Histone recognition and nuclear receptor co-activator functions of *Drosophila* Cara Mitad, a homolog of the N-terminal portion of mammalian MLL2 and MLL3

Chhavi Chauhan1,*, Claudia B. Zraly1,*,‡, Megan Parilla1, Manuel O. Diaz1,2 and Andrew K. Dingwall1,3,‡

**SUMMARY**

MLL2 and MLL3 histone lysine methyltransferases are conserved components of COMPASS-like co-activator complexes. In vertebrates, the paralogous MLL2 and MLL3 contain multiple domains required for epigenetic reading and writing of the histone code involved in hormone-stimulated gene programming, including receptor-binding motifs, SET methyltransferase, HMG and PHD domains. The genes encoding MLL2 and MLL3 arose from a common ancestor. Phylogenetic analyses reveal that the ancestral gene underwent a fission event in some Brachycera dipterans, including *Drosophila*, the C-terminal SET domain is encoded by *cmi* (*cara mitad* (*cmi*)) gene, which encodes the previously undiscovered N-terminal region consisting of PHD and HMG domains and receptor-binding motifs. The *cmi* gene is essential and its functions are dosage sensitive. CMI associates with TRR, as well as the EcR-USP receptor, and is required for hormone-dependent transcription. Unexpectedly, although the CMI and MLL2 PHDf3 domains could bind histone H3, neither essential and its functions are dosage sensitive. CMI associates with TRR, as well as the EcR-USP receptor, and is required for hormone-dependent transcription. Unexpectedly, although the CMI and MLL2 PHDf3 domains could bind histone H3, neither showed preference for trimethylated lysine 4. Genetic tests reveal that *cmi* is required for proper global trimethylation of H3K4 and that hormone-stimulated transcription requires chromatin binding by CMI, methylation of H3K4 by TRR and demethylation of H3K27 by the demethylase UTX. The evolutionary split of MLL2 into two distinct genes in *Drosophila* provides important insight into distinct epigenetic functions of conserved readers and writers of the histone code.

**KEY WORDS:** *Drosophila*, Co-activator, Histone, Hormone, Patterning

**INTRODUCTION**

Nuclear receptors (NRs) function as transcription factors that respond to cellular signals to initiate new gene expression programs (King-Jones and Thummel, 2005) and have essential roles in embryonic development, growth and differentiation. NRs collaborate with co-factors (>300) that provide important enzymatic and regulatory functions (Lonard and O’Malley, 2007; Lonard and O’Malley, 2006). Co-factors can be activators or repressors and are typically recruited to gene promoters through associations with receptors (Bulykho and O’Malley, 2011). Some co-factors direct changes in the epigenetic environment of target genes by direct covalent chromatin modification or nucleosome remodeling. Co-activators are recruited in a ligand-dependent manner, whereas unliganded receptors often associate with corepressors. Co-activators exist in large complexes required for the transcription of genes that are regulated by at least 48 vertebrate NRs, including retinoic acid receptor (RAR) (Goo et al., 2003; Lee et al., 2006), estrogen receptor (ER) (Mo et al., 2006), liver-X-receptor (LXR) (Lee, S. et al., 2008), farnesoid-X-receptor (FXR) (Kim et al., 2009), as well as a co-activator for p53 (Lee et al., 2009). Disruptions of both NRs and their co-regulators have been linked to many cancers and developmental disorders (Lonard et al., 2007; Sonoda et al., 2008).

Hormone signaling pathways in *Drosophila melanogaster* rely on two primary hormones, the steroid hormone 20-hydroxyecdysone (20HE) and sesquiterpenoid juvenile hormone (JH; JHE – FlyBase), and 18 receptors representing all major conserved nuclear receptor subfamilies (King-Jones and Thummel, 2005). *Drosophila* Ecdysone Receptor (EcR) is an FXR/LXR ortholog, whereas its heterodimeric partner Ultraspiracle (USP) is an RXR ortholog. *Drosophila* Trithorax-related (TRR) is a co-activator of EcR-USP. TRR is a histone lysine methyltransferase (HMT) that trimethylates histone 3 only on 4 (H3K4me3) and TRR functions are essential for activating ecdysone-regulated genes (Sedkov et al., 2003). TRR is closely related to another *Drosophila* protein, Trithorax (TRX), which regulates homeotic (Hox) gene expression through similar methyltransferase activity (Schuetzengruber et al., 2007; Simon and Tamkun, 2002). The mammalian counterparts of TRR are MLL2 (also known as ALR or MLL4) and MLL3 (also known as HALR). MLL2 and MLL3 are enormous (5537 aa and 4911 aa, respectively), with multiple conserved domains, including histone methyltransferase (SET) domain, five plant homeodomain (PHD) zinc fingers, an HMG-I binding motif, LXXLL NR binding motifs and FY-rich regions (Prasad et al., 1997). Through the SET domain, both MLL2 and MLL3 directly methylate histone H3 to mediate transcription activation (Issaeva et al., 2007; Vicent et al., 2011).

MLL2 and MLL3 are components of large SET1/COMPASS-like co-activator complexes (Eisenberg and Shilatifard, 2010; Miller et al., 2001; Nagy et al., 2002) that are required for NR-directed gene regulation (Goo et al., 2003; Issaeva et al., 2007; Lee et al., 2006; Mo et al., 2006). These complexes have important human disease connections, including developmental disorders and cancers. *MLL2*
and MLL3 are mutated in many Kabuki syndrome patients (Ng et al., 2010; Paulussen et al., 2011). MLL2 is frequently mutated in childhood medulloblastomas (14%) (Parsons et al., 2011), follicular lymphoma (89%) and diffuse large B-cell lymphoma (32%) (the two most common forms of non-Hodgkin lymphoma) (Morin et al., 2011), suggesting that MLL2 and MLL3 COMPASS-like complex activities have important epigenetic gene regulatory roles that normally function to inhibit cancer progression.

Proteins that co-purify with the MLL2 include ASH2, RBBP5 (RBQ3), DPY30, WDR5, adaptor protein ASC2, PTIP, PAI1 and histone demethylase UTX (Cho et al., 2007; Issaeva et al., 2007; Lee, S. et al., 2008; Mo et al., 2006). Recently, TRR was found in Drosophila COMPASS-like complexes (Mohan et al., 2011). Despite functional similarities, TRR is much smaller than MLL2 or MLL3 with homology limited to the C-terminal SET domain portion (Sedkov et al., 2003). TRR lacks the N-terminal PHD and HMG domains that might contribute to chromatin binding. MLL2-related family members are always encoded by large single genes in species other than Brachycera dipterans (A.K.D. and M.O.D., unpublished). To further our studies on epigenetic regulation of ecdysone target genes (Zraly et al., 2006), we searched for Drosophila genes that could encode a protein highly related to the N-terminal half of MLL2 and identified a single open reading frame (CG5591). We named the gene cara mitad (cmi; translated as ‘dear half’). Although cmi is unlinked to trr in the genome, our genetic studies using null mutants, in vivo depletion and overexpression revealed functions for cmi as a nuclear receptor co-factor necessary for hormone-regulated gene expression. Unexpectedly, the CMI type 3 PHD finger (PHDF3) (Chang et al., 2010; Park et al., 2010; Wang et al., 2010) was found to accommodate non-methylated, mono- and dimethylated H3K4, rather than trimethylated H3K4. Moreover, CMI-dependent activation also required demethylation functions of UTX, suggesting that NR-stimulated transcription involved at least three steps: binding of H3K4me1/2 by CMI, trimethylation of H3K4 by TRR and demethylation of H3K27 by UTX. The intriguing possibility that COMPASS-like functions in NR-directed transcription are associated with two independent proteins in flies suggests that recognition and binding to modified histones is a distinct step, separate from the epigenetic modification associated with other enzymes in the complex. This presents a unique opportunity to examine functions of histone recognition/binding and covalent histone tail lysine modifications as separate and essential features of NR-directed activation.

MATERIALS AND METHODS

Fly strains and genetic manipulations

Genetic crosses were performed on yeast-cornmeal-glucose medium at 25°C. Mutations in cmi (CG5591) were generated by Δ2-3 transposase-mediated excision of a F-element (IEdpn2/JE0642) residing ~380bp 5’ to the cmi transcript initiation site. Lethal phase analysis of cmi mutants was carried out as described (Zraly et al., 2002). Expression of epitope-tagged HA-cmi and short hairpin RNAi constructs (cmi-IR) was carried out by crosses with GAL4 driver strains at 18°C, 25°C or 29°C as indicated (Dietzl et al., 2007; Lam and Thummel, 2000). shRNA transgenic lines were obtained from Drosophila stock centers (Bloomington Drosophila Stock Center, Vienna Drosophila RNAi Center and National Institute of Genetics Fly Resource Center).

CMI antibodies, GST pulldowns and co-immunoprecipitation

A peptide corresponding to CMI amino acids (aa) 749–762 was used to generate rabbit polyclonal antisera (GenScript, Piscataway, NJ, USA). Preparation of embryo extracts and western analysis were performed using standard protocols (Zraly et al., 2003). Portions of CMI (XV, aa 647–1482; XS, aa 647–1071) and MLL2 (aa 1503–1560) were fused to GST in pGEX4T (Amersham/GE Healthcare, Piscataway, NJ, USA). Glutathione-S-transferase (GST) fusion proteins and control GST (C) were bound to agarose resin in equivalent amounts as assessed by Coomassie Blue gel staining. GST-pulldown assays were performed as described (Zraly et al., 2004). Wild-type Drosophila OregonR embryo extracts (500 μg) were treated with 10^-6 M 20HE for 1 hour at 4°C prior to incubation with immobilized GST proteins. Beads were washed extensively and bound proteins detected by immobiloblot using antibodies to EcR (Talbot et al., 1993), USP (Christianson et al., 1992) and TRR (Sedkov et al., 2003). Input lanes represent 20% (100 μg) of starting material. Extracts prepared from whole embryos or staged larvae/pupae containing P[GawB69B-GAL4]/UAS[GAL-ecr-HA-cmi] were immunoprecipitated with anti-HA antibodies (Abcam) (Brumby et al., 2002; Dingwall et al., 1995; Zraly et al., 2004). Extracts were pre-incubated with 20HE prior to co-immunoprecipitation.

ECoRE-lacZ assays

Transgenic flies carrying a 7xEcRE (Ecdysone Response Element) derived from the Hsp27 gene linked to a lacZ reporter (Koelle et al., 1991; Kozlova and Thummel, 2003; Talbot et al., 1993) were used to monitor CMi transactivation functions, either by immobiloblot with monoclonal antibodies to β-galactosidase or by quantitative CPRG assays (Seroude et al., 2002; Simon and Lis, 1987).

Chromatin immunoprecipitation (ChIP)

Whole-animal extracts prepared from early pupae expressing HA-CMI were used for chromatin immunoprecipitation assays according to manufacturer’s protocols (Millipore, Temecula, CA, USA), with rabbit anti-HA and anti-IgG (Abcam). qPCR reactions were carried out in triplicate. Fold enrichment was measured against IgG negative control and values normalized to ChIP input.

Histone analysis and peptide pulldown assays

Bulk cellular histones were purified from Drosophila S2 cells and staged pupae according to manufacturer’s protocols (Active Motif, Carlsbad, CA, USA). Western blots were performed using antibodies detecting modified or bulk histones. Band intensities were determined using ImageJ (NIH) analysis software, with H3K4me3 level values reflective of comparison to ‘GAL4 driver alone’ control (P[GawB69B] after normalizing to total H3 level for each sample.

Purified GST-PHD3b fusions (2 μg) were incubated with 1 μg biotinylated histone peptides (Millipore) in binding buffer (50 mM Tris-HCl pH 7.5, 150 mM or 300 mM NaCl, 0.1% NP40) overnight at 4°C. After 1 hour incubation with streptavidin beads (Pierce), immobilized complexes were collected by centrifugation, washed extensively with binding buffer, resuspended in 2× Laemmli sample buffer (BioRad, Hercules, CA, USA), boiled and loaded onto 12% SDS-PAGE gels. Western analysis was performed using antibody to GST (Abcam). GST-PHD3 fusions without histone peptides served as negative controls.

RNA analysis

Total RNA from 100 L2 larvae (~0–8 hours prior to L3 molt) of each genotype was prepared using the RNaqueous extraction system (Applied Biosystems, Foster City, CA, USA) and analyzed by semi-quantitative RT-PCR with HotStar Taq DNA polymerase (Qiagen, Valencia, CA, USA) using annealing temperatures appropriate for each primer pair. SYBR Green quantitative real-time PCR on reverse transcriptase reactions (qRT-PCR) was performed in triplicate using GoTaq qPCR Master mix (Promega). Levels of mRNA were analyzed using the comparative Ct method. qRT-PCR primers were used that spanned at least one intron when appropriate corresponding to usp, E75B (Eip75B – FlyBase), the EcR common region and specific E74 (Eip74E – Flybase) isoforms. The ribosomal protein gene rp49 (Rpl32 – FlyBase) was used as a reference standard.

Analysis of circulating hemocytes

Wandering L3 larvae were washed in water, dried briefly and punctured in mid-abdomen to remove circulating hemocytes (<1 μl). Hemolymph from five larvae of each genotype was suspended in 10 μl PBS, air-dried on
poly-lysine coated slides and fixed for 15 minutes in 4% formaldehyde in PBS, rinsed in PBS, then mounted in aquamount with DAPI. Circulating hemocyte counts were obtained from at least five independent measurements at each larval stage (ten fields counted and averaged) to obtain average cell counts per field. Experiments were repeated at least three times.

RESULTS

**Drosophila CMI is the counterpart to the N-terminus of the MLL family of nuclear receptor co-activators**

Key structural features of the vertebrate MLL co-activator family (Fig. 1) include a histone methyltransferase (SET) domain responsible for methylation of H3K4, PHD fingers thought to recognize and bind histone tails, NR binding motifs, FY-rich regions and an HMG domain (implicated in DNA binding) (Eissenberg and Shilatifard, 2010). *Drosophila* TRR is related to the carboxyl-terminal region of MLL2 and MLL3 and is required for hormone-dependent gene expression (Sedkov et al., 2003). Because TRR lacks the N-terminal PHD and HMG domains that might contribute to chromatin binding, we searched for *Drosophila* genes that might encode these features and identified a single homolog, CG5591, which we named cmi (cara mitad; Flybase FBgn0259119). The *cmi* gene potentially encodes for a 163 kDa protein with similar structure to the N-terminus of the MLL family proteins, including two PHD clusters, an HMG domain and three potential NR-binding sequences (Fig. 1A). CMI appears to be equally related to MLL2 and MLL3, suggesting a common ancestral origin. MLL-related homologs exist in the nematode *Caenorhabditis elegans*, primitive metazoan *Nematostella vectensis* and *Tribolium castaneum*, an insect of the order Coleoptera; however, these appear to be single proteins, implying that CMI and TRR split during the evolution of insects. The *cmi* gene resides on the second chromosome at 60A9, whereas *trr* is on the X chromosome at 2B14, suggesting a progenitor fission event. Cluster analysis based on the conserved HMG domain reveals potentially important common functions among the MLL family of transcription activators, perhaps in gene regulation through chromatin binding (Fig. 1B). SET domain cluster analysis also indicates strong evolutionary conservation (Fig. 1C).

Similar to *trr*, the *cmi* (CG5591) transcript is widely expressed with strong maternal contribution and uniform zygotic expression. Transcription of *cmi* increases during the early pupal stage, coincident with rising ecdysone hormone titers and the onset of metamorphosis (supplementary material Fig. S1). We found by immunostaining that CMI protein is broadly distributed in embryos and larval imaginal disc tissues (our unpublished observations).

**The *cmi* gene is essential for normal development**

We sought to determine whether *cmi* function was essential by generating null alleles and in vivo RNAi (Fig. 2). First, transposase-mediated excision of a closely linked *P*-element produced two mutant alleles of *cmi* (*cmi*¹ and *cmi*²). The *P[EPgy-2]EY06424* insertion is viable whereas the excision mutants are lethal as homozygotes and in combination with a large chromosomal deletion (*Df[2R]or-BR11*) that removes *cmi*. Both *cmi* mutations were sequenced and found to result in deletion of the *cmi* promoter and portions of the *cmi* open reading frame (Fig. 2A). qRT-PCR analyses of homozygous mutant larvae verified that no *cmi* mRNA was produced from either allele (Fig. 2B); however, RNA was still detectable from a divergently transcribed gene (CG5339/drev) in the *cmi*² mutant. We generated a full-length *cmi* cDNA with a hemagglutinin (HA) epitope-tag at the N-terminus of the open reading frame for rescue analyses. Expression of HA-cmi under GAL4 control in transgenic flies was able to restore viability.
Fig. 2. Proper CMI expression levels are required for normal development. (A) The cmi gene region, location of the P[EpGy2]ey06424 insertion and genomic deletions in the cmi P-excision mutants. RT-PCR primer sites are indicated by arrows. (B) Transcript analyses of cmi and drev in wild-type (OregonR), P[ey06424] and cmi mutant (cmi1 and cmi2) larvae. The ribosomal gene rp49 served as an RNA control. Quantitative RT-PCR (right panel) confirmed the complete loss of cmi mRNA in the cmi1 mutant. Asterisk indicates no mRNA detected. (C) Retained and malformed mouth hooks in cmi mutant larvae. Mutants display malformed hooks in L2 and occasionally L2 hooks are retained in L3. (D) Necrotic patches in cmi mutant larvae (arrows). Hemocyte counts were analyzed in wild-type (WT; OregonR, OR), cmi-IR, trr-IR and HA-cmi staged larvae (before/after ecdysone pulse). Shown are average hemocyte counts per field. Significant differences (P<0.01) were only observed comparing OR and HA-cmi hemocyte counts at both stages. (E) Depletion and overexpression of cmi in Drosophila embryos following P[GawB]69B-GAL4-dependent RNAi (cmi-IR) and overexpression (HA-cmi). CMI was detected using αCMI or αHA antibodies. Expression of the HA-cmi transgene was determined at varying growth temperatures (°C). Tubulin levels served as loading controls (αTub). (F) Expression of cmi in cultured S2 cells following cmi and trr RNAi. Knockdown of cmi was verified by western blot and qRT-PCR. (G) Proper CMI levels are essential for normal development. Left panels: Positions of veins in wild-type fly wings (OregonR) and upon cmi depletion (cmi-IR) and overexpression (HA-cmi) using the wing-specific C765-GAL4 driver. L2-L5, longitudinal veins; ACV, anterior crossvein; PCV, posterior crossvein. Note the shortening of veins with reduced cmi (arrowheads) and ectopic veins with elevated cmi levels. Right panels: Modest overexpression of HA-cmi at 22°C leads to additional pattern defects including ommatidial disruptions (arrows) and ectopic sex combs (arrow) on the male distal prothoracic tarsi (middle panels, enlarged in inset). RNAi depletion of cmi in developing eye discs using ey-GAL4 at 25°C results in reduced eye size and missing antennae (arrow, lower right). Error bars indicate s.e.m.
to cmi/Df(2R)or-BR11 hemizygotes, indicating that the lethality observed in the cmi mutant resulted from loss of cmi function. Second, we generated a cmi inverted repeat RNAi transgene (cmi-IR) and expressed it together with Dicer-2 to augment the RNAi effect (Dietzl et al., 2007) in vivo under the control of a GAL4-inducible promoter. Fig. 2E,F shows the levels of CMI upon cmi knockdown and overexpression both in vivo and in vitro using S2 cells. Widespread knockdown of cmi in the presence of Dicer-2 resulted in complete lethality (data not shown), whereas knockdown without Dicer-2 or knockdown in restricted patterns produced partial lethality and a broad range of developmental defects. Thus, the cmi gene is essential and the P-excision mutant alleles are likely to be null alleles.

To elucidate cmi functions in development, we performed a lethal phase analysis using both cmi alleles. The majority of homozygous cmi1 and cmi2 mutants failed to survive beyond the second (L2) or early third (L3) larval instar stages; however, substantial maternal contribution of cmi mRNA (supplementary material Fig. S1) might mask requirements during embryogenesis. Examination of dying larvae revealed that many had retained larval mouth hooks from the L2 molt and in some cases L2 and L3 mouth hooks were malformed (Fig. 2C), indicating a defect in hormone-dependent cuticle shedding. Mutant larvae exhibited sluggish movement, even though they retained normal olfactory functions, touch sensitivity and phototaxis (data not shown). Feeding cmi mutant larvae with ecdysone did not affect the phenotypes, suggesting that cmi functions downstream of ecdysone hormone production or secretion.

The cmi null mutant larvae eventually stopped moving and appeared to attempt pupariation. They failed to evert their spiracles, frequently turned brown but never produced a pupal cuticle, consistent with disrupted ecdysone hormone signaling associated with reduced EcR function (Bender et al., 1997). Many of the cmi larvae also developed interior necrotic patches (Fig. 2D), suggesting defects in larval hematopoiesis (Dearolf, 1998). Circulating larval hemocytes were isolated and counted from blue gut (BG) and clear gut (CG) third instar larvae, stages corresponding to before (BG) or during (CG) peak ecdysone titers. shRNAi depletion of cmi or trr using the P[GawB]69B-GAL4 driver led to insignificant changes in the number of circulating hemocytes at the CG stage. By contrast, overexpression of HA-cmi resulted in an approximately twofold increase in hemocyte count in both the BG and CG larvae, with no increase in the number of lamellocytes, indicating that the elevated hemocyte numbers were not directly related to ecdysone titers (Fig. 2D).

A requirement for cmi in adult tissues was determined using shRNAi depletion and overexpression. We confirmed GAL4-dependent cmi shRNAi knockdown and activation of the HA-cmi transgene by immunoblot (Fig. 2E). Widespread depletion and high level overexpression of cmi result in organellar lethality at various developmental stages (not shown). Targeted depletion of cmi in the wing imaginal disc resulted in shortened wing veins (cmi-IR; Fig. 2G) that could be enhanced by co-expression of Dicer-2 (Curtis et al., 2011; Dietzl et al., 2007) Expression of cmi-IR in the developing eye imaginal disc resulted in reduced eye size and loss of head and antennal structures (Fig. 2G). By contrast, modest overexpression of HA-cmi led to opposite pattern defects in the wing, as well as improper cuticle pigmentation, eye tumors and ectopic sex combs on the distal tarsi of male prothoracic legs (HA-cmi; Fig. 2G). Therefore, CMI levels are dose-sensitive for normal development and are under stringent regulation.

**Fig. 3. CMI functions in the ecdysone response pathway.** (A) qRT-PCR expression of ecdysone pathway regulators in late staged L2 wild-type and cmi mutant larvae. (B) Ecdysone-dependent induction of E74A and E74B transcription is reduced upon depletion of cmi in S2 cells and usp in vivo. E74 transcripts were reduced upon RNAi knockdown of cmi and upon GAL4-dependent expression of usp-IR in early pupae compared with the P[GawB]69B-GAL4 driver alone. (C) cmi interacts genetically with trr, EcR and usp. Shown are wings from wild-type (OregonR), cmi1 and trr4 heterozygotes and combinations of a cmi1 null allele with dominant-negative EcR or hypomorphic trr alleles. Incomplete veins are indicated by arrows. Ectopic veins associated with overexpressed HA-cmi are suppressed by simultaneous usp depletion. (D) Reduced cmi enhances the eye phenotype associated with a Bar mutation. The reduced eye phenotype of a heterozygous B1 female (left panel) is exacerbated in the presence of a heterozygous cmi1 mutation (right panel). The arrow lengths in both panels are the same. Error bars indicate s.e.m.

**CMI functions in hormone-dependent gene regulation**

Given the likely roles for cmi in hormone-regulated development, we examined several ecdysone primary response genes in cmi1 homozygous larvae using quantitative RT-PCR
Ecdysone receptor (EcR), ultraspiracle (usp) and E75B were unaffected, whereas the early ecdysone inducible gene Eip74EF showed reduced expression. Eip74EF produces two major protein isoforms, E74A and E74B, with common C-terminal ETS DNA-binding domains (Burtis et al., 1990). E74A and E74B are produced from overlapping transcription units and are coordinately regulated (Fletcher et al., 1995; Karim and Thummel, 1991). We confirmed the downregulation of E74 using RNAi knockdown of cmi in cultured S2 cells with added ecdysone and in vivo depletion of USP, a ligand-dependent activator of Eip74EF, at the early pupal (EP) stage (Fig. 3B; refer to Fig. 4D diagram).

TRR serves as a co-activator of EcR-dependent gene activation by virtue of its conserved SET methyltransferase (Sedkov et al., 2003). Like cmi null allele mutants, trr mutants are fully recessive (Sedkov et al., 1999) (Fig. 3C, lower left panel). Heterozygous combinations of cmi and a weak trr viable allele (trr<sup>v</sup>) produced flies with defects in wing patterning, as did combinations of cmi and an EcR dominant-negative mutant (Fig. 3C). The most prominent pattern disruptions appeared as shortened veins. Widespread depletion of usp is lethal, whereas shRNAi depletion in wing intervein cells using a blistered-GAL4 driver produces a crumpled wing phenotype (supplementary material Fig. S2). Simultaneous expression of HA-cmi with shRNAi usp in wing imaginal discs using the C765-GAL4 driver produced normal wings, indicating that the HA-cmi phenotype (shown in Fig. 2G) is suppressed by reducing USP levels and that, conversely, the effect of moderate depletion of USP is suppressed by elevated CMI levels (Fig. 3C, bottom right panel), suggesting functional association.
Normal eye development requires ecdysone pathway components, including the EcR-USP heterodimer (Cherbas et al., 2003; Gibb and Mc Keown, 2002) and the ecdysone-inducible zinc-finger transcription factor Broad-Complex (BR-C; BR – FlyBase) (Brennan et al., 2001). Progression of the morphogenetic furrow and ommatidial differentiation depend on ecdysone (Brennan et al., 1998). The B\textsuperscript{1} mutation displays a reduced eye size resulting from incomplete morphogenetic furrow progression that is enhanced by mutations in ecdysoneless and trr (Sedkov et al., 1999). In a similar fashion, a heterozygous cmi\textsuperscript{1} allele strongly enhanced the B\textsuperscript{1} phenotype (Fig. 3D), implicating CMI as a co-activator of EcR-USP functions.

**CMI is a nuclear receptor co-activator**

Portions of CMI with or without the NR-binding motifs were fused to GST and used to examine interactions with EcR, USP and TRR in embryo extracts (Fig. 4A). CMI interacted with both EcR and USP; and the most robust interaction involved the receptors and the largest CMI C-terminal fusion that contained the NR motifs. Little interaction was observed with CMI fusions lacking the HMG domain and NR-binding sites. TRR was previously shown to interact with both receptors (Sedkov et al., 2003); therefore, we investigated next whether TRR was part of the complexes formed between CMI and EcR-USP. We found that TRR was able to associate preferentially with a portion of CMI containing the HMG domain and NR interaction motifs (Fig. 4A). To confirm the interactions in vivo, we performed co-immunoprecipitation (CoIP) using the HA-tagged full length CMI (Fig. 4B). Immune complexes with CMI were enriched with TRR and USP, suggesting that they formed complexes in vivo. EcR showed a similar enrichment with CMI (not shown). As a control, the Brahma (SWI-SNF) complex component SNR1 did not appear to form complexes directly with CMI in this assay. During preparation of this manuscript, Shilatifard and colleagues (Mohan et al., 2011) identified a COMPASS-like complex that contained both TRR and CMI. In their report, they identified CMI as LPT (Lost PHD fingers of TRR).

Next, we tested for the ability of CMI to function as a transcription co-activator for EcR-USP using an in vivo lacZ reporter assay (Talbot et al., 1993). The Hsp27 gene EcR response element (EcRE) fused to the lacZ gene (7xEcRE-lacZ) responds to changes in hormone levels as a reporter for nuclear receptor-regulated transcription activity. Mid-embryonic (8-16 hours after egg laying) extracts have detectable EcRE-lacZ expression (Fig. 4C) resulting from EcR-USP activation in response to elevated ecdysone titer (Kozlova and Thummel, 2003; Maróy et al., 1988). lacZ levels were strongly upregulated in response to GAL4-driven expression of HA-cmi, whereas a control protein (SNR1) showed no change. Quantitative β-galactosidase (CPRG) assays confirmed that lacZ levels were elevated ~1.7-fold, consistent with a co-activator function.

To determine whether the CMI co-activator function was hormone dependent, we measured β-galactosidase activity from the EcRE-lacZ reporter during the larval-pupal transition stages (Fig. 4D). Quantitative assays were performed using extracts from staged animals with P[GawB]69B-GAL4-driven HA-cmi and cmi-IR. There was little lacZ produced in wild-type (EcRE-lacZ), HA-cmi or cmi-IR animals during the larval BG stage. lacZ levels increased during subsequent stages presumably in response to hormone-dependent activation. Differences in lacZ expression were first apparent at the CG stage when ecdysone titers peak, with HA-cmi animals showing substantially higher levels (fivefold vs BG) compared with wild-type (twofold vs BG). This trend continued in the EP stage; however, by the late prepupal stage (LP), widespread overexpression of HA-cmi caused lethality. By contrast, lacZ levels were lower in cmi-IR animals at all subsequent stages compared with wild type, indicating that CMI is required for ecdysone-stimulated transcription of the EcRE-lacZ reporter.

Based on our observed genetic and physical interactions between CMI and EcR-USP, we next sought to determine whether CMI was recruited to ecdysone response elements in vivo (Fig. 4E). Chromatin immunoprecipitation using whole animal extracts prepared from early pupae expressing HA-CMI and antibodies to the HA epitope revealed that CMI was strongly enriched at the endogenous Hsp27 EcRE, but not at another ecdysone inducible gene (Eig71Eh-i promoter) (Wright et al., 1996; Zraly et al., 2006) that lacks an EcRE, or at a control (rp49) gene. We conclude that CMI is a hormone-dependent co-activator of nuclear receptor-stimulated transcription.

**CMI is a chromatin-binding protein**

The CMI protein contains two PHD finger clusters (Fig. 1A) that are strongly conserved among MLL family proteins and are closely related to similar PHD fingers found in a variety of chromatin-binding proteins involved in transcription regulation (Fortschegger and Shiekhattar, 2011; Yap and Zhou, 2010). The CMI cluster ‘b’ PHD fingers (Fig. 5A) appear to be the most highly conserved and similar PHD fingers have been shown in some instances to recognize and bind histone H3 N-terminal tails, with distinct specificities. An alignment of type 3 PHD fingers (PHD3) (Chang et al., 2010; Park et al., 2010; Wang et al., 2010) similar to CMI is shown in Fig. 5A, including those known to bind H3K4me2/3 in vitro. A fusion of the CMI PHD3.b to GST was used to determine whether it was able to bind histones in vivo and in vitro. Analysis of native histones obtained from cultured S2 cells revealed that the PHD3.b was able to bind histone H3, but not H4, H2A or H2B (Fig. 5B; data not shown). We confirmed this result by generating a mutant version of the CMI PHD3.b that carried an amino acid change (W680A) predicted to disrupt histone binding (Pena et al., 2006). Surprisingly, neither wild type nor mutant GST-CMI PHD3.b fusions showed preferential binding to native H3K4me3. Instead, the wild type but not the mutant GST fusion was capable of binding bulk histone H3 as well as H3K4me2 (supplementary material Fig. S3). Unlike vertebrate MLL1 and Drosophila TRX, the CMI PHD3.b was unable to bind the cyclophilin CYP33, which is implicated in controlling MLL and TRX functions (Anderson et al., 2002; Fair et al., 2001; Horn et al., 2010; Park et al., 2010; Wang et al., 2010). Modified and unmodified histone H3 N-terminal peptides (aa 1-21) fused to biotin were used in pulldown assays to determine whether the PHD3.b of CMI could recognize and bind H3K4me3 in vitro (Fig. 5C). Similar to the native histones, the CMI PHD3.b was able to bind unmodified histone H3 and mono- and dimethylated H3K4 but not H3K4me3 peptides, and the W680A mutant was unable to bind any of the biotinylated peptides. To test whether this was a conserved feature of the MLL family PHD3.b domains, we generated a GST fusion of the MLL2 PHD3.b (aa 1503-1560) and found a pattern of histone peptide recognition similar to CMI. As a positive control, the PHD3 of the LID demethylase demonstrated strong preference for H3K3me2/3 (Li et al., 2010). Thus, similar to other closely related PHD3 domains, the CMI and MLL2 PHD3 domains appear to be able to accommodate and bind to H3K4 with methylated sidechains. However, despite their similarity to the MLL1 and LID PHD3 domains, the CMI and MLL2 PHD3.b
domains do not show a preference for recognition and binding of H3K4me3, suggesting that their primary recognition target might be another substrate.

We sought next to determine whether loss of CMI would impact chromatin methylation in vivo. Previous studies indicated that in vivo depletion of Caenorhabditis set-16 (Fisher et al., 2010) and Drosophila trr in S2 cells (Ardehali et al., 2011) resulted in diminished H3K4 trimethylation as determined by immunoblot. Depletion of trr in Drosophila L3 wing discs also appeared to result in modest reductions in H3K4me3 levels by immunofluorescence (Mohan et al., 2011). Widespread depletion of trr in early pupae using shRNAi resulted in reduced global trimethylation of H3K4 (Fig. 5D), consistent with reductions in methyltransferase activity (Sedkov et al., 2003). Importantly, widespread depletion of cmi also resulted in a decreased global appearance of H3K4me3 marks, suggesting that histone binding by CMI might help to stabilize H3K4me3 in vivo. To our knowledge, the CMI type 3 PHD finger is the first example of a PHD3 that can accommodate H3K4me0/1/2 marks, but not H3K4me3, suggesting a distinctly different role in transcription regulation.

**CMI biological functions depend on other co-activator complex factors**

In vertebrates, MLL2 and MLL3 are found in large Set1/COMPASS-like complexes (Eissenberg and Shilatifard, 2010). These include common subunits that are highly conserved (Ruthenburg et al., 2007), including ASH2, WDR5, RBBP5 and DPY30. The MLL2-MLL3 complexes contain additional subunits that appear to be specific for hormone-dependent gene regulation, including the histone H3K27 demethylase UTX(KDM6A)-UTY, nuclear transactivation protein PTIP, NCOA6 and PA1. We therefore questioned whether CMI biological functions are dependent on other putative COMPASS-like components. Moderate depletion of cmi in the developing wing displayed weak pattern disruptions that were enhanced by simultaneous expression of Dicer-2 (Fig. 6A,B), indicating that the phenotypes were dose-sensitive for cmi function. Similarly, depletion of trr (trr-IR; Fig. 6G) resulted in shortened longitudinal and crossveins. Co-expression of cmi-IR and trr-IR resulted in further shortening of the veins, indicative of a genetic cooperation (Fig. 6C). By contrast, co-expression of ash2-IR or wds-IR (Hallson et al., 2012) produced little enhancement of the cmi-IR phenotype (Fig. 6D), suggesting that ASH2 and WDS are not dose limiting in this assay. Co-expression of Utx-IR with cmi-IR produced the strongest enhancing interaction (Fig. 6F). Modest depletion of ash2, wds and Utx independently under identical conditions had no obvious wing phenotype (not shown).

Having established genetic relationships between putative complex co-factors through enhancing cmi loss-of-function phenotypes, we next sought to verify the interactions by testing whether limiting these factors would suppress a HA-cmi overexpression phenotype (Fig. 6H). Co-expression of HA-cmi with trr-IR resulted in almost complete suppression of the HA-cmi phenotype (Fig. 6I). Importantly, the wings displayed the trr loss-of-function phenotype (Fig. 6G), suggesting that although the HA-cmi phenotype was dependent on trr, the overexpression of cmi could not compensate for loss of trr function. Reductions of ash2

---

**Fig. 5. CMI is a chromatin-binding protein and is required for proper histone methylation in vivo.**

(A) PHD type 3 finger comparisons and schematic of the GST-CMI PHD3 construct with the zinc-coordinating sequences underlined. The CMI PHD3 regions are highlighted in the alignment. Residues important for the binding of PHD3 domains to H3K4me3 are underlined and indicated by asterisks. Amino acids important for interaction between PHD3 of MLL1 and cyclophilin CYP33 are shaded. Hs, Homo sapiens; D, Drosophila melanogaster; Ce, Caenorhabditis elegans. (B) CMI PHD3.b region associates with native histone H3. Native histones prepared from S2 cells were incubated with similar amounts of solubilized GST alone or GST-CMI PHD3.b immobilized on glutathione agarose resin. Precipitated proteins were probed by immunoblot for histones H3 and H4, and CYP33. (C) The CMI and MLL2 PHD3.b binds non-methylated H3, H3K4me1 and H3K4me2. Biotinylated histone peptides were incubated with similar amounts of GST-CMI PHD3.b, GST-CMI PHD3.b (W680A), GST-ALR/MLL2 and GST-LID. Bound GST fusions were analyzed by immunoblot with αGST antibodies. (D) CMI is required for normal levels of H3K4me3 in vivo. Native histones were prepared from early pupae expressing GAL4 alone (P[GawB]69B-GAL4 driver) and trr-IR or cmi-IR transgenes. Extracts were probed by immunoblot. The values shown under each lane represent band intensities compared with driver control normalized to H3 levels.
and wds moderately suppressed the effects of overexpressed HA-cmi (Fig. 6J,K), whereas reduction of Utx showed almost complete suppression (Fig. 6L). Significantly, in these assays TRR is present (except for trr-IR), allowing us to examine effects of altering cmi levels independent of the H3K4 methyltransferase activity. We conclude that UTX H3K27 demethylase activity is required along with histone binding functions of CMI and the H3K4 methyltransferase activity of TRR to activate nuclear receptor target genes.

**DISCUSSION**

Although the precise roles of proteins directly participating in nuclear receptor signaling remain largely speculative, many are thought to regulate transcription through effects on chromatin (Bannister and Kouzarides, 2011). The MLL2 and MLL3 co-activators function to epigenetically decode or modify histone lysine residues and provide activation functions for NR signaling at target genes (Eissenberg and Shilatifard, 2010). In Drosophila, CMI and TRR together have a single MLL family homolog. To our knowledge, this is the first example of an evolutionary ‘splitting’ of an epigenetic regulator involved in nuclear receptor signaling, whereby the essential gene regulatory functions of one protein have been parsed into two distinct proteins. CMI forms complexes with TRR, associates directly with hormone receptors and interacts with other putative COMPASS-like components, suggesting that Drosophila contains a functional counterpart to the mammalian ASCOM-MLL2 nuclear receptor co-activator complex (our results) (Mohan et al., 2011).

The MLL histone lysine methyltransferases (KMTs) can be divided into two conserved groups, the MLL1-MLL4(2) and MLL2(4)/ALR-MLL3/HALR subfamilies. Each MLL member is capable of forming related discrete complexes with several common components. The MLL-based complexes activate transcription in part through methyltransferase activity on histone H3 Lys4 residues within promoter-associated nucleosomes (Eissenberg and Shilatifard, 2010; Martin and Zhang, 2005). It might be partial functional overlap between MLL2 and MLL3 (Lee, J. et al., 2008), however, they are not redundant with the MLL1-MLL4 subfamily. The SET-domain methyltransferase activity of the MLL proteins is essential for transcription activation through histone lysine methylation, but the precise biological role of PHD fingers remains somewhat elusive. Closely related PHD fingers bind H3K4me3/2, the product of the methyltransferase activity. Within the context of a single protein, such as MLL1, the PHD3 recognition and binding of H3K4me3 is required for transcription activation of target genes (Chang et al., 2010).

Our findings that CMI and TRR function coordinately in a COMPASS-like complex suggest that cmi and trr probably split from a common ancestor. Gene-protein fusions are four times more common than fissions, perhaps reflecting a simpler genetic event (Kummerfeld and Teichmann, 2005). In cases in which fissions occur, it has been suggested that many involve subunits of multimeric complexes in which the two independent proteins interact physically (Kummerfeld and Teichmann, 2005). The process of splitting into two independent genes might involve gene duplication with subsequent partial degeneration, as has been observed in the monkey king (mkg) gene family in Drosophila (Wang et al., 2004).

The notion that a large protein contains domains that function both together and independently is not without precedent. TRX and MLL1 are cleaved by a specific protease, taspase-1. The two ‘halves’ interact with each other in a functional complex, but there is evidence that the N-terminal TRX peptide (TRX-N) binds chromatin without its TRX-C partner in transcribed regions of Hox genes (Schuettengruber et al., 2009; Schwartz et al., 2010). Transcription factor TFIIA and herpes simplex virus host cell factor (HCF1) are cleaved during maturation, with both halves necessary for a functional product (Hsieh et al., 2003). There is presently no evidence that MLL2 or MLL3 are cleaved or processed.
An important question is whether both the chromatin-binding and methyltransferase functions of the MLL family are required for transcription activation. Our data indicate that depletion of trr can suppress the effects of overexpressing cmi, suggesting that the activation potential of CMI depends on TRR methyltransferase activity. Similarly, simultaneous depletion of cmi and trr produces stronger phenotypes than depletion of either alone, indicative of cooperation on similar gene targets. Moreover, in vivo depletion of cmi results in reduced global H3 trimethylation, despite a functional trr gene.

### Developmental functions of CMI

Phenotypes associated with changes in CMI levels reveal important functions in hormone-regulated development. The larval defects in molting, morphogenetic furrow progression and necrosis associated with a cmi null allele, similar to trr (Sedkov et al., 2003), are consistent with impaired hormone signaling. Similarly, depletion of MLL2 in HeLa cells using siRNA led to reduced expression of genes known to be important for development and trimethylation of H3K4 was reduced at some promoters (Issaeva et al., 2007). Knockdown of MLL2 in MCF-7 cells impaired estrogen receptor (ERα) transcription activity and inhibited estrogen-dependent growth (Mo et al., 2006). Inactivation of the murine Mll3 resulted in stunted growth and reduced PPARγ-dependent adipogenesis with increased insulin sensitivity (Lee, J. et al., 2008). Perhaps reflecting synonymous functions in Drosophila, cmi/CG5591 was found to be important for regulating muscle triglyceride levels, suggesting conserved adipogenic functions (Pospisilik et al., 2010). Furthermore, CG5591 (cmi) is involved in phagocytosis (Stroschein-Stevenson et al., 2006) and regulation of caspase functions in response to cellular stress (Yi et al., 2007), implicating cmi in immune-cell regulation. The increased hemocyte number associated with elevated CMI suggests functions in hemocyte development, perhaps as an effect of chromatin remodeling (Badenhorst et al., 2002) or signaling (JAK/STAT, Hedgehog, Notch pathways) (Crozatier and Meister, 2007; Hou et al., 2002). It was previously shown that trr was important for Hedgehog (HH)-dependent signaling during eye development (Sedkov et al., 2003) and our cmi overexpression and depletion data are consistent with that possibility. However, the dosage-dependent cmi wing phenotypes are not consistent with changes in HH signaling (Baker, 2007), raising the possibility that cmi and trr are important for other growth and signaling pathways in wing development, including Decapentaplegic (DPP/TGFβ) and Wingless (WG/WNT) pathways (Blair, 2007).

Several steps are involved in activation of hormone-responsive target genes, including methylation of H3K4 by the MLL2-MLL3 COMPASS-like complex and displacement of demethylases (Vicent et al., 2011). Reduced cmi function resulted in lower hormone-responsive enhancer activation and genetic interactions between cmi, trr and Utx revealed that chromatin binding by CMI was important for gene activation in vivo. Furthermore, RNAi depletion of Utx suppressed HA-cmi overexpression wing phenotypes, suggesting that demethylation of H3K27 is a prerequisite for activation of some hormone target genes. This is supported by genetic evidence from C. elegans that indicated both histone H3K4 methylation by SET-16 (MLL2/MLL3 ortholog) and H3K27 demethylation by UTX-1 were required for attenuation of RAS signaling in the vulva (Fisher et al., 2010; Li and Kelly, 2011) and MLL2-MLL3 complex-related components were required for proper germ line development (Li and Kelly, 2011). Our genetic epistasis data reveals that Utx, trr and cmi functions are all required for activation in Drosophila.

Unexpectedly, the CMI PHD3.b showed binding to mono- and dimethylated H3K4, rather than trimethylated H3K4 (Sanchez and Zhou, 2011). Although CMI contains two PHD3 domains in two clusters similar to MLL3, MLL2 contains one PHD3 most closely related to the CMI and MLL3 PHD3.b domains. The second cluster appears in all isoforms of MLL3, whereas the N-terminal ‘α’ cluster is optional (Ruault et al., 2002). Additionally, the ‘β’ cluster is more closely related to the PHD cluster found in other MLL family proteins (Eisenberg and Shilatifard, 2010). PHD modules are thought to bind histones and present tail residues to the modifying enzyme subunits or stabilize those enzymes with their substrates (Musselman and Kutateladze, 2011). Recently, RNAi knockdown of trr in S2 cells was shown to affect H3K4 mono-, di-, and trimethylation, revealing widespread functions in regulating methylation in vivo (Ardehali et al., 2011) and suggesting that loss of TRR might destabilize the co-activator complex leading to de-protection of H3K4 methylation. One possibility is that CMI binds mono- and dimethylated H3K4 to prevent demethylation and stabilize TRR to allow for hormone-stimulated methylation and gene activation. CMI might disengage to allow for removal of methylation marks as hormone levels decrease and gene transcription is reduced. In contrast to MLL1-TRX function in maintenance of active gene transcription, CMI and TRR might be required for NR-targeted gene activation in response to temporally restricted hormone-dependent genome reprogramming.

### Acknowledgements

We thank A. Mazo, Y. Sedkov, R. Barrio, J. Secombe, C. Thummel for reagents, including antibodies and fly mutants; V. Bodapati for phylogenetic analyses; the Bloomington Drosophila Stock Center, Vienna Drosophila RNAi Center and National Institute of Genetics Fly Resource Center for GAL4 driver and RNAi transgenic lines.

### Funding

This work was supported by grants from the National Institutes of Health [1PO1-CA105049-01 to M.O.D.] and National Science Foundation [MCB0818620, MCB1122001 to A.K.D.]. Deposited in PMC for release after 12 months.

### Competing interests statement

The authors declare no competing financial interests.

### Supplementary material

Supplementary material available online at http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.076687/-/DC1

### References


