Critical Functions Specified by the MLL CXXC Domain Determine Leukemogenic Capacity

Noah Warren Birch
Loyola University Chicago

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

Part of the Biochemistry Commons

Recommended Citation
https://ecommons.luc.edu/luc_diss/504

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 © 2013 Noah Warren Birch
LOYOLA UNIVERSITY CHICAGO

CRITICAL FUNCTIONS SPECIFIED BY THE MLL CXXC DOMAIN DETERMINE LEUKEMOGENIC CAPACITY

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

PROGRAM IN MOLECULAR AND CELLULAR BIOCHEMISTRY

BY
NOAH WARREN BIRCH
CHICAGO, ILLINOIS
MAY 2013
ACKNOWLEDGEMENTS

The work presented within this dissertation would not have been possible without the generous assistance of the many people who have supported my training. I would like to thank the Graduate School of the Health Sciences Division and the Stritch School of Medicine (S.S.O.M.) for the opportunity to participate in the MD/PhD program at Loyola University Chicago. I would like to thank the faculty, staff, and students of the Molecular & Cellular Biochemistry (MCB) graduate program for providing a supportive and collegial environment for basic science training. In particular, I would like to thank the core MCB faculty—Mary Manteuffel, Ph.D., Richard Schultz, Ph.D., William Simmons, Ph.D., Michael Collins, Ph.D., and Alan Frankfater, Ph.D.—for their thoughtful guidance and dedication to teaching at each stage of my training in addition to the administrative support provided by Ashyia Paul and Elayne Grzeda. I am thankful for the support of the current and former MD/PhD program administrators: Charles Hemenway M.D., Ph.D., Ameet Kini, M.D., Ph.D., Donna Buczek, and Ela Stekala.

I would like to thank my mentor, Nancy Zeleznik-Le, Ph.D., for welcoming me into her laboratory, for generously sharing her scientific knowledge and skills, and for her constant encouragement and enthusiasm. I am grateful for the supportive and productive training environment which Dr. Zeleznik-Le provided while serving as an inspiring role model for my future career in biomedical research. I would also like to thank the members of my dissertation committee—Manuel Diaz, M.D., Charles Hemenway, M.D.,
Ph.D., Richard Schultz, Ph.D., William Simmons, Ph.D., and John Bushweller, Ph.D.—
for strengthening my scientific and academic work through prudent guidance and
constructive feedback. I would like to thank the current and former members of the
Zeleznik-Le laboratory including Laurie Risner, Ph.D., Nicholas Achille, Zhen-Biao Xia,
Ph.D., Yousaf Mian, Alyson Lokken, Shubin Zhang, M.D., Relja Popovic, Ph.D., Sonia
Brockway, and Adam Marek for teaching me new techniques, for troubleshooting help,
and for creating an enjoyable environment in which to work and train.

I would like to thank the laboratories and staff of the Hematologic Malignancies
Program and Oncology Institute, especially the members of the Hemenway, Diaz, Zhang,
Dingwall, and Bocchetta laboratories, Jiwang Zhang, M.D., Ph.D., Ute Osmers, Ph.D.,
Lorelei Hacholski, M.S.W., and Lynn Walter for sharing equipment, resources, and ideas
and for providing valuable administrative assistance. I would like to acknowledge our
valued collaborators, John Bushweller, Ph.D., and Aravinda Kuntimaddi (University of
Virginia) and Ming-Ming Zhou, Ph.D., and Shiraz Mujtaba, Ph.D. (Mount Sinai School
of Medicine) for their productive interactions and thank Kevin FitzGerald, S.J., Ph.D.
(Georgetown University) for providing the full length MLL2 construct used in this work.

I would like to recognize the funding sources which supported this dissertation
work including the National Heart, Lung, and Blood Institute of the National Institutes of
Health, the American Society of Hematology, and the S.S.O.M. MD/PhD program.

Finally, I would like to thank my family and friends who have inspired this
exciting journey, supported each new endeavor with patience and encouragement, and
taught me to appreciate that “further study is needed” for it is only when we recognize
where our current understanding ends that the opportunity for progress truly begins.
For my parents
There is an important difference between the merely ‘experimental’ and a genuine experiment. The one may be a feeling for novelty. The other is rationally based upon experience seeking a better way.

—Frank Lloyd Wright
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii  
LIST OF TABLES ix  
LIST OF FIGURES x  
ABSTRACT xii  

CHAPTER 1: INTRODUCTION 1  

CHAPTER 2: LITERATURE REVIEW 6  
The Mixed Lineage Leukemia (MLL) gene and leukemia 6  
Hematopoiesis 8  
The hematopoietic stem cell niche 10  
HOX cluster genes and MLL 13  
The role of MLL in transcriptional regulation 17  
MLL2 is the closest homolog of MLL 23  
The MLL CXXC Domain 26  

CHAPTER 3: METHODS 34  
Cloning of MLL/MLL2-AF9 fusion constructs for *in vitro* colony assay 34  
Generation of and cloning of mutant MLL2 CXXC domain constructs 40  
Production of Retrovirus 43  
Determination of Retrovirus Titer 44  
Isolation of murine bone marrow cells 46  
CD117 positive selection of murine bone marrow progenitor cells 46  
Retroviral transduction of c-kit positive marrow progenitor cells 48  
Methylcellulose Colony Assay 49  
Cytospin for cell morphology 51  
RNA Isolation 52  
Generation of cDNA 52  
Quantitative RT-PCR 53  
Generation of GST fusion repression domain constructs 54  
Bacterial over-expression of GST fusion MLL- and MLL1/2 repression domain constructs 55  
Purification of GST fusion repression domain constructs 56  
*In vitro* transcription and translation 57  
*In vitro* protein binding assay 58  
MTT Assay 60  
Statistical Analysis 60
CHAPTER 4: RESULTS

Aim 1: Determine how differences in the MLL CXXC domain and the MLL2 CXXC domain contribute to different transformation capacities of MLL/MLL2-AF9 fusion proteins

Comparison of co-repressor and co-activator protein interactions with MLL and MLL1/2 repression domains

Determination of residues within the CXXC domain that alter co-repressor binding

Comparison of co-activator PAF1 binding to residues of the MLL and MLL2 CXXC domains

Functional comparison of the capacity of MLL and MLL2 CXXC domain residues to transform bone marrow progenitor cells in vitro

Determination of titers for retroviruses expressing the fusion proteins MLL residues introduced into the MLL/MLL2-AF9 fusion protein showed variable ability to rescue transformation capacity in an in vitro colony assay

MLL residues showed variable ability to rescue increased Hoxa9 expression when introduced into the MLL/MLL2-AF9 fusion protein in an in vitro colony assay

Aim 2: Determine the role of Cysteine 1188 in the regulation of the MLL CXXC domain binding to non-methylated CpG DNA

Treatment of MLL-AF9 and MLL(C1188A)-AF9 cells with agents that can modify thiol groups

Growth of MLL-AF9 and MLL(C1188A)-AF9 cells under conditions of varying oxygen concentration

CHAPTER 5: DISCUSSION

REFERENCES

VITA
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Occurrence of common genetic abnormalities involving <em>MLL</em></td>
<td>7</td>
</tr>
<tr>
<td>2. Cloned DNA constructs showing restriction enzyme cutting sites and priming sites which were used for subcloning and sequencing</td>
<td>35</td>
</tr>
<tr>
<td>3. MLL and MLL2 cloning and sequencing primers</td>
<td>36</td>
</tr>
<tr>
<td>4. Site-directed mutagenesis primers used to introduce individual MLL residues into the MLL2 CXXC domain</td>
<td>41</td>
</tr>
<tr>
<td>5. Rat1a virus titers of wild type and mutant MLL/MLL2-AF9 retroviral fusion constructs</td>
<td>72</td>
</tr>
<tr>
<td>6. DNA binding affinities of isolated wild type and mutant CXXC domains</td>
<td>94</td>
</tr>
<tr>
<td>7. Summary of functional <em>in vitro</em> transformation and DNA-binding data</td>
<td>101</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure                                                                                      Page

1. Hematopoiesis diagram                                                                    9

2. The HOX genes are an evolutionarily conserved set of genes involved in development and hematopoiesis 14

3. Structural schematic of MLL, a MLL fusion, and MLL2 proteins                              18

4. The MLL CXXC domain is conserved across species, and similar domains are found within other proteins 27

5. Comparison of the MLL and MLL2 CXXC domains                                              28

6. Comparison of amino acid differences within the MLL and MLL2 CXXC domains to be studied in Aim 1. 32

7. HDAC1 binds to both the MLL repression domain and the equivalent MLL1/2 repression domain   62

8. HPC2 binds to both the MLL repression domain and the equivalent MLL1/2 repression domain   63

9. CtBP binds to both the MLL repression domain and the equivalent MLL1/2 repression domain   64

10. PAF1 binds to both the MLL repression domain and the equivalent MLL1/2 repression domain  67

11. The MLL2 CXXC domain can transform primary bone marrow cells when swapped into a MLL-AF9 fusion protein 69

12. Experimental design of in vitro colony assay                                              71

13. Rat1a virus titers of MLL/MLL2-AF9 retroviral fusion constructs                           73

14. In vitro self-renewal and proliferation capacity of MLL and MLL2 residues in the context of an AF9 fusion oncoprotein 75
15.  Round 3 colony and cell morphology  76
16.  *Hoxa9* expression from Round 1 colony assay cells  80
17.  Effect of thiol modifying agents on proliferation of MLL-AF9 versus MLL(C1188A)-AF9 transformed cells  83
18.  Effect of thiol modifying agents on self-renewal and proliferation of MLL-AF9 versus MLL(C1188A)-AF9 transformed cells  86
19.  Effects of varying oxygen concentration on colony forming ability of MLL-AF9 versus MLL(C1188A)-AF9 transformed cells  89
20.  Structure-function relationship between residues of the CXXC domain  93
21.  Models of the MLL and MLL2 CXXC domains showing differential DNA binding and equivalent binding to cofactor proteins  102
ABSTRACT

The Mixed Lineage Leukemia (MLL) gene can participate in chromosomal translocations which generate a fusion protein leading to aggressive, acute leukemia. These translocations may occur spontaneously or following previous treatments with particular chemotherapy agents. Compared to other types of acute leukemia, MLL-associated leukemias have a poor prognosis. A better understanding of how MLL fusion proteins contribute to leukemia is necessary in order to develop more effective treatments.

MLL fusion proteins retain the amino-terminal portion of MLL including the CXXC DNA-binding domain while the carboxy-terminal portion of the oncoprotein is comprised of the fusion partner. The MLL CXXC domain is a zinc-coordinating structure which binds to non-methylated DNA within the promoter region of target genes. The CXXC domain is located within a functional repression domain which is known to interact with cofactor proteins that modulate target gene expression. In my dissertation project, I investigated the functional role of amino acids within the CXXC domain to determine how specific residues contribute to the leukemogenic capacity of MLL fusion proteins.

The closest homolog of MLL, MLL2 (alternatively named MLL4), also contains a similar CXXC domain, yet an artificial MLL2 fusion protein is unable to transform cells in vitro. Knowing that the CXXC domain is a critical component of MLL fusion proteins...
proteins, I hypothesized that specific amino acid differences between MLL and MLL2 CXXC domains account for differences in leukemogenic capacity. To test this hypothesis, the MLL2 CXXC domain was cloned into the context of the well-studied MLL-AF9 fusion protein to generate an artificial MLL/MLL2-AF9 chimera. Amino acid substitutions within the MLL2 CXXC domain were then introduced to restore residues to the MLL amino acid sequence within this synthetic chimera. Retroviral transduction of bone marrow progenitor cells with the wild type or mutant chimeric MLL/MLL2-AF9 fusions was performed to measure proliferation capacity in a methylcellulose colony assay. Colony forming ability and cell proliferation were then measured along with observations in cellular morphology and target gene expression to determine the functional contributions of specific residues within the CXXC domain of the fusion proteins. *In vitro* binding studies were performed between MLL or synthetic MLL/MLL2 repression domain protein fragments and cofactor proteins to determine if there were differences in protein interactions which could account for the observed functional differences. By comparing the residues of the MLL and MLL2 CXXC domains, critical amino acids were identified on both the DNA-binding surface and on the opposite, non-DNA-contact surface of the CXXC domain that contribute to leukemogenic transformation.

Cysteine 1188 of the MLL CXXC domain is the only non-zinc-coordinating cysteine residue within the CXXC domain. This residue is critically positioned on the DNA-binding surface, and our laboratory has previously shown that mutation of this residue can disrupt the ability of the CXXC domain to bind to DNA. With a critical location on the DNA-binding surface and potential susceptibility to post-translational
modification, I hypothesized that the side chain of Cys1188 may be physiologically altered to regulate DNA-binding affinity. Transformed MLL-AF9 progenitor cells were treated with modifying agents or grown under conditions of varying oxygen concentration, and cell viability and colony forming ability were measured. The MLL-AF9 cells showed a modest increase in susceptibility to parthenolide treatment in an in vitro colony assay compared to control cells suggesting that Cys1188 may be altered through redox regulation. Growth under conditions of varying oxygen levels, however, showed no significant differences in proliferation of wild type or control cells. These results build on the initial studies of Cys1188 and the MLL CXXC domain conducted in our lab and provide valuable direction for future investigations which may eventually allow for therapeutic targeting of the CXXC domain in these aggressive forms of MLL-associated leukemia.
CHAPTER 1
INTRODUCTION

The Mixed Lineage Leukemia (MLL) gene is involved in translocations with more than 60 possible fusion partners leading to aggressive, acute leukemias with poor prognosis. The resulting chimeric fusion protein retains the N-terminal portion of MLL including the CXXC DNA-binding domain, while the C-terminal portion of the oncoprotein is comprised of the fusion partner\(^1,2\). The CXXC domain is found within a functionally-defined repression domain and binds to non-methylated CpG DNA\(^3,5\). This CXXC DNA-binding domain contains a clustering of eight cysteines which coordinate two zinc ions and is essential for the transformation capacity of cells expressing an MLL fusion protein\(^4\). The closest homolog of MLL, MLL2 (alternatively named MLL4), also contains a similar CXXC domain which preferentially binds non-methylated CpG DNA but with lower affinity than the MLL CXXC domain (Bushweller laboratory, University of Virginia, unpublished)\(^6,7\). Interestingly, cells expressing an artificial MLL2 fusion protein do not transform primary bone marrow cells and require the substitution of the MLL CXXC domain plus the ‘post-CXXC’ residues—a region comprising the majority of the MLL repression domain—in order to become immortalized \textit{in vitro}\(^7\). I hypothesized that amino acid differences between the CXXC domain or the repression domain of MLL and MLL2 must contribute to these functional differences. Specific amino acids along the DNA contact interface may play a role in differential DNA-
binding affinity. In addition, our lab has previously shown that the repression domain of MLL (containing the CXXC domain) binds to class I histone deacetylases (HDACs 1 and 2), polycomb group proteins (HPC2 and BMI-1), and the co-repressor C-terminal-binding protein (CtBP)\(^8\). Recently, a co-activator protein, PAF1, was also shown to bind to this same region of MLL\(^9, 10\). Differential binding of these cofactor proteins to MLL and MLL2 may affect overall function but might also contribute to CXXC DNA-binding affinity. Cysteine 1188, the only non-zinc-coordinating cysteine residue within the CXXC domain, is critically positioned on the DNA-binding surface and adjacent to one of the conserved cysteines involved in zinc coordination\(^11\). Previous studies conducted in our lab showed that mutation of Cys1188 to aspartate completely abrogated the ability of the CXXC domain to bind DNA and prevented both in vitro immortalization and development of leukemia in mice when introduced in the context of a MLL fusion protein\(^11\). In contrast, mutation of Cys1188 to alanine retained MLL’s ability to bind CpG DNA, as well as to immortalize in vitro and in vivo in the context of the fusion protein\(^11\). With a critical location on the DNA-binding surface of the CXXC domain and susceptibility to redox regulation, I hypothesized that the side chain of Cys1188 may be physiologically altered to regulate DNA-binding affinity, allowing this residue to function as a molecular switch that regulates the binding of MLL to its non-methylated CpG DNA targets.

**Hypothesis:** Specific amino acid residues of the MLL and MLL2 CXXC domains function to alter DNA binding, cofactor binding, and susceptibility to post-translational modification—critical features of MLL that determine leukemogenic capacity.
Aim 1: Determine how differences in the MLL CXXC domain and the MLL2 CXXC domain contribute to the different transformation capacities of MLL/MLL2-AF9 fusion proteins.

MLL and MLL2 are paralogs with high sequence homology. Both proteins contain similar functional domains including the CXXC DNA-binding domain with eight conserved Zn-coordinating cysteine residues. With such similarity, it is intriguing that MLL2 is unable to compensate for the embryonic lethality observed in Mll -/- mice\(^{12}\). Moreover, MLL2 has not been associated with translocations causing leukemia like MLL. The CXXC domain and the ‘post-CXXC’ residues of the larger repression domain (residues 1149-1337 of MLL) have been shown to account for the differences in leukemogenic capabilities of these proteins\(^{7}\). However, the specific amino acid residues within the repression domain which account for this difference have not yet been identified. The repression domain of MLL has previously been shown to bind to cofactor proteins but it is not currently known if these same co-repressors bind to MLL2\(^{8-10}\). I hypothesized that specific amino acid differences between MLL and MLL2 contribute to variable DNA-binding affinities of the CXXC domains and to variable recruitment of cofactor proteins which regulate transcription of MLL/MLL2 target genes. The first aim of my dissertation proposed to explore how these specific amino acid differences affect the critical functioning of the CXXC domain and alter the leukemogenic capacity of these translocations.
Aim 2: Determine the role of Cysteine 1188 in the regulation of the MLL CXXC domain binding to non-methylated CpG DNA.

Cysteine 1188 of the MLL CXXC domain is the only non-zinc-coordinating cysteine residue within the CXXC domain and is highly conserved across vertebrate orthologs of MLL. Our lab showed that this residue is critically positioned on the DNA-binding surface and adjacent to one of the conserved cysteines involved in zinc coordination\textsuperscript{11}. Mutation of Cys1188 to aspartate completely abrogated the ability of the CXXC domain to bind DNA even though it is not involved in direct interactions with DNA\textsuperscript{11}. Zinc finger motifs containing a zinc ion tetrahedrally coordinated to cysteine or histidine groups are relatively common within DNA-binding proteins and can be subject to redox regulation at free cysteine residues to alter protein function\textsuperscript{13}. The hypoxic niche (in which hematopoiesis takes place and leukemia stem cells can reside) may protect Cys1188 of the MLL CXXC domain from oxidation. With a critical location on the DNA-binding surface of this domain and susceptibility to post-translational modification, I hypothesized that the side chain of Cys1188 may be physiologically altered to regulate DNA-binding affinity, allowing this residue to function as a molecular switch that regulates the binding of MLL to its non-methylated CpG DNA targets.

It is not well understood how the balance of gene activation and repression is functionally regulated during normal hematopoiesis or MLL-associated leukemia. The CXXC DNA-binding domain and the ‘post-CXXC’ residues of the MLL repression domain have been shown to be essential for the transformation capacity of MLL fusion proteins\textsuperscript{2,7,11}. By studying specific amino acid residues of the MLL CXXC domain, the goal of my dissertation was to better understand functional features that contribute to
MLL’s ability to bind to DNA and to cofactor proteins and the potential post-translational modifications which we hypothesized are critical determination features for MLL-associated leukemia. My work serves as a starting point for therapeutic targeting of the MLL CXXC domain to alter leukemogenic capacity and, ultimately, lead to better clinical outcomes for these devastating forms of acute leukemia.
CHAPTER 2
LITERATURE REVIEW

The Mixed Lineage Leukemia (MLL) gene and leukemia

The Mixed Lineage Leukemia (MLL) gene was first identified for its involvement in translocations at the 11q23 locus associated with acute leukemia\(^1\). The MLL gene (also called MLL1, HRX, HTRX, or ALL-1) derives its name from the clinical presentation associated with its cytogenetic aberration: patients with these chromosomal rearrangements develop the pathological phenotype of acute lymphoid leukemia (ALL), acute myeloid leukemia (AML), or a mixed lineage leukemia\(^15\)-\(^18\). MLL genetic abnormalities account for 5-10% of ALL or AML cases\(^19\),\(^20\). The incidence of MLL-associated leukemia is significantly higher, though, in two distinct populations of leukemia cases: de novo infant leukemias and secondary leukemias which develop following chemotherapy with high doses of DNA topoisomerase II inhibitors\(^21\),\(^22\). Furthermore, MLL leukemias are particularly aggressive compared to other types of leukemia and are associated with relatively dismal outcomes with five-year survival rates of only 20-30% in infants\(^20\),\(^23\)-\(^25\).

The most common genetic abnormality associated with the MLL gene is an in-frame, chromosomal translocation which results in a chimeric protein expressing the amino-terminal end of the MLL protein fused to the carboxy-terminus of a partner protein\(^26\)-\(^28\). To date, over sixty different fusion partners have been identified in MLL
leukemia cases\textsuperscript{28}. The type of MLL-associated leukemia that develops depends in part on the particular fusion partner involved in the MLL oncoprotein. Table 1 shows the incidence of the most common fusion partners and the associated leukemia phenotype. Internal duplication of various exons of \textit{MLL} can also lead to the development of leukemia\textsuperscript{29-31}. In these instances, a separate partner protein is not fused to the oncoprotein; instead, specific functional domains of the MLL protein, including the CXXC domain, are duplicated to alter normal function. These genetic aberrations of the \textit{MLL} gene disrupt normal hematopoiesis causing the bone marrow to become overwhelmed with the production of immature leukemic cells. These rapidly proliferating cells crowd out the production of normal blood cells leading to anemia, bleeding disorders, and potentially to an increased susceptibility to infection. Eventually, this uncontrolled production of undifferentiated blast cells develops into end-organ failure and, ultimately, death in patients if the aggressive leukemia is not effectively treated.

<table>
<thead>
<tr>
<th>Genetic Alteration</th>
<th>Resultant Oncoprotein</th>
<th>Incidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>t(9;11)(p22;q23)</td>
<td>MLL-AF9/MLLT3</td>
<td>36%</td>
</tr>
<tr>
<td>t(10;11)(p12q23)</td>
<td>MLL-AF10/MLLT10</td>
<td>13%</td>
</tr>
<tr>
<td>PTD</td>
<td>(amplification of internal domains within MLL)</td>
<td>10%</td>
</tr>
<tr>
<td>t(6;11)(q27;q23)</td>
<td>MLL-AF6/MLLT4</td>
<td>9%</td>
</tr>
<tr>
<td>t(11;19)(q23;p13.1)</td>
<td>MLL-ELL</td>
<td>8%</td>
</tr>
<tr>
<td>t(11;19)(q23;p13.3)</td>
<td>MLL-ENL/MLLT1</td>
<td>8%</td>
</tr>
<tr>
<td>Other</td>
<td>i.e. MLL-AF1q/MLLT11, MLL-AF17/MLLT6</td>
<td>15%</td>
</tr>
<tr>
<td>t(4;11)(q21;q23)</td>
<td>MLL-AF4/ASF1</td>
<td>56%</td>
</tr>
<tr>
<td>t(11;19)(q23;p13.3)</td>
<td>MLL-ENL/MLLT1</td>
<td>16%</td>
</tr>
<tr>
<td>t(9;11)(p22;q23)</td>
<td>MLL-AF9/MLLT3</td>
<td>10%</td>
</tr>
<tr>
<td>t(10;11)(p12q23)</td>
<td>MLL-AF10/MLLT10</td>
<td>3%</td>
</tr>
<tr>
<td>PTD</td>
<td>(amplification of internal domains within MLL)</td>
<td>&lt; 1%</td>
</tr>
<tr>
<td>Other</td>
<td>i.e. MLL-AF6/MLLT4, MLL-EPS15/AF1P</td>
<td>15%</td>
</tr>
</tbody>
</table>

PTD: Partial Tandem Duplication
Hematopoiesis

MLL is involved in the normal process of hematopoiesis through the regulation of specific target genes (i.e. Hox cluster genes discussed later) which can also give rise to forms of acute leukemia when the expression of these target genes becomes misregulated. Hematopoiesis is the process by which the mature cells of the blood system develop from a series of immature, precursor cells (Figure 1). The hematopoietic stem cell (HSC) is the precursor cell from which all subsequent hematopoietic cells originate. HSCs are distinguished by the unique properties of all adult stem cells: long-term self-renewal which allows for the lifelong repopulation of the blood cells and pluripotency which enables HSCs to give rise to all cell types of the blood system including myeloid, lymphoid, and erythroid/megakaryocyte lineages. HSCs are identified by the presence or absence of specific cell surface markers (for mouse: c-kit⁺, Sca1⁺, Lin⁻, Flk2⁻, CD34⁻, and Slamf1⁺; and for human: CD34⁺, Lin⁻, CD38⁻, CD90⁺, and CD45RA⁻)³⁴.

When stimulated to divide, an HSC can undergo asymmetrical division resulting in two different cell types: one cell exhibiting the characteristics of an HSC that maintains the existing HSC pool and one daughter cell that can undergo maturation into a multipotent HSC with limited short-term repopulating capacity. This short-term HSC then gives rise to common erythroid, myeloid, and/or lymphoid progenitor cells that express a unique set of cell surface markers which distinguish the various cell lineages. Eventually, these progenitor cells decrease the expression of immature cell-surface markers (i.e. c-kit⁺) and differentiate into erythroid (red blood cells or platelet-producing megakaryocytes), myeloid (neutrophils, monocytes/macrophages, eosinophils, and basophil cells), or lymphoid (natural killer cells and B or T lymphocytes) cells. The
The process by which undifferentiated hematopoietic stem cells (HSCs) give rise to all types of fully-differentiated blood cells. (Cells of the erythroid lineage—red blood cells and platelet-producing megakaryocytes—have been omitted from this diagram for simplicity.)
differentiated blood cells are released from the bone marrow to circulate throughout the body or to travel to distant lymphoid tissues like the spleen, thymus, and lymph nodes where they function as mature cells or (in the case of lymphocytes) undergo further maturation.

The hematopoietic stem cell niche

The environment in which the blood cells develop influences hematopoiesis. Extracellular factors from this environment can stimulate hematopoietic cells to proliferate, differentiate, or mobilize through various signaling mechanisms. During embryonic development, hematopoiesis takes place in different locations as the fetus generates new tissues capable of cultivating the blood cells. Initially, hematopoiesis takes place in blood islands of the yolk sac surrounding the developing embryo before moving to the aortic-gonadal-mesonephric region. Once the tissues of the liver have developed, blood cell development moves to the fetal liver and finally to the bone marrow where hematopoiesis continues throughout adulthood\(^\text{35}\). In each of these locations, the maturing hematopoietic cells require a particular set of conditions that make up a niche called the HSC niche.

The primary location for hematopoiesis in newborns and adults is the bone marrow found within the medullary cavities of long bones and the cancellous tissues of the skull, pelvis, and vertebrae. The marrow has a distinct architecture consisting of a fibrous network and stromal cells which support the hematopoietic cells. The HSCs are found deep within the marrow cavities in the periendosteal region where they are in contact with osteoblasts which line the bone surfaces. Contacts made between these cells regulate signaling through N-cadherin and Notch1 and are critical for the retention of
HSCs in the niche and the maintenance of their pluripotent potential.\textsuperscript{36-38} Interspersed throughout the marrow are supportive stromal cells. These cells secrete soluble factors like G-CSF, GM-CSF, SCF (also known as kit ligand), IL-6, and CXCL12 which are known to regulate proliferation, differentiation, and mobilization of hematopoietic cells.\textsuperscript{39-42} Nitric oxide (NO) from stromal cells has also been shown to modulate hematopoiesis and the mobilization of progenitor cells.\textsuperscript{43-45}

Other factors like oxygen availability affect the cells within the HSC niche, and I hypothesized that oxygen levels may also affect MLL function during leukemogenesis.\textsuperscript{46, 47} Sinusoidal capillaries weave throughout the marrow to supply oxygen and nutrients to the niche and to remove metabolic byproducts. The partial pressure of oxygen from gross marrow aspirates from superior iliac crests of healthy donors has been measured to be 55 mm Hg in one report (or about 7\% compared to 21\% O\textsubscript{2} in typical ambient air of cell-culture incubators).\textsuperscript{48} Given the three-dimensional architecture and cellular complexity of the bone marrow, however, it is reasonable to envision microenvironments within the marrow in which there are variable levels of oxygen. The stroma surrounding the sinusoidal capillaries form layers in which hematopoietic cells found further away from the capillary would be exposed to lower levels of oxygen as the oxygen diffuses away from the vessel and is depleted by other intervening, metabolically active cells. In fact, the oxygen levels in bone marrow have been mathematically modeled to demonstrate that a range of oxygen pressures likely exists in the marrow given the presence of intervening acellular fibers, single cells, or colonies of proliferating cells that separate hematopoietic cells from the sinusoidal capillaries.\textsuperscript{49, 50} These models suggest that the less mature, progenitor cells are positioned in the regions with lowest oxygen
tensions (<1% O$_2$) while the more differentiated cells would be found closer to the capillaries with higher oxygen levels (6-7% O$_2$).

This association between regional oxygen levels in the bone marrow and the distribution of hematopoietic cells has since been determined in vivo using mice that have been administered fluorescent Hoechst 33342 (Ho) dye to measure diffusion within the marrow cavities$^{51}$. The authors found HSCs to be enriched in the regions with the lowest perfusion of Ho dye suggesting that oxygen would be less available to these cells. In fact, treatment with the hypoxia marker, pimonidazole, showed selective targeting of progenitor cells which were residing in the regions of least perfusion compared to the more mature hematopoietic cells within areas of greater perfusion. These observations suggest that the hematopoietic cells are distributed in an organized manner within the marrow which corresponds to the oxygen availability within the HSC microenvironment. Subsequent studies in vivo have also confirmed that the hematopoietic precursor cells lie in regions with low oxygen tension while cells that are more actively proliferating in the presence of growth factors are positioned near the capillary sinuses of the marrow$^{37,52}$. The effect of low oxygen tension on hematopoietic cells has been measured in vitro. Progenitor cells grown in culture under hypoxic conditions (1% O$_2$) have shown increased proliferation potential and increased colony forming ability$^{53-55}$. These findings suggest that a hypoxic bone marrow environment may be an advantageous environment for undifferentiated hematopoietic cells where they can maintain a self-renewing pool of stem cells$^{56}$.

The hypoxic niche, however, has also been shown to benefit leukemic stem cells which rely on similar niche properties for continuous proliferation and evasion of
therapeutic treatments. With a conserved cysteine residue within the critical CXXC DNA-binding domain (described in detail below), I hypothesized that the hypoxic niche would benefit leukemia cells expressing an MLL fusion protein. Under conditions of hypoxia, this cysteine—with a sulfhydryl side chain susceptible to post-translational modification—could be more readily kept in a reduced state which would allow the fusion protein to bind to its DNA targets and subsequently maintain gene expression and the undifferentiated cell state. In contrast, exposure to higher levels of oxygen could potentially lead to oxidation of the thiol side chain that would disrupt MLL’s ability to bind to DNA causing the cells to decrease proliferation and undergo differentiation.

**HOX cluster genes and MLL**

During hematopoiesis, a specific set of gene regulatory events occur that allow a cell to progress from a pluripotent HSC to a fully-differentiated blood cell of a particular lineage. One set of highly regulated genes called the HOX cluster genes were first implicated in the process of hematopoiesis when progenitor cells were found to display unique expression patterns of subsets of HOX genes: the expression of HOX genes was decreased when cells underwent differentiation toward erythroid or myeloid lineages. The name of the HOX genes is derived from a conserved DNA-binding domain (the homeobox domain) found in homeotic proteins which were first implicated in *Drosophila* development. Groups of multiple HOX genes form gene clusters that have been found to be evolutionarily conserved in metazoans from the fruit fly to chickens and humans suggesting that this set of homeobox-containing genes plays a critical role during development. Subsequent studies in mice and humans confirmed the developmental role of HOX genes (*Hox* in mice) in embryonic patterning where differential expression
Figure 2. The HOX genes are an evolutionarily conserved set of genes involved in development and hematopoiesis.

The HOX cluster genes are conserved throughout evolution from fruit flies (*Drosophila melanogastor*) to humans and are organized in a pattern which corresponds to their role in the determination of segmental identity during development. (This figure was reproduced from the book *Your Inner Fish* by Neil Shubin with permission from the illustrator, Kalliopi Monoyios, University of Chicago).\(^7\)
of various HOX family members in spinal tissue corresponds to the determination of segmental identity\textsuperscript{66-68}. Differential expression of these homeobox-containing genes from flies to humans helps to establish polarity in the embryo so that an anterior-posterior orientation can develop along the central axis\textsuperscript{69, 70}. The spatial orientation of these genes closely corresponds with their role in determining segmental identity, and this gene structure has been evolutionarily conserved in eukaryotes from invertebrates to vertebrates (Figure 2)\textsuperscript{71-73}. In contrast to invertebrates which have homebox genes arranged linearly in a single cluster, the 39 HOX genes in humans are organized in four distinct clusters—HOXA, HOXB, HOXC, and HOXD—each containing 9-11 HOX genes. This clustered arrangement allows each subset of genes to be under the same transcriptional regulatory control for coordinated, temporospatial expression\textsuperscript{74, 75}.

A critical role for MLL in Hox cluster gene regulation was demonstrated in heterozygous Mll +/- and null Mll -/- mice\textsuperscript{12}. Mll -/- mice are embryonic lethal by day E10.5 while the heterozygous mutants survive through birth with defects in axial skeletal formation associated with the misexpression of Hoxa7 and Hoxc8 compared to wild type Mll +/- mice. These defects demonstrate haploinsufficiency (Mll +/- mice with loss of a single allele exhibited phenotypic defects) and indicate an important role for MLL in the regulation of Hox genes during development. Furthermore, heterozygous Mll +/- mice showed decreased levels of erythrocytes, hemoglobin, and platelets suggesting the loss of Mll also impairs hematopoiesis\textsuperscript{12}. It was subsequently shown that loss of both Mll alleles prevents the sustained expression of Hox genes leading to the observed developmental defects and impaired embryonic hematopoiesis in both the yolk sac and fetal liver\textsuperscript{76-78}. More recent findings using a loss-of-function mutant have confirmed the role of MLL in
adult hematopoiesis in which postnatal mice are unable to sustain normal blood cell development beyond three weeks following targeted gene excision. Taken together, these findings show that MLL plays an important role in maintaining the proper temporospatial regulation of Hox cluster genes that is required during both embryonic development and during fetal and adult hematopoiesis.

The HOXA9 gene is a well-studied HOX cluster gene with regard to its role in hematopoiesis and MLL leukemia. The initial association between HOXA cluster genes and hematopoiesis was suggested from observations that undifferentiated progenitor cells show higher HOXA cluster gene expression levels than more mature hematopoietic cells. Subsequent studies demonstrated a role for HOXA9 in the transformation of myeloid cells through a t(7;11)(p15;p15) translocation which generates a leukemogenic NUP98-HOXA9 fusion transcript and through the cooperation of MEIS1 with HOXA9. Together Hoxa9 and Meis1 function as a heterodimeric transcription factor to bind to DNA targets through the homeobox domain. The functional significance of this cooperativity has been verified in vivo through the transplantation of primary bone marrow cells that overexpress both Hoxa9 and Meis1. These cells lead to overt leukemia in recipient mice within three months. The role of Meis1 in MLL associated leukemia was further verified by knocking down Meis1 in MLL-AF9 transformed hematopoietic cells resulting in delayed leukemia progression in vivo.

The transformation of primary hematopoietic cells by MLL fusion proteins was shown to be mediated through upregulation of Hoxa cluster genes which subsequently cooperate with Meis1. Specifically, a MLL oncoprotein is required for maintaining the expression of Hoxa7 and Hoxa9 in primary bone marrow cells to sustain immortalization.
in vitro. This critical role of MLL in HOX gene regulation was further demonstrated during embryoid body differentiation in which embryonic stem cells are stimulated in vitro to undergo coordinated differentiation to form hematopoietic progenitor cells followed by mature hematopoietic cells. Mll null progenitor cells show marked reduction in Hoxa, Hoxb, and Hoxc cluster genes compared to Mll heterozygous cells during embryoid body differentiation\(^88\). These Mll deficient embryoid body cells are capable of proliferating like the heterozygous counterparts but show reduced colony forming ability in vitro suggesting that positive regulation of Hox genes by Mll is necessary during the early stages of hematopoiesis in order to maintain the progenitor population from which the mature blood cells arise.

**The role of MLL in transcriptional regulation**

Regulation of HOX gene expression is multifaceted involving large transcriptional complexes which can bind to multiple promoters within each cluster of genes, alternative splicing of the various gene transcripts, and regulation by microRNAs and long non-coding transcripts\(^75,89,90\). The detailed mechanism by which MLL directly regulates the expression of HOX cluster genes is actively being studied and has been shown to involve multiple domains of the large MLL protein (431 kDa) (Figure 3A).

The MLL protein is a homolog of *Drosophila* trithorax and is a member of the SET (Suppressor of variegation, Enhancer of zeste, Trithorax) family of histone methyltransferases\(^1,91\). With the intrinsic enzymatic activity of its SET domain, MLL catalyzes trimethylation of lysine 4 on histone 3 (H3K4me3), a modification which is associated with active gene transcription\(^92\). Interestingly, homozygous deletion of the exons encoding the SET domain from both Mll alleles leads to decreased H3K4
Figure 3. Structural schematic of MLL, a MLL fusion, and MLL2 proteins.

(A) The MLL protein is comprised of multiple functional domains. Leukemia-causing fusion proteins necessarily retain the N-terminal domains of MLL including the menin-binding domain, AT hooks, and CXXC DNA-binding domain. The repression and transactivation domains of MLL are highlighted by red and green boxes, respectively. The conserved Taspase I cleavage sites are indicated by arrowheads.

(B) MLL2 shares the same organization and types of functional domains as MLL.
methylation in *Hox* promoters and decreased expression of several *Hox* genes; this deletion, however, does not recapitulate the embryonic lethality of the *Mll* null mice\(^93\). This finding suggests that the histone methyltransferase (HMT) activity of MLL’s SET domain helps to promote *Hox* gene activation but that the HMT activity is not the only function of MLL that contributes to gene regulation during development. The ability of MLL to regulate *HOX* gene transcription during normal hematopoiesis or during misregulated leukemogenesis is also dependent on other functional domains of MLL and on MLL’s ability to localize to the nucleus\(^2, 80, 94, 95\).

MLL is proteolytically cleaved by Taspase I separating MLL into N- and C-terminal fragments which associate through internal FYRC and FYRN domains and localize to the nucleus\(^96, 97\). The N-terminal fragment of MLL can bind directly to DNA through several domains: a set of three AT hooks and a zinc-binding domain called the CXXC domain (discussed in detail below)\(^3, 98\). These AT hook motifs recognize AT rich or bent regions of DNA, and the CXXC domain recognizes non-methylated CpG dinucleotides\(^2, 3, 5, 11\). These DNA sequences recognized by MLL are relatively ubiquitous, so targeting of MLL to specific genes must occur through associations with other proteins that can provide greater specificity. Interaction with menin (and the subsequent recruitment of lens epithelium-derived growth factor, LEDGF, a transcriptional co-regulator) and associations with E2F and host cell factor (HCF)-1 have been shown to help target MLL to specific regions of DNA and facilitate transformation\(^99-104\). Through its interaction with menin, MLL functions in a larger complex consisting of Ash2L, WDR5, Rbp5, Dpy30, and HCF to regulate gene transcription\(^105\). The critical menin-binding and DNA-binding domains are found within
the N-terminal portion of MLL and are required for the leukemogenic activity of MLL fusion proteins which induce immortalization and subsequently cause leukemia (Figure 3A)\(^2,4,94\).

MLL also contains a group of PHD finger domains which are commonly found in chromatin-associating proteins. The third PHD finger of MLL binds to both H3K4me3 histones and to the peptidyl-prolyl cis-trans isomerase Cyclophilin 33 (Cyp33) through the RNA-recognition motif (RRM) domain\(^{106-108}\). Together these interactions help MLL regulate target Hox gene expression through the prolyl isomerization activity of Cyp33 which decreases MLL’s ability to bind to the H3K4me3 mark on histone tails\(^{106,109}\). In MLL fusions, this third PHD finger is necessarily absent so that Cyp33 cannot be recruited to repress the transcriptional activation of target genes facilitated by the fusion protein\(^{110}\).

The CXXC domain of MLL is located within a larger region of MLL termed the repression domain (amino acids 1100-1400) (Figure 3A). The repression domain was functionally named for its ability to suppress reporter gene expression in contrast to the activation domain in the C-terminus of MLL (residues 2340-3123) which showed increased expression in the same assay\(^3\). Our lab has previously shown that the repression domain directly interacts with co-repressor proteins—histone deacetylases 1 and 2 (HDACs 1 & 2), polycomb group proteins (HPC2 & BMI-1), and C-terminal-binding protein (CtBP)—although the individual residues of MLL with which they interact have not yet been determined\(^8\). HDACs 1 & 2 have enzymatic activity for removing acetyl groups from lysine residues on histone tails. This restores a positive charge to the side chain of histone lysines and increases histone affinity for the negatively
charged phosphate backbone of the DNA to effectively maintain a closed, heterochromatin state.

The activating functions of MLL (a member of the Trithorax group of proteins) are opposed by the repressive functions of Polycomb group proteins\(^{111}\). The interplay between Trithorax and Polycomb group (PcG) proteins provides tight regulation of gene expression that is necessary for proper segmental development in the embryo\(^{112}\). In the adult, the balance between these two groups of activator and repressor complexes helps to maintain proper gene expression during hematopoiesis. BMI-I and HPC2 are both members of the polycomb repressive complex 1 (PRC1) which recognizes the repressive H3K27me3 mark placed by the PRC2 complex\(^{113}\). HPC2/CBX4 is the human polycomb group 2 protein that was first identified by Satijn \textit{et al.} as a repressor protein with a highly conserved C-terminal binding domain\(^{114}\). HPC2 functions as a core component of the Polycomb Repressive Complex 1 (PRC1) along with BMI-1, RING1, and polyhomeotic homolog (Ph) proteins. With its distinct chromodomain, HPC2 recognizes and binds H3K27me3 marks that are catalyzed by EZH2 of the PRC2 complex and recruits RING1 (and subsequently, the rest of the PRC1 complex) through the C-terminal domain to repress transcription\(^{111, 114-116}\). The polycomb group proteins function to establish and maintain a heterochromatic state that represses gene transcription. CtBP can cooperate with PcG proteins to repress transcriptional activity through its direct interaction with HPC2\(^{117}\). CtBP has been implicated in tumorigenesis through the repression of tumor suppressor genes\(^{118, 119}\).

In MLL-associated leukemia, however, the balanced transcriptional regulation of MLL target genes is disrupted by the fusion of a partner protein to MLL. The N-terminus
of MLL is often fused to the C-terminal sequence of a protein, commonly a partner protein that is found in transcription elongation complexes\textsuperscript{28}. The recruitment of a fused transcriptional coactivator to MLL target genes is thought to cause uninhibited gene expression through the subsequent recruitment of factors which mediate transcriptional elongation\textsuperscript{120-122}. This gene activation through the fused partner protein overcomes the transient associations of the repressive PcG complexes leading to gene misregulation and ultimately leukemia\textsuperscript{123,124}.

The CXXC domain and ‘post CXXC’ residues of MLL were recently shown to interact with the coactivator Polymerase Associated Factor (PAF) protein\textsuperscript{9,10}. PAF is a component of the PAF complex (PAFc) comprised of PAF1, RTF1, Ctr9, Leo1, and Cdc73 protein subunits which were identified through association with the COMPASS complex containing members of the Trithorax family\textsuperscript{125}. PAFc cooperates with elongating RNA polymerase II to enhance transcription and was shown to be necessary for introducing activating epigenetic marks (i.e. methylation of H3K4 and H3K79) during active gene transcription\textsuperscript{125-127}. In MLL fusions, the PAF1 subunit directly interacts with the residues flanking the CXXC domain and mutation within the CXXC domain can disrupt this interaction\textsuperscript{9,10}. PAFc was shown in these studies to enhance \textit{Hox} gene expression mediated by a MLL-AF9 fusion, and disruption of the PAF1 interaction resulted in decreased recruitment of the MLL fusion to \textit{Hox} gene promoters, a subsequent decrease in \textit{Hox} expression, and loss of transformation capacity \textit{in vitro}.

MLL functions in the context of a large network of proteins that are associated with transcriptional activation and elongation, and association with Polycomb Group proteins helps to moderate the transcriptional activation of MLL target genes. Within
these larger transcriptional complexes, wild type MLL functions to maintain active gene transcription through the HMT activity of the SET domain and by protecting unmethylated promoter regions of target genes via the CXXC domain further facilitating gene expression\textsuperscript{92,128,129}. In contrast, MLL fusions alter this carefully regulated transcription to cause overexpression of homeotic genes that maintain hematopoietic cells in an undifferentiated state. It has recently been shown that the MLL-AF9 fusion protein requires a normal allele of \textit{MLL} to maintain transcription of gene targets\textsuperscript{130}. Together the wild type MLL protein and the MLL fusion protein alter hematopoiesis by localizing to and driving increased expression of a subset of genes that are normally regulated by wild type MLL\textsuperscript{131,132}.

**MLL2 is the closest homolog of MLL**

MLL (also known as MLL1) is the founding member of the MLL family of SET domain-containing H3K4 histone methyltransferases of which there are five members (numbered 1 to 5). The closest homolog of MLL is MLL2, whose gene is located at chromosomal band 19q13.1\textsuperscript{6}. (Note that MLL2 is officially named MLL4 by the Human Genome Organization Gene Nomenclature Committee, though, in the literature these names have been shared with a less homologous member of the MLL family known as ALR-1; MLL2—as I will refer to the closest homolog of MLL that was studied in this dissertation—has also less commonly been referred to as TRX2, HRX2, KMT2b, WBP7, and KIAA0304.) While the \textit{MLL2} transcript is shorter than MLL (8.5 versus 12.5 kb), both genes are encoded by 37 exons and retain the same number and orientation of functional domains including the AT hooks, CXXC DNA-binding domain, PHD fingers, and SET domain (Figure 3B)\textsuperscript{6,133}. The large MLL2 protein (290 kDa) is proteolytically
cleaved by Taspase I in the same manner as MLL\textsuperscript{134}. Together, these findings suggest that a gene duplication occurred in vertebrates and that the duplicated genes have subsequently evolved into the two closely related proteins MLL and MLL\textsuperscript{2}\textsuperscript{6}.

Like MLL, MLL2 is ubiquitously expressed throughout the body\textsuperscript{6,135}. However, MLL2 is not a functionally redundant paralog of MLL. MLL and MLL2 knockout mice are both independently embryonic lethal with death by day E10.5 versus E11.5, respectively, suggesting that loss of one of these MLL family members cannot be compensated by the other\textsuperscript{136}. Embryos of Mll2 null mice show growth retardation, delayed differentiation, and increased cellular apoptosis\textsuperscript{136,137}. Decreases in Hoxb1 and Mox1 expression were observed in somites surrounding the neural tube of Mll2 null embryos, reminiscent of the axial changes observed in development of Mll deficient embryos\textsuperscript{12,136}.

MLL2 associates with a similar transcriptional activation complex as MLL to regulate Hox gene transcription. Like MLL, this complex includes menin, Ash2L, Rbbp5, WDR5, CpG binding protein (CGBP/CFP1), and Dpy30\textsuperscript{105,138,139}. Of importance, MLL and MLL2 do not co-exist simultaneously within the same purified transcriptional complexes, further confirming distinct roles for these two paralogs\textsuperscript{140}. In a study of global H3K4me3 patterns in mouse embryonic fibroblasts (MEFs), Mll -/- cells showed decreased H3K4me3 within the Hoxa and Hoxb cluster and no change in trimethylation within the Hoxd cluster\textsuperscript{129}. When menin was knocked out in MEFs, a decrease in trimethylation within the Hoxa and Hoxb loci was observed with a significant reduction of H3K4me3 within the Hoxd locus. Since MLL and MLL2 both interact with menin during transcriptional regulation, it seems that H3K4 trimethylation within the
Hoxd locus does not depend on MLL (H3K4me3 persisted in Mll -/- mice) but instead is dependent on another histone methyltransferase—likely MLL2—which also requires menin for H3K4 trimethylation. Other studies looking at the effects of the loss of Mll2 at specific loci have shown decreased expression of Hoxb1, -2, and -5 and Mox1 (with unaffected Hoxa1 expression) and decreased H3K4me3 in the promoter regions of Bcl2 and Magoh2136, 137, 141. Broader screening for MLL2 target genes has yet to be conducted to better understand how the role of MLL2 is distinct from that of MLL. It is also not known whether MLL2 interacts with the same co-repressor proteins that were previously identified by our lab to interact with the MLL repression domain (HDAC1/2, HPC2, BMI-1, and CtBP)8. The first aim of my dissertation was to determine if these interactions are also conserved between these two homologs.

Given the structural and functional similarities between MLL and MLL2, it is intriguing that only two reports to date have suggested involvement of MLL2 in tumorigenesis133, 142. Several solid tumor types have shown gene amplification of MLL2 (pancreatic or glioblastoma cell lines) or overexpression of MLL2 protein (invasive breast and colon cancer cell lines) compared to normal tissue. Interestingly, MLL2 translocations have not yet been found in leukemia. When expressed by retrovirus, an experimentally derived MLL2-ENL fusion containing the same critical N-terminal domains found in MLL fusions (Menin-binding domain, AT hooks, and CXXC domain) was unable to transform primary hematopoietic cells in vitro7. Transformation capacity of these cells could be rescued by the chimeric fusion protein only when sequences that included the MLL CXXC domain were introduced in place of the MLL2 sequence (MLL residues 1050/1149 to 1337). This finding suggests that there are specific differences in
amino acids within the CXXC domains of MLL and MLL2 that alter the ability of the fusion proteins to transform hematopoietic cells. The first aim of my dissertation focuses on these specific amino acid differences between the CXXC domains of MLL and MLL2 to determine which, if any, of the residue differences contribute to the leukemogenic capacity of MLL fusion proteins.

The MLL CXXC Domain

The CXXC DNA-binding domain of MLL was first identified as a region having homology to a domain found in DNA methyltransferases\(^98\). This domain is found in the N-terminus of both MLL and MLL2 and shown to bind to unmethylated CpG dinucleotides\(^2,5,11\). This domain is necessarily retained in MLL fusion proteins that have the capacity to generate leukemia and helps to maintain active expression of target \textit{Hox} genes by preventing repressive promoter methylation\(^2,5,11,128\). The CXXC domain is named for the spacing of eight cysteine residues within this domain which coordinate two zinc molecules (Figures 4 and 5). The spacing of these cysteines is conserved across MLL homologs of vertebrate and invertebrate species and also found in MLL2 and other mammalian proteins. Four cysteines each coordinate one zinc ion through their reduced, sulphydryl sidechains to generate a stable, tetrahedral Zn-binding motif. This motif varies from a classical Zn-finger structure in that the coordinating cysteine residues are not sequentially organized (Figures 4A and 5B)\(^11,143\).

The CXXC domain is a crescent-shaped structure in which the interior surface forms direct interactions with its target DNA (Figure 5). This domain preferentially binds to non-methylated CpG DNA sequences through electrostatic interactions and hydrogen bonding\(^5,11\). The CXXC domain is essential for \textit{in vitro} transformation of cells
Figure 4. The MLL CXXC domain is conserved across species, and similar domains are found within other proteins.

(A) Sequence alignments of MLL CXXC domains from various eukaryotes and other CXXC-domain containing proteins showing highly-conserved spacing of zinc-coordinating cysteines (yellow) and the non-zinc-coordinating cysteines aligning with Cys1188 of hMLL (red). The coordination of the first (1) and second (2) zinc ions by the conserved cysteine residues is indicated below. (Alignments courtesy of Manuel Diaz, MD, Loyola University Chicago.)

(B) Phylogenetic tree showing the homology between various mammalian CXXC domains (reproduced from Figure 1C of Frauer et al.144).
Figure 5. Comparison of the MLL and MLL2 CXXC domains.

A

(B) Sequence alignment of the MLL (aa1147-1203) and MLL2 (aa958-1014) CXXC domains highlighting zinc-coordinating cysteines in yellow, Cys1188 of MLL in red, and non-conserved residues of interest in blue.

(B) Solved solution structure of the MLL CXXC domain (green ribbon) in complex with DNA (ball and stick) from PDB ID 2KKF.

(C) Threaded model of the MLL2 CXXC domain based on the solved MLL structure.

(D) Amino acids that were selected to study in Aim 1 by introducing the MLL residue into the MLL2 CXXC domain.

(A) Sequence alignment of the MLL (aa1147-1203) and MLL2 (aa958-1014) CXXC domains highlighting zinc-coordinating cysteines in yellow, Cys1188 of MLL in red, and non-conserved residues of interest in blue. (B) Solved solution structure of the MLL CXXC domain (green ribbon) in complex with DNA (ball and stick) from PDB ID 2KKF. (C) Threaded model of the MLL2 CXXC domain based on the solved MLL structure. (D) Amino acids that were selected to study in Aim 1 by introducing the MLL residue into the MLL2 CXXC domain.
expressing the MLL-ENL fusion protein. Mutation in any of the eight, Zn-coordinating residues inhibits oncogenic capacity of these cells due to destabilization of the protein structure and loss of DNA binding. Based on the NMR structure of this domain solved both in the presence and absence of CpG-containing DNA, Laurie Risner, PhD, in our lab has recently introduced point mutations into the MLL CXXC domain which do not disrupt folding of the domain. When mutations of the MLL-AF9 CXXC domain were introduced into residues that directly interacted with DNA (R1154A, K1185A, Q1187A), a decrease in DNA-binding affinity was observed compared to the wild type MLL CXXC domain in conjunction with a loss of in vitro transformation capacity of these mutant MLL-AF9 fusions in colony forming assays. This evidence demonstrates that MLL fusion proteins require the DNA binding capacity of the CXXC domain for the development of acute leukemia.

There is one cysteine residue—Cysteine 1188—in the MLL CXXC domain which does not coordinate with the zinc molecules (highlighted in red in Figures 4-6). It is located adjacent to a conserved, Zn-coordinating cysteine (Cys1189) and is critically positioned on the surface of the DNA-binding interface. The thiol group of a cysteine side chain is particularly susceptible to biochemical changes via redox reactions and post-translational modification. Much research has been reported on the regulation of transcription factor binding to DNA. In many instances, this regulation occurs at sites of cysteine residues within a Zn-coordinating structure. Zinc-binding motifs containing a zinc ion tetrahedrally coordinated to cysteine or histidine groups are relatively common within DNA-binding proteins including SP1, NF-κB, p53, TFIIH, estrogen receptor, and estrogen related receptor-2 (ERR-2).
transcription factors have been subject to modulation by redox reactions as determined by site-directed mutagenesis studies and treatments with oxidizing or reducing agents leading to altered DNA-binding affinity or altered target gene expression\textsuperscript{13,147,151}. The DNA-binding domains of p53, TFIH p44, and ERR-2 all contain an additional cysteine residue which does not coordinate zinc and physiologically modulate protein binding to DNA to function as a molecular switch\textsuperscript{152-154}.

The non-Zn-coordinating cysteine of the MLL CXXC domain, Cys1188, is conserved in the MLL orthologs of vertebrates but is noticeably absent in more distant, invertebrate species (Figure 4A). Interestingly, epigenetic regulatory mechanisms (i.e. those involved with methylation of gene promoters) have simultaneously diverged during eukaryotic evolution\textsuperscript{155-158}. These parallel evolutionary phenomena support the hypothesis that Cys1188 in the CXXC domain of MLL may be functioning as a molecular switch to regulate the binding of CXXC to its \textit{HOX} gene targets.

Cysteine 1188 does not directly interact with DNA through hydrogen bonds or salt bridges but is located on the DNA-binding surface of the MLL CXXC domain\textsuperscript{11}. When this residue was mutated to alanine (C1188A), the transduced MLL-AF9 construct was able to immortalize hematopoietic progenitor cells with efficiency similar to wild type MLL-AF9 in a colony formation assay\textsuperscript{11}. The DNA-binding affinity was also found to be equal to that of the wild type CXXC domain. In contrast, when this Cys1188 residue was mutated to aspartate (C1188D), a repulsive negative charge was introduced into the domain which disrupted DNA binding resulting in a loss of colony forming ability, even though the domain retained appropriate folding. \textit{In vivo} transplantation of bone marrow cells expressing MLL(C1188A)-AF9 or MLL(C1188D)-AF9 into mice
confirmed the *in vitro* results showing that functional binding of the MLL CXXC domain to DNA is required for the development of leukemia. Of note, mutation of Cys1188 to inert alanine (which is not susceptible to regulation by post-translational modifications) led to a shorter latency of leukemia development compared to the wild type Cys1188 (median survival of about 40 versus 90 days)\textsuperscript{11}. With a critical position on the DNA-binding surface and a sulfhydryl side group susceptible to post-translation modification, this Cys1188 residue may serve a functional role within the MLL CXXC domain to regulate MLL’s binding to *HOX* gene targets.

This non-zinc-coordinating cysteine residue is conserved in the MLL2 CXXC domain. There are, however, a number of other amino acids which differ between MLL and MLL2 within the CXXC domain (Figure 5). The goal of the first aim of my dissertation was to study the functional differences between the residues of the MLL2 CXXC domain and those of the MLL CXXC domain to understand how specific residues of MLL contribute to leukemia. Based on the alignments of the MLL and MLL2 CXXC domains (Figure 5A), the solved CXXC domain structure of MLL (Figure 5B), and the threaded model of the MLL2 CXXC domain (Figure 5C), strategic amino acid substitutions were introduced into the MLL2 CXXC domain (Figure 5D). Eleven of the thirteen substitutions that were studied are positioned between the conserved cysteine residues which coordinate the zinc ions. Since the structural cysteine residues allow the MLL CXXC domain to form a stable domain, I predicted that these non-conserved residues of the MLL2 CXXC domain would be found in a similar position within the MLL2 CXXC domain yet alter the overall DNA- or protein-binding function(s) of the domain. Two additional amino acid substitutions of the CXXC domains that were also
Figure 6. Comparison of amino acid differences within the MLL and MLL2 CXXC domains to be studied in Aim 1.

Four surfaces of the MLL CXXC domain colored according to the sequence alignment above in which blue residues represent amino acid substitutions that were introduced into the MLL2 CXXC domain to restore the residue(s) to the MLL sequence.
studied (A964R and R1011M) are positioned outside of the boundaries defined by the conserved, structural cysteines. These residues are found in close proximity to one another within the MLL CXXC domain and near the DNA. They were predicted to alter DNA-binding affinity, so these substitutions were also introduced into the MLL2 CXXC domain for studying their contribution to MLL function. The specific amino acid differences that were studied in my dissertation are highlighted in blue on the structural model of the MLL CXXC domain in Figure 6. I predicted that amino acid differences between the MLL and MLL2 CXXC domains would contribute to variable DNA-binding affinities or to variable interactions with other cofactor proteins depending on the location of the specific residues within the CXXC domain. Amino acid substitutions on the interior surface of the domain may alter DNA-binding affinity. In contrast, substitutions on the outer surface of the domain may alter binding of cofactor proteins to affect overall function of the fusion protein or may indirectly contribute to altered CXXC DNA-binding affinity through allosteric effects. The goal of my first project aim was to explore how specific amino acid differences between the MLL and MLL2 CXXC domains affect the critical functioning of the CXXC domain and alter the transformation capacity of leukemogenic fusion proteins.
CHAPTER 3

METHODS

Cloning of MLL/MLL2-AF9 fusion constructs for in vitro colony assay

The wild type MLL2 CXXC domain was PCR amplified from a pcDNA4-His-MaxB plasmid containing the entire exons 1-37 of MLL2 (NCBI Accession Number NM_014727, courtesy of Kevin FitzGerald, S.J., PhD, Georgetown University) using a high-fidelity, cloned Pfu DNA polymerase (Stratagene, #600159) and using primers with AatII and BamHI restriction sites on the ends of the forward and reverse primers (Table 2, A). (Priming sites of cloning primers are depicted in Table 2, and the primer sequences are listed in Table 3.) The blunt end PCR product consisting of the MLL2 CXXC domain (nucleotides 2872-3042) with the flanking restriction sites was ligated into a pCR-Blunt holding vector in a 10µL reaction volume for 1 hour at 16°C using T4 DNA Ligase (Fermentas, #EL0014). 1µL of the ligation reaction mixture was transformed into 50µL of chemically competent OneShot® Top10 Competent E.coli cells according to the Zero Blunt® PCR Cloning Kit protocol (Invitrogen, #K2750; #C4040). The accuracy of the MLL2 CXXC insert was verified by DNA sequencing (ACGT, Inc., Wheeling, IL) using the M13R primer (Table 2, B; sequencing primers and priming sites are indicated in gray). The pCR-Blunt-MLL2-CXXC construct was purified using the GeneJET™ Plasmid Miniprep Kit (Fermentas, #K0503) as were all subsequent plasmids used in the cloning of the fusion constructs. All plasmid concentrations following
Table 2. Cloned DNA constructs showing restriction enzyme cutting sites and priming sites which were used for subcloning and sequencing.

<table>
<thead>
<tr>
<th>Name</th>
<th>Plasmid</th>
<th>Construct</th>
<th>Cloning Primer(s)</th>
<th>Sequencing Primer(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>MLL2 exons 1-37 (NCBI NM_014727)</td>
<td>pcDNA4-His-MaxB 5.3kb plasmid + 8.5kb insert</td>
<td>EcoRV</td>
<td>MLL2 CXXC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MLL2 'post-CXXC'</td>
</tr>
<tr>
<td>B</td>
<td>MLL2 CXXC (or mt MLL2 CXXC)</td>
<td>pCR-Blunt 3.5kb plasmid + 171bp insert</td>
<td>AatII, BamHI, Site-directed mutagenesis</td>
<td>M13R</td>
</tr>
<tr>
<td>C</td>
<td>MLL2 'post-CXXC'</td>
<td>pCR-Blunt 3.5kb plasmid + 480bp insert</td>
<td>BamHI, SalI</td>
<td>M13R</td>
</tr>
<tr>
<td>D</td>
<td>5' MLL</td>
<td>pMCS5 2.9kb plasmid + 4.2kb insert</td>
<td>EcoRI, AatII, BamHI, SalI</td>
<td>MLL 1119F</td>
</tr>
<tr>
<td></td>
<td>5' MLL/MLL2(CXXC)</td>
<td>pMCS5 2.9kb plasmid + 4.2kb insert</td>
<td>EcoRI, AatII, BamHI, SalI</td>
<td>MLL 1119F</td>
</tr>
<tr>
<td>F</td>
<td>5' MLL/MLL2(wt/mt CXXC+ 'post-CXXC')</td>
<td>pMCS5 2.9kb plasmid + 4.1kb insert</td>
<td>EcoRI, AatII, BamHI, SalI</td>
<td>MLL 1119F</td>
</tr>
<tr>
<td>G</td>
<td>MLL-AF9</td>
<td>pMSCVneo 6.5kb plasmid + 4.5kb insert</td>
<td>EcoRI, SalI, BglII</td>
<td>MLL RD</td>
</tr>
<tr>
<td>H</td>
<td>MLL/MLL2(CXXC)-AF9</td>
<td>pMSCVneo 6.5kb plasmid + 4.5kb insert</td>
<td>EcoRI, SalI, BglII</td>
<td>MLL 1119F</td>
</tr>
<tr>
<td>I</td>
<td>MLL/MLL2(wt/mt CXXC+ 'post-CXXC')-AF9</td>
<td>pMSCVneo 6.5kb plasmid + 4.4kb insert</td>
<td>EcoRI, SalI, BglII</td>
<td>MLL1/2 RD</td>
</tr>
<tr>
<td>J</td>
<td>MLL or MLL1/2 Repression Domain</td>
<td>pGS-21a 6.2kb plasmid + 900 or 800bp insert, resp.</td>
<td>EcoRV/NruI, SalI</td>
<td>T7 Term</td>
</tr>
</tbody>
</table>
## Table 3. MLL and MLL2 cloning and sequencing primers

<table>
<thead>
<tr>
<th>Cloning Primers</th>
<th>Nucleotides and Sites Details</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Amplicon</strong></td>
<td><strong>Primers (5’ to 3’)</strong></td>
</tr>
<tr>
<td><strong>MLL2 CXXC</strong></td>
<td>For: GGG ACG TCG AGG CAA GAA GAT GCG CAT GGC TCG ATG &lt;br&gt;Rev: GGG GAT CCT TCC ATC TTC CGA GCC TCT ATT TTG TC</td>
</tr>
<tr>
<td><strong>MLL2 ‘post-CXXC’</strong></td>
<td>For: GGG GAT CCC GAC TGG CTA AAA AAG GCC GG &lt;br&gt;Rev: GGG TCG ACA CGG ACG CGG TGC ACA CC</td>
</tr>
<tr>
<td><strong>MLL RD</strong></td>
<td>For: GGT CGC GAG AAC GAG AAA AGA TTT TGT CTT CC &lt;br&gt;Rev: GGG TCG ACT CTG ATC CTG TGG ACT CC</td>
</tr>
<tr>
<td><strong>MLL 1/2 RD</strong></td>
<td>For: GGT CGC GAG AAC GAG AAA AGA TTT TGT CTT CC &lt;br&gt;Rev: GGG TCG ACA CGG ACG CGG TGC ACA CC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sequencing Primers</th>
<th>Nucleotides and Sites Details</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Amplicon</strong></td>
<td><strong>Primers (5’ to 3’)</strong></td>
</tr>
<tr>
<td><strong>M13R (universal)</strong></td>
<td>GGA AAC AGC TAT GAC CAT G</td>
</tr>
<tr>
<td><strong>MLL 1119F</strong></td>
<td>CAG AAG ATG CTG AAC CTC TTG CTC C</td>
</tr>
<tr>
<td><strong>AF9 Rev</strong></td>
<td>GGG TCG ACA CGG ACG CGG TGC ACA CC</td>
</tr>
<tr>
<td><strong>T7 Term (universal)</strong></td>
<td>GCT AGT TAT TGC TCA GCG G</td>
</tr>
</tbody>
</table>

Miniprep isolation were determined using 2µL of eluted plasmid DNA with a NanoDrop 2000 (Thermo Scientific) prior to submitting sample(s) for sequencing or prior to subsequent digestion.

A novel construct encoding the N-terminal fragment of MLL (nucleotides 1-4212) was previously generated by Laurie Risner, PhD. This construct contains 5’-AatII and 3’-BamHI restriction sites flanking the nucleotide sequence encoding the MLL CXXC domain for conveniently swapping in mutant or alternative sequences (while still maintaining the original reading frame of MLL: NCBI Accession Number L04284)\(^1\) (Table 2, D). This 5’-fragment of MLL (pMCS5 vector) and the sequence encoding the MLL2 CXXC domain (pCR-Blunt vector) were subjected to double digestion with AatII (New England Biolabs, NEB) and BamHI (NEB) restriction enzymes in 40µL reaction volumes at 37°C for 1-2 hours in NEB Buffer #4 to release the respective sequence encoding the CXXC domains from the vectors. Both reactions were heat inactivated at 65°C for 15 minutes. Calf intestinal alkaline phosphatase, CIP (NEB, #MO290S), was
added to the digested pMCS5 vector containing the digested 5'-MLL fragment (35 minute incubation at 37°C) to prevent self-ligation of the vector by removing the terminal 5'-phosphate. The digestion products were then separated by electrophoresis on a 1% agarose gel or a 4% NuSieve® 3:1Agarose (BiWhittaker Molecular Analysis—BMA, #50090) gel at 90V for 30min. The respective bands were gel purified with the Qiaex® II Gel Extraction Kit (Qiagen, #20051) and ligated together for 2 hours at 14°C to create a 5’ construct encoding the chimeric MLL/MLL2(CXXC) domain (Table 2, E). 5µL of the ligation reaction mixture was used to transform 50µL of chemically competent DH5α E.coli cells according to supplier’s protocol (NEB, #C2988J). This construct was sequenced with the MLL 1119F primer which anneals at MLL nucleotide 3355 just upstream of sequence encoding the CXXC domain to ensure that the swap of the MLL and MLL2 sequences was successful and to ensure that the correct reading frame was maintained through the restriction sites.

The ‘post-CXXC’ region of MLL2 (nucleotides 3043-3522) was cloned from the full length MLL2 construct using high fidelity PCR amplification as described above, this time, with forward and reverse primers containing BamHI and SalI restriction sites, respectively (Table 2, A). The blunt PCR product was ligated into a pCR-Blunt vector and transformed into Top10® chemically competent E.coli as described above. The insert was sequenced to verify its accuracy using the M13R primer (Table 2, C).

To exchange the ‘post-CXXC’ region of MLL and MLL2, the chimeric pMCS5-5’ MLL/MLL2(CXXC) construct and the pCR-Blunt-MLL2 ‘post-CXXC’ constructs were doubly digested with BamHI (NEB) and SalI (NEB) in the same reaction conditions as above and the products were separated on a 1% agarose or 4% NuSieve® 3:1Agarose
gel by electrophoresis (90V for 30min). The appropriate fragments were then gel purified and ligated at 14°C for 4 hours to form the pMCS5-5’ MLL/MLL2(CXXC + ‘post-CXXC’) construct (Table 2, F). 2µL of the ligation mixture was used to transform 50µL of DH5α *E.coli* cells (NEB). The plasmid DNA was purified, and the insert accuracy was verified by sequencing with the MLL 1119F primer.

The wild type fusion construct, pMSCVneo-MLL-AF9, was previously cloned in our lab by Laurie Risner, PhD. This construct was used for subsequent studies as the highly-transforming, positive control and also served as the source of the 300-bp AF9 fusion partner for the MLL/MLL2 constructs (Table 2, G). The 3’-end of AF9 (NCBI Accession Number AK301474)—which is fused to a sequence encoding a FLAG tag at the 3’ end—was digested from the pMSCV-MLL-AF9 with SalI (Promega) and BglII (NEB) in a 30µL reaction volume with Buffer H (Pharmacia) overnight at 37°C. This digestion reaction was heat inactivated as done previously, the products separated by electrophoresis on a 2% agarose gel (90V for 45 minutes), and gel purified for use in a three-way ligation reaction.

The empty vector, pMSCVneo (used for retroviral expression in mammalian packaging cells), was linearized by double digestion with EcoRI (Gibco) and BglIII (NEB) in a 20µL reaction volume with Buffer H (Pharmacia) overnight at 37°C. This digestion reaction was heat inactivated and treated with CIP as before to prevent self-ligation of the vector. The linearized plasmid was purified through electrophoresis and gel extraction. The 5’-MLL/MLL2(CXXC ± ‘post-CXXC’) nucleotide sequence was released from the pMCS5 vector by digestion with EcoRI (Gibco) and SalI (Invitrogen) in a 50µL reaction volume with Buffer H (Pharmacia) for 2-3 hours at 37°C. The
digestion products were purified by gel extraction following electrophoresis on a 0.8% agarose gel for 45 minutes for optimal separation between the 4.2 kb insert and the ~3kb plasmid (Table 2, E & F).

To complete the cloning of the final fusion constructs, a three-way ligation reaction was conducted in which the EcoRI/BglII-linearzied pMSCVneo plasmid (6.5kb), the Sall/BglII-digested fragment of AF9 (300bp), and one of the EcoRI/Sall-digested 5’-MLL/MLL2(CXXC ± ‘post-CXXC’) fragments were combined in a 1:5:2 molar ratio with T4 DNA ligase (Fermentas) in a 10µL reaction volume overnight (nearly 24 hours) at 14°C. 1µL of three-way ligation mixture was added to 20µL of freshly thawed ElectroMAX™ DH5α-E™ E.coli cells (Invitrogen, #11319-019) while on ice. This cell and ligation mixture was then transferred to a pre-chilled, disposable microelectroporation chamber (Whatman Biometra, #11608-031) within the Chamber Safe component for the Cell-Porator electroporation apparatus (Gibco BRL Life Technologies, #11604). With the Voltage Booster set at 4kΩ and the Pulse Control / Power Supply set at 330kF capacitance, “Low Ω,” “Fast” charge rate, and switched to “Charge,” the voltage on the unit was increased above 310V. The “Charge” switch was flipped to “Arm” allowing the voltage to decrease until exactly 310V was reached at which point the cells were stimulated with an electrical pulse by pressing the trigger button. The electrically transformed cells were then transferred to 1mL of pre-warmed S.O.C. media, incubated for 1 hour at 37°C while shaking at 225rpm, and plated on LB agar with ampicillin (100µg/mL) selection overnight to allow transformed colonies to grow. The few colonies that formed were picked from the plate and screened for the correct insert first by PCR amplification using Taq DNA Polymerase (GenScript,
and secondly via bi-directional sequencing using the proximal MLL 1119F and distal AF9 Rev primers to ensure sequence coverage across all restriction site junctures of the large, 4.5kb fusion constructs (Table 2, H and I).

All PCR reactions were performed using a PCR Express Thermal Cycler (Thermo Hybaid, #HBPX 110). Agarose gel electrophoresis of constructs was conducted using the Sub-Cell® GT Agarose Gel Electrophoresis System (BioRad) and a low-voltage power supply (Thermo EC, #EC250-90). All small-scale centrifugations were conducted in a benchtop Spectrafuge 16M (Labnet, #C0160), and ligation reactions were chilled in a benchtop MicroCooler II™ (Boekel, #260010).

**Generation of and cloning of mutant MLL2 CXXC domain constructs**

Site-directed mutagenesis was performed on the isolated MLL2 CXXC domain within the pCR-Blunt plasmid (Table 2, B) using the QuickChange Site-Directed Mutagenesis Kit (Stratagene, #200518) in order to introduce residues of the MLL CXXC domain into the MLL2 CXXC domain. Mutagenesis primers are listed in Table 4. Performing the mutagenesis reaction on the isolated domain was advantageous as the small insert size facilitated convenient sequence verification and ensured that unexpected mutations did not occur during the PCR amplification steps of the mutagenesis reaction. A high fidelity *PfuTurbo* DNA polymerase provided with the kit was used to ensure accuracy of the nucleotide substitution(s).

The kit protocol was followed with some exceptions: 500ng of template pCR-Blunt-MLL2 CXXC dsDNA was used for the PCR reactions (instead of the recommended 5-50ng); the elongation step was run for 4 minutes to allow extension of the full 3.7kb construct; and the PCR reaction was run for 14 cycles. In the case of two
Table 4. Site-directed mutagenesis primers used to introduce individual MLL residues into the MLL2 CXXC domain.

<table>
<thead>
<tr>
<th>Amino Acid Substitution (MLL2 → MLL)</th>
<th>Codon Conversion</th>
<th>Primers (5’ to 3’)</th>
</tr>
</thead>
</table>
| A964R                               | GCT → CGT        | For: GCA AGA AGA TGC GCA TGC GTC GAT GTG GAC ACT GTC  
|                                     |                  | Rev: GAC AGT GTC CAC ATC GAC GCA TGC GCA TCT TCT TGC |
| H968Q                               | CAC → CAG        | For: ATG GCT CGA TGT GGA CAG TGT CGG GGC  
|                                     |                  | Rev: GCC CCG ACA CTG TCC ACA TCG AGC CAT |
| R970P                               | CGG → CCG        | For: GTG GAC ACT GTC CGG GCT GCC TAC G  
|                                     |                  | Rev: CGT AGG CAG CCC GGA CAG TGT CCA C |
| V975P                               | GTG → CCG        | For: CGG GGC TGC CTA CGT CCG CAG GAC TGT G  
|                                     |                  | Rev: CAC AGT CCT GCG GAC GTA GGC AGC CCC G |
| LRVQ973-6QVPE                       | CTA+GTG+CAG+CTA+CGG+GAG | For: GGA CAC TGT CGG GGC TGC CAA GTT CCG GAG GAC TGT GGG TCC TGT GT  
|                                     |                  | Rev: ACA CAG GAC CCA CAG TCC TCC GGA ACT TGG CAG CCC CGA CAG TGT CC |
| V982T                               | GTC → ACC        | For: AGG ACT GTG GGT CCT GTA CCA ACT GCC TAG ACA AGC  
|                                     |                  | Rev: GCT TGT CTA GGC AGT TGG TAC AGG ACC CAC AGT CCT |
| P993R                               | CCT → CGT        | For: CCA AGT TTG GGG GCC GTA ACA CCA AGA AGC A  
|                                     |                  | Rev: TGC TTC TTG GTG TTA CGG CCC CCA AAC TTG G |
| T995I                               | ACC → ATC        | For: TTG GGG GCC CTA ACA TCA AGA AOC AGT GCT G  
|                                     |                  | Rev: CAG CAC TGC TTC TTG ATG TTA GGG CCC CCA A |
| V1001K                              | GTA → AAA        | For: CAC CAA GAA GCA GTG CTG TAA ATA CCG GAA GTG TGA CAA A  
|                                     |                  | Rev: TTT GTC ACA CTT CCG GTA TTT ACA GCA CTG CTT CTT GTG G |
| VY1001-2KM                          | GTA+TAC → AAA+ATG | For: GCC CTA ACA CCA AGA AGC AGT GCT GTA AAA TGC GGA AGT GTG ACA AAA TAG AGG  
|                                     |                  | Rev: CCT CTA TTT TTG TAC ACT TCC GCA TTT TAC AGC ACT GCT TCT TGG TGT TAG GCC |
| R1011M                              | CGG → ATG        | For: GTG TGA CAA AAT AGA GGC TAT GAA GAT GGA AGT ATC CCA C  
|                                     |                  | Rev: CAC ACT GTT TTA TCT CCG ATA CTT CTA CCT TCC TAG GGG G |
mutagenesis reactions (T995I and R1011M), the PCR reaction was set up with temperature decrements (“touchdown mode”) in which the annealing temperature was begun at 62°C (higher than the recommended 55°C) for four cycles to facilitate the specific annealing of the mutant primers, with the annealing temperature then decreased 4°C every four cycles until the annealing temperature of 50°C was reached for the final six cycles (18 total cycles): 4 cycles at 62°C, 4 cycles at 58°C, 4 cycles at 54°C, and 6 cycles at 50°C. Following the 1 hour digestion of all reactions with DpnI at 37°C to remove the residual wild type template from the reaction mixture, 50µL of XL1-Blue supercompetent cells (provided in kit) were transformed with 1µL of the mutagenesis PCR product and plated on LB agar with kanamycin (25µg/mL) selection. Individual colonies from each mutagenesis reaction were grown, plasmid DNA purified from each clone, and constructs sequenced using the M13R primer to verify that the correct mutation was present with no other alterations in the DNA sequence.

The sequence-verified clones were then subjected to double digestion (like the wild type pCR-Blunt-MLL2 CXXC construct above) using AatII and BamHI (NEB or Fermentas FastDigest®) restriction enzymes in 40µL reaction volumes and ~5ng of mutant MLL2 CXXC plasmid DNA per reaction. The digested, mutant MLL2 CXXC constructs were gel purified and swapped into the pMCS5-5’MLL/MLL2(wt. CXXC + ‘post-CXXC’) construct (Table 2, F) using the AatII and BamHI restriction sites as was done to exchange the wild type MLL2 CXXC sequence above. The mutant constructs were subsequently subcloned into the larger fusion construct like the wild type version using the same restriction enzyme sites and three-way ligation technique described above. The constructs were sequenced at each step to verify the accuracy of the inserts including
the presence of the desired mutation(s) and the maintenance of the necessary reading frame through all restriction site junctures that make up the final pMSCVneo-MLL/MLL2(mt CXXC + ‘post-CXXC’) AF9 constructs (Table 2, I).

**Production of Retrovirus**

Plasmid DNA of the pMSCVneo (negative control), pMSCVneo-MLL-AF9 (positive control), and pMSCV-MLL/MLL2(wt./mt. CXXC ± ‘post-CXXC’) AF9 constructs was purified from 250mL overnight bacterial cultures in LB selection media using the Plasmid Purification Midi Kit (Qiagen, #12145). Centrifugation steps were performed using a high-capacity, floor model centrifuge (Beckman, #J2-MI) and the corresponding JA-10 and JA-20 rotors. The final DNA pellet was dissolved in 10mM Tris, pH 8.0 and the concentration determined using a NanoDrop 2000.

Five plates of Phoenix™ Ecotropic virus-packaging cells (Allele Biotech. & Pharm. Inc., #ABP-RVC-10002) were plated at 3x10⁶ cells per 10cm² dish for each construct on Day 1. These cells were allowed to grow overnight in HyClone® Dulbecco’s Modified Eagle’s Medium, DMEM/High Glucose (Thermo Scientific, #SH30022.01), with 1% penicillin/streptomycin and 10% fetal bovine serum (DMEM/1%P/S/10%FBS) at 37°C in a Heraeus incubator to achieve 50-60% confluence by the time of transfection on Day 2. Each plate of cells was transfected with 25µg plasmid DNA using the calcium phosphate CalPhos™ Mammalian Transfection Kit (ClonTech, #631312). Cells were incubated overnight at 37°C in 5% CO₂ with culture media containing the DNA plasmid and calcium phosphate complexes. This calcium phosphate-containing media was removed after ~16 hours on Day 3 and replaced with fresh media lacking calcium phosphate. The cultures were transferred to 32°C and 5%
CO₂ and the virus-containing media was collected from and replaced for each plate at 24-hour intervals for two days (Days 4 and 5). The media from plates containing identical constructs was pooled together on Day 5. The retroviral packaging cells were harvested in 1x sterile phosphate buffered saline (PBS) for detection of stable fusion protein production via Western Blot at a later date. The cells were pelleted in aliquots of 10⁷ cells in a Beckman GS-6KR centrifuge at 1500 rpm and 10°C and snap frozen on dry ice for storage at -80°C.

The pooled media containing the virus particles of each construct type was added to a primed Centricon® Plus-70 filter (Millipore, #UFC10008) for concentrating the virus supernatant. The viral media was filtered via centrifugation at 1640rpm and 25°C in a Beckman GS-6KR centrifuge for 1.5-2 hours until all but 1-2mL had entered the filter. The filters were inverted onto a collection reservoir and centrifuged in a benchtop centrifuge (Sorvall, #RT6000B) for 10 minutes at 3000rpm and 25°C to release the concentrated media containing the virus particles. The volume of the virus supernatant was adjusted to a final 7x virus concentration by adding the appropriate volume of HyClone® Roswell Park Memorial Institute (RPMI)-1640 media (Thermo Scientific, #SH30027.01) with 1% P/S and 20% FBS. The concentrated supernatants were separated into aliquots of 750µL, snap frozen on dry ice, and stored at -80°C.

**Determination of Retrovirus Titer**

The concentration of retrovirus was determined by infecting Rat1a fibroblast cells with serially diluted virus supernatant and was measured by the ability of the cells to form colonies in the presence of antibiotic selection. In this manner, colony formation would indicate effective viral uptake and expression of the desired plasmid through
acquired resistance to the antibiotic. A high number of colony forming units (CFUs) corresponds with a high virus titer, especially when the high number of colonies is formed in a well where the virus had been present in a low dilution. To measure the virus titer, a standardized number of 85,000 Rat1a cells was plated in each well of a 6-well plate (i.e. 1 6-well plate per retroviral construct), and the cells were allowed to grow overnight at 37°C and 5% CO₂ in DMEM/1%P/S/10%FBS. A series of 10-fold virus dilutions were made ranging from 10⁻² to 10⁻⁶ in 1-mL volumes of DMEM/10%FBS with 16µg/mL polybrene (without the 1% P/S that was previously used). Polybrene, a positively charged polymer, was added to the media to increase viral uptake efficiency by helping to neutralize repulsive negative charges on the cell surfaces. The day-old culture media from each well was replaced with this 1-mL of fresh media containing a diluted virus: 10⁻², 10⁻³, 10⁻⁴, 10⁻⁵, 10⁻⁶, or 1mL of polybrene-containing media with no virus. The cultures were returned to 37°C for 4 hours to allow the cells to take up the virus after which time 1ml of fresh DMEM/10%FBS without polybrene or P/S was added to each well. The cultures were allowed to grow overnight at 37°C and 5% CO₂. The following morning, the media was replaced with 2mL of fresh DMEM/10%FBS, and 1 mg/mL of HyClone® G418 antibiotic (Thermo Scientific, #SV30069.1) was added to each well to select for Rat1a cells which took up the retrovirus and express the neomycin resistance gene from the pMSCV constructs. Cultures were incubated at 37°C and 5% CO₂ for twelve days while changing the media as needed (usually every fourth day) and maintaining the G418 antibiotic selection for the duration of the culture period. At the end of the twelve days of selection, the media was removed from each well, and methylene blue dye (1.7µg/mL in methanol) was added to each well to stain Rat1a
colonies. The number of colonies in each well was counted as a colony forming unit (CFU), and the virus titer for each construct was determined by the average number of colonies in each well of serially diluted virus.

**Isolation of murine bone marrow cells**

Six- to twelve-week old C57Bl/6, male mice (Taconic, #C57BL/6NTac) were euthanized via CO$_2$ inhalation. The femurs and tibias were collected and cleaned using sterile surgical technique and placed in RPMI/2%FBS. The marrow cavities of the bones were flushed into a culture dish with this RPMI/2%FBS using a 1mL, 25-gauge, Eclipse™ Injection Needle with Luer-Lok™ Syringe (BD, #305780). The cells were pelleted by centrifugation (Beckman GS-6KR) at 1500rpm and 4°C for 5 minutes and resuspended in sterile, chilled 1x red cell lysis buffer (10x red cell lysis buffer: 1.55M NH$_4$Cl, 120mM NaHCO$_3$, 1.3mM EDTA). Resuspended cells were incubated at room temperature with periodic mixing for about 5 minutes until the red tint was no longer visible. This mixture was then passed through a nylon 70µm cell strainer (BD Falcon, #352350) to remove any debris. The remaining cells (the viable white blood cells from the bone marrow) were pelleted, resuspended in sterile 1x PBS, and counted using a hemocytometer counting chamber (Hausser Scientific, #3200) and 0.4% trypan blue stain (Gibco, #15250-061). The mice were handled and housed in compliance with LUMC Comparative Medicine Facility and I.A.C.U.C. policies.

**CD117 positive selection of murine bone marrow progenitor cells**

The isolated white blood cells from the mouse bone marrow underwent selection for the c-kit (CD117) cell surface marker using the EasySep® Mouse CD117 Positive Selection Kit (Stem Cell Technologies, #18757). Specifically, isolated white blood cells
were resuspended in sterile PBS/2%FBS to a final concentration of 1x10^8 cells/mL in a round-bottom Fisherbrand® 12x75mm polypropylene culture tube (Fisher, #14-956-10). The CD117 PE Labeling Reagent (containing c-kit antibody) was added to the cells at 50µL per 1-mL volume of cells, mixed, and allowed to incubate for 15 minutes at room temperature. Next, the EasySep® PE Selection cocktail containing a bispecific tetrameric antibody complex (recognizing both c-kit/CD117 and dextran) was added at 70µL per 1-mL volume of cells, mixed, and incubated at room temperature for 15 minutes. Resuspended Magnetic Nanoparticles coated with dextran were then added at 50µL per 1-mL volume of cells, mixed, and incubated at room temperature for 10 minutes. The suspension volume was brought up to 2.5mL by adding PBS/2%FBS and the tube was set inside the EasySep® Magnet (Stem Cell Technologies, #18000) without the tube cap and was incubated for 5 minutes at room temperature to magnetically separate the c-kit positive bone marrow progenitor cells. Next, while the tube remained inside the magnet, the supernatant was smoothly decanted in a single motion leaving the magnetically-separated c-kit positive cells within the tube. The tube was removed from the magnet, and the cells were washed with 2.5mL PBS/2%FBS and, again, incubated within the magnet without the tube cap for 5 minutes. This wash and incubation step was repeated for a total of four separations. The washed cells, having been positively selected based on c-kit/CD117 cell-surface expression, were counted and resuspended in RPMI/10%FBS/1%P/S which was supplemented with 0.05 mM β-mercaptoethanol, 50ng/mL SCF, 10ng/mL IL-3, and 10ng/mL IL-6 at a final concentration of 10^6 cells/mL. The cells were incubated in a humidified round-bottom, 96-well plate overnight at 37°C and 5% CO₂.
Retroviral transduction of c-kit positive bone marrow progenitor cells

The spinoculation method for was used for transducing c-kit positive bone marrow progenitor cells. 30,000 cells which had been selected on the previous day were added to a mixture of 750µL of concentrated viral supernatant, 7.5mM HyClone® HEPES solution (Thermo Scientific, #SH30237.01), 4µg/mL polybrene (also called hexadimethrine bromide from Sigma, #H9268), and RPMI/10%FBS to a final volume of 5mL in a conical 12x75mm culture tube (VWR, #60818-102). (Two aliquots—1.5mL total—of virus supernatant were used for the V975P and V982T constructs since these retroviruses had lower yields in the 10^4 CFU range. The total spinoculation volume was 10mL for these constructs, and the volumes of the other components were doubled accordingly to maintain the concentrations. Spinoculation with the empty vector, pMSCVneo retrovirus was conducted using half a volume of supernatant, 375µL, in a total volume of 1mL.) The spinoculation mixture was briefly mixed by vortexing and then subjected to centrifugation (Beckman GS-6KR) at 3000rpm and 33°C for 4 hours. Following spinoculation, the supernatant was removed, and the cells were resuspended in 200µL of fresh RPMI/10%FBS/1%P/S supplemented with 0.05mM β-mercaptoethanol, 50ng/mL SCF, 10ng/mL IL-3, and 10ng/mL IL-6. The cells were incubated in a humidified round-bottom, 96-well plate overnight at 37°C and 5% CO₂. The following day, the cells underwent a second spinoculation using the same volumes and components as the previous day. After this second spinoculation, the retrovirally transduced cells were plated in methylcellulose as described below.
Methylcellulose Colony Assay

Transduced c-kit positive hematopoietic progenitor cells were plated at 1,000-10,000 cells per 35mm culture plate depending on the observed transformation capacities of the various retroviral fusion constructs. Cells were added to MethoCult® M3234 Methylcellulose Medium for Mouse Cells (Stem Cell Technologies, #03234) which had been supplemented with glutamine, 50ng/mL SCF, 10ng/mL IL-3, 10ng/mL IL-6, and 10ng/mL GM-CSF. For the first round of plating immediately following viral transduction, 1.25mg/mL of HyClone® G418 antibiotic (Thermo Scientific, #SV30069.1) was added to the semi-solid media to select for those progenitor cells which were effectively transduced with the retrovirus and stably expressing the plasmid of interest. This mixture of cells, methylcellulose, cytokines, and antibiotic were thoroughly vortexed to generate a homogeneous culture. A 16-gauge, 1.5” blunt-end needle (Popper, #SBN16X1-1/2) was added to a 3ml syringe with Luer-Lok™ Tip (BD, #309657) and both were primed by flushing with culture medium to remove all residual air. 1.1mL of culture medium containing the cell/methylcellulose/cytokine mixture ± antibiotic was added to each of two tissue culture-treated 35-mm dishes (CytoOne, #CC7682-3340). The plates were gently rocked to evenly distribute the culture media throughout the plate. The smaller 35-mm dishes were added to a larger humidified chamber and incubated at 37°C for six days to allow for colony formation.

At the end of 6-days, the number of colonies on each plate was counted, reported as colony forming units (CFUs), and all CFUs were normalized to an initial number of 10,000 cells plated per dish. Each construct had two plates of cells grown in parallel, and the number of CFUs was averaged between these two plates for the final, normalized
CFU count of each round. Pictures of representative colony morphology were taken on a Leica DMIL microscope at 4x objective magnification with a Canon PowerShot S40 digital camera and acquired with Canon Utilities ZoomBrowser EX v3.2 software.

Cells from each round of plating were harvested by diluting the semi-solid methylcellulose media in sterile PBS and transferring to a conical tube for centrifugation. The number of viable cells was counted on a hemocytometer by trypan blue staining and reported as a total cell number normalized to 10,000 cells initially plated. 1,000-10,000 cells transduced with each construct were re-plated in methylcellulose media without G418 antibiotic for subsequent rounds (at least 3) to determine replating and proliferative potential of each fusion construct. Remaining cells from each round were frozen for viability in bone marrow freezing media (RPMI/47.5%FBS/7.5%DMSO) at a rate of -1°C/minute in an isopropanol Cryo 1°C Freezing container (Nalgene, #5100).

Colony assays for hydrogen peroxide (Sigma, #U8879-50TAB), parthenolide (Sigma, #P0667), and spermine nonoate (Sigma, #S150) toxicity assays were performed in the same manner as described above using MLL-AF9 and MLL(C1188A)-AF9 cells that had been previously transformed\textsuperscript{11}. These cells were treated overnight with varying concentrations of compound prior to being plated in methylcellulose at 1,000-2,000 cells per plate and incubated for 5-7 days.

For the colony assays with varying oxygen concentrations, untreated MLL-AF9 and MLL(C1188A)-AF9 cells were plated in methylcellulose with the appropriate cytokines in equal numbers and placed inside an airtight, humidified, modular incubator chamber (Billups-Rothenberg, #MIC-101). ( Cultures grown under 21% O\textsubscript{2} were not placed in an airtight chamber and were instead simply placed in an incubator containing
room air: 21% O$_2$/5% CO$_2$/74% N$_2$) The chamber was then flushed with air containing various oxygen concentrations (1, 5, or 60% O$_2$ with 5% CO$_2$ and balanced by N$_2$ from MedOx, #G1803961, G1811061, and G1806151, respectively) using a Harris two-stage regulator, and the gases were allowed to equilibrate for 1-2 hours before the chambers were flushed with the same gas for a second time to ensure a stable oxygen concentration. Air chambers were placed in a 37°C incubator for 5-6 days to allow for cell growth and colony formation. Following this incubation period, colony numbers were analyzed and compared.

**Cytospin for cell morphology**

Approximately 50,000 cells were suspended in less than 400µL and fixed to coated Shandon Cytoslides® (Thermo Scientific, #5991056) using Single Cytology Funnels (Fisher Scientific, #10-354) and centrifuged in a Cytospin 4 (Thermo Scientific) at 1000rpm for 4 minutes with medium acceleration. Cytopsun slides were allowed to air dry overnight at room temperature. Cells were stained with HEMA-FAST 3 Step Stain (Exaxol Chemical Corp., #SH6427G, #SH6427R, #SH6427B). Cover slips were affixed to the slides with Richard-Allen Scientific Cytoseal™ 60 (Thermo Scientific, #8310-4) to preserve the integrity of the samples. Photos of cell morphology were obtained on an Olympus Provis AX80 microscope at 40x objective magnification using a multi control unit (Olympus, #U-MCB) and a Q-Imaging Retiga-4000R digital camera (Model RET-4000R-F-M-12-C). Images were acquired using the QCapture Plug-In and Adobe Photoshop v6.0 software. Scale bars were determined using a stage micrometer (Graticules LTD) with millimeter gradations (LUMC Core Imaging Facility).
RNA Isolation

RNA from frozen, Round 1, colony assay cells (8.5x10⁴ to 5.82x10⁶ cells) was isolated with TRI Reagent® (Sigma, #T9424). Specifically, the frozen cells were quickly thawed at 37°C and washed with pre-warmed RPMI media to remove DMSO. The washed cells were lysed in 500µL TRI Reagent® for 5 minutes at room temperature. 100µL of chloroform was added to the mixture was shaken, and incubated at room temperature for 10 minutes after which point the solution was centrifuged at 12,000xg and 4°C for 15 minutes to separate mixture into three distinct layers. The upper, clear aqueous layer containing the RNA was transferred to a new tube, being careful not to transfer the precipitated DNA from the adjacent, middle layer. 250µL isopropanol was added to the isolated aqueous layer and incubated at room temperature for 10 minutes. The solution was centrifuged at 12,000xg and 4°C for 10 minutes. The supernatant was discarded leaving behind the precipitated RNA pellet which was washed with 1mL 75% molecular-grade ethanol (Absolute 200 Proof, AAPER Alcohol, #02A02UA) and resuspended by vortexing. After centrifugation at 7,500xg and 4°C for 5 minutes, the supernatant was removed, and the pellet was air dried for ~10 minutes. The RNA pellet was dissolved in RNase-free dH₂O (DEPC-treated dH₂O, de-gassed, and autoclaved). The concentration of the total isolated RNA was measured using the NanoDrop 2000.

Generation of cDNA

The isolated RNA from Round 1 colony assay cells was reverse transcribed into cDNA using a High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, #4368814). 10µL of RNA from each construct (not exceeding 2µg of RNA) was added to reaction mixtures containing 10x RT Buffer, 25x dNTP Mix, and 10x RT Random
Primers with or without (control to detect DNA contamination) MultiScribe™ Reverse Transcriptase according to the supplier’s protocol. The cDNA was generated via RT-PCR in a PCR Express Thermal Cycler using the cycling parameters described in the kit.

**Quantitative RT-PCR**

Quantitative Reverse-Transcriptase PCR—or Real Time PCR—was used to determine gene expression levels of *Hoxa9* bone marrow progenitor cells from Round 1 of the methylcellulose colony assay. The qRT-PCR reaction mixture was prepared using the TaqMan® Gene Expression Assay kit (Applied Biosystems). Each 20µL qRT-PCR reaction mixture was performed in triplicate with 4µL RT-PCR product (either cDNA or the no RT control products), 10µL 2x Gene Expression Master Mix (AB, #4324018), and 1µL of either 20x Gene Expression Assay for *Hprt1* (endogenous control, Mm01545399_m1) or *Hoxa9* (Mm00439364_m1). Triplicate samples (including, non-template controls) were distributed among the wells of a 0.2mL Thin Wall, Non-Skirted, 96-well PCR plate (Dot Scientific Inc., #650-PCR) and sealed with ThermalSeal RT™ Optically Transparent Sealing Film (Dot Scientific Inc., #TSRT2100). The qRT-PCR reaction was run on an Applied Biosystems 7300 Real-Time PCR System using the following cycling parameters: 50°C for 2 minutes, 95°C for 10 minutes, and then 40 cycles alternating between 95°C for 15 seconds and 60°C for 1min. The data was acquired with 7300 System software (Applied Biosystems) and analyzed using the $2^{-\Delta\Delta C_t}$ method\(^{159}\). The relative *Hoxa9* expression levels were normalized to endogenous *Hprt1* expression levels.
**Generation of GST fusion repression domain constructs**

The nucleotide sequences encoding the MLL and MLL1/2 repression domains (RD) were PCR amplified from pMSCV-MLL-AF9 and pMSCV-MLL/MLL2(wt. CXXC + ‘post-CXXC’) -AF9 constructs, respectively, using the cloning primer sequences listed Table 3. The forward primer contains a NruI restriction site which forms a blunt end when digested with NruI and could subsequently be ligated to the EcoRV blunt end of the GST expression vector, pGS-21a (GenScript, #SD0121-10ug). The reverse primers for cloning the RD constructs contain the SalI site that was used previously. The RD sequences were PCR-amplified using a Phusion® high-fidelity DNA polymerase (Finnzymes/NEB, #F530S), were ligated into the pCR-Blunt holding vector, and transformed into OneShot® Top10 chemically competent *E.coli* using the Zero Blunt® PCR Cloning kit described previously. The accuracy of the amplified RDs was verified by sequencing with the M13R primer.

Then, the purified pCR-Blunt-MLL-RD or –MLL1/2 -RD constructs were digested simultaneously with NruI (NEB) and SalI (NEB) in a 50µL reaction volume containing NEB Buffer #3 for 1.5 hours at 37°C. The digested RD fragments were separated by electrophoresis on a 1% agarose gel (20 minutes at 90V) and gel purified. The empty GST expression vector, pGS-21a, was digested with EcoRV (Amersham) and SalI (NEB) in a 50µL reaction volume containing NEB Buffer #3 for 2 hours at 37°C. This digest reaction was heat-inactivated at 65°C for 15 minutes and 2µL of CIP (NEB) was added to prevent self-ligation of the linearized vector which was subsequently purified through gel extraction. The digested RD constructs were ligated with the linearized pGS-21a GST expression vector in a 10µL reaction with T4 DNA ligase.
(Fermentas) for 1 hour at 16°C to generate the wild type repression domain constructs of pGS21a-MLL-RD and pGS21a-MLL1/2-RD (Table 2, J). The inserts encoding the RDs were sequence-verified using the universal T7 terminator primer.

The sequences encoding the mutant MLL2 CXXC domains (A964R, R970P, LRVQ973-6QVPE, V1001K, and R1011M) were swapped into the pGS21a-MLL1/2 repression domain construct in exchange for the wild type nucleotide sequence of MLL2 CXXC domain using the same unique AatII and BamHI restriction sites as were previously used to swap these domains during the cloning of the AF9 fusion constructs. The sequences of each mutant repression domain were verified using the T7 terminator primer.

**Bacterial over-expression of GST fusion MLL- and MLL1/2 repression domain constructs**

The pGS21a repression domain constructs were transformed into chemically competent BL21(DE)pLysS cells (Novagen, #69451-3) to maximize expression of the GST fusion constructs. Briefly, 1µL of plasmid DNA was added to 20µL of thawed BL21(DE)pLysS cells in a 15-mL round bottom culture tube on ice and incubated for 5 minutes. Cells were heat shocked at 42°C for 30 seconds and transferred back to ice for 2 minutes before adding 80µL of room temperature S.O.C. medium and shaking at 37°C and 225rpm for 1 hour. Transformed cells were plated on LB agar plates for ampicillin selection and grown overnight to allow bacterial colonies to form.

A freshly plated colony of BL21(DE)pLysS cells expressing each pGS21a-MLL or –MLL1/2(wt./mt CXXC + ‘post-CXXC’) repression domain construct was picked and used to inoculate 100mL of LB broth with ampicillin (100µg/mL) and chloramphenicol
(15µg/mL, Sigma, #C0378) overnight in a shaker at 225rpm and 37°C. The following morning, the 100-mL cultures were diluted 1:10 by adding 900mL LB Broth with ampicillin but no chloramphenicol. The cultures were grown for about 3 hours until an O.D.₆₀₀ of at least 0.5 was reached (Beckman DU® 530 Life Science UV/Vis Spectrophotometer). Protein production was then induced by adding 1mM IPTG (Research Products Int. Corp., #367-93-1) for 3 hours with continued shaking of the cultures at 225rpm and 37°C. Cultures were then pelleted at 6000xg in a JA-10 rotor of the high capacity Beckman J2-MI centrifuge, and the pellets were frozen at -20°C.

**Purification of GST fusion repression domain constructs**

Frozen bacterial pellets were thawed on ice and resuspended in 60mL Lysis Buffer (50mM Tris, pH 8.0, 300mM NaCl, 1mM EDTA) with protease inhibitor cocktail (1 Roche tablet, #04 693 132 001, per 50mL buffer) and PMSF (1:1000). Lysozyme (1mg/mL) was added directly from powder to the resuspension mix followed by Triton X-100 to a final 1% concentration. The lysis mixture was inverted frequently while incubating on ice for 5 minutes. The lysate was then sonicated two times for 20 seconds while on ice with the Output Control set to 5 (Branson Sonifier 250). The lysate was transferred to centrifugation tubes and spun at 9000rpm and 4°C for 30 minutes in the JA-20 rotor of the Beckman J2-MI centrifuge. The supernatant was transferred to a fresh tube and 500µL of a 50% slurry of swollen and washed Glutathione-agarose beads (Sigma, #CL-4B) was added to the supernatant. The solution was incubated end-over-end at low to medium speed on a Rugged Rotor (Glas-Col®) at 4°C for 2 hours to allow the Glutathione-agarose beads to bind the GST fusion constructs. The beads were then pelleted and washed four times with cold Wash Buffer (lysis buffer including the 1%
Triton X-100, protease inhibitors, and PMSF). Washed beads were resuspended in a volume of this wash buffer to create a 50% slurry of beads. All steps were performed on ice or at 4°C.

Purified GST proteins were detected by SDS-PAGE (4% stacking and 10% resolving acrylamide gel) using the BioRad Mini-PROTEAN® 3-Cell System followed by Coomasie staining or by Western Blot with a rabbit polyclonal anti-GST primary antibody (Abcam, #ab19256) and an ECL™ donkey anti-rabbit secondary antibody conjugated to Horseradish Peroxidase (GE Healthcare, #NA934V). The blot was developed with SuperSignal® West Pico Chemiluminescent Substrate (Thermo Scientific, #34078) and exposed to CL-XPosure™ Film (Thermo Scientific, #34090), or the chemiluminescence was detected using a Fujifilm LAS-3000 imaging system and Image Reader LAS-3000 software. Band intensities were analyzed with Multi Gauge v3.0 software (Fujifilm).

In vitro transcription and translation

HDAC1 (pcDNA3-HDAC1-FLAG Glycerol Stock #562/563), HPC2 (pcDNA3-T7-HPC2, Glycerol Stock #658), and PAF1 (pcDNA3-PAF1, Addgene #11059, Glycerol Stock #1136) proteins were expressed using the T×T® T7 Coupled Reticulocyte Lysate System (Promega, #L4610). Plasmid DNA was purified using the Midi Kit from Qiagen to be used as the template for the IVTT reaction. The reaction components were combined according to the supplier’s protocol using 1.0µg template DNA, the amino acid mixture minus methionine, RNase Inhibitor (Applied Biosystems, #AM2682), and L-[^35]S]-Methionine, aqueous solution (PerkinElmer, #NEG709A500UC) to radioactively
label the IVTT proteins for detection via phosphorimaging. The reaction was incubated at 30°C for 90 minutes.

5µL of each IVTT reaction mixture was electrophoresed on an SDS-polyacrylamide gel (4% stacking, 10% resolving) to compare the reaction efficiencies before the IVTT proteins would be used for protein binding studies. The electrophoresed gel was dried on filter paper (BioRad gel dryer, #583). The radioactive gel was exposed to a blanked Imaging Plate™ (Fujifilm, #BAS-III) for 4 to 12 hours within a Standard BAS Cassette (Fujifilm, #2040). The signal from the Imaging Plate™ was, then, detected with a Typhoon 8600 Variable Mode Imager and Typhoon Scanner Control v1.0 software (Molecular Dynamic Inc., Amersham) using the Storage Phosphor acquisition mode with a pixel size of 50 microns. ImageQuant TL v2005 software (Amersham) was used to analyze signal intensities generated by the $^{35}$S-labeled proteins.

**In vitro protein binding assay**

Equivalent amounts of purified GST repression domain proteins were combined with 20 to 30µL of HDAC1, HPC2, or PAF1 IVTT reaction product depending on the expression efficiency of the $^{35}$S-labeled protein. The binding volume was brought up to 100µL for all reactions by adding Wash Buffer (described above). GST-tagged repression domain proteins and IVTT proteins were incubated together on an end-over-end rotor at low to medium speed at 4°C for 2 hours. Beads were then washed 4 times while on ice with chilled Wash Buffer (1mL per wash). Interacting proteins were eluted from the beads by adding SDS loading buffer to the washed beads and boiling at 95°C for 5-10 minutes. The eluted proteins were separated by SDS-PAGE and bands were detected as described above. Detected bands of interacting $^{35}$S-labeled IVTT proteins
(HDAC1, HPC2, or PAF1) were quantified relative to the GST-MLL-RD binding intensity and normalized to the amount of each GST input.

FLAG-tagged CtBP was overexpressed in 293T cells by transient transfection using the calcium phosphate kit described previously. 48 hours after transfection, cells were collected, pelleted, and snap frozen on dry ice in aliquots of $10^7$ transfected cells, and stored at -80°C. When needed, cell aliquots were thawed on ice and resuspended in 500µL IPH Buffer (50mM Tris, pH 8.0, 150mM NaCl, 5mM EDTA, 0.5% NP-40, and protease inhibitor cocktail 1:200). The cell lysates were incubated on ice for 10 minutes and then centrifuged at >14,000rpm and 4°C for 30 minutes. The supernatant was snap frozen on dry ice and stored at -80°C until used in the *in vitro* binding study. Prior to the binding experiments, equal amounts of each GST-tagged repression domain protein were incubated on ice for 15 minutes with 200µL of Z’Buffer (25mM HEPES, pH 7.5, 12.5mM MgCl₂, 150mM KCl, 0.1% NP-40, 20mM Tris pH 8.0), BSA (1mg/mL), PI cocktail and PMSF (1:1000). 25µL of CtBP-FLAG-containing cell lysate was added to each GST-repression domain protein and incubated end-over-end at low to medium speed at 4°C for 2 hours. Glutathione beads were washed four times with 1mL chilled NETN buffer (100mM NaCl, 1mM EDTA, 0.5% NP-40, 20mM Tris pH 8.0). Bound proteins were eluted from the washed beads by the addition of SDS loading buffer and boiling for 5-10 minutes at 95°C. Proteins were separated by SDS-PAGE and detected by Western Blot using a primary monoclonal mouse anti-FLAG M2 antibody (Sigma, #F3165) and a secondary ECL™ sheep anti-mouse secondary antibody conjugated to Horseradish Peroxidase (Amersham, #NA931V). Band intensities of CtBP were calculated relative to the GST-MLL RD interaction and normalized to the amount of GST inputs.
MTT Assay

The MTT assay was used as a measure of cell proliferation and viability. The MLL-AF9 and MLL(C1188A)-AF9 transformed cell lines had previously been generated in our lab by Laurie Risner, PhD, and were cultured in RPMI/20%FBS/1%P/S with 50ng/mL SCF, 10ng/mL IL-3, 10ng/mL IL-6, and 10ng/mL GM-CSF. The optimal cell numbers for the MTT assay were determined by serially diluting the cells (0 to 400,000 cells) in 100µL of media per well of a 96-well flat-bottom plate in triplicate. MTT reagent, thiazolyl blue tetrazolium bromide (Sigma, #M5655), was added to cells at a final concentration of 0.5 mg/mL and incubated at 37°C for 6 hours after which purple formazan crystals were visible in the cultures. Crystals were dissolved with the addition of 100µL of solubilization solution (10% Triton X-100, 0.1mM HCl, isopropanol to volume). Absorbance of the reaction mix was measured at a wavelength of 550nm on a fluorescent plate reader and analyzed with Omega software. For toxicity assays, 30,000 MLL-AF9 or MLL(C1188A)-AF9 cells were pre-treated with varying doses of hydrogen peroxide (Sigma, #U8879-50TAB), parthenolide (Sigma, #P0667), or spermine nonoate (Sigma, #S150) and incubated overnight in a 96-well plate at 37°C and 5% CO₂. The following morning, MTT reagent was added as above. All MTT reactions were set up using a multi-channel pipet and measured in triplicate.

Statistical Analysis

The data presented represent the averages of multiple experiments with error bars representing one standard deviation unless otherwise noted. Statistical significance, where calculated, was determined by the two-tailed Student’s T-test for independent variables with the threshold for significance set at p < 0.05.
Aim 1: Determine how differences in the MLL CXXC domain and the MLL2 CXXC domain contribute to different transformation capacities of MLL/MLL2-AF9 fusion proteins

Comparison of co-repressor and co-activator protein interactions with MLL and MLL1/2 repression domains

The MLL repression domain (amino acids 1100-1400) has previously been shown to interact with transcriptional co-repressor proteins (HDAC1, HPC2, and CtBP) and co-activator, PAF1, but it is unknown if the closely related homolog, MLL2 (also known as MLL4), interacts with these proteins. A Glutathione-S-transferase (GST) pull-down assay was conducted to determine if these cofactors also interact with residues of the MLL2 repression domain. A GST-tagged construct was generated in which the ‘pre-CXXC’ residues of MLL (amino acids 1100-1151) were fused to the CXXC and ‘post-CXXC’ residues of MLL2 (amino acids 958-1172) which align with the MLL repression domain (Figure 6A). This MLL 1/2 repression domain construct was expressed in bacteria and purified with immobilized glutathione-agarose. Equal amounts of immobilized GST protein were incubated with cofactor proteins which were expressed by in vitro transcription and translation, IVTT, (HDAC1, HPC2, or PAF1) or by transient transfection in 293T cells (CtBP). Following incubation and washing, bound proteins were electrophoresed on a SDS-PAGE gel, and the interactions were detected by phosphorimaging (35S-labeled HDAC1 and HPC2) or Western blot (FLAG-tagged CtBP).
Figure 7. HDAC1 binds to both the MLL repression domain and the equivalent MLL1/2 repression domain.

(A) GST fusion constructs generated for pulldown assays (not drawn to scale). GST—isolated GST domain (negative control); MLL—MLL repression domain (amino acids 1100-1400) fused to GST; MLL 1/2—chimeric repression domain consisting of MLL ‘pre-CXXC’ residues (MLL amino acids 1100-1151) and the MLL2 CXXC and ‘post-CXXC’ residues aligning with the MLL repression domain (MLL2 amino acids 958-1172); mt. MLL 1/2—chimeric MLL 1/2 repression domain with amino acid substitution(s) which restore MLL2 CXXC domain residue to corresponding MLL residue.

(B) Phosphorimaging analysis of 4-5 independent GST pulldown assays showing relative $^{35}$S-labeled HDAC1 binding. (Mean ± 1 S.D.) A representative phosphorimage is shown below along with a Western Blot (bottom) of GST inputs. Detected bands of interacting HDAC1 were set relative to the intensity of the MLL band and normalized to the amount of each input GST protein. (p > 0.05 by Student’s T-test for all interactions compared to MLL1/2 binding, *except GST control.)
Figure 8. HPC2 binds to both the MLL repression domain and the equivalent MLL1/2 repression domain.

(A-C) Binding results for HPC2 and the various repression domain constructs from 3 independent experiments. Results are relative to MLL binding and normalized to input GST protein. (D) Average HPC2 binding results from the three independent experiments (Mean ± 1 S.D., p > 0.05 for all interactions).
Figure 9. CtBP binds to both the MLL repression domain and the equivalent MLL1/2 repression domain.

Average binding intensities from two independent GST pulldown assays showing CtBP-FLAG binding relative to MLL binding and normalized to input GST protein (Mean ± range). A representative Western blot from one binding reaction is shown below along with a Western Blot (bottom) of GST input proteins.
HDAC1, HPC2, and CtBP all bound specifically to the MLL 1/2 repression domain and the wild type MLL repression domain (positive control) with minimal or no binding to GST alone (negative control) (Figures 7-9). HDAC1 and CtBP showed similar levels of binding to the repression domain constructs containing either the wild type MLL or the wild type MLL 2 CXXC domain (Figures 7 and 9). In three independent experiments, however, HPC2 showed increased binding to the MLL 1/2 repression domain construct compared to the MLL repression domain (Figure 8). When the HPC2 binding results were averaged from each experiment, though, the dynamic range of the binding intensities from individual experiments prevented this difference of HPC2 binding to the MLL and MLL1/2 RDs from achieving statistical significance (p > 0.05 by Student’s T-test for all interactions).

**Determination of residues within the CXXC domain that alter co-repressor binding**

I next wanted to determine if residues unique to MLL —when introduced into the MLL2 CXXC domain—would reduce or enhance binding of co-repressor proteins compared to binding of the wild type MLL2 CXXC domain. To test the role of specific residues on protein binding, five constructs were generated in which amino acid substitutions were introduced into the MLL2 CXXC domain to convert this domain into the MLL amino acid sequence at each site: A964R, R970P, LRVQ973-6QVPE, V1001K, and R1011M (Figure 6A). (These specific residues were chosen for these binding assays based on the results of in vitro transformation and Hoxa9 expression shown later in the results: Figure 16) Interaction of the co-repressors with these mutant MLL2 CXXC domain constructs was performed out in the same manner as with the wild type constructs
above. HDAC1, HPC2, and CtBP were still able to interact with each of the hybrid MLL 1/2 constructs (Figures 7-9). However, the substitution of some MLL amino acids into the MLL2 CXXC domain did alter the relative binding intensities of these proteins compared to the wild type MLL or MLL2 CXXC domain-containing constructs. Specifically, the R970P and R1011M substitutions both showed decreased binding of HDAC1 compared to the wild type MLL CXXC domain. Additionally, the V1001K substitution reduced HPC2 binding from the wild type MLL1/2 domain levels toward the lower binding levels of the wild type MLL domain. The differences in binding intensities, however, were not determined to be statistically significant (p < 0.05 by Student’s T-test).

**Comparison of co-activator PAF1 binding to residues of the MLL and MLL2 CXXC domains**

The PAF1 subunit of the polymerase associated factor complex was recently shown to interact directly with the N-terminus of MLL through the residues which flank the CXXC domain in a DNA-independent manner\(^\text{10}\). Muntean et al showed that the interaction of PAF1 with an MLL fusion enhances transcription of leukemogenic target genes like *Hoxa9* and *Meis1*\(^\text{10}\). Simultaneously, work by Milne et al showed that mutation of Arg1153 to alanine within the MLL CXXC significantly decreases binding of PAF1 to the first half of the MLL repression domain (amino acids 1101-1250) and also decreases expression of *Hoxa9*\(^9\). The authors also showed that the residues of MLL2 which align with the first half of the MLL repression domain are unable to bind to PAF1. Taken together, these data suggest there are critical differences between MLL and MLL2
Figure 10. PAF1 binds to both the MLL repression domain and the equivalent MLL1/2 repression domain.

Average phosphorimaging analysis from three independent GST pulldown assays showing relative $^{35}$S-labeled PAF1 binding. A representative phosphorimage is shown below along with a Western Blot (bottom) of GST inputs (Mean ± 1 S.D.). Results are presented relative to MLL binding intensity and normalized to input GST proteins. (p > 0.05 by Student’s T-test for all interactions compared to MLL1/2 binding, *except GST control.)
within the CXXC domain and/or repression domain which affect the interaction of PAF1 and recruitment of transcriptional activating complexes to MLL target genes.

To better understand the role of specific residues within the MLL CXXC domain, PAF1 binding with the MLL, MLL 1/2, and mutant MLL 1/2 repression domain constructs was compared in the same GST pull-down assay as described above. $^{35}$S-labeled PAF1 directly interacted to the same degree with repression domain constructs containing both the MLL and the MLL2 CXXC domains. Introduction of MLL residues (A964R, R970P, LRVQ973-6QVPE, V1001K, or R1011M) into the MLL2 CXXC domain did not enhance or diminish PAF1 binding suggesting that the residues of the MLL and MLL2 CXXC domains do not show differential binding for PAF1 (Figure 10).

**Functional comparison of the capacity of MLL and MLL2 CXXC domain residues to transform bone marrow progenitor cells in vitro**

A previous study by Bach *et al* showed that introduction of both the MLL CXXC domain and ‘post-CXXC’ domain were necessary to confer a transforming capacity to an artificial MLL2-ENL fusion protein suggesting that there are critical residues within this region of MLL that are required for the oncogenic function of a MLL fusion protein. A previous study by Bach *et al* showed that introduction of both the MLL CXXC domain and ‘post-CXXC’ domain were necessary to confer a transforming capacity to an artificial MLL2-ENL fusion protein suggesting that there are critical residues within this region of MLL that are required for the oncogenic function of a MLL fusion protein. A previous study by Bach *et al* showed that introduction of both the MLL CXXC domain and ‘post-CXXC’ domain were necessary to confer a transforming capacity to an artificial MLL2-ENL fusion protein suggesting that there are critical residues within this region of MLL that are required for the oncogenic function of a MLL fusion protein. A previous study by Bach *et al* showed that introduction of both the MLL CXXC domain and ‘post-CXXC’ domain were necessary to confer a transforming capacity to an artificial MLL2-ENL fusion protein suggesting that there are critical residues within this region of MLL that are required for the oncogenic function of a MLL fusion protein.

Sequence alignment of the MLL and MLL2 CXXC domains shows conservation of the structural zinc-coordinating cysteines and the KFGG motif which are necessary for proper domain folding and recognition of non-methylated CpG DNA targets suggesting that other residue within the CXXC domain must account for the observed functional differences (Figure 12A).

In order to understand the critical role of specific amino acids within the CXXC domain, a chimeric fusion protein was generated in which the CXXC domain of MLL2
Figure 11. The MLL2 CXXC domain can transform primary bone marrow cells when swapped into a MLL-AF9 fusion protein.

(A) MLL-AF9 and MLL/MLL2(CXXC)-AF9 fusion protein schematics.
(B) Colony forming units (CFUs) and colony morphology from third plating.
(C) Cell numbers and cell morphology from third plating.
(D) Hoxa9 expression levels relative to MSCV (negative control) & normalized to Hprt expression.

p-values determined by T-test for comparison to MLL levels (**p<0.01, ***p<0.001)
was exchanged for that of MLL in the context of the well-studied MLL-AF9 fusion protein. This synthetic MLL/MLL2(CXXC)-AF9 fusion was able to transform primary hematopoietic cells in an *in vitro* colony assay in which a large number of tight colonies were formed in the third round of plating (Figure 11). These colonies were comprised of blast-like cells similar to those of MLL-AF9 cells. Quantitative RT-PCR showed increased levels of *Hoxa9* compared to the negative control (MSCV-transduced) cells, although the expression levels were lower than that of MLL-AF9 cells (*p* < 0.01 by Student’s T-test). The ability of the MLL2 CXXC domain to transform bone marrow progenitor cells suggests that MLL2 retains some important residues within the CXXC domain that are also present in MLL and are able to contribute to a transforming capacity of the chimeric protein.

In order to be able to study the gain-of-function effects of individual CXXC domain residues, I needed to generate another fusion construct which had lower transforming potency. A chimeric fusion protein was cloned in which the ‘post CXXC’ residues of MLL2 were swapped into MLL-AF9 along with the MLL2 CXXC domain to create a MLL/MLL2-AF9 fusion protein (Figure 12B). In this context, I hypothesized that the wild type MLL/MLL2-AF9 construct—lacking the critical CXXC and ‘post-CXXC’ domain residues from MLL—would be non- or minimally-transforming *in vitro* and would serve as a baseline control. I further hypothesized that introduction of individual amino acid substitutions in this context (13 positions highlighted in blue in the sequence alignment of Figure 12) to restore the MLL sequence in the MLL2 CXXC domain could potentially rescue the MLL-fusion-transforming function (Figure 12B). I
Figure 12. Experimental design of *in vitro* colony assay.

(A) Amino acid sequence alignment of the MLL and MLL2 CXXC domains with conserved, Zn-coordinating cysteines highlighted in yellow and amino acid differences highlighted in blue.

(B) Retroviral fusion constructs used in the *in vitro* colony assay.

(C) Experimental schematic for *in vitro* colony assay.
predicted that those MLL residues which have a critical role in the oncogenic capacity of an MLL fusion protein would rescue transformation capacity from the baseline levels of the wild type MLL/MLL2-AF9 cells in an in vitro methylcellulose colony assay.

**Determination of titers for retroviruses expressing the fusion proteins**

To express the fusion proteins in bone marrow progenitor cells, retroviruses containing the respective fusion constructs in the MSCVneo vector were generated as described in the methods section above. To determine the virus titer of each supernatant, Rat1a cells were infected with serial dilutions of supernatant. After selection with G418, the number of surviving Rat1a colonies was counted as colony forming units (CFUs) to determine viral titers for each fusion construct (Figure 13). The virus titers ranged from 2.5x10^4 to 2.6x10^6 CFUs per milliliter of supernatant (Table 5). In order to normalize the amount of virus used for each transduction, larger supernatant volumes of the low-titer

<table>
<thead>
<tr>
<th>Retroviral Construct</th>
<th>Virus Titer (CFUs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSCV</td>
<td>1 x 10^6</td>
</tr>
<tr>
<td>MLL-AF9</td>
<td>2 x 10^5</td>
</tr>
<tr>
<td>w.t. MLL/MLL2-AF9</td>
<td>8.5 x 10^5</td>
</tr>
<tr>
<td>A964R</td>
<td>1.4 x 10^6</td>
</tr>
<tr>
<td>H968Q</td>
<td>8.0 x 10^5</td>
</tr>
<tr>
<td>R970P</td>
<td>2.6 x 10^6</td>
</tr>
<tr>
<td>V975P</td>
<td>5.2 x 10^4</td>
</tr>
<tr>
<td>LRVQ973-6QVPE</td>
<td>1.3 x 10^6</td>
</tr>
<tr>
<td>V982T</td>
<td>2.5 x 10^4</td>
</tr>
<tr>
<td>P993R</td>
<td>7.5 x 10^5</td>
</tr>
<tr>
<td>T995I</td>
<td>2.4 x 10^5</td>
</tr>
<tr>
<td>V1001K</td>
<td>4.9 x 10^5</td>
</tr>
<tr>
<td>VY1001-2KM</td>
<td>1.2 x 10^6</td>
</tr>
<tr>
<td>R1011M</td>
<td>1.1 x 10^6</td>
</tr>
</tbody>
</table>
Figure 13. Rat1a virus titers of MLL/MLL2-AF9 retroviral fusion constructs.
constructs were used during spinoculation so that an effective infection reaction could take place on the order of $10^5$-$10^6$ CFUs per spinoculation reaction.

**MLL residues introduced into the MLL/MLL2-AF9 fusion protein showed variable ability to rescue transformation capacity in an *in vitro* colony assay**

An *in vitro* colony assay was performed to determine the functional importance of individual residues within the CXXC domain. Bone marrow progenitor cells were isolated from C57Bl/6 mice via c-kit positive selection and were transduced with similar amounts of either wild type or mutant MLL/MLL2-AF9 retroviral constructs. The MSCVneo empty vector retrovirus was used as a negative control, and infection with the well-characterized MLL-AF9 construct served as a positive control. Cells were grown in the presence of G418 for the first round of plating to select for cells infected by retrovirus which express the MSCVneo vector neomycin resistance gene. After the first round of growth, colony and cell numbers were counted and then the cells were serially replated for at least three rounds or until the cultures no longer produced colonies (Figure 12C). The numbers of colonies and cells were quantified in the third round as a measure of self-renewal and cell proliferation capacity, respectively. When non-transformed cells—like those infected with MSCVneo empty vector—were grown in methylcellulose in the presence of growth factors, the bone marrow progenitor cells were stimulated to undergo differentiation. Eventually, the cells lost their proliferative capacity forming fewer and more diffuse colonies in early re-plating and failed to form any colonies by Round 3 (Figures 14 and 15). At the third round of plating, the non-transformed cells showed a differentiated morphology and generated diffuse colonies (Figure 15). In contrast, cells transduced with the highly transforming MLL-AF9 formed large numbers of dense
Figure 14. *In vitro* self-renewal and proliferation capacity of MLL and MLL2 residues in the context of an AF9 fusion oncoprotein.

(A) Colony forming units (CFUs) and (B) Cell numbers from Round 3 of colony assay. (Data represent the mean ± 1 S.D. from at least three independent experiments; p-values determined by T-test for comparison to MLL/MLL2-AF9 levels)

(C) Correlation between CFUs and cell numbers from the control, wild type, and mutant fusion constructs used in the methylcellulose colony assay.
Figure 15. Round 3 colony and cell morphology.
colonies by Round 3 which continued to form colonies through a fifth round. The tight colonies of MLL-AF9 transduced cells had distinct borders and are comprised of blast-like cells with large nuclei and little cytoplasm (Figure 15).

By the third round, the number of colonies formed by cells transduced by wild type MLL/MLL2-AF9 was decreased 7-fold relative to the MLL-AF9 transformed cells, and the total cell number was decreased 10-fold compared to the MLL-AF9 transduced cells (Figure 14). These colonies more closely resembled those transduced with the empty vector negative control with a more diffuse colony morphology and comprised of cells with a differentiated morphology (Figure 15).

The thirteen individual point mutant MLL/MLL2-AF9 constructs showed variable capacities to immortalize c-kit positive progenitor cells \textit{in vitro}. No single amino acid substitution was able to completely restore MLL/MLL2-AF9 transformation capacity back to the levels of MLL-AF9. The measures of transformation capacity—colony formation and cell proliferation—showed a positive correlation when MLL residues were introduced into the MLL2 CXXC domain whereby those residues which showed higher colony forming ability also showed an increase in total cell number (Figure 14C). The amino acid substitutions of A964R, R970P, LRVQ973-6QVPE, and VY1001-2KM showed the highest capacities to rescue transformation (with at least a 40% increase in self-renewal and cell proliferation) compared to the MLL/MLL2-AF9 construct. The substitutions of P993R, T995I, and V1001K showed intermediate (20-40%) rescue of transformation ability while H968Q, V975P, V982T, and R1011M substitutions showed little or no (<20%) capacity to rescue cellular self-renewal or proliferation from baseline
levels of the wild type construct (Figure 14). Colony and cellular morphologies of these mutant constructs also closely correlated with transformation capacities. Those constructs causing the highest self-renewal capacity produced dense, tight colonies while those exhibiting little capacity for self-renewal generated small numbers of primarily diffuse colonies. The dense colonies were comprised of blast-like cells resembling MLL-AF9 transformed cells while the diffuse colonies consisted of mostly differentiated cells with large, vacuolated cytoplasm and relatively small nuclei. The constructs with intermediate transformation capacities produced a mixture of dense and diffuse colonies consisting of both undifferentiated and differentiated cells (Figure 15). These results indicate that there are individual amino acids of MLL that are able to functionally restore the phenotype of the MLL-AF9 fusion protein when introduced into the MLL2 CXXC domain. These specific residues are critical for the transformation capacity of the CXXC domain and the leukemogenic capacity of MLL fusion proteins.

**MLL residues showed variable ability to rescue increased Hoxa9 expression when introduced into the MLL/MLL2-AF9 fusion protein in an in vitro colony assay**

*Hoxa9* is a well-studied target gene of MLL. During normal hematopoiesis, regulated transcription allows *Hoxa9* expression levels to decrease as a cell differentiates toward a more mature lineage and decreases the highly proliferative capacity of the immature precursor cells. In leukemia cells harboring an MLL fusion protein, expression of the *Hoxa9* gene is upregulated contributing to the maintenance of an undifferentiated state and to an unlimited capacity for proliferation and self-renewal of hematopoietic progenitor cells. As an independent test for transformation capacity, transduced cells from the first round of the methylcellulose colony assay were collected to measure *Hoxa9*
expression. RNA from these cells was isolated and reverse transcribed to generate cDNA. Primers for the *Hoxa9* locus were then used to amplify this specific mRNA transcript and levels were quantified via qRT-PCR.

As expected, the bone marrow progenitor cells expressing MLL-AF9 showed high levels of *Hoxa9* expression (17-fold increase) compared to the empty vector MSCVneo (negative control). The wild type MLL/MLL2-AF9 construct expressing the MLL2 CXXC and ‘post-CXXC’ residues showed a 10-fold reduction in *Hoxa9* expression levels compared to MLL-AF9. The thirteen constructs encoding MLL amino acid substitutions within the MLL2 CXXC domain showed variable capacities to restore high *Hoxa9* expression from the baseline levels of wild type MLL/MLL2-AF9 (Figure 16A). The expression levels of *Hoxa9* positively correlated with colony formation and cell proliferation ability in which higher transforming constructs also exhibit higher levels of *Hoxa9* expression (Figure 16B and 15C). However, in contrast to the colony formation and cell proliferation numbers, some specific MLL amino acid substitutions (A964R and LRVQ973-6QVPE) were able to completely restore *Hoxa9* expression to levels observed in MLL-AF9 transduced progenitor cells (p > 0.05 by Student’s T-test compared to MLL-AF9 *Hoxa9* expression levels). The MLL substitutions of R970P, V1001K, and VY1001-2KM showed intermediate (40-60%) rescue of target gene expression compared to the wild type level. The constructs showing weak transformation capacity also showed low levels (<40%) of *Hoxa9* expression compared to MLL-AF9: H968Q, V975P, V982T, P993R, T995I, and R1011M. The ability of specific residues to restore *Hoxa9* expression independently verifies the functional importance of the
**Figure 16. Hoxa9 expression from Round 1 colony assay cells.**

(A) Quantification of *Hoxa9* expression via qRT-PCR performed in triplicate using RNA isolated from colony assay cells from first plating (Round 1). Data was normalized to *Hprt* mRNA levels and set relative to empty vector MSCV (negative control) *Hoxa9* levels (Mean ± 1 S.D; p-values were calculated by Student’s T-test for comparison to MLL/MLL2-AF9 expression levels). Note that MSCV and R1011M values represent a statistically significant lower level of *Hoxa9* expression compared to the MLL-AF9 expression level in contrast to the other constructs which show increased *Hoxa9* expression levels. Also note that A964R and LRVQ973-6QVPE mutants showed complete restoration of *Hoxa9* expression to MLL-AF9 levels (*‡* indicates p > 0.05 by Student’s T-test when comparing expression to MLL-AF9 levels).

(B) Correlation of *Hoxa9* expression levels with (B) colony forming units (CFUs) and (C) cell numbers relative to MLL-AF9.
residues that also rescued colony formation and cell proliferation ability. Those residues which restored Hoxa9 expression levels to MLL-AF9 levels (or close to the same level) are important residues of the MLL CXXC domain contributing to the ability of MLL fusion proteins to transform hematopoietic progenitor cells.

**Aim 2: Determine the role of Cysteine 1188 in the regulation of the MLL CXXC domain binding to non-methylated CpG DNA**

**Treatment of MLL-AF9 and MLL(C1188A)-AF9 cells with agents that can modify thiol groups**

Cysteine 1188 is the only non-zinc-coordinating cysteine residue within the CXXC domain of MLL. This residue is positioned on the DNA-binding surface of the domain in the structure published by Cierpicki et al\textsuperscript{11}. I hypothesized that post-translational modification of the thiol side chain of Cys1188 through oxidation or nitrosylation would inhibit the binding of the CXXC domain to its non-methylated target DNA. Loss of DNA binding by the MLL CXXC domain would lead to decreased transformation capacity by the MLL fusion protein in an \textit{in vitro} colony assay\textsuperscript{11}. In order to test this hypothesis, MLL-AF9-transformed cells were treated with increasing amounts of several agents that are known to oxidize (hydrogen peroxide and parthenolide) or nitrosylate (spermine nonoate) free thiol groups. As a control, a transformed cell line expressing mutant MLL-AF9 fusion protein was used in which the cysteine at residue 1188 was replaced with alanine\textsuperscript{11}. The methyl side chain of alanine is not reactive, so this specific amino acid position would remain unaffected by treatment with thiol-reactive agents. The CXXC domain containing this C1188A mutation was shown by our collaborators to bind to non-methylated CpG DNA with equal affinity to the wild type
domain\textsuperscript{11}. This MLL(C1188A)-AF9 fusion protein was also shown by our lab to transform bone marrow progenitor cells similar to MLL-AF9\textsuperscript{11}. Should oxidation or nitrosylation modify the cysteine residue in a way that interferes with binding of the CXXC domain to DNA, the wild type MLL-AF9 cells should show a decreased transformation capacity compared to the mutant MLL(C1188A)-AF9 cells.

A colorimetric, cellular proliferation assay using MTT reagent was performed to characterize potential the effects of oxidation or nitrosylation between the wild type MLL-AF9 and mutant MLL(C1188A)-AF9 transformed cell lines. In the MTT assay, viable cells are able to internalize the tetrazolium salt reagent (yellow color) and reduce the compound in the mitochondria to a formazan salt product (purple color) in a manner that is dependent on the number of cells and the metabolic activity of the viable cells in the culture. Those cells which are viable and rapidly proliferating will reduce more tetrazolium reagent to formazan to produce a darker purple color and result in an increased absorbance reading at 570 nm. Wild type MLL-AF9 and mutant MLL(C1188A)-AF9 cells were serially diluted (0 to 200,000 cells per well) in a 96-well plate in triplicate to determine relative MTT activity of untreated cells and to determine the optimal number of cells for each treatment assay. Independent of the number of cells, the wild type MLL-AF9 cells showed higher absorbance readings indicative of higher metabolic activity and greater cellular proliferation of the wild type cells compared to the mutant MLL(C1188A)-AF9 cells (Figure 17A). 30,000 cells was determined to be optimal for treatment as this number of cells had an absorbance reading near the middle of the linear portion of the cell titration curve. This allowed for variation in absorbance
Figure 17. Effect of thiol modifying agents on proliferation of MLL-AF9 versus MLL(C1188A)-AF9 transformed cells.

(A) MTT assay absorbance readings for increasing wild type (W.T.) MLL-AF9 and MLL(C1188A)-AF9 cell number as a function of cell viability.

(B) Experimental design for treating cells and measuring cell viability. Relative cell viability of W.T. MLL-AF9 or MLL(C1188A)-AF9 transduced cells following treatment with (C) hydrogen peroxide, (D) parthenolide, (E) spermine nonoate. Cell viability data is set relative to the untreated cell viability of MLL-AF9 or MLL(C1188A)-AF9 transduced cells, respectively (Mean ± 1 S.D.).
values depending whether cells would exhibit increased or decreased proliferation during the treatments. The MTT results are presented as cell viability relative to untreated cells transduced with wild type MLL-AF9 or mutant MLL(C1188A)-AF9.

Cells transduced with wild type (w.t.) MLL-AF9 or MLL(C1188A)-AF9 were treated with increasing amounts of hydrogen peroxide (0-14mM), parthenolide (0-10µM), or spermine nonoate (0-729µM). (The ranges of doses used were based on those reported for other studies in the literature\textsuperscript{161-167}.) Cells were incubated with the compounds in culture medium for 18 hours followed by MTT assay measurement (Figure 17B).

Treatment with all reagents showed a dose-dependent decrease in viable cell numbers in both transformed cell types (Figure 17C-E). The effects of hydrogen peroxide and spermine nonoate on cell viability were similar for both the wild type and the mutant C1188A cells at all doses. This was also true at higher doses of parthenolide treatment (5-10µM). Lower doses of parthenolide (1-4µM) showed a greater decrease in cell viability of the mutant C1188A cells compared to the wild type cells. This was contrary to my initial hypothesis and suggest that a more sensitive assay would be necessary to investigate the potential functional effects of oxidation or nitrosylation on Cys1188.

Next, I performed a methylcellulose colony assay as an independent method for measuring the longer-term effects on transformation capacity of these cells following treatment. The same thiol-modifying agents from above were used to determine if any selective effect on self-renewal or proliferation could be observed. Either wild type MLL-AF9 or mutant MLL(C1188A)-AF9 transduced cells were treated as before for 18
hours in liquid culture with increasing doses of hydrogen peroxide (0-200\(\mu\)M),
parthenolide (0-5\(\mu\)M), or spermine nonoate (0-32\(\mu\)M). After incubation, cells and culture
media containing the compounds were plated in semi-solid methylcellulose. Cells were
incubated for five days and colony and cell numbers were then quantified and compared
(Figure 18A). As with the MTT assay, hydrogen peroxide, parthenolide, and spermine
nonoate treatments showed dose-dependent effects on both the wild type and mutant-
transduced cells with increasing doses showing fewer numbers of colonies and cells
(Figure 18B-D). Parthenolide treatment also showed dose-dependent toxicities for both
cell types. In this longer-term assay, a modest decrease in colony and cell numbers was
observed in the wild type MLL-AF9 transduced cells compared to the mutant
MLL(C1188A)-AF9 transduced cells at intermediate doses of parthenolide (1.5 and
2.5\(\mu\)M) (Figure 18C, note asterisks). At these intermediate doses of parthenolide, the
self-renewal and cell proliferation capacities of wild type MLL-AF9 transduced cells
were more susceptible to parthenolide treatment than the mutant MLL(C1188A)-AF9
cells. This result suggests that the cysteine 1188 residue may have a functional role in
contributing to the colony forming ability of MLL fusion proteins since the only
difference between these two cell types is a cysteine versus alanine at this position within
the CXXC domain. It is possible that Cys1188 is modified by intermediate doses of
parthenolide in a way that alters CXXC DNA-binding affinity ultimately prevents
cellular immortalization.
Figure 18. Effect of thiol modifying agents on self-renewal and proliferation of MLL-AF9 versus MLL(C1188A)-AF9 transformed cells.

(A) Experimental schematic for treating wild type (W.T.) MLL-AF9 or MLL(C1188A)-AF9 cells with thiol-modifying agents. Dose-dependent effects of (B) hydrogen peroxide, (C) parthenolide, or (D) spermine nonoate treatment on colony forming ability—measured as colony forming units, CFUs—(left) and cellular proliferation (right). Results represent the averages of three independent experiments and are set relative to the CFUs or cell numbers for untreated MLL-AF9 or MLL(C1188A)-AF9 transduced cells, respectively. (Mean ± 1 S.D., p-values determined by Student’s T-Test)
Growth of MLL-AF9 and MLL(C1188A)-AF9 transduced cells under conditions of varying oxygen concentration

The modest effects on the colony forming ability of MLL-AF9 transduced cells compared to those transduced with MLL(C1188A)-AF9 following parthenolide treatment suggested that the cysteine 1188 residue of MLL’s CXXC domain may be susceptible to oxidation by parthenolide’s induction ROS which could oxidize the thiol side chain of Cys1188 to abrogate DNA-binding and decrease colony forming potential. To further test whether potential oxidation of cysteine 1188 could alter transformation capacity, the MLL-AF9 and MLL(C1188A)-AF9 transformed cells were grown under conditions of varying oxygen concentration. I hypothesized that lower oxygen concentrations (1-5% O₂) would mimic the hypoxic bone marrow niche in which hematopoietic stem and progenitor cells reside. In this environment, cells undergo less aerobic respiration resulting in fewer metabolic ROS which could potentially oxidize the free thiol group of Cys1188. In contrast, the higher experimental concentrations of oxygen (21-60% O₂) would expose the wild type MLL-AF9 and mutant MLL(C1188A)-AF9 cells to higher levels of ROS, in turn, inducing the potential oxidation of Cysteine 1188 (reversible or irreversible) and alter the transformation capacity of the oncogenic MLL fusion protein through altered DNA-binding affinity. I hypothesized that a hypoxic environment like the bone marrow niche would maintain the free thiol side group of Cys1188 in a reduced state allowing the MLL CXXC domain to bind to DNA while an environment with increased oxygen would increase the potential for thiol oxidation of Cys1188 leading to disruption of DNA-binding and ultimately decreasing the transforming ability of the MLL-AF9 fusion protein.
At all concentrations of oxygen (1-60% O₂), the control MLL(C1188A)-AF9 cells showed a lower number of colonies compared to the wild type MLL-AF9 cells (Figure 19A). This finding corresponded with the initial cell titration of this Aim 2 in which the wild type MLL-AF9 transduced cells showed greater MTT activity (i.e. indicating a greater overall cell viability and/or greater metabolic activity) than the mutant MLL(C1188A)-AF9 transduced cells (Figure 17A). In order to account for this inherent difference, the colony forming units (CFUs) were normalized to growth under room air (21% O₂) conditions (Figure 19B). This normalized data revealed only slight fluctuations in colony forming ability from 1 to 21% O₂ with no statistically significant difference between the wild type and mutant cells. Both transformed cell types showed a marked decrease in numbers and sizes of colonies under conditions of 60% O₂ demonstrating that both cell types are sensitive to hyperoxia, though, the effects were not more pronounced for the wild type MLL-AF9 cells compared to the control MLL(C1188A)-AF9 cells as I had predicted. Given that the treatment of MLL-AF9 transduced cells with thiol modifying agents and growth under varying oxygen concentrations generated only modest results, more sensitive experiments will need to be conducted in the future to ultimately determine whether Cys1188 of MLL may be subject to redox regulation which would allow this residue to function as a molecular switch that modulates DNA-binding affinity of the CXXC domain.
Figure 19. Effects of varying oxygen concentration on colony forming ability of MLL-AF9 versus MLL(C1188A)-AF9 transformed cells.

(A) Effects of varying oxygen concentration on colony forming ability—measured as colony forming units, CFUs—of wild type MLL-AF9 or MLL(C1188A)-AF9 cells.

(B) Colony forming ability (measured as CFUs) set relative to CFUs under 21% oxygen concentration for each respective cell type: W.T. MLL-AF9 and MLL(C1188A)-AF9.

Data represent averages from two to eight independent experiments (Mean ± 1 S.D.); p-values were determined by Student’s T-test comparing CFUs of W.T. MLL-AF9 to MLL(C1188A)-AF9 at each oxygen level (*, **, ***) or comparing CFUs of 21% oxygen to CFUs at other oxygen levels (+, ++, +++).
CHAPTER 5
DISCUSSION

Through a complex set of interactions involving DNA, histones, and other transcription factors, MLL fusion proteins participate in the epigenetic misregulation of HOX genes in leukemias with rearrangements of the MLL gene. The N-terminal portion of MLL—including the CXXC domain and ‘post-CXXC’ residues—is essential for the immortalizing capabilities of MLL fusions. It is not completely understood, though, which specific residues within the CXXC domain of MLL are necessary for leukemogenesis, nor how they mediate this function. The goal of my dissertation project was to identify critical amino acids of the MLL CXXC domain in the context of MLL fusion protein-mediated cell immortalization. From my studies, we could gain a better understanding of the mechanism by which MLL causes leukemia and ultimately identify residues which could be targeted by future therapies for MLL associated leukemias.

I used MLL2, the closest homolog of MLL, as a model to identify functionally critical residues of the CXXC domain within the context of the well-characterized MLL-AF9 fusion protein. Interestingly, MLL2 has not been found in chromosomal translocations associated with leukemia, and the N-terminus of MLL2 cannot functionally substitute for MLL in cell immortalization assays. To study differences between the specific residues of the MLL and MLL2 CXXC domains, I first generated a fusion protein in which the MLL2 CXXC domain was swapped into MLL-AF9 to
generate a synthetic MLL/MLL2(CXXC)-AF9 fusion protein. This chimeric fusion protein—in which only the CXXC domains of MLL and MLL2 were exchanged—was still capable of transforming primary hematopoietic cells. This finding suggested that there is enough homology between these two domains that the MLL and MLL2 CXXC domain can functionally substitute for one another and that there are likely other critical residues outside of MLL’s CXXC domain that are able to compensate for amino acid differences introduced by the MLL2 CXXC domain. The MLL and MLL2 CXXC domains share about 60% homology including the highly conserved zinc-coordinating cysteine residues found in the MLL CXXC domain that are essential for the proper folding and stability of the domain. Even with this degree of homology, though, differences in DNA-binding affinity have been observed by our collaborators at the University of Virginia. The isolated MLL2 CXXC domain shows nearly a four-fold decrease in binding affinity for the same non-methylated CpG dinucleotides as the isolated MLL CXXC domain (Table 6, Bushweller Lab, unpublished data) suggesting that the amino acid differences between the MLL and MLL2 CXXC domains have functional significance.

In order to more effectively study the role of these potentially critical residues within MLL’s CXXC domain, additional amino acids comprising the ‘post-CXXC’ region of MLL2 were introduced into the chimeric MLL/MLL2-AF9 fusion protein along with MLL2’s CXXC domain. The amino acid sequence of the MLL2 ‘post-CXXC’ region aligns with the ‘post-CXXC’ residues of MLL which constitute the repression domain which our lab has previously shown binds to co-repressor proteins HDAC1, CtBP, HPC2, and BMI-1. More recently, it was determined that the MLL CXXC and
post-CXXC domain could also interact with the PAF coactivator complex. With only about 30% homology between MLL and MLL2 in this larger region, it was possible to reduce transformation ability by removing even more MLL residues and was perhaps also likely to alter the functional binding of co-repressor or co-activator proteins that could ultimately contribute to the differential leukemogenic potential of the fusion protein. The transformation capacity of this MLL/MLL2-AF9 fusion protein was now significantly reduced by nearly seven-fold compared to that of MLL-AF9. Using this wild type chimeric protein, which showed only limited transformation capacity in vitro, I now had a baseline from which I could attempt to restore colony forming ability and cellular proliferation through the introduction of individual residues of MLL.

Thirteen amino acid substitutions were introduced into the CXXC domain of the synthetic MLL/MLL2-AF9 fusion which restored the residues to the MLL sequence. In vitro colony assay results showed variable capacities of these residues to rescue colony formation and cellular proliferation compared to MLL-AF9. There was a positive correlation between the ability of a point mutation to restore cellular transformation and to rescue the level of Hoxa9 expression: Those constructs that generated more colonies and cell numbers by the third round of plating in vitro also exhibited higher levels of Hoxa9 expression after only one round of plating in vitro. The relationship between the functional transformation data and the structure of the CXXC domain is shown in Figure 20. The residues in Figure 20 are colored according to their ability to restore (in part or in whole) leukemogenic capacity of the chimeric fusion protein. Residue substitutions that were able to restore function are located on both the DNA-contact surface of the CXXC domain and on the opposite, non-DNA contacting surface. The residue
Figure 20. Structure-function relationship between residues of the CXXC domain

(A) Sequence alignment of the MLL and MLL2 CXXC domain highlighting the substituted residues by their ability to rescue Hoxa9 expression levels within the context of the MLL/MLL2-AF9 fusion protein.

(B) The MLL CXXC domain in four orientations showing amino acids as a space-filled model and colored according to the ability of the residue(s) to rescue Hoxa9 expression levels within the MLL/MLL2-AF9 fusion protein. (DNA is shown as a ball-and-stick model.)
substitutions on the DNA-contact surface (i.e. A964R, P993R, T995I, V1001K, VY1001-2KM) could play a direct role in altering DNA-binding affinity. In contrast, those residues showing altered function that are on the opposite surface (i.e. R970P and LRVQ973-6QVPE) could play an indirect role on DNA-binding affinity through allosteric effects by altering the conformation of the domain or could potentially alter interactions with other proteins through changes in conformation or surface charges.

*In vitro* binding studies were conducted to see if any of these amino acid substitutions within the CXXC domain could alter DNA binding or binding to other cofactor proteins previously shown to interact with the MLL repression domain. Aravinda Kuntimaddi (from the Bushweller laboratory at the University of Virginia) performed preliminary studies on isolated CXXC domain to test DNA-binding affinity (Table 6). The A964R substitution within the MLL2 CXXC domain was able to almost completely restore DNA binding affinity to the level of MLL. This finding along with the functional colony assay data with this substitution confirms the importance of CXXC domain binding to DNA that was previously shown by our labs\(^\text{11}\). In the NMR solution structure of the MLL CXXC domain complexed with DNA, this residue (arginine 1153 in MLL) was not shown to form a direct interaction with the DNA. The flanking residues, however, both contact the phosphate backbone so R1153 of MLL may help to stabilize these interactions to increase DNA-binding affinity\(^\text{11}\).

<table>
<thead>
<tr>
<th>CXXC Domain</th>
<th>Kₐ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>wt MLL</td>
<td>90</td>
</tr>
<tr>
<td>wt MLL2</td>
<td>340</td>
</tr>
<tr>
<td>A964R MLL2</td>
<td>96</td>
</tr>
<tr>
<td>R1011M MLL2</td>
<td>170</td>
</tr>
</tbody>
</table>

Data provided by Aravinda Kuntimaddi & John Bushweller, PhD (University of Virginia)
The DNA binding affinity of the R1011M MLL2 CXXC mutant was also determined by the Bushweller lab. This substitution showed a fifty percent restoration of DNA binding affinity, yet in contrast to the A964R substitution, this partial increase in affinity did not manifest in a rescue of transformation capacity. In fact, the R1011M substitution showed no increase in colony forming ability and a decrease in \textit{Hoxa9} expression compared to the wild type fusion protein. This finding suggests that there may be a minimum DNA-binding affinity of the CXXC domain necessary for conferring the capacity for transformation to the fusion proteins. Additional DNA-binding studies will need to be performed to determine if altered DNA binding accounts for the functional rescue observed for any of the other residue substitutions (i.e. R970P, LRVQ973-6QVPE, or V1001K). I hypothesize that the V1001K substitution, in particular, could increase DNA-binding affinity, as the lysine residue in this position (K1190 of MLL) was previously shown to form an electrostatic interaction with the phosphate backbone of DNA\textsuperscript{11}. Reintroducing the positively charged side group of lysine in this position would likely enhance the ability of the MLL2 CXXC domain to bind to DNA and account for the partial functional restoration of both the V1001K and the VY1001-2KM substitutions.

I performed \textit{in vitro} binding studies between co-repressors or PAF1 and the isolated wild type or mutant repression domains using a subset of the constructs from the functional colony assay studies. The three constructs showing the greatest rescue of function (colony formation, cell proliferation, and \textit{Hoxa9} expression) were chosen—A964R, R970P, and LRVQ973-6QVPE—along with a substitution showing intermediate restorative effects (V1001K). The R1011M substitution was also studied for comparison...
to the wild type MLL1/2 repression domain since this mutant construct showed equally low transformation capacity as the wild type MLL/MLL2-AF9 fusion protein in the methylcellulose colony assay.

The MLL repression domain showed similar binding to co-repressors HDAC1 and CtBP compared to the chimeric MLL1/2 repression domain containing ‘pre-CXXC’ residues from MLL and residues from MLL2 comprising the CXXC domain and ‘post-CXXC’ region. This construct was the equivalent “repression domain” present in the fusion constructs used in the colony assay and was cloned for this in vitro binding study so that I could determine whether altered co-repressor binding could account for differences in any of the functional data observed with the studied chimeric fusions. Each of the five amino acid substitutions that were tested showed relatively similar amount of binding to HDAC1 and CtBP compared to wild type MLL or MLL1/2 repression domain constructs. The minimal reduction in HDAC1 binding that was observed for the R970P construct compared to the MLL1/2 construct (p > 0.05) cannot alone account for the observed functional changes in the colony assay, however, because the R1011M construct (which showed no functional rescue) also exhibited a slight decrease in HDAC1 binding, though, the differences in binding were not statistically significant (p > 0.05). Likewise, the increased CtBP binding by V1001K was comparable to the increase observed with the R1011M mutant protein.

The in vitro binding studies using the repression domain constructs showed enhanced binding of HPC2 to the MLL1/2 repression domain compared to the MLL repression domain. The wild type MLL1/2 construct showed at least a two-fold increase in binding to HPC2 compared to the MLL repression domain; however, when the results
of multiple experiments were averaged together, this binding difference did not achieve statistical significance due to the dynamic range of the band intensities between experiments. The amino acid substitutions observed in the binding studies (A964R, R970P, LRVQ973-6QVPE, V1001K, and R1011M) showed no difference (p > 0.05) in binding when compared to the MLL1/2 repression domain. This finding suggests that HPC2 binding does not account for the functional differences observed for the fusion proteins that were studied in the in vitro transformation assay. Future interaction studies between MLL and other cofactor proteins will need to be performed to account for any functional differences between the mutant constructs which are not explained solely by differences in DNA-binding affinity. Our laboratory has previously shown that the polycomb group protein BMI-1 also interacts with the MLL repression domain. Differential binding of BMI-1 to the mutant repression domain constructs has yet to be explored and could potentially explain the functional differences observed for some of the mutant MLL/MLL2-AF9 constructs. Screening for other proteins which interact with the MLL repression domain may also yield additional candidates for future in vitro binding studies.

Recent studies have implicated a role for the polymerase associated factor complex (PAFc) in MLL leukemia. MLL directly interacts with PAFc through an interaction with the PAF1 subunit in the regions flanking the CXXC domain. Sequential deletion studies determined the minimal interaction domain to include residues of the MLL ‘pre-CXXC’ (residues 1115-1154) and the MLL ‘post-CXXC’ (residues 1209-1299) region which together are found within the repression domain of MLL (functionally defined as MLL residues 1100-1400). PAFc interacts with the MLL-
AF9 fusion protein to help in the recruitment of the fusion and to enhance *Hoxa9* expression. Mutation and deletion studies showed that loss of the PAFc interaction prevented recruitment of MLL or MLL fusions to target gene promoters and for transformation *in vitro*\(^9,10\).

My *in vitro* binding studies confirmed that the MLL repression domain directly interacts with the PAF1 protein subunit of PAFc. Using the chimeric MLL1/2 repression domain construct made up of the analogous repression domain from the minimally transforming fusion protein, I found that PAF1 is able to bind directly in equal amounts compared to the MLL repression domain. It was previously shown that a region of MLL2 comprised of the residues flanking the CXXC domain was unable to bind to PAF1 (presumably residues 900 to roughly 1042 of MLL2; the actual residues are not reported in the publication)\(^9\). The difference in findings might be attributed to the ‘pre-CXXC’ residues which, in my construct, are from MLL and were shown by Muntean *et al.* to be essential for binding PAF1 or perhaps also attributed to the extended ‘post-CXXC’ residues from MLL2 in my constructs which would align with the full repression domain from MLL\(^10\). The structure and function of the MLL and MLL2 ‘post-CXXC’ regions have yet to be determined. Future studies will be necessary to understand the specific role of this region in normal MLL function and in the functioning of leukemogenic fusion proteins.

Of particular interest is a R1153A mutation (R1153 of MLL mutated to alanine) studied by Milne *et al.* in an interaction between PAF1 and a mutant MLL construct comprised of the first half of the MLL repression domain (MLL residues 1101-1250)\(^9\). This is the exact opposite mutation which I had generated within the MLL2 CXXC
domain—A964R—in which an alanine is present in this position in MLL2 and an arginine in MLL. The authors show that introducing this single R1153A mutation into MLL disrupts the interaction of PAF1 even when critical residues flanking the MLL CXXC domain are still present as were shown to be necessary for interaction with PAF1 by Muntean et al.⁹,¹⁰. At the same time, the authors speculate that this mutation does not alter DNA-binding affinity based on the findings of an electrophoretic mobility shift assay (EMSA) using the MLL2 CXXC domain which has an alanine in this position⁷. In contrast to this hypothesis, though, our collaborators have now shown through a quantitative method (NMR titration) that an alanine residue in this position of the CXXC domain does indeed alter DNA binding affinity (Table 6) and the reverse mutation in the MLL2 CXXC domain (A964R) nearly completely restores DNA affinity levels to that of the MLL CXXC domain.

When individual amino acid substitutions from MLL were introduced into the MLL1/2 repression domain, I observed no altered PAF1 binding for any of the isolated GST-tagged constructs (p > 0.05). The absence of a change in binding is likely due to the equal binding of PAF1 to both the wild type MLL and the wild type MLL1/2 repression domains so that introduction of MLL residues into the MLL1/2 construct would not enhance or disrupt binding to any great extent. This finding of unaltered PAF1 binding between the wild type MLL and MLL1/2 repression domains and the mutant repression domains is significant given the previous findings that PAF1 recruitment is necessary for leukemogenesis. This finding demonstrates that the PAF1 interaction with MLL—while it may be necessary—is not sufficient to confer a transforming capacity to an MLL fusion protein.
Furthermore, the findings of PAF1 binding to all repression domain constructs along with the *in vitro* DNA-binding data of our collaborators help to distinguish the importance of PAF complex recruitment from the ability of MLL to bind to its non-methylated DNA targets. The wild type MLL and MLL1/2 repression domain constructs are both able to bind to PAF1, suggesting that the PAF complex would be subsequently recruited to target loci. However, my functional colony assay data showed differences in transformation abilities of the MLL-AF9 and MLL/MLL2-AF9 fusion proteins as the latter formed only minimal colonies comprised of differentiated cells. Significantly, the MLL/MLL2-AF9 transduced bone marrow cells showed only minimal increase in *Hoxa9* expression levels compared to the negative control even though PAF1 is able to bind to the chimeric repression domain and would be expected to enhance gene transcription through recruitment of PAFc, as reported previously. There were significant differences observed in DNA-binding affinities, however, between the wild type MLL and MLL2 CXXC domains and the A964R CXXC domain. The A964R mutation nearly restored DNA binding affinity while showing no alteration of PAF1 binding to the repression domain construct *in vitro* (i.e. no increased binding to PAF1 as might have been expected with this reverse mutation according to the previous report by Milne *et al.*) demonstrating that DNA binding of the CXXC domain mediated by the MLL R1153 residue is of paramount importance for leukemogenesis and recruitment of PAFc may be a subsequent step in the process of transformation. A chromatin immunoprecipitation experiment looking at co-recruitment of PAF and the fusion protein to the *Hoxa9* promoter would further substantiate this conclusion.
A summary of the results from my studies comparing residues of the MLL and MLL2 CXXC domains are shown Table 7. Taken together, the results suggest a model whereby specific residues on both the inner DNA-contact surface and on the opposite, non-DNA-contact surface of the MLL CXXC domain (i.e. R1153 and QVPE1162-5 of MLL) are critical for the transformation capacity of an MLL fusion protein (Figure 21). The arginine 1153 residue of MLL has now been determined by our collaborators to increase DNA-binding affinity when introduced into the MLL2 CXXC domain (A964R substitution). In contrast to the MLL, the CXXC domain of MLL2 shows a lower affinity for non-methylated DNA. This difference can be attributed to the absence of a positive charge at this position in MLL2 (A964 of MLL2 versus R1153 of MLL) which could form an electrostatic interaction with the negatively charged phosphate backbone of the DNA to enhance binding. The role of the critical residues on the outer CXXC domain surface (i.e. P1159 and QVPE1162-5 of MLL) has yet to be determined as these residues

### Table 7. Summary of functional *in vitro* transformation and DNA-binding data.

<table>
<thead>
<tr>
<th>Residue</th>
<th>Transformation</th>
<th>Hoxa9 Expression</th>
<th>DNA binding</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLL</td>
<td>++++</td>
<td>++++</td>
<td>+++</td>
</tr>
<tr>
<td>MLL1/2</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>A964R</td>
<td>++</td>
<td>++++</td>
<td>+++</td>
</tr>
<tr>
<td>H968Q</td>
<td>+</td>
<td>++</td>
<td>ND</td>
</tr>
<tr>
<td>R970P</td>
<td>+++</td>
<td>+++</td>
<td>ND</td>
</tr>
<tr>
<td>V975P</td>
<td>+</td>
<td>++</td>
<td>ND</td>
</tr>
<tr>
<td>LRVQ973-6QVPE</td>
<td>++</td>
<td>++++</td>
<td>ND</td>
</tr>
<tr>
<td>V982T</td>
<td>+</td>
<td>+</td>
<td>ND</td>
</tr>
<tr>
<td>P993R</td>
<td>++</td>
<td>+</td>
<td>ND</td>
</tr>
<tr>
<td>T995I</td>
<td>++</td>
<td>+</td>
<td>ND</td>
</tr>
<tr>
<td>V1001K</td>
<td>++</td>
<td>+++</td>
<td>ND</td>
</tr>
<tr>
<td>VY1001-2KM</td>
<td>++</td>
<td>+++</td>
<td>ND</td>
</tr>
<tr>
<td>R1011M</td>
<td>+</td>
<td>-</td>
<td>++</td>
</tr>
</tbody>
</table>

Symbols indicate approximate 2-3 fold (+), 4-5 fold (++), or >5 fold (+++) increase in functional activity compared to baseline MLL1/2 levels (+), a relative reduction (-) in activity compared to MLL1/2, or data not determined (ND)
Figure 21. Models of the MLL and MLL2 CXXC domains showing differential DNA binding and equivalent binding to cofactor proteins.

(A) Model of the MLL CXXC domain (green space-filling model) showing binding with high affinity to DNA (ball-and-stick model) dependent on R1153 of MLL (A964 in MLL2).

(B) Model of MLL2 CXXC domain (blue space-filling model) showing weaker binding to DNA (ball-and-stick model) through A964 of MLL2 (R1153 in MLL). (Note that both CXXC domains show similar binding to cofactor proteins HDAC1, HPC2, CtBP, and PAF1. A patch of 4 residues (MLL QVPE1162-1165 and MLL2 LRVQ973-6) which showed high functional restoration of colony forming ability and Hoxa9 expression are highlighted in red on the opposite, non-DNA contact surface.)
did not significantly alter binding of cofactor proteins (HDAC1, HPC2, CtBP, and PAF1) in a way that corresponded to the observed transformation capacities. It is possible that DNA-binding studies of the isolated domains may reveal that residue differences on the outer surface may have an impact on DNA binding affinity perhaps through allosteric changes within the CXXC domain. In both models of Figure 21, it is the ability of the CXXC domain to bind to DNA that allows for the recruited PAF complex to enhance transcription of target genes and ultimately contribute to leukemogenesis.

The second aim of my dissertation focused on the role of MLL’s Cys1188, the only non-zinc coordinating cysteine residue within the CXXC domain. This residue is conserved in the CXXC domains of vertebrate MLL orthologs and also in the MLL2 CXXC domain. The proliferation and colony forming ability of MLL-AF9 transduced cells was measured following treatment with various thiol modifying agents. MTT assays and colony formation assays showed dose-dependent toxicities of MLL-AF9 (and control MLL(C1188A)-AF9) cells treated with hydrogen peroxide, parthenolide, and spermine nonoate.

Hydrogen peroxide is a strong oxidizing agent capable of undergoing the Fenton reaction in which reactive oxygen species are generated. These free radicals can, in turn, oxidize free thiol groups in a reversible (i.e. induction of disulfide bond formation) or an irreversible (i.e. sulfonic acid formation) manner. Several reports have shown that growth factor stimulation of hematopoietic cells induces the production of intermediate levels of reactive oxygen species (ROS) that activate signaling pathways to mediate cell cycle progression implicating a role for ROS in normal hematopoiesis and leukemia development\textsuperscript{161, 162, 168}. However, at higher levels like those observed with treatment of
high concentrations of hydrogen peroxide, ROS is damaging to DNA, proteins, and lipids causing an induction of apoptosis. Increases in programmed cell death at high levels of hydrogen peroxide treatment likely accounted for the decreased proliferation and colony forming ability of both the MLL-AF9 and control MLL(C1188A)-AF9 cells in the MTT and colony formation assays.

Parthenolide is a sesquiterpene lactone derived from the feverfew plant \textit{(Tanacetum parthenium)} which has been used for aboriginal medicinal purposes ranging from treatment of headaches to inflammatory disorders\textsuperscript{169}. The lactone ring and epoxide group of parthenolide are susceptible to nucleophilic attack by reactive side chains including sulflydryl groups. Parthenolide has recently gained interest as an anti-cancer compound for its ability to induce formation of intracellular reactive oxygen species and mediate apoptosis through oxidative stress\textsuperscript{163, 170, 171}. This increase in ROS could then cause oxidation of free thiol groups like Cys1188 of the MLL CXXC domain and potentially interfere with DNA binding. Interestingly, parthenolide has also been shown to decrease DNA methylation by DNA methyltransferase 1 (DNMT1) through alkylation of a free thiol side group within the catalytic site\textsuperscript{172}. It is possible that parthenolide could form a similar covalent adduct to MLL’s Cys1188 through a Michael’s reaction which would sterically interfere with the binding of the CXXC domain to its non-methylated DNA targets.

The MTT assays and colony forming assays showed dose-depdendent toxicities of parthenolide on MLL-AF9 and the control MLL(C1188A)-AF9 cells. Unlike hydrogen peroxide and spermine nonoate, however, the MLL-AF9 cells showed a modest sensitivity to parthenolide treatment compared to the mutant control cells at intermediate
treatment concentrations. Since the only difference in these cell lines is the Cys1188 residue, it can be concluded that this non-zinc-coordinating cysteine residue has a minor effect on cell proliferation and colony formation in the context of the in vitro colony assay. The modest affect of parthenolide on Cys1188 of the MLL CXXC domain may be associated with the indirect increase of intracellular ROS or with a direct binding of the lactone ring to the thiol side group. The exact mechanism for this effect was not determined in my colony assay studies. It would be possible to determine if a covalent parthenolide adduct is present through tandem mass spectrometry in which the MLL-AF9 fusion protein could be purified from a large number of treated cells using the fusion protein’s FLAG tag. If a spectral shift corresponding to parthenolide’s molecular weight is observed compared to untreated MLL-AF9 protein, it can be concluded that parthenolide covalently modifies Cys1188 to perturb the transformation capacity of MLL-AF9 cells. In contrast, the lack of a spectral shift from treated MLL-AF9 protein might indicate that the generation of ROS could reversibly oxidize Cys1188 to alter fusion protein function. An in vitro DNA-binding study could also be performed to determine the direct DNA-binding effects of parthenolide treatment on the isolated MLL CXXC domain. Ultimately, the toxicity of parthenolide on MLL-AF9 cells is likely mediated through programmed cell death pathways as previously reported for many AML cell types.165

Treatment of MLL-AF9 and control cells with spermine nonoate showed dose-dependent cellular toxicities for both cell lines and, like hydrogen peroxide treatment, no significant differences in effects were observed between proliferation and colony formation of the either cell type. Spermine nonoate is a nitric oxide (NO) donor that has
previously been shown to have anti-proliferative effects on tumor cells through inhibition of cell cycle progression and induction of cellular necrosis\textsuperscript{166, 167}. Treatment of MLL-AF\textsuperscript{9} wild type and mutant cell lines did not show differential sensitivities to spermine nonoate treatment which was equally toxic at high concentrations. While moderate amounts of NO are known to stimulate the mobilization of hematopoietic cells, higher levels of NO appear to be cytotoxic to leukemia cells in the \textit{in vitro} MTT and colony assays\textsuperscript{45}.

The treatment of MLL-AF\textsuperscript{9} cells with thiol modifying agents and growth of these cells under varying oxygen levels yielded only modest results. It still remains to be more clearly determined whether Cys1188 is susceptible to redox regulation that would allow this residue to function as a molecular switch to modulate binding of the CXXC domain to DNA. Treatment of cells with a wider array of oxidizing agents may confirm the modest effects observed by parthenolide, though, a more sensitive assay would be useful for determining altered effects on Cys1188 as global treatment effects (i.e. oxidation of free thiol groups throughout the cell and induction of oxidative stress) may have masked the specific effects of an individual cysteine residue. An \textit{in vitro} DNA-binding study in which the isolated CXXC domain is treated with possible modifying agents could measure the direct DNA-binding effects of any potential Cys1188 changes. In addition, tandem mass spectrometry could be used to screen for possible modifications of Cys1188 from purified MLL-AF9 fusion proteins. Any spectral shifts associated with the Cys1188 residue could be evaluated to determine what modification(s) may be present at this particular position. If and when it is definitively determined how potential Cys1188 modifications can alter the function of the CXXC domain, the susceptible side chain of
Cys1188 could be targeted with therapies to alter functions of the leukemogenic MLL fusion protein.

Ultimately, the CXXC domain of MLL may be an ideal target for future therapies for MLL-associated leukemias. The requirement of this domain for transformation of hematopoietic cells suggests that disruption of critical interactions within this domain would offer some hope for patients. By comparing the residues of the MLL and MLL2 CXXC domains, critical amino acids were identified on both the DNA contact surface and the opposite, non-DNA contact surface that contribute to leukemogenic capacity. Small molecule inhibitors against the DNA-binding surface of the MLL CXXC domain would be particularly effective if the target includes the critical R1153 residue leading to disruption of DNA binding. In addition, leukemia treatments that target the hypoxic bone marrow niche may also selectively target the leukemic stem cell populations that give rise to the burden of leukemia cells and subsequent relapse of disease. Currently, the therapies for MLL-associated leukemia are limited, resulting in poor clinical outcomes. Knowledge of specific residues within the MLL CXXC domain which are critical to the leukemogenic capacity of MLL fusions may offer improved strategies for treating these aggressive forms of leukemia.
REFERENCES


135. Butler LH, Slany R, Cui X, Cleary ML, Mason DY. The HRX proto-oncogene product is widely expressed in human tissues and localizes to nuclear structures. Blood 1997 May 1;89(9):3361-70.


VITA

The author, Noah Birch, was born to Warren and Kathleen Birch and was raised in Homer Glen, IL. He graduated from Lockport Township High School (Lockport, IL) in 2002 after which Noah attended the Illinois Institute of Technology, IIT (Chicago, IL). While at IIT, he conducted research in the Department of Biological, Chemical, and Physical Sciences for three years. Under Thomas Irving, Ph.D., he developed a protocol for the purification of myosin from *Drosophila melanogastor*. For the next two years, he worked under Hyun-soon (Joy) Chong, Ph.D., synthesizing novel acyclic, macrocyclic, and bimodal ligands for use as MRI contrast agents and as cancer therapeutics. With Dr. Chong’s mentorship, Noah determined that he would apply to a combined MD/PhD program in order to pursue his dual interests in clinical medicine and biomedical research. In his final semester at IIT, he also worked with Jialing Xiang, M.D., Ph.D., to generate adenovirus for inducible studies of the androgen receptor. Noah also worked as a Certified Nursing Assistant in the Emergency Department of Palos Community Hospital (Palos Heights, IL) during his summer and winter breaks throughout college. In May of 2006, Noah graduated from IIT with a Bachelor of Science degree in Molecular Biochemistry and Biophysics, an Applied Music Minor in Percussion and Viola Studies, and a Certificate in Leadership Studies.

Noah entered the MD/PhD program at Loyola University Chicago in July 2006. After completing his first two years of medical school coursework, Noah began his
graduate school work in the Molecular and Cellular Biochemistry program and joined the laboratory of Nancy Zeleznik-Le, Ph.D., within the Hematologic Malignancy Program of Loyola’s Cardinal Bernardin Cancer Center. Under Dr. Zeleznik-Le’s guidance, Noah studied the molecular mechanisms of MLL-associated leukemia and explored novel therapeutics. He presented his work at the National MD/PhD Student Conference, the annual meeting of the American Association for Cancer Research (AACR), and the joint meeting of the American Society for Clinical Investigation (ASCI) and the Association of American Physicians (AAP). During his graduate studies, Noah received competitive awards from the National Institutes of Health’s National Heart, Lung, and Blood Institute (Ruth L. Kirschtein National Research Service Award for Predoctoral MD/PhD Fellows) and the American Society of Hematology (2009 & 2010 Trainee Research Awards). Noah served as co-President of the biomedical Graduate Student Council and is a member of AACR and the American Physician Scientist Association (APSA).

After completing his dissertation work, Noah will complete his training in the S.S.O.M. MD/PhD program with two years of clinical rotations. He will graduate from the MD/PhD program in 2013 when he will continue clinical training during residency. Noah intends to pursue the career path of a physician scientist in which he will study the molecular basis of disease in the laboratory while diagnosing and treating patients in the clinical setting.