2014

The Role of AF9 and AF9-Mediated Protein Interactions in Hematopoiesis and Leukemogenesis

Alyson Anne Lokken

Loyola University Chicago

Follow this and additional works at: https://ecommons.luc.edu/luc_diss

Part of the Molecular Biology Commons

Recommended Citation

Lokken, Alyson Anne, "The Role of AF9 and AF9-Mediated Protein Interactions in Hematopoiesis and Leukemogenesis" (2014). Dissertations. 1281.
https://ecommons.luc.edu/luc_diss/1281

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Dissertations by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License.

Copyright © 2014 Alyson Anne Lokken
LOYOLA UNIVERSITY CHICAGO

THE ROLE OF AF9 AND AF9-MEDIATED PROTEIN INTERACTIONS IN HEMATOPOIESIS AND LEUKEMOGENESIS

A DISSERTATION SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL IN CANDIDACY FOR THE DEGREE OF DOCTOR OF PHILOSOPHY PROGRAM IN MOLECULAR BIOLOGY

BY

ALYSON ANNE LOKKEN

CHICAGO, IL

AUGUST 2014
ACKNOWLEDGEMENTS

I would first like to express my sincerest appreciation to my mentor, Dr. Nancy Zeleznik-Le, for giving me the opportunity to develop as a scientist under her guidance. I learned not only how to be a skilled biologist at the bench, but also how to think critically about what I was doing and why I was doing it. I learned that persistence, patience, and a positive attitude are required to succeed as a scientist. I know that she has prepared me well to take the next step in my career, and for that I am very grateful.

Additionally, I would like to thank my committee chair, Dr. Dingwall, and my committee members Dr. Manuel Diaz, Dr. Charles Hemenway and Dr. Jiwang Zhang, for their support throughout my doctoral training. They were always available for discussing not only my latest successes and failures in lab, but for casual conversation and a good laugh. I could not have asked for a better committee to guide me through this process.

I’d also like to acknowledge Dr. John Bushweller and the members of his lab at the University of Virginia, especially Aravinda Kuntimaddi, for their contributions to this project. Having such friendly experts in the structural aspects of this project made for a wonderful experience in cross-country scientific collaboration.

I would like to thank the director of the Molecular Biology program, Dr. Denning, for his guidance, as well as the students in the program and our invaluable secretary,
Donna Buczek, for their support throughout my time as a graduate student. Thanks to Pat Simms and Veronica Volgina in the FACs lab for their assistance and conversation. A very special thank you to Lorelei Halcholski and the members of the Gene Regulation and Epigenetics group—having a positive atmosphere in which to work meant so much to me. There was never a shortage of people to chat with in the hallway and that made even the hardest days much more enjoyable.

A huge thank you to the members of the Zeleznik-Le lab—Nick, Yousaf, Laurie, Noah, Sonia, Adam, and Shubin. I may be partial, but I don’t think there is a better lab at Loyola. From experimental troubleshooting to lunchtime conversation to weekend potlucks, you all have made this experience such an enjoyable one. I look forward to sharing many more fun times with all of you!

Finally, I wish to thank my friends and family. Even at the hardest times, I have always had the most amazing group of people supporting me. From West Salem to St. Paul to Chicago, I have managed to make some incredible friends and I want to thank them for always being there for me, and for being the funniest people I know.

My sisters, Amy and Annika, have been such great role models for me throughout my entire life. They put up with an annoying little sister and helped me become who I am today. My brothers-in-law, Scott and Shane, and my wonderful nieces and nephews Grace, Liam, Audrey and Hudson, have really made our family complete.

To my dad, who was always beaming with pride over the successes of his daughters. He once told me that he thought about going into research but decided not to because things just never worked out often enough, and I thought about that and smiled during every failed experiment. And lastly to my mom, who has been my greatest support
and my biggest inspiration. Her encouragement, listening ear, positive attitude and guidance are immeasurable. I could not have done this without her.
To my family
# TABLE OF CONTENTS

**ACKNOWLEDGEMENTS**

ii

**LIST OF TABLES**

ix

**LIST OF FIGURES**

x

**LIST OF ABBREVIATIONS**

xiii

**ABSTRACT**

xv

**CHAPTER ONE: INTRODUCTION**

1

**CHAPTER TWO: LITERATURE REVIEW**

4

- Hematopoiesis
  4
- The Hematopoietic Hierarchy of Differentiation
  5
- Transcriptional Control of Differentiation
  7
- Chromatin and Epigenetics
  15
- MLL
  18
  - MLL Protein Domains
    20
  - MLL Target Genes
    24
- MLL Leukemia
  26
- AF9
  30
  - AF9 Homolog ENL
    38
  - AF9 and ENL in Transcriptional Elongation
    39
- AF9 Interacting Proteins
  42
  - AF4
    42
  - DOT1L
    45
  - CBX8
    49
  - BCoR
    53

**CHAPTER THREE: METHODS**

56

- Cloning of Full-Length AF9
  56
- Site-Directed Mutagenesis to Introduce Mutations into Full-Length AF9
  58
- Cloning of the MSCVneoMLL-AF9(D546R) Fusion Protein
  60
- Cloning of AF9(D546R) 376-568 and 470-568 into pCMV3Xflag
  61
- Retrovirus Production, Concentration and Titering
  61
- Immunoprecipitation and Western Blot Analysis
  63
- RNA Isolation and cDNA Synthesis
  65
- Reverse Transcription PCR
  65
- Quantitative Reverse Transcription-PCR (qRT-PCR)
  66
- Genomic DNA Isolation
  68
- Af9 Genotyping PCR
  68
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary Mouse Bone Marrow Isolation</td>
<td>69</td>
</tr>
<tr>
<td>c-Kit Positive Selection of Mouse Bone Marrow Progenitor Cells</td>
<td>69</td>
</tr>
<tr>
<td>Retroviral Transduction</td>
<td>70</td>
</tr>
<tr>
<td>Colony Replating Assay</td>
<td>70</td>
</tr>
<tr>
<td>Genetic Complementation Assay</td>
<td>71</td>
</tr>
<tr>
<td>Murine Colony-Forming Cell (CFC) Assays</td>
<td>72</td>
</tr>
<tr>
<td>CFC Assays with $Af^{fl/fl}$ c-Kit+ Bone Marrow Progenitor Cells</td>
<td>72</td>
</tr>
<tr>
<td>CFC Assays with YFP-AF9 Transduced $Af^{fl/fl}$ Bone Marrow Progenitor Cells</td>
<td>74</td>
</tr>
<tr>
<td>Cytospins and Staining</td>
<td>75</td>
</tr>
<tr>
<td>Surface Marker Staining by Flow Cytometry</td>
<td>75</td>
</tr>
</tbody>
</table>

**CHAPTER FOUR: RESULTS**

**AIM ONE:** To determine which AF9 complex interactions are necessary for normal hematopoietic development

- Cloning and Stable Expression of full length AF9 point mutants 77
- Interaction of AF9(D546R) with DOT1L and AF4 82
- *In vivo* Deletion of AF9 Does Not Alter Levels of Hematopoietic Lineage-Specific Transcription Factors in Af9-Inducible Knockout Mouse Bone Marrow Cells 89
- Loss of AF9 Does Not Alter Hematopoietic Lineage Differentiation *in vitro* 91
- Re-expression of Wild-Type AF9 or AF9 Mutants S565A and S565D Does Not Alter *in vitro* Lineage Differentiation of Af9-Inducible Knockout Bone Marrow Progenitor Cells 102

**AIM TWO:** To determine whether AF9 interactions with AF4 and Dot1l are necessary for leukemogenesis

- AF9 Interaction with AF4 Is Required for *in vitro* Leukemic Transformation 106
- Genetic Complementation of MLL-AF9(D544R) Restores *in vitro* Transformation 110
- AF9 Interaction with DOT1L is Required for *in vitro* Leukemic Transformation 114
- MLL-AF9 Mutants E506R, I538A and K557E Do Not Alter the *in vitro* Leukemic Transformation Ability of the MLL-AF9 Oncogenic Fusion 118
- MLL-AF9 Fusion Proteins Resulting from Two Different Breakpoints in AF9 Are Equal in Transforming Potential 121
- Productive Transcription Elongation is Altered in MLL-AF9(D544R and D546R)-Expressing Bone Marrow Cells, but not MLL-AF9 Mutants E506R, I538A or K557E-Expressing Cells 124

**CHAPTER FIVE: DISCUSSION**

**Aim One:** The Role of AF9 in Hematopoiesis

- AF9(D546R) Interaction with AF4 and DOT1L 128
- The Role of AF9 in Hematopoietic Differentiation 134
  - Erythroid Lineage Differentiation 134
Myeloid Lineage Differentiation
Megakaryocytic Lineage Differentiation
Re-Expression of AF9 in Af9-null Bone Marrow Progenitor Cells
Af9 is Not a Critical Regulator of Adult Murine Hematopoiesis
Aim Two: The Role of AF9 in Leukemogenesis
MLL-AF9(D544R) Mutation Decreases Leukemic Transformation Potential
Transformation is Rescued by Restoration of the AF4-AF9 Salt Bridge Interaction
MLL-AF9(D546R) Mutation Reduces Leukemic Transformation Potential
MLL-AF9 Mutants E506R, I538A and K557E Do Not Alter Transformation Ability
Transformation Potential of MLL-AF9 Fusions Resulting from Two Different Breakpoints in AF9
Transcription Elongation in Cells Expressing MLL-AF9 Mutants
Model of AF9 Interaction with AF4 and DOT1L in Leukemogenesis

REFERENCES

VITA
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Primer Sequences for Cloning of MSCVYFP.AF9.3xFLAG</td>
<td>57</td>
</tr>
<tr>
<td>2. Primer Sequences for Site Directed Mutagenesis</td>
<td>60</td>
</tr>
<tr>
<td>3. Primer Sequences for Cloning of MSCVneoMLL-AF9(D546R).3xFLAG</td>
<td>61</td>
</tr>
<tr>
<td>4. Primer Sequences for Cloning AF9(D546R) 376-568 and 470-568</td>
<td>61</td>
</tr>
<tr>
<td>5. Primer Sequences for Amplification of AF5 and AF9 Mutant Sequence</td>
<td>66</td>
</tr>
<tr>
<td>6. Primer Sequences for Erythroid and Myeloid Lineage Transcription Factors</td>
<td>67</td>
</tr>
<tr>
<td>7. Primer Sequences for the Exon 1, Exons 8/9 and Exon 12 of Meis1</td>
<td>68</td>
</tr>
<tr>
<td>8. Primers Sequences for Genotyping of Af9 Mice</td>
<td>69</td>
</tr>
<tr>
<td>9. FACs Antibodies for Erythroid and Myeloid Surface Markers</td>
<td>76</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Weissman Model of Hematopoietic Hierarchy</td>
<td>6</td>
</tr>
<tr>
<td>2. Adolfsson Model of the Hematopoietic Hierarchy</td>
<td>8</td>
</tr>
<tr>
<td>3. Protein Domains of MLL and MLL Fusion Proteins</td>
<td>23</td>
</tr>
<tr>
<td>4. Protein Domains of AF9 and ENL</td>
<td>32</td>
</tr>
<tr>
<td>5. AF9-Containing Protein Complexes</td>
<td>34</td>
</tr>
<tr>
<td>6. Protein Domains of AF4</td>
<td>43</td>
</tr>
<tr>
<td>7. Protein Domains of DOT1L</td>
<td>46</td>
</tr>
<tr>
<td>8. Protein Domains of CBX8</td>
<td>50</td>
</tr>
<tr>
<td>9. Protein Domains of BCoR</td>
<td>54</td>
</tr>
<tr>
<td>10. Engineering and Stable Expression of Full Length AF9 Point Mutants</td>
<td>80</td>
</tr>
<tr>
<td>11. Interaction of AF9(D546R) fragments 470-568 and 376-568 with DOT1L</td>
<td>85</td>
</tr>
<tr>
<td>12. Interaction of AF9(D546R) fragments 470-568 and 376-568 with AF4</td>
<td>88</td>
</tr>
<tr>
<td>13. Expression of Lineage-Specific Transcription Factors in Af9 Inducible Knockout Mouse Bone Marrow Cells</td>
<td>91</td>
</tr>
<tr>
<td>14. Erythroid Colony Formation by Af9 Inducible Knockout Bone Marrow Progenitor Cells</td>
<td>92</td>
</tr>
<tr>
<td>15. Surface Marker Staining of Af9 Inducible Knockout Bone Marrow Progenitor Cells Plated for Erythroid Lineage Differentiation</td>
<td>94</td>
</tr>
</tbody>
</table>
16. Myeloid Colony Formation by Af9 Inducible Knockout Bone Marrow Progenitor Cells

17. Surface Marker Staining of Af9 Inducible Knockout Bone Marrow Progenitor Cells Plated for Myeloid Lineage Differentiation

18. Acetylcholinesterase-positive Megakaryocyte Staining of Af9 Inducible Knockout Bone Marrow Progenitor Cells

19. Erythroid Colony Formation in Af9 Inducible Knockout Bone Marrow Progenitor Cells Re-Expressing AF9 Wild-Type, S565A or S565D Mutants

20. Surface Marker Staining of Af9 Inducible Knockout Bone Marrow Progenitor Cells Re-Expressing Wild-type or Mutant AF9 Plated for Erythroid Lineage Differentiation

21. MLL-AF9(D544R) Colony Assay

22. Hoxa9 and Meis1 Expression in MLL-AF9(D544R) Transduced Bone Marrow Progenitor Cells

23. MLL-AF5(K717D) Colony Assay and Hoxa9 and Meis1 Expression

24. Genetic Complementation of MLL-AF5(K717D) by AF9(D544R): Colony Assay and Expression of Hoxa9 and Meis1

25. MLL-AF9(D546R) Colony Assay

26. Hoxa9 and Meis1 Expression in MLL-AF9(D546R) Transduced Bone Marrow Progenitor Cells


28. Hoxa9 and Meis1 Expression in MLL-AF9 Point Mutant-Transduced Bone Marrow Progenitor Cells

29. MLL-AF9 (376-568) or (475-568) Colony Assay

30. Hoxa9 and Meis1 Expression in Bone Marrow Cells Transduced with MLL-AF9(376-568) or (475-568)

31. Transcriptional Elongation of Meis1 in MLL-AF9 Wild-Type or Mutant Transformed Bone Marrow Progenitor Cells
32. Conservation of a Hydrophobic Motif in the AF9 Interacting Regions of AF4/AF5q31, DOT1L, BCoR and CBX8 129

33. AF9(WT or D546R) Binding to DOT1L 130

34. AF9(WT or D546R) Binding to AF4 132

35. Model of AF9 Interaction with AF4 and DOT1L at MLL Target Genes 156
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABL</td>
<td>Acute Biphenotypic Leukemia</td>
</tr>
<tr>
<td>AHD</td>
<td>ANC1 Homology Domain</td>
</tr>
<tr>
<td>ALL</td>
<td>Acute Lymphoid Leukemia</td>
</tr>
<tr>
<td>AML</td>
<td>Acute Myeloid Leukemia</td>
</tr>
<tr>
<td>ANC1</td>
<td>Actin Non-Complementing gene 1, budding yeast Homology Domain</td>
</tr>
<tr>
<td>CLP</td>
<td>Common Lymphoid Progenitor</td>
</tr>
<tr>
<td>CMP</td>
<td>Common Myeloid Progenitor</td>
</tr>
<tr>
<td>E/Meg</td>
<td>Erythrocyte/Megakaryocyte</td>
</tr>
<tr>
<td>GM</td>
<td>Granulocyte/Monocyte</td>
</tr>
<tr>
<td>GMP</td>
<td>Granulocyte-Macrophage Progenitor</td>
</tr>
<tr>
<td>HMT</td>
<td>Histone Methyltransferase</td>
</tr>
<tr>
<td>HSC</td>
<td>Hematopoietic Stem Cells</td>
</tr>
<tr>
<td>MEP</td>
<td>Megakaryocyte-Erythroid Progenitor</td>
</tr>
<tr>
<td>MLL</td>
<td>Mixed Lineage Leukemia</td>
</tr>
<tr>
<td>MPP</td>
<td>Multipotent Progenitor</td>
</tr>
<tr>
<td>PcG</td>
<td>Polycomb Group</td>
</tr>
<tr>
<td>P-TEFb</td>
<td>Positive Transcription Elongation Factor</td>
</tr>
<tr>
<td>Pol II</td>
<td>RNA Polymerase II</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>SEC</td>
<td>Super Elongation Complex</td>
</tr>
<tr>
<td>TrxG</td>
<td>Trithorax Group</td>
</tr>
</tbody>
</table>
ABSTRACT

The AF9 gene is one of the most common chromosomal translocation partners of the *MLL (Mixed Lineage Leukemia)* gene in MLL leukemia. Wild-type AF9 is a member of the pTEFb transcription elongation complex, and interacts with gene regulatory proteins such as AF4/AF5q31, H3K79 histone methyltransferase DOT1L, Polycomb repressive protein Pc3/CBX8 and BCL-family co-repressor protein BCoR. These interactions are retained in the oncogenic MLL-AF9 fusion protein, and may be required for leukemic transformation.

Using bone marrow progenitor cells isolated from conditional *Af9* knockout mice, we examined *in vitro* differentiation of hematopoietic progenitor cells to the erythroid, myeloid and megakaryocytic lineages in the presence or absence of *Af9*. Based on previously published studies, we hypothesized that loss of *Af9* would result in a loss of erythroid colony formation. However, our results indicate no significant difference in erythroid, myeloid or megakaryocytic colony formation ability in cells expressing or null for *Af9*. Expression of transcription factor genes associated with erythroid or myeloid lineage commitment were also unaltered in cells null for *Af9* compared to control. There was, however, a statistically significant decrease in cell surface expression of early erythroid cell surface marker Ter119 in *Af9*-null cells compared to control mice, indicating a potential role for Af9 in early erythroid lineage commitment.
To examine the role of AF9-mediated protein interactions in MLL-AF9-mediated leukemic transformation, we engineered point mutants in the C-terminus of AF9 in the context of MLL-AF9 fusion protein predicted to disrupt specific protein-protein interactions. Amino acid substitutions were selected in collaboration with the laboratories of Dr. Hemenway and Dr. Bushweller based on structural studies and previously published results.

Hematopoietic progenitor cells were isolated from the bone marrow of wild type mice and retrovirally transduced with constructs encoding wild type or mutant MLL-AF9 fusion proteins, or empty vector control. *In vitro* colony assay experiments indicate that mutants MLL-AF9(D544R) and MLL-AF9(D546R) dramatically reduced colony formation ability and expression of MLL target genes *Hoxa9* and *Meis* compared to MLL-AF9(WT). Coimmunoprecipitation studies show that the D546R mutation in the context of amino acids 376-568 of AF9 significantly reduces binding to Dot1l while maintaining interaction with Af4, complementing biophysical studies performed in the Bushweller lab at the University of Virginia. Previous work by Dr. Bhavna Malik indicate that the D544R mutation in the context of AF9 amino acids 376-568 disrupts binding to Af4 but maintains binding to Dot1l.

These results indicate that disruption of AF9 interaction with AF4 and DOT1L decreases colony formation and target gene expression in *in vitro* leukemia assays. These interactions are attractive targets for designing small molecule inhibitors for the treatment of MLL leukemia.
CHAPTER ONE

INTRODUCTION

Blood is composed of four components: plasma, platelets, red blood cells, and white blood cells. Together, these cells are responsible for delivering oxygen to tissues and forming the immune system that protects against invading pathogens. All blood cells arise from a hematopoietic stem cell (HSC) that has the ability to differentiate to any mature hematopoietic cell.

Leukemia is a cancer of the blood or bone marrow, resulting from the accumulation of abnormal hematopoietic cells. DNA damaging mutations that occur in early hematopoietic stem or progenitor cells cause these abnormal cells to proliferate without properly differentiating to the necessary effector cells of the hematopoietic system. In 2013, over 48,000 people in the United States were diagnosed with leukemia, and over 23,000 people died from the disease\(^1\). Leukemia is also the most common cancer in children and adolescents, accounting for one-third of all cancer-related deaths in individuals 15 years old or younger.

Patients with leukemia involving the *MLL* (*Mixed Lineage Leukemia*) gene develop leukemia that is particularly aggressive and correlated with a poor prognosis\(^2\). MLL leukemias result from the chromosomal translocation of the *MLL* gene with one of over 79 fusion proteins\(^3\). This results in the expression of an oncogenic fusion protein that drives leukemic development of either acute myeloid (AML) or lymphoid (ALL)
phenotype. Understanding the mechanisms of leukemogenesis is necessary in order to develop targeted therapeutics to treat this heterogeneous disease.

AF9 is one of the most common fusion partners in MLL leukemia\(^4\). The wild-type AF9 protein is a member of the P-TEFb elongation complex that aids RNA Polymerase II during gene transcription\(^5\)-\(^7\). The C-terminus of AF9 directly interacts with at least four proteins: AF4\(^8\) and DOT1L\(^9\), which are considered activators of gene expression, and CBX8\(^10\) and BCoR\(^11\), associated with repression of gene targets. The overall goal of this project was to first, understand the role of Af9 in normal hematopoiesis, and second, to examine which protein interactions mediated by Af9 are required for leukemic development.

In the first aim of this study, we utilized a conditional Af9 knockout mouse to elucidate the role of Af9 in normal hematopoietic lineage commitment. Based on previously published data, we hypothesized that Af9 is required for erythroid and megakaryocytic differentiation, and that loss of Af9 in this system would result in a loss of these cells. To test this hypothesis, we isolated hematopoietic progenitor cells from the conditional Af9 knockout mouse and induced deletion of the Af9 gene \textit{in vitro}. These cells were then cultured \textit{in vitro} under conditions which promoted erythroid, myeloid or megakaryocytic lineage differentiation. Examining colony formation, expression of lineage-specific transcription factors, and cell surface marker analysis, we concluded that contrary to our hypothesis, Af9 expression is not required for normal hematopoiesis \textit{in vivo}. 
The second aim of this study examined the role of AF9 interactions with Dot1l and Af4 in leukemic development. In collaboration with structural biologists in the laboratory of Dr. John Bushweller at the University of Virginia, we engineered point mutations in the C-terminus of AF9 predicted to specifically disrupt binding to one protein while maintaining interaction with the other. These mutants were expressed in the oncogenic MLL-AF9 fusion protein that has been shown to transform bone marrow and cause rapid leukemia development when introduced into mice.

Results from this study indicate that AF9 interactions with Af4 and Dot1l are critical for in vitro transformation ability of an MLL fusion. Amino acid substitutions disrupting either Af4 or Dot1l binding to AF9 result in a loss of leukemic potential as measured by colony formation and MLL target gene expression. These structurally-informed point mutations identified critical residues mediating interactions between AF9 and AF4 or DOT1L that can be used to develop targeted therapeutics in the treatment of MLL leukemia.
CHAPTER TWO

LITERATURE REVIEW

Hematopoiesis

Hematopoiesis is the process of generating the blood and immune cells found within an organism from a common pluripotent hematopoietic stem cell (HSC). The hematopoietic system consists of three main groups of cells: red blood cells (erythrocytes), white blood cells, and platelets. Red blood cells compose the majority of blood, with 2-3 million new cells being produced in the bone marrow per second\(^\text{12}\). White blood cells are sub-divided into three types: lymphocytes, monocytes, and granulocytes. Lymphocytes include T and B cells that together coordinate the immune response to foreign pathogens. Monocytes, which mature into macrophages once they escape the blood and enter into a specific tissue, are responsible for phagocytosing invading pathogens as a first line of innate immune defense. Granulocytes include neutrophils, basophils and eosinophils. These cells have enzyme-containing granules that, when released, aid in digesting pathogens. Platelets are cell fragments shed from megakaryocytes that are responsible for forming blood clots. Maintaining the balance of hematopoietic cells circulating in the body is essential for delivering oxygen to tissues, protecting the organism from invading pathogens, and in maintaining hematopoiesis.
The Hematopoietic Hierarchy of Differentiation

All hematopoietic cells originate from HSCs in the bone marrow. The “true” HSC is the long-term hematopoietic stem cell, or LT-HSC. These cells are found at a very low frequency in the bone marrow, and exhibit lifelong self-renewal capability and differentiation potential to all hematopoietic lineages. The short-term hematopoietic stem cell (ST-HSC) still exhibits multilineage differentiation potential, but has a decreased capacity for self renewal.

Differentiation from an HSC to a progenitor cell requires a loss of self-renewal ability concurrent with an activation of genes that enforce differentiation. According to the classical model of hematopoiesis described by the Weissman group (Figure 1)\textsuperscript{13}, the earliest progenitor cell capable of differentiating to all hematopoietic lineages is the Multipotent Progenitor, or MPP\textsuperscript{13}. The Common Lymphoid Progenitor (CLP) and Common Myeloid Progenitor (CMP) arise from the MPP. As their names suggest, the CLP can form all cells of the lymphoid lineage but lose all myeloid potential, whereas the CMP can give rise to all cells from the myeloid lineage but lack lymphoid differentiation potential\textsuperscript{13}.

B and T cells terminally differentiate from the CLP, while the CMP gives rise to two additional progenitor cells. The Megakaryocyte/Erythroid Progenitor (MEP) gives rise to megakaryocytes and erythrocytes, or red blood cells (Figure 1). The Granulocyte/Monocyte Progenitor (GMP) gives rises to the granulocytes and monocytes\textsuperscript{13}, although a second study has supported the existence of a shared macrophage and dendritic cell precursor, the Macrophage/Dendritic Progenitor (MDP)\textsuperscript{14}. 
Figure 1. Weissman Model of the Hematopoietic Hierarchy

Canonical model of hematopoietic differentiation from the LT-HSC to terminally differentiated hematopoietic cell\textsuperscript{13}. Key transcription factors required for lineage specification are indicated.
While this is considered the classical model of hematopoiesis, a second model proposed by Adolfsson and colleagues is gaining supporting evidence\textsuperscript{15}. On the basis of $Flt3$ expression levels, a surface marker associated with early hematopoietic progenitor cells, these authors propose that the MEP arises directly from the ST-HSC without transit through the CMP (Figure 2)\textsuperscript{15}. In this case, the ST-HSC can give rise to either a Lymphoid Primed Multipotent Progenitor (LMPP) or the MEP directly. The LMPP terminally differentiates to cells of the lymphoid and myeloid lineage but lacks all erythroid potential, whereas the MEP gives rise to only megakaryocytes and red blood cells.

This hypothesis is supported by studies of a $SpiI$ knockout mouse. $SpiI$ encodes for the Pu.1 transcription factor necessary for myeloid lineage differentiation. Mice deficient for Pu.1 exhibit a loss of the CLP and CMP, but have a normal MEP compartment and normal numbers of erythroid cells\textsuperscript{16}. This supports the hypothesis that the erythroid lineage does not arise from a shared progenitor with the myeloid lineage, but diverts directly from an earlier progenitor or stem cell.

**Transcriptional Control of Hematopoietic Differentiation**

While the precise hematopoietic hierarchy is still under investigation, the mechanisms by which cells are signaled to differentiate is well understood, especially in the murine system. Genome-wide expression analyses as well as knock-out mouse models have identified several master regulators of hematopoietic differentiation.
Revised model of hematopoietic differentiation as described by Adolfsson et al. Contrasting to the Weissman model, the ST-HSC gives rise to a lymphoid-myeloid primer progenitor cell or a megakaryocytic/erythroid progenitor cell, diverting all megakaryocytic/erythroid development from a shared myeloid progenitor.
As mentioned previously, Pu.1 is required for myeloid lineage differentiation. The Pu.1 protein is a member of the ETS family of transcription factors. This protein binds to purine-rich DNA sequences termed the PU-box in the promoter of genes involved in myeloid differentiation. Deletion of Spi1 in mice results in a fatal loss of B cells and macrophages\textsuperscript{17, 18}.

Pu.1 is required not only for the maturation of HSCs to CMPs and CLPs (Figure 1), but for the repopulation of HSCs in chimeric mice transplanted with fetal liver cells\textsuperscript{16, 19, 20}. Spi1 expression is high in HSCs, CMPs, B cells, monocytes and granulocytes, and is downregulated during erythroid and T cell differentiation\textsuperscript{21}.

The maturation of the CMP to the GMP is controlled by expression of Cebp\alpha (Figure 1). CCAAT/Enhancer Binding Protein alpha, encoded by the Cebp\alpha gene, is a basic leucine zipper (bZIP) domain transcription factor that binds in a sequence-specific manner to DNA to regulate the transcription of myeloid differentiation genes. Cebp\alpha is expressed in HSCs, myeloid progenitor cells and granulocytes, but is not expressed in macrophages\textsuperscript{22, 23}. Genetic knockout of this gene results in normal numbers of CMPs but a lack of GMPs\textsuperscript{24, 25}, indicating a critical role for this protein in the differentiation of CMPs to GMPs.

GMPs have the potential to differentiate to granulocytes or monocytes. This decision is controlled by the expression of interferon-\gamma (IFN\gamma)-responsive transcription factor 8 (Irf8) (Figure 1). While the mechanism by which Irf8 controls myeloid differentiation is not well understood, results from a knockout mouse model highlight the
importance of this gene in monocytic versus granulocytic differentiation. Mice null for \textit{Irf8} had increased numbers of granulocytes but fewer macrophages, and developed a myeloproliferative syndrome resembling human chronic myeloid leukemia (CML)\textsuperscript{26}. Re-expression of \textit{Irf8} in myeloid progenitors isolated from the \textit{Irf8}\textsuperscript{-/-} mice shifted differentiation back toward macrophages\textsuperscript{27}.

Final maturation from the GMP stage to terminally differentiated granulocytes is controlled by the expression of \textit{Gfi1} and \textit{Cebp\varepsilon} (Figure 1). \textit{Gfi1} is expressed in HSCs, neutrophils, and early B and T cells\textsuperscript{28, 29}. Neutrophil differentiation was completely blocked in mice lacking \textit{Gfi1} expression\textsuperscript{28, 29}. Similarly, \textit{Cebp\varepsilon} is expressed in myeloid and lymphoid cells, and mice null for \textit{Cebp\varepsilon} also exhibited abnormal granulocyte maturation\textsuperscript{30}.

The divergent branch stemming from the CMP is the MEP, giving rise to erythrocytes and megakaryocytes (Figure 1). As mentioned above, recent studies have also supported the emergence of the MEP directly from the ST-HSC, bypassing a shared progenitor with granulocyte/macrophage potential altogether (Figure 2)\textsuperscript{15}. There are several major transcription factors involved in the regulation of erythroid/megakaryocytic differentiation, but the GATA family of transcription factors plays a central role. The members of this family bind to a consensus DNA sequence through a conserved C-terminal zinc finger motif. A second conserved zinc finger is found at the N-terminus and stabilizes binding to the DNA.
*Gata2* is highly expressed in HSCs and early progenitors, but is downregulated as cells begin to differentiate to erythroid or megakaryocytes. As *Gata2* levels decrease, *Gata1* levels increase. This is termed the GATA-switch, and is required for terminal differentiation of the MEP.

The N-terminal zinc finger of GATA1 has been shown to interact with several proteins including FOG-1 (Friend of Gata-1). FOG-1 is also a zinc finger-containing transcription factor, and exhibits an expression pattern that mimics GATA1. Both proteins are required for erythroid and megakaryocytic development, as is evident by knock-out gene studies in mice. *Gata1*−/− mice are embryonic lethal between E10.5-11.5, and die of lack of primitive erythropoiesis. Conditional deletion of *Gata1* in mice resulted in thrombocytopenia and a delay in megakaryocyte differentiation, as well as a lack of red blood cells. Similarly, *Fog*−/− null mice exhibit delayed megakaryocytic development. It is the interaction of Gata1 and Fog-1 that is thought to downregulate expression of *Gata2* to initiate terminal differentiation of progenitor cells. Interestingly, mutations in the N-terminus of GATA1 that prevent interaction with FOG-1 have been found in patients with X-linked disorders of megakaryocytic development.

After the GATA switch occurs, additional transcription factors are necessary to promote differentiation to erythrocytes or megakaryocytes. While the Gata1/Fog-1 complexes are necessary for lineage specification of both erythrocytes and megakaryocytes, it is the presence of additional factors that determine the specificity. Interaction of Gata1/Fog-1 with the Ets family transcription factors Fli1 and Tel are essential for megakaryopoiesis, whereas interaction with zinc finger transcription factors
factor Eklf (Klf1) is essential for erythropoiesis (Figure 1)\textsuperscript{48}. Fli-1 and Eklf directly interact with and inhibit each other, as Eklf binding to Fli-1 alters the conformation of the Fli-1 DNA binding domain, preventing it from binding to its target genes \textsuperscript{49}.

More recent studies have added additional levels of complexity to the transcriptional regulation of erythroid and megakaryocytic differentiation. Chromatin immunoprecipitation (ChIP) studies followed by massively parallel sequencing (ChIP-seq) in human erythroleukemic cell lines have shown that GATA1 binding most often occurs not at promoters, but at distal enhancer elements \textsuperscript{50-52}. Interestingly, studies of the chromatin occupancy of GATA1-associated transcription factor EKLF indicate the vast majority of binding also occurs more than 10kb away from known genes, but within 1kb of GATA1 localization \textsuperscript{53}. Further analysis of chromatin conformation using the circularized chromatin conformation capture (4C) technique suggests that EKLF assists in localizing actively transcribed chromatin to transcription factories \textsuperscript{54}. Additionally, FOG-1 has been shown to immunoprecipitate with the NuRD chromatin remodeling complex\textsuperscript{55}, providing yet another mechanism for GATA1-FOG1 complexes to regulate target gene expression.

Divergent from myeloid differentiation from the MPP is the lymphoid branch of the hematopoietic hierarchy. MPP differentiation to the common lymphoid progenitor (or the lymphoid-primed multipotential progenitor according to the Adolfsson model) correlates with an increase in expression of both Flt3 and of the \(\alpha\) chain of the interleukin 7 receptor (IL-7R\(\alpha\))\textsuperscript{56-59}, although signaling via IL-7R is not required for the generation of CLPs. Mice deficient in Flt3 have HSCs that are unable to differentiate to
myeloid and lymphoid cells, and mice deficient for Il7Ra expression show a three-fold
decrease in the number of CLPs\textsuperscript{60,61}(Figure 1).

Additional factors thought to be involved in the maturation to the CLP/LMPP
from the MPP include Pu.1 and Ikaros (Figure 1). As mentioned previously, mice null for
Spi1, the gene encoding the Pu.1 protein, die during embryogenesis or perinatally, and
exhibit defects in myeloid and lymphoid differentiation\textsuperscript{17,18}. These mice do not have
detectable T cells in the thymus, nor do they have B cells in the liver, suggesting that
Pu.1 acts at the levels of a progenitor common to both B and T cell development\textsuperscript{17,62}.

Ikaros is a zinc finger-containing transcription factor that, like Eklf, is associated
with the NuRD chromatin remodeling complex as well as the SWI-SNF complex\textsuperscript{63,64}. While Ikaros has been shown to be both an activator and a repressor\textsuperscript{65,66}, its activity is
essential for lymphoid differentiation. Mice lacking Ikaros completely lack B cells and
have impaired fetal T cell development, although adult T cell differentiation is
unaffected\textsuperscript{67}.

B cell differentiation occurs through several steps, each of which requires a
specific transcriptional regulatory program. B cells proceed through a series of
maturation stages, from the early lymphoid progenitor to the pre-pro-B cells, to the pro-B
cells, and finally to the pre-B cell prior to mature B cells. The major player in
commitment to the B cell lineage is Pax5, which is stably expressed throughout all stages
of B cell maturation\textsuperscript{68}(Figure 1). Loss of Pax5 expression results in an arrest of B cell
development at the pro-B stage\textsuperscript{69}. Pax5 has been shown to both activate target genes
required for B cell maturation such as \textit{Lef1} and \textit{Irf8}, and to repress target genes including \textit{Flt3}, the previously described CLP-associated marker \cite{70,71}.

\textbf{T lymphopoiesis is equally complex. T-progenitor cells mature in the thymus, and are marked by expression of CC-chemokine receptor 9 (CCR9), stem-cell antigen 1 (SCA1), and FLT3 \cite{72,74}. Early T cell progenitors mature through several steps of differentiation, from the early T-cell lineage progenitor (ETP), to the double CD4, CD8 negative 2a (DN2a) cells, then to DN2b, DN3a, DN3b, DN4, double CD4, CD8 positive (DP), to the T cell receptor positive DP cells, and finally to terminally differentiated CD4+ or CD8+ T cells. Each maturation stage requires expression of a specific combination of transcription factors, but initiation of the entire T lineage differentiation program requires Notch signaling (Figure 1). Expression of constitutive Notch leads to T cell-Acute Lymphoblastic Leukemia (T-ALL) in mice, whereas knockout of Notch target gene \textit{Rbpj} results in a loss of all T cell development \cite{75,76}. Additional transcription factors play important roles at later stages of T cell differentiation, including GATA3, MYB, and TCF1 \cite{77,83}.

It is clear that differentiation from hematopoietic stem cell to a terminally differentiated cell is very complex, involving an intricate network of tightly regulated transcription factors that turn on and off target genes necessary for lineage maturation. More recent studies have begun to examine the mechanisms by which gene expression is controlled at the epigenetic level, or how DNA accessibility affects which genes can be turned on and off by the above described transcription factors. The mechanisms that
regulate these epigenetic modifications are as critical as the transcription factors that control proper hematopoietic development.

**Chromatin and Epigenetics**

The human body is composed of over 200 different cell types that work in concert to support a living human being. Each of these cells has 46 chromosomes containing over 3 billion base pairs of DNA. Stretched out end-to-end, the amount DNA in each cell would reach over two meters. In order to efficiently pack this much DNA within the nucleus of every cell in the body, a complex organization system has evolved.

Chromatin refers to the packaging of DNA around proteins in the nucleus of a cell. At the most basic level, approximately 146 base pairs of DNA wrap around a core of eight histone proteins: two each of histone H2A, H2B, H3, and H4. The DNA wraps around the octamer core approximately 1.7 times. Between the DNA-wrapped histone proteins are approximately 20-60 base pairs of DNA that can associate with a fifth histone protein, the linker histone H1, forming a nucleosome. When viewed by electron microscopy, this level of DNA organization, approximately 10nm in diameter, resembles beads on a string with the DNA-wrapped histone octamer being the bead and the spacer DNA being the string. Nucleosomes further fold upon each other forming 30nm chromatin fiber up to 300nm fibers.

The extent to which nucleosomes are compacted, and how tightly the DNA is wrapped around the histone octamer, determines, in part, whether genes are actively transcribed or not. Only loosely packed, accessible DNA can be targeted by the DNA transcription machinery and transcribed. Tightly wound, compacted DNA is not
accessible to the transcription machinery and therefore is not active. These regions are termed euchromatin and heterochromatin, respectively.\(^{85}\)

Chromatin can be compacted and expanded by ATP-dependent chromatin remodeling complexes which contain proteins with ATPase enzymatic domains. The energy from ATP hydrolysis is used to restructure or shift the nucleosome along the DNA, or to expel nucleosomes altogether to make the DNA accessible to the transcription machinery.

The stable inheritance of a phenotype resulting from changes in chromosome structure without alterations in the DNA sequence itself is termed epigenetics.\(^{86}\) Proteins involved in placing epigenetic marks can be activators that turn on gene expression, or repressors that turn off gene expression. It is through epigenetics that cells can adopt specific functions. As discussed above, each cell in the body contains exactly the same DNA. It is the combination of genes that are expressed and repressed that gives a cell identity and function.

One mechanism of epigenetic regulation is though modification of the histone proteins themselves. Histone proteins have flexible N-terminal tails that extend beyond the octamer core\(^ {84}\) that can be post-translationally modified by proteins.\(^ {87}\) Such modifications include acetylation, phosphorylation, methylation, ubiquitylation, sumoylation, and ADP ribosylation. Modification of histone tails can serve different purposes, from changing the charge in the tail to alter DNA-protein or nucleosome-nucleosome interactions, or to providing a binding surface for proteins to localize to the histone and carry out additional functions.
The residue of the histone tail that is modified, and the modification itself can determine the functional outcome \(^{85}\). While these marks are vast and varied, a few have been very well studied. For example, di- or tri-methylation of histone H3 at lysine 9 (H3K9me2/3) within the gene body, and tri-methylation of H3K27 in the promoter region are associated with transcriptionally repressed genes. Tri-methylation of H3K4 in the promoters of genes is associated with transcriptional activation\(^ {88}\). Understanding the function associated with these marks, and the proteins that place these marks, is critical for understanding complex gene regulation.

A second mechanism of epigenetic regulation is DNA methylation. DNA methylation occurs on the fifth carbon of cytosine rings in CG dinucleotides \(^ {89}\). Cytosines located 5’ to guanine nucleotides, termed CpG for cytosine-phosphate backbone-guanine, are represented less than statistically predicted in the genome \(^ {90}\). Methylated cytosines can spontaneously undergo a deamination reaction, resulting in a transition mutation to thymine. This process has been evolutionarily selected against, thus the CG dinucleotide exhibits a decreased representation throughout the genome.

CpGs are, however, found at an observed-to-expected ratio of >60% in the promoters of approximately 70% of human protein coding genes \(^ {91}\). Clusters of CG dinucleotides making up >50% GC content in a 300-3,000 base pair range of DNA are termed CpG islands. These clusters are mostly unmethylated, and the genes associated with these promoter regions are transcribed\(^ {92}\). Methylation of these CpGs by DNA methyltransferase enzymes is associated with transcriptional repression, and these methyl marks are maintained throughout cell division by maintenance DNA methyltransferase
enzymes, thus preserving the silencing of the gene\textsuperscript{93}. DNA methylation is often exploited by cancer cells, in which the promoters of tumor suppressor genes can be methylated and therefore silenced. Conversely, the promoters of developmental genes that are silenced as cells differentiate may become unmethylated, allowing for expression of genes that should otherwise be turned off.

**MLL**

The MLL (Mixed Lineage Leukemia) protein is a member of the Trithorax group of proteins that positively regulate gene transcription by modifying chromatin structure, among other functions. MLL is an ortholog of the *Drosophila melanogaster* Trithorax protein, so named for the transformation of abdominal segments that occurs when the gene is mutated\textsuperscript{94}. Further research identified a role for trithorax in the positive regulation of homeotic genes that determine the spatial body plan during development\textsuperscript{95}.

Studies in mice further elucidated a role for *Mll* in development and hematopoiesis. A lacZ reporter gene was inserted in-frame in exon 3 of mouse *Mll*, located just downstream of the AT-hooks shown to be required for *trithorax* activity in *Drosophila*, resulting in a loss of Mll activity. β-galactosidase staining indicated that *Mll* is widely expressed in the developing embryo from the first time point examined (E7.0) into adulthood, with highest expression levels found in the nervous system and the somites that give rise to mature cells of the mesoderm lineage. These regions are regulated by the expression of *Hox* genes, indicating an overlap between *Mll* expression and regions regulated by *Hox* genes in the developing mouse\textsuperscript{96}. 
Wild-type, $Mll^{+/+}$ and $Mll^{-/-}$ mice were generated to study the function of Mll. $Mll^{-/-}$ mice died during embryonic development by E10.5, indicating that expression of $Mll$ is essential for viability. Further examination of the $Mll^{+/+}$ mice showed that $Mll$ is a haploinsufficient gene; expression from only one allele does not produce enough protein product to achieve a wild-type phenotype. These mice were smaller at birth compared to the wild-type mice and had retarded growth. In addition, the $Mll^{+/+}$ mice also exhibited skeletal abnormalities. Both anterior and posterior homeotic transformations were present, similar to those seen in $Hox$-deficient mice\(^96\).

As Trithorax group (TrxG) proteins are known to positively regulate $Hox$ expression, whole-mount embryo hybridization and RNA \textit{in situ} hybridization were used to assess any alterations in $Hox$ gene expression upon loss of $Mll$. Reduced boundaries of expression were evident for Mll targets $Hoxc9$ and $Hoxa7$ in the $Mll^{+/+}$ mice, and expression was absent altogether in E10.5 embryos from $Mll^{-/-}$ mice, indicating that Mll is required for the proper maintainence of Hox gene expression\(^96\).

$Mll^{+/+}$ mice also exhibited defects in hematopoiesis such as anemia, a reduction in platelets, and reduced numbers of B-cells\(^96\). The role of Mll in hematopoiesis was further studied by \textit{in vitro} differentiation of yolk sac progenitor cells isolated from wild-type, $Mll^{+/+}$ and $Mll^{-/-}$ mice. Cells lacking Mll produced fewer colonies containing smaller cells than heterozygous and wild-type controls, indicating that Mll is required for maintenance of the both the numbers of hematopoietic progenitors and for their proper differentiation into mature hematopoietic cells, particularly the myeloid and macrophage lineages\(^97\).
Conditional loss of *Mll* in the hematopoietic lineage did not alter the steady state levels of mature blood cells in adult mice. However, when cultured in methylcellulose, the conditional *Mll* knockout cells exhibited a reduced ability to form myeloid and lymphoid colonies compared to control cells. When injected into lethally irradiated recipient mice, the conditional knockout cells failed to reconstitute the bone marrow, indicating that Mll is required for adult HSC self-renewal.

Similarly, hematopoietic lineage-restricted inducible loss of *Mll* in adult mice caused rapid and fatal cytopenia within 19 days. Loss of *Mll* in the hematopoietic lineage resulted in a loss of hematopoietic stem and early progenitor cells, but more mature, lineage-committed progenitor cells were unaffected. A recent study identified a network of genes specifically regulated by MLL in HSCs. These genes, including *Prdm16*, *Pbx1*, and *Hoxa9*, are involved in the self-renewal and proliferative capacity of HSCs. The results of these studies indicate a critical role for MLL in the maintenance of hematopoietic stem cells.

**MLL Protein Domains**

It is thought that MLL regulates expression of genes necessary for development and hematopoiesis, at least in part, through its chromatin modifying ability. The C-terminus of MLL contains a histone methyltransferase domain with specificity for histone H3 lysine 4. Trimethylation of H3K4 is associated with positive gene expression, although it is still debated whether the placement of this mark initiates gene transcription, or if it is just associated with genes that are already being activity transcribed.
There are many other important protein domains of MLL that govern its role in gene regulation (Figure 3A). The *MLL* gene, located at chromosome 11 band q23, encodes a full length protein of 3,969 amino acids that is post-translationally cleaved by taspase-1 into a 320 kilo-Dalton (kDa) N-terminal and 180 kDa C-terminal fragment that remain associated non-covalently via the FYRN and FYRC domains \(^{103,104}\) (Figure 3B).

The N-terminus of MLL contains a motif that binds the protein Menin. Menin directly binds LEDGF, a protein with a PWWP DNA binding domain. MLL, through its interaction with Menin, is bridged to LEDGF and recruited to target DNA sequences. Loss of the Menin binding motif of MLL in the context of a MLL fusion protein results in decreased transformation ability \(^{105}\).

Three A-T hooks are located just downstream from this region. A-T hooks recognize and bind to bent, A-T rich DNA sequences \(^{106}\). This region has been shown to interact with SET, protein phosphatase 2A (PP2A) and GADD34 proteins \(^{107,108}\). The A-T hooks precede two speckled nuclear localization domains. These domains are required for localizing MLL to punctuate structures within the nucleus \(^{109,110}\).

The Repression Domain of MLL was identified in a study in which this region of the protein was shown to have repressive activity in a reporter assay \(^{106}\). Additional studies have found that corepressor proteins such as CtBP, BMI-1 and CBX4 (HPC2) bind to this region, as does Histone Deacetylase 1/2 (HDAC1/2) \(^{111}\). Recent studies have shown that the Polymerase Associated Factor (PAF) complex, associated with gene activation, also binds to this region \(^{112,113}\).
Located within the Repression Domain of MLL is the CXXC domain, so named for the conserved spacing of eight cysteine resides. The CXXC domain was identified for its homology to the CXXC domain of the DNA Methyltransferase 1 (DNMT1) protein\textsuperscript{114}. Studies from our lab and others have shown that the CXXC domain binds to unmethylated CpG DNA, protecting it from DNA methylation. The ability of the CXXC domain to bind unmethylated DNA is required for transformation by the MLL-AF9 fusion protein in MLL leukemia\textsuperscript{115,116}.

Downstream of the CXXC domain is an atypical bromodomain and four plant homodomain (PHD) zinc fingers. The second of the four PHD fingers is involved in homodimerization of MLL\textsuperscript{117}. The third PHD finger has been shown to bind to H3K4 trimethyl marks or the cyclophilin CYP33 in a mutually exclusive manner\textsuperscript{118,119}. The ability of MLL to both “read” histone marks and “write” the H3K4 methyl mark on histones is thought to promote the maintenance of histone modifications.

The C-terminus of MLL contains a transactivation domain (TAD) that is associated with coactivator protein CBP binding\textsuperscript{120}. And, as mentioned previously, the most carboxy-terminal protein domain in the wild type MLL protein is the SET domain with H3K4 histone methyltransferase ability. Together, the multiple functional domains of MLL provide the protein with a multitude of ways to participate in gene regulation, both with activating and repressive functions.
Figure 3. Protein Domains of MLL and MLL Fusion Proteins

(A) Domain schematic of the human MLL protein with protein domains and corresponding amino acids indicated. Alternate names and chromosomal location are noted. (B) After cleavage by Taspase 1, the N-terminal and C-terminal fragments of MLL associate through the FYRN and FYRC domains. (C) MLL fusion proteins express the N-terminus of MLL through the breakpoint cluster region and the C-terminus of one of 79 known fusion partner proteins.
MLL Target Genes

The *HOX* genes are very well-studied gene targets of MLL. HOX proteins are DNA-binding transcription factors that bind DNA via a homeodomain located at their C-terminus. HOX proteins function to regulate body segment identity as well as determine embryonic cell fate. Expression of HOX genes is positively regulated by the Trithorax group of proteins including MLL, whereas expression is repressed by the Polycomb group of proteins\textsuperscript{96, 121, 122}.

These genes are highly evolutionarily conserved, so much so that loss of a *Hox* gene in *Drosophila melanogaster* could be rescued by re-expression of the chicken ortholog\textsuperscript{123}. In humans and mice, the *HOX* genes are divided into four clusters: A, B, C and D, located on four separate chromosomes. Within these clusters, there are 13 paralogous groups totaling 39 genes in whole\textsuperscript{124, 125}. The *HOX* genes are regulated spatially and temporally such that the genes located 3’ in the clusters are transcribed earlier in development and at the anterior end of the organism. Conversely, genes located 5’ in the clusters are transcribed later in development and in the posterior end\textsuperscript{124}.

*HOX* genes are also critical regulators of hematopoiesis. Genes in the A, B and C cluster are highly expressed in hematopoietic stem and early progenitor cells, and expression of these genes is decreased as the cells differentiate and mature. Overexpression of *HOX* genes such as *HOXA9* is a characteristic of certain types of leukemia, and coexpression of *Hoxa9* and *Meis1* is sufficient to transform normal mouse myeloid progenitor cells\textsuperscript{126}.
MEIS1, another target of MLL, is a member of the three amino acid loop extension (TALE) group of proteins that also contain a homeodomain. Like HOX proteins, MEIS1 plays a role in development and hematopoiesis, although its role is less well understood. Similar to HOX genes, MEIS1 expression is highest in early hematopoietic stem and progenitor cells, and expression decreases as cells differentiate. Meis1−/− mice die at E14.5 and exhibit a lack of megakaryocytes and subsequently, extensive hemorrhaging. MEIS1, like HOXA9, is overexpressed MLL-rearranged leukemias.

In addition to MEIS1 and HOX genes, more recent genome-wide studies have identified a role for MLL as a potential master regulator of gene transcription. Chromatin immunoprecipitation studies have examined the binding of MLL at loci throughout the genome. One of the earliest reports found MLL bound at the promoter of 5,186 genes in U937 monocytic cells with coincident binding of RNA Polymerase II at 90% of the genes, and H3K4 trimethylation at 92% of the genes. However, a separate study used Mll+/+ and Mll−/− murine embryonic fibroblasts (MEFs) to further examine the correlation between Mll localization, H3K4me3 and gene activation. The authors found peaks of H3K4me3 at 10,041 genes in the wild type cells, with a loss of H3K4me3 at only 525 promoters, or approximately 5% of genes, in the Mll−/− cells. These results suggest that Mll maybe be responsible for H3K4me3 at only a subset of genes.

Another study comparing the localization of MLL and MLL fusion proteins found wild-type MLL bound at the promoter and 5’ region of 2,595 genes. While the exact number of genes regulated by MLL is still being examined, it is clear that MLL plays a
critical role in regulating specific target genes involved in development and hematopoiesis.

**MLL Leukemia**

The transcriptional activation role of MLL is exploited in a specific type of leukemia termed MLL Leukemia. Patients with MLL leukemia express oncogenic fusion proteins that result in the aberrant transcription of MLL target genes such as the HOX genes described above (Figure 3C). Additionally, amplifications of the MLL gene are found in patients with AML and myelodysplasia, and partial tandem duplication of amino-terminal exons of MLL are found in approximately 10% of all AML cases $^{135,136}$.

In MLL fusion genes, a double strand break in the DNA occurs within the breakpoint cluster region (BCR) of the MLL gene. As these breaks are deleterious to a cell, mechanisms of DNA repair are employed. In this case, non-homologous end joining (NHEJ) fuses the broken MLL gene to one of over 79 translocation partner genes $^3$(Figure 3C). The cause for breakage of MLL within the BCR, as well as the cause for breakage within the fusion partner, is not well understood. While there is little sequence homology within the breakpoint cluster regions of both MLL and the known fusion partners, the BCRs share several structural elements including topoisomerase II cleavage sites, DNAse I hypersensitivity sites and scaffold attachment regions that serve to associate proteins involved in transcription and replication $^{137-140}$. These regions may be more susceptible to breakage and subsequent aberrant repair by NHEJ.

MLL translocations are found in approximately 10% of all Acute Myeloid or Lymphoid Leukemias diagnoses, but the incidence of MLL leukemia is higher in infants.
Approximately 70% of infant ALL and 35-50% of infant AML patients express MLL fusion genes. While the five year event free survival (EFS) in infants with ALL without MLL recombination is 74-96%, infants with ALL with a MLL recombination have a five year EFS of only 19%.

MLL fusions are also present in a subset of leukemias that arise after treatment for a primary cancer, termed therapy-related MLL leukemia. Cells require expedient replication of DNA for cell division to occur. As the double helix unwinds to separate the strands for replication, tension builds in the DNA proximal to the replication fork and if not relieved, would halt the DNA polymerase. Topoisomerase enzymes introduce breaks in the DNA phosphodiester backbone to allow for unwinding of the DNA to alleviate the tension, and then reanneal the cut DNA back together. Topoisomerase enzymes play a similar role in RNA transcription.

Topoisomerase II inhibitors such as etoposide, doxorubicin and daunorubicin are often administered to cancer patients to block the activity of the topoisomerase enzyme in rapidly dividing cancer cells, notably preventing the reannealing of broken DNA strands. This triggers the DNA damage response pathway and induces apoptotic cell death. The 11q23 locus has been shown to be highly susceptible to double strand breaks, and inhibition of topoisomerase enzymes prevents repair at this locus. Aberrant repair of the broken DNA by NHEJ can then result in the production of oncogenic fusion proteins. Approximately 5-10% of all MLL leukemias are therapy-related, almost all of which are AML.
While almost 80 MLL fusion partners have been characterized, over 85% of all MLL leukemias involve one of 6 partner genes: AF4, AF9, ENL, AF10, ELL and AF6. These proteins belong to one of three gene families, all of which are involved in transcription regulation. Of note, new nomenclature has been instituted for genes in these families: AF4/AFF1, AF9/MLLT3, ENL/MLLT1, AF10/MLLT10, and AF6/MLLT4. However, as much of the literature regarding these genes references the original gene names, they will be used throughout this dissertation.

Results from several studies support the hypothesis that the fusion partner dictates the leukemic phenotype. For example, MLL-AF9 fusion proteins are highly associated with AML, whereas MLL-AF4 fusions often given rise to ALL or acute biphenotypic leukemia (ABL). Retroviral transduction of MLL-AF9 transforms both early hematopoietic progenitors and committed myeloid progenitors, and when transplanted into mice, results in proliferation and accumulation of myeloblasts. Similarly, a knock-in MLL-AF9 mouse model in which expression of the MLL-AF9 fusion protein is under control of the endogenous Mll promoter also leads to development of acute myeloid leukemia in the majority of the mice.

Conversely, knock-in of the MLL-AF4 fusion resulted in mice with a mixed lymphoid and myeloid phenotype, with a predominance of B220+ or CD19+ B lymphoid cells. These mice developed hematologic malignancies with a median of 520 days compared to 220 days for mice with a knock-in of the MLL-AF9 fusion. These results support the hypothesis that the fusion partner is responsible for dictating the leukemic
phenotype. However, additional studies suggest that the microenvironment is the critical regulator of myeloid versus lymphoid leukemic fate.

Wei et al examined the leukemic phenotype of MLL-AF9-transduced CD34+ human cord blood cells cultured under myeloid or lymphoid cytokine conditions in vitro and in vivo. When cultured in the presence of myeloid cytokines, cells expressed myelomonocytic surface markers. However, switching these myeloid-conditioned, long term-cultured MLL-AF9 cells to media promoting erythroid or lymphoid development failed to differentiate the cells, as indicated by decreased surface marker expression of lineage-specific markers compared to control CD34+ cells. Interestingly, when MLL-AF9 transduced CD34+ cells were cultured under conditions to promote B cell differentiation, the cells exhibited a mixed lineage phenotype, suggesting that the cytokine environment plays a critical role in determining the leukemic phenotype of these cells.

To examine the role of the microenvironment on leukemogenesis in vivo, MLL-AF9-expressing CD34+ cells were either directly injected after retroviral transduction into mice, or were cultured for an extended period of time in the presence of myeloid or lymphoid cytokines prior to injection into mice. Non-obese Diabetic/Severe Combined Immunodeficiency (NOD-SCID) or NOD-SCID β2-microglobulin−/− mice injected with MLL-AF9-transduced cells cultured under myeloid conditions only developed AML with a latency of 10 weeks. Injection with cells cultured under lymphoid conditions resulted in mice with AML, B-ALL or Acute Biphenotypic Leukemia (ABL) with a latency of approximately 13.5 weeks. Direct injection of newly-transduced cells caused mice to
develop B-ALL or ABL, but no mice developed AML. However, injection of these same cells into mice transgenic for IL-3, SCF and GM-CSF expression resulted in only AML development in mice within 7 weeks. Similarly, injection of MLL-AF9 cell lines cultured under myeloid or lymphoid conditions into these transgenic mice resulted in development of AML exclusively within 5 weeks. These results suggest that it is not the fusion partner per se, but the microenvironment in which the fusion-expressing cell is located that determines the leukemic phenotype 149.

While the interplay between the function of the partner brought into MLL fusion genes and the microenvironment in which they are expressed is still being elucidated, it is clear that MLL fusion proteins are key drivers of leukemogenesis. MLL-AF9 fusions in particular are highly leukemogenic, and understanding the mechanism by which AF9 functions in wild-type cells is necessary understanding how it contributes to misregulated hematopoiesis in MLL leukemia.

AF9

The AF9 gene, located at chromosome band 9p22, is fused to MLL in approximately 25% of de novo AML in children and 2-5% of all AMLs 4. Patients with MLL-AF9 translocations have a median survival of four years 4. Understanding the function of AF9 is necessary to understanding its role in MLL leukemia in order to develop efficacious therapeutics.

AF9 is a member the YEATs (Yaf9, ENL, AF9, Taf1, Sas5) domain-containing family of proteins, so named for the conservation of the domain in both yeast (Yaf9, Taf1, Sas5) and mammalian (AF9, ENL) proteins. AF9 and ENL are highly homologous
to each other, sharing approximately 50% amino acid identity overall and 80% identity in conserved N and C terminal regions\textsuperscript{150}, but these proteins show little homology to other mammalian gene families (Figure 4A).

The YEATs domain is located at the N-terminus of both AF9 and ENL (Figure 4A, B). This domain is often found in proteins with chromatin-modifying ability\textsuperscript{151}. As is further discussed in the review on transcription elongation, the YEATS domain of AF9 and ENL interact with the Polymerase Associated Factor complex (PAF\textsubscript{c}), which recruits the Super Elongation Complex (SEC) to poised polymerases, facilitating productive elongation of nascent RNA\textsuperscript{152,153}.

A nuclear localization signal (NLS) is located downstream of the YEATs domain and allows for translocation of the protein from the cytoplasm to the nucleus. This region is also proximal to a serine/proline rich region. The C-terminus of AF9 encodes a 94 amino acid AHD domain [actin non-complementing gene 1 budding yeast (Anc1) homology domain] involved in protein-protein interactions\textsuperscript{154}. This region of AF9 is highly conserved with ENL, and to a lesser degree, to the yeast ANC1\textsuperscript{155}. It has been shown to interact with at least four other proteins: AF4, DOT1L, CBX8 and BCoR in mammals\textsuperscript{8-11} (Figure 5).
Figure 4. Protein Domains of AF9 and ENL

Protein domain schematic of (A) AF9 and (B) ENL with alternate names and chromosomal location noted. Lines below the protein structure indicate the approximal fragment of the protein present in MLL fusions. (C) ) Solved NMR structure of AF4 peptide in complex with the C-terminal AHD of AF9 (figure from Leach 2013). (D) Amino acid sequence of a segment of the AF9 region known to interact with AF4, DOT1L, CBX8 and BCoR is highly conserved in homolog ENL. The bolded aspartic acid mediates an electrostatic interaction with AF4 family members.
Recent studies by Leach et al showed that the AHD of AF9 is highly disordered with a structure that is almost entirely random coil with a minimal beta structure in its unbound state\textsuperscript{156}. The addition of an AF4 peptide corresponding to the known AF9-interacting sequence \textsuperscript{157} resulted in a structural rearrangement of the AHD, forming the alpha-beta structure shown in Figure 4C. A beta strand in AF4 residues 761-766 is located next to the beta hairpin of AF9 residues 535-546, together forming a three-stranded antiparallel beta sheet. AF9 also forms three alpha helices around the AF4 peptide\textsuperscript{156}(Figure 4C).

At least one electrostatic interaction between AF9(D544) and AF4(K764) exists (Figure 4D, 6B). This interaction is the focus of further studies presented in this dissertation. The interface between AF9 and AF4 is otherwise primarily hydrophobic, mediated by several aliphatic residues within the region. Interestingly, while the sequences of the proteins that interact with AF9 are highly dissimilar, DOT1L, CBX8 and BCoR proteins all have conserved spacing of the aliphatic residues present in AF4 that mediate the AF4-AF9 interaction. Coexpression of AF9 and any one of the four binding partners resulted in highly similar chemical shifts, suggesting that AF9 adopts a similar structure when bound to any of the four proteins, and that the binding partners all share a common binding site on AF9\textsuperscript{156}. These results support the hypothesis that the C-terminus of AF9 remains highly disordered until binding to a partner protein. As the binding partners of AF9 are considered both transcriptional activators (AF4, DOT1L) and repressors (CBX8, BCoR), it is plausible that the interaction of AF9 with one of these proteins may allow for differential regulation of genes targeted by AF9.
The AHD of AF9 has been shown to interact with at least four known proteins: CBX8, DOT1L, BCoR and AF4/AF5q31. Shaded regions of the interacting protein denote approximate region of AF9 interaction. CBX8 is a member of the PRC1 repressive complex. DOT1L is the only known H3K79 histone methyltransferase. BCoR has been shown to associate with BCL-6 to repress gene targets. AF4/AF5q31 are associated with P-TEFb, and with additional proteins (not shown) form the Super Elongation Complex (SEC).

To further investigate the functional role of Af9, a constitutive knockout mouse was engineered by disrupting the second exon of Af9 with a lacZ-neo cassette. Af9<sup>−/−</sup> mice were normal and fertile<sup>158</sup>. Breeding with these mice produced homozygous offspring at
the expected ratio, indicating that the homozygous null Af9 mice survive embryogenesis. However, 50% of these Af9+/− mice were found dead at birth or died within hours after. The remaining surviving mice were severely runted and died within two weeks. These mice did not exhibit any gross external malformations, nor did they have internal changes in organ structure. Hematopoiesis appeared normal in the Af9+/− mice as assessed by histology and flow cytometry158.

Further examination of whole-skeleton preparations showed that the homozygous mutants exhibited axial skeletal abnormalities, lacking a pair of floating ribs and the associated vertebrae. The Af9+/− mice also had deformed sternums, and often had extra cervical vertebrae compared to wild type controls. These malformations, in addition to other subtle changes, indicated that anterior homeotic transformation of the axial skeletal had occurred in mice lacking Af9 expression comparable to the phenotype of Mll+/− mice96. Lumbar and caudal vertebrae, and the appendicular skeleton were normal158.

β-galactosidase staining of embryos taken from heterozygous mice between E10.5 and E14.5 indicated widespread Af9 expression in the mouse including the skeletal system, neural system, gut, forming genital organs, in the caudal hindbrain and at the midbrain-hindbrain border. Slight differences in the expression pattern of B-galactosidase were seen in the homozygous null embryos, with decreased staining at the anterior cervical region compared to the heterozygous embryos, as well as more broad lateral expression in this region compared to the Af9+/− embryos158.

The widespread expression pattern of Af9 in developing tissues and the homeotic transformations seen upon loss of Af9 suggested a role for Af9 as a master regulator of
early embryonic development, similar to Mll. As proper expression of Hox genes is critical for normal development, the authors examined the expression patterns of several Hox genes in wild-type, heterozygous and homozygous null Af9 embryos. The expression of Hoxd4 in Af9+/− embryos was reduced and did not reach the anterior boundary compared to wild-type or Af9+/− E9.5 embryos. This altered expression pattern was not seen with other Hox genes. From this study, the authors concluded that Af9 is required for normal embryonic development, at least in part by Hox gene regulation.

Pina and colleagues sought to further elucidate the role of Af9 in hematopoietic development. Global transcriptome analysis of long term culture-initiating cells (LTC-IC) was performed, and identified AF9 as one of the most highly upregulated genes in the quiescent human LT-HSC, consistent with other published microarray data. The expression of AF9 was further analyzed in the various progenitor compartments of human cord blood. Results indicated that AF9 expression was highest in the HSC, and that expression diminished at each stage of differentiation with a 2-3 fold reduction from HSC to CMP and again to GMP, suggesting a role for AF9 in early lineage commitment.

Forced expression of AF9 by lentiviral transduction in human CD34+CD38− cord blood cells resulted in an increase in erythroid progenitors (BFU-E) at the expense of the granulocyte/macrophage progenitors (CFU-GM). The earlier multipotent CFU-GEMM population that gives rise to the CFU-GM was unaffected. Overexpression of AF9 proteins lacking the YEATS domain or AHD, respectively, produced colony numbers similar to vector control, indicating that the entire full-length protein is necessary to
mediate the expansion of the erythroid progenitor compartment. Similarly, overexpression of AF9 resulted in an increase in CD41+ colonies in a megakaryocyte-CFC assay. Conversely, knockdown of AF9 with short hairpin RNAs produced an almost complete loss of BFU-E and CFU-GEMM and a reduction in CD41+ colonies, suggesting that AF9 plays a critical role in early erythroid/megakaryocyte (E/Meg) lineage fate determination159.

To test the effect of AF9 overexpression in vivo, the authors injected NOD/SCID β2-microglobulin-/- mice with CD34+ cord blood cells transduced with lentivirus encoding GFP-tagged AF9. Two weeks after intratibial injection, engraftment of GFP-positive cells within the erythroid (CD71+), megakaryocyte (CD41+) and lymphoid/myeloid (CD45+) compartments was examined. The results correlated with the in vitro data, indicating an increase in GFP-positive cells in the erythroid and megakaryocytic populations upon forced expression of AF9 with a concurrent decrease in GFP-positive cells in the lymphoid/myeloid compartments159.

Expression of transcription factors associated with the E/Meg lineage such as GATA1 and GATA2 were increased upon overexpression of AF9 in CD34+ cord blood cells, while expression of transcription factors associated with the GM lineage (PU.1, CEBPA, GFI1) were decreased upon forced expression of AF9159. Together, the results of this study support a role for AF9 in early hematopoietic commitment to the erythroid and megakaryocytic lineage.
AF9 Homolog ENL

Interestingly, the AF9 homolog ENL (MLLT1) shows an opposing pattern of expression despite highly conserved sequence similarity throughout regions of the protein. ENL is expressed in most tissues with the exception of the liver, with the highest expression levels found in the testis and brain. In terms of hematopoietic expression, ENL is highly expressed in lymphoid cells but not in myeloid cells, the opposite pattern of AF9.\textsuperscript{150}

The ENL protein is structurally very similar to AF9, sharing 50% of amino acid sequence overall, with an increase to 80% identity within the YEATS and AHD domains (Figure 4B, 4D). However, the 5’ regulatory region of Enl shows no sequence similarity to the Af9 promoter region, suggesting that the expression of these genes is distinctly regulated. The 5’ regulatory region of Enl does not contain a consensus TATA element, but includes several Sp1 binding sites as well as potential N-myc, c-Ets, c-Rel and GATA binding sites.\textsuperscript{150}

Enl-deficient mice were engineered by insertion of a neomycin-resistance cassette with a stop codon in reverse orientation into exon six of Enl, resulting in a protein less than 40% of full length.\textsuperscript{150} Heterozygous mice did not exhibit any developmental defects; however, crosses with these mice did not yield any homozygous null offspring, indicating an essential role for Enl in embryonic development. Genotyping analysis of developing embryos using timed matings indicated the absence of Enl\textsuperscript{−/−} mouse embryos as early as E8.5, the earliest time point examined.\textsuperscript{150}
*Drosophila melanogaster* express a gene called *ear* (*Enl/Af9*-related) that is equally homologous to both human AF9 and ENL, suggesting that these genes arose from a common ancestor. Like AF9 and ENL, *ear* is highly expressed in early *Drosophila* embryos, consistent with the role of this protein and the human homologs in early transcriptional regulation\(^{163}\). However, it is possible that AF9 and ENL diverged from a single gene to fulfill separate roles in development.

**AF9 and ENL in Transcriptional Elongation**

Both AF9 and ENL are found in protein complexes associated with transcriptional elongation\(^5\)\(^-\)\(^7\). During transcription, RNA Polymerase II (Pol II) is recruited to target gene promoters and initiates transcription of approximately 12 nucleotides before it becomes paused (promoter-proximal pausing)\(^{164}\). Elongation is halted by two factors: DSIF [5,6-dichloro-1-β-d-ribofuranosylbenzimidazole (DRB) sensitivity-inducing factor] and NELF (negative elongation factor)\(^{165}\)\(^,\)\(^166\). Genome-wide ChIP-seq studies have shown that the major peak of Pol II is at about +50bp, with Pol II also found at low levels downstream of the poised promoter peak, and absent in the -2000 to -10,000 region\(^{167}\). The distribution of DSIF follows that of Pol II, and indeed DSIF has been shown to track along the gene with Pol II. NELF is primarily found only at the poised Pol II peak\(^{167}\). Association of Pol II with DSIF and NELF has been shown to significantly reduce the elongation rate of Pol II in *in vitro* transcription assays\(^{168}\).

Poised Pol II can be released to initiate productive elongation through the association with P-TEFb\(^{169}\). P-TEFb is composed of CDK9 and a cyclin T1/2 subunit.
Binding of cyclin T1/2 to CDK9 induces a conformational change in CDK9 that, along with phosphorylation of CDK9 threonine 186, is required for activation of the kinase. P-TEFb then phosphorylates the serine 2 residue of the C-terminal domain (CTD) of the largest subunit of Pol II, as well as the hSpt5 subunit of DSIF and the E subunit of NELF to release inhibition by DSIF and NELF allowing for productive elongation. At this point, additional factors associated with Pol II act to stabilize the complex and promote elongation at a rate of 3.8kb/min for up to 2 million base pairs. These factors include transcription factor IIS (TFIIS), which helps Pol II restart after backtracking on the DNA, TFIIF and ELL 1-3, which stimulate the elongation rate of Pol II, as well as the PAF complex and the Super Elongation Complex (SEC) defined below.

The members of the SEC were first identified through studies of viral gene activation upon HIV infection. P-TEFb, prior to recruitment to Pol II, is sequestered by binding to the 7SK snRNP. In HeLa cells, approximately 90% of P-TEFb is bound in an inactive form to the 7SK snRNP complex. This complex includes the RNA-binding protein HEXIM 1 or 2, the La-related protein LARP7, and the 7SK methyl phosphate capping enzyme (MePCE). HEXIM1/2 binds the 7SK snRNP as well as P-TEFb, and LARP7 and MePCE act to stabilize this complex. The HIV Tat protein releases P-TEFb from the 7SK snRNP and recruits it to the poised Pol II at the HIV LTR through a mechanism not yet well understood. Tandem affinity purification of CDK9 and Tat first identified SEC components AF9, ENL, ELL2, and AF5q31, all of which are known fusion partners in MLL leukemia.
Subsequent biochemical analyses using different methodologies have identified additional components and distinct complex formations composed of the proteins in the SEC. Biswas et al biochemically purified proteins associated with AF9, AF4 and ELL and identified the associated proteins by mass spectrometry. Reconstitution of these complexes using purified proteins lead the authors to identify distinct complexes involved in transcription regulation. First, AF4 or AF5q31 directly interact with Cyclin T1 of P-TEFb. AF4/AF5q31 then recruits AF9, an interaction that will be further discussed. P-TEFb links ELL and ELL-Associated Factors 1 and 2 (EAF1/2) to the AF4-P-TEFb complex. DOT1L was shown to interact with AF9 in a complex distinct from the AF4-P-TEFb complex, which aligns with the finding that DOT1L and AF4 bind to the same region of AF9\(^6\)(Figure 5).

A separate study by Yokoyama and colleagues purified endogenous AF4 from leukemic K562 cells and identified AF5q31, ENL, CDK9 and Cyclin-T1 in the immunoprecipitates. The authors termed this complex the AEP for AF4 family/ENL family/P-TEFb complex\(^7\). Unlike the Biswas study, DOT1L was not found to be present in this complex.

A third study by Mueller et al immunoprecipitated Flag-ENL in HEK293 cells and detected 15 associated proteins with ENL. These included CDK9, Cyclin T1, AF5q31 and AF4, as did the other studies. Additional associated proteins included BCoR and CBX8 which will be further discussed, as well as nine other proteins. This complex was termed the EAP for ENL Associated Proteins\(^5\).
While the techniques used and the final components identified between these studies have some differences, it is clear that multi-protein complexes including AF9 are involved in the intricate regulation of RNA transcription. Understanding the role of AF9, and those proteins with which AF9 interacts, is necessary for understanding the regulation and misregulation of gene expression in normal and neoplastic development.

**AF9 Interacting Proteins**

As indicated earlier, multiple proteins have been shown to interact with AF9, including AF4, DOT1L, CBX8 and BCoR. These interactions are mediated by the C-terminal AHD of AF9 which is retained in MLL-AF9 fusion proteins.

**AF4**

AF4 (AFF1) is a member of the AFF/ALF family of nuclear proteins including AF5q31, LAF4, and FMR2. These proteins contain conserved N and C-terminal homology domains, a serine/proline-rich transactivation domain and an AF4/LAF4/FMR2 (ALF) homology domain (Figure 6A). The ALF domain, specifically residues 273-289, interacts with E3 ubiquitin ligase proteins of the *Drosophila seven in absentia* (sina) homologues (Siah)-1a and Siah-2a family.

A specific mutation in the Siah-binding motif of Af4 was identified in a screen of chemically mutagenized mice. Male mice were injected with *N*-ethyl-*N*-nitrosurea (ENU) which targets spermatagonial stem cells. These mice were then bred, and offspring with dominant motor phenotypes were screened to identify the mutated gene. Mice with jerky, ataxic gaits, termed “robotic mice” were found to express Af4 protein with a V280A mutation in the Siah-binding motif of Af4. These mice also exhibited a
rapid loss of Purkinje cells in the cerebellum and developed complete, bilateral cataracts\textsuperscript{186}. The mutation expressed in the robotic mouse reduces binding between AF4 and the SIAH E3 ubiquitin ligases, significantly decreasing AF4 protein turnover and subsequently leading to an accumulation of AF4 protein. This mouse has therefore acted as a useful model of AF4 overexpression.

**Figure 6. Protein Domains of AF4**

(A) Protein domain schematic of AF4 with alternate names and chromosomal location noted. Line below the protein structure indicates the approximal fragment of the protein present in MLL fusions. (B) The AF9 interacting sequence is highly conserved among AF4 family members. A conserved lysine (indicated in bold) mediates an electrostatic interaction with AF9 residue D544.
As previously discussed, AF4 directly interacts with both the kinase and cyclin subunits of P-TEFb\(^5\). Specifically, AF4 residues 1-308 are necessary for binding to Cyclin T1\(^188\). AF4 is a direct phosphorylation target of P-TEFb, a modification which was shown to decrease the transcriptional activity of AF4 in a reporter assay\(^187\). Recent studies have shown that AF4 is also present in the 7SK snRNP complex that sequesters inactive P-TEFb, as well as a member of the P-TEF-b-AFF1(AF4)-BRD4 (PAB) complex that delivers active P-TEFb to target genes\(^188\). These authors argue that AF4 should be considered a member of all P-TEFb complexes, not just the P-TEFb-associated SEC.

AF4 residues 766-779, highly conserved in all AF4 family members, interact with AF9 and ENL (Figure 6B). Binding to either one of these proteins protects AF4 from SIAH-mediated degradation\(^187\). Inhibition of the AF4-AF9 interaction by a synthetic AF4 mimetic peptide results in necrotic cell death of MLL leukemia cells, the mechanism of which is not yet understood\(^157,189\).

Isnard and colleagues engineered a constitutive \(Af4\) knockout mouse to examine the role of Af4 in development\(^190\). Mice heterozygous for the null allele appeared normal and were fertile. Homozygous mice were born at the expected proportion and 80% of those \(Af4^{+/}\) mice appeared normal. The remaining 20% were significantly smaller, the smallest of which did not survive beyond ten days. These mice exhibited defects in lymphoid development with decreased thymic cellularity, a reduction in CD4+CD8+ T cells, and a reduction in pre-B and mature B cells. All other hematopoietic compartments
appeared normal. Interestingly, MLL-AF4 leukemias are of lymphoid phenotype in 95% of the cases\(^{191}\), further supporting a role for AF4 in lymphoid development.

**DOT1L**

The DOT1L (Dot1-like, KMT4) protein was identified through its N-terminal sequence homology to the yeast Dot1 protein, which was discovered in a genetic screen as a suppressor of telomeric silencing (Dot1, disruptor of telomere silencing)\(^{192}\). This region of shared homology encodes a conserved methyltransferase domain. DOT1L is the only member of the class I methyltransferase family, in contrast to all other known histone methyltransferases (HMTs) of the class V family. The Class V methyltransferases, including MLL, contain a conserved SET domain at the C-terminus and target residues on the protruding histone tail\(^{193,194}\). DOT1L is the only identified non-SET domain-containing HMTase and exhibits specificity for methylation of histone H3 at lysine 79 (H3K79), a residue located within the globular structure of the H3 protein\(^{195-197}\) (Figure 7). DOT1L is also the only enzyme identified thus far that methylates H3K79, and knockout of Dot1 in yeast and *Drosophila*, and Dot1l in mice all result in a global loss of H3K79 methylation\(^{197-199}\).

Studies in yeast suggest that Dot1 is a distributive enzyme—depositing both mono-, di-, and tri-methyl marks on H3K79 after one round of catalysis\(^{200}\). Most of the class V methyltransferases are processive enzymes, requiring methyl marks to be placed sequentially from mono to di to trimethyl\(^{201}\). These results also suggest that all methylation states of H3K79—mono, di and trimethyl—mark loci of active transcription\(^{200}\). Genome-wide studies support the association of H3K79 methylation with
highly transcribed genes\textsuperscript{202}, with hypomethylation of H3K79 associated with regions of repressed chromatin\textsuperscript{195, 197}.

**Figure 7. Protein Domains of DOT1L**

![Protein Domain Schematic of DOT1L](image)

Protein domain schematic of DOT1L with alternate names and chromosomal location noted.

However, there are specific examples in which H3K79 marks repressed genes, including transcription factors *Tbr1* and *CTGF*, and the epithelial sodium channel gene *ENaCα*. Overexpression of AF9, which binds to DOT1L residues 479-1222, results in hypermethylation of H3K79 in the promoter region of *ENaCα* and reduced *ENaCα* expression. The requirement of the Af9-Dot1l interaction in *ENaCα* promoter H3K79 methylation was further supported by knock-down of Af9, which resulted in a decrease in H3K79 methylation at the *ENaCα* promoter region and subsequent increase in gene transcription\textsuperscript{9}.

*Dot1l* expression is ubiquitous and essential for development, as *Dot1l<sup>−/−</sup>* mice do not survive beyond E10.5. These embryos exhibit cardiovascular defects including an
enlarged heart and decreased vasculature. Conditional cardiac-specific deletion of Dot1l caused dilated cardiomyopathy, at least in part due to the loss of Dot1l-mediated H3K79 methylation and decreased expression of dystrophin, a protein necessary for supporting cardiac smooth muscle cells from mechanical stress during contraction.

Embryos from Dot1l−/− mice also exhibit severe anemia, indicating a role for Dot1l in erythropoiesis. Loss of Dot1l expression resulted in delayed development of erythroid, but not myeloid or lymphoid, progenitor cells. Expression of Gata2, which as discussed previously is required for maturation of the HSC to the CMP, was downregulated in Dot1l-deficient mice. Pu.1 expression, associated with myeloid lineage differentiation, was significantly increased in Dot1l null mice. These results suggest that H3K79 methylation by Dot1l regulates the lineage commitment of early progenitor cells. In a separate study, conditional loss of Dot1l resulted in pancytopenia, including loss of HSCs and lineage progenitor cells.

Additionally, DOT1L has been shown to play a role in regulating heterochromatin formation at telomeres. DNA silencing at telomeres is mediated by the recruitment of Sir (silent information regulator) proteins. These proteins bind only to unmethylated H3K79, and methylation at this histone residue precludes Sir binding and affects heterochromatin formation.

DOT1L is also involved in DNA double strand break repair. The DNA damage response protein 53BP1 binds to methylated H3K79 at DNA double strand breaks, and either mutation of the H3K79 residue or knockdown of DOT1L inhibits recruitment of 53BP1 to sites of DNA damage. In regard to cell cycle progression, DOT1L H3K79
methylation levels have been shown to fluctuate during the cell cycle in both yeast and humans\cite{204, 209}. \textit{DotI1}-null embryonic stem cells show defects in proliferation and are blocked in the G2/M phase of the cell cycle\cite{210}. DOT1L is also involved in transcriptional elongation. As previously discussed, Dot1l has been found to be in complex with proteins associated with the P-TEFb elongation complex\cite{5, 187} including AF9. Studies by Biswas suggest that the DOT1L-AF9 complex is mutually exclusive of the AF9-P-TEFb complex\cite{6}.

In MLL leukemias expressing the MLL-AF9 fusion, it is hypothesized that the AF9-DOT1L interaction aberrantly recruits the histone methyltransferase to MLL target genes, promoting H3K79 methylation and productive elongation of MLL target genes associated with leukemogenesis such as \textit{Hoxa9} and \textit{Meis1}. Indeed, loss of \textit{Dot1l} expression in MLL-AF9 transformed bone marrow cells results in a loss of colony formation and \textit{Hoxa9} expression compared to wild-type cells\cite{211}.

This finding extends beyond MLL-AF9 fusions to other MLL-rearranged fusion proteins; loss of \textit{Dot1l} in MLL-GAS7 or MLL-AFX, but not E2a-PBX-transformed cells also results in a decrease in colony formation, although not as dramatically as with the MLL-AF9 expressing cells\cite{211}. Genome-wide ChIP-seq studies indicate that of 139 genes targeted by the MLL-AF9 fusion protein, 120 genes are enriched in H3K79 dimethylation\cite{212}. Loss of \textit{Dot1l} in MLL-AF9 leukemia cells resulted in a specific decrease in expression of genes targeted by MLL-AF9, and these cells were unable to induce leukemia when transplanted into mice\cite{212}.
Targeting DOT1L has thus been identified as a potential therapeutic approach in patients with MLL-rearranged leukemias. Several inhibitors have been developed, including one compound engineered by Epizyme Inc. This inhibitor, EPZ004777, competes for binding with the DOT1L cofactor S-adenosylmethionine (SAM) to a specific pocket in the DOT1L structure. SAM is the substrate for DOT1L; upon binding, DOT1L catalyzes the methylation of lysine 79 at histone H3 and releases S-adenosyl-L-homocysteine (SAH) as the product. Inhibitors that prevent SAM binding to the DOT1L SAM-binding pocket would thus prevent methylation of target residues by DOT1L. While all histone methyltransferases use SAM as the methyl group donor, EPZ004777 is highly selective for DOT1L over other histone methyltransferases\textsuperscript{213}.

Treatment of leukemic cell lines with EPZ004777 resulted in a specific decrease in H3K79 dimethylation without altering other histone modifications. Expression of MLL target genes \textit{HOXA9} and \textit{MEIS1} were reduced in treated cells compared to control, whereas genes associated with hematopoietic differentiation were increased. Mice injected with leukemic MV4-11 cells harboring the MLL-AF4 fusion gene exhibited a significant increase in median survival when treated with the DOT1L inhibitor compared to control treated mice\textsuperscript{213}. A derivative of this compound is now in phase I clinical trials.

\textbf{CBX8}

AF9 was also shown to interact with Polycomb group (PcG) protein CBX8 (HPC3) in a yeast two-hybrid screen and subsequently validated by other techniques\textsuperscript{10}. Further characterization of this interaction mapped amino acids 202-333 of Cbx8 binding to the C-terminal 94 amino acids of AF9 (Figure 8). Members of the CBX family encode
a conserved N-terminal chromodomain and a C-terminal C-box involved in interaction with other proteins such as RING1A/B \(^{214-216}\). The AF9 interacting region is located between these two domains, in a region of the Cbx8 protein that shows no homology to other Cbx family members\(^{10}\).

**Figure 8. Protein Domains of CBX8**

Protein domain schematic of DOT1L with alternate names and chromosomal location noted.

The Polycomb group (PcG) of proteins act to silence gene targets in opposition to the activating Trithorax group of proteins, which includes MLL. Polycomb proteins exert their function through two complexes: the Polycomb Repressive Complex (PRC) 1 and PRC2. In the canonical model of PRC-mediated repression, PRC2 containing EED, SUZ12 and EZH2 is recruited to target genes. EZH2, a SET-domain containing histone methyltransferase, then catalyzes methylation of histone H3 lysine 27. This mark, associated with gene repression\(^{217-219}\), then provides a docking site for PRC1 which is composed of PHC, RING1A/B, BMI1 or MEL18, and a CBX protein (CBX 2, 4, 7 or
8)\textsuperscript{220, 221}. The exact composition of this complex is tissue-and differentiation-stage specific\textsuperscript{222-224}. RING1 is an E3 ubiquitin ligase that catalyzes mono-ubiquitination of H2A at lysine 119, which is a mark associated with transcriptional repression. Additional proteins such as DNA methyltransferase (DNMTs) or histone deacetylases (HDACs) can also be recruited to mediate additional silencing.

Maintaining the proper balance between the activity of the Trithorax group of proteins and the Polycomb group of proteins is necessary for appropriate self-renewal and differentiation of tissues\textsuperscript{219}. The HOX genes, as previously described, are very well studied targets of TRX and PcG. The promoters of these and other genes involved in development are often marked by both the TRX-associated H3K4 methylation and PRC2-associated H3K27 methylation marks. These bivalent genes are poised for either rapid activation or repression following removal of one of the two marks. Polycomb group proteins also play a role in cell cycle checkpoints. For example, members of both PRC1 and PRC2 bind to and repress the $p^{16ink4a}/p^{19Arf}$ locus at the G1/S checkpoint\textsuperscript{220}. Studies have also identified roles for PcG proteins in the DNA damage response pathway, as H2AK119Ub accumulates at DNA damage foci\textsuperscript{225, 226}. As both PcG and TRX proteins play critical roles in development and homeostasis, it is not surprising that PRC and TRX proteins are often mutated in cancer.

The role of CBX proteins in hematopoiesis and leukemogenesis has been a rapidly emerging area of study. Klauke and colleagues showed the overexpression of Cbx7 enhances the self-renewal of hematopoietic stem cells and leads to leukemia, whereas overexpression of Cbx 2, 4, or 8 leads to stem cell exhaustion. PRC complexes
containing difference Cbx proteins are differentially recruited to a subset of target genes\textsuperscript{227}.

In regard to MLL leukemias, CBX8 has been shown to be required for MLL-AF9 mediated transformation and leukemogenesis. Tan and colleagues retrovirally transduced bone marrow cells with MLL-AF9 fusion proteins encoding mutations that prevented binding to Cbx8. Cells expressing these mutants were unable to serially replate in a colony forming assay compared to the wild-type MLL-AF9 fusion\textsuperscript{228}. These results were further confirmed by expressing the MLL-AF9 fusion in bone marrow isolated from wild-type or conditional Cbx8 knockout mice. Loss of Cbx8 resulted in a loss of colony formation of MLL-AF9 transformed cells compared to control cells still expressing Cbx8. Loss of transformation ability was associated with a decrease in activation of a luciferase reporter gene under the control of the Hoxa9 promoter. While CBX8 is generally thought to be a repressive protein, these results support a role for CBX8 in transcriptional activation in MLL leukemia, possibly through association with the histone acetyltransferase (HAT) TIP60\textsuperscript{228}.

AF9 has also been shown to be present in a complex with additional members of the PRC1 family. Immunoprecipitation of FLAG-tagged AF9 in HEK293 cells showed association with endogenous PRC1 proteins CBX8, RING1A, RING1B and BMI1. Knockdown of CBX8 with shRNA resulted in decreased association of AF9 with RING1B and BMI1, indicating that interactions are mediated by the AF9-CBX8 complex\textsuperscript{229}.
Gel filtration experiments indicated that other AF9 binding partners, such as DOT1LL, AF4 and BCoR, are not present in the AF9-PRC1 complex. Overexpression of CBX8 results in decreased recruitment of DOT1LL to the ENaCα locus, which as discussed, is mediated by AF9. As H3K79 methylation is associated with repression of this gene, overexpression of CBX8 resulted in increased expression of ENaCα. Co-overexpression of CBX8 with AF9 restored levels to control. Conversely, shRNA-mediated knockdown of CBX8 resulted in a decrease in ENaCα expression\textsuperscript{229}. These results highlight the importance of the distinct functions of AF9 that depend on the proteins with which it is associated, the last known of which is BCoR.

**BCoR**

The BCL-6 Corepressor (BCoR) protein was identified as a transcriptional repressor through its interaction with the oncoprotein BCL-6\textsuperscript{230}(Figure 9). In a reporter gene assay, coexpression of BCoR and BCL-6 resulted in potentiation of the transcriptional repression mediated by BCL-6\textsuperscript{230}. These proteins are found to localize at the same BCL-6 target genes in germinal center B cells\textsuperscript{231}.

BCoR itself is widely expressed throughout mouse and human cells, in both embryonic development and in adult tissues\textsuperscript{230, 232}. High levels of expression are present in the developing tooth primordium, eye, limb buds, branchial arches and nervous system tissues\textsuperscript{232}, all of which are abnormal in patients with X-linked Oculofaciocardiodental Syndrome (OFCD) characterized by mutations in BCoR\textsuperscript{233, 234}.

Whole exome sequencing of a patient with acute myeloid leukemia with a normal karyotype (CN-AML) identified a clonal somatic mutation in BCoR\textsuperscript{235}. Further analysis
of 553 CN-AML patients found that BCoR mutations were present in 3.8% of the cases. These mutations often resulted in reduction or loss of BCoR expression. AML patients with BCoR mutations had 7.5% Event Free Survival (EFS) compared to 41.4% of patients with wild-type BCoR, with 25.6% and 56.7% Overall Survival, respectively\textsuperscript{235}.

**Figure 9. Protein Domains of BCoR**

A yeast two-hybrid screen seeking to identify proteins that interact with the C-terminal 94 amino acids of AF9 identified BCoR as a potential interacting protein\textsuperscript{10}. Further analysis identified the minimal AF9 binding domain to be located at BCoR amino acids 1127-1251(Figure 9). Within this region is a stretch of 34 amino acids that is present in two, but not all four mouse BCoR splice variants. Only isoforms containing this region were able to bind to AF9, whereas all isoforms can bind to BCL-6 as the regions affected by alternative splicing are located downstream of the BCL-6 binding domain. These results suggest that the interaction of specific BCoR isoforms with AF9
can modulate the transcriptional activation activity of AF9, both in the context of the full length protein and perhaps in the context of MLL-AF9 fusions.

Collectively these studies all indicate a critical role for AF9 and AF9 interacting partners AF4, DOT1LL, CBX8 and BCOR in gene regulation and development (Figure 5). The function of interactions mediated by AF9 is less well understood. Here, we seek to elucidate the role of AF9 and AF9-mediated interactions in normal hematopoiesis and leukemogenesis.
CHAPTER THREE

METHODS

Cloning of Full-length AF9

To clone full-length AF9 into a FLAG vector, the 3XFLAG tag was PCR amplified from the p3xFLAG vector (courtesy of Dr. C. Hemenway) with the high fidelity Phusion polymerase (Finnzymes/Thermo Scientific) using a forward primer including a XhoI site and a reverse primer with XbaI and BamHI sites, as well as a stop codon sequence (1U/50μl reaction Phusion Polymerase, 10mM dNTPs, 10uM each forward and reverse primer, 5x HF Buffer)(Table 1). The amplicon was digested with XhoI overnight in NEB Buffer 4, followed by digestion with BamHI the next day for an additional two hours. The digested amplicon was then purified by gel extraction with the Qiaex II Gel Extraction kit (Qiagen). The pMSCV.YFP vector (courtesy of Dr. J. Zhang) was digested in parallel with the insert with XhoI overnight in NEB Buffer 3, followed by a two hour digestion the next day with BglII. The digested vector was treated with 2μl calf intestine alkaline phosphatase (CIP) enzyme (NEB), incubated at 37°C for 30 minutes and gel purified.

The 3xFLAG fragment was ligated into the digested MSCV.YFP vector overnight at 14°C at a 3:1 insert-to-vector molar ratio. The following day, 1μl of the ligation reaction was transformed into DH5α chemically competent cells. The ligation mix was added to 50μl of cells on ice, gently flicked 5-6 times, and placed on ice for 30 minutes.
Bacteria were heat shocked at 42°C for 30 seconds and returned to the ice for 5 minutes. 950μl of room temperature SOC was added to the cells, and they were incubated at 37°C, rotating at 225 rpm for 1 hour. The transformed bacteria were plated on LB plates containing 100μg/mL ampicillin and placed in a 37°C incubator overnight.

The following day, isolated transformants were inoculated into 5mL of LB broth containing ampicillin (100μg/ml) and were grown overnight shaking at 225 rpm, 37°C. The plasmid DNA was isolated from the bacteria using the GeneJet Thermo Scientific Miniprep Kit and was screened by PCR using the MSCV vector forward sequencing primer and the reverse primer used for cloning. Positive clones verified by sequencing.

Full length AF9 was PCR amplified from the p3xFLAG.AF9 vector (courtesy of Dr. C. Hemenway) with a forward primer containing an EcoRI restriction site and a reverse primer including a XhoI restriction site (Table 1). The PCR product was digested sequentially with XhoI overnight followed by the addition of EcoRI enzyme the next day for an additional two hours. The enzymes were heat inactivated for 10 minutes at 65°C and the PCR product was purified by gel extraction (Qiagen Gel Extraction Kit).

Table 1. Primer Sequences for Cloning of MSCV.YFP.AF9.3xFLAG

<table>
<thead>
<tr>
<th>Amplicon</th>
<th>Primer Sequence (5’→3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3xFLAG</td>
<td>Forward: CCGCTCGAGGACTACAAAGACCATG</td>
</tr>
<tr>
<td></td>
<td>Reverse: CGGGATCCTCTCTAGACTAGTCATCGTCATC</td>
</tr>
<tr>
<td>AF9</td>
<td>Forward: GGAATTCATGGCTAGCCTG</td>
</tr>
<tr>
<td></td>
<td>Reverse: CCGCTCGAGGATGTCCAGATG</td>
</tr>
</tbody>
</table>

To clone full length AF9 into the newly constructed MSCV.YFP.3xFLAG vector, the above mentioned digested AF9 was ligated into MSCV.YFP.3xFLAG vector digested
with EcoRI and XhoI as previously described, yielding the final MSCV.YFP.AF9.3xFLAG vector. The final construct was screened by digestion with EcoRI, and positive clones were verified by sequencing.

**Site-Directed Mutagenesis to Introduce Mutations into Full-length AF9**

Primers were designed according to the Agilent Quik-Change Site Direct Mutagenesis kit recommended protocol to introduce one of eight single amino acid changes in full length AF9. The MSCV.YFP.AF9.3xFLAG plasmid was used as a template in a PCR reaction containing 1μl of each site directed mutagenesis primer (Table 2), 5μl 10x reaction buffer, 1μl dNTP mix, 1μl PfuTurbo and water to 50μl total volume. A control reaction using pWhiteScript template and provided primers was assembled in parallel. The reactions were incubated in a ThermoHybaid PCR Express Thermocycler for one cycle at 95°C for 30 seconds, followed by 12-18 cycles of 95°C for 30 seconds, 55°C for 1 minute, and 68°C for 8 minutes or 5 minutes for the mutagenized construct and the pWhiteScript control, respectively. After the PCR was completed, the samples were cooled to 37°C followed by digestion with DpnI for 1 hour at 37°C.

DpnI-treated mutagenized DNA (1μl) was transformed into 50μl XL1-Blue cells (Agilent). The mixture was swirled and incubated on ice for 30 minutes. The samples were then heat shocked at 42°C for 45 seconds and returned to the ice for 2 minutes. 500μl pre-warmed SOC was added to the reaction and incubated at 37°C, shaking at 225 rpm. After 1 hour, the transformed bacteria were plated on LB plates containing
100μg/mL ampicillin supplemented with 2% X-Gal and IPTG for blue-white selection. Plates were placed in the 37°C incubator overnight.

The next day, 5ml of LB broth containing ampicillin was inoculated with a single colony and grown overnight at 37°C shaking at 225rpm. After 12-16 hours of growth, DNA was extracted from the transformants and screened by double digestion with XhoI and EcoR1 to digest AF9 from the vector. Clones were sent for sequence verification.

After confirming the presence of the desired mutation, AF9 was cut out of the vector and ligated into the MSCV.YFP.3xFLAG backbone. The plasmid DNA was incubated with XhoI enzyme, BSA and NEBuffer 4 in a 50μl reaction overnight at 37°C. The following day, EcoRI enzyme was added to the reaction mixture and the digestion proceeded for 2 hours at 37°C. The digested product was run on a 2% gel and extracted.

The digested mutant AF9 was ligated into MSCV.YFP.3xFLAG vector digested with XhoI and EcoRI, CIP-treated and gel extracted as described previously. The mutant full-length AF9 was ligated into MSCV.YFP.3xFLAG in a 2μl insert:1μl vector 10μl total reaction. Ligation reactions were placed in a 16°C water bath overnight.

The following day, 1μl of the ligation reaction was transformed into DH5α chemically competent cells as previously described. Transformants were selected and inoculated into 5ml LB broth containing ampicillin. Cultures were grown at 37°C shaking at 225rpm overnight. After 12-16 hours of growth, DNA was extracted from the bacteria. Clones were screened by EcoRI restriction enzyme digestion and sent for sequencing using the MSCV Reverse primer.
### Table 2: Primer Sequences for Site Directed Mutagenesis

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Nucleotide Change</th>
<th>Primer Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>E506R</td>
<td>GAA → AGA</td>
<td>Sense: tgaatgacgacaglctctagatagactgtgtagctcagc&lt;br&gt;Antisense: ggaagctctccctctcttaggtatgcctgtcactaaatca</td>
</tr>
</tbody>
</table>
| I538A    | ATC → GCC        | Sense: tgaactctataagaaacgagcttcatgccccaaacaaacatatttgttttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtg
reactions was transformed into DH5α electrocompetent cells by electroporation. Transformants were screened by PCR with primers amplifying the MLL-AF9 fusion junction, and positive clones were verified by sequencing.

Table 3. Primer Sequences for Cloning of MSCVneoMLLAF9(D546R)3xFLAG

<table>
<thead>
<tr>
<th>Primer</th>
<th>Primer Sequence (5’→3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF9 Forward</td>
<td>AAAGCGGTCGACTTTAAGATTCTTGAGTTGAAAGTCCAATTAAAG</td>
</tr>
<tr>
<td>1x FLAG Reverse</td>
<td>AAAGGATCCCTCATTGTGTCATCGTCTTTTGTAGTCGGAGATGTCCAG</td>
</tr>
</tbody>
</table>

**Cloning of AF9(D546R) 376-568 and 470-568 into pCMV3xFLAG**

The region containing amino acids 376-568 and 470-568 of AF9 were PCR amplified from MSCV.YFP.AF9(D546R) with an AF9 forward primer containing an EcoRI restriction site and a reverse AF9 primer containing a KpnI restriction site (Table 4). Digested amplicons were isolated by gel purification and ligated into the pCMV3xFLAG vector (courtesy of Dr. C. Hemenway) digested with EcoRI and KpnI for two hours followed by CIP treatment, heat inactivation and gel purification. Ligation reactions and transformations were performed as described above. Transformants were screened using EcoRI digest, and positive clones were sent for sequencing for verification.

Table 4. Primer Sequences for Cloning AF9(D546R) 376-568 and 470-568

<table>
<thead>
<tr>
<th>Primer</th>
<th>Primer Sequence (5’→3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF9(376-568) Forward</td>
<td>GCGAATTCATCTTGAAACAACCAGTCCCT</td>
</tr>
<tr>
<td>AF9(470-568) Forward</td>
<td>GCGAATTCATTACTAAAACCCAAACAAC</td>
</tr>
<tr>
<td>AF9 Reverse</td>
<td>ATGGGTACCTCAGGATGTCCAGATGT</td>
</tr>
</tbody>
</table>

**Retrovirus Production, Concentration, and Titering**

Five 10cm tissue culture dishes of Phoenix Ecotropic cells were seeded at 2 x 10^6 cells in DMEM supplemented with 10% fetal bovine serum and 1%
penicillin/streptomycin for each retrovirus construct. The following day, the media was changed 2-3 hours before calcium phosphate transfection with 22.5μg retroviral construct and 2.5μg pCL-Eco packaging plasmid following the manufacturers’ protocol (ClonTech). Sixteen hours post-transfection, the transfection media was removed and 7ml of DMEM + 10% FBS was added. Cells were placed at 32°C, 5% CO₂ overnight. The following day, the supernatant was collected, 7ml of fresh media was added and the cells were returned to the 32°C incubator. The viral supernatant was again harvested the following day and pooled with the previously collected supernatant. The supernatant was spun to pellet any cell debris and concentrated 6x using Centricon Plus 70 filters (Millipore) by centrifugation at 1640rpm, 33°C in a swing bucket JS 75 rotor for approximately two hours. Concentrated virus was aliquoted, snap frozen, and stored at -80°C.

To determine titer of the prepared retroviruses, 2.5x10⁴ Rat1a cells were seeded per well of a 6-well tissue culture plate. The following day, the virus preparation was serially diluted 1:2 up to 1:1x10⁶ in DMEM + 10% FBS supplemented with 16μg/mL polybrenene and added to the cells where they were incubated at 37°C, 5% CO₂. After two hours, an additional milliliter of DMEM + 10% FBS was added to the cells to bring the final volume to 2ml and cells were returned to the 37°C incubator. The following day, the media was removed and replaced with 2ml DMEM 10% FBS.

For cells transduced with plasmid encoding a neomycin resistance gene, media was supplemented with 0.5mg/ml G418. Selection was maintained for ten days at which time media was removed, the cells were washed with PBS and stained with methylene
blue (0.1g of methylene blue in 60ml methanol) for 5 minutes. Plates were rinsed in diH2O and air dried, after which colonies were enumerated.

Cells transduced with plasmid encoding yellow fluorescence protein (YFP) and an untransduced control were removed from the plate by trypsinization 48 hours after retroviral transduction and resuspended in PBS + 2% FBS. The cell suspension was passed through a mesh filter to achieve a single-cell suspension, and the percentage of YFP+ cells was assessed using flow cytometry.

**Immunoprecipitation and Western Blot Analysis**

Phoenix Ecotropic cells transfected with full-length AF9 point mutants to prepare virus as described above were harvested four days after transfection. Cell pellets were lysed in 1x SDS Sample Buffer and boiled at 95°C for 5 minutes. Protein lysates were resolved on a 10% separating, 4% stacking SDS-PAGE gel at 150 Volts. Proteins were transferred onto a nitrocellulose membrane using the Bio-Rad Trans-Blot Turbo Transfer System, 1.5mm gel transfer program. Membranes were blocked in 5% milk in Tris Buffered Saline + 0.05% Tween (TBS-t) for a minimum of one hour. AF9 proteins were detected using a mouse anti-FLAG M2 antibody (Sigma) at a 1:3000 dilution in 5% milk for one hour to overnight at 4°C. Membranes were washed four times for five minutes each in TBS-t, and then incubated with secondary anti-mouse or anti-rabbit antibody (GE Healthcare Life Sciences) at a 1:8000 dilution in milk for two hours. Membranes were washed as described and developed using ECL Substrate (Thermo).

MLL-AF9-transformed c-kit+ bone marrow cells frozen from colony assay experiments (as subsequently described) at the end of week one were thawed and plated
in RPMI + 10% fetal bovine serum, 10ng/ml IL-3, 10ng/ml IL-6, 100ng/ml SCF and 10ng/ml GM-CSF overnight. The following day, cells were lysed in 1x SDS sample buffer followed by boiling at 95°C for 5 minutes. Lysates were resolved on an 8% SDS-PAGE gel at 150 volts and transferred on the BioRad Trans-Blot Turbo Transfer System using the program for high molecular weight proteins. Membranes were blocked for two hours and incubated with primary mouse anti-FLAG M2 antibody overnight, and secondary anti-mouse antibody for two hours the following day as described above. Membranes were developed using Super-Signal West Dura substrate (Thermo).

For coimmunoprecipitation experiments, HEK293 cells were cotransfected with plasmids encoding either FLAG-tagged full length AF9 or C-terminal fragments of AF9 (wild type or D546R mutant, amino acid residues 376-568 or 470-568) and one of five GFP-tagged proteins: full length murine Dot1L, murine Dot1L amino acids 479-659, human DOT1L amino acids 828-1095, human AF4 RI3A- fragment, corresponding to amino acids 647-871 of AF4 or human AF4 amino acids 755-777 using calcium phosphate transfection according to the manufacturer’s protocol (ClonTech). Cells were harvested 72 hours after transfection and lysed in 1ml IP lysis buffer (500mM Tris-HCl pH 7.4, 150mM NaCl, 0.5% Triton-X 100 (v/v) and freshly added Protease Inhibitor Cocktail and 1μM DTT). Lysates were sonicated with five pulses at output four, twice, before centrifugation at 12,000rpm for 20 minutes to remove the insoluble fraction. 950μl of prepared lysate was incubated with 20μl slurry FLAG M2 beads (Sigma) blocked in Z-Buffer (25mM HEPES pH 7.5, 12.5mM MgCl2, 150mM KCl, 0.1% NP-40 20% Glycerol, BSA [1mg/mL], and freshly added Protease Inhibitor Cocktail 1:1000)
overnight, rotating at 4°C. The remaining 50µl were saved for Western Blot to determine input amounts. The following day, the bead-bound proteins were pelleted by centrifugation at 1500 rpm, 4°C for one minute, followed by four-30 second washes with 1mL NETN buffer (100mM NaCl, 1mM EDTA, 0.5% NP-40, 20mM Tris-HCl pH 8.0, and freshly added Protease Inhibitor Cocktail 1:1000). Proteins were eluted by boiling the beads in SDS sample buffer at 95°C for 5 minutes. After a brief centrifugation to pellet the beads, the eluate was resolved on a 10% gel as described above.

RNA Isolation and cDNA Synthesis

RNA was isolated from cells using TriReagent following the manufacturers’ protocol (Sigma). RNA was treated with DNaseI to remove any genomic DNA contamination according to the manufacturers’ protocol (Fermentas). Complimentary DNA was synthesized using the ABI High Capacity cDNA Synthesis kit (Applied Biosystems/Life Technologies). Up to 2µg of RNA was reverse transcribed in a reaction following the manufacturers’ protocol. cDNA was stored at -20°C.

Reverse Transcription-PCR

RNA was isolated from cells expressing combinations of MLL-AF5(WT or K717D) and AF9(WT or D544R) in the genetic complementation experiments (subsequently described) at the end of week four, and cDNA was prepared as described. RT-PCR followed by sequencing was performed to verify the presence or absence of the mutation in these cells. Primers used to amplify these regions are in Table 5. PCR reactions were prepared with 1x Buffer, 1x Q Solution, 200µM dNTPs, 0.2µM forward and reverse primers, 0.25U/50µl Taq Polymerase and 100ng of cDNA template or 10ng
plasmid control. PCR cycling conditions were as follows: 95°C 2 minutes, followed by 30 cycles of 95°C 30sec/58.3°C 1 min/72°C 1 min, and final extension at 72°C for 10 min. PCR products were resolved on a 0.8% agarose gel at 90V. Products were sent for sequencing with their respective forward primers after clean-up using the Qiagen PCR Clean-up Spin Column kit.

Table 5. Primer Sequences for Amplification of AF5 and AF9 Mutant Sequence

<table>
<thead>
<tr>
<th>Amplicon</th>
<th>Primer Sequence (5’→3’)</th>
</tr>
</thead>
</table>
| AF9 (WT or D544R) | F: AAAGCGGTGAAGTTTGGAAAGTGAAAAGTCCAATAAAG  
|                   | R: AAAGGATCCACTTACCCATGTCATGTCATTTGTCAGTC                  |
| AF5 (WT or K717D) | F: CTTCCTTCTTCTACAAACTC  
|                   | R: CGTCTCCTTGGCTGTTACT                                      |

Quantitative Reverse Transcription-PCR (qRT-PCR)

TaqMan probe technology was used to assess expression of murine *Hoxa9, Meis1* and *Hprt* control gene expression (*Hoxa9*: Mm00439364_m1, *Meis1*: Mm00487664_m1, *Hprt*: Mm01545399_m1; Applied Biosystems/Life Technologies). Reaction mixtures containing 1μl 20x Gene Expression TaqMan probe, 10μl of 2x Gene Expression Assays master mix (Applied Biosystems/Life Technologies), 2μl of undiluted prepared cDNA and water to 20μl volume were prepared for each sample in triplicate. qRT-PCR was performed on an Applied Biosystems 7300 Real Time PCR System using the standard cycling program of one cycle at 50°C for 2 minutes, one cycle at 95°C for 10 minutes, and 40 cycles of 95°C for 15 seconds/60°C for 1 minute. Data were analyzed using the ABI Prism 7300 software. Relative target gene expression was normalized to *Hprt* expression using the 2^ΔΔCt method\(^{236}\). Experiments were repeated a minimum of three times, and data are shown as an average of all experiments with the standard error of the
mean. Student’s t-test was performed to determine statistical significance using GraphPad Prism software.

Levels of mRNA encoding transcription factors associated with erythroid/megakaryocyte (Gata1, Gata2) and granulocyte/macrophage (Pu.1, Cebpa) lineages were assessed using SYBR Green technology (Table 6). Reaction mixtures containing 10μl SYBR Green Mix (BioRad), 1μl 10uM forward primer, 1μl 10uM reverse primer, 1μl cDNA diluted 1:3 in H2O to 20μl total volume were prepared in triplicate for each sample. The standard cycling protocol was used with the addition of the dissociation step. Target gene expression was normalized to Gapdh expression and analyzed as described above.

Table 6. Primer Sequences for Erythroid and Myeloid Lineage Transcription Factors

<table>
<thead>
<tr>
<th>Amplicon</th>
<th>Primer (5’→3’)</th>
</tr>
</thead>
</table>
| Gapdh    | Forward: GTGAGGCCGGTGCTGAGTAT  
           | Reverse: TCATGAGCCCTTCCACAATG |
| Gata1    | Forward: TCACAAGATGAATGGTCAGA  
           | Reverse: TGGTCGTGGACAGTGTAGT |
| Gata2    | Forward: CCCCAGCGAGGCTGTCTT  
           | Reverse: TCAGACGACAACCACCACTT |
| Pu.1     | Forward: CCCCCACCCCGCTCTACACCAG  
           | Reverse: GAGGGTGTGGCGGTGGGGG |
| Cebpa    | Forward: GTAACCTTTGTGCCCTTGGGACT  
           | Reverse: GGAAGCGGAATCCTCCCAAATA |

*Gata1* primer sequence23, Gata2 and *Pu.1* primer sequences52

Expression of *Meis1* was analyzed using primers specifically designed to amplify the 5’ exon 1, exon 8/9, and 3’ exon 12 (Table 7). The protocol for qRT-PCR with SYBR Green as described above was followed. Additionally, the *Meis1* TaqMan probes described were also used in this assay to amplify exons 6/7.
Table 7. Primer Sequences for Exon 1, Exon 8/9 and Exon 12 of Meis1

<table>
<thead>
<tr>
<th>Amplicon</th>
<th>Primer (5’→3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’ Meis1 (Exon 1)</td>
<td>Forward: CCTGTGCAACACACACTTTAC</td>
</tr>
<tr>
<td></td>
<td>Reverse: GGGAGATTAGAGAGGGGTTA</td>
</tr>
<tr>
<td>Meis1 Exons 8/9</td>
<td>Forward: GTGACGATGATGACCTTGATAA</td>
</tr>
<tr>
<td></td>
<td>Reverse: CAGAAGGGAAGGGGTGTGTTAG</td>
</tr>
<tr>
<td>3’ Meis1 (Exon 12)</td>
<td>Forward: CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC</td>
</tr>
<tr>
<td></td>
<td>Reverse: TTCCCATCCTTTCCCCGATTAC</td>
</tr>
</tbody>
</table>

**Genomic DNA isolation**

Genomic DNA was isolated from cell pellets using the Puregene Genomic DNA Purification Kit (Gentra) according to the manufacturers protocol. DNA pellets were resuspended in 50-250μl DNA Hydration Solution and incubated at 65°C for 1 hour, followed by continued hydration overnight at room temperature. Genomic DNA was stored at -20°C.

**Af9 Genotyping PCR**

To assess *in vitro* 4-OHT-mediated deletion of floxed *Af9* in the hematopoietic progenitor colony forming cell (CFC) assays, polymerase chain reaction was performed using primers that anneal to genomic DNA sequences located up and downstream of the flanked loxP sites in the *Af9*fl/fl mouse (Table 8). Approximately 100ng of genomic DNA was amplified in a 25μl reaction containing 1X Buffer, 200μM dNTPs, 0.25μM each primer, and 1U/50ul Taq Polymerase. The PCR reaction was cycled at 94°C for 5 min, followed by 40 cycles of 94°C for 30 seconds, 57°C for 30 seconds, and 72°C for 90 seconds. A final 10 minutes extension at 72°C completed the reaction. PCR products were resolved on a 0.8% agarose gel at 90V.
Table 8. Primer Sequences for Genotyping of Af9 Mice

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’→3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Af9 Forward</td>
<td>AAACAATAGCTTTGGCAGCG</td>
</tr>
<tr>
<td>Af9 Reverse</td>
<td>CCTAGGCAGCGAGGAAGTTT</td>
</tr>
</tbody>
</table>

Primary Mouse Bone Marrow Isolation

C57Bl/6, Af9(Mllt3)\textsuperscript{fl/fl}, or Af9(Mllt3)\textsuperscript{fl/fl} B6.129-Gt(Rosa)Sor\textsuperscript{tm1(cre/ERT2)Tyj}/J mice were sacrificed by CO\textsubscript{2} inhalation. Femurs and tibias were collected in RPMI + 10% FBS. Bone marrow was flushed from the bones using a 25g 1.5” needle into a 6-well plate containing additional RPMI + 10% FBS. Bones were flushed until all bone marrow had been removed and bones appeared white. Bone marrow was collected in a conical tube and centrifuged at 1500 rpm for 5 minutes.

c-kit Positive Selection of Mouse Bone Marrow Progenitor Cells

Isolated bone marrow cells were resuspended in 2ml PBS + 2% FBS at 100x10\textsuperscript{6} cells/ml in 5ml polypropylene tubes for c-kit+ selection using the Stem Cell EasySep Mouse CD117 Positive Selection kit. All antibody incubations were conducted in the tissue culture hood at room temperature for the indicated amounts of time.

CD117 antibodies were added to the cells at 70μl/ml and incubated for 15 minutes. EasySep PE Selection cocktail was then added at 70μl/ml for an additional 15 minutes. Magnetic nanoparticles were mixed by pipeting five times, and were then added at 50μl/ml to the cell suspension for 10 minutes. After the final incubation, the cells were brought to a total 2.5ml volume with PBS + 2% FBS. Cells were gently mixed by pipeting 2-3 times and then placed in a chilled EasySep magnet without the cap for 5 minutes. While holding the tube in the magnet to retain nanoparticle-bound cells, the tube
was inverted and the supernatant was poured off. The tube was removed from the magnet and 2.5ml PBS + 2% FBS was added. Cells were resuspended by gently pipeting 2-3 times and replaced back in the magnet for 5 minutes. The separation protocol was repeated for a total of four-5 minute separations. After the final separation, cells were counted and plated overnight in a round-bottom 96 well plate in RPMI + 10% FBS, 1% penicillin/streptomycin supplemented with 0.05mM β-mercaptoethanol, 100ng/ml SCF, 10ng/ml IL-3 and 10ng/ml IL-6.

**Retroviral Transduction**

c-kit+ cells were transduced with retroviruses in RPMI 10% FBS supplemented with 100mM HEPES and 4μg/mL polybrene by centrifugation at 3000 rpm, 33°C for four hours on one or two days. The specific protocol was adjusted based on the experimental design as described below.

**Colony Replating Assay**

For MLL-AF9 point mutants, 30,000 c-kit+ cells were transduced with 750μl high titer MSCVneoMLL-AF9(wild type or mutant) virus in a 1mL total volume (750μl viral supernatant, 7.5mM HEPES, 4μg/ml Polybrene and RPMI 10% FBS to volume) for two consecutive days. After the first spinoculation, cells were plated overnight in RPMI + 10% FBS, 1% penicillin/streptomycin supplemented with 0.05mM β-mercaptoethanol, 100ng/ml SCF, 10ng/ml IL-3 and 10ng/ml IL-6 in 96 well plates. The spinoculation procedure was repeated on the second day, after which cells were plated at 10,000 cells/1.1mL in methylcellulose supplemented with glutamine, 100ng/ml SCF, 10ng/ml
IL-3, 10ng/ml IL-6, 10ng/ml GM-CSF, 1.25mg/ml G418 and RPMI in 35mm dishes. The dishes were placed in a larger 15cm dish humidified with an uncovered 35mm dish containing H2O and placed in a 37°C, 5% CO2 incubator for seven days. After one week, colonies were counted using a Leica DMIL model microscope (Wetzlar, Germany) and enumerated as colonies/10,000 cells. Photographs of the colonies were taken using a Nikon Diaphot inverted fluorescence microscope with an Olympus DP21 digital camera.

For replating, the cells were collected in tubes containing 20mL of 1x PBS, washed one time, resuspended at an appropriate volume and counted. One-to-ten thousand cells were replated as described above for a total of four times. Experiments were repeated three to six times and statistical significance was assessed using a Student’s t-test with p<0.05.

**Genetic Complementation Assay**

c-kit+ bone marrow progenitors were isolated from C57Bl/6 mice as described above. 50,000 cells were retrovirally transduced with a combination of AF9(WT or D544R) and MLL-AF5(WT or K717D) virus in a 2ml total transduction volume, three tubes per construct. Each spinoculation tube contained 50,000 cells, 750μl of MSCVYFPFAF9 virus, 750μl MSCVneoMLL-AF5 virus, polybrene and HEPES as described. The virus combinations were as follows: MLL-AF5(WT) + AF9(WT), MLL-AF5(K717D) + AF9(WT), MLL-AF5(WT) + AF9(D544R), or MLL-AF5(K717D) + AF9(D544R).
Cells were centrifuged at 3000rpm, 33°C for four hours on two consecutive days. At the end of the last spin, cells were plated in methylcellulose with G418 for one week to select for cells expressing the MSCVneoMLL-AF5(WT or K717D) construct.

After one week, cells were harvested from methylcellulose and sorted for YFP+ using the FACsAria III instrument (BD Biosciences). The YFP+G418R cells were divided evenly and plated in methylcellulose for a colony replating assay as described above.

**Murine Colony-Forming Cell (CFC) Assays**

**CFC Assays with Af9fl/fl c-kit+ Bone Marrow Progenitor Cells**

For experiments using untransduced c-kit+ bone marrow cells, cells were harvested and immediately pre-treated with 200nM 4-hydroxytamoxifen (Sigma) or ethanol control in a 96-well plate for four hours at 37°C. After the four hour pre-treatment, cells were plated for myeloid, erythroid or megakaryocytic lineage differentiation.

For erythroid lineage differentiation, 10,000 cells in RPMI + 2% FBS supplemented with 10ng/ml IL-3, 50ng/ml SCF, 3U/ml rhEPO, 1% P/S, 100nM 4-OHT or ethanol control, were plated in 2.4mL aliquots methylcellulose (M3234, Stem Cell) to a total volume of 3.3mL. 1.5mL, or approximately 4,500 cells, were plated in two-35mm dishes and placed in a covered 15cm plated with a 35mm uncovered dish of water for humidity. Erythroid colonies were enumerated per 10,000 cells plated and photographed at day four.
For myeloid lineage differentiation, cells were plated in 900μl RPMI + 10% FBS, 10ng/ml IL-3, 10ng/ml IL-6, 50ng/ml SCF, 10ng/ml GM-CSF, and 100nM 4-OHT or ethanol control. The cell suspension was added to 2.4ml methylcellulose aliquots supplemented to yield a total volume of 3.3ml. 1.1ml of the methylcellulose suspension containing approximately 6,800 cells was plated in 35mm dishes. Cells were placed along with an uncovered 35mm dish of water in a 15cm plate in a 37°C incubator for 12 days. Colonies were enumerated per 10,000 cells plated.

For megakaryocytic lineage differentiation, 11,000 cells in 300μl IMDM containing 100nM 4-OHT or ethanol control were added to a 1.7mL Megacult-C medium containing 50 ng/ml rhTPO, 10ng/ml rmIL-3, and 20ng/ml rhIL-6 (Stem Cell). 1.2ml cold collagen was added to the Megacult-C cell suspension and quickly mixed. 750μl of the suspension, or approximately 2500 cells, was added per chamber of a double chamber slide (Stem Cell Double Chamber Slide Kit #04963). The slides were placed in a 15cm plate along with an uncovered 35mm dish of water for humidity in a 37°C, 5% CO₂ incubator for 7-10 days.

At the end of the incubation period, the slides were removed from the incubator and dehydrated using filter paper layered on top of mesh spacers (Stem Cell Double Chamber Slide Kit #04963). After the filter paper was fully saturated, it was removed and the slides with mesh spacer were placed in a container of ice cold acetone for five minutes, gently releasing the mesh from the slide with slight agitation. Slides were air dried for 15 minutes and either stained immediately or stored in the dark at 4°C for up to a month.
If refrigerated, slides were brought to room temperature prior to staining. Acetylcholinesterase stain was prepared by dissolving 10mg of acetyhiocholiniodide (Sigma #A5751) in 0.1M sodium phosphate buffer, pH 6.0. Under constant stirring, the following were added in order: 1ml 0.1M sodium citrate, pH 6.0; 2ml 30mM copper sulphate, pH 6.0; 2ml 5 mM potassium ferricyanide solution (ACS Reagent #244023). Slides were completely covered with the staining solution and incubated for 5 hours at room temperature in a covered chamber. The staining solution was removed and slides were fixed in 95% ethanol for 10 minutes, followed by rinsing with lukewarm water. The slides were then air dried before counterstaining with Harris’ hematoxylin solution (Sigma 107K4356) for 30 seconds. Slides were rinsed with lukewarm water, air dried, and stored in covered containers at room temperature. Achetylcholinesterase-positive colonies were counted, and slides were photographed using an AMG Evos XL Core microscope (Fischer Scientific).

CFC Assays with YFP-AF9 Transduced \textit{Af9}^{\text{fl/fl}} Bone Marrow Progenitor Cells

For CFC assays requiring expression of YFP-tagged AF9 constructs, 300,000 c-kit+ cells were transduced with 750μl MSCV.YFP, MSCV.YFP.AF9(WT) or MSCV.YFP.AF9(mutant) virus in a 1mL total volume as described above. Cells were centrifuged at 3000 rpm, 33°C for 4 hours, and plated at $10^6$ cells/ml in RPMI + 10% FBS, 1% penicillin/streptomycin supplemented with 0.05mM β-mercaptoethanol, 100ng/ml SCF, 10ng/ml IL-3 and 10ng/ml IL-6 in 96-well plates overnight.
Approximately 30 hours later, cells were collected in PBS + 2% FBS and sorted for YFP+ cells using the FACS Aria Instrument (Becton Dickinson).

Immediately after sorting, YFP+ cells were plated in cytokine-supplemented media with 200nM 4-hydroxytamoxifen (Sigma) or ethanol control in a 96-well plate for four hours at 37°C. After the four hour pre-treatment, cells were plated for myeloid, erythroid or megakaryocytic lineage differentiation as described above.

**Cytospins and Staining**

Cells were harvested from methylcellulose, washed and resuspended at 20,000 cells/200ul in 1x PBS. Cells were spun at 1000rpm for 4 minutes using the CytoSpin4 ThermoSci cytospin onto charged microscope slides (Premiere, CA Scientific Co). The cells were air dried and stained with Hema 3 (Fisher Scientific).

**Surface Marker Staining by Flow Cytometry**

Cells collected from CFC assays for myeloid or erythroid lineage were resuspended in 500μl PBS + 2% FBS, and divided amongst two polystyrene FACs tubes. A separate tube of approximately 50,000 cells was resuspended in 200μl PBS + 2% FBS for an unlabeled control. Compensation controls were prepared by adding one drop of either α-Rat α-Hamster Igκ positive or negative compensation beads (BD Bioscience) to 100μl 1x PBS. Specific antibody dilutions were prepared, and master mixes were prepared to assess surface marker expression of erythroid and myeloid lineage markers (Table 9).

1μl of each antibody (eBioscience) was added to the appropriate compensation tubes, and the antibody cocktails were added to the samples. Tubes were kept on ice in
the dark for 20 minutes. 500μl of PBS + 2% FBS was added to each tube, gently mixed, and centrifuged at 1500 rpm for 7 minutes. The supernatant was carefully decanted and the cells were resuspended in the residual volume. Secondary conjugate antibodies were added to the appropriate tubes and incubated on ice in the dark for 20 minutes. An additional 500μl PBS + 2% FBS was added to wash the cells, gently mixed, and cells were spun down at 1500 rpm for 7 minutes. Compensation beads and antibody-bound cells were resuspended in 250μl PBS + 2% FBS and kept on ice in the dark until analyzed using the FACs CantoII or LSRFortessa instruments (BD Biosciences). Cell surface marker expression was analyzed using FlowJo (TreeStar Inc., Oregon).

Table 9. FACs Antibodies for Erythroid and Myeloid Surface Markers

<table>
<thead>
<tr>
<th>Erythroid Lineage Cocktail</th>
<th>Myeloid Lineage Cocktail</th>
</tr>
</thead>
<tbody>
<tr>
<td>APC-CD117</td>
<td>APC-CD117</td>
</tr>
<tr>
<td>PE-CD71</td>
<td>PE-CD41</td>
</tr>
<tr>
<td>Biotin-Ter119</td>
<td>Biotin-CD11b</td>
</tr>
<tr>
<td>SA-PECy7</td>
<td>SA-APCCy7</td>
</tr>
<tr>
<td></td>
<td>PE-Cy7-Gr1</td>
</tr>
</tbody>
</table>
AIM 1: To determine which AF9 complex interactions are necessary for normal hematopoietic development

Cloning and Stable Expression of full length AF9 point mutants

The C-terminus of AF9 has been shown to interact with four proteins, each of which is associated with gene activation or repression. We sought to elucidate the function of these protein interactions in normal hematopoiesis and leukemogenesis by engineering point mutations in AF9 predicted to either destabilize the protein, block its interaction with one or more known AF9 binding partners, or modify a potential phosphorylation site.

Initial structural studies of AF9 from the Bushweller lab at the University of Virginia identified four potential residues as mutation targets. The C-terminal region of AF9 that has been shown to interact with multiple proteins is composed of three alpha helices and one beta strand (Figure 10A). Alpha helices generally follow the $i + 4$ rule, in which the amine group of each amino acid donates a hydrogen bond to the carboxyl group of the amino acid four residues upstream$^{238}$. There are approximately 3.6 residues per turn, and most commonly three helix turns correspond to ten amino acids total$^{239}$.

Three amino acid residues, E506, K557, and L562, were selected as they are located within α-helices in the AF9 C-terminus. Glutamate 506 is present in helix one,
and lysine 557 and leucine 562 are located in helix three (Figure 10A). Helix one and three are in close proximity to each other, and can form a potential protein-protein interaction interface.

Polar amino acids often participate in salt bridges with amino acids of complementary charge, creating stable electrostatic bonds that are important for protein stability. Both E506 and K557 are charged residues that were mutated to an amino acid of the opposite charge. Negatively charged glutamate 506 was mutated to positively charged arginine, and positively charged lysine 557 was mutated to negatively charged glutamate. We hypothesized that reversing the charge would disrupt the structure of the helix and potential charge-charge interactions with other proteins at the helix interface.

Leucine 562 is a hydrophobic amino acid buried within helix three (Figure 10A). Hydrophobic amino acids are often found in the core of protein structures, isolated from the aqueous environment. Mutation to small, slightly polar alanine is expected to disrupt the proper folding of the helix, altering protein interactions mediated by the AF9 structure in this region.

A fourth residue chosen for mutation was isoleucine at amino acid 538 (Figure 8A). Isoleucine is a hydrophobic amino acid that prefers to be buried within hydrophobic cores. Isoleucine has two branches at the β-carbon position, making it a bulkier structure than amino acids with a single branch. Because of this, it is not often incorporated into α-helices, but is commonly found in β-sheets. Indeed, I538 is located at the C-terminal end of β strand one in the AF9 structure. β strand one and β strand two form an antiparallel β-sheet connected by a loop in the C-terminus of AF9. Mutation of I538 to
alanine was predicted to disrupt the folding of the β sheet, and therefore potentially altering any protein-protein interactions mediated by this secondary structure.

Serine 565 was identified as a potential phosphorylation site. It is located at the C-terminus of the AF9 protein, three amino acids from the end of the protein (Figure 8A). Two mutations were introduced to study the potential functional relevance of this residue: an alanine substitution mimicking an unphosphorylable residue, or an aspartate substitution mimicking a constitutively phosphorylated serine.

After the structure of AF9 was further resolved in complex with an AF4 peptide in 2013, additional structurally-informed amino acid mutational targets in AF9 were identified. The D544R mutation was first identified through previous work in Dr. Hemenway’s laboratory. They mapped the interacting region of AF9 to AF4 amino acids 760-770. This region is highly conserved in all AFF family members, including a conserved lysine residue (Figure 6B). AF4(K764) resulted in a loss of binding to AF9 in a yeast two-hybrid assay. Screening a mutagenized AF9 cDNA library for clones that could restore binding to the AF4(K764) mutant identified an aspartic acid at residue 544 of AF9. Supporting structural studies in collaboration with the Bushweller laboratory characterized the AF9(D544)-AF4(K764) interaction as forming a stable salt bridge. Binding studies determined a dissociation constant 0.17mM for the interaction of AF9 with AF4. Mutation of the aspartic acid at position 544 in AF9 to aspartic acid reduced the binding to wild type AF4 to Kd=7.39nM, indicating that the AF9-AF4 interaction mediated by residues D544-K764 is necessary for high affinity binding of the two proteins.
Figure 10. Engineering and Stable Expression of Full Length AF9 Point Mutants

A

Structure of AF9 solved in complex with AF4\textsuperscript{156} (PDB ID 2LM0) with point mutations examined in these studies indicated. N- and C-terminal ends as well as the three alpha helices are noted. (B) Detection of FLAG-tagged full length AF9 mutants in Phoenix Ecotropic cells by western blot with Actin loading control.
AF9 aspartate 546, two residues C-terminal to the aspartate critical for interaction with AF4, was identified as mediating interaction with DOT1L (Figure 10A). Similar to the AF9-AF4 salt bridge interaction, AF9 is thought to interact with DOT1L via a charge-charge interaction. There are three binding regions for AF9 on DOT1L: amino acids 628-653, 863-878 and 878-900. These interaction sites have a conserved sequence of LxL/I/VxIxI/LxxL/V, clustering together five hydrophobic residues within a 10 amino acid region (unpublished data from the Bushweller lab). Structural studies identified AF9 residue D546 as a potential candidate to mutate to inhibit binding of AF9 at all three DOT1L interfaces.

To express the full length AF9 point mutants in bone marrow cells, the plasmids were packaged for retroviral transduction. For virus production, ecotropic Phoenix cells, engineered from the human embryonic kidney (HEK 293) cell line to stably express the retroviral packaging genes, were transfected with the AF9 plasmids. The supernatant containing retroviral particles was collected and concentrated. Virus titers were determined by serially diluting the virus and transducing Rat1a cells. Acceptable titers were in the 1x10^6 to 1x10^7 virus particle/mL range.

To determine whether the mutant AF9 proteins were stably expressed, the transfected ecotropic Phoenix cells were harvested and protein lysates were prepared to assess expression of the AF9 point mutants. As some of the mutants were predicted to disrupt local protein folding, it was important to show that introduction of amino acid substitutions in AF9 did not affect the stability of the protein. An anti-FLAG antibody was used to detect the AF9 point mutants, and Actin was used as a control for the amount
of protein loaded in the western blot. Mutants D544R, I538A, S565A and S565D all expressed comparably to the wild type AF9 protein (Figure 10B). A slight reduction in expression of D546R, E506R, K557E and L562A mutants was evident, however proteins are readily detected. All AF9 mutants are stably expressed and are of the expected size of 65 kilodaltons, indicating that introducing point mutants into AF9 did not result in instability and degradation, and that these mutants can be used to assess the function of interactions mediated at these residues in the following experiments.

**Interaction of AF9(D546R) with DOT1L and AF4**

The binding between AF9 mutants D544R, E506R, I538A, K557E, L562A and S565D in the context of the C-terminal 376-568 amino acids of AF9 to DOT1L and AF4 was assessed by coimmunoprecipitation studies by a former graduate student in the Hemenway lab, Dr. Bhavna Malik. The interaction between AF9 expressing the D546R mutation and AF4 or DOT1L had not yet been studied, and so I sought to examine the functional importance of this residue.

The D546R point mutation was predicted to specifically disrupt binding to DOT1L while maintaining interaction with AF4; however, it was important to examine functionally how this mutation would affect binding between AF9 and these interacting partners. Coimmunoprecipitation studies were conducted to examine the binding of C-terminal AF9 residues 470-568 and 376-568 to fragments of DOT1L and AF4. Approximately 75% of the MLL-AF9 fusions expressed in patients with MLL leukemia include AF9 residues 376-568. However, 12% of patients express a MLL-AF9 fusion including only amino acids 475-568 of AF9. Introduction of this MLL-AF9(475-
fusion protein by homologous recombination into the endogenous *Mll* locus resulted in AML development in mice, indicating that this smaller 94 amino acid AF9 fragment is sufficient for transformation. AF9 fragments 376-568 and 470-568 were tested for binding to AF4 and DOT1L in order to elucidate any potential differences in binding mediated by the 192 amino acid versus the 98 amino acid C-terminal tail of AF9.

FLAG-tagged AF9(470-568) or AF9(376-568), either wild type or containing the D546R mutation, were coexpressed in HEK-293 cells with various fragments of GFP-tagged AF4 or DOT1L to examine the interaction between these proteins. The AF9 protein was immunoprecipitated using anti-FLAG antibody-conjugated agarose beads, and then FLAG-purified proteins were probed by western blots with an anti-GFP antibody.

As mentioned above, unpublished binding studies from the Bushweller lab have identified three sites in DOT1L that interact with AF9: amino acids 628-653, 863-878, and 878-900. The D546R mutation in AF9 is predicted to disrupt binding to all three DOT1L sites, but coimmunoprecipitation studies were necessary to examine the interaction of this AF9 mutant with DOT1L. Constructs encompassing either the most N-terminal interacting site [mDot1(479-659)] or the two downstream sites [hDOT1L(828-1095)], in addition to a construct expressing the full length Dot1l protein, were used to assess binding to both C-terminal fragments of AF9, wild type or the D546R mutant. These sites are present and highly homologous between both human and mouse DOT1L proteins.
The AF9(470-568) D546R mutant exhibits more than a 70% reduction in binding to full length Dot1l or the DOT1L (828-1095) construct encompassing the two most C-terminal interaction sites compared to wild type AF9 (Figure 11A). Coexpression with the Dot1l fragment containing the upstream AF9 interacting sequence (479-659) increased binding 5-fold with the D546R mutant as compared to wild type. Of note, coexpression of any DOT1L construct with AF9(D546R) mutant resulted in decreased FLAG-AF9 detected in the input. All three DOT1L constructs had similar input amounts regardless of coexpression with wild type or D546R mutant AF9 (Figure 11A).

Similar results were obtained with the larger AF9(376-568) D546R mutant fragment (Figure 11B). Binding to full length Dot1l and DOT1L(828-1095) was decreased with the D546R mutant 30% and 80%, respectively, as compared to wild type. Interaction with the more N-terminal mDot1l(479-659) fragment was not significantly different between AF9(376-568) wild type or D546R mutant. Coexpression of AF9(D546R) with full length Dot1l or DOT1L(828-1095) resulted in decreased AF9 levels in the input, whereas coexpression with Dot1l(479-659) did not alter AF9 levels. The input levels for all three GFP-tagged Dot1l constructs remained consistent (Figure 9B).

From these results, we conclude that AF9 residue D546 is involved in mediating interaction with DOT1L, specifically interaction with the full length protein or the DOT1L fragment encompassing two of the three AF9 binding sites (863-878 and 878-900), but not the first site (628-653).
Figure 11. Interaction of AF9(D546R) fragments 470-568 and 376-568 with DOT1L

A
Coimmunoprecipitation studies of AF9 fragments 470-568 (A) or 376-568 (B) and DOT1L constructs full-length mDotl, mDot1l (479-659) and hDOT1L(828-1095). Figures are representative of two independent experiments. Pixel analysis was performed using ImageJ software. Immunoprecipitated GFP-tagged DOT1L proteins were normalized to FLAG-AF9 input. * = p<0.05, ** = p<0.02.
Because AF9 is known to interact with AF4 amino acids 768-777\(^{157}\), a construct containing this minimal interaction domain (AF4 755-777) was used to assess the effect of the D546R mutation on the AF9-AF4 interaction. A second AF4 fragment containing a slightly larger region of AF4, including the AF9 interaction domain but lacking the bipartite nuclear localization signal (amino acids 647-871), was also used to examine AF9(D546R) binding.

No statistically significant difference in binding of AF9(470-568) WT or D546R to the larger AF4 fragment (647-871) was evident, but the D546R mutant shows an almost 8-fold increase in binding to the AF4 755-777 fragment encompassing the minimal AF9 interaction domain, despite a loss of detectable AF9 in the input (Figure 12A).

No statistically significant differences in binding of AF9(376-568) WT or D546R were present with either AF4 construct, despite trends towards a slight reduction in binding with the D546R mutant compared to wild type (Figure 12B). Again, there is a reduction in detectable AF9(D546R) in the input samples of HEK293 cells cotransfected with AF4(755-777). In contrast to the AF9(470-568) fragment, there is also a reduction in AF9(376-568) D546R mutant when coexpressed with the larger 647-871 AF4 fragment (Figure 12B).

From these results, we conclude that mutation of AF9 D546 to arginine does not significantly abrogate binding of AF9 to AF4. An exception to this is the interaction of the smaller AF9 fragment (470-568) with the minimal AF4 fragment (755-777), in which binding of AF9(D546R) is statistically increased compared to wild type.
Figure 12. Interaction of AF9(D546R) fragments 470-568 and 376-568 with AF4

A

Coimmunoprecipitation studies of AF9 fragments 470-568 (A) or 376-568 (B) and AF4(647-871) or AF4(755-777). Figures are representative of two independent experiments. Pixel analysis was performed using ImageJ software. Immunoprecipitated GFP-tagged AF4 proteins were normalized to FLAG-AF9 input., ** = p<0.02.
In vivo Deletion of Af9 Does Not Alter Levels of Hematopoietic Lineage-Specific Transcription Factors in Af9-Inducible Knockout Mouse Bone Marrow Cells

The role of AF9 has not been well studied beyond its function in MLL fusion proteins. As reviewed in the previous chapter, the Af9 knockout mouse dies perinatally and exhibits skeletal defects as well as altered expression of developmental genes such as Hoxd4. The cause of death was not further examined in these mice. An independent group overexpressed or knocked down AF9 using lentiviral expression vectors in CD34+ human cord blood cells. They showed that overexpression of AF9 resulted in an expansion of erythroid and megakaryocytic colonies in a colony forming cell (CFC) assay. Conversely, knockdown of AF9 resulted in a complete loss of this compartment. These two studies together indicate a critical role for Af9 in embryonic development, as well as in differentiation of hematopoietic cells.

Our laboratory, in conjunction with Dr. Hemenway’s group, sought to further assess the role of Af9 in an inducible knockout mouse system. Crosses were designed by Dr. Zeleznik-Le and Dr. Hemenway. The initial rounds of breeding produced mice that were heterozygous or homozygous for the Af9 allele flanked by two loxP sites. Additional crosses with B6.129-Gt(ROSA)Sor\textsuperscript{tm1(cre/ERT2)/Tyl/J} mice, which express a tamoxifen-inducible Cre recombinase, yielded Af9\textsuperscript{fl/fl Cre/+} mice. When tamoxifen is administered to these mice by intraperitoneal injection, the tamoxifen is processed to the bioactive 4-OHT molecule which then binds to the estrogen response (ER) element in the ubiquitously expressed Rosa26 promoter. Cre recombinase expression is then turned on in all tissues, and the enzyme binds to and catalyzes DNA recombination between the
paired loxP sites, excising the *Af9* gene. This system allows for conditional deletion of floxed region of *Af9* at specific time points in all tissues.

Homozygous floxed-*Af9* and homozygous floxed-*Af9* Cre mice were administered tamoxifen by I.P. injection at specific time points. Only mice with the Cre allele will excise the floxed *Af9* gene upon tamoxifen treatment. Mice were sacrificed at set times after injection, and peripheral blood and bone marrow were assessed for hematopoietic development. Other members of the laboratory examined surface marker expression in the peripheral blood and bone marrow of these mice, while I assessed transcript levels of lineage-specific transcription factor genes in bone marrow cells. Based on previous studies, we hypothesized that *Af9* is a critical regulator of erythroid/megakaryocyte (E/Meg) lineage commitment, and that loss of *Af9* would result in a decrease or loss of differentiation of bone marrow cells to this lineage.

As discussed in the previous chapter, expression of specific transcription factors is necessary for hematopoietic stem and progenitor cells differentiation to mature blood cells. Expression of *Gata1* and *Gata2* is required for progenitor cells to differentiate into the E/Meg lineage, whereas expression of *Pu.1* and *Cebpa* is associated with granulocyte/macrophage (GM) lineage commitment (Figure 1). If *Af9* is required for E/Meg lineage commitment, we expect to see a decrease in levels of mRNA encoding transcription factors associated with differentiation to this lineage (*Gata1* and *Gata2*), with a possible compensatory increase in transcription factors associated with GM lineage commitment.
We do not see any significant changes in the transcript levels of erythroid (Gata1, Gata2) or myeloid (Pu.1, Cebpa) lineage transcription factor genes in mice expressing Af9 compared to those null for Af9 expression (Figure 13).

Figure 13. Expression of Lineage-Specific Transcription Factors in Af9 Inducible Knockout Bone Marrow Cells

Transcript levels of erythroid lineage-associated transcription factor genes Gata1 and Gata2 and myeloid lineage-associated transcription factor genes Pu.1 and Cebpa in the bone marrow of Af9 fl/fl or Af9 fl/fl cre/+ mice after intraperitoneal injection with tamoxifen. RNA was isolated from nine mice of each genotype, and cDNA was analyzed in triplicate. Expression was normalized to Gapdh, and error bars indicate standard deviation.

Loss of Af9 Does Not Alter Hematopoietic Lineage Differentiation in vitro

In addition to the ongoing in vivo studies in our laboratory, I assessed the differentiation ability of bone marrow cells lacking Af9 in vitro. Rather than administering tamoxifen to mice in vivo to induce deletion of Af9, c-kit+ bone marrow progenitor cells were isolated from Af9 fl/fl cre/+ mice that had not been exposed to
tamoxifen. Cells were then plated in semi-solid methylcellulose media with cytokines which potentiate differentiation of the erythroid, myeloid or megakaryocytic lineages with or without 4-OHT, the bioactive derivative of tamoxifen, to induce deletion of Af9.

To examine the ability of cells to differentiate along the erythroid lineage following loss of Af9, bone marrow progenitor cells from Af9<sup>fl/fl Cre/+</sup> mice were plated in methylcellulose with IL-3, IL-6 and EPO in the presence or absence of 4-OHT to examine the role of Af9 in erythroid cell development. Loss of Af9 did not alter the ability of bone marrow progenitor cells to differentiate to the erythroid lineage in this system. There was no difference in the number of colonies formed in cells expressing or lacking Af9 (Figure 14).

**Figure 14. Erythroid Colony Formation by Af9 Inducible Knockout Bone Marrow Progenitor Cells**

Erythroid colony formation in c-kit+ bone marrow progenitor cells isolated from Af9<sup>fl/fl cre/+</sup> mice. Cells were treated with 4-OHT or ethanol control in vitro and plated in methylcellulose in the presence of cytokines to promote differentiation to the erythroid lineage. Colonies were enumerated per 10,000 cells plated. Experiments were repeated in duplicate three times.
FACs analysis was used to examine changes in expression of cell surface markers associated with erythroid development, including CD71 and TER119, upon Af9 deletion. CD71 is a transmembrane glycoprotein present on the surface of proliferating cells, reticulocytes and erythroid precursor cells. TER119 is a cell surface antigen specific to erythroid cells, expressed on proerythroblasts through mature erythrocytes, but not expressed on earlier CFU-E and BFU-E cells. There was a significant reduction in expression of CD71 in cells lacking expression of Af9 compared to wild type, although there were no significant changes in expression of TER119. Levels of CD117, or the c-Kit receptor present on all hematopoietic progenitor cells, also remained unchanged in wild-type cells or those lacking Af9 (Figure 15A).

In addition, I assessed the surface marker expression of myeloid lineage-associated cell surface markers, hypothesizing that any alterations in erythroid lineage differentiation and surface marker expression may result in a compensatory response in the myeloid lineage differentiation and surface marker expression. CD41 is the alpha subunit of the gpIIb/gpIIIa complex expressed on the surface of platelets and megakaryocytes. Gr-1 is a myeloid differentiation antigen expressed on granulocytes and monocytes. CD11b, also known as Mac-1, is a transmembrane protein expressed on most myeloid cells including granulocytes, monocytes, and dendritic cells. There were no significant changes in expression of the myeloid surface markers CD41, Gr-1 or CD11b in cells with or without Af9 (Figure 15B). We conclude based on these results that Af9 is not required for the formation of erythroid colonies in vitro despite a reduction in CD71 surface marker staining in cells lacking Af9 expression.
Figure 15. Surface Marker Staining of Af9 Inducible Knockout Bone Marrow Progenitor Cells Plated for Erythroid Lineage Differentiation

A

Unstained  +EtOH  +4-OHT

CD117

0.219  44.1  41.2

CD71

0.351  40.7  66.3

TER119

0.302  41  42.8

MFI

+EtOH  +4OHT

CD117  CD71  Ter119
Surface marker staining of \( A^f9^{fl/fl \ cre/+} \) mouse bone marrow progenitor cells plated for erythroid lineage differentiation. Unstained cells were used as a control, and the average Mean Fluorescence Intensity (MFI) was calculated using the geometric mean of fluorescence intensity of three experiments, \( * = p<0.05 \). (A) Expression of erythroid lineage specific markers Ter119 and CD71, along with the hematopoietic progenitor marker CD117 (B) Expression of CD117 and myeloid lineage-specific markers CD41, Gr-1 and CD11b.

Next we examined the differentiation of inducible \( A^f9 \) knockout cells to the myeloid lineage by plating c-kit+ bone marrow cells from \( A^f9^{fl/fl \ cre/+} \) mice in cytokines to promote myeloid lineage differentiation: IL-3, IL-6, SCF and GM-CSF. Cells were plated in the presence of either 4-OHT or ethanol, and colony forming unit-granulocyte macrophage (CFU-GM) colony formation was assessed after 12 days.

There was no significant change in colony formation of bone marrow cells with or without Af9 expression (Figure 16). Similarly, there was no significant change in myeloid-associated surface marker expression in these cells (Figure 17A). We also examined the expression of erythroid lineage surface markers in these cells,
hypothesizing that any alterations in the myeloid lineage could be compensated by increases or decreases in differentiation to the erythroid lineage and vice versa. We did not see any changes in expression of erythroid-associated surface markers (Figure 17B). From these results, we conclude that Af9 does not play a role in myeloid lineage differentiation in vitro.

Figure 16. Myeloid Colony Formation by Af9 Inducible Knockout Bone Marrow Progenitor Cells

Myeloid colony formation in c-kit+ bone marrow progenitor cells isolated from Af9^{fl/fl}cre/+ mice. Cells were treated with 4-OHT or ethanol control in vitro and plated in methylcellulose in the presence of cytokines to promote differentiation to the myeloid lineage. Colonies were enumerated per 10,000 cells plated. Experiments were repeated in duplicate three times.
Figure 17. Surface Marker Staining of Af9 Inducible Knockout Bone Marrow Progenitor Cells plated for Myeloid Lineage Differentiation

A

Unstained  +EtOH  +4-OHT

CD117

0.137  3.8  5.23

CD41

0.166  0.201  0.104

Gr-1

0.27  40.8  30.5

CD11b

0.44  72.9  69.1
Surface marker staining of Af9\textsuperscript{fl/fl} cre/+ mouse bone marrow progenitor cells plated for myeloid lineage differentiation. Unstained cells were used as a control, and the average Mean Fluorescence Intensity (MFI) was calculated using the geometric mean of fluorescence intensity of three experiments. Expression of hematopoietic progenitor marker CD117 and (A) myeloid lineage-specific markers CD41, Gr-1 and CD11b or (B) erythroid lineage specific markers CD71 and TER119.

We assessed the ability of bone marrow progenitor cells with or without Af9 to differentiate to the megakaryocytic lineage \textit{in vitro}. We hypothesized that loss of Af9 would result in a loss of differentiation to both erythroid and megakaryocytic lineages. To test this, cells were plated in chamber slides in a semi-solid media supplemented with TPO, IL-6 and IL-3, and either 4-OHT to induce recombination and deletion of Af9 or ethanol control. After seven days, the slides were fixed and stained for acetylcholinesterase activity (Figure 18A). Mouse megakaryocyte and early megakaryocyte progenitors express the acetycholinesterase enzyme. When incubated with acetylcholine substrate, the enzyme releases Cu\textsuperscript{2+} ions which precipitate as colorless copper thiocoline. Subsequent incubation with ammonium sulfide converts the colorless copper thiocoline molecules to brown-colored copper sulfate\textsuperscript{242}. The slides also contain
myeloid cells which do not stain brown. Acetylcholinesterase-positive megakaryocytes were counted, and no significant difference was seen between cells expressing or lacking Af9 (Figure 18B). As these cells are fixed on the slides, surface marker staining was not assessed. We conclude based on these assays that Af9 is not required for megakaryocyte lineage commitment in *vitro*.

**Figure 18. Acetylcholinesterase-Positive Megakaryocyte Staining of Af9 Inducible Knockout Bone Marrow Progenitor Cells**
(A) Representative image of acetylcholinesterase-positive megakaryocytes among acetylcholinesterase-negative myeloid cells (left, 4X magnification). Higher magnification (10X) of positively-staining megakaryocytes (right). (B) Quantification of Acetylcholinesterase-positive \textit{in vitro} differentiated megakaryocytes from Af9\textsuperscript{fl/fl} Cte/+ mouse bone marrow progenitor cells treated with 4-OHT or ethanol control. After plating in semi-solid medium in the presence of cytokines to promote megakaryocytic differentiation, chamber slides were dehydrated and stained for acetylcholinesterase activity. Acetylcholinesterase-positive colonies were counted and enumerated per 10,000 cells plated. Experiments were conducted in duplicate three times.

Together these results show that loss of Af9 \textit{in vitro} does not alter lineage commitment of c-kit+ bone marrow cells to the myeloid or megakaryocytic lineages under these conditions. Erythroid colony formation was not affected by the loss of Af9, although there was a reduction in the early erythroid marker CD71.

**Re-expression of Wild-Type AF9 or AF9 Mutants S565A and S565D Does Not Alter \textit{in vitro} Lineage Differentiation of Af9-Inducible Knockout Bone Marrow Progenitor Cells**

The first \textit{in vitro} differentiation studies I conducted were based on the early hypothesis that loss of Af9 would result in a loss of cells differentiating to the erythroid and megakaryocyte compartment. Prior to obtaining the above described results, we tested whether re-expression of wild-type AF9 could rescue any defects in lineage commitment. Further, we wanted to see if expression of two AF9 mutants, S565A and S565D, would alter the ability of cells to differentiate to the erythroid lineage. These mutants were selected based on preliminary results from an earlier experiment that indicated the serine residue at 565 may be critical in mediating progenitor cell maturation (data not shown).
Bone marrow progenitor cells were isolated from \( \textit{Af9}^{fl/\text{Cre/+}} \) mice, and were transduced with retrovirus encoding AF9(WT), AF9(S565A), AF9(S565D) or vector alone. As described previously, serine 565 near the C-terminus of AF9 was identified as a possible phosphorylation site. A mutation to alanine would prevent this residue from being phosphorylated, whereas mutation to an aspartic acid mimics constitutive phosphorylation. All constructs were YFP tagged, allowing for sorting of YFP+ cells 36 hours after transduction. YFP+ cells were plated in methylcellulose in the presence of cytokines to promote erythroid lineage differentiation with or without 4-OHT, as described previously.

Cells transduced with empty vector showed no difference in erythroid colony formation when treated with or without 4-OHT (Figure 19). This is consistent with the results shown in figure 12. Re-expression of AF9(WT) or either mutant (S565A or S565D) also did not significantly alter colony formation (Figure 19). The cells were also subjected to surface marker analysis, examining both erythroid and myeloid-associated surface markers as described previously. Consistent with colony formation data, there were no significant differences in expression of surface markers associated with erythroid (Figure 20A) or myeloid (Figure 20B) lineage upon loss of Af9, or re-expression of AF9 WT or mutants compared to the ethanol-treated vector control.
Erythroid colony formation by c-kit+ progenitor cells isolated from the bone marrow of Af9$^{fl/fl}$ Cre/+ mice. Cells were transduced with retroviruses encoding YFP-tagged empty MSCV vector, full length wild-type AF9, or full length AF9 mutants S565A or S565D. YFP+ cells were plated in methylcellulose in the presence of cytokines to promote erythroid lineage differentiation, and 4-OHT to induce deletion of Af9 unless otherwise noted. Colonies were counted after four days and enumerated per 10,000 cells. The experiment was conducted in duplicate.

From these results, we confirm that loss of Af9 does not alter the ability of c-kit+ murine bone marrow cells to form erythroid colonies \textit{in vitro} under the conditions tested. Additionally, re-expression of wild-type AF9, or AF9 mutants S565A or S565D does not alter erythroid lineage commitment compared to the control cells.
Figure 20. Surface Marker Staining Profiles of Af9 Inducible Knockout Bone Marrow Progenitor Cells Re-Expressing Wild-type or Mutant AF9 plated for Erythroid Lineage Differentiation

A

Surface marker staining of retrovirally transduced Af9\textsuperscript{fl/fl} cre/+ mouse bone marrow progenitor cells plated for erythroid lineage differentiation. Unstained cells were used as a control, and the average Mean Fluorescence Intensity (MFI) was calculated using the geometric mean of fluorescence intensity. Expression of hematopoietic progenitor marker CD117 and erythroid lineage specific markers CD71 and TER119 (A) or myeloid lineage-specific markers CD41, Gr-1 and CD11b (B) was calculated.
AIM 2: To determine whether AF9 interactions with Af4 and Dot1l are necessary for leukemogenesis

AF9 Interaction with AF4 Is Required for in vitro Leukemic Transformation

As discussed previously, AF9 has been shown to interact with AF4 in part through AF9 residue D544 and AF4 residue K764. These charged amino acids form a salt bridge that promotes stable bonding between the two proteins with a $K_d$ of 0.17nM$^{156}$. Mutation of AF9 residue D544 to the positively charged arginine reduces this binding 43-fold to 7.39nM$^{240}$. Additionally, coimmunoprecipitation studies conducted in Dr. Hemenway’s laboratory indicated a loss of binding to AF4 when AF9 residue D544 is mutated to arginine$^{240}$. Here, we sought to assess the functional relevance of the AF9-AF4 interaction mediated by this salt bridge using an in vitro leukemic transformation assay.

When normal bone marrow is plated in semi-solid methylcellulose media, the bone marrow progenitor cells will form colonies within the first week, but then will continue to differentiate and undergo cell cycle arrest. Hence, these untransformed cells will not form colonies after replating. When oncogenic fusion proteins such as MLL-AF9 are introduced into bone marrow progenitor cells, the cells will be blocked from differentiating and will continue to form colonies upon subsequent replating. We hypothesized that the interaction of AF4 and AF9 in the context of the leukemia-causing MLL-AF9 fusion is critical for the ability of the oncoprotein to transform bone marrow cells. Loss of the AF9-AF4 interaction through mutation of the aspartic acid at residue
544 in AF9 to an arginine would result in a reduction in colony forming ability in the in vitro colony formation assay.

As previously discussed, AF9 interacts with AF4 which binds to and stabilizes the pTEFb elongation complex. It is possible that AF9 in the context of the MLL-AF9 fusion recruits the pTEFb elongation complex to MLL target genes through interaction with AF4, resulting in aberrant transcription of genes that promote leukemogenesis. Loss of this AF9-AF4 interaction would no longer bring the pTEFb elongation complex to target genes, resulting in a loss of expression of genes required for the leukemic transformation program, and loss of serial replating ability.

MLL-AF9 wild type, MLL-AF9(D544R) mutant or empty MSCVneo vector were retrovirally expressed in c-kit+ bone marrow cells isolated from wild type mice. Transduced cells were plated in methylcellulose media in the presence of cytokines to promote growth. Colonies were counted after one week prior to replating up to four times. Results indicate that while MLL-AF9(WT) transformed cells continue to form colonies throughout the four weeks of the assay, there is a significant reduction in colony formation in bone marrow cells expressing the MLL-AF9(D544R) fusion (Figure 21A). Loss of colony formation was not a result of a lack of MLL-AF9(D544R) expression, as this fusion protein was detected in cells harvested at the end of one week (Figure 21C).

In addition to forming fewer colonies compared to wild-type, the MLL-AF9(D544R)-expressing bone marrow cells form more diffuse colonies (Figure 21B). The cells within those colonies appear more differentiated, with an increased cell size and the presence of granules characteristic of differentiated cells in the myeloid lineage.
MLL-AF9(WT)-expressing bone marrow cells appear blast-like; they are smaller in size, lack granules, and have a high nucleus-to-cytoplasm ratio upon staining with Wright-Giemsa (Figure 21B).

Figure 21. MLL-AF9(D544R) Colony Assay

(A) Colony formation by c-kit+ progenitor cells expressing MLL-AF9(WT), MLL-AF9(D544R) or vector control throughout four weeks of replating. Colony assays were conducted in duplicate a total of six times. (B) Representative colony photographs at 4X and 10X magnification (top), and cell morphology by cytospin followed by Wright-Giemsa staining (bottom). (C) Western blot of bone marrow cells expressing empty vector, MLL-AF9(WT) or MLL-AF9(D544R) isolated after one week.
As previously mentioned, leukemic transformation by MLL fusion proteins is associated with an increase in expression of MLL target genes such as *Meis1* and *Hoxa9*. While MLL-AF9(WT) transduced bone marrow cells show a 92-fold increase in expression of *Hoxa9* and 43-fold increase in expression of *Meis1* at week 1 compared to the MSCVneo control at week one, cells transduced with MLL-AF9(D544R) show a statistically significant reduction in expression, with only 13-fold increase in *Hoxa9* expression and 15-fold increase in *Meis1* expression compared to MSCVneo control (Figure 22).

**Figure 22. Hoxa9 and Meis1 Expression in MLL-AF9(D544R) Transduced Bone Marrow Progenitor Cells**

Relative expression of MLL target genes *Hoxa9* (A) and *Meis1* (B) in bone marrow progenitor cells transduced with retroviruses encoding MLL-AF9(WT), MLL-AF9(D544R) or empty vector following one week in methylcellulose. Expression was normalized to *Hprt* and set relative to MSCVneo vector control. Three independent experiments were each performed in triplicate. Error bars indicate standard deviation.
These results indicate that the AF9-AF4 interaction mediated by the salt-bridge formed between AF9 residue D544 and the lysine at 764 of AF4 is required for efficient \textit{in vitro} leukemic transformation.

**Genetic Complementation of MLL-AF9(D544R) Restores \textit{in vitro} Transformation**

From the previous findings, we concluded that the AF9-AF4 interaction mediated by the AF9 D544 residues is required for transformation \textit{in vitro}. We therefore hypothesized that coexpression of the MLL-AF9(D544R) mutant with an AF4 mutant substituting an aspartic acid for the positively charged lysine would restore the saline interaction and would rescue transformation ability \textit{in vitro}.

Since MLL-AF4 cannot transform bone marrow cells and we did not have access to full-length AF4 cDNA, we utilized the highly conserved AF4 homolog AF5 to further examine the requirement for the AF9-AF4/5 salt bridge interaction in leukemic transformation. Functionally, AF4 or AF5 are necessary to stabilize the pTEFb elongation complex, and can bind to AF9 or the AF9 homolog ENL. AF4 K764 is conserved in all AFF family members, and functionally corresponds to amino acid K717 in AF5 (Figure 6B). Fluorescence anisotropy studies with fragments of AF9 and AF5 indicate comparable affinity of binding between AF9 and AF5, and AF9 and AF4 (kD of 0.2nM vs. 0.17nM, respectively)\textsuperscript{240}.

We hypothesized that reversing the positive charge at AF5 K717 to an aspartic acid with a negatively charged side chain would restore binding to the AF9(D544R)
mutant with a positive charge, thus rescuing colony forming ability and target gene expression.

We first wanted to assess the ability of MLL-AF5(WT or K717D) to transform bone marrow progenitor cells. We hypothesized that MLL-AF5(WT) would transform cells, and that MLL-AF5(K717D) would not transform cells as it would not form a salt bridge with wild type AF9. As expected, MLL-AF5(WT) transduced bone marrow cells continually replate, whereas MLL-AF5(K717D) expressing cells show a reduction in colony formation (Figure 23A). This corresponds to a decrease in Meis1 expression in MLL-AF5(K717D)-expressing cells as compared to wild-type (Figure 23B). MLL-AF5(WT) expressing cells exhibit greater than forty-fold increase in Meis1 expression over vector control at the end of week one whereas MLL-AF5(K717D) expressing cells express only a 27-fold increase. Interestingly, Hoxa9 expression was not altered in cells expressing the wild type or mutant fusion (Figure 23C). From these results, we conclude expression of the MLL-AF5(K717D) mutant in c-kit+ progenitor cells severely impairs the transformation ability of the MLL-AF5 fusion protein.

To test whether genetic complementation of the AF9(D544R) mutation could rescue in vitro transformation ability, c-kit+ bone marrow progenitor cells were co-transduced with YFP-tagged AF9(WT or D544R) and G418-resistant MLL-AF5(K717D) constructs. Cells were selected in G418 for one week before sorting for YFP+ cells to select for cells expressing both constructs.
Figure 23. MLL-AF5(K717D) Colony Assay and Hoxa9 and Meis1 Expression

(A) Colony formation by c-kit+ progenitor cells expressing MLL-AF5(WT), MLL-AF5(K717D) or vector control throughout three weeks of replating. Quantitative RT-PCR of (B) Meis1 and (C) Hoxa9 expression in bone marrow cells isolated after one week in methylcellulose. Expression was normalized to Hprt and set relative to MSCVneo vector control. Three independent experiments were each performed in triplicate. Error bars indicate standard deviation.
Figure 24. Genetic Complementation of MLL-AF5(K717D) by AF9(D544R): Colony Assay and Expression of Hoxa9 and Meis1

(A) Colony formation by c-kit+ progenitor cells co-expressing MLL-AF5(K717D) and AF9(WT or D544R). Quantitative RT-PCR of (B) Meis1 and (C) Hoxa9 expression in bone marrow cells isolated after one week in methylcellulose. Expression was normalized to Hprt and set relative to MSCVneo vector control. Three independent experiments were each performed in triplicate. Error bars indicate standard deviation.
While cells coexpressing MLL-AF5(K717D) and AF9(WT) are unable to serially replate in methylcellulose, coexpression of both MLL-AF5(K717D) and AF9(D544R) mutants restored colony formation ability (Figure 24A). This increase in colony formation corresponds to an increase in *Meis1* expression compared to the MLL-AF5(K717D)/AF9(WT) expressing cells (Figure 24B). *Hoxa9* expression was significantly reduced in the cells co-expressing both mutant constructs (Figure 24C). This was a very surprising result and will be discussed further in the Discussion chapter.

These results indicate that the electrostatic interaction between AF9 and AF4/AF5 is necessary for *in vitro* leukemic transformation. Loss of the interaction through introduction of charge-repulsion in AF9 can be rescued by expression of a complementary charge in the AF5/AF4.

**AF9 Interaction with DOT1L is Required for *in vitro* Leukemic Transformation**

As previously mentioned, AF9 interacts with DOT1L at three regions: DOT1L 628-653, 863-878 and 878-900. We hypothesized that mutation of aspartate D546 in AF9 to arginine would reverse the charge at this location, preventing AF9 from interacting with DOT1L. Structural studies from the Bushweller lab indicate that this amino acid substitution reduces binding of AF9 to DOT1L at all three binding interfaces (unpublished data).

To assess the relevance of the AF9-DOT1L interaction in leukemic transformation, c-kit+ bone marrow progenitor cells were retrovirally transduced with MLL-AF9 fusions, either wild type or D546R mutant. We predicted that cells expressing
the wild type fusion would serially replate for the duration of the experiment, whereas the MLL-AF9(D546R)-expressing cells would not. In cells expressing the wild type fusion gene, the AF9-DOT1L interaction would be maintained and Dot1l would be recruited via AF9 interaction to MLL target genes, leading to the deposition of the H3K79 methyl mark within the gene body of genes required for leukemic transformation. H3K79me2 is associated with expression of MLL fusion target genes, and loss of H3K79me2 is associated with a decrease in MLL target gene expression\textsuperscript{212}. If the AF9-DOT1L interaction is inhibited in cells expressing MLL-AF9(D546R), then DOT1L cannot be recruited, and there would be a reduction in H3K79 methylation potentially coinciding with a decrease in expression of target genes necessary for transformation.

Colony assays indicate that the MLL-AF9(WT) fusion-expressing cells continually replate for the duration of the experiment, whereas cells expressing the D546R mutant fusion have a statistically significant reduction in colony formation (Figure 25A). Both the colony and cell morphology were drastically different, with the MLL-AF9(WT) expressing bone marrow cells forming very tight, compact colonies comprised of mainly blast cells, while the MLL-AF(D546R) mutant expressing cells formed very few, very diffuse colonies composed of differentiated myeloid cells (Figure 25B). To assure that both fusion proteins were equally expressed, western blot analysis was performed showing that MLL-AF9(WT) and (D546R) protein expression was detectable in bone marrow cells after week one (Figure 25C).
Figure 25. MLL-AF9(D546R) Colony Assay

(A) Colony formation by c-kit+ progenitor cells expressing MLL-AF9(WT), MLL-AF9(D546R) or vector control throughout four weeks of replating. The experiment was repeated eight times in duplicate. (B) Representative colony photographs at 4X and 10X magnification (top), and cell morphology by cytospin followed by Wright-Giemsa staining (bottom). (C) Western blot of bone marrow cells expressing empty vector, MLL-AF9(WT) or MLL-AF9(D546R) isolated after one week.
If recruitment of DOT1L to MLL-AF9 target genes is required for successful elongation of transcripts, we would expect to see a decrease in *Hoxa9* and *Meis1* transcript levels in cells expressing the D546R mutant. Quantitative reverse transcription PCR of bone marrow cells isolated after week one showed a significant reduction in *Hoxa9* (Figure 26A) and *Meis1* (Figure 26B) expression in MLL-AF9(D546R)-expressing cells compared to MLL-AF9(WT).

**Figure 26. Hoxa9 and Meis1 Expression in MLL-AF9(D546R) Transduced Bone Marrow Progenitor Cells**

![Graph A](image1)

*Figure A.* Relative expression of MLL target genes *Hoxa9* (A) and *Meis1* (B) in bone marrow progenitor cells transduced with retroviruses encoding MLL-AF9(WT), MLL-AF9(D546R) or empty vector following one week in methylcellulose. Expression was normalized to *Hprt* and set relative to MSCVneo vector control. Three independent experiments were each performed in triplicate. Error bars indicate standard deviation.
From these studies, we can conclude that disruption of the AF9-DOT1L interaction by charge reversal of AF9 aspartate residue 546 prevents *in vitro* leukemic transformation.

**MLL-AF9 Mutants E506R, I538A and K557E Do Not Alter the *in vitro* Leukemic Transformation Ability of the MLL-AF9 Oncogenic Fusion**

To test the functional effects of AF9 mutants E506R, I538A and K557E, all predicted to disrupt the structure of AF9 itself, c-kit+ bone marrow progenitors were retrovirally transduced with the empty MSCVneo vector which is not expected to transform cells, MLL-AF9(WT) which is known to transform progenitor cells, and the experimental mutants MLL-AF9(E506R), MLL-AF9(I538A) or MLL-AF9(K557E). If protein interactions mediated by secondary structure imposed by E506, I538, or K557 of AF9 are required for transformation, then mutation of these residues will result in a decrease in colony formation.

There was no significant difference in the ability of MLL-AF9(WT) or mutant fusions to transform bone marrow progenitor cells over the course of the experiment (Figure 27A). Colony morphology was similar between all constructs, with the exception of cells expressing MLL-AF9(I538A) which exhibited a slightly more diffuse morphology and ragged edges compared to the tight, compact colonies formed by cells expressing the other constructs (Figure 26B). The individual cells within these colonies also appear slightly more differentiated than the MLL-AF9(WT) control cells (Figure 26B).
Figure 27. MLL-AF9 Mutants E506R, I538A and K557E Colony Assay

(A) Colony formation by c-kit+ progenitor cells expressing MLL-AF9(WT), MLL-AF9(E506R), MLL-AF9(I538A), MLL-AF9(K557E) or vector control throughout four weeks of replating. Experiments were conducted in duplicate five to six times. (B) Representative colony photographs at 4X and 10X magnification (top), and cell morphology by cytopsin followed by Wright-Giemsa staining (bottom).
Target gene expression was also unaffected by expression of the mutants compared to wild type. Neither *Hoxa9* nor *Meis1* levels were statistically reduced in mutant-expressing cells compared to wild-type MLL-AF9-expressing cells (Figure 28).

**Figure 28. Hoxa9 and Meis1 Expression in MLL-AF9 Point Mutant-Transduced Bone Marrow Progenitor Cells**

Relative expression of MLL target genes *Hoxa9* (A) and *Meis1* (B) in bone marrow progenitor cells transduced with retroviruses encoding MLL-AF9(WT), MLL-AF9(E506R), MLL-AF9(I538A), MLL-AF9(K557E) or empty vector following one week in methylcellulose. Expression was normalized to *Hprt* and set relative to MSCVneo vector control. Three independent experiments were each performed in triplicate. Error bars indicate standard deviation.

Based on these results, we conclude that residues E506, I538 and K557 are not critical for AF9 structure or mediating protein interactions required for *in vitro* leukemic transformation.
MLL-AF9 Fusion Proteins Resulting from Two Different Breakpoints in AF9 Are Equal in Transforming Potential

In MLL leukemia, the N-terminus of MLL breaks at the breakpoint cluster region and fuses with one of over 79 fusion partners. To date, five distinct breakpoint regions have been identified in AF9\textsuperscript{154, 241, 243}, one of the most common MLL fusion partners.

The most frequent breakpoint in AF9, occurring in approximately 75% of patients with t(9;11), results in a MLL fusion protein expressing amino acids 376-568 of AF9. The second most common breakpoint, occurring in 12% of patients, results in a shorter fragment of AF9 including amino acids 475-568\textsuperscript{154, 241, 243}. Knock-in studies of MLL-AF9(475-568) under the control of endogenous Mll regulatory elements results in AML development in mice, indicating that the 93 amino acids of AF9 present in this fusion are sufficient for transformation\textsuperscript{147}.

Because there is a one hundred amino acid difference between the fragment of AF9 most commonly found in patients and the construct that is used to study the function of the fusion in the laboratory, we sought to compare the constructs side-by-side in a colony formation assay. The presence or absence of additional residues could differentially modulate protein-protein interactions, and thus \textit{in vitro} studies may not be an accurate representation of what is happening \textit{in vivo}.

To assess this, we transduced cells with retrovirus encoding MLL-AF9 fusions, either with the 192 C-terminal amino acids found in MLL patients (376-568) or the shorter 475-568 fragment, as well as a MSCVneo empty vector control. There were no statistically significant changes in colony formation during the first three rounds of
plating. By week four, the shorter fusion gained a statistically significant advantage (Figure 30A). There were no detectable changes in colony or cell morphology (Figure 30B), and there was no significant change in *Hoxa9* or *Meis1* expression in cells expressing the MLL-AF9(376-568) or MLL-AF9(475-568) fusion (Figure 29A).

**Figure 29. MLL-AF9(376-568) and (475-568) Colony Assay**

(A) Colony formation by c-kit+ progenitor cells expressing wild-type MLL-AF9(376-568), MLL-AF9(475-568) or vector control throughout four weeks of replating. (B) Representative colony photographs at 4X and 10X magnification (top), and cell morphology by cytospin followed by Wright-Giemsa staining (bottom).
Figure 30. *Hoxa9* and *Meis1* Expression in Bone Marrow Cells Transduced with MLL-AF9(376-568) or (475-568)

Relative expression of MLL target genes *Hoxa9* (A) and *Meis1* (B) in bone marrow progenitor cells transduced with retroviruses encoding wild-type MLL-AF9(376-568), MLL-AF9(475-568) or empty vector following one week in methylcellulose. Expression was normalized to *Hprt* and set relative to MSCVneo vector control. Three independent experiments were each performed in triplicate. Error bars indicate standard deviation.
From these results, we conclude that there is no significant difference in the transformation ability of the MLL-AF9 fusion resulting from two different breakpoint regions of AF9 at weeks one, two and three. There is a significant increase in colony formation in cells expressing MLL-AF9(475-568) which will be further discussed.

**Productive Transcription Elongation is Altered in MLL-AF9(D544R and D546R)-expressing Bone Marrow Cells, but not MLL-AF9 Mutants E506R, I538A or K557E-expressing Cells**

We hypothesized that specific point mutations in the C-terminus of AF9 would disrupt binding to one or more of the proteins known to interact with AF9. AF4, DOT1L and CBX8 have all been identified as members of P-TEFb-associated complexes. To examine whether these point mutations affect productive elongation, we used primers designed to amplify four regions of the *Meis1* gene. The *Meis1* genomic region is approximately 138 kilobases (kb), encoding a 3.3kb mRNA transcript containing 12 exons. We hypothesized that the proper and stable assembly of the P-TEFb elongation complex is required for productive elongation of this long gene. Mutations in AF9 that disrupt binding to AF4, DOT1L or CBX8 would disrupt the formation or function of the Super Elongation Complex, and would prevent productive elongation as measured by quantitative real-time PCR using primer sets that span the length of the transcript.

Primer sets were used to assess transcription of four regions of *Meis1* in MLL-AF9 wild type or mutant-transduced bone marrow cells after one week in methylcellulose. One primer set was located in exon 1, 21 nucleotides after the transcription start site (Figure 31A). The second primer set spanned exons 6 and 7, the third primer set spanned exons 8 and 9, and the final primer set was located in the last exon of *Meis1*. 
Figure 31. Transcriptional Elongation of *Meis1* in MLL-AF9 Wild-Type or Mutant Transformed Bone Marrow Progenitor Cells

A) *Meis1* locus encoding 12 exons in a 138 kb region (UCSC Genome Browser). Primer sets were designed to amplify regions within exon 1, exons 6/7, exons 8/9 or exon 12. B) Quantitative RT-PCR for expression of *Meis1* with primers amplifying exon 1, exon 6/7, exon 8/9 or exon 12. Expression was normalized to Gapdh and set relative to MLL-AF9(WT). Three independent experiments were each performed in triplicate. Error bars indicate standard error.
If AF9 interaction with AF4, DOT1L, or CBX8 is required for RNA polymerase II to successfully track along the length of the Meis1 gene, then expression of mutants disrupting these interactions would result in reduced expression of Meis1 at the 3’ end of the gene compared to the 5’ end.

Bone marrow cells transduced with MLL-AF9(D544R) or (D546R) exhibit decreased expression of Meis1 when amplified using primers for exon 1 compared to MLL-AF9(WT)-transduced cells. MLL-AF9 mutants E506R, I538A, and K557E show a modest reduction in Meis1 expression using this primer set. Levels of Meis1 remain low in MLL-AF9(D544R or D546R) expressing cells when using primers for exon 6/7, exon 8/9 and exon 12 compared to wild type. Expression levels of Meis1 as measured by the exon 6/7 primers indicate comparable levels of expression in the E506R, I538A and K557E mutants compared to MLL-AF9(WT) expressing cells. Analyzing transcript levels further down the Meis1 gene with the exon 8/9 primer set indicates a modest reduction in cells expressing MLL-AF9(E506R) whereas the I538A and K557E mutants are similar to wild type. Expression at the 3’ end of Meis1 using the exon 12 primer set indicates no change in expression in the E506R, I538A or K557E mutants compared to wild type.
CHAPTER FIVE

DISCUSSION

Aim One: The Role of Af9 in Hematopoiesis

In the first aim of my project, I sought to examine the role of Af9 in hematopoietic development, and to determine which protein interactions mediated by Af9 were required for normal hematopoiesis. As mentioned previously, the C-terminal AHD of AF9 interacts with at least four proteins: AF4, DOT1L, CBX8 and BCoR. In collaboration with the laboratory of Dr. Bushweller at the University of Virginia, specific point mutations were identified that were predicted to disrupt binding to one partner protein while maintaining binding with the others. Using these AF9 mutants, we sought to elucidate which AF9-mediated protein interactions are required for proper hematopoietic development.

All point mutations were readily expressed, with mutants D544R, I538A, S565A and S565D expressing similarly to wild type AF9. The mutants D546R, E506R, K557E and L562A were expressed at levels slightly less than wild-type. While this reduction is minimal, it is not unexpected considering the location of these mutations within the structured AF9 protein.

Residues E506, K557 and L562 are all present in alpha helices one or three of AF9 (Figure 8). The formation of alpha helical structure is highly sequence-dependent,
and disruption of the structure could result in misfolding and potential degradation. The D546 residue was identified as mediating a potential salt bridge interaction with DOT1L.

It is possible that AF9 must bind to another protein to adopt structure and prevent degradation, and that mutation of this residue may inhibit the stabilizing interaction. The physiological role of intrinsically disordered proteins (IDPs) is not well understood, and we do not know if other proteins, such as heat shock proteins, are required for the stabilization of proteins like AF9 in vivo.

**AF9(D546R) Interaction with AF4 and DOT1L**

As part of her dissertation studies, Dr. Bhavna Malik examined the interaction of AF9 point mutants with AF4, DOT1L and CBX8. She co-expressed the C-terminal 194 amino acids of AF9 (376-568), wild type or mutants D544R, E506R, K557E, L562A, S565A or S565D with AF4, DOT1L or CBX8. The D546R mutation was later identified as a potential disruptor of binding between AF9 and DOT1L, and so I examined the interaction of AF9 wild type or D546R mutant with AF4 and DOT1L by coimmunoprecipitation.

The C-terminal fragment of AF9 most often found in patients with MLL-AF9 rearrangement (376-568) or a smaller fragment (470-568) that is sufficient to transform bone marrow cells and induce leukemia in a MLL-AF9 knock-in mouse model was coexpressed\textsuperscript{147} with full length Dot1l, Dot1l 479-659, or DOT1L 828-1095. As mentioned previously, the hydrophobic motif present in all AF9 binding partners (Figure 32) is repeated three times in the DOT1L protein. The 479-659 construct encompasses
one site, whereas the 828-1095 fragment includes the two C-terminal hydrophobic motifs (referred to as sites two and three, Figure 33).

**Figure 32. Conservation of a Hydrophobic Motif in the AF9 Interacting Regions of AF4/AF5q31, DOT1L, BCoR and CBX8**

<table>
<thead>
<tr>
<th>Conserved Hydrophobic Motif</th>
<th>Published AF9 Interaction Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF4: 751-PLRDTPPPQSIMKTEEDLLSRIQPQPGKG-780</td>
<td>766-779</td>
</tr>
<tr>
<td>AF5q31: 705-PLSPEPDPPRLIVKEEDNLLTRIPGKPYKE-734</td>
<td>720-733</td>
</tr>
<tr>
<td>DOT1L SITE 1: 628-SQISEKQRHCELQSFYVEKSHQEQELL-657</td>
<td></td>
</tr>
<tr>
<td>SITE 2: 855-YGSSGELTSTPSLPLSTVQPKLVSIP-884</td>
<td>828-1094</td>
</tr>
<tr>
<td>SITE 3: 869-IPLSTVPQKLPVLSAVLPSRAERAR-898</td>
<td></td>
</tr>
<tr>
<td>BCoR: 1194-EDPHYSELTNKVCYELTGHPPKQRHLLH-1223</td>
<td>1127-1251</td>
</tr>
<tr>
<td>CBX8: 323-DMGAAQGGRPSLLLAPVPARTILGDPDEESWS-352</td>
<td>202-333</td>
</tr>
</tbody>
</table>

Conserved LxVxIxL/VxxL/V/I hydrophobic motif in AF9 interacting partners as identified by Leach et al.

Coimmunoprecipitation studies show that expression of the AF9(D546R) mutant, either in the context of the larger AF9 C-terminal fragment (376-568) or the shorter fragment (470-568), results in reduced binding to full length Dot1l and to the fragment of DOT1L encompassing the two C-terminal hydrophobic motifs (Figure 33). These results indicate that the binding between AF9 and DOT1L is mediated in part by AF9 residue D546.

Binding between AF9(376-568) and a Dot1l fragment including the site one hydrophobic motif was not significantly affected by the D546R mutant, whereas expression of AF9(D546R) in the context of the shorter AF9 fragment resulted in an increase in binding to the Dot1l fragment encompassing site one (Figure 33). The increase in binding between the shorter fragment of AF9(D546R) and Dot1l(479-659), which is also a trend but statistically insignificant with the larger AF9(D546R) fragment,
maybe be a result of different affinity binding between Dot1l site 1 versus sites 2 and 3. Higher-affinity interactions may be more drastically affected by the introduction of a mutation within the region, whereas lower affinity interactions may not be affected. Why the D546R mutation in the context of the short AF9 fragment enhances binding over wild type AF9 could be a result of enhanced affinity of the D546R mutant to a heat shock protein or other protein necessary to stabilize AF9, which is reduced upon inclusion of the additional amino acids in the larger fragment of AF9.

**Figure 33. AF9(WT or D546R) Binding to DOT1L**

Summary of immunoprecipitation studies with DOT1L full length, 479-659 or 828-1095 and AF9(WT or D546R) fragments 470-568 and 376-568. Repeated hydrophobic motifs shown in Figure 32 are termed sites 1, 2 and 3.

Indeed, work by Lin et al showed that AF9 interacts with HSP90 and associated co-chaperones HSP70 and p60/HOP\textsuperscript{244}. The AF9-HSP90 interaction was identified using amino acids 475-568 of AF9, the same region of AF9 that interacts with AF4, DOT1L, CBX8 and BCoR. As discussed, structural studies suggest that AF9 binding to one of
these four partner proteins is mutually exclusive. It would be informative to compare binding of AF9 and HSP90 by HSQC spectral analysis to the HSQC spectra of AF9 in complex with AF4/DOT1L/CBX8/BCoR. If all five AF9-interacting proteins bind to the same residues of AF9 and induce the same structure, we would envision a sequential binding model in which AF9 is translated in the cytoplasm, is transported into the nucleus by through association with HSP90, and then is released by HSP90 to bind AF4, DOT1L, CBX8 or BCoR.

Expression of the D546R mutant does not alter the binding to AF4. Coexpression of the same AF9 fragments (376-568 or 470-568), wild type or mutant, with two fragments of AF4 encompassing either the minimal AF9 interaction region (755-777) or a slightly larger region including the AF9 binding domain (647-871), did not result in any changes in binding between AF9 and AF4 (Figure 34). The exception to these results is seen upon coexpression of the shorter 470-568 AF9(D546R) fragment with the minimal AF4 fragment 755-777, which resulted in a statistically significant increase in binding. As with binding to Dot1l, it appears that expression of the mutant in the 98 amino acid fragment of AF9, but not the 192 amino acid fragment, enhances binding with the minimal fragment of AF4 (or Dot1l, as discussed above). It may be more energetically favorable for smaller proteins to interact\textsuperscript{245}. Smaller proteins could have a lower degree of entropy, and thus require lower activation energy for the conformational change to occur upon binding to a partner protein. The electrostatic charges present also alter entropy, and this could explain why the mutant D546R binds better to the minimal fragment of AF9 compared to wild type AF9. Biochemical analyses would have to be
conducted to determine the thermodynamics of binding between the longer and short fragments of AF9 (wild type or D546R) with AF4 and DOT1L fragments.

**Figure 34. AF9(WT or D546R) Binding to AF4**

Summary of immunoprecipitation studies with AF4 fragments 755-777 and 647-871 and AF9(WT or D546R) fragments 470-568 and 376-568.

In all of these coimmunoprecipitation studies, I overexpressed all proteins, either full length or fragments of proteins. Coimmunoprecipitation of endogenous proteins may show different results, as the additional domains in these proteins may also mediate binding of AF9 to a partner protein. However, the fragments examined here are the amino acids of AF9 present in MLL fusion proteins, and thus these studies are relevant in examining the role of these protein interactions in MLL-AF9 leukemias.

Overall, these results highlight the importance of the AF9 D546 residue in mediating interaction between AF9 and DOT1L. The structural studies from the Bushweller lab identified a conserved hydrophobic motif present in all AF9 binding partners, and suggested that AF9 binding to any of the four proteins AF4, DOT1L, CBX8
or BCoR is mediated through this motif. However, the selective disruption of binding to DOT1L but not AF4 with the AF9(D546R) residue indicates that interaction between these two proteins is mediated, at least partially, through a region distinct from the hydrophobic motif.

Of note, coexpression of AF9(D546R) with any of the four fragments used in these immunoprecipitation experiments resulted in decreased detectable AF9(D546R) protein in the input compared to wild-type, despite transfecting the same amount of DNA. This result was reproducible and was evident even when the amount of plasmid DNA transfected was adjusted in an attempt to normalize the input amounts, suggesting that AF9(D546R), in the context of the larger 376-568 or the shorter 470-568 fragment, is less stable than the same fragments of wild type AF9. This mutant was also expressed at a slightly lower level than wild type in the context of the full length protein as seen in Figure 8B. As discussed, this mutant may be less stable than wild type AF9 as a result of decreased affinity binding to a protein required for stabilization such as a heat shock protein. A recent study has also shown that the intrinsically disordered C-terminus of DNA glycosylase NEIL1 is stabilized through intramolecular, primarily electrostatic, interactions. It is possible that the D546R mutation prevents potential intramolecular interactions from occurring, and thus the protein is degraded.

As reviewed by Dyson and Wright, many transcriptional regulators are intrinsically disordered proteins (IDPs). The process of coupled folding and binding, in which an unstructured protein (or unstructured region of a protein) adopts a structure upon interaction with a target, can allow for rapid and promiscuous protein-protein
interaction. In the case of AF9, the ability to bind both activators and repressors could allow for differential regulation of gene targets depending on other, yet unknown signals in the cell. As disordered proteins are relatively unstable, they could be degraded without chaperoning by another protein. Whether AF9 is complexed to other proteins in vivo that promote its stability is unknown but would be an interesting question to pursue.

Of the studies published on the complex components of the SEC, only the Biswas group identified proteins by biochemical purification of AF9. The authors overexpressed FLAG-HA-tagged AF9 and examined associated proteins. We could interrogate additional AF9-interacting proteins, and determine which of the proteins are required for AF9 stability by knocking down candidates proteins and examining the levels of AF9.

**The Role of Af9 in Hematopoietic Differentiation**

To examine the role of Af9 in in vitro hematopoietic differentiation, I isolated c-kit+ progenitor cells from the bone marrow of Af9flo/flo Cre/+ mice and plated these cells in semi-solid media to promote the differentiation to erythroid, myeloid or megakaryocytic lineages in the presence or absence of 4-OHT. We hypothesized that loss of Af9 upon 4-OHT treatment would result in a loss of erythroid and megakaryocyte colony formation, with a possible compensatory increase in myeloid lineage differentiation. However, results show that Af9 is not required for in vitro differentiation to erythroid, megakaryocyte, or myeloid lineages under the conditions tested here.

**Erythroid Lineage Differentiation**

While the number of erythroid colonies formed was not different in the presence or absence of Af9, there is a significant decrease in the mean fluorescence intensity (MFI)
of CD71 expression in cells null for Af9 compared to the ethanol control treated cells. CD71 is marker of early erythroid precursor cells, so the decrease in MFI suggests that while these cells are still expressing CD71, the level of expression is reduced upon loss of Af9. Af9 may therefore play a role in the differentiation to erythroid precursor cells, but this reduction is either not sufficient to alter overall colony formation ability, or it is occurring in only a subset of cells that are not required for colony formation. There was no change in Ter119 expression, and expression overall is low, which is consistent with previous reports that Ter119 is expressed later in erythroid differentiation than the time points examined here.

Examination of myeloid markers CD41, Gr-1 and CD11b in cells plated for erythroid lineage differentiation did not indicate significant differences in cells expressing or lacking Af9. CD41 levels are low in both control and 4-OHT-treated cells, which is expected as CD41 is expressed on the surface of platelets and megakaryocytes, both mature cell types that would not be present at this early four day time point. We would expect to see CD41 expression, as well as expression of other myeloid surface markers, appear by twelve days in culture. Levels of Gr-1 and CD11b are elevated in these erythroid cytokine-cultured bone marrow cells, with an average MFI of approximately 2500 relative units for Gr-1 and CD11b, but levels are not as high as in Af9\(^{fl/fl}\)\ cre/+ plated for myeloid differentiation in the presence or absence of 4-OHT with average MFIs of 3500 and 4000 for Gr-1 and CD11b, respectively. This indicates that despite being cultured in the presence of cytokines to promote erythroid lineage commitment, there is at least a subset of cells that differentiate to the myeloid lineage.
Interestingly, while the differences in MFI between ethanol and 4-OHT treated cells is not significant, there is a trend toward increased myeloid marker expression in cells lacking Af9. This would support our hypothesis that a loss of Af9 results in a loss of erythroid lineage commitment (decreased MFI of CD71) with concurrent compensation by myeloid lineage differentiation (increased MFI of Gr-1 and CD11b). Additional replicates may elucidate whether this is a real effect, or if it is just variability between experiments.

**Myeloid Lineage Differentiation**

There was no significant difference in myeloid colony formation in cells expressing or lacking Af9. Examination of surface marker expression does not indicate any statistically significant differences in the percent of positive cells or the mean fluorescence intensity of either myeloid (CD41, Gr-1, CD11b) or erythroid (CD71, Ter119) expression. Compared to cells plated for erythroid lineage differentiation, these myeloid-directed cells express higher levels of CD41, Gr-1 and CD11b. In terms of erythroid marker expression in these cells, we do see expression of CD71 in both control and 4-OHT-treated cells, but the average MFI for these cells is less than half of that in cells plated for erythroid lineage differentiation. Ter119 expression is elevated in the myeloid-cultured cells as compared to cells cultured in erythroid cytokines. This is likely a result of the later time point at which these cells were harvested, as Ter119 is expressed later in erythroid development than the time point at which the erythroid cultured cells were harvested.
The presence of erythroid markers on these cells suggests that at least a subset of cells are able to differentiate toward the erythroid lineage and acquire erythroid lineage markers, despite being cultured in the presence of cytokines to promote myeloid lineage differentiation. This may partially be due to the c-kit+ selection of murine bone marrow cells for plating in these assays. C-kit is expressed on HSCs as well as committed progenitor cells, including the MEP and the GMP. Using this protocol, the cells plated were a heterogeneous population of stem and progenitor cells, which would express different surface markers depending on the stage of differentiation at the time of isolation. I could examine differences in surface marker expression and colony formation using lineage (Lin)-depleted cells, or isolating Lin-Scal+c-kit+ cells (LSK) which enrich for HSCs.

**Megakaryocytic Lineage Differentiation**

As loss of Af9 does not alter erythroid or myeloid colony formation, it also does not alter megakaryocyte differentiation as detected by acetylcholinesterase staining. The method that I used to differentiate and quantify megakaryocyte differentiation does not allow for surface marker analysis, but based on colony formation alone, Af9 is not required for megakaryocytic development. Consistent with these in vitro studies on the role of Af9 in hematopoietic lineage differentiation, I did not see alterations in the levels of mRNA encoding erythroid or myeloid transcription factors in the bone marrow of Af9\textsuperscript{fl/fl} or Af9\textsuperscript{fl/fl Cre/+} mice treated with tamoxifen. These results combined suggest that Af9 is not a master regulator of murine erythroid/megakaryocytic lineage commitment.
Re-expression of AF9 in Af9-null Bone Marrow Progenitor Cells

In addition to examining the role of Af9 in hematopoietic development, I proposed to examine the role of specific Af9-partner protein interactions in hematopoietic development by re-expressing full length AF9 point mutants in Af9 floxed Cre+ c-kit+ bone marrow cells. We originally hypothesized that loss of Af9 would result in a loss of erythroid and megakaryocytic colony formation that could be rescued by re-expression of wild type AF9. These experiments were the first I conducted before obtaining the results described above, but these experiments support the findings I reported.

Vector transduced cells formed the same number of colonies when plated for erythroid lineage commitment in the presence or absence of 4-OHT, just as untransduced cells formed the same number of colonies with or without 4-OHT treatment. Re-expression of wild-type AF9 or AF9 mutants S565A or S565D did not alter colony forming ability. While there was no phenotype to rescue in this experiment, since loss of Af9 did not alter colony formation, these results also suggest that overexpression of AF9 does not alter colony formation. Overexpression of AF9 in CD34+ CD38- human cord blood resulted in an increase in the E/Meg compartment\(^{159}\), whereas here I saw no change upon retroviral transduction of AF9 in c-kit+ cells. I did not compare the levels of overexpressed AF9 with endogenous levels of Af9, but I would expect that high-titer retroviral transduction of AF9 under the control of the Murine Stem Cell Virus promoter would result in very high expression of AF9 in these cells, likely over that of endogenous levels and therefore comparable to an overexpression system.
In the study by Pina and colleagues, lentiviral knockdown of AF9 in human cord blood resulted in a loss of erythroid progenitor colony formation. Here, complete genetic loss of \( Af9 \) in murine c-kit+ bone marrow cells as detected by genotyping PCR does not result in decreased colony formation. Overexpression of AF9 in human CD34+CD38- cells caused an expansion of erythroid/megakaryocytic cells, but re-expression of AF9 in \( Af9^{-/-} \) c-kit+ cells does not alter colony formation. These results indicate that Af9 does not play a role in hematopoietic differentiation in this system.

**Af9 is Not a Critical Regulator of Adult Murine Hematopoiesis**

The results described refute our hypothesis that Af9 is a critical regulator of erythroid and megakaryocytic lineage commitment. This is contrary to the study by Pina and colleagues\(^{159} \), but may align with the report of hematopoiesis in the constitutive \( Af9 \) knockout mouse. Collins et al reported that hematopoiesis was normal\(^{158} \), but no data were provided so it is impossible to compare our findings.

The discrepancy between our results and that of the Pina study could be a result of differences in human versus murine hematopoiesis. Those authors used human cord blood cells to examine the role of AF9 whereas we utilized a conditional \( Af9 \) knockout mouse. There are many similarities in hematopoietic development between mice and humans, but there are also many differences. A recent study compared the transcriptional profile of early, intermediate and late erythroid progenitor cells in humans and mice. These researchers found that at each stage of differentiation, the top 500 most highly expressed genes in human were orthologous to only 30% of the top 500 most highly expressed genes in mice\(^{248} \). Phenotypically, erythropoiesis was not altered, but these
results suggest that the transcriptional regulation of this process is highly divergent between humans and mice. This is something to keep in mind when looking to translate studies of mouse cells to the human system. The use of humanized mouse models, in which human hematopoietic cells are grafted into mice, may provide a more translational approach to studying hematopoiesis.\textsuperscript{249}

Of note in regard to human hematopoiesis, global transcriptome analysis showed that \textit{AF9} expression is highest in the HSC, and levels decrease by half by the CMP stage\textsuperscript{159}. If \textit{AF9} is a critical player in erythroid/megakaryocytic development, we would expect high levels throughout differentiation to the CMP and MEP. However, this may support the Adolfsson model of hematopoiesis in which the MEP arises directed from the ST-HSC without transit through a CMP stage\textsuperscript{15}. High levels of \textit{AF9} could support differentiation directly to the erythroid/megakaryocytic lineage from the ST-HSC, playing a role in earlier hematopoietic lineage commitment rather than terminal differentiation from a later progenitor.

The source of the cells in which the role of \textit{Af9} was assessed may also play an important role. The Pina et al study isolated human CD34+CD38\textsuperscript{-} cord blood cells\textsuperscript{159}. We isolated murine c-kit\textsuperscript{+} bone marrow cells. In addition to isolating cells from different species, there are several differences in cord blood versus bone marrow stem cells. Patients transplanted with umbilical cord blood stem cells exhibit a low severity Graft vs. Host Disease (GVHD) compared to patients transplanted with bone marrow stem cells, although the patients receiving cord blood stem cells had a delayed engraftment\textsuperscript{250}. Umbilical cord blood also contains the largest percentage of CD34+CD38\textsuperscript{-} stem cells.
compared to bone marrow or granulocyte colony-stimulating factor-mobilized peripheral blood stem cells\textsuperscript{251, 252}. A study published by the Staal lab directly compared the expression profile of CD34+ cells isolated from umbilical cord blood and bone marrow, and identified 51 differentially expressed genes\textsuperscript{253}. These genes were involved in proliferation, apoptosis, differentiation and engraftment capacity, confirming that stem cells isolated from different sources may have different characteristics.

Another reason we did not see an effect on hematopoietic differentiation \textit{in vitro} upon loss of \textit{Af9} may be a result of compensation by \textit{Af9} homolog \textit{Enl}. As previously discussed, \textit{Af9} and \textit{Enl} are highly homologous in their N and C-terminal regions. While these proteins are generally expressed in different tissues (discussed in Chapter One), it is possible that loss of \textit{Af9} in hematopoietic cells could result in a compensatory increase in \textit{Enl} expression.

To test this hypothesis, we could transduce \textit{Af9}^{fl/fl} \textit{Cre/+} mice with control (scrambled) or \textit{Enl} shRNAs and plate the cells with or without 4-OHT in cytokines for erythroid, myeloid or megakaryocytic lineage commitment. If \textit{Enl} is compensating for loss of \textit{Af9} as a regulator of hematopoietic differentiation, then genetic loss of \textit{Af9} combined with shRNA-mediated knockdown of \textit{Enl} may result in a loss of E/Meg colonies, as we initially hypothesized. In addition to examining hematopoietic colony formation, we could also examine surface marker expression and expression of hematopoietic lineage transcription factors as previously described.

Whether AF9 and ENL play redundant or distinct roles within the cell is an interesting question. The proteins are highly homologous, especially in the N and C
terminal regions. AF9 and ENL are interchangeable members of the SEC, however they show distinct tissue expression patterns. In leukemia, the MLL-AF9 fusion protein results in a myeloid leukemia, whereas MLL-ENL fusion proteins often present as biphenotypic or lymphoid leukemia. If the proteins are so similar, and the fusion partner dictates the leukemic phenotype in MLL recombined leukemias, we would expect MLL-AF9 and MLL-ENL to transform cells similarly. Further studies elucidating the role of these two homologs is needed to identify their function in both hematopoiesis and leukemogenesis.

**Aim Two: The Role of AF9 in Leukemogenesis**

In my second aim, I sought to determine whether protein interactions mediated by the C-terminus of AF9 in the context of the MLL-AF9 fusion protein were necessary for *in vitro* leukemic transformation. The first mutant fusion gene examined was MLL-AF9(D544R). As previously described, this mutation is predicted to disrupt an electrostatic interaction mediated by AF9 aspartic acid D544 and AF4 lysine 764 by reversing the charge at the AF9 544 residue. Coimmunoprecipitation studies using the 376-568 amino acid C-terminal fragment of AF9 shows that expression of the D544R mutant reduces binding to AF4 compared to wild type, while maintaining interaction with DOT1L and CBX8.240

**MLL-AF9(D544R) Mutation Decreases Leukemic Transformation Potential**

To assess the function of this mutation *in vitro*, I transduced c-kit+ bone marrow cells with retrovirus encoding the MLL-AF9 fusion, either wild type or D544R. Colony formation was significantly reduced in cells expressing the MLL-AF9(D544R) fusion compared to wild-type, and the colonies that did form were smaller and more diffuse than
the wild type control. The cells within these colonies were more differentiated than the blast-like MLL-AF9(WT) expressing cells. Expression of MLL target genes *Hoxa9* and *Meis1* were also reduced in these cells at the end of week one. Both *Hoxa9* and *Meis1* are shown to be upregulated in MLL leukemias, and are required for transformation\(^{254}\).

The results indicate that the AF9 interaction with AF4 specifically is required for *in vitro* leukemic transformation. Because the D544R mutant was shown by coimmunoprecipitation studies to maintain interaction with DOT1L and CBX8\(^{240}\), this suggests that the selective loss of interaction with AF4 is responsible for decreased transformation ability.

These findings are further supported by binding data from the Bushweller lab, indicating that the dissociation constant between fragments of wild type AF9 and AF4 is 0.17nM. Mutation of wild type AF9 to D544R results in a decrease in binding affinity to a kD of 9.6nM. Charge reversal of AF4 lysine residue K717 to positively charged asparagine restored binding to AF9(D544R) in colocalization studies conducted by Dr. Ming-jin Chang, supporting the hypothesis that this salt bridge interaction is required for binding between AF9 and AF4\(^{240}\).

**Transformation is Rescued by Restoration of the AF4-AF9 Salt Bridge Interaction**

We hypothesized that like the colocalization studies, complementary charge reversal of AF4 K764 to glutamate could rescue the loss of colony formation by the MLL-AF9(D544R) mutant-expressing cells. The original experimental design was to coexpress MLL-AF9(D544R) with AF4(WT or K764), but overexpression of AF4 was toxic to cells. MLL-AF4 does not transform progenitor cells\(^7\), so coexpression of MLL-
AF4(WT or K764E) with AF9(WT or D544R) was not an option. We therefore utilized the AF4 homolog AF5q31 which, as shown in Figure 6B, is highly homologous to AF4. AF4 K764 is equivalent to AF5q31 K717, a conserved lysine in members of the AF4 family that mediates the salt bridge interaction with AF9. We do not have access to the AF5 clone, and thus our experimental design was to coexpress MLL-AF5q31(WT or K717D) with full length AF9 (WT or D544R).

We confirmed the findings by Yokoyama that MLL-AF5 efficiently transforms bone marrow progenitor cells. Expression of MLL-AF5(K717D) results in reduced colony forming ability, supporting our model that the charge reversal introduced by this amino acid substitution prevents interaction with AF9 at amino acid D544. We also examined MLL target gene expression in these cells and saw a significant reduction in Meis1 expression in MLL-AF5(K717D)-expressing cells compared to wild type. Interestingly, Hoxa9 was not altered between these mutants.

As far as I am aware, there have been no studies published showing differential expression of Hoxa9 and Meis1 in colony assays with MLL fusion proteins. Hoxa9 and Meis1 are specifically upregulated in ALL cases with MLL translocations compared to non-MLL recombined ALLs\(^{255,256}\). Additionally, transcriptional analysis of Hoxa9 and Meis1 show upregulation of both genes in 13 of 14 examined ALL cases with MLL-AF4 translocations, and eight of eight AML patients with MLL recombination\(^{257}\). A previous study showed that expression of Hoxa9 or Meis1 alone could not transform bone marrow cells and cause leukemia, but coexpression of Hoxa9 and Meis1 resulted in leukemia development and death in mice with a median latency of 67 days\(^{258}\). Based on these
findings, we expected to see an increase in *Hoxa9* and *Meis1* in cells expressing MLL-AF5(WT), with a decrease in expression of both genes in cells expressing MLL-AF5(K717D), correlating with colony formation. However, that was not what we saw.

The N-terminus of MLL that is present in the fusion proteins directs the fusion to gene targets, and it is thought that the C-terminal fusion partner provides a gain-of-function that results in aberrant expression of gene targets. I expect that both MLL-AF5 wild type and K717D are both being recruited to gene targets, and that could be confirmed with chromatin immunoprecipitation studies. It is therefore likely that the additional components being recruited to these loci are either different, or are differentially required, between MLL-AF5 wild type and mutant expressing cells.

In terms of the general recruitment to target loci, AF5 and AF4 interchangeably participate in the P-TEFb elongation complex. Both proteins bind directly to Cyclin T1 or T2 of the P-TEFb dimer; residues 1-308 of AF4 and 1-300 of AF5 have been shown to mediate this interaction\(^{183, 188}\). In the MLL fusion protein, the fusion point within AF4 occurs around amino acid 360, thus the MLL-AF4 fusion does not retain Cyclin T1 binding. AF4 and family members homo- and hetero-dimerize with other members of the family via their C-terminal homology domain, with a preference for heterodimerization\(^7\). Therefore, in the example of the experiment discussed here, MLL recruits the MLL-AF5 fusion to the target gene. AF5 then homodimerizes with full length AF5 or heterodimerizes with AF4. The recruited AF4/5 can then bind to Cyclin T1 that interacts with CDK9 to form P-TEFb, which binds to RNA Pol II. The recruited AF4/5, as well as the portion of AF5 in the MLL fusion, can also bind to AF9/ENL through the minimal
AF9/ENL interaction domain conserved in these proteins. AF9/ENL stabilize the P-TEFb elongation complex and aid in its processivity during transcription.

The presence of the AF5(K717D) mutation disrupts the salt bridge interaction between AF5 and AF9, and that would occur regardless of the loci to which it is targeted. Therefore, the most likely explanation is that the proper formation of this SEC is required at Meis1 to a greater degree than it is at Hoxa9. Meis1 is a much longer gene than Hoxa9, encompassing a genomic region of 138 kilobases compared to only 5.6 kilobases for Hoxa9. As RNA polymerase transcribes a gene, it often stalls and requires the components of the SEC to overcome the stall and continue to process along the template DNA. Without the proper formation of this complex, the Polymerase may be unable to complete transcription of Meis1, whereas the association with SEC components may be less necessary at Hoxa9.

In regard to the restoration of the salt bridge interaction by charge reversal on both AF9 and AF5, coexpression of MLL-AF5(K717D) with AF9(WT) did not transform bone marrow progenitor cells, whereas coexpression of both mutants MLL-AF5(K717D) with AF9(D544R) resulted in increased colony formation through the third week of replating. Corresponding to this serial replating ability was an increase in Meis1 expression compared to the MLL-AF5(K717D) + AF9(WT) expressing cells. Hoxa9 expression again seems to be differentially regulated from Meis1 expression in this system. The hypothesis described above is applicable in this example as well, although I would not necessarily expect to see a decrease in Hoxa9 expression in cells expressing both mutants compared to expressing MLL-AF5(K717D) + AF9(WT); if proper
formation of the SEC is not required for elongation of *Hoxa9*, I would expect to see no changes in the levels.

A possible explanation for the decrease in *Hoxa9* expression in cells expressing both MLL-AF5(K717D) and AF9(D544R) compared to cells expressing MLL-AF5(K717D) and AF9(WT) may be a result of simply overexpressing two mutants that can bind to endogenous SEC proteins required for transcription at *Hoxa9*. Perhaps overexpression of both an AF5 mutant and an AF9 mutant act in a dominant negative manner, binding to endogenous proteins AF4/AF5, AF9/ENL, or DOT1L, and preventing formation of the SEC that may be required for the high levels of transcription occurring at the *Hoxa9* locus.

These results are perplexing, but the presence of the mutations was verified in the cDNA used in these samples. RNA was isolated from three independent experiments and the results were consistent. ChIP analysis for SEC components such as AF4/AF5, AF9/ENL and P-TEFb components CDK9 and Cyclin T1/2 at the *Meis1* and *Hoxa9* loci may shed light on any differential regulation between these two genes previously thought to play the similar roles in leukemogenesis.

These findings highlight the requirement for the AF9-AF4/AF5 salt bridge interaction in leukemic transformation. As previously described, AF9 interacts with a region of hydrophobic residues in AF4 (also conserved in AF5, Figure 32). This hydrophobic region is present in other AF9 interacting proteins DOT1L, CBX8 and BCoR. These results show that mutation of AF9 at aspartic acid D544 selectively disrupts binding to AF4, while maintaining binding to DOT1L and CBX8 (BCoR binding was not
assessed). These results show that the AHD of AF9 likely interacts with proteins through both the hydrophobic motif shared by all proteins, but also through additional interactions that may be specific to each partner protein.

AF4 K764 is located within the hydrophobic motif conserved in AF9 binding partners (Figure 32). While the charge-reversal studies support the conclusion that the electrostatic interaction between AF9 and AF4/AF5 is required for transformation, an alternative hypothesis is that mutation of K764E, or K717D in the case of AF5q31, disrupts the organization of the hydrophobic motif. Hydrophobic residues often associate in the core of proteins to remain isolated from the aqueous environment. Disrupting the charge on an amino acid within this motif may alter the packing altogether, losing the binding interface with AF9, and that could result in the loss of transformation. However the complementation experiment suggests that this is not the case, because coexpression of AF5(K717D) with AF9(D544R) restores transformation ability. If the hydrophobic motif was disrupted by the AF5(K717D) mutation, coexpression of AF9(D544R) would not be able to rescue the folding of that motif.

Disruption of AF9 binding to AF4 specifically by this mutation provides an attractive target to develop small molecule inhibitors. The Hemenway lab has designed and tested AF4 mimetic peptides that disrupt the AF4-AF9 interaction. These peptides show cytotoxicity specifically in MLL fusion-expressing cell lines, although the mechanism is not well understood\textsuperscript{157,189,259}.

It would be informative to examine whether the salt bridge interaction mediated by AF9 homolog ENL with AF4 family members is also necessary for leukemic
transformation. AF9 and ENL are highly homologous, and AF9 aspartate 544 is conserved in ENL (Figure 4D). As either AF9 or ENL can participate in the P-TEFb elongation complex, and both are fusion partners of MLL, it is likely that an electrostatic interaction between ENL and AF4 family members would also be required for leukemogenesis.

**MLL-AF9(D546R) Reduces Leukemic Transformation Potential**

The AF9 D546R mutant was predicted to specifically disrupt binding to DOT1L while maintaining binding to AF4 and CBX8. Coimmunoprecipitation studies confirmed that expression of AF9(376-568) D546R reduced binding to DOT1L compared to the wild type AF9 fragment, while maintaining interaction with AF4 (CBX8 binding studies are ongoing). Similar to the D544-K764 electrostatic interaction between AF9 and AF4 discussed above, AF9 D546 likely mediates an electrostatic interaction with a negatively charged residue in DOT1L.

To test the effect of this mutation on in vitro transforming ability, c-kit+ progenitor cells were retrovirally transduced with MLL-AF9, wild type or D546R mutant. Cells expressing MLL-AF9(D546R) showed decreased colony formation compared to the wild-type fusion, and the colonies that did form were very small and diffuse. The cells within these colonies were differentiated compared to the blast morphology in the MLL-AF9(WT)-expressing cells. Both *Hoxa9* and *Meis1* expression were significantly reduced in MLL-AF9(D546R) expressing cells compared to the wild type control, indicating that the interaction mediated by AF9 residue D546 is required for leukemic transformation *in vitro*.
Identification of the residue on DOT1L mediating the interaction with AF9 D546 would allow for similar complementation studies conducted with the D544R mutant to confirm that it is indeed a salt bridge interaction between these two proteins that, when disrupted, results in a loss of transforming potential.

The D546 residue is two amino acids downstream of the D544 residue that mediates AF4 interaction, yet these mutants seem to be selective for disrupting binding to DOT1L and AF4, respectively. This is entirely plausible as structural studies indicate that AF9 likely interacts with only one protein at a time. The AHD of AF9 is a small stretch of 94 amino acids, and studies of the chemical shift of AF9 with AF4, DOT1L, CBX8 or BCoR suggest that all four proteins bind to the same region of AF9 via a conserved hydrophobic motif. Biochemical purification studies have also shown that AF9 and DOT1L exist in a complex distinct from AF9 and AF4 and the other PTEF-b proteins, further supporting the hypothesis that AF9 interacts singly with one of the four identified proteins to mediate specific effects within the cell.

It would be very useful to examine the effects of a double AF9(D544R/D546R) mutant in both binding studies and in in vitro colony formation assays. While MLL-AF9(D544R)-expressing cells had reduced colony formation ability compared to wild-type, and expression of MLL-AF9(D546R) in bone marrow cells almost completely abolished colony formation, there were colonies at the end of each of four weeks, indicating that transformation was not completely inhibited in the cells. This may be because the D544R-mutant cells can retain binding to Dot1L, and the D546R-mutant cells
retain binding to Af4. The double mutant would presumably be unable to bind to both proteins, and thus we may see a complete loss of colony formation.

**MLL-AF9 Mutants E506R, I538A and K557E Do Not Alter Transformation Ability**

Expression of additional MLL-AF9 mutants E506R, I538A and K557E did not result in reduced colony formation compared to MLL-AF9(WT). There were no changes in target gene expression, nor any gross alterations in colony or cell morphology. Of note, however, is the MLL-AF9(I538A) mutation. This mutant was predicted to disrupt binding to both AF4 and DOT1L. We expected to see a loss of colony formation ability in bone marrow progenitors expressing this fusion, although results showed that expression of MLL-AF9(I538A) transforms cells to the same degree as MLL-AF9(WT). The colonies formed by MLL-AF9(I538A)-expressing cells, however, often had a dense core surrounded by more diffuse cells that themselves appeared partially differentiated compared to the wild type control.

Unpublished coimmunoprecipitation studies by Dr. Malik indicate that mutation of AF9 residue I538 to alanine results in a loss of binding to AF4 and DOT1L, but this mutant enhances binding to CBX8 compared to AF9(WT). If this is the case, it is possible that the enhanced AF9-CBX8 interaction is leading the recruitment of CBX8 to genes targeted by the MLL-AF9 fusion. CBX8 may then deposit repressive histone marks to turn off gene expression, which could explain the partially differentiated phenotype. In our studies, expression of MLL-AF9(I538A) does not result in a decrease in *Hoxa9* or *Meis1* expression in bone marrow cells at week one, but perhaps we would see a decrease if we examined later time points, or if we examined other MLL target genes. It would be
interesting to examine the effect of this mutant in vivo, as we would expect leukemia to develop based on the colony assay results, but the phenotype of the leukemia or the time of onset may be different compared to MLL-AF9(WT) control.

Transformation Potential of MLL-AF9 Fusions Resulting from Two Different Breakpoints in AF9

I also examined the transformation ability of MLL-AF9 fusions containing either 93 or 194 amino acids of AF9. To date, five breakpoint regions in AF9 have been identified in patients with t(9;11)(p22;q23) rearrangements. Of the reported cases in the literature, approximately 75% of MLL-AF9 fusions expressed in patients include amino acids 375-568 of AF9 in frame with 5’ MLL. These region was first identified by Nakamura and colleagues as site A. Found less frequently, in approximately 12% of patients with t(9;11), is a MLL fusion including amino acids 475-568 of AF9. A knock-in mouse model in which MLL-AF9(475-568) was introduced via homologous recombination into the endogenous Mll locus resulted in development of AML in mice, indicating that this shorter C-terminal region of AF9 is sufficient for transformation.

The AHD of AF9 that mediates protein-protein interactions is fully present in both the long and short form of the MLL-AF9 fusions. However, it was possible that the 100 amino acid difference between these two fusions could mediate additional interactions that would alter transformation ability. I tested this in vitro by expressing the long and short fusions in bone marrow cells, and assessing serial replating ability. I did not see a statistically significant difference in colony formation ability between the cells expressing either fusions until week four, at which point MLL-AF9(475-568)-expressing
cells formed more colonies than cells expressing the longer form. It would be interesting to examine whether the enhanced colony formation ability in cells expressing the short form fusion would continue for subsequent replatings. In regard to human leukemia, it would be interesting to examine the survival rates of patients harboring MLL-AF9 fusions resulting from the different AF9 breakpoint regions.

Neither *Hoxa9* nor *Meis1* expression levels were significantly different between cells expressing MLL-AF9 fusions with the long or short AF9 C-terminal tail at the end of week one. This result is consistent with the colony formation between cells expressing these constructs at the end of week one. However, we may expect to see increased *Hoxa9* and *Meis1* levels at week four, the time point in which the MLL-AF9(475-568)-expressing cells exhibited enhanced colony forming ability over MLL-AF9(376-568) expressing cells.

It would be interesting to examine the affinity of binding between the long and short form of the C-terminal AF9 tail with known AF9-interacting partners. If the shorter form is indeed more transforming, perhaps this region binds to certain “activating” proteins with higher affinity than does the longer form, or the longer form may bind to proteins with repressive function with higher affinity. In terms of the coupled folding of intrinsically disordered proteins, it may require less energy to induce structure of a 93 amino acid fragment as compared to 194 amino acids, making it easier for this smaller fragment to bind to proteins that interact with AF9 and confer an “activating” function.

Elucidating any differences in protein interactions between the long and short C-terminus of AF9 would be useful in understanding the mechanisms of transformation
mediated by these different MLL-AF9 fusion proteins, and could potentially inform different treatment options for patients expressing the long or short MLL-AF9 fusion.

**Transcription Elongation in Cells Expressing MLL-AF9 Mutants**

All of the point mutants were engineered to disrupt binding of AF9 to AF4, DOT1L, CBX8 or BCoR. We hypothesized that loss of AF9 interaction with one or more of these proteins would result in decreased colony formation and target gene expression, and indeed that was what we saw. The D544R mutation, which disrupts binding to AF4, and the D546R mutation, which disrupt binding to DOT1L, both resulted in decreased colony formation. To assess whether the interaction of AF9 with AF4 or DOT1L is required for productive elongation of the RNA Polymerase II complex, we designed primer sets spanning the length of the 138 kb Meis1 genomic region, hypothesizing that the SEC would be unable to assemble or function properly to promote productive elongation.

We saw reduced expression of *Meis1* at all four genomic regions in cells transduced with the D544R or D546R mutants compared to the wild type control, supporting our model that decreased colony formation and target gene expression is a result of disrupting required AF9-mediated protein interactions in the SEC. Interestingly, we saw increased expression levels at the 3’ end of the gene compared to the 5’ end in cells expressing either of these mutants. This could be a result of primer efficiency; perhaps the exon 6/7, exon 8/9 and exon 12 primer sets amplify their target more efficiently than the exon 1 primer set. The 5’ end of the transcript could also be degraded. If the RNA Pol II complex is stalled for too long, the RNA Polymerase II complex can be
ubiquitylated and targeted for degradation\textsuperscript{260}. Capping enzymes, as well as splicing factors and proteins responsible for 3’ polyadenylation often exert their functions cotranscriptionally with RNA Pol II\textsuperscript{261, 262}. A study in yeast showed that 5’ capping is an early event in transcriptional elongation\textsuperscript{263}. It could therefore be possible that RNA Polymerase II cannot successfully elongate target gene transcripts in cells expressing the MLL-AF9 D544R or D546R mutant. As a result, the nascent transcript is uncapped and is degraded from the 5’ end.

The MLL-AF9 E506R, I538A and K557E mutants also show decreased expression of \textit{Meis1} using the exon 1 primer set. These results suggest that all MLL-AF9 point mutants studied here exhibit defects in transcription initiation. It would be interesting to examine RNA Pol II occupancy throughout the Meis1 locus in cells transduced with these MLL-AF9 mutants. We could also assess the phosphorylation state of the CTD of RNA Pol II to discern whether Pol II is stalled at initiation or is in the productive elongation stage.

\textbf{Model for AF9 Interaction with AF4 and DOT1L in Leukemogenesis}

Overall, these studies have highlighted the requirement for AF9 interaction with Af4 and Dot1l in \textit{in vitro} models of leukemogenesis in the murine system. The use of structurally-informed point mutations identified critical interactions that could potentially be targeted by small molecule inhibitors to treat patients with MLL leukemia.

I used genetic mutation of AF9 in the context of the MLL-AF9 fusion protein to disrupt binding with Af4 or Dot1l. We hypothesize that disrupting the AF9-AF4 interaction prevents P-TEFb from being recruited and/or stabilized at MLL target gene
promoters, resulting in decreased expression of genes required for leukemogenesis (Figure 35). We do see a decrease in target gene expression, but it would be interesting to see if the SEC is being recruited to these loci upon loss of AF9-AF4 binding.

Figure 35. Model of AF9 Interaction with AF4 and DOT1L at MLL Target Genes

MLL fusions are recruited to MLL target genes via the N-terminal DNA binding activity of MLL. The C-terminal AF9/ENL portion of the fusion protein recruits SEC proteins AF4/AF5 which interact with P-TEFb to stabilize RNA polymerase. Downstream within the gene body, AF9 interacts with DOT1L, which deposits methyl marks at H3K79. As shown, mutations that disrupt AF9 interaction with AF4/AF5 or DOT1L result in decreased target gene expression. These protein-protein interactions are therefore attractive targets for developing therapeutics for the treatment of MLL leukemias.

Similarly, we see that loss of AF9-DOT1L interaction results in reduced colony formation, decreased target gene expression and differentiation of the leukemic blast cells. However, the mechanism is not understood. DOT1L is the H3K79 histone methyltransferase, so we hypothesize that loss of AF9-DOT1L interaction results in a loss of H3K79 methylation at genes targeted by MLL-AF9. ChIP-seq studies addressing this question are ongoing by our collaborators at the University of Virginia.
Furthermore, the function of H3K79 methylation is not well understood. This mark is often enriched in gene bodies and DOT1L has been identified in several studies as a member of the Super Elongation Complex, thus it is hypothesized the DOT1L and H3K79 methylation play a role in transcriptional elongation. This mark is associated with actively transcribed genes in genome-wide studies in Drosophila and mammalian cells. At least one study in yeast, and an additional ChIP-seq study using human CD4+ T cells, failed to find this correlation. In fact, H3K79me3 was associated with gene repression in these CD4+ T cells. It is clear from our studies and others, especially those utilizing Dot1l mice, that Dot1l is required for leukemic transformation by MLL fusions, although more work needs to be done to determine the mechanism.

MLL-rearranged leukemias are very complex, but the use of structurally-informed mutational analyses to dissect the mechanisms of leukemogenesis has proved to be a very helpful method in studying the function of these fusion proteins. It is our hope that this knowledge can be used in developing small molecule inhibitors and targeted therapeutics to disrupt interactions required for leukemic transformation.
REFERENCES


111. Xia, Z.B., Anderson, M., Diaz, M.O. & Zeleznik-Le, N.J. MLL repression domain interacts with histone deacetylases, the polycomb group proteins HPC2 and BMI-


VITA

Alyson Lokken was born in West Salem, WI on July 12th, 1986 to Ron and Diane Lokken as the youngest of three girls. She graduated from West Salem High School in 2004 and began her scientific career at the University of St. Thomas (St. Paul, Minnesota) that fall. Her first research experience was in a canoe examining the community dynamics of *Chrysemys picta*, the common painted turtle, in the lakes of Minnesota. She transitioned from ecology to molecular biology research in her final two years of college in the laboratory of Dr. Jennifer Cruise before graduating with a Bachelor of Science in Biology in 2008.

In August of 2008, Alyson joined the Program in Biochemistry and Molecular Biology at Loyola University Medical Center. Shortly thereafter, she joined the laboratory of Dr. Nancy J. Zeleznik-Le, where she studied the role of specific protein-protein interactions in hematopoiesis and leukemogenesis.

While at Loyola, Alyson was a student representative on the Graduate Student Council as well as co-President for a year. She also served as a mentor to several summer students and as a tutor for first year students in the Core curriculum.

After completing her Ph.D., Alyson will pursue a postdoctoral position in the field of epigenetics, with the goal of teaching biology at an undergraduate university.