The presented images resemble images of PC::GFP at the same embryonic stage published in[112].

4.5 Quantification of GFP and PcG protein molecule numbers in early embryos

Live Imaging of fluorescently tagged molecules is a widely used method to gain qualitative insight into biological processes, however quantification of molecule numbers, which requires calibration, has been difficult, limiting the interpretation of *in vivo* imaging studies. To quantify GFP signals and to obtain biophysical information *in vivo,* virus like particles (VLPs) have been found to be a useful tool [143, 144].

Noninfectious and double-layered VLPs consisting of two capsid proteins from the rotavirus, VP2 and VP6, have a strictly icosahedrical symmetry and stochiometry and can be co-expressed in a baculovirus system [143]. As each VLP consists of 120 GFP-VP2 fusion molecules these particles have been used for calibrating GFP fusion protein molecules in living cells [144].

In order to quantify the players of the PRE-mediated epigenetic memory mechanism, we employed VLPs for quantifying GFP fusion protein molecules. Numbers of endogenous proteins were then estimated and via the expression ratio of GFP-fusion to endogenous

proteins in transgenic embryos of 2-3 hours age (quantified as described in 4.1 and [112, 115]) for PcG, TrxG and DNA-binding protein molecules in the blastoderm embryo.

VLPs at a concentration of 2mg/ml (2,4 $x10^{13}$ particles/ml) were obtained as a kind gift from *A. Charpilienne*, UMR CNRS-INRA Virologie Moleculaire et Structurale, France. VLPs were imaged with laser-scanning confocal microscopy under the same conditions as blastoderm embryos by *J. Fonseca* in our lab. Imaging and Quantification are explained in the figure legend of Fig. 29.

Negative stain electron microscopy was employed to confirm the integrity of the VLP sample used (Fig. 28).

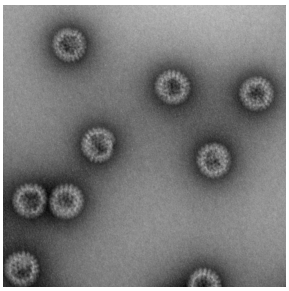

Figure 28: Negative staining of virus like particles (VLPs)

VLPs were diluted with 20mM Pipes, 10mM $CaCl_2$ to a concentration of $3x10^7$ particles/ml and were negatively stained for electron microscopy with Morgagni 268 using protocols from the IMBA/IMP EM facility. The image was taken at 36000x magnification.

	# of GFP fusion/nucleus	# of endogenous protein /nucleus	concentration (uM)
DSP1	3188967.00	1782778.97	0.96
PHO	652.00	3571.06	0.18
PC	83124.00	12668.23	0.13
E(Z)	8213.00 */ 75445.00 **	43615.89* / 43489.19**	0.70* / 0.21**
ASH1	30411.00	8688.86	22.64

Figure 29: Quantification of GFP and endogenous protein in the nucleus of blastoderm embryos

Nuclei from blastoderm embryos were imaged using the Zeiss LSM 510 laser scanning confocal microscope with a 63x oil immersion objective and 2.6x zoom. Scaling was x: 0.051μm, y: 0.051μm, z: 0.15μm. Laser intensity and detector gain varied depending on the intensity obtained from different GFP fusion proteins. VLPs obtained from Charpilienne [143] at 2mg/ml were

diluted 1:20 in H_2O and imaged under respective conditions. Images of nuclei and VLPs were deconvolved using Hygens Professional and an acquired PSF and signals were thresholded with Imaris image processing software. The average nuclear volumes and respective sum intensities of indicated transgenic lines on the left panel were used for further calculations. Signals from VLPs were binned $1x10^4$, and the mode of the distribution was taken as the intensity of 1VLP, which consists of 120 GFP molecules. Confocal microscopy and Image Processing were performed by *Joao Fonseca*.
The I/nucleus was defined as the thresholded sum intensity value/nuclear volume obtained from Imaris, subtracted by the total background intensity in the nuclear volume. From that value, the number (#) of GFP fusion protein molecules per nucleus was calculated by dividing the product of (I/nucleus)*120 with the mode intensity value of 1VLP under the respective conditions measured with Imaris and correcting with the background factor ((sum intensity-background intesity)/sum intesity).
The number (#) of endogenous protein per nucleus was calculated by dividing the number of GFP fusion protein molecules with the product of the ratio of the transgene/endogenous protein from quantified western blots of 2-3 hour embryo collections and the factor difference to wildtype levels. These numbers are plotted in Fig. 30.
The concentration (nM) of endogenous protein levels in the nucleus were calculated by converting the number of molecules/nucleus to molecules/μm^3 and by dividing by the avogadro constant to give mol/μm^3. These values are expressed as μM concentrations.
Transgenic lines examined in this study were: *pα1Tubulin-GfpDsp1, pα1Tubulin-Gfppho, pPc-PcGfp, pE(z)-GfpE(z)* (*), *pα1Tubulin-GfpE(z)*(**), *pα1Tubulin-Gfpash1*

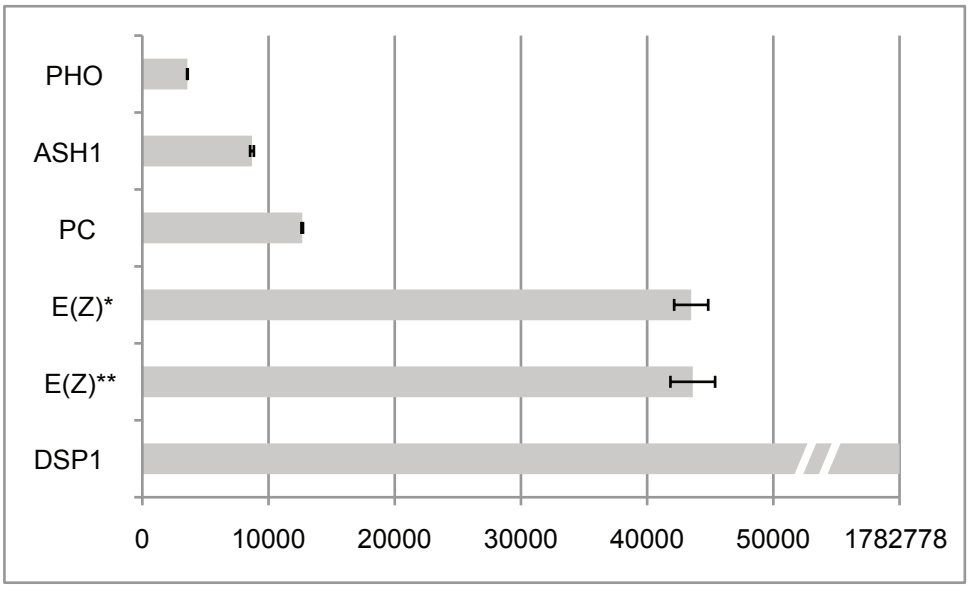

Figure 30: Numbers of endogenous PcG, TrxG and DNA binding proteins in the nucleus of blastoderm embryos

(Fig.30)
The calculated numbers of endogenous molecules per nucleus from Fig. 29 were plotted as a chart. Error values reflect standard error of measurements from several nuclei.
*… *pE(z)-GfpE(z)*, **… *pα1Tubulin-GfpE(z)*

With this method for GFP signal calibration, numbers for endogenous protein molecules spanning a range of 4 orders of magnitude were estimated. (Fig. 29 and 30)
The DNA binding protein DSP1 was calculated to be the most abundant of the measured proteins, has also other biological roles beyond PcG silencing and has many targets, which may be reflected by the measured high number of DSP1 molecules.
The PcG protein PC was also estimated to be abundant, in large excess over the estimated number of PREs in the genome, which is about 200-400 as assessed by chromatin IP in embryos [75]. PHO was less abundant than PC, present at approximately the same number of molecules as there are PHO binding sites in PREs [71]. Based on the calculation of Ringrose et al 2003, there are approximately ten PHO sites per PRE and genome wide ChIP estimates 200-400 PREs per genome [75].
ASH1 and E(Z) were also highly abundant, present at more molecules per nucleus than there are PREs in the genome [71].
Therefore, these numbers are likely to reflect the differences in the biological requirement for different players of the PcG/TrxG/DNA-binding proteins mediated mechanism and the abundance of target sites in the genome.

4.6 Fluorescent tagging of a PRE for live imaging

4.6.1 PRE tagging strategy and transgenic fly lines

Quantitative live imaging experiments on the fluorescently tagged PRC1 PcG proteins PC and PH on individual bands on polytene chromosomes in larval salivary gland nuclei have shown that the proteins have different kinetics at different loci, independently of locus size and intensity [115]. This is consistent with a previous study indicating different complex stabilities at different target genes [116]. Both of these studies imply a different PcG complex composition and suggest that activation and repression are dynamically controlled by chemical equilibria between functional complexes and their targets.
A fundamental experiment to elucidate the function of repressing PcG complexes, activating TrxG complexes and recruiting DNA-binding proteins at PREs is to study PcG, TrxG and DNA-binding protein kinetics at a single PRE locus of defined activity. Therefore, I established a fluorescently tagged transgenic PRE to study the kinetics of functional complexes of the selected GFP fusion proteins (4.1) in context of PRE activity.
For this study, the well-defined 3.2kb fragment of the *Fab-7* PRE of the homeotic *BX-C* complex was selected and a switchable PRE design as described in [6] was constructed

(Fig. 31). The Fab7 PRE was cloned downstream of an upstream activating sequence (*UAS*) [145, 146] to control the transcriptional state of the PRE [101] and the corresponding reporter gene *miniwhite* [146], which also served as transformation marker. It has been shown using similar constructs, that ectopic and transient activation of the Fab7 PRE via the UAS element in embryogenesis is sufficient to switch the PRE to the active state, which is maintained epigenetically until adulthood [6, 7]. Upstream of the PRE a *LacO* sequence array was cloned to visualize the transgenic locus, as a fluorescently tagged LacI repressor protein can bind the *LacO* operator sequence [123, 124]. To minimize influence of the LacO sites on the PRE, a 9.8kb yeast spacer DNA was placed between the LacO repeats and the Fab7 PRE [147, 148]. Three different versions of this transgene were constructed, with 2.5kb (64 copies), 5kb (128 copies) and 10kb (256copies) of the *LacO* repeat. Only transgenic flies with the 2.5kb array were obtained (Fig. 31 B). To test for the influence of the *LacO* array and the spacer DNA on *Fab7* PRE activity, control transgenes without a LacO array, and additionally without the spacer DNA were generated (Fig. 31 C and D).

Other transgenes contained the mCherry::LacI fusion protein under control of strong house keeping gene promoter sequences (*Actin5C*, *α1Tubulin* and *Ubiquitin63E*) A truncated version of the LacI protein was used lacking sequences for tetramer formation, in order to minimize interference [123]. The construct also contained the SV40 NLS coding sequence C-terminally for nuclear targeting. Surprisingly, only *pubiquitin-mCherryLacI* transgenic flies were obtained, whereas neither *α1Tubulin* nor *Actin5C* promoter driven transgenic flies were found. In order to avoid overlaying the *miniwhite* eye phenotype readout for PRE activity in recombined flies that carry both *pubiquitin-mCherryLacI* and *Fab7* transgenes (Fig. 31 A+B, A+C, A+D), pKC27 vectors for *mCherryLacI* transgenes did not contain *miniwhite* as transformation marker. Instead, transgenes were screened by single fly PCR.

All constructs were cloned in the pKC27 vector and site-specific integration employing the ΦC31 integrase (see 4.1 and 3. Materials and Methods) enabled all experiments and controls to be performed at the same genetic locus, thus avoiding variable effects of genomic position on PRE transgene behaviour. The *mCherryLacI* transgenes were injected into embryos carrying the 43.4 landing site on chromosome 2L at position 38E3, and all Fab7 containing transgenes were inserted into the 43.16a landing site at position 46E1 on chromosome arm 2R. For generating flies carrying both a Fab7 PRE transgene and the *pubiquitin-mCherryLacI* on the same chromosome, flies carrying one of each

transgene were crossed and the progeny was screened for *miniwhite* expression and for the presence of mCherry by PCR.

(For details on cloning, transgenesis and screening see Materials and Methods.)

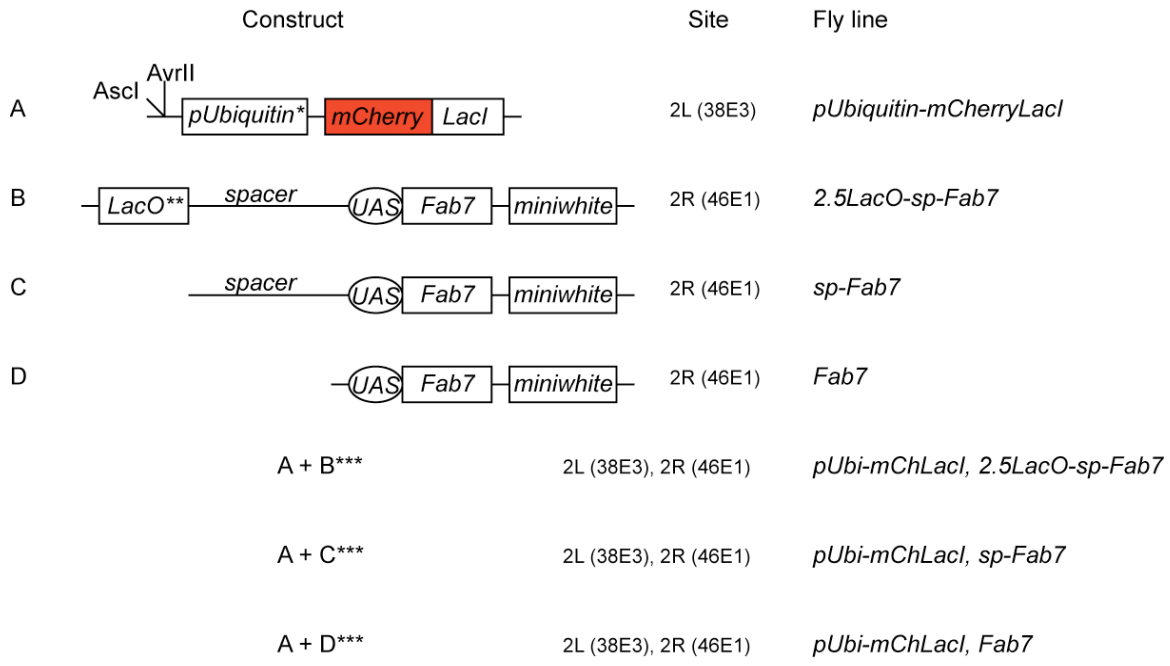

Figure 31: Schematic representation of established transgenic fly lines for fluorescent PRE-tagging

Constructs in (A-D) were cloned into a pKC27 vector backbone and were integrated into 43.4 (A) and 43.16a landing sites (B-D) with site-specific integration by the ΦC31 integrase.

The construct in (A) carries two unique restriction sites for inserting the DNA for a GFP fusion protein to combine fluorescent PRE-tagging with live imaging of a selected PcG protein (4.1). Three different promoters were cloned for mCherry::LacI fusion protein expression (*… the *ubiquitin63E*, the *Actin5c*, or the α*1Tubulin* promoter). Constructs were cloned analogously to GFP fusion proteins.

Construct (B) was cloned with 2.5kb, 5kb or 10kb LacO repeats 10kb upstream of the Fab7 PRE/TRE (ref.) sequence separated by yeast spacer DNA. For inducing PRE/TRE activity by additionally introduced Gal4, an upstream activating sequence (*UAS*) was cloned. The complemented *miniwhite* gene was used as transformation marker and reporter for PRE/TRE activity in the eye. Constructs were cloned to make control fly lines that lack LacO repeats (C) and the spacer DNA (D).

***For fluorescent locus-tagging and control experiments, flies carrying the mCherry-tagged repressor LacI (A) were recombined with flies that have the transgenic Fab7 element (B-D). For details see Materials and Methods.

4.6.2 Analysis of the transgene PRE function

PC binds the transgenic PRE

In order to test whether the transgenic Fab7 PRE in the 43.16a landing site can recruit PcG proteins, polytene chromosome immunostainings of larval salivary gland squashes were performed with an anti-PC antibody and as described in Materials and Methods.

The landing site 43.16a is inserted at position 2R 46E1, which lies between the *engrailed* PRE and a binding site at position 46C. The empty landing site locus did not show PC binding (Fig 32 A). When the transgene 3.2kb *Fab7* core PRE was inserted, a new PC band appeared at this position, as shown in (Fig. 32 B) for preparations from *2.5LacO-sp-Fab7* larvae.

Therefore, the transgene Fab7 PRE is able to recruit the PC protein.

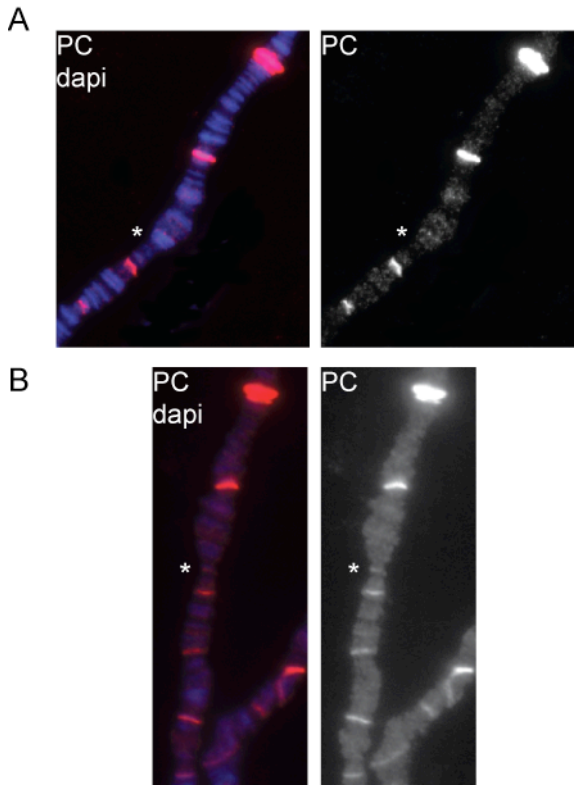

Figure 32: Immunostaining of larval salivary gland polytene chromosomes reveals a new PC binding site at the transgenic PRE locus

(Fig. 32)
Salivary gland squashes from third instar larvae were prepared and PC binding on polytene chromosomes was detected with a rabbit Anti-PC/anti-rabbit alexa647 antibody staining (red). DNA was stained with DAPI. The asterisks in (B) indicate the additional PC binding site on chromosome arm 2R at position 46E1, the site of the 43.16a landing site, whereas there is no PC detected in the empty landing site (*Helena Okulski*, (A)).

The transgenic PRE silences the target reporter gene

PREs are known to silence their target genes via the recruitment and function of PcG proteins. The antagonizing TrxG proteins bind the same stretches of DNA but maintain the transcriptional active state of the target gene. In recent studies, it has been shown that PcG and TrxG proteins can also bind simultaneously [75]. Therefore the presence of silencing or activating complexes might not be the only reason for the PRE activity on its target gene. In transgenic assays performed in our laboratory, we have shown with site-specific integration by the ΦC31 integrase that different PREs behave differently in the same landing site and that the same PRE acts differently in different landing sites (Helena Okulski, Mutational and Functional Analysis of Polycomb- and Trithorax Response Elements in Drosophila melanogaster, Diploma thesis, University of Vienna, 2009, and Birgit Druck, Qualitative- and Quantitative Comparison of Polycomb-/Trithorax Response Elements in Drosophila melanogaster, Diploma thesis, FH Campus Wien, 2007). Thus, these sequences can act as silencing and activating regulatory DNA, influenced by the genomic position of the transgene, and might not be silencing elements by default.

In order to investigate PRE activity of the 3.2kb Fab7 fragment in the 43.16a landing site, and the influence of the *UAS* sequence, the spacer DNA from yeast, and the *LacO* repeat and mCherry::LacI binding on its activity, I investigated the *miniwhite* reporter gene activity, shown by the eye colour, as a readout of PRE function. All transgenic fly lines used in this study had a mutated *white*$^{-}$ background on the X chromsome and would have had a white colour without transgenic *miniwhite* expression, and carried transgenes homozygously.

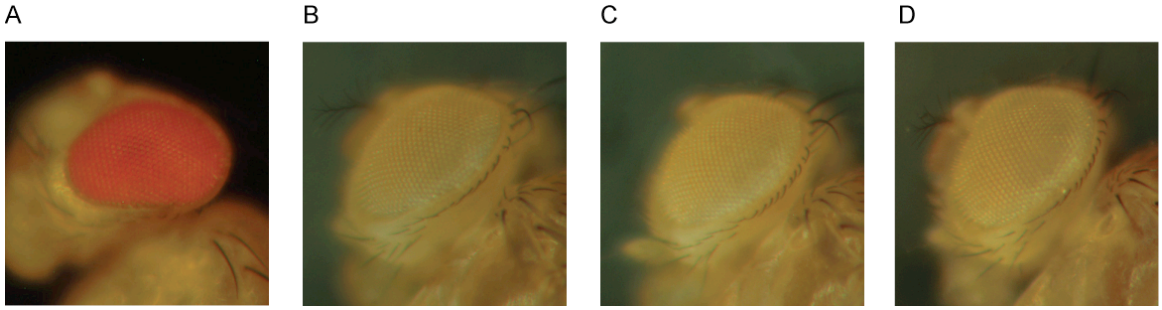

Figure 33: *Miniwhite* expression in transgenic adult flies

Male adult flies of the same age were collected and the reporter gene expression of *miniwhite* was documented in homozygous 43.16a fly lines with site-specifically inserted *miniwhite* only, as control (A), *UAS-Fab7-miniwhite* (B), *2.5kb LacO-spacer-UAS-Fab7-miniwhite* (C) and additionally homozygously expressed *mCherry::LacI* integrated at 43.4 (D).

In flies that carried site-specifically inserted pKC27 vectors with *miniwhite* only in the 43.16a landing site, transgenic *miniwhite* was expressed, and adult flies had red eyes (Fig. 33 A). However, when *miniwhite* was under *Fab7* PRE control (Fig. 31 D and Fig. 33 B), the PRE silenced the transgene, and adult flies had almost totally white eyes. This phenotype did not change in the presence of the spacer DNA and 2.5kb *LacO* array (Fig. 31 B and C, Fig. 33 C). In recombined flies that expressed the *mCherryLacI* under *pUbiquitin63E* control in addition to carrying the *LacO* tagged PRE reporter construct (Fig. 31 A+B), the binding of LacI to *LacO* could theoretically disrupt the silenced state of the PRE. However, as can be seen in (Fig 33 D), this was not the case, as the eyes of flies carrying both components were identical in colour to those carrying the tagged PRE contruct alone. To complete a correct conclusion, this experiment did not give the proof that LacI is binding to the *LacO* operater sequence.

Taken together these results demonstrate that the 3.2kb *Fab7* PRE silences the target gene in its transgenic position and therefore potentially serves as an activateable PRE for future studies. Furthermore, fluorescent tagging with the LacI-*LacO* system does not interfere with PRE activity.

However, switching the 3.2kb *Fab7* PRE to an activating state in this system was not performed in the course of this work, and remains as an objective for future studies.

Binding of a Gal4 transcriptional activator to the *UAS* sequence entails PRE activation and the phenotypic red eye color due to the activated *miniwhite* reporter in a *white*$^{-}$ mutant background. A *Fab7* transgenic PRE can be switched from a transcriptionally silent to an active state by crossing to a driver line carrying a heat-shock inducible Gal4 transcription

factor. Heat shock is performed during embryogenesis, and results in robust and permanent activation of the PRE, which is epigenetically maintained, even after removal of the Gal4 activator. This enables the transition of a silent to an active PRE. The activated state is maintained throughout mitosis and meiosis [6, 7, 101, 149].

To switch the fluorescently tagged and silencing *Fab7* PRE, which has been established in this work, recombined lines (Fig 31 A+B, A+C, A+D) will be further crossed to temporary and inducible expressing heatshock Gal4 driver lines, enabling the PRE to be switched to an active state. For the setup of this experiment, a driver line that encodes Gal4 on the third chromosome is needed. As those available are marked with *white* and have dark red eyes, they cannot be used in the switching assay whose readout is the *miniwhite* activity controlled by the PRE. The construction of a 3x*P3 GFP* marked *hsGAL4* transgene is currently underway in our laboratory and will enable the switching experiment to be performed without interference from a second copy of *white*.

4.6.3 Live imaging of the tagged PRE in larval salivary glands

In embryos, the PcG proteins have been reported to form about 100 foci in the nucleus. In larval wing discs, these foci coalesce to form about 30-40 sites, whilst in larval salivary gland nuclei, individual loci are visible in the microscope as distinct bands [6, 67, 112, 115]. Thus each of these tissues potentially offers the opportunity to study the interaction of PcG/TrxG proteins with a single transgenic locus. However, to study GFP fusion protein kinetics at the visually tagged transgenic *Fab7* PRE, the best tissue in Drosophila is the third instar larval salivary gland containing huge polytene chromosomes, which make it possible to distinguish single bands *in vivo* (4.2.3, 4.3.3 and 4.4.3).

To investigate whether mCherry::LacI expression under *pUbiquitin63E* control and the size of a 2.5kb LacO (64x) repeat are sufficient to visualize the transgenic *Fab7* locus in the landing site 43.16a, a stable recombined line was established carrying both the tagged PRE and the mCherry::LacI construct (Fig. 31 A+B). Salivary gland nuclei were dissected and immediately imaged with laser-scanning confocal microscopy (see 3. Materials and Methods). In most nuclei, the transgenic locus was easily detected as a single band with settings for live imaging (Fig. 34). No background fluorescence was observed under the settings used, indicating that mCherry::LacI is neither strongly overexpressed nor very stable.

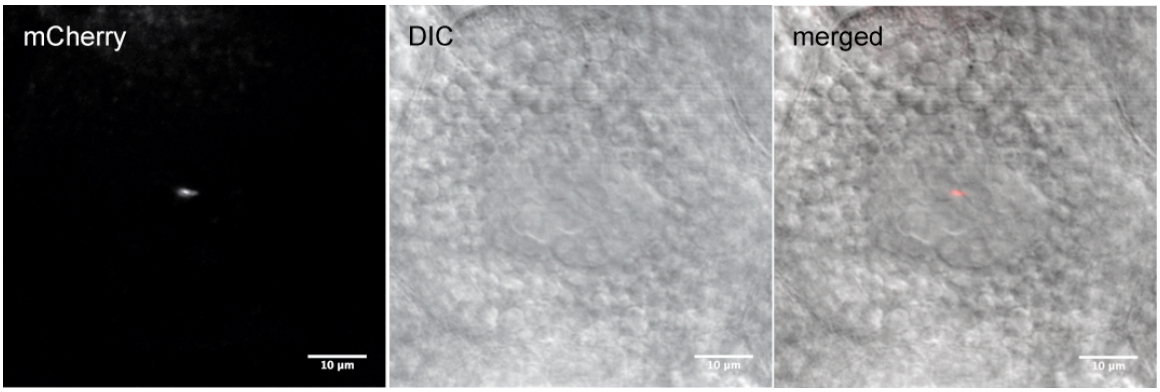

Figure 34: Live Imaging of the fluorescently by mCherry::LacI-*LacO* recognition tagged trangenic *Fab7* PRE

mCherry::LacI distribution a in living salivary gland nucleus is shown of *pbi-mCherryLacI, 2.5LacO-sp-Fab7* homozygously transgenic larvae. The image was taken with a 63x/1.4NA oil immersion objective. 561nm excitation with 13% laser output and a pinhole of 272µm was used. The corresponding DIC image is shown in the middle, and the merge on the right. The white bars represent 10 µm.

In summary, the strategy designed to visually tag the *Fab7* PRE in a transgenic locus gives good results in salivary gland nuclei under conditions for live imaging and will enable in future experiments to distinguish this one PRE of defined transcriptional state out of many binding loci of GFP fusion proteins introduced in this work. (See 5.6 Future Perspectives).

5 Discussion

5.1 Transgenic fly lines expressing GFP-tagged members of the PcG

In the present work transgenic fly lines were established, which express a selected member of the PcG that is fluorescently tagged with GFP, to be able to study the dynamic behavior of the PcG and TrxG mediated epigenetic memory mechanism *in vivo* with advanced microscopy. Proteins of the PcG and TrxG with different functions and involved in different complexes were chosen to be investigated.

The focus of this study is on PcG proteins, which maintain the transcriptional silent state of a target gene. However, also a TrxG member associated with the memory of transcriptional activation was initially selected to be able to investigate the potential PcG/TrxG equilibrium [115, 116]. In a diploma thesis, which I supervised, GFP-tagged ASH1 expressing lines have been established and thoroughly analyzed by Eva Dworschak (*Generation and Characterization of ash1-egfp Transgenic Fly Lines for Live Imaging of ASH1*, at Fachhochschule Campus Wien, 2009). The TrxG member *ash1* was selected, because it functions exclusively in the epigenetic memory mechanism, contrary to other TrxG proteins that are involved in general transcriptional processes and chromatin remodelling (see 2.2.2). Compared to TRX, which is also a specific histone methyl transferase establishing activation-correlated histone marks, it is smaller (approximately 250kDa for ASH1 [45, 150] to 370kDa for TRX [150]) and therefore better amenable to molecular cloning techniques and transgenesis. Furthermore, recent publications on *ash1* in *Drosophila* have proposed a central role of ASH1 in TrxG mediated epigenetic memory [55, 151], however no kinetic studies to investigate its behavior *in vivo* have been performed.

E(z) was selected for this study as it is the catalytic subunit of the PRC2 complex, methylating repression-associated histone marks (see 2.2.1). The kinetics of the PRC1 proteins PC and PH have been studied, giving interesting insights into their behavior and leading to a model for PcG silencing in which repression is allowed by mass-action chemical equilibria [115, 116]. The PRC2 complex has been proposed to have a different function in the epigenetic silencing mechanism than the PRC1, namely establishing

repressive histone marks on H3K9 and H3K27 and therefore, designating chromatin for repression (see also 2.3.2). Therefore, this study investigated the PRC2 member *E(z)*.
As the PRC1 and 2 core complexes as well as the known TrxG proteins do not bind PREs directly, they are thought to be recruited to their target sites by specific DNA binding factors recognizing short sequence motifs in PREs. Two of the known DNA binding proteins, PHO and DSP1 were selected on the basis of recent studies [10, 39, 64, 131] to generate GFP fusion proteins in this work.

To study the proteins of interest by live imaging, I aimed to have physiological fusion protein expression levels and sufficiently detectable signal intensities under live imaging conditions. Therefore, different constructs of the selected proteins tagged with GFP were cloned, which had different promoters for driving fusion protein expression. The ubiquitous and constitutative promoters of the *Actin5c* and the *α1Tubulin* genes were selected for strong transgenic fusion protein expression. Both promoters are commonly employed to drive high expression of the binary Gal4-UAS system in all tissues and at all developmental stages [152] (http://flystocks.bio.indiana.edu/Browse/misc-browse/gal4.php). Additionally, the presumed endogenous regulatory regions were cloned. As the endogenous promoters are ill defined, the genomic DNA sequence upstream of the first codon up to the gene region of the adjacent gene according to GBrowse (http://flybase.org/cgi-bin/gbrowse/dmel/) was amplified, sequenced and cloned to drive expression of the transgene. Similarily, 700bp upstream of the endogenous *Pc* gene have been used for driving PC::GFP, which ensured appropriate levels of fusion protein expression [112]. In the case of transgenic PH::GFP expression, the *Ph* promoter has not been characterized. Instead, the *Pc* promoter was used in addition to a UAS regulatory sequence to drive transgene expression at physiological levels in salivary gland nuclei, but not in embryos and larval tissues, therefore the Gal4-UAS system was employed to enhance protein expression to levels applicable for live imaging [115]. In this study the endogenous regulatory sequence was not tested.

Transgenic fly lines of almost all constructs in the landing site were obtained. Three *GfpPho* lines with transgene expression under the *pho*, the *Actin5c* and the *α1Tubulin* promoter were viable when homozygous for the transgene. In the screen for transgenic GFP::DSP1 and GFP::ASH1 no lines expressing the fusion protein under *Actin5c* promoter control were found. Interestingly, GFP::E(Z) under *Actin5c* promoter control resulted in a stable fly line only with the transgene in a heterozygous state. Few

homozygous flies were observed, however could not be bred as a line. This suggests that the *Actin5c* promoter drives the fusion protein expression to a level that did not permit proper development to adulthood. Interestingly, the *α1Tubulin* promoter is also a strong ubiquitous and constitutative promoter, however the fact that viable homozygous lines were obtained suggests that the *α1Tubulin* drives transgene expression to a more physiological level than the *Actin5c* promoter in the transgenic locus of this study.
Fly lines for all tagged proteins of interest with endogenous regulatory sequences could be established. Except for GFP::DSP1 under the presumed *Dsp1* promoter, stocks of all lines could be generated that were homozyous for the transgene. Driving GFP::DSP1 by the presumed endogenous promoter resulted only in heterozygous transgenic flies. This could be due to an overdose of functional GFP::DSP1, because the cloned *Dsp1* regulatory sequence drives transgene expression to a high level (see 4.2 and 5.4). This endogenous regulatory sequence was the only cloned presumed promoter that gave strong transgene expression (see 5.3 and 5.4). Endogenous regulatory sequences of *pho* and *E(z)* were not strong enough promoters to drive expression levels suitable for live imaging, and in the case of *pho* the endogenous promoter driven transgene was not sufficient to rescue the lack of endogenous PHO (see 5.5 and 5.2).
In summary, the strategy to clone different promoters to test for optimal transgene expression levels was successful in establishing a set of hetero- and homozygously stable fly lines of each fusion protein of interest.

The phiC31 integration tool was used to make transgenic flies with unique and specific transgenic sites carrying a cassette for fusion protein expression. This method was selected out of three available methodologies for transgenesis in Drosophila [153-155]. A transgenic construct carrying the recognition site *attB* can be specifically inserted into a landing site, which contains the *attP* recognition site for the provided phiC31 integrase and, which is unique in the respective landing site fly line. This method is rapid and efficient, if an established library of different characterized landing site fly lines, which are generated by P-element transgenesis, is available. P-element transgenesis uses vectors that are based on the P transposon and generates different fly lines that have unpredictable and sometimes multiple insertions of the transgene. This transgenesis approach was not used directly in this study, because insertion sites have to be subsequently mapped and characterized. The novel technology of homologous recombination in Drosophila was also not employed, as it consists of the time consuming sequence of first random P-element transgenesis and in consecutive fly generations

mobilization of the transgene to the homologous locus. Therefore, the most efficient method to obtain characterized and comparable transgenic flies was selected for this study. Here, I have performed a rapid screen to identify the best expression strategy by generating many comparable transgenic flies expressing GFP-tagged fusion proteins of interest, and selected promoters that drive the transgene to a level optimal for both genetic rescue and live imaging. For future extensions of this project, GFP-tagged proteins that showed a rescue can be selected for homologous recombination to avoid the presence of endogenous protein in kinetic experiments.

5.2 Rescue experiments reveal functional fusion proteins

The N-terminally tagged GFP fusion proteins expressed in established transgenic fly lines (see 5.1) were tested for their functionality in a rescue experiment, in which the fusion protein has to replace the respective endogenous protein. For this, the transgenic lines carrying a fusion protein cassette on the second chromosome were crossed into a homozygously lethal mutant background on another chromosome, and the functionality was quantified according to the occurrence of viable flies. For that purpose, marked (balancer) chromosomes that also encode the respective endogenous proteins have to be crossed into the transgenic lines. Then, these lines are crossed to fly lines carrying the homozygously lethal mutant of the endogenous protein. Subsequently the viable progeny that carry both the homozygously lethal mutation and the transgenic fusion protein, that therefore substitutes the endogenous protein, confirm fusion protein functionality. In this work, GFP::PHO and GFP::E(Z) have been shown to be functional fusion proteins by this experimental procedure. As recently demonstrated by Eva Dworschak in the lab, also GFP::ASH1 is a functional fusion protein that fully rescued *ash1*^{10}*/Df(3L)* and *ash1*^{22}*/Df(3L)* transheterozygotes.

The *pho* rescue experiment was performed with the *pα1Tubulin-Gfppho* and the *ppho-Gfppho* lines. Marked 4th chromosome transgenic *pph-Gfppho* and pα*1*Tubulin-*Gfppho* lines were established, however, flies of the *Actin5c-Gfppho* line could not be crossed to the balancer *ci*^{D}*/ey*^{D} line, because *ey*^{D} was not apparent in the transgenic background. Therefore, the line could not be established with marked chromosomes to distinguish mutants in the final cross. As mentioned earlier (4.3.1), PcG proteins regulate eye development [139] and the PHO protein has been detected at the *eyeless* gene [75], therefore is likely involved in its regulation. Overexpression with this strong promoter (see

5.1, 5.4) could suppress the mutation to an extent that it loses the typical phenotype. Rescue experiments with *pα1Tubulin-Gfppho;;ci^{D}/ey^{D}* and *ppho-Gfppho;;ci^{D}/ey^{D}* lines revealed results that were dependent on the expression strategy and correlated level of transgenic fusion protein. Whereas the endogenous *pho* regulatory sequence did not rescue the homozygous lethality of the *pho^{1}* mutation, the much stronger *α1Tubulin* promoter gave a full rescue when homozygously expressing transgenic GFP::PHO. Only three outliers of more than 400 counted flies were rescued in the experiment with *ppho-Gfppho*. Interestingly, they were all males. This could be connected to the expression of the attached *miniwhite* cassette, which also gives a stronger eyecolour in males due to the X chromosome origin of *white*, or may be simply coincidence. Definitely, the presumed *pho* promoter does not drive transgenic expression to levels sufficient to overcome lethality by lack of functional endogenous PHO. Transgenic GFP::PHO expression under the *α1Tubulin* promoter control gave a full rescue in homozygous and approximately a quarter of a full rescue in heterozygous transgenics. This result demonstrates that GFP::PHO is a functional variant of PHO and shows that not only the structural effects of tagging but also the level of transgene expression significantly influence the outcome of a genetic rescue experiment. The result of the rescue experiment with *pα1Tubulin-Gfppho* did not show a meaningful difference between the counted numbers of genotypes between males and females, therefore this difference in expression strength between males and females is not perceivable with strong promoters.

GFP::E(Z) was also shown to be a functional fusion protein that can replace endogenous E(Z) in experiments with transgenic *pE(z)-GfpE(z)* and *pα1Tubulin-GfpE(z)* lines. When crossed to third chromosome balancers in the preceeding step of the rescue experiment, the *pActin5c-GfpE(z);Dr/Tm6c* line was not viable, which may be caused by the genetic load of the *pActin5c* driven transgene with the balancer chromosome and *Dr* mutation in combination. Expression of GFP::E(Z) by the *α1Tubulin* promoter showed a full rescue in homozygous and interestingly also heterozygous transgenes. In the case of E(Z) either lower levels of GFP::E(Z) are sufficient or the *GfpE(z)* transcript and/or GFP::E(Z) protein are more stable and therefore can replace endogenous E(Z) functions for a longer period of time than GFP::PHO. Interestingly, although the presumed endogenous promoter sequence results in a full rescue in the homozygous condition, the heterozygotes were not rescued, suggesting that the exact expression level could be critical at a certain stage or throughout development. Both N-terminally GFP-tagged PHO and E(Z) are functional PcG proteins and show specific dependencies on the expression strategies tested.

By the same experimental procedure the functionality of GFP::DSP1 was tested. Similar to *GfpE(z)* fly lines, stocks with first chromosome balancers and transgenes driven by either the endogenous regulatory sequence or the *α1Tubulin* promoter could be established. In contrast, stocks with the same first chromosome balancer and the transgene driven by the *Actin5c* promoter could not be established, because the *Binscy,w/C(I)DX,y*^{1}*,f*^{1}*;pActin5c-GfpDsp1* progeny was not viable. Two lines for rescue experiments were considered to be sufficient to obtain a statistically significant result if GFP::DSP1 is a functional fusion protein. However, lethality was temperature-independent, most likely due to an additional homozygously lethal mutation on the *Dsp1*^{1} chromosome. Nevertheless, the *GfpDsp1* allele could not rescue *Dsp1*^{1} temperature-sensitive homozygous lethality. Therefore, the functionality of GFP::DSP1 could neither be confirmed nor refuted. Unfortunately, the *Dsp1*^{1} mutation is the only lethal mutation of *Dsp1* known, therefore, there are no more alleles available to perform the rescue experiment with. For future studies, the chromosome of *Dsp1*^{1} could be crossed out to exchange the genetic background on this chromosome.

Additionally to the rescue experiments, double stainings of polytene chromosomes were performed to test whether the endogenous proteins and the GFP-fusion proteins bind to the same target loci. GFP::DSP1 expression under the *α1Tubulin* promoter showed that GFP::DSP1 participates in correct complex formation, since GFP and DSP1 signals almost completely overlapped and could be mapped to previously characterized PcG target sites (see 4.2.1). Unfortunaltely, this could not be demonstrated for GFP::PHO and GFP::E(Z) because of unsatisfactory antibody qualities.

Recently, both colocalization (see *Generation and Characterization of ash1-egfp Transgenic Fly Lines for Live Imaging of ASH1,* diploma thesis by Eva Dworschak at Fachhochschule Campus Wien, 2009) and fusion protein functionality in a rescue experiment (data not shown) could be ascertained for GFP::ASH1. Therefore, future investigation in demonstrating the functionality of N-terminally GFP-tagged DSP1 is likely to be successful.

Previous rescue experiments on GFP tagged PcG fusion proteins investigated the functionality of PC::GFP [112] and PH::GFP [115]. To determine the functionality of the PC fusion protein, the transgenic lines were crossed to different *Pc* alleles in a similar procedure as performed in this work. The authors found that the transgene rescued mutations in the chromodomain, but not in the C-terminal part of *Pc* or null mutations. They showed that target specificity was retained, however the GFP tag interfered with the functional domain located at the C-terminus [112]. Therefore, it could be possible that a so

far unknown aspect of PC was impaired or that the repression complex could not properly accommodate all PC proteins containing GFP [112, 115]. PH::GFP was able to rescue a homozygous null mutation and is therefore totally competent for repression [115]. The rescue experiments for the N-terminally tagged GFP::PHO and GFP::E(Z) in this study demonstrated functional fusion proteins that fully rescued homozygously lethal mutations until adulthood.

5.3 Quantification of transgenic expression show complex regulation

Quantification of western blots was performed to investigate fusion protein levels and transgenic effects on endogenous protein levels in crude protein extracts of 0-12h embryos. Additionally, crude exptracts of third instar larval salivary glands and wing imaginal discs were used for western blot analysis, however, except for the preparation of larval salivary glands of GfpDsp1 transgenic fly lines, no interpretable results were obtained. This was due to differences in antibody qualities for this experimental setup, which includes protein extraction of whole tissues. This method is fast and implies little loss of material, but it contains impurities and the probability of background from cytoplasmic protein is high. Hence, different protein extraction and purification strategies such as nuclear protein isolation of larger samples could increase the quality of western blots obtained particularily with anti-PHO and anti-E(Z) antibodies.

Transgenic and endogenous protein levels in embryos and salivary glands were investigated, because they were of interest for live imaging studies and needed for VLP-based determination of molecule numbers (4.5). In this sudy, overnight collections of embroys were used, which represent a mixture of ages from 0-12 hours, giving average values over this developmental time window. For further investigation into developmentally dependent changes of protein expression in the embryo, western blot analysis of staged embryos is currently being performed in our lab. Additionally, preparations of other tissues of interest for future live imaging studies could give better results for quantifying protein levels in transgenic lines.

A further limitation of the Western blot quantification reported here is that the fusion protein level has to be high enough to be detectable by the method of choice. In this study, analog chemiluminescence films and the digital Kodak Molecular Imager were employed for signal detection (See 3. Materials and Method and 4.2.2). The signals of PHO and E(Z) on western blots were not strong enough for detection by the Kodak

Molecular Imager and could not be employed for quantifying band intensities by this approach. The DSP1 protein was detectable. Both methods gave similar results for GFP::DSP1 levels in embroys and salivary glands. GFP::PHO was also hardly detectable after appropriate exposure times on chemiluminescence films that let distinguish bands from background. Due to non-specific staining of the anti-PHO antibody, bands could not be detected after too long exposure times from background smear. Importantly, the more similar the protein level of transgenic and endogenous protein the more accurate is comparative quantification, due to limitations in the dynamic range of both detection methods. If the intensity level in pixels of the stronger of the two compared bands reaches its maximum, the amount of protein loaded cannot be correlated anymore. On the other hand, if the intensity in the stronger band is in the range of graylevels, the fainter band can be too weak for detection. Therefore, the amount of GFP::DSP1 to DSP1 in embryos of pDsp1/CyO flies or the amount of endogenous E(Z) compared to GFP::E(Z) in *pE(Z)* and *pαTubulin* lines, may be less than estimated (4.2.2 and 4.4.2).
The ratios used for calculation of molecule numbers shown in Fig 29 were measured by Philipp Steffen after this study was completed, overcoming the problems encountered in this study by introducing the following optimization steps: 1) staged embryos were used, and 2) fluorescence instead of chemiluminescence for detection was measured.

Additionally to Western blot analysis for quantifying protein levels, real-time PCR on cDNA preparations of 0-12 hour embryo collections weas performed to investigate transcript levels of *Gfp*, *pho* and *E(z)*. GFP::PHO was difficult to quantify by Western blot and therefore transgenic transcription was ascertained. Transcription on *E(z)* was examined to gain more insight into the observed upregulation of endogenous *E(Z)* in transgenic embryos. *Dsp1* was not investigated at the transcriptional level since the Western results were robust, and are most informative for quantification of molecule numbers in live imaging. Transcription of *Dsp1* transgenes could be investigated in future extensions of this study.

In embryos of 0-12 hours age a high amount of endogenous DSP1 is detected on Western blot analysis. In comparison, the amount of DSP1 in salivary glands was significantly lower in the same experiment. The level of GFP::DSP1 was lower than the endogenous protein level, especially in the heterozygous *pDsp1* line. Exceptionally, the *α1Tubulin* promoter drives fusion protein expression in salivary glands at levels higher than DSP1 levels. Accordingly, the cloned endogenous promoter sequence of 2100bp

upstream of the coding region of *Dsp1* drives transgene expression detectably in Westernblot analysis. Generally, GFP::DSP1 levels in the homozygous *pα1Tubulin* line were more than two times higher than in the heterozygous *pDsp1* line, however *pDsp1* was not homozygous viable. Potentially, the high amount of GFP::DSP1 could affect a regulatory sequence present in the presumed endogenous promoter region, but does not effect expression by the *α1Tubulin* promoter. Interestingly, there was a small transgenic effect of up or down regulation for endogenous DSP1. In *pDsp1* samples DSP1 levels were higher, and in *pα1Tubulin* DSP1 levels were lower than DSP1 in wildtype samples. The high presence of GFP::DSP1 in the *pα1Tubulin* line could lead to downregulation of endogenous DSP1. As the promoter of *Dsp1* is undefined, complex regulatory sequences could have been cloned within the 2.1kb fragment. Although total DSP1 and GFP::DSP1 in transgenic flies increased approximately 1.5 – 2 times no phenotypic effects of overexpression were observed.

In transgenic *GfpPho* embryos, GFP::PHO levels were low compared to endogenous PHO. Similarily to GFP::DSP1 in embryos, the fusion protein level was below the endogenous protein level. Under *ppho* control no fusion protein nor *Gfp* transcript was detected, therefore the cloned regulatory region of 4.65kb was despite its large size not sufficent for transgene expression at that genomic position and may not include the necessary information for the proper *pho* expression pattern. Also, GFP::PHO may not be as stable and faster degraded than PHO. Comparing protein expression levels, under the *Actin5c* promoter expression is less than approximately 6 times lower than under the *α1Tubulin* promoter. Remarkably, the relation is the opposite with transcript levels. The *Gfppho* transcript level is higher in *pActin5c* than in *pα1Tubulin* transgenic flies.

Interestingly, the endogenous E(Z) level in embryos varied strongly depending on the promoter driving expression in the transgenic lines. We observed strong upregulation in the heterozygous *pActin5c* line, whereas E(Z) overexpression was significantly less in the *pE(z)* and *pα1Tubulin* lines. In contrast, GFP::E(Z) expression in the *pActin5c/CyO* line was lower than in the other lines, which showed strong fusion protein expression with levels remarkably higher than the endogenous E(Z) level. Compared to embryo preparations of other GFP fusion protein expressing lines, the presumed endogenous and the *α1Tubulin* promoters drove expression to levels significantly above endogenous protein levels. Interestingly, in all *GfpE(z)* transgenic lines the sum of functional E(Z) and GFP(E(Z) was approximately 20-fold increased, representing a massive total overexpression. Remarkably, the transcript levels did not correlate to the protein levels. Quantitative real-time PCR detected no *Gfp* transcript and a slight increase of *E(z)*

transcript over wildtype endogenous *E(z)* levels in *p(Ez)* transgenic embryos. This was in line with an increase of endogenous protein in these samples. Although there was no *GfpE(z)* transcript detected, GFP::E(Z) fusion protein levels were substantial. This result standing alone could be explained with high protein stability. However, in the *pActin5c* line the *E(z)* transcript level was high correlating with an observed upregulation of endogenous E(Z). The *GfpE(z)* transcript level was higher than in *pE(z)*, but less GFP::E(Z) fusion protein was detected than in *pE(Z)* samples. The *E(z)* transcript level in *pα1Tubulin* embryos was also strongly increased due to upregulation of endogenous transcription, but the increase of endogenous E(Z) was not comparatively high.

In all transgenic lines the strongest GFP fusion protein expression was obtained with the *α1Tubulin* promoter. However, in analyzed real-time PCR samples most transcript was detected under the *Actin5c* promoter. The promoter sequences could include a region that is transcribed with the coding region and this sequence could contain structural information, which may influence the transcript stability and translation efficency.
Interstingly, the largest cloned endogenous promoter fragment was *ppho* with 4650bp in size, but this did not drive transgenic transcription to detectable levels in RT PCR and western blot analysis. *pDsp1* of 2100bp and *pE(z)* of 400bp were significantly smaller but contained the information to strongly drive fusion protein expression.
In summary, it is clear from this analysis that different transgenic constructs and expression strategies have different impact on actual transcript and protein levels, which overall do not correlate and allow straightforward conclusions. This could indicate that these proteins are controlled by specific and complex regulatory systems. Also, the genomic position and surrounding chromatin structure of the transgenic locus is probably influencing the expression of the fusion proteins.
It is important to mention that in this study extra copies of at least partially functional variants of PcG proteins were homo- and heterozygously added at an arbitrarily selected genomic position. It is likely that fusion protein expression shows totally different features when the *Gfp* tag is inserted by homologous recombination. As this will give a situation most similar to the endogenous expression, this strategy should be considered for future studies. However, it cannot be guaranteed that expression levels are high enough for live imaging studies. The approach used in this study offers the possibility for a relatively fast screen of different designs for fusion protein tagging and expression. By this method I was able to establish various transgenic lines of functional fusion proteins and analyze optimal expression levels for live imaging studies.

5.4 GFP-tagged PcG proteins enable the investigation of various aspects of the PcG mediated epigenetic memory mechanism by live imaging

Third instar larval salivary glands and preblastoderm embryos were employed for live imaging experiments in this study. Fluorescence intensities of GFP fusion proteins in the obtained stably transgenic fly lines were tested under live imaging conditions. This implicates mild conditions like low laser power, fast line speed and a fast experimental setup to examine the living sample. In this work the technical challenge of studying GFP fusion proteins in whole living organisms and tissues was taken. Contrary to single cells in culture, the observed cells are within their natural environment and represent the best condition to observe physiological processes.

Polytene chromosomes in larval salivary glands offer the possibility to study the differences of the dynamic behavior of PcG proteins at different sites of action [115]. Also, this tissue will be the sample of choice for later studies on a visually tagged transgenic PRE site of defined transcriptional status, as discussed in chapter 5.5.
GFP::DSP1 in salivary glands of third instar larvae of both stable transgenic fly lines (*pα1Tubulin* and *pDsp1/CyO*) showed discrete bands with different fluorescence intensities on polytene chromosomes. As observed in previous studies, DSP1 and GFP::DSP1 specifically bound to approximately 200 binding sites [64, 156].
In *Gfppho* larvae, both constitutative promoters drove expression in salivary glands to a level that was appropriate for live imaging. The endogenous promoter sequence was not sufficient for detection in that tissue, which was suspected because of undetectable transcript and protein levels in embryos. It is possible that detectable levels of GFP::PHO may exist in other tissues or at other developmental stages. Interestingly, in all samples a cytoplasmic signal was detected. Cellular PHO distribution was further tested by larval stainings using an anti-PHO antibody, which revealed no difference of PHO signal between wildtype and transgenic samples and therefore this cytoplasmic signal was not an overexpression artefact. Neither GFP::DSP1 nor GFP::E(Z) showed a cytoplasmic distribution in salivary glands.
GFP::E(Z) was strongly enough expressed in salivary gland nuclei for live imaging experiments in all three transgenic lines. Correlating to the expression levels in embryos, the strongest signals were obtained with *E(z)* and *α1Tubulin* promoters.

Generally, nucleoli were free of GFP-tagged PcG fusion protein signal in all samples. This correlates with the distribution of PC::GFP in somatic nuclei. Interestingly, in the male germ line PC::GFP was associated with the nucleolus, however the function of this subnuclear fraction remains elusive [157] (reviewed in [158]).
Both *Actin5c* (homo- and heterozygous) and *α1Tubulin* promoters drove GFP expression in salivary glands suitable for live imaging with low laser intensities and small pinhole sizes. Expression strategies of using endogenous promoters of *Dsp1* and *E(z)*, but not of *pho*, were successful for driving the respective fusion protein in salivary glands.

In this study, whole preblasoderm embryos were investigated with live imaging microscopy to examine the distribution of GFP fusion proteins during mitosis. During massive mitotic chromosomal reorganization, the vast majority of non-histone proteins dissociate from the chromosomes [142]. However, at least a small fraction of PcG and TrxG proteins could remain on chromatin to reestablish the gene expression pattern of target genes after cell division. As described in detail in 4.2.4, the first divisions of large early embryonic nuclei occur synchonously, and therefore are a good sample of choice for examining the distribution of GFP-PcG proteins during mitosis. This experiment requires the expression of GFP fusion proteins to levels detectable for live imaging at that early stage of development.
Time-lapse microscopy of PC::GFP [112] during synchronous nuclear divisions at early embryo development showed that during mitosis the bulk of PC::GFP was depleted from the chromosomes and distributed into the surrounding cytoplama, though a small detectable fraction remained associated. The decrease in fluorescence intensity started at the onset of mitosis and proceeded until anaphase, after which a steady recovery of fluorescence into the nuclear area was observed. A previously published study [107] employed antibody staining against the PcG proteins PC, PH and PSC in blastoderm embroys did not show a chromatin-associated fraction of PC or PH, but dissociation of all three proteins during prophase, dispersion of protein signals and sequential reassociation of PSC in metaphase and PC and PH in anaphase (PH) to telophase (PH and PC). Therefore, live imaging of fluorescent fusion proteins in combination with thorough analysis of fusion protein expression and function may be a more instructive method to study the mitotic behavior of PcG proteins than antibody-dependent immunofluorescence studies on fixed tissue.
Live imaging of GFP fusion proteins in this study at synchronous nuclear divisions in the preblastoderm embryo monitored the strong association of the DNA binding factor DSP1

of the PcG group throughout mitosis. Similarily to PC::GFP in [112], GFP::DSP1 was detectable from cycle 10 on, so the distribution during the last four synchronous divisions could be captured. Before cycle 10 eventual GFP signals would be hidden in background fluorescence of yolk and yolk nuclei. If GFP::DSP1 is already expressed at earlier stages will be answered by Western blot analysis of staged embryo samples. Strikingly, a large quantity of GFP::DSP1 stayed on mitotic chromosomes. Therefore, DSP1 is likely one of the factors that stay attached at target genes to recruit other members of the PcG mediated epigenetic memory mechanism after nuclear division to maintain the target gene expression state in the daughter cells. Alternatively, the observed association of GFP::DSP1 during early embryogenesis may not reflect its function as a PcG protein, but a more general function of DSP1 as a high mobility group (HMG1/2) protein [10, 130, 159], which have global genomic functions in establishing active or inactive chromatin domains [160]. Interestingly, GFP::DSP1 is depleted of mitotic chromosomes in sensory organ precursor cells (*Joao Fonseca*, unpublished observation), therefore DSP1 binding to mitotic chromosomes could be essential at certain developmental stages and tissues.

Disappointingly, fluorescence intensities of GFP::PHO were too weak for examining its mitotic distribution. Interphase nuclei after cycle 10 showed faintly visible GFP signals after deconvolution, but the gained information were too little to answer if PHO remains on mitotic chromosomes. Alternatively, N-terminally tagged *pho* constructs without the 3'UTR sequence, and C-terminally tagged *pho* constructs were used to establish transgenic fly lines, which were not analyzed so far, but could be used for future studies to examine the distribution of PHO during mitosis.

GFP::E(Z) was also detectable from cycle 10 on, but this fusion protein did not give as strong signal as GFP::DSP1. After deconvolution signal from a minor fraction of GFP::E(Z) was faintly detected on mitotic chromosomes. Therefore, although almost all GFP::E(Z) dissociated, some could stay anchored at chromatin and might be necessary for early PRC2 function after nuclear division.

Recent data in the lab also showed that GFP::ASH1 resided on mitotic chromosomes in embryos and larval sensory organ precurser cells, and therefore ASH1 could represent an anchoring factor for TrxG protein complexes to maintain the transcriptional active state at transcribed target genes.

Interestingly, as observed in [107] and with live imaging experiments in the presented work and extended experiments in the lab, the dispersion of dissociated protein fractions differed between the observed proteins. Whereas PSC was clearly excluded from metaphase mitotic chromosomes [107], other proteins were not excluded from this domain

[107] (and this study). PC disperses into the cytoplasm after prophase [107, 112] and PH remained visible outside the condensed chromatin volume until telophase. GFP::E(Z) stayed near chromosomal territory. This could imply, that for having an organized and economical reestablishment of PcG and TrxG mediated chromatin structures to maintain the transcriptional pattern at target genes during the cell cylcle, some PcG and TrxG proteins stay to a major or minor extent attached to chromatin for a sequential and hirarchical recruitment of functional complexes. These anchoring factors could have key roles in establishing repressing or activating structures. Furthermore, the observed differences in dispersion suggest also nuclear and cytoplasmic organization of PcG proteins during mitosis, and definitely should come under detailed and quantitative investigation in future studies to understand the establishment, function and maintenence of the PcG and TrxG mediated epigenetic memory.

5.5 Transgenic fluorescent PRE tagging generates a single locus for qualitative and quantitative live imaging

To address the function of PcG complexes, TrxG complexes and DNA binding proteins at specific PREs I designed an experimental setup to study GFP fusion protein kinetics at a single RFP-tagged locus of defined activity. This strategy will allow the exmination of a dynamic regulation model of the memory mechanism at target genes by repressive PcG and activating TrxG protein complexes. So far, it has been shown that different loci display different kinetics of the PcG proteins PC and PH in vivo [115], and that binding of PcG complexes vary at different target sites depending on methylation status and transcriptional activity in salivary glands [116], implying different complex stabilities. However, assigning differences of PcG binding to functional differences at PREs could not be attained in the live imaging study of [115] since individual loci cannot be identified in whole unfixed salivary gland nuclei. To achieve this, a transgenic PRE at a genomic locus that is not a PcG targeted region was established in this work to build an artificial system to study kinetic properties of individual components of PcG complexes with respect to activating and repressing PRE function with quantitative live imaging. An alternative strategy would be to visually tag an endogenous PRE, however, this would require considerably more effort to set up by homologous recombination and could be too complex for interpretation, as many endogenous loci contain several PREs in close proximity to each other, that is below the resolution of microscopical methods.

To be able to distinguish the repressive and the activating state of the employed *Fab7* PRE, which is a well-studied PRE of the *Hox* complex, a transcriptional switch was essential. Therefore, the inherent function of the cloned PRE at the transgenic locus had to be ascertained. As a simple readout of this, the reporter gene *miniwhite* showing the extent of its expression in the eye by the level of eye colour pigment was chosen. The prerequisite according to the model of PRE function was that the transgenic *Fab7*, acts as silencer by default. This was shown in transgenic adults, strongly repressing the eye colour referred by the *miniwhite* reporter gene in *white* mutant background. To verify this in salivary glands and other tissues, real-time PCR on *miniwhite* transcription can be performed. For transcriptional switching, the inducible activating sequence *UAS* was added to the transgene design. This should enable the ectopic and stable switch from the silenced to the activating state by temporary Gal4 expression, which activates the *UAS* [6, 7]. In this study, the proof of switching the visually tagged transgenic *Fab7* could not be performed due to the lack of heatshock Gal4-driver lines that are not coupled to *white* or *miniwhite* as transformation markers, but will be performed in the continuation of this work, using GFP marked heatshock Gal4 lines that have recently been generated in the lab.

Also, the visual tagging with specific GFP::LacI-*LacO* binding could influence the function of the transgenic *Fab7* PRE activity. In this work, a 10kb large spacer DNA has been chosen from yeast to distance the *LacO* sequence from the UAS-Fab7-miniwhite cassette. Alternatively, boundary elements like insulator sequences were considered, however the pairing ability of these elements could generate pairing artefacts of PRE activity and *miniwhite* expression like enhanced pairing sensitive silencing. The *miniwhite* reporter gene assay was successfully employed to rule out the influence of the *UAS*, *LacO* and spacer DNA sequences on transgenic *Fab7* PRE function.

Necessary for live imaging, the expression of mCherry::LacI has to be strong enough to tag the transgenic locus, which should not be masked by signal from free (unbound) mCherry. Furthermore, the size of the *LacO* repeat has to be large enough to bind as many mCherry::LacI molecules to generate a detectable signal for live imaging. Therefore, mCherry::LacI expression and the size of the *LacO* sequence have to be in optimal balance. In salivary glands, a 2.5 kb LacO array visualized the transgenic locus as a single band on polytene chromosomes under live imaging conditions. However, in diploid tissue of larval imaginal discs and in larval brain, this tagging was not detectable. Therefore, bigger size transgenics have to be generated or higher mCherry::LacI expression has to be achieved to study the locus in these tissues. In one injection trial,

only 2.5kb transgenic lines were obtained. The huge repetitive sequences of 5kb and 10kb could decrease the chance of integration into the genome.

5.6 Future perspectives

The work presented here constitutes the basis for a variety of future live imaging studies on PcG proteins in Drosophila melanogaster. The generated GFP fusion protein expressing fly lines are promising tools to study functional protein complexes of the PcG mediated epigenetic memory mechanism in various tissues and in developmental context. Using quantitative methods such as FRAP (Fluorescence recovery after photobleaching) diffusion constants and binding rates of GFP tagged proteins to chromatin can be assessed as published for PC and PH [115]. This information can be gained from differentiating and mitotic tissue. By crossing the established transgenic lines to available red fluorescently tagged histone H2B [161], chromatin can be visualized and protein kinetics in chromatic and nucleoplasmic domains can be separately analyzed. These approaches would enable more detailed and quantitative information on the behavior of GFP tagged PcG proteins during the cell cycle than with qualitative time-lapse microscopy as shown in this work. Furthermore, the tools presented here enable the determination of absolute molecule numbers, the monitoring of interrelations between components of the PcG mediated epigenetic mechanism and changes during the development.

To study PcG protein kinetics to correlate complex dynamics to PRE activity, the transgenic flies expressing GFP fusion proteins established in this work will need to be combined with flies that carry the visually tagged transgenic Fab7 PRE. To obtain fly lines that stably carry more than two transgenic loci and if required also homozygously and which can be easily distinguished, the approach of combining transgenic elements into cassettes has been taken. Therefore, the protein expression cassettes of GFP fusion proteins and the mCherry-tagged LacI can be combined by molecular cloning, and additional transgenic lines can be established by injection of the combined construct into the same landing site. Single fly PCR for mCherry to screen for transgenic flies avoids the need of the transformation marker *(mini)white* and recombination with *LacO*-tagged *Fab7* generates stable transgenes on one chromosome with all components to study GFP-tagged E(Z), DSP1, PHO and ASH1 at a single PRE of defined activity.

To demonstrate the switching ability of the transgenic *Fab7*, temporary Gal4 transcription factor binding to the *UAS* regulatory region has to be epigenetically maintained by the

PRE, giving activation of the reporter gene *miniwhite* during several rounds of mitosis. Transient Gal4 induced by heatshock in embroys resulting in transgenic PRE activation maintained until adulthood has been established in [6, 7], but needs to be confirmed in the transgenic position and set up used here including large sequences of multiple *LacO* repeats and spacer DNA added in this work. Therefore, available driver lines with *white* or *miniwhite* are not useful because of covering the intrinsic readout for PRE activity. After the establishment of *3xP3 Gfp* marked *hsGal4* driver lines in the lab, the epigenetic memory of the *UAS* switch can be ascertained. To this end, heat-shock inducible Gal4 transcription factor will be expressed during embryogenesis. The epigenetically maintained activation by the transgenic *Fab7* PRE of the *miniwhite* reporter gene will be detected in the adult fly eye. However, for further studying of the dynamics of GFP fusion proteins at the transgenic PRE, driver lines with GFP expression as a transformation marker may be problematic, as they present a second source of GFP besides the GFP-tagged PcG proteins of interest. Therefore, the examination of GFP fusion protein kinetics at the silent transgenic Fab7 can be performed straightforwardly, however the comparison to the active state requires a temporary source of Gal4 that does not interfere with the setup of this transgenic system by background of (*mini)white* or *Gfp* expression. As the *3xP3* promoter is specific for expression in the eye [162], this may not be a problem in larval salivary glands, however, leaky expression will have to be ruled out by expression analysis in larvae.

6 Conclusions

In this thesis work, the generation and characterization of tools for live imaging studies of the PcG protein mediated epigenetic memory mechanism is presented. The PcG protein E(Z) and the DNA binding factors DSP1 and PHO have been fluorescently tagged with GFP and transgenic lines expressing the GFP fusion proteins under various promoters have been established. The functionality of GFP fusion variants of PHO and E(Z) has been demonstrated and GFP::DSP1 has been shown to bind endogenous target loci. Transgenic expression levels and effects on endogenous protein levels have been quantitatively analyzed.
Additionally, transgenic flies have been established that carry a visually tagged transgenic PRE to track a single and specific PRE. In combination with the GFP tagged PcG fusion proteins this presents a device to qualitatively and quantitatively study PcG protein behavior at a single locus in respect to its function.

Live Imaging of GFP-tagged fusion proteins in this study and extended work has revealed that different proteins have different chromatic and nucleoplasmic distribution during mitosis and DSP1 and the TrxG ASH1 remain on mitotic chromosomes during early embryogenesis. Quantification of fluorescence signal intensities in blastoderm embryos enabled the calculation of molecule numbers of proteins involved in the PcG and TrxG mediated epigenetic memory mechanism.

7 Abbreviations

ash1, 2	absent, small or homeotic discs 1, 2
ATP	Adenosine triphosphate
B	Bar
bp	base pairs of DNA
BRM	Brahma
BSA	Bovine Serum Albumin
ChIP	Chromatin Immunoprecipitation
CHRASCH	Chromatin associated silencing complex for homeotics
ciD	cubitus interruptus dominant
CyO	Curl of Oister
D	distal
ddH2O	(double) distilled water
Da	Dalton
Df	deficiency
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleotide-triphosphate
Dr	Drop
dSfmbt	Drosophila Scm-related gene containing four mbt domains
E.Coli	Escherichia coli
EDTA	ethylenediaminetetraacetic acid
GAF	GAGA factor
gDNA	genomic DNA
(E)GFP	(Enhanced) Green Fluorescent Protein
EtOH	Ethanol
eyD	eyeless dominant
E(z)	Enhancer of zeste
FRAP	Fluorescence Recovery After Photobleaching
HxKy	Lysine y of Histone x, e.g. H3K27
HMTase	Histonemethyltransferase
HP1	heterochromatin protein 1
kb	Kilobase pairs of DNA
l	litre
LB	L(uria) B(ertani)
µ	micro
m	meter
M	Mega
MOR	Moira
N	nano
NaCl	Sodium chloride
ncRNA	non-coding RNA
p	proximal
P	P-element
PBS	Phospate Buffered Saline
Pc (-l)	Polycomb (-like)
PcG	Polycomb-Group
PCR	Polymerase Chain Reaction
PH	Polyhomeotic
PhoRC	PHO repressive complex
PRC	Polycomb Repressive Complex
PRE	Polycomb Responsive Element
psc	posterior sex combs
psq	pipsqueak
RNA	Ribonucleic Acid
RT	room temperature
rtPCR	real-time PCR
Sb	Stubble
Ser	Serate
SET	SU(VAR), E(Z) and TRX

SU(VAR)	Suppressor of variegation
SU(Z)12	Suppressor of zeste 12
SWI/SNF	SWItch/Sucrose NonFermentable
TAC	trothorax activating complex
TAE	Tris-Acetate-EDTA
Tb	Tubby
TBP	TATA binding protein
TE	Tris-EDTA
Tr	Transgene
TRE	Trithorax Responsive Element
TrxG	Trithorax-Group
VLPs	Virus like particles
W	white
wt	Wild-type
y	yellow

8 References

1. Ptashne, M.a.G., A., *Genes and Signals.* Cold Spring Harbor New York: Cold Spring Harbor Laboratory Press, 2002.
2. C.D.Allis, T.J., D.Reinberg *Epigenetics.* CSHL Press, 2007.
3. Ringrose, L. and R. Paro, *Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins.* Annu Rev Genet, 2004. **38**: p. 413-43.
4. Lewis, E.B., *A gene complex controlling segmentation in Drosophila.* Nature, 1978. **276**(5688): p. 565-70.
5. Maeda, R.K. and F. Karch, *The ABC of the BX-C: the bithorax complex explained.* Development, 2006. **133**(8): p. 1413-22.
6. Cavalli, G. and R. Paro, *The Drosophila Fab-7 chromosomal element conveys epigenetic inheritance during mitosis and meiosis.* Cell, 1998. **93**(4): p. 505-18.
7. Cavalli, G. and R. Paro, *Epigenetic inheritance of active chromatin after removal of the main transactivator.* Science, 1999. **286**(5441): p. 955-8.
8. Maurange, C. and R. Paro, *A cellular memory module conveys epigenetic inheritance of hedgehog expression during Drosophila wing imaginal disc development.* Genes Dev, 2002. **16**(20): p. 2672-83.
9. Jones, R.S. and W.M. Gelbart, *Genetic analysis of the enhancer of zeste locus and its role in gene regulation in Drosophila melanogaster.* Genetics, 1990. **126**(1): p. 185-99.
10. Decoville, M., et al., *DSP1, an HMG-like protein, is involved in the regulation of homeotic genes.* Genetics, 2001. **157**(1): p. 237-44.
11. Muller, J. and J.A. Kassis, *Polycomb response elements and targeting of Polycomb group proteins in Drosophila.* Curr Opin Genet Dev, 2006. **16**(5): p. 476-84.
12. Schuettengruber, B., et al., *Genome regulation by polycomb and trithorax proteins.* Cell, 2007. **128**(4): p. 735-45.
13. Klymenko, T., et al., *A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities.* Genes Dev, 2006. **20**(9): p. 1110-22.
14. Levine, S.S., et al., *The core of the polycomb repressive complex is compositionally and functionally conserved in flies and humans.* Mol Cell Biol, 2002. **22**(17): p. 6070-8.
15. Muller, J. and P. Verrijzer, *Biochemical mechanisms of gene regulation by polycomb group protein complexes.* Curr Opin Genet Dev, 2009. **19**(2): p. 150-8.
16. Shao, Z., et al., *Stabilization of chromatin structure by PRC1, a Polycomb complex.* Cell, 1999. **98**(1): p. 37-46.
17. Saurin, A.J., et al., *A Drosophila Polycomb group complex includes Zeste and dTAFII proteins.* Nature, 2001. **412**(6847): p. 655-60.
18. Francis, N.J., et al., *Reconstitution of a functional core polycomb repressive complex.* Mol Cell, 2001. **8**(3): p. 545-56.
19. Mulholland, N.M., I.F. King, and R.E. Kingston, *Regulation of Polycomb group complexes by the sequence-specific DNA binding proteins Zeste and GAGA.* Genes Dev, 2003. **17**(22): p. 2741-6.
20. Bannister, A.J., et al., *Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain.* Nature, 2001. **410**(6824): p. 120-4.

21. Fischle, W., et al., *Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains.* Genes Dev, 2003. **17**(15): p. 1870-81.
22. Gambetta, M.C., K. Oktaba, and J. Muller, *Essential role of the glycosyltransferase sxc/Ogt in polycomb repression.* Science, 2009. **325**(5936): p. 93-6.
23. Wang, H., et al., *Role of histone H2A ubiquitination in Polycomb silencing.* Nature, 2004. **431**(7010): p. 873-8.
24. Grimm, C., et al., *Structural and functional analyses of methyl-lysine binding by the malignant brain tumour repeat protein Sex comb on midleg.* EMBO Rep, 2007. **8**(11): p. 1031-7.
25. Grimm, C., et al., *Molecular recognition of histone lysine methylation by the Polycomb group repressor dSfmbt.* EMBO J, 2009. **28**(13): p. 1965-77.
26. Huang, D.H., et al., *pipsqueak encodes a factor essential for sequence-specific targeting of a polycomb group protein complex.* Mol Cell Biol, 2002. **22**(17): p. 6261-71.
27. Huang, D.H. and Y.L. Chang, *Isolation and characterization of CHRASCH, a polycomb-containing silencing complex.* Methods Enzymol, 2004. **377**: p. 267-82.
28. Muller, J., et al., *Histone methyltransferase activity of a Drosophila Polycomb group repressor complex.* Cell, 2002. **111**(2): p. 197-208.
29. Czermin, B., et al., *Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites.* Cell, 2002. **111**(2): p. 185-96.
30. Tie, F., et al., *The Drosophila Polycomb Group proteins ESC and E(Z) are present in a complex containing the histone-binding protein p55 and the histone deacetylase RPD3.* Development, 2001. **128**(2): p. 275-86.
31. Wu, C.T., et al., *Homeosis and the interaction of zeste and white in Drosophila.* Mol Gen Genet, 1989. **218**(3): p. 559-64.
32. Jones, R.S. and W.M. Gelbart, *The Drosophila Polycomb-group gene Enhancer of zeste contains a region with sequence similarity to trithorax.* Mol Cell Biol, 1993. **13**(10): p. 6357-66.
33. Jones, C.A., et al., *The Drosophila esc and E(z) proteins are direct partners in polycomb group-mediated repression.* Mol Cell Biol, 1998. **18**(5): p. 2825-34.
34. Tie, F., T. Furuyama, and P.J. Harte, *The Drosophila Polycomb Group proteins ESC and E(Z) bind directly to each other and co-localize at multiple chromosomal sites.* Development, 1998. **125**(17): p. 3483-96.
35. Brown, J.L., et al., *The Drosophila Polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1.* Mol Cell, 1998. **1**(7): p. 1057-64.
36. Oktaba, K., et al., *Dynamic regulation by polycomb group protein complexes controls pattern formation and the cell cycle in Drosophila.* Dev Cell, 2008. **15**(6): p. 877-89.
37. Mohd-Sarip, A., et al., *Pleiohomeotic can link polycomb to DNA and mediate transcriptional repression.* Mol Cell Biol, 2002. **22**(21): p. 7473-83.
38. Mohd-Sarip, A., et al., *Synergistic recognition of an epigenetic DNA element by Pleiohomeotic and a Polycomb core complex.* Genes Dev, 2005. **19**(15): p. 1755-60.
39. Wang, L., et al., *Hierarchical recruitment of polycomb group silencing complexes.* Mol Cell, 2004. **14**(5): p. 637-46.
40. Smith, S.T., et al., *Modulation of heat shock gene expression by the TAC1 chromatin-modifying complex.* Nat Cell Biol, 2004. **6**(2): p. 162-7.

41. Kennison, J.A. and J.W. Tamkun, *Dosage-dependent modifiers of polycomb and antennapedia mutations in Drosophila.* Proc Natl Acad Sci U S A, 1988. **85**(21): p. 8136-40.
42. Kennison, J.A., *The Polycomb and trithorax group proteins of Drosophila: trans-regulators of homeotic gene function.* Annu Rev Genet, 1995. **29**: p. 289-303.
43. Collins, R.T., et al., *Osa associates with the Brahma chromatin remodeling complex and promotes the activation of some target genes.* EMBO J, 1999. **18**(24): p. 7029-40.
44. Crosby, M.A., et al., *The trithorax group gene moira encodes a brahma-associated putative chromatin-remodeling factor in Drosophila melanogaster.* Mol Cell Biol, 1999. **19**(2): p. 1159-70.
45. Papoulas, O., et al., *The Drosophila trithorax group proteins BRM, ASH1 and ASH2 are subunits of distinct protein complexes.* Development, 1998. **125**(20): p. 3955-66.
46. Mohrmann, L., et al., *Differential targeting of two distinct SWI/SNF-related Drosophila chromatin-remodeling complexes.* Mol Cell Biol, 2004. **24**(8): p. 3077-88.
47. Daubresse, G., et al., *The Drosophila kismet gene is related to chromatin-remodeling factors and is required for both segmentation and segment identity.* Development, 1999. **126**(6): p. 1175-87.
48. Srinivasan, S., et al., *The Drosophila trithorax group protein Kismet facilitates an early step in transcriptional elongation by RNA Polymerase II.* Development, 2005. **132**(7): p. 1623-35.
49. Janody, F., et al., *Two subunits of the Drosophila mediator complex act together to control cell affinity.* Development, 2003. **130**(16): p. 3691-701.
50. Denis, G.V. and M.R. Green, *A novel, mitogen-activated nuclear kinase is related to a Drosophila developmental regulator.* Genes Dev, 1996. **10**(3): p. 261-71.
51. Gutierrez, L., et al., *The Drosophila trithorax group gene tonalli (tna) interacts genetically with the Brahma remodeling complex and encodes an SP-RING finger protein.* Development, 2003. **130**(2): p. 343-54.
52. Machado, C. and D.J. Andrew, *D-Titin: a giant protein with dual roles in chromosomes and muscles.* J Cell Biol, 2000. **151**(3): p. 639-52.
53. Beisel, C., et al., *Histone methylation by the Drosophila epigenetic transcriptional regulator Ash1.* Nature, 2002. **419**(6909): p. 857-62.
54. Lachner, M., R.J. O'Sullivan, and T. Jenuwein, *An epigenetic road map for histone lysine methylation.* J Cell Sci, 2003. **116**(Pt 11): p. 2117-24.
55. Sanchez-Elsner, T., et al., *Noncoding RNAs of trithorax response elements recruit Drosophila Ash1 to Ultrabithorax.* Science, 2006. **311**(5764): p. 1118-23.
56. Petruk, S., et al., *Trithorax and dCBP acting in a complex to maintain expression of a homeotic gene.* Science, 2001. **294**(5545): p. 1331-4.
57. Hughes, C.M., et al., *Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus.* Mol Cell, 2004. **13**(4): p. 587-97.
58. Yokoyama, A., et al., *Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression.* Mol Cell Biol, 2004. **24**(13): p. 5639-49.
59. Hodgson, J.W., B. Argiropoulos, and H.W. Brock, *Site-specific recognition of a 70-base-pair element containing d(GA)(n) repeats mediates bithoraxoid polycomb group response element-dependent silencing.* Mol Cell Biol, 2001. **21**(14): p. 4528-43.

60. Schwendemann, A. and M. Lehmann, *Pipsqueak and GAGA factor act in concert as partners at homeotic and many other loci.* Proc Natl Acad Sci U S A, 2002. **99**(20): p. 12883-8.
61. Brown, J.L., et al., *The Drosophila pho-like gene encodes a YY1-related DNA binding protein that is redundant with pleiohomeotic in homeotic gene silencing.* Development, 2003. **130**(2): p. 285-94.
62. Busturia, A., et al., *The MCP silencer of the Drosophila Abd-B gene requires both Pleiohomeotic and GAGA factor for the maintenance of repression.* Development, 2001. **128**(11): p. 2163-73.
63. Mahmoudi, T., et al., *GAGA facilitates binding of Pleiohomeotic to a chromatinized Polycomb response element.* Nucleic Acids Res, 2003. **31**(14): p. 4147-56.
64. Dejardin, J., et al., *Recruitment of Drosophila Polycomb group proteins to chromatin by DSP1.* Nature, 2005. **434**(7032): p. 533-8.
65. Brown, J.L., et al., *An Sp1/KLF binding site is important for the activity of a Polycomb group response element from the Drosophila engrailed gene.* Nucleic Acids Res, 2005. **33**(16): p. 5181-9.
66. Blastyak, A., et al., *Efficient and specific targeting of Polycomb group proteins requires cooperative interaction between Grainyhead and Pleiohomeotic.* Mol Cell Biol, 2006. **26**(4): p. 1434-44.
67. Bantignies, F. and G. Cavalli, *Cellular memory and dynamic regulation of polycomb group proteins.* Curr Opin Cell Biol, 2006. **18**(3): p. 275-83.
68. Beuchle, D., G. Struhl, and J. Muller, *Polycomb group proteins and heritable silencing of Drosophila Hox genes.* Development, 2001. **128**(6): p. 993-1004.
69. Simon, J., et al., *Elements of the Drosophila bithorax complex that mediate repression by Polycomb group products.* Dev Biol, 1993. **158**(1): p. 131-44.
70. Chan CS, R.L., Pirrotta V., *A Polycomb response element in the Ubx gene that determines an epigenetically inherited state of repression.* EMBO J, 1994. **13:2553-64**.
71. Ringrose, L., et al., *Genome-wide prediction of Polycomb/Trithorax response elements in Drosophila melanogaster.* Dev Cell, 2003. **5**(5): p. 759-71.
72. Tolhuis, B., et al., *Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in Drosophila melanogaster.* Nat Genet, 2006. **38**(6): p. 694-9.
73. Negre, N., et al., *Chromosomal distribution of PcG proteins during Drosophila development.* PLoS Biol, 2006. **4**(6): p. e170.
74. Schwartz, Y.B., et al., *Genome-wide analysis of Polycomb targets in Drosophila melanogaster.* Nat Genet, 2006. **38**(6): p. 700-5.
75. Schuettengruber, B., et al., *Functional anatomy of polycomb and trithorax chromatin landscapes in Drosophila embryos.* PLoS Biol, 2009. **7**(1): p. e13.
76. Katsani, K.R., M.A. Hajibagheri, and C.P. Verrijzer, *Co-operative DNA binding by GAGA transcription factor requires the conserved BTB/POZ domain and reorganizes promoter topology.* EMBO J, 1999. **18**(3): p. 698-708.
77. Lanzuolo, C., et al., *Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex.* Nat Cell Biol, 2007. **9**(10): p. 1167-74.
78. Pirrotta, V., et al., *Distinct parasegmental and imaginal enhancers and the establishment of the expression pattern of the Ubx gene.* Genetics, 1995. **141**(4): p. 1439-50.
79. Mihaly, J., et al., *In situ dissection of the Fab-7 region of the bithorax complex into a chromatin domain boundary and a Polycomb-response element.* Development, 1997. **124**(9): p. 1809-20.

80. Barges, S., et al., *The Fab-8 boundary defines the distal limit of the bithorax complex iab-7 domain and insulates iab-7 from initiation elements and a PRE in the adjacent iab-8 domain.* Development, 2000. **127**(4): p. 779-90.
81. Milne, T.A., et al., *MLL targets SET domain methyltransferase activity to Hox gene promoters.* Mol Cell, 2002. **10**(5): p. 1107-17.
82. Jacobs, J.J., et al., *The oncogene and Polycomb-group gene bmi-1 regulates cell proliferation and senescence through the ink4a locus.* Nature, 1999. **397**(6715): p. 164-8.
83. Boyer, L.A., et al., *Polycomb complexes repress developmental regulators in murine embryonic stem cells.* Nature, 2006. **441**(7091): p. 349-53.
84. Lee, T.I., et al., *Control of developmental regulators by Polycomb in human embryonic stem cells.* Cell, 2006. **125**(2): p. 301-13.
85. Sing, A., et al., *A vertebrate Polycomb response element governs segmentation of the posterior hindbrain.* Cell, 2009. **138**(5): p. 885-97.
86. Woo, C.J., et al., *A region of the human HOXD cluster that confers polycomb-group responsiveness.* Cell. **140**(1): p. 99-110.
87. Poux, S., D. McCabe, and V. Pirrotta, *Recruitment of components of Polycomb Group chromatin complexes in Drosophila.* Development, 2001. **128**(1): p. 75-85.
88. Papp, B. and J. Muller, *Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins.* Genes Dev, 2006. **20**(15): p. 2041-54.
89. Cao, R. and Y. Zhang, *The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3.* Curr Opin Genet Dev, 2004. **14**(2): p. 155-64.
90. Kuzmichev, A., et al., *Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3.* Mol Cell, 2004. **14**(2): p. 183-93.
91. Ringrose, L., *Polycomb comes of age: genome-wide profiling of target sites.* Curr Opin Cell Biol, 2007. **19**(3): p. 290-7.
92. Mutskov, V. and G. Felsenfeld, *Silencing of transgene transcription precedes methylation of promoter DNA and histone H3 lysine 9.* EMBO J, 2004. **23**(1): p. 138-49.
93. Christensen, J., et al., *RBP2 belongs to a family of demethylases, specific for tri- and dimethylated lysine 4 on histone 3.* Cell, 2007. **128**(6): p. 1063-76.
94. Smith, E.R., et al., *Drosophila UTX is a histone H3 Lys27 demethylase that colocalizes with the elongating form of RNA polymerase II.* Mol Cell Biol, 2008. **28**(3): p. 1041-6.
95. Eissenberg, J.C., et al., *The trithorax-group gene in Drosophila little imaginal discs encodes a trimethylated histone H3 Lys4 demethylase.* Nat Struct Mol Biol, 2007. **14**(4): p. 344-6.
96. Lee, N., et al., *The trithorax-group protein Lid is a histone H3 trimethyl-Lys4 demethylase.* Nat Struct Mol Biol, 2007. **14**(4): p. 341-3.
97. Fischle, W., Y. Wang, and C.D. Allis, *Binary switches and modification cassettes in histone biology and beyond.* Nature, 2003. **425**(6957): p. 475-9.
98. Fitzgerald, D.P. and W. Bender, *Polycomb group repression reduces DNA accessibility.* Mol Cell Biol, 2001. **21**(19): p. 6585-97.
99. Dellino, G.I., et al., *Polycomb silencing blocks transcription initiation.* Mol Cell, 2004. **13**(6): p. 887-93.
100. Dejardin, J. and G. Cavalli, *Chromatin inheritance upon Zeste-mediated Brahma recruitment at a minimal cellular memory module.* EMBO J, 2004. **23**(4): p. 857-68.
101. Schmitt, S., M. Prestel, and R. Paro, *Intergenic transcription through a polycomb group response element counteracts silencing.* Genes Dev, 2005. **19**(6): p. 697-708.

102. Klymenko, T. and J. Muller, *The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins.* EMBO Rep, 2004. **5**(4): p. 373-7.
103. Poux, S., et al., *The Drosophila trithorax protein is a coactivator required to prevent re-establishment of polycomb silencing.* Development, 2002. **129**(10): p. 2483-93.
104. Hansen, K.H., et al., *A model for transmission of the H3K27me3 epigenetic mark.* Nat Cell Biol, 2008. **10**(11): p. 1291-300.
105. Margueron, R., et al., *Role of the polycomb protein EED in the propagation of repressive histone marks.* Nature, 2009. **461**(7265): p. 762-7.
106. Li, G., G. Sudlow, and A.S. Belmont, *Interphase cell cycle dynamics of a late-replicating, heterochromatic homogeneously staining region: precise choreography of condensation/decondensation and nuclear positioning.* J Cell Biol, 1998. **140**(5): p. 975-89.
107. Buchenau, P., et al., *The distribution of polycomb-group proteins during cell division and development in Drosophila embryos: impact on models for silencing.* J Cell Biol, 1998. **141**(2): p. 469-81.
108. Miyagishima, H., et al., *Dissociation of mammalian Polycomb-group proteins, Ring1B and Rae28/Ph1, from the chromatin correlates with configuration changes of the chromatin in mitotic and meiotic prophase.* Histochem Cell Biol, 2003. **120**(2): p. 111-9.
109. Cheutin, T., et al., *Maintenance of stable heterochromatin domains by dynamic HP1 binding.* Science, 2003. **299**(5607): p. 721-5.
110. Festenstein, R., et al., *Modulation of heterochromatin protein 1 dynamics in primary Mammalian cells.* Science, 2003. **299**(5607): p. 719-21.
111. Chen, D., et al., *Condensed mitotic chromatin is accessible to transcription factors and chromatin structural proteins.* J Cell Biol, 2005. **168**(1): p. 41-54.
112. Dietzel, S., et al., *The nuclear distribution of Polycomb during Drosophila melanogaster development shown with a GFP fusion protein.* Chromosoma, 1999. **108**(2): p. 83-94.
113. Rastelli, L., C.S. Chan, and V. Pirrotta, *Related chromosome binding sites for zeste, suppressors of zeste and Polycomb group proteins in Drosophila and their dependence on Enhancer of zeste function.* EMBO J, 1993. **12**(4): p. 1513-22.
114. Strutt, H. and R. Paro, *The polycomb group protein complex of Drosophila melanogaster has different compositions at different target genes.* Mol Cell Biol, 1997. **17**(12): p. 6773-83.
115. Ficz, G., R. Heintzmann, and D.J. Arndt-Jovin, *Polycomb group protein complexes exchange rapidly in living Drosophila.* Development, 2005. **132**(17): p. 3963-76.
116. Ringrose, L., H. Ehret, and R. Paro, *Distinct contributions of histone H3 lysine 9 and 27 methylation to locus-specific stability of polycomb complexes.* Mol Cell, 2004. **16**(4): p. 641-53.
117. Ng, H.H. and A. Bird, *Histone deacetylases: silencers for hire.* Trends Biochem Sci, 2000. **25**(3): p. 121-6.
118. Zhang, H., et al., *SUMO modification is required for in vivo Hox gene regulation by the Caenorhabditis elegans Polycomb group protein SOP-2.* Nat Genet, 2004. **36**(5): p. 507-11.
119. Voncken, J.W., et al., *MAPKAP kinase 3pK phosphorylates and regulates chromatin association of the polycomb group protein Bmi1.* J Biol Chem, 2005. **280**(7): p. 5178-87.
120. Lippincott-Schwartz, J. and G.H. Patterson, *Development and use of fluorescent protein markers in living cells.* Science, 2003. **300**(5616): p. 87-91.

121. Tripoulas, N., et al., *The Drosophila ash1 gene product, which is localized at specific sites on polytene chromosomes, contains a SET domain and a PHD finger.* Genetics, 1996. **143**(2): p. 913-28.
122. Rozovskaia, T., et al., *Trithorax and ASH1 interact directly and associate with the trithorax group-responsive bxd region of the Ultrabithorax promoter.* Mol Cell Biol, 1999. **19**(9): p. 6441-7.
123. Robinett, C.C., et al., *In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition.* J Cell Biol, 1996. **135**(6 Pt 2): p. 1685-700.
124. Vazquez, J., A.S. Belmont, and J.W. Sedat, *Multiple regimes of constrained chromosome motion are regulated in the interphase Drosophila nucleus.* Curr Biol, 2001. **11**(16): p. 1227-39.
125. Reed, B.H., S.C. McMillan, and R. Chaudhary, *The preparation of Drosophila embryos for live-imaging using the hanging drop protocol.* J Vis Exp, 2009(25).
126. Protocols, S., *Transgenesis Techniques.* Humana Press, 2009.
127. Tsien, R.Y., *The green fluorescent protein.* Annu Rev Biochem, 1998. **67**: p. 509-44.
128. Cormack, B.P., R.H. Valdivia, and S. Falkow, *FACS-optimized mutants of the green fluorescent protein (GFP).* Gene, 1996. **173**(1 Spec No): p. 33-8.
129. Groth, A.C., et al., *Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31.* Genetics, 2004. **166**(4): p. 1775-82.
130. Mosrin-Huaman, C., et al., *DSP1 gene of Drosophila melanogaster encodes an HMG-domain protein that plays multiple roles in development.* Dev Genet, 1998. **23**(4): p. 324-34.
131. Fritsch, C., et al., *The DNA-binding polycomb group protein pleiohomeotic mediates silencing of a Drosophila homeotic gene.* Development, 1999. **126**(17): p. 3905-13.
132. Carrington, E.A. and R.S. Jones, *The Drosophila Enhancer of zeste gene encodes a chromosomal protein: examination of wild-type and mutant protein distribution.* Development, 1996. **122**(12): p. 4073-83.
133. Decoville, M., et al., *HMG boxes of DSP1 protein interact with the rel homology domain of transcription factors.* Nucleic Acids Res, 2000. **28**(2): p. 454-62.
134. Zink, B. and R. Paro, *In vivo binding pattern of a trans-regulator of homoeotic genes in Drosophila melanogaster.* Nature, 1989. **337**(6206): p. 468-71.
135. Fauvarque, M.O., V. Zuber, and J.M. Dura, *Regulation of polyhomeotic transcription may involve local changes in chromatin activity in Drosophila.* Mech Dev, 1995. **52**(2-3): p. 343-55.
136. Canaple, L., et al., *The Drosophila DSP1 gene encoding an HMG 1-like protein: genomic organization, evolutionary conservation and expression.* Gene, 1997. **184**(2): p. 285-90.
137. Weigmann, K., et al., *FlyMove--a new way to look at development of Drosophila.* Trends Genet, 2003. **19**(6): p. 310-1.
138. Foe, V.E. and B.M. Alberts, *Studies of nuclear and cytoplasmic behaviour during the five mitotic cycles that precede gastrulation in Drosophila embryogenesis.* J Cell Sci, 1983. **61**: p. 31-70.
139. Janody, F., et al., *A mosaic genetic screen reveals distinct roles for trithorax and polycomb group genes in Drosophila eye development.* Genetics, 2004. **166**(1): p. 187-200.
140. Girton, J.R. and S.H. Jeon, *Novel embryonic and adult homeotic phenotypes are produced by pleiohomeotic mutations in Drosophila.* Dev Biol, 1994. **161**(2): p. 393-407.

141. Breen, T.R. and I.M. Duncan, *Maternal expression of genes that regulate the bithorax complex of Drosophila melanogaster.* Dev Biol, 1986. **118**(2): p. 442-56.
142. Martinez-Balbas, M.A., et al., *Displacement of sequence-specific transcription factors from mitotic chromatin.* Cell, 1995. **83**(1): p. 29-38.
143. Charpilienne, A., et al., *Individual rotavirus-like particles containing 120 molecules of fluorescent protein are visible in living cells.* J Biol Chem, 2001. **276**(31): p. 29361-7.
144. Dundr, M., et al., *Quantitation of GFP-fusion proteins in single living cells.* J Struct Biol, 2002. **140**(1-3): p. 92-9.
145. Fischer, J.A., et al., *GAL4 activates transcription in Drosophila.* Nature, 1988. **332**(6167): p. 853-6.
146. Pirrotta, V., *Vectors for P-mediated transformation in Drosophila.* Biotechnology, 1988. **10**: p. 437-56.
147. Ringrose, L., et al., *Comparative kinetic analysis of FLP and cre recombinases: mathematical models for DNA binding and recombination.* J Mol Biol, 1998. **284**(2): p. 363-84.
148. Ringrose, L., et al., *Quantitative comparison of DNA looping in vitro and in vivo: chromatin increases effective DNA flexibility at short distances.* EMBO J, 1999. **18**(23): p. 6630-41.
149. Zink, D. and R. Paro, *Drosophila Polycomb-group regulated chromatin inhibits the accessibility of a trans-activator to its target DNA.* EMBO J, 1995. **14**(22): p. 5660-71.
150. Sedkov, Y., et al., *The bithorax complex is regulated by trithorax earlier during Drosophila embryogenesis than is the Antennapedia complex, correlating with a bithorax-like expression pattern of distinct early trithorax transcripts.* Development, 1994. **120**(7): p. 1907-17.
151. Petruk, S., et al., *Association of trxG and PcG proteins with the bxd maintenance element depends on transcriptional activity.* Development, 2008. **135**(14): p. 2383-90.
152. O'Donnell, K.H., C.T. Chen, and P.C. Wensink, *Insulating DNA directs ubiquitous transcription of the Drosophila melanogaster alpha 1-tubulin gene.* Mol Cell Biol, 1994. **14**(9): p. 6398-408.
153. Rubin, G.M. and A.C. Spradling, *Genetic transformation of Drosophila with transposable element vectors.* Science, 1982. **218**(4570): p. 348-53.
154. Venken, K.J. and H.J. Bellen, *Transgenesis upgrades for Drosophila melanogaster.* Development, 2007. **134**(20): p. 3571-84.
155. Venken, K.J. and H.J. Bellen, *Emerging technologies for gene manipulation in Drosophila melanogaster.* Nat Rev Genet, 2005. **6**(3): p. 167-78.
156. Salvaing, J., et al., *Corto and DSP1 interact and bind to a maintenance element of the Scr Hox gene: understanding the role of Enhancers of trithorax and Polycomb.* BMC Biol, 2006. **4**: p. 9.
157. Chen, X., et al., *Tissue-specific TAFs counteract Polycomb to turn on terminal differentiation.* Science, 2005. **310**(5749): p. 869-72.
158. Ringrose, L., *Polycomb, trithorax and the decision to differentiate.* Bioessays, 2006. **28**(4): p. 330-4.
159. Lehming, N., et al., *Chromatin components as part of a putative transcriptional repressing complex.* Proc Natl Acad Sci U S A, 1998. **95**(13): p. 7322-6.
160. Bianchi, M.E. and A. Agresti, *HMG proteins: dynamic players in gene regulation and differentiation.* Curr Opin Genet Dev, 2005. **15**(5): p. 496-506.
161. Mayer, B., et al., *Quantitative analysis of protein dynamics during asymmetric cell division.* Curr Biol, 2005. **15**(20): p. 1847-54.

162. Horn, C., B. Jaunich, and E.A. Wimmer, *Highly sensitive, fluorescent transformation marker for Drosophila transgenesis.* Dev Genes Evol, 2000. **210**(12): p. 623-9.

Curriculum vitae

Cornelia Gänger

Vienna, Austria
cornelia.gaenger@imba.oeaw.ac.at
cornelia.gaenger@gmx.at

Date of Birth: 29.06.1982
Nationality: Austrian

Education and Qualifications

Since 9/2009
Wiener Goldschmiedeakademie

4/2006 – 3/2010
Vienna BioCenter-PhD program /IMBA, Vienna University
Winter selection 2005/2006
Project management workshop
Presentation skills course

10/2000 - 11/2005
University of Vienna
Undergraduate studies in Microbiology and Genetics
Major subject: Immunology
Minor subjects: Structural Biology, Developmental Biology

9/1992 – 6/2000
Grammar School with the main emphasis on Natural Sciences
Tulln, Lower Austria
Excellent Certificate of final exams

9/1988 – 6/1992
Elementary School
Gablitz, Lower Austria

Work Experience

4/2006 – 3/2010
IMBA-Institute of Molecular Biotechnology of the Austrian Academy of Sciences, Vienna
PhD student in the Epigenetics group of Dr. Leonie Ringrose
Lab techniques in molecular biology and cloning, working with Drosophila melanogaster, fluorescence microscopy and in vivo imaging
Supervision of a diploma student

10/2004 – 10/2005
IMP-Institute of Molecular Pathology, Vienna

Diploma student in the structural biology group of Dr. Tim Clausen
Thesis: "Structure-Function Analysis on the ClpC Chaperone and Its Adapter Protein MecA"
Lab techniques in cloning, protein purification and biochemistry, protein crystallography, data evaluation and presentation

Since 09/2000
Various jobs, to mention a few
Komparserie der Wiener Staatsoper
Seefestspiele Mörbisch
Vienna Life Ball (volunteer)
Business Coaching Partners, Regus, Bosch (office)
Private childcare and coaching lessons for school children

Skills

German (mother tongue), English (fluent), French (basic)
Microsoft Office,
Führerschein B