Genetic Dissection Reveals the Role of Ash1 Domains in Counteracting Polycomb Repression

Eshagh Dorafshan, Tatyana G. Kahn, Alexander Glotov, Mikhail Savitsky,1 and Yuri B. Schwartz2
Department of Molecular Biology, Umeå University, 90187 Umeå, Sweden
ORCID ID: 0000-0003-4790-3920 (Y.B.S.)

ABSTRACT Antagonistic functions of Polycomb and Trithorax proteins are essential for proper development of all metazoans. While the Polycomb proteins maintain the repressed state of many key developmental genes, the Trithorax proteins ensure that these genes stay active in cells where they have to be expressed. Ash1 is the Trithorax protein that was proposed to counteract Polycomb repression by methylating lysine 36 of histone H3. However, it was recently shown that genetic replacement of Drosophila histone H3 with the variant that carried Arginine instead of Lysine at position 36 did not impair the ability of Ash1 to counteract Polycomb repression. This argues that Ash1 counteracts Polycomb repression by methylating yet unknown substrate(s) and that it is time to look beyond Ash1 methyltransferase SET domain, at other evolutionary conserved parts of the protein that received little attention. Here we used Drosophila genetics to demonstrate that Ash1 requires each of the BAH, PHD and SET domains to counteract Polycomb repression, while AT hooks are dispensable. Our findings argue that, in vivo, Ash1 acts as a multimer. Thereby it can combine the input of the SET domain and PHD-BAH cassette residing in different peptides. Finally, using new loss of function alleles, we show that zygotic Ash1 is required to prevent erroneous repression of homeotic genes of the bithorax complex in the embryo.

Embryonic development is controlled by genes encoding morphogens and transcriptional regulators. These genes need to be switched on in correct cells at appropriate time and their expression, or lack thereof, has to be maintained as the embryo grows and cells continue to divide. Polycomb and Trithorax systems are critical to maintain the expression status of developmental genes (Schwartz and Pirrotta 2007; Schuettengruber et al. 2017). Protein components of the Polycomb system assemble in complexes that act as epigenetic repressors. While mechanistic details of the repression are still being worked out, studies in fruit flies indicate that tri-methylation of Lysine 27 of histone H3 (H3K27) by one of the Polycomb complexes (PRC2) is essential for the process (Pengelly et al. 2013; McKay et al. 2015; Coleman and Struhl 2017; Laprell et al. 2017).

The Trithorax system counteracts Polycomb repression to ensure that developmental genes repressed by Polycomb complexes in certain cell types are not erroneously shut down in cells where they have to remain active (Poux et al. 2002; Klymenko and Muller 2004). Genetic evidence from studies on Drosophila melanogaster indicate that the Trithorax (Trx) and Absent, small, or homeotic discs 1 (Ash1) proteins are critical components of the system (Poux et al. 2002; Klymenko and Muller 2004). Of the two proteins, Ash1 is particularly interesting. Unlike Trx, which binds Polycomb regulated genes regardless of their transcriptional state (Schuettengruber et al. 2009; Schwartz et al. 2010), Ash1 binds and forms extensive chromatin domains exclusively when Polycomb regulated genes are transcriptionally active and Polycomb repression is impaired (Schwartz et al. 2010; Kharchenko et al. 2011; Huang et al. 2017). Trx and Ash1 are incorporated in distinct complexes whose composition is yet to be fully characterized (Petruk et al. 2001; Mohan et al. 2011; Huang et al. 2017; Schmahling et al. 2018) but both have SET domains and can methylate histone H3 (Smith et al. 2004; Tanaka et al. 2007; An et al. 2011; Yuan et al. 2011; Tie et al. 2014). In vitro experiments indicate that the histone methyltransferase activity of PRC2 is inhibited by prior methylation of histone H3 tail at Lysine 4 (H3K4) or Lysine 36 (H3K36) (Schmitges et al. 2011; Yuan et al. 2011).
et al. 2011; Voigt et al. 2012), the sites that can be methylated by Trx and Ash1. From this, it was proposed that Trx and Ash1 counteract Polycomb repression by inhibiting PCR2 catalytic activity via H3K4 and H3K36 methylation (Schmitges et al. 2011; Yuan et al. 2011; Voigt et al. 2012). Unsurprisingly, most recent studies of Trx and Ash1 have focused on mechanistic details of histone H3 methylation by their SET domains (Dorighi and Tamkun 2013; Tie et al. 2014; Huang et al. 2017; Schmahling et al. 2018; Dorafshan et al. 2019; Hou et al. 2019; Lee et al. 2019).

Although the “PC2-inhibition” model explains the antagonistic relation between Polycomb and Trithorax systems, it is at odds with results of in vivo studies. Thus, experiments of Hödl and Basler (Hödl and Basler 2012), as well as our recent work (Dorafshan et al. 2019), showed that genetic replacement of Drosophila histone H3 with variants that carry Arginine instead of Lysine at positions 4 or 36 does not impair the ability of Trx and Ash1 to counteract Polycomb repression. Since intact SET domains of both proteins are required to antagonize the repression (Stassen et al. 1995; Dorafshan et al. 2019), this argues that Trx and Ash1 counteract Polycomb system by methylating, yet unknown, non-histone proteins. To find these substrates, it is time to characterize other evolutionary conserved domains of these proteins which, so far, received little attention. Are these domains critical for Trx and Ash1 function? Do they act in concert or as separate modules? Which of them, if any, are critical to bind chromatin? Here we used the power of Drosophila genetics to address these questions for the protein domains of Ash1. Our experiments demonstrate that Ash1 requires each of the BAH, PHD and SET domains to counteract Polycomb repression, while AT hooks are dispensable. Our complementation analyses suggest that, in vivo, Ash1 acts as a multimer and can combine the input of SET and PHD or BAH domains residing in different monomers. Finally, using new loss of function ash1 alleles, we demonstrate that its zygotic function is required to prevent erroneous repression of homeotic genes of the bithorax complex in the embryo.

MATERIALS AND METHODS

Plasmid construction

Ash1 expressing constructs were assembled by Gateway LR recombination (Invitrogen) between an entry construct containing corresponding OneStRep-tagged CDS and the destination vector pWattB-Ubi-DEST. To generate entry constructs with truncated Ash1 CDS, corresponding deletions were introduced into the entry construct containing full-length Ash1 CDS (pENTR1A-OneStRep-Ash1FL) as described below. The pENTR1A-OneStRep-Ash1FL and pWattB-Ubi-DEST constructs are described in (Dorafshan et al. 2019).

To generate pENTR1A-OneStRep-Ash1ΔSET construct, deltaSET_AB and deltaSET_CD fragments, flanking the SET domain, were amplified from pENTR1A-OneStRep-Ash1FL using primers deltaSET_A, deltaSET_B, and deltaSET_C, deltaSET_D. For corresponding oligonucleotide sequences see Table S1. pENTR1A-OneStRep-Ash1FL was digested with BstZ17I and SphI and deltaSET_AB and deltaSET_CD fragments were introduced to the linear plasmid using InFusion. pENTR1A-OneStRep-Ash1ΔPHD was sequenced using ASH1_seq8, ASH1_seq9, and ASH1_seq12 primers.

To generate pENTR1A-OneStRep-Ash1ΔBAH construct, two fragments (upstream deltaBAH_AB and downstream deltaBAH_CD) flanking BAH domain were amplified from pENTR1A-OneStRep-Ash1FL using primers deltaBAH_A, deltaBAH_B, and deltaBAH_C, deltaBAH_D. pENTR1A-OneStRep-Ash1FL was digested with ClaI and XhoI restriction enzymes. Linear vector (8.7 kb) was extracted from the gel and used in the InFusion reaction together with fragments deltaBAH_AB and deltaBAH_CD. The result pENTR1A-OneStRep-Ash1ΔBAH construct was sequenced using ASH1_seq13 and ASH1CN_Cpwd primers.

To obtain pENTR1A-OneStRep-Ash1ΔAT construct, all three AT hooks were deleted sequentially. Two fragments (deltaAT_E upstream and deltaAT_GH downstream) flanking the second AT hook were amplified from pENTR1A-OneStRep-Ash1FL using primers deltaAT_E, deltaAT_F, and deltaAT_G, deltaAT_H. pENTR1A-OneStRep-Ash1FL was digested with SacI and BstZ17I restriction enzymes. Linear vector (9 kb) was extracted from the gel and used in the InFusion reaction together with fragments deltaAT_EF and deltaAT_GH to yield the pENTR1A-OneStRep-Ash1ΔAT2 construct. Next, two fragments (deltaAT_EI upstream and deltaAT_HJ downstream) flanking the third AT hook were amplified from pENTR1A-OneStRep-Ash1ΔAT2 using primers deltaAT_E, deltaAT_I, deltaAT_H, deltaAT_J. pENTR1A-OneStRep-Ash1ΔAT2 was digested with SacI and BstZ17I restriction enzymes. Linear vector (9 kb) was extracted from the gel and used in the InFusion reaction together with fragments deltaAT_EI and deltaAT_HJ to obtain pENTR1A-OneStRep-Ash1ΔAT23. Finally, two fragments (deltaAT_AB upstream and deltaAT_CD downstream) flanking the first AT hook were amplified from pENTR1A-OneStRep-Ash1ΔAT23 using primers deltaAT_A, deltaAT_B, and deltaAT_C, deltaAT_D. pENTR1A-OneStRep-Ash1ΔAT23 was digested with SacI. Linear vector (9 kb) was extracted from the gel and used in the InFusion reaction together with fragments deltaAT_AB and deltaAT_CD to get the final pENTR1A-OneStRep-Ash1ΔAT construct. The pENTR1A-OneStRep-Ash1ΔAT2 and pENTR1A-OneStRep-Ash1ΔAT23 plasmids were sequenced using ASH1_seq4, ASH1_seq5, ASH1_seq6, and ASH1_seq7 primers. The pENTR1A-OneStRep-Ash1ΔAT was further sequenced using ASH1_seq3 and ASH1_seq14 primers.

To generate pWattB-U6.2-2targets plasmid, the target sequences for CRISPR/Cas9-mediated cleavage (SET-CRSP-1 and poSET-CRSP-1) were selected using Cas9 Target Finder software (http://www.shigen.nig.ac.jp/lyly/nigfly/cas9/index.jsp). SET-CRSP-1 is located within ash1 postSET domain 243bp apart. Corresponding pairs of complementary oligonucleotides (SET-CRSP-1.1 and SET-CRSP-1.2) and (poSET-CRSP-1.1 and poSET-CRSP-1.2) were ordered from Sigma-Aldrich and annealed to obtain double stranded SET-CRSP-1 and poSET-CRSP-1 fragments. The DNA of the pBF-U6.2B and pBF-U6.2 targets (kind gifts from Dr. Maria Kim) was digested with BbsI restriction enzyme and the linear products ligated with SET-CRSP-1 and poSET-CRSP-1 fragments resulting in pBF-U6.2B-SET and pBF-U6.2-poSET constructs, respectively. pBF-U6.2-poSET construct was digested with EcoRI and NotI and the resulting 517bp fragment (U6.2-poSET) isolated and ligated with the pBF-U6.2B-SET construct, digested with the same enzymes. This yielded the pBF-U6.2-2targets constructs. Following this, the DNA of the pBF-U6.2B-2targets construct was digested with SacI. The resulting 1.2kb SacI fragment (containing U6-2B_2targets) was

3802 | E. Dorafshan et al.
blunt-end ligated with DNA of the pWattB plasmid (Savitsky et al. 2016) digested with XhoI, which yielded pWattB-U6.2-2targets construct.

**Fly strains**

w1; ash1Δ/+, P[w+Δm11]; FRT[w+]2A/TM6C,Stbl, Tb1 (Bloomington Drosophila Stock Center, 24161), w1; U6.2-2targets (Bloomington Drosophila Stock Center, 7945), and w1; ash1Δ/+, P[w+Δm11]; FRT[w+]2A/TM6C, Stbl, Tb1 flies (Bloomington Drosophila Stock Center, 7945), and w1; ash1Δ/+, P[w+Δm11]; FRT[w+]2A/TM6C, Stbl, Tb1 (Steffen et al. 2013) were re-balanced over TM3, Ser, or Act-GFP+mw balancer. Oregon R flies (kind gift from Dr. Jan Larsson) were used as wild-type in all experiments unless stated otherwise. y2,cho2,v1; attP40[nos-Cas9]/CyO fly strain (Kondo and Ueda 2013) was used as a source of Cas9. w1; If/CyO; MKRS/TM6, Tb was used to balance CRISPR/Cas9-edited chromosomes.

**Fly transformation**

All constructs were injected in pre-blastoderm embryos by BestGene Inc. The Ash1 expressing (full-length and truncated) transgenes were injected in the y1; M[vas-int.Dm]ZH-2A w1; M[3xP3-RFP(attP)ZH-51C] (24482) strain. The pWattB-U6.2-2targets construct was injected in the y1; w67c23, P[CaryP]attP40 strain.

**CRISPR/Cas9-mediated genome editing**

The editing strategy described in (Kondo and Ueda 2013) was used to generate deletions within Ash1 SET domain. Flies of the y1; w67c23, pWattB-U6.2-2targets strain, expressing two gRNA, were crossed to y2,cho2,v1; attP40[nos-Cas9]/CyO; y1 w67c23; If/CyO; MKRS/TM6, Tb and the progeny individually screened for editing events by PCR with ash1_seq8 and deltaSET_D primers which amplify 499bp fragment from unedited chromosomes and 262bp fragment from chromosomes with precise deletion.

**Polytene chromosome preparation and immunostaining**

Salivary glands were dissected from 3rd instar larvae, preparation and immunostaining of polytene chromosomes was performed as described elsewhere (Eggett et al. 2004). Images were taken with Zeiss Axioskop microscope equipped with Plan-Apochromat 63x/1.40 oil DIC M27 objective, filters set (63HE for red channel, 38HE for green channel and 49 for DAPI) and AxioCam MR R3 camera. Images were processed with ZenPro software (v2.3, Zeiss) and mounted in ImageJ (v1.42h). For the list of antibodies see Table S2.

**Fly cuticle preparation and embryo immunostaining**

Flies were boiled in 10% KOH for 10 min, incubated in distilled water for 30 min, and dehydrated in 70% ethanol and 99% ethanol for 10 min each. Ethanol was removed and replaced with glycerol. Cuticles were incubated for 30 min in glycerol, then dissected under the stereo microscope and mounted on glass slide in glycerol. Embryos were immunostained as described in (Dorafshan et al. 2019). For the list of antibodies see Table S2.

**Chromatin immunoprecipitation (ChIP)**

ChIP and qPCR analysis were performed as described (Kahn et al. 2016) except that crosslinked material was sonicated in 4ml of 10mM Tris-HCl pH 8.0, 1mM EDTA pH 8.0 for 45 min with Branson 450 Digital Sonifier (45 cycles of 20 sec ON – 40 sec OFF). The isolated ChIP material was re-suspended in 400 μl of DNase free water and 4 μl were used for each quantitative PCR reaction. The antibodies used are listed in Table S2 and the ChIP amplicons are listed in Table S3.

**Western blot**

Nuclear extracts from whole 3rd instar larvae were separated on a 6% SDS-PAGE and blotted to PVDF membrane for 3 hr at 200mA. The same extracts were separated on a 15% SDS-PAGE and stained with Coomassie to be used as loading control. Primary and secondary antibodies were diluted in 1xPBS with 1% BSA and 0.05% Tween-20. For the list of antibodies, see Table S2.

**Data availability**

Strains and plasmids are available upon request. The authors affirm that all data necessary for confirming the conclusions of the article

Figure 1  Generation of Ash1 transgenes. (A) Transgenic Ash1 variants. Four different mutant variants were generated starting from the full-length Ash1 cDNA (Ash1FL). SET (dark green box), PHD (yellow box) and BAH (gray box) domains were deleted in Ash1ΔSET, Ash1ΔPHD and Ash1ΔBAH variants, respectively. In the Ash1ΔAT variant all three AT hooks (orange boxes) were deleted. Brackets show corresponding deleted regions. (B) All Ash1 Open Reading Frames from above were cloned in a vector containing Ubi-p63E promoter, an attB integration site, and a mini-white gene as a selection marker. Each construct was integrated in the same attP landing site (ZH-51C), using phiC31-mediated recombination.
are present within the article, figures, and tables. Supplemental material available at FigShare: https://doi.org/10.25387/g3.9884120.

RESULTS

To investigate how individual domains enable Ash1 to counteract Polycomb repression, we made an array of transgenic constructs (Figure 1). Each construct contained either the full length Ash1 Open Reading Frame (ORF) or truncated Ash1 ORFs lacking specific domains and attB site for phiC31-mediated site-specific recombination (Bischof et al. 2007). All constructs were driven by Ubiquitin (Ubi-p63E) promoter (Butcher et al. 2004) and integrated in the same 51C genomic site on chromosome 2L.

To test how well various transgenic Ash1 variants counteract Polycomb repression, we decided to use segment-specific expression of homeotic genes Ubx and Abd-B as a readout. The two genes are part of the bithorax complex gene cluster and specify the identity of the third thoracic and the first abdominal segments and the last four abdominal segments, respectively (Lewis 1978). Ubx and Abd-B are classic targets of Polycomb/Trithorax regulation and alteration in their expression patterns leads to morphological abnormalities that are easy to detect.

When ash1 function is compromised, the expression of Ubx and Abd-B is stochastically lost which, in turn, leads to partial transformation of corresponding segments toward the anterior fate (Shearn 1989; Klymenko and Muller 2004; Huang et al. 2017; Schmahling et al. 2018). Of the published ash1 alleles, ash122 and ash1Df(3L)Exel9011 (hereafter referred to ash19011) are the most severe. ash122 is a point mutation that converts Glutamin 129 into an early stop codon and has been reported as a null allele (Tripoulas et al. 1996). The ash19011 allele is the 53kb deletion that spans the entire ash1 gene region as well as 13 other genes (Parks et al. 2004). ash122/ash19011 animals die at early pupal stage with less than 10% of the animals producing enough adult cuticle to score homeotic transformations (Dorafshan et al. 2019). The hypomorphic ash121 allele is a substitution of Glutamic acid 1365 to Lysine within the Associated With SET (AWS) domain (Tripoulas et al. 1994). ash122/ash121 animals develop to pharate adult stage and about 12% survive as adults (Dorafshan et al. 2019). All mutant adult flies show haltere to wing and third leg toward second leg transformations, reflecting partial loss of the Ubx gene expression (Figure S1A). They also show transformations of the 5th and 6th abdominal segments toward anterior fate caused by partial loss of the Abd-B gene expression (Figure S1B).

Figure 2 Complementation of ash122/ash19011 mutations with transgenic Ash1 variants. (A) Reciprocal crossing schemes to combine different Ash1 transgenes with the ash122/ash19011 mutations. (B) The above crosses are expected to yield four phenotypically distinct progeny classes. Class I and Class II flies contain two copies of the Ash1 transgene while Class III and Class IV flies bear the marker Cy mutation and only one copy of the transgene. Class I and III flies have ash122/ash19011 background while Class II and IV flies are heterozygous for only one of the mutant ash1 alleles and carry Ser and GFP markers. (C) Histograms show the viability of ash122/ash19011 mutant flies supplemented with different Ash1 transgenes. The viability is plotted as percent ratios between the number of Class I and Class II progeny multiplied by two (for two copies of transgenic constructs) or as percent ratios between the number of Class III and Class IV progeny multiplied by two (for one copy of transgenic constructs). The ability of different Ash1 mutant transgenes to restore the viability of ash122/ash19011 flies was compared to that of the Ash1FL transgene and the significance of the deviation evaluated with chi-square test (df = 1). n = total number progeny counted for each cross.
To assess whether our transgenic system is capable to supply enough Ash1 protein, we introduced the transgene expressing full-length Ash1 (Ubi-Ash1FL) into the ash122/ash19011 mutant background. Two copies of the Ubi-Ash1FL transgene fully restored the viability of ash122/ash19011 flies (Figure S2), which showed no homeotic transformations, were fertile and could be maintained as a stock.

**Ash1 requires BAH, PHD and SET domains to counteract Polycomb repression**

To test whether Ash1 variants lacking specific protein domains are able to counteract Polycomb repression, we crossed flies bearing one copy of a transgenic Ash1 variant on the second chromosome and the ash19011 allele on the third chromosome with flies containing one copy of the same transgene Ash1 variant on the second chromosome and the ash122 allele on the third chromosome (Figure 2A). To control for maternal effects, each cross was repeated in reciprocal setup. Four distinct classes are expected in the progeny of such cross (Figure 2B). If a transgenic Ash1 fully compensates for the loss of endogenous Ash1 protein, in the progeny, the number of trans-heterozygous ash122/ash19011 animals, supplemented with one or two copies of the Ash1 transgene should be one half of the number of flies heterozygous for ash1 mutation (Figure 2B). We also expect transgenic ash122/ash19011 animals to have no homeotic transformations.

Similar to Ubi-Ash1FL, one or two copies of the Ash1 transgene lacking AT-hooks (Ubi-Ash1ΔAT) restored the viability and proper expression pattern of the homeotic genes (Figures 2C, 3A-B). In contrast, the transgenes lacking the BAH, PHD, or SET domains (Ubi-Ash1ΔBAH, Ubi-Ash1ΔPHD and Ubi-Ash1ΔSET) failed to complement ash122/ash19011 mutations. Although two copies of the transgenes could partially restore the viability of the ash122/ash19011 flies, one copy was not sufficient (Figure 2C). More important, all ash122/ash19011 flies “rescued” with two copies of the transgenes displayed obvious homeotic transformations (Figure 3A-B). Western-blot analysis showed that the Ubi-Ash1ΔBAH, Ubi-Ash1ΔPHD and Ubi-Ash1ΔSET transgenes produce at least as much protein as the transgene expressing the full-length Ash1 variant (Figure 3C-D). This rules out the trivial possibility that the Ash1 variants lacking BAH, PHD, or SET domains are less stable. Overall, we conclude that Ash1 requires BAH, PHD and SET domains, but not AT-hooks, to counteract Polycomb repression of the Ubx and AbdB genes.

**New alleles suggest that ash122 is a hypomorph mutation**

The observation that two copies of Ubi-Ash1ΔBAH, Ubi-Ash1ΔPHD and Ubi-Ash1ΔSET transgenes can restore the viability of ash122/ash19011 flies was unexpected. Ash1 protein or mRNA deposited in the embryo by the heterozygous mother may be sufficient for the mutant embryos to develop to adult stage. However, the ash122/ash19011 flies supplemented with two copies of the Ubi-Ash1ΔSET or Ubi-Ash1ΔPHD transgenes are fertile and, although too weak to establish the permanent stock, can be interbred and propagated for up to three generations. This argues that maternal contribution is not the reason for their viability. Alternatively, Ash1 may have two distinct functions, one necessary to counteract Polycomb repression, another required for viability. Perhaps, the methyltransferase activity of the SET domain or the functions of the BAH and PHD domains are needed for the former but are dispensable for the viability. Finally, ash122 mutation may retain some wild-type function. This mutation introduces a premature stop codon at position 129 of the open reading frame, which truncates it to encode for a short polypeptide that lacks all conserved domains (Tripoulas et al. 1996). However, recent evidence indicates that, in many *Drosophila* mRNAs, ribosomes can read through single stop codon and produce small but functional amounts...
of polypeptides that incorporate amino acids encoded downstream of supposed translation termination sites. (Steneberg and Samakovlis 2001; Dunn et al. 2013). Therefore, the premature stop codon of the ash122 allele may be read through yielding small amount of the full-length protein, whose function may be boosted by large amounts of truncated Ash1 proteins.

To discriminate between the two possibilities, we generated new ash1 alleles using CRISPR/Cas9 system (Gratz et al. 2014). To this effect, we designed guide RNAs to target Cas9 endonuclease to a site within the ash1 ORF that encodes the SET domain and to another site just downstream of the post-SET domain (Figure 4A). When cut at the two designated sites and repaired by the non-homologous end joining, break at positions corresponding to amino acid 1459 (Met) and 1539 (Leu) within the ash1 ORF. In the screen for the deletion, two alleles were recovered. In the ash117F allele 79 amino acid are deleted and 6 amino acids around the deletion breakpoint are changed (marked with red rectangles). In the ash13M allele, 81 amino acid are deleted and a frameshift before the deletion breakpoint changed 4 amino acids (red rectangles) and introduced 4 stop codons (gray rectangles). The frameshift continues after the deletion breakpoint.

In the schematic of the corresponding DNA sequences, the nucleotides highlighted in blue indicate positions of the guide RNAs and the nucleotides highlighted in orange represent insertions and deletions (indels). (B) Crossing scheme to test the complementation of ash13M/ash19011 mutations with ash1 transgenes. Female flies homozygous for an ash1 transgene on the second chromosome and heterozygous for ash19011 allele on the third chromosome are crossed to males with the same second chromosome but heterozygous for ash13M allele on the third chromosome. In the progeny, two classes are expected. While both classes are homozygous for an ash1 transgene, Class I (ash1-) flies are trans-heterozygous for ash13M and ash19011 alleles while Class II (ash1+) contains a wild-type ash1 allele present on the balancer TM3, Ser, Act-GFP chromosome. The cross is expected to yield twice as many Class II flies. (C) The ability of truncated Ash1 proteins to restore the viability of ash19011/ash13M flies was compared to that of the full-length transgenic Ash1 (Ash1FL). While the Ash1ΔAT transgene fully restored the viability, neither Ash1ΔBAH nor Ash1ΔSET transgenes yielded any viable flies and the Ash1ΔPHD transgene complemented ash1 loss of function inefficiently. n = the total number of progeny counted, P = probability that the observed difference in the survival rate of the wild-type (Ash1FL) and truncated ash1 transgenes is caused by chance (as evaluated by chi-square test, df = 1) and % viability indicates the ClassI/ClassII ratio multiplied by two. (D) Hometric phenotypes of the adult ash1122/ash1121 (control) and Ash1ΔPHD; ash19011/ash13M flies. For representative images of strong and mild abdominal transformations see Figure S3. Y-axes display the percent of flies with corresponding transformation. n = the total number of flies scored.
the edited DNA would carry an ORF encoding for the catalytically inactive protein that lacks half of the SET domain and the entire post-SET domain (Figure 4A). Screening for such deletion, we recovered two new alleles. One corresponded to the desired “in-frame” deletion of 79 amino acids, which we called ash1M (Figure 4A). Another deletion, dubbed ash1M, was larger (81 amino acids) and caused translation frame shift that introduced multiple successive stop codons in the middle of the SET domain (Figure 4A). The only possible product of the ash1M allele is a truncated Ash1 protein that lacks all but AT-hook domains. Importantly, the heterozygous ash1M/+ flies are viable, fertile and show no homeotic transformations, which argues that ash1M is not a dominant negative mutation.

Unlike ash122/ash19011 flies supplemented with two copies of the UbI-Ash1ΔSET transgene, homozygous ash122 or trans-heterozygous ash122/ash19011 flies die as third instar larvae or at early pupal stage. This indicates that, when the possible source of the wild-type protein is excluded, the Ash1 SET domain is required for viability. Attempts to complement the trans-heterozygous ash19011/ash1M mutations (likely zygotic null) with two copies of various Ash1 transgenes further corroborate this notion (Figures 1A, 4B). In contrast to the ash122/ash19011 complementation results (Figure 2C), the UbI-Ash1ΔBAH and UbI-Ash1ΔSET transgenes failed to restore the viability of the ash19011/ash1M mutants although the UbI-Ash1FL and UbI-Ash1ΔAT transgenes were still able to do so (Figure 4C). These observations argue that the ash1M allele is more severe than ash122 and that the BAH, or SET domain-active protein is modular and acts as a multimer.

Interallelic complementation indicates that Ash1 protein is modular and acts as a multimer

BAH, PHD and SET domains are all required for Ash1 to counteract Polycomb repression. While SET domain likely functions by methylating lysines (Tanaka et al. 2007; Huang et al. 2017; Schmahling et al. 2018) of yet unknown substrate(s) (Dorafshan et al. 2019), the roles of PHD and BAH domains are less clear. Genome-wide mapping in cultured Drosophila cells indicates that Ash1 binds weakly to multiple genomic sites, often located within long 5’ introns of transcriptionally active genes (Kharchenko et al. 2011; Huang et al. 2017). It also binds strongly to a couple dozen regions, many of which correspond to known Polycomb regulated genes caught in

Figure 5 Stochastic loss of Abd-B expression in the central nervous system of the ash1 mutant embryos. Immunostaining of ash1 mutant embryos with antibodies specific to Abd-B (red) shows reduced Abd-B expression in both homozygous ash1M (A) and trans-heterozygous ash19011/ash122 (B) mutant embryos. Heterozygous embryos, where ash1 mutant allele is combined with the TM3,SB,e,Kr::GFP balancer, serve as a control. In the control embryos Abd-B is expressed in parasegments 14-10 (marked with white brackets) in a gradient that slopes toward anterior pole. The loss of Abd-B signal, although seen in both mutants, is more pronounced in homozygous ash1M embryos compared to trans-heterozygous ash19011/ash122 counterparts. In both (A) and (B), the control and mutant embryos were stained together and separated by strong GFP immunostaining (green) of the Bolwig’s organs (marked with white arrows). The embryos are oriented with anterior poles facing up. The scale bars correspond to 50 μm.
transcriptionally active state (Schwartz et al. 2010; Kharchenko et al. 2011; Huang et al. 2017). PHD and BAH domains may be involved in binding of Ash1 to either or both types of regions.

To investigate this possibility, we performed Chromatin Immunoprecipitation (ChIP) and immunostained polytene chromosomes with antibodies against Ash1. Ideally, we would have liked to examine the binding of various transgenic Ash1 variants in animals completely devoid of endogenous Ash1 protein (i.e., on ash19011/ash13M background). Unfortunately, such larvae proved too laborious to collect in numbers sufficient for ChIP. Therefore, we resorted to do the assays in ash122/ash19011 larvae (strong hypomorphs). As illustrated by Figure 6, in the control ash122/ash19011 larvae, ChIP signals and the polytene chromosome staining are reduced to background. However, both are restored to wild-type levels in larvae supplemented with two copies of transgenes expressing either the full-length or the PHD- or SET-deficient Ash1 proteins. This argues that the PHD- and

![Figure 6](image-url)

**Figure 6** Truncated Ash1 proteins can still bind chromatin. (A) Chromatin from the wild-type, ash19011/ash122, Ash1FL; ash19011/ash122 (Ash1FL), Ash1ΔPHD; ash19011/ash122 (Ash1ΔPHD), Ash1ΔBAH; ash19011/ash122 (Ash1ΔBAH), and Ash1ΔSET; ash19011/ash122; (Ash1ΔSET) third instar larvae was used for immunoprecipitation with the antibodies against Ash1. Histograms display the mean of the two independent experiments (n = 2) with dots indicating individual experimental results. noc, elf4a, hth, emc, rap, eff and Su(z)2 represent loci strongly bound by Ash1 in wild-type larval cells. An intergenic region on chromosome 3R (intergenic) and constitutively expressed Taf4 gene were used as negative controls. (B) Representative pictures of the distal part of the polytene X chromosome from the third instar larvae of the same genotypes as in (A) immunostained with antibodies against Ash1 protein (red). Immunostaining with antibodies against an unrelated BEAF-32 protein (green) was used as positive control. Staining with DAPI (blue) was used to map positions of characteristic polytene chromosome bands (yellow lines) according to Bridges nomenclature (Bridges 1935). No distinct bands are visible on the chromosomes of the ash19011/ash122 mutants stained with anti-Ash1 antibodies, but the immunostaining pattern is restored in larvae supplemented with ash1 transgenes. The chromosomes from Ash1ΔBAH; ash19011/ash122 larvae show weaker anti-Ash1 staining suggesting that Ash1ΔBAH protein may bind chromatin less well compared to other tested variants. All images were acquired at the same magnification.
SET-deficient proteins can still bind strongly to the de-repressed Polycomb regulated genes and weakly to multiple sites throughout the genome. Polytene chromosomes from Ash1ΔBAH; ash1^{9011}/ash1^{22} larvae show weaker anti-Ash1 staining suggesting that Ash1ΔBAH protein may bind chromatin less well compared to other tested variants. However, ChIP experiments did not detect obvious reduction in binding of the BAH-deficient protein so more work is needed to conclude whether BAH domain contributes to Ash1 binding to chromatin.

To gain further insight into the contribution of individual domains, we considered two possible scenarios. In the first scenario, the SET, PHD and BAH domains work in concert as parts of the same Ash1 molecule. Possibly, enabling one specific molecular property, for example, efficient lysine methylation by the SET domain. Alternatively, BAH, PHD and SET domains may act as independent modules each contributing its own distinct function. If Ash1 molecules form multimers (or at least a dimer), the second scenario allows for the functional Ash1 protein to be composed from a mixture of distinct individually truncated Ash1 polypeptides. The latter could be genetically tested by interallelic complementation. In such cases, highlighted by classical experiments using Ash1 variants, Class I flies are homologous for ash1^{T7F} mutation while Class II contains a wild-type ash1 allele (ash1^{+}). The cross is expected to yield twice as many Class II flies. (B) The ability of truncated Ash1 proteins to restore the viability of ash1^{T7F} mutants was compared to that of the full-length transgenic Ash1 (Ash1FL). While the Ash1ΔAT, Ash1ΔBAH and Ash1ΔPHD transgenes restore the viability to the same extent as Ash1FL, the Ash1ΔSET transgene does not. Note that even the viability of the Ash1FL; ash1^{T7F} flies remains at ~20% of the expected, which suggests that the ash1^{T7F} chromosome carries additional unrelated recessive mutations that affect fly fitness.

**DISCUSSION**

This study leads to three main conclusions. First, we have shown that, in addition to the methyltransferase activity of the SET domain, the functions of the PHD and BAH domains are necessary for Ash1 to counteract Polycomb repression and support fly viability. Unexpectedly, evolutionarily conserved AT-hook domains appear not critical for either, although, we cannot exclude that they contribute to some aspect of Ash1 function not detected by our experiments. Our findings urge more effort to understand molecular properties of the Ash1 PHD and BAH domains. In other chromatin regulators both kinds of domains are known to interact with N-terminal tails of histones H3 or H4. Depending on variations in their amino acid sequences, PHD and BAH domains may have higher affinities to tails methylated at specific Lysines or

**Figure 7** PHD- and BAH-deficient Ash1 proteins complement ash1^{T7F} mutation when provided in trans. (A) Crossing scheme to test the complementation of ash1^{T7F} mutation with ash1 transgenes. Female flies homozygous for an ash1 transgene on the second chromosome and heterozygous for ash1^{T7F} allele on the third chromosome were crossed to males with the same genotype. In the progeny, two classes are expected. While both classes are homologous for an ash1 transgene, Class I flies are homologous for ash1^{T7F} mutation (ash1^{+}) while Class II contains a wild-type ash1 allele (ash1^{+}). The cross is expected to yield twice as many Class II flies. (B) The ability of truncated Ash1 proteins to restore the viability of ash1^{T7F} mutants was compared to that of the full-length transgenic Ash1 (Ash1FL). While the Ash1ΔAT, Ash1ΔBAH and Ash1ΔPHD transgenes restore the viability to the same extent as Ash1FL, the Ash1ΔSET transgene does not. Note that even the viability of the Ash1FL; ash1^{T7F} flies remains at ~20% of the expected, which suggests that the ash1^{T7F} chromosome carries additional unrelated recessive mutations that affect fly fitness.
Arginines (Sanchez and Zhou 2011; Yang and Xu 2013). In addition, some of the PHD domains can bind DNA or RNA (with dissociation constants comparable to those for interactions with histones) (Weaver et al. 2018) and many BAH domains mediate interactions with non-histone proteins (Yang and Xu 2013). Curiously, it was recently reported that BAH domains of the Arabidopsis thaliana proteins SHORT LIFE (SHL) and EARLY BLOTTING IN SHORT DAYS (EBS) can specifically interact with the histone H3 tail methylated at Lysine 27 (Li et al. 2018). Should the BAH domain of Ash1 possess similar property, it would give Ash1 a way to discriminate between “regular” active genes and genes that just overcome Polycomb repression and still retain its hallmark, the histone H3 tri-methylated Lysine 27 (Schwartz et al. 2010).

Second, from the results of the interallelic complementation experiments, we conclude that the SET domain of Ash1 and the combination of its PHD and BAH domains (PHD-BAH cassette) represent two molecularly separable modules. This implies that many biochemical properties of the PHD-BAH cassette can be studied in isolation from the SET domain using shorter truncated Ash1 peptides. This is particularly advantageous given technical difficulties of working with the full-length protein of Ash1 size (over 250kDa). Our complementation experiments suggest that the PHD and BAH domains need to be on the same molecule to function properly. Yet, we note that small fraction of the ash19011 allele survives to adult stage when supplemented with the transgene expressing full-length Ash1FL (as evaluated by chi-square test). % viability indicates the ClassI/ClassII ratio multiplied by two. The same crossing scheme to test the complementation of ash19011/ash13M mutants with different combinations of ash1 transgenes. Flies heterozygous for one of the ash1 transgenes on the second chromosome and heterozygous for ash19011 allele, balanced over TM3,Ser,Act-GFP third chromosome, were crossed with flies homozygous for another ash1 transgene and heterozygous for the ash13M allele, balanced over TM3,Ser,Act-GFP third chromosome. The same crossing scheme was used for all combinations of Ash1ΔBAH, Ash1ΔPHD and Ash1ΔSET transgenes. In the progeny, two classes are expected. While both classes are trans-heterozygous for a combination of the ash1 transgenes, Class I flies are also trans-heterozygous for ash13M and ash19011 alleles (ash1) while Class II contains a copy of the wild-type ash1 allele (ash1+). The cross is expected to yield twice as many Class I flies. (B) Different combinations of truncated ash1 transgenes were compared for their ability to complement the lethality of the ash19011/ash13M mutants. The complementation by two copies of the Ash1FL transgene was used as positive control. n = the total number of progeny counted, P = probability that the observed difference in the survival rate of the ash1 mutant flies supplemented with corresponding combinations of ash1 transgenes is caused by chance (as evaluated by chi-square test).

Suppose Ash1 acts as a dimer. The equilibrium concentration of the example protein produced by the hypomorphic ash1 allele is poor, just 10% of that seen with the transgene expressing full-length Ash1. Nevertheless, this observation suggests that, while the PHD and BAH domains act in cooperative manner, the latter is still partially functional even in the absence of the PHD domain.

The third conclusion from this study is that in vivo Ash1 acts as a multimer. Additional biochemical studies are required to define the exact Ash1 tertiary structure. Regardless, the multimeric nature of the Ash1 protein is consistent with the dose dependent behavior of ash1 mutations (Shearn 1989; Tripoulias et al. 1994). Moreover, it provides an explanation of why the function of the small amount of wild-type protein produced by the hypomorphic ash1 allele is boosted by defective Ash1 molecules produced from the transgenic constructs. Suppose Ash1 acts as a dimer. The equilibrium concentration of the
Figure 9 Trans-complementation between Ash1 domains. In vivo, Ash1 molecules form dimers (as shown on the schematics) or multimers. The SET domain of Ash1 can function as an autonomous module. Therefore, Ash1 molecules that lack the SET domain (Ash1ΔSET) can combine with molecules that lack BAH- or PHD-domains to yield functional Ash1 complex (A, B). In contrast, a complex composed of the Ash1 molecules lacking either BAH- or PHD-domain appears ineffective (C). This suggests that BAH and PHD domains need to be part of the same molecule to function properly or may need to cooperate to allow multimerization.

It has been commonly assumed that the early premature stop codon of the ash122 allele completely abolishes ash1 function. Therefore, when it was found that the stochastic loss of the homeotic gene expression is detected only in ash122 homozygous embryos derived from the ash122 mutant germ cells, this was taken to indicate that maternally supplied Ash1 protein is sufficient to maintain the expression of homeotic genes throughout embryonic development. Since ash122 mutant animals derived from ash122 mutant germ cells survived until late larval stage, it seemed that the ash1 function was less critical than that of the trithorax (trx) gene whose mutants die during embryogenesis. Our observations argue that the lack of detectable changes in the homeotic gene expression in the ash122 embryos produced by the heterozygous mothers is due to small amount of function provided by the ash122 allele and that such changes are readily detectable in mutants with combination of stronger ash1 alleles. This, in turn, suggests that the phenotype of the true maternal and zygotic loss of ash1 function is still unknown and that the question of whether trithorax and ash1 functions are equally important is still open.

ACKNOWLEDGMENTS
We thank Dr. Maria Kim and Dr. Jan Larsson for gifts of fly strains and plasmids. We are grateful to Dr. Vincenzo Pirrotta for the gift of anti-Ash1 antibody. This work was supported in part by grants from Swedish Research Council to YBS and grants from Knut and Alice Wallenberg Foundation and Kempestiftelserna to EpiCoN (YBS co-PI). Stocks obtained from the Bloomington Drosophila Stock Center (NIH P40OD018337) were used in this study. The anti-Abd-B 1A2E9, monoclonal antibody developed by S. Celniker and the anti-BEAF-32 monoclonal antibody developed by P. Schell were obtained from the Developmental Studies Hybridoma Bank, created by the NICHD of the NIH and maintained at the University of Iowa, Department of Biology, Iowa City, IA 52242. Author contributions: ED and YBS conceived the project. ED performed all genetic experiments and analyzed homeotic phenotypes and embryonic Abd-B expression. TGK performed ChIP experiments and characterized the expression of Ash1 transgenes. AG analyzed Ash1 binding to polytene chromosomes. MS helped to analyze homeotic phenotypes. ED and YBS wrote the manuscript with input from all authors.

LITERATURE CITED
An, S., K. J. Yeo, Y. H. Jeon, and J. J. Song, 2011 Crystal structure of the human histone methyltransferase ASH1L catalytic domain and its implications for the regulatory mechanism. J. Biol. Chem. 286: 8369–8374. https://doi.org/10.1074/jbc.M110.203380
Bischof, J., R. K. Maeda, M. Hediger, F. Karch, and K. Basler, 2007 An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. Proc. Natl. Acad. Sci. USA 104: 3312–3317. https://doi.org/10.1073/pnas.0611511104
Bridges, C. B., 1935 Salivary chromosome maps with a key to the banding of the chromosomes of Drosophila melanogaster. J. Hered. 26: 60–64. https://doi.org/10.1093/oxfordjournals.jhered.a190422
Butcher, R. D., S. Chodagam, R. Basto, J. G. Wakefield, D. S. Henderson et al., 2004 The Drosophila centrosome-associated protein CP190 is essential for viability but not for cell division. J. Cell Sci. 117: 1191–1199. https://doi.org/10.1242/jcs.00979
Coddington, A., and J. R. Fincham, 1965 Proof of Hybrid Enzyme Formation in a Case of Inter-Allelic Complementation in Neurospora Crassa. J. Mol. Biol. 12: 152–161. https://doi.org/10.1016/S0022-2836(65)80289-3
Coleman, R. T., and G. Struhl, 2017 Causal role for inheritance of H3K27me3 in maintaining the OFF state of a Drosophila HOX gene. Science 356: eaai8236. https://doi.org/10.1126/science.aai8236
Dorafshan, E., T. G. Kahn, A. Glotov, M. Savitsky, M. Walther et al., 2019 Ash1 counteracts Polycomb repression independent of histone H3 lysine 36 methylation. EMBO Rep. 20: e46762. https://doi.org/10.15252/embr.201846762
Dorighi, K. M., and J. W. Tamkun, 2013 The trithorax group proteins Kismet and Ash1 promote H3K36 dimethylation to counteract Polycomb group repression in Drosophila. Development 140: 4182–4192. https://doi.org/10.1242/dev.095786
Dunn, J. G., C. K. Foo, N. G. Belletier, E. R. Gavis, and J. S. Weissman, 2013 Ribosome profiling reveals pervasive and regulated stop codon readthrough in Drosophila melanogaster. eLife 2: e01179. https://doi.org/10.7554/eLife.01179
Egger, H., A. Gortchakov, and H. Saumweber, 2004 Identification of the Drosophila interband-specific protein Z4 as a DNA-binding zinc-finger protein determining chromosomai structure. J. Cell Sci. 117: 4253–4264. https://doi.org/10.1242/jcs.01292
Gratz, S. J., F. P. Ukken, C. D. Rubinstein, G. Thiede, L. K. Donohue et al., 2014 Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila. Genetics 196: 961–971. https://doi.org/10.1534/genetics.113.160713
Hödl, M., and K. Basler, 2012 Transcription in the absence of histone H3.2 and H3K4 methylation. Curr. Biol. 22: 2253–2257. https://doi.org/10.1016/j.cub.2012.10.008
Hou, P., C. Huang, C. P. Liu, N. Yang, T. Yu et al., 2019 Structural Insights into Stimulation of Ash1’s H3K36 Methyltransferase Activity through Mrg15 Binding. Structure 27: 837–845 e833.
