The Polymerase Associated Factor Complex-Protein Arginine Methyltransferase Axis: A Mechanism of Acute Leukemia

by

Justin Tyler Serio

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Molecular and Cellular Pathology) in the University of Michigan 2017

Doctoral Committee:

Assistant Professor Andrew G. Muntean, Chair
Collegiate Professor Gregory R. Dressler
Associate Professor Maria E. Figueroa, University of Miami
Associate Professor Ivan P. Maillard
DEDICATION

To all those who pursue research to deepen our understanding of the world and improve how we live, communicate, and create.

To the patients and their families, past, present, and future fighting cancer; may your spirit in the face of hardship inspire us to overcome these diseases.
ACKNOWLEDGEMENTS

First and foremost, I would like to thank my mentor Dr. Andrew Muntean for his continued support during my graduate program. His measured guidance, availability, and willingness to patiently work through setbacks both experimental and developmental have been invaluable. I feel I have grown as a scientist, presenter, and teacher through the mentorship he has provided. This extends to the Muntean laboratory which has been an ideal place to learn and work with many dedicated and collaborative members over my time here. I want to personally thank, James Ropa, Wei Chen, Lili Chen, Maria Mysliwski, and Nirmalya Saha for their help with my training, their various expertise, intellectual discussion, and comradery through the highs and lows of graduate school.

I am also grateful for the support of my thesis committee, Drs. Gregory Dressler, Maria Figueuroa, and Ivan Maillard who have been excellent resources throughout my project while challenging me to broaden my scientific perspective. The MCP program for creating a highly enthusiastic environment among the graduate students and I’m thankful to Drs. Zaneta Nikolovska-Coleska, Nick Lukacs, and our tireless program administrator Laura Labut for all they do for us. To the community of fellow graduate students in biomedical sciences for being an amazing group of talented young scientists who have become dear friends over the years. Your friendship is one of the main reasons I’ve made it through this degree. I know our futures are bright and I can’t wait to see what we will accomplish.

My aspirations to pursue scientific research are undoubtedly due to my high school biology teacher Dr. Felicia Ledogar (Doc L), who was by far the most engaging educator I had the good fortune to learn from. Her creative approach to communicating science inspired me to make research, science communication, and education a priority in my
career and volunteering projects. Additionally, many thanks to the faculty in the Biochemistry Department at Brandeis University for their passion for teaching and encouragement for me to pursue my graduate studies. Special thanks to my program advisor Dr. Daniel Oprian for directing me to consider Michigan’s PhD programs. Also, I’m am grateful to my undergraduate PIs Drs. Gregory Petsko and Dagmar Ringe for providing a proactive opportunity and environment to learn technical skills and spark an interest in pursuing research as a career.

I am forever indebted to my parents Donna and Ralph and my extended family for their unwavering support during my PhD. Whether through providing advice, celebrating milestones or being there to listen, they always help move me forward towards my goals and my personal development in life. They constantly encourage me to open doors and build bridges, which has instilled in me a strong sense of team play and collaboration that I believe is an integral part of making impactful scientific discoveries and improving the world. Most importantly, I would like to thank my incredible fiancée Molly Nathanson who was willing to endure the challenges of being long-distance for over 4 years while we completed our degrees. You help me laugh through my frustrations, center me when I lose focus, and inspire me to be better at everything I do. Your love and friendship are the greatest joys in my life. Thank you for being there for me every step of this journey and those still to come.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEDICATION</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>xi</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>xiv</td>
</tr>
<tr>
<td><strong>Chapter 1: Introduction</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 The Hematopoietic System</td>
<td>1</td>
</tr>
<tr>
<td>1.1.1 Hematopoiesis</td>
<td>1</td>
</tr>
<tr>
<td>1.1.2 Lineage Specific Transcription Factors &amp; Differentiation</td>
<td>3</td>
</tr>
<tr>
<td>1.2 Epigenetic Regulation</td>
<td>6</td>
</tr>
<tr>
<td>1.2.1 The Epigenetic Regulatory Machinery</td>
<td>6</td>
</tr>
<tr>
<td>1.2.2 Epigenetic Antagonism: Polycomb and Trithorax Group Proteins</td>
<td>8</td>
</tr>
<tr>
<td>1.2.3 Epigenetic Writers, Erasers, and Readers</td>
<td>11</td>
</tr>
<tr>
<td>1.3 Transcriptional and Epigenetic Dysregulation in Leukemias</td>
<td>14</td>
</tr>
<tr>
<td>1.3.1 Dysregulation of Hematopoiesis and Chromosomal Translocation</td>
<td>14</td>
</tr>
<tr>
<td>1.3.2 The 11q23 Breakpoint and MLL-Fusion Leukemias</td>
<td>16</td>
</tr>
<tr>
<td>1.3.3 The Consistency of the MLL Breakpoint and the Retention of the N-Terminus</td>
<td>19</td>
</tr>
<tr>
<td>1.4 The Polymerase Associated Factor Complex (PAFc)</td>
<td>20</td>
</tr>
<tr>
<td>1.4.1 The Function of the PAF Complex</td>
<td>20</td>
</tr>
<tr>
<td>1.4.2 The Origins of PAF Complex</td>
<td>23</td>
</tr>
<tr>
<td>1.4.3 The PAF Complex in Cancer</td>
<td>25</td>
</tr>
<tr>
<td>1.5 Summary and Goals</td>
<td>28</td>
</tr>
</tbody>
</table>

**Chapter 2: The Function and Transcriptional Targets of the PAFc in Acute Myeloid Leukemia** 29
## 2.1 Introduction

2.1.1 The MLL-PAFc Interaction 29
2.1.2 Distinct Requirements of the MLL-PAFc Interaction in Normal and Malignant Cells 30

## 2.2 Results

2.2.1 Efficient Excision of Cdc73 in MLL-AF9 Leukemic Cells 31
2.2.2 Loss of Cdc73 Induces Differentiation of MLL-AF9 cells 32
2.2.3 Loss of Cdc73 Alters the Epigenetic Landscape in Leukemic Cells 34
2.2.4 Gene Expression Alterations and Pathway Analysis in Cdc73 Excised MA9 Cells 35
2.2.5 Comparison of Cdc73 Excision and Knockdown Genome Wide Expression 39
2.2.6 Evaluated Genes and Other Potential Downstream Targets 40

## 2.3 Methods

2.3.1 Cell Line Generation and Culture Conditions 42
2.3.2 Genotyping 43
2.3.3 Differentiation, Cell Cycle, and Apoptosis Staining 44
2.3.4 RNA Extraction, RT-PCR, and qPCR 44
2.3.5 Histone Peptide Binding Assay and FLAG Elution 45
2.3.6 Histone Extraction and Western Blotting 46
2.3.7 RNA Sequencing and Analysis Pipeline 46
2.3.8 ChIP-Sequencing and Analysis Pipeline 47

## 2.4 Summary & Discussion

## Chapter 3: Characterization of the PAFc-Prmt Axis in Acute Leukemia

### 3.1 Introduction

3.1.1 The Prmt Family of Epigenetic Modifiers 51
3.1.2 The Relevance of Prmts in Leukemia 54

### 3.2 Results

3.2.1 Modulation of Prmts Alters Growth Potential of MLL-AF9 Leukemic Cells 56
3.2.2 Inducible Knockdown of Prmt5 Disrupts Leukemic Maintenance 58
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.3 Prmt5 is Necessary for AML in vivo</td>
<td>61</td>
</tr>
<tr>
<td>3.2.4 Prmt5 Enzymatic Activity in Leukemia and the PAFc-Prmt5 Axis</td>
<td>64</td>
</tr>
<tr>
<td>3.2.5 Regulation of the Prmt5 Locus</td>
<td>67</td>
</tr>
<tr>
<td>3.2.6 Enzymatic Activity of Prmt5 is Required for Leukemogenesis in vivo</td>
<td>70</td>
</tr>
<tr>
<td><strong>3.3 Methods</strong></td>
<td></td>
</tr>
<tr>
<td>3.3.1 shRNA Generation and Validation</td>
<td>74</td>
</tr>
<tr>
<td>3.3.2 Competition Assay</td>
<td>74</td>
</tr>
<tr>
<td>3.3.3 Colony Formation Assay</td>
<td>75</td>
</tr>
<tr>
<td>3.3.4 Prmt5 Re-expression Cell Line Generation</td>
<td>75</td>
</tr>
<tr>
<td>3.3.5 siRNA Mediated PAFc Knockdown</td>
<td>76</td>
</tr>
<tr>
<td>3.3.6 ChIP</td>
<td>76</td>
</tr>
<tr>
<td>3.3.7 Luciferase Assay</td>
<td>77</td>
</tr>
<tr>
<td>3.3.8 Human Leukemic Cell Line Culture Conditions</td>
<td>77</td>
</tr>
<tr>
<td>3.3.9 Prmt5 Inhibition Assay</td>
<td>77</td>
</tr>
<tr>
<td>3.3.10 In Vivo Leukemic Latency Model</td>
<td>78</td>
</tr>
<tr>
<td>3.3.11 Cytospin and Morphology Assessment</td>
<td>78</td>
</tr>
<tr>
<td><strong>3.4 Summary &amp; Discussion</strong></td>
<td>79</td>
</tr>
<tr>
<td><strong>Chapter 4: Summary, Conclusions, and Future Directions</strong></td>
<td>83</td>
</tr>
<tr>
<td><strong>4.1 Summary</strong></td>
<td>83</td>
</tr>
<tr>
<td><strong>4.2 Conclusions and Future Directions</strong></td>
<td>84</td>
</tr>
<tr>
<td>4.2.1 The PAFc Regulation of Epigenetic Modifiers and Differentiation Pathways</td>
<td>84</td>
</tr>
<tr>
<td>4.2.2 Regulation of the Prmt5 Locus Through Hoxa9 and Stat5</td>
<td>86</td>
</tr>
<tr>
<td>4.2.3 The Role of Id2 Downstream of the PAFc-Prmt5 Axis</td>
<td>90</td>
</tr>
<tr>
<td>4.2.4 Targeting the PAFc in MLL-Fusion Leukemias</td>
<td>93</td>
</tr>
<tr>
<td><strong>APPENDIX</strong></td>
<td>94</td>
</tr>
<tr>
<td><strong>REFERENCES</strong></td>
<td>96</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Lineage Specific Transcription Factors Regulate Blood Differentiation</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>The Gata1-PU.1 Lineage Decision</td>
<td>6</td>
</tr>
<tr>
<td>1.3</td>
<td>Functional Antagonism Between MLL and PRC2 Complexes</td>
<td>9</td>
</tr>
<tr>
<td>1.4</td>
<td>The Mutational Landscape in AML</td>
<td>15</td>
</tr>
<tr>
<td>1.5</td>
<td>MLL-Fusions Lose Their C-Terminal Domains Upon Translocation</td>
<td>17</td>
</tr>
<tr>
<td>1.6</td>
<td>Interactors of the WT MLL and MLL-Fusion Complexes</td>
<td>18</td>
</tr>
<tr>
<td>1.7</td>
<td>Functional Roles of the PAFc</td>
<td>21</td>
</tr>
<tr>
<td>1.8</td>
<td>Model of the PAFc Internal Structure and Interaction With RNAPII</td>
<td>24</td>
</tr>
<tr>
<td>2.1</td>
<td>The MLL-PAFc Interaction Interface</td>
<td>30</td>
</tr>
<tr>
<td>2.2</td>
<td><em>Cdc73</em> Can be Conditionally Excised in MA9 Leukemic Cells</td>
<td>32</td>
</tr>
<tr>
<td>2.3</td>
<td>Excision of <em>Cdc73</em> Induces Differentiation in MA9 Leukemic Cells</td>
<td>33</td>
</tr>
<tr>
<td>2.4</td>
<td>Effects of <em>Cdc73</em> Excision on Leukemic Phenotypes</td>
<td>34</td>
</tr>
<tr>
<td>2.5</td>
<td>The Global Epigenetic Landscape is Altered Upon Cdc73 Excision</td>
<td>35</td>
</tr>
<tr>
<td>2.6</td>
<td>RNA-Sequencing of MA9-<em>Cdc73</em>fl-CreER Leukemic Cells</td>
<td>36</td>
</tr>
<tr>
<td>2.7</td>
<td>GAGE Analysis from RNA-Sequencing of <em>Cdc73</em> Excised Leukemic Cells</td>
<td>37</td>
</tr>
<tr>
<td>2.8</td>
<td>The Prmt Family is Broadly Downregulated Upon <em>Cdc73</em> Excision</td>
<td>39</td>
</tr>
<tr>
<td>2.9</td>
<td><em>Cdc73</em> KO &amp; <em>PAF1</em> KD Datasets Identify Conserved <em>Prmt</em> Target Genes</td>
<td>40</td>
</tr>
<tr>
<td>2.10</td>
<td>PAFc Related Genes From RNA-Sequencing</td>
<td>41</td>
</tr>
<tr>
<td>3.1</td>
<td>The Function and Pathways Utilized by the Prmt Family</td>
<td>53</td>
</tr>
<tr>
<td>3.2</td>
<td>Constitutive Overexpression &amp; Knockdown of <em>Prmt5</em> in MA9 Leukemic Cells</td>
<td>57</td>
</tr>
<tr>
<td>3.3</td>
<td>Inducible <em>Prmt5</em> Knockdown Reduces Fitness of MA9 Leukemic Cells</td>
<td>59</td>
</tr>
<tr>
<td>3.4</td>
<td>Inducible Knockdown of <em>Prmt5</em> Alters Cell Cycle, but not Differentiation</td>
<td>60</td>
</tr>
<tr>
<td>3.5</td>
<td>MA9 Leukemic Cells Demonstrate Selective Pressure for Prmt5</td>
<td>61</td>
</tr>
</tbody>
</table>
Figure 3.6: Knockdown of *Prmt5* Extends Disease Latency in Hoxa9/Meis1 Driven Leukemia

Figure 3.7: MA9 Leukemia Model Shows Selective Pressure to Maintain *Prmt5* Expression

Figure 3.8: There is a Selective Pressure for *Prmt5* in *vivo*

Figure 3.9: The Catalytic Activity of Prmt5 is Required for MA9 Leukemic Cells

Figure 3.10: Evidence of the PAFc-Prmt5 Axis in MA9 Leukemic Cells

Figure 3.11: ChIP-sequencing Dataset Analysis at the *PRMT5* Locus

Figure 3.12: The PAFc is a Transcriptional Regulator of the *Prmt5* Locus

Figure 3.13: Localization of PAFc and MLL to Known PAFc Target Locus *Meis1*

Figure 3.14: Chemical Inhibition of PRMT5 *in vivo* Extends Disease Latency in an MA9 Mouse Model

Figure 3.15: Verification of Leukemia in Prmt5 Inhibitor Latency Assay

Figure 3.16: EPZ015666 Treatment Hinders Human Derived Leukemic Cell Growth

Figure 3.17: PRMT5 Inhibition Offsets MA9 Leukemic Cell Growth *in vitro*

Figure 3.18: Model of Downstream Regulatory Arms of the PAFc in MLL-Fusion Leukemia

Figure 4.1: *Prmt* Expression is not Universally Altered by Cdc73 Excision in Leukemic Cells

Figure 4.2: HOXA9 is Methylated in Leukemic Cells

Figure 4.3: The Hoxa9-Prmt5 Feedback Loop in AML

Figure 4.4: Id2 May be a Key Downstream Target of the PAFc-Prmt5 Axis

Figure 4.5: PAFc-Stat5 Coordination May Regulate *ld* Genes
LIST OF TABLES

Table 1: Classes of Epigenetic Modifiers 7
Table 2: Combination of MLL-AF9 & MLL-ENL Overlap w/ Cdc73 Excision Dataset 38
Table 3: The Prmt Family in Cancer 54
Table 4: Prmt shRNA Sequences 74
Table 5: MLL-AF9 Targets Overlap w/ Cdc73 Excision Dataset 94
Table 6: MLL-ENL Targets Overlap w/ Cdc73 Excision Dataset 95
# LIST OF ABBREVIATIONS

AML – Acute myeloid leukemia  
PAFc – Polymerase associated factor complex  
RBCs – Red blood cells  
HSCs – Hematopoietic stem cells  
AGM – Aorta-gonad-mesonephros region  
EHT – Endothelial to Hematopoietic Transition  
BM – Bone marrow  
MPP – Multipotent progenitor  
CMP – Common myeloid progenitor  
CLP – Common lymphoid progenitor  
GMP – Granulocyte/Macrophage progenitor  
MEP – Megakaryocyte/erythroid progenitor  
SCF – Stem cell factor  
c-Kit – CD117  
HSPCs – Hematopoietic stem and progenitor cells  
iPSCs – Induced pluripotent stem cells  
Runx1 – Runt related transcription factor 1  
Tal1 – T-Cell acute lymphocytic leukemia 1  
LT-HSCs – Long-term hematopoietic stem cells  
PU.1 – Purine rich box 1  
CBP – CREB binding protein  
CEBPα – CCAAT/enhancer binding protein alpha  
PTMs – Post-translational modifications  
Me – Methylation  
Ac – Acetylation  
Ub – Ubiquitylation  
HMTs – Histone methyltransferases  
HATs – Histone acetyltransferases  
HDMs – Histone demethylases  
HDACs – Histone deacetylases  
TrxG – Trithorax group  
PcG – Polycomb group proteins  
Hox – Homeobox  
HP1 – Heterochromatin-associated protein 1  
PRC1/2 – Polycomb repressive complex 1/2  
COMPASS – Complex proteins associated with Set 1  
H3K4 – Histone 3 lysine 4
H3K27 – Histone 3 lysine 27
H2AK119 – Histone 2A lysine 119
SAM – S-adenosyl methionine
SET – Su(var)3-9, enhancer of zeste and trithorax
LSD1 – Lysine specific demethylase 1
UTX – Ubiquitously transcribed tetratricopeptide repeat, X chromosome
JmjD3 – Jumonji domain containing 3
PHD – Plant homeodomain
MOF – Males absent on the first homolog
ASH2L – ASH2-like protein
Wdr5 – WD repeat domain 5
Rbp5 – RB binding protein 5
Hoxa9 – Homeobox a9
Meis1 – Myeloid ecotropic viral integration site 1
MeCom – MDS1 and EVI1 complex locus
JHDM1A – JmjC domain-containing histone demethylation protein 1A
NuRD – Nucleosome remodeling complex
CoREST – REST corepressor 1
Brd4 – Bromodomain containing 4
P-TEFb – Positive transcription elongation factor b
ALL – Acute lymphoid leukemia
CML – chronic myeloid leukemia
CLL – chronic lymphoid leukemia
BCR – Breakpoint cluster region
ABL – Abelson tyrosine-protein kinase 1
Pbx1 – Pre-B-cell leukemia homeobox 1
HLF – Hepatic leukemic factor
KMT2A – Lysine methyltransferase 2A
SEC – Super elongation complex
RNAPII – RNA polymerase II
DSIF – DRRB sensitivity inducing factor
Dot1L – Dot1 like protein
Alu – Arthrobacter luteus
PTD – Partial tandem duplication
Ledgf – Lens epithelium-derived growth factor
CTD – C-terminal domain
NELF – Negative elongation factor
Cdc73 – Cell division cycle 73
Rtf1 – Resolves TBP function 1
H2BK120 – histone H2B lysine 120
FACT – Facilitates chromatin transcription
CstF – Cleave stimulatory factor
CPSF – Cleave and polyadenylation specificity factor
Chd1 – Chromodomain Helicase DNA Binding Protein 1
Ctr9 – Cln Three Requiring 9
Stat3 – Signal Transducer and Activator of Transcripton 3
Paf1 – Polymerase associated factor 1
Leo1 - Left open reading frame 1
Wdr61 – WD repeat containing 61
CSC – Cancer stem cell
Hrpt2 – Hyperparathyroidism 2
HPT-JT – Hyperparathyroidism jaw tumor
Pttn11 – Protein-tyrosine phosphatase non-receptor type 11
Suv39h1 – Suppressor of variegation 3-9 homolog
MA9 – MLL-AF9
GSEA – Gene set enrichment analysis
GAGE - Generally applicable gene set enrichment
Prmt – Protein arginine methyltransferase
MEP50 – Methylosome protein 50
hESC – Human embryonic stem cell
Gfi1 – Growth factor independent 1
ABSTRACT

Transcriptional and epigenetic mechanisms are pivotal to the maintenance of gene programs responsible for cellular homeostasis. Like many cancers, acute myeloid leukemia (AML) is a disease associated with transcriptional and epigenetic dysregulation. Chromosomal translocations involving the 11q23 locus harboring the H3K4 methyltransferase Mixed Lineage Leukemia 1 (MLL1), are found in AML and generate oncogenic fusion proteins with deregulated transcriptional potential. The Polymerase Associated Factor complex (PAFc) is an epigenetic co-activator complex that makes direct contact with MLL fusion proteins and wild type MLL. Disruption of this interaction impedes leukemic progression and diminishes MLL-fusion driven target gene expression. However, the role of the PAFc in AML remains incompletely understood. Conditional excision of the PAFc subunit Cdc73, relieved the differentiation block of AML cells, induced global epigenetic changes, and halted leukemic growth and cell cycling. To mechanistically assess these phenotypes, we explored the transcriptional targets regulated by the PAFc that facilitate leukemia by performing RNA-sequencing after conditional loss of Cdc73. We found Cdc73 promotes expression of an early hematopoietic progenitor gene program that prevents differentiation. Among the target genes, we confirmed the protein arginine methyltransferase (Prmt) family is broadly downregulated upon excision of Cdc73. Phenotypic analysis revealed that Prmt1, Prmt4, and Prmt5 knockdown could hinder leukemic progression in vitro. We determined that Prmt5 is a direct target that is positively regulated by a transcriptional unit that includes the PAFc, MLL1, Hoxa9, and Stat5 in leukemic cells. We observed reduced Prmt5-mediated H4R3me2s following excision of Cdc73 placing this histone modification downstream of the PAFc and revealing an epigenetic axis between the PAFc and Prmt5. Knockdown or pharmacologic inhibition of Prmt5 causes a G1 arrest and reduced proliferation resulting in extended leukemic disease latency in vivo. Notably, we observed a strong selective pressure to maintain Prmt5 in vitro and in vivo.
supporting its role in leukemic progression. Additionally, chemical inhibition of PRMT5 disrupted the proliferation of a range of human derived leukemic cell lines suggesting a broader relevance for a PAFc-Prmt5 axis beyond MLL-fusion AML. Overall, we demonstrate the PAFc-Prmt5 axis facilitates leukemic development and is a potential therapeutic target for AMLs.
CHAPTER 1
Introduction
1.1 The Hematopoietic System

1.1.1 Hematopoiesis
The blood system is an extensive and highly regulated network of specialized cell types that are necessary for many functions within an organism. Some of the activities of differentiated blood cells include the transmission of oxygen across tissues, the repair of vascular damage, and protection from invasive pathogens. Hematopoiesis is the process of blood system generation during embryonic development and is subsequently maintained in the adult through the differentiation cycles of hematopoietic stem cells (HSCs). This process has been studied extensively in model organisms and begins around embryonic day E7.5 in mice undergoing two distinct developmental waves called primitive and definitive hematopoiesis (1). Primitive hematopoiesis occurs in the yolk sac fueling the swift expansion of an embryo through the transient generation of erythroid cells giving rise to red blood cells (RBCs). Additionally, myeloid lineage cell types like macrophages are observed during the primitive wave (2). These primitive macrophages are difficult to distinguish from their definitive counterparts while primitive erythroid cells are marked by the expression of embryonic globin genes and nucleation (3). Definitive hematopoiesis is the second wave of blood development also emerging in the yolk sac around embryonic day E8.5 in mice. Cells produced during this wave go on to populate the blood system post-birth and through adulthood including definitive erythroid progenitors, lymphoid progenitors, and several myeloid lineage cell types (4–6). The definitive phase is also marked by the generation of HSCs beginning on roughly day E10.5 in mice and E32 in humans within the aorta-gonad-mesonephros region (AGM) in a process called the endothelial to hematopoietic transition (EHT) (7–9). While the HSC is observed in many embryonic compartments during development, HSCs in
adult hematopoiesis reside in the bone marrow (BM) for the remainder of an organism’s lifespan (10,11).

The HSC is a pluripotent cell capable of both self-renewal and differentiation into lineage specific progenitors and ultimately specialized blood cells. This pluripotent capacity is evidenced by the ability of a single HSC to repopulate the entire blood system (12–14). HSCs are maintained over the life of an organism and must self-renew to compensate for cells lost to differentiation, egress or apoptosis due to environmental stresses on the hematopoietic system. These self-renewal divisions occur either symmetrically consisting of 2 new HSCs or asymmetrically maintaining 1 HSC and producing a less potent progenitor cell (15). The differentiation process progresses from an HSC to a multipotent progenitor (MPP) then splits into the myeloid or lymphoid lineage with the emergence of either a common myeloid progenitor (CMP) or a common lymphoid progenitor (CLP). CLPs can then differentiate into pre-B and pre T committed cell types ultimately generating B-cells and T-cells respectively. The CMP alternatively, can further specialize into granulocyte/macrophage progenitors (GMPs) or megakaryocyte/erythroid progenitors (MEPs) ultimately producing specialized myeloid cells like monocytes or red blood cells respectively (Figure 1.1) (16). Interestingly, several reports reviewed here (17) observe that subpopulations of HSCs can directly differentiate to MEPs suggesting that progenitor cells like CMPs are not universally required to generate other progenitors like the MEP.

Differentiation of an HSC requires a fundamental change in the gene programs that help maintain its multipotent state. Cells in the hematopoietic compartment express cell surface receptors allowing extracellular communication that can influence differentiation. Signaling ligands like the cytokine stem cell factor (SCF) can bind its receptor CD117 (c-Kit), a marker of hematopoietic stem and progenitor cells (HSPCs), to help maintain their cell state (18,19). These cytokines can be used in vitro to induce differentiation of
induced pluripotent stem cells (iPSCs) along several hematopoietic lineages (20–22). However, once a signal is received, what determines the appropriate gene programs that result in cellular differentiation? To connect these external signals to downstream gene programs and induce differentiation requires the coordinated activity of lineage specific transcription factors.

1.1.2 Lineage Specific Transcription Factors & Differentiation
The various cells of the hematopoietic compartment express lineage specific transcription factors that drive lineage fate decisions (23–26). The expression of lineage specific gene programs that includes lineage specific transcription factors helps to facilitate HSC differentiation or self-renewal.

Transcription Factors and Lineage Fate
Decades of knockout mouse model studies have demonstrated that loss of transcription factors can result in the depletion of specific cell populations within the blood system. These studies revealed that specific transcription factors help govern HSC differentiation into various hematopoietic lineages or the self-renewal of HSCs. A key distinction of mouse knockout models in hematopoiesis is whether they are tested during embryonic development or during steady-state adult hematopoiesis. In the case of development, many lineage transcription factors are essential for the emergence of HSCs. For example, knock out of runt related transcription factor 1 (Runx1) demonstrated it is critical for the EHT and the production of HSCs by day E12.5 (27). Interestingly, conditional knockouts of Runx1 in the adult show no major defects in HSC stability (28). Despite limited effects on HSCs in this context, conditional deletion of Runx1 does perturb megakaryocyte and platelet development albeit does not lead to lethality (29). Similarly, other transcription factors like the helix-loop-helix protein T-cell Acute Lymphocytic Leukemia 1 (Tal1) also known as SCL, is an essential protein in primitive hematopoiesis. Tal1 knockout mice show embryonic lethality by day E9.5 and are devoid of most hematopoietic cell populations from many lineages (30,31). However, like Runx1, Tal1 conditional knockouts in adults appeared dispensable for the self-renewal or repopulation capacity of HSCs (32–34). Other studies demonstrate,
specific subpopulations like quiescent long-term hematopoietic stem cells (LT-HSCs) have high expression of Tal1 and there is a dose dependent effect on HSCs between heterozygotes and knockout mice (34). However, these defects stabilize over time suggesting Tal1 is necessary, but not sufficient for self-renewal of HSCs.

![Figure 1.1: Lineage Specific Transcription Factors Regulate Blood Differentiation](image)

The GATA protein family consists of key transcription factors involved in erythropoiesis, and is shown to coordinate with Runx1 and Tal1 during EHT fate decisions in development (35,36). Gata1 and Gata2 are both embryonic lethal around day E10.5 resulting in major erythroid defects and faulty HSC turnover and maintenance respectively (37–39). Lastly, Purine Rich Box 1 (PU.1) is another transcription factor essential to myelopoiesis and lymphopoiesis. PU.1 knockout mice are also embryonic lethal at roughly day E17.5 showing reduced HSPCs and deficiencies in the myeloid and lymphoid compartments with limited effects on erythroid cells (40–42). Conditional
knockouts in adult mice show defective monocytes and lymphocytes with expansion of the granulocytic lineage. This indicates PU.1 helps direct cells towards CLP progeny and monocytic cells over granulocytic and MEP progeny. These mouse model results place PU.1 and Gata1/2 as important determinants of the CMP lineage fate decision (Figure 1.1). These in vivo models elucidated that lineage fate decisions could be influenced by transcription factors during hematopoiesis.

Transcriptional Control of Lineage Specification

Many lineage specific transcription factors direct fate decisions by functionally antagonizing downstream gene programs. One example of this balance is the transcriptional antagonism of Gata1 and PU.1. All Gata family members can bind the PU.1 differentiation regulator with Gata1 having the highest affinity (43–45). Gata1 is primarily associated with erythropoiesis and binds a conserved WGATAR DNA motif to activate target genes with CREB binding protein (CBP) (46,47). Conversely, PU.1 is associated with myelopoiesis, binds cJun and recognizes its own DNA motif GAGGAA at target loci (48). Normally, equal expression of these lineage factors prevents lineage skewing by both proteins functionally inhibiting the other through binding events. During fate decisions, Gata1 or PU.1 expression elevates restricting the opposing transcription factors gene programs by co-binding these loci and differentiating accordingly (Figure 1.2) (49,50). This co-binding cell fate regulation also occurs between PU.1 and CCAAT/Enhancer Binding Protein Alpha (CEBPα) within the myeloid lineage. High PU.1 protein level can skew myeloid progenitors towards macrophage differentiation while high CEBPα levels induce granulocyte development (Figure 1.1) (51). These transcriptional balances functionally bridge extracellular inputs to gene program activation, which can drive differentiation. As stated, cellular receptors allow the transmission of extracellular signals into changes in gene programs through these lineage transcription factors. This includes regulation of which cell surface markers remain on newly differentiated cells. For example, when an MEP emerges from a CMP, Gata1 can also downregulate the KIT locus which lowers the prevalence of c-Kit cell surface receptor abrogating responsiveness to SCF (52,53). The reverse observation occurs with Tal1 which activates c-Kit preventing progenitor cells from differentiating.
through responsiveness to SCF (54). To execute the complex functions of living organisms such as lineage specification, every cell must coordinate their actions in a temporal and spatial manner. Recent work has demonstrated that epigenetic modifiers can collaborate with lineage transcription factors to help cells regulate and provide a cellular memory of their differentiation gene programs.

![Diagram: The Gata1-PU.1 Lineage Decision](image)

**Figure 1.2: The Gata1-PU.1 Lineage Decision** Schematic depicting the Gata1 transcription factor bound through its WGATAR motif at erythroid target genes with CBP (top-left). Upon myelopoiesis, the binding of PU.1 to Gata1 interrupts erythroid gene programs (bottom-left). PU.1 binds its characteristic GAGGAA motif with cJun to activate myeloid genes (Top-right). During erythropoiesis, GATA1 binds PU.1 to disrupt myeloid gene expression (bottom-right). Modified from Burda et al. Leukemia 2010.

### 1.2 Epigenetic Regulation

#### 1.2.1 The Epigenetic Regulatory Machinery

Epigenetics is an essential layer of transcriptional regulation responsible for some critical functions including access to gene targets and communicating where transcriptional complexes and cofactors should assemble (55,56). Prior to the epigenetic paradigm, heritable changes to the genome were limited to sequence specific variation such as mutation, deletion or recombination (57). However, further study identified that non-sequence related chromosomal changes were also heritable demonstrated by X-inactivation and genomic imprinting (58–60). This regulatory
process acting on the DNA sequence was named epigenetics and helps to organize DNA, modulate gene transcription, and facilitate a cellular memory to maintain differentiation programs and homeostasis (55). Each cell must accommodate the three-billion DNA nucleotide pairs of our genome within their nucleus in the form of 22 pairs of chromosomes and two sex chromosomes. In order to organize DNA and regulate genes, 147 base pairs of DNA is wound around a scaffolding protein octamer known as a nucleosome. These nucleosomes are comprised of 2 sets of 4 histone protein subunits H2A, H2B, H3, and H4 (61). DNA-nucleosome bound structures are known as chromatin and can exist as euchromatin, an open conformation that allows DNA accessibility and gene activation or heterochromatin, a closed conformation restricting access and repressing gene transcription (62). Chemical moieties can be placed on both the DNA itself and histone proteins and are referred to as post-translational modifications (PTMs). The epigenetic machinery comes in three primary flavors, readers, writers, and erasers that bind, add or remove PTMs respectively (Table 1) (63). Major PTMs include methylation (me), acetylation (Ac), ubiquitylation (Ub) and others (64). Despite the role of DNA PTMs in gene regulation and disease, for the purposes of this work we will focus primarily on histone modifications.

<table>
<thead>
<tr>
<th>Epigenetic Type</th>
<th>Enzyme Types</th>
<th>Example Proteins</th>
<th>Targets</th>
<th>Functional Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Readers</td>
<td>Bromodomain, Chromodomain, PHD Finger, WD40 Repeat</td>
<td>Brd4, Chd1, Hp1, Smarc2, Med16, Wdr61</td>
<td>Acetylated, Methylated, Lysines</td>
<td>Complex Stability Chromatin Remodeling Trx Factor Binding</td>
</tr>
<tr>
<td>Writers</td>
<td>HATs, HMTs, DNMTs, Phosphorylases, E2/E3 Ligases</td>
<td>MOZ, p300, GCN5, MLL, Set1, Dot1L, Prmt5, DNMT3a, Rad6/Bre1</td>
<td>H3K4, H3K36, H3K79, H4K16, H4R3, H2BK120 Cytosine, Ser/Thr/Tyr</td>
<td>Transcriptional Activation Transcriptional Repression Protein Degradation Protein-Protein Interaction</td>
</tr>
<tr>
<td>Erasers</td>
<td>HDACs, HDMs, Phosphatases</td>
<td>HDAC1-11, Sirt1-7, Lsd1, Jmjcd2a, Jaird1, Jmjcd6, Utx, Shp2</td>
<td>H3K4, H3K9, H4K16, H4R3, H3K27</td>
<td>Transcriptional Activation Transcriptional Repression Protein-Protein Interaction</td>
</tr>
</tbody>
</table>

Enzymes that place, remove or bind PTMs are known as epigenetic modifying proteins.
and are categorized based on the moiety they affect and how they interact with the modification. The major groups of the histone modifiers are histone methyltransferases (HMTs), histone acetyltransferases (HATs), histone demethylases (HDMs), and histone deacetylases (HDACs) (Table 1). Each group of these enzymes play distinct roles in regulation of the genome and can reinforce or diminish gene activity depending on which site and modifications are involved. A classic example of this behavior between epigenetic regulatory proteins is the opposing functions of the Trithorax (TrxG) and Polycomb group proteins (PcG).

1.2.2 Epigenetic Antagonism: Polycomb and Trithorax Group Proteins
Originally identified through mutational screens in Drosophila, the PcG genes were categorized based on disrupting the development of sex combs (65). Alternatively, mutations in the TrxG genes phenocopy deficiencies in Homeobox (Hox) genes required for proper body plan development including the thoracic region, hence the name trithorax (66,67). The PcG genes are known to be phenotypically related to TrxG genes when these mutants were found to enhance Homeobox (Hox) gene expression patterning (68). The epigenetic role of PcG and TrxG was speculated after Paro and Hogness discovered that a PcG protein had a chromodomain similar to heterochromatin-associated protein1 (HP1) (69). It was later found that many of the Drosophila PcG and TrxG proteins existed in multi-protein complexes such as polycomb repressive complex 1/2 (PRC1/2) and Complex Proteins Associated with Set1 (COMPASS) respectively. Further investigation determined these complexes had methyltransferase activity capable of depositing methyl groups on histone H3 lysine 4 (H3K4me3) by COMPASS in S. cerevisiae and Drosophila (70–72). The PRC1/2 complexes were also found to have methyltransferase activity, placing methyl groups on histone H3 lysine 27 (H3K27me3) and ubiquitin groups on histone H2A lysine 119 (H2AK119Ub) in Drosophila respectively (73).

In the mammalian system, PRC1/2 and COMPASS are important for the modification of H2AK119Ub, H3K27me3, and H3K4me3 and are carried out by subunits of these complexes including Ring2, EZH2, and Set1 or MLL1 respectively (74,75). These HMTs
are epigenetic writers that utilize S-adenosyl methionine (SAM) as a substrate to transfer methyl groups to histone tail target sites (76). Methyltransferase activity is dependent on Su(var)3-9, Enhancer-of-zeste and Trithorax (SET) domains to catalyze these reactions and is found within these epigenetic modifiers. Requirements for these epigenetic modifiers are evidenced by MLL1 deletion in mouse embryos leading to embryonic lethality at day E10.5 (77). Hematopoietic knockouts also are embryonic lethal between days E12.5-E16.5 due to the disruption of definitive hematopoiesis in fetal livers and the AGM (78,79). In adult hematopoiesis, conditional MLL1 null mice exhibit bone marrow failure from HSC exhaustion and are moribund within 21 days post induction (80,81). Similarly, EZH2 deletion leads to embryonic lethality at day E7.5 due to primitive hematopoietic failure. However, in the adult, only defects in lymphopoiesis are observed and may suggest functional differences of EZH2 between these developmental stages (82).

**Figure 1.3: Functional Antagonism Between MLL and PRC2 Complexes** Schematic depicting the MLL complex placement of H3K4me3 at active gene loci (green circles) and the PRC2 complex placement of H3K27me3 to repress gene loci (red circles). These modifications and complexes generally prevent the function of the other.

The MLL1 driven H3K4me3 and EZH2 driven H3K27me3 modifications are functionally relevant in the hematopoietic system to help drive fate decisions. Target genes of the lineage transcription factors are observed to have bivalent modification of H3K4me3 and H3K27me3 in HSCs and MPPs (83,84). These genes remain in a poised state until lineage specification, driven by a lineage transcription factor, resolves bivalency through several mechanisms including HDMs function on the H3K4 or H3K27 histone site. If the
given gene must be inactivated, lysine specific demethylase 1 (LSD1) can demethylate H3K4 while the PRC2 will further compact this locus through deposition of H3K27me3 (85). Conversely, at an activated gene, ubiquitously transcribed tetratricopeptide repeat, X chromosome (UTX) or Jumonji domain containing 3 (Jmd3) will demethylate H3K27 while the MLL or COMPASS complexes further modify the locus with H3K4me3 (86). Generally, H3K4me3 is a characteristic modification associated with gene activation while H3K27me3 is associated with gene repression (Figure 1.3).

The MLL1 H3K4 methyltransferase is a large multi-domain protein containing DNA binding AT hooks, several protein/histone binding Plant homeodomains (PHD), a transactivation domain, a CXXC DNA binding domain, and a catalytic SET domain (72,87–93). MLL1 forms a complex with several epigenetic modifiers and coactivators to carry out its methyltransferase activity. Wild-type (WT) MLL complexes consist of the histone acetyltransferase Males absent on the first homolog (MOF) which catalyzes H4K16Ac, ASH2-Like Protein (ASH2L) an HMT that mono- and di-methylates H3K4, WD Repeat Domain 5 (WDR5) and DYP30 which align H3 for proper epigenetic modification, and RB Binding Protein 5 (RBBP5) which stabilizes all complex members (94–96). Additionally, MLL is known to be recruited to target loci through an interaction with the polymerase associated factor complex (PAFc) to activate its target genes like Homeobox a9 (Hoxa9) and Myeloid Ecotropic Viral Integration Site 1 (Meis1) around its N-terminal CXXC domain (92,97).

MLL1 conditional knockout mice are embryonic lethal at day E10.5 and demonstrate a variety of effects in the hematopoietic compartment including depleted fetal liver size, cellularity, and capacity to grow in vitro or in vivo (98). Adult HSPCs are critically dependent on MLL1 and those lacking this epigenetic modifier cannot regulate cell fate decisions and differentiation properly (80). Mechanistically, this was attributed to MLL1 target genes responsible for self-renewal and quiescent maintenance like Hoxa9 and MDS1 and EVI1 Complex Locus (MeCom) (99). A good example of lineage specific transcription factors cooperating with epigenetic proteins is the MLL1-Runx1 interaction. MLL1 and Runx1 are known interacting partners, allowing the stabilization of the Runx1
protein to avoid degradation by the proteasome (100). The H3K4me3 modification catalyzed by MLL1, further facilitates the Runx1 driven activation of the PU.1 gene at co-bound sites. Regulation of PU.1 expression level is critical for myeloid, lymphoid, and erythroid differentiation as well as HSC maintenance (101). These findings implicate MLL1 as a crucial epigenetic modifier in HSC maintenance and lineage specification. Beyond poised differentiation genes, the other classes of epigenetic modifiers can help promote active or inactive transcriptional states to regulate cell functions like differentiation.

1.2.3 Epigenetic Writers, Erasers, and Readers

As introduced above, epigenetic proteins can be subdivided based on their function of placing modifications, removing modifications or recognizing modifications referred to as writers, erasers, and readers respectively. These categories of proteins act on many histone sites and often demonstrate similar antagonism to the Polycomb and Trithorax complexes to facilitate or impede gene expression.

Writers

Examples of epigenetic writers include Polycomb and Trithorax group proteins described above, which are essential to delineating where active and repressed gene loci are within the genome. The HMTs and HATs can modulate many sites beyond H3K4 and H3K27 such as H3K9, H3K36, H3K79, and H4K16. Each modification correlates with various transcriptional states such as gene elongation for H3K36, H3K79 methylation, and H4K16 acetylation or gene repression for H3K9 methylation (102). Similar to the antagonism of H3K4me3 and H3K27me3, acetylation modifications on H3K9 and H3K27 can prevent these sites from becoming methylated limiting heterochromatin formation and promoting active gene expression (103,104). In fact, the acetyl modification is known to chemically alter the charge distribution on histone sites destabilizing the nucleosome and allowing DNA to be more accessible for transcription (105). More specifically, the H4K16Ac modification is observed to influence the accessibility of genes by preventing chromatin compaction and chromatin remodeler binding through this charge destabilizing mechanism (106). This observation may
explain why acetylation is primarily associated with active loci. Beyond lysine residues, arginines can also be methylated and influence transcriptional states, but will be discussed in detail in Chapter 3. When cells need to reverse the activity at a given locus they must first clear modifications through eraser enzymes.

**Erasers**

Two prevalent types of eraser enzymes within the cell are the HDM and HDAC families. The first eraser enzyme observed was LSD1, which removes methyl modifications on the H3K4 residue revealing that epigenetic modifications were reversible (85). Several years later, further investigation identified JmjC Domain-Containing Histone Demethylation Protein 1A (JHDM1A) another HDM acting on the H3K36 site (86). Many lysine methylation sites such as H3K4, H3K9, H3K27, and H3K36 can all be demethylated except for H3K79 which to date has no known demethylase (107). The enzymatic activity of HDACs act on similar sites like H3K9 and H3K27 and usually correspond with gene repression. Deacetylation is commonly observed in cancers and HDAC inhibitors are of particular interest for therapeutic development with roughly 300 completed clinical trials to date in a variety of cancer types on clinicaltrials.gov (108). Similar to MLL1 and EZH2, the HDM and HDAC enzymes often act in complexes. For example, HDAC1 and HDAC2 which are members of both the Nucleosome Remodeling Deacetylase (NuRD) complex and the REST corepressor 1 (CoREST) complex (109,110). Additionally, LSD1 is known to complex with both NuRD and CoREST to demethylate the H3K4 site repressing genes at certain loci or the H3K9 site leading to gene activation at other targets (111,112). These complexes can then deacetylate and demethylate in tandem illustrating the coordination of epigenetic modifiers to act as molecular switches, shifting gene loci between active and repressed states as needed by the cell. The ability to switch between states at gene loci requires complexes and cofactors to localize properly which involves the function of reader proteins.

**Epigenetic Readers**

Epigenetic reader proteins function by interpreting epigenetic marks placed on chromatin. They are essential to communicate with transcription factor complexes
allowing cofactor recruitment, further epigenetic modification, and ultimately proper gene regulation (113). Reader proteins utilize several protein domains that can bind residues (like histone residues) that contain specific methylation or acetylation patterns. These include the acetylation reading bromodomain, the YEATS domain and the multitude of methylation reading domains such as PHD, Tudor, WD40, and chromodomain (114,115). One of the most well characterized reader proteins is bromodomain-containing 4 (Brd4), which was first identified in mice as a chromosomal binding component (116). Its two conserved bromodomains allow it to bind acetylated lysines on chromatin and are thought to help maintain acetylation levels in the cell (117,118). Highlighting the importance of epigenetics on transcriptional control in differentiation, the Vakoc group has demonstrated that many hematopoietic lineage specific transcription factors interact with epigenetic modifying enzymes to further regulate their specified gene programs (119). The authors highlight the requirement of p300 and Brd4 to activate gene targets of CEBPα and PU.1. Recent evidence also suggests Brd4 can help recruit positive transcription elongation factor b (P-TEFb) to allow transcriptional elongation in certain contexts (120,121). Other transcriptional regulators are known to coordinate with Brd4 including the PAFc, which will be described in detail in section 1.4 (122). Recently, Brd4 has shown promise as a therapeutic target in many cancers including acute myeloid leukemia and is being pushed towards clinical trials (123–125).

Beyond Brd4, many epigenetic enzymes like MLL1, require reader domains to identify and respond to the epigenetic landscape at target loci. MLL1 is shown to recognize H3K4me3 through one of its PHD domains and may allow the complex to propagate this modification near active promoters (126). Similarly, the MLL1 cofactor MOF, is able to read the H3K4me3 modification through its chromodomain identifying active gene loci and utilizing its acetyltransferase activity on the neighboring H4K16 site (127). Brd4 is known to recognize the H4K16Ac modification and is a proposed mechanism of P-TEFb and Brd4 recruitment to active loci further demonstrating that these epigenetic modifiers can coordinate during transcription (128). This demonstrates like MLL1, that perturbation of epigenetic modifiers can greatly influence the steady state of the
hematopoietic system through the action of lineage transcription factors. Many DNA altering events like mutation or amplification of these lineage factors and the epigenetic machinery can permanently dysregulate the blood system generating leukemia.

1.3 Transcriptional and Epigenetic Dysregulation in Leukemias

1.3.1 Dysregulation of Hematopoiesis and Chromosomal Translocation

Leukemias often present with genomic alteration to epigenetic modifiers and lineage transcription factors (129). When an HSPC accumulates mutations, it can transform into a hyperproliferative leukemic cell unable to properly differentiate into specialized cell types due to a differentiation block (130). These phenotypic changes lead to the common symptoms of the disease including anemia from lack of RBCs, vulnerability to infection from lack of lymphocytes, and poor wound healing from lack of platelets. The two major criteria stratifying leukemias are its lineage, and rate of onset (acute or chronic). Typically, chronic leukemias develop more slowly while acute leukemias are rapid with both types culminating in a “blast crisis” or accumulation of blast cells in the blood system. Clinically, a patient is classified as having leukemia if they possess roughly 20-30% of either myeloblasts or lymphoblasts in the bone marrow or peripheral blood (131). Although there are many clinical categories of leukemia, these criteria generally define the acute myeloid (AML), acute lymphoid (ALL), chronic myeloid (CML), and chronic lymphoid (CLL) subtypes.

Interestingly, many lineage specific transcription factors can be dysregulated in leukemias by mutation or chromosomal rearrangement. Runx1 for example, is mutated in 10-20% of adult AML or ALLs leading to improper regulation of the lymphoid lineage gene program and poor clinical outcome (132). Additionally, CEBPa mutants are present in about 10% of AML patients and are thought to skew myeloid progenitors into a preleukemic state while deregulating many pathways including the cell cycle (Figure 1.4) (129,133). Chromosomal translocations also serve as mechanisms to disrupt transcription factor networks and complexes in leukemic cells. These molecular events are caused by double strand breaks in DNA and the swapping of two broken
chromosomal segments with one another. These rearrangements periodically result in two gene segments fusing ultimately producing a fusion protein after translation. Work in the 1960’s by Hungerford and Nowell in CML patients identified an abnormal shortened chromosome subsequently named the Philadelphia chromosome for the location of its discovery (134). This observation spurred investigation into the Philadelphia chromosome’s functional role in CML. Decades of study highlighted by Janet Rowley’s observation that this was a reciprocal exchange of two chromosomal segments,

Figure 1.4: The Mutational Landscape in AML AML presents with a wide range of mutations in many pathways surrounding transcription. These include many factors specific to transcriptional and epigenetic regulation, upstream signaling pathways, chromatin organizers like cohesion and the splicing factors. Chen et al. Nat Gen 2013.

identified the breakpoint cluster region-abelson murine leukemia viral oncogene homolog 1 (BCR-ABL) fusion contained within the Philadelphia chromosome. These reciprocal translocations resulting in BCR-ABL fusions were later found to be the primary driver of a majority of CML and some ALL or AML patients reviewed here (135,136). The presence of these translocations allows clinicians to diagnose and stratify leukemias based on their cytogenetic profile and led to the discovery of other fusion proteins. These include PML-RARα, a fusion protein found in acute promyelocytic
leukemia resulting from translocation of chromosomes 15 and 17, while the NUP98-HOXA9 fusion, from translocations between chromosomes 7 and 11, and the CBFβ-MYH11 fusion, from an inversion within chromosome 16 are found in AML patients (137).

Many translocation events involve hematopoietic lineage specific transcription factors or their complex members. These include the t(8:21) Runx1 or AML1-ETO fusion which is observed in roughly 20% of infant AML and blocks normal myelopoiesis leading to the accumulation of myeloblasts (138). The transcriptional dysregulation of GATA1, CEBPα, and PU.1 by this fusion protein is the underlying cause for blocking proper differentiation (139–141). The Tal1 complex DNA binding component E2A is also involved in translocations, such as t(1:19) or t(17:19) which generate potent fusion proteins (142). Its two most common partners are Pre-B-Cell Leukemia Homeobox 1 (Pbx1) and hepatic leukemic factor (HLF). These proteins are a known cofactor of major leukemia associated genes Hoxa9 and Meis1, and a regulator of Tal1 complex member Lmo2 and pro survival gene Bcl2 respectively. Again, these fusion chimeras disrupt lineage specific differentiation through aberrant Hoxa9/Meis1 or Lmo2/Bcl2 programs respectively and ultimately generate ALL through immortalization of lymphoid progenitors (143). In addition to lineage transcription factors, chromosomal translocations also involve epigenetic regulators including MLL1.

1.3.2 The 11q23 Breakpoint and MLL-Fusion Leukemias
Another chromosomal region involved in frequent translocations was discovered on chromosome 11q23 (144,145). Several studies identified the gene bridging the breakpoint region of this translocation leading to several names including HRX, MLL, ALL-1 or HRX1 (146–149). Now known as MLL1 and Lysine methyltransferase 2A (KMT2A), this gene is fused to one of over 70 potential gene partners observed after translocation as of 2013 (150). Roughly 80% of infant ALL, 10% of adult ALL, and 4% of adult AML contain MLL1 translocations (137,151,152). The MLL-rearranged leukemic subtype maintains one of the worst outcomes in patients and requires a deeper mechanistic understanding to develop effective therapies (153).
Upon an 11q23 translocation, MLL1 is severed at a highly consistent breakpoint region that removes most of the WT gene including its catalytic SET domain and fuses it to another severed gene (Figure 1.5) (146,154). The most common partners are AF9, AF4, and ENL all of which are coactivators of transcription involved with activating complexes like DotCom or the super elongation complex (SEC) (155). Despite losing its catalytic SET domain, MLL-fusions still aberrantly activate transcription of their downstream target genes like Hoxa9 and Meis1. However, some reports detail that WT MLL produced by the remaining wild type allele also localize to targets with MLL-fusion chimeras and compensate for this catalytic deficiency (97,156). Intriguingly, the WT

![Image of Figure 1.5: MLL-Fusions Lose Their C-Terminal Domains Upon Translocation](image)

**Figure 1.5: MLL-Fusions Lose Their C-Terminal Domains Upon Translocation**

A) Schematic demonstrating how fusion genes are generated by a chromosomal translocation. B) Protein diagram of WT MLL (top) and its protein domains. The MLL Fusion protein (bottom) is displayed with its retained protein domains and the most common fusion partners are listed in the post breakpoint region.

MLL1 protein was found to be dispensable in leukemias while family member MLL2,
another H3K4 methyltransferase was essential (157,158). Furthermore, in MLLΔSET knockout mice, no major defects are observed implying the catalytic activity of MLL may not necessarily be essential in both normal and malignant states (159).

Regarding the oncogenic activity of the MLL fusion protein, many common fusion partners are members of two primary transcriptional coactivator complexes the SEC or DotCom (Figure 1.6). The SEC has many subunits including AF4, ELL, AF9, ENL, and P-TEFb, which allow MLL-fusions to stably remain bound at gene targets over-activating their expression (160). P-TEFb is an essential cofactor of the RNA Polymerase II (RNAPII) transcriptional machinery required to allow transition to the elongation phase of transcription. By phosphorylating DRBB Sensitivity Inducing Factor (DSIF), P-TEFb helps release RNAPII from a paused state facilitating active transcription. The second complex that can be recruited by MLL-fusions centers on DOT1-Like Protein (Dot1L), which deposits the H3K79me2 modification, a known marker of gene elongation (161). This complex is also known to utilize P-TEFb for MLL-fusion target gene activation. Some of these cofactors, when not fused to MLL, like ENL and AF9 contain YEATS domains which can read acetylated lysine residues such as H3K9Ac, reinforcing the binding of the fusion complex and propagating active transcription (162–164). Additionally, Dot1L is recruited by the AF9 segment of the MLL-

**Figure 1.6: Interactors of the WT MLL and MLL-Fusion Complexes** WT MLL complex composed of N-terminal subunits LEDGF and Menin and C-terminal subunits MOF, RBBP5, ASH2L, and WDR5 (Left). MLL-Fusion complexes alter the binding partners at target genes by localization of the EAP cofactors (Right). Modified from Slany, R. Haematologica, 2009.
localization was correlated with increased leukemic transformation potential (169). Both enzymatic and protein-protein interaction inhibitors of Dot1L have furthered its role in MLL-fusion AML while clinical trials are ongoing to test its clinical efficacy (ClinicalTrials.gov Identification Number: NCT02141828)(170,171).

1.3.3 The Consistency of the MLL Breakpoint and the Retention of the N-Terminus

Clinical observation shows that a ~8kb region from exon 9 through exon 11 of the MLL gene accounts for 95% of all MLL breakpoints (150). Investigation into the breakpoint cluster region (BCR) identifies topoisomerase II and DNase I hypersensitivity sites as potential inducers of these chromosomal breaks (172,173). Many leukemic patients with chromosomal translocations have these sites present within both genes involved in the subsequent fusion and were thought to explain why breakages commonly occurred within the BCR (174–176). Recent evidence has suggested Endonuclease G is a contributor to MLL-fusion breakage after genotoxic or replication stress (177). Another proposed mechanism for chromosomal breaks is through Arthrobacter luteus (Alu) elements. These sequences are observed within the BCR of the MLL gene and were originally thought to explain why MLL partial tandem duplications (MLL-PTD) and BCR-ABL fusions were generated (178–180). Due to frequency of Alu elements in the genome, the field speculated that this might explain why MLL breakpoints had a variety of gene fusion partners (180). Functional hypotheses as to why breaks in the MLL gene tend to occur within the BCR have focused on certain MLL domains. For example, the addition of WT MLL PHD domains into the MLL-AF9 oncoprotein was sufficient to prevent leukemic transformation (181). This may explain why no breakpoints occur after the PHD region of MLL. Conversely, disruption of the N-terminal CXXC domain was sufficient to hinder MLL-AF9 leukemic progression, while perturbing this domain in other WT MLL family members did not affect their function (93,182).

In addition to the importance of the CXXC domain, three N-terminal cofactors, Menin, Lens Epithelium-Derived Growth Factor (Ledgf), and the PAFc are invariably retained after MLL rearrangement (Figure 1.6). Menin was originally found to complex with MLL to regulate p27 and p18 kinase inhibitors and is considered a tumor suppressor in
endocrine tumors (183,184). Use of sequential deletion constructs and phenotypic analyses demonstrated that Menin binds MLL fusions on the N-terminus, and disruption of this interaction can impede leukemic growth, induce differentiation, reduce Meis1 expression, and stall disease progression (185,186). This paradoxically demonstrates a tumor suppressor being essential for the oncogenic function of a leukemic driver like MLL-fusions. Further investigation by the Cleary group showed Menin facilitated the binding of Ledgf, a coactivator of autoimmune response and HIV replication (187,188). This study demonstrated that Ledgf like Menin, was crucial for transformation and maintenance of MLL-rearranged leukemias (189). For these reasons, both Menin and Ledgf are being evaluated for their therapeutic utility in AML (190–192).

The PAFc is a multi-subunit complex that through interaction with RNAPII, can regulate transcription (193). Our group and others through mass spectrometry binding screens, found the PAFc directly interacts with wild-type MLL and MLL-fusions on the pre-CXXC and RD2 region upstream of the breakpoint cluster region (92,97). An MLL-AF9 deletion construct within the preCXXC-RD2 region could abrogate leukemic transformation by blocking this interaction, limiting the colony formation capacity of bone marrow cells and exhibiting a differentiated cellular morphology. Additionally, combinatorial siRNA knockdown of the PAFc subunits restricted the localization of MLL to primary gene targets like Hoxa9 implicating the PAFc as a necessary cofactor for MLL target gene expression. These findings implicate the PAFc, like Menin and Ledgf, in leukemic transformation and maintenance and further support the requirement for MLL fusions to retain regions surrounding the CXXC domain. Interestingly, these N-terminal interactors of MLL-fusions are known to bind RNAPII or its subunits and may explain the aberrant gene activation at target loci in these cancers (194,195). To better understand the mechanisms of MLL-fusion leukemia, the functional role of the MLL-PAFc interaction necessitates further investigation.

1.4 The Polymerase Associated Factor Complex

1.4.1 The Function of the PAF Complex
The PAFc plays an important role in the initiation, elongation, and termination steps of transcription through several key interactions. Primarily, the association of the PAFc with RNAPII occurs through its C-terminal domain (CTD) and tethers the complex facilitating the recruitment of key transcriptional activators and cofactors. The PAFc associates with the transcriptional start site through the polyA tail at genes undergoing transcription (196–198). Many transcriptional complexes are needed to maintain the PAFc at enhancers and active genes for the duration of this process. The PAFc is initially recruited to paused RNAPII by interacting with DSIF (199). P-TEFb phosphorylates the Spt5 subunit of DSIF, which is thought to direct PAFc recruitment (200,201). This phosphorylation event also ejects the transcriptional regulator Negative Figure 1.7: Functional Roles of the PAFc The PAFc plays key roles in chromatin remodeling, gene elongation, 3' end processing, termination and several developmental pathways. Chd1 is a PAFc interactor and reorganizes chromatin by ejecting nucleosomes allowing the RNAPII machinery access to DNA templates (Bottom Left). Many epigenetic modifiers listed are directly or indirectly reliant on the PAFc to localize to proper gene loci to drive transcription (Top Left). Poly(A) polymerase and endonucleases CSTF and CPSF require the PAFc to prepare mRNA transcripts for termination and cleavage for RNA export respectively (Top Right). The Wnt, Hedgehog, and Notch pathways are connected to the PAFc through transcription factors such as β-Catenin, Gli1 and the Notch-ICD (Bottom right).
Elongation Factor (NELF), allowing RNAPII to transition from a paused state into the elongation phase.

A recent study provides strong evidence that the Cell Division Cycle 73 (Cdc73) subunit interacts directly with RNAPII through its GTPase-like domain while the Resolves TBP Function1 (Rtf1) subunit is responsible for binding phosphorylated DSIF reinforcing RNAPII binding during the end of proximal pausing (193). However, this function of Rtf1 is more uncertain in the mammalian PAFc and is thought to have independent functions (202). Once recruited, the PAFc is responsible for localization and stabilization of many epigenetic modifying enzymes that further promote active transcription (Figure 1.7). Studies in yeast demonstrated that the ubiquitylation on histone H2B lysine 120 (H2BK120Ub), H3K4me3, and H3K79me2 deposition are reliant on the PAFc (203,204). H2BK120Ub, is catalyzed by the E2 and E3 ubiquitin ligases Rad6 and Bre1 and further study revealed that H2BK120Ub corresponds with increased methylation of H3K4, H3K36, and H3K79 deposited by Set1 or MLL1, Set2, and Dot1L respectively (205). The Bre1 subunit is shown to directly interact with the Paf1 and Cdc73 subunits of the PAFc to localize to target sites and carry out its ubiquitylation function in yeast. (206).

As mentioned above, MLL1 is localized through direct interaction with the PAFc while Dot1L is indirectly recruited through interaction with P-TEFb and the SEC or the MLL fusion complex. Furthermore, the Facilitates Chromatin Transcription (FACT) complex's interaction stability with RNAPII is dependent on the PAFc (207). H2BK120Ub was also observed to be essential for the function of the FACT complex through the PAFc (208).

Finally, 3’ end processing complexes such as cleavage stimulation factor (CstF) and cleavage and polyadenylation specificity factor (CPSF) are needed to cleave mRNA from RNAPII and have been shown to require the PAFc (Figure 1.7) (209).

There are other reports linking the PAFc to chromatin remodeling as well. The Rtf1 subunit was shown by Simic et al. to be necessary for Chd1 recruitment allowing proper nucleosome reorganization ahead of a transcribing RNAPII complex (Figure 1.7) (210). Dysfunctional chromatin remodeling by improper Chd1 localization can impede transcription due to the lack of accessible chromatin as RNAPII moves along a locus.
Additionally, the PAFc can bind linker histone H1 and through a cooperative mechanism with Cul4a to ubiquitylate H4K31 promoting gene activation (211). This suggests the PAFc can bridge epigenetic modification to chromatin reorganization and transcriptional activation. Lastly, using the Cln Three Requiring 9 (Ctr9) subunit as a surrogate for the PAFc, Youn et al. demonstrated the IL-6 signaling pathway requires the PAFc for proper Signal transducer and Activator of transcription 3 (Stat3) localization and target gene activation (212). This again demonstrates the PAFc as a bridge between external cell signals and transcriptional activation through transcription factors. Additionally, despite STAT3 having DNA binding capability it still requires the function of the PAFc to activate its gene targets and remain at target loci. Overall, the PAFc is a major player in transcriptional and epigenetic regulation bridging chromatin organization, gene elongation, 3’ end processing, and transcription factor localization and stability.

1.4.2 The Origins of the PAF Complex

Our understanding of transcriptional regulation has grown to include many core components, co-factors, and epigenetic enzymes. To better understand transcription, the Jaehning and Burton groups performed RNAPII binding assays in yeast to identify novel cofactors required for transcription (194). Their strategy utilized the evolutionarily conserved C-terminal repeat sequence YSPTSPS of RNAPII to identify critical binding partners. Two of the interactors of the CTD of RNAPII were a 65kDa protein they named polymerase-associated factor 1 (Paf1) and a 50kDa protein that was identified as Cdc73. These two factors became the first two subunits of the PAFc. Interestingly, these proteins were not observed binding to core RNAPII subunits implying the CTD was the primary interaction site between the PAFc and RNAPII. Further analysis using affinity chromatography and mass spectrometry, showed the PAFc contained 5 proteins including Ctr9, Left Open Reading Frame 1 (Leo1), Rtf1, and in human cells, WD Repeat Containing 61 (Wdr61) also known as Ski8 (213–216).

A critical focus of the field after discovering the PAFc, was determining how the complex formed and what inter-subunit dependencies existed. Extensive work by the Roeder group provided a comprehensive assessment of PAFc member interactions using
human subunits expressed in HeLa cells (217). They observed that Paf1 bound all complex members, Cdc73 bound all subunits except Wrd61, and Ctr9 was critical for Wdr61 complex association. Additionally, yeast PAFc is known to require Cdc73 and Rtf1 to properly bind the CTD of RNAPII or localize to chromatin, but does not prevent complex formation (218,219). Conversely, the Paf1 and Ctr9 subunits are integral to formation of the PAFc in yeast (220). A recent study using electron microscopy and BS3 crosslinking provides new insights into the structure of the PAFc and its interaction with RNAPII. Xu et al. suggests that Ctr9 functions as the core of the PAFc bridging Cdc73 with the Leo1 and Paf1 subunits (Figure 1.8)(193). They continue by confirming Rtf1 interacts with phosphorylated DSIF and observed no interaction with the Cdc73 subunit, which was previously speculated in yeast (219).

**Figure 1.8: Model of the PAFc Internal Structure and Interaction With RNAPII**

A) BH3 crosslinking screen showing the contact points between PAFc subunits. These confirm previously reported interactions shown by Leo1 (blue), Paf1 (pink) and PAFc (orange). Known structural elements like low complexity regions (green), the Rtf1 Plus3 domain (red), the Cdc73 GTPase-like domain (purple), and the Ctr9 Tetratricopeptide repeats (yellow) are also shown. B) Model generated from cryo-electron microscopy showing PAFc (mesh, top) interacting with RNAPII (ribbon, bottom). DNA is shown bound to RNAPII in blue as well as the produced mRNA in red. Several subunits of RNAPII are highlight including the Rpb3 subunit in red shown interacting with the Cdc73 subunit of the PAFc in purple.

Interestingly, the PAFc members are not essential in yeast, but are critical for viability in higher organisms (207,221,222). In the mammalian system, Ctr9 deficient pre-
Implantation embryos from mice were reported to have many developmental defects within the inner cell mass affecting endodermal differentiation, global H3K36 methylation, and pluripotency genes like Oct4 (223). These effects were also observed with Rtf1 deficiency suggesting the PAFc is necessary for early embryonic development. RNAi screens in C. elegans also demonstrate all PAFc subunits are essential for epidermal cell structure during development as well (224). Additionally, germline deletion of Cdc73 in mice are embryonic lethal while conditional deletion mice do not survive beyond embryonic day E6.5 (222). Similarly, Cdc73 conditional knockouts in adult mice, led to cell death and systemic organ failure. Embryonic lethality was also confirmed in Drosophila using knockouts of the Cdc73 homolog Hyrax (225). Past work suggested that the PAFc was only required for the transcription of a subset of yeast genes, but recent evidence argues that the PAFc localizes to a broad range of genetic loci in the human system and may broadly reduce RNAPII transcriptional output (193,226–228). Still, this remains an open question in the mammalian and leukemic systems.

1.4.3 The PAF Complex in Cancer

Beyond its roles in transcription and epigenetic modification, the PAFc and its various subunits have been implicated in a variety of cancer types. Interestingly, there are reports of a role of PAFc subunits in cancer progression with both tumor suppressor and oncogenic observations. It is therefore critical to understand the molecular mechanisms involving the PAFc in specific cell contexts. For example, PAF1 also known as PD2, is overexpressed in pancreatic cancer and the Batra group has demonstrated potential mechanisms this subunit utilizes to promote disease progression. Dey et al. using Panc1 and MiaPaCa pancreatic cancer cell lines, observed that knockdown of Paf1 reduced expression of chromatin remodeler Chd1 and MLL1. Paf1 and Chd1 were found to interact in both the nuclear and cytoplasmic compartment and nuclear localization of Chd1 was dependent on Paf1 (229). This provides evidence that the PAFc or its individual subunits have multiple roles in facilitating chromatin remodeling to allow transcriptional activation. Follow up work showed that Paf1 was highly expressed in cancer stem cell (CSC) populations and co-localized with the characteristic stemness
factors Oct3/4 (230). Lastly, overexpression of Paf1 was observed in human pancreatic tissue and increased the prevalence of CSC’s accelerating oncogenic progression (231).

Germline inactivating mutations in the CTR9 subunit have recently been described to predispose individuals to Wilms tumor through familial mutational profiling (232). Unlike the PAF1 subunit in pancreatic cancer, CTR9 appears to be a tumor suppressor in this context. However, estrogen receptor (ER) positive luminal breast cancer expresses high levels of CTR9 and leads to poorer patient outcomes compared to ER negative patients. Zeng & Xu go on to show that Ctr9 is required to maintain ERα protein levels, estrogen signaling driven gene programs, and can sensitize MCF7 breast cancer cell line to tamoxifen treatment (233). This again elucidates a role for PAFc subunits in connecting signaling receptors like IL-6 and ERα to transcriptional regulation in the nucleus.

The CDC73 subunit coded by the Hyperparathyroidism 2 (HRPT2) gene, is named for its role in hyperparathyroidism-jaw tumors which has 46 unique germline and 33 unique somatic mutations observed across patients to date (234). More than half of these mutations occur in exons 1, 2, and 7 in the N-terminal segment of the gene and the vast majority result in gene truncations from nonsense or frameshift mutations. Intriguingly, Menin mutations are also common in familial hyperparathyroidism-jaw tumors (HPT-JT), and are mutually exclusive with Cdc73 mutations (235,236). This may indicate a functional redundancy of Menin and Cdc73 in this context, whereas in MLL-fusion leukemia they appear to coordinate maintaining the disease. There are many functions of the Cdc73 disrupted by these truncation or missense mutations at highly conserved residues. Parathyroid specific deletion of one Cdc73 allele has been utilized to model parathyroid carcinoma in mice as well (237). Selvarajan et al. speculated that Cdc73 might act as a strong biomarker of breast cancer. They observed high levels of Cdc73 correlated with reduced tumor size, invasiveness, and tumor stage (238).

Despite its tumor suppressor role in HPT-JT and breast cancer, Cdc73 is overexpressed or functions as an oncogene in several cancer types. Takahashi et al.
demonstrated a key role for Cdc73 in gastric cancer through its dephosphorylation by Shp2 phosphatase also known as protein-tyrosine phosphatase non-receptor type 11 (Ptpn11). Unphosphorylated Cdc73 acts a molecular shuttle interacting with β-Catenin in the cytoplasm localizing it to Wnt gene targets. Conversely, phosphorylated Cdc73 favors interaction with suppressor of variegation 3-9 homolog 1 (SUV39H1) and suppressed target genes like cell cycle regulator, cyclin D1 and transcription factor c-myc (239). This finding may indicate Cdc73 is also methylated and recognized by SUV39H1’s chromodomain to facilitate binding and prevent activation of Wnt target genes. Since many epigenetic modifiers have non-histone targets, elucidating the role of this potential methylation may broaden our understanding of Cdc73 and the PAFc in cancer. Recent work by Xiang et al. in hepatocellular carcinomas recapitulated Takahashi’s findings showing a Shp2 induced interaction between Cdc73 and β-Catenin. Similar to gastric cancer, nuclear localization of β-Catenin was driven by Cdc73 inducing the Wnt pathway and promoted hepatocytes to dedifferentiate into CSCs. Additionally, high Shp2 expression was shown to positively correlate with resistance to chemotherapy (240). These studies not only demonstrate the context dependency of the Cdc73 subunit in cancer, but also argue that this may actually operate as a molecular switch facilitated by PTMs.

As a complex, the PAFc is linked with reduced survival in non-small cell lung cancer patients. Zhi et al. shows a correlation between expression of the PAFc and c-Myc a well-known oncogene. They go on to show that c-Myc expression is dependent on the PAFc similar to other cancer types (241). However, there are other reports that the PAFc can repress the c-myc locus. Lin et al. showed Cdc73 or Paf1 knockdown could amplify the expression of c-myc and improve cell growth in several human cell types (242). Again, it appears there is a context dependency with the PAFc and c-myc regulation. Due to the broad role of c-myc in cancer, determining the breadth of mechanisms dependent on the PAFc will greatly reinforce our understanding of its function. Lastly, past work from our group and others have demonstrated the PAFc is essential for MLL-fusion localization and target gene activation in acute leukemias. (92). These data point to context dependent functions for both the PAFc and its components.
to act as oncogenes and tumor suppressors requiring a more thorough understanding of the molecular mechanisms of this transcriptional regulator.

1.5 Summary and Goals

To maintain proper functionality and homeostasis of the hematopoietic system, blood cells must balance self-renewal and differentiation events through transcriptional control of downstream gene programs. The inter-connected network of lineage specific transcription factors governs this system by resolving fate decisions skewing cells to differentiate to the necessary lineage. This includes cooperation with epigenetic modifying enzymes like MLL1, which can help control the transcriptional state of target genes. Dysregulation of the differentiation process results in transformation events generating leukemia and are specifically tied to transcriptional and epigenetic regulatory complexes and their function. The MLL-fusions are generated through chromosomal translocations and remain one of the most aggressive diagnoses among AMLs (153). The PAFc is a six-subunit protein complex able to interact with RNAPII and MLL-fusion proteins to facilitate localization and hyperactivation of target genes promoting leukemia. Currently, our understanding of the PAFc and its role in leukemia remains incomplete. To this end, we aimed to uncover the molecular mechanisms downstream of the PAFc that promote leukemogenesis. Our work described hereafter identified the Prmt family of HMTs as targets of the PAFc in MLL-AF9 driven AML. Further characterization revealed the PAFc directly regulated the Prmt5 locus along with several transcription factors including the MLL-AF9 fusion protein, Stat5, and Hoxa9. Our subsequent analysis of the PAFc-Prmt5 axis demonstrated its relevance in the progression of leukemic disease latency and in several human derived leukemic cell lines. Illustrating how the PAFc regulates transcriptional and epigenetic targets like Prmt5 broadens our understanding of how leukemic transformation occurs, how it is maintained, and how we can best negate it through therapeutic development and intervention.
CHAPTER 2
The Function and Transcriptional Targets of the PAFc in Acute Myeloid Leukemia

2.1 Introduction

2.1.1 The MLL-PAFc Interaction
Epigenetic modifiers are an integral part of gene regulation commonly disrupted in acute leukemias. AML and ALL can present with chromosomal translocations resulting in gene fusions containing epigenetic modifiers such as the MLL1 H3K4 methyltransferase. MLL translocations at the 11q23 locus account for roughly 35-50% of pediatric AML and 80% of infant ALL (152,243). The over-activation of MLL target genes such as Hoxa9 and Meis1 induced by these MLL-rearrangements, promote blocks in differentiation and facilitate hyperproliferation generating leukemia. The interaction of MLL and MLL-fusions with a transcriptional regulator called the PAFc is shown to influence the over-activation of MLL target genes promoting leukemogenesis (92). Since the PAFc functions to recruit epigenetic modifiers such as MLL1 and through the MLL complex cofactor Dot1L, an understanding of the MLL-PAFc interaction could elucidate essential molecular mechanisms in acute leukemia.

Several domains of the MLL protein are critical for its ability to transform cells upon MLL-rearrangement including the Ledgf-Menin interaction region and the CXXC domain (244). Deletion of the CXXC domain was shown to abrogate leukemic transformation, but the underlying mechanism describing its role in leukemogenesis required further characterization (182). To probe for MLL1 protein interactions that utilized the CXXC domain that may be necessary to promote leukemic transformation, Milne et al. and Muntean et al. performed binding screens across the N-terminus of MLL and identified an interaction with the PAFc (92,97). This work determined that the MLL-PAFc interaction occurred in a DNA independent manner and centered on the pre-CXXC and RD2 region of the MLL protein upstream of the breakpoint cluster (Figure 2.1). Interrogation of the PAFc in this interaction revealed that the PAF1 and CTR9 subunits -
bound to this pre CXXC and RD2 region respectively and were essential for the proper activation of MLL target genes like Hoxa9 in both a wild-type MLL and MLL-fusion contexts. The PAFc is observed traversing the Hoxa9 locus and if knocked down, reduces the localization of the MLL protein and subsequent activating modifications like H3K4me3 and H3K79me2. Phenotypically, deletion of the MLL-PAFc interaction surface led to an increase in hematopoietic differentiation, impaired cellular proliferation, and impeded leukemic transformation in vitro. Leukemic latency assays in vivo also demonstrated that deletion constructs of the MLL-PAFc interaction surface could nullify leukemogenesis when injected into sub-lethally irradiated recipient mice (Figure 2.1).

**Figure 2.1: The MLL-PAFc Interaction Interface** The interaction surface of the PAFc and MLL was narrowed to the CXXC region just upstream of the conserved MLL breakpoint cluster (left). In vivo latency models in mice reveal deletion of MLL-PAFc interface can impede leukemic progression compared to MLL-fusion control (right).

2.1.2 Distinct Requirements of the MLL-PAFc Interaction in Normal and Malignant Cells

With the establishment of the MLL-PAFc interaction, our group asked how direct disruption of this binding event affected both leukemic cells and normal bone marrow. To answer this question, an MLL dominant negative fragment (MLL-DN) consisting of the pre-CXXC and part of the CXXC domain of MLL was generated which could stably bind Paf1 like WT MLL (245). Competitive binding assays in vitro show a dose dependent inhibition of Paf1 binding to WT MLL in the presence of the MLL-DN fragment indicating it acts as a dominant negative construct. Overexpression of this fragment impeded the progression of leukemic cells in vitro suggesting the MLL-PAFc
interaction is important for leukemogenesis and could serve as a chemical tool to evaluate this interaction \textit{in vivo}.

Our group wondered if this interaction was also necessary for normal blood cells or specific to leukemic cells. When using our MLL-DN fragment in both non-competitive and competitive reconstitution assays, the disruption of the MLL-PAFc interaction was tolerated \textit{in vivo} and showed no disadvantage to controls. This suggested that disruption of the MLL-PAFc interaction was not essential for normal hematopoiesis while critically hindering leukemic transformation and thus, may be an ideal therapeutic target for AML. Our group also showed that the conditional excision of the Cdc73 subunit of the PAFc is sufficient to reduce the growth of MLL-AF9 (MA9) cells and to a lesser extent, leukemic cells harboring the E2A-HLF fusion, a non-MLL rearranged AML. This suggested the PAFc may play roles in leukemogenesis beyond MLL-fusion driven AMLs.

Drawing on this previous data, we sought to better understand the specific molecular mechanisms downstream of the PAFc that were facilitating these leukemic phenotypes. Several open questions remained that could bridge this gap between PAFc function and leukemic progression. These include what cellular phenotypes are affected by loss of PAFc subunit Cdc73, how does the patterning of epigenetic modification change in this context, and what transcriptional programs are reliant on this complex to facilitate leukemogenesis in MA9 leukemic cells? Below, we aim to answer these questions to broaden our understanding of leukemic transformation and evaluate the PAFc as an essential regulator in MLL-rearranged AML.

\section*{2.2 Results}

\subsection*{2.2.1 Efficient Excision of Cdc73 in MLL-AF9 Leukemic Cells}
To evaluate Cdc73 dependent molecular and cellular phenotypes in leukemic cells, we utilized a \textit{Cdc73}^{fl/fl} mouse to generate MLL-AF9 AML cell lines transduced with a tamoxifen (4OHT) inducible CreER (MA9-\textit{Cdc73}^{flfl}-CreER). We separately generated
control cell lines lacking a floxed Cdc73 allele, but expressing CreER (MA9-CreER). The floxed allele includes loxP sites flanking exon 2. Activation of CreER leads to excision of the second exon of the Hrpt2 gene leading to a premature stop codon and a truncated transcript (222). We confirmed knockout of the gene by PCR and reduction in protein level by western blot (Figure 2.2A-B). By titrating 4OHT, we determined a concentration of 4OHT that would demonstrate effective excision of the Cdc73 allele while displaying no toxic side effects on control MA9-CreER cells. We confirmed that control MA9-CreER cells showed no toxicity inherently or up to a 7.5nM dose of 4OHT, but effectively excised Cdc73 in MA9-Cdc73fl-CreER cells. Using this dosage in proliferation assays MA9-Cdc73fl-CreER cells showed a distinct loss of cell proliferation after Cdc73 was excised while control MA9-CreER cells and EtOH treated samples grew normally. This recapitulated the findings shown previously from our group (245).

**Figure 2.2:** Cdc73 Can be Conditionally Excised in MA9 Leukemic Cells  A) DNA gel of MA9-CreER and MA9-Cdc73fl-CreER leukemic cells treated with 4OHT over time. Directed PCR primers display excised, floxed, and wild-type bands. B) Western blot depicting the quantity of Cdc73 and β-Actin in MA9-CreER and MA9-Cdc73fl-CreER treated with EtOH or 4OHT at 48 hours.

2.2.2 Loss of Cdc73 Induces Differentiation of MLL-AF9 Cells

Previous work from our group showed that disruption of the MA9-PAFc interaction through MLL-AF9 deletion constructs impeded transformation potential of the fusion proteins (92). We sought to explore the cellular phenotypes associated with Cdc73 excision that would help elucidate its role in leukemia. Upon 4OHT treatment, MA9-Cdc73fl-CreER cells displayed morphologic changes consistent with myeloid differentiation including elevated cytoplasmic to nuclear ratios, increased granularity, and reduced c-Kit and increased Cd14 cell surface expression (Figure 2.3B-C). We performed qPCR analysis of genes associated with myeloid differentiation (Mmp8, Mmp12, Id2, Ltf, and Cd80), which showed significant upregulation upon Cdc73
excision (Figure 2.3D) (246,247). From this data, we conclude that Cdc73 and the PAFc are essential factors that maintain differentiation blocks in leukemic cells.

Due to the proliferative defects observed from our previous work in Cdc73 excised cells, we speculated cell cycle or apoptotic pathways may be affected (245). To that end, we measured how our MA9-Cdc73fl-CreER cells cycled after 4OHT treatment. Cell cycle analysis revealed Cdc73 excision results in a G1 phase cell cycle block while changes in apoptosis were mild (Figure 2.4A-B). Next, we investigated the role of Cdc73 and the PAFc in colony forming assays, which reflect leukemic stem cell potential. We pre-treated our MA9-Cdc73fl-CreER cells with 4OHT and measured their growth in methylcellulose media. We observed a significant decrease in the colony formation capacity of MA9 cells lacking Cdc73 by raw counts and visually by INT staining and light microscopy (Figure 2.4C). Based on the above phenotypes, Cdc73 is necessary for leukemic cells to progress through the cell cycle and maintain colony formation capacity.

Figure 2.3: Excision of Cdc73 Induces Differentiation in MA9 Leukemic Cells A) Scheme used to generate MA9-Cdc73fl-Cre ER and MA9-CreER cell lines. B) Morphology changes of MA9 cells at 100x magnification. Scale bars = 20µm. C) Intensity of Cd14 and c-Kit cell surface markers measured by flow cytometry. An isotype is used as a non-specific binding control. D) Expression of a panel of differentiation associated genes measured by qPCR scaled to EtOH control.
2.2.3 Loss of Cdc73 Alters the Epigenetic Landscape in Leukemic Cells

The PAFc is a known epigenetic regulatory complex tasked with facilitating many PTMs on histones such as H2B120Ub, H3K4me3, and H3K79me2 (248,249). We analyzed the global histone landscape in MA9-Cdc73fl-CreER cells treated with EtOH or 4OHT and detected a reduction of H3K4me3, H3K79me2, and an almost complete ablation of H2BK120ub following excision of Cdc73 (Figure 2.5A). This data implied Cdc73 and likely the entire PAFc is necessary to maintain proper placement of these epigenetic marks related to transcriptional elongation as previously described in yeast and other cell types (206,248).

Due to the recent observation that the PAFc localizes to many regions of the genome and its effect on epigenetic modifications, we wondered if there was a preference for
Cdc73 binding specific histone proteins or modifications. Several studies show members of the PAFc are capable of direct chromatin binding including PAF1 through H3R17 methylation, and Leo1 through histone H3. (250,251). Using biotinylated histone tail peptides, we performed immunoprecipitation using lysates from 293 cells and observe Cdc73 can bind histones H2A, H3, and H4 to similar degrees, but displays less affinity for histone H2B (Figure 2.5B). Our observations suggest Cdc73 is a necessary component of the PAFc to facilitate global epigenetic patterning. Since epigenetic modification is closely associated with transcriptional activation and the PAFc is a known transcriptional regulator in this process, reduction of Cdc73 could potentially alter the expression of many genes.

![Figure 2.5: The Global Epigenetic Landscape is Altered Upon Cdc73 Excision](image)

A) MA9-\textit{Cdc73fl-CreER} cells were treated with 7.5nM EtOH or 4OHT for 48 hours. \(\sim2\times10^6\) cells were harvested after 72 hours and subjected to histone extraction then probed with the indicated antibodies. H3 and β-Actin are used as loading controls. B) Biotinylated histone tails were bound to streptavidin beads and immunoprecipitated with lysates from 293 cells. Scramble peptide is used as a binding control and lysate is used as an internal control.

### 2.2.4 Gene Expression Alterations and Pathway Analysis in Cdc73 Excised MA9 Cells

After observing that many phenotypes associated with AML are altered upon Cdc73 excision, we asked what gene programs were controlled by Cdc73 that may contribute to slowing proliferation, halting cell cycle, and blocking the differentiation of MA9 cells. To that end, we performed RNA-sequencing using our MA9-\textit{Cdc73fl-CreER} and MA9-CreER cell lines. To control for expression changes associated with 4OHT treatment or inherent differences between cell lines, we used two controls. First, MA9-\textit{Cdc73fl-CreER} cells were treated with either EtOH or 4OHT for 48 hours. Second, MA9-
Cdc73fl-CreER and MA9-CreER cells were both treated with 4OHT for 48 hours. Plotting expression changes using both experimental procedures resulted in remarkably similar changes and a correlation of 0.98 (Figure 2.6B). This indicates the vast majority of changes in gene expression are not in response to 4OHT or cell line differences, but rather the excision of Cdc73. We observed that only a subset of genes were significantly changed (fold-change > 1.5, FDR q-value < 0.01) and that genes were both upregulated and downregulated upon Cdc73 excision perhaps pointing to a role for the PAFc in both activation and repression (Figure 2.6C) (252).

Figure 2.6: RNA-Sequencing of MA9-Cdc73fl-CreER Leukemic Cells A) Scheme for RNA-Seq experiment setup. B) Scatterplot comparing the Log2 fold changes of MA9-Cdc73fl-CreER, 4OHT treated cells to either our EtOH or CreER control. A linear correlation of 0.98 was calculated using R software. C) MA plot depicting the log2 fold changes between EtOH and 4OHT treated MA9-Cdc73fl-CreER cells (FC > 1.5, q-value < 0.01). Significantly altered genes are shown in red while non-significant genes are shown in black. D) GSEA comparing MA9-Cdc73fl-CreER cells to MA9-CreER control. Green curve depicts either positive or negative correlation with listed gene set and vertical hashes represent individual genes in the set.
Analysis of our data revealed 1835 upregulated and 2151 downregulated genes were significantly changed 48 hours following Cdc73 excision. To determine whether the genes that were affected by loss of Cdc73 are associated with specific gene programs or processes, we performed unbiased pathway analysis. Cdc73 excision corresponds with several gene pathways associated with differentiation such as immune and defense response and osteoclast differentiation showing enrichment in our upregulated gene set. Conversely, we see negative enrichment of methyltransferase activity and transfer of methyl groups associated with the downregulated gene set (Figure 2.7). To explore specific gene programs regulated by the PAFc in leukemic cells we also utilized Gene Set Enrichment Analysis (GSEA). Consistent with the differentiation phenotype described above, excision of Cdc73 increased expression of a gene program associated with myeloid development while control MA9-CreER cells expressed an early hematopoietic progenitor program along with a Hoxa9/Meis1 gene program characteristic of MLL-fusion driven leukemias (Figure 2.6D). Interestingly, 25 of 129 direct MLL-AF9 target genes and 42 of 165 direct MLL-ENL targets showed significant
expression changes upon loss of Cdc73 in our MA9-Cdc73fl-CreER cells (Table 2) (253,254). These data are suggestive of PAFc functions that are both cooperative and independent of MLL fusion proteins. Consistent with the pathway analysis, expression of a methyltransferase activity program was enriched in genes that were downregulated in MA9-Cdc73fl-CreER cells following 4OHT treatment (Figure 2.6D & Figure 2.7). Among the genes altered in the methyltransferase activity program were several members of the Prmt family. Thus, we investigated the fold change of Prmt genes with an RPKM > 5 following excision of Cdc73 and observed that most were downregulated (Figure 2.8A).

To validate our RNA-sequencing results, we tested the expression of several Prmt family members by qPCR using our MA9-Cdc73fl-CreER cells. We determined that upon 4OHT treatment, Prmt1, Prmt4, and Prmt5 are noticeably downregulated by more than two-fold (Figure 2.8B). Together, this data demonstrated that Cdc73 is involved in the maintenance of progenitor cell state, Hoxa9/Meis1 gene programs, and influences the expression of many methyltransferases including the Prmt family.
2.2.5 Comparison of Cdc73 Excision and Paf1 Knockdown Genome Wide Expression

Because we utilized Cdc73 as a surrogate for the PAFc, we aimed to confirm if these expression changes were specific to Cdc73 or dependent on the PAFc. To answer this concern, we compared our mouse Cdc73 excision expression data set with a publicly available knockdown dataset from human MLL-AF9 driven leukemic THP-1 cells targeting another PAFc subunit PAF1 and discovered both unique and overlapping genes from the up and downregulated groups (Figure 2.9A) (228). Interestingly, among the genes that were commonly downregulated in both data sets were several Prmt genes. In fact, PRMT1, PRMT3, and PRMT5 were significantly downregulated following loss of either Cdc73 or PAF1 (q<0.05) (Figure 2.9B). These data implied that not only are these genes most likely regulated by the PAFc versus a CDC73 specific target, but that this regulation is conserved between mouse and human leukemic cells. Since our previous findings only involved leukemic cells, we wondered how Prmt expression compared to normal blood cells. Using murine MA9 leukemic cells and normal murine HSPCs from the bone marrow, we compared the expression of the Prmt family using RNA sequencing. We observed that most Prmts are upregulated in MA9 driven leukemic cells compared to normal bone marrow HSPCs and that Prmt5 showed one of the highest percentage increases in expression (Figure 2.9C). These data demonstrate many genes differ between loss of various PAFc subunits implying they have complex

---

**Figure 2.8: The Prmt Family is Boradly Downregulated Upon Cdc73 Excision**

A) Log$_2$ fold change of Prmt family members with an RPKM >5 from RNA-sequencing. Expression is relative to MA9-Cdc73fl-CreER EtOH treated cells for each gene. B) Validation of Prmt1, Prmt4, and Prmt5 expression decrease after Cdc73 excision by qPCR in MA9-Cdc73fl-CreER cells.
independent functions. However, a cluster of genes is similarly altered including Prmt1, Prmt3, and Prmt5 that this data suggests are conserved PAFc target genes.

### 2.2.6 Evaluated Genes and Other Potential Downstream Targets

Our RNA-seq sequencing generated 4000 altered genes in our upregulated and downregulated datasets. To gain a better understanding of general transcriptional changes from the loss of Cdc73, we investigated several specific genes such as PAFc components, PAFc interactors, Hox genes, MLL targets, and potential downstream targets relevant in leukemia. Upon Cdc73 excision we see no change in Ctr9, Rtf1, and Wdr61 expression, but upregulation of the Paf1 and Leo1 subunits. Rnf20/40, the genes

---

**Figure 2.9: Cdc73 KO & PAF1 KD Datasets Identify Conserved Prmt Target Genes**

A) Venn diagrams depict both overlapping and distinct genes are up and down regulated with excision of Cdc73 and knockdown of PAF1. Comparison of our Cdc73 excision dataset with a PAF1 KD data set in human THP-1 (MLL-AF9) leukemic cells displayed 210 overlapping upregulated genes and 255 overlapping down-regulated genes. B) Table shows Prmt family gene expression changes following knockdown of PAF1 in THP1 cells and excision of Cdc73 in MA9 cells. Blue rows indicate genes, which displayed significant expression changes (q-value < 0.01) in both Cdc73 and PAF1 datasets. C) The Prmt family is over-expressed in murine MA9 leukemic cells relative to normal LinKit+ bone marrow. Cutoffs of FDR < 0.05, Fold Change > 2.

<table>
<thead>
<tr>
<th>Gene IDs</th>
<th>THP-1 cells (human)</th>
<th>MLL-AF9 cells (mouse)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PAF1 KD</td>
<td>Ctrl</td>
</tr>
<tr>
<td>PRMT3</td>
<td>PAF1 KD</td>
<td>40.69</td>
</tr>
<tr>
<td>PRMT1</td>
<td>PAF1 KD</td>
<td>127.85</td>
</tr>
<tr>
<td>PRMT3</td>
<td>PAF1 KD</td>
<td>4.48</td>
</tr>
<tr>
<td>PRMT10</td>
<td>PAF1 KD</td>
<td>2.15</td>
</tr>
<tr>
<td>PRMT6</td>
<td>PAF1 KD</td>
<td>22.59</td>
</tr>
<tr>
<td>PRMT2</td>
<td>PAF1 KD</td>
<td>20.05</td>
</tr>
<tr>
<td>PRMT7</td>
<td>PAF1 KD</td>
<td>17.62</td>
</tr>
<tr>
<td>CARM1</td>
<td>PAF1 KD</td>
<td>28.03</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>RPMK</th>
<th>LKs</th>
<th>MA9</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>150</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

40
encoding for BRE1, and *Supt4a/5*, the genes encoding for the DSIF protein subunits, are upregulated while we see no change in the epigenetic acetylation reader *Brd4*. This implies that our observed gene expression changes are likely directly due to Cdc73 and not due to indirect function of other PAFc subunits or cofactors. Many MLL1 target genes are downregulated including *Eya1, Pbx3, Prdm16, Ruvbl2*, and *Lin28a*.

![Expression changes between 4OHT treated MA9-CreER (black bars) and 4OHT treated MA9-Cdc73fl-CreER (red bars) by RPKM. Highlighted are PAFc subunits and interactors, MLL targets and potentially new PAFc targets.](image)

**Figure 2.10: PAFc Related Genes From RNA-Sequencing** Expression changes between 4OHT treated MA9-CreER (black bars) and 4OHT treated MA9-Cdc73fl-CreER (red bars) by RPKM. Highlighted are PAFc subunits and interactors, MLL targets and potentially new PAFc targets.

Additionally the Hoxa cluster is generally upregulated except for *Hoxa1* and *Hoxa2* upon excision while the leukemia associated genes *Hoxa9* and *Hoxa10*, are not changed significantly. We also observe several potential target genes of interest including *Flt3, Ccnd2, Deptor, Parp2, Creb1*, and *Runx3* in our downregulated cluster (Figure 2.10). This group of potential targets has been implicated in T-ALL, CML, and AML and may provide insights into the mechanisms of the MLL-fusion subtype or regulatory networks.
of the PAFc (255–258). The methyltransferase gene set identified by gene set enrichment analysis (GSEA) and our pathway analysis shows a decrease of PRC2 complex members Ezh2 and Suz12, DNA methyltransferases Dnmt1 and Dnmt3a, and H3K9 methyltransferases Suv39h1 and Suv39h2. Several of these components have previously been reported to bind the PAFc or Cdc73 directly and are relevant in leukemia suggesting they could be part of alternative PAFc regulatory mechanisms in AML (239,259).

2.3 Methods

2.3.1 Cell Line Generation and Culture Conditions
Cell lines were derived from either wild type C57Bl/6 mouse bone marrow (BM), Cdc73fl/fl mouse BM or TetON mouse BM. Mice were subjected to an intraperitoneal injection of 10mg/mL 5-FU and allowed 5 days for hematopoietic expansion. During that time, Plat-E viral packaging cells were split 1:10 into BioCoat plates with DMEM + 10% FBS and no antibiotics. Plates were labeled with the vector being transduced into the mouse BM. Plat-E cells were transfected by mixing 576uL OptiMEM + 18uL FuGene6 reagent incubating for 5 mins then adding 6uL of desired plasmid (1ug/uL). This solution after a 15 minute incubation was added dropwise to Plat-E plates and incubated overnight. The following day the Plat-E media was changed to 6mL IMDM +15% SCFBS +1% P/S. The same day (5 days post 5-FU injection) mice were euthanized and femurs and tibias were extracted, separated from muscle and connective tissue, and placed into PBS after removing bone caps. The bone marrow was flushed using ~4mL of ice cold IMDM +2% SCFBS + 1% P/S per bone. Cells were kept on ice, pipetted to break up debris and subjected to the Easy Step Mouse Hematopoietic Progenitor Cell Enrichment kit. After the isolation is complete cells are resuspended into IMDM + 15% SCFBS + 1% P/S + IL-3 and SCF and incubated overnight. Cells were than virally transduced by combining ~1mL of isolated hematopoietic progenitors, ~6mL of Plat-E media and 1:2000 dilution of Polybrene. The solution was spinoculated by centrifugation at 3200rpm for 90mins. Plat-E media was replaced for repeat spinoculation the following day. After the spin the transduced cells were kept in a 24 well plate and mixed with
IMDM +15% SCFBS +1% P/S and incubated overnight. The following day cells were spinoculated again as stated above and allowed to recover by incubation for 2 days. After this time, cells were subjected to selection in either Puromycin, Neomycin or Hygromycin as needed and slowly weaned off SCF. After about 1-week cells have selected and should be stable in culture. Each cell line is then frozen in multiple vials and stored at -80°C for future use. All leukemic and bone marrow derived cells were cultured in IMDM + 15% Stem Cell FBS and 0.1% IL-3.

2.3.2 Genotyping
Each cell line or mouse needs to be verified to prove it contains the correct genetic manipulation or background. We need to isolate genomic DNA from either a cell line or a live mouse to obtain this information. For cell lines, we harvest ~1 million cells and use the DNAeasy kit to purify gDNA. For mice, we obtain a tail snip and dissolve the tail in 50mM NaOH while boiling at 95°C for 1-2hrs. Once gDNA is obtained we perform PCR using the following amplification strategy with Taq polymerase as described by the manufacturer's instructions.

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>94°C</td>
<td>1.5 mins</td>
</tr>
<tr>
<td>94°C</td>
<td>30 secs</td>
</tr>
<tr>
<td>55°C</td>
<td>30 secs</td>
</tr>
<tr>
<td>68°C</td>
<td>1.5 mins + 3 sec /cycle</td>
</tr>
<tr>
<td>68°C</td>
<td>5 mins</td>
</tr>
<tr>
<td>4°C</td>
<td>Hold</td>
</tr>
</tbody>
</table>

30 cycles

The PCR reactions are run on a 1.5% agarose gel using 8ul of 6X high molecular weight stain at 120V for 30 mins. The bands for the status of Cdc73 in the mice or cells is determined by 3 bands, the WT band at ~271bp, the floxed band at ~304bp and the excised band at ~370bp using the following primers. An Hrpt2 forward primer: 5’-TCCTTTCCATTGTGCAAGCTGGTGTG-3’, an Hrpt2 reverse primer: 5’-TGCCAGTGGAACAACTCATCCTA-3’, and an Hrpt2 floxed primer: 5’-ATTCCAACTGGCTTCCAAGCAG-3’ were used.
**2.3.3 Differentiation, Cell Cycle, and Apoptosis Staining**

MA9-\(\text{Cdc73fl-CreER}\) cells were plated at 250K to 500K cells in 6 well plates and treated with 7.5nM EtOH or 4OHT and grown for 48 hours. Cells were counted and \(\sim 1-2 \times 10^6\) cells per sample were washed with PBS and diluted to the same density in flow buffer (PBS + 2% FBS + 2mM EDTA) for differentiation and apoptosis assays. Cells were incubated in a 96 well plate with either no stain, isotype or stain of c-Kit, Cd14 or Cd11b antibody for 1 hour covered with foil at 4°C or at RT for 15 minutes for the Annexin V-DAPI co-stain for apoptosis experiments. After incubation, cell surface markers were washed 3X and resuspended in 200uL flow buffer and read on an LSRII Flow Cytometer. Cell cycle experiments were washed in PBS and 1mL of cell suspension was combined with 9mL ice cold EtOH dropwise while vortexing and stored in the -20°C for at least 24 hours. Cells were washed with flow buffer and resuspended with RNase A (100ug/mL) in flow buffer and split into non-stain and DAPI stain wells in a 96 V-bottom plate at 200uL each. Cells were incubated in the dark at 37°C for 30 minutes then read on an LSRII Flow Cytometer. Data was analyzed using FlowJo X and ModFit Software for cell cycle.

**2.3.4 RNA Extraction, RT-PCR, and qPCR**

\(\sim 1 \times 10^6\) cells were harvested, washed in PBS and pellets were resuspended in 1mL of Trizol reagent and pipetted until completely dissolved. Samples were incubated at room temperature for 5 mins and 200uL of chloroform was added. Samples were vortexed for 15 seconds incubated at RT for 3 minutes and centrifuged at 12,000g for 15 minutes at 4°C. Transfer upper aqueous phase into fresh 1.5mL Eppendorf tube and add 500uL of isopropanol. Incubate for 10 minutes at RT then centrifuge at 12,000g for 10 minutes at 4°C. If lower cell number was used, 1µL of Glycogen was added to see the RNA pellet. Supernatant was removed and pellet was washed with 75% EtOH followed by centrifuging at 7500g for 5 minutes at 4°C. Pellets were air dried and resuspended in \(\sim 25-50\)uL of Ultrapure DNase/RNase free water. RNA quantity and quality was measured on a nanodrop plate reader in the Bachman Lab by adding 1.5uL per sample. RNA should read out at a 260/280 absorbance ratio of 2.0 while DNA should read out at 1.8. This excludes possibility of gDNA contamination. Once the concentration and purity
is determined we generated cDNA using the SuperScript III First Strand Synthesis Kit following the manufacturer’s instructions.

After the PCR reaction is complete the samples are diluted up to 100uL with Ultrapure DNase/RNase free water and stored in the -20°C freezer. When ready to setup the qPCR, samples are thawed and a master mix is made for each gene being tested on the plate. The master mix consists of SYBR Green Master Mix, forward and reverse primers, and cDNA template. The mixtures are setup in 96 well plates, sealed in optical film, spun down, and run on a 7500 Real Time PCR System plate reader.

2.3.5 Histone Peptide Binding Assay and FLAG Elution
293T cells were transfected with 6uL of either MigR1 empty vector, FLAG-Cdc73 or FLAG-Cdc73-Y>F mutant. Cells were then lysed in lysis buffer (BC-300, 0.1% NP-40, 1:1000 dilution Aprotinin, 1:500 dilution PMSF, 1:100 phosphatase inhibitor, Roche protease inhibitor tablet) for 1 hr rotating at 4°C. Lysate samples were taken from supernatant and M2-Flag beads were pre-washed with TBS 2X. Lysate was added to washed beads and rotated 2 hrs at 4°C. Supernatant was washed with BC-300 + 0.1% NP-40 3X then incubated with 7.5uL Flag-peptide for 15 mins. Solution was spun and supernatant saved for histone binding and 100uL taken as Elutant sample. Next 12uL streptavidin beads were incubated with 1uL histone peptide per sample in BC-300 + 0.1% NP-40 for 2 hrs. After incubation FLAG eluted supernatant was combined with streptavidin-histone bound peptide and rotated for 2 hrs at 4°C. Sample was washed 3X with BC-300 +0.1% NP-40 and then resuspended in 50-100uL of loading buffer per sample. Lysate, eluant and histone samples were run on a 4-15% Mini-Protean TGX gel and run at 150V for ~45 minutes and transferred onto a nitrocellulose membrane in 1X Transfer Buffer +10% MeOH. Transfer was done at 25V for 1.75 hours. Membranes were blocked with 5% milk for 1 hr. Primary αCdc73, αPaf1, αCtr9 and αFlag antibodies were applied to membrane and rocked overnight at 4°C Membranes were washed 3X in TBST, incubated with αRabbit secondary antibody for 1 hr, and washed 3X in TBST. We incubated the membrane with SuperSignal West Pico Chemiluminescent substrate for ~2-4 minutes and visualized the protein content was on a ChemiDoc XRS.
2.3.6 Histone Extraction and Western Blotting

Histone Extractions were performed by resuspending cell pellets in 250uL nuclear extraction buffer (10mM Tris-HCl, 10mM MgCl2, 25mM KCl, 1% Triton X-100, 8.6% Sucrose and 0.5 Roche protease inhibitor tablet) followed by incubation on ice for 10 minutes. Cells were centrifuged at 10,000rpm for 1 min at 4°C. Pellet was resuspended in 100uL 0.4N H2SO4 and rotated at 4°C for 1 hr. Samples were centrifuged at 10,000g for 10 minutes at 4°C. Supernatant was transferred to new 1.5mL Eppendorf tube and 10x volume of ice cold acetone was added. Supernatant was placed at -20°C overnight and centrifuged 10,000g for 10 minutes. Acetone was removed and the pellet was resuspended in 60-80uL of ddH2O. Protein content was calculated using a BCA quantitation kit and samples were diluted to equivalent concentrations in SDS loading buffer + 2-Mercaptoethanol. Samples were loaded into a 4-15% Mini-Protean TGX gel and run at 150V for ~45 minutes and transferred onto a nitrocellulose membrane in 1X Transfer Buffer +10% MeOH. Transfer was done at 25V for 1.75hours. Membranes were blocked with either 5% Milk, 2% BSA or 5% BSA. Antibodies include: αCdc73, αPaf1, αPrmt5, αH3, αH3K4me3, αH3K79me2, αH2BUb, αR-Me2s, αH4R3me2s, αActin, αGAPDH. Antibodies were diluted in 5% Milk, 2% BSA or 5% BSA and added to membranes rocking overnight at 4°C. Membranes were washed with 1X TBST 3X at room temperature. Secondary antibody for αRabbit IgG or αMouse IgG were applied to the membranes for 1 hour and washed again 3X in 1X TBST. Chemi-luminescence signals were detected using SuperSignal West Pico Chemiluminescent substrate for ~2-4 minutes and imaged on a ChemiDoc XRS.

2.3.7 RNA Sequencing and Analysis Pipeline

MA9-Cdc73fl-CreER and MA9-CreER cells were treated with either 7.5nM EtOH or 4OHT in culture and harvested at both 24 and 48 hours, in duplicate. RNA was purified from cells using the Qiagen RNeasy Mini Kit. All samples were processed by the University of Michigan Sequencing Core. Briefly, poly-A RNA was isolated and reverse transcribed to cDNA, which was then end repaired and poly-A tailed. Adapters were ligated and the libraries were sequenced using the Illumina HiSeq Libraries platform.
Raw reads were decoded and converted to fastq files. For previously published datasets, RNA-sequencing raw reads were downloaded from project number SRP048744 in the GEO database and converted to fastq format using the SRA toolkit. For all RNA-sequencing analyses, quality of reads were confirmed using FastQC, and reads were mapped to the reference genome (mm10 for mouse samples and hg38 for human samples) using Bowtie2 and Tophat2.1. Transcripts were assembled, assigned to genes and were quantified using the tool suite Cufflinks2.1. Gene Set Enrichment Analysis (GSEA) and Generally Applicable Gene Set Enrichment (GAGE) were used to search for enriched gene programs and pathway analysis. R3.2.3 and RStudio were used for data framing and graph production.

2.3.8 ChIP Sequencing and Analysis Pipeline

Raw read files from project number SRP048744 were downloaded from the GEO database and were converted to fastq format using the SRA toolkit. Fastq reads were mapped to the reference genome (hg38) using Bowtie2. Mapped reads were filtered for a mapQ value of >10 using Samtools and sorted by chromosome for peak calling. Mapped read files were converted to BedGraph format using Bedtools. Peaks were called using MACS2 using the sorted mapped read files and using the sorted mapped read input file as control. Default parameters for building a model and setting a significance threshold for calling a peak were used. Additional arguments used were: -B --call-summits. Called peaks and read pileups were viewed using the Integrated Genome Browser9.0.

2.4 Summary & Discussion

Our previous work has established the PAFc as a binding partner to MLL-fusions that promotes MLL target gene expression and leukemogenesis by aiding in MLL-fusion protein localization to target genes (92). However, the regulatory mechanisms specifically utilized by the PAFc in leukemic cells remain poorly understood. Using the Cdc73 subunit of the PAFc, we asked what cellular phenotypes and transcriptional programs are dependent on the PAFc in MA9 leukemic cells. To investigate these
questions, we developed leukemic cell lines containing conditional knockout alleles of the \( \textit{Cdc73} \) subunit to evaluate the role of the PAFc in leukemic cells. Our data illustrates the genetic requirement of \( \textit{Cdc73} \) to maintain hyperproliferation, active cell cycling, and importantly, helps maintain a differentiation block characteristic of leukemic cells. Supporting the PAFc’s role as an epigenetic regulator, loss of \( \textit{Cdc73} \) alone is sufficient to alter the global epigenetic landscape reducing many activating histone modifications including H3K4me3 and H2B Ub. Comparing EtOH to 4OHT treated MA9-\( \textit{Cdc73} \)fl-CreER cells, we observe several thousand up and downregulated gene expression changes by RNA-Sequencing. This finding aligned with yeast studies, suggesting that \( \textit{Cdc73} \) and the PAFc are not universally required for transcription, but rather a subset of gene targets (227). The prevalence of both up and downregulated genes in our dataset may also imply the PAFc plays roles in gene repression, but would require further study. Bioinformatic analysis revealed upregulation in developmental and differentiation gene programs, but downregulation in metabolic and methyltransferase gene programs upon \( \textit{Cdc73} \) excision. These results again support \( \textit{Cdc73} \) and the PAFc blocking differentiation in leukemic cells.

Due to the requirement of the MLL-PAFc interaction and the correlation of Hoxa9/Meis1 upregulated gene programs in MA9 leukemic cells, we would expect a significant overlap of MLL and PAFc gene targets. Interestingly, only a small subset of PAFc and MLL target genes overlap and even display variation between MLL fusion partners MLL-AF9 and MLL-ENL (Table 2, Appendix 1 – Table 5 & Table 6). These findings support the PAFc having MLL independent functions that may promote leukemic progression in MLL-rearranged cells. Comparison of PAFc and MLL targets in various fusion contexts may reveal novel transcriptional pathways that could explain why different fusion partners have varied leukemic latencies \textit{in vivo} (260). Further, our RNA sequencing datasets overlapping \( \textit{Cdc73} \) knockout and PAF1 knockdown provide evidence that the PAFc subunits can have complex independent transcriptional targets and function. These transcriptional variations may explain why some subunits can be pro-oncogenic like Paf1 in pancreatic cancer while others appear to act as tumor suppressors like \( \textit{Cdc73} \) in HPT-JT (231,234)
We speculated that excision of Cdc73 might lead to expression effects on PAFc subunits and known cofactors previously observed using deletion strains in yeast which could potentially cloud our results (218,261). However, the PAFc subunit expression is relatively stable along with cofactors Brd4 and epigenetic modifiers like Bre1. These findings suggest the Cdc73 subunit is predominantly responsible for the expression changes from our data set. Further, the expression of Cdc73 seems to have less of an effect on other subunits in the mammalian system compared to observations in yeast.

Within our methyltransferase activity gene list, we also observe many epigenetic modifying proteins associated with gene repression being downregulated upon Cdc73 excision. These include PRC2 complex members EZH2 and Suz12, H3K9 methyltransferase Suv39h1 as well as DNMT3A and DNMT3B. If the PAFc is directly regulating these genes, our data implies there is general de-repression at many gene targets due to the downregulation of PRC2 components and DNMTs. Beyond its activating role with MLL and MLL-fusions at target genes like Hoxa9 and Meis1, this observation supports the PAFc may also regulate repressive complexes and their target genes. Additionally, since gene expression of the complex subunits are stable, the PAFc may still form properly, but may lack the ability to remain localized to target genes without Cdc73 to tether to RNAPII. Directed ChIP at known downstream targets may reveal a disruption of PAFc formation and localization while ChIP-sequencing may demonstrate if most of the complex no longer co-localizes throughout the genome. Further experimentation looking at the protein level and co-binding of PAFc subunits may elucidate the complexes ability to form in MLL-fusion leukemic cells to further define why these cells are susceptible to the absence of Cdc73 and subsequent functionality of the PAFc.

Overall through these analyses, we identified major transcriptional pathways downstream of the PAFc that may direct leukemic progression. We investigated the Prmt family of epigenetic modifiers as potential downstream targets of the PAFc in MLL-
driven leukemias to better understand transcriptional mechanisms of leukemogenesis and broaden the role of the PAFc as an epigenetic regulator.
CHAPTER 3
Characterization of the PAFc-Prmt Axis in Acute Leukemia

3.1 Introduction

3.1.1 The Prmt Family of Epigenetic Modifiers
Our RNA-sequencing data in MA9 leukemic cells after Cdc73 excision yielded thousands of significantly altered genes. Many of these genes are involved in differentiation, epigenetic modification or are commonly associated with leukemia. Among this dataset, we noted that a majority of the Prmt family was significantly down regulated upon Cdc73 excision suggesting the PAFc may be an upstream regulator of these epigenetic modifiers. Additionally, a breadth of cancers commonly present with overexpression of the Prmt genes and in some cases, rely on their enzymatic function to promote oncogenesis (262–264). These findings have led to the development of small molecule inhibitors for these enzymes and their role in leukemic cells is beginning to be defined (265,266). Due to their overexpression in cancer, it is critical to understand how the Prmts are regulated to better design targeted therapies. Many mechanistic questions remain open including what are the transcriptional regulators of the Prmts and what is their specific role in MLL driven leukemic cells? By understanding the role of the Prmts in leukemia, we may gain insight into the molecular mechanisms downstream of the PAFc.

Similar to lysine methylation, arginine methylation plays a critical role in many cellular processes. Arginine methylation was first observed modifying several proteins harvested from human myelin in the 1970’s (267,268). However, it wasn’t until the 1990’s that the first arginine methyltransferase was identified in yeast and given the name protein arginine methyltransferase 1 (Prmt1) (269). The Prmt family can catalyze mono-methylation and two functionally distinct di-methylation modifications called
symmetric dimethylarginine (SDMA) and asymmetric dimethylarginine (ADMA). Despite their similar structure, these two dimethyl modifications have drastically different transcriptional outcomes. ADMA is highly correlated with gene activation while SDMA corresponds with gene repression. However, recent studies have suggested that SDMA can also activate genes in a cell type and context dependent manner (270). Most Prmts place activating asymmetric marks on histone targets while Prmt5 is the predominant symmetric dimethylase acting to repress gene loci (271). Some of the histone sites modified by Prmts include H3 arginine 2 (H3R2) through Prmt1, Prmt5, and Prmt6, H3 arginine 8 (H3R8) through Prmt2 and Prmt5, H3 arginine 17 (H3R17) through Prmt4, and H4 arginine 3 (H4R3) through Prmt1 and Prmt5 (272). Many of these modifications act on similar residues such as H4R3 and H3R2 and help regulate transcriptional states like the antagonism of polycomb and trithorax group enzymes. Additionally, many Prmt proteins have non-histone substrates that influence transcription factors, coactivators, DNA damage proteins, signaling kinases, and the splicing machinery reviewed here (Figure 3.1) (271).

Loss of many Prmt family members in knockout mouse models does not prevent viability and have limited defects post birth that are resolved by adulthood (271). In contrast, Prmt1, Prmt4, and Prmt5 are critical for development and their knockout results in either embryonic lethality or death shortly after birth. This suggests that no compensatory mechanism exists for these enzymes from other Prmt family members. Pawlak et al. demonstrated that Prmt1 activates in the embryo between day E6.5 and E8.5 during the formation of the neural plate. Subsequent crosses of Prmt1 +/- parents, revealed a resorption rate of ~25% marking Prmt1-/- as embryonic lethal. Interestingly, asymmetric arginine activity and prevalence in the cell was reduced by roughly 85% establishing Prmt1 as the primary ADMA enzyme. Follow up work by Yu et al. reconfirmed this embryonic lethality of Prmt1 by day E7.5 with an inducible CreER system (273,274).

Evaluation of Prmt4 +/- embryos by Yadav et al. showed gradual lethality. Starting around E18.5-19.5 ratios of Prmt4 +/- pups begin to die out while no pups reached the weaning
stage. Physically, the null pups show stunted growth and presented with deficient lung inflation if born by caesarian. Strikingly, even Prmt4<sup>+/−</sup> mice show non-mendelian ratios compared to Prmt4<sup>+</sup>+ implying loss of one allele may still be sufficiently detrimental to development shortly after birth (275). Tee et al. evaluated Prmt5<sup>−/−</sup> mice to understand how this enzyme regulated early fate decisions in ES cells. Prmt5<sup>−/−</sup> embryos were not viable past day E6.5 and even Prmt5<sup>−/−</sup> blastocysts harvested at day E3.5 could not form ES cells in culture. The authors go on to show Prmt5 localization around day E6.5 is essential for maintenance of pluripotency gene programs and provides a mechanism explaining the Prmt5<sup>−/−</sup> embryonic lethal phenotype (276). Additionally, hematopoietic specific conditional knockouts of Prmt5 in adult mice induce pancytopenia in the BM due to its catalytic function (277).

![Diagram of Prmt Family Functions](image)

**Figure 3.1: The Function and Pathways Utilized by the Prmt Family** Non-histone targets include developmental and cancer related transcription factors, DNA damage proteins and components of the splicing machinery. Modified from Yang & Bedford Nat Rev Cancer 2013.

The most highly expressed family member Prmt1, has roles ranging from facilitating macrophage differentiation through PPARγ during the immune response as well as regulating cell cycle through Btg1 in kidney cells (278,279). Prmt4 similarly affects a range of pathways including promoting Myb target genes to regulate differentiation in hematopoietic cells and dampening metabolic pathways through Mdh1 in the pancreas (280,281). Several Prmts can also antagonize the methyltransferase activity of MLL1.
One of the PHD domains on MLL1 shows reduced binding affinity to H3K4me3 modifications if H3R2me2a, deposited by Prmt1 and Prmt6 is present (282). Furthermore, through interaction with Cdk4, Prmt5 regulates entry into S phase in hepatocellular and mammary tissue while Prmt1 and Prmt4 show limited effects. (283,284).

3.1.2 The Relevance of Prmts in Leukemia
Within the last decade there has been a noticeable focus on the Prmt family in cancer spanning many tissues and subtypes (Table 3). The literature demonstrates that many Prmt family members are overexpressed in cancers and often correspond with poor patient outcome (Figure 2.9C) (285). Additionally, many of the non-histone targets of the Prmt family members have primary roles in oncogenesis and/or maintenance. These include p53 a commonly mutated tumor suppressor, WNT/β-catenin signaling predominant in gastric and colon cancers, and Hoxa9, a primary component in MLL-driven leukemias (286–288).

### Table 3: The Prmt Family in Cancer

<table>
<thead>
<tr>
<th>Prmt</th>
<th>Type</th>
<th>Mark</th>
<th>Cancer Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prmt1</td>
<td>ADMA</td>
<td>H4R3</td>
<td>AML, Breast, Lung, Colon</td>
</tr>
<tr>
<td>Prmt2</td>
<td>ADMA</td>
<td>H3R8</td>
<td>Breast</td>
</tr>
<tr>
<td>Prmt3</td>
<td>ADMA</td>
<td>Non-Histone</td>
<td>Breast</td>
</tr>
<tr>
<td>Prmt4</td>
<td>ADMA</td>
<td>H3R17, H3R26</td>
<td>AML, Breast, Colon</td>
</tr>
<tr>
<td>Prmt5</td>
<td>SDMA</td>
<td>H3R2, H4R3, H3R8</td>
<td>AML, Lymphoma, Lung, Colon</td>
</tr>
<tr>
<td>Prmt6</td>
<td>ADMA</td>
<td>H3R2, H2AR29</td>
<td>Bladder, Lung</td>
</tr>
<tr>
<td>Prmt7</td>
<td>Monomethylation</td>
<td>H4R3, H2AR3, H3R2</td>
<td>Breast</td>
</tr>
<tr>
<td>Prmt8</td>
<td>ADMA</td>
<td>Unknown</td>
<td>Melanoma, Ovarian</td>
</tr>
<tr>
<td>Prmt9</td>
<td>Unknown</td>
<td>Unknown</td>
<td>No</td>
</tr>
</tbody>
</table>

Several Prmts are described to play roles in leukemias. For example, Prmt1 is most well characterized in AMLs bearing the AML1-ETO fusion where a direct protein-protein
interaction facilitates activation of downstream target genes (289). Additional work shows Prmt1 is essential for MLL-EEN driven leukemias and interacts with this fusion by complexing with Sam68 and CBP (290). Recently, the So group has shown Prmt1 is important for leukemic progression of several MLL-fusions such as MLL-GAS7 and MLL-EEN. They demonstrate that Prmt1 deletion or knockdown in MLL-GAS7 latency models can almost completely abrogate leukemogenesis in vivo and reduces its localization to and activation of MLL target genes Hoxa9 and Meis1 in vitro (263). Prmt4 is relevant in breast cancer where it methylates the SWI/SNF complex subunit BAF155 leading to activation of oncogenes like c-myc and the promotion of metastasis and cell migration (291). In leukemic cells, Prmt4 is shown to regulate myeloid differentiation through methylation of the RUNX1 transcription factor (264). This methylation event can induce RUNX1 to repress a miRNA, miR223, which can feedback to repress Prmt4 expression. This antagonism between Prmt4 through Runx1 and miR223 may reveal a mechanism how the unregulated overexpression of Prmt genes can help maintain leukemic phenotypes. Supporting this claim, the knockdown of Prmt4 in AML1-ETO driven mouse models, was sufficient to double the median leukemic disease latency.

The SDMA Prmt5 enzyme is linked to several solid tumors including prostate, squamous cell carcinoma, and colorectal cancer (292–294). Prmt5 is most well characterized in lymphomas repressing the RB tumor suppressor family and synergizing with gain of function cyclin D1 mutations to promote oncogenesis (295,296). In this context, many known oncogenes are thought to require Prmt5 including the Notch1 ICD, c-Myc, and the MLL-AF9 fusion in AML which appear to show strong selective pressure to re-express Prmt5 during in vivo latency assays. However, it has yet to be shown that reduction of Prmt5 expression in an AML can delay leukemic latency. Furthermore, chemical inhibitors of Prmt5 have shown efficacy on mantle cell lymphoma cell lines and can abrogate tumor formation in mouse xenograft models in vivo (Table 3) (265). In the case of AML, Tarighat et al. has demonstrated that ectopic overexpression of Prmt5 can further amplify leukemic cell growth and quicken leukemic disease latency in vivo (266). They also show that Prmt5 chemical inhibition can induce differentiation and diminish
the expression of Flt3 target genes in human MLL-fusion driven THP-1 and MV4-11 cells.

Work performed concurrently with our study from Dr. Stephen Nimer and colleagues demonstrated that conditional deletion of Prmt5 in MLL-AF9 leukemic cells extended leukemic disease latency in competitive and non-competitive settings (297). They go on to show that inhibition of Prmt5 could induce differentiation, halt leukemic cell growth and extend leukemic disease latency in vivo. Clinically, Prmt5 is gaining traction with several trials underway evaluating its methylation of epidermal growth factor receptor (EGFR) in colorectal cancers after cetuximab treatment and direct inhibition of Prmt5 enzymatic activity in solid tumors and non-Hodgkin’s lymphoma (ClinicalTrials.gov Identifier:NCT02022995 & NCT02783300).

Prmt genes are overexpressed in many cancer types and several studies performed concurrently with our group have established roles for the Prmts in AML. Our RNA-sequencing data demonstrates the expression of many Prmt family members is sensitive to the PAFc through Cdc73 excision. However, it remains unclear if the PAFc directly regulates the Prmt gene family and by what mechanism. For reasons outlined in the results section to follow, we ultimately narrow our focus to Prmt5 as a downstream target regulated by the PAFc. Moving forward we aimed to thoroughly evaluate the role of Prmt5 in MLL-fusion leukemias to understand this regulatory mechanism.

3.2 Results

3.2.1 Modulation of Prmts Alters Growth Potential of MLL-AF9 Leukemic Cells
With the observation that Cdc73 excision broadly downregulates the expression of the Prmt family, we sought to evaluate how the PAFc regulated these targets in MLL-fusion AML. To investigate the role of Prmt genes in leukemia, we narrowed our evaluation to Prmt1, Prmt4, and Prmt5 based on several criteria: 1) the identification of Prmt1 and Prmt5 in both PAF1 and Cdc73 expression analyses (Figure 2.9A-B) 2) genetic deletion of Prmt1, Prmt4, and Prmt5 shows embryonic lethality in mice suggesting non-
overlapping function, and 3) reported roles of all three genes in various cancers including leukemia (Table 3). To understand the role of these Prmts we asked what the effects of modulating their expression had on MA9 leukemic cells. We generated

**Figure 3.2: Constitutive Overexpression & Knockdown of Prmt5 in MA9 Leukemic Cells**

A) Verification of Prmt knockdown using two shRNAs for Prmt1, Prmt4, and Prmt5 in pSM2c vectors. Samples scaled to shRenilla expression level. B) Constitutive Prmt knockdown reduces leukemic growth. Renilla is used as an shRNA control significance taken on Day 4. C) Constitutive knockdown of Prmt4 and Prmt5 reduces colony formation of leukemic cells. Colony plates shown on the right and scale bars = 20μm. D) Overexpression of Prmt5 increases growth potential of leukemic cells. Puro-Neo used as transduction control and MA9-Neo is used as vector control. E) Overexpression of Prmt5 increases colony formation of leukemic cells. All samples compared to MA9-Puro or shRenilla sample using 2-Way ANOVA unless specified otherwise above All * = p-value < 0.05, ** = p-value <0.01, **** = p-value < 0.0001.
constitutively active pSM2C-shRNA vectors targeting these Prmts and validated them in MA9 cells by qPCR (Figure 3.2A). Knockdown of Prmt1, Prmt4, and Prmt5 results in a significant decrease in cell number of MA9 transformed cells compared to shRenilla controls (Figure 3.2B). However, a significant decrease of colony formation was only observed following knockdown of Prmt4 and Prmt5 (Figure 3.2C). We next generated overexpression constructs of our selected Prmts to see if they could synergize with MA9 driven growth and colony formation. Interestingly, only overexpression of Prmt5 significantly increased both proliferation and colony formation in MA9 cells compared to empty vector controls (Figure 3.2D-E). Since knockdown and overexpression of Prmt5 showed significant phenotypic effects and we identified this gene in both PAF1 and Cdc73 expression data sets, we focused on the role of Prmt5 in leukemic cells and its regulation by the PAFc.

3.2.2 Inducible Knockdown of Prmt5 Disrupts Leukemic Maintenance

Beyond cell growth and colony formation, we wondered if Prmt5 knockdown affected other phenotypes including differentiation and global epigenetic patterning. To more precisely assess the function of Prmt5 in leukemic cells, we generated an inducible TetON Prmt5 knockdown system. Upon treating these inducible cells with doxycycline (Dox) we can activate an RFP reporter cassette and our shRNA of interest targeting Prmt5. We observed reduced mRNA and protein levels of Prmt5 following shRNA induction by Dox treatment (Figure 3.3A-B). We performed competition assays using MA9-TetON-shPrmt5 or control MA9-TetON-shRenilla cell lines co-cultured with parental MA9-TetOn cells. Using RFP as a readout for activated shRNA, cells with Prmt5 knockdown were outcompeted by parental MA9-TetON cells while those expressing the shRenilla construct remained stable through the experiment (Figure 3.3C). This illustrates a competitive growth disadvantage in leukemic cells with reduced Prmt5 expression.

Similarly, to our constitutive vectors, growth assays show knockdown of Prmt5 causes a proliferative defect in MA9 cells using this inducible system (Figure 3.4A). Interestingly, c-Kit and Cd14 cell surface expression only mildly change after Prmt5 knockdown
suggesting the differentiation of leukemic cells observed following excision of \textit{Cdc73}, (Figure 2.3B-D) is not facilitated by \textit{Prmt5} (Figure 3.4B). However, reduction of \textit{Prmt5} does lead to G1 cell cycle arrest similarly to \textit{Cdc73} excision suggesting \textit{Prmt5} may promote cell cycle progression in leukemic cells downstream of the PAFc (Figure 3.4C).

Because knockdown of \textit{Prmt5} expression did not completely halt the growth of MA9 cells, we asked whether this might be due to escapees. To this end, we maintained MA9-TetON-\textit{shPrmt5} cells in Dox for 9 days and monitored \textit{Prmt5} protein level

\textbf{Figure 3.3: Inducible \textit{Prmt5} Knockdown Reduces Fitness of MA9 Leukemic Cells}  

\textbf{A)} Verification of \textit{Prmt} knockdown using two shRNAs for \textit{Prmt1}, \textit{Prmt4}, and \textit{Prmt5} in inducible pTRMPV vectors. Samples were measured at 48 hrs after 1μg/mL Dox treatment. Samples were scaled to β-Actin and normalized to vehicle treated cells.  

\textbf{B)} Verification of \textit{Prmt} knockdown after Dox treatment by Western blot. β-Actin is used as a loading control. C) Loss of \textit{Prmt5} reduces cell fitness. Cells were plated in duplicate in a 3:1 ratio of parent MA9-TetON cells to MA9-TetON sh\textit{Prmt5} knockdown cells and treated as in (A). RFP signal was read every two days by flow cytometry. D) Gating of RFP readout is shown at Day 2 and Day 12 for sh\textit{Renilla} and sh\textit{Prmt5}.
western blot. Prmt5 protein level does recover after initial knockdown suggesting a selective pressure to maintain Prmt5 expression in leukemic cells (Figure 3.5).

**Figure 3.4: Inducible Knockdown of Prmt5 Alters Cell Cycle, but not Differentiation**

A) Prmt5 knockdown reduces cell proliferation. Cells were treated with 1 μg/mL Dox once per day and grown for 4 days. ShRenilla is used as a control. **** = p-value < 0.0001 by Two-way ANOVA. B) Prmt5 knockdown has minimal effect on differentiation. Cells were treated as in (A) and harvested at 48 hours, incubated with c-Kit or Cd14 antibody and analyzed by flow cytometry. C) Loss of Prmt5 leads to G1 cell cycle arrest. Cells were treated as in (A) and stained with DAPI and analyzed by flow cytometry. * = p-value < 0.05, *** = p-value < 0.001 by two-tailed student's t-test.
Prmt5 is Necessary for AML in vivo

With our phenotypic data demonstrating a role for Prmt5 in MA9 leukemic cells, we next asked how loss of Prmt5 affects AML disease latency in vivo. To examine the role of Prmt5 in a broader group of AML, we utilized our inducible knockdown system in leukemic cells driven by overexpression of MLL-fusion targets Hoxa9 and Meis1, which are overexpressed in ~50% of AML (298). Here, we transduced primary TetON Hoxa9/Meis1 leukemic cells with TRMPV-shPrmt5 vectors and injected them into syngeneic recipient mice followed by knock down of Prmt5 with Dox treatment. We verified that an aliquot of cells from our injection pool showed activated shRNA and knocked down Prmt5 prior to injection (Figure 3.6A). Dox mediated Prmt5 knockdown significantly extended disease latency by 4.5 days compared to our non-Dox treated group (Figure 3.6B). Leukemic cell infiltration in the liver and spleen was comparable between both groups as was the spleen weight at the moribund state (Figure 3.6C-D).

We repeated this assay using MA9-TetON cells. A pool of cells taken prior to injection confirmed knockdown of Prmt5 as expected (Figure 3.7A). Oddly, we see no statistical latency extension between our treatment groups (Figure 3.7B). We observe a significant reduction in spleen weight in the Dox treated group at the moribund state, but comparable levels of leukemic tissue infiltration, and Prmt5 expression (Figure 3.7C-D). The long disease latency in the MLL-AF9 model (Median of 48 days), may have allowed these cells to re-express Prmt5 and mask a latency extension phenotype.
To that end we wondered if there were escapees *in vivo* like we observed *in vitro*. Both the MA9 cells and Hoxa9/Meis1 cells are GFP+ allowing us to read out the prevalence of leukemic cells in the bone marrow. Looking at our Hoxa9/Meis1 overexpressed leukemic cells, we observe similar levels of...
GFP+ in both treatment groups, confirming comparable leukemic burden in the bone marrow (Figure 3.8A). Further, the Prmt5 expression at the moribund state is also similar suggesting a selective pressure to re-express Prmt5 (Figure 3.8B). Similarly, MA9 leukemic cells display comparable levels of both GFP+ and Prmt5 expression at the moribund state between treatment groups (Figure 3.8C-D). These data indicate

Figure 3.7: MA9 Leukemia Model Shows Selective Pressure to Maintain Prmt5 Expression A) Primary MA9-TetON-shPrmt5 cells treated with 1μg/mL Dox for 48 hours prior to injection. Prmt5 expression measured by qPCR to confirm shRNA. RFP images shown on the right depict active shRNA. Scale bars = 20μm B) Survival curve of Dox- and Dox+ groups. Dox treatment given in food (625 mg/kg) and water (2mg/mL) from day 4 until day 56. N=7 per group. C) Spleen weights of both groups at moribund state. * = p-value< 0.05 by student’s t-test. Images shown below E) Tissue sections of liver (40X) and spleen (20X) at moribund state for both groups. Scale bars = 50μm for liver and 100μM for spleen
Prmt5 plays a role in the leukemic progression of AMLs displaying overexpression of the Hoxa9/Meis1 gene program such as MLL-fusion leukemias.

3.2.4 Prmt5 Enzymatic Activity in Leukemia and the PAFc-Prmt5 Axis

We next asked whether the enzymatic activity of Prmt5 was required for the effects on leukemic progression. We transduced our MA9-TetOn-shPrmt5 cells with either wild-type PRMT5 (WT-P5) or a catalytic dead PRMT5 mutant (CD-P5) as described (299), and subjected them to growth assays. While the growth defect in MA9-TetON-shPrmt5 cells after Prmt5 knockdown is rescued by WT-P5 expression, the CD-P5 mutant did not; suggesting the enzymatic activity of Prmt5 is necessary for MLL-AF9 leukemic cell growth (Figure 3.9A). Given the selective pressure for Prmt5 and its ability to amplify leukemic growth when overexpressed, we wondered if overexpression of Prmt5 could rescue the growth defects in MA9 leukemic cells after excision of Cdc73. 

![Figure 3.8: There is a Selective Pressure for Prmt5 in vivo](image)

**A)** Leukemic prevalence in the BM at moribund state of Hoxa9/Meis1 latency model between non-transduced (NT), shPrmt5, and shPrmt5+Dox groups. Percentage of GFP was measured by flow cytometry and compared using student’s t-test. **B)** Prmt5 expression level of both treatment groups is not significantly changed at moribund state using student’s t-test. **C)** GFP+ leukemic cells prevalence in the BM of mice at moribund state for non-Dox and Dox+ groups. **D)** Expression level of Prmt5 of non-Dox and Dox+ groups at moribund state using student’s t-test.
To that end we transduced MA9-Cdc73fl-CreER cells with our WT-Prmt5 and CD-Prmt5 constructs to generate stable cell lines. Upon 4OHT treatment, we observe that our empty vector control shows halted cell growth as expected however, overexpression of WT-Prmt5 or CD-Prmt5 are unable to rescue excision of Cdc73 (Figure 3.9B). This result suggests Prmt5 is only one downstream PAFc target contributing to leukemic progression.

To directly assess the enzymatic output of Prmt5 we investigated the global histone modification patterning in MA9 cells following knockdown. Western blot analysis
revealed global reduction of pan SDMA and H4R3me2s following knockdown of Prmt5 as expected with little change on H3K4me3, H3K79me2, and H2Bub (Figure 3.10A). We similarly detected reduced H4R3me2s and pan SDMA after Cdc73 excision consistent with a PAFc-Prmt5 axis (Figure 3.10B). This effect was specific since H4R3 asymmetric methylation (H4R3me2a) catalyzed by Prmt1 and Prmt6 remained...
unchanged (Figure 3.10C). Additionally, knockdown of the PAF1 subunit in HeLa cells also reduces Prmt5 protein level consistent with the PAFc regulating Prmt5 expression (Figure 3.10D). This data illustrates two primary findings. 1) that Prmt5 is functionally downstream of the PAFc in MA9 leukemic cells and 2) that MLL-AF9 leukemic cells require the PRMT5 protein, and its enzymatic activity to maintain and promote oncogenesis in vitro.

3.2.5 Regulation of the Prmt5 Locus

In yeast, the PAFc has been shown to regulate the transcription of about 20% of the genome (227). Similarly, in human cells, reduced expression of PAF1 led to only modest changes in gene expression suggesting the PAFc governs specific gene programs. The Prmt family including Prmt5 was identified as a PAFc sensitive target that we now show phenocopies many Cdc73 excision effects on MA9 leukemic cells. However, the exact transcriptional regulation of Prmt5 by the PAFc was still unknown. To investigate whether Prmt5 is directly regulated by the PAFc we analyzed ChIP-sequencing data from the GEO database, targeting several PAFc components generated from the human MLL-AF9 driven AML cell line, THP-1 (228). We observed distinct called binding peaks of the PAFc, represented by CDC73 and LEO1 that overlap with RNAPII at the transcriptional start site of the PRMT5 gene (Figure 3.11).

Figure 3.11: ChIP-Sequencing Dataset Analysis at the PRMT5 Locus ChIP-sequencing tracks at the PRMT5 locus in human THP-1 leukemic cells. Tracks for RNAPII, CDC73, and LEO1 are shown. Genes marked in teal are moving right to left while those in red are moving left to right.
To confirm this enrichment in MA9 driven mouse cells, we performed ChIP-qPCR experiments on the *Prmt5* locus in MA9-*Cdc73fl*-CreER leukemic cells with and without Cdc73 excision (Figure 3.12A). Our findings show Cdc73, Leo1, WT MLL1, and MLL-AF9 were enriched at the *Prmt5* locus. Excision of Cdc73 reduces enrichment of

---

**Figure 3.12: The PAFc is a Transcriptional Regulator of the Prmt5 Locus**

A) Scheme showing the sites on the *Prmt5* locus that were tested by ChIP-qPCR. B) MA9-*Cdc73fl*-CreER cells were treated with either 7.5nM EtOH or 4OHT and grown for 48 hours. ~30x10^6 cells were harvested and ChIP was performed at the promoter, transcriptional start site (TSS), and Intron 3 of the *Prmt5* locus for CDC73, LEO1, Flag tagged MLL-AF9 fusion, MLLc, and H3K4me3. A site 20kb upstream of *Prmt5* locus was used as a negative control. All values are normalized to input or H3 and IgG was used as a non-specific binding control.

C) Stat5 and Hoxa9 localize to *Prmt5* locus. Hoxa9/Meis1-*Cdc73fl*-CreERT2 cells were treated as described in (B)
all four proteins suggesting Cdc73 is necessary for the localization of these proteins to the \textit{Prmt5} locus (Figure 3.12B), like the known MLL1 and PAFc target gene \textit{Meis1} (Figure 3.13A-B). Loss of Cdc73 also leads to a decrease in H3K4me3 consistent with the decrease in transcriptional activity and \textit{Prmt5} gene expression reduction (Figure 3.12B). Consistent with the PAFc regulating \textit{Prmt5} expression, a \textit{Prmt5}-Luciferase reporter was activated in a dose dependent manner to increasing levels of PAFc in the

\textbf{Figure 3.13: Localization of PAFc and MLL to Known PAFc Target Locus Meis1} A) Diagram showing the location of ChIP primers on the \textit{Meis1} locus B) CDC73, LEO1, Flag-MLL-AF9, MLLc, and STAT5 ChIP at Exon1 and Intron8 of the \textit{Meis1} locus. All samples are scaled to input and IgG is used as a binding control. C) PAFc and MLL-AF9 synergize at \textit{Prmt5} locus. Luciferase readout is normalized to MSCV only control (first bar). ** = p-value < 0.01, *** = p-value < 0.001. P-values determined by student t-test and compare MLL-AF9 only bar (3\textsuperscript{rd}) to increasing PAFc bars (4\textsuperscript{th} – 6\textsuperscript{th}).
presence of MLL-AF9 (Figure 3.13C). Recent work in CML illustrates STAT5 is a direct regulator of Prmt5 downstream of the BCR-ABL fusion protein (300). Further, significant overlap has been reported for Stat5 and Hoxa9 binding sites throughout the genome (301). Thus, we analyzed binding of both Hoxa9 and Stat5 at the Prmt5 locus. To test this, we generated 4OHT inducible Cdc73 floxed leukemic cell lines transformed with Myc-tagged Hoxa9 and HA-tagged Meis1. We observed Stat5 binding at the Prmt5 transcriptional start site that was strongly dependent on Cdc73 (Figure 3.12D). Hoxa9 was also enriched at the Prmt5 locus, but was less responsive to loss of Cdc73 (Figure 3.12D). This suggested that Hoxa9 is not dependent on the PAFc at this locus and may act as a pioneering factor to help localize other transcription factors and cofactors to regulate Prmt5. These data support that Prmt5 is directly regulated by the PAFc, which promotes expression in conjunction with other transcription factors such as Stat5 and Hoxa9 in leukemic cells.

3.2.6 Enzymatic Activity of Prmt5 is Required for Leukemogenesis in vivo

PRMT5 inhibitors are currently in clinical trials for the treatment of solid tumors and Non-Hodgkin’s lymphoma (ClinicalTrials.gov Identifier: NCT02783300). Due to the requirement of Prmt5 enzymatic activity in vitro, we hypothesized AML would respond to

![Figure 3.14: Chemical Inhibition of PRMT5 in vivo Extends Disease Latency in an MA9 Mouse Model](image)

A) Sub-lethally irradiated mice were intravenously injected with 100K primary MLL-AF9 cells (N=7 per dose group) or WT BM cells (N=4 per dose group). At Day 4 mice were subjected to vehicle or EPZ015666 treatment (200mg/kg) BID for 14 days by oral gavage (200μL). ** = p-value < 0.01 by Wilcoxon test. C) Mice dosed with PRMT5 Inhibitor have reduced spleen weight at moribund stage. Spleens were harvested from leukemic mice photographed and weighed. **** = p-value< 0.0001 by student’s t-test.

Prmt5 chemical inhibition in an MLL-AF9 mouse model. Recent reports in mantle cell
lymphoma (MCL) have characterized a potent Prmt5 inhibitor called EPZ015666, which we utilized in our MA9 model (265). Using primary MA9 leukemic cells, we injected sub-lethally irradiated syngeneic recipient mice and dosed them B.I.D with EPZ015666 or vehicle for 14 days by oral gavage. The median survival of mice receiving MA9 cells was extended by 4.5 days with EPZ015666 treatment similar to Prmt5 knockdown (Figure 3.14A). While all moribund mice succumbed to leukemia, a reduction in spleen weight and leukemic infiltration in EPZ015666 treated mice points to a less progressed disease compared to the vehicle treated group (Figure 3.14B & Figure 3.15A-B).

Figure 3.15: Verification of Leukemia in Prmt5 Inhibitor Latency Assay A) Images of mouse spleens at moribund stage after treatment with vehicle or EPZ015666 in either wild-type BM or MLL-AF9 BM injected mice. B) Images of BM (100X), liver (50X), sternum (50X), and spleen (20X) for mice at moribund stage. Scale bars = 20μm for BM, Liver and Sternum. Scale bars = 100μm for spleen. C) Mice weight for duration of EPZ01556 inhibitor dosing schedule. EPZ015566 showed mild weight loss in control animals (left panel). Mice weight drops steadily as leukemia develops in MA9 groups (right panel) D) Western depicting H4R3me2s mark at moribund state for vehicle and EPZ015666 treated mice. H3 is used as a loading control.
Potentially, these mice died early due to drug toxicity, but none of our control mice dosed with EPZ015666 died during the dosing schedule and displayed only minor weight loss (Figure 3.15C). However, it remains possible that these mice are weakened by the EPZ015666 dosing schedule and are more susceptible to lower leukemic burdens. Furthermore, we see no difference in H4R3me2s methylation at the moribund state between both treatment groups that likely reflects a strong selective pressure for Prmt5 catalytic activity once the drug is removed (Figure 3.15D). These results imply

Figure 3.16: EPZ015666 Treatment Hinders Human Derived Leukemic Cell Growth Human leukemic cell lines show reduced proliferation upon treatment with PRMT5 inhibitor EPZ015666. U937 (CALM-AF10), NB4 (PML-RARα), and MonoMac6 (MLL-AF9) cell lines were dosed with EPZ015666 at indicated concentrations. Cells numbers were normalized, reseeded, and retreated every 4 days over a 12-day time course. Cell number represents total cells accumulated drawn from the rate of growth between each reseeding. IC50 values were calculated at Day 12 and H4R3me2s is shown at Day 4 for the 10, 3, 1, and 0μM doses.

the enzymatic activity of Prmt5 plays an important role in MLL-fusion AML progression
and its chemical inhibition may be an effective option for patients with high expression levels of Prmt5.

To understand how chemical inhibition of Prmt5 activity affects human leukemic cell growth, we treated several human leukemic cell lines with several concentrations of the PRMT5 inhibitor EPZ015666 and measured proliferation. CALM-AF10 driven U937, PML-RARα driven NB4, and MLL-AF9 driven MonoMac6 cells all display a dose dependent decrease in proliferation and H4R3me2s demonstrating multiple leukemic subtypes are affected by PRMT5 inhibition (Figure 3.16). Notably, NB4 cells, which display low methylthioadenosine phosphorylase (MTAP) expression, also show the greatest sensitivity to PRMT5 inhibition (Figure 3.17A). This is remarkably consistent with recent reports revealing cancer cell dependence on PRMT5 with the loss of the MTAP enzyme (302,303). Similarly, mouse MA9 leukemic cells showed a dose

![Figure 3.17: PRMT5 Inhibition Offsets MA9 Leukemic Cell Growth in vitro](image)

A) MTAP expression in U937, NB4, and MonoMac6 cells as determined by qPCR scaled to GAPDH. B) Increasing doses of EPZ015666 inhibit the growth of mouse derived MA9 cells. DMSO was used as a control (0nM). IC50 value is shown below at Day 6. C) H4R3me2s methylation is decreased in a dose dependent manner with 1, 0.1, and 0μM of EPZ015666 at Day 3. H3 is used as a loading control.
dependent decrease in proliferation and H4R3me2s following Prmt5 inhibition (Figure 3.17B-C). Overall, this data provides translational evidence that human leukemic cell lines are susceptible to Prmt5 inhibition and may implicate this enzyme in many AMLs driven by various fusion oncoproteins.

3.3 Methods

3.3.1 shRNA Generation and Validation

ShRNA knockdown systems were utilized in both a constitutive and inducible manner depending on our experimental question. For the constitutive assays, we used the pSM2C vector while for the inducible assays we used the TRMPV vector. First, shRNA sequences were acquired from previous literature or from the RNAi consortium through the Broad Institute. Once the CSV file for the gene of interest is found we identified the 21mer target sequence. Full shRNA hairpins were designed as described (304) and the sequences generated were then checked that the Xho1 5’ and EcoR1 3’ cut sites were still intact. Using Invitrogen’s primer design tool we generated each miR30 116mer shown below (Table 4).

| shPrmt1.1   | 5'-gctgtgtacagtggagcagctgacatcaacctaggtagctgtgctgatgttgactgtgctgctgtgctgctggg-3' |
| shPrmt1.2   | 5'-gctgtgtacagtggagcagctgacatcaacctaggtagctgtgctgatgttgactgtgctgctgtgctgctggg-3' |
| shPrmt4.1   | 5'-gctgtgtacagtggagcagctgacatcaacctaggtagctgtgctgatgttgactgtgctgctgtgctgctggg-3' |
| shPrmt4.2   | 5'-gctgtgtacagtggagcagctgacatcaacctaggtagctgtgctgatgttgactgtgctgctgtgctgctggg-3' |
| shPrmt5.1   | 5'-gctgtgtacagtggagcagctgacatcaacctaggtagctgtgctgatgttgactgtgctgctgtgctgctggg-3' |
| shPrmt5.2   | 5'-gctgtgtacagtggagcagctgacatcaacctaggtagctgtgctgctgtgctgctggg-3' |

3.3.2 Competition Assay

MA9-TetON-shPrmt5 knockdown cells were cocultured in 12 well plates with MA9-TetON cells in a fixed 3:1 ratio. Cocultures of MA9-TetON-shRenilla and MA9-TetON, and MA9-TetON monocultures were used as a control. The wells were all dosed with
1 μg/mL Doxycycline once per day over a 12-day time course and passaged as necessary. Every 2 days roughly 250K cells were harvested, washed in Flow Buffer (PBS + 2% FBS + 1mM EDTA) and run on an LSRII flow cytometer. The cells containing our shRNA vector will fluoresce green and after dox treatment change to red. Using the PE channel, we measured the ratio of the PE+ or RFP population vs the PE- or non-RFP population. The ratios were scaled to the first reading on day 2 to avoid any plating errors between wells.

3.3.3 Colony Formation Assay

Mouse BM was harvested from mice 5 days after 5-FU (3mg) treatment and Lin-/c-Kit+ cells were isolated using a Mouse Hematopoietic Progenitor Isolation Kit. BM cells were transduced with MSCV-neo-MLL-AF9 retrovirus packaged using Plat-E cells followed by two days of spinoculation as described previously. Cells were selected in G418 (neomycin 1mg/ml) for 1 week and allowed to recover. MA9 BM was then re-transduced with pSM2C-shPrmt5 knockdown or pSM2cPrmt5 overexpression vectors and control vectors by the same procedure then selected in puromycin (100ng/mL). Cells were selected for one day and plated at 20,000 cells in methylcellulose media in 35mm dishes. Colonies were scored after 1 week. For Prmt overexpression assays cells were re-plated for 3 rounds.

3.3.4 Prmt5 Re-Expression Cell Line Generation

Using the QuikChange IIXL Site Directed Mutagenesis Kit, we generated both a WT-Prmt5 and CD-Prmt5 re-expression construct following the manufacturer's instructions. R367A and R368A mutations were made on Prmt5 to generate a catalytically dead construct as described (305). These two constructs were TOPO cloned and inserted into MigR1 vectors and sequence verified from both ends of the plasmid and within the sequence to confirm the mutation. The validation primers used are as follows: Prmt5 forward primer: 5'-GGAATTCCGCCACCATGGAACAAAAACTCATCTCAGAAG-3', Prmt5 reverse primer 5'-CCGCTCGAGGGCAGGAGCTAGAGGCCCAATGGTATATG-3', and Prmt5 mid primer 5'-GCTCCTACCTCCAATACCTGG-3'. Constructs were then viral
transduced into MA9-TetON-shPrmt5 cell lines using Plat-E viral packaging cells as previously described in chapter 2.

3.3.5 siRNA Mediated PAFc Knockdown
HeLa cells were treated with siRNA against PAF1 at a confluency of approximately 40-50% using oligofectamine reagent (Invitrogen) following the manufacturer’s protocol. Non-targeting siRNA was used as a control. SiRNA targeting a non-specific sequence (D-001810-01-05) and PAF1 (L-020349-01-0005) were acquired from GE Healthcare. Knockdown was verified by western blot analysis using whole cell extracts, which were prepared in 2X Sample Buffer (Roche).

3.3.6 ChIP
30x10^6 MA9-Cdc73fl-CreER cells treated for 48 hours with either 7.5nM EtOH or 4OHT, harvested and centrifuged into pellets. Cell pellets were crosslinked in 1% paraformaldehyde and lysed in SDS lysis buffer (+aprotinin, PMSF, Protease inhibitor tablet) then incubated on ice for 10 minutes. Sample was sheared through a 27G needle and subject to sonication on a BioRuptor XL. Cells were sonicated for 20 minutes in TPX tubes (30 seconds On, 30 seconds Off). Cells were centrifuged at 15,000rpm for 20 minutes and supernatant was pooled and combined with dilution buffer (+ aprotinin, PMSF, Roche Protease Inhibitor tablet). Input and fragmentation samples were taken (1%) and frozen for later use. The remaining volume was distributed to individual 1.5mL Eppendorf tubes and combined with 30uL of pre-washed Protein G Dynabeads. Samples were pre-cleared by rotating for 1 hour at 4°C. Beads were magnetized out of solution and supernatant was transferred to new tube and combined with 4ug of the corresponding antibody. Samples were incubated at 4°C overnight and washed for 3 minutes with Low Salt, High Salt, LiCl and TE buffers and then eluted in EB buffer (10% SDS, 48 mg NaHCO_3, ddH_2O) at 42°C for 30 minutes. Crosslinking was reversed by combining supernatant with 10uL NaCl at 65°C water bath overnight. DNA was purified using PCR purification kit. All qPCR was performed using a ratio of 10ul SYBR Green, 1uL Forward and Reverse Primer, 7uL ddH20 and 1uL DNA template on a 7500 Fast Real-Time PCR Thermocycler. Antibodies used for
ChIP include αCDC73, αLEO1, αMLLc, αFlag, αH3, αH3K4me3. All qPCR primer sequences generated using Primer3 online tool.

3.3.7 Luciferase Assay
293 cells were split into well plates seeded at 100K cells. Cells were transfected by incubating 150uL OptiMEM + 4.5uL FuGene6 reagent for 5 mins followed by adding 50ng Renilla, 400ng promoter-luciferase construct, and 1200ng of transcription factor of interest. Transfection mixture was incubated for 20 minutes and 293 culture media was changed to Opti-MEM + 0.5% FBS without antibiotics. 50uL of transfection mixture was added to 3 wells per sample and incubated at 37°C for 48 hours. Cells are then washed with PBS 2X and lysed by adding 200uL of passive lysis buffer and rocked for 15 minutes at RT. 20uL from each well are transferred to luminometer tubes and read for 10 secs with a 0.2 sec incubation time per sample. The ratio of renilla signal to luciferase determines the relative activity of the locus under each condition.

3.3.8 Human Leukemic Cell Line Culture Conditions
U937 cells were cultured in RPMI + 10% FBS. NB4 cells were cultured in RPMI w/ 2mM L-glutamine + 10% FBS. MonoMac6 cells were cultured in RPMI + 10% FBS, Non-Essential Amino Acids, Na-Pyruvate and 10ug/mL human insulin. 293 cells were cultured in DMEM + 10% FBS. All media contained 1% P/S unless its removal was temporarily required by experimental protocol.

3.3.9 Prmt5 Inhibition Assay
MA9 cells were plated at 10K cells in 2.5mL IMDM + 15% Stem Cell FBS + IL-3 in a 12 well plate and subjected to various doses of EPZ015666 (10uM, 3.33uM, 1.11uM, 370.4nM, 123.5nM, 41.2nM, 13.7nM, and 0nM) in triplicate. Cells were counted every 3rd day, retreated with EPZ015666 and reseeded at original density for a total of 12 days. DMSO was added up to the volume for a 10uM dose for each well to control for toxicity. Human derived leukemic U937, NB4, and MonoMac6 cells were cultured as previously described and plated as shown above for MA9 cell experiment. Cells were counted every 4 days over 12-day time course.
3.3.10 In Vivo Leukemic Latency Model and Histology

Primary leukemic BM cells were generated by injecting Hoxa9/Meis1 transduced TetOn BM cells into syngeneic C57Bl/6 recipient mice. Primary leukemic cells were harvested and transduced with either TRMPV-shRenilla or shPrmt5 packaged retrovirus. Transduced cells were selected in 200 ug/mL hygromycin and allowed to recover for 9 days. 250K cells were injected intravenously into the tail vein of sublethally irradiated (600 Rads) C57BL/6 recipient mice. Mice were given Baytril water (Putney) for 3 days after irradiation and rested for 4 days to allow for engraftment. On day 4 post-injection mice were put on a diet of Doxycycline food (625mg/kg) and water (2mg/mL) for the duration of the experiment. For Prmt5 chemical inhibition in vivo, 100K primary MA9 leukemic BM cells were intravenously injected into the tail vein of sub-lethally irradiated mice (600 Rads). On Day 4 post injection, mice began a dosing schedule of EPZ015666 by oral gavage at 200 mg/kg BID for 14 days. Mice were divided into 4 groups: no cell injection + vehicle, no cell injection + drug, cell injection + vehicle and cell injection + drug. Mice were weighed every day and euthanized at the moribund stage. Sternum, thymus, liver, spleen, peripheral blood, femurs and tibia tissue samples were stored in 10% formolin and processed by the In Vivo Animal Core. Sternums were decalcified and all tissues were cut, stained with Hematoxylin and Eosin and processed onto slides. Tissues were imaged on a BX41 Upright Microscope with DP71 1.45Mp camera at 100X for BM, 100X for liver, and 50X for spleen. 100X images were taken with Resolve immersion oil. RNA from BM cells was collected in Trizol for expression analysis.

3.3.11 Cytospin and Morphology Assessment

Cells were diluted to 500K / mL and 100µL was mixed with 300µL of PBS in 1.5 mL Eppendorf tubes. Slides were prelabeled with sample name, cell number, and date. Slides were inserted into cytospin holder covered by paper and solution funnel. The slide holders were placed in a Cytospin rotator and spun at 800 rpm for 5 minutes. Slides were removed carefully as to not disturb the cell spot. Slides were then subjected to the Hema3 staining kit for fixation, staining, and washing. Slides were dried for ~10 mins and then covered with Permount and a coverslip. Images were taken on a BX41
Upright Microscope with DP71 1.45Mp camera at 100X magnification using Resolve immersion oil.

3.4 Summary & Discussion

The overexpression of the Prmt family and its epigenetic role in a broad range of cancers, necessitates a deeper understanding of how these genes are regulated. Downregulation of the Prmt family in Cdc73 excised MA9 leukemic cells, suggested the PAFc may be an upstream regulator in this context. We discovered that Prmt1, Prmt4, and Prmt5 knockdown can all slow cell growth while Prmt4 and Prmt5 specifically can reduce colony formation capacity of MA9 leukemic cells (Figure 3.2B-C) These results were consistent with studies in AML cells harboring other oncogenic drivers like AML1-ETO and MLL-GAS7, suggesting that the PAFc regulation of Prmts may be more broadly applicable in AML (264,266,289).

Further phenotypic analysis shows that Prmt5 is relevant in regulating cell cycle and cellular growth, but not differentiation (Figure 3.4). A recent study demonstrates that Prmt5 inhibition can impede MA9 leukemic cell cycling like our findings, while conversely sees an induction of myeloid differentiation (297). This disparity in differentiation could be due to shRNA escapees that express sufficient Prmt5 to maintain blocked differentiation which would not influence cells subjected to a chemical inhibitor. Loss of Prmt5 or excision of Cdc73 reduces H4R3me2s and total SDMA bridging the PAFc to Prmt5 enzymatic function (Figure 3.8). Cdc73 excision cells appear to have a stronger reduction of H4R3me2s despite a higher Prmt5 expression level compared to Prmt5 knockdown cells. We speculated other target genes of the PAFc may be facilitating this effect and looked into our RNA sequencing data. We observe that methylosome protein 50 (MEP50) an integral cofactor of Prmt5, is also significantly downregulated upon Cdc73 excision. Prmt5 enzymatic activity is heavily reliant on MEP50 and the reduction of MEP50 or inhibition of Prmt5 would have a similar functional consequence on Prmt5 enzymatic output (306). Since Cdc73 excision and Prmt5 inhibition showed an induction of differentiation this may explain the
phenotypic disparity compared to Prmt5 knockdown cells. It's also possible that measuring our Prmt5 knockdown cells at a later timepoint would yield an induction of differentiation and should be investigated. This data points to a PAFc-Prmt5 axis that through certain transcriptional pathways governing key cellular processes like differentiation and epigenetic regulation maintains leukemia (Figure 3.18).

The PAFc localizes to and is necessary to maintain several transcription factors at the Prmt5 locus implying it is a direct regulator of this gene (Figure 3.12). One of these observed cofactors was Hoxa9, a major downstream target gene of wild-type MLL and MLL-fusion proteins. Since Hoxa9 was not reliant on Cdc73 and the PAFc to localize to Prmt5 we speculated it my act as pioneering factor. To test this hypothesis, we can knockdown Hoxa9 in MA9 leukemic cells and measure the enrichment of PAFc subunits and Stat5 at the Prmt5 locus. Furthermore, if Hoxa9 does indeed regulate Prmt5 similar to the PAFc we may also observe a reduction in Prmt expression, protein level, and global H4R3me2s modification. Performing these experiments upon Stat5 knockdown would also further confirm that this transcription factor is critical for Prmt5 regulation. Work by Miglori et al. identified that Prmt5 deposition of H3R2me2s corresponds with active gene loci and interestingly helps MLL-complex member WDR5 localize to active sites (307). This suggests that wild-type MLL complexes may also utilize Prmt5 enzymatic function to activate their target genes. Furthermore, since MLL-fusion leukemias are sensitive to Prmt5 knockdown or inhibition, this may support a necessary role for wild-type MLL in this context. Since MLL and MLL-fusion proteins enriched at the Prmt5 locus, their regulation of Prmt5 may maintain this function in leukemic cells as well. Further investigation into the Prmt5 and H3R2me2s enrichment at MLL-target genes may bolster this hypothesis.

Importantly, both shRNA knockdown and chemical inhibition of Prmt5 can extend leukemic disease latency in mice and a selective pressure to maintain Prmt5 expression is observed from both in vitro and in vivo experiments (Figure 3.11 - 3.14) This observation paired with the requirement of Prmt5 in normal bone marrow suggests that excessive inhibition of Prmt5 may require a precise therapeutic window to avoid toxicity.
as speculated from our latency experiment (Figure 3.14, 3.15) (277). Currently, the EPZ01566 Prmt5 inhibitor used here is undergoing clinical trial in lymphoma and solid tumor patients (ClinicalTrials.gov NCT0278330). Despite potential toxicity, our latency assay demonstrates at the moribund state H4R3me2s is equivalent between vehicle and EPZ015666 treated mice (Figure 3.15D). This suggests once the drug is removed the selective pressure for functional Prmt5 may quickly reactivate leukemogenesis. This may indicate that Prmt5 inhibition alone is not sufficient to extensively halt leukemic progression and combinatorial treatment may be more clinically utile. Since the PAFc downregulates several members of the Prmt family, targeting this complex directly may have stronger effects on leukemic cells without depleting Prmt expression excessively, preserving normal bone marrow cells. Additionally, our observation of both the PAFc and MLL-fusions at the Prmt5 locus may implicate an inhibitor of the MLL-PAFc interaction in disrupting this leukemic mechanism. Using our MLL-DN construct and retesting Prmt5 levels and functional output would help determine this possibility. Therefore, the effects on Prmt5 and its function downstream of this interaction warrant further investigation.

![Diagram](image-url)

**Figure 3.18: Model of Downstream Regulatory Arms of the PAFc in MLL-Fusion Leukemia**

Schematic depicting the regulation of Hoxa9/Meis1 and Prmt5 by the PAFc leading to the maintenance of leukemic pathways.

Overall, we identify Prmt5 as a downstream target gene of the PAFc and may have implications in other leukemic subtypes and solid tumors. The work presented here
confirms that MLL-driven AML is another potential cancer type to benefit from chemical inhibition of Prmt5 and establishes this enzyme as an essential contributor to leukemogenesis.
Chapter 4
Summary, Conclusions, and Future Directions

4.1 Summary

The underlying objective of this study was to understand the PAFc specific role in MLL-fusion AML. The PAFc is a multi-subunit transcriptional and epigenetic regulator complex known to interact with many epigenetic modifying enzymes like MLL1. Chromosomal rearrangements can result in the fusion of MLL1 to gene partners generating aggressive acute leukemias with poor prognosis affecting both adults and infants. Previous study has illustrated the PAFc-MLL interaction to be integral for leukemic progression and maintenance. However, the PAFc is known to have context dependent roles in cancer based on the subunit evaluated and warrants further inquiry in AML. Using a conditional excision system for the Cdc73 subunit, we observe impeded leukemic progression through several phenotypes including differentiation induction and altered global epigenetic patterning. We aimed to determine the transcriptional programs governed by the PAFc through RNA-sequencing and identified many altered pathways including gene-sets related to differentiation and methyltransferase activity. Among the methyltransferase cluster we noted the Prmt family was broadly downregulated suggesting they were downstream targets of the PAFc. At the time, very little was known about how the Prmts were regulated or their involvement in AML despite their oncogenic role in many cancer types. Our work revealed a PAFc-Prmt5 mechanistic axis through comparisons of Cdc73 excised and Prmt5 knockdown leukemic cells. Further, we find the PAFc directly regulates Prmt5 along with other factors like MLL, Stat5, and Hoxa9. Importantly, Prmt5 and its enzymatic activity are needed for leukemic progression in vivo marked by shRNA mediated knockdown and chemical inhibition mouse models. This demonstrates that targeting Prmt5 therapeutically may be clinically viable. Overall, our study has provided a deeper understanding of Prmt5 and the PAFc in MLL-fusion AML.
and revealed many potential PAFc regulated mechanisms in leukemia detailed below.

4.2 Conclusions & Future Directions

4.2.1 The PAFc Regulation of Epigenetic Modifiers and Differentiation Pathways

Here we provide evidence that the PAFc-Prmt5 axis is a molecular mechanism for the progression and maintenance of AML. However, our RNA-sequencing data set revealed many potential downstream pathways of the PAFc relevant in AML including differentiation and methyltransferases (Figure 2.6D & Figure 2.7). The PAFc is already known to facilitate histone modification through the recruitment of epigenetic enzymes such as Rad6/Bre1 and MLL through direct interactions (92,206,248). Our RNA-sequencing data shows that Bre1 and MLL1 expression level remain unchanged after Cdc73 excision. This implies that the PAFc is not directly responsible for the expression of these modifiers and thus they should still be able to carry out their function. These data taken together identify that the global reduction in H3K4me3 and H2BUb observed in this study is most likely due to improper localization of these enzymes to their target loci through the PAFc.

Our RNA sequencing data also demonstrates that other methyltransferases like Dnmt3a, Ezh2, and PRC2 cofactor Suz12 are downregulated upon Cdc73 excision. This suggests that these repressive complex components and enzymes may be regulated like the Prmts through the PAFc. ChIP-sequencing experiments in MA9 leukemic cells will elucidate if the PAFc localizes to these targets and serves as a regulator. If so, perhaps excision of Cdc73 would not only deplete the PAFc localization to these targets, but also the PRC2 modification H3K27me3 globally. Additionally, testing the effects of Cdc73 excision on the localization of Dnmt3a and PRC2 proteins would reinforce these epigenetic changes. Bisulfite sequencing experiments could detect methylcytosine deposited by Dnmt3a and further identify what gene targets are repressed and reliant on Cdc73 and the PAFc through these methyltransferases. Since our upregulated pathway dataset and phenotypic analysis show an induction in differentiation it's possible the genes in these datasets are repressed through the PAFc.
regulation of Dnmt3a and PRC2. In fact, Dnmt3a and Dnmt3b are shown to methylate the loci of Oct4 and Nanog during the differentiation of hESCs to repress pluripotency genes (308). Additionally, the PAFc is known to activate the Oct4 and Nanog genes in hESCs to maintain pluripotency (259). Rigbolt et al. went on to show that during PMA induced differentiation of hESCs, Dnmt3a and Dnmt3b bind the PAFc through the PAF1 and CD73 subunits. Taken together this establishes a potential role for the PAFc in the repression of gene targets that may help maintain leukemia. Observation of a PAFc-Dnmt3a/b interaction and their colocalization at genes in our differentiation cluster by ChIP or ChIP sequencing would further bolster this hypothesis.

Additionally, some of our findings suggest the PAFc regulatory pathways may apply more broadly than MA9 leukemias. We have tested if knockout of Cdc73 affects Prmt1, Prmt4, and Prmt5 in several leukemic cell lines including AML1-ETO, Hoxa9/Meis1, and E2A-HLF. Increased Prmt5 expression in AML1-ETO cells suggests the PAFc regulation of Prmt5 is not universal or may be influenced by other factors downstream of this fusion protein. Testing the PAFc-Prmt5 axis further in this cell line may bolster our understanding of PAFc mediated transcriptional regulation and Prmt5 regulatory mechanisms. We also observed normal hematopoietic progenitors show decreased Prmt5 expression after Cdc73 excision implying this pathway is not cancer cell specific (Figure 4.1). Since both Cdc73 and Prmt5 knockouts are embryonic lethal, the PAFc-Prmt5 axis may play a key role in normal cells. Comparison of normal bone marrow and

Figure 4.1: Prmt Expression is not Universally Altered by Cdc73 Excision in Leukemic Cells A) qPCR comparison of MA9 cells to normal BM measuring expression level of Cdc73 and Prmt5. All measurements are normalized to the EtOH control reading and scaled by β-Actin. B) qPCR comparing MA9, AML-ETO, Hoxa9/Meis1, and E2A-HLF leukemic cell lines for Prmt family expression level. All samples are scaled to β-Actin.
MA9 leukemic cells after Cdc73 excision or Prmt5 knockdown, may elucidate differentially regulated pathways or genes reliant on the PAFc-Prmt5 axis. These findings could broaden our understanding of the epigenetic regulatory role carried out by the PAFc.

Beyond Prmt5, preliminary ChIP-sequencing data demonstrates that the Prmt4 locus is occupied by PAFc components and may also be directly regulated. Previous work using histone peptides and purified PAFc subunits, shows that the PAFc has high binding affinity for the H3R17me2a modification placed by Prmt4 over the H4R3me2a modification placed by Prmt1 in MCF7 breast cancer cells (250). The authors provide evidence that the PAFc utilizes the H3R17me2a modification to localize to ER target genes. This mechanism may also be relevant in MLL-fusion leukemic cells and warrants further investigation. Alternatively, the PAFc may be able to sense the activity or expression level of Prmt4 through the H3R17me2a modification and may regulate Prmt4 expression accordingly. Our preliminary findings using histone peptide binding assays in 293 cells show that the Cdc73 subunit of the PAFc can bind an H3R17 methylated peptide although the specificity of this interaction is not entirely clear.

### 4.2.2 Regulation of the Prmt5 Locus Through Hoxa9 and Stat5

The relevance of the Prmt family in cancer provides many avenues of therapeutic intervention. The data presented here identifies Prmt5 as a direct downstream target of the PAFc in the context of MLL-AF9 driven leukemias. Our model describes the PAFc as a regulator of both major MLL targets like Hoxa9 and Meis1, and Prmt5 (Figure 3.18). Previous reports identify a role for Stat5 in CML downstream of the BCR-ABL fusion (300). Our data shows Stat5 was enriched at the Prmt5 locus and was dependent on the PAFc (Figure 3.12). This suggests that the PAFc may regulate Prmt5 through Stat5 in CML in addition to AML. To investigate this axis more thoroughly we need to evaluate if modulation of Stat5 can affect Prmt5 expression or function downstream. Using shRNA against Stat5 we can measure Prmt5 expression level, protein content and deposition of H4R3me2s in MA9 leukemic cells. If Stat5 is a legitimate regulator of Prmt5, we would expect to see these experiments reduce Prmt5 and its enzymatic
output in the cell similar to excision of $Cdc73$. We also observe $Stat5$ expression level is unchanged after $Cdc73$ excision suggesting loss of its enrichment may be due to localization defects or potential depleted phosphorylation by an upstream kinase. Confirmation of a Stat5-PAFc interaction by CoIP would further reinforce the regulation of $Prmt5$ and the reliance of Stat5 enrichment on the PAFc.

Interestingly, another Stat family member Stat3, is shown to require the PAFc to help localize it to target genes downstream of the IL-6 signaling pathway through direct interaction (212). The Stat3 and Stat5 reliance on the PAFc for localization, despite possessing DNA binding domains, suggests they require other cofactors like the PAFc to remain stably bound to target loci. Inhibition of Stat5 in our MA9 leukemic cells may also reveal similar phenotypes as our $Prmt5$ knockdown or $Cdc73$ excision cells and would warrant investigation. A recent paper demonstrated a selective Stat5 compound could abrogate the growth of several human derived leukemic cell lines and may be due to downregulation of $Prmt5$ (309). Alternatively, the Jak/STAT pathway has been shown to facilitate myeloproliferative neoplasms through JAKV617F mutations (310). Interestingly, these mutants could phosphorylate Prmt5 protein preventing binding to its cofactor MEP50 and reducing its methyltransferase activity. Comparison of JAKV617F mutant AML cells and our MLL-AF9 leukemic cells may identify why the PAFc doesn’t universally maintain Prmt5 expression. Additionally, measuring the phosphorylation status of Prmt5 protein after $Cdc73$ excision may further explain the reduction of its enzymatic output in our MA9 cells.

Our disease latency model and ChIP data suggests $Prmt5$ is also regulated by Hoxa9 and may be relevant in all high Hox expressing leukemic subtypes. Since roughly 50% of AML presents with overactivated Hox transcriptional programs, understanding the interplay between Hoxa9 and Prmt5 may reveal critical leukemic mechanisms applicable to many AML subtypes (298). Recent literature shows that Prmt5 can methylate HOXA9 protein in endothelial cells and the site of methylation has been identified (288). This methylation event was sufficient to stabilize the Hoxa9 protein on the E-selectin locus and amplify its expression. Interestingly, preliminary data provided
by the Rual lab shows that in leukemic cells, Hoxa9 protein is indeed methylated on arginine 141 slightly contrasting with methylation of arginine 140 in endothelial cells (Figure 4.2). Since Prmt5 is noticeably upregulated in murine AML cells, perhaps Hoxa9 can further enhance the expression of its downstream gene targets through Prmt5 mediated methylation. Taken together, this data may suggest that Hoxa9 and Prmt5 operate in a positive feedback loop downstream of the PAFc in leukemic cells. The posttranslational methylation of Hoxa9 by Prmt5 may amplify Hoxa9 transcriptional activity increasing its ability to maintain leukemia. Since the Prmt5 locus from our

![Figure 4.2: HOXA9 is Methylated in Leukemic Cells](image)

**Figure 4.2: HOXA9 is Methylated in Leukemic Cells** Representative MS/MS spectrum showing the monomethylation at R141 of HOXA9. Tryptic peptides obtain upon in-gel digestion of HOXA9 were resolved on a reverse phase column and collision induced dissociation spectra were obtained using Orbitrap mass spectrometer. A MS/MS spectrum corresponding to 141RmeGDCcamPTLDHTLSLTDYACcamGSPPVDR166 of HOXA9 (precursor [M+H]+3 = 973.46 m/z) is shown above. Observed b- and y-ions are indicated. Ccam = carbamidomethyl-Cysteine; Rme = Methyl-Arginine. Courtesy of the Rual Lab.

ChIP data is bound by Hoxa9 it may positively feedback to maintain higher transcriptional activity through the maintenance of high Prmt5 expression (Figure 4.3). To test this hypothesis, several experiments could be undertaken. We know that ectopic overexpression of Hoxa9 and Meis1 in bone marrow cells are sufficient to generate an aggressive leukemia in mouse models (311). To establish if Hoxa9 methylation is
important in leukemogenesis, we can overexpress our R141A mutant along with Meis1 in normal BM and inject these cells into sub-lethally irradiated mice. If the mutant Hoxa9 group develops leukemia more slowly or is unable to generate a disease compared to WT Hoxa9, we could conclude that methylation of Hoxa9 is critical to leukemogenesis. Follow up experiments using a tagged R141A mutant of Hoxa9 may further describe this mechanism. Measuring the enrichment of tagged mutant Hoxa9 at known Hoxa9 target loci including Prmt5 would determine if Hoxa9 methylation is required for its localization.

**Figure 4.3: The Hoxa9-Prmt5 Feedback Loop in AML** Diagram depicting the PAFc as an upstream regulator of Prmt5 and Hoxa9. Post-translational methylation of Hoxa9 by Prmt5 may help Hoxa9 feedback to maintain levels of Prmt5 expression. This loop may contribute to the maintenance of AMLs.

Additionally, both Hoxa9 and Prmt5 are shown to be important in maintaining the function of HSCs in normal bone marrow. For example Hoxa9 expression level is critical to drive hESC differentiation into HSCs during blood system development (312). Further, overexpression of Hoxa9 in normal HSCs leads to increased self-renewal and stem cell expansion while knockdown was shown to deplete self-renewal of HSCs (313). Mechanistically, Rad21 a component of the cohesion complex, was shown to interact with the PRC2 complex and localize it to the Hoxa9 gene for transcriptional silencing revealing it may help regulate Hoxa9 mediated HSC maintenance (314). Prmt5 is most highly expressed in HSPC compartments and knockouts have several phenotypes in HSCs including depletion of cell number and activated cell cycling of LT-HSCs leading to pancytopenia (277). It is possible that the PAFc as observed in MA9 leukemic cells might also govern the maintenance of Hoxa9 and Prmt5 in HSCs. The
PAFc is connected to maintaining pluripotency genes like Oct4 and Nanog in hESCs and may also reinforce HSC homeostasis through Hoxa9 and Prmt5 expression (259). The Hoxa9-Prmt5 feedback loop may be applicable in this non-leukemic HSC compartment to maintain self-renewal programs as well and would be intriguing to evaluate. Isolating HSCs from human or mouse bone marrow and measuring methylated Hoxa9 by mass spectrometry as performed above would provide evidence this loop may apply in normal blood. Directed ChIP at the Prmt5 and Hoxa9 loci in HSCs for the PAFc would further bolster this possible mechanism. If this feedback loop is real, it may broaden our knowledge about how a regulatory cycle in normal HSCs helps maintain homeostasis in the hematopoietic system while dysregulation of Hoxa9 and Prmt5 can hyperactivate this process and transform and/or maintain leukemia.

4.2.3 The Role of Id2 Downstream of the PAFc-Prmt5 Axis

To further evaluate the PAFc-Prmt5 axis in AML, we wanted to test how the expression of Prmt5 target genes would change due to either Cdc73 excision or Prmt5 knockdown. One of the differentiation genes we used to confirm differentiation induction in Cdc73 excised leukemic cells (Figure 2.3D) was Id2. This gene was selected based on several factors from the literature. Huang et al. demonstrated that Prmt5 is an essential regulator of oligodendrocyte differentiation through Id2 repression (315). This implicated Id2 as a direct target of Prmt5. Additionally, Id2 is a distinct marker of active myeloid differentiation in normal and leukemic cells evidenced by its expression level in more mature subtypes of AML and prevalence in granulocytes and macrophages (316). However, it is important to note that Id2 has opposing functions in normal and cancerous tissue depending on the cellular context. For example, high Id2 expression in triple negative breast cancer correlated with poor outcomes and lineage transcription factor growth factor independent 1 (Gfi1) can repress Id2 to allow for proper erythropoiesis (317,318).

In our AML system, we measured the expression level of Id2 after Prmt5 knockdown and compared it to our Cdc73 excision cell lines. The results revealed that Id2 was upregulated by roughly 2-fold with excision of Cdc73 or knockdown of Prmt5 (Figure
4.4A-B). Furthermore, we asked if knockdown of Prmt5 in our in vivo latency assay would affect Id2 expression. Interestingly, at the moribund state, we see a statistical increase of Id2 expression in our Dox treated group (Figure 4.4C). This data suggests that although having varied disease latencies these two groups display similar Prmt5 expression, perhaps the increased expression of Id2 may be a by-product of delayed leukemic disease progression. In fact, one study found that Id2 expression not only positively correlated with myeloid differentiation and improved clinical outcomes in AML patients, but its overexpression could delay leukemic disease latency and restrict colony forming capacity in vitro (319). In CML, Id2 levels are also reduced while family member Id1 is highly expressed and regulated by the Stat5 transcription factor (320). Since the Stat5 regulation of Prmt5 seems conserved from CML to AML and is reliant on the PAFc for localization to target genes, perhaps Id1 may be influenced by the PAFc as well. Taken together, the opposing functions of Id1 and Id2 in CML may be regulated by the PAFc-Stat5 interaction upstream and be relevant in AML (Figure 4.5). In fact, Id1 is highly overexpressed in AMLs including MLL-AF9 driven human THP-1 and MOLM13

Figure 4.4: Id2 May be a Key Downstream Target of the PAFc-Prmt5 Axis A) Expression of Prmt5 and Id2 in MA9-Cdc73fl-CreER cells treated with 7.5nM EtOH or 4OHT at 48 hours. B) Expression of Prmt5 and Id2 in MA9TetON-shPrmt5 cells treated with vehicle or 1ug/mL Dox at 48 hours. C) Id2 expression of BM extracted from secondary Hoxa9/Meis1 leukemic mice at moribund stage with and without Prmt5 knockdown. * = p-value < 0.05.
cells correlating with poor overall survival (321,322). Directed ChIP at the *ld1* and *ld2* loci for the PAFc, Stat5, Prmt5, repressive modifications like H4R3me2s and activating modifications like H3K4me3 will help determine if this mechanism exists similar to our PAFc-Prmt5 axis.

Oddly, despite seeing increased *ld2* expression upon knockdown of *Prmt5*, we do not see a corresponding phenotypic activation of differentiation by flow cytometry (Figure 3.4B). This perhaps indicates *Id2* de-repression after *Prmt5* knockdown is not sufficient to fully induce differentiation. Testing the expression of our panel of differentiation genes between vehicle and Dox-treated MA9-shPrmt5 leukemic cells may elucidate if this is the case. We know that *Cdc73* excision also reduces the expression of Prmt5 cofactor MEP50 from our RNA-sequencing experiment. If multiple genes are needed to fully induce differentiation perhaps *Id2* is more sensitive to minor changes in Prmt5 functionally while others can only be de-repressed if Prmt5 is functionally blocked. A recent paper from Dr. Stephen Nimer and colleagues observe that chemical inhibition of Prmt5 does induce differentiation and similar to disruption or reduction of the Prmt5-MEP50 interaction may again explain why this disparity occurs (297). Overall, our

**Figure 4.5: PAFc-Stat5 Coordination May Regulate Id Genes** Diagram depicting the potential role for the PAFc as an upstream regulator of Id genes through Stat5 mediated transcriptional activation. The control of *Id1* and *Id2* may help maintain cells in a leukemic state.
preliminary findings reinforce the negative correlation between \( \text{Id2} \) expression and leukemic progression and may implicate Prmt5 as a regulator of \( \text{Id2} \) in AML downstream of the PAFc.

4.2.4 Targeting the PAFc in MLL-Fusion Leukemias

Our data suggests that multiple Prmt family members may be directly regulated by the PAFc besides Prmt5 demonstrated in this study. Prmt1, Prmt4, and Prmt5 all have been linked to acute leukemia and are also shown to be embryonic lethal in knockout models\(^{(273,275,276)}\). Therapies that heavily decrease Prmt protein function may not be tolerated by patients as was mildly observed in our \textit{in vivo} latency assay using Prmt5 inhibitors due to effects on normal cells (Figure 3.14-15). However, upstream inhibition of the PAFc may serve to reduce \( \text{Prmt} \) expression as we have demonstrated upon \( \text{Cdc73} \) excision enough to reduce leukemic progression without preventing normal cell function. We have already demonstrated that disruption of the MLL-PAFc interaction is detrimental to leukemic cells, but is tolerated by the normal hematopoietic compartment (Figure 2.1)\(^{(92,245)}\). Since MLL fusion proteins were detected at the \( \text{Prmt5} \) locus it is possible chemical inhibition of the MLL-PAFc interaction may already influence \( \text{Prmt} \) family expression level and should be investigated. Our lab performed several preliminary fluorescent binding screens using fragments of the PAFc subunit PAF1 and the CXXC-RD2 region of the MLL-fusion. We were only able to narrow the binding area to about 50aa of MLL. This however, made compound design against this protein-protein interaction quite challenging. Despite these setbacks, follow up work disrupting this interaction may have a twofold benefit against the maintenance of AML since PAFc regulates both \( \text{Hoxa9} \) and \( \text{Prmt5} \), which have MLL-fusion enrichment at their loci. If this approach is unsuccessful, inhibiting PAFc localization to the Prmt loci may also broadly reduce \( \text{Prmt} \) expression without dysregulating normal cells.
### APPENDIX

Tables of target gene overlap between MLL-fusion targets and *Cdc73* excision data described in chapter 2

<table>
<thead>
<tr>
<th>human_gene_ID</th>
<th>mouse_gene_ID</th>
<th>RPKM_Cdc73_excision</th>
<th>RPKM_control</th>
<th>Fold_Change</th>
<th>q_value(cufflinks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EYA1</td>
<td>Eya1</td>
<td>0</td>
<td>2.63496</td>
<td>0.275106191</td>
<td>0.000165472</td>
</tr>
<tr>
<td>HPGD</td>
<td>Hpgd</td>
<td>387.631</td>
<td>878.753</td>
<td>0.441750128</td>
<td>0.000165472</td>
</tr>
<tr>
<td>PMS1</td>
<td>Pms1</td>
<td>1.9027</td>
<td>4.98044</td>
<td>0.485365625</td>
<td>0.000165472</td>
</tr>
<tr>
<td>UBE3A</td>
<td>Ube3a</td>
<td>13.3133</td>
<td>26.2734</td>
<td>0.524808055</td>
<td>0.000165472</td>
</tr>
<tr>
<td>TNPO1</td>
<td>Tnpo1</td>
<td>44.1443</td>
<td>84.8003</td>
<td>0.526155503</td>
<td>0.000165472</td>
</tr>
<tr>
<td>PHTF1</td>
<td>Phtf1</td>
<td>6.58748</td>
<td>12.975</td>
<td>0.542932379</td>
<td>0.0615731</td>
</tr>
<tr>
<td>RREB1</td>
<td>Rreb1</td>
<td>10.2714</td>
<td>19.2878</td>
<td>0.555575272</td>
<td>0.000165472</td>
</tr>
<tr>
<td>DCUN1D1</td>
<td>Dcun1d1</td>
<td>16.4876</td>
<td>30.3454</td>
<td>0.557900043</td>
<td>0.000165472</td>
</tr>
<tr>
<td>ZBTB1</td>
<td>Zbtb1</td>
<td>7.09449</td>
<td>13.2288</td>
<td>0.568860721</td>
<td>0.000165472</td>
</tr>
<tr>
<td>CNOT6L</td>
<td>Cnot6l</td>
<td>19.038</td>
<td>34.0634</td>
<td>0.571479092</td>
<td>0.000165472</td>
</tr>
<tr>
<td>CDK6</td>
<td>Cdk6</td>
<td>22.6461</td>
<td>39.6816</td>
<td>0.581248034</td>
<td>0.000165472</td>
</tr>
<tr>
<td>ERBB2IP</td>
<td>Erbb2ip</td>
<td>22.9919</td>
<td>40.2504</td>
<td>0.581616178</td>
<td>0.000165472</td>
</tr>
<tr>
<td>MTOH</td>
<td>Mtoh</td>
<td>70.1118</td>
<td>115.258</td>
<td>0.611672315</td>
<td>0.000165472</td>
</tr>
<tr>
<td>MBNL1</td>
<td>Mbnl1</td>
<td>108.905</td>
<td>169.471</td>
<td>0.644713764</td>
<td>0.00012136</td>
</tr>
<tr>
<td>SESN1</td>
<td>Sesn1</td>
<td>7.7245</td>
<td>12.5303</td>
<td>0.644812015</td>
<td>0.000165472</td>
</tr>
<tr>
<td>ORMDL1</td>
<td>Ormd1</td>
<td>17.0434</td>
<td>26.9074</td>
<td>0.646545361</td>
<td>0.0004572</td>
</tr>
<tr>
<td>METTL5</td>
<td>Mettl5</td>
<td>11.9904</td>
<td>18.9277</td>
<td>0.651876534</td>
<td>0.570835</td>
</tr>
<tr>
<td>MYB</td>
<td>Myb</td>
<td>83.3229</td>
<td>127.798</td>
<td>0.654691067</td>
<td>0.00099168</td>
</tr>
<tr>
<td>ZBTB25</td>
<td>Zbtb25</td>
<td>6.95428</td>
<td>11.1366</td>
<td>0.655396075</td>
<td>0.00124606</td>
</tr>
<tr>
<td>RAP1GDS1</td>
<td>Rap1gds1</td>
<td>24.336</td>
<td>37.33</td>
<td>0.660996608</td>
<td>0.00124606</td>
</tr>
<tr>
<td>PURB</td>
<td>Purb</td>
<td>34.1082</td>
<td>21.5209</td>
<td>1.558916384</td>
<td>0.000165472</td>
</tr>
<tr>
<td>HOXA7</td>
<td>Hoxa7</td>
<td>23.4742</td>
<td>12.6907</td>
<td>1.787651471</td>
<td>0.000165472</td>
</tr>
<tr>
<td>CDKN1B</td>
<td>Cdkn1b</td>
<td>21.9856</td>
<td>7.79673</td>
<td>2.612970956</td>
<td>0.000165472</td>
</tr>
<tr>
<td>EPM2AIP1</td>
<td>Epm2aip1</td>
<td>13.6106</td>
<td>4.39465</td>
<td>2.708349939</td>
<td>0.000165472</td>
</tr>
<tr>
<td>LPAR6</td>
<td>Lpar6</td>
<td>12.5759</td>
<td>2.95553</td>
<td>3.432131724</td>
<td>0.000165472</td>
</tr>
</tbody>
</table>

Shown above are the 25/129 genes directly bound by MLL-AF9 (human cells) that are significantly affected by loss of *Cdc73* (mouse cells). 20 genes were downregulated (dark gray) upon loss of *Cdc73*; 5 genes were upregulated (white) upon loss of *Cdc73*. 
Shown above are the 42/165 genes directly bound by MLL-ENL that are significantly affected by loss of Cdc73. 33 genes were downregulated (dark gray) upon loss of Cdc73; 9 genes were upregulated (white) upon loss of Cdc73.
REFERENCES


50. Stopka T, Amanatullah DF, Papetti M, Skoultchi AI. PU.1 inhibits the erythroid program by binding to GATA-1 on DNA and creating a repressive chromatin structure. EMBO J. 2005 Nov 2;24(21):3712–3723.


133. Pulikkan JA, Tenen DG, Behre G. C/EBPα deregulation as a paradigm for leukemogenesis. Leukemia. 2017 Jul 19;


172. Stanulla M, Wang J, Chervinsky DS, Thandla S, Aplan PD. DNA cleavage within the MLL breakpoint cluster region is a specific event which occurs as part of higher-order chromatin fragmentation during the initial stages of apoptosis. Mol Cell Biol. 1997 Jul;17(7):4070–4079.


