INTERNSHIP ARTICLE

A Comparison of Drosophila CMI to Human MLL2/ALR Reveals Homologous Histone Binding and Recognition Preferences

Riva Trivedi[1], Claudia Zraly[2], and Andrew Dingwall[2]*
Student[1]: The Illinois Mathematics and Science Academy, 1500 Sullivan Road, Aurora, IL 60506
Mentor[2]: Loyola University Stritch School of Medicine, 2160 South 1st Avenue, Maywood, IL 60153
*Corresponding author: adingwall@lumc.edu

Abstract

The MLL/ALR family consists of enormous multi-domain proteins found in large co-activator complexes (called COMPASS-like complexes) involved in nuclear receptor dependent gene transcription. These complexes activate gene transcription through binding and covalent modification of histone residues. Mutational loss of the human MLL/ALR genes has been implicated in developmental disorders and cancers. The Drosophila MLL/ALR homolog is a single representative split into two genes during evolution, known as cmi and trr, with each encoding for conserved portions essential for transcription regulation. To further define the functional relationship between cmi and trr, we used in vivo knock-down experiments using conditional shRNAi transgenes. We found depletion in cmi and trr gene levels greatly affected development of Drosophila, and resulted in lowered global histone lysine methylation. Different "loss of function" phenotypes were also noted, such as defects in wing vein development. To determine whether the fly and human MLL/ALR share similar histone binding properties, we cloned the human ALR/MLL2 PHD3 finger by PCR using reverse transcribed mRNA, performed protein expression analysis, and determined histone binding preference using whole genomic chromatin and histone peptide pulldown analyses. We found the human MLL/ALR genes and the Drosophila homologs share similar histone binding and recognition preferences.

Introduction

Cancers are caused by gene mutations followed by uncontrolled cell growth, with the rate of cell growth determining whether a cancer is benign or malignant/metastatic. Cancers are formed based on different environmental and genetic factors. This explains why even after countless years of research in effort to developing a "cure", there is no single remedy that can be discovered to solve every problem. However, investigating the potential causes and effects that cause cancer to form can make it possible to learn more about the issue. This can be achieved by examining distinctive genetic effects in offspring based on mutations and modifications in their DNA. In this project, genes particularly known for forming cancerous diseases were studied.

The genetic material of a eukaryotic cell, known as DNA, is contained in the nucleus in structures called chromosomes. Because they are so long, chromosomal DNA strands must be compacted, twisted several times, and wrapped around histone octamers known as nucleosomes. These nucleosomes also twist into larger structures of chromosomes. The chromosomal DNA, along with proteins that package it in the nucleus of the cell, are often referred to as chromatin. Chromatin's function is to stabilize and compact the DNA, prevent DNA damage, as well as regulate gene expression and DNA replication[1].

Histones have specific amino acid residues which can be covalently modified through acetylation, methylation, and phosphorylation, known as the "histone code", which can lead to different effects in gene regulation[2]. Specifically, the methylation of Histone H3 on lysine four (H3K4) can affect 'epigenetic memory'. The epigenetic marks are placed onto the histone tail residues (e.g., trimethylated lysine 4 or H3K4me3) as a consequence of transcription and are sometimes heritable through successive cell generations[3].

In humans, the paralogous ALR/MLL2 and HALR/MLL3 proteins are found as core components of COMPASS-like transcription coactivator complexes[4,5]. MLL2 and MLL3 are enormous (both are encoded by genes nearly 15,000 base pairs in length) and contain multiple functional domains, including a methyltransferase enzyme which puts methyl groups distinctively onto H3K4, nuclear receptor binding motifs and regions known as PHD (plant homeodomain) zinc fingers that are thought to mediate protein-protein interactions. The histone lysine methylation appears to generally correlate positively with cell growth and endurance and is associated with active gene transcription[6]. In addition to the enzymatic modification of histones, the MLL2 and MLL3 proteins also contain as many as 7 clustered PHD finger domains within the N-terminal portion. These PHD finger domains are each small ~60 amino acid portions of the protein that in some cases have been shown to recognize and bind to specific histone tail modifications. The single Drosophila MLL2/MLL3 counterpart is split into two proteins known as Cara Mitad (CMI) and Trithorax-related (TRR). The CMI protein is highly related to the N-terminal portion, while the TRR protein is related to the C-terminal portion of both MLL2 and MLL3[7]. CMI and TRR have been found in the same COMPASS-like complex in Drosophila[7]. The CMI protein in Drosophila directly corresponds to the part of MLL2/MLL3 containing these PHD finger domains. In fact, the specific type 3 fingers (PHDf3) in humans and Drosophila have a very close conservation of amino acid sequences[7].

The Drosophila cmi and trr genes are essential for organismal development, with mutant phenotypes reminiscent of those associated with loss of hormone signaling[7,8]. It is already understood that the mammalian MLL2 and MLL3 genes both encode for H3K4 methyltransferases which function as epigenetic transcriptional activators during the fetal developmental period[9,10]. Recent studies have illustrated the effects of mutations in the

The Journal of Experimental Secondary Science

human MLL2 and MLL3 and murine Mll3 genes, with connections to dramatic changes in early development, cell survival, and diseases such as cancer. In the past several years, the genomes from patients with a variety of cancers have been sequenced to identify mutations correlated with disease. These studies have identified a significant number of somatic mutations in the MLL2 and MLL3 genes associated with non-Hodgkin lymphoma[11]. A similar study on patients with transitional cell carcinoma (the most common form of bladder cancer) also found a large number of mutations in chromatin remodeling genes, including MLL2[12]. A rapidly growing number of publications have reported loss of MLL2 or MLL3 in association with childhood brain cancers (medulloblastomas), kidney, prostate and breast cancers, as well as pancreatic and gastric cancers. Additionally, other studies suggest Kabuki syndrome, a developmental disorder known to be linked to skeletal and cognitive defects, plus cardiac and immunological disorders, are also caused by mutations in MLL2[13,14]. These studies have dramatically increased our understanding of how misregulation of histone modifications has caused great maladies that could lead to human cancer. Factors which moderate the additions, eliminations, and readings of modifications are causing this misregulation[15].

The strong similarity between the human MLL2 and MLL3 and fruit fly CMI proteins is strong evidence that they are functional homologs. In order to examine this directly, we wished to test the hypothesis that the conserved PHDf3 domains would recognize and bind similar histone tail modifications. We first cloned the MLL2 PHDf3 region and then characterized the histone binding preference of both the MLL2 and CMI PHDf3 fingers. Due to its easily manipulative genetic tools, the fruit fly Drosophila melanogaster is a wonderful model system for investigating the effects of histone modifications during development. Using the Drosophila genetic system, we investigated the role of CMI in the regulation of epigenetic modifications. We individually depleted CMI and TRR in Drosophila using conditional expression of silencing RNAs, and observed the effects on the progeny to test whether the PHD finger and methyltransferase have essential functions in development.

Generally, by exploring both biochemical and genetic factors, we gained quantitative and qualitative results to steer us in answering how the structures and mechanisms for recognition of specifically modified histones are similar and/or different between the Drosophila CMI and human MLL2/3 PHD3 fingers. They could share analogous properties, or they could be unlike each other. We hypothesized the structures and mechanisms would be similar due to their highly conserved sequences.

Materials and Methods

Cloning of the MLL2 PHDf3 region, bacterial expression and purification. The MLL2 PHDf3 region was cloned and expressed as a bacterial glutathione-S (GST) fusion protein. Due to the presence of an intron within the genomic DNA encoding for the PHDf3, we obtained complementary DNA (cDNA) from NIH293T human cell line. Specific Polymerase Chain Reaction (PCR) primers were used to amplify and isolate the specific finger region (amino acids 1503-1560). The primers were designed to include a portion of the PHD finger sequence with sequences containing unique restriction sites (BamHI and EcoRI) that allowed for cloning and manipulation. Amplification was performed using a proofreading polymerase to minimize the possibility of amplification errors. The PCR product was then cloned into pBlueScript II SK (Fermentas, Inc.) and sequenced. We created a bacterial fusion with the finger and glutathione-S-transferase (GST) in the vector pGEX4T (Amersham, Inc.) using the engineered restriction sites (Figure 1A). The GST:PHDf3 protein was estimated to be 32.3 kilodaltons (kD). In order to test for inducible expression of the fusion protein, bacteria containing the GST fusion plasmid were grown at 30°C, induced with IPTG, electrophoresed, and then run on 10% SDS PAGE protein gels for 75-90 minutes at 125 volts. Proteins were visualized by staining with coomassie blue. A 0-4 hour time course was run to measure induction (Figure 1B). In order to isolate the soluble protein fusion, proteins were extracted using the B-Per reagent (Pierce/Thermo-Fisher Scientific, Inc.) and soluble (Sol) and insoluble (Insol) protein was collected. The soluble GST:PHD fusion was purified away from other contaminating proteins using glutathione agarose resin. Soluble bacterial extract was added to Glutathione-agarose (Glu-agarose) resin (Sigma) and the GST:PHDf3 protein was eluted from the resin using free glutathione (Sigma). Varying amounts of the purified protein were examined by 10% SDS-PAGE gel analysis for 75-90 minutes at 125 volts (Figure 1C). *Histone binding analysis.* Native histones were prepared from either Drosophila or human cultured cells as described[7] and incubated with similar amounts of the CMI7 or MLL2 PHDf3 fusion proteins. Bound histone was analyzed by immunoblot with αH3 antibodies (Abcam). A Western blot was run on the purified protein to illustrate it bound to Histone 3, ensuring its functionality (Figure 2 and Chauhan et al., 2012[7]). We used modified histone peptide arrays to test binding preferences of both this PHDf3 as well as that of Drosophila[7]. Briefly, biotinylated histone peptides (Millipore, Inc.) were incubated with similar amounts of GST-CMI PHDf3.b, GST-CMI PHDf3.b (W680A-mutant version that is predicted to not bind histones), GST-ALR/MLL2 and GST-LID (control PHDf3). Bound GST fusions were analyzed by immunoblot with αGST antibodies (Abcam). *Genetic studies in Drosophila.* We used the 'Driver-Responder' system for gene-specific depletion frequently used in Drosophila melanogaster fruit flies, to address the role of certain genes in the regulation of epigenetic modifications[16,17,18]. Genetic combinations were made by using an engrailed GAL4 driver which activated a shRNAi (obtained from the Bloomington Drosophila Stock Center and the Vienna Drosophila Stock Center) to deplete certain genes. The driver serves as a transcription factor that "drives" the expression of the transgene carrying the short hairpin in a certain time and location within the developing fly. In our case, we depleted cmi and trr to observe the effects of losing either the PHD finger or the methyltransferase in the progeny of the Drosophila.

Results

As a first step in determining whether the human MLL2 and Drosophila CMI proteins were functionally related, we cloned and expressed a portion of the MLL2 protein that carried the conserved PHDf3 region. This PHD domain was predicted to recognize and bind to specifically modified histone residues based

The Journal of Experimental Secondary Science

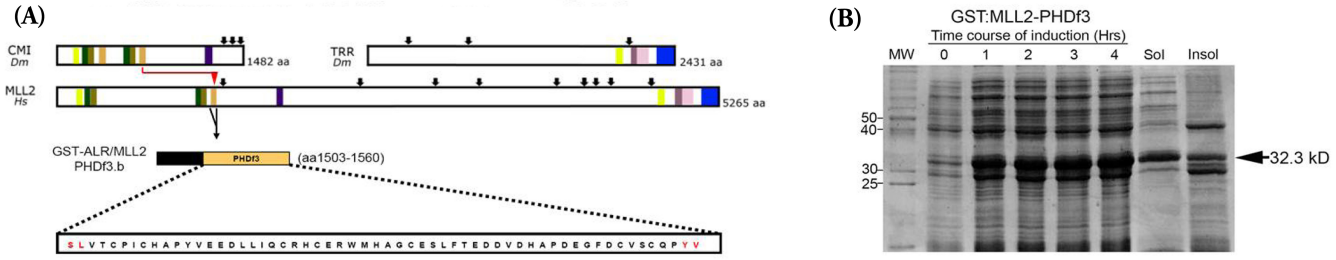

Figure 1. Cloning and expression of the human MLL2 PHDf3. A) Diagram and amino acid sequence of MLL2 PHDf3 (aa1503-1560) in the bacterial expression vector pGEX4T (GST). **B)** Time course of induction/expression of the GST:PHDf3 using IPTG. The GST:PHDf3 protein is 32.3 kilodaltons (kD). The GST:PHDf3 fusion was soluble in this system. **C)** Purification of the GST:PHDf3 using glutathione-agarose resin.

on analyses of similar PHD fingers from unrelated proteins. We found that both the CMI and MLL2 PHDf3 domains were capable of binding to native histone H3 (Figure 2 and Chauhan et al., 2012[7]). Unexpectedly, after running our histone peptide array, we were able to see both PHDf3 proteins bind nonmethylated, monomethylated, and dimethylated H3K4, confirming they had conserved binding properties (Figure 2 and Chauhan et al., 2012[7]). However, similar PHDf3 domains from other proteins, including human MLL1, BPTF, NURF301, ING2/4 and Drosophila LID, were shown to bind to H3K4 carrying three methyl groups (H3K4me3), suggesting our PHD fingers had distinctly different recognitions properties[19].

We next examined the in vivo phenotypic consequences of depleting cmi, as well as its conserved methyltransferase partner trr. We found that conditional depletion of cmi and trr gene levels specifically in the developing wing tissue caused truncations in certain wing veins (Figure 3). Furthermore, several additional "loss of function" phenotypes were noted in the Drosophila, such as held-out wings and no flight capability (Table 1).

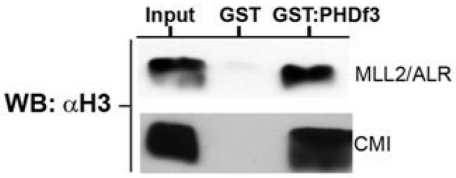

Figure 2. The PHDf3 domains of CMI and MLL2 share conserved histone binding properties. The CMI and MLL2 GST fusion proteins bind to native histone H3. Western blot of GST-pulldown analysis showing that the PHDf3 domains of both CMI and MLL2 are able to bind to histone H3.

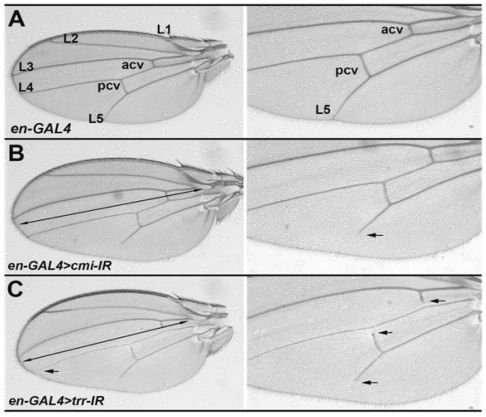

Figure 3. Depletion of CMI and TRR leads to specific vein pattern defects. A) Wild type Drosophila wing. Shown are the positions of the normal veins, including the longitudinal (L1-L5), posterior crossvein (pcv) and anterior crossvein (acv). The photos on the left were obtained at 63X magnification. Panels on the right at 100X views of the same wings. **B)** CMI depleted wing with an incomplete L5 vein (arrow in right panel). The arrow in the left panel represents the boundary between the anterior portion of the wing (top) and posterior portion (bottom). **C)** TRR depleted wing with an incomplete L5 vein, anterior cross wing, and posterior cross vein (arrows in right panels).

Discussion

The focus of this project was how the structures and mechanisms for recognition of specifically modified histones are similar and/or different between the Drosophila CMI and human MLL2/3 PHD3 fingers. We hypothesized both structures and mechanisms would share similar properties due to their highly conserved sequences. As both the CMI and the MLL2/ALR PHDf3 proteins bound to similarly modified H3K4, we can conclude they have comparable histone binding and recognition preferences[7]. Additionally, by observing the vein patterning defects in the fruit flies, we could conclude that in vivo depletion of cmi and trr gene levels had a significantly negative effect on the development of Drosophila, and resulted in a reduction of global histone lysine methylation. Also, the "loss of function" phenotypes add to our conclusion that depleting these genes leads to developmental disorders.

As we know cmi and trr genetically interact in flies[7] and depleting their gene products using RNAi leads to patterning defects and

The Journal of Experimental Secondary Science

Table 1. Drosophila Wing Phenotypes Associated with Loss of CMI/TRR Function

Genotype	Num. of flies	Percent expressing phenotype			
		Held-out wings	*Inc. L5*	*Inc. PCV*	*Inc. L4*
trr-IR;enGal4 (e16E)	402	98.97[1]	88.75[2]	8.18[3]	2.50[4]
Dicer;cmi-IR/enGal4(e16E)	299	99.00	99.67	0	0

[1]Based on 292 flies scored
[2]Based on 320 flies scored
[3]Based on 220 flies scored
[4]Based on 60 flies scored

a global reduction in the trimethylation of H3K4, we may infer cmi must be required for trr to function and activate its target genes. In other words, loss of the PHD finger affects the ability for the methyltransferase to function. This signifies rather than operating through independent parallel pathways to regulate gene expression, they seemingly function through the same hormone dependent pathway. Moreover, if the trimethylation of lysine 4 is normal when cmi is depleted, then cmi would not be required for trr to methylate target genes. However in this case, cmi function is necessary for trimethylation. This leads us to believe that CMI may be required to bring in trr function to the target genes. Therefore, the recognition of and binding to histones is essential in targeting specific genes. The methyltransferase is significant in histone regulation, and the PHD finger is evidently required for it to fulfill its function. Thus, mutations affecting these proteins could have harsh effects on growth and development if the correct genes are not targeted or are misregulated. This is supported by recent findings that several developmental disorders and cancers are associated with loss of MLL2 and MLL3[13,14].

Therefore, the data supports our hypothesis. It is important to note, since the CMI PHDf3 finger is involved in recognition of certain histone tail peptides, and the MLL2 finger has the same recognition pattern, due to the high amino acid sequence conservation between the two, we can extrapolate information we examined from Drosophila developmental phenotypes and connect it to human diseases[7]. Tests can be performed on Drosophila to further investigate the relationship of this PHD finger with the developmental maladies, potentially resulting in an expansion of our knowledge regarding growth disorders.

References

1. Kouzarides, T. (2007). Chromatin modifications and their function. Cell, 128(4), 693-705.

2. Bannister, A. J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Research, 21, 381-395.

3. Li, T., and Kelly, W. G. (2011). A role for set1/MLL-related components in epigenetic regulation of the Caenorhabditis elegans germ line. PLoS Genet, 7(3), e1001349.

4. Eissenberg, J.C. and Shilatifard, A. (2010). Histone H3 lysine 4 (H3K4) methylation in development and differentiation. Developmental Biology, 339(2), 240-249.

5. Mohan, M., Herz, H. M., Smith, E. R., Zhang, Y., Jackson, J., Washburn, M. P., Florens, L., Eissenberg, J. C. , Shilatifard, A. (2011). The COMPASS family of H3K4 methylases in Drosophila. Molecular and Cellular Biology, 31: 4310-4318.

6. Issaeva, I., Zonis, Y., Rozovskaia, T., Orlovsky, K., Croce, C. M., Nakamura, T., … Canaani, E. (2006). Knockdown of ALR (MLL2) reveals ALR target genes and leads to alterations in cell adhesion and growth. Molecular and Cellular Biology, 27(5), 1889-1903.

7. Chauhan, C., Zraly C. B., Parilla, M., Diaz, M. O., Dingwall, A. K. (2012). Histone recognition and nuclear receptor co-activator functions of Drosophila Cara Mitad, a homolog of the N-terminal portion of mammalian MLL2 and MLL3. Development, 139(11), 1997-2008.

8. Sedkov, Y., Cho, E., Petruk, S., Cherbas, L., Smith, S. T., Jones, R. S., Cherbas, P., Canaani, E., Jaynes, J. B., and Mazo, A. (2003). Methylation at lysine 4 of histone H3 in ecdysone-dependent development of Drosophila. Nature, 426(6962), 78-83.

9. Lee, J., Saha, P. K., Yang, Q. H., Lee, S., Park, J. Y., Suh, Y., Lee, S. K., Chan, L., Roeder, R. G., Lee, J. W. (2008). Targeted inactivation of MLL3 histone H3-Lys-4 methyltransferase activity in the mouse reveals vital roles for MLL3 in adipogenesis. Proc Natl Acad Sci, U.S.A., 109(45), 19229-19234.

10. Lee, J., Kim, D. H., Lee, S., Yang, Q. H., Lee, D. K., Lee, S. K., Roeder, R. G., Lee, J. W. (2009). A tumor suppressive coactivator complex of p53 containing ASC-2 and histone H3-lysine-4 methyltransferase MLL3 or its paralogue MLL4. Proc Natl Acad Sci, U.S.A.,106(21), 8513-8518

11. Morin, R. D., Mendez-Lago, M., Mungall, A. J., Goya, R., Mungall, K. L., Corbett, R. D., … Marra, M. A. (2011). Frequent mutation of histone-modifying genes in non-hodgkin lymphoma. Nature, 476, 298-303.

12. Gui, Y., Guo, G., Huang, Y., Hu, X., Tang, A., Gao, S., … Cai, Z. (2011). Frequent mutations of chromatin remodeling genes in transitional cell carcinoma of the bladder. Nature Genetics, (43)9, 875-878.

13. Ng, S. B., Bigham, A.W., Buckingham, K. J., Hannibal, M. C., McMillin, M. J., Gildersleeve, H. I., … Shendure, J. (2010). Exome sequencing identifies MLL2 mutations as a cause of kabuki syndrome. Nature Genetics, (42)9, 790-793.

14. Paulussen, A. D. C., Stegmann, A. P. A., Blok, M. J., Tserpelis, D., Posma-Veltor, C., Detisch, Y., … Schrander-Stumpel, C. T. R. M. (2010). MLL2 mutation spectrum in 45 patients with kabuki syndrome. Human Genome Variation Society, 32, E2018-E2025.

15. Chi, P., Allis, C. D., Wang, G. G. (2010). Covalent histone modifications — miswritten, misinterpreted and mis-erased in human cancers. Nature Reviews Cancer, 10, 457-469.

16. Brand, A. H., Manoukian, A. S., and Perrimon, N. (1994). Ectopic expression in Drosophila. Methods Cell Biol, 44, 635-654.

17. Brand, A. H. and Perrimon, N. (1993). Tageted gene expression as a means of altering cell fates and generating dominant phenotypes. Development, 118, 401-415.

18. St Johnston, D. (2002). The art and design of genetic screens: Drosophila melanogaster. Nature Rev Genet, 3, 176-188.

19. Lallous, N. and Ramón-Maiques, S. (2011). Chromatin recognition protein modules: the PHD finger. eLS. DOI: 10.1002/9780470015902.a0023176

Acknowledgements

I would like to thank the IMSA Student Inquiry and Research Program for allowing me to pursue this opportunity. This work was supported by a grant from the National Science Foundation (MCB1122001).

The Journal of Experimental Secondary Science