Biochemical Purification and Functional Characterization of a Novel Trithorax-Group Protein Complex

Sigrun Schmähling

Dissertation
Ludwig-Maximilians-Universität München
2019

Biochemical Purification and Functional Characterization of a Novel Trithorax-Group Protein Complex

Dissertation

submitted for the degree of

Doctor of Natural Sciences (Dr. rer. nat.)

to the Faculty of Biology

of the Ludwig-Maximilians-Universität München

by

Sigrun Schmähling

born in Bingen, Germany

This work was carried out under the supervision of Dr. Jürg Müller at the Max Planck Institute of Biochemistry in Martinsried, Germany.

First reviewer:	Prof. Dr. Barbara Conradt
Second reviewer:	Prof. Dr. Peter Becker
Date of submission:	January 29, 2019
Date of oral examination:	June 27, 2019

Statutory Declaration / Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass meine Dissertation selbstständig und ohne unerlaubte Hilfsmittel angefertigt worden ist.
Die vorliegende Dissertation wurde weder ganz, noch teilweise einer anderen Prüfungskommission vorgelegt.
Ich habe zu keinem früheren Zeitpunkt versucht, eine Dissertation einzureichen oder an einer Doktorprüfung teilzunehmen.

München, den 08.07.2019

Sigrun Schmähling

Parts of this work have been published in the following article:

Schmähling, S., Meiler, A., Lee, Y., Mohammed, A., Finkl, K., Tauscher, K., Israel, L., Wirth, M., Philippou-Massier, J., Blum, H., Habermann, B., Imhof, A., Song, J.-J. and Müller, J. (2018). Regulation and function of H3K36 di-methylation by the trithorax-group protein complex AMC. *Development* **145**, dev163808.
https://dev.biologists.org/content/145/7/dev163808

The article is included in the appendix of this dissertation.

Abstract

The trithorax group (trxG) of genes encodes diverse transcriptional regulators that play a pivotal role in development of animals and plants. The characteristic property of trxG proteins is to maintain active transcription of their target genes by antagonizing the action of specific transcriptional repressors, the Polycomb group (PcG), at these genes. Together, PcG and trxG proteins form a system of counterbalancing transcriptional regulators that controls the correct spatial and temporal expression of developmental regulator genes in the body plan of multicellular organisms. The canonical target genes of the PcG/trxG system in animals are *HOX* genes. *HOX* genes encode another group of transcription factors that determine and preserve the identity of body structures in their specific expression domains. Loss of trxG function at an actively transcribed *HOX* gene results in the silencing of this gene by PcG proteins. The phenotypic consequences are morphological abnormalities in the body plan known as the classical homeotic phenotypes. In this thesis, I studied the trxG protein Absent, small, or homeotic discs 1 (Ash1). Genetic studies originally identified Ash1 as a trxG member in *Drosophila* based on its homeotic mutant phenotype. At the molecular level, Ash1 functions as a histone methyltransferase (HMTase) with specificity for di-methylation of lysine 36 in histone H3 (H3K36me2). The deposition of H3K36me2 is generally believed to be the means of Ash1 to oppose PcG action at target genes. However, the Ash1 protein alone has only poor HMTase activity. Structural studies have shown that the Ash1 catalytic domain is actually auto-inhibited. This led to the suggestion that Ash1 may be controlled by an unknown activation mechanism, possibly involving hitherto undiscovered interactors of Ash1.

As first step in this study, I purified and characterized Ash1 complexes from *Drosophila* to investigate the regulation and function of Ash1 HMTase activity. The purifications of Ash1 revealed that the protein is tightly associated with two other proteins: the *Drosophila* orthologue of human MORF4-related gene on chromosome 15 (MRG15) and the 55 kDa subunit of Chromatin assembly factor 1 (Caf1-55). Reconstitution of this assembly with recombinant proteins confirmed that Ash1, MRG15 and Caf1-55 together form a stable complex that I named AMC. Further analysis of the interactions in AMC showed that Ash1 binds MRG15 via a conserved FxLP-motif. Intriguingly, the interaction with MRG15 proved to strongly stimulate the Ash1 catalytic activity for H3K36 di-methylation on recombinant nucleosomes *in vitro*.

In order to investigate the MRG15 function *in vivo*, I generated an *MRG15* deletion allele in *Drosophila*. To learn about Ash1 HMTase function in transcription, an *ash1* mutant carrying a point mutation in the catalytic Ash1 domain that disrupts its HMTase activity was made. Genetic

analyses of both mutants revealed that the *MRG15* null and the *ash1* catalytic phenotypes are very similar and highly specific. In particular, *MRG15* null and *ash1* catalytic mutants complete embryonic and larval development and undergo metamorphosis, but exhibit severe homeotic transformations caused by the loss of expression of multiple *HOX* genes in the adult stage. Remarkably, neither of the two mutants showed any other obvious morphological aberrations. These results strongly suggest that MRG15 is also important for AMC enzymatic activity *in vivo*. Moreover, the clearly homeotic mutant phenotype identifies *MRG15* as a novel trxG gene.

Apart from AMC, *Drosophila* possesses two other H3K36 HMTases, Nuclear receptor-binding SET domain protein (NSD) and SET domain-containing protein 2 (SET2). I assessed the contribution of AMC to the global amount of cellular H3K36me2 and found that, consistent with the specific phenotypes of *ash1* and *MRG15* mutants, H3K36me2 bulk levels were not detectably reduced in these mutants. However, applying chromatin immunoprecipitation assays, I detected a strong reduction in H3K36me2 levels at the AMC target gene *Ultrabithorax (Ubx)* in mutants lacking Ash1.

The findings reported in this work advance the understanding of the Ash1 HMTase function on a biochemical as well as on a genetic level. Ash1 was purified in a novel complex, AMC, that was shown to classify as a trxG complex. The discovery that the AMC subunit MRG15 is critically required for the Ash1 HMTase activity resolves how Ash1 becomes catalytically stimulated. The observation of loss of *HOX* gene expression in the *ash1* catalytic mutant is consistent with the decrease in H3K36me2 levels at *Ubx* in the absence of Ash1. Both analyses together reveal H3K36 di-methylation of *HOX* gene chromatin as the mechanism by which AMC maintains active *HOX* gene expression.

Zusammenfassung

Proteine der Trithorax-Gruppe (trxG) sind Regulatoren der Gentranskription mit entscheidenden Funktionen in der Entwicklung von Tieren und Pflanzen. Ihre charakteristische Eigenschaft ist es die Transkription von Zielgenen aufrecht zu erhalten, indem sie den Aktivitäten der transkriptionsreprimierenden Proteine der Polycomb-Gruppe (PcG) entgegensteuern. Zusammen bilden PcG- und trxG-Proteine ein ausgewogenes System antagonistisch wirkender Transkriptionsregulatoren, das für die zeitlich und örtlich korrekte Expression entwicklungsregulatorischer Gene im Körperbauplan vielzelliger Organismen verantwortlich ist. Die kanonischen Zielgene des PcG/trxG-Systems in Tieren sind *HOX*-Gene. *HOX*-Gene codieren wiederum für Transkriptionsfaktoren, die innerhalb ihrer Expressionsdomänen die Identität von Körperstrukturen determinieren und erhalten. Der Verlust der Funktion von trxG-Proteinen an einem aktiv transkribierten *HOX*-Gen führt zur „Übernahme“ durch PcG-Proteine, die das betreffende Gen „stilllegen“. Morphologische Anomalien im Körperbauplan, die klassischen homöotischen Phänotypen, sind die Folge. In meiner Doktorarbeit habe ich das trxG-Protein „Absent, small, or homeotic discs 1“ (Ash1) erforscht, das ursprünglich in genetischen Studien anhand seines homöotischen Mutationsphänotyps in *Drosophila* identifiziert wurde. Auf der molekularen Ebene agiert Ash1 als Histon-Methyltransferase (HMTase), die spezifisch das Lysin 36 des Histons H3 dimethyliert (H3K36me2). Es wird vermutet, dass Ash1 PcG-Proteinen entgegenwirkt, indem es das Chromatin an gemeinsamen Zielgenen mit H3K36me2 modifiziert. Isoliertes Ash1-Protein weist alleine allerdings nur schwache HMTase-Aktivität auf. Strukturstudien haben gezeigt, dass der Zugang zur katalytischen Tasche in Ash1 durch ein autoinhibitorisches Element blockiert wird. Dies lässt vermuten, dass die enzymatische Aktivität von Ash1 über einen unbekannten Mechanismus stimuliert wird, an dem bisher nicht beschriebene Ash1-Bindungspartner beteiligt sein könnten.

Im ersten Schritt meiner Doktorarbeit reinigte ich Ash1-Komplexe aus *Drosophila* auf und charakterisierte diese, um die Regulation und Funktion der HMTase-Aktivität von Ash1 zu untersuchen. In den Aufreinigungen wurden mit Ash1 zwei weitere Proteine stark angereichert, die als spezifische Bindungspartner in Frage kamen: das *Drosophila*-Ortholog des humanen „MORF4-related gene on chromosome 15“ (MRG15) und das „55 kDa subunit of Chromatin assembly factor 1“ (Caf1-55). Rekonstitutionsexperimente mit rekombinanten Proteinen zeigten, dass Ash1, MRG15 und Caf1-55 in der Tat zusammen einen stabilen Komplex formen, der AMC genannt wurde. Wie weitere Analysen der Interaktionen in AMC aufdeckten, bindet Ash1 MRG15

über ein konserviertes FxLP-Motiv. Bemerkenswerterweise stimulierte diese Wechselwirkung *in vitro* die katalytische Aktivität von Ash1 H3K36 zu di-methylieren.

Mit dem Ziel die Funktion von MRG15 auch *in vivo* zu erforschen, generierte ich ein Deletionsallel von *MRG15* in *Drosophila* zur Charakterisierung von *MRG15*-Nullmutanten. Um die Funktion der HMTase-Aktivität von Ash1 in der Transkription zu untersuchen, wurde zudem eine *ash1*-Mutante mit einer Punktmutation in der katalytischen Ash1-Domäne, die zu Inaktivität führt, hergestellt (*ash1*-katalytische Mutante). In genetischen Analysen beider Mutanten zeigte sich, dass der *MRG15*-Null- dem *ash1*-katalytischen Phänotyp stark ähnelt und beide sehr spezifisch sind. Sowohl die *MRG15*-Null- als auch die *ash1*-katalytischen Mutanten beenden die Larvenstadien und durchlaufen die Metamorphose, weisen aber im Adultstadium starke homöotische Transformationen auf, die auf den Expressionsverlust mehrerer *HOX*-Gene zurückgehen. Bemerkenswerterweise wurden in keiner der beiden Mutanten weitere morphologische Anomalien neben den homöotischen Transformationen gefunden. Zusammengenommen legen die Ergebnisse der genetischen Analysen nahe, dass MRG15 auch *in vivo* eine wichtige Rolle für die enzymatische Aktivität von AMC spielt. Darüber hinaus identifiziert sein eindeutig homöotischer Mutationsphänotyp *MRG15* als neues trxG-Gen.

Neben AMC verfügt *Drosophila* über zwei weitere HMTasen, die spezifisch H3K36 methylieren: das „Nuclear receptor-binding SET domain protein“ (NSD) und das „SET domain-containing protein 2“ (SET2). Ich analysierte den Beitrag von AMC zur Gesamtheit an zellulärem H3K36me2 und stellte dabei fest, dass die H3K36me2-Gesamtmenge in *ash1*-Null- und *MRG15*-Nullmutanten in Übereinstimmung mit ihren spezifischen Phänotypen nicht im detektierbaren Rahmen reduziert ist. Am AMC-Zielgen *Ultrabithorax (Ubx)* dagegen zeigten Chromatin-Immunpräzipitations-Experimente eine starke Reduktion der H3K36me2-Level in *ash1*-Nullmutanten.

Die Ergebnisse dieser Doktorarbeit bringen das Verständnis der Funktion der Ash1-HMTase auf der biochemischen wie auch auf der genetischen Ebene wesentlich voran. Ash1 wurde in einem zuvor unbekannten Protein-Komplex, AMC, aufgereinigt, von dem die genetischen Analysen seiner Untereinheiten zeigten, dass er als trxG-Komplex eingestuft werden kann. Die Entdeckung, dass die AMC-Untereinheit MRG15 für die HMTase-Aktivität von Ash1 entscheidend ist, erklärt, wie Ash1 stimuliert wird. Die Ergebnisse, dass *ash1*-katalytische Mutanten einen starken Verlust an *HOX*-Gen-Expression aufweisen und dass die H3K36me2-Level am *Ubx*-Gen in Abwesenheit von Ash1 wesentlich reduziert sind, bestätigen einander. Beide Analysen zusammengenommen zeigen, dass die Dimethylierung von H3K36 im Chromatin von *HOX*-Genen der Mechanismus ist, über den der AMC-Komplex PcG-Proteinen entgegenwirkt und die Expression von *HOX*-Genen aufrechterhält.

Acknowledgments

First and foremost, I would like to thank Jürg for giving me the opportunity to carry out my PhD in his laboratory and to work on a very exciting and multifaceted project by which I could really grow and learn a lot. Many tanks to you, Jürg, also for supervision and invaluable scientific input, especially towards the end!

The atmosphere in the lab was and is great and the enthusiasm for science contagious. Thank you all for contributing to this and for sharing your little tips and tricks in the experimental work!

Maria Gambetta introduced me to the 'fly world'. Maria, thank you for never running out of patience when discussing my countless questions in genetics and beyond!

Reinhard Kalb was always ready to share his extensive knowledge in biochemistry with me. Reinhard, I am truly grateful to you for having been unconditionally helpful and supportive at innumerous occasions!

The project would not have gone so far without Katja Finkl, who is supporting us in the *Drosophila* work, and Katharina Tauscher, who is running the insect cell culture in the lab. Thank you both loads for all your help!

To Aynur Kaya Copur and Jacques Bonnet I would like to say that their help to reach the finish line is very much appreciated!

I would also like to thank my collaborators Prof. Axel Imhof and the people in his team, Dr. Lars Israel and Marc Wirth, who performed the mass spectrometric analysis for my project! Axel Imhof helped a lot by discussing the data over and over and was also a member of my thesis advisory committee (TAC). Many thanks as well to the other members of my TAC, namely Dr. Jürg Müller, Prof. Peter Becker and Prof. Barbara Conradt, who was also my official supervisor at university! To Prof. Barbara Conradt, Prof. Peter Becker, Dr. Bettina Bölter, Prof. John Parsch, Prof. Marc Bramkamp and Prof. Nicolas Gompel I am grateful that they agreed to be part of my thesis defense committee and to read and correct this thesis.

Table of Contents

1 Introduction

1.1 Polycomb-group and trithorax-group transcriptional regulators in development

Metazoans and the majority of plants develop through cell divisions and cell differentiations from a totipotent unicellular zygote into multicellular organisms that are composed of a broad spectrum of various cell types. Certain cell types together assemble to higher-order structures, such as organs or body segments. The latter are perceived as serially repeated units with related characteristics along the anterior-posterior body axis of bilaterian animals. Each structure in the body plan of multicellular organisms is dedicated to a specific function that defines its own identity. How these identities are determined at the onset of development and how they are maintained during the entire life span are the major questions in developmental biology. Intriguingly, the different cell types, that represent the basal building blocks of the body plan, all share the same genotype. By now, we know that it is the specific set of genes expressed out of all genes of a given genotype that makes the difference and that is characteristic for each biological structure, e.g. cell types, organs or segments. But what are the regulatory mechanisms behind and what are their means to ensure that in each structure only the corresponding identity-forming gene-set is expressed? Answers to the latter point emerged with more and more detailed understanding of epigenetic mechanisms as means to establish heritable transcriptional states beyond the DNA sequence.

The focus of the work presented here lies on the protein Absent, small, or homeotic discs 1 (Ash1). Ash1 is a member of the Polycomb group/trithorax group (PcG/trxG) system that is constituted of transcriptional regulators responsible for the maintenance of the identity of body structures in animal and plant development. PcG and trxG proteins control by epigenetic mechanisms which identity-defining genes are expressed in a given structure and which are not. In this manner, they regulate a large variety of developmental processes, such as vegetative growth, flower formation and flowering time in plants as well as cell proliferation and stem cell maintenance in mammals and preservation of segment identity in bilaterian animals (reviewed in Pu & Sung, 2015; Grossniklaus & Paro, 2014). Segment identity is maintained by PcG and trxG proteins through transcriptional regulation of the *HOX* gene family. *HOX* genes are the canonical targets of the PcG/trxG system in animals and shall be introduced in detail in the following.

1.1.1 Characteristics of *HOX* genes, the classical target genes of the PcG/trxG system

Originally, *HOX* genes were identified in *Drosophila melanogaster* and have been studied most extensively in this model organism. They will therefore be described here using mainly *Drosophila** as an example. The *Drosophila* body is divided into segments from the anterior to the posterior end of the body axis. Along the anterior-posterior body axis, *HOX* genes are expressed in restricted specific domains, that are nearly superimposable with the body segments, under the regulation of the PcG/trxG system and other transcription factors (described below in para 1.1.2). Figure 1 shows the *HOX* gene expression domains in a post-gastrulation *Drosophila* embryo. In their domains, *HOX* genes function as transcriptional regulators themselves that control the transcription of genes which ultimately form the identity of the corresponding segment. For correct morphogenesis, segment identities need to be preserved from the embryo up to the adult stage by maintenance of the pattern of *HOX* gene expression domains throughout development (compare embryo and adult fly in Figure 1).

Besides their segment-specific expression and function, members of the *HOX* gene family have one structural characteristic in common, the presence of a homeobox in their sequence. The homeobox is a conserved element encoding for the DNA-interacting homeodomain. Via the homeodomain, HOX proteins bind to *cis*-regulatory adenine-thymine rich elements in their target genes (Beachy *et al*, 1988; Samson *et al*, 1989; Noyes *et al*, 2008; Berger *et al*, 2008).

Within the *Drosophila* genome, *HOX* genes are organized in two homeotic clusters, the Antennapedia complex (ANT-C) and the bithorax complex (BX-C) (Lewis, 1978; Kaufman *et al*, 1980; Lewis *et al*, 1980). The ANT-C comprises the genes *labial* (*lab*), *proboscipedia* (*pb*), *Deformed* (*Dfd*), *Sex combs reduced* (*Scr*) and *Antennapedia* (*Antp*); the BX-C contains the genes *Ultrabithorax* (*Ubx*), *abdominal A* (*abd-A*) and *Abdominal B* (*Abd-B*). Interestingly, the order of the *HOX* genes on the chromosome, with the exception of *pb*, corresponds to the order of the *HOX* gene expression domains along the anterior-posterior body axis, as sketched in Figure 1 and as described by Harding *et al*, 1985 and Akam, 1987. This phenomenon is known as spatial collinearity.

Sequences and the clustered organization of *HOX* genes are highly conserved from insects to vertebrates. Vertebrates possess several copies of one homeotic cluster, that contains sub-clusters orthologous to both *Drosophila* clusters, ANT-C as well as BX-C. The copies of the vertebrate cluster have been suggested to originate from gene duplication of an ancestral HOX cluster common with insects (Graham *et al*, 1989; Duboule & Dollé, 1989). In the genome of most

* The terms "*Drosophila*" or "fly" refer to *Drosophila melanogaster* throughout this work.

vertebrates, including mouse and human, four copies of the homeotic cluster are present, HOXA, HOXB, HOXC and HOXD.

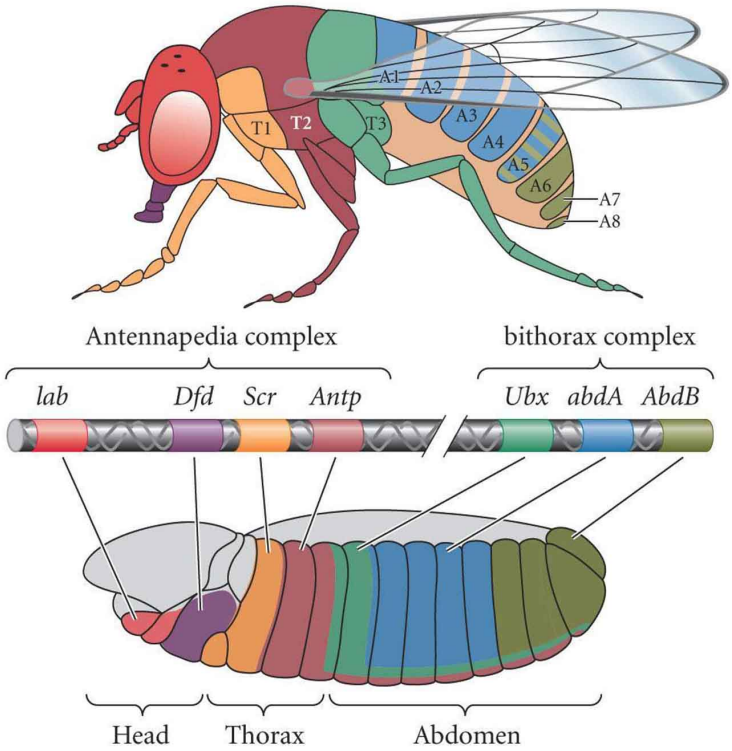


Figure 1. *HOX* gene expression domains and body segments in *Drosophila*. *HOX* genes preserve segment identity in bilaterian animals. For correct morphogenesis, the characteristic pattern of *HOX* gene expression domains along the body axis needs to be maintained throughout development, from the embryonic (bottom) up to the adult stage (top). In the *Drosophila* genome, *HOX* genes are organized collinear to their expression domains in two clusters: the Antennapedia complex and the bithorax complex (middle). T1 to T3: thoracic segments 1 to 3; A1 to A8: abdominal segments 1 to 8. The illustration has been taken from Gilbert, 2006.

1.1.2 Regulation of *HOX* gene expression in development

According to current knowledge, three transcription regulatory systems are engaged in the definition of the precisely bounded *HOX* gene expression domains: Besides the PcG/trxG system, these are internal cross-regulatory mechanisms among *HOX* genes and regulation by segmentation genes.

Segmentation genes initiate segmentation by setting up body segments and their polarity in the first stages of *Drosophila* embryogenesis. Subsequently, during cellular blastoderm formation, combinations of sub-groups of the segmentation genes, the gap and the pair-rule

genes, induce identity formation of each pre-set segment by activating expression of the corresponding *HOX* genes (Irish *et al*, 1989). The gap and the pair-rule genes are replaced shortly after by proteins of the PcG/trxG system as *HOX* gene regulators. The PcG proteins are the transcription silencing factors in this system. They repress their target genes in segments, where they need to be inactive for identity preservation (Figure 2). In contrast, the trxG proteins protect *HOX* genes and other target genes from PcG-mediated silencing and maintain their transcription active at locations where their function is required. By their concerted, antagonizing action, PcG and trxG proteins maintain the patterning of the *HOX* gene expression domains along the anterior-posterior body axis from the post-blastoderm embryo up to the adult stage. *HOX* genes themselves contribute to the patterning, apart from a few exceptions, by repressing other *HOX* genes, whose expression domains are located anterior to theirs (Duboule & Morata, 1994). The cross-regulation among *HOX* genes was first described in the studies Morata & Kerridge, 1982 and Struhl, 1982.

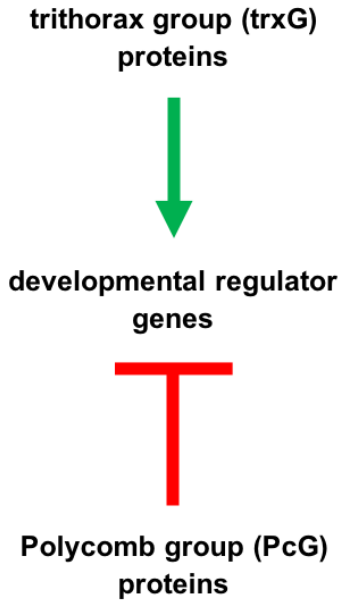


Figure 2. Transcription regulation by the PcG/trxG system.
PcG proteins repress transcription of *HOX* genes and other PcG/trxG target genes, whereas trxG proteins maintain their transcription active. Which action dominates at which gene depends on the respective segment identity.

1.1.3 Homeotic phenotypes of *ash1* mutants and PcG/trxG mutants in general

Mutations in all regulatory systems of *HOX* genes and in *HOX* genes themselves may lead to either loss or gain of function of certain *HOX* genes and consequently result in adaptation of the identities of the neighboring segments by the affected segments. The characteristic morphological transformations as part of the segment identity changes are known as homeotic transformations. While homeotic transformations in a *HOX* gene mutant only occur in a specific segment, mutants of transcriptional regulators of *HOX* genes generally exhibit a spectrum of such aberrant morphologies in various segments due to their role of controlling the entire homeotic gene cluster(s) Ant-C and/or BX-C.

The thorough analysis by Lewis, 1978 of homeotic transformations in a *Drosophila* mutant of the PcG group eponym *Polycomb (Pc)* led to a first understanding of the concept of PcG-

mediated transcription regulation: Ed Lewis described partial transformation of the thoracic and the first seven abdominal segments toward the identity of the eighth abdominal segment in homozygous *Pc* mutant larvae. From this observation, Ed Lewis derived the proposal that *Pc* encodes for a universal repressor of all genes in BX-C. In the following years, the easily extrinsically detectable homeotic transformations became the standard diagnostic tool for the discovery of other PcG proteins and, notably, for the discovery of the first universal positive regulators of the *HOX* genes in ANT-C and BX-C, the trxG proteins. The eponym of the trxG family, the *trithorax* (*trx*) gene, was identified and characterized in a *trx* mutant exhibiting transformations of the posterior abdominal segments toward the identity of anterior abdominal segments, and of the third thoracic segment towards the second (Ingham & Whittle, 1980; Ingham, 1981).

An allele of *ash1*, the trxG member whose function is investigated in this work, was first isolated and mapped in a screen for mutants with imaginal disc defects (Shearn *et al*, 1971; Shearn & Garen, 1974). Imaginal discs are a type of insect larval primordia for adult structures. Homeotic transformations of imaginal disc-derived tissues in *ash1*-mutant adults reminiscent of *trx* mutants were first reported in Shearn *et al*, 1987: Shearn and colleagues described transformations of genitalia towards leg identity and of structures of the third thoracic segment towards second thoracic segment identity, as for example the third leg adapting second leg morphology and the balancing organ haltere adapting wing morphology. In a follow-up study (Tripoulas *et al*, 1994), *ash1*-mutant transformations of structures derived from another type of larval precursor cells, the histoblasts, were shown: Abdominal segments were partially developed into anterior abdominal segments. Molecular evidence that loss of *HOX* gene expression is indeed responsible for the aberrant morphology in *ash1* mutants was provided in LaJeunesse & Shearn, 1995. The transformations described above, that are characteristic for *ash1* mutants will also be analyzed in *Drosophila* mutants in this study.

In general terms, mutants of members of the trxG family are characterized by a *HOX* gene loss-of-function phenotype displaying mostly posterior-to-anterior segment transformations. PcG mutants, in contrast, show ectopic expression of *HOX* genes, that results in mostly anterior-to-posterior segment transformations. The directionality of the transformations in the PcG/trxG mutants along the body axis is determined by the characteristic of *HOX* genes described in the previous paragraph to be negatively regulated by other *HOX* genes and not by PcG proteins in domains located posterior to their normal expression domain.

Whereas PcG proteins are active suppressors of *HOX* gene expression, the trxG proteins are often defined in the literature as factors, that maintain rather than initially activate PcG/trxG target gene transcription. This view is based on studies that showed specific *HOX* gene loss-of-

function phenotypes in mutants of the trxG genes *trx* or *ash1* alone, but demonstrated that *HOX* gene expression is almost restored to wild-type levels in double mutants of *trx* or *ash1* and members of the PcG family (Ingham, 1983; Klymenko & Müller, 2004). Moreover, *HOX* genes were misexpressed outside of their normal expression domains in the trxG/PcG double mutants, just like it is observed in the corresponding PcG single mutants. Thus, *trx* and *ash1* seem to positively regulate transcription by blocking establishment of PcG-mediated repression, which is a dispensable function in a PcG-mutant background. Yet, this mechanism does not apply to all members of the trxG family. A subset of the trxG proteins has been isolated in screens for dominant repressors of *Pc* mutant phenotypes (Kennison & Tamkun, 1988; Fauvarque *et al*, 2001), which implies a direct transcription activating function of these factors. Many of the dominant repressors of *Pc* mutants exhibit diverse mutant phenotypes beyond homeotic transformations and have therefore been suggested not to be restricted to transcription regulation of the classic PcG/trxG targets, but to fulfill rather global roles in gene regulation (reviewed in Kassis *et al*, 2017).

1.1.4 PcG/trxG target genes other than *HOX* genes and Ash1 chromatin binding genome-wide

HOX genes are the classic target genes of the PcG/trxG system and a role in *HOX* gene transcription regulation is generally seen as the prerequisite for classification of a specific factor as PcG or trxG protein. However, this does not rule out the possibility, that PcG/trxG members might also be involved in the transcription control of genes other than *HOX* genes.

A variety of genome-wide studies have indeed shown binding of PcG and trxG proteins to numerous other sites in addition to the *HOX* clusters (Lee *et al*, 2006; Boyer *et al*, 2006; Schwartz *et al*, 2006; Tolhuis *et al*, 2006; Nègre *et al*, 2006; Oktaba *et al*, 2008; Schwartz *et al*, 2010; Kockmann *et al*, 2013; Huang *et al*, 2017; Kwong *et al*, 2008). Actual functional evidence from genetic studies demonstrating transcriptional regulation by PcG/trxG members is only available for a subset of the bound non-*HOX* genes. Among these are mainly developmental regulator genes such as *even-skipped*, *engrailed* (Dura & Ingham, 1988; Americo *et al*, 2002), *hedgehog* (Maurange & Paro, 2002), *apterous*, *pannier*, *teashirt*, *Distall-less* and *Dorsocross* (Oktaba *et al*, 2008). Another group of genes that were found to be transcriptionally controlled by PcG/trxG proteins are cell cycle regulators, namely *Cyclin A* and *Cyclin B* (Martinez *et al*, 2006; Oktaba *et al*, 2008). Moreover, auto-regulatory mechanisms have been reported for the PcG family: The expression rate of genes encoding for the PcG proteins Posterior sex combs (Psc) and Suppressor of zeste 2 (Su(z)2) is dependent on PcG factors (Park *et al*, 2012). It is important to note that some

of the above listed non-*HOX* genes, that are here called PcG/trxG targets, might actually be under the control of only individual PcG or trxG proteins fulfilling additional functions outside of the PcG/trxG system.

The trxG protein Ash1 and its human homologue ASH1-like (ASH1L) have even been suggested to act as global transcription factors, that co-activate their target genes regardless of gene type and function (Gregory *et al*, 2007; Kockmann *et al*, 2013). Gregory and colleagues found ASH1L enriched at active housekeeping genes just as well as at active tissue-specific genes by chromatin immunoprecipitation-quantitative polymerase chain reaction (ChIP-qPCR) analyses. Kockmann and colleagues based their view on Ash1 as a general transcription regulator on genome-wide ChIP-DNA sequencing (ChIP-seq) analysis in *Drosophila* Schneider's 2 cell line-derived S2-DRSC cells, that showed Ash1-binding in the promoter region of virtually all active genes. In the light of these results, the puzzling question is, how the suggested global role of Ash1 could match with its specific homeotic mutant phenotype as described in the previous paragraph. Other studies then argue rather in favor of gene-specific functions of Ash1. Immunohistochemical stainings of *Drosophila* polytene chromosomes revealed about 100 sites bound by Ash1 (Tripoulas *et al*, 1996; Srinivasan *et al*, 2008). Nonetheless, these are still many more sites than there are *HOX* genes. In another genome-wide binding ChIP-chip study in Schneider's 2 cell line-derived Sg4 cells, only 50 genomic Ash1-binding sites in form of broad 10 to 75 kb-long chromatin domains encompassing gene bodies have been reported (Schwartz *et al*, 2010). Extensive Ash1-binding domains of a similar nature at about 400 genes in S2-DRSC cells were found recently by ChIP-seq as published in Huang *et al*, 2017. Transcriptomic analyses by Huang and colleagues revealed that only 18 out of the 400 Ash1-bound genes were actually downregulated (> 2-fold) in consequence of Ash1 knockdown. Most of the downregulated genes were not annotated or were developmental regulator genes, but not *HOX* genes. *HOX* genes are known to be transcriptionally silent in the analyzed S2-DRSC cell line (Cherbas *et al*, 2011). In the study Schmähling *et al*, 2018, in which parts of this thesis are published, the transcriptome of *Drosophila* thoracic imaginal discs was analyzed. *ash1*-null mutant and wild-type imaginal discs showed differential expression of about 600 genes by a factor greater than or equal to 2 and of about 300 genes by a factor greater than or equal to 4.

1.2 Polycomb-group/trithorax-group complexes and their molecular functions

Most PcG and trxG proteins that were first genetically identified in *Drosophila* have been subsequently purified in large multiprotein complexes (see Figure 4 for an overview). The means of PcG and trxG family members to control transcription are epigenetic mechanisms and the assembly of the PcG or trxG proteins to multimeric complexes provides an efficient way to fulfill the complex tasks in epigenetic regulation (e.g. complex binding and targeting to chromatin and regulation of enzymatic activities).

1.2.1 Polycomb-group complexes

PcG complexes act mainly as histone modifiers and readers. The Pc protein, whose mutant genotype analysis marked the beginning of PcG/trxG research (para 1.1.2), was purified in the Polycomb repressive complex 1 (PRC1) by Shao and colleagues (Shao *et al*, 1999). Pc comprises a chromodomain with a binding preference for tri-methylated histone H3 lysine K27 (H3K27me3) (Min *et al*, 2003; Fischle *et al*, 2003). The PRC1 subunit Sex comb extra (Sce) was shown to act as an E3 ligase, that mono-ubiquitinates lysine K118 in histone H2A (H2AK118ub1) in *Drosophila* and H2AK119 in mammals (Wang *et al*, 2004a). The role of H2A mono-ubiquitination deposited by Sce in PcG-mediated transcriptional silencing is not well understood. The repression of *HOX* and other canonical PcG target genes by PRC1 occurs independently of H2A ubiquitination in *Drosophila* and mice (Pengelly *et al*, 2015; Illingworth *et al*, 2015). What is established in the literature is, that PRC1 can alter chromatin structure: The PRC1 subunits Psc, its paralogue Su(z)2 and most likely also the Polyhomeotic (Ph) proteins are involved in PRC1-mediated chromatin compaction *in vitro*, which was, notably, even observed in the absence of histone tails (Francis *et al*, 2004; King *et al*, 2005; 2002).

The PRC1-like complex dRing-associated factors (dRAF) shares two subunits with PRC1, namely Psc respectively Su(z)2, and the E3 ligase Sce (Lagarou *et al*, 2008). In distinction from PRC1, dRAF contains the Lysine-specific demethylase 2 (dKDM2), which specifically removes the active chromatin mark H3K36me2 (Lagarou *et al*, 2008).

The dimeric Polycomb repressive deubiquitinase (PR-DUB) complex contains the subunits Additional sex combs (Asx) and Calypso. Calypso deubiquitinates H2AK118ub1 and is dependent on the presence of Asx for its catalytic activity (Scheuermann *et al*, 2010). The mechanism of PR-DUB-mediated transcription silencing is not yet understood. It has been suggested to counterbalance unconfined H2AK118 mono-ubiquitination by the PRC1 subunit Sce, but PR-DUB

activity is only required for repression at a subset of PRC1 target genes (Gutiérrez *et al*, 2012; Scheuermann *et al*, 2010).

Besides PRC1, the Polycomb repressive complex 2 (PRC2) is the most intensively studied PcG complex. PRC2 was originally purified in parallel by several groups from *Drosophila* embryos (Czermin *et al*, 2002; Müller *et al*, 2002) as well as from mammalian cells (Cao *et al*, 2002; Kuzmichev *et al*, 2002). The means of PRC2 to silence transcription is methylation of lysine K27 in histone H3 (H3K27) (Pengelly *et al*, 2013). As catalytic subunit of PRC2, Enhancer of zeste (E(z)) was identified and shown to be the only HMTase in *Drosophila* with specificity for H3K27 (Ebert *et al*, 2004; Ketel *et al*, 2005). E(z) catalyzes mono-, di- and tri-methylation of H3K27 (H3K27me1/2/3), but strictly requires the presence of the PRC2 subunits Suppressor of zeste 12 (Su(z)12) and Extra sex combs (Esc) for its activity (Cao *et al*, 2002; Czermin *et al*, 2002; Müller *et al*, 2002; Nekrasov *et al*, 2005). The 55 kDa subunit of Chromatin assembly factor 1 (Caf1-55) is generally also considered as a PRC2 core subunit but does not have an impact on the efficiency of the PRC2 HMTase activity. It has been suggested that Caf1-55 contributes to the binding of PRC2 to nucleosomes (Nekrasov *et al*, 2005). Pre-existing histone modifications on substrate nucleosomes are integrated into the regulation of PRC2 activity by allosteric mechanisms (Schmitges *et al*, 2011; Yuan *et al*, 2011; Margueron *et al*, 2009; see also para 1.3.1). These allosteric mechanisms have been suggested to be mediated by Su(z)12, Esc and the mammalian Esc homologue Embryonic ectoderm development protein (EED). The tetrameric PRC2 core complex purifies with additional subunits, that are either the Polycomblike (Pcl) protein or the proteins Jumonji, AT rich interactive domain 2 (Jarid2) and Jing together (Nekrasov *et al*, 2007; Herz *et al*, 2012). These cofactors have been shown to be involved in PRC2 binding to specific subsets of target genes (Savla *et al*, 2008; Kim *et al*, 2009; Herz *et al*, 2012) and to promote H3K27me3 deposition by PCR2 *in vitro* (Choi *et al*, 2017) and *in vivo* (Nekrasov *et al*, 2007).

The only PcG complex, that does not possess histone-modifying activity is the Pho-repressive complex (PhoRC). PhoRC is a dimeric complex formed by the proteins Scm-like with four MBT domain-containing protein 1 (Sfmbt) and Pleiohomeotic (Pho) or its paralogue Pho-like (Phol) (Klymenko *et al*, 2006). Pho and Phol are the only PcG proteins known to bind to DNA in a sequence specific manner (Brown *et al*, 1998; 2003). The binding sites of Pho/Phol are located within cis-regulatory elements, called Polycomb response elements (PREs), that are found upstream of or within PcG/trxG target genes. PREs are essential for inheritance of PcG-mediated transcriptional silencing in *Drosophila* (Laprell *et al*, 2017; Coleman & Struhl, 2017). In addition to PhoRC, PREs are enriched for subunits of PRC1, PRC1-like complexes and PRC2 (Schwartz *et al*, 2006).

Based on the findings described above, a model for PcG-mediated transcription repression involving hierarchical recruitment of the PcG complexes in *Drosophila* had been proposed (Wang *et al*, 2004b): The PhoRC complex, that is bound to PREs, would recruit PRC2 via interactions with Esc and E(z). PRC2 would then decorate neighboring nucleosomes with the H3K27me3 mark, that would in turn be bound by the chromodomain of the PRC1 subunit Pc and thereby recruit PRC1 to chromatin. Finally, PRC1 might trigger chromatin compaction at the target gene. By now, this model has been challenged by recent studies, especially regarding its universality (reviewed in Kassis *et al*, 2017). The data in Frey *et al*, 2016 for example suggest that PRC1 can be recruited independently of PRC2 to PREs by direct interactions between the sterile α-motif (SAM) domains of Sex comb on midleg (Scm) in PRC1 and of Sfmbt in PhoRC. Overall, it seems more likely that various mechanisms are employed by PcG proteins to silence different target genes. One aspect of transcription regulation by the PcG, that is for certain, is that *HOX* genes are decorated with PRC2-deposited H3K27me3 in the transcriptionally silent state in body segments where they are controlled by the PcG and not by other *HOX* genes (Papp & Müller, 2006; Bowman *et al*, 2014).

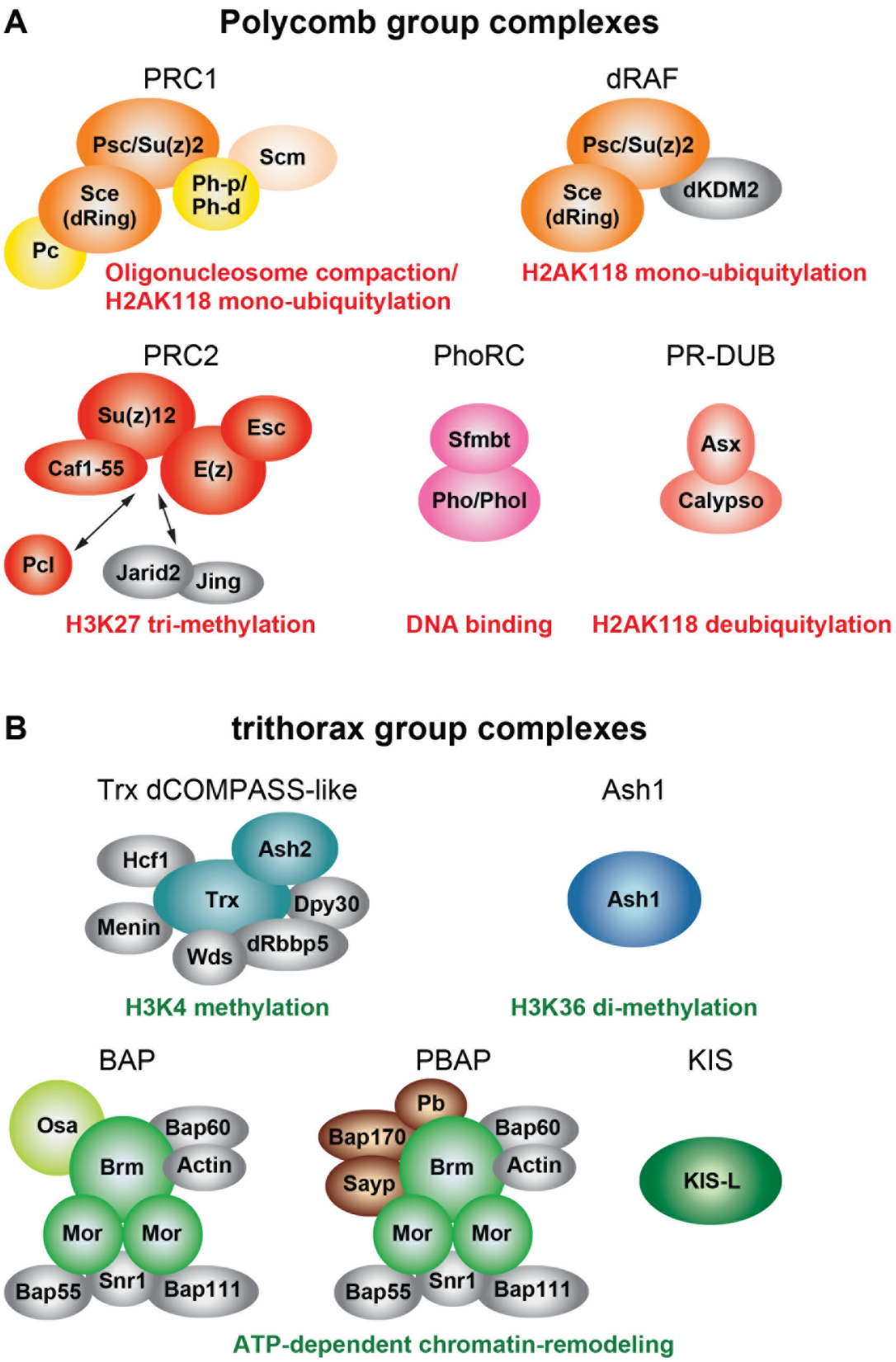


Figure 3. PcG and trxG complexes in *Drosophila* and their chromatin modifying or binding functions. (A) Complete overview on purified PcG complexes. Subunits in color have mutant PcG homeotic phenotypes, subunits in grey are biochemically identified complex members. Subunits shared by PRC1 and dRAF are coloured orange. Scm is present in substoichiometric amounts in PRC1. Pcl and the Jarid2/Jing dimer are mutually exclusive PRC2 subunits. *Continued on next page.*

(B) Selection of trxG complexes. Only the chromatin remodelers and histone modifiers as best studied trxG complexes are shown. Subunits in blue are proteins classified as trxG members due to their mutant homeotic phenotypes, subunits in grey co-purify but do not exhibit trxG mutant phenotypes. Stable interactors forming complexes with Ash1 or KIS-L could not be confirmed yet. Subunits in green are classified trxG proteins due to their functions as dominant suppressors of *Pc* mutants. Common subunits of the related BAP and PBAP complexes are of the same green shade. The protein Osa purifies exclusively with BAP and subunits in brown with PBAP. For abbreviations, see list in para 6.1. This illustration has been modified from Kassis *et al*, 2017.

1.2.2 trithorax-group complexes

Compared to the PcG of proteins, the trxG is much more heterogeneous in phenotypes as well as in molecular mechanisms of transcription regulation.

trxG proteins were found to oppose PcG-mediated silencing by adenosine triphosphate(ATP)-dependent chromatin-remodeling, chromosome cohesion, covalent histone modification and mediation of crosstalk between transcription factors and RNA polymerase II. Like PcG proteins, trxG proteins act as subunits of large multimeric assemblies. trxG complexes that are either chromatin remodelers or histone modifiers have been most extensively studied and are shown in Figure 4B.

The broad spectrum of mechanisms is in line with the diversity of trxG phenotypes. Only a subset of the trxG proteins has actually been identified based on a specific homeotic mutant phenotype like Ash1 and Trx (para 1.1.3). The majority was found in numerous screens for genetic interactors with various known PcG or trxG mutants as reviewed in Kassis *et al*, 2017. In the following, I will focus on two different groups of trxG proteins that have been known longest and are therefore most established. These are, on one hand, the trxG members that were discovered by their *HOX* gene loss-of-function phenotypes and, on the other hand, the trxG members that were found indirectly in a screen for dominant suppressors of *Pc* mutants (Kennison & Tamkun, 1988). Table 1 provides an overview on these factors, that are discussed in detail below.

Table 1. trxG proteins in *Drosophila* and their human orthologues. This table has been modified from Kennison & Tamkun, 1988 and Schuettengruber *et al*, 2011 and updated according to Kassis *et al*, 2017. Only members of the trxG family are shown, whose mutants exhibit a clear *HOX* gene loss-of-function phenotype or act as dominant suppressors of *Pc* mutants. For abbreviations not explained in this table, see list in para 6.1. For references, see text.

***Drosophila* trxG protein**	**Human orthologue(s)**	***Drosophila* complex**	**Function**
Histone-modifying complexes			
Ash1 (Absent, small, or homeotic discs 1)	ASH1L	this study	H3K36 di-methylation
Trx (Trithorax)	MLL1, MLL2	Trx complex (dCOMPASS-like)	H3K4 methylation
Ash2 (Absent, small, or homeotic discs 2)	ASH2L	Trx, Trr and Set1 complexes (dCOMPASS-family)	Complex integrity or stability of HMTase activities of dCOMPASS family
Utx (Ubiquitously transcribed tetratricopeptide repeat, X chromosome)	KDM6A, KDM6B, UTY	Trr complex (dCOMPASS-like)	H3K27me3 demethylation
ATP-dependent chromatin-remodeling complexes			
Brm (Brahma)	BRM, BRG1	BAP/PBAP (SWI/SNF family)	ATP hydrolysis
Mor (Moira)	SMARCC1, SMARCC2	BAP/PBAP (SWI/SNF family)	Complex integrity
Osa (Osa)	ARID1A, ARID1B	BAP (SWI/SNF family)	DNA binding
KIS-L (KISMET-L)	CHD7	-	ATP hydrolysis
Other functions			
Fsh-S (Female sterile (1) homeotic short)	BRD2, BRD4, BRDT	-	H3KAc binding, possible recruitment of transcription elongation factors, serine-threonine kinase
Kto (Kohtalo)	TRAP230	Mediator module	Promotion of enhancer/promoter interactions
Skd (Skuld)	TRAP240	Mediator module	
Vtd (Verthandi), alternative name: Rad21	Rad21	Cohesin	Mediation of interactions between insulators, enhancers, promoters in interphase
Nipped-B	NIPBL	Kollerin	Cohesin loading on interphase chromosomes

The trxG protein and ATPase Brahma (Brm) was purified from *Drosophila* in two different complexes, the Brahma-associated protein complex (BAP) and the Polybromo-containing BAP complex (PBAP) (Dingwall *et al*, 1995; Papoulas *et al*, 1998; Mohrmann *et al*, 2004). Both are ATP-dependent chromatin remodelers, that bind specifically acetylated histones via bromodomains in their subunits, such as in Brm. Besides complex-specific subunits, BAP and PBAP have seven subunits in common (Mohrmann *et al*, 2004; Chalkley *et al*, 2008). One of them is the trxG protein Moira (Mor) that has been suggested to form the functional core of BAP and PBAP together with Brm (Crosby *et al*, 1999; Phelan *et al*, 1999). The DNA-binding trxG protein Osa only purifies in BAP. With this subunit composition, BAP is highly related to the yeast Switch/sucrose nonfermenting (SWI/SNF) complex, whereas PBAP is considered to be the homologous complex to yeast remodels the structure of chromatin (RSC).

The trxG protein KISMET-L (KIS-L) is a member of the chromodomain helicase DNA-binding (CHD) family, a subcategory of the SNF2 family of ATP-dependent chromatin remodelers (Daubresse *et al*, 1999; Therrien *et al*, 2000). KIS-L contains all domains characteristic of CHD proteins, that are tandem chromodomains with binding specificities for methylated histones and a SNF2-like ATPase domain. Therefore, KIS-L most likely positively regulates *HOX* gene transcription by altering nucleosome positioning upon ATP hydrolysis. Possible stable biochemical KIS-L interactors have not been identified yet.

The Skuld (Skd) and the Kohtalo (Kto) protein are both subunits of the dissociable cyclin-dependent kinase 8 module that is part of the mediator complex (Allen & Taatjes, 2015). Mutations in *skd* and *kto* were isolated as dominant suppressors of the homeotic transformations that occur in Pc heterozygotes (Kennison & Tamkun, 1988). Intriguingly, genetic screens for mutants that show Polycomb-like phenotypes in homozygotes also led to the identification of loss-of-function mutations in *skd* and *kto* (Gaytán de Ayala Alonso *et al*, 2007). The mammalian homologue of Kto has been proposed to be involved in chromosome loop formation between enhancers and promoters in dependence of the cohesin complex (Apostolou *et al*, 2013). Cohesin itself is best known for controlling chromosome segregation in mitosis and meiosis, but has also been shown to affect transcription in interphase cells by regulating interactions between insulators, enhancers and promoters (reviewed in Dorsett & Merkenschlager, 2013). Interestingly, mutations in the cohesin subunit Verthandi (Vtd), also known as Rad21, were isolated as well as suppressors of the homeotic phenotypes of Pc heterozygotes (Kennison & Tamkun, 1988; Hallson *et al*, 2008). The same was observed for mutations in Nipped-B (Hallson *et al*, 2008), which is a subunit of the kollerin complex that loads cohesion on interphase chromosomes. In general, mutations that cause a decrease of chromosome-bound cohesin give rise to trxG phenotypes. But

mechanistically, the interdependence of cohesin with the PcG/trxG system in transcription regulation is not understood to date.

Even though Kto, Vtd and Nipped-B have been classified as trxG proteins, their roles in transcription are clearly not limited to the regulation of PcG/trxG target genes. Rather global roles in transcription activation have also been ascribed to Brm and the other trxG proteins involved in chromatin remodeling due to the wide range of gene expression defects in their corresponding loss-of-function mutants (Armstrong *et al*, 2002; Moshkin *et al*, 2007; Srinivasan *et al*, 2008). In contrast, histone modifying trxG proteins, such as Trx or Ash1, are generally considered as factors that fulfill functions specific for and rather restricted to the PcG/trxG system, albeit this view has also been sporadically challenged (Kockmann *et al*, 2013; Gregory *et al*, 2007).

Methylation of histone H3 lysine K4 (H3K4), a modification typically found at actively transcribed genes, has been shown to be catalyzed by three HMTases in *Drosophila melanogaster*: SET domain-containing protein 1 (dSET1), Trithorax-related (Trr) and Trx (reviewed in Kassis *et al*, 2017). The three H3K4 HMTases have been purified in distinct multiprotein complexes, that are all related in subunit composition to the yeast COMPASS complex (COMPASS: Complex of proteins associated with SET1) (Mohan *et al*, 2011). The trxG protein Absent, small, or homeotic discs 2 (Ash2) exists in all three *Drosophila* COMPASS complexes and has been proposed to be required for enzymatic activity or complex stability (Dehé *et al*, 2006; Dou *et al*, 2006). The COMPASS complex with dSET1 as core (dCOMPASS) does not contain other proteins with a specific mutant *HOX* gene loss-of-function phenotype apart from Ash2. dSET1 itself has been shown to be responsible for the bulk levels of H3K4me2 and -me3 (Ardehali *et al*, 2011; Hallson *et al*, 2012) indicating a genome-wide function of dCOMPASS. The trxG protein Utx is a subunit of a COMPASS-type complex that contains the H3K4 methyltransferase Trr (Trr dCOMPASS-like complex) (Cho *et al*, 2007; Mohan *et al*, 2011). Utx functions as the only H3K27me3 demethyltransferase in *Drosophila* (Smith *et al*, 2008) (Utx: Ubiquitously transcribed tetratricopeptide repeat, X chromosome). The activity of Utx has shown to be important for regulation of *HOX* gene expression in the very early stages of embryonic development during the onset of zygotic transcription (Copur & Müller, 2013; 2018). Since Trr and the other complex subunits do not exhibit mutant homeotic phenotypes, Utx appears to antagonize PcG action independently of other Trr dCOMPASS-like complex activities. The Trx protein, the eponym of the trxG, is embedded in the Trx dCOMPASS-like complex and is considered in the literature as the only H3K4 methyltransferase in *Drosophila* with a homeotic mutant phenotype (reviewed in Kassis *et al*, 2017). The human Trx homologues are the Mixed lineage leukemia protein 1 (MLL1) and 2 (MLL2). The catalytic domain in Trx is the classic SET domain, that was first identified in this protein (Mazo *et al*, 1990) and later found to be conserved in many other protein

methyltransferases (SET: Su(var)3-9, Enhancer-of-zeste and Trithorax). The HMTase activity of Trx has been suggested to be its critical function *in vivo* (Smith *et al*, 2004). Results of more recent studies indicate that Trx deposits H3K4me1 and me2, but not H3K4me3 (Tie *et al*, 2014; Rickels *et al*, 2016). H3K4 methylation has been shown to antagonize PcG action in a direct manner by inhibiting the catalytic activity of PRC2 allosterically (Schmitges *et al*, 2011).

1.2.2.1 Ash1 and possible Ash1 binding partners

Like Trx, the *Drosophila* Ash1 protein and its mammalian homologues modify histones as SET domain HMTases and positively regulate transcription. However, the mechanism of Ash1 and its potential interaction partners have long remained enigmatic. At the beginning of the work presented here, no Ash1-containing multiprotein complexes had been purified while most other PcG and trxG proteins had already been shown to exist and act in multimeric assemblies (Figure 4). In 2013, Kockmann and colleagues reported that both the short and the long isoforms of the Female sterile homeotic (Fsh) protein, Fsh-S and Fsh-L, physically interact with Ash1 (Kockmann *et al*, 2013). In fact, Fsh-S had been classified as a trxG protein before, when it was shown to be involved in *HOX* gene regulation (Digan *et al*, 1986; Chang *et al*, 2007). The human homologue of Fsh, the bromodomain-containing protein 4 (BRD4), binds via its bromodomains to acetylated histones at enhancers and promoters and has been suggested to recruit the positive transcription elongation factor b to these regulatory sequences (Jonkers & Lis, 2015). How Ash1 could be involved in Fsh function is not known. The interaction between Fsh and Ash1 described by Kockmann and colleagues was identified by tandem affinity purification of tagged Ash1 protein from *Drosophila* S2-DRSC cells. In a reverse approach, in immunoaffinity purifications of transgenic Fsh-S from S2-DRSC cells, Ash1 did not co-purify (Chang *et al*, 2007). Hence, further investigation, also in other biological systems, is needed to clarify if Ash1 interacts with Fsh and whether the two proteins form a stable complex.

The HMTase activity of Ash1 is in the focus of this work and therefore introduced in separate paragraphs in the following.

1.3 Specificity, function and regulation of Ash1 HMTase activity

The substrate specificity and the methylation state generated by Ash1 have been controversial for a long time. Ash1 was first believed to mainly methylate H3K4 (Byrd & Shearn, 2003; Gregory *et al*, 2007). By now it is established that Ash1 and its mammalian homologues exclusively mono- and di-methylate lysine K36 of histone H3 (H3K36me1/2). Ash1-dependent

H3K36 mono- and di-methylation *in vitro* was first shown by HMTase assays on substrate histone H3 lysine residue mutants and by mass spectrometric analysis in the study Tanaka *et al*, 2007. Confirming data sets were published in Yuan *et al*, 2011 and An *et al*, 2011. *In vivo* evidence was provided by Dorighi & Tamkun, 2013, who observed co-localization of Ash1 with H3K36me2 and a reduction of H3K36me2, but not H3K36me3, in *ash1* mutants at *Drosophila* polytene chromosomes. The current knowledge on how the H3K36me2 mark deposited by Ash1 opposes PcG-mediated transcriptional silencing, is summarized in the following.

1.3.1 Role of Ash1 in opposing transcriptional silencing by the Polycomb group

H3K36me2/me3 and H3K27me3 rarely coexist on the same histone H3 tail of nucleosomes (Yuan *et al*, 2011; Voigt *et al*, 2012) or in the same chromosome bands in *Drosophila* polytene chromosomes (Dorighi & Tamkun, 2013). Together, these observations suggest crosstalk between H3K36me2, catalyzed by Ash1 amongst other HMTases, and PRC2-deposited H3K27me3, that might be the means of Ash1 to counteract PcG action. Methylation of H3K27 has been shown to be essential for PRC2-mediated transcriptional silencing (Pengelly *et al*, 2013). HMTase assays with PRC2 on premodified nucleosomes revealed that H3K36me2 and -me3 marks indeed directly inhibit H3K27 mono-, di- and tri-methylation by PRC2 (Schmitges *et al*, 2011; Yuan *et al*, 2011). The inhibition was observed with PRC2 complexes from *Drosophila*, plants and humans. In order to affect H3K27 methylation, the H3K36 methylation mark needs to be present *in cis*, i.e. on the same histone H3 tail that contains the K27 substrate site (Schmitges *et al*, 2011; Voigt *et al*, 2012). The molecular mechanism behind this PRC2 inhibition has not been deciphered yet, but Schmitges and colleagues report findings indicating that H3K36 methylation might regulate the catalytic PRC2 subunit E(z) allosterically via the PRC2 subunit Su(z)2.

In the light of these *in vitro* data on PRC2 regulation, it is well conceivable but remains to be demonstrated that inhibition of the catalytic activity of PRC2 by di-methylating H3K36 is the actual mechanism Ash1 employs to counteract PcG-mediated transcriptional silencing. However, various findings in *Drosophila* point towards this scenario. Polytene chromosomes in *ash1* mutants exhibit a decrease in total H3K36me2 levels, while H3K27me3 levels are elevated (Srinivasan *et al*, 2008; Dorighi & Tamkun, 2013). In the study Papp & Müller, 2006, the distribution of the H3K27me3 mark along the *HOX* gene *Ubx* was analyzed. Papp and Müller found that in tissues where *Ubx* is actively transcribed, H3K27me3 is highly enriched in the upstream control region of *Ubx*, but strongly diminished in the gene body. Analysis of the same tissues in *ash1* mutants showed that H3K27me3 is present not only in the upstream control region but also at the *Ubx* promoter and all along the coding region, resembling the H3K27me3 distribution

pattern in tissues where the *Ubx* gene is repressed in *wild type* (Papp & Müller, 2006). It appears that in the transcriptionally silent state of *Ubx*, the H3K27me3 mark spreads from the upstream PRE, where PRC2 is bound, into the coding region. Consistent with this gain of repressive H3K27me3, *Ubx* expression in the analyzed *ash1* mutant tissues was strongly decreased in the study by Papp and Müller.

Based on these results, Ash1 was proposed to function by the following mechanism to antagonize PcG action (Papp & Müller, 2006; Schmitges *et al*, 2011): H3K36me2 deposited by Ash1 would act as a barrier that prevents spreading of the H3K27me3 mark past the promoter into the coding region. In *ash1* mutants, in contrast, H3K36me2 levels would be reduced and PRC2 would therefore no longer be inhibited, resulting in H3K27me3 deposition in this part of the gene to permit PcG repression by mechanisms not understood to date (see also para 1.2.1). However, whether Ash1 indeed acted by depositing H3K36me2 and did so across the entire coding region was not addressed at the time. Moreover, the actual binding profile of Ash1 has remained highly controversial (para 1.1.4).

1.3.2 Ash1-generated H3K36me2 in the context of genome-wide H3K36 methylation

H3K36 methylation marks actively transcribed genes in eukaryotic species ranging from yeast to *Drosophila* and humans as reviewed in Wagner & Carpenter, 2012. The two methylation states, H3K36me2 and H3K36me3, are enriched in chromatin comprising the entire length of the coding regions of active genes with a frequently observed bias of H3K36me2 towards the middle part and a bias of H3K36me3 towards the 3' ends (Bell *et al*, 2007; Pokholok *et al*, 2005; modENCODE data sets ID6388 and ID4950).

To date, various roles have been ascribed to H3K36 methylation in transcription (Venkatesh & Workman, 2013; Wagner & Carpenter, 2012). All evidence taken together points towards two major functions of H3K36 methylation. These are, on one hand, antagonizing PRC2-mediated transcriptional silencing as described above in para 1.3.1, and, on the other hand, quality control of transcription by preventing spurious intragenic transcription initiation through histone deacetylation by the Reduced potassium dependency-3 small (Rpd3S) complex. The yeast Rpd3S complex contains the chromo barrel domain protein ESA1-associated factor 3 (Eaf3). Structural studies showed that Eaf3 and its metazoan homologue MORF4-related gene on chromosome 15 (MRG15) interact with H3K36me2/me3 via their chromo barrel domains (Sun *et al*, 2008; Xu *et al*, 2008; Zhang *et al*, 2006). In line with these studies, Rpd3S is thought to bind specifically to H3K36 methylated sites through Eaf3 and to maintain a hypoacetylated state at

these sites. Loss of Rpd3S function in yeast leads to transcription initiation from within the coding region of genes in the sense and antisense direction (Carrozza et al, 2005; Joshi & Struhl, 2005; Li *et al*, 2007).

In *Saccharomyces cerevisiae*, a single HMTase, the SET domain-containing protein 2 (Set2), generates all existing H3K36 methylation marks (me1, me2 and me3) and is thought to be targeted to transcribed regions by association with the phosphorylated C-terminal domain of elongating RNA polymerase II (RNAP II) (Krogan *et al*, 2003; Xiao *et al*, 2003; Venkatesh & Workman, 2013). H3K36 methylation in metazoans is more complex: Their genomes encode several different H3K36 HMTases, each one responsible for catalysis of a distinct methylation state. In *Drosophila*, three H3K36 HMTases have been identified including SET2, Ash1 and the Nuclear receptor-binding SET domain protein (NSD) (Bell *et al*, 2007). While SET2 has been suggested to generate the bulk of H3K36me3, NSD appears to be responsible for catalyzing the bulk of H3K36me2 (Bell *et al*, 2007; Larschan *et al*, 2007). The contribution of Ash1 to H3K36 methylation in *Drosophila* with respect to quantity as well as genomic sites has remained poorly understood.

1.3.3 Architecture of the Ash1 SET domain and regulation of its activity

Ash1 contains characteristic conserved domains, qualifying this protein as both a chromatin writer and a chromatin reader. These domains all cluster in the Ash1-C-terminal half (Figure 4). The classic protein methyltransferase domain SET is followed by a bromodomain, a plant homeodomain (PHD) and a bromo-adjacent homology (BAH) domain (CD-search published in Marchler-Bauer *et al*, 2015). The Ash1 SET domain consists of the subdomains Associated With SET (AWS), core SET and post-SET (An *et al*, 2011; Rogawski *et al*, 2015). The alignment in Figure 4 demonstrates the high conservation of the Ash1 SET domain sequence across species among Ash1 orthologues and with the SET domains of other H3K36me2 HMTases like human NSD1 and *C. elegans* Maternal-effect sterile protein 4 (MES4).

Albeit clear genetic evidence was missing when I started my thesis, HMTase activity was considered to be the key physiological function of Ash1. It is probably for that reason that from all Ash1 domains, only the SET domain has been analyzed in greater detail by biochemical and structural approaches to shed light on the Ash1 catalytic mechanism and its regulation. An et al. reported the structure of the human ASH1L SET domain including all subdomains (An *et al*, 2011). This structure revealed that the entrance of the substrate binding pocket in the core SET domain is blocked by a loop formed by parts of the post-SET domain. The observation of high structural flexibility of this loop together with altered ASH1L enzymatic activity upon loop mutation

prompted An and colleagues to propose that this loop may fulfill a crucial regulatory function by auto-inhibiting Ash1 catalytic activity. Such a regulatory mechanism may be a common feature of H3K36 HMTases, since highly similar inhibitory post-SET loop conformations have been reported for human SET2 (Zheng *et al*, 2012) and human NSD1 (Qiao *et al*, 2011). For NSD1, molecular dynamics simulations predicted that binding to nucleosomes opens up the substrate binding pocket by changing the position of the inhibitory loop (Qiao *et al*, 2011). In line with such simulations, NSD1 is active on nucleosomes (Qiao *et al*, 2011). In the case of Ash1, in contrast, substrate binding does not suffice to trigger efficient histone methylation. *Drosophila* Ash1 or human ASH1L alone exhibit very weak or no detectable HMTase activity at all on nucleosomes (An *et al*, 2011; Rogawski *et al*, 2015; unpublished experiments by the Jürg Müller laboratory). This suggested that Ash1 binding partners might be required to permit the enzyme to efficiently methylate histones in nucleosomes, possibly by re-positioning the auto-inhibitory loop. Other loop-independent regulatory mechanisms mediated by Ash1 interactors are also conceivable.

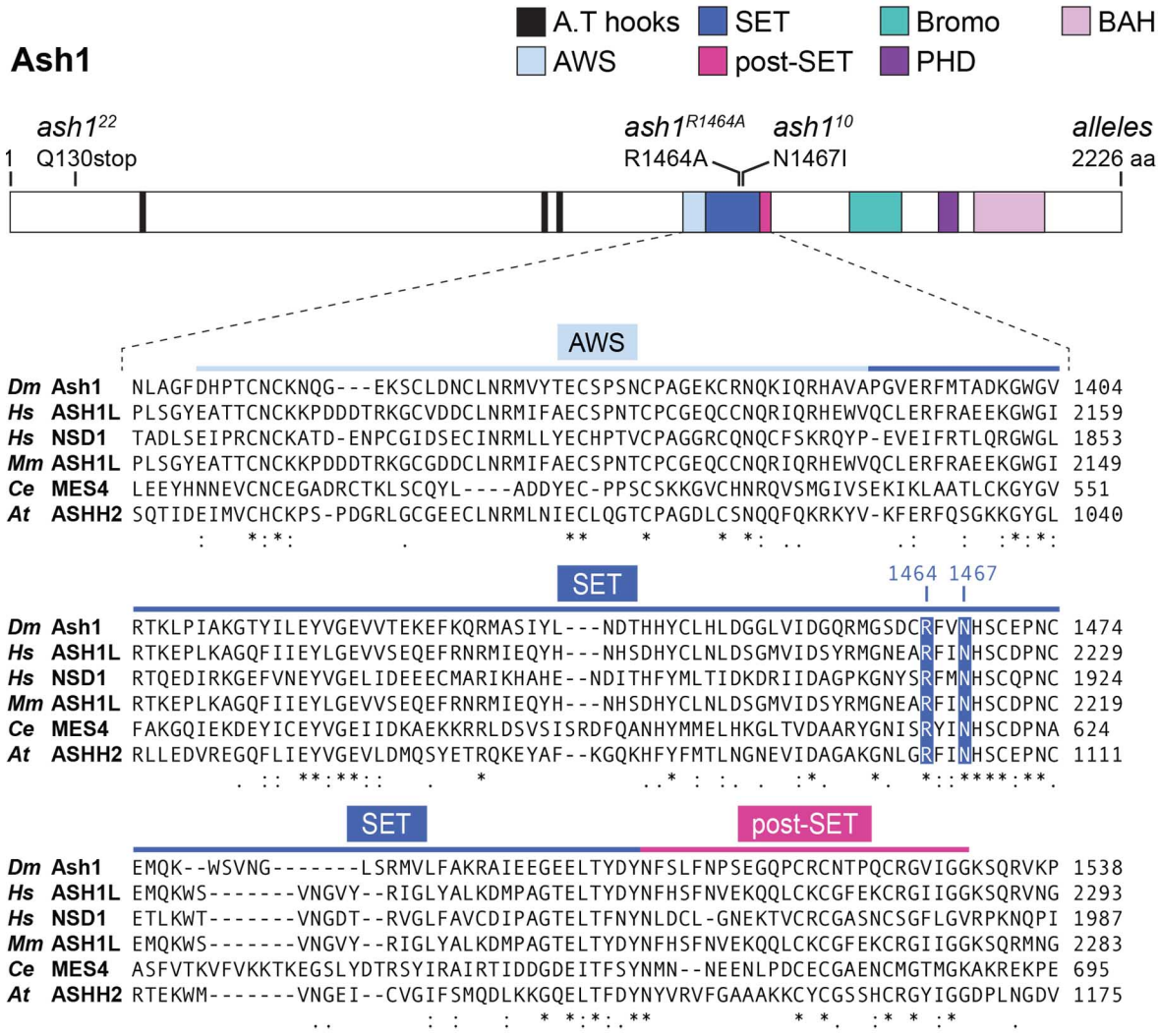


Figure 4. Ash1 domain architecture and SET domain alignment with mutations in *ash1* mutant alleles indicated. The catalytic domain of Ash1, SET, consists of three subdomains: AWS, SET and post-SET. They are highly conserved among Ash1 orthologues (*Hs/Mm* ASH1L, *At* ASHH2) and with other H3K36me2 HMTases (*Hs* NSD1, *Ce* MES4). The alignment was performed with Clustal Omega (EMBL-EBI).

The Ash1 domain scheme shows also mutations of *ash1* mutant alleles used in this study. The *ash1*^{22} allele contains a premature stop codon, Q130stop, at the Ash1 N-terminus and has been reported in Tripoulas *et al*, 1996 to be a functional null allele. The *ash1*^{R1464A} and the *ash1*^{10} alleles bear point mutations of highly conserved residues in the SET domain (mutated residues marked in blue in alignment). See para 3.5 for detailed description and analysis of these alleles.

A.T hook: adenine-thymine hook; AWS: Associated With SET; SET: Su(var)3-9, Enhancer-of-zeste and Trithorax; PHD: plant homeodomain; BAH: bromo-adjacent homology; aa: amino acids; *Dm*: *Drosophila melanogaster*; *Hs*: *Homo sapiens*; *Mm*: *Mus musculus*; *Ce*: *Caenorhabditis elegans*; *At*: *Arabidopsis thaliana*; ASH1L: ASH1-like; MES4: Maternal-effect sterile protein 4; ASHH2: ASH1 homologue 2.

1.4 Aims of this work: Understanding the regulation and function of the Ash1 HMTase in transcription

As discussed in the previous paragraphs, *ash1* has been classified as a member of the trxG of genes based on its homeotic mutant phenotype. Genetic analyses of *ash1* mutants and *ash1* PcG double mutants strongly suggested that Ash1 maintains transcription active at PcG/trxG target genes by specifically antagonizing transcription-silencing PcG actions. However, many questions were open concerning the molecular mechanisms underlying Ash1 function. It has been established that Ash1 possesses HMTase activity for di-methylation of H3K36. And the H3K36me2 modification has been shown to inhibit H3K27 methylation by the PcG complex PRC2 *in vitro*. But if the deposition of H3K36me2 is indeed the mechanism Ash1 employs to antagonize PcG-mediated transcriptional repression, remained to be demonstrated. Moreover, the two other H3K36 HMTases in *Drosophila*, SET2 and NSD, have been proposed to generate the bulk of H3K36 methylation. Whether Ash1 also contributes to H3K36 methylation genome-wide or if it modifies chromatin rather restrictively at PcG/trxG target genes as its mutant phenotype suggests, was another unresolved, controversial question in the field. Importantly, it was also not understood how the Ash1 HMTase function becomes activated in the first place. The SET domain of isolated Ash1 has been shown to be auto-inhibited.

The first objective of this thesis as part of the overall goal to learn about the regulation and function of the Ash1 HMTase was to identify Ash1 binding partners by purifying Ash1 from *Drosophila*. Ash1 protein complexes were then reconstituted *in vitro* from recombinant Ash1 and the newly found Ash1 complex subunits to probe their HMTase activity. These studies uncovered that one of the binding partners strongly stimulates the catalytic efficiency of Ash1. In order to assess the relevance of this stimulatory mechanism for Ash1 activity *in vivo*, the phenotypes of mutants for *ash1* and its interactor were analyzed. Importantly, the complementation of the genetic analyses with molecular studies in these mutants permitted also the investigation of the function of Ash1-generated H3K36me2 in *HOX* gene expression and the contribution of Ash1 to the bulk of H3K36me2 in *Drosophila*.

2 Materials and Methods

2.1 DNA analysis and cloning techniques

2.1.1 PCR amplification and PCR product purification

Polymerase chain reaction (PCR) amplification of specific DNA sequences was performed with the 'Phusion High-Fidelity PCR Master Mix with HF Buffer' (New England Biolabs).

Apart from site-directed mutagenesis reactions (para 2.1.6), component concentrations in the PCR reactions and thermocycling conditions were chosen according to manufacturer's instructions. One reaction contained 20 – 200 pg template. For preparative PCR, 50 µl reactions and for analytical PCR, 25 µl reactions were pipetted. PCR products were analysed by agarose gel electrophoresis and purified with the 'QIAquick PCR Purification Kit' or the 'QIAquick Gel Extraction Kit', both from Qiagen.

2.1.2 Separation of DNA and mononucleosomes by agarose gel electrophoresis

Agarose gel electrophoresis was either performed under denaturing or under native conditions. In the denaturing approach, DNA fragments from a PCR reaction or restriction digest were separated dependent on their size to analyze the length or amount of a specific fragment and, where necessary, to purify it from the gel. In native conditions, separation is not only size- but also conformation-dependent. This approach was chosen to test the quality of reconstituted mononucleosomes for HMTase assays by separating correctly assembled mononucleosomes from free DNA and from nucleosomes with an incorrect stoichiometry.

For denaturing electrophoresis, an agarose gel consisting of 0,8 – 1,5% (weight/volume) agarose (Sigma) and 0,01% (volume/volume) 'GelRed Nucleic Acid Gel Stain' (Biotum) in 1x TBE was prepared (1x TBE: 89 mM tris(hydroxymethyl)aminomethane (Tris); 89 mM boric acid; 2 mM ethylenedinitrilo-tetraacetic acid (EDTA)). In each case, the percentage of agarose was adjusted to the size of the DNA fragments that were to be separated. DNA samples were mixed with 6x DNA loading dye (Fermentas) and loaded on the gel along with appropriate size markers. The gel run took place in 1x TBE at 90 V for the time necessary for DNA separation.

To analyze mononucleosomes by native electrophoresis, a 1% (w/v) agarose gel (SeaKem ME Agarose, Biozym) in 0,4x TBE without a DNA stain such as 'GelRed Nucleic Acid Gel Stain' was poured. Approximately 3% of each nucleosome reconstitution were taken for the electrophoretic

quality check and supplemented with 8% glycerol instead of loading dye prior to gel loading. The gel was run in 0,4x TBE for at least 1 hr and afterwards stained for 20 min in 0,4x TBE in a light-protected box on a shaker.

In the last step, DNA was visualized in gel documentation systems with either an ultraviolet radiation or a blue light source.

2.1.3 Restriction endonuclease digest and fusion of DNA fragments

If conventional restriction cloning was performed, PCR product respectively insert and vector, that were going to be fused, were digested with endonucleases from New England Biolabs. 50 µl digest reactions containing 1-3 µg of insert or vector, 10 U per 1 µg DNA of enzymes and enzyme-specific buffer were set up. Digests took place overnight at a temperature corresponding to the respective enzyme's optimum activity. As purification method of DNA from the digest mixture, gel extraction with the 'QIAquick Gel Extraction Kit' (Qiagen) was chosen to also remove undigested plasmid. Digested and purified insert and vector were ligated with the Quick Ligation Kit from New England Biolabs according to manufacturer's instructions.

The most frequently used cloning method in this study was not restriction cloning but In-fusion cloning with the 'In-fusion HD Plus EcoDry kit' from Clontech. In this method, only the vector is digested with endonucleases and no ligation is performed. Complementary 15 nucleotides long 5'-overhangs at the termini of the previously linearized vector and the insert are generated and annealed in one step by the In-Fusion enzyme mix.

2.1.4 Transformation, amplification and purification of DNA constructs

Standard transformation was performed with DH5α competent cells (Thermo Fisher), whereas plasmids larger than 15 kb were transformed into One Shot OmniMAX 2 $T1^R$ *E. coli* from Invitrogen. For both bacterial strains, the same chemical transformation protocol was followed. 50 µl of competent bacteria were thawn on ice. Then, 2 µl of either ligation mixture or In-Fusion reaction (para 2.1.3) were added to the bacteria. After an incubation step on ice for 20 min, the transformation mixture was heat-shocked for 45 seconds at 42°C and then immediately placed back on ice for 2 min. The bacteria were topped up with 250 µl lysogeny broth (LB) without antibiotics and grown for 1 hour at 37°C and 600 rounds per minute (rpm) in a thermoshaker. Next, 10 – 70 µl of the bacteria culture were plated on an LB agar plate, that had been supplemented with the appropriate antibiotic, and incubated overnight at 37°C.

To amplify plasmids prior to purification, 4 ml LB media with antibiotics were inoculated with a single clone picked from the LB agar plate and the bacteria culture was grown overnight at

37°C with agitation in a bacteria shaker. Plasmids were purified from the culture using the 'Plasmid Mini Kit' from Qiagen.

If DNA constructs were intended for transformation into *Drosophila*, 10 ml overnight cultures were set up and plasmids were purified with the 'PureLink HiPure Plasmid Miniprep Kit' (Invitrogen) to obtain better yields of higher purity. For preparation of the bacterial artificial chromosome (BAC) CH322-147P9, 10 ml of bacteria culture were supplemented with 10% arabinose after overnight incubation for induction of high-copy BAC amplification and incubation was continued for another 5 hours (hrs) at 37°C.

2.1.5 DNA sequencing

DNA constructs were pre-mixed with sequencing primers in the Mix2Seq Kit (eurofins Genomics) according to manufacturer's instructions and sent for sequencing to eurofins Genomics. Sequencing primers used in this study are listed in Table 4, Table 5 and Table 6.

2.1.6 Site-directed mutagenesis

For PCR-based site-directed mutagenesis of plasmids, two 30-40 nucleotides long mutagenesis primers, complementary to each other and with the intended nucleotide substitutions in their center, were designed. PCR reactions were prepared with the 'Phusion High-Fidelity PCR Master Mix with HF Buffer' (New England Biolabs) introducing the following modifications to the manufacturer's protocol: Primer concentration was decreased to 0,1 µM to avoid dimers of the long mutagenesis primers, template concentration was increased to 30 ng/50 µl since amplification in this case is only linear. Thermocycling conditions were always set as follows regardless of template size and specific primer melting temperature:

Initial denaturation	98°C	30 sec
25 cycles	98°C	15 sec
	53°C	18 sec
	72°C	10 min
Final extension	72°C	7 min

After amplification of the DNA construct in the thermocycler, 10 U of DpnI were directly added to 50 µl PCR reaction to digest the methylated template. The digest was carried out for 3 hrs at 37°C. Then, purification of the PCR product directly followed by transformation were performed as described in para 2.1.1 and 2.1.4.

2.1.7 DNA constructs and their generation

DNA constructs in this study were either cloned in preparation of the generation of *Drosophila melanogaster* mutants or in preparation of baculovirus-based protein expression in insect cells. Table 2 provides a list of these constructs and points out relevant features of the corresponding vector backbones (for explanations of abbreviations see list in para 6.1).

Table 2. Overview of DNA constructs used in this work and description of their vector backbones.

Construct	Vector	Vector application and features	Reference/source vector
pCaSpeR-*tub*-NTAP-ash1	pCaSpeR-*tub*-NTAP	Fly transformation and expression vector; *P* element; *white* marker gene; *α-tubulin1*-promoter; N- or C-terminal TAP tag	Thummel & Pirrotta, 1991; modified by Klymenko *et al*, 2006
pCaSpeR-*tub*-ash1-CTAP	pCaSpeR-*tub*-CTAP		
pUMR-FLAP- ash1^{R1464A}	pUMR-FLAP	Fly transformation and expression vector; integration into *attP* landing site via *attB* site; *α-tubulin1*-promoter	Gambetta & Müller, 2014; Pengelly *et al*, 2015
pUMR-FLAP-ash1^{wt}			
pW35-MRG15$^{\Delta}$	pW35	Fly transformation; *P* element *mini-white* gene; *FRT*s; I-*Sce*I sites	Gong & Golic, 2003
BAC CH322-160G6	*attB*-P[acman]-CM^{R}-BW	Fly transformation; integration into *attP* landing site via *attB* site; encodes MRG15	Venken *et al*, 2009, BACPAC Resources Center
pFastBac1-Flag-Ash1$_{C}$	pFastBac1-Flag	Transfer of cDNA to bacmid in preparation of recombinant protein expression; Flag-tag N-terminal to MCS	pFastBac1 vector purchased at Invitrogen and modified by Raquel Matos
pFastBac-HT-Ash1$_{C}$	pFastBac-HT	Transfer of cDNA to bacmid in preparation of recombinant protein expression; TEV-cleavable 6xhistidine(His)-tag N-terminal to MCS	Invitrogen
pFastBac-HT-Ash1$_{C}^{RxRP}$			
pFastBac-HT-Ash1$_{C}^{R1464A}$			
pFastBac-HT-Caf1-55			
pFastBac1-StrepII-Caf1-55	pFastBac1-StrepII	Transfer of cDNA to bacmid in preparation of recombinant protein expression; Strep-tag StrepII N-terminal to MCS	pFastBac1 vector purchased at Invitrogen and modified by Reinhard Kalb
pFastBac1-StrepII-Esc			
pFastBac1-StrepII-MRG15			
pFastBac1-StrepII-Msl3			

2.1.8 Cloning of TAP-ash1 expression constructs pCaSpeR-*tub*-NTAP-ash1 and pCaSpeR-*tub*-ash1-CTAP

Arif Mohammed, a former postdoc in our laboratory, cloned the complete $Ash1_{1-2226}$ open reading frame into the *Drosophila* transformation vectors pCaSpeR-*tub*-NTAP and pCaSpeR-*tub*-CTAP, respectively. Both vectors had been modified from pCaSpeR vectors by Tanya Klymenko (Klymenko *et al*, 2006). N- or CTAP stands for N- or C-terminal tandem affinity purification tag.

2.1.9 Cloning of Ash1 transgenes *ash1*^{R1464A} and *ash1*^{wt} into the pUMR-FLAP vector

Genomic *ash1* coding and flanking sequences corresponding to the genome coordinates BDGP R6.14 chr3L:19,600,040...19,590,604 were subcloned from BAC CH322-147P9 (Venken *et al*, 2009) into the pUC19 vector by conventional restriction cloning (BDGP R6.14: Berkeley *Drosophila* genome project release 6 update 14 genome assembly).

In the *TBP-associated factor 6 (Taf6)* coding sequence, present in this fragment, the ATG initiation codon and four downstream alternative start codons specified in Table 3 were mutated to TAA stop codons in the case of both the *ash1*^{R1464A} and the *ash1*^{wt} transgene. For the *ash1*^{R1464A} transgene, the AGG codon for Arg_{1464} was converted into GCG. Mutagenesis primers used are given in Table 3. All mutagenesis reactions were performed with the 'QuikChange Lightning Multi Site-Directed Mutagenesis Kit' from Agilent Technologies according to manufacturer's instructions.

In the final cloning step, both transgenes were cloned into the pUMR-FLAP vector, which had been modified to contain an *attachment site on bacterial DNA* (*attB* site) by Reinhard Kalb in our laboratory. The finalized constructs were sequenced in-depth with primers listed in Table 4. For injection of these constructs into *Drosophila* see para 2.3.2.

Table 3. Mutagenesis primer for mutating *Taf6* start codons and *ash1* SET domain coding sequence for generation of the *ash1*^{wt} and *ash1*^{R1464A} transgenes. *Table continued on next page.*

Name	Sequence (5' to 3')	Codons changed
ash1_R1464A_F	CAAGAATGGTTGACAAACGCACAATCGCTGCCCATCCG	AGG into GCG in ash1 → R1464A
Taf6_M1Stop_R	GCTCGGTTTCGACGGTTTTCCACTTTATTCCCAGGTTTTTACAAATAAACA	ATG into TAA in Taf6 → M1Stop
Taf6_M26Stop_R	GCTCTCCGCGATCACCTTTTAGGACTCCGCCGAG	ATG into TAA in Taf6 → M26Stop

Name	Sequence (5' to 3')	Codons changed
Taf6_M80Stop_R	CATTTCGCACCTTAAGGGATTAGTCGATGTCCCGCAC TGAG	ATG into TAA in Taf6 → M80Stop
Taf6_M175Stop_R	CGCATCTTTGTTTAGGCCTTGATCTTACTTAATAACTG GATTGACCGAGTC	ATG into TAA in Taf6 → M175Stop
Taf6_M280Stop_R	GATTATCCAGAAGCGCACGAACTTAGCGCATGAGGT AAATAAGCAAC	ATG into TAA in Taf6 → M280Stop

Table 4. Sequencing primers used to verify the pUMR-FLAP-ash1^{wt}/-ash1^{R1464A} constructs.

Name	Sequence (5' to 3')	Name	Sequence (5' to 3')
-1286_R	CTTCTCCAGAAACAGCGAAGG	+3178_R	CAGCTAAACAGAAGGTGTCGG
-957_R	CTGCCGCATCTTTGTTTAG	+3576_R	CCACTGCAATCCCTAAGAGAGAC
-829_F	GTTTGTTGTGGAGGGAGTGC	+3972_R	GATGGCATCGATCCCAATAC
-366_R	AGCTGCTCGGCCAGTGTG	+4397_R	TCCTGCGAGCCTTCTACCAC
-14_F	ACGTGATGTCTACGCGACAG	+4794_R	CGAACTAAGCTACCCATTGCG
+369_R	GCGTAAGTGTGCTGTAGGTAATG	+5162_R	TGTCATTGGTGGCAAGTCG
+567_F	CGTTTCCAGAACCTTTGCTG	+5962_R	AAGCGCTGCAGTCCCTGAAG
+785_R	GTTAGGAGTCGTAGCAACCAAG	+6374_R	CCTGGAGGAATTTACCGAAG
+1165_R	GCGATACGGAGGACACAAC	+6782_R	GGAATGCAACGATGAGGATC
+1571_R	CATGGCGTATCTGGACAAGC	+7156_R	GGATGCAACTGCTGTCCATG
+1963_R	GCTACCGGAGTGATTGCAAG	+7551_R	CCATGTAAGATGTAACGGTGATAG
+2356_R	GTATGTGCAGCAGCTATGTGTC	+8152_F	GCAAATGAATTTTTGCGATAC
+2787_R	CAAAGCAATGACAGTTCCAGTC		

2.1.10 Cloning of the pW35-MRG15$^{\Delta}$ construct for generation of the *MRG15$^{\Delta}$* deletion allele

The *MRG15$^{\Delta}$* allele was generated by the ends-out targeting technique (Gong & Golic, 2003), by which the region of interest is removed by replacement with the *mini-white* (*w*) marker gene by homologous recombination. To this end, sequences homologous to the genomic regions up- and downstream of *MRG15* had to be cloned into the pW35 vector up- and downstream of its *mini-white* gene and are referred to as 5' and 3' homology arms from here on.

In the *MRG15$^{\Delta}$* allele, the major part of exon 2, the complete exon 3 and the major part of exon 4 of *MRG15* (BDGP R6.14 chr3R: 15,276,676...15,277,889) were deleted, but not the entire open reading frame in order to not damage genes overlapping with *MRG15*. Accordingly, in the initial cloning step, sequences corresponding to BDGP R6.14 chr3R: 15,277,890...15,281,975 and BDGP R6.14 chr3R: 15,272,003…15,276,675 were amplified as 5' and 3' homology arms from BAC

CH322-160G6 (Venken *et al*, 2009). The 5' homology arm was subcloned into the pCR 2.1-TOPO TA vector (Invitrogen) to then mutate the initiation ATG of *MRG15* to ATC by site-directed mutagenesis using primers

5'-CTATTATTGAAAATAAAATCGGAGAAGTAAAACCCGCTAAAG-3' and

5'-CTTTAGCGGGTTTTACTTCTCCGATTTTATTTTCAATAATAG-3'.

Both, the 5' and the 3' homology arm, were subsequently cloned into pW35 with the 'In-fusion HD Plus EcoDry kit' from Clontech according to manufacturer's instructions and the sequence of the resulting construct pW35-MRG15$^{\Delta}$ was verified with primers listed in Table 5.

The succeeding *in vivo* work is described in para 2.3.7.

Table 5. Sequencing primers used to verify the pW35-MRG15$^{\Delta}$ construct and the homology arms including their endogenous flanking regions in the *MRG15$^{\Delta}$* deletion allele

Primer 5' homology arm (HA)		**Primer 3' homology arm (HA)**	
Name	**Sequence (5' to 3')**	**Name**	**Sequence (5' to 3')**
5'pw35_F	ACATCCACTTAACGTATGCTTGC	3'pw35_F	GCTCTCTTTGTGTGCGTGTG
5'HA_1R	TGACAAGGTCCCCAGAATCG	3'HA_1R	CCTTAAATTCCTCGTAAAGAACAGC
5'HA_2R	CGCGACTTGAACGGATTGTG	3'HA_2F	GCTGGTTTCTGGGCAAGTTC
5'HA_3R	GCGAAATCAACAAAGGGGTGC	3'HA_2R	AGAAACCAGCGAGCCTCTTC
5'HA_4R	CCGCTTTCTCTCTGCCTCTC	3'HA_3R	TTGGATCGCAGTGGTTGAGG
5'HA_5R	CAACGTCATGTCTGCGAACG	3'HA_4R	GATTGCAAAGAAGGGTGGCG
5'HA_6R	TTGCCAGTCCGAGGTTGATG	3'HA_5R	GCAGTAAACAAGCCGCTGAG
5'HA_7R	GCTCAGACAAGTGCGCAATC	3'HA_6R	TTACCCATTCACCTGCGTCC
5'HA_8R	AAAGACTGTGCGTAGGGTGC	3'HA_7R	TCATGCCCAACGAGTTCCAG
5'HA_9R	CAGAACTTGCTCGCACTTGG	3'HA_8R	ATGAGCCGAACATCAGGTGG
5'HA_10R	TTGCAGTGTAACCGCATG	3'HA_9R	CCAACGCCAAGCTGAATCAC
5'HA_11R	GGAAGATCAGCCGACGACAG	3'HA_10R	GACAATGCACGCCTTGATCG
5'HA_12F	AACCCGTTCCCCTGCAAATG	3'HA_11R	GACGACAACAAGCAGCAGTG
5'pw35_R	CAAAGTGCAACTGAAGGCGG	3'HA_12R	CGTCAAGCAGGAGACCGATG
		3'pw35_R	AAGTGATAGAGCCTGAACCAG

2.1.11 Cloning of cDNA into pFastBac vectors for baculovirus-based expression

All pFastBac constructs used, and with the exception of pFastBac1-Flag-Ash1$_C$ generated in the course of this study, are listed in Table 2. For proteins to be expressed with an N-terminal hexa-histidine tag (His-tag), the respective cDNAs were cloned into the the pFastBac-HT vector (Invitrogen); for proteins to be expressed with an N-terminal Strep-tag, cDNAs were cloned into the pFastBac1-StrepII vector. Cloning steps were performed with the 'In-fusion HD Plus EcoDry kit' from Clontech according to manufacturer's instructions.

In the case of the Ash1 constructs, the cDNA of the C-terminus of Ash1, $Ash1_C$ (aa 1041 - 2226), was amplified from a pFastBac1-Flag-$Ash1_C$ construct that had been cloned previously by Raquel Matos, a former Phd student in our laboratory. The Caf1-55-cDNA was a gift from the Christoph Müller laboratory, EMBL, Heidelberg. Each, the $Ash1_C$-cDNA and the full-length Caf1-55-cDNA, were integrated into a pFastBac-HT vector, that had been linearized with KasI and HindIII beforehand. This pFastBac-HT-$Ash1_C$ construct was then directly used as template to generate the pFastBac-HT-$Ash1_C^{RxRP}$ and the pFastBac-HT-$Ash1_C^{R1464A}$ constructs by site-directed mutagenesis. For the FxLP to RxRP mutations in the pFastBac-HT-$Ash1_C^{RxRP}$ construct the mutagenesis primers

5'-CTCCGAAGGACTGAAATGGACCGTGAGCGACCTTACGACATTTGGTG-3' and

5'-CACCAAATGTCGTAAGGTCGCTCACGGTCCATTTCAGTCCTTCGGAG-3' were used.

The mutagenesis primers to make the pFastBac-HT-$Ash1_C^{R1464A}$ construct were

5'-CAGCGGATGGGCAGCGATTGTGCGTTTGTCAACCATTCTTGC-3' and

5'-GCAAGAATGGTTGACAAACGCACAATCGCTGCCCATCCGCTG-3'

The cDNAs for Esc, MRG15 and Msl3 were obtained from the 'DGRC cDNA library GOLD Parts 1-3, v2.1.1' (Stapleton *et al*, 2002). Full-length Caf1-55-, full-length Esc-, full-length MRG15- and full-length Msl3-cDNAs were cloned into a pFastBac-StrepII vector, that had been digested with KpnI and SalI for the MRG15 construct and with KasI and HindIII for the other constructs.

All pFastBac constructs were sequenced thoroughly with primers in Table 6.

Table 6. Sequencing primers to verify the pFastBac constructs. The number in the primer name indicates primer distance in base pairs from the start codon. MRG stands for MRG15.

Name	Sequence (5' to 3')	Name	Sequence (5' to 3')
pFastBac_F	TGTTCGCCCAGGACTCTAGC	Caf1_1068F	GAGCAGAGTACGGAGGATGC
pFastBac_R	GCAGGCTCTAGATTCGAAAGC	Esc_356F	TCCGGATCCCGATGAAGTAT
Ash1_2995F	CCACTGCAATCCCTAAGAGAGAC	Esc_750F	CACAAGATCGAACTGTCGAA
Ash1_3391F	GATGGCATCGATCCCAATAC	Esc_1150F	ACAATTCGCGTAGCGTGG
Ash1_3816F	TCCTGCGAGCCTTCTACCAC	MRG_71R	TCCACGAACAGAGTGTTTGC
Ash1_4213F	CGAACTAAGCTACCCATTGCG	MRG_370F	ATAGCAACACCTCGCAATCG
Ash1_4581F	TGTCATTGGTGGCAAGTCG	MRG_480R	GGCGGTGCTGTTAGTAGTCG
Ash1_4923F	CTCCACACCATCTTCTCCTTC	MRG_965F	CGTTTATGGCAACCTCCTTG
Ash1_5318F	AAGCGCTGCAGTCCCTGAAG	MRG_1170R	CGTGAGTAGGTTCTGCATGG
Ash1_5730F	CCTGGAGGAATTTACCGAAG	Msl3_20F	GACACCGCTCTTTCACAAGG
Ash1_6138F	GGAATGCAACGATGAGGATC	Msl3_582F	CCACAGGAGGATCGCATTATG
Caf1_179F	GGATGGCAAGGACTACTCGG	Msl3_869R	CCGTCTACCACCTCCTTCAG
Caf1_331R	CTCCAAACTCGCCCTTCTCG	Msl3_1141F	TCTACATTGTTACCGCCAGC
Caf1_638F	CACACCCAAGGAGCATAGGG		

2.2 Antibodies

Table 7. Antibodies and their corresponding dilutions used in this study.
IgG: immunoglobulin G; WB: western blot; IF: immunofluorescence; DSHB: Developmental Studies Hybridoma Bank; ChIP: chromatin immunoprecipitation; HRP: horseradish peroxidase.

	Specificity	**Usage**	**Host species / clonality**	**Source / reference**
Primary antibodies	Ash1 (aa 443-769)	1:3000 in WB	Rabbit polyclonal	Schmähling *et al*, 2018
	Ogt	1:5000 in WB	Rabbit polyclonal	H-300, Santa Cruz, sc-32921
	MRG15 (full-length)	1:3000 in WB	Rabbit polyclonal	Kusch *et al*, 2004
	E(z) (full-length)	1:7000 in WB	Rabbit polyclonal	Gambetta *et al*, 2009
	Caf1-55 (full-length)	1:50.000 in WB	Rabbit polyclonal	α-Nurf55 in Gambetta *et al*, 2009
	Spt5	1:50.000 in WB	Guinea pig polyclonal	Saunders *et al*, 2003
	H3K36me2	1:250 in WB	Rabbit monoclonal	C75H12 / Cell Signaling
	H3K36me3	1:750 in WB in Figure 11	Rabbit monoclonal	D5A7 / Cell Signaling
	H3K36me3	1:1000 in WB in Figure 16	Rabbit polyclonal	ab9050 / Abcam
	H4	1:200.000 in WB	Rabbit polyclonal	ab10158 / Abcam
	Ubx	1:30 in IF	Mouse monoclonal	FP3.38 / DSHB
	Abd-B	1:200 in IF	Mouse monoclonal	1A2E9 / DSHB
	H3K36me2	ChIP (3 µl for 1 ml IP with chromatin from haltere and 3rd leg discs from 60 larvae)	Rabbit polyclonal	ab9049 / Abcam
Secondary antibodies	Rabbit IgG	1:5000 in WB	Donkey polyclonal	HRP anti-rabbit IgG / Amersham Biosc. (NA934)
	Guinea pig IgG	1:5000 in WB	Goat polyclonal	HRP anti-guinea pig IgG / Santa Cruz (sc-2438)
	Mouse IgG	1:500 in IF	Goat polyclonal	Cy3 anti-mouse IgG / Jackson ImmunoResearch

2.3 *Drosophila* genetics

2.3.1 Fly husbandry

Fly stocks and crosses were cultivated in *Drosophila* breeding vials or bottles on standard agar medium composed of cornmeal, soy flour, yeast and treacle supplemented with phosphoric acid and methyl paraben as fungicides. The incubation conditions were set to 65% relative humidity and a temperature of either 25°C or 18°C depending on how fast progeny needed to be obtained.

For mass production of embryos to generate large-scale nuclear extracts for tandem affinity purifications (TAP) (para 2.5.1.1), populations of fly strains carrying *ash1* transgenes were raised in breeding bottles and transferred to large-sized cages for the period of embryo collection. To obtain a sufficient amount of the transgenic embryos, 12 cages were necessary, each filled with 25 – 30 g of adult flies. Embryos were collected from food plates. This food in petri dishes was made of agar medium supplemented with apple juice, treacle and methyl paraben and, when dry, a portion of yeast paste added on top. At the end of embryo collection, adult flies were famished and a new population was started by expanding the transgenic strains in breeding bottles. Wild-type embryos (*Oregon-R*) were collected in the laboratory of Peter Becker (BMC, LMU, Munich).

2.3.2 Transformation of DNA constructs into *Drosophila*

The DNA sequences to be integrated into the *Drosophila* genome from the TAP-constructs (pCaSpeR-*tub*-NTAP-ash1 and pCaSpeR-*tub*-ash1-CTAP; para 2.1.8) and the pW35-MRG15$^{\Delta}$-construct (para 2.1.10) were *P* elements. All three constructs were injected into the germ line of wild-type *Drosophila melanogaster* (*w*^{1118}) by standard methods, where the *P* elements were incorporated into the genome as originally described in Rubin & Spradling, 1982. Fly stocks not showing an unspecific phenotype due to the *P* element insertion site were established and the *P* elements were mapped.

The pUMR-FLAP-ash1^{R1464A}- and the pUMR-FLAP-ash1^{wt}-construct (para 2.1.9) as well as the BAC CH322-160G6 (Table 2) carried *attB* sites for integration into *attachment sites on phage DNA* (*attP* landing sites) by PhiC31 integrase-mediated site directed transgenesis (Groth *et al*, 2004). The constructs were injected into the germ line of *Drosophila melanogaster*, pUMR-FLAP-ash1^{R1464A} and pUMR-FLAP-ash1^{wt} into a strain containing the landing site *VK37* (Venken *et al*, 2006), CH322-160G6 into a strain with the landing site *ZH-51C* (Bischof *et al*, 2007). See Table 8 for the genotypes of the landing site strains.

2.3.3 Genotypes of *Drosophila* strains used in this study

Table 8. List of genotypes of *Drosophila* strains used in this study. As indicated, some strains were generated in the context of this work, others obtained from outside sources.

Genotype	Usage	Source / Reference
Oregon-R	control for TAP	-
w ; + ; + (w^{1118})	injection; control in Figure 11 - 22	-
w ; if-1/CyO	Establishment of transgenic stocks	-
y^1w ; + ; Dr^1/TM6C		-
y^1w ; sp^1/CyO ; TM2/TM6B		-
y^1w ; + ; Dr^1/TM3, twi::EGFP		-
w ; + ; $ash1^{22}$, FRT2A/ TM3, twi::EGFP	Ash1 mutant analysis	Tripoulas *et al*, 1994
w ; P{α-tub1-NTAP-ash1} ; $ash1^{22}$, FRT2A	'NTAP-Ash1' and 'Ash1-CTAP' strains for TAP	Schmähling *et al*, 2018
w ; P{α-tub1-ash1-CTAP}/CyO ; $ash1^{22}$, FRT2A		
w hsp70-flp ; + ; P{ovoD1-18}3L, P{FRT(w^{hs})}2A/TM2/TM6B	Generation ash1 maternal minus mutants	Norbert Perrimon, Harvard Medical School
w ; + ; $ash1^{10}$/TM6C	Ash1 catalytic mutant analysis	Tripoulas *et al*, 1994
y^1 P{nos-phiC31\int.NLS}X ; PBac{y[+]-attP-3B}VK00037	landing site strain	Venken *et al*, 2006
w ; $ash1^{R1464A}$ (VK37) ; $ash1^{22}$, FRT2A/TM6B	Ash1 catalytic mutant analysis	Schmähling *et al*, 2018
w; $ash1^{wt}$(VK37) ; $ash1^{22}$, FRT2A		
y^1 w ; + ; P{lacW}$MRG15^{j6A3}$/TM3, Sb^1	MRG15 mutant analysis	Spradling *et al*, 1999
y^1 w/Dp(2;Y)G, P{hs-hid}Y ; P{70FLP}11, P{70I-SceI}2B, sna^{Sco}/CyO, P{hs-hid}4	Generation MRG15 deletion allele	Yang Hong, University of Pittsburgh School of Medicine
w ; P{$MRG15^{\Delta}$} ; $MRG15^{+}$	*P[donor]* strain	Schmähling, unpublished
w ; + ; $MRG15^{\Delta}$/TM6C	MRG15 deletion	Schmähling *et al*, 2018
w ; + ; Df(3R)BSC741/TM3, twi::EGFP	complementation test; MRG15 mutant analysis	Kevin Cook, Bloomington Drosophila Stock Center
y^1 M{vas-int.Dm}ZH-2A w^{1118} ; M{3xP3-RFP.attP'}ZH-51C	landing site strain	Bischof *et al*, 2007
w ; CH322-160G6 (ZH-51C)/CyO	Rescue of *$MRG15^{\Delta}$*	Schmähling, unpublished

2.3.4 Crosses performed to obtain the analyzed genotypes

The *ash1*-null mutants devoid of Ash1 maternal load are called *ash1*$^{22\ m-\ z-}$ in this work and had the genotype *w ; + ; ash1*$^{22\ m-\ z-}$*, FRT2A* (*FRT*: *flippase recognition target site*). In the first step to obtain these mutants, *w ; + ; ash1*^{22}*, FRT2A/ovo*^{D}*, FRT2A* females with germ line clones homozygous for *ash1*^{22}*, FRT2A* were generated applying the autosomal FLP-DFS (flippase-dominant female sterile) technique described in Chou & Perrimon, 1996. These *w ; + ; ash1*^{22}*, FRT2A/ovo*^{D}*, FRT2A* females were then crossed with *w ; + ; ash1*^{22}*, FRT2A/TM3, twi::EGFP* males. Part of the progeny were the wanted *w ; + ; ash1*$^{22\ m-\ z-}$*, FRT2A* mutants, that could be distinguished and separated from the other viable genotype in the progeny (*w ; + ; ash1*$^{22\ m-}$*, FRT2A/ TM3, twi::EGFP*) by lack of GFP expression to perform phenotypic analyses.

ash1-catalytically inactive mutants devoid of maternal Ash1 are named *ash1*$^{R1464A\ m-z-}$ in this work and had the genotype *w ; ash1*$^{R1464A\ m-\ z-}$ *(VK37)/+ ; ash1*^{22}*, FRT2A*. They were generated by crossing *w ; ash1*$^{R1464A\ m+\ z-}$ *(VK37) ; ash1*^{22}*, FRT2A* females with *w ; + ; ash1*^{22}*, FRT2A/TM3, twi::EGFP* males.

The *ash1*$^{10\ m-\ z-}$ flies had the genotype w ; + ; *ash1*$^{10\ m-\ z-}$*/ash1*^{22}*, FRT2A* and were the progeny of w ; + ; *ash1*$^{10\ m+\ z-}$*/ash1*^{22}*, FRT2A* mothers and *w ; + ; ash1*^{10}*/TM6C* fathers.

The genotype of the *MRG15* mutant flies called *MRG15*$^{j6A3\ m+\ z-}$ was *w ; + ; P{lacW}MRG15*$^{j6A3\ m+\ z-}$*/Df(3R)BSC741* (*Df*: deficiency allele) generated by crossing flies from the strains *w ; + ; P{lacW}MRG15*^{j6A3}*/TM3, Sb*^{1} and *w ; + ; Df(3R)BSC741/TM3, twi::EGFP* with each other.

The *MRG15*-null mutant flies *MRG15*$^{\Delta\ m+\ z-}$ had the genotype *w; + ; MRG15*$^{\Delta\ m+\ z-}$*/Df(3R)BSC741* and were derived from crosses between flies from the strains *w ; + ; MRG15*$^{\Delta}$*/TM6C* and *w ; + ; Df(3R)BSC741/TM6C*. In order to obtain *MRG15*-null mutants devoid of MRG15 maternal load, *w ; + ; MRG15*$^{\Delta\ m+\ z-}$*/Df(3R)BSC741* females were crossed further with *w ; + ; Df(3R)BSC741/TM3, twi::EGFP* males.

2.3.5 Adult cuticle preparations for microscopy

Adult *Drosophila melanogaster* of the respective genotype were collected and stored in 70% ethanol. Prior to dissection, ethanol had to be replaced by soaking the carcasses in PBT overnight at 4°C (PBT: 0,2% (v/v) 'TWEEN 20' from Sigma-Aldrich in phosphate-buffered saline (PBS)). On the next day, cuticles of the thoracic and abdominal segments were dissected as cleanly as possible in PBT in a glass well. As mounting medium, a 1:1 mixture of lactic acid and Hoyer's solution (30 g gum arabic and 200 g chloral hydrate in 28% (v/v) glycerol) had been prepared. Each dissected cuticle was transferred to a drop of mounting medium on a microscope

slide and covered with a cover slip. To flatten the mounted tissue, a weight was put on the cover slip and the microscopy sample was kept at room temperature overnight without moving. Remaining tissue that was not part of the cuticle was dissolved by the mounting medium.

Microscopic analysis of the adult cuticles was performed with an 'Axio Scope.A1' microscope. The pictures were taken with bright-field illumination and Nomarski contrast settings under the '10x/0.3 Ph1 EC Plan-Neofluar' objective with the attached camera 'AxioCam MRm' (microscope, objective and camera from Zeiss).

2.3.6 Immunofluorescent staining of larval tissues

Third instar larvae were bisected in ice-cold PBS. The anterior parts were inverted; guts and fat bodies were removed and the carcasses with imaginal discs respectively brains still attached were transferred to an Eppendorf tube. For fixation, carcasses were incubated in 4% (w/v) formaldehyde in PBS for 20 min. The fixation step as well as the washing and blocking steps described in the following were performed at room temperature with agitation. After fixation, carcasses were washed twice for 10 min in PBT and blocked for 1 hr in BBT (1% (w/v) bovine serum albumin (BSA) and 0,1% (v/v) 'Triton X-100' from Sigma-Aldrich in PBS). BBT was exchanged four times during blocking. The subsequent incubation with a primary antibody directed against Ubx (FP3.38, DSHB, diluted 1:30) or Abd-B (1A2E9, DSHB, diluted 1:200) took place in BBT overnight at 4°C and was followed by six 10 min-washes in BBT. Next, carcasses were incubated in parallel with the 1:500 diluted fluorescently-labeled secondary antibody 'Cy3 anti-mouse IgG' (Jackson ImmunoResearch) and 50 ng/ml Hoechst 33342 DNA stain in BBT overnight at 4°C and, as in all steps from here onwards, protected from light. After two 10 min-washes in BBT and four 10 min-washes in PBT, imaginal discs respectively brains were separated from the carcasses in PBT and mounted on microscope slides in Fluoromount-G (Southern Biotech). The immunofluorescent stainings were analyzed with the confocal laser scanning microscopes 'TCS SP8' from Leica (haltere and 3rd leg imaginal discs) or the 'LSM 780' from Zeiss (brains). Pictures were taken with the corresponding cameras of these microscopes under the 'HC PL APO CS2 20x/0.75 IMM' objective (haltere and 3rd leg imaginal discs) or under the 'LD LCI PL APO 25x/0.8 IMM Korr DIC M27' objective (brains). All pictures shown in this work represent the average of four sequential scans in the same focal plane.

2.3.7 Generation of the *MRG15$^{\Delta}$* deletion allele

The fly transformation vector pW35 used to generate the *MRG15$^{\Delta}$* deletion allele is a *P* element vector and harbors *FRTs* and I-*Sce*I recognition sites as relevant elements for ends-out

targeting (Gong & Golic, 2003). The *FRT* and I-*Sce*I sites flank the homology arms that border in turn the *mini-white* marker gene.

First, the pW35-MRG15$^{\Delta}$ construct bearing sequences up- and downstream of the *MRG15* gene as homology arms (see para 2.1.10) was injected and integrated into the *Drosophila melanogaster* genome as a *P* element as outlined in para 2.3.2. The respective transgenic flies were kept as *P[donor]* strain. Then, flies trans-heterozygous on chromosome two for the *P* element and for a chromosome carrying the *hs-FLP* and the *hs-I-SceI* transgenes, that encode the flippase and the I-*Sce*I endonuclease under the heat-shock promoter of *hsp70*, were generated. Expression of flippase and I-*Sce*I was induced by three subsequent heat-shocks of the trans-heterozygous larvae for 1 hr at 37°C, 24, 48 and 72 hrs post egg laying, in order to excise the *P* element and enable recombination. When these candidate recombinants reached adult stage, they were individually crossed to *Df(3R)BSC741/TM6B* adults in a complementation test. *Df(3R)BSC741* is a deficiency allele uncovering the entire *MRG15* gene on chromosome three. Out of approximately 150 crosses, multiple independent targeting events were isolated, that manifested themselves in a Trithorax mutant phenotype similar to *ash1* mutants. One of these alleles, *MRG15$^{\Delta}$*, was chosen for thorough sequencing analysis. First, genomic DNA from *MRG15$^{\Delta}$* homozygotes, the *P[donor]* strain and the wild-type strain (*w^{1118}*) was isolated (para 2.3.8) and analyzed by PCR with the primer pairs listed in Table 9, followed by agarose gel electrophoresis. Primer pairs U and D were designed in a way that an amplification product could only be obtained in case of specific recombination and therefore correct disruption of *MRG15*. The PCR products generated with U and D on the genomic DNA of *MRG15$^{\Delta}$* were cloned into the pCR 2.1-TOPO TA vector and sequenced full-length with primers in Table 5. The primers specific for the pW35 vector were replaced by the commercial M13 forward and reverse primers (Invitrogen) specific for pCR 2.1-TOPO TA.

After the molecular analysis of the *MRG15$^{\Delta}$* allele, a genetic rescue assay was set up to test if the Trithorax mutant phenotype of the *MRG15$^{\Delta}$/Df(3R)BSC741* trans-heterozygotes is rescued by the BAC *CH322-160G6*, which encodes MRG15. To this end, *w ; +/CyO ; MRG15$^{\Delta}$/TM2* flies were crossed with *w ; CH322-160G6 (ZH-51C)/CyO ; Df(3R)BSC741/TM2* flies and the adult phenotype of the progeny *w ; CH322-160G6 (ZH-51C)/CyO ; MRG15$^{\Delta}$/Df(3R)BSC741* was compared with the phenotype of the progeny *w ; +/CyO ; MRG15$^{\Delta}$/Df(3R)BSC741*.

Table 9. Primer pairs used to verify correct disruption of *MRG15* in the *MRG15*$^{\Delta}$ allele

Name	Sequence (5' to 3')	Purpose
U_R	AGCGGTCGTGGATCAGATTTC	Verification of correct disruption upstream (U) of *MRG15* gene
U_F	AGCAACGAGAATAGAGTGCCG	
D_R	CGCTGCATGAATTAGCTTGGC	Verification of correct disruption downstream (D) of *MRG15* gene
D_F	TTTCTTGATCTGGACCTCGGC	
W_R	CATCGGTACTGTCCCATCCG	Double check of *mini-white* region (W)
W_F	GCCAAGCTAATTCATGCAGCG	
E_R	CGCGACTTGAACGGATTGTG	Negative control. Primer pair specific for endogenous (E) *MRG15* genomic locus.
E_F	GGCGGTGCTGTTAGTAGTCG	

2.3.8 Purification of genomic DNA from adult *Drosophila*

Fifty adult flies of the respective genotype were collected in 300 µl Solution A (0,1 M Tris-HCl pH9; 0,1 M EDTA pH8; 1% (w/v) sodium dodecyl sulfate (SDS)) and squashed manually with a micro-pestle (Eppendorf). To foster further homogenization and to denature proteins, the suspension was incubated at 70°C for 30 min. SDS was then removed by adding 70 µl of 8 M potassium acetate and incubation on ice for 30 min. In the next step, insoluble tissue particles were pelletized and discarded. DNA in the soluble fraction was precipitated by adding 0,5x volume of isopropanol and centrifugation at 21.000 x g for 5 min. The DNA pellet was washed with 70% ethanol and subsequently resuspended in 100 µl of 10 mM Tris-HCl pH8,5. RNA in the sample was then digested by adding 15 mU of Ribonuclease (Roche, #11119915001) and incubation for 5 min at 37°C.

To the end of further purifying the genomic DNA, phenol-chloroform extraction was performed using 2 ml MaXtract High Density Tubes (Qiagen) according to manufacturer's instructions. For DNA precipitation, 1 µl of 20 mg/ml glycogen, 0,1x volume of 3 M sodium acetate pH5,2 and 2,5x volumes of 100% ethanol were added to the aqueous phase of the phenol-chloroform extraction (1x volume) and mixed by inversion. The sample was then incubated for 30 min at -80°C and centrifuged at 21.000 x g and 4°C for 20 min. The obtained DNA pellet was washed with 70% ethanol. Lastly, the purified genomic DNA was taken up in 50 µl 10 mM Tris-HCl pH8,5.

2.4 Recombinant protein expression using baculoviruses

All recombinant proteins used in this work were expressed in insect cells via baculovirus infection. To this end, baculoviruses bearing the respective cDNAs were generated with the 'Bac-to-Bac Baculovirus Expression System' from Invitrogen. Table 10 lists all baculoviruses that have been produced in the course of this study.

Table 10. Baculoviruses generated in this study. Proteins marked as full-length lack methionine 1.

Ash1 viruses	Caf1-55 and MRG15 viruses	Esc and Msl3 viruses
Flag-Ash1$_{C\ (1041 – 2226)}$	6xHis-Caf1-55$_{full\text{-}length}$	StrepII-Esc$_{full\text{-}length}$
6xHis-Ash1$_{C\ (1041 – 2226)}$	StrepII-Caf1-55$_{full\text{-}length}$	StrepII-Msl3$_{full\text{-}length}$
6xHis-Ash1$_{C}^{RxRP}{}_{(1041 – 2226)}$	StrepII-MRG15$_{full\text{-}length}$	
6xHis-Ash1$_{C}^{R1464A}{}_{(1041 – 2226)}$		

2.4.1 Cultivation of insect cells

Two different insect cell lines were used for baculovirus-based protein expression dependent on the precise application: Sf21 cells for generation of viruses (IPLB-Sf21 AE, ovarian tissue, Spodoptera frugiperda, Invitrogen, Cat no.12682-019) and High Five cells for protein expression (BTI-TN-5B1-4 ovarian tissue, Trichoplusia ni, Invitrogen, High Five Frozen cells, P/N 51-4005). Sf21 cells were cultivated in EX-CELL TiterHigh animal-component free medium (Sigma-Aldrich) at a density of 7 – 10 x 10^5 cells/ml; High Five cells were grown in Express Five SFM (1x) serum free medium (Gibco, Life Technologies) supplemented with 18 mM L-glutamine at a density of 4 – 8 x 10^5 cells/ml. Both cell lines were maintained as 500 ml suspension cultures in '3 l disposable polycarbonate Erlenmeyer flasks' (Corning) with vent caps at 27°C at 90 rpm. Cell counts, viability and morphology were regularly monitored with the cell viability analyzer 'Vi-cell CR' from Beckman coulter.

2.4.2 Virus generation with the 'Bac-to-Bac Baculovirus Expression System'

The 'Bac-to-Bac Baculovirus Expression System' (Invitrogen; Ciccarone *et al*, 1998) comprises several steps from generation of recombinant bacmids by site-specific transposition in *E. coli* to bacmid transfection into insect cells and production of the first passage of recombinant baculoviruses.

2.4.2.1 Transposition of cDNA from pFastBac to bacmid

The cDNA from which the respective protein of interest was going to be expressed was cloned into the pFastBac1 vector that was coding for the chosen protein tag in its multiple cloning site as described in para 2.1.11.

Next, the cDNA was to be integrated into a baculovirus shuttle vector, the bacmid, by transformation of the cloned pFastBac1 construct into DH10EMBacY *E. coli* cells, which already contained the bacmid and a helper plasmid required for transposition. To this end, 100 µl of DH10EMBacY *E. coli* were thawn on ice; 2 µl of the pFastBac1 construct purified with the 'Plasmid Mini Kit' from Qiagen were added, gently mixed with the bacteria and incubated for 30 min on ice. Then, the transformation mixture was heat-shocked for 45 sec at 42°C and placed back on ice for 2 min. Next, the bacteria suspension was topped up with 900 µl of room temperature SOC (super optimal broth with catabolite repression) and incubated at 37°C for 3 - 4 hrs with agitation. In this time window, transposition took place. Afterwards, bacteria were spun down at 845 x g for 4 min, resuspended in 200 µl LB and plated in different amounts (e.g. 20 and 160 µl) on selective LB agar medium for blue-white screening (selective LB agar medium: 50 µg/ml kanamycin, 10 µg/ml tetracycline, 34 µg/ml chloramphenicol, 7 µg/ml gentamicin, 100 µg/ml 5-Bromo-3-indolyl β-D-galactopyranoside (Bluo-Gal), 40 µg/ml Isopropyl β-D-1-thiogalactopyranoside (IPTG)). The bacterial plates were incubated protected from light at 37°C for at least 36 hrs until blue and white colonies could be observed. For confirmation of candidate recombinants, several white colonies expected to contain the bacmid with the cDNA integrated were re-stroken on another selective plate along with a blue colony as control.

2.4.2.2 Bacmid purification

Three independent white colonies containing the recombinant bacmid were inoculated into 2,5 ml LB medium with antibiotics (50 µg/ml kanamycin, 10 µg/ml tetracycline, 34 µg/ml chloramphenicol, 7 µg/ml gentamicin) and incubated at 37°C for 20 hrs with agitation. The next day, bacteria only from cultures that had grown dense were processed further and pelletized at 21.000 x g for 3 min. Each bacterial pellet was resuspended in 300 µl P1 buffer (Plasmid Mini Kit, Qiagen). For cell lysis, 300 µl P2 buffer (Plasmid Mini Kit, Qiagen) were added, the suspension carefully inverted and incubated for 5 min at room temperature. Then, the lysate was neutralized by topping up with 300 µl P3 (Plasmid Mini Kit, Qiagen), again careful inversion of the suspension and incubation for 5 min on ice. Cellular debris were removed in a centrifugation step at 21.000 x g and 4°C for 10 min. The pellet was discarded, the supernatant with the bacmid DNA mixed with 800 µl isopropanol by inversion of the Eppendorf tube and incubated for 10 min on ice. The DNA

was precipitated by centrifuging at 21.000 x g and 4°C for 20 min. After discarding the supernatant, 70% ethanol was added and the sample centrifuged again at 21.000 x g and room temperature for 10 min. The pellet was air-dried at 37°C. Once the DNA appeared transparent, it was resuspended in 5 mM Tris-HCl pH 8,5 and stored at 4°C until transfection of Sf21 cells.

2.4.2.3 Transfection of Sf21 cells with bacmid

In preparation of transfection, SF21 cells were seeded into three independent wells in a 6-well plate per recombinant bacmid and into two additional wells that were going to serve as negative controls. More precisely, 8 x 10^5 SF21 cells in 2 ml EX-CELL TiterHigh medium were plated per well followed by incubation of the well plate for 30 – 60 min at 27°C without agitation to allow cells to settle and adhere. In the meantime, the two components of the transfection mixture, bacmid DNA and transfection reagent were prepared: First, for each transfection reaction separately, 1000 ng bacmid DNA (concentration after purification usually 1000 ng/1 µl) were diluted in 100 µl EX-CELL TiterHigh medium and gently mixed. Second, a master mix with 8 µl of Cellfectin II (Thermo Fisher) and 92 µl EX-CELL TiterHigh medium per transfection reaction was pipetted. 100 µl of this master mix were added to each bacmid dilution. This transfection mixtures were gently mixed and incubated for 15 – 30 min at 27°C. Next, the 200 µl transfection mixtures were pipetted dropwise to their allocated wells. To one negative control well only Cellfectin II and no DNA was added, to the second negative control only 200 µl EX-CELL TiterHigh medium. The SF21 cells were incubated in the transfection mixtures for 3 – 5 hrs at 27°C without agitation to allow uptake of the bacmid DNA. Then, the medium in the wells was replaced with fresh EX-CELL TiterHigh medium. In a further incubation step for 96 hrs at 27°C, the first virus passage, P1, was produced. Once the expected cytopathic effect was observed in the Sf21 cells, the supernatant containing P1 was harvested and stored in falcon tubes at 4°C until further amplification.

2.4.3 Virus amplification

For the purpose of obtaining a baculoviral stock of high titer and volume, all P1 stocks generated from one recombinant bacmid were passaged further.

50 ml of free-floating 4 x 10^5 Sf21 cells in an Erlenmeyer flask were infected with 2,5 ml of one P1 stock and incubated for 96 hrs at 27°C and 90 rpm. The cell suspension was then centrifuged at 930 x g for 15 min. The supernatants containing the P2 viruses were harvested. On the base of the pellet size, it was estimated which supernatant most likely had the highest viral titer. Only this supernatant was passaged further and used to infect 250 - 500 ml free-floating 4 x

10^5 Sf21 cells in a 1:100 volume ratio. This cell suspension was incubated for 72 hrs at 27°C and 90 rpm and then centrifuged in 500 ml centrifuge flasks from Corning at 930 x g for 15 min; the supernatant representing the P3 viral stock was collected and stored at 4°C. The P3 stock was the standard passage applied for protein expression.

2.4.4 Recombinant protein expression

For preparative protein expression, 500 ml free-floating High Five cells at a density of 8 x 10^5 cells/ml were infected with one up to three different P3 viruses simultaneously for coexpression and incubated for 72 hrs at 27°C and 90 rpm. Then, the cell suspension was centrifuged at 930 x g for 15 min, the supernatant was discarded and the cell pellet with the recombinant proteins snap-frozen in liquid nitrogen and stored at -80°C until protein purification.

If a new P3 virus was used or if specific combinations of recombinant proteins were coexpressed for the first time, small scale P3 titration tests were set up prior to preparative protein expression. In these titration tests, cells were infected with different virus dilutions and different combinations of virus dilutions. It was then tested by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie Blue staining which dilution led to satisfying expression levels and which combination resulted in more or less equal expression levels among coexpressed proteins.

2.5 Cell extract preparation and protein purification

2.5.1 Purification of Ash1 complexes from Drosophila

2.5.1.1 Large-scale nuclear extract preparation from Drosophila embryos

As starting material for nuclear extract preparation for TAP (para 2.5.1.2), 10 – 14 hrs old embryos were collected from a population of wild-type, NTAP-Ash1 or Ash1-CTAP flies (para 2.3.1) on several successive days to accumulate in total at least 40 g (dry weight dechorionated embryos). 40 g of embryos suffice for approximately three TAPs. Until collection was completed, already taken food plates with embryos were stored up to three days at 4°C to arrest development. In the case of the Ash1-CTAP flies, embryos homo- and heterozygous for the transgene might have been mixed. The chromosome bearing the Ash1-CTAP transgene is homozygous lethal at an undetermined developmental stage. In the next step, all embryos together were transferred from the food plates to a three sieves system with decreasing pore sizes (0,75 µm > 0,375 µm > 0,125 µm). After thorough washes with cold tap water in this sieves

system, embryos were transferred to a beaker with 20% bleach in embryo wash solution (0,7% NaCl; 0,04% 'TWEEN 20' from Sigma-Aldrich) and incubated for 3 min for dechorionation. Bleach was washed off from the embryos with cold tap water in the 0,125 µm sieve. Then, embryos were put on a 150 mm Millipore filter in a Büchner funnel on top of a vacuum device. All liquid was removed by applying vacuum so that the dry weight of the embryos could be determined.

The following steps of the nuclear extract preparation were all carried out at 4°C. The dechorionated embryos were transferred to a glass dounce homogenizer (LSC LH-21 homogenizer, Yamato) and topped up with 1 ml buffer NU1 per 1 g embryos (Buffer NU1: 15 mM HEPES pH 7,6; 10 mM KCl; 5 mM $MgCl_2$; 0,1 mM EDTA pH 7,9; 0,5 mM triethylene glycol diamine tetraacetic acid (EGTA) pH 7,9; 20% glycerol; 350 mM sucrose; 1 mM dithiothreitol (DTT); 1 mM 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF); 1x complete ('complete, EDTA–free Protease Inhibitor Cocktail') from Roche). Homogenization was performed by moving the mortar 30 times up and down during a pestle rotation of 1500 – 2000 rpm. The obtained lysate was filtered through two layers of miracloth. Next, for the purpose of isolating the nuclei, the lysate was centrifuged at 9600 x g and 4°C for 30 min. The supernatant was discarded and the pelletized nuclei were resuspended in 0,5 ml low salt buffer per 1 g embryos (low salt buffer: 15 mM 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES) pH 7,6; 20 mM KCl; 1,5 mM $MgCl_2$; 0,2 mM EDTA pH 7,9; 20% glycerol; 1 mM DTT; 1 mM AEBSF; 1x complete).

The nuclei suspension (1x volume) was transferred to falcon tubes and mixed with 1x volume of high salt buffer (15 mM HEPES pH 7,6; 800 mM KCl; 1,5 mM $MgCl_2$; 0,2 mM EDTA pH 7,9; 20% glycerol; 1 mM DTT; 1 mM AEBSF; 1x complete). In the following incubation step on a turning wheel for 30 min at 4°C, nuclei were lysed homogenously. Then, the soluble nuclear extract was separated from insoluble chromatin and lipids by ultracentrifugation at 182.000 x g and 4°C for 1 hour. The middle fraction out of three fractions representing soluble nuclear extract was collected and dialyzed overnight in 'Snake Skin Dialysis Tubing' with a 3,5 kDa cut-off (Thermo Fisher) against NE200 buffer (15 mM HEPES pH 7,6; 200 mM KCl; 1,5 mM $MgCl_2$; 0,2 mM EDTA pH 7,9; 20% glycerol; 1 mM DTT).

Protein concentration in the ready nuclear extracts after dialysis ranged from 10 – 15 mg/ml. This meant that 7 – 11 mg of nuclear protein could be obtained per 1 g dechorionated embryos. Extract quality was tested by probing for trxG and PcG proteins in western blots. The nuclear extracts were aliquoted, snap-frozen in liquid nitrogen and stored at -80 C until TAP.

2.5.1.2 Tandem affinity purification

Prior to TAP, 'IgG Sepharose 6 Fast Flow affinity resin' (GE Healthcare) was cross-linked with dimethyl pimelimidate. 200 µl of cross-linked beads were transferred into a 10 ml 'Poly-Prep Chromatography Column' from Bio-Rad. To remove the bead storage buffer and non-cross-linked immunoglobulins G, beads were washed in four consecutive steps: First, in 1 ml 0,5 M acetic acid, second, in 5 ml PA buffer, third, in 1 ml 0,5 M acetic acid and fourth, in 5 ml PA buffer (10 mM Tris-HCl pH 8; 150 mM NaCl; 0,1% 'IGEPAL CA-630' from Merck; 2 mM $MgCl_2$; 0,1 mM EDTA; 0,5 mM DTT). Next, beads were equilibrated in 10 ml PA buffer for 10 – 30 min on a turning wheel. TAP was always performed in parallel from the nuclear extracts of transgenic NTAP-Ash1 or Ash1-CTAP flies and from the nuclear extract of the wild-type control (para 2.5.1.1). The protein concentrations of both nuclear extracts, that ranged between 10 – 15 mg/ml, were adjusted to approximately the same concentration. Then, 10 ml of each extract were loaded on a column with equilibrated IgG beads. The following binding and wash steps of the TAP were performed as described in Klymenko *et al*, 2006. For elution, Calmodulin beads ('Calmodulin Affinity Resin', Agilent) were transferred from the columns to Eppendorf tubes. In total, three elution steps were performed. In each step, Calmodulin beads were incubated with 200 µl of the EGTA-containing buffer CE for 30 min at 4°C in a thermoshaker at 1000 rpm and then centrifuged for 10 min at 3000 x g and 4°C. Supernatants representing the eluates were taken off and stored unpooled until further processing by SDS-PAGE, silver staining (para 2.7.1 and 2.7.2) and mass spectrometry (para 2.7.4). CE buffer was composed of 10 mM Tris-HCl pH 8; 150 mM NaCl; 0,1% 'IGEPAL CA-630'; 1 mM $MgCl_2$; 2 mM EGTA; 1 mM imidazole pH 8 and 10 mM β-mercaptoethanol.

2.5.2 Larval tissue extract preparation for western blot analysis

Larval cuticles with imaginal discs attached were dissected out of 3rd instar larvae of a given genotype in ice-cold PBS supplemented with protease inhibitors (0,5x complete) as described in para 2.3.6. The wing, haltere and 3rd leg imaginal discs and, when indicated in the figure legend, also brains were separated from the carcasses, taken up in 0,1% 'IGEPAL CA-630' in PBS and transferred into a siliconized Eppendorf tube. The supernatant of the dissected tissues, which had settled at the tube bottom after short centrifugation at 500 x g, was replaced with 100 µl 1,5x 'NuPAGE LDS Sample Buffer' (Thermo Fisher) per discs from 30 carcasses. Next, the sample was sonicated in the 'Bioruptor Next Gen' (Diagenode, sonication program: 6x 30 sec on and 30 sec off at high-energy settings) and subsequently incubated for 3 min at 75°C. Tissue debris were then pelletized in a centrifugation step for 10 min at 21.000 x g. The supernatant representing the

total cellular extract was further analyzed by SDS-PAGE and western blot (para 2.7.1 and 2.7.5) for either Ash1, MRG15 or H3K36me2.

2.5.3 Purification of recombinant Ash1 complexes from insect cells

2.5.3.1 Insect cell extract preparation by freeze-thaw cycling

Prior to Strep- or His-affinity purification, insect cells harvested after baculovirus-based recombinant protein expression as described in para 2.4.4 were lysed in freeze-thaw cycles. To this end, 20 ml lysis buffer were pipetted on a frozen cell pellet from 500 ml High Five cell suspension culture at a density of 0,8 x 10^6 cells/ml. Volumes and cell numbers given here correspond to preparative cell extractions and protein purifications to perform biochemical assays afterwards. The lysis buffer in the case of subsequent Strep-affinity purification, Strep-Buffer A, was composed of 20 mM Tris-HCl, pH 8; 300 mM KCl; 2 mM $MgCl_2$; 15% glycerol; 10 µM $ZnSO_4$; 0,1% 'IGEPAL CA-630' (Merck); 1 mM DTT; 1 mM AEBSF and 1x complete; the lysis buffer in the case of subsequent His-affinity purification, His-Buffer A, was composed of 20 mM Tris-HCl, pH 8; 300 mM KCl; 4 mM $MgCl_2$; 5 mM imidazole, pH 8; 5% glycerol; 10 µM $ZnSO_4$; 0,05% 'IGEPAL CA-630'; 4 mM β-mercaptoethanol; 1 mM AEBSF and 1x complete. The suspension from cells and lysis buffer in a falcon tube was thawn completely in a water bath at room temperature with agitation on a magnetic stirrer and then snap-frozen in liquid nitrogen. This procedure was repeated three times. Thereafter, the cell lysate was centrifuged at 75.600 x g for 1 hour. The supernatant was collected, cleared further by filtering in 'Millex-SV 5 µm' sterile filters (Merck) and subjected to protein purification (para 2.5.3.2 and 2.5.3.3).

2.5.3.2 Strep-Tactin affinity purification

In preparation of Strep-affinity purification, 'Strep-Tactin Sepharose' beads (IBA) were washed and equilibrated with 20 column volumes (CV) Strep-Buffer A (see para 2.5.3.1 for composition). Then, 2,5 ml beads in a 'Glass Econo-Column' (Bio-Rad) with a three-way cock were loaded with cleared cell lysate generated from 500 ml High Five cell culture (0,8 x 10^6 cells/ml) as described in para 2.5.3.1. The binding step for Strep-tagged proteins was performed at a low flow-rate (<0,5 ml/min) by releasing the lysate by gravity flow. Also, all subsequent steps were executed by gravity flow. After binding, beads were first washed with 20 CV Strep-Buffer B (20 mM Tris-HCl, pH 8; 300 mM KCl; 15% glycerol; 0.1% 'IGEPAL CA-630' (Merck); 1 mM DTT; 1 mM AEBSF; 1x complete) and then with 30 CV Strep-Buffer C (20 mM Tris-HCl, pH 8; 300 mM NaCl; 15% glycerol; 1 mM DTT; 1 mM AEBSF). Finally, bound protein was eluted with d-desthiobiotin in

three subsequent steps, each step with 1 CV Strep-Buffer D (20 mM Tris-HCl, pH 8; 150 mM NaCl; 10% glycerol; 0,5 mM DTT; 1 mM AEBSF; 1x complete and 5 mM d-desthiobiotin-NaOH, pH 8).

The three eluates were not pooled before controlling purity and concentration of the purified material by SDS-PAGE and Coomassie Blue staining. Typically, eluates from Strep-affinity purification needed to be concentrated in centrifugal filter units 'Amicon Ultra-4', 10 kDa cut-off (Merck) before proceeding to SDS-PAGE for presentation of purified complexes or histone methyltransferase assays (para 2.6.3 and 2.7.1).

2.5.3.3 His-affinity purification

For the purpose of purifying 6xHis-tagged proteins, gravity flow chromatography with 'Ni-NTA Agarose' (Qiagen) as matrix was performed at 4°C. Per cell lysate generated from 500 ml High Five cell culture (0,8 x 10^6 cells/ml) as described in para 2.5.3.1, 1 ml Ni-NTA beads was used. Beads were transferred into a 'Glass Econo-Column' from Bio-Rad, washed with ddH_2O and then equilibrated with His-Buffer A (see para 2.5.3.1 for composition). In the subsequent binding step, cleared cell lysate was loaded on the beads and released by gravity flow. Next, two wash steps were performed, one with 20 column volumes (CV) His-Buffer B (20 mM Tris-HCl pH 8; 300 mM KCl; 10 mM imidazole pH 8; 5% glycerol; 0,05% 'IGEPAL CA-630' (Merck); 4 mM β-mercaptoethanol; 1 mM AEBSF; 1x complete) and the following wash step with 30 CV His-Buffer C (20 mM Tris-HCl pH 8; 300 mM NaCl; 20 mM imidazole pH 8; 5% glycerol; 0,05% 'IGEPAL CA-630'; 4 mM β-mercaptoethanol; 1 mM AEBSF; 0,5x complete). After washing, bound protein was eluted with 250 mM imidazole in three steps, each step in 1,3 ml (1,3 CV) His-Buffer D (20 mM Tris-HCl pH 8; 300 mM NaCl; 250 mM imidazole pH 8; 5% glycerol; 0,05% 'IGEPAL CA-630'; 4 mM β-mercaptoethanol; 1 mM AEBSF; 1x complete).

Eluates were not pooled, but checked separately for yield of purified protein by SDS-PAGE and Coomassie Blue staining (para 2.7.1 and 2.7.3). Without further concentration, SDS-PAGE for presentation of purified complexes could be performed.

2.6 Histone methyltransferase assays on mononucleosomes

HMTase assays were performed with different Ash1 complexes in comparison on recombinant *Drosophila* mononucleosomes. The protocol for generation of these mononucleosomes, from expression and purification of recombinant histones to reconstitution and purification of histone octamers and lastly the nucleosome assembly, was modified from Luger *et al*, 1999 and is described in more detail in the following.

2.6.1 Reconstitution of *Drosophila* octamers

Archana Prusty, a former PhD student in our laboratory, had expressed wild-type *Drosophila* histones in BL21(DE3) PlysS *E. coli*, had purified them through a 'HiTrap SP Sepharose FF' ion exchange column (GE Healthcare) and had lyophilized histones for long-term storage.

In the first step of the octamer reconstitution, I dissolved lyophilized histones by incubating 2 mg of each core histone in 500 µl unfolding buffer (20 mM Tris-HCl pH 7,5; 7 M guanidine-HCl; 10 mM DTT) for 1 hr at room temperature with gentle agitation. Then, all four different core histones were mixed and incubated for another 45 min at room temperature. For octamer formation, the histone mixture was transferred into a 3 ml 'Slide-A-Lyzer Dialysis Cassette' with a 3,5 kDa cut-off (Thermo Fisher) and dialyzed overnight against high salt-refolding buffer, which was exchanged three times during dialysis (refolding buffer: 10 mM Tris-HCl pH 7,5; 2 M NaCl; 1 mM EDTA pH 8; 5 mM β-mercaptoethanol). Next, newly reconstituted octamers were spun at 9.400 x g and 4°C for 10 min to remove possible precipitates. The supernatant was loaded onto and pumped through a 'Superdex 200 10/300 GL' size exclusion column (GE Healthcare), that had been equilibrated in Refolding Buffer beforehand. 400 µl fractions were collected and analyzed by SDS-PAGE and Coomassie Blue staining (para 2.7.1 and 2.7.3). Fractions showing equal amounts of all four core histones in Coomassie Blue staining were considered to be homogenous, i.e. to contain almost exclusively octamers, and stored at 4°C for maximal two months until nucleosome assembly.

2.6.2 Assembly of *Drosophila* mononucleosomes

First, a 215 base pair(bp)-long DNA fragment bearing the nucleosome positioning 601 sequence (Lowary & Widom, 1998) was amplified from a pUC19-601 construct, that had been cloned by Reinhard Kalb, a former postdoc in our laboratory. In the chosen amplicon, the 147 bp-nucleosome binding sequence was flanked by a 40 bp-long DNA overhang on the 5' site and a 28 bp-long DNA overhang on the 3' side. To generate preparative amounts of the 215 bp-601 DNA fragment, 96x 50 µl PCR reactions with 30 ng template were performed in a well plate following standard protocols (para 2.1.1.). After PCR product purification with the 'QIAquick PCR Purification Kit' from Qiagen, DNA was concentrated by sodium acetate/ethanol precipitation: 0,1x sample volume of sodium acetate (pH 5,2) and 2,5x volume of ethanol were added to the DNA. After centrifugation for 40 min at 21.000 x g and 15°C, DNA pellets were washed with ethanol and then resuspended in seven times less volume 10 mM Tris-HCl (pH 8,5) than the DNA solution input volume of the acetate/ethanol precipitation.

Assembly of nucleosomes from 215 bp-601 DNA and recombinant octamers generated as described in para 2.6.1 was carried out by diluting the salt content slowly in gradient dialysis. First, small-scale nucleosome test reconstitutions with varying mass ratios of octamers to DNA were performed and tested in native agarose gel electrophoresis (para 2.1.2) to find conditions for homogenous nucleosome preparation with ideally all octamers and all DNA engaged in correctly assembled mononucleosomes. The mass concentration of the octamers was here determined presuming that an optical density of 0,42, when measuring absorbance at 280 nm, indicates a *Drosophila* octamer concentration of 1mg/ml. Typically, the test reconstitutions showed that a mass ratio of octamers to the 215 bp-601 DNA of 0,85 leads to a homogenous preparation. The actual nucleosome assembly reaction was then prepared with the following composition corresponding to Refolding Buffer conditions: 1x TE (10 mM Tris-HCl, 1 mM EDTA, pH 8); 0,3 mg/ml DNA; 2 M NaCl; 0,26 mg/ml octamers. The nucleosome assembly reaction was transferred to a dialysis tube from the 'Mini Dialysis Kit, 1 kDa cut-off' (GE Healthcare) and placed in a beaker with Refolding Buffer on a magnetic stirrer at 4°C. Applying a peristaltic pump system, the Refolding Buffer was slowly diluted with 1x TE buffer at a flow rate of maximal 1,5 ml/min. After five times dilution, when it could be assumed that salt concentration in the assembly reaction had decreased from 2 M to 0,4 M and nucleosomes had formed, dialysis tubes were transferred to 1x TE buffer at 4°C and dialysis was continued for further 4 hrs. After completion of dialysis, the quality of the obtained nucleosomes was controlled by native agarose gel electrophoresis (para 2.1.2). Mononucleosomes were stored at 4°C for maximal two weeks before use in HMTase assays.

2.6.3 Histone methyltransferase reactions

All Ash1 complexes tested for their catalytic activity in the HMTase assays described in this paragraph had been isolated by Strep-Tactin affinity purification (para 2.5.3.2), snap-frozen in aliquots in liquid nitrogen, stored at -80°C and thawn on ice only once, shortly before pipetting the HMTase reactions. To the end of estimating the concentration of the enzyme Ash1 in each Ash1 complex purification, one aliquot had been sacrificed for a dilution series that was compared with a BSA standard by SDS-PAGE and Coomassie staining (para 2.7.1, 2.7.3).

The HMTase reactions were performed with approximated 33 or 100 nM Ash1 complex and 80 µM *S*-adenosyl methionine on 400 nM reconstituted *Drosophila* mononucleosomes (para 2.6.2) in a reaction buffer containing 65 mM Tris-HCl pH 8,5 at 25°C; 78 mM NaCl; 2,5 mM $MgCl_2$; 0,23 mM EDTA pH 8; 1 mM DTT; 1 mM β-mercaptoethanol; 2,6 mM d-desthiobiotin; 5% glycerol; 0,5 mM AEBSF and 0,5x complete. The samples were incubated for 3 hrs at 25°C without

agitation. After 3 hrs, reactions were stopped by adding 13 µl of 4x 'NuPAGE LDS Sample Buffer' (Thermo Fisher) to 40 µl reaction volume and denatured for 10 min at 70°C.

As means of analysis of the nucleosome substrates, an antibody-based approach was chosen. The HMTase reactions were first separated by SDS-PAGE and then subjected to western blotting (para 2.7.1 and 2.7.5).

2.7 Protein analysis

2.7.1 SDS polyacrylamide gel electrophoresis

SDS polyacrylamide gel electrophoresis was performed to separate proteins in a size-dependent manner for downstream analysis. Two different electrophoresis buffer systems were applied in this work, the Tris-Glycine system and the Bis-Tris system. Tris-Glycine polyacrylamide gels were self-cast following standard protocols based on the original report Laemmli, 1970. Bis-Tris polyacrylamide gels were precast and purchased from Thermo Fisher either as 'NuPAGE Bis-Tris Gels' or as 'Bolt Bis-Tris Plus Gels'. The particular buffer system and gel acrylamide concentration were chosen (1) dependent on the size of the protein or the range of sizes of several proteins to be separated and visualized, (2) dependent on the required sharpness of protein bands and (3) dependent on the sample volume. Wells of Bolt gels are three-dimensional and therefore accommodate a higher sample volume than conventional gels. Which specific polyacrylamide gel was taken for which experiment is indicated in the figure legends. In all cases, samples were denatured by incubation for 10 min at 70°C in 1x 'NuPAGE LDS Sample Buffer' (Thermo Fisher) prior to loading. Gels were run in the 'XCell SureLock Mini-Cell Electrophoresis System' (Thermo Fisher) at 150 – 180 V. As running buffer for the Tris-Glycine system standard 1x Laemmli running buffer was used, as running buffer for the Bis-Tris system either MES SDS buffer in case of separation of smaller proteins or MOPS SDS buffer in case of separation of bigger proteins were chosen (both buffers from Thermo Fisher).

Following SDS-PAGE, proteins in the gel were either visualized by silver or Coomassie Blue stainings (para 2.7.2 and 2.7.3), digested in preparation of mass spectrometric analysis (para 2.7.4) or subjected to western blot assays (para 2.7.5).

2.7.2 Silver staining of tandem affinity purified proteins

10% of the same TAP eluate that was analyzed by mass spectrometry were separated by SDS-PAGE (previous paragraph) and silver-stained for illustration purposes.

All steps of the silver staining described in the following took place on a shaker at room temperature. First, the polyacrylamide gel was fixed overnight in fixing solution (40% ethanol, 10% acetic acid). Then, acetic acid was removed in three wash steps in 30% ethanol for 20 min each. Next, the gel was sensitized for staining by incubation for 1 min in freshly prepared 0,02% (w/v) sodium thiosulfate followed by three washes in deionized water for 20 sec each. Staining took place in pre-cooled 0,2% (w/v) silver nitrate for 1 hour. After three quick 20 sec-washes in water, the gel was developed in developing solution (3% (w/v) sodium carbonate; 0,05% formaldehyde; 0,0004% (w/v) sodium thiosulfate). Development was manually terminated when the obtained staining was considered to be sufficient (after 5 – 15 min development) by transferring the gel shortly into water, then into 0,5% glycine for 5 min. Lastly, the gel was washed in water for 30 min and stored in 1% acetic acid at 4°C.

2.7.3 Coomassie Blue staining

After SDS-PAGE, gels were washed three times for 10 min in deionized water to rinse out SDS and then stained in 'PageBlue Protein Staining Solution' (Thermo Fisher) for minimum 60 min to overnight with agitation. To destain, gels were washed for 5 min in water with agitation.

2.7.4 Mass spectrometric analysis of tandem affinity purified proteins

For mass spectrometric analysis, 90% of the most concentrated eluate out of three eluates of each TAP (para 2.5.1.2) were separated on a 4 - 12% Bis-Tris polyacrylamide gel and stained with Coomassie Blue (para 2.7.1 and 2.7.3). The entire gel lane was then excised and subdivided into ten smaller slices. From this step onwards, samples were processed by the Axel Imhof laboratory (BMC, LMU, Munich). Each gel slice was destained, alkylated with iodacetamide and digested with trypsin as described in Barth *et al*, 2014. The obtained peptides were then isolated from the gel matrix by acid extraction. In addition, 90% of a less concentrated eluate of the Ash1-CTAP purification were alkylated and digested directly in solution without previous separation by SDS-PAGE. From this point onwards, all samples were processed in the same manner. Each sample was concentrated by vacuum evaporation and taken up in 0.1% TFA. 50% of each of the digested TAP eluates was injected into an 'Ultimate 3000 HPLC system' from Thermo Fisher. Samples were then processed, loaded into and measured in the LTQ Orbitrap mass spectrometer (Thermo Fisher) as described in Barth *et al*, 2014. The resulting raw data were analyzed with the Andromeda algorithm of the MaxQuant protein analysis package (version 1.5.3.30) against the Flybase dmel-all-translation-r5.32.fasta database including reverse sequences and contaminants. The acquired protein identifications were accepted provided that at

least one unique peptide had been detected. Peptide intensities were then converted into Intensity based absolute quantification (iBAQ) values to obtain a quantitative measurement of the identified proteins (see Schwanhäusser *et al*, 2011 for details on this quantification method). For easier presentation, iBAQ values were log2-transformed and subsequently missing values were imputed from a random distribution centered at 1/3 x log2 of the experimental data.

Table 12 provides a list of the Ash1, MRG15 and Caf1-55 peptides identified in the Ash1-CTAP in-solution digest, the Ash1-CTAP in-gel digest and the NTAP-Ash1 in-gel digest.

2.7.5 Western blot

In all western blot experiments performed in this study, proteins separated by SDS-PAGE were transferred from polyacrylamide gels to membranes by wet electroblotting in Towbin transfer buffer in the 'TE 22 Mini Tank Transfer Unit' (GE Healthcare). The Towbin buffer was composed of 25 mM Tris; 192 mM glycine; 0,05% (w/v) SDS and 20% (v/v) methanol. For blotting histones, as done for the analysis of the HMTase assays (para 2.6.3), the general protocol had to be modified to also operate for small-sized proteins and for the anti-histone antibodies used. These specific modifications are pointed out in the following description.

After completion of the SDS-PAGE and in preparation of the transfer, the nitrocellulose membrane from GE Healthcare ('Hybond ECL', 0,45 µm; for histones 'Amersham Protran Premium', 0,2 µm) was pre-wetted shortly in ddH_2O. Then, the polyacrylamide gel and the membrane were soaked in Towbin Buffer for 10 min. Next, the transfer cassette with gel and membrane and then the transfer unit comprising the transfer cassette, the tank filled with Towbin buffer and the lid with electric leads were assembled according to the instructions of GE Healthcare, the manufacturer of the transfer unit. The transfer was performed at 4°C with agitation on a magnetic stirrer and commonly for 90 min at 90 V, for histones first for 10 min at 90 V followed by 30 min at 60 V.

After transfer, the membrane was shortly washed in PBS (for histones: in TBS) and blocked for 1 hr up to overnight in milk blocking solution (5% (w/v) skim milk (AppliChem, #A0830) and 0,2% (v/v) 'TWEEN 20' (Sigma-Aldrich) in PBS) at 4°C with gentle agitation. For histone blots, blocking was always performed overnight in BSA blocking solution (5% (w/v) BSA and 0,1% (v/v) 'TWEEN 20' in TBS). Then, the individual primary antibodies were diluted in milk respectively BSA blocking solution in the case of histone blots (see Table 7 for precise dilutions). Incubation of the membrane (parts) with the primary antibody(ies) took place overnight at 4°C with gentle agitation with the exception of the anti-H4 antibody (ab10158, Abcam). The anti-H4 antibody was added to anti-H3K36me2 or anti-H3K36me3 antibody solution (C75H12 and D5A7,

Cell Signaling) on the membrane on the next day for an incubation time of only 1 -2 hrs to avoid unspecific binding of this antibody. Incubation with the primary antibodies was followed by three wash steps, each for 10 min in 0,2% (v/v) 'TWEEN 20' in PBS with agitation. Histone blots were washed in 0,05% (v/v) 'TWEEN 20' in TBS. Incubation with the respective secondary antibody (see Table 7 for dilutions) was performed for 2 – 4 hrs in milk blocking solution with agitation at room temperature, in the case of histone blots in BSA blocking solution at 4°C. Then, the membrane was washed exactly as described for washes after the primary antibody incubation step.

For western blot development, the detection reagent 'Amersham ECL' (RPN2209) or 'Amersham ECL Select' (RPN2235) from GE Healthcare was pipetted on the membrane, the choice of the enhanced chemiluminescence (ECL) solution being dependent on the expected strength of bands. Immediately after application of the detection reagent, the emerging chemiluminescence on the membrane was visualized either by exposure to 'Amersham Hyperfilm ECL' (GE Healthcare) films with subsequent development in the X-ray film processor 'OPTIMAX 2010 NDT' (Protec) or by a charge-coupled device (CCD) camera in the imager 'ImageQuant LAS 4000' (GE Healthcare).

2.8 Chromatin immunoprecipitation followed by quantitative PCR

The chromatin immunoprecipitation assay coupled to qPCR was performed from haltere and 3rd leg imaginal discs of wild-type and *ash1*$^{22\ m+z-}$ 3rd instar larvae and directed against H3K36 di-methylation at the *Ubx* gene. The protocol for fixation of imaginal discs, preparation of chromatin and immunoprecipitation described in the following was modified from Orlando *et al*, 1997.

2.8.1 Dissection and fixation of haltere and third leg imaginal discs

Wild-type or *ash1*$^{22\ m+z-}$ larvae were processed in batches of 60 animals, accounting for one biological replicate, from dissection onwards.

Third instar larvae were bisected horizontally in ice-cold PBS. The anterior part was inverted, gut and fat body were removed. The remaining cuticle with discs still attached was transferred to an Eppendorf tube. Sixty of these carcasses were accumulated and then fixed altogether for exactly 15 min in 1 ml crosslinking solution (1% (w/v) methanol-free formaldehyde (Thermo Fisher, #28906); 50 mM Hepes pH 8; 100 mM NaCl; 1 mM EDTA; 0,5 mM EGTA). Crosslinking was performed at room temperature with gentle agitation as all following steps in

Stop solution and wash buffers as well. During the course of fixation, Crosslinking Solution was exchanged four times. Fixation was then terminated by replacing Crosslinking Solution with 1ml stop solution (125 mM glycine and 0,01% (v/v) 'Triton X-100' in PBS). Stop Solution was exchanged five times during a 10-min incubation. Two 10-min wash steps followed, the first one in Wash Buffer A (10 mM Hepes pH 8; 1 mM EDTA; 0,5 mM EGTA; 0,25% (v/v) 'Triton X-100') and the second one in Wash Buffer B (10 mM Hepes pH 8; 1 mM EDTA; 0,5 mM EGTA; 0,01% (v/v) 'Triton X-100'; 200 mM NaCl). In both washes, the respective buffer was renewed four times.

Fixed carcasses were stored in Wash Buffer B at 4°C for maximal three days. After separation of the imaginal discs from the carcasses in Wash Buffer B, haltere and 3rd leg imaginal discs from 60 larvae were taken up in 130 µl sonication buffer (10 mM Hepes pH 8; 1 mM EDTA; 0,5 mM EGTA; 0,1% (w/v) SDS; 1 mM AEBSF; 1x complete). The 130 µl disc-samples, each representing one biological replicate, were snap-frozen in liquid nitrogen and stored at -80°C until preparation of chromatin.

2.8.2 Chromatin preparation from imaginal discs and immunoprecipitation

From chromatin preparation onwards, all biological replicates (three wild-type and three *ash1*$^{22\ m+z-}$ batches), that were to be directly compared in this ChIP assay, were processed in parallel.

The discs in Sonication Buffer, dissected, fixed and frozen as described in para 2.8.1 were thawn on ice and sonicated one by one in 'AFA Fiber microTUBES' in the 'S220 Focused-ultrasonicator' from Covaris (sonication program: 105W/2%/30 min) at 4°C. After sonication, each 130 µl disc-sample respectively biological replicate was supplemented with 0,5% (v/v) N-lauroylsarcosine and incubated for 10 min at 4°C on a rotating wheel. Insoluble tissue debris were removed in a 10-min centrifugation step at 21.000 x g and 4°C. The supernatants with the soluble chromatin were transferred to dialysis tubes (one tube per replicate) from the 'Mini Dialysis Kit, 1 kDa cut-off' (GE Healthcare) and dialyzed overnight at 4°C against dialysis buffer (10 mM Tris-HCl pH 8; 1 mM EDTA; 0,5 mM EGTA; 4% glycerol). Dialysis buffer was changed twice in the first few hours of dialysis. Chromatin preparation was completed with centrifugation of the dialyzed chromatin at 21.000 x g and 4°C for 20 min to remove possible precipitates.

In a first step of the immunoprecipitation, 10 µl chromatin were saved from each replicate as input material that would be needed later on to determine the DNA amount before affinity purification. Until parallel processing of input and immunoprecipitation (IP) samples from Ribonuclease digest onwards, input material was stored at 4°C. The volume of the remaining major part of the chromatin was topped up from approximately 100 µl per replicate respectively

IP sample to 1 ml and thereby adjusted to RIPA-140 buffer conditions (10 mM Tris-HCl pH 8; 140 mM NaCl; 1 mM EDTA; 1% 'Triton X-100'; 0,1% SDS; 0,1% sodium deoxycholate; 1x complete; 1 mM AEBSF). For the purpose of pre-clearing the chromatin from components, that bind unspecifically to the purification matrix going to be used, 40 µl of a 50% slurry of 'Protein A Sepharose CL-4B' (GE Healthcare) were added to each IP sample. The Protein A sepharose beads had been equilibrated in RIPA-140 buffer on a rotating wheel at 4°C for 1,5 hrs beforehand. Until stated otherwise, all incubation and wash steps described in the following took also place on a rotating wheel at 4°C. The IP samples with beads were incubated for 1 hr and then centrifuged for 2 min at 1500 x g. The supernatants containing the pre-cleared chromatin were taken off and incubated with 3 µl anti-H3K36me2 antibody (ab9049, Abcam) per IP sample overnight. Next, 40 µl of 50% slurry of Protein A sepharose were added to each IP and incubation was continued for 4 hrs to capture the chromatin-antibody complexes. Samples were then centrifuged for 2 min at 1500 x g and the supernatants discarded. Beads were washed in several consecutive steps by incubation in 1 ml wash buffer for 10 min with rotation in each step and centrifugation for 2 min at 1500 x g for buffer changes. The first wash was performed in RIPA-140 buffer followed by four more stringent washes in RIPA-500 buffer (10 mM Tris-HCl pH 8; 500 mM NaCl; 1 mM EDTA; 1% 'Triton X-100'; 0,1% SDS; 0,1% sodium deoxycholate; 1x complete; 1 mM AEBSF); then, one wash in LiCl buffer was carried out (LiCl buffer: 250 mM LiCl; 10 mM Tris-HCl pH 8; 1 mM EDTA pH 8; 0,5% (v/v) 'IGEPAL CA-630' (Merck); 0,5% (w/v) sodium deoxycholate) and lastly two washes in 1x TE buffer. From now on, beads and inputs, that had been taken before the IP, were processed in parallel. Beads of each biological replicate were resuspended in 100 µl 1x TE; the volumes of inputs were topped up to 100 µl with 1x TE. Each sample was supplemented with 10 µg/ml Ribonuclease (Roche, #11119915001) and then incubated for 30 min at 37°C and 300 rpm in an Eppendorf thermoshaker. To release proteins from chromatin, 0,5% (w/v) SDS and 0,5 mg/ml proteinase K were added and samples were digested for 10 hrs at 37°C and 300 rpm. In a subsequent incubation step for 6 hrs at 65°C and 300 rpm, formaldehyde cross-links were reversed. Next, samples were vortexed vigorously to detach DNA from the beads and spun at 21.000 x g for 1 min. For purification of DNA from the supernatants, the 'MinElute PCR Purification Kit' (Qiagen) was used according to manufacturer's instructions. DNA from each sample was eluted in 500 µl 10 mM Tris-HCl pH 8,5 and stored at -20°C.

2.8.3 Quantitative PCR

The relative DNA amounts of the respective target sequences in the IP and in the input samples were determined by qPCR in the 96-well real-time PCR instrument 'StepOnePlus Real-Time PCR System' (Applied Biosystems) using SYBR Green I as DNA stain. 20 µl reactions containing 1x 'Power SYBR Green 2x MasterMix' (Thermo Fisher), 250 – 500 nM forward and reverse primers and template DNA were prepared. The appropriate primer concentrations not leading to primer dimer formation had been determined beforehand as well as the appropriate dilutions of template DNA from the IP or input samples, that would lead to amplification in the linear range. A DNA standard series was generated and measured in parallel with the IP and input reactions of all three replicates of both genotypes, *wild type* and *ash1*$^{22\ m+z-}$, on the same 96-well plate. From the measured data, the percentage of input DNA of the respective target region recovered after precipitation was calculated using the 'StepOne Software v2.3' (Applied Biosystems).

See Table 11 for location of the amplified target regions and sequences of the corresponding primer pairs.

Table 11. qPCR primer pairs used for amplification of DNA sequences precipitated in ChIP against H3K36me2. Positions of the middle nucleotides in the amplified target regions are given relative to the corresponding gene transcription start site. See list in para 6.1 for abbreviations.

Gene	Position	Forward primer (5' to 3')	Reverse primer (5' to 3')
Ubx	0	TCCAATCCGTTGCCATCGAACGAAT	TTAGGCCGAGTCGAGTGAGTTGAGT
	+0,8 kb	AATTGGTTTCCAGGGATCTGC	ATCCAAAGGAGGCAAAGGAAC
	+1,2 kb	GGCAGTCCTGTTTGTAGGCT	GAGTCCCTATGCCAACCACC
	+34,4 kb	GTCCTGGCCAAGGCAAATATT	CGAAAGGAGAACGGAGAATGG
	+75,5 kb	TTCGTTCAGCTCCTTGATCG	ACAGACATACACCCGCTACC
	+76,9 kb	GCTCGTTGGATCCACTAAACT	TGAGCCGTTAATTGATCGTGAG
wg	+7,3 kb	GTCCGGATCGTGTACAGTGA	GCTGCATTCGGACTAACTGG
tsh	+6,8 kb	TGCCTCGTCTGTTTTAAGTGC	TTGGTGGATGAGTTGGATGGA
Lam	+1,3 kb	AGTGCGTGGAAACTGAATCG	ACCACGCCTTTTGTCTCTTC
dpr12	+6,2 kb	TCAAGCCGAACCCTCTAAAAT	AACGCCAACAAACAGAAAATG
	+39,1 kb	CCGAACATGAGAGATGGAAAA	AAAGTGCCGACAATGCAGTTA

3 Results

3.1 Caf1-55 and MRG15 co-purify with Ash1 from *Drosophila*

In this first section, I shall describe the purification of Ash1 protein complexes from *Drosophila* embryos using tandem affinity purification (TAP), a strategy that has previously been successfully applied for homogenous purification of multimeric PcG protein complexes (e.g. Klymenko *et al*, 2006; Nekrasov *et al*, 2007).

3.1.1 Generation and testing of TAP-Ash1 transgenic strains

In a first step, Arif Mohammed, a former postdoc in the Jürg Müller laboratory, had generated fly strains carrying Ash1 transgenes. One strain contained a transgene encoding full-length Ash1 fused to a TAP tag at its N-terminus (NTAP-Ash1) and another strain a transgene encoding full-length Ash1 with a C-terminal TAP tag (Ash1-CTAP). In both transgenes, the protein was expressed under the control of an *α-tubulin1*-promoter. The TAP tags themselves were composed of a Protein A-moiety, the tag for the first TAP purification step, a *Tobacco Etch Virus* (TEV) cleavage site and a Calmodulin binding peptide (CBP), the tag for the second purification step (Rigaut *et al*, 1999; Figure 5A).

To ensure that the TAP-Ash1 fusion proteins were physiologically functional, the capacity of the Ash1 transgenes to rescue the *ash1*-null mutant phenotype was analyzed. Tripoulas and colleagues had generated the ethyl methanesulfonate (EMS) allele $ash1^{22}$ and reported it to be an *ash1* null allele (Tripoulas *et al*, 1994; 1996; Figure 4). *Drosophila* homozygous for $ash1^{22}$ die in the late pupal stage at the end of metamorphosis and display severe homeotic transformations. Both the NTAP-Ash1 and the Ash1-CTAP transgenes were crossed into an $ash1^{22}$ homozygous background and found to rescue these animals into viable and fertile adults with wild-type morphology, that could be kept as healthy stocks (see Figure 5B for their genotypes). In the following, animals with these genotypes will be called NTAP-Ash1 and Ash1-CTAP flies.

Consistent with the successful rescue tests, the western blot in Figure 5C shows that full-length NTAP-Ash1 or full-length Ash1-CTAP proteins were expressed from their corresponding transgenes. In this western blot, nuclear extracts of wild-type, NTAP-Ash1 and Ash1-CTAP embryos had been probed against Ash1. Although it was expected that the TAP-Ash1 proteins were over-expressed under the *α-tubulin1*-promoter and although they displayed stronger bands

than the endogenous Ash1 in this western blot, it has to be noted, that no statement can be made here about the expression levels of the TAP-Ash1 proteins. The reason for this is that the heavy chains of the α-Ash1 antibody and the secondary antibody bound with high affinity to the Protein A moiety in the TAP-tag in addition to the specific antigen binding. The interaction between the heavy chain and Protein A is a property of all immunoglobulins G. Figure 5C also shows that Ash1 was prone to degradation in the nuclear extract preparation.

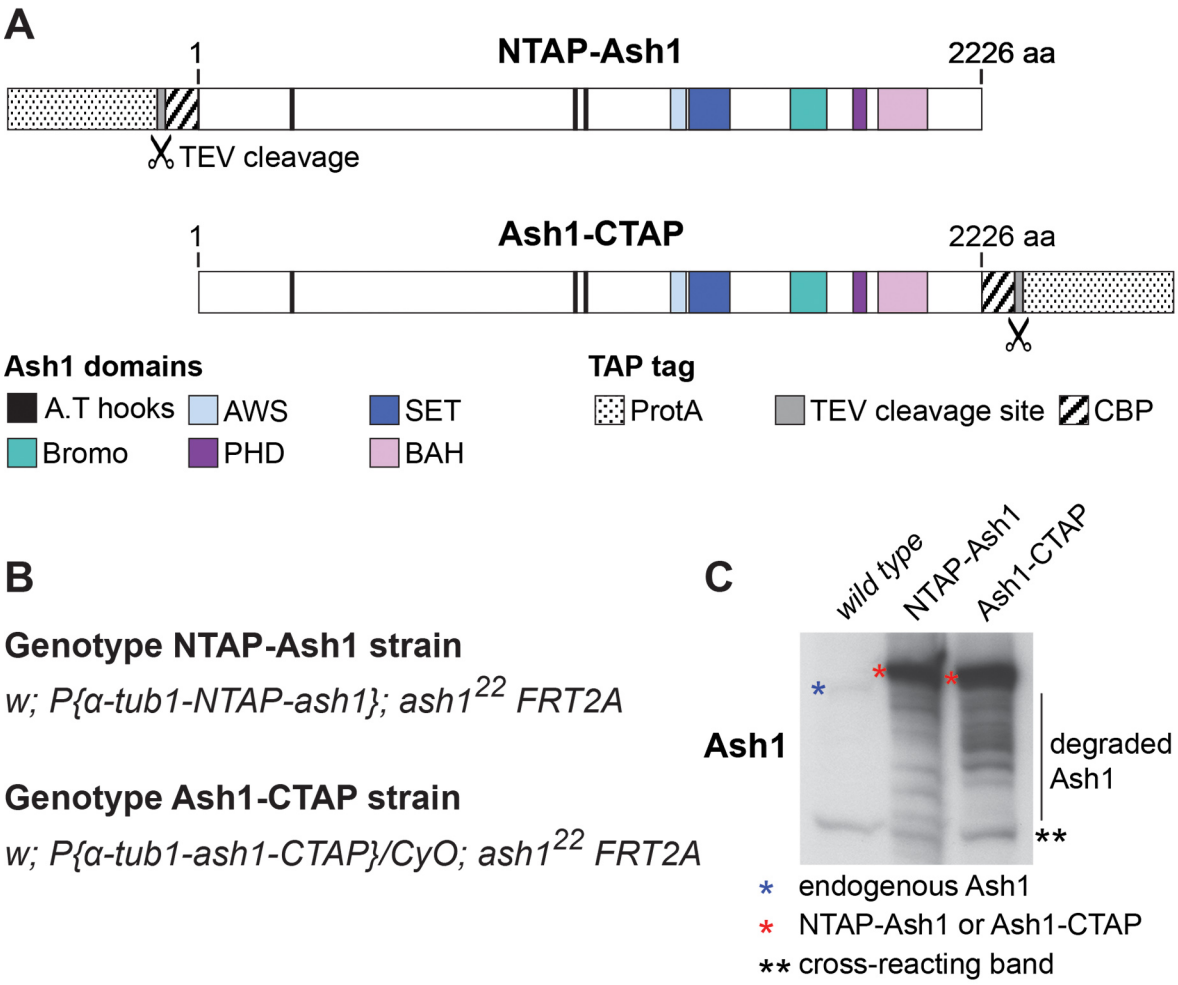


Figure 5. Architecture and expression of TAP-tagged Ash1 fusion proteins in transgenic strains. **(A)** Schematic representation of the N-terminally and the C-terminally TAP-tagged Ash1 fusion proteins (NTAP-Ash1 and Ash1-CTAP). The tags are drawn out of scale. TAP-tag composition: Protein A moiety (ProtA), TEV cleavage site, Calmodulin binding peptide (CBP). **(B)** Genotypes of transgenic fly strains from which Ash1 was TAP-purified. Both, the NTAP-Ash1 and the Ash1-CTAP strain, did not express endogenous Ash1 due to their *ash1*^{22} homozygous background. The chromosome bearing the Ash1-CTAP transgene was homozygous lethal. **(C)** Western blot probed against Ash1 showing that NTAP-Ash1 and Ash1-CTAP proteins were expressed full-length (red asterisk) in embryonic nuclear extracts of NTAP-Ash1 and Ash1-CTAP strains (genotypes as in (B)). The non-specific band marked with two black asterisks served here as a loading control showing that comparable amounts of wild-type and TAP-Ash1 nuclear extracts had been loaded. The intensity differences between the TAP-Ash1 bands and the endogenous Ash1 band (blue asterisk) are not proportional to differences in expression levels due to additional binding of the antibodies' heavy chains to the Protein A-moiety in the TAP tag.

3.1.2 Performance and analysis of purifications of TAP-tagged Ash1 proteins

Ash1 and its binding partners were purified by TAP from embryos of the NTAP-Ash1 flies as well as of the Ash1-CTAP flies. Both fly strains only expressed the respective transgenic Ash1 protein, but no endogenous Ash1, that could have possibly interfered with the co-purification of Ash1-associated proteins (see Figure 5B for genotypes).

The NTAP-Ash1 and the Ash1-CTAP strains were expanded, embryos collected and large-scale nuclear extracts generated. Then, TAPs were performed from TAP-Ash1 and wild-type nuclear extracts in parallel. Part of the purified material was subjected to SDS-PAGEs and visualized by silver stainings (Figure 6A). The protein band patterns in the silver stainings demonstrate that proteins could be successfully enriched by purification of both, NTAP-Ash1 as well as Ash1-CTAP, in contrast to the wild-type controls.

Next, the very same eluates, that are shown in the silver stainings in Figure 6A, were analyzed by mass spectrometry in the Axel Imhof laboratory for identification and quantification of the purified proteins. As a quantitative measurement, the iBAQ factors of each protein in the NTAP-Ash1 respectively Ash1-CTAP eluate and in the corresponding wild-type (wt) eluate were determined (see para 2.7.4 for details on the iBAQ quantification). Then, the ratio of both iBAQ factors (iBAQ(TAP-Ash1/wt)) was formed to express protein enrichment in the eluates. Figure 6B presents in a scatter plot the enrichment of each protein by the NTAP-Ash1 purification versus the enrichment by the Ash1-CTAP purification. Three proteins are high-lighted in red in the scatter plot that stand out from the other identified proteins due to their high enrichment in both purifications: the bait Ash1, Caf1-55 and MRG15 (see Table 12 for identified peptides). These are the proteins I focused on in the follow-up work. All other top protein hits were only co-purified in high quantities in one of the two TAPs. Also, the specificity of many other high hits was questionable, such as of ribosomal proteins or of constituents of the cytoskeleton that had probably been introduced into the TAPs as cytoplasmic impurities during nuclear extract preparation (Figure 6B, hits labeled in grey). Fsh, the only protein hitherto reported as a binding partner of Ash1, had been identified as the most enriched protein after the bait in Ash1 purifications from *Drosophila* S2-DRSC cells (Kockmann *et al*, 2013). However, in the TAP purifications I performed here from *Drosophila* embryos, Fsh (Figure 6B, labeled in black) was only quantified as about the 300th most abundant (NTAP-Ash1 TAP) or as the 50th most abundant (Ash1-CTAP TAP) protein hit. Therefore, I do not consider Fsh as a possible stable Ash1 complex subunit in the following.

The western blots presented in Figure 6C are complementary to the mass spectrometric analysis. The input materials of the NTAP-Ash1 TAP (NTAP-Ash1 and wild-type nuclear extracts)

and the corresponding TAP eluates were blotted and probed with antibodies as indicated. Ash1, MRG15 and Caf1-55 were detected in the NTAP-Ash1 but not in the wild-type eluate. The PcG proteins O-glycosyltransferase (Ogt) and E(z) served as negative controls. Both were neither found by mass spectrometry nor by western blot in the NTAP-Ash1 or Ash1-CTAP eluates.

In conclusion, the results of the NTAP-Ash1 and the Ash1-CTAP purifications together strongly suggested that Caf1-55, MRG15 and Ash1 interact biochemically. Whether these interactions are specific and direct, was investigated in the next step in reconstitution assays with recombinant proteins.

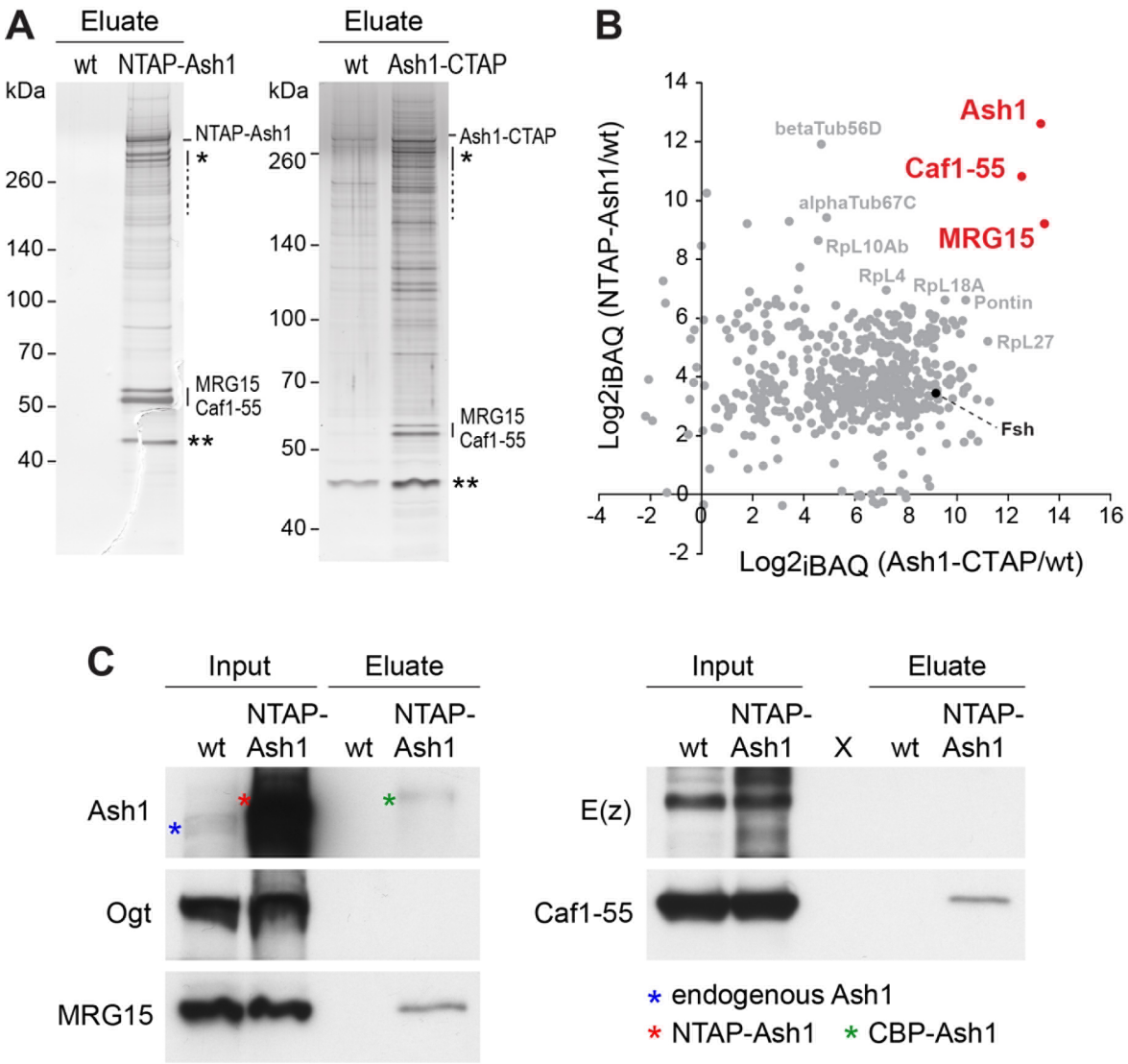


Figure 6. Analysis of TAP of NTAP-Ash1 and Ash1-CTAP from *Drosophila* embryos.
(A) Proteins purified by TAPs from nuclear extracts of wild-type (wt) and NTAP-Ash1 embryos (left) and wild-type and Ash1-CTAP embryos (right), that were separated on '4-12% NuPAGE Bis-Tris Gels', then visualized by silver stainings. Protein identity of labeled bands is only assumed since entire lanes, not single bands, were analyzed by mass spectrometry. Probable Ash1 degradation products are indicated with one asterisk. The band labeled with two asterisks is considered as non-specific, since it had been detected before in several mock purifications and here in the wild-type control of the Ash1-CTAP purification. **(B)** Scatter plot representation of protein enrichments in NTAP-Ash1 eluates versus protein enrichments in Ash1-CTAP eluates as determined by mass spectrometric analysis of the in-gel digests. Protein enrichments are expressed as $\log_2$-transformed ratios of the iBAQ factors of each protein in the TAP-Ash1 eluate and in the corresponding wild-type control. Caf1-55 and MRG15 (red) stand out as factors highly co-enriched with both, NTAP-Ash1 and Ash1-CTAP. Proteins written in grey possibly co-purified unspecifically, Fsh (black) has been previously reported as an Ash1 interactor. iBAQ: Intensity based absolute quantification value; see para 6.1 for protein names written out. **(C)** Confirmation of MRG15 and Caf1-55 as co-purifying factors with NTAP-Ash1 by western blot. The 'Input' material of TAPs (nuclear extracts of wild-type and NTAP-Ash1 embryos) was analyzed versus the corresponding TAP eluates ('Eluate'). *Continued on next page*.

The red asterisk marks NTAP-Ash1 in the transgenic extracts. The strong NTAP-Ash1 signal was partially caused by unspecific binding of the IgG antibodies to the Protein A moiety of the TAP tag in addition to the antigen binding. The blue asterisk marks endogenous Ash1 in the wild-type control and the green asterisk CBP-Ash1 in the eluate, which remained after TEV cleavage of the TAP tag after the first TAP purification step (Figure 5A). Ogt and E(z) are negative controls.

3.2 Recombinant Ash1, Caf1-55 and MRG15 co-exist in a complex

The mass spectrometric analysis of the TAPs identified Caf1-55 and MRG15 as promising candidates for Ash1 complex subunits. Both candidates were already well-studied chromatin-associated proteins, which have been identified as subunits of various other nuclear multiprotein complexes. Caf1-55 is a core-subunit of the PcG complex PRC2, the histone chaperone complex CAF-1, the Nucleosome remodeling and deacetylase complex (NuRD), the Nucleosome remodeling factor complex (NURF), as well as of Histone acetyltransferase 1 (HAT1) complexes and of histone deacetylase (HDAC) complexes (reviewed in Suganuma *et al*, 2008). As sketched in Figure 7, Caf1-55 bears a WD40 β-propeller by which it binds to soluble H3/H4 heterodimers and to unmodified H3 in the context of nucleosomes (Tagami *et al*, 2004; Nowak *et al*, 2011; Schmitges *et al*, 2011). MRG15 has been reported to interact with the human polypyrimidine tract-binding (PTB) protein that is involved in alternative splicing (Luco *et al*, 2010). The chromatin-remodeling and HAT complex Tip60 in *Drosophila* and its homologous complex NuA4 in yeast have been shown to contain MRG15 (Kusch *et al*, 2004) respectively the MRG15 yeast homologue Eaf3 (Eisen *et al*, 2001; Krogan *et al*, 2004) as stable subunit (Tip60: Tat-interactive protein 60 kDa complex; NuA4: Nucleosome acetyltransferase of histone H4 complex). The functions of the Tip60 complex are manifold and range from roles in DNA repair and apoptosis to transcriptional regulation as reviewed in Sapountzi *et al*, 2006. Moreover, the MRG15 orthologue Eaf3 has been purified as a member of the HDAC complex Rpd3S (Carrozza *et al*, 2005; Keogh *et al*, 2005). MRG15 bears two conserved domains, the chromo barrel and the MRG domain (Figure 7). In the light of the H3K36 methyltransferase activity of Ash1, the binding specificities of the chromo barrel domains of MRG15 and Eaf3 for H3K36me2/3 are highly interesting (Zhang *et al*, 2006; Sun *et al*, 2008; Xu *et al*, 2008). The MRG domain, the hallmark of MRG protein family members, is involved in protein-protein interactions (Bowman *et al*, 2006; Xie *et al*, 2012; 2015).

With the aim to reconstitute and to investigate the suggested interactions of Ash1 with Caf1-55 and MRG15, binding studies were performed by purifying the corresponding recombinant

proteins, that had been co-expressed in various combinations using baculoviruses in High Five insect cells. Of Ash1, the C-terminal half from amino acid 1041 to 2226 encompassing all conserved domains ($Ash1_C$) was expressed. Caf1-55 and MRG15 were each expressed as full-length proteins. See Figure 7 for domain architecture of Ash1, $Ash1_C$, Caf1-55 and MRG15.

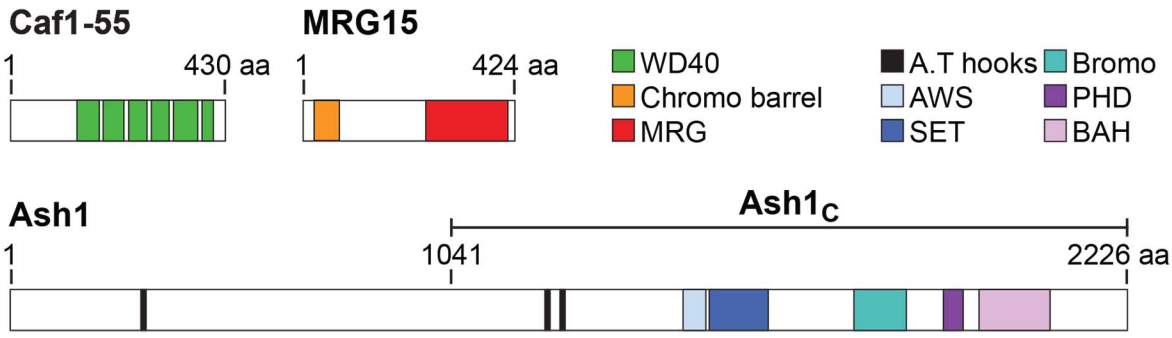


Figure 7. Domain architecture of *Drosophila* Caf1-55, MRG15 and Ash1. For *in vitro* reconstitution assays, full-length Caf1-55, full-length MRG15 and the C-terminal part of Ash1 labeled as $Ash1_C$ were recombinantly expressed. Chromo: chromatin organization modifier; MRG: MORF4-related gene; WD40: Tryptophan-aspartic acid 40.

3.2.1 Interactions of Ash1 with Caf1-55 and MRG15 can be reconstituted

To test if Ash1 and Caf1-55 physically interact with each other, $Ash1_C$ was co-expressed with Strep-tagged Caf1-55 and, in a separate co-expression, with full-length Strep-tagged Esc. I had decided on the PRC2 subunit Esc as a negative control, since Esc also bears a conserved WD40 domain like Caf1-55. When Strep-affinity purification was then performed from High Five cell extracts of the $Ash1_C$-Strep-Caf1-55 co-expression, $Ash1_C$ was co-isolated. But with Strep-Esc as bait in the pull-down assay from the $Ash1_C$-Strep-Esc co-expression, $Ash1_C$ could not be co-purified. Therefore, Ash1 does interact directly and specifically with Caf1-55. See Figure 8A for Coomassie Blue stainings of the isolated proteins. On a side note, the tags of proteins are only mentioned in the text above and in the following, when they serve as affinity tags for the corresponding purification. All tags are indicated in the legend of Figure 8.

Next, Strep-affinity purifications were performed from a co-expression of $Ash1_C$ and full-length Strep-MRG15 and in parallel from a co-expression of $Ash1_C$ and full-length Strep-Male-specific lethal-3 (Msl3) to verify the proposed Ash1:MRG15 interaction. Msl3 is also a member of the MRG protein family and, as such, contains a conserved MRG domain just like MRG15. However, $Ash1_C$ only clearly co-purified with MRG15 and not with Msl3 (Figure 8B). Though MRG15 seemed more enriched than Ash1 In the Ash1-MRG15 purification in the Coomassie Blue

staining in (Figure 8B), but the stoichiometric ratio of Ash1 to Strep-MRG15 in the eluate equaled the ratio of the expression levels of both proteins in the input material.

Overall, these pull-down assays showed that Ash1 interacts directly and specifically with both, Caf1-55 and MRG15.

3.2.2 In the trimeric complex, Caf1-55 and MRG15 interact with Ash1, but not with each other

Figure 8A and B show that Ash1 interacts with Caf1-55 in the absence of MRG15 as well as with MRG15 in the absence of Caf1-55, but it was yet an open question if also all three proteins together form a complex. Towards an answer to this question, His-affinity purifications were performed from a co-expression of all three proteins of interest and, in parallel, from a co-expression of His-Caf1-55 and MRG15. His-tagged Caf1-55 served as bait in both purifications. The pull-down from the latter co-expression of His-Caf1-55 and MRG15, resulted in the isolation of His-Caf1-55 alone (Figure 8C). However, in the pull-down from the co-expression of His-Caf1-55, $Ash1_C$ and MRG15, all three proteins were purified, notably, also MRG15 (Figure 8C). Hence, Ash1, Caf1-55 and MRG15 can co-exist in a complex. Though, MRG15 was enriched in sub-stoichiometric amounts, which is most likely due to greater difficulties to purify MRG15 than to purify Caf1-55 with the protocol used here. In Coomassie Blue stainings of aliquots taken in all purification steps (not presented here), I observed that a large fraction of the over-expressed MRG15 protein stays behind with the cell debris from the insect cell extract.

Taken together, the results of all purifications presented in Figure 8, show that Ash1, Caf1-55 and MRG15 form a complex *in vitro*. Within this complex, Ash1 acts as the scaffolding unit binding to Caf1-55 and MRG15, whereas Caf1-55 and MRG15 do not physically interact with each other. I named the complex AMC corresponding to its subunits (Figure 8D).

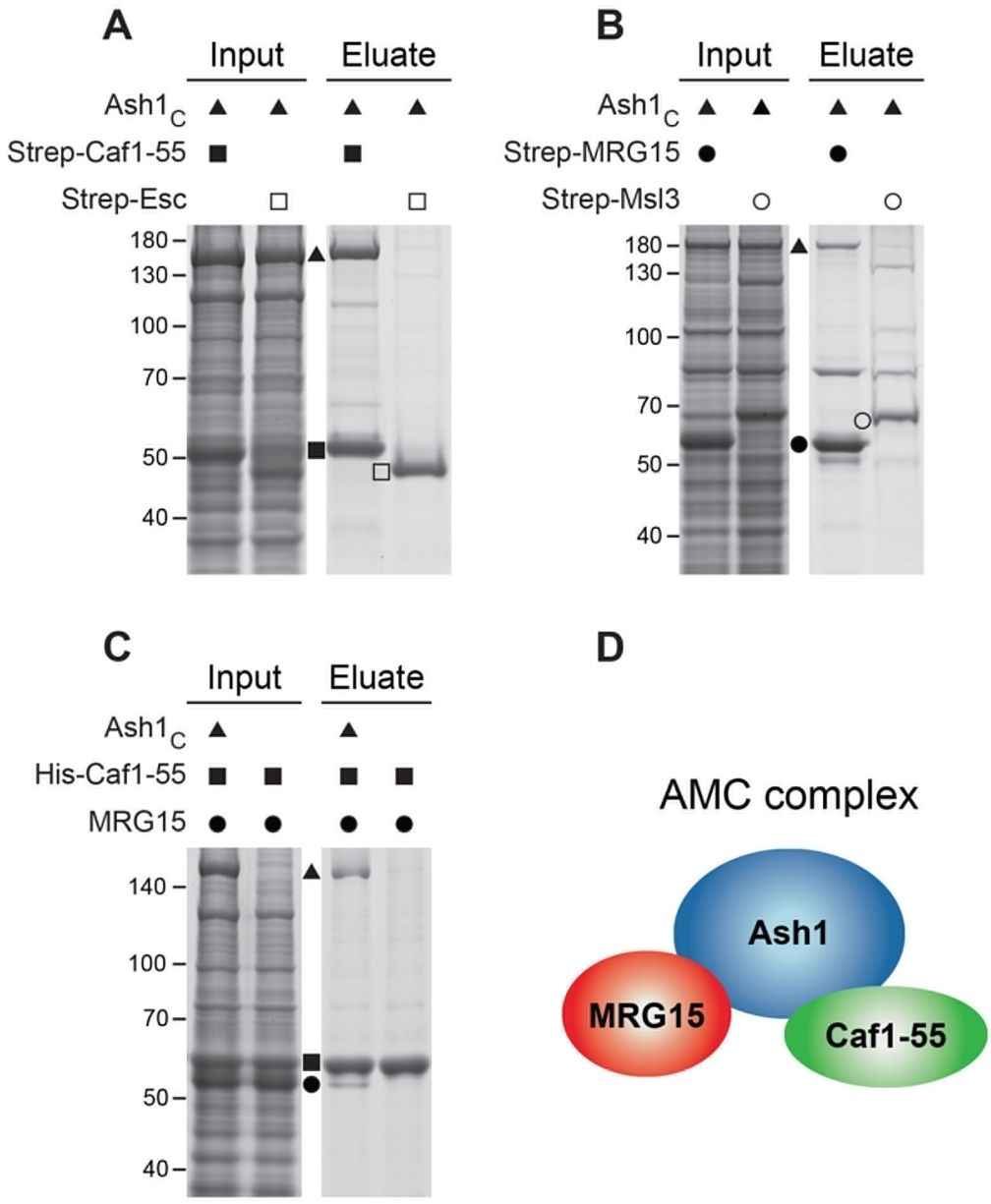


Figure 8. Investigation of biochemical interactions among Ash1, Caf1-55 and MRG15.
(A-C) Total High Five cell lysates from co-expressions of the indicated recombinant proteins (Inputs) and affinity-purified material (Eluates) obtained from inputs, visualized by SDS-PAGE and Coomassie Blue stainings. $Ash1_C$ was Flag-tagged. **(A)** Strep-purifications of co-expressions of $Ash1_C$ and Strep-Caf1-55 and of $Ash1_C$ and Strep-Esc showing that $Ash1_C$ binds to Caf1-55, but not to Esc, albeit both bear a WD40 domain. PA gel: 10% Laemmli. **(B)** Strep-purifications of co-expressions of $Ash1_C$ with MRG domain proteins, $Ash1_C$ with Strep-MRG15 and $Ash1_C$ with Strep-Msl3, revealing that $Ash1_C$ interacts with MRG15, but not with Msl3. PA gel: 8% Bolt Bis-Tris Plus. **(C)** His-purifications of co-expressions of $Ash1_C$, His-Caf1-55 and MRG15 and, in parallel, of His-Caf1-55 and MRG15 showing that here Caf1-55 and MRG15 do not interact, but together with $Ash1_C$ they can co-exist in a complex. In this gel, Caf1-55 migrated slower than MRG15 due to its polyhistidine-tag and the different SDS-PAGE buffer system used. PA gel: 8% Laemmli. MRG15 was Strep-tagged. **(D)** Schematic representation of the Ash1 complex, that is named AMC and was reconstituted from Ash1, Caf1-55 and MRG15.

3.3 Ash1 binds to MRG15 via a conserved FxLP motif

The binding assay presented in the following is a continuation of the *in vitro* reconstitution of the interactions in the AMC complex described in para 3.2, but with the aim to investigate the binding of Ash1 to MRG15 in more detail. As described in the previous paragraph, MRG15 has been purified as a subunit of various multiprotein complexes other than AMC. Interestingly, the different MRG15 interactors in these complexes all share a conserved FxLP sequence which has been suggested to be required for the binding of MRG15. In particular, Xie and colleagues uncovered in structural studies that the PHD finger protein 12 (Pf1) binds via its FxLP motif to the MRG domain of MRG15 in the mammalian Rpd3S/Sin3S complex and showed an analogous interaction between the FxLP motif of the MRG/MORF4L-binding protein (MRGBP) and the MRG15 MRG domain in the mammalian NuA4/Tip60 complex (Xie *et al*, 2012; 2015). Intriguingly, I found the FxLP motif also conserved in Ash1 in an alignment of several Ash1 orthologues (Figure 9). The Ash1 FxLP motif is located in direct vicinity of the Ash1 SET domain.

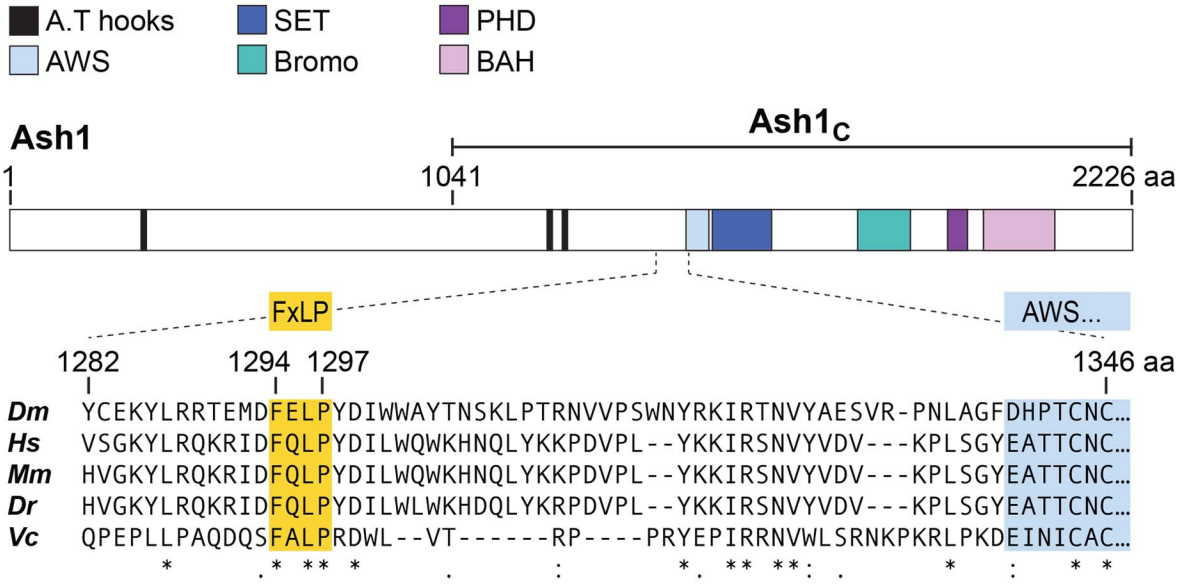


Figure 9. Identification of a conserved FxLP motif in Ash1 characteristic for MRG domain binding proteins. Alignment of the *Drosophila* Ash1 segment containing the FxLP motif with orthologous Ash1 sequences. The Ash1 FxLP motif is located approximately 50 aa upstream of the AWS domain and 100 aa upstream of the SET domain. *Dr*: *Danio rerio*; *Vc*: *Volvox carteri*. Alignment was done with Clustal Omega (EMBL-EBI).

The next consequential step was to test whether the identified FxLP motif in Ash1 is engaged in binding to MRG15 like the FxLP motifs of Pf1 and MRGBP. I generated a baculovirus strain from which mutant $Ash1_C$ protein with the FxLP sequence altered to RxRP was expressed ($Ash1_C^{RXRP}$). Next, Strep-affinity purifications were performed from co-expressions of $Ash1_C^{RXRP}$ and Strep-MRG15 and from co-expressions of wild-type $Ash1_C$ and Strep-MRG15 as control. As shown in Figure 10, $Ash1_C$ successfully purified along with MRG15, whereas only MRG15 and no $Ash1_C^{RXRP}$ was detected in the eluate of the $Ash1_C^{RXRP}$-Strep-MRG15 purification. The RxRP-mutation had disrupted the binding to MRG15. In other words, the Ash1 FxLP motif is essential for the interaction of Ash1 with MRG15.

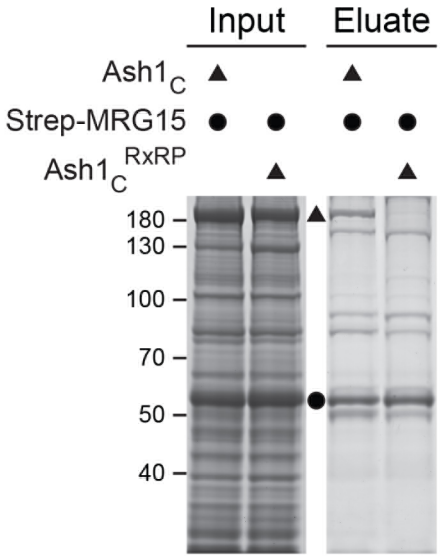


Figure 10. Verification of Ash1 binding to MRG15 via the FxLP motif. Total High Five cell lysates from co-expressions of Strep-MRG15 with wild-type $Ash1_C$ and with mutant $Ash1_C^{RxRP}$ (Inputs) and Strep-purified material (Eluates) obtained from inputs, visualized by SDS-PAGE and Coomassie Blue stainings. Wild-type $Ash1_C$ co-purified with Strep-MRG15, whereas $Ash1_C^{RxRP}$ could not bind to MRG15 and was not detected in the eluate. PA gel: 8% Bolt Bis-Tris Plus. $Ash1_C$ and $Ash1_C^{RxRP}$ were His-tagged.

3.4 Ash1 exhibits high H3K36 di-methyltransferase activity in complex with MRG15

Ash1 possesses very low or no HMTase activity at all, when it is overexpressed, isolated and tested alone. As in the case of human ASH1L, this is most likely due to the presence of the conserved auto-inhibitory loop in Ash1 (An *et al*, 2011). After having identified Caf1-55 and

MRG15 as Ash1 interactors, I now went on to investigate in HMTase assays whether Caf1-55 and MRG15 are able to alter the enzymatic activity of Ash1.

The HMTase assays were performed on recombinant mononucleosomes as substrates. The enzymatic complexes analyzed were the $Ash1_C$:Caf1-55 and the $Ash1_C$:MRG15 assemblies that have been presented before in Figure 8A and B and, in addition, a mutant $Ash1_C^{R1464A}$:MRG15 complex. $Ash1_C^{R1464A}$:MRG15 had, like the other complexes, been purified from co-expressions from baculoviruses in High Five cells. The mutated arginine R1464 is located in the SET domain of Ash1 in an RXXNHS motif that is highly conserved among SET domain HMTases (Figure 4). Mutation of this residue in other HMTases like in *Drosophila* E(z) (Müller *et al*, 2002) or in yeast Set2 (Strahl *et al*, 2002) resulted in catalytic inactivity. The structure of human ASH1L shows that this arginine stabilizes the orientation of the *S*-adenosyl methionine-binding loop. Modeling of *Drosophila* Ash1 based on the ASHL structure in the Elena Conti laboratory (MPIB) suggested that this function of the arginine R1464 is lost upon mutation to alanine without disrupting the SET domain conformation. The phenotype of *ash1*^{R1464A} mutants in *Drosophila* is presented in the following paragraphs. Here, the $Ash1_C^{R1464A}$:MRG15 complex was used to control detection of only Ash1-specific HMTase activity in the HMTase assays. As detection method for histone methylation, western blotting was chosen. The HMTase reactions were analyzed with antibodies against H3K36me2 and H3K36me3. I did not analyze H3K36me1, since none of the tested antibodies exhibited satisfying specificity.

The results of the western blots are shown in Figure 11 along with Coomassie Blue stainings of the enzymatic complexes separated by SDS-PAGE. The current consensus in the literature is that Ash1 catalyzes specifically di-methylation of K36 in histone H3 (para 1.3). Therefore, I shall first discuss the western blot against H3K36me2. Based on this western blot, $Ash1_C$ in complex with Caf1-55 has not been active (Figure 11; lanes 2, 3). Only bands resulting from unspecific antibody binding are visible in lanes 2 and 3 on the membrane probed against H3K36me2, like in the mock reaction (lane 1). The mock reaction contained no enzymatic complexes, but mononucleosomes. However, remarkably, considerable levels of H3K36me2 are detected in the reactions with the $Ash1_C$:MRG15 complex (lanes 4, 5). The lack of detectable H3K36me2 levels in the reactions with the mutant $Ash1_C^{R1464A}$:MRG15 assembly (lanes 6, 7) proves that the H3K36me2 modifications in the $Ash1_C$:MRG15 reactions have been generated by the Ash1 SET domain and not by MRG15 or other factors in these assays. Lastly, none of the complexes tested here demonstrated H3K36 tri-methyltransferase activity.

The results of the HMTase reactions taken together, strongly suggest that MRG15 potently enhances Ash1 enzymatic activity. In addition, these experiments represent a further confirmation of the long-debated catalytic specificity of Ash1 for H3K36 di-methylation (see also para 1.3).

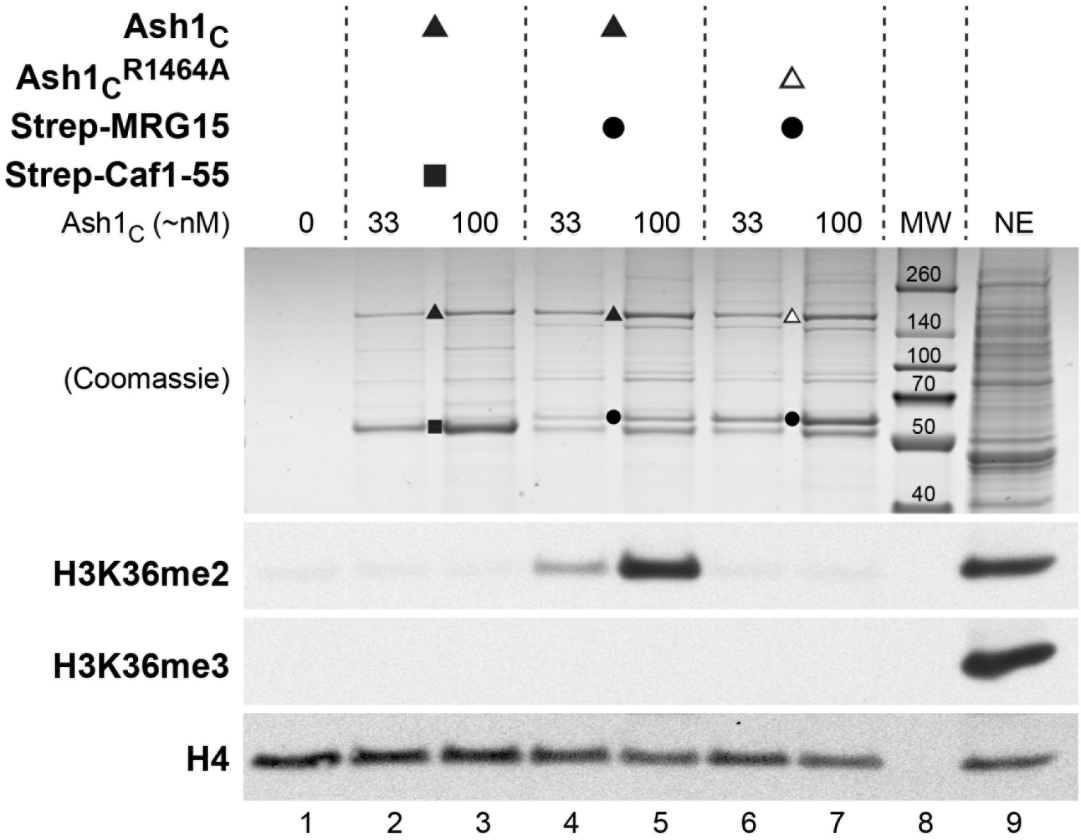


Figure 11. Comparative analysis of H3K36 di- and tri-methylation activity of Ash1:Caf1-55 and Ash1:MRG15 complexes. HMTase reactions were performed in parallel with *Drosophila* $Ash1_C$:Caf1-55 (lane 2, 3), $Ash1_C$:MRG15 (lane 4, 5) and $Ash1_C^{R1464A}$:MRG15 (lane 6, 7) complexes as well as no enzyme as mock (lane 1) on recombinant *Drosophila* mononucleosomes. 50% of the reactions were analysed for H3K36me2, the other 50% for H3K36me3 by SDS-PAGEs and western blotting. Shown here, is the Coomassie Blue staining of the upper part of the gel from the H3K36me2-analysis visualizing the enzymatic complexes (10% Bolt Bis-Tris Plus gel). The bottom part of the gel with the nucleosome substrates was blotted and the membrane probed against H3K36me2 and H4 as loading control. From the H3K36me3-analysis, only the blot against H3K36me3 is presented. Each Ash1 complex was tested for activity in two concentrations. The Ash1 complex concentrations had been normalized before by estimating $Ash1_C$ concentrations from comparison with a Coomassie Blue-stained BSA standard. $Ash1_C$ and $Ash1_C^{R1464A}$ proteins were His-tagged.
The $Ash1_C$:Caf1-55 and the SET domain-mutant $Ash1_C^{R1464A}$:MRG15 complex did not exhibit H3K36 di-methylation or tri-methylation activity (lanes 2, 3, 6, 7). The wild-type $Ash1_C$, when in complex with MRG15, was highly active for H3K36 di-methylation, but not for H3K36 tri-methylation (lanes 4, 5). The *Drosophila* embryonic nuclear extract (NE) in lane 9 serves as control for the western blot analysis and shows that the blot against H3K36me3 technically worked. MW (lane 8): molecular weight standard in kDa.

3.5 Catalytic *ash1* mutants show a trxG-like homeotic phenotype

Analysis of multiple EMS-induced alleles of *ash1* in *Drosophila* established that *ash1* mutants in general display a trxG phenotype (Shearn *et al*, 1987; Tripoulas *et al*, 1994). But a systematic comparative *in vivo* analysis of defined loss-of-function mutations in individual Ash1 domains remained to be performed. For that reason, the contribution of the different domains to Ash1 function cannot be assessed to date. The key role is generally ascribed to the Ash1 SET domain since histone methylation is the only Ash1 function described in the literature. In a set of genetic experiments presented in this paragraph, I analyzed the phenotype of a *Drosophila* Ash1 SET domain mutant only expressing catalytic inactive Ash1 protein in direct comparison with an *ash1*-null mutant. The alleles employed for this purpose were (1) $ash1^{R1464A}$, (2) $ash1^{wt}$, (3) $ash1^{10}$ and (4) $ash1^{22}$. The mutations of all alleles have already been introduced in Figure 4. In brief, the $ash1^{R1464A}$ and the $ash1^{10}$ alleles both bear a SET domain point mutation of a well conserved residue and encode full-length Ash1 protein. The R1464A mutation has been shown in the previous paragraph 3.4 to be responsible for catalytic inactivity of the recombinant $Ash1_C^{R1464A}$ protein *in vitro*. The $ash1^{10}$ allele contains an N1467I mutation in close proximity to the R1464A mutation. But it has not been tested if the enzymatic activity of the Ash1 protein expressed from the $ash1^{10}$ allele is impaired or abolished. The $ash1^{22}$ allele is considered as an *ash1* null allele due to a premature stop codon created by mutation of the Q130 codon in *ash1* (Tripoulas *et al*, 1996). The $ash1^{10}$ and $ash1^{22}$ alleles are isolates from EMS screens (Shearn *et al*, 1987; Tripoulas *et al*, 1994), whereas the $ash1^{R1464A}$ allele has been generated in the course of this study along with its control allele $ash1^{wt}$. Both, $ash1^{R1464A}$ and $ash1^{wt}$ are transgenes that were cloned from genomic *ash1* sequences and subsequently inserted into the same *attP* site of a specific landing site strain. For the precise genotypes of all *Drosophila* mutants with these alleles presented in the following including a description how they were generated, please see para 2.3.3 and 2.3.4. Animals derived from mothers homozygous for $ash1^{22}$ were considered to be devoid of maternally deposited wild-type Ash1 and are marked as m-, animals having been provided with maternal Ash1 are marked as m+. Mutants without the endogenous wild-type *ash1* gene carry the label z-, standing in this case for no wild-type Ash1 protein expressed from the zygotic genome.

Before characterizing the *ash1* mutant phenotypes, I performed western blot analyses to ensure that full-length Ash1 protein is expressed from the transgenic $ash1^{R1464A}$ allele. Specifically, flies called $ash1^{R1464A\ m+\ z-}$, that carried the $ash1^{R1464A}$ allele in a homozygous $ash1^{22}$-null mutant background, were tested along with homozygotes for the $ash1^{22}$-null allele ($ash1^{22\ m+\ z-}$ animals).

As shown in the western blot in Figure 12, the *ash1*$^{22\ m+\ z-}$ flies did indeed not express detectable levels of Ash1. Since total extracts from dividing tissues, brains and imaginal discs, from 3rd instar larvae were analyzed, the maternally deposited Ash1 protein was not traceable neither in these animals. However, in the extracts from the *ash1*$^{R1464A\ m+\ z-}$ flies, wild-type levels of full-length Ash1 protein were detected, that could have only been generated from the SET domain mutant transgene as the only *ash1* allele in these animals. Therefore, the *ash1*^{R1464A} allele is functional and suitable for analysis of the *ash1*-catalytically inactive mutant phenotype.

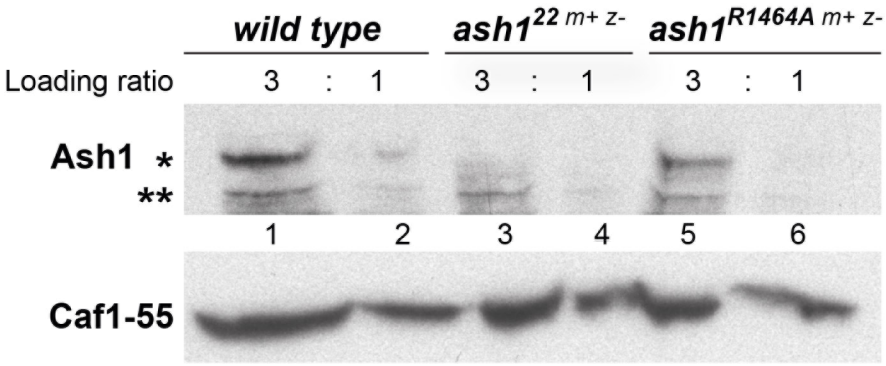


Figure 12. Analysis of *ash1*^{22} and *ash1*^{R1464A} alleles for Ash1 protein expression. Western blot of total extracts from brains and imaginal discs of wild-type, *ash1*$^{22\ m+\ z-}$ and *ash1*$^{R1464A\ m+\ z-}$ 3rd instar larvae. The upper membrane part was probed against Ash1, the bottom part against Caf1-55 as loading control. Full-length Ash1 protein is marked with one asterisk, a cross-reacting band is marked with two asterisks and served as additional loading control here. Ash1 protein was not detectable in the extract of the reported *ash1*-null mutant *ash1*$^{22\ m+\ z-}$, whereas *ash1*$^{R1464A\ m+\ z-}$ flies expressed full-length Ash1 at wild-type levels from the mutant transgene.

3.5.1 The catalytic activity of Ash1 is required for viability

In a first step to characterize the *ash1*-catalytic mutant phenotype versus the *ash1*-null mutant phenotype, the lethality of the corresponding mutants was analyzed. I collected a specific number of 1st and 2nd instar larvae of each genotype and counted the fraction of flies, that developed from this pool of larvae into pupal stage, as well as the fraction, that eclosed as adults from the pupal case. Females and males were not quantified separately, since the survival rates throughout the developmental stages of all *ash1* mutants analyzed did not show a gender-specific bias. The results are presented in the histogram in Figure 13.

Homozygotes for the *ash1*^{22} allele but with maternally deposited wild-type Ash1 protein (*ash1*$^{22\ m+\ z-}$) were compared with *ash1*^{22} homozygotes devoid of maternal Ash1 (*ash1*$^{22\ m-\ z-}$). I point this out, since all previous studies on the requirement of Ash1 during development analyzed *ash1* mutants that had been provided with wild-type Ash1 protein by their *ash1*^{22} heterozygous

mothers. Here, the severest possible *ash1*-null mutant, $ash1^{22\ m-\ z-}$, that neither contained zygotically expressed nor maternally deposited Ash1, was derived from mothers with $ash1^{22}$ homozygous germ cells. Surprisingly, the majority of these $ash1^{22\ m-\ z-}$ flies developed as far as pupal stage, just like $ash1^{22\ m+\ z-}$ animals did (Figure 13, data bars 2, 3). But fewer $ash1^{22\ m-\ z-}$ flies than $ash1^{22\ m+\ z-}$ flies completed metamorphosis and reached the pharate adult stage (last pupal stage with adults in the pupal case, not quantified here).

The transgenic catalytic-mutant allele $ash1^{R1464A}$ and its control, the transgenic $ash1^{wt}$ allele, were crossed into a homozygous $ash1^{22}$ mutant background. These flies, named $ash1^{R1464A}$ and $ash1^{wt}$ in the following text and in Figure 13, were directly comparable with each other. They contained the respective transgene as the only *ash1* allele, which allowed for analyses of the specific phenotypes of both transgenes. As a first result, the $ash1^{wt}$ transgene rescued fully the $ash1^{22}$ mutant phenotype. The $ash1^{wt}$ flies developed into fertile adults with wild-type morphology, which could be kept as a strain. Interestingly, the SET domain-mutant $ash1^{R1464A}$ flies developed into the adult stage as well. The first generation, $ash1^{R1464A\ m+\ z-}$ animals, were sick but fertile. This permitted the generation of $ash1^{R1464A\ m-\ z-}$ flies, that were only provided with mutant $Ash1^{R1464A}$ protein as maternally loaded and zygotically expressed Ash1 protein. Again, a fraction of these $ash1^{R1464A\ m-\ z-}$ animals reached adulthood and eclosed from the pupal case (data bars 5). Therefore, lethal stages and survival rates of $ash1^{R1464A\ m-\ z-}$ and $ash1^{wt}$ flies were very similar as determined here, i.e. counting flies as adults once they eclosed the pupal case. However, for a complete description of the lethality of $ash1^{R1464A}$ mutants, it has to be noted that the fraction of the $ash1^{R1464A\ m-\ z-}$ flies, that hatched after metamorphosis, died without exception within one to two days. No strain could be maintained. Moreover, $ash1^{R1464A\ m-\ z-}$ and also $ash1^{R1464A\ m+\ z-}$ adults exhibited severe homeotic transformations which will be described in detail in the following paragraph. In comparison with the $ash1^{22}$-null mutants, that invariably died in the pupal stage (data bars 3), the phenotype of the $ash1^{R1464A}$ flies was weaker.

Taken together, the main finding of this experiment is that the catalytic activity of Ash1 is required for viability over more than one generation. This result is further supported by the phenotype of the $ash1^{10\ m-\ z-}$ flies that were trans-heterozygous for the second Ash1 SET domain-mutant allele $ash1^{10}$ and for $ash1^{22}$. These mutants were derived from trans-heterozygous $ash1^{10\ m+\ z-}$ mothers and from fathers heterozygous for the $ash1^{10}$ allele. A fraction of the $ash1^{10\ m-\ z-}$ mutants hatched from the pupal case (data bars 6). But these $ash1^{10\ m-\ z-}$ adults exhibited homeotic transformations and died within one to two days, just like the $ash1^{R1464A\ m-\ z-}$ flies did.

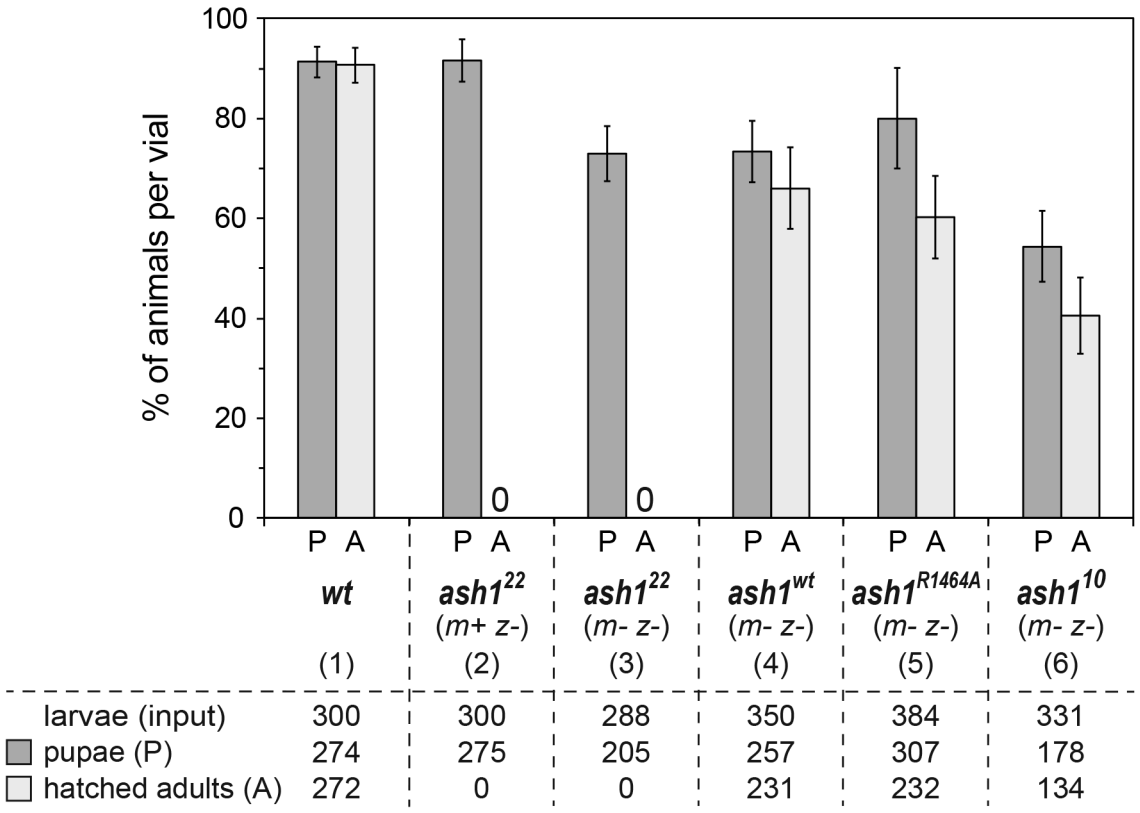


Figure 13. Survival rates of *ash1*-null and *ash1*-catalytic mutants throughout development. The indicated numbers of 1st and 2nd instar larvae (input) of each specified genotype were collected, distributed among at least three different breeding vials and reared. The numbers of these larvae developing into pupal stage (P) and the numbers of adults hatching from the pupal case (A) were determined. In the histogram, the percentages of P and A from the inputs are given as mean values ± standard deviation, that were derived from the numbers counted in the three or more independent vials for each genotype. *ash1*-null mutants, $ash1^{22\ m+\ z-}$ and $ash1^{22\ m-\ z-}$, developed up to pupal stage, but did not hatch (0). A fraction of the *ash1*-catalytic mutants $ash1^{R1464A\ m-\ z-}$ and $ash1^{10\ m-\ z-}$ eclosed from the pupal case, though neither of them could be maintained as a strain.

3.5.2 Adult *ash1*-catalytic mutants exhibit homeotic transformations like *ash1*-null mutant

Ash1 has been reported to regulate transcription of *HOX* genes in development as described in paragraph 1.1.3. Consequently, impairment of Ash1 function in *ash1* mutants manifests itself in homeotic morphological transformations.

In this paragraph, I compare the morphological transformations of an *ash1*-catalytically inactive mutant in type and severity with the transformations of a known *ash1*-null mutant. In particular, the morphology of the thoracic segments T2 and T3 and of the male posterior

abdomen of *wild type*, *ash1*$^{22\ m-\ z-}$, *ash1*$^{R1464A\ m-\ z-}$ and *ash1*$^{10\ m-\ z-}$ flies was studied in adult cuticle sections. Microscopic images of characteristic cuticles of each genotype are shown in Figure 14.

The thorax sections (Figure 14, T2/T3 segment) revealed morphological transformations of structures in the T3 segment toward T2 identity in all *ash1* mutants analyzed. These were namely partial transformations of the balancing organ haltere in T3 into T2-wing tissue (dorsal view, compare rows 2 – 4 with row 1) and transformations of T3-hypopleurite (hp) tissue into T2-sternopleurite (sp) tissue (lateral view, rows 2 – 4). In the male abdomen of all *ash1* mutants (Abdomen, dorsal view), abdominal segments 5 and 6 (A5 and A6) were transformed toward A4 identity as indicated by either complete (row 2) or partial patchy loss (row 3, 4) of pigmentation in the affected segments. Furthermore, additional abdominal A7 segment structures formed in *ash1* mutants (row 2 - 4).

Altogether, the *ash1*-catalytic mutants *ash1*$^{R1464A\ m-\ z-}$ and *ash1*$^{10\ m-\ z-}$ showed the same posterior-to-anterior homeotic transformations as the *ash1*-null mutant *ash1*$^{22\ m-\ z-}$. However, the severity of the transformations was on average higher in *ash1*$^{22\ m-\ z-}$ flies than in the *ash1*-catalytic mutants, an observation in line with the earlier lethality of the *ash1*-null mutant (para 3.5.1). The images of *ash1*$^{22\ m-\ z-}$ in Figure 14 are not even truly representative as for the other *ash1* mutants. They show rather milder examples of the *ash1*$^{22\ m-\ z-}$ phenotype since adult cuticles could only be mounted from the pharate adult stage onwards. The majority of the *ash1*$^{22\ m-\ z-}$ flies died in an earlier pupal stage (personal observation). It is important to be precise about the *ash1*$^{22\ m-\ z-}$ phenotype, since, as mentioned before, the severest possible *ash1*-null phenotype that is also devoid of maternally loaded Ash1 protein has not been described before. Comparison of the phenotypes of the SET domain mutants *ash1*$^{R1464A\ m-\ z-}$ and *ash1*$^{10\ m-\ z-}$ among each other showed that both developed homeotic transformations of comparable severity.

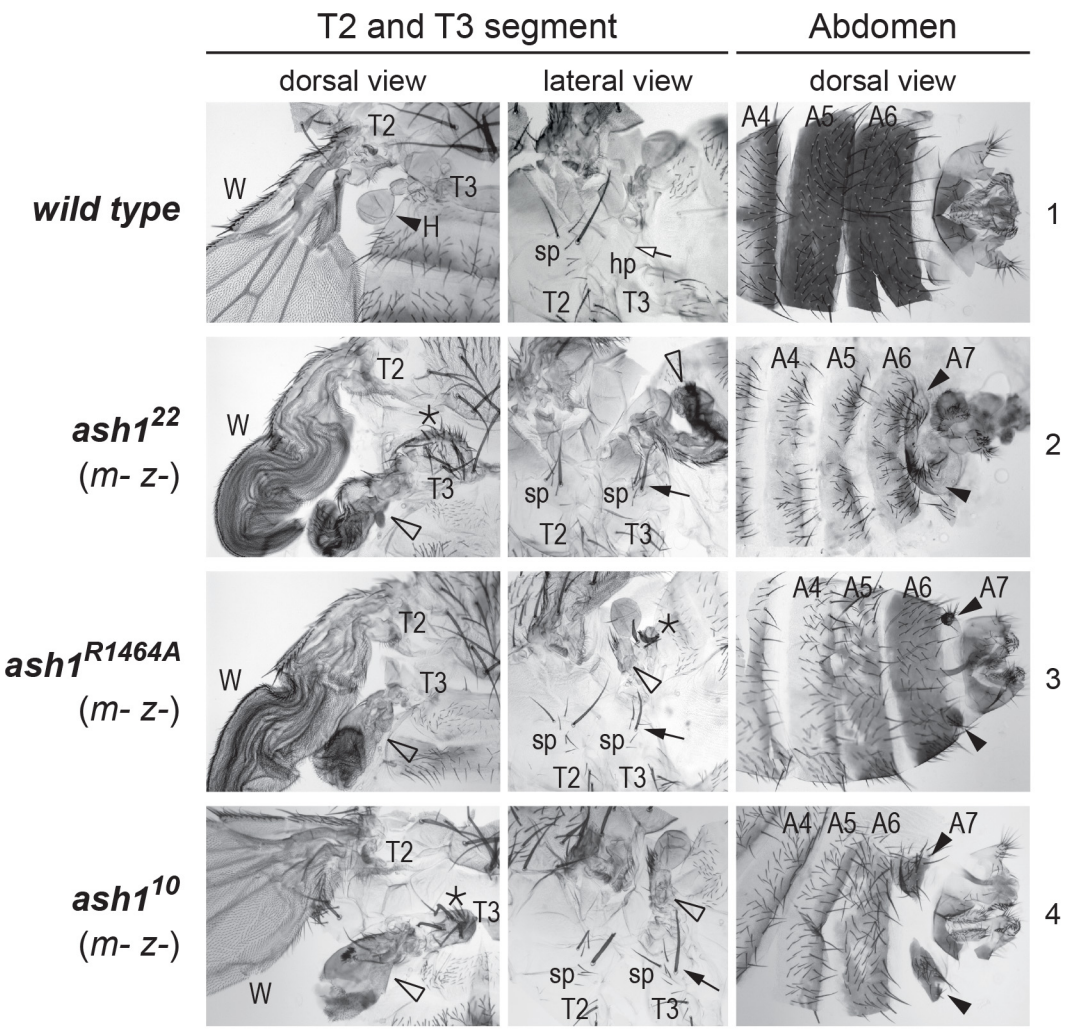


Figure 14. Comparison of morphological transformations in *ash1*-null and *ash1*-catalytic mutant adults. Adult (row 1, 4) or pharate adult (row 2, 3) cuticle preparations of the indicated genotypes with dorsal and lateral views of T2 and T3 thoracic segments and dorsal view of male posterior abdominal segments A4 to A6. Dorsal view of T2/T3: Wild-type flies form the haltere organ (H, black arrowhead) in T3 posterior to the wing (W) in T2. In the *ash1*-null mutant ($ash1^{22\ m-\ z-}$) and the *ash1*-catalytic mutants ($ash1^{R1464A\ m-\ z-}$ and $ash1^{10\ m-\ z-}$), the haltere was partially transformed into wing tissue (empty arrowheads). The asterisks in $ash1^{22\ m-\ z-}$ and $ash1^{10\ m-\ z-}$ images indicate meta- to mesonotum transformations. Lateral view of T2/T3: The sternopleurite (sp) tissue constitutes part of T2, the hypopleurite (hp) tissue part of T3 in *wild type*. The $ash1^{22\ m-\ z-}$, the $ash1^{R1464A\ m-\ z-}$ and the $ash1^{10\ m-\ z-}$ flies exhibited hp-to-sp transformations, which were apparent by sp bristles, that are characteristic for T2-sp tissue and were formed in T3 in the *ash1* mutants (black arrows). Dorsal view of abdomen: Segments A5 and A6 are fully pigmented in wild-type males. In $ash1^{22\ m-\ z-}$ mutants, A5 and A6 pigmentation was almost completely lost, in $ash1^{R1464A\ m-\ z-}$ and $ash1^{10\ m-\ z-}$ flies partially in a patchy pattern. Additional A7 segment structures developed fully in $ash1^{22\ m-\ z-}$ flies or rudimentary in $ash1^{R1464A\ m-\ z-}$ and $ash1^{10\ m-\ z-}$ flies (black arrowhead).
The severity of a specific transformation varied in mutants of one genotype (e.g.: haltere-to-wing transformations in row 3, empty arrowheads, or meta- to mesonotum transformations marked with asterisks in rows 2-4, dorsal versus lateral views). But taken the analysis of n>40 flies of each genotype into account, all presented transformations were on average severer in $ash1^{22\ m-\ z-}$ mutants than in the *ash1*-catalytic mutants. $ash1^{R1464A\ m-\ z-}$ and $ash1^{10\ m-\ z-}$ mutants were comparably strongly affected. T3-to-T2 transformations are indicative of loss of *Ubx*, A5/A6-to-A4 transformations of loss of *Abd-B* expression.

3.5.3 Ash1 catalytic activity is essential for normal expression of the *HOX* genes *Ubx* and *Abd-B*

The homeotic transformations described for *ash1*-null and *ash1*-catalytic mutants in the previous paragraph are indicative of loss of expression of specific *HOX* genes: the transformation of the T3 thoracic segment towards T2 identity of loss of *Ubx* and the transformation of the A5 and A6 abdominal segments towards A4 identity of loss of *Abd-B*.

As a matter of fact, loss of *Ubx* expression in $ash1^{22}$ and other *ash1* EMS mutants has already been demonstrated (LaJeunesse & Shearn, 1995; Klymenko & Müller, 2004). With the experiment presented in this paragraph, I aimed at providing molecular evidence for the presumable loss of *HOX* gene expression in *ash1*-catalytic mutants. Then, the results were to be compared with loss of *HOX* gene expression in the *ash1*-null mutant $ash1^{22\ m-\ z-}$. I probed the primordial tissues of the adult T3 segment in larvae, namely the haltere and 3rd leg imaginal discs, of wild-*type*, $ash1^{22\ m-\ z-}$, $ash1^{R1464A\ m-\ z-}$ and $ash1^{10\ m-\ z-}$ flies with an antibody against Ubx. In wild-type haltere and 3rd leg discs, *Ubx* was expressed in all cells (Figure 15). The *ash1*-null and both *ash1*-catalytic mutants, in contrast, showed patchy, irregular loss of *Ubx* expression in both T3 discs. It is well established, that *Ubx* drives and maintains T3 segment identity (Figure 1). Thus, it can be concluded here, that the loss of *Ubx* expression in the *ash1* mutant larval T3 imaginal discs represents the cause for the morphological *ash1* mutant transformations of parts of the T3 segment into characteristic adult T2 structures as described in the previous paragraph and shown in Figure 14. Consistently, the extent of *Ubx* expression loss in the different *ash1* mutants (Figure 15, Ubx, row 2 – 4) paralleled the severity of the homeotic T3 transformations in the corresponding adults (Figure 14, T2/T3 segment, dorsal view, row 2 – 4): $ash1^{22\ m-\ z-}$ flies were severest, $ash1^{R1464A\ m-\ z-}$ and $ash1^{10\ m-\ z-}$ flies on average equally severely affected.

Besides the thorax, the adult abdominal epidermis exhibited homeotic transformations in the *ash1*-null and *ash1*-catalytic mutants, which are indicative of loss of *Abd-B* expression (Figure 14). The primordia for the abdominal epidermis are histoblasts. This tissue is difficult to dissect. Hence, I analyzed *Abd-B* expression in the *ash1* mutants in the larval central nervous system (CNS), which expresses *Abd-B* in its posterior cells in *wild type* (Figure 15, posterior CNS, row 1). The $ash1^{22\ m-\ z-}$, $ash1^{R1464A\ m-\ z-}$ and $ash1^{10\ m-\ z-}$ flies showed all patchy loss of *Abd-B*. In the fraction of the cells, that retained *Abd-B* expression, the levels of the HOX protein were reduced (row 2 – 4).

In summary, the analyses of the *ash1*-catalytic phenotype with regard to lethality (para 3.5.1), morphology (para 3.5.2) and *HOX* gene expression (this para) led to three main observations: (1) Ash1 catalytic activity is required for viability over more than one generation, (2) the *ash1*-catalytically inactive phenotype is homeotic and (3) the *ash1*-catalytically inactive

phenotype is weaker than the *ash1*-null phenotype. I showed the HOX loss-of-function syndrome of *ash1*-catalytic mutants by testing two independent Ash1 SET domain-mutant alleles, $ash1^{R1464A}$ and $ash1^{10}$. The R1464A mutation abolished enzymatic activity of Ash1 *in vitro* within the detection limit (para 3.4), the N1467I mutation of the $ash1^{10}$ allele has not been tested in this regard. But both alleles, in absence of a wild-type *ash1* locus, caused a homeotic phenotype, almost identical in type and severity, that I consider to be the *ash1*-catalytically inactive phenotype. Morphological aberrations other than homeotic transformations were not observed. The fact that the *ash1*-catalytically inactive phenotype is weaker than the *ash1*-null phenotype could have a technical or a biological reason. Despite its inactivity *in vitro*, the Ash1 protein bearing the R1464A mutation could possess residual activity *in vivo*. In that case, the $ash1^{10}$ allele would not be a true *ash1*-catalytically inactive allele neither. The possible biological explanation is that Ash1 fulfills other, so far undescribed functions in addition to H3K36 di-methylation.

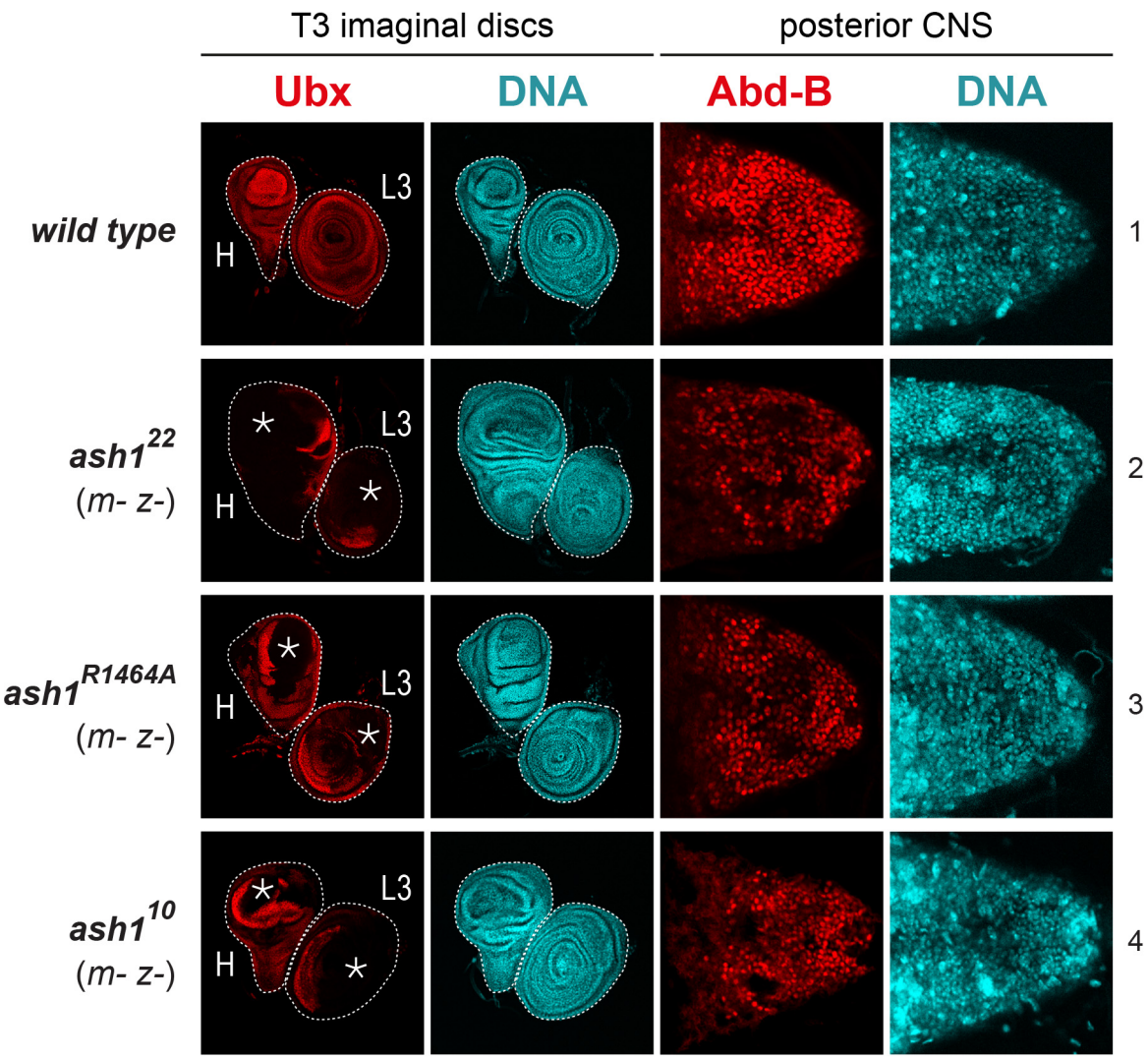


Figure 15. Comparison of expression of the *HOX* genes *Ubx* and *Abd-B* in *ash1*-null and *ash1*-catalytic mutant larvae. Immunostainings of the T3 primordia haltere (H) and 3rd leg (L3) imaginal discs against Ubx (red) and of the posterior CNS against Abd-B (red). Co-staining of DNA visualized nuclei (cyan). In total, tissues from n>40 3rd instar larvae of each indicated genotype were analyzed. Wild-type H and L3 discs expressed *Ubx* in all cells, whereas *Ubx* expression in H and L3 discs of all *ash1* mutants, $ash1^{22\ m-\ z-}$, $ash1^{R1464A\ m-\ z-}$ and $ash1^{10\ m-\ z-}$, was partially lost in a patchy irregular pattern. *ash1* mutants differed in the extent of tissue area showing loss of *Ubx* expression (asterisks). On average, $ash1^{22\ m-\ z-}$ flies were severest, $ash1^{R1464A\ m-\ z-}$ and $ash1^{10\ m-\ z-}$ flies equally severely affected. Moreover, the sizes of the H discs in the *ash1* mutants compared to the size of the wild-type H disc were enlarged indicating transformation into wing primordial tissue (all images cropped to scale). *Abd-B* was expressed in the posterior cells of the CNS of wild-type larvae. The posterior CNS of $ash1^{22\ m-\ z-}$, $ash1^{R1464A\ m-\ z-}$ and $ash1^{10\ m-\ z-}$ mutants showed in some tissue areas patchy loss and in others reduction of *Abd-B* expression.

3.6 The *MRG* mutant phenotype resembles the *ash1* mutant phenotype

In the beginning of this work, I showed by TAP from *Drosophila* (para 3.1), by binding studies with recombinant proteins and by HMTase assays (para 3.2, 3.3, 3.4) that the MRG15 protein physically interacts with Ash1 and probably strongly enhances Ash1 HMTase activity *in vitro*. The consequential question, if and to what extent MRG15 is important for the enzymatic activity of Ash1 *in vivo*, was addressed by the genetic experiments presented in the following. Specifically, the *MRG15*-null mutant phenotype was analyzed and compared with the homeotic *ash1*-catalytically inactive phenotype, which had been characterized in depth in the previous paragraph 3.5. Moreover, I would like to point out, that the analysis of the *MRG15*-null mutant phenotype per se was already of interest since the genetics of *MRG15* has been poorly studied in *Drosophila* so far.

In a first step, I tested if the only available endogenous *MRG15* mutant allele at that time, $MRG15^{j6A3}$, would meet the requirement for the planned analyses to be a null allele. The $MRG15^{j6A3}$ allele has been generated in the course of the 'BDGP gene disruption project' (Spradling *et al*, 1999). It carries an 11 kb-long *P* element insertion in the second *MRG15* exon, that disrupts the chromo barrel domain (Figure 16A). Hypothetically, the expression of a truncated MRG15 protein, which bears the complete MRG domain, is possible beginning from alternative start codons downstream of the *P* element in $MRG15^{j6A3}$. The MRG domain most likely harbors the binding site of the Ash1 FxLP motif (para 3.3) and its presence could suffice for formation of Ash1:MRG15 complexes. To test for MRG15 expression from the $MRG15^{j6A3}$ allele, western blot analysis was performed on total extracts from imaginal discs of $MRG15^{j6A3\ m+\ z-}$ 3rd instar larvae. $MRG15^{j6A3\ m+\ z-}$ animals contained maternally loaded but not zygotically expressed wild-type MRG15. These flies were trans-heterozygous for the $MRG15^{j6A3}$ allele and the *Df(3R)BSC741* allele, that is a chromosomal deletion in which the entire *MRG15* gene and neighboring genes are removed. When the western blot was probed with an MRG15-specific antibody, no full-length MRG15 protein was detected in the $MRG15^{j6A3\ m+\ z-}$ extracts (Figure 16B). The MRG15 protein and RNA loaded maternally into the embryo is expected to be degraded or diluted out by cell division in imaginal discs as late as 3rd instar larval stage. But, notably, one band appeared repeatedly in the $MRG15^{j6A3\ m+\ z-}$ extract that was not visible in *wild type*. This band could, based on its migration behavior in the gel, represent truncated MRG15 expressed from one of the alternative start codons in the $MRG15^{j6A3}$ allele marked in Figure 16A. In addition to MRG15, the $MRG15^{j6A3\ m+\ z-}$ extracts were analyzed for H3K36me3 by western blot. The binding

specificity of the MRG domain of MRG15 to H3K36me3 could possibly indirectly affect the levels of the histone modification. However, in the *MRG15*$^{j6A3\ m+\ z-}$ extracts tested here, the H3K36me3 bulk levels were comparable to the levels in *wild type*. Altogether, the main conclusion of the analysis of the *MRG15*^{j6A3} allele was that its identity as a null allele is questionable.

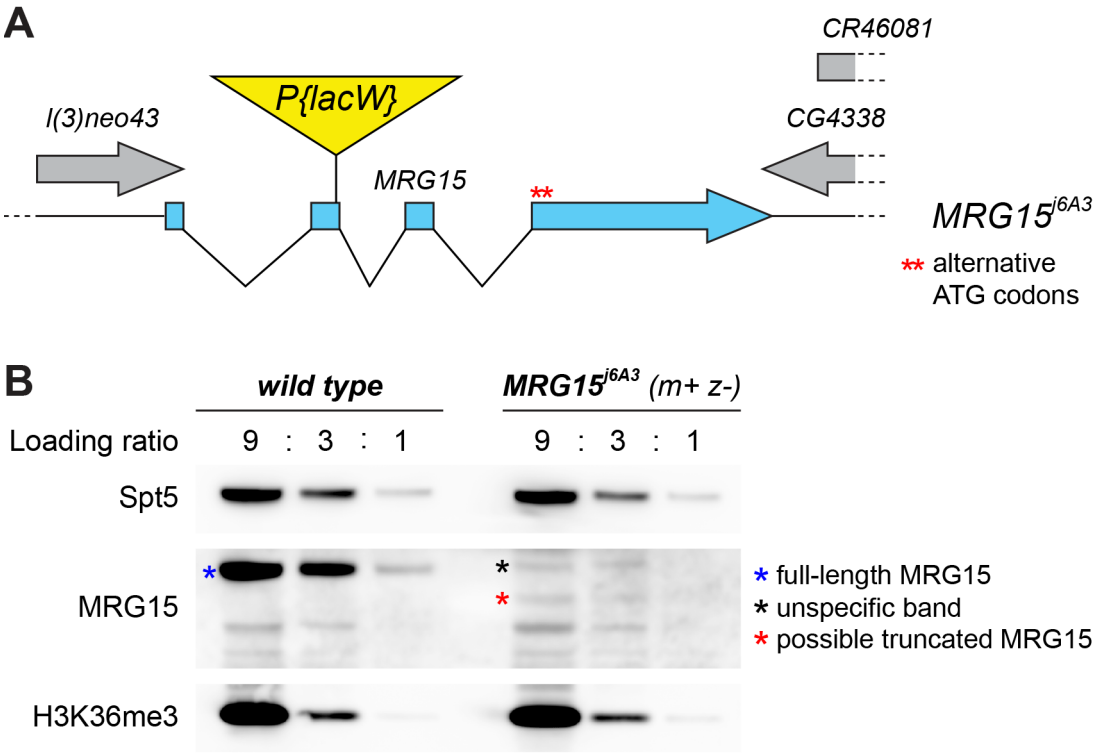
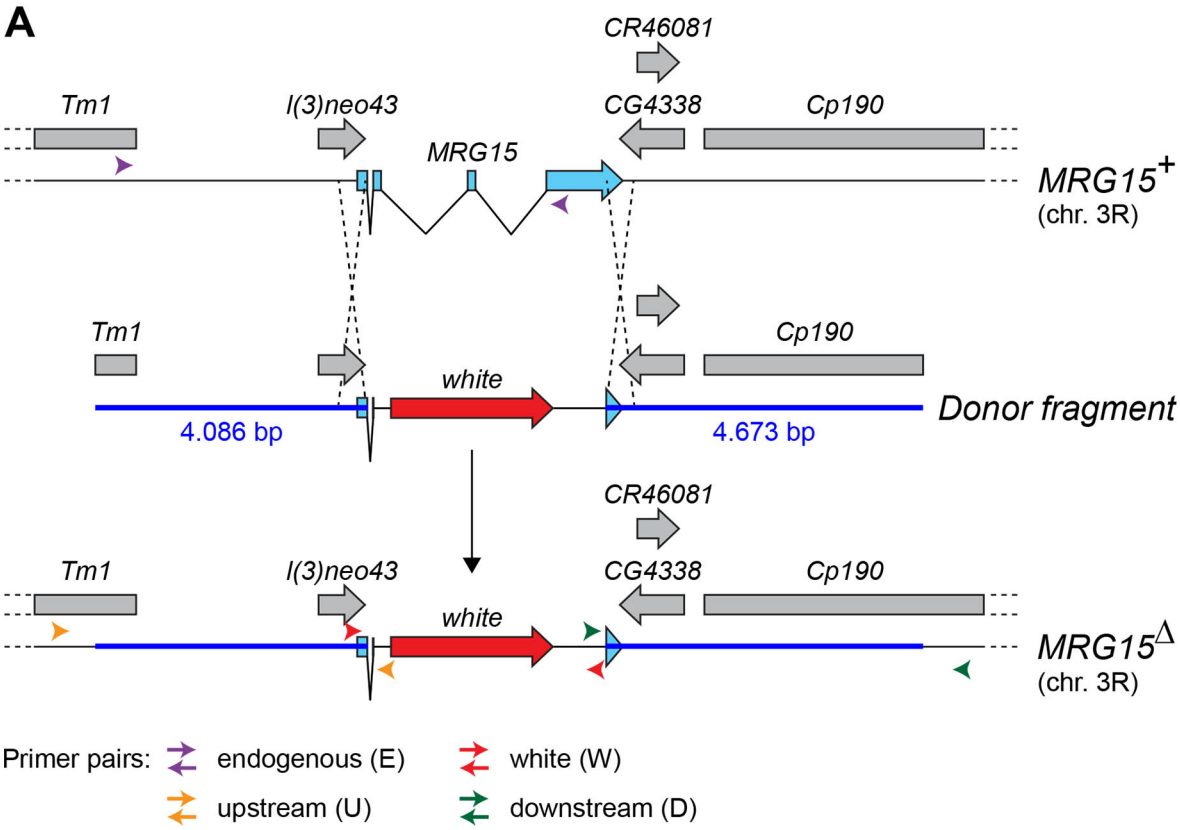


Figure 16. Structure and expression of the *MRG15*^{j6A3} allele. (A) Schematic representation of the *MRG15*^{j6A3} allele and neighboring genes. In *MRG15*^{j6A3}, the *MRG15* gene is disrupted by the insertion of the *P* element *P{lacW}* in the second exon. Alternative start codons downstream of *P{lacW}* are indicated with red asterisks. **(B)** Western blot on serial dilutions of total extracts from T2 (wing) and T3 imaginal discs of wild-type and *MRG15*$^{j6A3\ m+\ z-}$ 3rd instar larvae. The aim was to test the *MRG15*^{j6A3} allele for *MRG15* expression products. Full-length MRG15 protein (blue asterisk in *wild type*) was not found in the *MRG15*$^{j6A3\ m+\ z-}$ extract. The band marked with a black asterisk in the *MRG15*$^{j6A3\ m+\ z-}$ extract is unspecific and runs higher than full-length MRG15 as was more clearly observable in western blots after longer SDS PAGE runs (not shown here). The red asterisk labels a band that was repeatedly detected by the MRG15 antibody only in the *MRG15*$^{j6A3\ m+\ z-}$ extract and not in *wild type*. This band could represent a truncated MRG15 protein expressed from one of the alternative start codons shown in the scheme in (A). Spt5 served as loading control. Bulk levels of H3K36me3 were not affected in the *MRG15*^{j6A3}-mutant extracts.

Since it could not be proven, that the *P* element insertion allele $MRG15^{j6A3}$ is a functional null allele, I generated a molecularly defined null allele, $MRG15^{\Delta}$, in which the major part of the *MRG15* gene, including 90% or more of the MRG domain and of the chromo barrel domain, was deleted. As deletion method, ends-out targeting according to Gong & Golic, 2003 was chosen. How ends-out targeting works, how the technique was applied to generate the $MRG15^{\Delta}$ allele and how the deletion allele was then verified, is sketched in Figure 17A and described in detail in the 'Materials and Methods' section, in the paragraphs 2.1.10 and 2.3.7. In brief, the *MRG15* sequence to be removed was replaced by the *mini-white* marker gene by homologous recombination. Recombination occurred between sequences of the target genomic region and homologous sequences (homologous arms), that flanked the *mini-white* gene in a donor fragment. The donor fragment had originally been an inserted *P* element, that was excised from a *P[donor]* allele before recombination.

Lastly, the obtained $MRG15^{\Delta}$ allele was tested by $MRG15^{\Delta}$-specific PCR amplification reactions on genomic DNA from $MRG15^{\Delta}$ homozygotes (Figure 17B), by sequencing and by a genetic rescue assay (para 2.3.7). All tests together showed that the *MRG15* gene has been deleted successfully and that the generated allele causes a phenotype specific for *MRG15*. The $MRG15^{\Delta}$ allele was therefore considered as suitable for characterization of the *MRG15*-null mutant phenotype and for comparison with the *ash1*-catalytically inactive phenotype. The respective experiments are presented in the following paragraphs.



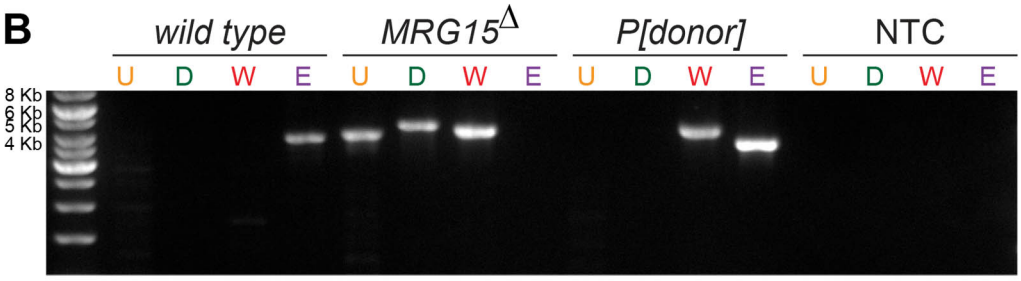


Figure 17. Generation and verification of the *MRG15*$^{\Delta}$-deletion allele. (A) Scheme of the generation of the *MRG15*$^{\Delta}$ allele by the ends-out targeting technique. The major part of the *MRG15* gene in the wild-type *MRG15* genomic locus (top) was replaced with the *mini-white* marker gene provided by an excised *P* element (donor fragment, middle). The mechanism behind *MRG15* replacement was homologous recombination between homologous arms flanking the *mini-white* gene in the donor fragment and homologous sequences in the wild-type *MRG15*$^{+}$ genomic locus. The resulting *MRG15*$^{\Delta}$-deletion allele is shown at the bottom. **(B)** Quality control of the *MRG15*$^{\Delta}$-deletion allele by *MRG15*$^{\Delta}$-specific PCR reactions on genomic DNA from *wild type*, *MRG15*$^{\Delta}$ homozygotes and the *P[donor]* strain, which beared the non-excised donor fragment on chromosome 2. The specificity of all primer pairs used was as illustrated in (A) and as follows: endogenous (E) for the wild-type *MRG15*$^{+}$ genomic locus, white (w) for the *mini-white* gene, upstream (U) and downstream (D) for the *MRG15*$^{\Delta}$ allele. The major part of the wild-type *MRG15* gene was successfully replaced by *mini-white*, and the homologous arms from the donor fragment had recombined correctly based on the size of the U- and D-PCR products from the *MRG15*$^{\Delta}$ genomic DNA. NTC: no template control; chr.: chromosome.

3.6.1 *MRG15*-null mutant flies die throughout all developmental stages, but may reach adulthood

Preliminary analyses of the generated $MRG15^{\Delta}$-null allele (Figure 17) showed that a considerable fraction of the $MRG15^{\Delta}$ homozygotes completes development up to adult stage. But flies trans-heterozygous for the $MRG15^{\Delta}$ allele and the *Df(3R)BSC741* deletion, that uncovers *MRG15* and neighboring genes, exhibited higher survival rates than the $MRG15^{\Delta}$ homozygotes, possibly due to second site mutations on the chromosome with the $MRG15^{\Delta}$ deletion. For that reason, the systematic analyses of the lethality and of further phenotypic aspects described in the paragraphs 0 and 3.6.3 of *MRG15*-null mutants were performed with the $MRG15^{\Delta}/Df(3R)BSC741$ trans-heterozygotes, for simplicity called $MRG15^{\Delta}$ flies in the following.

Survival rates were determined of $MRG15^{\Delta\ m+\ z-}$ as well as of $MRG15^{\Delta\ m-\ z-}$ flies in the pupal and in the adult stage (see the histogram in Figure 18 for results). Both genotypes were devoid of zygotically expressed MRG15, $MRG15^{\Delta\ m+\ z-}$ flies contained maternally deposited MRG15, $MRG15^{\Delta\ m-\ z-}$ flies did not. As already indicated above, more than 50% of the analyzed $MRG15^{\Delta\ m+\ z-}$ mutant larvae survived into adult stage. They exhibited a mutant phenotype, but were fertile, so that $MRG15^{\Delta\ m-\ z-}$ progeny could be obtained without generation of germ line clones. Remarkably, a small fraction of these $MRG15^{\Delta\ m-\ z-}$ flies, again, underwent metamorphosis and eclosed from the pupal case (Figure 18), just like the *ash1*-catalaytic inactive mutants, $ash1^{R1464A\ m-\ z-}$, did (Figure 13). But it was impossible to maintain flies of both genotypes, $MRG15^{\Delta\ m-\ z-}$ or $ash1^{R1464A\ m-\ z-}$, as stocks. Less than 20% of the collected $MRG15^{\Delta\ m-\ z-}$ larvae actually reached adulthood. The majority of these *MRG15*-null mutants died before in all stages throughout development. Therefore, the survival rate of the $MRG15^{\Delta\ m-\ z-}$ flies was considerably lower than the survival rate of the $ash1^{R1464A\ m-\ z-}$ flies (60% of 1st and 2nd instar larvae hatched as adults, Figure 13). A possible reason for this could be AMC-independent functions of MRG15 as MRG15 has been purified with several other complexes that fulfill crucial roles for viability like the Tip60 complex (para 3.2). But overall, in light of these Ash1-independent functions of MRG15, the *MRG15*-null phenotype in terms of lethality is surprisingly weak.

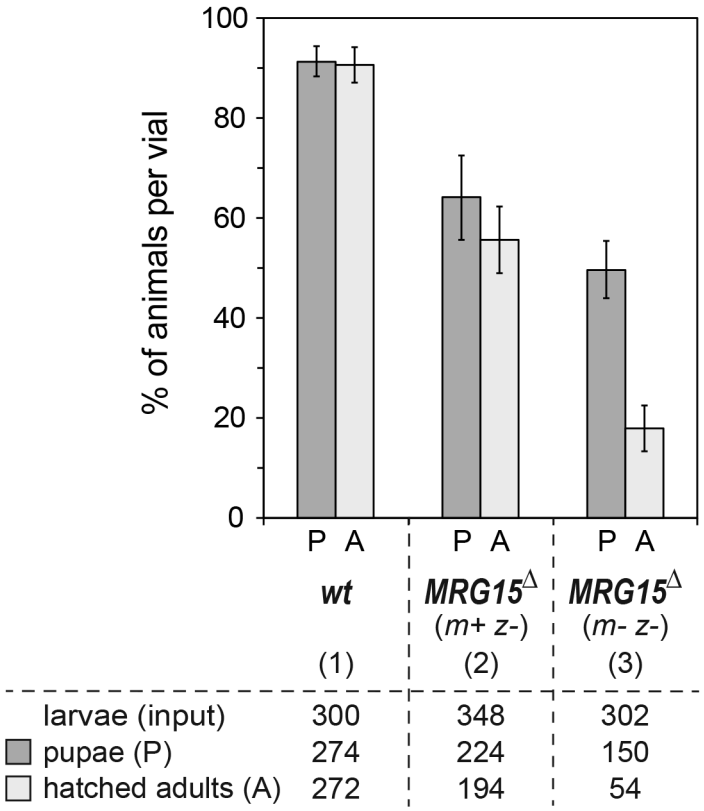


Figure 18. Survival rates of *MRG15*$^{\Delta}$-null mutants throughout development. This assay was set up exactly as described in Figure 13 for the *ash1* mutants. A specific number of 1st and 2nd instar larvae of each indicated genotype (input) was collected. Then the fractions of flies from this input, that developed into pupae (P) and that eclosed as adults from the pupal case (A) were counted. The numbers of P and A of each genotype are presented as mean percentages from the input in the histogram. For calculation of the mean percentages and their standard deviations, see Figure 13, legend. Fractions of the *MRG15*$^{\Delta\ m+\ z-}$- and, notably, of the *MRG15*$^{\Delta\ m-\ z-}$-null mutants developed up to pupal stage and even hatched as adults. But the majority, more than 80%, of the *MRG15*$^{\Delta\ m-\ z-}$ flies died before, throughout development from larval to adult stage. *MRG15*$^{\Delta}$ flies were trans-heterozygous for the *MRG15*$^{\Delta}$ allele and the chromosomal deletion *Df(3R)BSC741*.

3.6.2 Adult *MRG15*-null mutant exhibits trxG-like homeotic transformations

ash1-catalytic and -null mutant adults exhibit a trxG-like homeotic phenotype with posterior-to-anterior morphological transformations as described in detail previously in para 3.5.2. The survival of a fraction of the $MRG15^{\Delta\ m-\ z-}$ flies into adult stage permitted also analysis of their adult cuticles. This analysis revealed high similarities of the *MRG15*-null mutant phenotype with the homeotic *ash1*-mutant phenotypes (Figure 14, Figure 19). The thoracic segment T3 of $MRG15^{\Delta\ m-\ z-}$ flies transformed toward T2 identity: The T3-haltere organ partially developed into T2-wing tissue (Figure 19, T2/T3 segment, dorsal view) and the T3-hypopleurite tissue into T2-sternopleurite tissue (T2/T3 segment, lateral view). In the male $MRG15^{\Delta\ m-\ z-}$ abdomen (Figure 19, Abdomen), the A5 segment partially adopted A4 identity. An additional male abdominal segment A7 was formed.

All described trxG-like homeotic transformations were observed in each of the analyzed $MRG15^{\Delta\ m-\ z-}$ flies (n>40), but the degree of severity varied. The adult cuticle sections in Figure 19 display the estimated average $MRG15^{\Delta\ m-\ z-}$-null mutant phenotype. For a more comprehensive view on the *MRG15*-null mutant phenotype I would like to add that already the $MRG15^{\Delta\ m+\ z-}$ flies, which had been supplied with MRG15 maternal load, showed homeotic transformations (not presented here). These transformations were less pronounced than in their progeny devoid of MRG15 maternal load, the $MRG15^{\Delta\ m-\ z-}$ flies.

In relation to the morphological transformations in *ash1* mutants, the mean $MRG15^{\Delta\ m-\ z-}$-null mutant phenotype was very similar to, but slightly milder than the *ash1*-catalytically inactive phenotype of the $ash1^{R1464A\ m-\ z-}$ flies (Figure 19, row 2 vs. row 3).

The most surprising observation made in the analysis of the *MRG15*-null mutant adult morphology was the specificity of the phenotype. Apart from the described trxG-like homeotic transformations, no other aberrations were detected in the $MRG15^{\Delta\ m-\ z-}$-adult cuticulae despite the association of MRG15 with several nuclear multi-protein complexes that fulfill quite varied functions (para 3.2).

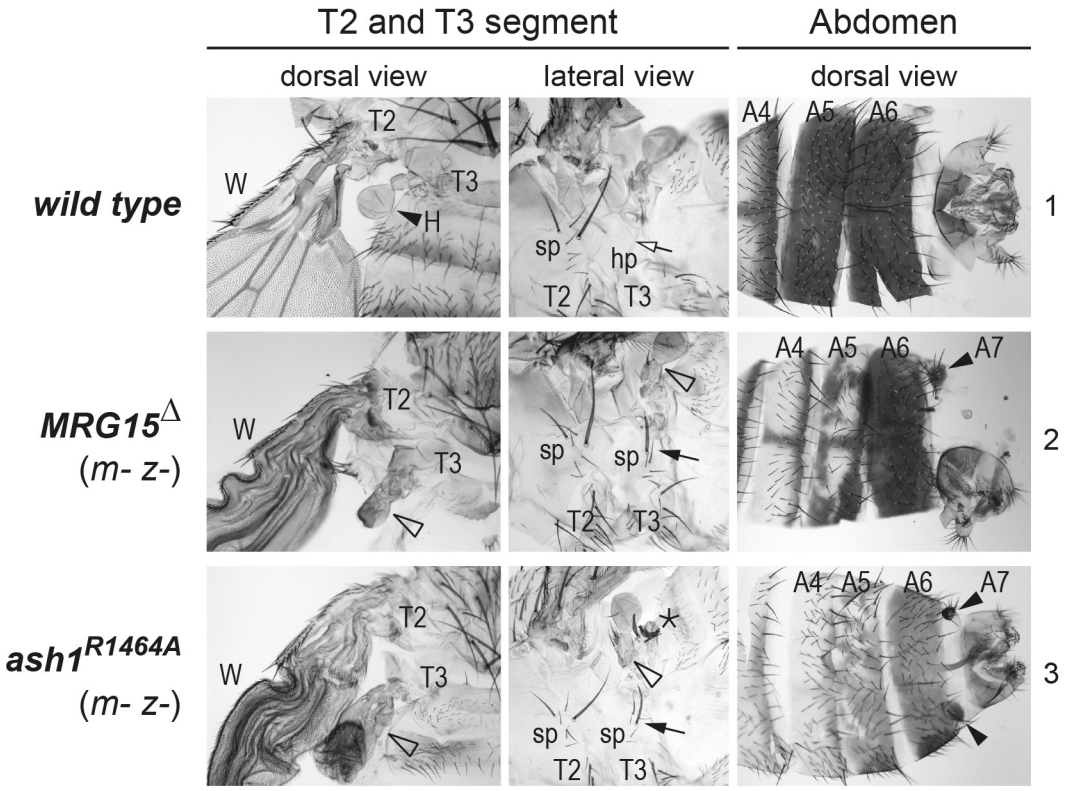


Figure 19. Analysis of adult morphological transformations in *MRG15*$^{\Delta}$-null mutant and comparison with *ash1*-catalytically inactive phenotype. Adult (row 1) or pharate adult (row 2, 3) cuticle preparations of indicated genotypes with dorsal and lateral views of T2 and T3 thoracic segments and dorsal view of male posterior abdominal segments A4 to A6. Pictures of wild-type and *ash1*$^{R1464A\ m-\ z-}$ cuticles are taken from Figure 14, since the same body segments were analyzed for transformations here. Dorsal view of T2/T3: The haltere organ in T3 marked with H and a black arrowhead in *wild type* showed partial transformation into T2-wing tissue in *MRG15*$^{\Delta\ m-\ z-}$ flies (row 2, empty arrowhead), like the haltere organ in *ash1*$^{R1464A\ m-\ z-}$ flies did (row 3, empty arrowhead). Lateral view of T2/T3: The hypopleurite (hp) tissue in T3 (empty arrow in *wild type*) was partially transformed into T2-sternopleurite (sp) tissue in *MRG15*$^{\Delta\ m-\ z-}$ flies as visible by sp bristles developed in T3 of the MRG15 mutants (row 2, black arrow). As described before, the *ash1*$^{R1464A\ m-\ z-}$ flies also exhibited hp-to-sp transformations and in some cases like here meta- to mesonotum transformations (row 3, asterisk). Dorsal view of abdomen: In the *MRG15*$^{\Delta\ m-\ z-}$ mutant, just like in the *ash1*$^{R1464A\ m-\ z-}$ mutant, pigmentation in the abdominal segment A5 and in some cases in A6 was partially lost in a patchy pattern (row 2), which indicates transformation of A5 (and A6) toward A4 identity. *MRG15*$^{\Delta\ m-\ z-}$ flies also developed an additional abdominal segment A7 (row 2, black arrowhead), like *ash1*$^{R1464A\ m-\ z-}$ flies did (row 3, black arrowhead). The *MRG15*$^{\Delta\ m-\ z-}$-null mutant flies exhibited the same kinds of trxG-like homeotic transformations of slightly milder severity as the *ash1*-catalytically inactive mutant flies *ash1*$^{R1464A\ m-\ z-}$ (n>40 animals per genotype analyzed).

3.6.3 Expression of *Ubx* and *Abd-B* is partially lost in *MRG15*-null mutant tissue

The homeotic transformations that were observed in the third thoracic segment T3 or in the posterior abdominal segment A5 in *MRG15*$^{\Delta}$-null mutant adults (para 3.5.2) suggested loss of expression of the *HOX* genes *Ubx* or *Abd-B* in the primordia of these body segments, as shown before in paragraph 3.5.3 for the *ash1* mutants. Consequently, in an analogous manner to the immunostainings of the *ash1* mutants (Figure 15), the primordia of the T3 segment, the haltere and 3rd leg imaginal discs of *MRG15*$^{\Delta\ m-\ z-}$ 3rd instar larvae were probed with an antibody against Ubx. In addition, the CNS tissues of *MRG15*$^{\Delta\ m-\ z-}$ flies were stained against Abd-B. The CNS was chosen for technical reasons in place of the primordial tissues of the posterior abdominal segments to analyze *Abd-B* expression. Representative pictures of the immunostainings against the HOX proteins in *MRG15*$^{\Delta\ m-\ z-}$ primordial larval tissues are set into direct comparison to corresponding immunostainings of the *ash1*-catalytically inactive flies *ash1*$^{R1464A\ m-\ z-}$ in Figure 20. They show that *Ubx* expression was indeed lost in haltere and 3rd leg imaginal discs of *MRG15*$^{\Delta\ m-\ z-}$ flies in a patchy and irregular pattern. The Abd-B stainings revealed either reduction or patchy loss of *Abd-B* expression in the posterior CNS cells of the *MRG15*$^{\Delta\ m-\ z-}$ mutant. Thus, taken the loss of *HOX* gene expression in *MRG15*$^{\Delta\ m-\ z-}$ flies together with the homeotic morphological transformations in the same mutant described in paragraph 3.5.3, it can be stated here that the *MRG15*-null phenotype is homeotic.

In the *ash1*$^{R1464A\ m-\ z-}$ flies, loss of *Ubx* and *Abd-B* expression in T3 discs and in the CNS as described before in paragraph 3.5.3 and shown in Figure 20 occurred in a similar manner to the expression loss in *MRG15*$^{\Delta\ m-\ z-}$ flies. The severity of the phenotype was only slightly milder in *MRG15*$^{\Delta\ m-\ z-}$ flies. Overall, taken not only the loss of *HOX* gene expression and the resulting adult homeotic transformations (para 3.5.2) but also the partial survival into adults of both mutants into account (para 3.5.1), the *MRG15*$^{\Delta\ m-\ z-}$-null phenotype was very similar to the *ash1*-catalytically inactive phenotype of the *ash1*$^{R1464A\ m-\ z-}$ flies. This resemblance of phenotypes together with the likely boosting effect of MRG15 on Ash1 enzymatic activity *in vitro* strongly indicate that MRG15 also plays an important role for the HMTase activity of Ash1 *in vivo*. The role in Ash1 activity might even be the major cellular function of MRG15 in light of the relative weakness and the specificity of the *MRG15*-null mutant phenotype for loss of *HOX* gene function.

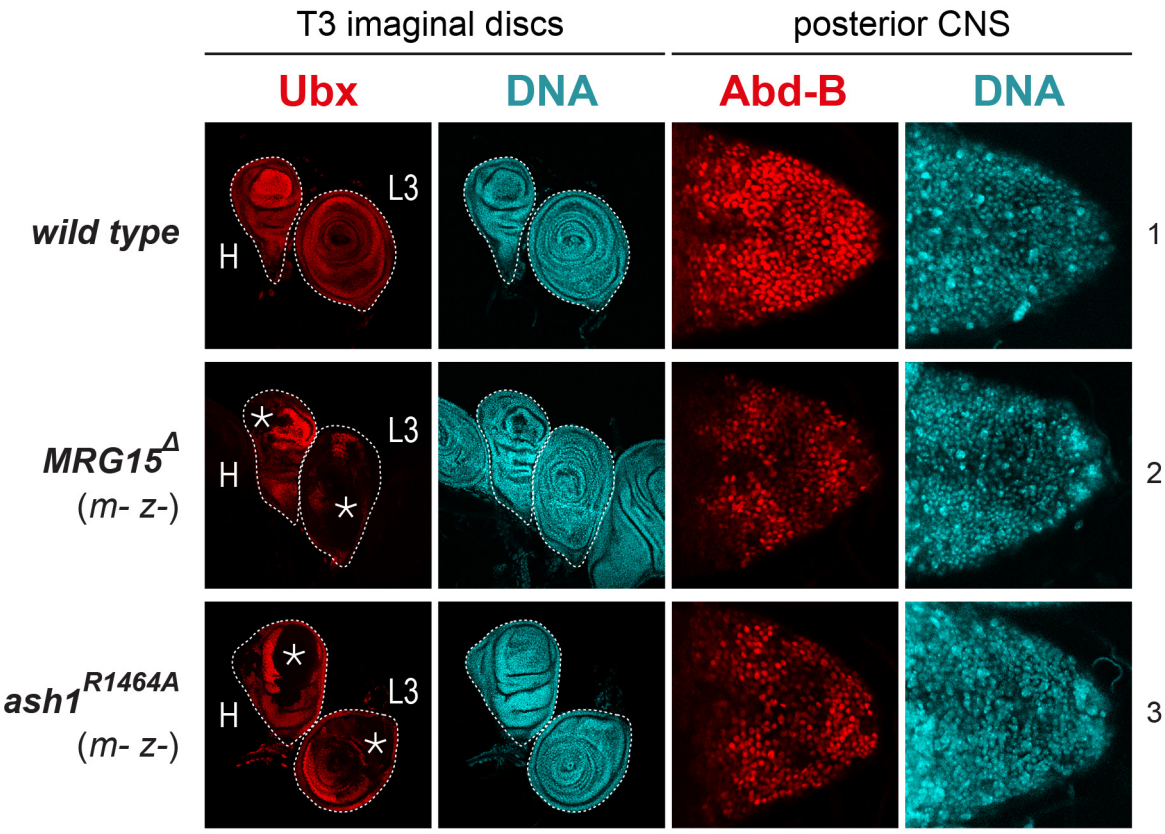


Figure 20. Expression of the *HOX* genes *Ubx* and *Abd-B* in *MRG15*$^{\Delta}$-null mutant and, in comparison, in *ash1*^{R1464A}-catalytically inactive mutant. T3 primordia haltere (H) and 3rd leg (L3) imaginal discs of more than 40 3rd instar larvae of each indicated genotype were probed with an antibody against Ubx (red). The posterior CNS tissues of these larvae were probed with an antibody against Abd-B (red). DNA-specific staining visualized nuclei (cyan). The pictures of wild-type and *ash1*$^{R1464A\ m-\ z-}$ tissues were overtaken from Figure 15. In *wild type*, Ubx was detected over the entire tissue area in all cells of the T3 imaginal discs H and L3. In contrast, *MRG15*$^{\Delta\ m-\ z-}$ flies showed a patchy, irregular loss of *Ubx* expression in H and L3 discs, just as *ash1*$^{R1464A\ m-\ z-}$ flies did. Expression of *Abd-B*, that was detected in the posterior CNS cells in *wild type*, was in *MRG15*$^{\Delta\ m-\ z-}$ flies either lost in a patchy pattern or reduced like in *ash1*$^{R1464A\ m-\ z-}$ flies. The mean extent of loss of *Ubx* and *Abd-B* expression, assessed by the size of tissue area affected (row 2/3, Ubx, asterisk), was slightly weaker in the *MRG15*$^{\Delta\ m-\ z-}$ mutant than in the *ash1*$^{R1464A\ m-\ z-}$ mutant.

3.7 H3K36 di-methylation levels are decreased at the Ash1-target gene *Ubx*, but not genome-wide

The identity of Ash1 as a transcriptional regulator of *HOX* genes is indisputable. Moreover, genetic analyses presented in the paragraphs 3.5 and 3.6 showed on one hand, that mutants of the Ash1 interactor MRG15 exhibit a homeotic phenotype and, on the other hand, that it is the capacity of Ash1 to di-methylate H3K36, that is crucial for *HOX* gene regulation. However, whether the function of Ash1 is restricted to *HOX* genes is a controversial matter in the literature (para 1.1.4). It has been claimed that Ash1 acts as a genome-wide transcription factor (Gregory *et al*, 2007; Kockmann *et al*, 2013). In that case, Ash1 would probably contribute substantially to the total levels of H3K36me2. To test this assumption, I generated total cellular extracts from the wing (T2) and the haltere and 3rd leg (T3) larval imaginal discs of *wild type*, of the *ash1*-null mutant flies $ash1^{22\ m+\ z-}$ and of the *MRG15*-null mutant flies $MRG15^{\Delta\ m-\ z-}$. The $ash1^{22\ m+\ z-}$ flies were homozygous for the $ash1^{22}$-null allele, and the $MRG15^{\Delta\ m-\ z-}$ flies trans-heterozygous for the deletion alleles $MRG15^{\Delta}$ and *Df(3R)BSC741*. The total cellular extracts were analyzed for H3K36me2 by western blot. As shown in Figure 21, the H3K36me2 bulk levels were not detectably decreased, neither in the *ash1*-null nor in the *MRG15*-null mutant, which points to a rather gene-specific function of Ash1 and not to a genome-wide role in transcription.

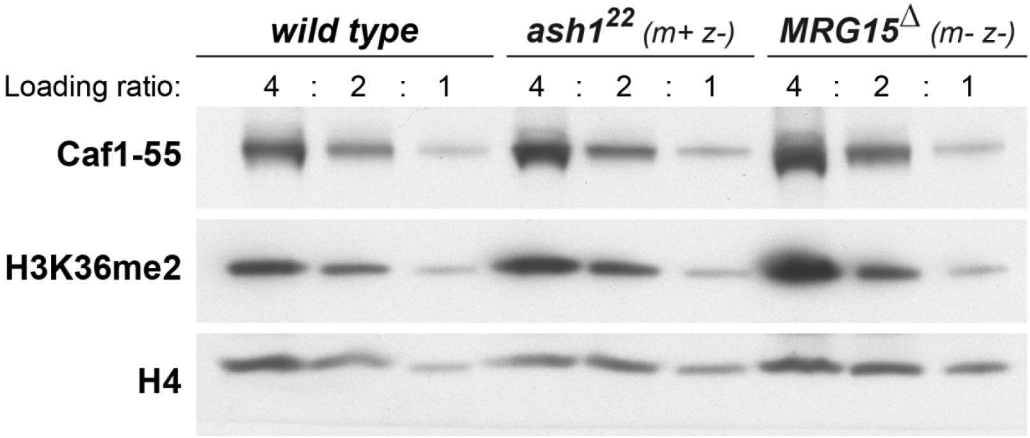


Figure 21. Analysis of H3K36me2 bulk levels in $ash1^{22}$- and $MRG15^{\Delta}$-null mutants.
Western blot on serial dilutions of total cell extracts from T2 and T3 imaginal discs of wild-type, $ash1^{22\ m+\ z-}$ and $MRG15^{\Delta\ m-\ z-}$ 3rd instar larvae. The membrane was probed against H3K36me2 and, as loading controls, against Caf1-55 and H4. Genome-wide H3K36me2 bulk levels were not decreased in $ash1^{22}$- and $MRG15^{\Delta}$-null mutants.

After not having observed an effect on genome-wide H3K36me2 by the absence of Ash1, the next questions asked were if and how the H3K36me2 landscape at the single-gene level at the well-known Ash1 target gene *Ubx* is affected in *ash1* mutants. To this end, a ChIP assay with an antibody against H3K36me2 was performed on chromatin isolated from the T3 imaginal discs of wild-type and $ash1^{22\ m+\ z-}$ flies. The T3 discs were chosen as tissues to be analyzed due to their homogenous *Ubx* expression in all cells (Figure 15). Levels of H3K36me2 were then determined by qPCR at one or more locations in the transcribed regions of *Ubx* and the control genes *wingless* (*wg*), *teashirt* (*tsh*), *Lamin* and *defective proboscis extension response 12* (*dpr12*) (Figure 22). Work from the Jürg Müller laboratory not presented here had shown that *wg*, *tsh* and *Lamin* are expressed, but not regulated by Ash1 in T3 discs (Schmähling *et al*, 2018), whereas *dpr12* was chosen as an example for an inactive gene in the analyzed tissues. Regarding the gene regions to test for H3K36me2, one has to bear in mind that it is not known where Ash1 acts along the gene body. But at many *Drosophila* genes, H3K36me2 accumulates from the middle stretch towards the 3' end (Bell *et al*, 2007). Based on this observation and on ChIP-seq profiles of H3K36me2 in wild-type *Drosophila* embryos (modENCODE ID 6388), the regions to be amplified by qPCR were selected (see Figure 22 and Table 11 for precise amplicon locations). At *Ubx*, also amplicons in the 5' region were analyzed since H3K36me2 has been shown to inhibit H3K27me3 deposition by PRC2 when present on the PRC2 substrate histone H3 tail (Schmitges *et al*, 2011; Voigt *et al*, 2012). The H3K27me3 mark has been suggested to spread at *Ubx* in a linear manner from the upstream control region, where it is detected independent of transcription, into the 5' part of the gene body in the transcriptionally silent state (Papp & Müller, 2006).

The results of the ChIP-qPCR assay are presented in the histogram in Figure 22 as percentages of the chromatin precipitated with the H3K36me2 antibody from the input chromatin. In line with the role of H3K36me2 as a hallmark for actively transcribed chromatin, all genes expressed in the T3 discs were enriched for H3K36me2 in their coding regions in wild-type tissue. At the *Ubx* gene in *wild type*, H3K36me2 was analyzed and detected in the very 3' end as well as in the first exon, but not at the TSS, which is likely to be located within a nucleosome-poor region while transcription is ongoing. The low H3K36me2 levels at the inactive gene *dpr12* were considered as background. In the $ash1^{22\ m+\ z-}$-null mutant tissues, H3K36me2 levels did not differ significantly from wild-type levels at the actively transcribed genes *wg*, *tsh* and *Lamin*, whose expression is independent from Ash1. However, at the Ash1 target gene *Ubx*, H3K36me2 was strongly decreased at the 5' and at the 3' end in the *ash1*-null mutants. The fact that H3K36me2 was decreased but not completely eliminated at *Ubx* is remarkable and, as a matter of discussion, might provide insight into the mechanism by which Ash1 regulates gene expression. The residual

H3K36me2 at *Ubx* was most likely generated by the other *Drosophila* H3K36 di-methyltransferase, NSD, in the absence of Ash1. A contribution by maternal Ash1, that is loaded into the embryo, to H3K36me2 levels at *Ubx* in *ash1*$^{22\ m+\ z-}$ flies at larval stage is improbable after that many cell divisions.

In summary, the ChIP-qPCR analysis demonstrated that the H3K36me2 levels along the coding region of the actively transcribed Ash1 target gene *Ubx* are decreased in absence of Ash1.

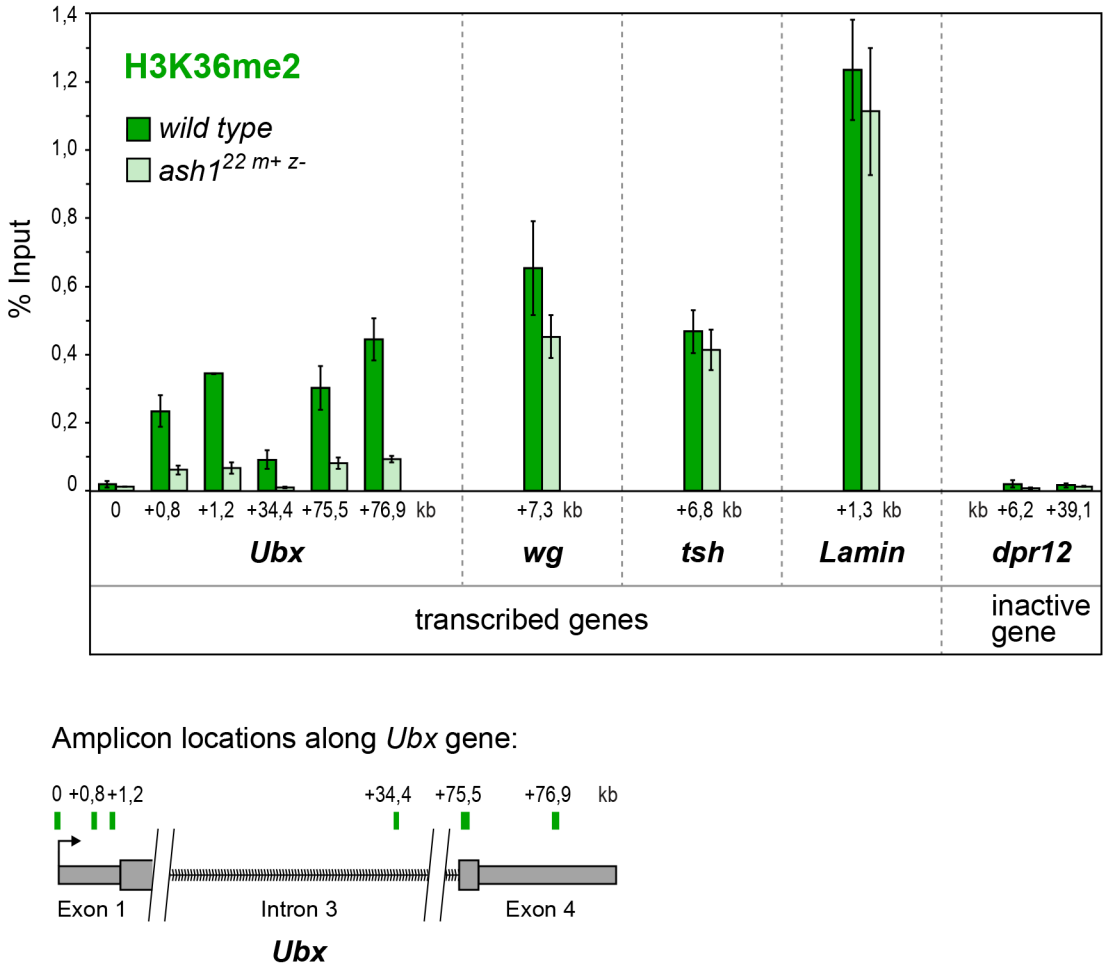


Figure 22. Analysis of H3K36me2 levels at the Ash1 target gene *Ubx* in *ash1*^{22}-null mutants. The histogram shows H3K36me2 enrichment at the genes *Ubx*, *wg*, *tsh*, *Lamin* and *dpr12* in chromatin, that was isolated from T3 imaginal discs of wild-type (dark-green bars) and *ash1*$^{22\ m+\ z-}$ (light-green bars) 3rd instar larvae and then analyzed by ChIP-qPCR. At each gene, one or several regions were amplified. The positions of the amplicons are given as distances from the TSS of each gene in kilobases (kb) in the histogram and for *Ubx*, in addition, at the schematic *Ubx* gene representation and in Table 11. H3K36me2 levels shown here are mean values derived from three independent biological replicates of both genotypes and are presented as percentages of input chromatin precipitated at each region. The error bars illustrate standard deviations.
All actively transcribed genes were enriched for H3K36me2 in their coding regions in wild-type T3 discs, the inactive gene *dpr12* was not. But solely *Ubx*, as the only Ash1-regulated gene presented, displayed decreased H3K36me2 levels at its first and at its last exon in *ash1*^{22}-null mutant tissues. Notably, H3K36me2 at *Ubx* was not completely eliminated.

4 Discussion

Within the PcG/trxG family, the trxG protein and H3K36 di-methyltransferase Ash1 has been a rather enigmatic member. Ash1 plays an essential role in development. It maintains *HOX* gene expression active by antagonizing PcG-mediated silencing of transcription. But little was understood about the molecular mechanisms of Ash1 action and how they are regulated. A major limitation in investigating Ash1 mechanisms has been the lack of knowledge about the composition of probable Ash1 protein complexes in both flies and mammals. In contrast, most other PcG and trxG proteins had already been purified and shown to function in defined multiprotein complexes (reviewed in Kassis *et al*, 2017).

In this study, I purified and characterized Ash1 complexes from *Drosophila* embryos. Moreover, I analyzed the regulation of the catalytic activity of Ash1 and its role in development in comparison with the function of the MRG15 protein, one of the newly identified Ash1 interactors. These investigations led to the following main findings: First, Ash1 forms a stable multimeric complex with the chromatin readers Caf1-55 and MRG15. This novel trxG complex is named AMC. Second, Ash1 interacts with MRG15 via a conserved FxLP motif, that is localized in close vicinity to the Ash1 catalytic domain. Third, MRG15 is strongly suggested by my results to enhance the enzymatic activity of Ash1 to di-methylate H3K36. Fourth, Ash1 enzymatic activity is the main Ash1 function required for *HOX* gene expression. Fifth, *MRG15* mutants exhibit a very specific phenotype characterized by loss of *HOX* gene expression. MRG15 can therefore be classified as a novel trxG protein. Sixth, in the absence of AMC, genome-wide H3K36me2 bulk levels are not detectably diminished, but H3K36me2 in the chromatin of the AMC target gene *Ubx* is strongly decreased. These main findings are discussed in depth in the following.

4.1 AMC complex purification and reconstitution

In this work, two independent TAP purification approaches from *Drosophila* embryos using either N- or C-terminally tagged Ash1 proteins as baits led to the isolation of Ash1 in complex with MRG15 and Caf1-55 (para 3.1, Figure 6B). My data do not confirm an earlier report that identified the trxG member Fsh as a stable binding partner of Ash1 (Kockmann *et al*, 2013).

In pull-down assays with recombinant proteins, I could verify the physical interactions of Ash1 with MRG15 as well as with Caf1-55 and show that all three proteins together form a complex *in vitro*, the AMC complex (Figure 8).

4.1.1 AMC complex formation *in vitro* and *in vivo*

The reconstitution of AMC suggests that Ash1 most likely functions as the scaffolding unit to which MRG15 and Caf1-55 bind independently. At least under the standard conditions used here in the corresponding pull-down assay from insect cells overexpressing His-Caf1-55 and MRG15 (Figure 8C), MRG15 and Caf1-55 do not interact directly with each other. Structural studies are required in order to obtain a definite picture of the interactions among the subunits in AMC. In light of the indication that MRG15 and Caf1-55 do not physically assemble without Ash1, it cannot be formally ruled out that Ash1 exists *in vivo* in various separate complexes rather than in AMC. These separate Ash1 complexes might either exclusively contain MRG15 or exclusively Caf1-55. I did not attempt gel filtration or density gradient centrifugation assays to clarify the nature of the *in vivo* Ash1 complexes. However, in the TAP purification of C-terminally tagged Ash1 from embryos, MRG15 and Caf1-55 were isolated in nearly stoichiometric amounts compared to Ash1 (see mass spectrometric data in Figure 6B), which indicates the presence of all three proteins in the same Ash1 complex. Speculating about the function of MRG15 and Caf1-55 with regard to Ash1, it is well conceivable that both complement each other while acting in one complex. As shown in this study, MRG15 likely stimulates the enzymatic activity of Ash1 and the human and yeast MRG15 homologues have been reported to bind with their chromo barrel domains to H3K36me2, the catalytic product of Ash1 (Zhang *et al*, 2006; Sun *et al*, 2008; Xu *et al*, 2008). Caf1-55, in contrast, binds unmodified histone H3, the substrate of Ash1 (Nowak *et al*, 2011; Schmitges *et al*, 2011). Moreover, work by the Bing Zhu laboratory (Chinese Academy of Sciences, Beijing) has provided further arguments in support of AMC as a stable complex (Huang *et al*, 2017): In their independent study, they also co-purified MRG15 and Caf1-55 together with Ash1 from the *Drosophila* tissue culture cell-line S2-DRSC. Importantly, Huang and colleagues showed in addition that AMC is conserved in mammals by purifying the human orthologous complex from the HEK293 cell-line. The conservation of the interaction between Ash1 and MRG15 has also been demonstrated in Schmähling *et al*, 2018 by complex reconstitution from human ASH1L and MRG15.

4.1.2 Interaction interfaces between Ash1 and MRG15

Investigation of the Ash1:MRG15 interaction in more detail revealed that an FxLP motif, located in Ash1 in direct vicinity to its SET domain, is essential for the binding of Ash1 to MRG15 (Figure 10). MRG15 belongs to the MRG protein family. The FxLP motif has already been shown to be highly conserved among binding partners of MRG protein family members and to interact with their conserved MRG domains (Xie *et al*, 2012; 2015). The MRG domain of another MRG protein,

Msl3, was reported to be bound by an FxLP-containing motif in the Male-specific lethal-1 (Msl1) protein in the dosage-compensation complex (Kadlec *et al*, 2011; Xie *et al*, 2015). However, Ash1 interacts specifically with MRG15 and not with Msl3 despite the conservation of its FxLP sequence (Figure 8B). The likely explanation is a second binding interface that might mediate specificity. Xie and colleagues demonstrated that the FxLP-motif is essential but not sufficient for high-affinity interactions with MRG domains (Xie *et al*, 2012). They identified a second MRG domain binding motif in addition to FxLP in the MRG15 interactors Pf1 and MRGBP as well as in the Msl3 interactor Msl1 (Xie *et al*, 2015). Notably, this second binding motif was not conserved in sequence. Whether further interaction sites in Ash1 besides FxLP contribute to binding to MRG15 remains to be uncovered by structural studies.

4.2 Molecular functions of AMC subunits

4.2.1 Stimulation of Ash1 HMTase activity by MRG15

I analyzed the enzymatic activity of Ash1 in dependence of the presence of MRG15 or Caf1-55 *in vitro* (Figure 11). The Ash1:MRG15 complexes were highly active for H3K36 di-methylation. The Ash1:Caf1-55 complexes, in contrast, did not generate detectable H3K36me2 levels. These results therefore suggest that only MRG15 but not Caf1-55 enhances Ash1 HMTase activity in AMC. However, it should be noted that I have not been able to purify sufficient amounts of the corresponding Ash1 protein alone, and could therefore not perform HMTase reactions to quantify the stimulatory effect of MRG15 on Ash1 activity. Biochemical analyses with defined amounts of Ash1:MRG15 and Ash1 alone using shorter Ash1 fragments encompassing the SET domain were carried out by Ji-Joon Song and co-workers and are reported in Schmähling *et al*, 2018. These analyses suggest that full-length MRG15 enhances H3K36 methylation by Ash1 about 30-fold. Ash1 activity stimulation by full-length MRG15 by the same order of magnitude was also determined by Huang and colleagues in their independently performed experiments (Huang *et al*, 2017). Moreover, the enhancement of Ash1 activity by MRG15 is conserved in the human ASH1L:MRG15 complex (Schmähling *et al*, 2018). Thus, the data in this thesis, in Schmähling *et al*, 2018 and in Huang *et al*, 2017 together strongly indicate that Ash1 does not function as an HMTase on its own but needs to be integrated in the AMC complex in order to methylate H3K36.

The molecular mechanism that MRG15 applies to enhance Ash1 activity remains to be determined. However, it is tempting to speculate that MRG15 may somehow release the auto-inhibitory loop formed by the Ash1 post-SET domain, that blocks access to the substrate binding pocket (An *et al*, 2011). Concerning Ash1 domains required for Ash1 enzymatic action, it was

demonstrated that a minimal Ash1 construct, which encompasses the SET domain region including the AWS, core-SET and post-SET subdomains as well as the preceding FxLP motif, is sufficient to methylate histones and to become stimulated by MRG15 (Schmähling *et al*, 2018).

Within MRG15, the MRG domain alone suffices for enhancing Ash1 enzymatic activity (Huang *et al*, 2017; Schmähling *et al*, 2018). The second conserved domain of MRG15, the chromo barrel, in contrast, does not affect the efficiency of Ash1, when tested alone. However, when analyzed in the presence of the MRG domain, the chromo barrel domain further augments the enhancing effect of the MRG domain on Ash1 action (Huang *et al*, 2017). In the face of this result, it is interesting to recall that the MRG15 chromo barrel domain possesses binding specificity toward H3K36me2/me3. The chromo barrel domain might therefore contribute to the stimulation of Ash1 enzymatic activity by the MRG domain by mediating a positive feedback loop: In this scenario, AMC would start decorating individual nucleosomes with H3K36me2, the MRG15 subunit would then bind with its chromo barrel domain to nucleosomes carrying H3K36me2 and thereby augment the efficiency of H3K36 methylation by AMC. Possible mechanisms behind the proposed stimulating effect of the MRG15 chromo barrel-H3K36me2 interaction could be induction of allosteric conformational changes in the AMC complex or correction of the positioning of AMC on nucleosomes. The latter point may, for example, result in better accessibility of unmodified substrate H3K36 residues on neighboring nucleosomes to the active site of Ash1. This would imply simultaneous binding of AMC to one nucleosome that is already modified with the own catalytic product and to a neighboring nucleosome with an unmodified histone substrate, as has in fact recently been shown for another HMTase complex, PRC2. Poepsel et al. published a near-atomic resolution structure of human PRC2 bound to a bifunctional di-nucleosome (Poepsel *et al*, 2018). In this structure, the PRC2 subunit EED interacts with an H3K27me3-modified nucleosome while a domain in EZH2 adjacent to the catalytic SET domain binds to a second nucleosome displaying an unmodified histone H3 tail, which is in this configuration optimally positioned for methylation. Previously, PRC2 has been shown to be allosterically activated by its own catalytic product, H3K27me3 (Hansen *et al*, 2008; Margueron *et al*, 2009). Poepsel and colleagues suggest that their structure provides the molecular explanation for the allosteric activation of PRC2. In Poepsel *et al*, 2018, they propose that the PRC2-dinucleosome configuration may form at borders of H3K27me3-modified to unmodified chromatin domains and thereby may empower spreading of this histone mark. Whether the HMTase activity of AMC is indeed also enhanced by a positive feedback loop and whether this feedback loop would be mediated by an AMC-dinucleosome configuration analogous to the PRC2-dinucleosome structure is currently a matter of speculation, but well conceivable. Importantly in

this context, the enhancing contribution of the MRG15 chromo barrel domain to the Ash1 HMTase activity has been observed on oligonucleosomes as substrates, not on mononucleosomes (Huang *et al*, 2017). The oligonucleosomes were homogenously unmodified. Continuative HMTase assays with substrate oligonucleosomes that contain both, H3K36me2-pre-modified and unmodified nucleosomes, could be interesting experiments to analyze the role of the MRG15 chromo barrel domain in AMC function further. However, the essential MRG15 element for stimulation of the Ash1 catalytic activity is, in conclusion, the MRG domain as the data in the studies Huang *et al*, 2017 and Schmähling *et al*, 2018 collectively demonstrate. The binding of the MRG domain to the FxLP motif and perhaps other Ash1 sites may directly change the conformation of the Ash1 active site, possibly by displacement of the auto-inhibitory loop.

The here discussed findings from my thesis, Schmähling *et al*, 2018 and Huang *et al*, 2017 collectively are the first evidence that MRG15 is able to directly modulate the enzymatic activity of an associated HMTase. Previous studies on the role of MRG15 in chromatin-modifying complexes other than AMC focused mainly on analyzing its contribution to complex function by binding H3K36 methylation on chromatin (Joshi & Struhl, 2005; Carrozza *et al*, 2005; Keogh *et al*, 2005). MRG15 or its yeast homologue Eaf3 have for example been identified as stable subunits of the yeast HDAC complex Rpd3S, as well as of the homologous chromatin-remodeling and HAT complexes NuA4 (yeast) and Tip60 (*Drosophila*) (see para 3.2 for references). Unraveling the molecular mechanism that MRG15 employs in AMC to enhance the enzymatic efficiency of Ash1 in future studies might also lead to a better understanding of the regulation of the catalytic activities of the other MRG15-containing complexes. Interestingly, it has already been reported for the NuA4 HAT complex that Eaf3 depletion results in localized decreased histone acetylation levels but, notably, not in reduced binding of the complex to its target regions (Steunou *et al*, 2016).

In the case that MRG15 proves to stimulate Ash1 by re-positioning the auto-inhibitory loop, the corresponding mechanism might be of general interest for understanding regulation of other H3K36 HMTases like SET2, that bear an analogous auto-inhibitory element (see also para 1.3.3).

4.2.2 Possible AMC-specific functions of Caf1-55

The functions of Caf1-55 in the context with AMC were not investigated in this study with the exception of the analysis of the HMTase activity of the Ash1:Caf1-55 complexes. As already discussed in the previous paragraph, the results of the corresponding HMTase assays (Figure 11) suggest that Caf1-55 does not stimulate Ash1 enzymatic activity. However, it is conceivable that in

AMC, when MRG15 is also present, Caf1-55 has an indirect impact on the Ash1 catalytic efficiency: Caf1-55 could stabilize the integrity of the AMC complex or its positioning on nucleosomes and thereby support the Ash1 enzymatic action. But could fulfillment of these accessory functions be the major role of Caf1-55 in AMC?

As subunit of other chromatin-modifying complexes like PRC2, NuRD, NURF etc. (reviewed in Suganuma *et al*, 2008), Caf1-55 has been suggested to mediate complex targeting to histone substrates and/or complex anchoring on nucleosomes. On a side note, the role of Caf1-55 as a histone chaperone in the CAF-1 complex is again different and not of relevance here. The suggestions for Caf1-55 functions in targeting or anchoring of complexes to chromatin have circulated in the literature for a long time. They were originally based on the binding specificity of Caf1-55 for histone H4 as observed in the nineties already (Vermaak *et al*, 1999; Tie *et al*, 2001). Subsequent biochemical and structural studies suggested that Caf1-55 rather binds to nucleosomes via interaction of its WD40 domain with the very N-terminus of unmodified histone H3 (Nowak *et al*, 2011; Schmitges *et al*, 2011). Histone H3 tails with K27me3 or K9me3 marks are also bound, but not in a specific manner that would indicate a significance of the modifications for Caf1-55 functions like histone crosstalk. Within the PRC2 complex, Caf1-55 in combination with Suppressor of zeste 12 (Su(z)12) was found to be critical for stable binding of PRC2 to nucleosomes (Nekrasov *et al*, 2005). However, puzzlingly, the authors observed efficient H3K27 methylation of mononucleosomes by a trimeric PRC2 sub-complex lacking Caf1-55 and only consisting of Esc, E(z) and Su(z)12. Nekrasov and colleagues proposed thereupon that the contribution of Caf1-55 to PRC2 binding to nucleosomes may be of importance on more complex nucleosome targets, like chromatin *in vivo*. There, Caf1-55 could be responsible for the correct positioning of PRC2 on nucleosomes for efficient histone methylation.

Taken together, the available data on the function of Caf1-55 in chromatin-modifying complexes is surprisingly limited in the face of its existence in various important complexes. The popular assumption that Caf1-55 anchors protein complexes on nucleosomes is still rather speculative and might be too simplistic, albeit it is unlikely to be completely wrong. Therefore, I speculate here that in the context with AMC, Caf1-55 contributes to the binding of the complex to nucleosomes, but without determining binding specificity to certain chromatin domains, since unmodified and various modified histone H3 tails are bound by the Caf1-55 WD40 domain in an unselective manner. Like the trimeric PRC2 sub-complex, Ash1 in complex with MRG15 does not require Caf1-55 for efficient methylation of histones in mononucleosomes (Figure 11). However, a function of Caf1-55 in correct orientation of AMC on chromatin is well conceivable.

4.3 The homeotic mutant phenotypes of *ash1* and *MRG15*

The catalytic activity of Ash1 is boosted by its binding partner MRG15 *in vitro*. But is MRG15 also important for Ash1 enzymatic function *in vivo*? And is the Ash1 H3K36 di-methyltransferase activity actually the key function by which Ash1 regulates transcription of *HOX* genes? I addressed these questions in genetic experiments and found that *ash1*-catalytically inactive mutants and *MRG15* mutants exhibit homeotic phenotypes similar to each other. This implies, first, that the Ash1 HMTase activity is indeed essential for normal *HOX* gene expression and, second, that MRG15 most likely contributes substantially to efficient H3K36 di-methylation by Ash1 in the AMC complex *in vivo* as well. The regulation of *HOX* gene expression by the enzymatic function of AMC is, moreover, directly corroborated by ChIP experiments establishing that AMC is critically required for high-level H3K36me2 at *Ubx*.

The ChIP analyses of H3K36me2 will be discussed in the following paragraphs 4.4 and 4.5. Here, I shall focus on my genetic analyses of the homeotic transformations and the lethality of *ash1*-null, *ash1*-catalytically inactive and *MRG15*-null mutants as described in the paragraphs 3.5 and 3.6. Notably, these mutants were devoid of maternally deposited and zygotically expressed wild-type Ash1 or MRG15 protein, which ensured comparability and allowed for characterization of the true null phenotypes. Concerning the *ash1*-null mutants, I found that the hitherto undescribed phenotype of flies lacking maternally deposited and zygotically expressed wild-type Ash1 (*ash1*-null $^{m- z-}$) is, surprisingly, only slightly more severe than the phenotype of flies lacking only zygotic expression of *ash1* (*ash1*-null $^{z-}$): The homeotic transformations of *ash1*-null $^{m- z-}$ flies were stronger than of *ash1*-null $^{z-}$ flies, but both mutants developed up to pupal stage (Figure 13, personal observations). This indicates that Ash1 function is not important in early embryogenesis and, notably, also not during the onset of zygotic transcription when *HOX* gene expression is particularly sensitive to perturbations in the PcG/trxG system. However, Ash1 is essential for viability during the later stages of development.

The analysis of the catalytically inactive *ash1* phenotype showed that the lack of Ash1 SET domain function leads to loss of expression of the *HOX* genes *Ubx* and *Abd-B* and, consequently, to homeotic transformations, exactly like they are observed in *ash1*-null mutants (Figure 14, Figure 15). In direct comparison with the *ash1*-null mutant phenotype, the severity of the transformations was slightly weaker in the *ash1*-catalytically inactive mutants. Furthermore, about 50% of the *ash1*-catalytically inactive mutants developed even beyond the pupal stage (Figure 13), but died invariably one to two days after eclosing from the pupal case. The slightly weaker *ash1*-catalytically inactive phenotype in comparison to *ash1*-null mutants may account for

another function of Ash1 in *HOX* gene regulation in addition to di-methylating H3K36. However, importantly, this data shows for the first time that Ash1 enzymatic activity is indeed essential for normal *HOX* gene expression. It thereby strongly supports previous suggestions that H3K36 di-methylation is the major means of Ash1 to oppose transcriptional silencing of *HOX* genes by PcG members (see Schmitges *et al*, 2011 and para 1.3.1).

The characterization of the *MRG15*-null mutant phenotype was in the context of this study primarily relevant to assess the importance of MRG15 for AMC function *in vivo*. Nonetheless, a clear genetic analysis of *MRG15* was also of more general interest. The *MRG15*-mutant phenotype had been poorly studied in *Drosophila*. The available knowledge was mainly limited to specific aspects concerning the role of MRG15 in the Tip60 complex (Kusch *et al*, 2004; 2014). Moreover, the only previously identified *MRG15* allele, $MRG15^{j6A3}$, which has been isolated in a screen for *P* element insertions that cause lethality when homozygous (Spradling *et al*, 1999), turned out to be viable when analyzed in trans with a chromosomal deficiency uncovering *MRG15* (personal observation). Further analysis suggested that a truncated MRG15 protein is expressed from $MRG15^{j6A3}$ (Figure 16B). In my genetic analysis of *MRG15*, I first generated a definite *MRG15*-null allele by deleting almost the entire *MRG15* coding region (Figure 17). I found using this allele that *MRG15*-null mutants, especially in the absence of MRG15 maternal load, die in a very heterogeneous manner throughout all stages of development (Figure 18, personal observation). Nevertheless, a small fraction of these *MRG15*-null mutants even undergoes metamorphosis and ecloses as adults. When analyzing the morphology of these *MRG15*-null mutant adults, I found anterior-to-posterior homeotic transformations but no other obvious defects (Figure 19). When analyzing *MRG15*-null mutant larvae, I detected loss of *Ubx* and *Abd-B* expression in line with the nature of the homeotic transformations observed in adults (Figure 20). Altogether, the *MRG15*-null mutant phenotype is very similar to the phenotype of *ash1*-catalytically inactive mutants with respect to the kinds and severity of homeotic transformations as well as the partial survival up to the adult stage. This genetic analysis strongly suggests that MRG15 enhances Ash1 HMTase activity *in vivo*.

Considering that MRG15 is also a subunit of various other chromatin-associated multiprotein complexes like Tip60 or Rpd3S, the high specificity of the phenotype and the partial viability of the *MRG15*-null mutants are very surprising. Based on the specific homeotic phenotype, MRG15 even appears to fulfill its major function in the AMC complex and seems not to be essential for the other complexes with which it associates. Nevertheless, this is not a definite statement since I focused in my analysis on the external morphology of the *MRG15* mutants. Further studies are needed to analyze for phenotypic defects in internal structures like

organs or in physiological processes that might be caused by lack of MRG15. Also, as discussed above, a considerable fraction of the *MRG15*-null mutant animals dies before the pupal stage. This could actually be due to roles of MRG15 in complexes other than AMC, since *ash1*-null mutants quite consistently develop into pupae.

In a nutshell, the genetic analyses of the *MRG15*-null mutants uncover that these animals have a pronounced *HOX* loss-of-function syndrome. This phenotype classifies MRG15 as a novel trxG protein.

4.4 The specificity of AMC function in transcriptional regulation

The specificity of Ash1 in transcriptional regulation has been a matter of debate with sporadic and often controversial reports in the literature as discussed in paragraph 1.1.4. Some studies have suggested that Ash1 functions as a global transcription factor that is involved in regulation of the major part if not all active genes in the genome (Gregory *et al*, 2007; Kockmann *et al*, 2013). Others, in contrast, have found that Ash1 works more restrictively in a gene-specific manner (Schwartz *et al*, 2010; Huang *et al*, 2017; Schmähling *et al*, 2018). But even these latter reports still disagree substantially about the number of genes that are regulated by Ash1 as the identified Ash1 targets in *Drosophila* range from a dozen (Huang *et al*, 2017) to a few hundred genes (Schmähling *et al*, 2018). A possible and simple explanation for this discrepancy could be the use of different biological systems in these studies, i.e. larval imaginal discs in one case (Schmähling *et al*, 2018) and tissue culture cells in the other (Huang *et al*, 2017). Interestingly, all studies on this topic found independently from each other that many but not all Ash1 targets are developmental regulator genes and that of these, by far not all are *HOX* genes. However, the mutant phenotypes of *ash1*, as they were analyzed in this thesis (Figure 14) and as they have been described previously (Shearn *et al*, 1987; Tripoulas *et al*, 1994), as well as the phenotype of *MRG15* mutants (Figure 19), are remarkably specific with characteristic homeotic transformations, but no other defects. Nonetheless, it is somewhat difficult to imagine that within the Ash1 target genes only loss of expression of *HOX* genes should have phenotypic consequences. Reduced levels of the Ash1-regulated non-*HOX* genes may perhaps affect the formation of internal structures or physiological processes in AMC mutants, which could not have been detected in my analyses of external adult morphology. However, my analyses certainly corroborate that *HOX* genes are major targets of the AMC complex. Altogether, a global function of Ash1 in transcription as has been suggested is very unlikely according to my results, not only due to the specificity of the AMC mutant phenotypes but also due to their relative mildness. As

discussed in the previous paragraph, AMC mutants develop up to pupal stage (*ash1*-null mutants) or even beyond (*ash1*-catalytically inactive and *MRG15*-null mutants).

The *ash1*-catalytically inactive phenotype has established that the Ash1 capacity to di-methylate H3K36 is required for normal Ash1-dependent *HOX* gene expression (previous paragraph). Therefore, H3K36me2 levels are a better indicator for either a global or a more specific AMC function than the number of Ash1 binding sites. Ash1 binding sites have been used as parameter for assessing Ash1 specificity in transcriptional regulation in the study Kockmann *et al*, 2013. When I analyzed here the H3K36me2 bulk levels in total cell extracts from *Drosophila* imaginal discs by western blot, I found that these levels are not globally decreased in *MRG15*-null and *ash1*-null mutants (Figure 21). I would like to recall at this point that there are two other H3K36 HMTases in *Drosophila* besides Ash1, NSD and SET2. NSD also catalyzes H3K36 di-methylation, whereas SET2 has been suggested to generate only H3K36me3 and, in doing so, to be dependent on NSD or Ash1 pre-modifying substrates with H3K36me2 (Bell *et al*, 2007). Accordingly, the H3K36me2 bulk levels in the larval extracts presented in Figure 21, have, in all likelihood, been generated by NSD. NSD has already been shown previously to be responsible for the bulk of H3K36me2 in *Drosophila* (Bell *et al*, 2007). And taken the data from Bell and colleagues together with my analysis, NSD is most likely the only global *Drosophila* H3K36 di-methyltransferase, while Ash1 might modify chromatin at specific target genes. In order to test this assumption, I had analyzed the enrichment of H3K36me2 at the established Ash1-target gene *Ubx* in *Drosophila* haltere and 3rd leg imaginal discs (T3 discs). The immunostainings performed against Ubx in this study have shown loss of *Ubx* expression in *ash1* mutants in the very same tissues (Figure 15). When analyzing H3K36me2 at *Ubx*, I found that in the absence of Ash1 the levels of the histone modification were significantly decreased throughout the gene body from the 5' to the 3' end, although not down to background levels (Figure 22). The same observations, strong reduction but not elimination of H3K36me2, have also been made at other Ash1-regulated genes in *ash1*-null mutants, while the H3K36me2 levels at further actively transcribed genes remained unchanged (Schmähling *et al*, 2018). Therefore, Ash1 most likely di-methylates H3K36 at specific target genes. One might argue, that the decrease of H3K36me2 in the absence of Ash1 is an expectable side-effect in consequence of transcription stop at the corresponding genes. Such an assumption could be made based on the fact that the yeast H3K36 HMTase Set2 is known to be recruited to chromatin by elongating RNAP II (Venkatesh & Workman, 2013). The studies Bell *et al*, 2007 and Stabell *et al*, 2007 have suggested a homologous mechanism for the recruitment of *Drosophila* SET2. However, in *Drosophila*, SET2 is thought to generate only H3K36me3 and there is no indication that the targeting of NSD or Ash1 to their genomic sites of action is RNAP II-

mediated. In fact, Ash1 has been reported to associate to chromatin in the absence of elongating RNAP II as well (Dorighi & Tamkun, 2013). Importantly, I observed H3K36me2 reduction in *ash1*-null mutants also at the very 5' end of the *Ubx* gene, which further supports the suggestion that this effect occurred independently of the absence of elongating RNAP II and as a specific result of loss of Ash1 enzymatic function. Moreover, the loss of *HOX* gene expression in the *ash1*-catalytically inactive mutant (Figure 15) implies that the decrease of the catalytic Ash1 product, H3K36me2, at Ash1 target genes in *ash1* mutants is the cause and not a consequence of the stop of transcription.

The general question discussed in this paragraph was, how specific does Ash1 function in transcription? In summary, all data from this study concerning this question contradict a genome-wide role of Ash1 in transcriptional regulation and strongly suggest that Ash1, as core subunit of the AMC complex, controls a restricted set of target genes. Data supporting this suggestion encompass the results of the genetic analyses of *ash1* and *MRG15* mutants that revealed weak but very specific phenotypes and the results of the analyses of the H3K36me2 levels in total in the cell and at single genes that revealed a decrease in the levels of H3K36me2 only gene-specifically. The main target genes of AMC in *Drosophila* are apparently *HOX* genes since loss of *HOX* gene expression in AMC mutants results in clear phenotypic manifestations.

4.5 The mechanism of transcriptional regulation by AMC catalytic activity

At *HOX* and possibly other target genes, PcG and trxG proteins form a counterbalancing system of transcriptional regulators that oppose each other in function. PcG proteins silence transcription whereas trxG proteins positively regulate transcription. Whose action is prevalent, depends on the spatial context in metazoans. I showed in this study that the trxG member Ash1 most likely functions as core subunit of the AMC complex and, as such, indeed maintains transcription active by di-methylation of H3K36 as has been suspected previously (discussed in para 4.3 and 4.4). As underlying mechanism has been proposed that the H3K36me2 mark allosterically inhibits the enzymatic activity of the PcG complex PRC2 to tri-methylate H3K27 (Schmitges *et al*, 2011). Enrichment of H3K27me3 marks at the chromatin spanning the gene body induces transcriptional silencing of *HOX* genes through so far unknown means (Papp & Müller, 2006; Pengelly *et al*, 2013). The suggestion of an Ash1 mechanism involving inhibition of the PRC2 function as HMTase was based on corresponding enzymatic *in vitro* assays (Schmitges *et al*, 2011; Yuan *et al*, 2011). The consequential question was, if this is indeed the mechanism by which Ash1-

generated H3K36me2 keeps transcription active *in vivo*. In the only study which addressed this question so far (Papp & Müller, 2006), the H3K27me3 levels at *Ubx* in T3 imaginal discs were analyzed. In these tissues, the activating trxG function prevails in *wild type*, H3K27me3 is not enriched in the gene body and *Ubx* is transcribed. In the absence of Ash1, in contrast, H3K37me3 levels are increased across the gene body and *Ubx* expression is lost, according to Papp and Müller. The findings of my work, that the transcribed region of *Ubx* in T3 imaginal discs is decorated with H3K36me2 and that these levels of H3K36me2 are decreased in *ash1*-null mutants (Figure 22), complements the study Papp & Müller, 2006. Both studies together strongly support the suggestion that H3K36me2 deposited by Ash1 in AMC indeed inhibits the PRC2 activity to tri-methylate H3K27 at Ash1 target genes. Due to limitations in the amount of chromatin from *ash1*-null mutant imaginal discs, I could not test the enrichment of H3K27me3 in parallel to H3K36me2 in my experiment. But the *Ubx* amplicons at which Papp and Müller monitored increase of H3K27me3 in T3 discs of *ash1*-null mutants are partially identical with the *Ubx* amplicons at which I observed decrease of H3K36me2 in the same *ash1*-null mutant tissues (Table 11).

The finding that H3K36me2 levels at *Ubx* (Figure 22) and other Ash1 target genes (Schmähling *et al*, 2018) are only decreased but not eliminated in *ash1*-null mutants is highly interesting concerning possible AMC mechanisms. AMC apparently cooperates with the global HMTase NSD in H3K36 di-methylation at its target genes. NSD is the enzyme most likely to have generated the H3K36me2 marks that were still detected in the absence of Ash1. This finding also suggests that AMC actually acts on top of pre-existing basal H3K36me2 levels and, therefore, on chromatin that might be already transcriptionally active. This scenario would be in line with genetic analyses having revealed that the loss of *Ubx* expression in T3 discs of *ash1* mutants is restored in *ash1* PRC2 double mutants (Klymenko & Müller, 2004). The phenotypes of the *ash1* PRC2 double mutants prove that transcription of *Ubx* is per se possible without Ash1. The view of Ash1 as a factor that rather upholds transcription in the presence of PcG members, in contrast to primarily activating it, was deduced from these genetic experiments.

The histone modification enrichments at *Ubx* in the study Papp & Müller, 2006 and in this work (Figure 22) represent the average levels from all cells in the *ash1*-null mutant T3 imaginal discs. In fact, *ash1*-null mutant T3 imaginal discs and as well T3 discs that are mutant for *MRG15* lose *Ubx* expression in a conspicuous all-or-none patchy pattern over the entire tissue area: In some regions of the discs, the *Ubx* expression levels seem to equal wild-type levels whereas in other regions, *Ubx* is erased (Figure 15, Figure 20). The pattern of these regions suggests that the cells within one region of a specific *Ubx* expression state are clones. Accordingly, the homeotic transformations of *ash1*- or *MRG15* mutant adult structures form in a patchy fashion as well

(Figure 14, Figure 19). This all-or-none loss of *Ubx* expression in the mutant T3 discs is likely also reflected by the histone modification landscape at the *Ubx* gene body: H3K36me2 levels might be diminished and H3K27me3 levels highly enriched in clonal patches where *Ubx* is not transcribed and vice versa in clonal patches of cells actively transcribing *Ubx*. Interestingly, in order to keep the Ubx protein levels up, it appears to suffice in this scenario if the H3K36me2 levels and with that the transcriptionally active state is maintained at only one out of the two *Ubx* alleles on the two homologous chromosomes per nucleus, since no clonal patches exhibiting medium *Ubx* expression levels were observed in the *ash1* or *MRG15* mutant T3 discs. In the cells not expressing *Ubx*, in contrast, both alleles would have lacked H3K36me2 and in consequence been 'switched off'.

Taken together, (1) the patchy all-or-none loss of *Ubx* expression in AMC mutants, (2) the H3K36me2 decrease as well as the H3K27me3 increase along the transcribed region of *Ubx* in *ash1* mutants and (3) the inhibition of PRC2 enzymatic activity by H3K36me2 suggest the following model as to how the AMC complex functions to maintain transcription active at specific genes. This model is mainly based on data collected on *Ubx*. But it might well apply to the regulation of other Ash1 target genes as well, whose chromatin is decorated with H3K36me2 in the transcriptionally active and with H3K27me3 in the transcriptionally silent state. In this model, AMC targets genes that already exhibit basal levels of H3K36me2/me3 in their transcribed regions generated by NSD and SET2. The initial binding of AMC to the chromatin of these genes and the tethering of the complex to nucleosomes while modifying histones might be a cooperative process to which MRG15, Caf1-55 and various Ash1 domains contribute. The MRG15 chromo barrel domain might mediate binding of AMC to nucleosomes pre-modified with H3K36me2/me3 (para 4.2.1), the WD40 domain of Caf1-55 to nucleosomes with unmodified histone H3 tails (para 4.2.2). The bromodomain, the PHD finger and the BAH domain of Ash1 might mediate histone crosstalk by interacting with histone modifications other than H3K36me2/me3. While bound to AMC target genes, the complex would di-methylate H3K36 in chromatin along the gene body and thereby augment the pre-existing H3K36me2 levels. On the already partially decorated chromatin, the proposed AMC-dinucleosome configuration (see para 4.2.1), in which the complex engages simultaneously a modified nucleosome with the MRG15 chromo barrel domain and an unmodified histone H3 tail with the Ash1 SET domain region, might empower maximum AMC efficiency. The deposited H3K36me2 marks then most likely maintain transcription active by allosteric inhibition of the PRC2 activity to tri-methylate H3K27me3. For effective inhibition of PRC2 action, the H3K36me2 marks at the 5' end of the gene body of *Ubx* or other Ash1 target genes might be especially important: Papp and Müller showed in their analysis of the chromatin

at *Ubx*, that H3K27me3 spreads in a linear manner from the upstream PRE, the PRC2 binding site, across the upstream control region and, in the transcriptionally silent state, further into the gene body. But in the transcriptionally active state, H3K36me2 might form a barrier that prevents the linear spreading of H3K27me3 from nucleosome to nucleosome into the transcribed region. Since the H3K36me2 mark is required to be present *in cis* on the same histone H3 tail as the PRC2 substrate residue K27 for inhibition (para 1.3.1), the H3K36me2 barrier ideally starts from the 5' end of the gene body. The puzzling part of the model described so far is, that according to the *ash1* mutant *HOX* gene loss-of-function phenotypes, H3K36me2 deposited by Ash1 seems to be essential for antagonizing PRC2 function, albeit NSD apparently also generates H3K36me2 at Ash1 target genes. I hypothesize here that the H3K36me2 levels need to pass a certain threshold to form a barrier that is actually effective against H3K27me3 invasion or, in other words, to efficiently inhibit PRC2. In this scenario, the specific function of Ash1 in AMC would be to augment the pre-existing H3K36me2 levels at the AMC target genes, so that this critical threshold is reached. Specifically at *Ubx*, this threshold might lie just above the average H3K36me2 levels NSD alone generates. Therefore, in some cells, the basal H3K36me2 levels deposited by NSD at at least one of the two *Ubx* alleles in AMC mutants would actually suffice to inhibit PRC2 there and to establish a stable transcriptionally active state, in which *Ubx* would be expressed at wild-type levels. In contrast, in other cells, the H3K36me2 levels at both *Ubx* alleles would not reach the critical threshold without AMC, which would lead to establishment of a stable transcriptionally silent state. In the following cell divisions, the corresponding transcriptional states would be clonally propagated resulting in the patchy pattern of the all-or-none loss of *Ubx* expression in T3 imaginal discs.

In a nutshell, this model of the mechanism of transcription regulation by Ash1 suggests that Ash1 antagonizes transcriptional silencing by the PcG as enzymatic core of the AMC complex. In the model, the specific function of AMC is to augment the number of H3K36me2-modified nucleosomes in the gene body of the target genes beyond a certain threshold that allows effective inhibition of PRC2 enzymatic activity.

The major contributions to this model by my PhD thesis are the following findings, that have been discussed in depth in the previous paragraphs: Ash1 indeed regulates transcription by the means of H3K36 di-methylation. Together with MRG15 and Caf1-55, Ash1 forms the stable trxG complex AMC. Ash1 is most likely active as HMTase *in vivo* as subunit of AMC and not alone.

5 References

Akam M (1987) The molecular basis for metameric pattern in the Drosophila embryo. *Development* **101:** 1–22

Allen BL & Taatjes DJ (2015) The Mediator complex: a central integrator of transcription. *Nat. Rev. Mol. Cell Biol.* **16:** 155–166

Americo J, Whiteley M, Brown JL, Fujioka M, Jaynes JB & Kassis JA (2002) A complex array of DNA-binding proteins required for pairing-sensitive silencing by a polycomb group response element from the Drosophila engrailed gene. *Genetics* **160:** 1561–1571

An S, Yeo KJ, Jeon YH & Song J-J (2011) Crystal structure of the human histone methyltransferase ASH1L catalytic domain and its implications for the regulatory mechanism. *J. Biol. Chem.* **286:** 8369–8374

Apostolou E, Ferrari F, Walsh RM, Bar-Nur O, Stadtfeld M, Cheloufi S, Stuart HT, Polo JM, Ohsumi TK, Borowsky ML, Kharchenko PV, Park PJ & Hochedlinger K (2013) Genome-wide chromatin interactions of the Nanog locus in pluripotency, differentiation, and reprogramming. *Cell Stem Cell* **12:** 699–712

Ardehali MB, Mei A, Zobeck KL, Caron M, Lis JT & Kusch T (2011) Drosophila Set1 is the major histone H3 lysine 4 trimethyltransferase with role in transcription. *EMBO J.* **30:** 2817–2828

Armstrong JA, Papoulas O, Daubresse G, Sperling AS, Lis JT, Scott MP & Tamkun JW (2002) The Drosophila BRM complex facilitates global transcription by RNA polymerase II. *EMBO J.* **21:** 5245–5254

Barth TK, Schade GOM, Schmidt A, Vetter I, Wirth M, Heun P, Thomae AW & Imhof A (2014) Identification of novel Drosophila centromere-associated proteins. *Proteomics* **14:** 2167–2178

Beachy PA, Krasnow MA, Gavis ER & Hogness DS (1988) An Ultrabithorax protein binds sequences near its own and the Antennapedia P1 promoters. *Cell* **55:** 1069–1081

Bell O, Wirbelauer C, Hild M, Scharf AND, Schwaiger M, MacAlpine DM, Zilbermann F, van Leeuwen F, Bell SP, Imhof A, Garza D, Peters AHFM & Schübeler D (2007) Localized H3K36 methylation states define histone H4K16 acetylation during transcriptional elongation in Drosophila. *EMBO J.* **26:** 4974–4984

Berger MF, Badis G, Gehrke AR, Talukder S, Philippakis AA, Peña-Castillo L, Alleyne TM, Mnaimneh S, Botvinnik OB, Chan ET, Khalid F, Zhang W, Newburger D, Jaeger SA, Morris QD, Bulyk ML & Hughes TR (2008) Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. *Cell* **133:** 1266–1276

Bischof J, Maeda RK, Hediger M, Karch F & Basler K (2007) An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. *Proc. Natl. Acad. Sci. U.S.A.* **104:** 3312–3317

Bowman BR, Moure CM, Kirtane BM, Welschhans RL, Tominaga K, Pereira-Smith OM & Quiocho FA (2006) Multipurpose MRG domain involved in cell senescence and proliferation exhibits structural homology to a DNA-interacting domain. *Structure/Folding and Design* **14:** 151–158

Bowman SK, Deaton AM, Domingues H, Wang PI, Sadreyev RI, Kingston RE & Bender W (2014) H3K27 modifications define segmental regulatory domains in the Drosophila bithorax complex. *eLife* **3:** e02833

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

Brown JL, Fritsch C, Mueller J & Kassis JA (2003) The Drosophila pho-like gene encodes a YY1-related DNA binding protein that is redundant with pleiohomeotic in homeotic gene silencing. *Development* **130:** 285–294

Brown JL, Mucci D, Whiteley M, Dirksen ML & Kassis JA (1998) The Drosophila Polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1. *Mol. Cell* **1:** 1057–1064

Byrd KN & Shearn A (2003) ASH1, a Drosophila trithorax group protein, is required for methylation of lysine 4 residues on histone H3. *Proc. Natl. Acad. Sci. U.S.A.* **100:** 11535–11540

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

Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, Shia W-J, Anderson S, Yates J, Washburn MP & Workman JL (2005) Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123:** 581–592

Chalkley GE, Moshkin YM, Langenberg K, Bezstarosti K, Blastyak A, Gyurkovics H, Demmers JAA & Verrijzer CP (2008) The transcriptional coactivator SAYP is a trithorax group signature subunit of the PBAP chromatin remodeling complex. *Mol. Cell. Biol.* **28:** 2920–2929

Chang Y-L, King B, Lin S-C, Kennison JA & Huang D-H (2007) A double-bromodomain protein, FSH-S, activates the homeotic gene ultrabithorax through a critical promoter-proximal region. *Mol. Cell. Biol.* **27:** 5486–5498

Cherbas L, Willingham A, Zhang D, Yang L, Zou Y, Eads BD, Carlson JW, Landolin JM, Kapranov P, Dumais J, Samsonova A, Choi J-H, Roberts J, Davis CA, Tang H, van Baren MJ, Ghosh S, Dobin A, Bell K, Lin W, et al (2011) The transcriptional diversity of 25 Drosophila cell lines. *Genome Research* **21:** 301–314

Cho Y-W, Hong T, Hong S, Guo H, Yu H, Kim D, Guszczynski T, Dressler GR, Copeland TD, Kalkum M & Ge K (2007) PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. *J. Biol. Chem.* **282:** 20395–20406

Choi J, Bachmann AL, Tauscher K, Benda C, Fierz B & Müller J (2017) DNA binding by PHF1 prolongs PRC2 residence time on chromatin and thereby promotes H3K27 methylation. *Nat. Struct. Mol. Biol.* **24:** 1039–1047

Chou TB & Perrimon N (1996) The autosomal FLP-DFS technique for generating germline mosaics in Drosophila melanogaster. *Genetics* **144:** 1673–1679

Ciccarone VC, Polayes DA & Luckow VA (1998) Generation of Recombinant Baculovirus DNA in E.coli Using a Baculovirus Shuttle Vector. *Methods Mol. Med.* **13:** 213–235

Coleman RT & Struhl G (2017) Causal role for inheritance of H3K27me3 in maintaining the OFF state of a Drosophila HOX gene. *Science* **356:** eaai8236

Copur O & Müller J (2013) The histone H3-K27 demethylase Utx regulates HOX gene expression in Drosophila in a temporally restricted manner. *Development* **140:** 3478–3485

Copur Ö & Müller J (2018) Histone Demethylase Activity of Utx Is Essential for Viability and Regulation of HOX Gene Expression in Drosophila. *Genetics* **208:** 633–637

Crosby MA, Miller C, Alon T, Watson KL, Verrijzer CP, Goldman-Levi R & Zak NB (1999) The trithorax group gene moira encodes a brahma-associated putative chromatin-remodeling factor in Drosophila melanogaster. *Mol. Cell. Biol.* **19:** 1159–1170

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

Daubresse G, Deuring R, Moore L, Papoulas O, Zakrajsek I, Waldrip WR, Scott MP, Kennison JA & Tamkun JW (1999) The Drosophila kismet gene is related to chromatin-remodeling factors and is required for both segmentation and segment identity. *Development* **126:** 1175–1187

Dehé P-M, Dichtl B, Schaft D, Roguev A, Pamblanco M, Lebrun R, Rodríguez-Gil A, Mkandawire M, Landsberg K, Shevchenko A, Shevchenko A, Rosaleny LE, Tordera V, Chávez S, Stewart AF & Géli V (2006) Protein interactions within the Set1 complex and their roles in the regulation of histone 3 lysine 4 methylation. *J. Biol. Chem.* **281:** 35404–35412

Digan ME, Haynes SR, Mozer BA, Dawid IB, Forquignon F & Gans M (1986) Genetic and molecular analysis of fs(1)h, a maternal effect homeotic gene in Drosophila. *Developmental biology* **114:** 161–169

Dingwall AK, Beek SJ, McCallum CM, Tamkun JW, Kalpana GV, Goff SP & Scott MP (1995) The Drosophila snr1 and brm proteins are related to yeast SWI/SNF proteins and are components of a large protein complex. *Mol. Biol. Cell* **6:** 777–791

Dorighi KM & Tamkun JW (2013) The trithorax group proteins Kismet and ASH1 promote H3K36 dimethylation to counteract Polycomb group repression in Drosophila. *Development*

Dorsett D & Merkenschlager M (2013) Cohesin at active genes: a unifying theme for cohesin and gene expression from model organisms to humans. *Curr. Opin. Cell Biol.* **25:** 327–333

Dou Y, Milne TA, Ruthenburg AJ, Lee S, Lee JW, Verdine GL, Allis CD & Roeder RG (2006) Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat. Struct. Mol. Biol.* **13:** 713–719

Duboule D & Dollé P (1989) The structural and functional organization of the murine HOX gene family resembles that of Drosophila homeotic genes. *EMBO J.* **8:** 1497–1505

Duboule D & Morata G (1994) Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends Genet.* **10:** 358–364

Dura JM & Ingham P (1988) Tissue- and stage-specific control of homeotic and segmentation gene expression in Drosophila embryos by the polyhomeotic gene. *Development* **103:** 733–741

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

Eisen A, Utley RT, Nourani A, Allard S, Schmidt P, Lane WS, Lucchesi JC & Cote J (2001) The yeast NuA4 and Drosophila MSL complexes contain homologous subunits important for transcription regulation. *J. Biol. Chem.* **276:** 3484–3491

Fauvarque MO, Laurenti P, Boivin A, Bloyer S, Griffin-Shea R, Bourbon HM & Dura JM (2001) Dominant modifiers of the polyhomeotic extra-sex-combs phenotype induced by marked P element insertional mutagenesis in Drosophila. *Genet. Res.* **78:** 137–148

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

Francis NJ, Kingston RE & Woodcock CL (2004) Chromatin compaction by a polycomb group protein complex. *Science* **306:** 1574–1577

Frey F, Sheahan T, Finkl K, Stoehr G, Mann M, Benda C & Müller J (2016) Molecular basis of PRC1 targeting to Polycomb response elements by PhoRC. *Genes & Development* **30:** 1116–1127

Gambetta MC & Müller J (2014) O-GlcNAcylation prevents aggregation of the Polycomb group repressor Polyhomeotic. *Developmental Cell* **31:** 629–639

Gambetta MC, Oktaba K & Müller J (2009) Essential role of the glycosyltransferase sxc/Ogt in polycomb repression. *Science* **325:** 93–96

Gaytán de Ayala Alonso A, Gutiérrez L, Fritsch C, Papp B, Beuchle D & Müller J (2007) A genetic screen identifies novel polycomb group genes in Drosophila. *Genetics* **176:** 2099–2108

Gilbert SF (2006) Developmental Biology Eighth Edition. Sunderland, MA: *Sinauer Associates, Inc.*

Gong WJ & Golic KG (2003) Ends-out, or replacement, gene targeting in Drosophila. *Proc. Natl. Acad. Sci. U.S.A.* **100:** 2556–2561

Graham A, Papalopulu N & Krumlauf R (1989) The murine and Drosophila homeobox gene complexes have common features of organization and expression. *Cell* **57:** 367–378

Gregory GD, Vakoc CR, Rozovskaia T, Zheng X, Patel S, Nakamura T, Canaani E & Blobel GA (2007) Mammalian ASH1L is a histone methyltransferase that occupies the transcribed region of active genes. *Mol. Cell. Biol.* **27:** 8466–8479

Grossniklaus U & Paro R (2014) Transcriptional silencing by polycomb-group proteins. *Cold Spring Harb Perspect Biol* **6:** a019331–a019331

Groth AC, Fish M, Nusse R & Calos MP (2004) Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. *Genetics* **166:** 1775–1782

Gutiérrez L, Oktaba K, Scheuermann JC, Gambetta MC, Ly-Hartig N & Müller J (2012) The role of the histone H2A ubiquitinase Sce in Polycomb repression. *Development* **139:** 117–127

Hallson G, Hollebakken RE, Li T, Syrzycka M, Kim I, Cotsworth S, Fitzpatrick KA, Sinclair DAR & Honda BM (2012) dSet1 is the main H3K4 di- and tri-methyltransferase throughout Drosophila development. *Genetics* **190:** 91–100

Hallson G, Syrzycka M, Beck SA, Kennison JA, Dorsett D, Page SL, Hunter SM, Keall R, Warren WD, Brock HW, Sinclair DAR & Honda BM (2008) The Drosophila cohesin subunit Rad21 is a trithorax group (trxG) protein. *Proc. Natl. Acad. Sci. U.S.A.* **105:** 12405–12410

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

Harding K, Wedeen C, McGinnis W & Levine M (1985) Spatially regulated expression of homeotic genes in Drosophila. *Science* **229:** 1236–1242

Herz H-M, Mohan M, Garrett AS, Miller C, Casto D, Zhang Y, Seidel C, Haug JS, Florens L, Washburn MP, Yamaguchi M, Shiekhattar R & Shilatifard A (2012) Polycomb repressive complex 2-dependent and -independent functions of Jarid2 in transcriptional regulation in Drosophila. *Mol. Cell. Biol.* **32:** 1683–1693

Huang C, Yang F, Zhang Z, Zhang J, Cai G, Li L, Zheng Y, Chen S, Xi R & Zhu B (2017) Mrg15 stimulates Ash1 H3K36 methyltransferase activity and facilitates Ash1 Trithorax group protein function in Drosophila. *Nature Communications* **8:** 1649

Illingworth RS, Moffat M, Mann AR, Read D, Hunter CJ, Pradeepa MM, Adams IR & Bickmore WA (2015) The E3 ubiquitin ligase activity of RING1B is not essential for early mouse development. *Genes & Development* **29:** 1897–1902

Ingham P & Whittle R (1980) Trithorax: A new homoeotic mutation of Drosophila melanogaster causing transformations of abdominal and thoracic imaginal segments. I. Putative role during embryogenesis. *Molec. Gen. Genet.* **179:** 607–614

Ingham PW (1981) Trithorax: A new homoeotic mutation of Drosophila melanogaster. II. The role of trx^{+} after embryogenesis. *Wilehm Roux Arch Dev Biol* **190:** 365–369

Ingham PW (1983) Differential expression of bithorax complex genes in the absence of the extra sex combs and trithorax genes. *Nature* **306:** 591–593

Irish VF, Martinez-Arias A & Akam M (1989) Spatial regulation of the Antennapedia and Ultrabithorax homeotic genes during Drosophila early development. *EMBO J.* **8:** 1527–1537

Jonkers I & Lis JT (2015) Getting up to speed with transcription elongation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **16:** 167–177

Joshi AA & Struhl K (2005) Eaf3 Chromodomain Interaction with Methylated H3-K36 Links Histone Deacetylation to Pol II Elongation. *Mol. Cell* **20:** 971–978

Kadlec J, Hallacli E, Lipp M, Holz H, Sanchez-Weatherby J, Cusack S & Akhtar A (2011) Structural basis for MOF and MSL3 recruitment into the dosage compensation complex by MSL1. *Nat. Struct. Mol. Biol.* **18:** 142–149

Kassis JA, Kennison JA & Tamkun JW (2017) Polycomb and Trithorax Group Genes in Drosophila. *Genetics* **206:** 1699–1725

Kaufman TC, Lewis R & Wakimoto B (1980) Cytogenetic analysis of chromosome 3 in Drosophila melanogaster: The homoeotic gene complex in polytene chromosome interval 84a-B. *Genetics* **94:** 115–133

Kennison JA & Tamkun JW (1988) Dosage-dependent modifiers of polycomb and antennapedia mutations in Drosophila. *Proc. Natl. Acad. Sci. U.S.A.* **85:** 8136–8140

Keogh M-C, Kurdistani SK, Morris SA, Ahn SH, Podolny V, Collins SR, Schuldiner M, Chin K, Punna T, Thompson NJ, Boone C, Emili A, Weissman JS, Hughes TR, Strahl BD, Grunstein M, Greenblatt JF, Buratowski S & Krogan NJ (2005) Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* **123:** 593–605

Ketel CS, Andersen EF, Vargas ML, Suh J, Strome S & Simon JA (2005) Subunit contributions to histone methyltransferase activities of fly and worm polycomb group complexes. *Mol. Cell. Biol.* **25:** 6857–6868

Kim H, Kang K & Kim J (2009) AEBP2 as a potential targeting protein for Polycomb Repression Complex PRC2. *Nucleic Acids Research* **37:** 2940–2950

King IFG, Emmons RB, Francis NJ, Wild B, Müller J, Kingston RE & Wu C-T (2005) Analysis of a polycomb group protein defines regions that link repressive activity on nucleosomal templates to in vivo function. *Mol. Cell. Biol.* **25:** 6578–6591

King IFG, Francis NJ & Kingston RE (2002) Native and recombinant polycomb group complexes establish a selective block to template accessibility to repress transcription in vitro. *Mol. Cell. Biol.* **22:** 7919–7928

Klymenko T & Müller J (2004) The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO Rep* **5:** 373–377

Klymenko T, Papp B, Fischle W, Köcher T, Schelder M, Fritsch C, Wild B, Wilm M & Müller J (2006) A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities. *Genes & Development* **20:** 1110–1122

Kockmann T, Gerstung M, Schlumpf T, Xhinzhou Z, Hess D, Beerenwinkel N, Beisel C & Paro R (2013) The BET protein FSH functionally interacts with ASH1 to orchestrate global gene activity in Drosophila. *Genome Biol.* **14:** R18

Krogan NJ, Baetz K, Keogh M-C, Datta N, Sawa C, Kwok TCY, Thompson NJ, Davey MG, Pootoolal J, Hughes TR, Emili A, Buratowski S, Hieter P & Greenblatt JF (2004) Regulation of chromosome stability by the histone H2A variant Htz1, the Swr1 chromatin remodeling complex, and the histone acetyltransferase NuA4. *Proc. Natl. Acad. Sci. U.S.A.* **101:** 13513–13518

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

Kusch T, Florens L, Macdonald WH, Swanson SK, Glaser RL, Yates JR, Abmayr SM, Washburn MP & Workman JL (2004) Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions. *Science* **306:** 2084–2087

Kusch T, Mei A & Nguyen C (2014) Histone H3 lysine 4 trimethylation regulates cotranscriptional H2A variant exchange by Tip60 complexes to maximize gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **111:** 4850–4855

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

Kwong C, Adryan B, Bell I, Meadows L, Russell S, Manak JR & White R (2008) Stability and dynamics of polycomb target sites in Drosophila development. *PLoS Genet* **4:** e1000178

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

Lagarou A, Mohd-Sarip A, Moshkin YM, Chalkley GE, Bezstarosti K, Demmers JAA & Verrijzer CP (2008) dKDM2 couples histone H2A ubiquitylation to histone H3 demethylation during Polycomb group silencing. *Genes & Development* **22:** 2799–2810

LaJeunesse D & Shearn A (1995) Trans-regulation of thoracic homeotic selector genes of the Antennapedia and bithorax complexes by the trithorax group genes: absent, small, and homeotic discs 1 and 2. *Mechanisms of Development* **53:** 123–139

Laprell F, Finkl K & Müller J (2017) Propagation of Polycomb-repressed chromatin requires sequence-specific recruitment to DNA. *Science* **356:** 85–88

Larschan E, Alekseyenko AA, Gortchakov AA, Peng S, Li B, Yang P, Workman JL, Park PJ & Kuroda MI (2007) MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism. *Mol. Cell* **28:** 121–133

Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, Johnstone SE, Cole MF, Isono K-I, Koseki H, Fuchikami T, Abe K, Murray HL, Zucker JP, Yuan B, Bell GW, Herbolsheimer E, Hannett NM, Sun K, et al (2006) Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **125:** 301–313

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

Lewis RA, Wakimoto BT, Denell RE & Kaufman TC (1980) Genetic analysis of the Antennapedia gene complex (Ant-C) and adjacent chromosomal regions of Drosophila melanogaster. II. Polytene chromosome segments 84A-84B1,2. *Genetics* **95:** 383–397

Li B, Gogol M, Carey M, Pattenden SG, Seidel C & Workman JL (2007) Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription. *Genes & Development* **21:** 1422–1430

Lowary PT & Widom J (1998) New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276:** 19–42

Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM & Misteli T (2010) Regulation of alternative splicing by histone modifications. *Science* **327:** 996–1000

Luger K, Rechsteiner TJ & Richmond TJ (1999) Expression and purification of recombinant histones and nucleosome reconstitution. *Methods Mol. Biol.* **119:** 1–16

Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C & Bryant SH (2015) CDD: NCBI's conserved domain database. *Nucleic Acids Research* **43:** D222–6

Margueron R, Justin N, Ohno K, Sharpe ML, Son J, Drury WJ, Voigt P, Martin SR, Taylor WR, De Marco V, Pirrotta V, Reinberg D & Gamblin SJ (2009) Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* **461:** 762–767

Martinez A-M, Colomb S, Déjardin J, Bantignies F & Cavalli G (2006) Polycomb group-dependent Cyclin A repression in Drosophila. *Genes & Development* **20:** 501–513

Maurange C & Paro R (2002) A cellular memory module conveys epigenetic inheritance of hedgehog expression during Drosophila wing imaginal disc development. *Genes & Development* **16:** 2672–2683

Mazo AM, Huang DH, Mozer BA & Dawid IB (1990) The trithorax gene, a trans-acting regulator of the bithorax complex in Drosophila, encodes a protein with zinc-binding domains. *Proc. Natl. Acad. Sci. U.S.A.* **87:** 2112–2116

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

Mohan M, Herz H-M, Smith ER, Zhang Y, Jackson J, Washburn MP, Florens L, Eissenberg JC & Shilatifard A (2011) The COMPASS family of H3K4 methylases in Drosophila. *Mol. Cell. Biol.* **31:** 4310–4318

Mohrmann L, Langenberg K, Krijgsveld J, Kal AJ, Heck AJR & Verrijzer CP (2004) Differential targeting of two distinct SWI/SNF-related Drosophila chromatin-remodeling complexes. *Mol. Cell. Biol.* **24:** 3077–3088

Morata G & Kerridge S (1982) The role of position in determining homoeotic gene function in Drosophila. *Nature* **300:** 191–192

Moshkin YM, Mohrmann L, van Ijcken WFJ & Verrijzer CP (2007) Functional differentiation of SWI/SNF remodelers in transcription and cell cycle control. *Mol. Cell. Biol.* **27:** 651–661

Müller J, Hart CM, Francis NJ, Vargas ML, Sengupta A, Wild B, Miller EL, O'Connor MB, Kingston RE & Simon JA (2002) Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell* **111:** 197–208

Nekrasov M, Klymenko T, Fraterman S, Papp B, Oktaba K, Köcher T, Cohen A, Stunnenberg HG, Wilm M & Müller J (2007) Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes. *EMBO J.* **26:** 4078–4088

Nekrasov M, Wild B & Müller J (2005) Nucleosome binding and histone methyltransferase activity of Drosophila PRC2. *EMBO Rep* **6:** 348–353

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

Nowak AJ, Alfieri C, Stirnimann CU, Rybin V, Baudin F, Ly-Hartig N, Lindner D & Muller CW (2011) Chromatin-modifying complex component Nurf55/p55 associates with histones H3 and H4 and Polycomb Repressive Complex 2 subunit Su(z)12 through partially overlapping binding sites. *J. Biol. Chem.* **286:** 23388–23396

Noyes MB, Christensen RG, Wakabayashi A, Stormo GD, Brodsky MH & Wolfe SA (2008) Analysis of homeodomain specificities allows the family-wide prediction of preferred recognition sites. *Cell* **133:** 1277–1289

Oktaba K, Gutiérrez L, Gagneur J, Girardot C, Sengupta AK, Furlong EEM & Müller J (2008) Dynamic regulation by polycomb group protein complexes controls pattern formation and the cell cycle in Drosophila. *Developmental Cell* **15:** 877–889

Orlando V, Strutt H & Paro R (1997) Analysis of chromatin structure by in vivo formaldehyde cross-linking. *Methods* **11:** 205–214

Papoulas O, Beek SJ, Moseley SL, McCallum CM, Sarte M, Shearn A & Tamkun JW (1998) The Drosophila trithorax group proteins BRM, ASH1 and ASH2 are subunits of distinct protein complexes. *Development* **125:** 3955–3966

Papp B & Müller J (2006) Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. *Genes & Development* **20:** 2041–2054

Park SY, Schwartz YB, Kahn TG, Asker D & Pirrotta V (2012) Regulation of Polycomb group genes Psc and Su(z)2 in Drosophila melanogaster. *Mechanisms of Development***:** 1–12

Pengelly AR, Copur Ö, Jäckle H, Herzig A & Müller J (2013) A histone mutant reproduces the phenotype caused by loss of histone-modifying factor Polycomb. *Science* **339:** 698–699

Pengelly AR, Kalb R, Finkl K & Müller J (2015) Transcriptional repression by PRC1 in the absence of H2A monoubiquitylation. *Genes & Development* **29:** 1487–1492

Phelan ML, Sif S, Narlikar GJ & Kingston RE (1999) Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. *Mol. Cell* **3:** 247–253

Poepsel S, Kasinath V & Nogales E (2018) Cryo-EM structures of PRC2 simultaneously engaged with two functionally distinct nucleosomes. *Nat. Struct. Mol. Biol.* **25:** 154–162

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

Pu L & Sung ZR (2015) PcG and trxG in plants - friends or foes. *Trends Genet.* **31:** 252–262

Qiao Q, Li Y, Chen Z, Wang M, Reinberg D & Xu R-M (2011) The structure of NSD1 reveals an autoregulatory mechanism underlying histone H3K36 methylation. *J. Biol. Chem.* **286:** 8361–8368

Rickels R, Hu D, Collings CK, Woodfin AR, Piunti A, Mohan M, Herz H-M, Kvon E & Shilatifard A (2016) An evolutionary conserved epigenetic mark of polycomb response elements implemented by Trx/MLL/COMPASS. *Mol. Cell* **63:** 318–328

Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M & Séraphin B (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* **17:** 1030–1032

Rogawski DS, Ndoj J, Cho HJ, Maillard I, Grembecka J & Cierpicki T (2015) Two loops undergoing concerted dynamics regulate the activity of the ASH1L histone methyltransferase. *Biochemistry* **54:** 5401–5413

Rubin GM & Spradling AC (1982) Genetic transformation of Drosophila with transposable element vectors. *Science* **218:** 348–353

Samson ML, Jackson-Grusby L & Brent R (1989) Gene activation and DNA binding by Drosophila Ubx and abd-A proteins. *Cell* **57:** 1045–1052

Sapountzi V, Logan IR & Robson CN (2006) Cellular functions of TIP60. *Int. J. Biochem. Cell Biol.* **38:** 1496–1509

Saunders A, Werner J, Andrulis ED, Nakayama T, Hirose S, Reinberg D & Lis JT (2003) Tracking FACT and the RNA polymerase II elongation complex through chromatin in vivo. *Science* **301:** 1094–1096

Savla U, Benes J, Zhang J & Jones RS (2008) Recruitment of Drosophila Polycomb-group proteins by Polycomblike, a component of a novel protein complex in larvae. *Development* **135:** 813–817

Scheuermann JC, de Ayala Alonso AG, Oktaba K, Ly-Hartig N, McGinty RK, Fraterman S, Wilm M, Muir TW & Müller J (2010) Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB. *Nature* **465:** 243–247

Schmähling S, Meiler A, Lee Y, Mohammed A, Finkl K, Tauscher K, Israel L, Wirth M, Philippou-Massier J, Blum H, Habermann B, Imhof A, Song J-J & Müller J (2018) Regulation and function of H3K36 di-methylation by the trithorax-group protein complex AMC. *Development* **145:** dev163808

Schmitges FW, Prusty AB, Faty M, Stützer A, Lingaraju GM, Aiwazian J, Sack R, Hess D, Li L, Zhou S, Bunker RD, Wirth U, Bouwmeester T, Bauer A, Ly-Hartig N, Zhao K, Chan H, Gu J, Gut H, Fischle W, et al (2011) Histone methylation by PRC2 is inhibited by active chromatin marks. *Mol. Cell* **42:** 330–341

Schuettengruber B, Martinez A-M, Iovino N & Cavalli G (2011) Trithorax group proteins: switching genes on and keeping them active. *Nat. Rev. Mol. Cell Biol.* **12:** 799–814

Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W & Selbach M (2011) Global quantification of mammalian gene expression control. *Nature* **473:** 337–342

Schwartz YB, Kahn TG, Nix DA, Li X-Y, Bourgon R, Biggin M & Pirrotta V (2006) Genome-wide analysis of Polycomb targets in Drosophila melanogaster. *Nat. Genet.* **38:** 700–705

Schwartz YB, Kahn TG, Stenberg P, Ohno K, Bourgon R & Pirrotta V (2010) Alternative epigenetic chromatin states of Polycomb target genes. *PLoS Genet* **6:** e1000805

Shao Z, Raible F, Mollaaghababa R, Guyon JR, Wu CT, Bender W & Kingston RE (1999) Stabilization of chromatin structure by PRC1, a Polycomb complex. *Cell* **98:** 37–46

Shearn A & Garen A (1974) Genetic control of imaginal disc development in Drosophila. *Proc. Natl. Acad. Sci. U.S.A.* **71:** 1393–1397

Shearn A, Hersperger E & Hersperger G (1987) Genetic studies of mutations at two loci of Drosophila melanogaster which cause a wide variety of homeotic transformations. *Developmental biology***:** 1–12

Shearn A, Rice T, Garen A & Gehring W (1971) Imaginal disc abnormalities in lethal mutants of Drosophila. *Proc. Natl. Acad. Sci. U.S.A.* **68:** 2594–2598

Smith ER, Lee MG, Winter B, Droz NM, Eissenberg JC, Shiekhattar R & Shilatifard A (2008) Drosophila UTX is a histone H3 Lys27 demethylase that colocalizes with the elongating form of RNA polymerase II. *Mol. Cell. Biol.* **28:** 1041–1046

Smith ST, Petruk S, Sedkov Y, Cho E, Tillib S, Canaani E & Mazo A (2004) Modulation of heat shock gene expression by the TAC1 chromatin-modifying complex. *Nat. Cell Biol.* **6:** 162–167

Spradling AC, Stern D, Beaton A, Rhem EJ, Laverty T, Mozden N, Misra S & Rubin GM (1999) The Berkeley Drosophila Genome Project gene disruption project: Single P-element insertions mutating 25% of vital Drosophila genes. *Genetics* **153:** 135–177

Srinivasan S, Dorighi KM & Tamkun JW (2008) Drosophila Kismet regulates histone H3 lysine 27 methylation and early elongation by RNA polymerase II. *PLoS Genet* **4:** e1000217

Stabell M, Larsson J, Aalen RB & Lambertsson A (2007) Drosophila dSet2 functions in H3-K36 methylation and is required for development. *Biochemical and Biophysical Research Communications* **359:** 784–789

Stapleton M, Carlson J, Brokstein P, Yu C, Champe M, George R, Guarin H, Kronmiller B, Pacleb J, Park S, Wan K, Rubin GM & Celniker SE (2002) A Drosophila full-length cDNA resource. *Genome Biol.* **3:** RESEARCH0080

Steunou A-L, Cramet M, Rossetto D, Aristizabal MJ, Lacoste N, Drouin S, Côté V, Paquet E, Utley RT, Krogan N, Robert F, Kobor MS & Côté J (2016) Combined action of histone reader modules regulates NuA4 local acetyltransferase function but not its recruitment on the genome. *Mol. Cell. Biol.* **36:** 2768–2781

Strahl BD, Grant PA, Briggs SD, Sun Z-W, Bone JR, Caldwell JA, Mollah S, Cook RG, Shabanowitz J, Hunt DF & Allis CD (2002) Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Mol. Cell. Biol.* **22:** 1298–1306

Struhl G (1982) Genes controlling segmental specification in the Drosophila thorax. *Proc. Natl. Acad. Sci. U.S.A.* **79:** 7380–7384

Suganuma T, Pattenden SG & Workman JL (2008) Diverse functions of WD40 repeat proteins in histone recognition. *Genes & Development* **22:** 1265–1268

Sun B, Hong J, Zhang P, Dong X, Shen X, Lin D & Ding J (2008) Molecular basis of the interaction of Saccharomyces cerevisiae Eaf3 Chromo domain with methylated H3K36. *J. Biol. Chem.* **283:** 36504–36512

Tagami H, Ray-Gallet D, Almouzni G & Nakatani Y (2004) Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* **116:** 51–61

Tanaka Y, Katagiri Z-I, Kawahashi K, Kioussis D & Kitajima S (2007) Trithorax-group protein ASH1 methylates histone H3 lysine 36. *Gene* **397:** 161–168

Therrien M, Morrison DK, Wong AM & Rubin GM (2000) A genetic screen for modifiers of a kinase suppressor of Ras-dependent rough eye phenotype in Drosophila. *Genetics* **156:** 1231–1242

Thummel CS & Pirrotta V (1991) New pCasper P-element vectors. *Dros. Inf. Newslett.*

Tie F, Banerjee R, Saiakhova AR, Howard B, Monteith KE, Scacheri PC, Cosgrove MS & Harte PJ (2014) Trithorax monomethylates histone H3K4 and interacts directly with CBP to promote H3K27 acetylation and antagonize Polycomb silencing. *Development* **141:** 1129–1139

Tie F, Furuyama T, Prasad-Sinha J, Jane E & Harte PJ (2001) 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* **128:** 275–286

Tolhuis B, de Wit E, Muijrers I, Teunissen H, Talhout W, van Steensel B & van Lohuizen M (2006) Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in Drosophila melanogaster. *Nat. Genet.* **38:** 694–699

Tripoulas N, LaJeunesse D, Gildea J & Shearn A (1996) The Drosophila ash1 gene product, which is localized at specific sites on polytene chromosomes, contains a SET domain and a PHD finger. *Genetics* **143:** 913–928

Tripoulas NA, Hersperger E, La Jeunesse D & Shearn A (1994) Molecular genetic analysis of the Drosophila melanogaster gene absent, small or homeotic discs1 (ash1). *Genetics* **137:** 1027–1038

Venkatesh S & Workman JL (2013) Set2 mediated H3 lysine 36 methylation: regulation of transcription elongation and implications in organismal development. *Wiley Interdiscip Rev Dev Biol* **2:** 685–700

Venken KJT, Carlson JW, Schulze KL, Pan H, He Y, Spokony R, Wan KH, Koriabine M, de Jong PJ, White KP, Bellen HJ & Hoskins RA (2009) Versatile P[acman] BAC libraries for transgenesis studies in Drosophila melanogaster. *Nature Publishing Group* **6:** 431–434

Venken KJT, He Y, Hoskins RA & Bellen HJ (2006) P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. *Science* **314:** 1747–1751

Vermaak D, Wade PA, Jones PL, Shi YB & Wolffe AP (1999) Functional analysis of the SIN3-histone deacetylase RPD3-RbAp48-histone H4 connection in the Xenopus oocyte. *Mol. Cell. Biol.* **19:** 5847–5860

Voigt P, LeRoy G, Drury WJ, Zee BM, Son J, Beck DB, Young NL, Garcia BA & Reinberg D (2012) Asymmetrically modified nucleosomes. *Cell* **151:** 181–193

Wagner EJ & Carpenter PB (2012) Understanding the language of Lys36methylation at histone H3. *Nat. Rev. Mol. Cell Biol.* **13:** 115–126

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

Wang L, Brown JL, Cao R, Zhang Y, Kassis JA & Jones RS (2004b) Hierarchical recruitment of polycomb group silencing complexes. *Mol. Cell* **14:** 637–646

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

Xie T, Graveline R, Kumar GS, Zhang Y, Krishnan A, David G & Radhakrishnan I (2012) Structural basis for molecular interactions involving MRG domains: implications in chromatin biology. *Structure* **20:** 151–160

Xie T, Zmyslowski AM, Zhang Y & Radhakrishnan I (2015) Structural Basis for Multi-specificity of MRG Domains. *Structure* **23:** 1049–1057

Xu C, Cui G, Botuyan MV & Mer G (2008) Structural basis for the recognition of methylated histone H3K36 by the Eaf3 subunit of histone deacetylase complex Rpd3S. *Structure/Folding and Design* **16:** 1740–1750

Yuan W, Xu M, Huang C, Liu N, Chen S & Zhu B (2011) H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J. Biol. Chem.* **286:** 7983–7989

Zhang P, Du J, Sun B, Dong X, Xu G, Zhou J, Huang Q, Liu Q, Hao Q & Ding J (2006) Structure of human MRG15 chromo domain and its binding to Lys36-methylated histone H3. *Nucleic Acids Research* **34:** 6621–6628

Zheng W, Ibáñez G, Wu H, Blum G, Zeng H, Dong A, Li F, Hajian T, Allali-Hassani A, Amaya MF, Siarheyeva A, Yu W, Brown PJ, Schapira M, Vedadi M, Min J & Luo M (2012) Sinefungin derivatives as inhibitors and structure probes of protein lysine methyltransferase SETD2. *J. Am. Chem. Soc.* **134:** 18004–18014

6 Appendix

6.1 Abbreviations

A.T hook	adenine-thymine hook
A1 to A8	abdominal segments 1 to 8
aa	amino acids
abd-A	*abdominal A*
Abd-B	*Abdominal B*
AEBSF	4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride
alphaTub67C	Tubulin alpha-4 chain
AMC	Ash1:MRG15:Caf1-55 complex
ANT-C	Antennapedia complex
Antp	*Antennapedia*
ARID1A/1B	AT-rich interactive domain-containing protein 1A/1B
Ash1	Absent, small, or homeotic discs 1
Ash1-CTAP/NTAP-Ash1	Ash1 fusion protein with C-/N-terminal TAP tag
ASH1L	ASH1-like
Ash2	Absent, small, or homeotic discs 2
Asx	Additional sex combs
attB site	*attachment site on bacterial DNA*
attP site	*attachment site on phage DNA*
AWS	associated with SET domain
BAC	bacterial artificial chromosome
BAH	bromo-adjacent homology
BAP	Brahma-associated protein complex
Bap55/60/111/170	Brahma-associated protein 55/60/111/170 kD
BDGP R6	Berkeley *Drosophila* genome project release 6 genome assembly
betaTub56D	Tubulin beta-1 chain
bp	base pair
BRD2/4	Bromodomain-containing protein 2/4
BRDT	Bromodomain testis-specific
BRG1	Brahma homolog 1
Brm	Brahma
BSA	bovine serum albumin
BX-C	bithorax complex
Caf1-55	55 kDa subunit of Chromatin assembly factor 1
CBP	Calmodulin binding peptide
CHD7	Chromodomain-helicase-DNA-binding protein 7
ChIP	chromatin immunoprecipitation
ChIP-seq	ChIP-DNA sequencing

chr3R/L	right/left arm of chromosome 3
Chromo	chromatin organization modifier domain
COMPASS	Complex of proteins associated with SET1
complete	EDTA–free protease inhibitor cocktail from Roche
ddH_2O	double-distilled water
Df	deficiency allele
Dfd	*Deformed*
Dm or *Drosophila*	*Drosophila melanogaster*
dpr12	*defective proboscis extension response 12*
dRAF	dRing-associated factors
DTT	dithiothreitol
E(z)	Enhancer of zeste
Eaf3	ESA1-associated factor 3
EED	Embryonic ectoderm development
EMS	ethyl methanesulfonate
Esc	Extra sex combs
FLP	flippase
FRT	*flippase recognition target site*
Fsh	Female sterile homeotic
H disc/H organ	haltere disc/organ
H3K36me1/2/3	histone H3 lysine 36 mono/di/tri-methylation
H3KAc	histone H3 lysine acetylation
HAT	histone acetyltransferase
Hcf1	Host cell factor 1
HDAC	histone deacetylase
His-tag	hexa-histidine tag
HMTase	histone methyltransferase
hp	hypopleurite
hrs	hours
Hs	*Homo sapiens*
Hsp70	Heat shock protein 70
iBAQ	Intensity based absolute quantification
IgG	immunoglobulin G
Jarid2	Jumonji, AT rich interactive domain 2
kb	kilobases
kDa	kilodalton
KDM2/6A/6B	Lysine-specific demethylase 2/6A/6B
KIS-L	KISMET-L
Kto	Kohtalo
L3 disc	3rd leg disc
lab	*labial*
LB	lysogeny broth
m+	possesses maternally loaded protein/RNA

MCS	multiple cloning site
MLL1/2	Mixed lineage leukemia protein 1/2
Mm	*Mus musculus*
Mor	Moira
MRG15	MORF4-related gene on chromosome 15
MRGBP	MRG/MORF4L-binding protein
Msl3	Male-specific lethal-3
NIPBL	Nipped-B-like
NSD	Nuclear receptor-binding SET domain
NuA4	Nucleosome acetyltransferase of histone H4
NuRD	Nucleosome remodeling and deacetylase complex
NURF	Nucleosome remodeling factor
Ogt	O-glycosyltransferase
ORF	open reading frame
P1-3	baculovirus passages 1-3
PA	polyacrylamide
PAGE	polyacrylamide gel electrophoresis
para	paragraph(s)
pb	*proboscipedia*
Pb	Polybromo
PBAP	Polybromo-containing BAP
Pc	*Polycomb*
PcG	Polycomb group of genes
Pcl	Polycomblike
Pf1	PHD finger protein 12
Ph-d	Polyhomeotic-distal
Ph-p	Polyhomeotic-proximal
PHD	plant homeodomain
Pho	Pleiohomeotic
Phol	Pleiohomeotic-like
PhoRC	Pho-repressive complex
PR-DUB	Polycomb repressive deubiquitinase
PRC1	Polycomb repressive complex 1
PRC2	Polycomb repressive complex 2
PRE	Polycomb response element
Psc	Posterior sex combs
qPCR	quantitative polymerase chain reaction
Rbbp5	Retinoblastoma binding protein 5
RNAP II	RNA polymerase II
Rpd3S	Reduced potassium dependency-3 small
RpL10Ab	Ribosomal protein L10Ab
RpL18A	60S ribosomal protein L18A
RpL27	Ribosomal protein L27

RpL4	60S ribosomal protein L4
S2-DRSC cells	Schneider's 2-*Drosophila* RNAi Screening Center cells
SAM	sterile α-motif
Sayp	Supporter of activation of yellow protein
Sce	Sex comb extra
Scm	Sex comb on midleg
Scr	*Sex combs reduced*
SET	Su(var)3-9, Enhancer-of-zeste and Trithorax
SET1	SET domain-containing protein 1
SET2	SET domain-containing protein 2
Sfmbt	Scm-like with four MBT domain-containing protein 1
Skd	Skuld
SMARCC1/2	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily C member 1/2
Snr1	Snf5-related 1
sp	sternopleurite
Su(z)12	Suppressor of zeste 12
Su(z)2	Suppressor of zeste 2
SWI/SNF	Switch/sucrose nonfermenting
T1 to T3	thoracic segments 1 to 3
TAP	tandem affinity purification
TEV	*Tobacco Etch Virus*
Tip60	*Tat-interactive protein 60 kDa*
TRAP230/240	Thyroid hormone receptor-associated protein complex 230/240
Trr	Trithorax-related
trx	*trithorax*
trxG	trithorax group of genes
tsh	*teashirt*
tub	*α-tubulin1*-promoter
Ubx	*Ultrabithorax*
UTY	Ubiquitously-transcribed TPR protein on the Y chromosome
v/v	volume per volume
Vtd	Verthandi
w	*white* gene
w/v	weight per volume
WD40 β-propeller	tryptophan-aspartic acid 40 β-propeller
Wds	Will die slowly
wg	*wingless*
wt	*wild type*
z-	gene not zygotically expressed

6.2 List of Figures

6.3 List of Tables

6.4 Supplementary data: peptides found in TAP-Ash1 eluates

Table 12. Ash1, MRG15 and Caf1-55 peptide sequences identified by mass spectrometry in eluates of NTAP-Ash1 and Ash1-CTAP purifications.

	Ash1-CTAP (in-solution digest)	Ash1-CTAP (in-gel digest)	NTAP-Ash1 (in-gel digest)
Ash1			
ADIDADNYQCER		+	
AIQSIKDSYEQQK	+		+
AKEETIVQTAVPR	+	+	+
AEVESPIISAIDIKEDTK	+		+
AIEEGEEITYDYNFSIFNPSEGQPCR	+		
EQAEAAPQPPPKSEPEIRPAK	+		+
CICGIYK		+	
CICGIYKDEGIMIQCSK		+	
CIDAQTAQEQPIDISYIISGR		+	
CMVWQHTECTK		+	
DEGIMIQCSK		+	
DICSAMETIK		+	
DIPIKDESGK		+	+
DISSAVAVAK		+	
DKNIPQYQSTIIQDFMEK		+	
DSPIVPIKVTPPPIIPIEASPDEDVIR		+	+
EEIQIDPIWR		+	+
EETIVQTAVPR		+	
EIDVNKKFR			+
EIPIEEFTEEGHR		+	
EKPVQPVTVEEIPPEIPVSQEEIDAEAEAK		+	+
EQAEAAPQPPPK		+	+
ERDSPIVPIK		+	+
EVDREIPIEEFTEEGHR		+	+
EVISSEEEPGK	+	+	+
EVISSEEEPGKIAVK	+		+
FMTADKGWGVR	+		+
FVNHSCEPNCEMQK	+	+	
FYPNEVVR	+	+	
GGSISATNPDNFISK	+	+	+
GISAPADATAVHVVTPVAPNK	+	+	+
GITQAVHDPEIEKMAK	+	+	+
GRPMECNDEDHCYICEIR	+	+	
GTYIIEYVGEVVTEKEFK	+	+	+
GVIGGKSQR			+
HAVAPGVER			+
HIIEQPTSVSGAGSSASNSPIR	+	+	

	Ash1-CTAP (in-solution digest)	Ash1-CTAP (in-gel digest)	NTAP-Ash1 (in-gel digest)
HYIITPGERPPAEVAFANGK	+	+	+
IASYVQIVEIIGDSESIQSFKPK		+	+
IDMAYIDKR	+		+
IGSTAATSKVEFR	+		+
IDSIPTEHDPIPASESHNPGPQDYASCSESSEDK		+	
IENVIITMK		+	+
IIDISPSSICSIK		+	
IITEIEIITSTFNSR		+	+
INAEAWAAAAAAAK		+	+
INESVITK		+	+
INRTGFPTVR			+
IPDGIDPNTNFSCK		+	
IQDDRITGSSGK			+
IQATIAAPSPAQQITINGGGPASTISK		+	
IQPISEKEK			+
IQRPQTPAR			+
ISAIRPTIGVVATK			+
ISVVAIER		+	+
IVVDNNSISGGK	+	+	+
KQKTEIDVGAGPTTMHK	+		
KVSVEQQTTAVIDEHEPEFDPDDEPIQSIRETR	+		
KVVPTVPAPGNSGPAINESADSGVISTTSTTQSTTPSPK	+		
IASSSGISK			+
MQNENAVPTGSIPIASSSKPK	+	+	+
MSDIITTVSSKK	+	+	+
MVIFAK		+	+
MVYTECSPSNCPAGEK	+	+	
NAHKNPAETDSITDQSSQSK	+	+	+
NIEAGTQPK			+
NIPQYQSTIIQDFMEK		+	+
NPAETDSITDQSSQSK		+	+
NREQAEAAPQPPPK		+	
NVVPSWNYR		+	+
QAGKDISSAVAVAK		+	+
QFNTFIVR		+	
QGDAVYVIR	+	+	+
QKTEIDVGAGPTTMHK		+	
QPVIEEPPPTPPPQQK		+	+
QSIMPPPAK		+	+
RVESDTEDTTVEGSFRK		+	
RFYPNEVVR		+	+
RIDSIPTEHDPIPASESHNPGPQDYASCSESSEDK		+	
RPSTPSSPSIAAQISAICSPR		+	
RTEMDFEIPYDIWWAYTNSK		+	

	Ash1-CTAP (in-solution digest)	Ash1-CTAP (in-gel digest)	NTAP-Ash1 (in-gel digest)
RVESDTEDTTVEGSFR		+	
SATQFSVQR	+	+	+
SDTDGIRMR			+
SGYVSDYGSVR			+
SINIDSK			+
SIQAQVEQGHYKTPQEFDDHMQQIFVEAK		+	+
SIKSATQFSVQR			+
SIKSATQFSVQRSDTDGIR			+
SMSVGAASGTGASTTICSK	+	+	
SIPTTSASK			+
SQAQFNAR			+
SQSNDSSSPDDHKIPIK			+
SRIENVIITMK	+		+
SSAASMCSSYVSGVSR		+	
SSADDTVEDQDIIQIAGISIGQSSEESNEYISKPSIK		+	+
SSNNVNVQAAPNPPIDCERVPQAGEAR	+	+	
STASTKSQAQFNAR			+
SYAPHDVDPSIIK	+	+	+
TEIDVGAGPTTMHK	+	+	+
TEMDFEIPYDIWWAYTNSK		+	
TGGNIIIK			+
TNVYAESVRPNIAGFDHPTCNCK		+	
TPQEFDDHMQQIFVEAK		+	+
TQMIGQTVNAK		+	+
TYIVAGIFSNHYK	+	+	
VESDTEDTTVEGSFR			+
VIAAKSGYVSDYGSVR			+
VIYPPPR			+
VKKTYIVAGIFSNHYK		+	
VKPIPAVEAKPSGEGISGR		+	+
VPQAGEARETFVAR			+
VSIYEVVPIEIVIGR	+	+	+
VSVEQQTTAVIDEHEPEFDPDDEPIQSIR	+	+	+
VVETIIHK	+	+	+
WSVNGISR		+	+
YVTTGQYFGR		+	+
YYISIMR		+	+
MRG15			
AIPQVAIDINFSKGDR	+	+	
ASTPSKDSNTSQSTASSTPTTSAGPGSK	+		+
DSNTSQSTASSTPTTSAGPGSK		+	
FAIGGGEVIK		+	
EDPAAAETTEEEGPVAPK		+	

	Ash1-CTAP (in-solution digest)	Ash1-CTAP (in-gel digest)	NTAP-Ash1 (in-gel digest)
GEVKPAKVENYSTGTDANTIFVDGER	+		
GWDVGVAGMK		+	
HPDTPISEIYGSFHIIR		+	
HYITDDWYAVVR		+	
IGPNSTIVFEVEIK		+	
IIEIPAK	+	+	+
IKIPDEIKHYITDDWYAVVR	+	+	
IQSNNKTFDSIIK	+		
IVDQVVGK		+	
IVDQVVGKGEEAK	+		
MSEQRPSITGSDVAEKPIPPTTTPSTPTTEPAPCVESEEAYAAK		+	
NSSIFFSMSNFINVDPEYVR		+	+
NWDEWVPENR		+	+
QEYIVATVTK		+	
SEAGSTGTTTTNSTANSTTSR	+	+	+
SFHISGVAIDKGQEAK	+	+	
SMFWGINMKPER		+	
TKPDATPVEYYIHYAGWSK	+	+	+
TQYADVMQK	+	+	+
VENYSTGTDANTIFVDGER		+	+
VICFHGPIIYEAK	+	+	
VITCPPHMAYGAR		+	
VSVYYIGR	+		
VTVQQISEQYIAHK		+	+
YSQTIIK		+	
Caf1-55			
APAVGIDIGTTYSCVGVFQHGK		+	
ARFEEINADIFR		+	
ATIDEDNIK	+	+	
EIEGVCNPIITK		+	
ETAEAYIGK		+	
FDDAAVQSDMK		+	
FEEINADIFR		+	
FEISGIPPAPR		+	
GEFGGFGSVCGK		+	
HPSKPEPSGECQPDIR		+	
HWPFEVVSADGKPK		+	
IEIEIK		+	
IEVTYKDEK		+	+
IGEEQSTEDAEDGPPEIIFIHGGHTAK	+	+	+
IHSFESHKDEIFQVQWSPHNETIIASSGTDR		+	
IHVWDISK		+	+
IIIGTHTSDEQNHIIIASVQIPSEDAQFDGSHYDNEK		+	

	Ash1-CTAP (in-solution digest)	Ash1-CTAP (in-gel digest)	NTAP-Ash1 (in-gel digest)
IINEPTAAAIAYGIDK	+	+	
IIQDIFNGK		+	
IMIWDTR		+	
IVTHFVQEFK		+	
KFDDAAVQSDMK		+	
MKETAEAYIGK	+	+	
NGIESYCFNMK		+	
NIFTGHTAVVEDVAWHIIHESIFGSVADDQK		+	
NQVAMNPTQTIFDAK	+	+	
PSHTVDAHTAEVNCISFNPYSEFIIATGSADK		+	
QKEIEGVCNPIITK		+	
QTQTFTTYSDNQPGVIIQVYEGER		+	
RIHVWDISK		+	
SDNAAESFDDAVEER		+	+
SINPDEAVAYGAAVQAAIIHGDK		+	
STAGDTHIGGEDFDNR		+	
SVIHDIVIVGGSTR		+	
TFFPEEISSMVITK		+	
TPSSDVIVFDYTK		+	
TTPSYVAFTDTER	+	+	
TVAIWDIR		+	+
TVTNAVITVPAYFNDSQR	+	+	
VDRSDNAAESFDDAVEER	+		
VEIIANDQGNR		+	
VINEEYKIWK			+
WIDANQIADKEEYEHR		+	
YMPQNACVIATK		+	

6.5 Curriculum Vitae

EDUCATION AND QUALIFICATIONS

07/2011 – 06/2019	**PhD** at Max Planck Insitute of Biochemistry, Martinsried Research group: Dr. Jürg Müller Focus: Epigenetic transcriptional regulation in *Drosophila melanogaster* Fellow of the graduate program "**International Max Planck Research School for Molecular Life Sciences**"
10/2006 – 09/2010	***Diplom* in Biochemistry** (equiv. to MSc) at *Freie Universität*, Berlin Final grade: 1 (grading from best to worst: 1 - 5)
10/2007 – 07/2008	**Year abroad** in Buenos Aires, Argentina Laboratory internships at *Universidad de Buenos Aires*
10/2004 – 07/2006	***Vordiplom* in Biochemistry** (approx. equiv. to BSc) at *Freie Universität* Final grade: 1 (grading from best to worst: 1 - 5)
10/2003 – 07/2004	**Studies in Romance Languages** at *Johannes Gutenberg-Universität*, Mainz
03/2003	***Abitur*** (equiv. to high school diploma) at *Sebastian Münster-Gymnasium*, Ingelheim Final grade: 1,3 (grading from best to worst: 1,0 – 6,0)
09/1999 – 12/1999	**Student exchange** with Lethbridge Collegiate Inst., Lethbridge, Canada

PUBLICATIONS

- Schmähling, S., Meiler, A., Lee, Y., Mohammed, A., Finkl, K., Tauscher, K., Israel, L., Borath, M., Philippou-Massier, J., Blum, H., Habermann, B., Imhof, A., Song, J.-J., and Müller, J. **Regulation and function of H3K36 di-methylation by the trithorax-group protein complex AMC.** Development 145, dev163808 (2018).

- Lee, Y., Yoon, E., Cho, S., Schmähling, S., Müller, J. and Song, J.-J. **Structural basis of MRG15-mediated activation of the ASH1L histone methyltransferase by releasing an auto-inhibitory loop.** Structure, in press.

CONFERENCES

- EMBL conference "Transcription and Chromatin", EMBL, Heidelberg, 2018
  Poster presentation
- "Chromatin Day", SFB 1064 meeting – Chromatin Dynamics, Martinsried, 2017. **Talk**
- 4DCellFate meeting, Lisbon, Portugal, 2016. **Short talk**
- EMBL conference "Transcription and Chromatin", EMBL, Heidelberg, 2016
  Poster presentation
- Keystone Symposium, "Chromatin and Epigenetics", Whistler, Canada, 2016
  Poster presentation
- Spetses Summer School, "Chromatin and Systems Biology", Spetses, Greece, 2013
  Poster presentation

RESEARCH EXPERIENCE

07/2011 – 06/2019	**PhD** at Max Planck Institute of Biochemistry, Martinsried Supervision: Dr. Jürg Müller Title: "Biochemical purification and functional characterization of a novel trithorax-group protein complex" Principal methods: Protein purification**,** histone methyltransferase assays, baculovirus expression system in insect cell culture, ChIP-qPCR, RNA-Seq, *Drosophila* genetics
01/2010 – 09/2010	**Project to obtain *Diplom* degree** at Robert Koch Institute, Berlin Supervision: Dr. Sebastian Voigt Title: "Characterization of the putative immunomodulator e127 of the rat cytomegalovirus" Principal methods: Cell and viral culture, virus purification by density gradient centrifugation, ELISA, flow cytometry, quantitative PCR
11/2008 – 12/2008	**Internship** at Max Delbrück Centre, Berlin Research group: Dr. Frank Rosenbauer Subject: Long distant regulation of the PU.1 promoter by an insulator element in hematopoiesis
04/2008 – 05/2008	**Internship** at *Universidad de Buenos Aires*, IByME Research group: Prof. Alejandro F. De Nicola Subject: Neuroprotective effects of progesterone and allopregnanolone in the Wobbler mouse model for motor neuron diseases
11/2007 – 12/2007	**Internship** at *Universidad de Buenos Aires*, INGEBI Research group: Dr. Martin Vazquez Subject: Splice factors of the trans-spliceosome in *Trypanosoma cruzi*

RESEARCH EXPERIENCE

05/2007 – 06/2007	**Internship** at Institute of Medical Genetics, Charité, Berlin Supervisor: Dr. Marc Trimborn Subject: The BRIT1 protein in cell cycle control
01/2007 – 02/2007	**Internship** at *Freie Universität*, Berlin Research group: Prof. Gerhard Multhaup Subject: Comparative analysis of keratins in tortoise shells using MALDI-TOF/TOF mass spectrometry

© 2018. Published by The Company of Biologists Ltd | Development (2018) 145, dev163808. doi:10.1242/dev.163808

RESEARCH ARTICLE

Regulation and function of H3K36 di-methylation by the trithorax-group protein complex AMC

Sigrun Schmähling[1], Arno Meiler[2], Yoonjung Lee[3], Arif Mohammed[1], Katja Finkl[1], Katharina Tauscher[1], Lars Israel[4], Marc Wirth[4], Julia Philippou-Massier[5], Helmut Blum[5], Bianca Habermann[2], Axel Imhof[4], Ji-Joon Song[3] and Jürg Müller[1,*]

ABSTRACT

The *Drosophila* Ash1 protein is a trithorax-group (trxG) regulator that antagonizes Polycomb repression at HOX genes. Ash1 di-methylates lysine 36 in histone H3 (H3K36me2) but how this activity is controlled and at which genes it functions is not well understood. We show that Ash1 protein purified from *Drosophila* exists in a complex with MRG15 and Caf1 that we named AMC. In *Drosophila* and human AMC, MRG15 binds a conserved FxLP motif near the Ash1 SET domain and stimulates H3K36 di-methylation on nucleosomes. *Drosophila MRG15*-null and *ash1* catalytic mutants show remarkably specific trxG phenotypes: stochastic loss of HOX gene expression and homeotic transformations in adults. In mutants lacking AMC, H3K36me2 bulk levels appear undiminished but H3K36me2 is reduced in the chromatin of HOX and other AMC-regulated genes. AMC therefore appears to act on top of the H3K36me2/me3 landscape generated by the major H3K36 methyltransferases NSD and Set2. Our analyses suggest that H3K36 di-methylation at HOX genes is the crucial physiological function of AMC and the mechanism by which the complex antagonizes Polycomb repression at these genes.

KEY WORDS: Trithorax group, Ash1, MRG15, Histone H3K36 methylation, *Drosophila*

INTRODUCTION

In organisms ranging from yeast to humans, the chromatin spanning the transcribed region of active genes is modified by di- and tri-methylation of lysine 36 in histone H3 (H3K36me2/3). Although nucleosomes in the 5′ regions of transcribed genes are predominantly di-methylated at H3K36, nucleosomes in the 3′ region of genes mainly carry the H3K36me3 modification (Bell et al., 2007; Pokholok et al., 2005). Among the different roles ascribed to H3K36me2/3 (Venkatesh and Workman, 2013), there is accumulating evidence for two principal mechanisms by which this modification impacts on gene transcription. First, studies in yeast revealed that the H3K36me2/3 mark is recognized by the chromo barrel domain of the Eaf3 subunit (Sun et al., 2008; Xu et al., 2008) of the Rpd3S complex that, by deacetylating nucleosomes in the transcribed region, suppresses initiation of transcription at intragenic sites (Venkatesh and Workman, 2013). H3K36me2/3 is thus thought to participate in the quality control of transcription by preventing production of unwanted transcripts. Second, in metazoans, the H3K36me2/3 modification allosterically inhibits the histone methyltransferase (HMTase) activity of Polycomb Repressive Complex 2 (PRC2) and thereby prevents PRC2 from depositing H3K27me3 on H3K36me2/3-modified nucleosomes (Schmitges et al., 2011; Yuan et al., 2011). H3K36me2/3 was therefore proposed to protect transcriptionally active genes from becoming tri-methylated at H3K27 and thereby getting repressed by the Polycomb system. This antagonism between H3K36me2/3 and PRC2 is thought to be particularly crucial at developmentally regulated genes that, although active in some cells, are at the same time repressed by Polycomb in other cells of the body (Gaydos et al., 2012; Klymenko and Müller, 2004).

In yeast, all H3K36 di- and tri-methylation is generated by a single histone methyltransferase, Set2, that associates with the phosphorylated form of elongating RNA polymerase II (Krogan et al., 2003; Venkatesh and Workman, 2013; Xiao et al., 2003). In metazoans, SET2 is responsible for generating the bulk of H3K36me3, whereas NSD generates the bulk of H3K36me2 (Bell et al., 2007; Gaydos et al., 2012; Larschan et al., 2007).

Higher metazoans contain an additional SET-domain HMTase that di-methylates H3K36, called Ash1 (An et al., 2011; Dorighi and Tamkun, 2013; Tanaka et al., 2007; Yuan et al., 2011). The *absent, small, or homeotic discs 1* (*ash1*) gene in *Drosophila* was first identified because of the phenotype of *ash1* mutants, which developed into pharate adults and showed homeotic transformations in several body segments (Shearn et al., 1987). The similarity of the homeotic phenotypes of *ash1* and *trithorax* mutants led to the classification of *ash1* as a trithorax-group (trxG) regulator (Shearn, 1989). As expected from the phenotype, *ash1* mutants show loss of expression of multiple HOX genes within their normal expression domains (LaJeunesse and Shearn, 1995). However, in *ash1* mutants that also lack PRC2, HOX gene expression is restored to normal levels and, in addition, these double mutants also show widespread misexpression of HOX genes, similar to PRC2 single mutants (Klymenko and Müller, 2004). This suggested that, at least at HOX genes, Ash1 is not required for transcriptional activation per se but is needed to antagonize instalment of Polycomb repression. It is important to note that in the wild type, PRC2 and other Polycomb group (PcG) protein complexes are bound at target genes, not only in the cells in which these genes are repressed but also in the cells in

[1]Max-Planck Institute of Biochemistry, Laboratory of Chromatin Biology, Am Klopferspitz 18, 82152 Martinsried, Germany. [2]Max-Planck Institute of Biochemistry, Computational Biology, Am Klopferspitz 18 82152 Martinsried, Germany. [3]Korea Advanced Institute of Science and Technology (KAIST), Department of Biological Sciences, Daejeon 34141, Korea. [4]Zentrallabor für Proteinanalytik, BioMedical Center, Ludwig-Maximilians-University Munich, Großhadernerstr. 9, 82152 Martinsried, Germany. [5]Laboratory for Functional Genome Analysis (LAFUGA), Gene Center, Ludwig-Maximilians-Universität München, Feodor-Lynen-Strasse 25, 81377 Munich, Germany.

*Author for correspondence (muellerj@biochem.mpg.de)

J.M., 0000-0003-3837-3817

Received 24 January 2018; Accepted 5 March 2018

Development (2018) 145, dev163808. doi:10.1242/dev.163808

which they are expressed (Bowman et al., 2014; Kang et al., 2017; Kwong et al., 2008; Langlais et al., 2012; Papp and Müller, 2006). Nevertheless, at active genes, PRC2 fails to tri-methylate H3K27 in their chromatin. In *ash1* mutants, however, PRC2 deposits H3K27me3 ectopically across the entire promoter and coding region (Papp and Müller, 2006). Such ectopic methylation by PRC2 in *ash1* mutants is also detected on polytene chromosomes, where several genomic sites show an increase in H3K27me3 immunofluorescence signal (Dorighi and Tamkun, 2013; Srinivasan et al., 2008). Moreover, genome-wide ectopic H3K27 tri-methylation is also observed in *C. elegans* mutants lacking the NSD orthologue Mes-4 (Gaydos et al., 2012). Together with the above-mentioned finding that H3K36me2/3 inhibits H3K27 methylation by PRC2 on nucleosomes *in vitro* (Schmitges et al., 2011; Yuan et al., 2011), these observations collectively suggested that Ash1 keeps HOX and possibly also other target genes active by di-methylating H3K36 in the transcribed region of their chromatin and thereby preventing H3K27me3 deposition and instalment of Polycomb repression. However, several aspects that are central to this model have remained unresolved. First, the Ash1 protein alone shows only weak HMTase activity because its SET domain is auto-inhibited (An et al., 2011). This raises the question of how Ash1 catalytic activity becomes stimulated. Second, at least in *Drosophila* tissue culture cells, the bulk of H3K36me2 is generated by NSD (Bell et al., 2007) and it is not known where and to what extent Ash1 contributes to H3K36 di-methylation, in particular at HOX target genes. Third, it is not known whether Ash1 also regulates genes other than HOX genes during *Drosophila* development.

Here, we have biochemically purified Ash1 protein complexes from *Drosophila* and characterized their activity *in vitro* and in the developing organism. Our work reveals that Ash1 HMTase activity is activated by MRG15, a subunit of the identified Ash1 complex, and we show that this complex, rather than the Ash1 protein alone, is the active form of this H3K36 methyltransferase, both *in vitro* and *in vivo*. We show that Ash1 is not needed for global H3K36 di-methylation but is essential to generate normal H3K36me2 levels at HOX and other genes that we found to be de-regulated in *ash1* mutants. The specific homeotic phenotypes of mutants that lack Ash1 or Ash1 HMTase activity establish that H3K36 di-methylation at HOX genes is a key physiological function of the Ash1 protein complex for *Drosophila* morphogenesis.

RESULTS

Biochemical purification and reconstitution identify MRG15 and Caf1 as Ash1 complex subunits

To identify proteins that form stable assemblies with the Ash1 protein in *Drosophila*, we used a tandem affinity purification (TAP) strategy (Rigaut et al., 1999) and purified N- or C-terminally tagged Ash1 (NTAP-Ash1 or Ash1-CTAP, respectively) from embryonic nuclear extracts (Fig. 1A). The transgenes expressing NTAP-Ash1 or Ash1-CTAP in these assays rescued animals homozygous for the $ash1^{22}$ null mutation (Tripoulas et al., 1996) into morphologically normal and fertile adults, permitting the purification of these fusion proteins from animals lacking untagged endogenous Ash1. Mass spectrometric analyses of the purified material identified MRG15 and Caf1-55 (for simplicity referred to as Caf1) as the two major proteins that co-purified with both NTAP-Ash1 and Ash1-CTAP (Fig. 1B). In both purifications we failed to detect Fsh1, a protein that was previously reported to co-purify with Ash1 from *Drosophila* tissue culture cells (Kockmann et al., 2013).

We tested whether interactions of MRG15 and Caf1 with Ash1 could be reconstituted with recombinant proteins. MRG15, like its yeast orthologue Eaf3, contains a chromo barrel domain that binds H3K36me2/3 (Zhang et al., 2006) and an MRG domain (Fig. 2A). Structural studies on MRG15 had revealed that the protein uses its MRG domain to bind extended regions of its interaction partners via high-affinity interactions that are centred around a conserved FxLP motif in those partner proteins (Xie et al., 2012, 2015). Inspection of the *Drosophila* and vertebrate Ash1 protein sequences identified in each case a single FxLP motif in a conserved location, about 40 amino acid residues N-terminal to the AWS domain that precedes the catalytic SET domain (Fig. 2A). Using baculovirus expression vectors, we co-expressed *Drosophila* $Ash1_{1041\text{-}2226}$ ($Ash1_C$) with

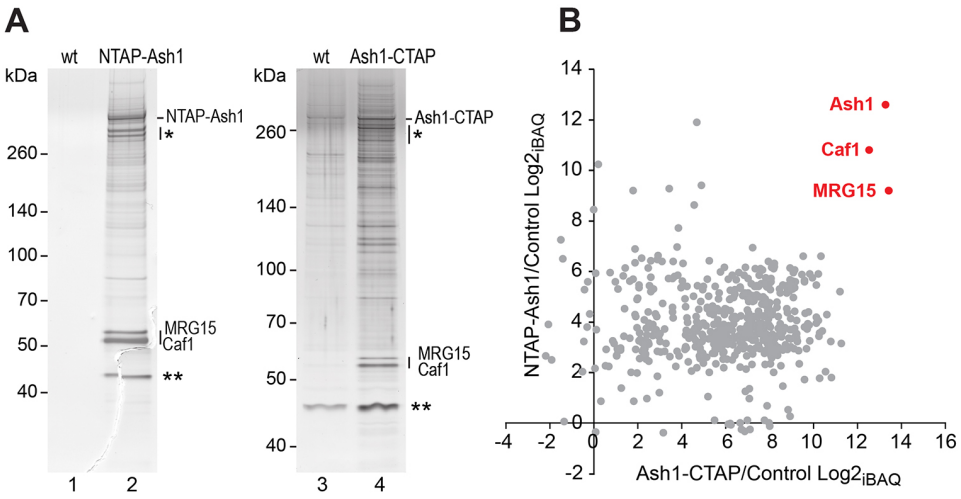


Fig. 1. Ash1 purified from *Drosophila* exists in a complex with MRG15 and Caf1. (A) Proteins isolated by tandem affinity purification (TAP) from wild-type (wt) and *α-tubulin1-NTAP-Ash1* embryos (left), and from wild-type and *α-tubulin1-Ash1-CTAP* embryos (right) separated on a 4-12% polyacrylamide gel and visualized by silver staining. The bands marked by an asterisk are TAP-Ash1 degradation products; the bands marked with two asterisks were considered as non-specific because they were also detected in several mock TAPs from the wild-type control (e.g. lane 3). (B) Scatterplot representation of log2 transformed $iBAQ_{Ash1}/iBAQ_{mock}$ ratios (Intensity Based Absolute Quantification) from the mass spectrometric data of the same NTAP-Ash1 and Ash1-CTAP purifications shown in A (see Tables S1 and S2).

Development (2018) 145, dev163808. doi:10.1242/dev.163808

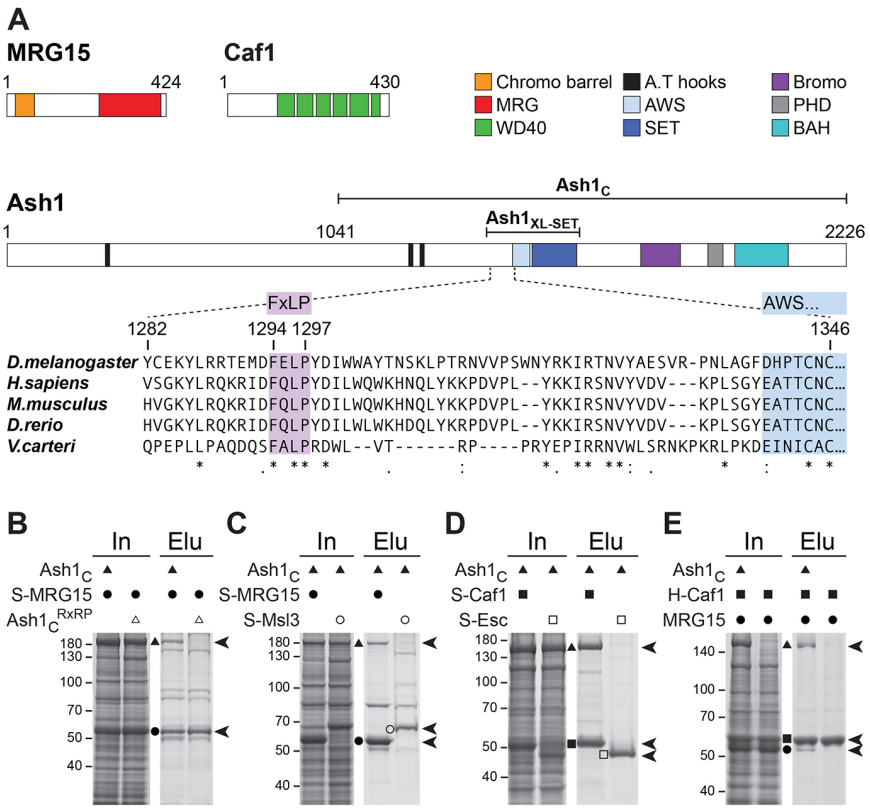


Fig. 2. Reconstitution of recombinant AMC and identification of the conserved Ash1 FxLP motif as the MRG15 interaction site. (A) Domain architecture of MRG15, Caf1 and Ash1 and alignment of Ash1 protein sequences harbouring the FxLP motif. (B-E) Total extracts (In) from Hi-5 cells co-expressing the indicated AMC subunits and material isolated from these cells (Elu) by Strep- (B-D) or His- (E) affinity purification, separated on 8% (B,C,E) or 10% (D) SDS polyacrylamide gels and visualized by Coomassie Blue staining. (E) Use of the His-tag on Caf1 and a different running buffer accounts for the slower migration behaviour of Caf1 relative to MRG15 in this experiment. See text for a description of the results.

Strep-tagged full-length MRG15 (S-MRG15) in Hi-5 insect cells and then performed Strep-affinity purification. This resulted in the isolation of an $Ash1_C$:MRG15 complex (Fig. 2B). In contrast, when we co-expressed S-MRG15 with Ash1C containing a mutation of the FxLP motif to RxRP, this mutant $Ash1_C^{RxRP}$ protein failed to co-purify with S-MRG15 (Fig. 2B). This suggests that the FxLP motif in $Ash1_C$ is crucial for interaction with MRG15. As control for the specificity of this interaction, we also tested for interaction of Ash1C with Strep-tagged Msl3 (S-Msl3). Msl3 is another MRG domain protein that uses a similar mode of interaction like MRG15 to bind a conserved FxLP motif in its binding partner MSL1 (Xie et al., 2015). However, strep affinity-purification from cells co-expressing $Ash1_C$ with S-Msl3 resulted in the isolation of S-Msl3 alone (Fig. 2C). We conclude that $Ash1_C$, via the conserved FxLP motif, directly and specifically interacts with MRG15.

We next investigated whether Caf1 directly interacts with $Ash1_C$. Caf1 is a WD40 β-propeller protein that is a core subunit of several chromatin-modifying complexes, including PRC2, and is required for cell viability (Anderson et al., 2011). Co-expression of $Ash1_C$ with S-Caf1 resulted in the isolation of an $Ash1_C$:Caf1 complex. In contrast, $Ash1_C$ did not co-purify with S-Esc, another WD40 β-propeller subunit of PRC2, used as control (Fig. 2D). $Ash1_C$ therefore directly and specifically interacts with Caf1.

Finally, we tested whether Caf1 and MRG15 also directly bind each other. His-affinity purification from cells co-expressing His-tagged Caf1 (H-Caf1) and MRG15 resulted in the isolation of H-Caf1 alone (Fig. 2E). When $Ash1_C$ was also co-expressed, a trimeric complex containing substoichiometric amounts of MRG15 could be isolated (Fig. 2E). This suggests that Caf1 and MRG15 do not interact directly and that $Ash1_C$ forms a scaffold to which both proteins bind independently of each other. In conclusion, these biochemical purification and reconstitution assays identify Ash1, MRG15 and Caf1-55 as subunits of a novel protein complex that we named AMC.

MRG15 stimulates H3K36 di-methylation in *Drosophila* and human AMC

Previous studies reported that C-terminal fragments of *Drosophila* or mammalian Ash1 protein containing the SET domain (Fig. 2A) have HMTase activity for dimethylation of H3K36 in nucleosomes (An et al., 2011; Eram et al., 2015; Miyazaki et al., 2013; Tanaka et al., 2007; Yuan et al., 2011). Structural and enzymatic studies on the isolated SET domain of human ASH1L found that this domain has only weak HMTase activity because it is auto-inhibited by a loop from the post-SET domain that blocks access to the substrate-binding pocket (An et al., 2011). In a first set of experiments, we performed HMTase assays with the $Ash1_C$:MRG15 or $Ash1_C$:Caf1 complexes described above (Fig. 2). As substrate, recombinant mononucleosomes were used and the reactions were monitored by western blot analysis with antibodies against H3K36me2 or H3K36me3. Under our experimental conditions, we were unable to detect HMTase activity in reactions with the $Ash1_C$:Caf1 complex (Fig. 3A, lanes 2-3) but, strikingly, the $Ash1_C$:MRG15 complex showed robust activity for generating H3K36me2 (Fig. 3A, lanes 4-5). Mutation of the Ash1 SET domain at Arg1464, a residue that stabilizes the orientation of the SAM-binding loop, abolished HMTase activity; the $Ash1_C^{R1464A}$:MRG15 complex failed to generate detectable levels of H3K36me2 (Fig. 3A, lanes 6-7). This control confirms that H3K36 di-methylation generated by the $Ash1_C$:MRG15 complex is catalysed by the Ash1 SET domain. We also note that the $Ash1_C$:MRG15 complex catalyses only H3K36 di- and not tri-methylation on nucleosomes (Fig. 3A). Taken together, these analyses show that association of MRG15 with Ash1 greatly enhances its catalytic

Development (2018) 145, dev163808. doi:10.1242/dev.163808

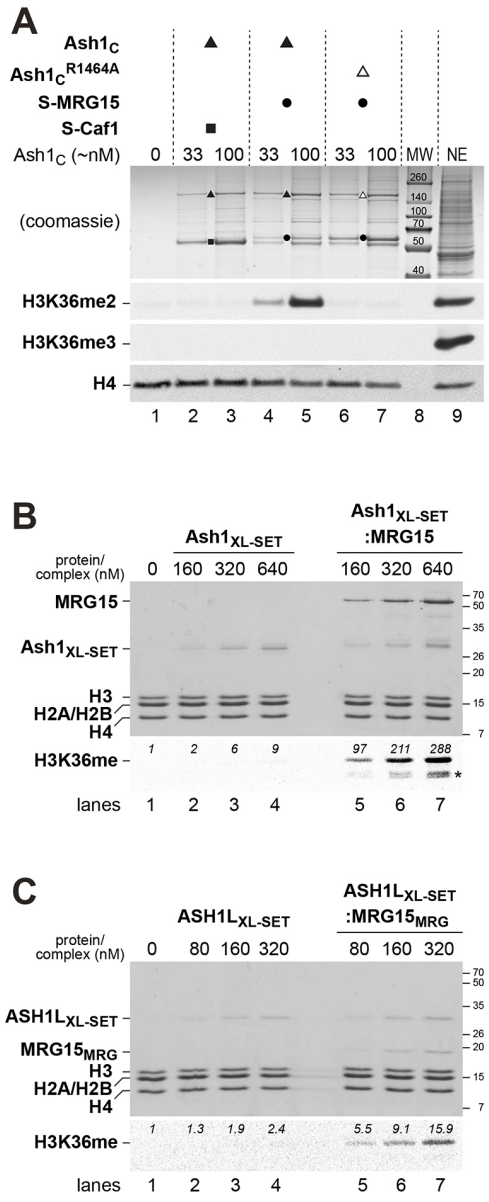


Fig. 3. H3K36 di-methylation by *Drosophila* Ash1 and human ASH1L is stimulated by MRG15. (A) HMTase reactions with recombinant *Drosophila* $Ash1_C$:Caf1 (lanes 2, 3), $Ash1_C$:MRG15 (lanes 4, 5) or $Ash1_C^{R1464A}$:MRG15 (lanes 6, 7) complexes and reconstituted recombinant mononucleosomes (400 nM in lanes 1-7) were separated on a 10% SDS-polyacrylamide gel; the upper part of the gel was stained with Coomassie Blue to visualize the enzyme complexes, the bottom part was analysed by western blotting with antibodies against H3K36me2 and, as a control, H4. The same reaction was also analysed with antibody against H3K36me3. Enzyme concentrations in the reactions were normalized by estimating $Ash1_C$ concentration relative to a Coomassie Blue-stained protein standard. *Drosophila* embryo nuclear extract (NE) in lane 9 served as control for western blot analysis. Lane 8: molecular weight marker (MW). (B) HMTase reactions with recombinant *Drosophila* $Ash1_{XL\text{-}SET}$ (lanes 2-4) or $Ash1_{XL\text{-}SET}$:MRG15 complex (lanes 5-7) on reconstituted recombinant *Xenopus* oligonucleosomes (320 nM in lanes 1-7). One part of the reaction was analysed on a 15% SDS-polyacrylamide gel to visualize proteins by Coomassie Blue staining (top), the other part of the reaction was separated on a 15% SDS-polyacrylamide gel, transferred to membrane and analysed by fluorography. PhotoShop software was used to quantify radioactive signal in the H3 band; this signal represents the sum of H3K36me1 and H3K36me2. Asterisk indicates H3 degradation products. (C) HMTase reactions with recombinant human $ASH1L_{XL\text{-}SET}$ (lanes 2-4) or $ASH1L_{XL\text{-}SET}$:$MRG15_{MRG}$ complex (lanes 5-7) on reconstituted recombinant *Xenopus* oligonucleosomes (320 nM in lanes 1-7). Reactions were analysed as in B.

activity for H3K36 di-methylation and it therefore appears that the AMC complex and not Ash1 alone is the active form of this HMTase.

We next wanted to quantify the stimulatory effect of MRG15 on Ash1 HMTase activity. To achieve this, we expressed and purified a short soluble fragment of the *Drosophila* Ash1 protein comprising the SET domain and the preceding amino acid sequences that include the FxLP motif ($Ash1_{1275\text{-}1522}$, called $Ash1_{XL\text{-}SET}$; Fig. 2A and Fig. 3B, lanes 2-4). Like $Ash1_C$, this $Ash1_{XL\text{-}SET}$ fragment also interacted with full-length MRG15 and could be purified as a stable $Ash1_{XL\text{-}SET}$:MRG15 complex (Fig. 3B, lanes 5-7). We performed HMTase assays on recombinant oligonucleosome arrays using S-[methyl-^{3}H] adenosylmethionine and quantified methyl-^{3}H incorporation into histone H3 using fluorography (Fig. 3B). The $Ash1_{XL\text{-}SET}$:MRG15 complex showed about 30-fold higher HMTase activity compared with $Ash1_{XL\text{-}SET}$ alone under our assay conditions (Fig. 3B, compare lanes 5-7 with lanes 2-4).

To extend these analyses, we also investigated the interaction between human ASH1L and MRG15. Specifically, we reconstituted and purified a recombinant minimal $ASH1L_{XL\text{-}SET}$:$MRG15_{MRG}$ complex containing an $ASH1L_{XL\text{-}SET}$ protein fragment ($ASH1L_{2035\text{-}2288}$) in complex with the MRG domain ($MRG15_{151\text{-}362}$, called $MRG15_{MRG}$) of human MRG15 (Fig. 3C, lanes 5-7). The human $ASH1L_{XL\text{-}SET}$:$MRG15_{MRG}$ complex showed about a sevenfold higher HMTase activity than the $ASH1L_{XL\text{-}SET}$ protein alone (Fig. 3C, compare lanes 5-7 with lanes 2-4). We conclude that the interaction between ASH1L and MRG15, and the stimulatory effect of MRG15 on ASH1L HMTase activity is conserved in humans. Moreover, these results also suggest that interaction of the MRG domain of MRG15 with an ASH1L fragment comprising only the SET domain and the preceding amino acid stretch with the FxLP motif (i.e. $ASH1L_{XL\text{-}SET}$) is sufficient to activate ASH1L HMTase activity by almost an order of magnitude.

AMC HMTase activity is required for viability and is crucial for HOX gene regulation in *Drosophila*

In the next set of experiments, we investigated the physiological role of AMC and its H3K36 di-methyltransferase activity in *Drosophila*. To this end, we generated animals that completely lacked Ash1 or MRG15 protein, or mutants expressing full-length but catalytically inactive Ash1 protein.

In the first experiment, we analysed the *ash1*-null mutant phenotype. Earlier studies had identified and characterized $ash1^{22}$ as a protein null mutation (Tripoulas et al., 1996). Consistent with these earlier reports, we were unable to detect Ash1 protein in larval extracts from $ash1^{22}$ homozygotes (Fig. S1). It is important to note that previous studies investigating the requirement of Ash1 during *Drosophila* development had analysed *ash1* mutant animals that were derived from heterozygous mothers and therefore contained maternally deposited wild-type Ash1 protein during the early stages of development (Shearn, 1989; Shearn et al., 1987; Tripoulas et al., 1996). The phenotype of *ash1*-null mutant animals that lack both maternally deposited and zygotically expressed Ash1 protein, in the following referred to as $ash1^{m-z-}$ mutants, has not been described. We generated $ash1^{22\ m-z-}$ mutant animals from females with $ash1^{22}$ mutant germ cells. Interestingly, these $ash1^{22\ m-z-}$ mutant animals

Development (2018) 145, dev163808. doi:10.1242/dev.163808

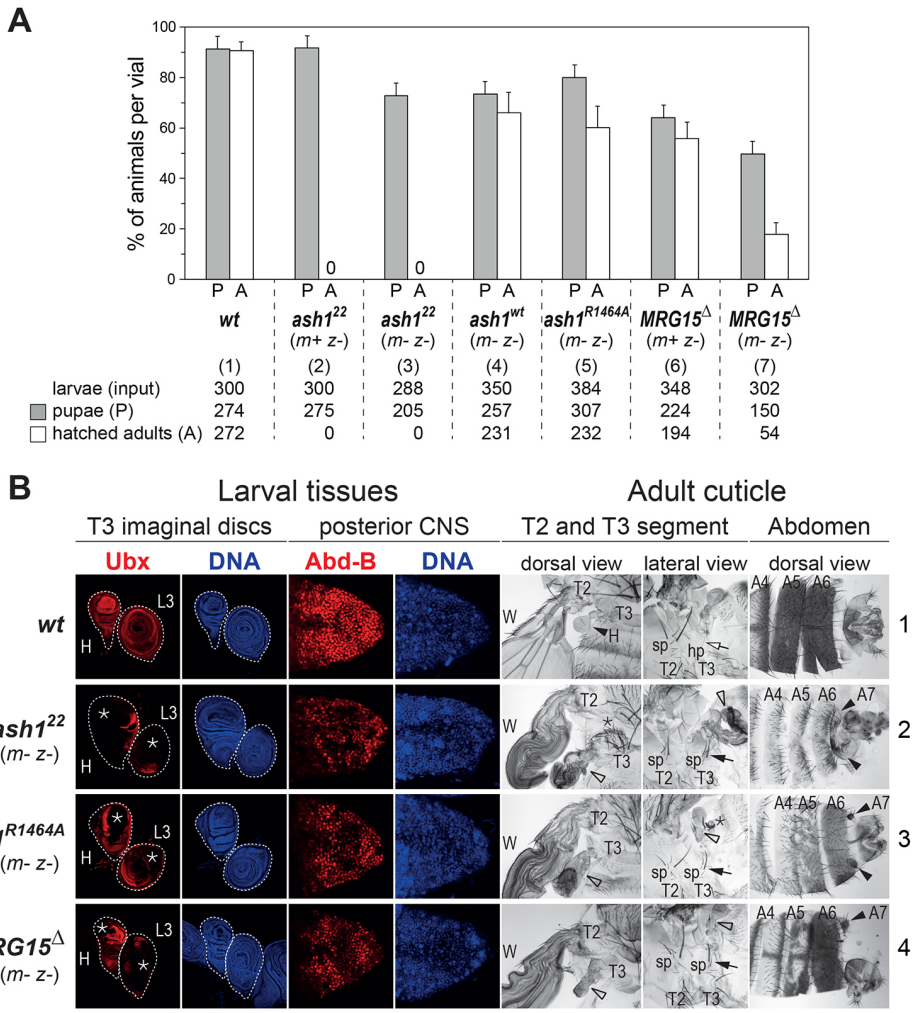


Fig. 4. Requirement of AMC HMTase activity for viability and HOX gene regulation. (A) Viability of *Drosophila* with different *ash1* and *MRG15* mutant genotypes. For each genotype [(1)-(7)], indicated numbers of 1st/2nd instar larvae (input) were isolated, distributed into at least three different food vials and reared. In each vial, the percentage of animals that formed pupae (grey bar) and eclosed from the pupal case (white bar) was determined. Histograms represent the mean±s.d. of these percentages in individual vials of a given genotype. $ash1^{22\ m+\ z-}$ and $ash1^{22\ m-\ z-}$ animals do not eclose from the pupal case (asterisk). The genotype of $MRG15^{\Delta m+\ z-}$ and $MRG15^{\Delta m-\ z-}$ animals shown here and in B is $MRG15^{\Delta}/Df(3R)BSC741$ (see text). (B) Lack of AMC function causes a specific HOX loss-of-function syndrome. Left: HOX gene expression analysis. Larval haltere (H) and third leg (L3) imaginal disc tissues stained using Ubx antibody (red) and co-stained with Hoechst to visualize nuclei (DNA, blue); CNS tissues stained using Abd-B antibody (red) and Hoechst (DNA, blue). In the wild type (wt), Ubx is expressed in all cells of the H and L3 disc. $ash1^{22\ m-\ z-}$, $ash1^{R1464A\ m-\ z-}$ and $MRG15^{\Delta m-\ z-}$ mutants show patchy loss of Ubx expression (asterisks) in irregular patterns in both discs; this phenotype is observed in all animals in all three genotypes (n>40) but comparison of discs from these larvae shows that the tissue area showing loss of expression is most extensive in $ash1^{22\ m-\ z-}$ mutants, somewhat less extensive in $ash1^{R1464A\ m-\ z-}$ mutants and even less extensive in $MRG15^{\Delta m-\ z-}$ mutants. In wild-type animals, Abd-B is expressed in all cells of the posterior CNS, whereas $ash1^{22\ m-\ z-}$, $ash1^{R1464A\ m-\ z-}$ and $MRG15^{\Delta m-\ z-}$ mutants all show patchy loss or reduction of Abd-B expression. Right: cuticle phenotype analysis. Preparations of cuticles from adults (row 1) or pharate adults (row 2-4) with dorsal and lateral views of T2 and T3 segment structures and dorsal views of the posterior abdomen, including the A4, A5 and A6 segments. Dorsal view T2 and T3: in wild-type animals, haltere disc tissues form the haltere (H, black arrowhead) in T3. $ash1^{22\ m-\ z-}$, $ash1^{R1464A\ m-\ z-}$ and $MRG15^{\Delta m-\ z-}$ mutants show haltere-to-wing transformations (empty arrowheads; wing in T2 marked as W in all cases), owing to loss of Ubx expression. The extent of this T3 to T2 transformation matches the extent of Ubx expression loss in the three mutant genotypes ($ash1^{22\ m-\ z-}$>$ash1^{R1464A\ m-\ z-}$>$MRG15^{\Delta m-\ z-}$). Lateral view T2 and T3: T3 to T2 transformation in $ash1^{22\ m-\ z-}$, $ash1^{R1464A\ m-\ z-}$ and $MRG15^{\Delta m-\ z-}$ mutants due to loss of Ubx expression in the 3L disc is manifested by transformation of the hypopleurite (hp) in T3 (empty arrow in wild-type animals) into sternopleurite (sp) tissue with sp bristles (black arrows), which are normally only found in T2. The haltere-to-wing (empty arrowhead) and meta- to mesonotum transformations (asterisk) are variable in animals of the same genotype (compare lateral and dorsal views). Dorsal view of abdomen: wild-type males show characteristic pigmentation in A5 and A6 that is almost completely lost in $ash1^{22\ m-\ z-}$ mutants or lost in a patchy pattern in $ash1^{R1464A\ m-\ z-}$ and $MRG15^{\Delta m-\ z-}$ mutants. A7 segment structures appearing in the three mutant genotypes are indicated.

developed up to the pupal stage and died as pharate adults, like $ash1^{22\ m+\ z-}$ mutants (Fig. 4A). $ash1^{22\ m-\ z-}$ pharate adults show a spectrum of anteriorly directed homeotic transformations that are similar to but slightly more severe than those of mutants lacking only zygotic expression of Ash1 (Shearn, 1989; Shearn et al., 1987). Specifically, $ash1^{22\ m-\ z-}$ mutant pharate adults showed transformation of the third (T3) to the second thoracic segment (T2) owing to widespread loss of expression of the HOX gene *Ultrabithorax* (*Ubx*) in larval haltere and third-leg imaginal disc primordia that form the T3 segment in adults (Fig. 4B, compare

Development (2018) 145, dev163808. doi:10.1242/dev.163808

rows 2 and 1). The abdominal segments A5 and A6 of *ash1*$^{22\ m-\ z-}$ pharate adults showed transformations towards A4, most noticeable by the loss of pigmentation in males (Fig. 4B, compare rows 2 and 1). Moreover, these males also develop an A7 segment that is normally suppressed in the wild type (Fig. 4B, compare rows 2 and 1). All these transformations are indicative of loss of expression of the HOX gene *Abdominal-B* (*Abd-B*) in the larval primordia of these adult structures. The requirement of Ash1 for normal expression of *Abd-B* is also apparent in the central nervous system (CNS), where *ash1*$^{22\ m-\ z-}$ mutant larvae show patchy loss of Abd-B expression (Fig. 4B, compare rows 2 and 1). Animals that completely lack Ash1 protein therefore show a specific HOX loss-of-function syndrome but, perhaps surprisingly, no other obvious morphological defects.

We then investigated whether loss of Ash1 HMTase activity is responsible for these phenotypes by analysing animals that expressed the Ash1^{R1464A} mutant protein instead of wild-type Ash1. As documented above, the R1464A mutation in Ash1 SET domain severely compromises HMTase activity (Fig. 3A). We generated animals that were homozygous for the *ash1*^{22} null mutation but carried a single transgene that expressed the Ash1^{R1464A} protein or, as a control, wild-type Ash1 protein from a genomic *ash1* fragment. In the control animals, the transgene-encoded wild-type Ash1 protein fully rescued *ash1*^{22} homozygotes into viable and fertile adults that were morphologically indistinguishable from wild-type *Drosophila* and could be maintained as a healthy strain (Fig. 4A). In contrast, the Ash1^{R1464A} mutant protein largely failed to rescue the homeotic phenotypes of *ash1*^{22} homozygotes but a substantial fraction of these animals nevertheless eclosed from the pupal case (Fig. 4A). The eclosed first generation, referred to as *ash1*$^{R1464A\ m+\ z-}$ mutants, were fertile and this permitted the generation of *ash1*$^{R1464A\ m-z-}$ animals in which not only zygotically expressed but also the maternally supplied Ash1 protein contained the R1464A point mutation. These *ash1*$^{R1464A\ m-z-}$ mutant animals showed loss of HOX gene expression and homeotic transformations almost as severe as *ash1*$^{22\ m-\ z-}$ mutant animals (Fig. 4B, compare rows 3 and 2). Moreover, the *ash1*$^{R1464A\ m-\ z-}$ mutant animals that eclosed from the pupal case invariably died after 1-2 days (Fig. 4A). The inability of the Ash1^{R1464A} protein to maintain normal HOX gene expression indicates that Ash1 HMTase activity is crucial for this process. The slightly less severe phenotype of *ash1*$^{R1464A\ m-\ z-}$ mutants and the finding that a fraction of these animals even eclose from the pupal case could be explained by low levels of residual HMTase activity of the Ash1^{R1464A} protein *in vivo* or, alternatively, by an HMTase-independent function of Ash1 in maintaining HOX gene transcription.

Finally, we tested the requirement of MRG15 for AMC function *in vivo*. To analyse *Drosophila* mutants lacking the MRG15 protein, we used homologous recombination (Gong and Golic, 2003) to generate *MRG15*$^{\Delta}$, an allele that deletes almost the entire MRG15-coding region (Fig. S2). Among the animals homozygous for *MRG15*$^{\Delta}$ that were derived from heterozygous parents, a substantial fraction developed into adults that eclosed from the pupal case (Fig. 4A) and showed mild HOX loss-of-function phenotypes in the adult epidermis as their only detectable morphological defect. Many of these *MRG15*$^{\Delta\ m+\ z-}$ animals died shortly after eclosing; better survival was observed in animals that were trans-heterozygous for *MRG15*$^{\Delta}$ and *Df(3R)BSC741*, another chromosomal deletion that removes the entire *MRG15* gene and additional flanking genes. For all experiments described below, we therefore used *MRG15*$^{\Delta}$/*Df(3R)BSC741* trans-heterozygous animals but for simplicity refer to them as *MRG15*$^{\Delta}$ mutants. *MRG15*$^{\Delta\ m+\ z-}$ surviving adults were fertile and produced *MRG15*$^{\Delta\ m-\ z-}$ progeny that lacked both maternally-deposited and zygotically-expressed MRG15 protein. A fraction of these *MRG15*$^{\Delta\ m-\ z-}$ mutant animals again developed into pupae and adults (Fig. 4A) that, intriguingly, showed loss of HOX gene expression and homeotic transformations that overall were almost as severe as those observed in *ash1*$^{R1464A\ m-\ z-}$ mutants (Fig. 4B, compare rows 4 and 3). This striking similarity of the *MRG15* null and *ash1* catalytic mutant phenotypes, together with the finding that MRG15 is required for efficient H3K36 di-methylation by Ash1 *in vitro* (Fig. 3), implies that MRG15 is also important for Ash1 HMTase activity *in vivo*. We note that the survival of *MRG15*$^{\Delta\ m-z-}$ mutant animals to adulthood and the specific homeotic phenotypes may seem surprising because MRG15 is also present in other chromatin protein complexes, such as the Tip60 complex (Kusch et al., 2004). It therefore appears that the role of MRG15 in AMC likely is the primary vital function of this protein in *Drosophila*. Finally, because of the specific homeotic phenotype of MRG15 mutants, we propose that MRG15 should, like Ash1, be classified as a trxG protein.

AMC is required for the regulation of several hundred genes

The analyses described above provide strong evidence that AMC activity is crucial for maintaining normal expression of HOX genes. To obtain a comprehensive overview of the genes that are regulated by this complex, we compared the transcriptome in imaginal disc tissues from *ash1*^{22} homozygous larvae (*ash1*$^{22\ m+\ z-}$) with that in the same tissues from wild-type larvae. To achieve this, we extracted RNA from hand-dissected batches of haltere and third-leg imaginal discs from the T3 segment, in the following referred to as T3 discs, and, in parallel, also from batches of wing imaginal discs from the T2 segment, referred to as T2 discs, from both wild-type and *ash1*^{22} homozygous larvae. Transcriptome sequencing was then performed on at least four independent biological replicates of each type of sample. Bioinformatic analyses of RNA-seq data from *ash1*^{22} mutant and wild-type larvae identified several hundred genes that were differentially expressed in both T3 and T2 discs. Specifically, about 300 genes are differentially expressed with a log2 fold change ≥2 and about 600 genes are differentially expressed with a log2 fold change ≥1 (Fig. 5A,B; Table S3). As expected from the analyses shown above (Fig. 4B), *Ubx* was downregulated more than fourfold in *ash1* mutant T3 discs (Fig. 5A) but was not detected as a differentially regulated gene in T2 discs (Fig. 5B) because it is not expressed in that tissue. Importantly, most other genes that were differentially expressed in T3 discs of *ash1* mutant and wild-type larvae were also differentially expressed in T2 discs (Fig. 5C). The altered expression of these genes in *ash1* mutants is therefore not a consequence of reduced expression of the transcription factor Ubx in T3 discs. In summary, these results show that Ash1 is required for the normal expression of a few hundred genes in addition to regulating the HOX genes.

These results raised the issue of whether Ash1 directly binds to these deregulated genes and whether Ash1 is required for dimethylation of H3K36 in their chromatin. Three separate studies have generated genome-wide Ash1 protein-binding profiles in different *Drosophila* tissue culture cell lines (Huang et al., 2017; Kockmann et al., 2013; Schwartz et al., 2010). Kockmann et al. reported that Ash1 binds in a sharply localized manner in the promoter region of most active genes in the cells they analysed (Kockmann et al., 2013), whereas Schwartz et al. reported that Ash1 is bound at about 50 genomic regions where it is associated with large chromatin domains that span on average about 10 kb

Development (2018) 145, dev163808. doi:10.1242/dev.163808

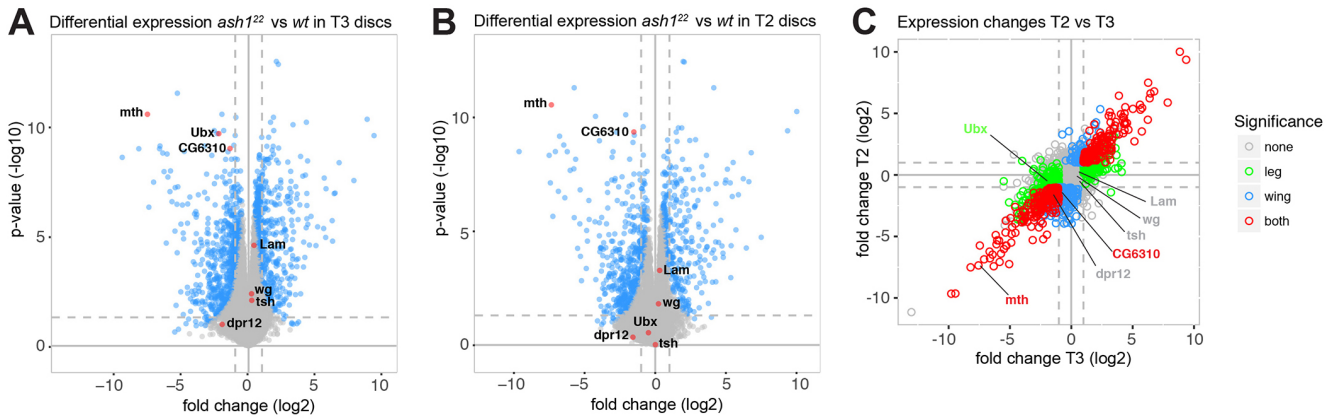


Fig. 5. Alteration of gene expression in *ash1* mutants. (A,B) Volcano plots of changes in gene expression in $ash1^{22\,m+\,z-}$ ($ash1^{-/-}$) larvae compared with wild-type larvae in T3 discs (A) and in T2 discs (B). Genes selected for the analyses documented in Fig. 6 are labelled. (C) Scatter plot comparing T2 and T3 by log2 fold change of gene expression in $ash1^{22\,m+\,z-}$ versus wild-type larvae, colour-coded by statistical significance in T2 or T3, or both (P<0.01, log2 fold change ≥2). There is high similarity in T3 and T2 tissues. Genes also labelled in A and B are indicated.

(Schwartz et al., 2010); Huang et al. identified around 500 Ash1-bound genes (Huang et al., 2017). To analyse Ash1 binding in the same tissues that we had used for our transcriptome analyses, we attempted to perform ChIP-seq experiments in imaginal discs. However, we failed to enrich sufficient amounts of chromatin by immunoprecipitation with Ash1 antibodies and were therefore unable to generate reliable Ash1 protein-binding profiles. In the following, we therefore focused our analysis on investigating how H3K36me2 levels are affected in mutants lacking AMC.

Bulk H3K36me2 levels are not diminished in mutants lacking AMC

In the first experiment, we assessed the contribution of AMC to total H3K36me2 levels in developing larvae. We analysed H3K36me2 bulk levels in imaginal disc tissues dissected from wild-type, $ash1^{22\,m+\,z-}$ or $MRG15^{\Delta\,m-\,z-}$ third instar larvae. As shown in Fig. 6A, H3K36me2 bulk levels were not detectably diminished in either of the two mutants. These results suggested that AMC might contribute to H3K36 di-methylation in a more gene-specific manner.

Ash1 is required for high-level H3K36 di-methylation at Ash1-regulated genes

We next analysed H3K36me2 levels at genes that we had found to be downregulated in *ash1* mutants. We performed chromatin immunoprecipitation (ChIP) assays with H3K36me2 antibodies on chromatin prepared from haltere and third leg imaginal discs (T3 discs) dissected from $ash1^{22}$ homozygous larvae ($ash1^{22\,m+\,z-}$) or from wild-type larvae and used real-time quantitative PCR (qPCR) to monitor H3K36me2 levels in the transcribed region of specific genes. These genes were: *methuselah* (*mth*), as an example of a gene that is strongly (almost 200-fold) downregulated in *ash1* mutants; *Ubx*, as a moderately (more than fourfold) downregulated gene; and *CG6310*, as a weakly (about twofold) downregulated gene (Fig. 5A). As control, we analysed H3K36me2 at the *wingless* (*wg*), *teashirt* (*tsh*) and *lamin* (*lam*) genes that are all expressed in T3 discs but are not downregulated in *ash1* mutants (Fig. 5A, Table S3). As an additional control, we analysed H3K36me2 at *dpr12*, a gene that is virtually inactive in T3 and T2 discs (Table S3). As illustrated in Fig. 6B, in wild-type animals, H3K36me2 is detected in the coding region of the expressed *Ubx*, *mth*, *CG6310*, *wg*, *tsh* and *lam* genes, but not at the inactive *dpr12* gene (Fig. 6B). In *ash1* mutants, H3K36me2 levels were strongly diminished at *Ubx* and *mth*, and were mildly reduced at *CG6310*, but not significantly changed at *wg*, *tsh* and *lam* (Fig. 6B). Nonetheless, even at the *Ubx* and *mth* genes, where H3K36me2 levels are strongly reduced, the modification is not completely abolished. This suggests that the residual H3K36 di-methylation at these genes is generated by other H3K36 HMTases, most likely by NSD and/or Set2. These HMTases are likely also responsible for H3K36me2 at the *wg*, *tsh* and *lam* genes, where the modification appeared undiminished in *ash1* mutants. It is important to keep in mind that even though the extent of H3K36me2 reduction appears to roughly match the extent to which expression of target genes is reduced in *ash1* mutants, the H3K36me2 reduction might not be uniform across the cell population. A case in point for this is *Ubx*, where expression in T3 discs of *ash1* mutants is lost in a mosaic all-or-none fashion (Fig. 4B). It is possible that, in cells showing loss of *Ubx* expression, H3K36me2 might be completely lost from *Ubx* chromatin and that the residual H3K36me2 ChIP signal at *Ubx* in *ash1* mutants (Fig. 6B) represents H3K36me2 at *Ubx* in those cells that retained normal levels of *Ubx* expression. In summary, these analyses show that genes that are downregulated in *ash1* mutants show a reduction, but not complete loss, of H3K36me2 in their chromatin.

DISCUSSION

Biochemical studies over the past have revealed that almost all PcG and trxG regulators originally identified through genetics are subunits of multi-protein complexes that modify chromatin; this has greatly helped to unravel the molecular mechanism of these proteins and to understand how they function (Kassis et al., 2017). Here, we have investigated the molecular interactions and the mechanism of action of the trxG protein Ash1 and its role in regulating gene expression in *Drosophila*. The work reported in this study leads to the following main conclusions. First, biochemical purifications from *Drosophila* show that Ash1 is the subunit of a multi-protein complex that contains MRG15 and Caf1: the AMC complex. Second, reconstitution of AMC using recombinant proteins uncovered that the MRG domain of MRG15 binds to a conserved FxLP motif next to the SET domain of Ash1, and that this interaction greatly enhances Ash1 catalytic activity for H3K36 di-methylation in nucleosomes. A recent study by Bing Zhu's lab

Development (2018) 145, dev163808. doi:10.1242/dev.163808

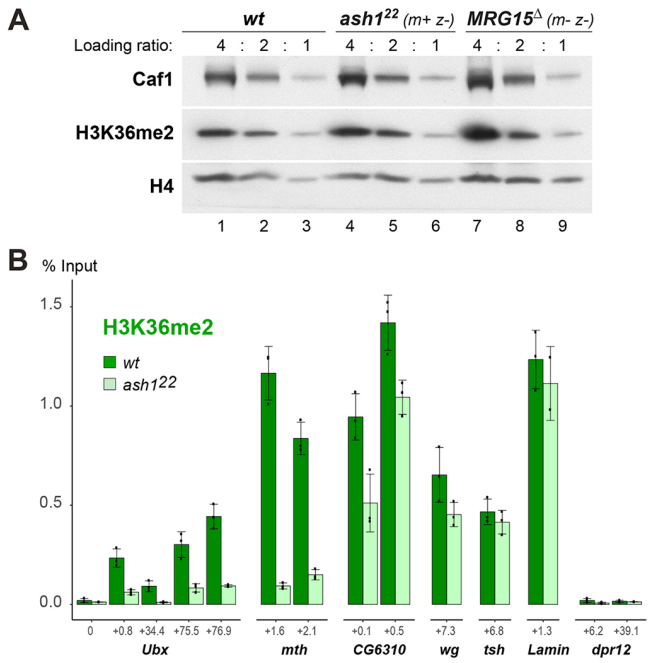


Fig. 6. Ash1 is required for normal H3K36me2 levels at HOX and other target genes. (A) H3K36me2 bulk levels are unchanged in *ash1* or *MRG15* null mutants. Western blot on serial dilutions (4:2:1) of total extracts from imaginal discs from wild-type (wt), *ash1*$^{22\ m+\ z-}$ and *MRG15*$^{\Delta m-\ z-}$ mutant larvae, probed with antibodies against H3K36me2 and, as loading control, histone H4 and Caf1. The genotype of the *MRG15*$^{\Delta m-\ z-}$ animals is *MRG15*$^{\Delta}$/*Df(3R)BSC741* (see text). (B) ChIP qPCR analysis in wild-type (dark-green bars) and in *ash1*^{22} homozygous (*ash1*^{22}, light-green bars) larvae, monitoring H3K36me2 levels at the *Ubx*, *mth*, *CG6310*, *wg*, *tsh*, *lam* and *dpr12* genes in T3 discs. At each gene, H3K36me2 was analysed at one or more regions, and for each qPCR, amplicon coordinates are indicated as distance in kb from the transcription start site, see also Table S5. For both genotypes, bars show ChIP signals from three independent ChIP reactions that were performed on three independently prepared batches of chromatin and are presented as a percentage of input chromatin precipitated at each region; dots show individual experimental results and error bars show standard deviation.

reported the purification of an identical Ash1-MRG15-Caf1 complex and they also found that MRG15 stimulates Ash1 enzymatic activity *in vitro* (Huang et al., 2017), providing independent support for these first two conclusions from our work. Third, our transcriptome analyses in developing *Drosophila* reveal that AMC is required for the normal expression of a few hundred genes in addition to the HOX genes. Fourth, we show that animals that lack Ash1 or MRG15, or contain a catalytically inactive version of Ash1 have the capacity to go through embryonic, larval and pupal development, and complete metamorphosis to differentiate into adults that show very specific homeotic phenotypes as their main morphological defect. Notably, this defines MRG15 as a novel trxG protein. Fifth, we find that AMC does not make a major contribution to the bulk of H3K36 di-methylation in *Drosophila* but that the complex is essential for generating wild-type levels of H3K36me2 at HOX and other target genes that are downregulated in *ash1* mutants. In the following, we shall focus on specific aspects of these findings.

Activation of Ash1 methyltransferase activity by MRG15

Our analyses and those of Huang et al. (2017) strongly suggest that AMC rather than the Ash1 protein alone is the active form of this HMTase and that the MRG15 subunit stimulates Ash1 HMTase activity via a mechanism that is conserved in flies and mammals. In particular, the data sets in Fig. 3 collectively suggest that interaction of the MRG domain with the FxLP motif preceding the Ash1/ASH1L SET domain increases the catalytic activity of this domain in both the fly and the human complex. Without structural information about this interaction, we can currently only speculate on the activation mechanism. An attractive possibility would be that MRG15 binding allosterically activates the SET domain by displacing the auto-inhibitory loop formed by the Ash1 post-SET domain (An et al., 2011) and thereby facilitates access of the H3K36 substrate lysine to the catalytic centre in the SET domain. In addition to this allosteric activation mechanism, MRG15 might also promote AMC activity through a second mechanism involving interaction of the MRG15 chromo barrel domain with nucleosomes that already carry the H3K36me2 and/or -me3 modification (Zhang et al., 2006). Specifically, in the chromatin of actively transcribed genes that contain H3K36me2/3-modified nucleosomes, interaction of MRG15 with such nucleosomes – perhaps in cooperation with interactions of the Ash1 bromodomain or PHD finger with other histone modifications – might permit AMC to di-methylate H3K36 in unmodified neighbouring nucleosomes more efficiently. This might be particularly crucial when histone H3 is exchanged in the wake of transcription or when newly synthesized octamers containing unmodified H3 are incorporated after DNA replication. In this context, it should also be recalled that the target genes *Ubx*, *mth* or *CG6310* still contain low levels of H3K36me2-modified nucleosomes in discs of *ash1*$^{22\ m+\ z-}$ mutant larvae. This residual H3K36me2 is unlikely to represent nucleosomes that were modified by maternally-deposited Ash1 protein. First, Ash1 protein is undetectable in these cells (Fig. S1) and, second, the replication-coupled dilution of parental nucleosomes in these dividing cells will require *de novo* methylation of newly incorporated H3 molecules at every S-phase. It therefore seems more likely that the low level of H3K36me2 at these genes in *ash1* mutants is generated by NSD and Set2. According to this view, Ash1 would thus act on top of an H3K36me2/3 landscape generated by these more globally acting H3K36-methylating enzymes. Furthermore, Ash1 association with polytene chromosomes was reported to depend on Kismet/CHD7, a nucleosome remodelling factor that is required for the transition from transcription initiation to transcriptional elongation (Srinivasan et al., 2008). Together, these observations all point to a scenario where AMC acts on chromatin that already is transcriptionally active and possibly already is at least partially decorated with H3K36me2/3.

Developmental and gene expression defects in mutants lacking AMC function

Our transcriptome analyses identified several hundred genes that are de-regulated in *ash1* mutants. This observation may seem surprising given that *ash1*-null and catalytic mutants nevertheless develop into pharate or even viable and fertile adults, respectively. The homeotic transformations in *ash1* mutants show that downregulation of HOX gene expression has a clear physiological consequence. We have not been able to detect any other obvious morphological defects in the epidermal structures of *ash1* mutants. However, it is possible that changes in the expression levels of AMC-regulated genes other than HOX genes cause morphological defects in internal structures or organs, or that they affect the physiology of mutant animals and could in this way impact on their survival and viability. Future studies will be needed to assess the consequences of altered expression levels of

Development (2018) 145, dev163808. doi:10.1242/dev.163808

the non-HOX genes regulated by AMC. Here, we shall focus our discussion on the role of AMC in regulating HOX genes where lack of the complex and the reduction of H3K36me2 have a clear physiological consequence.

Requirement for H3K36 dimethylation by AMC to counteract Polycomb repression

As discussed in the Introduction, genetic and molecular studies originally uncovered that Ash1 prevents H3K27 tri-methylation by PRC2 in the coding region of the *Ubx* gene in cells where this gene is expressed (Papp and Müller, 2006). Biochemical studies *in vitro* then established that PRC2 HMTase activity for H3K27 methylation is inhibited on nucleosomes carrying H3K36me2 (Schmitges et al., 2011; Yuan et al., 2011). Here, we now show that Ash1 is indeed required for deposition of normal levels of H3K36me2 in the *Ubx*-coding region in cells where *Ubx* is normally expressed (Fig. 6B). Together, this supports a model in which AMC di-methylates H3K36 in the *Ubx*-coding region and thereby antagonizes H3K27me3 deposition by PRC2 and the instalment of Polycomb repression at this gene. It remains to be investigated whether H3K36 di-methylation by AMC also antagonizes H3K27 tri-methylation at other Ash1-regulated genes. Inspection of the available H3K27me3 profiles in wild-type *Drosophila* embryonic, larval or adult cells, or in different tissue culture cell lines provides no evidence for presence of H3K27me3 at the *mth* or *CG6310* genes in any of these cells (www.modencode.org). At these genes, AMC might therefore preserve normal levels of expression through mechanisms other than counteracting Polycomb repression.

A conspicuous feature of mutants lacking AMC function is the all-or-none loss of HOX gene expression. Specifically, the patchy loss of *Ubx* and *Abd-B* expression and the patchy transformations in the adult epidermis of *ash1*-null, *ash1* catalytic inactive or *MRG15*-null mutants suggest that expression of these genes is lost in a stochastic fashion in a fraction of larval cells and, once lost, this 'OFF' state is then clonally propagated in their daughter cells (Fig. 4B). Conversely, in other cells, HOX gene expression appears to be maintained (Fig. 4B). How could this variegated loss of expression be explained? As discussed above, the low levels of H3K36me2 at *Ubx* in *ash1* mutants are likely generated by NSD and/or Set2. It is possible that, in the absence of AMC, H3K36 methylation by these other HMTases may suffice to sustain H3K36me2 levels on at least one of the two *Ubx* alleles on the two homologous chromosomes above the threshold level needed to antagonize PRC2. We imagine that this crucial threshold is, however, not reliably reached in all cells, and, as a consequence, gain of H3K27 tri-methylation by PRC2 on both *Ubx* alleles may result in a stable OFF switch that is then clonally propagated. According to this view, a key physiological role of the trxG protein complex AMC is to augment the number of H3K36me2-modified nucleosomes across the chromatin of active HOX genes to safeguard them from H3K27 tri-methylation by PRC2.

MATERIALS AND METHODS

Drosophila strains

Strains with the following genotypes were generated and/or used in this study:

w (wild type);
w; NTAP-Ash1; ash1^{22} FRT2A;
w; ash1-CTAP/CyO; ash1^{22} FRT2A;
w; ash1^{22} FRT2A/TM3 twi::EGFP;
w hsp70-flp; ovo^{D} FRT2A/TM2/TM6B;
w; ash1^{wt}(VK37); ash1^{22} FRT2A;
w; ash1^{R1464A} (VK37); ash1^{22} FRT2A/TM6B;
w; MRG15$^{\Delta}$/TM6C; and *w; Df(3R)BSC741/TM3 twi::EGFP*.

The genotypes of the animals shown in Fig. 4 were as follows: *ash1$^{22\ m-\ z-}$* animals were *w; ash1^{22} FRT2A* and derived from crossing *w; ash1^{22} FRT2A/ovo^{D} FRT2A* females (with germ line clones) with *w; ash1^{22} FRT2A/TM3 twi::EGFP* males; *ash1$^{R1464A\ m-\ z-}$* animals were *w; ash1^{R1464A} (VK37)/+; ash1^{22} FRT2A* and were derived from crossing *w; ash1^{R1464A} (VK37); ash1^{22} FRT2A* females with *w; ash1^{22} FRT2A/TM3 twi::EGFP* males; and *MRG15$^{\Delta\ m-\ z-}$* animals were *w; MRG15$^{\Delta}$/Df(3R)BSC741* and derived from crossing *w; MRG15$^{\Delta}$/Df(3R)BSC741* females with *w; Df(3R)BSC741/twi::EGFP* males.

Antibodies

Antibodies used in this study are listed in Table S4. Antibodies against Ash1$_{517\text{-}842}$ were raised in rabbits. The Ash1$_{517\text{-}842}$ epitope was expressed as a 6×His-tagged protein in *Escherichia coli* and purified under denaturing conditions; the same epitope was used for affinity purification of the antibody.

Tandem affinity purification (TAP) of Ash1 complexes

The *NTAP-Ash1* and *Ash1-CTAP* transgenes both contained the entire Ash1$_{1\text{-}2226}$ open reading frame in the previously described *Drosophila* transformation vectors CaSpeR-NTAP and CaSper-CTAP, respectively (Klymenko et al., 2006); plasmid maps are available on request. Tandem affinity purifications were performed from embryonic nuclear extracts of *w; NTAP-Ash1; ash1^{22} FRT2A* and *w; ash1-CTAP/CyO; ash1^{22} FRT2A* strains as described previously (Klymenko et al., 2006).

Mass spectrometric analysis of proteins isolated by TAP

Eluates from calmodulin beads after TAP were separated on 4-12% polyacrylamide gels. One part of the material was used for silver staining to visualize the proteins for illustration (Fig. 1A). For mass spectrometric analysis, the bulk part of the material was separated on the same type of gel, the entire gel lane was excised and subdivided into different slices that were then each alkylated and digested with trypsin as described (Barth et al., 2014). Peptides were collected by acid extraction, concentrated by evaporation and resuspended in 0.1% TFA. Fifty percent of the digested material was injected into an Ultimate 3000 HPLC system (Thermo-Fisher Scientific) and analysed as described previously (Barth et al., 2014). For protein identification, the raw data were analysed with the Andromeda algorithm of the MaxQuant protein analysis package (version 1.5.3.30) against the FlyBase dmel-all-translation-r5.32.fasta database, including reverse sequences and contaminants. For quantification, Intensity Based Absolute Quantification (iBAQ) values were calculated from peptide intensities. For presentation, values were log2-transformed and subsequently missing values were imputed from a random distribution centred at 1/3×log2 of the obtained experimental data (Table S1). Ash1, MRG15 and Caf1 peptides identified by mass spectrometry are listed in Table S2.

Expression and purification of recombinant proteins

Drosophila Ash1$_{C}$, and the complete coding sequences of MRG15, Msl3, Caf1 and Esc were cloned into pFastBac1 (ThermoFisher) with appropriate affinity-tag coding sequences at their N termini: StrepII-MRG15, StrepII-Msl3, StrepII-Caf1, StrepII-Esc, His6-Ash1$_{C}$, His6-Ash1$_{C}^{RxRP}$ and His6-Ash1$_{C}^{R1464A}$. Plasmids and viruses are available on request. Fig. 2B-D and Fig. 3 show Strep-Tactin affinity purifications. Insect cells were lysed in Strep-Buffer A [20 mM Tris-HCl (pH 8), 300 mM KCl, 2 mM $MgCl_2$, 15% glycerol, 10 µM $ZnSO_4$, 0.1% NP-40 substitute, 1 mM 1,4-dithiothreitol and protease inhibitors]. Cleared lysates were loaded on Strep-Tactin sepharose beads (IBA) and washed multiple times with Strep-Buffer A. Retained proteins were released in Strep-Buffer B [20 mM Tris-HCl (pH 8), 150 mM NaCl, 10% glycerol, 0.5 mM 1,4-dithiothreitol and protease inhibitors] supplemented with 5 mM d-desthiobiotin (Sigma). For affinity purification using the His-affinity tag in Fig. 2E, cells were lysed in His-buffer A [20 mM Tris-HCl (pH 8), 300 mM KCl, 4 mM $MgCl_2$, 5 mM imidazole, 5% glycerol, 10 µM $ZnSO_4$, 0.05% NP-40 substitute, 4 mM β-mercaptoethanol and protease inhibitors] and subjected to a Ni-affinity purification (Ni-NTA Agarose, Qiagen). After multiple washes with His-buffer A, proteins were eluted with His-Buffer B [20 mM Tris-HCl (pH 8), 300 mM NaCl, 250 mM imidazole, 5% glycerol, 0.05% NP-40 substitute, 4 mM β-mercaptoethanol and protease inhibitors].

For expression in *E. coli*, *Drosophila* Ash1$_{XL\text{-}SET}$ and human ASH1L$_{XL\text{-}SET}$ were cloned into a modified pET28a-TEV vector, containing an N-terminal

Development (2018) 145, dev163808. doi:10.1242/dev.163808

6×His-tag and a TEV protease site. *Drosophila* full-length MRG15 and human $MRG15_{MRG}$ domain were cloned into pET21a vector. $Ash1_{XL\text{-}SET}$ and $ASH1L_{XL\text{-}SET}$ proteins were expressed and $Ash1_{XL\text{-}SET}$:MRG15 and $ASH1L_{XL\text{-}SET}$:$MRG15_{MRG}$ complexes were co-expressed in *E. coli* BL21 (RILP) cells. The proteins were then purified on Ni-NTA (Qiagen) resin. Following treatment with TEV protease to remove the N-terminal 6×His-tag, the proteins were further purified by ion exchange, and size exclusion chromatography in a buffer containing 50 mM Tris HCl (pH 8.0) and 100 mM NaCl.

HMTase assays

Mononucleosome substrates were reconstituted with recombinant *Drosophila* histones and a 215 bp long 601 DNA fragment. HMTase reactions on mononucleosomes were performed in buffer containing 80 µM SAM, 65 mM Tris-HCl (pH 8.5), 78 mM NaCl, 2.5 mM $MgCl_2$, 0.23 mM EDTA (pH 8.0), 1 mM DTT, 1 mM β-mercaptoethanol, 2.6 mM d-desthiobiotin, 5% glycerol and protease inhibitors, and were incubated for 3 h at 25°C. For HMTase reactions on oligonucleosomes, nucleosomal arrays were assembled with recombinant *Xenopus* histones and G5E4 DNA. Reactions were performed in buffer containing 2 µM ^{3}H-SAM, 50 mM Tris-HCl (pH 9.0), 40 mM NaCl, 5 mM MgCl2 and 4 mM DTT, and were incubated for 80 min at 37°C. The reactions were analysed on a 15% SDS-polyacrylamide gel, visualized by Coomassie Blue staining or transferred to an immobilon-PSQ PVDF membrane (ISEQ00010, Millipore) and exposed to an image plate. The image plate was scanned with FLA-7000 (Fuji Film) for autoradiography.

Ash1 genomic transgenes

Transgenes containing genomic *ash1* fragments comprised BDGP R6.14 chr3L sequences 19,600,040…19,590,604, using BAC CH322-147P9 as template. In the *Taf6*-coding sequence present in this fragment, multiple ATG initiation codons were converted into stop codons. For the $ash1^{R1464A}$ transgene, the AGG codon for Arg_{1464} was mutated to GCG. The genomic fragments were cloned into a modified attB vector (pUMR-FLAP) and integrated at the attP site VK37 (BDSC 9752).

Generation of the *MRG15*$^{\Delta}$ deletion allele

MRG15$^{\Delta}$ was generated by replacing BDGP R6.14 chr3R: 15,276,676…15,277,889 with *miniwhite* marker gene using ends-out targeting with the pw35 vector (Gong and Golic, 2003). The 5′ homology arm (BDGP R6.14)chr3R:15,277,890…15,281,975) and the 3′ homology arm (BDGP R6.14 chr3R:15,272,003…15,276,675) shown in Fig. S2 were amplified from BAC CH322-160G6 and the initiation ATG of *MRG15* was mutated to ATC. The *MRG15*$^{\Delta}$ allele isolated and used throughout this study was selected among multiple independent targeting events after confirming that *MRG15* was disrupted (Fig. S2) and that sequences in the homology arms and flanking DNA were unaltered.

Immunostaining and adult cuticle preparations of *Drosophila*

Immunostaining of imaginal discs and preparations of adult cuticles were performed following standard protocols (Beuchle et al., 2001).

Western blot analyses on larval tissue extracts

Western blots were performed on total extracts prepared from pooled hand-dissected wing, haltere and 3rd leg imaginal discs of a given genotype (Copur and Müller, 2013).

Transcriptome analysis by NGS

RNA was isolated from independently prepared batches of hand-dissected haltere and third-leg imaginal discs (T3 discs; discs from eight larvae for each biological replicate) or wing discs (T2 discs; discs from eight larvae for each biological replicate) from wild-type or $ash1^{22}$ homozygous larvae, using the Direct-zol RNA mini Prep Kit (Zymo Research). After additional Agencourt AMPure XP purification (Beckman Coulter), isolated RNA was analysed on a Bioanalyzer (Agilent) using an RNA 6000 Nano Chip (Agilent). Non-degraded RNA from at least four independent biological replicates of each type of sample was used to construct sequencing libraries using sense mRNA-Seq Library Prep Kit (Lexogen). Quality-controlled and quantified libraries were sequenced on an HiSeq1500 system (Illumina) in the single-end mode (100 nt read length).

For analysis, trimmed and quality-filtered reads were mapped using the STAR aligner (Dobin et al., 2013) to the Ensembl genome annotation and *Drosophila* genome assembly dm3. Read counts were quantified using featurecounts (Liao et al., 2014) and differential gene expression calculated with limma using the voom transformation (Law et al., 2014).

ChIP analysis

Chromatin preparation from hand-dissected haltere and third-leg imaginal discs (T3 discs) from wild-type or $ash1^{22}$ homozygous larvae and ChIP analysis was performed as described (Laprell et al., 2017). For each biological replicate, chromatin prepared from the discs of 60 larvae were used as input material. ChIP was performed with polyclonal anti-H3K36me2 antibody (Abcam ab9049) and qPCR primers used for analysis are listed in Table S5.

Acknowledgements

We thank Bing Zhu for discussions.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: S.S., J.-J.S., J.M.; Methodology: S.S., A. Meiler, J.M.; Validation: S.S., A. Meiler, A.I., J.-J.S., J.M.; Formal analysis: S.S., A. Meiler, B.H., A.I.; Investigation: S.S., A. Meiler, Y.L., A. Mohammed, K.F., K.T., L.I., M.W., J.P.-M., J.-J.S.; Data curation: A. Meiler, J.P.-M.; Writing - original draft: S.S., J.M.; Writing - review & editing: S.S., J.M.; Visualization: S.S., A. Meiler, B.H., A.I., Y.L., J.-J.S., J.M.; Supervision: H.B., B.H., A.I., J.-J.S., J.M.; Project administration: A.I., J.-J.S., J.M.; Funding acquisition: A.I., J.-J.S., J.M.

Funding

Work in the lab of J.M. was supported by the Max-Planck-Gesellschaft and the Deutsche Forschungsgemeinschaft (SFB1064). Work in the lab of A.I. was supported by the Deutsche Forschungsgemeinschaft (SFB1064). Work in the lab of J.-J.S. was supported by the National Research Foundation of Korea (NRF) (NRF-2016R1A2B3006293 to J.-J.S.). Deposited in PMC for immediate release.

Data availability

The RNA-seq analysis data have been deposited in GEO under accession number GSE108120.

Supplementary information

Supplementary information available online at
http://dev.biologists.org/lookup/doi/10.1242/dev.163808.supplemental

References

An, S., Yeo, K. J., Jeon, Y. H. and Song, J.-J. (2011). Crystal structure of the human histone methyltransferase ASH1L catalytic domain and its implications for the regulatory mechanism. *J. Biol. Chem.* **286**, 8369-8374.

Anderson, A. E., Karandikar, U. C., Pepple, K. L., Chen, Z., Bergmann, A. and Mardon, G. (2011). The enhancer of trithorax and polycomb gene Caf1/p55 is essential for cell survival and patterning in Drosophila development. *Development* **138**, 1957-1966.

Barth, T. K., Schade, G. O. M., Schmidt, A., Vetter, I., Wirth, M., Heun, P., Thomae, A. W. and Imhof, A. (2014). Identification of novel Drosophila centromere-associated proteins. *Proteomics* **14**, 2167-2178.

Bell, O., Wirbelauer, C., Hild, M., Scharf, A. N. D., Schwaiger, M., MacAlpine, D. M., Zilbermann, F., van Leeuwen, F., Bell, S. P., Imhof, A. et al. (2007). Localized H3K36 methylation states define histone H4K16 acetylation during transcriptional elongation in Drosophila. *EMBO J.* **26**, 4974-4984.

Beuchle, D., Struhl, G. and Müller, J. (2001). Polycomb group proteins and heritable silencing of Drosophila Hox genes. *Development* **128**, 993-1004.

Bowman, S. K., Deaton, A. M., Domingues, H., Wang, P. I., Sadreyev, R. I., Kingston, R. E. and Bender, W. (2014). H3K27 modifications define segmental regulatory domains in the Drosophila bithorax complex. *Elife* **3**, e02833.

Copur, Ö. and Müller, J. (2013). The histone H3-K27 demethylase Utx regulates HOX gene expression in *Drosophila* in a temporally restricted manner. *Development* **140**, 3478-3485.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21.

Dorighi, K. M. and Tamkun, J. W. (2013). The trithorax group proteins Kismet and ASH1 promote H3K36 dimethylation to counteract Polycomb group repression in Drosophila. *Development* **140**, 4182-4192.
Eram, M. S., Kuznetsova, E., Li, F., Lima-Fernandes, E., Kennedy, S., Chau, I., Arrowsmith, C. H., Schapira, M. and Vedadi, M. (2015). Kinetic characterization of human histone H3 lysine 36 methyltransferases, ASH1L and SETD2. *Biochim. Biophys. Acta* **1850**, 1842-1848.
Gambetta, M. C., Oktaba, K. and Müller, J. (2009). Essential role of the glycosyltransferase Sxc/Ogt in Polycomb repression. *Science* **325**, 93-96.
Gaydos, L. J., Rechtsteiner, A., Egelhofer, T. A., Carroll, C. R. and Strome, S. (2012). Antagonism between MES-4 and Polycomb repressive complex 2 promotes appropriate gene expression in C. elegans germ cells. *Cell Rep.* **2**, 1169-1177.
Gong, W. J. and Golic, K. G. (2003). Ends-out, or replacement, gene targeting in Drosophila. *Proc. Natl. Acad. Sci. USA* **100**, 2556-2561.
Huang, C., Yang, F., Zhang, Z., Zhang, J., Cai, G., Li, L., Zheng, Y., Chen, S., Xi, R. and Zhu, B. (2017). Mrg15 stimulates Ash1 H3K36 methyltransferase activity and facilitates Ash1 Trithorax group protein function in Drosophila. *Nat. Commun.* **8**, 1649.
Kang, H., Jung, Y. L., McElroy, K. A., Zee, B. M., Wallace, H. A., Woolnough, J. L., Park, P. J. and Kuroda, M. I. (2017). Bivalent complexes of PRC1 with orthologs of BRD4 and MOZ/MORF target developmental genes in Drosophila. *Genes Dev.* **31**, 1988-2002.
Kassis, J. A., Kennison, J. A. and Tamkun, J. W. (2017). Polycomb and trithorax group genes in Drosophila. *Genetics* **206**, 1699-1725.
Klymenko, T. and Müller, J. (2004). The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO Rep.* **5**, 373-377.
Klymenko, T., Papp, B., Fischle, W., Köcher, T., Schelder, M., Fritsch, C., Wild, B., Wilm, M. and Müller, J. (2006). A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities. *Genes Dev.* **20**, 1110-1122.
Kockmann, T., Gerstung, M., Schlumpf, T., Xhinzhou, Z., Hess, D., Beerenwinkel, N., Beisel, C. and Paro, R. (2013). The BET protein FSH functionally interacts with ASH1 to orchestrate global gene activity in Drosophila. *Genome Biol.* **14**, R18.
Krogan, N. J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D. P., Beattie, B. K., Emili, A., Boone, C. et al. (2003). Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. *Mol. Cell. Biol.* **23**, 4207-4218.
Kusch, T., Florens, L., Macdonald, W. H., Swanson, S. K., Glaser, R. L., Yates, J. R., Abmayr, S. M., Washburn, M. P. and Workman, J. L. (2004). Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions. *Science* **306**, 2084-2087.
Kwong, C., Adryan, B., Bell, I., Meadows, L., Russell, S., Manak, J. R. and White, R. (2008). Stability and dynamics of polycomb target sites in Drosophila development. *PLoS Genet.* **4**, e1000178.
LaJeunesse, D. and Shearn, A. (1995). Trans-regulation of thoracic homeotic selector genes of the Antennapedia and bithorax complexes by the trithorax group genes: absent, small, and homeotic discs 1 and 2. *Mech. Dev.* **53**, 123-139.
Langlais, K. K., Brown, J. L. and Kassis, J. A. (2012). Polycomb group proteins bind an engrailed PRE in both the “ON” and “OFF” transcriptional states of engrailed. *PLoS ONE* **7**, e48765.
Laprell, F., Finkl, K. and Müller, J. (2017). Propagation of Polycomb-repressed chromatin requires sequence-specific recruitment to DNA. *Science* **356**, 85-88.
Larschan, E., Alekseyenko, A. A., Gortchakov, A. A., Peng, S., Li, B., Yang, P., Workman, J. L., Park, P. J. and Kuroda, M. I. (2007). MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism. *Mol. Cell* **28**, 121-133.
Law, C. W., Chen, Y., Shi, W. and Smyth, G. K. (2014). voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29.
Liao, Y., Smyth, G. K. and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930.
Miyazaki, H., Higashimoto, K., Yada, Y., Endo, T. A., Sharif, J., Komori, T., Matsuda, M., Koseki, Y., Nakayama, M., Soejima, H. et al. (2013). Ash1l methylates Lys36 of histone H3 independently of transcriptional elongation to counteract polycomb silencing. *PLoS Genet.* **9**, e1003897.
Papp, B. and Müller, J. (2006). Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. *Genes Dev.* **20**, 2041-2054.
Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Lee, T. I., Bell, G. W., Walker, K., Rolfe, P. A., Herbolsheimer, E. et al. (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122**, 517-527.
Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M. and Séraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* **17**, 1030-1032.
Schmitges, F. W., Prusty, A. B., Faty, M., Stützer, A., Lingaraju, G. M., Aiwazian, J., Sack, R., Hess, D., Li, L., Zhou, S. et al. (2011). Histone methylation by PRC2 is inhibited by active chromatin marks. *Mol. Cell* **42**, 330-341.
Schwartz, Y. B., Kahn, T. G., Stenberg, P., Ohno, K., Bourgon, R. and Pirrotta, V. (2010). Alternative epigenetic chromatin states of polycomb target genes. *PLoS Genet.* **6**, e1000805.
Shearn, A. (1989). The ash-1, ash-2 and trithorax genes of Drosophila melanogaster are functionally related. *Genetics* **121**, 517-525.
Shearn, A., Hersperger, E. and Hersperger, G. (1987). Genetic studies of mutations at two loci of Drosophila melanogaster which cause a wide variety of homeotic transformations. *Roux’s Arch. Dev. Biol.* **196**, 231-242.
Srinivasan, S., Dorighi, K. M. and Tamkun, J. W. (2008). Drosophila Kismet regulates histone H3 lysine 27 methylation and early elongation by RNA polymerase II. *PLoS Genet.* **4**, e1000217.
Sun, B., Hong, J., Zhang, P., Dong, X., Shen, X., Lin, D. and Ding, J. (2008). Molecular basis of the interaction of Saccharomyces cerevisiae Eaf3 chromo domain with methylated H3K36. *J. Biol. Chem.* **283**, 36504-36512.
Tanaka, Y., Katagiri, Z.-I., Kawahashi, K., Kioussis, D. and Kitajima, S. (2007). Trithorax-group protein ASH1 methylates histone H3 lysine 36. *Gene* **397**, 161-168.
Tripoulas, N., LaJeunesse, D., Gildea, J. and Shearn, A. (1996). The Drosophila ash1 gene product, which is localized at specific sites on polytene chromosomes, contains a SET domain and a PHD finger. *Genetics* **143**, 913-928.
Venkatesh, S. and Workman, J. L. (2013). Set2 mediated H3 lysine 36 methylation: regulation of transcription elongation and implications in organismal development. *Wiley Interdiscip. Rev. Dev. Biol.* **2**, 685-700.
Xiao, T., Hall, H., Kizer, K. O., Shibata, Y., Hall, M. C., Borchers, C. H. and Strahl, B. D. (2003). Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. *Genes Dev.* **17**, 654-663.
Xie, T., Graveline, R., Kumar, G. S., Zhang, Y., Krishnan, A., David, G. and Radhakrishnan, I. (2012). Structural basis for molecular interactions involving MRG domains: implications in chromatin biology. *Structure* **20**, 151-160.
Xie, T., Zmyslowski, A. M., Zhang, Y. and Radhakrishnan, I. (2015). Structural Basis for Multi-specificity of MRG Domains. *Structure* **23**, 1049-1057.
Xu, C., Cui, G., Botuyan, M. V. and Mer, G. (2008). Structural basis for the recognition of methylated histone H3K36 by the Eaf3 subunit of histone deacetylase complex Rpd3S. *Structure* **16**, 1740-1750.
Yuan, W., Xu, M., Huang, C., Liu, N., Chen, S. and Zhu, B. (2011). H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J. Biol. Chem.* **286**, 7983-7989.
Zhang, P., Du, J., Sun, B., Dong, X., Xu, G., Zhou, J., Huang, Q., Liu, Q., Hao, Q. and Ding, J. (2006). Structure of human MRG15 chromo domain and its binding to Lys36-methylated histone H3. *Nucleic Acids Res.* **34**, 6621-6628.