Characterizing the Role of L3MBTL3 in Megakaryopoiesis
Using Cell and Chemical Biology Approaches

by

Liubov Krichevsky

A thesis submitted in conformity with the requirements
for the degree of Master of Science
Graduate Department of Medical Biophysics
University of Toronto

© Copyright by Liubov Krichevsky 2014
Characterizing the Role of L3MBTL3 in Megakaryopoiesis Using Cell and Chemical Biology Approaches

Liubov Krichevsky

Master of Science
Medical Biophysics
University of Toronto
2014

Abstract
The histone code regulates a number of processes including transcription, DNA replication, and DNA repair. L3MBTL3 is a putative histone code reader that possesses methyl-lysine binding MBT domains. This study commences with the cellular characterization of UNC1215, the first potent and selective chemical probe targeting a methyl-lysine reader domain, the MBT domains of L3MBTL3. Once established as a potent, selective, and cell-active chemical probe, UNC1215 was used to study the biological function of L3MBTL3. L3MBTL3 was shown to bind histone H3 and to associate with the repressive BHC complex in HEK293 cells and in the erythroid leukemia K562 cell line. Possibly via its association with this complex, L3MBTL3 impairs megakaryopoiesis in a manner dependent on its MBT domains. Additionally, L3MBTL3 may promote erythroid development. In summary, this study utilized cell and chemical biology approaches to obtain insights into the mechanistic function of L3MBTL3, and uncover a role in repressing megakaryopoiesis.
Acknowledgments

First and foremost, I would like to thank my supervisor, Dr. Cheryl Arrowsmith for her continued support over the past several years. She accepted me into her lab as a summer student, co-supervised me during my undergraduate research project, and finally supported me during my Master's degree. I am very grateful for the research opportunities she provided me, as well as for her supervision, guidance, and direction. In addition, I would like to thank my committee members, Dr. Mathieu Lupien and Dr. Linda Penn, for their valuable suggestions and advice.

I would also like to thank past and present members of the Arrowsmith lab for creating a friendly lab environment, as well as for their continuous help, encouragement, and support. In particular, I would like to thank Shili Duan for all her help and support in the lab and Lilia Kaustov for all her encouragement. I would also like to thank members of the SGC, in particular past and present members of the Cell Biology group, including Dalia Barsyte, Tatiana Shatseva, Evelyne Lima Fernandes, Sandy Au and many others for sharing their expertise and for creating an enjoyable research environment. I would like to thank Natalie Nady for her past mentorship and for getting me interested in research when I was an undergraduate student. I would also like to thank Rowena Reyes and Rebecca Clare for administrative support. Finally, I would like to thank our numerous collaborators who made significant contributions to this project.

Last but not least, I would like to thank my family and friends for all their kindness, support, and words of encouragement, when I needed it most.
Table of Contents

Abstract .............................................................................................................. ii
Acknowledgements ........................................................................................ iii
Table of Contents ............................................................................................ iv-v
List of Figures ................................................................................................... vi
List of Abbreviations ....................................................................................... vii-viii

Chapter 1: Introduction ...................................................................................... 1
  1.1 Genetics versus Epigenetics ................................................................... 2
  1.2 Epigenetics: Origin and Definition ....................................................... 2
  1.3 Mechanisms Regulating the Epigenetic Phenotype of a Cell ............... 3
  1.4 Histone Post-Translational Modifications .......................................... 3
  1.5 The Histone Code Hypothesis .............................................................. 4
  1.6 Writers, Readers, and Erasers of the Histone Code ......................... 6
  1.7 The Malignant Brain Tumor Family ................................................... 7
  1.8 Histone Binding Activity of the MBT Family ..................................... 7
  1.9 MBT Proteins Play a Role in Diverse Developmental Processes and
      in Cancers .............................................................................................. 8
  1.10 Lethal(3)Malignant Brain Tumor-like Protein 3 (L3MBTL3) .......... 10
    1.10.1 Histone Binding Activity of L3MBTL3 ........................................ 10
    1.10.2 L3MBTL3 is Homozygously Deleted in Medulloblastoma .......... 10
    1.10.3 L3MBTL3 Plays a Role in Hematopoiesis .................................. 11
  1.11 Chemical Biology .................................................................................. 11
  1.12 Purpose and Rationale of the Thesis .................................................. 12
  1.13 References ........................................................................................... 12

Chapter 2: UNC1215 and Other Cellurally Active L3MBTL3-Specific
Chemical Probes .......................................................................................... 16
  2.1 Summary ............................................................................................... 17
  2.2 Introduction ........................................................................................... 17
  2.3 Results ................................................................................................... 19
    2.3.1 Development and in vitro Characterization of UNC1215 .......... 19
    2.3.2 UNC1215 is a Non-Cytotoxic, Cell Permeable Compound that
        Binds Exogenous Full-Length L3MBTL3 in Cells ......................... 21
    2.3.3 UNC1215 is a Potent Cellurally-Active L3MBTL3 Antagonist .... 23
    2.3.4 Dimerization of 3xMBT in Cells .................................................. 24
    2.3.5 Cellular Characterization of UNC1679 and UNC1967 ............... 25
  2.4 Discussion ................................................................................................ 27
  2.5 Materials and Methods ......................................................................... 28
  2.6 References .............................................................................................. 30
<table>
<thead>
<tr>
<th>Section</th>
<th>Summary</th>
<th>Introduction</th>
<th>Results</th>
<th>Discussion</th>
<th>Materials and Methods</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Summary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.2</td>
<td>Introduction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.3</td>
<td>Results</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.3.1</td>
<td>Flag-L3MBTL3 Co-immunoprecipitates Histone H3 from HEK293 Cells</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.3.2</td>
<td>Chromatin and Cellular Localization of L3MBTL3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.3.3</td>
<td>3xFlag-L3MBTL3 Binds Members of the BHC Complex in HEK293 Cells</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.3.4</td>
<td>L3MBTL3 Knockdown Results in Repression of Notch Pathway Genes in HFF Cells</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.3.5</td>
<td>K562 Cells Possess a Functional Repressive BHC Complex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.3.6</td>
<td>UNC1215-mediated 3xMBT Domain Inhibition does not Alter Expression of BHC Complex Target Genes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.3.7</td>
<td>3xFlag-L3MBTL3 Associates with Members of the BHC Complex in K562 Cells in a Manner Independent of the 3xMBT Domains</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.3.8</td>
<td>3xFlag-L3MBTL3 Impairs PMA-Induced Megakaryocytic Differentiation of K562 cells in a 3xMBT Domain Dependent Manner</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.4</td>
<td>Discussion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.4.1</td>
<td>Binding of L3MBTL3 to Chromatin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.4.2</td>
<td>The Association of L3MBTL3 with the BHC Complex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.4.3</td>
<td>The Mechanistic Role of L3MBTL3 in the BHC Complex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.4.4</td>
<td>L3MBTL3 - an Activator or a Repressor?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.4.5</td>
<td>The Role of L3MBTL3 in Megakaryocytic Differentiation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.4.6</td>
<td>The Role of L3MBTL3 in Erythroid Development</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.4.7</td>
<td>Conclusion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section</th>
<th>Summary</th>
<th>A Growing Understanding of the Mammalian MBT Family of Proteins</th>
<th>Novel Approaches in the Study of the Histone Code</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Thesis Summary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.2</td>
<td>A Growing Understanding of the Mammalian MBT Family of Proteins</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.3</td>
<td>Novel Approaches in the Study of the Histone Code</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.4</td>
<td>References</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
List of Figures

Chapter 1
Figure 1.1 MBT Family of Proteins 9

Chapter 2
Figure 2.1 Domain organization of L3MBTL3 18
Figure 2.2 UNC1215 binds the 3xMBT domains of L3MBTL3 in vitro 20
Figure 2.3 Mero-1215 co-localizes with GFP-L3MBTL3 foci in HEK293 cells 22
Figure 2.4 Biotinylated UNC1215 affinity purifies Flag-L3MBTL3 from cells 23
Figure 2.5 UNC1215 potently increases the cellular mobility of GFP-3xMBT in cells 24
Figure 2.6 Exogenous 3xMBT dimerizes in HEK293 cells regardless of the presence of UNC1215 25
Figure 2.7 UNC1679 and UNC1967 are cellularly-active L3MBTL3 probes 26

Chapter 3
Figure 3.1 Summary of hematopoietic abnormalities seen in L3MBTL3 knockout mice 34
Figure 3.2 L3MBTL3 associates with members of the BHC complex 35
Figure 3.3 L3MBTL3 binds histone H3 37
Figure 3.4 GFP fusions of various L3MBTL3 constructs display differential localization on chromatin 38
Figure 3.5 L3MBTL3 associates with members of the BHC complex 39
Figure 3.6 L3MBTL3 knockdown results in repression of Notch pathway genes 41
Figure 3.7 The erythroid leukemia K562 cell line possesses a functional BHC complex 43
Figure 3.8 3xFlag-L3MBTL3 associates with LSD1 and CoREST in K562 cells, in the presence or absence of PMA-induced megakaryocytic differentiation, independently of its 3xMBT domains 45
Figure 3.9 L3MBTL3 impairs PMA-induced megakaryocytic differentiation of K562 cells in a partially 3xMBT domain-dependent manner 47-48
Figure 3.10 3xFlag-L3MBTL3 may promote hemin-induced erythroid development of K562 cells 54
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3xMBT</td>
<td>Three Malignant Brain Tumor domains of L3MBTL3</td>
</tr>
<tr>
<td>53BP1</td>
<td>p53 binding protein 1</td>
</tr>
<tr>
<td>BHC</td>
<td>BRAF-HDAC</td>
</tr>
<tr>
<td>Bromo</td>
<td>Brahma organization modifier</td>
</tr>
<tr>
<td>BPTF</td>
<td>Bromodomain PHD Finger Transcription Factor</td>
</tr>
<tr>
<td>Chromo</td>
<td>Chromatin organization modifier</td>
</tr>
<tr>
<td>CD10</td>
<td>Cluster of differentiation 10</td>
</tr>
<tr>
<td>CD41</td>
<td>Cluster of differentiation 41</td>
</tr>
<tr>
<td>CD61</td>
<td>Cluster of differentiation 61</td>
</tr>
<tr>
<td>CD71</td>
<td>Cluster of differentiation 71</td>
</tr>
<tr>
<td>CMP</td>
<td>Common myeloid progenitor</td>
</tr>
<tr>
<td>CoREST</td>
<td>REST Corepressor 1</td>
</tr>
<tr>
<td>CRISPR</td>
<td>Clustered regularly interspaced short palindromic repeats</td>
</tr>
<tr>
<td>EMT</td>
<td>Epithelial to mesenchymal transition</td>
</tr>
<tr>
<td>FL</td>
<td>Full-length</td>
</tr>
<tr>
<td>FP</td>
<td>Fluorescence polarization</td>
</tr>
<tr>
<td>FRAP</td>
<td>Fluorescence recovery after photobleaching</td>
</tr>
<tr>
<td>GFI1b</td>
<td>Growth Factor Independent 1B</td>
</tr>
<tr>
<td>GFP</td>
<td>Green fluorescent protein</td>
</tr>
<tr>
<td>GMP</td>
<td>Granulocyte/macrophage progenitor</td>
</tr>
<tr>
<td>HBA</td>
<td>Hemoglobin alpha</td>
</tr>
<tr>
<td>HBG</td>
<td>Hemoglobin subunit gamma</td>
</tr>
<tr>
<td>HDAC1/2</td>
<td>Histone deacetylase 1/2</td>
</tr>
<tr>
<td>HMG20B</td>
<td>High mobility group 20B</td>
</tr>
<tr>
<td>HP1</td>
<td>Heterochromatin protein 1</td>
</tr>
<tr>
<td>IP/MS</td>
<td>Immunoprecipitation/mass spectrometry</td>
</tr>
<tr>
<td>ITC</td>
<td>Isothermal calorimetry</td>
</tr>
<tr>
<td>Kme1</td>
<td>Monomethylated lysine</td>
</tr>
<tr>
<td>Kme2</td>
<td>Dimethylated lysine</td>
</tr>
<tr>
<td>L(3)MBTL3</td>
<td>Lethal(3)Malignant Brain Tumor-like Protein 3</td>
</tr>
<tr>
<td>LSD1</td>
<td>Lysine-specific histone demethylase 1A</td>
</tr>
<tr>
<td>MBT</td>
<td>Malignant Brain Tumor</td>
</tr>
<tr>
<td>Meis1</td>
<td>Myeloid Ecotropic Viral Integration Site 1</td>
</tr>
<tr>
<td>MEP</td>
<td>Megakaryocyte/erythroid progenitor</td>
</tr>
<tr>
<td>Mero-1215</td>
<td>Merocyanine76 dye, conjugated to UNC1215</td>
</tr>
<tr>
<td>MNase</td>
<td>Micrococcal nuclease</td>
</tr>
<tr>
<td>PcG</td>
<td>Polycomb group</td>
</tr>
<tr>
<td>PHD</td>
<td>Plant Homeo Domain</td>
</tr>
<tr>
<td>PHF21A</td>
<td>PHD finger protein 21A</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>PMA</td>
<td>Phorbol 12-myristate 13-acetate</td>
</tr>
<tr>
<td>PRC1</td>
<td>Polycomb Repressive Complex 1</td>
</tr>
<tr>
<td>PRC2</td>
<td>Polycomb Repressive Complex 2</td>
</tr>
<tr>
<td>PTM</td>
<td>Post translational modification</td>
</tr>
<tr>
<td>RBPJ</td>
<td>Recombination signal binding protein for immunoglobulin kappa J region</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Real-time polymerase chain reaction</td>
</tr>
<tr>
<td>SAM</td>
<td>Sterile Alpha Motif</td>
</tr>
<tr>
<td>SFMBT1/2</td>
<td>Scm-Like With Four MBT Domains Protein 1/2</td>
</tr>
<tr>
<td>SPI1</td>
<td>Spleen focus forming virus (SFFV) proviral integration oncogene</td>
</tr>
<tr>
<td>UHRF1</td>
<td>Ubiquitin-like with PHD and ring finger domains 1</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction
1.1 Genetics versus Epigenetics

The 20th century was marked by a number of crucial discoveries in the fields of developmental biology and the study of heredity. Early in the 20th century, work done by Wilson, Boveri, and Sutton strongly suggested that chromosomes carried inheritance and developmental program information. A major breakthrough came in the 1940s and the 1950s through the work of Avery, MacLeod, and McCarty (1944) and Hershey and Chase (1952) who demonstrated that DNA, not protein, was the carrier of genetic information (1).

However, it was not long before scientists realized that some phenomena could not be explained solely by genetic inheritance. One such phenomenon is the process of differentiation and the maintenance of a mature and specialized cellular phenotype from one cell division to the next (2). Nuclear transplant experiments from differentiated cells into enucleated oocytes conducted by Laskey and Gurdon (1970) strongly suggested that somatic cells retain all the DNA required to direct development, indicating that the process of differentiation was not a result of alterations or deletions in the DNA sequence; therefore, something beyond the DNA sequence itself contributed to differentiation (1,3). Other examples of such unexplained phenomena include X chromosome inactivation and position effect variegation (1-2). These phenomena, which could not be explained on the basis of genetics alone, are today explained by the science of Epigenetics, which literally means “above genetics” (2,4).

1.2 Epigenetics: Origin and Definition

The term "epigenetics" was originally coined by Conrad Waddington in 1942 in an attempt to bridge the areas of genetics and developmental biology, which at the time were treated as distinct areas (2,5). For Waddington, epigenetics incorporated all the processes that controlled the development of a zygote into a mature organism (1). Over the years, as our understanding of the various molecular processes involved in development has evolved, the term "epigenetics" has undergone a number of revisions in an attempt to account for bizarre phenomena that could not be explained by genetics alone (1-2). Today, epigenetics is defined as "the study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence" (6).
1.3 Mechanisms Regulating the Epigenetic Phenotype of a Cell

Today, a number of mechanisms have been proposed to regulate and establish our epigenomes. The earliest proposed mechanism, that of DNA methylation, was put forward as far back as 1969 by Griffith and Mahler (7). Soon after in 1975, two separate studies proposed that DNA methylation plays a role in the silencing of X-chromosomes (8) and in differentiation (9), as well as that DNA methylation patterns are heritable - a key concept in epigenetics (8-9). Other mechanisms that are known to regulate cellular epigenomes include the post-translational modification of histones (the major focus of this thesis), introduction of histone variants, ATP-dependent chromatin remodeling, and non-coding RNAs (10).

1.4 Histone Post-Translational Modifications

Eukaryotic nuclei contain 2 meters of tightly folded DNA. DNA does not exist on its own. Rather, it is found in association with basic proteins called histones, consisting of the four core histones H2A, H2B, H3, and H4, as well as the linker histone H1. An H3-H4 tetramer associates with two H2A-H2B dimers to form an octamer. Subsequently, ~146 base pairs of DNA wrap around the octamer to form the nucleosome, the basic building block of chromatin. Nucleosomes are separated by 10-60 base pairs of linker DNA that is bound by histone H1, forming a 10nm chromatin fiber often referred to as having a "beads-on-a-string" configuration. These 10nm fibers are heavily folded and condensed to eventually form 100-400nm interphase chromatin (11,12).

Each of the core histones has a globular domain, as well as an N-terminal histone tail of 20-35 residues in length that protrudes out of the nucleosome (12). Residues found in the histone tails (and to a smaller extent those found in the globular part of the histones), are heavily modified by a number of post-translational modifications (PTMs). These modifications include acetylation of lysine residues, phosphorylation of serine and threonine residues, methylation of lysine and arginine residues, as well as ADP-ribosylation, ubiquitination, sumoylation, neddylation, biotinylation, proline isomerization, and others (4). The correlation of these modifications to transcriptional regulation is not a new idea. In fact, a link between histone acetylation and transcriptional activation has been made as early as 1964(13). Strong evidence for this
connection came in 1996 with the demonstration that a histone acetyltransferase isolated from *Tetrahymena* was homologous to a previously identified putative yeast transcriptional co-activator protein (14).

Although the link between histone post-translational modifications and transcriptional regulation was made, the mechanism connecting them was still unclear. Two different, but non-exclusive hypotheses have been proposed to explain how histone PTMs affect chromatin structure and downstream processes. The first hypothesis addresses the direct effect that histone modifications have on chromatin structure and on the strength of histone-histone and histone-DNA contacts (15). A commonly cited example is that acetylation of lysine residues of histones neutralizes their positive charge, thus weakening the interaction between histones and negatively charged DNA. This results in localized euchromatin (loose or unpacked chromatin) formation which can be accessed by DNA transcription machinery (4,15-16). Indeed, acetylation is often found in actively transcribed areas. Similarly, phosphorylation of serine and threonine residues introduces a net negative charge to histones, thus loosening nucleosomal packaging (4,16-17). Indeed, serine phosphorylation is seen at genes that are the earliest to be induced in response to mitogenic stimulation. Finally, some bulky modifications, such as ubiquitin, can disrupt overall chromatin folding (4).

While appealing, this hypothesis left a number of unanswered questions. Firstly, it did not explain the ability of small, non-bulky modifications, such as methylation, which do not change the charge of histones to elicit changes in chromatin structure. Secondly, the hypothesis failed to address the fact that serine phosphorylation is not always found in euchromatic areas; in fact, it is required for chromatin compaction during mitosis. The second hypothesis, referred to as the histone code hypothesis, provides an explanation that takes these issues into account (18-19).

**1.5 The Histone Code Hypothesis**

The histone code hypothesis proposes that post-translational modifications of histones form surfaces that can be bound or “read” by so-called “reader proteins.” These reader proteins are often found in multi-protein complexes and thus, may aid in the recruitment of other chromatin-modifying enzymes (e.g., acetyltransferases, deacetylases, methyltransferases, etc.), chromatin-
remodeling complexes (e.g., SWI/SNF), and/or DNA-modifying enzymes. In this manner, the recognition of a histone mark (or a combination of marks) by a reader protein will lead to alterations in chromatin structure and regulation of downstream processes, such as transcription (18-19).

An essential element of the histone code hypothesis is that PTMs can act sequentially or in combination to affect downstream processes. Thus, a particular PTM may have distinct effects depending on other PTMs found in the area (18-19). In support of this hypothesis, several proteins have been shown to have multiple reader domains. A recent example of this phenomenon is the BPTF (Bromodomain PHD Finger Transcription Factor) protein that has two PHD fingers and one bromodomain that is in close proximity to the second PHD finger. Bromodomains are readers of the acetyl-lysine mark while PHD fingers are capable of recognizing methylated lysines. BPTF was shown to bind mononucleosomes bearing a combination of the H3K4me3 (trimethylated lysine 4 of histone H3) and H4K16ac (acetylated lysine 16 of histone H4) marks more effectively than to nucleosomes bearing either mark in isolation (20).

In this manner, the histone code hypothesis accounts for the fact that a particular modification, such as methylation, can lead to multiple effects. For example, methylated lysines at positions K4 and K9 of histone H3 are recognized by distinct reader proteins which are part of different multiprotein complexes, resulting in distinct downstream effects on chromatin structure. Similarly, different degrees of lysine methylation at a particular position can be recognized by distinct readers, leading to distinct downstream effects (21-22). Alternatively, as explained with the BPTF example, the binding of a reader to a single PTM can be affected by the presence or absence of other neighbouring PTMs, meaning that a particular PTM can have several distinct effects depending on its environment (20). In a similar manner, the histone code hypothesis explains why in certain cases, phosphorylation can lead to chromatin compaction even though when charge alone is considered, they are predicted to lead to euchromatin formation. Thus, although there are some clear patterns as with acetylation leading to euchromatin formation, there is no universal correlation between a certain histone modification and chromatin state; the specific position of the modification and the presence of neighbouring PTMs are critical to the outcome (18-19).
Another attractive aspect of the histone code hypothesis is that in some cases, it helps to mechanistically explain a fundamental requirement of epigenetics - heritability through meiosis and/or mitosis. The heritability of the histone code is still a controversial aspect though several models have been proposed. For example, heterochromatin protein 1 (HP1), the reader of the H3K9me3 mark, binds to SUV39H1, the methyltransferase enzyme responsible for depositing the H3K9me3 mark, suggesting a mechanism by which the mark can spread across chromatin. Similarly, it serves as a model for the mitotic inheritance of the mark: HP1 "reads" the H3K9me3 mark and at the same time binds SUV39H1, bringing it into proximity with newly deposited histones on a newly transcribed DNA strand (23). A similar model has been described for the propagation of the H3K27me3 mark which is maintained by the Polycomb Repressive Complex 2 (PRC2) (24-25).

### 1.6 Writers, Readers, and Erasers of the Histone Code

The histone code is established by “writer” enzymes that deposit particular modifications onto histone residues. These include acetyltransferases, arginine and lysine-methyltransferases, and kinases, among others. Although once considered to be static, these marks can be removed by “eraser” enzymes, including deacetylases, demethylases, and phosphatases, respectively (4,18-19).

Readers of the histone code possess various structural domains that allow them to recognize specific modifications. For example, bromodomains, the first identified reader domains (26) bind acetylated lysine residues. In contrast to acetylation, methylation of lysine residues exists in three different states (mono-, di-, or tri-methylation), which form chemically and physically distinct interaction surfaces; therefore, a number of domains have evolved to recognize the different degrees of lysine methylation (21).

Chromodomains (chromatin organization modifier) of HP1 were the first domains identified to bind methylated lysine residues (23). Other methyl-lysine binding modules include PHD and WD40 domains (21). Additionally, through analysis of sequence homology to the chromo domain, several other methyl-lysine binding motifs were identified, including Tudor, PWWP, and
MBT (Malignant Brain Tumor) domains. Collectively, these four sets of domains are referred to as the Royal Family (27).

1.7 The Malignant Brain Tumor Family

The Malignant Brain Tumor (MBT) domain was originally identified as a motif consisting of 99-103 amino acids in the *Drosophila lethal(3)malignant brain tumor, dl(3)mbt*, gene. The name originated from the fact that temperature-sensitive *l(3)mbt Drosophila* mutants developed malignant transformations of optic neuroblasts (28). Over the next several years, the MBT domain was identified in two more *Drosophila* genes, *dScm* and *dSfmbt*. *dScm* is a Polycomb Group (PcG) protein since it is required for the regulation of homeotic gene expression during development. The MBT domain was also identified in 9 mammalian genes which are homologs of the 3 *Drosophila* genes. All MBT family members contain 2-4 repeats of the MBT domain, found in tandem. Most also have a C-terminal hetero- or homo-dimerization SAM domain, as well as zinc fingers (29).

1.8 Histone Binding Activity of the MBT Family

The *Drosophila* MBT protein mutants tied the MBT domains to both tumor suppression and development, but their function remained unknown until analysis of sequence homology indicated that MBT domains are evolutionary related to other methyl-lysine binding domains, such as the Tudor and Chromo domains (27). Thus, MBT domain proteins were hypothesized to act as histone code readers. Indeed, pulldown, fluorescence polarization, and various peptide array assays indicate that at least in vitro, MBT domains bind mono- and di-methylated lysine residues on histone peptides (30-31).

The crystal structures of MBT domains of several MBT proteins have been solved. They indicate that each domain consists of an N-terminal α-helical arm and a C-terminal β-barrel globular core. An arm of one MBT domain packs against the β-subunit of another MBT repeat. Thus, the set of MBT domains is required for proper folding. Interestingly, crystal structures indicate that only a
single MBT repeat is responsible for accommodating the methylated histone peptide although the entire set of MBT domains is required for proper folding of the repeats (29,32-34).

Crystal structures of MBT domains in complex with methylated histone peptides explain the discrimination of MBT domains against unmethylated and tri-methylated substrates. Methyl-lysine accommodating MBT repeats contain an "aromatic cage" which consists of three aromatic residues situated approximately perpendicularly to each other, forming three sides of the cage. In addition, the binding pocket consists of a cysteine residue at the bottom, as well as a conserved aspartic acid residue. The aromatic residues form cation-π interactions with the quaternary ammonium of the methylated lysine residue, thus stabilizing the structure. In addition, a highly conserved aspartic acid residue forms a hydrogen bond with the methyl-ammonium proton. This interaction cannot take place with the tri-methylated lysine due to the absence of the methyl-ammonium proton. In addition, the entry channel of the cage is narrow, thus sterically hindering the ability of the tri-methyl lysine to enter. Finally, because the surface of the pocket consists of hydrophobic amino acids, interaction with methylated lysine residues is preferred over the unmethylated lysine residue (29,32-34). Thus, the methylated lysine side chain forms extensive contacts with the MBT pocket. In contrast, histone residues adjacent to the lysine residue do not appear to interact significantly with the MBT domains, which explains why in vitro MBT proteins have been shown to interact with methylated lysine residues at multiple positions with comparable affinities and an apparent lack of sequence specificity (29,33).

1.9 MBT Proteins Play a Role in Diverse Developmental Processes and in Cancers

Several studies have been published describing the distinct roles that MBT domain proteins play in development, differentiation, and disease. For example, L3MBTL2, L3MBTL3, and SCML2 have been found to be homozygously deleted in a few cases of medulloblastoma (35) while L3MBTL4 may act as a tumor suppressor in breast cancer (36). MBTD1 plays a role in hematopoiesis and skeletal development (37). Recently, SFMBT1 was shown to play a role in spermatogenesis, myogenesis, and epithelial-to-mesenchymal transition (EMT) (38-40). Finally, L3MBTL3 plays a role in hematopoiesis, in particular in the myeloid lineage, via an unknown
mechanism (41). Several MBT proteins, including L3MBTL3, MBTD1, and L3MBTL2, are required during murine embryonic development (37,41-42) while others, such as L3MBTL1 appear to be dispensable with knockout mice surviving with no abnormalities (43), perhaps due to compensation by other MBT proteins. Finally, biochemical characterization of MBT proteins has indicated that they are found in distinct multi-protein chromatin-binding complexes (38-40,42,44). Figure 1.1 summarizes the various roles and complex associations of the MBT proteins.

Figure 1.1. MBT family of proteins. Mammals have 9 MBT domain containing proteins. MBT proteins play distinct roles in development (blue) and cancers (red), and associate with various chromatin-binding complexes (green). Adapted from: (http://apps.thesgc.org/resources/phylogenetic_trees/index.php?domain=MBT#options)
1.10 Lethal(3)Malignant Brain Tumor-like Protein 3 (L3MBTL3)

In this study, we chose to study the MBT protein, L3MBTL3. It is homologous to Drosophila's tumor suppressor \textit{l(3)mbt}. L3MBTL3 has three MBT domains, a C-terminal SAM domain, two zinc fingers of unknown function, and two nuclear localization signals (45).

1.10.1 Histone Binding Activity of L3MBTL3

The histone binding ability of L3MBTL3 has been studied \textit{in vitro} using peptide arrays and fluorescence polarization techniques. In the peptide array assay, modified histones are immobilized on a membrane and the membranes are then incubated with a tagged set of MBT domains, followed by Western blot analysis against the tag. Such experiments indicated that, compared to other MBT proteins, the MBT domains of L3MBTL3 display highly promiscuous binding activity and interact with a large number of modified peptides with no apparent sequence specificity. In contrast, fluorescence polarization experiments indicated a preference of the MBT domains of L3MBTL3 towards binding di-methylated peptides over mono-methylated ones and complete lack of binding to tri-methylated substrates. A slightly tighter binding affinity was reported towards the H4K20me2 mark over the H3K4me2 and H3K9me2 marks (79\,\mu M, 193\,\mu M, and, 356\,\mu M respectively) (31).

1.10.2 L3MBTL3 is Homozygously Deleted in Medulloblastoma

A 2009 study implicated L3MBTL3 in medulloblastoma, a malignant pediatric brain tumor. In this study, genomes of 212 medulloblastoma tumors were screened to find areas of genomic loss and gain. Interestingly, several abnormalities were found to center on epigenetic regulators, including histone lysine demethylases (JMJD2B, JMJD2B), histone lysine methyltransferases (EHMT1, SMYD4), and putative MBT domain methyl-lysine histone code readers (L3MBTL2, L3MBTL3, and SCML2). Although each individual abnormality occurred in a comparatively small percentage of tumors, this study suggested that misregulation of the epigenome plays a significant role in the development of medulloblastoma. In particular, L3MBTL3 was found to be homozygously deleted in one medulloblastoma patient. Further mechanistic studies indicated
that when L3MBTL3 was transfected into the DAOY medulloblastoma cell line, which has a homozygous deletion of L3MBTL3, the cells became growth-inhibited, thus implicating L3MBTL3 as a potential tumor suppressor in medulloblastoma (35).

1.10.3 L3MBTL3 Plays a Role in Hematopoiesis

A 2005 knockout mice study indicated that L3MBTL3 plays a role in definitive erythropoiesis. L3MBTL3 knockout mice die at a late embryonic stage (between E17.5 and E19.5) from anemia. In the absence of L3MBTL3, maturational progression down the myeloid lineage was impaired as seen by decreased levels of mature erythrocytes, granulocytes, and macrophages and increased levels of corresponding upstream progenitors. In contrast, the lymphoid lineage was largely unaffected. The development of other organs showed no detectable abnormalities. Thus, it was concluded that L3MBTL3 promotes maturational progression down the myeloid lineage (41).

Immunoprecipitation/mass spectrometry results obtained by our collaborators indicate that L3MBTL3 associates with the transcriptionally repressive BRAF-HDAC (BHC) complex. The BHC complex, in association with the transcription factor GFI1b, has been reported to promote the development of the erythroid and megakaryocytic lineages (46).

1.11 Chemical Biology

Chemical biology has emerged as a novel approach in the study of epigenetics. Chemical biology involves the use of chemical tools to aid the study of a particular biological system. Potent, specific, and cellularly-active chemical probes that target a particular protein or a particular protein domain allow for further in-depth characterization of the target. Prior to use, the chemical probe must be thoroughly characterized to demonstrate in vitro potency and specificity, identify its mechanism of action, and investigate its stability and activity in the cellular context (47). Once these parameters of a chemical probe are established, it can be faithfully used to investigate the biological role of the target. A recent example of a chemical probe against an epigenetic protein is JQ1, which targets BET bromodomains. The use of JQ1 has resulted in a much deeper understanding of the transcriptional program regulated by BET bromodomain proteins, such as
regulation of the c-myc oncogene (48-49). Thus, the development of specific and potent chemical probes is instrumental for understanding the biology and function of the target protein.

1.12 Purpose and Rationale of the Thesis

This work is an attempt to study the mechanistic details of the biological function of L3MBTL3, particularly its role in the development of the megakaryocytic lineage. We hypothesize that L3MBTL3 is a methyl-lysine histone code reader that plays a role in the regulation of hematopoiesis as a result of its association with the repressive BHC complex. Chapter 2 will describe the work that was done towards the in vitro development and cellular characterization of UNC1215, an L3MBTL3-specific chemical probe. Chapter 3 will focus on the cellular histone binding activity of L3MBTL3, as well as on its role in megakaryopoiesis which was investigated with the use of UNC1215. The final chapter will close with brief concluding remarks and future directions.

1.13 References

Chapter 2

UNC1215 and other Cellularly Active L3MBTL3-Specific Chemical Probes

This work was conducted in collaboration with Stephen Frye's laboratory. Cloning of all constructs was performed by Shili Duan of the Arrowsmith laboratory. Experiments involving the \textit{in vitro} characterization of UNC1215, UNC1679, and UNC1967, as well as cellular cytotoxicity assays were performed by members of Stephen Frye's laboratory. The crystal structure was obtained and solved by Nan Zhong and Wolfram Tempel of the Structural Genomics Consortium (SGC). FRAP experiments were conducted by Dalia Barsyte of SGC. I performed all the other experiments, including UNC1215 co-localization experiments, affinity purifications, cellular dimer formation assays, and UNC1967 and UNC1679 cellular competition experiments.

The work described herein is published in and reproduced in part with permission from:


and


\texttt{http://pubs.acs.org/articlesonrequest/AOR-Fq7KFp52m7IyN5QJAUG6}
2.1 Summary

L3MBTL3 belongs to the Malignant Brain Tumor (MBT) family of proteins, all of which possess a tandem set of 2-4 MBT domains. Due to their ability to recognize mono- or di-methylated lysine residues on histone peptides, MBT domain-containing proteins have been proposed to act as histone code readers. This chapter describes the development and cellular characterization of UNC1215, the first published potent and selective chemical probe against a methyl-lysine reader protein. In vitro, UNC1215 binds the MBT domains of L3MBTL3 (3xMBT) with a $K_d$ of 120nM and competitively displaces 3xMBT off methylated histone peptides with an $IC_{50}$ of 40nM. UNC1215 binds 3xMBT in a 2:2 mode that involves both the first and second MBT domains of L3MBTL3. UNC1215 shows minimal to no cross-reactivity with other MBT domain proteins, as well as with other epigenetic and non-epigenetic proteins. In cells, UNC1215 was found to be non-toxic and was demonstrated to bind exogenous L3MBTL3 with a similar dependency on the 3xMBT domains as seen in vitro. UNC1215 caused an increase in the cellular mobility of GFP-L3MBTL3 in a dose-dependent manner with an $EC_{50}$ of 50-100nM, thus confirming its potency in cells. Finally, this chapter closes with the cellular characterization of two additional L3MBTL3-specific antagonists that show a higher selectivity towards L3MBTL3 over its closest homolog, L3MBTL1.

2.2 Introduction

L3MBTL3 is a putative histone code reader that contains three Malignant Brain Tumor (3xMBT) repeats, a C-terminal homo- or hetero-dimerization SAM domain, as well C2C2 and C2H2 zinc fingers (Figure 2.1) (1). While L3MBTL3 has been implicated in medulloblastoma (2) and in hematopoiesis (3), its molecular mechanism of action and the contribution of the methyl-lysine binding 3xMBT domains to its function remain unknown. One way to study this question is through chemical biology techniques. Good quality chemical probes developed against a particular target can serve as invaluable tools in studying the target's biological function.
What defines a good quality chemical probe? Several parameters have been proposed by Frye in 2010 (4). Firstly, the chemical probe must display potency and selectivity in *in vitro* assays. Secondly, its mechanism of action, such as the site it binds or the part of the target it inhibits, must be understood. This can be achieved through crystal structure studies as well as other approaches. Thirdly, the physical and chemical properties of the chemical probe, such as its stability, must be thoroughly characterized; this is crucial so that the observed effects can be attributed to the effect of the chemical probe on the target protein, rather than to an off-target effect. Finally, the functionality of the chemical probe must be characterized in the context of cells (4).

Many of the above-listed parameters act as guidelines, rather than strict rules to which one must adhere. For example, unlike some other sources (5), these guidelines do not state a set value for the *in vitro* and cellular potency of the probe or a fold selectivity over other targets that a compound must achieve. While such strict rules would indeed be useful in that they would provide objectivity, they would eliminate, sometimes unnecessarily and ambiguously, potentially useful chemical probes (4,6). These guidelines, on the other hand, still ensure the high quality of chemical probes while providing some room for variability, depending on the particular context and biological question (4).

This chapter describes the development and the *in vitro* and cellular characterization of UNC1215, the first reported specific, potent and cellurally-active probe against a methyl-lysine reader domain, the 3xMBT domains of L3MBTL3. The chapter concludes with a brief cellular characterization of two additional L3MBTL3-specific probes, UNC1679 and UNC1967.

**Figure 2.1. Domain organization of L3MBTL3.** L3MBTL3 consists of three tandem MBT domains, two zinc fingers, and a C-terminal SAM domain. The zinc finger shown in yellow is a C2C2 zinc finger while the zinc finger shown in green is a C2H2 zinc finger. The SAM domain is a hetero- or homo-dimerization domain. The two vertical lines represent nuclear localization signals. The three MBT domains, shown in orange, are the methyl-lysine binding domains. In each of the first two MBT domains, a conserved aspartic acid residue that is predicted to be critical for methyl-lysine binding is indicated.
2.3 Results

2.3.1 Development and in vitro Characterization of UNC1215

UNC1215 is a dibasic compound containing two methyl-lysine mimics (Fig. 2.2a) that was identified in an AlphaScreen competition assay and was found to displace 3xMBT of L3MBTL3 from methylated histone peptides with a half-maximum inhibitory concentration ($IC_{50}$) of 40nM. UNC1079 is a piperidine analog that is structurally similar to UNC1215, but inactive as an antagonist of 3xMBT methyl-lysine binding (Fig. 2.2b). Thus, it can be used as a negative control in functional studies. Isothermal calorimetry (ITC) assays indicate that UNC1215 binds 3xMBT with a $K_d$ of 120nM (Fig. 2.2c). To assess which of the three MBT domains of L3MBTL3 plays a role in binding UNC1215 (and, by extension, to methylated histone peptides), mutants of the conserved aspartic acid residue, that is predicted to be critical for binding methylated lysine residues, were designed. A mutation in the first MBT domain (D274A) decreased the binding affinity to UNC1215 by roughly 30-fold while a mutation in the second MBT domain (D381A) completely abolished binding, as measured by ITC, indicating that while the second MBT domain is crucial for binding, the first domain also plays a supporting role (Fig. 2.2c).

AlphaScreen and ITC assays indicate that UNC1215 is specific towards the 3xMBT domains of L3MBTL3 over the MBT domains of its family members, including its closest homolog L3MBTL1 (which showed 50x and 80x less specificity towards UNC1215 than L3MBTL3, as measured by AlphaScreen and ITC, respectively). Similarly, other non-MBT domain containing methyl-lysine readers (including 53BP1 and UHRF1, which possess Tudor domains, and CBX7 which possesses a chromodomain) showed no or much weaker binding to UNC1215 than 3xMBT of L3MBTL3. In addition, a number of other epigenetic and non-epigenetic families of proteins were screened by various assays for cross-reactivity with UNC1215. In particular, UNC1215 was found to have no effect on the activity of histone methyltransferases, no or weak interaction with bromodomain proteins and lysine demethylases, no effect on a panel of 50 diverse kinases (except weak inhibition of FLT3), and no activity against or binding to G-Protein Coupled Receptors, except for weak antagonist activity against the Muscarinic Receptors M1 and M2. Thus, overall, UNC1215 was concluded to be a highly selective and potent compound that targets the 3xMBT domains of L3MBTL3.
Next, we obtained an X-ray crystal structure of 3xMBT in complex with UNC1215 at a resolution of 2.6Å, which readily explained the ITC data observations. The crystal structure revealed that the pyrroldine amine of UNC1215 that is meta to the aniline substituent binds the second MBT domain of L3MBTL3, while the pyrroldine amine in the ortho position binds the first MBT pocket of another 3xMBT molecule, thus bridging them together. In addition, another UNC1215 compound binds the two sets of 3xMBT molecules in a reciprocal manner, thus resulting in a UNC1215-3xMBT 2:2 dimer (Fig. 2.2d). This result is in agreement with the ITC data (Fig. 2.2c) that indicates that both the first and second MBT domains play a role in binding UNC1215.

Figure 2.2. UNC1215 binds the 3xMBT domains of L3MBTL3 in vitro. (a) Structure of UNC1215. (b) Structure of UNC1079, an inactive piperidine analog of UNC1215. (c) Isothermal calorimetry reveals that UNC1215 binds 3xMBT with a $K_d$ of 120nM in a manner that is partially dependent on the first MBT domain and fully-dependent on the second MBT domain. (d) UNC1215 induces dimerization of 3xMBT domains of L3MBTL3 by binding the first MBT domain of one 3xMBT molecule and the second MBT domain of another 3xMBT molecule; another UNC1215 molecule binds in a reciprocal fashion. The first MBT domain is in red; the second is in green; the third is in blue. This work was not performed by the author. Figures a-c) are reproduced with permission from James, L.I. et al. Nat. Chem. Biol. 2013.
2.3.2 UNC1215 is a Non-Cytotoxic, Cell Permeable Compound that Binds Exogenous Full-Length L3MBTL3 in Cells

Once the potency and specificity of UNC1215 were characterized \textit{in vitro}, we sought to study its cellular activity. UNC1215 was found to be non-toxic up to a concentration of 100µM and stable in media in the presence or absence of HEK293 cells for up to 72 hours.

Because all \textit{in vitro} experiments were performed with 3xMBT of L3MBTL3, an unanswered question remained whether UNC1215 would also bind to full length L3MBTL3. To test this in the context of cells, HEK293 cells were transfected with an N-terminal GFP-L3MBTL3 fusion construct, followed by an overnight treatment with UNC1215 conjugated to a merocyanine dye, mero76, through a hexadiamine linker to the aniline ring of UNC1215 (mero-1215). GFP-L3MBTL3 was found to localize into distinct nuclear foci as seen by confocal microscopy (Fig. 2.3a). Similarly, mero-1215 also localized into overlapping nuclear foci (Fig. 2.3a). As a negative control, HEK293 cells were treated with mero76 alone. The dye did not co-localize with the GFP-L3MBTL3 foci (Fig. 2.3b), indicating that the ability of mero-1215 to bind GFP-L3MBTL3 is not caused by the dye.

To further assess the binding mode of UNC1215 in the context of cells, HEK293 cells were transfected with GFP fusions of first or second MBT domain aspartic acid mutants, GFP-L3MBTL3 D274A or GFP-L3MBTL3 D381A, respectively. Mero-1215 was found to co-localize with the first domain mutant (Fig. 2.3c), but not with the second domain mutant (Fig. 2.3d), in agreement with the biochemical data which showed that the second MBT domain was absolutely required for binding UNC1215.

While the foci formed by GFP-L3MBTL3 are likely not physiologically-relevant (as C-terminal L3MBTL3-GFP fusions and N-terminal mCherry-L3MBTL3 fusions do not form nuclear foci, instead spreading homogeneously throughout the nucleus, Fig. 2.4a), the foci still served as a convenient tool for co-localization studies. However, to ascertain that the foci did not yield false-positive results, an additional binding experiment with a Flag-tagged L3MBTL3 construct (which does not form foci either, not shown) was performed. To this end, HEK293 cells were transfected with Flag-L3MBTL3, followed by incubation of the cell lysate with biotinylated-
UNC1215. Pulldown of the biotinylated compound using streptavidin beads successfully affinity purified Flag-L3MBTL3, an interaction which was successfully outcompeted with increasing amounts of untagged UNC1215 (Fig. 2.4b).

**Figure 2.3. Mero-1215 co-localizes with GFP-L3MBTL3 foci in HEK293 cells.** (a) Mero-1215 co-localizes with nuclear foci formed by full length GFP-L3MBTL3 transfected into HEK293 cells. (b) Mero76 dye does not form nuclear foci in HEK293 cells transfected with GFP-L3MBTL3 and therefore, does not contribute to co-localization with GFP-L3MBTL3. (c) Mero-1215 successfully co-localizes with foci formed by first MBT domain mutant GFP-L3MBTL3 D274A. (d) Mero-1215 does not co-localize with foci formed by the second MBT domain mutant GFP-L3MBTL3 D381A. Scale bars, 10μm; green is GFP-L3MBTL3, red is mero-1215, co-localization of red and green is in yellow, and blue is Hoeschst dye. This experiment was performed by the author. Adapted and reproduced with permission from James, L.I. et al. Nat. Chem. Biol. 2013.
2.3.3 UNC1215 is a Potent Cellularly-Active L3MBTL3 Antagonist

Because MBT proteins are thought to act as histone code readers via the action of their MBT domains, we hypothesized that UNC1215 would compete with methyl-lysine marks found on histones, resulting in the dissociation of L3MBTL3 from chromatin, and increased protein mobility. To test this, GFP fusions of the 3xMBT domains were transfected into HEK293 cells, followed by treatment with UNC1215 at various concentrations. We then performed fluorescence recovery after photobleaching (FRAP) experiments to assess whether UNC1215 affected the mobility of 3xMBT. Indeed, UNC1215 treatment caused a decrease in the recovery time post-photobleaching in a dose-dependent manner with an EC\textsubscript{50} of 50-100nM while UNC1079 had no

---

**Figure 2.4. Biotinylated UNC1215 affinity purifies Flag-L3MBTL3 from cells.** (a) N-terminal GFP-L3MBTL3 fusion proteins localize into nuclear foci while C-terminal L3MBTL3-GFP and N-terminal mCherry-L3MBTL3 fusion proteins do not. Scale bar, 10µm. (b) Immunoprecipitation of biotinylated UNC1215 pulls down Flag-L3MBTL3 from HEK293 cells. 1 and 10 equivalents of untagged UNC1215 (compared to biotin-UNC1215) outcompete the interaction. **These experiments were performed by the author.** Adapted and reproduced with permission from James, L.I. et al. Nat. Chem. Biol. 2013.
effect (Fig. 2.5). Interestingly, GFP-3xMBT domain mutants, D274A and D381A, also displayed more rapid mobility than the wildtype protein (not shown), providing further evidence to support the mechanism of action and binding mode of UNC1215, as well as the role of the first and second MBT domains in the function of the protein. Finally, the mobility of the full-length GFP fusion protein also increased as a result of UNC1215 treatment or MBT domain mutations, but to a lesser extent than the GFP-3xMBT fusion protein (not shown).

2.3.4 Dimerization of 3xMBT in Cells

The dimer crystal structure of 3xMBT in complex with UNC1215 suggested that L3MBTL3 may similarly dimerize in cells as a consequence of binding to its endogenous substrate. To evaluate this, HEK293 cells were co-transfected with Flag-3xMBT and GFP-3xMBT constructs and treated with UNC1215 or water as a control. Immunoprecipitations using an anti-Flag antibody demonstrated that Flag-3xMBT dimerizes with GFP-3xMBT regardless of the presence of UNC1215 (Fig. 2.6), suggesting that the endogenous substrate may indeed induce dimerization of 3xMBT in a similar fashion to UNC1215.

Figure 2.5: UNC1215 potently increases the cellular mobility of GFP-3xMBT in cells. HEK293 cells were transfected with a GFP-3xMBT fusion construct. FRAP assays indicate that UNC1215 treatment increases the cellular mobility of GFP-3xMBT in a dose dependent manner with an EC$_{50}$ of 50-100nM, while UNC1079 has no effect. This experiment was not performed by the author.
2.3.5 Cellular Characterization of UNC1679 and UNC1967

During the development of UNC1215, a deeper understanding of the structural and chemical requirements of a potent L3MBTL3 antagonist was achieved. This led to the development of two additional L3MBTL3-binding compounds, UNC1679 and UNC1967 (Fig. 2.7a). *In vitro* studies demonstrated that these compounds maintained highly potent binding towards L3MBTL3, in addition to having higher selectivity towards L3MBTL3 compared to other MBT-containing proteins, in particular the closest homolog L3MBTL1.

To assess the potency of UNC1679 and UNC1967 in the context of cells, HEK293 cells were transfected with GFP-L3MBTL3, followed by treatment with 5µM mero-1215 and a subsequent
incubation with 1 equivalent (relative to mero-1215) of untagged UNC1215, UNC1679, or UNC1967. All three untagged compounds displaced mero-1215 from GFP-L3MBTL3 foci with UNC1215 and UNC1967 appearing to be more potent than UNC1679 (Fig. 2.7b).

**Figure 2.7. UNC1679 and UNC1967 are celluarly-active L3MBTL3 probes.** (a) Structure of UNC1679 (left) and UNC1967 (right). (b) Mero-1215 co-localizes with nuclear foci formed by GFP-L3MBTL3 in HEK293 cells (top panel). 5µM of UNC1215, UNC1679, and UNC1967 displace mero-1215 from GFP-L3MBTL3 foci (second, third, and fourth panel, respectively), with UNC1215 and UNC1967 acting more potently than UNC1679. Scale bars, 10µm; green is GFP-L3MBTL3, red is mero-1215, yellow indicates co-localization, and blue is Hoeschst dye. This experiment was performed by the author. Adapted and reproduced with permission from James, L.I. et al. J. Med. Chem. 2013. Copyright 2014 American Chemical Society.
2.4 Discussion

This work describes the development and characterization of the L3MBTL3-targeting antagonist, UNC1215. Based on the presented results, we believe that UNC1215 has the properties of a good quality chemical probe. UNC1215 is a potent antagonist of L3MBTL3 methyl-lysine binding \textit{in vitro}, displaying an IC$_{50}$ of 40nM. Its specificity has been extensively characterized \textit{in vitro} and it has been shown to have no to limited cross-reactivity with a wide range of proteins, including other MBT proteins, as well as other epigenetic and non-epigenetic targets. The mechanism of action of UNC1215 has been characterized \textit{in vitro} using ITC experiments that showed complete dependence on the second MBT domain and partial dependence on the first MBT domain, a result which was explained by the obtained crystal structure. Further, the stability of the compound in various contexts (in media, with and without cells) has been confirmed, providing credibility to the fact that the obtained results are an effect of UNC1215-mediated L3MBTL3-inhibition rather than of an off-target effect caused by a UNC1215 degradation product. Finally, the cellular activity of UNC1215 has been confirmed via the described co-localization studies and the FRAP assay, demonstrating nanomolar potency. Thus, UNC1215 can confidently be used in future studies investigating the biological function of L3MBTL3 and in particular, the dependence of L3MBTL3 on its methyl-lysine binding domains.

While this study largely aimed to characterize UNC1215 as an L3MBTL3 specific chemical probe, UNC1215 has already proven to be a useful tool in providing interesting insight into the biology of L3MBTL3. Based on sequence alignment with other functional MBT pockets, the second MBT repeat of L3MBTL3 was predicted to possess methyl-lysine binding activity due to the presence of a conserved aromatic cage (7). In contrast, the third MBT repeat possesses a bulky arginine residue at the "floor" of the cage, which is expected to prevent methyl-lysine from inserting into the cage due to electrostatic repulsion and steric hindrance effects. However, while the first and third MBT pockets are not predicted to bind to methylated lysine residues, they may still serve to accommodate another substrate. Indeed, crystal structures of previous MBT proteins, such as of L3MBTL1 have shown that its non-methyl lysine binding MBT pockets are able to accommodate the morpholino ring of MES or the proline ring of the third MBT repeat (8). In the present study, the binding of UNC1215, not only into the second methyl-lysine functional MBT pocket, but also into the first MBT pocket, suggests that the first repeat may indeed form a binding surface for an endogenous ligand. Moreover, the crystal structure of
3xMBT with UNC1215 displaying a 2:2 binding mode (Fig. 2.2d) suggests that the endogenous substrate of L3MBTL3 may cause a similar dimerization of L3MBTL3. Indeed, over-expressed 3xMBT fusion proteins were found to co-immunoprecipitate together even in the absence of UNC1215 (Fig. 2.6).

FRAP assays indicated that while UNC1215 has an effect on both the 3xMBT and the full-length L3MBTL3 constructs (Fig. 2.5 and not shown), its effect is weaker in the context of the full-length protein. This suggests that other domains of L3MBTL3 (two zinc fingers and the SAM domain) also play a role in the localization and/or function of L3MBTL3 and its anchoring to chromatin (or potentially to another substrate).

Having demonstrated that UNC1215 is a cell-permeable, potent, and selective chemical probe targeting the 3xMBT domains of L3MBTL3, one can now confidently utilize it in other assays to further investigate the biological function of L3MBTL3.

### 2.5 Materials and Methods

**Cell lines and cell culture**

The HEK293 cell line was cultured in high glucose DMEM medium (Gibco 11960) supplemented with 10% FBS, 1% glutamine, and 1% penicillin-streptomycin.

**Cloning and mutagenesis**

L3MBTL3 cDNA was PCR amplified using MGC clone corresponding to NP_001007103.1 and subsequently cloned into pACGFP C3, pACGFP N3, or pmCherry (Clontech), or pCDNA3-Flag (kindly provided by Dr. Brian Raught) using standard cloning procedures. Site directed mutagenesis was performed with Quick change (Stratagene).
Isothermal calorimetry, AlphaScreen assays, X-ray crystallography, and cytotoxicity assays


Immunofluorescence

Images were acquired using a Quorum Spinning Disk Confocal microscope equipped with 405-nm, 491-nm, 561-nm and 642-nm lasers. Image analysis was performed using Volocity 5.4.1.

Co-Localization and displacement experiments with mero-1215

For mero-1215 co-localization experiments, 100,000 HEK293 cells per mL per well were seeded on coverslips in 12-well plates. Cells were then transfected with wildtype GFP-L3MBTL3 using GeneJuice reagent, according to manufacturer's instructions for 16–24 hours. This was followed by a 2 hour incubation at 37°C with 2.5μM merocyanine dye (mero76) or mero76-UNC1215 conjugate per mL per well. Prior to imaging, cells were treated with Hoechst 33342 dye for 30 minutes, washed with PBS and incubated in DMEM/F-12 without phenol red (11039).

For mero-1215 displacement experiments, 150,000 HEK293 cells were seeded on coverslips in 12-well plates. After 24 hours, the cells were transfected with GFP-L3MBTL3, and after another 16 hours they were treated with 5μM mero-1215. After a 3 hour incubation period, the cells were treated with 5μM of UNC1215, UNC1679, or UNC1967 for an additional 3 hours. Prior to imaging, the cells were stained with Hoechst 33342 dye, washed with PBS, and incubated with DMEM/F-12 media without phenol red (11039).

Affinity purification with biotin-UNC1215 conjugate

HEK293 cells were transiently transfected with Flag-L3MBTL3 using GeneJuice reagent according to manufacturer’s instructions. Approximately 10 million cells were used per immunoprecipitation. Frozen cell pellets were lysed in lysis buffer containing 50 mM Tris-HCl,
pH 8.0, 500mM NaCl, 2mM EDTA, 1% Triton X-100 and complete protease inhibitor cocktail (Roche). After vortexing and centrifugation at 13.2krpm at 4°C for 5min to pellet the DNA, the supernatant was diluted to 150mM NaCl (while keeping the concentrations of other reagents constant) and was precleared with Streptavidin Mag Sepharose magnetic beads (GE Healthcare) for 1 hour at 4°C. Streptavidin beads were incubated with 5nmol biotin-UNC1215 in 300μL IP buffer (50mM Tris-HCl, pH 8.0, 150mM NaCl, 2mM EDTA, 1% Triton X-100) and protease inhibitor cocktail (Roche) for 30 minutes at room temperature. Precleared lysate was added to the biotin-UNC1215 prebound streptavidin beads in the absence or presence of 5nmol or 50nmol cold, untagged UNC1215 overnight at 4°C. Three washes were performed with IP buffer, followed by elution with buffer containing 10mM Tris-HCl, pH 8.0, 150mM NaCl, 2mM EDTA, 1% Triton, 3% SDS, 5mM DTT and 15mM βME. Sigma M2 (F1804) Flag antibody was used in immunoblotting.

**Cellular dimerization experiments**

HEK293 cells were co-transfected with Flag-3xMBT and GFP-3xMBT using GeneJuice reagent, according to manufacturer's instructions. Next day, cells were treated with 2μM UNC1215 or water as a negative control. Cells were harvested 24 hours later and immunoprecipitation was conducted using the immunoprecipitation protocol described above (under "affinity purification with biotin-1215 conjugate). Anti-Flag or anti-IgG antibodies were used for immunoprecipitation. The anti-Flag antibody used in immunoprecipitations and in immunoblotting was Sigma M2, F1804; the anti-GFP antibody used in immunoblotting was Clontech, 632381.

### 2.6 References

Chapter 3

Characterization of the Chromatin Binding Activity of L3MBTL3 and its Role in Megakaryocytic Development

The work presented here was performed in collaboration with Jean Philippe Lambert (of the Gingras laboratory) and Edyta Marcon (of the Greenblatt laboratory) who performed the immunoprecipitations/mass spectrometry experiments. HEK293T-REx cells expressing 3xFlag-L3MBTL3 were generated by Jean Philippe Lambert. Cloning of all constructs was performed by Shili Duan of the Arrowsmith laboratory. All other experiments were performed by the author.
3.1 Summary

Lethal(3)Malignant Brain Tumor-like Protein 3 (L3MBTL3) is a putative methyl-lysine histone code reader that promotes the development of erythrocytes, granulocytes, and macrophages during murine embryonic development. Immunoprecipitations conducted in HEK293 cells over-expressing full-length L3MBTL3 or a construct composed only of the three methyl-lysine binding MBT (3xMBT) domains demonstrate association with histone H3. Chromatin fractionation experiments demonstrated that full-length L3MBTL3 localizes into non-digestible portions of chromatin while 3xMBT partially mis-localizes into the digestible portion. Immunoprecipitation/mass spectrometry (IP/MS) results obtained by our collaborators indicate that over-expressed L3MBTL3 associates with members of the chromatin binding, repressive BHC complex in HEK293 cells. Co-immunoprecipitations, followed by immunoblot analysis, were conducted to confirm some of these interactions in HEK293 cells and in an erythroid leukemia K562 cell line which inducibly over-expressed 3xFlag-L3MBTL3. These interactions remained stable during PMA-induced megakaryocytic differentiation of K562 cells and were independent of the 3xMBT domains. Finally, differentiation experiments indicated that L3MBTL3 represses certain aspects of PMA-induced megakaryocytic development of K562 cells in a partially 3xMBT domain-dependent manner.

3.2 Introduction

The Malignant Brain Tumor family of proteins consists of 3 members in *Drosophila* and 9 members in mammals (1). This chapter will focus on L3MBTL3, a poorly characterized MBT domain containing protein. L3MBTL3 contains three MBT domains (3xMBT), two zinc fingers, and a C-terminal SAM domain (2). The 3xMBT domains possess methyl-lysine binding activity, showing a preference towards di-methylated over mono-methylated lysine residues *in vitro* (3). L3MBTL3 knockout mice die at a late embryonic stage from anemia due to a maturational defect in the myeloid lineage (4). In particular, L3MBTL3 knockout embryos contained lower levels of mature erythrocytes, macrophages, and granulocytes, but higher levels of megakaryocytic/erythroid progenitors (MEP), granulocyte/macrophage progenitors (GMP) and common myeloid progenitors (CMP) than L3MBTL3 wildtype embryos (Figure 3.1). An effect
on the megakaryocytic lineage was not reported (4). The mechanism of this maturational impairment remains unknown.

![Diagram showing hematopoietic abnormalities](image)

**Figure 3.1: Summary of hematopoietic abnormalities seen in L3MBTL3 knockout mice.** L3MBTL3 knockout mice contain decreased levels of mature erythrocytes, granulocytes, and macrophages, but increased levels of upstream progenitors, including the MEP (megakaryocytic/erythroid progenitors), GMP (granulocyte/macrophage progenitors), and CMP (common myeloid progenitors) populations. The lymphoid lineage is largely unaffected (4).

Immunoprecipitation/mass spectrometry studies performed by our collaborators indicate that over-expressed 3xFlag-L3MBTL3 associates with members of the chromatin-binding, transcriptional repressor BRAF-HDAC (BHC) complex in HEK293 cells (Figure 3.2). The BHC complex, first identified in 2002 (5), consists of six core subunits including the histone deacetylases HDAC1 and HDAC2, the PHD domain protein PHF21A, the transcriptional corepressor CoREST (aka RCOR1), HMG20B (aka BRAF35), and the histone demethylase LSD1 (aka KDM1A). The BHC complex represses transcription as a result of the function of the HDACs, which deacetylate acetylated lysine residues, and of LSD1 which demethylates the activating histone H3 lysine 4 mono- and di-methyl marks (H3K4me1/2). CoREST is a transcriptional corepressor that is required for LSD1-mediated demethylation of nucleosomes (6-7). PHF21A recognizes unmethylated H3K4 via its PHD finger. A proposed model suggests that PHF21A helps to anchor LSD1 at newly demethylated H3K4 sites, thereby preventing their remethylation and re-activation of gene expression. Indeed, knockdown of PHF21A resulted in de-
repression of LSD1 target genes (8). Finally, HMG20B is an HMG domain non-specific DNA binding protein (5).

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Total Peptide Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>L3MBTL3</td>
<td>0</td>
</tr>
<tr>
<td>SAMD1</td>
<td>0</td>
</tr>
<tr>
<td>KDM1A</td>
<td>0</td>
</tr>
<tr>
<td>PHF21A</td>
<td>0</td>
</tr>
<tr>
<td>CoREST</td>
<td>0</td>
</tr>
<tr>
<td>HDAC1</td>
<td>0</td>
</tr>
<tr>
<td>HMG20B</td>
<td>0</td>
</tr>
<tr>
<td>RCOR3</td>
<td>0</td>
</tr>
<tr>
<td>HDAC2</td>
<td>0</td>
</tr>
<tr>
<td>SFMBT1</td>
<td>0</td>
</tr>
<tr>
<td>SFMBT2</td>
<td>0</td>
</tr>
<tr>
<td>HIST1H1E</td>
<td>0</td>
</tr>
<tr>
<td>HIST3H3</td>
<td>0</td>
</tr>
<tr>
<td>H2AFX</td>
<td>0</td>
</tr>
<tr>
<td>CSNK2A1</td>
<td>0</td>
</tr>
<tr>
<td>CSNK2A2</td>
<td>0</td>
</tr>
<tr>
<td>CSNK2B</td>
<td>0</td>
</tr>
<tr>
<td>BCLAF1</td>
<td>9</td>
</tr>
</tbody>
</table>

**Figure 3.2: L3MBTL3 associates with members of the BHC complex.** Immunoprecipitation of 3xFlag-L3MBTL3 from HEK293 cells, followed by mass spectrometry analysis identified BHC complex members as putative L3MBTL3 interactors (orange). Other putative interactors include the SAM domain containing protein, SAMD1, MBT domain-containing proteins SFMBT1 and SFMBT2, histones, casein kinsases, and BCLAF1.

The BHC complex plays a role in several distinct processes, such as in neuronal gene repression (5), and myeloid differentiation (9), and possibly in myogenesis (10) and epithelial-to-mesenchymal transition (11). The actual role depends on the transcription factor that the complex associates with. For example, when the BHC complex associates with the DNA binding protein Gfi1b, it promotes erythroid and megakaryocytic maturation (9). In fact, many members of the complex have been implicated in erythroid and megakaryocytic development. For example,
LSD1 knockdown in murine erythroleukemia cells and in the murine megakaryoblastic cell line, L8057, resulted in impaired erythroid and megakaryocytic differentiation, respectively (9). Similarly, CoREST knockdown in L8057 also impaired megakaryocytic development (9). Gfi1b knockout mice die from anemia and exhibit defects in erythropoiesis and in megakaryopoiesis (12). In contrast, HMG20B knockdown in mouse primary proerythroblasts and in a mouse fetal liver cell line resulted in spontaneous differentiation (13).

Based on this data, we hypothesized that L3MBTL3 is a methyl-lysine histone code reader that exerts its effects via association with the BHC complex, thereby regulating transcription at various loci. In particular, we hypothesized that when L3MBTL3 binds the BHC complex in the context of the transcription factor Gfi1b, it regulates erythroid and megakaryocytic differentiation.

To address this hypothesis, we began by studying the histone and chromatin association of L3MBTL3. Next, we confirmed the binding of L3MBTL3 to BHC complex members in HEK293 cells and subsequently in K562 cells, which we utilized as a model system for erythroid and megakaryocytic development. Finally, we investigated the effect that L3MBTL3 over-expression and UNC1215-mediated 3xMBT domain inhibition has on megakaryocytic differentiation.

### 3.3 Results

#### 3.3.1 Flag-L3MBTL3 Co-Immunoprecipitates Histone H3 from HEK293 Cells

To study the association of L3MBTL3 with histones, HEK293 cells were transiently transfected with a full-length (FL) Flag-L3MBTL3 construct or with a Flag-3xMBT construct and immunoprecipitations were conducted with a Flag antibody. Western blots indicate that both FL Flag-L3MBTL3 as well as Flag-3xMBT co-immunoprecipitated histone H3 preferentially over histone H4 (Figure 3.3). Histones were more efficiently co-immunoprecipitated by the 3xMBT construct than by the FL protein (Figure 3.3, compare lanes 6 and 9).
Next, we sought to study the chromatin localization of L3MBTL3 and the role of the MBT domains in more detail. To this end, HEK293 cells were transiently transfected with N-terminal GFP fusions of various L3MBTL3 constructs, including full-length (FL) wildtype L3MBTL3, FL L3MBTL3 with an aspartic acid to alanine mutation in the first or second MBT domain (D249A and D381A, respectively), or with a 3xMBT construct. The aspartic acid is a conserved residue in all functional MBT pockets and is critical for binding the methylated lysine. To assess the chromatin localization of the various L3MBTL3 mutants, we followed a previously published chromatin digestion protocol (Figure 3.4a) (14). Briefly, nuclei of transfected HEK293 cells were isolated and subjected to treatment with micrococcal nuclease (MNase) for increasing periods of time (10 minutes, 30 minutes, and 1 hour). MNase is a nuclease that preferentially cuts chromatin between nucleosomes. The digested contents are then centrifuged and separated into a soluble fraction, containing digested chromatin, and a pellet fraction containing undigested chromatin (Figure 3.4b).

**Figure 3.3: L3MBTL3 binds histone H3.** Both FL Flag-L3MBTL3 and Flag-3xMBT co-immunoprecipitate histone H3 from HEK293 cells. The Flag-3xMBT construct runs at ~50kDa together with the heavy chain of the antibody. This experiment was repeated twice.

### 3.3.2 Chromatin and Cellular Localization of L3MBTL3

Next, we sought to study the chromatin localization of L3MBTL3 and the role of the MBT domains in more detail. To this end, HEK293 cells were transiently transfected with N-terminal GFP fusions of various L3MBTL3 constructs, including full-length (FL) wildtype L3MBTL3, FL L3MBTL3 with an aspartic acid to alanine mutation in the first or second MBT domain (D249A and D381A, respectively), or with a 3xMBT construct. The aspartic acid is a conserved residue in all functional MBT pockets and is critical for binding the methylated lysine. To assess the chromatin localization of the various L3MBTL3 mutants, we followed a previously published chromatin digestion protocol (Figure 3.4a) (14). Briefly, nuclei of transfected HEK293 cells were isolated and subjected to treatment with micrococcal nuclease (MNase) for increasing periods of time (10 minutes, 30 minutes, and 1 hour). MNase is a nuclease that preferentially cuts chromatin between nucleosomes. The digested contents are then centrifuged and separated into a soluble fraction, containing digested chromatin, and a pellet fraction containing undigested chromatin (Figure 3.4b).
Western blots indicate that FL wildtype L3MBTL3 localizes exclusively into the non-digestible chromatin fraction (Fig. 3.4c, top panel). Both the first and second MBT mutants show a minor mis-localization into the soluble portion, with the second MBT mutant showing a slightly higher mislocalization in agreement with the known role of the second MBT domain as the major driver of the methyl-lysine binding activity (Fig. 3.4c, second and third panels). In contrast, the 3xMBT construct shows significant mis-localization into the soluble portion, illustrating the importance of the zinc fingers and the SAM domain in the relative localization of L3MBTL3 (Figure 3.4c).

Figure 3.4: GFP fusions of various L3MBTL3 constructs display differential localization on chromatin. (a) Experimental outline of the micrococcal nuclease chromatin digestion protocol. Adapted from Wu, S. et al. (14). (b) MNase-mediated chromatin digestion results in the formation of a soluble, digestible chromatin fraction and a non-digestible fraction, as seen on the agarose gel stained for DNA. (c) Full-length GFP-L3MBTL3 localizes exclusively into the non-digestible fraction. A mutation in the first or second MBT repeats results in a minor mislocalization of the protein into the soluble fraction. A construct composed of MBT domains only localizes both into the digestible and non-digestible chromatin fractions.
3.3.3 3xFlag-L3MBTL3 Binds Members of the BHC Complex in HEK293 Cells

To better understand the biological function of L3MBTL3, we sought to discover its protein interactors. Previously, our collaborators performed immunoprecipitations followed by mass spectrometry analysis on HEK293T-REx cells inducibly expressing 3xFlag-L3MBTL3. The detected peptides were found to correspond to various members of the chromatin-binding, repressive BHC complex including LSD1, CoREST, HDAC1, HDAC2, PHF21A, HMG20B, and RCOR3. Other peptides with high counts corresponded to SFMBT1, SFMBT2, SAMD1, BCLAF1, as well as several histones (Figure 3.2). In agreement with these results, several recent papers studying SFMBT1 found it in a complex with L3MBTL3, as well as with members of the BHC complex (10-11,15). To validate the interaction between L3MBTL3 and BHC complex members, we performed immunoprecipitations on HEK293T-Rex cells inducibly expressing 3xFlag-L3MBTL3 and found that immunoprecipitations of endogenous LSD1 and CoREST indeed co-purified 3xFlag-L3MBTL3 (Fig. 3.5). In contrast, a reciprocal immunoprecipitation with a Flag antibody failed to pull down LSD1 or CoREST, possibly due to the fact that comparatively higher levels of 3xFlag-L3MBTL3 are present in the cell lysate and a large fraction of it remains unassociated with any factors.

![Image of immunoprecipitation results]

**Figure 3.5: L3MBTL3 associates with members of the BHC complex.** Immunoprecipitations of endogenous LSD1 and CoREST pulled down 3xFlag-L3MBTL3 inducibly expressed by HEK293T-Rex cells. In contrast, immunoprecipitation of 3xFlag-L3MBTL3 with a Flag antibody failed to pull down endogenous LSD1 and CoREST. mIgG serves as the negative control for the Flag immunoprecipitation while rIgG serves as the negative control for the LSD1 and CoREST immunoprecipitations.
Having established the association of exogenous L3MBTL3 with LSD1 and CoREST, we sought to determine the role of L3MBTL3 in this complex. The MBT proteins have generally been characterized as transcriptional repressors ((10-11,16-17). We therefore hypothesized that knocking down L3MBTL3 would result in de-repression of BHC complex target genes. Unfortunately, knockdowns of L3MBTL3 using both siRNA and shRNA approaches have failed in a number of cell lines, including the Human Embryonic Kidney HEK293, the colon carcinoma HCT116, and the erythroid leukemia K562 cell lines (while side by side knockdowns of various positive control genes were successful). In contrast, knockdowns have been successful in the osteosarcoma U2OS and human foreskin fibroblasts HFF cell line. We presumed that perhaps the cell lines where knockdown was unsuccessful express comparatively low levels of L3MBTL3 transcript. To investigate this, RNA was extracted from cell pellets of various cell lines, reverse transcribed, and the resulting cDNA was used to measure transcript levels. The G401 rhabdoid tumor cell line, which expresses comparatively high levels of L3MBTL3, was used as a positive control. Results suggest that knockdown effectiveness does not depend on L3MBTL3 mRNA levels as HFF and U2OS cell lines, where knockdowns were successful, express roughly the same levels of L3MBTL3 transcript as HEK293 cells and lower levels than HCT116 cells, where knockdowns failed (Fig. 3.6a).

Since L3MBTL3 knockdowns were successful in the HFF cell line, we decided to use it for further investigation of the transcriptional regulation activity of L3MBL3. Perhaps the most widely studied function of the BHC complex is in repression of neuronal gene expression (5). We therefore attempted to measure levels of various neuronal genes (Snap25, Golga7, and KCNB1) after L3MBTL3 knockdown; however, results were inconsistent as levels of expression of these genes were very low, yielding poor PCR replicates (not shown).

Searching for another route, we noticed that the transcription factor RBPJ (aka CSL) was identified (albeit inconsistently) in some of the IP/MS experiments as a potential L3MBTL3 interactor. RBPJ is a DNA-binding protein that regulates the Notch pathway, which affects many processes ranging from hematopoiesis to neurogenesis (18). In the absence of Notch signaling, RBPJ inhibits Notch target genes by associating with a repressive complex consisting of LSD1, CoREST, as well as the histone deacetylase SIRT1 and the co-repressor CtBP1 (19). Thus,
although this complex differs from the BHC complex, they have several common members including LSD1 and CoREST, with which L3MBTL3 interacts. We therefore decided to test whether L3MBTL3 affects the Notch pathway, presumably via its association with CoREST, LSD1, and possibly RBPJ. To our surprise, L3MBTL3 knockdown in HFF cells resulted in repression of two Notch pathway genes, *Hes1* and *Notch3*, but not of *LSD1*, which was used as a negative control (Fig. 3.6b), suggesting that L3MBTL3 may act as a transcriptional activator in this context if the effect is direct. In support of this, LSD1 and CoREST were found to localize at the promoter-proximal RBPJ binding site of the *Hes1* gene in the lung fibroblast IMR90 cell line (19). However, conclusive evidence has not yet been obtained.

**Figure 3.6: L3MBTL3 knockdown results in repression of Notch pathway genes.**
(a) Measurement of L3MBTL3 transcript levels in several cell lines in order to determine if inability to knock down L3MBTL3 is a result of low L3MBTL3 expression revealed no such correlation. Results of two biological replicates are shown. Error bars indicate SEM.
(b) L3MBTL3 knockdown in HFF cells resulted in repression of Notch pathway genes, *Hes1* and *Notch3*. LSD1 transcript levels, which was included as a negative control, are unchanged as a result of L3MBTL3 knockdown. siNT refers to non-targeted siRNA; siL3MBTL3 refers to L3MBTL3-directed siRNA. Results are based on a single biological replicate.

**3.3.5 K562 Cells Possess a Functional Repressive BHC Complex**

Another developmental process the BHC complex has been shown to regulate is myeloid development, including differentiation of erythrocytes, megakaryocytes, granulocytes, and macrophages (9). This function of the BHC complex was of particular interest to us because
L3MBTL3 knockout mice displayed abnormalities in some of these lineages (5). We hypothesized that L3MBTL3 may regulate erythropoiesis and possibly, megakaryopoiesis, via its association with the BHC complex in context of the DNA binding protein Gfi1b. To study the role of L3MBTL3 in the development of these lineages, we chose to use the K562 erythroleukemia cell line which can be induced to differentiate down both the erythroid and megakaryocytic lineages (20).

First, we assessed whether the K562 cell line possesses a functional BHC complex. Reciprocal immunoprecipitations of LSD1 and CoREST were performed, followed by immunoblotting. LSD1 and CoREST were indeed found to be in contact with each other (Fig. 3.7a). Next, we assessed the functionality of the BHC complex by knocking down LSD1, followed by real-time PCR to assess the levels of expression of BHC complex target genes, Gfi1b, C-myb, and Meis1 (9,21). In addition, levels of the transcription factor, Spi1 were also measured. Spi1 was assumed to be a BHC complex target gene since, according to ENCODE data, LSD1, CoREST, HDAC1, and HDAC2 all co-localize to the promoter area of Spi1 in K562 cells (http://www.genome.ucsc.edu/ENCODE/). Indeed, LSD1 knockdown targeting two distinct regions of the LSD1 transcript resulted in de-repression of Gfi1b, Spi1, and Meis1, validating the repressive activity of the complex in K562 cells (Fig. 3.7b). The effect of LSD1 knockdown on C-myb transcript levels is inconsistent between the two distinct shRNA sequences, possibly due to an off-target effect (Fig. 3.7b).

### 3.3.6 UNC1215-Mediated 3xMBT Domain Inhibition does not Alter Expression of BHC Complex Target Genes

Once the functionality of the BHC complex in K562 cells was confirmed, we proceeded to study the role of L3MBTL3 in this complex. Since several approaches to achieve an L3MBTL3 knockdown in K562 cells have failed, we opted to use UNC1215 to study the role of the 3xMBT domains in the function of L3MBTL3. We hypothesized that if the 3xMBT domains were crucial for the transcriptional regulation activity of L3MBTL3, UNC1215 treatment would result in a change in expression of BHC complex target genes. A 3µM concentration of UNC1215 was presumed to be sufficient for this experiment since UNC1215 was previously shown to have an antagonistic effect on the foci formation of 3xMBT at an IC90 of ~1µM (2). In addition, a 3µM
concentration is not expected to cause non-specific off-target effects (2). Thus, K562 cells were treated with 3µM UNC1215 for 72 hours and samples were taken every 24 hours for RNA extraction and gene expression analysis. Overall, no difference or only a very minor decrease was seen in the expression levels of BHC complex target genes, Gfi1b, C-myb, and Spi1 at each timepoint as a result of UNC1215 treatment (Fig. 3.7c).

Figure 3.7: The erythroid leukemia K562 cell line possesses a functional BHC complex. (a) Endogenous LSD1 and CoREST interact in K562 cells as demonstrated by reciprocal immunoprecipitations, followed by Western blotting. (b) LSD1 knockdown in K562 cells using two different shRNA sequences de-represses BHC complex target genes, Gfi1b, C-myb, Spi1, and Meis1. Results are based on a single biological replicate. (c) UNC1215 treatment of K562 cells does not lead to a major change in gene expression of BHC complex target genes. Results are based on a single biological replicate.
3.3.7 3xFlag-L3MBT3 Associates with Members of the BHC Complex in K562 Cells in a Manner Independent of the 3xMBT Domains

Due to lack of availability of sensitive L3MBT3 antibodies and due to the comparatively low level of expression of the L3MBT3 transcript in K562 cells (Fig 3.6a), we generated K562 cells that inducibly express 3xFlag-L3MBT3 under the control of a doxycycline-responsive promoter. We then performed immunoprecipitations to assess whether L3MBT3 bound LSD1 and/or CoREST. Indeed, immunoprecipitations of endogenous LSD1 and CoREST (using two different antibodies against distinct epitopes) pulled down 3xFlag-L3MBT3 (Figure 3.8a, left). Next, to determine whether L3MBT3 remains associated with the complex during megakaryocytic differentiation, cells were incubated with PMA (phorbol 12-myristate 13-acetate) to induce megakaryocytic differentiation and immunoprecipitations were then performed. Differentiation was confirmed by measuring transcript levels of megakaryocytic markers, CD10 and CD61 (20,22,23) (Fig. 3.8b). Indeed, immunoprecipitations of endogenous LSD1 and CoREST pulled down 3xFlag-L3MBT3 during megakaryocytic differentiation (Fig. 3.8a, right). Finally, to investigate whether the association of L3MBT3 with the BHC complex is dependent on methyl-lysine recognition via 3xMBT domains, K562 cells were induced with PMA in the presence or absence of 3µM UNC1215. No differential binding was observed between L3MBT3, LSD1, and CoREST in UNC1215-treated versus untreated cells (Fig. 3.8c). Differentiation was confirmed by RT-PCR (Fig. 3.8d).

3.3.8 3xFlag-L3MBT3 Impairs PMA-Induced Megakaryocytic Differentiation of K562 Cells in a 3xMBT Domain Dependent Manner

Next, we sought to assess how L3MBT3 affects megakaryocytic differentiation. To this end, we treated inducible 3xFlag-L3MBT3- or empty vector-infected K562 cells with doxycycline, with a concurrent treatment of PMA or ethanol as the vehicle control. In addition, to dissect out the function of 3xMBT, these differentiation experiments were also conducted in the presence or absence of 3µM UNC1215. Samples were collected at various time points to evaluate the extent of differentiation by measuring transcript levels of PMA-induced megakaryocytic differentiation markers, CD10, CD61, and CD41 (Fig. 3.9a).
Figure 3.8: 3xFlag-L3MBTL3 associates with LSD1 and CoREST in K562 cells, in the presence or absence of PMA-induced megakaryocytic differentiation, independently of its 3xMBT domains. (a) Endogenous LSD1 and CoREST interact with inducibly-expressed 3xFlag-L3MBTL3 in K562 cells (left) and in K562 cells differentiated down the megakaryocytic lineage (right). (b) Transcript levels of CD10 and CD61 were measured to confirm PMA-induced megakaryocytic differentiation. (c) UNC1215 treatment does not affect the association of 3xFlag-L3MBTL3 with LSD1 or CoREST during PMA-induced megakaryocytic differentiation of K562 cells. (d) Transcript levels of CD10 and CD61 were measured to confirm PMA-induced megakaryocytic differentiation.
Interestingly, 3xFlag-L3MBTL3 over-expression impaired PMA-induced megakaryocytic differentiation of K562 cells as measured by significantly decreased levels of CD10 and CD61 transcripts at several time points (compare solid blue to solid green curves in Fig. 3.9b at 24hrs, 48hrs, and 96hrs; and in Fig. 3.9d at 48hrs and 96hrs; p < 0.05). UNC1215-mediated inhibition of endogenous L3MBTL3 resulted in significantly increased levels of CD10 and CD61 at several time points in empty-vector cells (compare solid blue to dashed blue curves in Fig. 3.9b at 48hrs, 72hrs, and 96hrs; and in Fig. 3.9d at 48hrs; p < 0.05). In contrast, UNC1215 inhibition of over-expressed 3xFlag-L3MBTL3 resulted in a significant increase of CD10 only at two time points (compare solid green to dashed green curves in Fig. 3.9b at 24hrs and 48hrs), and no significant increase of CD61 (Fig. 3.9d). Interestingly, UNC1215 caused a minor, but significant, increase in CD10 levels even in the absence of PMA treatment (in Fig. 3.9c, compare solid orange to dashed red at 48hrs, 72hrs, and 96hrs, and solid pink to dashed purple at 48hrs; p < 0.05). On the other hand, levels of the megakaryocytic marker CD41 were not affected by 3xFlag-L3MBTL3 over-expression or UNC1215 treatment (Fig. 3.9f). Thus, judging by CD10 and CD61 levels, L3MBTL3 impairs PMA-induced megakaryocytic differentiation of K562 cells in a partially 3xMBT domain-dependent manner.

Successful 3xFlag-L3MBTL3 over-expression was confirmed by measuring L3MBTL3 transcript levels (Fig. 3.9g). Doxycycline treatment of 3xFlag-L3MBTL3 infected K562 cells in the presence of ethanol leads to a roughly 10-fold increase in L3MBTL3 gene expression at each time point compared to non-infected cells (compare pink to orange curves). Interestingly, PMA treatment leads to a roughly 1.5-4 fold increase in the level of endogenous L3MBTL3 (compare blue to orange curves in Fig. 3.9g, any time point; p < 0.05). Similar results were previously seen with PMA treatment of another erythroleukemia cell line, TF-1, as well as with HL-60 and KG-1 cell lines (4). However, it is likely that this increase in L3MBTL3 transcript levels is not an L3MBTL3-specific effect; in fact, phorbol esters, such as PMA, often upregulate a number of mRNA transcripts, partially via mRNA stabilization (24). Similarly, a combination treatment of doxycycline and PMA results in a synergistic increase in L3MBTL3 transcript levels (Fig. 3.9g, green curve). L3MBTL3 transcript levels are independent of UNC1215 treatment.
RT-PCR to measure levels of PMA-induced megakaryocytic differentiation markers (CD10, CD61, and CD41)
3.4 Discussion

3.4.1 Binding of L3MBTL3 to Chromatin

The histone binding and chromatin localization assays illustrate a number of interesting points about the biology of L3MBTL3. Firstly, the immunoprecipitations assays confirm that FL L3MBTL3 and the 3xMBT domains alone are indeed capable of binding histones in the context of cells (although direct binding to histones in cells was not strictly demonstrated) (Fig. 3.3). Interestingly, FL L3MBTL3 and 3xMBT behaved in a similar manner in that both preferentially co-immunoprecipitated histone H3 over histone H4. Although this does not prove a direct association of L3MBTL3 with histone H3, such an association is likely since L3MBTL3 binds the BHC complex, which contains at least two histone H3-binding members: LSD1, which...
demethylates the H3K4me1/2 marks, uses a deep, negatively charged pocket to bind the positively charged histone H3 tail (25), and PHF21A which recognizes the unmethylated H3K4 mark via its PHD finger (8). Thus, as a result of its histone methyl-lysine recognition ability, L3MBTL3 may recognize H3K4me1/2 and aid in the recruitment of the BHC complex to chromatin; alternatively, L3MBTL3 may oppose the function of the complex by binding the H3K4me1/2 mark, protecting it from demethylation.

Surprisingly, 3xMBT co-immunoprecipitated histones more efficiently than FL (Fig. 3.3, second and third panels, compare lanes 6 and 9). This may seem counterintuitive because one would assume that the additional domains of L3MBTL3 would somehow aid in its association with chromatin. Several possibilities may explain this. The simplest explanation may be differential transfection efficiency and protein abundance of the two different L3MBTL3 constructs. Another non-mutually exclusive possibility is that 3xMBT may be a promiscuous binder and recognize a variety of methylation sites while FL L3MBTL3 may be more specific towards a particular methylation mark due to the presence of other domains which may confer additional binding specificity (either as a result of a direct interaction with chromatin or with other chromatin-binding proteins). In other words, this result suggests that MBT domains are not solely responsible for specifying histone association. A similar conclusion can be reached from the chromatin digestion experiments where the FL protein and the 3xMBT domain construct display distinct chromatin localization patterns (Fig. 3.4c). An alternative explanation may be that in the course of the cell lysis and extraction procedure (which involved sonication and benzonase-mediated DNA digestion), the nucleosomes were, at least partially, dissociated. Thus, it is possible that under these particular experimental conditions, the non-MBT domains of L3MBTL3 would not aid, and might even hinder the association of L3MBTL3 with histones.

Interestingly, 3xMBT preferentially co-immunoprecipitated histone H3 over histone H4, despite the fact that in vitro fluorescence polarization (FP) assays indicated preferential binding of 3xMBT to the H4K20me2 mark (3). The H4K20me2 mark is comparatively abundant (26), so it is not likely that this discrepancy can be attributed to the low levels of the mark in cells (although absolute abundance compared to the H3K4me2 mark is unknown). A possible explanation may involve additional PTMs that are found in cells, but absent in in vitro assays with histone peptides; for example, the presence of a PTM on a residue neighbouring H4K20 may interfere with binding of 3xMBT in cells. However, this is entirely speculative.
The chromatin digestion experiments demonstrate that full-length L3MBTL3 localizes exclusively into the MNase non-digestible portion of chromatin. Consistent with this, SFMBT1, an L3MBTL3 binding partner (15), displays a similar localization pattern (14). Finally, mutations in the first and second MBT domains cause only a minor mislocalization of L3MBTL3 into the soluble fraction, once again highlighting the importance of other L3MBTL3 domains. This result also agrees with FRAP data showing that UNC1215-mediated inhibition of L3MBTL3 leads to a somewhat increased mobility of L3MBTL3 (2). It would be interesting to investigate whether UNC1215 treatment of HEK293 cells would cause a similar mis-localization of FL L3MBTL3 into the digestible chromatin fraction.

3.4.2 The Association of L3MBTL3 with the BHC Complex

The binding of L3MBTL3 to BHC complex members was not unexpected as several recent papers have demonstrated that SFMBT1 binds BHC complex members and also binds L3MBTL3 (10,15). Interestingly, SFMBT1 was found to associate with members of the BHC complex, with the co-repressors CtBP1 and CtBP2 that act with LSD1 and CoREST in a distinct complex, with several MBT domain proteins (including L3MBTL3), with Polycomb Repressive Complex 1 (PRC1) members, and a number of other proteins. L3MBTL3, on the other hand, displayed a more selective set of interactors, with the BHC complex being the sole complex identified as a putative interactor. Future studies should investigate whether L3MBTL3 and SFMBT1 have cooperative or opposing roles in histone recognition and in the function of the BHC complex.

How L3MBTL3 associates with the BHC complex is not yet known. While SFMBT1 binds CoREST directly through a conserved DUF3588 motif of unknown function (11), it is still unclear whether L3MBTL3 binds the BHC complex directly or associates with it via SFMBT1, perhaps by virtue of SAM domain heterodimerization (27,28,29). Another possibility is that the 3xMBT domains of L3MBTL3 may be involved in recognition of one of the BHC complex members. However, because UNC1215 did not disrupt the association of L3MBTL3 with LSD1 or CoREST (Fig. 3.8c), the association of L3MBTL3 with the BHC complex does not solely depend on the 3xMBT domains; this serves as indirect evidence for a distinct substrate for the 3xMBT domains, likely methylated histones.
Finally, while this study has shown that L3MBTL3 associates with BHC complex members, association (direct or indirect) with Gfi1b has not been demonstrated due to technical difficulties with finding good quality Gfi1b antibodies. Future studies should attempt to address the question of identifying the transcription factor that guides the DNA-binding activity of the BHC-L3MBTL3 complex.

3.4.3 The Mechanistic Role of L3MBTL3 in the BHC Complex

The particular function of L3MBTL3 in the BHC complex has not yet been established. One putative function is that L3MBTL3 might aid in the recruitment of the complex to target loci by recognizing a particular methylation mark, likely the H3K4me1/2 mark. Such a role would enable the demethylation function of LSD1 by placing it in proximity to its target mark. However, phenotypically, LSD1 and L3MBTL3 have distinct functions. Whereas LSD1 knockdown led to impaired PMA-induced megakaryocytic differentiation of the megakaryoblastic L8057 cell line, L3MBTL3 UNC1215-mediated inhibition promoted PMA-induced megakaryocytic differentiation of K562 cells. Similarly, L3MBTL3 over-expression resulted in impairment of megakaryocytic differentiation. Thus, if the function of LSD1 to promote megakaryopoiesis in murine L8057 cells is conserved in K562 cells, a mechanistic role in which L3MBTL3 supports the function of LSD1 (i.e., a recruitment role) does not agree with the phenotypic results. Another potential function for L3MBTL3 would be to bind the H3K4me1/2 marks and shield them from demethylation by LSD1, thus, opposing the activity of LSD1. In this manner, L3MBTL3 would act as a negative regulator of the complex, a mechanistic role which would agree with its phenotypic function in respect to LSD1.

3.4.4 L3MBTL3 - an Activator or a Repressor?

While all currently characterized MBT family members have been shown to act as repressors of gene expression (1,10-11,16-17), the slight repressive effect that UNC1215-mediated L3MBTL3 inhibition had on the expression of BHC complex target genes in K562 cells (Fig. 3.7c) suggests a potential role of L3MBTL3 as an activator. Whether this repressive effect is significant is
uncertain. Additional experimental replicates may bring more clarity. However, a weak UNC1215-mediated transcriptional effect would not be unexpected as in all reported 3xMBT domain inhibition or 3xMBT domain mutational experiments (such as the chromatin localization experiments, FRAP assays measuring L3MBTL3 mobility (2), and megakaryopoiesis differentiation assays), interference with the 3xMBT domains did not cause a complete abrogation of function or localization.

L3MBTL3 knockdown resulted in repression of Notch pathway genes, Hes1 and Notch3, in HFF cells. In IMR90 cells, LSD1 and CoREST are found at the promoter-proximal site of the Hes1 gene (19). Future studies should investigate whether L3MBTL3 also localizes at the promoters of Hes1 and Notch3 in HFF cells. A positive result would suggest that L3MBTL3 acts as an activator of gene expression.

The fact that L3MBTL3 has a phenotypically distinct role from LSD1 in megakaryopoiesis also supports the possibility that L3MBTL3 may activate gene expression since LSD1 acts as a repressor of gene expression when targeting the activating H3K4me1/2 marks. Although LSD1 is also capable of demethylating the repressive H3K9me1/2 marks (and thus, acting as an activator), this activity was not reported in the context of the BHC complex, but in the context of the androgen receptor (30).

Finally, another piece of evidence that indirectly points towards a potentially activating role of L3MBTL3 is that fetal liver cells of L3MBTL3 knockout mice display a significant reduction in expression of the cyclin-dependent protein kinase inhibitor, p57^{KIP2} (4) in contrast to LSD1 knockdowns in HeLa cells that result in de-repression of p57^{KIP2} (31). However, whether these are effects of the direct activity of an LSD1-L3MBTL3 complex is unknown. Future work should aim to pinpoint the precise epigenetic role that L3MBTL3 possesses.

### 3.4.5 The Role of L3MBTL3 in Megakaryocytic Differentiation

Based on RT-PCR results measuring various markers of PMA-induced megakaryocytic differentiation of K562 cells, L3MBTL3 appears to inhibit certain aspects of megakaryocytic development in a 3xMBT domain-dependent manner. The altered levels of the CD10 and CD61
transcripts that result from L3MBTL3 over-expression or inhibition are likely an indirect effect since BHC complex members, LSD1, HDAC1, and HDAC2 do not localize to the CD10 and CD61 promoters in K562 cells, based on ENCODE data (http://www.genome.ucsc.edu/ENCODE/). Interestingly, L3MBTL3 over-expression and UNC1215 treatment only affected the CD10 and CD61 markers, but not the CD41 marker. It is possible that a distinct pathway regulates the expression of the CD41 marker that is independent of L3MBTL3 and the BHC complex.

Another question that remains unanswered is whether the decrease in megakaryocytic markers as a result of L3MBTL3 over-expression is due to a decreased differentiation state of the cell population or due to a decreased number of differentiated cells. Indeed, the number of adherent (and therefore, differentiated) cells in the 3xFlag-L3MBTL3 over-expressing group was lower than in the control (empty vector) group; however, this difference was not found to be statistically significant. Another approach to address this question would be to use Wright-Giemsa staining to compare the numbers of differentiated cells in the two cell populations.

The comparatively mild impairment that L3MBTL3 over-expression has on megakaryocytic differentiation may be due to the mixed population of 3xFlag-L3MBTL3 over-expressing K562 cells, with some cells over-expressing L3MBTL3 at high levels and others at low levels. A potential approach to dissect out the full magnitude of the role of L3MBTL3 would be to co-stain K562 cells with a Flag antibody and an antibody against a megakaryocytic marker to search for a predicted correlation between cells expressing high levels of Flag-L3MBTL3 and low levels of megakaryocytic markers, and vice versa.

3.4.6 The Role of L3MBTL3 in Erythroid Development

Based on the L3MBTL3 knockout mice phenotype, L3MBTL3 is predicted to promote erythroid development. Preliminary experiments with K562 cells over-expressing 3xFlag-L3MBTL3 in the presence of hemin, which induces erythroid differentiation, indicate that this may indeed be the case. Hemin-treated K562 cells that over-express 3xFlag-L3MBTL3 displayed slightly higher levels of erythroid differentiation markers, including γ-globin, α-globin, and CD71 than K562 cells with basal level of L3MBTL3 expression (compare red and purple bars in Fig. 3.10b, c, and
Fig. 3.10a confirms successful L3MBTL3 over-expression, which increases synergistically as a result of a combination treatment of doxycycline and hemin. Replicates need to be repeated for an assessment of significance to be made. If L3MBTL3 is indeed shown to promote erythroid differentiation, it would point to its plasticity as an epigenetic regulator: LSD1 has been shown to promote erythroid development of murine erythroleukemia cells (9); thus, in this context, L3MBTL3 and LSD1 would promote an identical phenotypic outcome, presumably by playing supportive mechanistic roles, indicating that L3MBTL3 may act as a transcriptional repressor in this context.

**Figure 3.10: 3xFlag-L3MBTL3 may promote hemin-induced erythroid development of K562 cells.** (a) L3MBTL3 transcript levels rise synergistically in the presence of doxycycline and hemin treatment. Hemin alone does not increase endogenous L3MBTL3 transcript levels. (b) γ-globin, (c) α-globin, and (d) CD71 transcript levels are slightly higher in hemin-treated K562 cells that over-express 3xFlag-L3MBTL3, compared to those that express empty vector. Results are based on a single replicate.
3.4.7 Conclusion

In conclusion, this study investigated the chromatin binding activity of L3MBTL3 and its role in megakaryopoiesis and erythropoiesis, utilizing UNC1215 as a tool to study the functional significance of the methyl-lysine binding activity of L3MBTL3.

3.5 Materials and Methods

Plasmids

L3MBTL3 cDNA was PCR amplified using MGC clone corresponding to NP_001007103.1 and subsequently cloned into pACGFP C3 or pCDNA3-Flag (kindly provided by Dr. Brian Raught) using standard cloning procedures. Site directed mutagenesis was performed with Quick change (Stratagene). pLVX-Tet3G and pLVX-TRE3G were purchased from Clontech. 3xFlag-L3MBTL3 was PCR amplified from 3xFlag-L3MBTL3-pCDNA5-FRT-TO (kindly provided by Jean-Philippe Lambert of the Gingras laboratory), and subsequently cloned into the pLVX-TRE3G vector using In-Fusion system (Clontech).

Cell Culture and Generation of Stable Cell Lines

K562 cells (American Type Culture Collection, ATCC) were grown in IMDM medium, HFF (ATCC), HEK293T cells (ATCC) and HEK293T-REx cells over-expressing 3xFlag-L3MBTL3 (kindly provided by Jean-Philippe Lambert of the Gingras laboratory) were grown in high glucose DMEM medium, all supplemented with 10% fetal bovine serum, 1% glutamine, and 1% penicillin-streptomycin.

Lentiviral pLVX-Tet3G and pLVX-TRE3G (empty vector or with 3xFlag-L3MBTL3) plasmids were co-transfected with packaging plasmids into HEK293T cells and lentiviruses were harvested 72 hours later. For infection, IMDM medium with 20% BIT was used, supplemented with 8µg/ml polybreme. K562 cells were infected with pLVX-Tet3G expressing virus, selected with 600µg/ml G418, and subsequently infected with the pLVX-TRE3G expressing virus,
followed by 2µg/ml puromycin selection. After selection was complete, infected cells were cultured in the presence of 300µg/ml G418 and 1µg/ml puromycin.

**L3MBTL3-Histones Immunoprecipitations**

Immunoprecipitations were conducted using a previously described protocol (32). Briefly, HEK293 cells were transiently transfected with Flag-L3MBTL3 using GeneJuice reagent, according to manufacturer's instructions. Cells were resuspended in lysis buffer (50mM Heps-NaOH, pH 8.0, 100mM KCl, 2mM EDTA, 0.1% NP40, and 10% glycerol, supplemented with Complete Roche protease inhibitors), followed by three freeze-thaw cycles on dry ice. Samples were then sonicated, followed by a 1 hour digestion at 4°C using benzonase. Samples were centrifuged at 13.2krmp for 20 minutes and 4°C. The supernatant was then pre-cleared and incubated with magnetic beads, pre-bound to Flag antibodies, overnight at 4°C. Next day, the beads were washed three times with lysis buffer, and eluted with 10mM Tris-HCl, pH 8.0, 150mM NaCl, 2mM EDTA, 1% Triton, 3% SDS, 5mM DTT and 15mM βME for 1 hour at 37°C.

**Chromatin Fractionation**

HEK293 cells were transfected with GFP-tagged L3MBTL3 constructs using GeneJuice reagent, according to manufacturer's instructions. Chromatin fractionation experiments were conducted as previously described (14,33-34) with some modifications. Briefly, ~25 million cells were lysed in 600ul buffer A (10mM Heps, pH 7.9, 10mM MgCl2, 0.34M sucrose, 10% glycerol) supplemented with 1mM DTT, 0.1% NP-40, and Roche Complete protease inhibitor cocktail for 8 minutes on ice, followed by a 5 minute centrifugation at 1,300xg, 4°C. The supernatant was then removed and the nuclear pellet washed in buffer A. The pellet was then resuspended in 300µl buffer A, supplemented with 1mM CaCl2, and 1 unit of micrococcal nuclease (Sigma) and incubated at 37°C. After 10, 30, and 60 minutes of incubation at 37°C, 100µl of digested nuclei were removed and the reactions were quenched by the addition of 5mM EGTA. Digested nuclei were then centrifuged at 1,300xg, 5min, 4°C, and the pellet was lysed with buffer B (3mM EDTA, 0.2mM EGTA, and 1mM DTT, and Roche Complete protease inhibitor cocktail) for 30
minutes at 4°C. Digested and non-digested fractions were then separated by centrifugation, the pellet (insoluble fraction) was washed once more with buffer B, and resuspended in high salt buffer (2.5M NaCl, 50mM Tris-HCl, pH 8.0, Roche Protease inhibitor cocktail) with a syringe.

**BHC Complex Immunoprecipitations**

HEK293T-REx cells were treated with 2µg/ml tetracycline for 24hrs prior to harvesting for immunoprecipitations. K562 cells, stably infected with pLVX-TRE3G 3xFlag-L3MBTL3, were treated with 1µg/ml doxycycline for 4hrs, in the presence or absence of UNC1215, prior to the addition of 10nM PMA for another 48hrs. Fresh doxycycline was added into the media every 24hrs. Fresh lysates were used for immunoprecipitation.

The immunoprecipitation protocol has been previously described (15). Briefly, cells were lysed in cell membrane lysis buffer (10mM Tris, pH 8.0, 10mM KCl, 1.5mM MgCl2, supplemented with Complete Roche protease inhibitors) at 4°C for 30min, followed by a 10min centrifugation at 13.2krpm to pellet the nuclei. The nuclei were then lysed in nuclear lysis buffer (20mM Tris, pH 8.0, 25% glycerol, 420mM NaCl, 1.5mM MgCl2, 0.2mM EDTA, and 0.3% NP-40, supplemented with Complete Roche protease inhibitors) for 30min at 4°C, followed by a 30min centrifugation at 13.2krpm to obtain nuclear extract. 300µg of nuclear extract were then incubated with appropriate antibodies overnight at 4°C. The lysate-antibody mix was incubated with Protein G Dynabeads for 4 hours at 4°C. The beads were then washed 4 times with nuclear lysis buffer and eluted with elution buffer (10mM Tris-HCl, pH 8.0, 150mM NaCl, 2mM EDTA, 1% Triton, 3% SDS, 5mM DTT and 15mM βME) for 1hr at 37°C.

**Antibodies**

The following antibodies were used: Flag (Sigma, F1804), LSD1 (Abcam, ab17721) and CoREST (Millipore 07-455) for immunoprecipitations and immunoblotting; CoREST (Bethyl Laboratories, A300-130A) for immunoprecipitations; Histone H3 (Millipore 05-928), and Histone H4 (Abcam, ab7311) for immunoblotting. The following fluorescently-conjugated
IRDye 800CW secondary antibodies were used: Goat anti-Mouse IgG (H + L) (Li-COR, 926-32210) and Goat anti-Rabbit (H + L) (Li-COR, 926-32211).

**Knockdowns**

L3MBTL3 knockdown in the HFF cell line was achieved using SMARTpool siRNA (Thermo Scientific, M-024729-01) and DharmaFECT transfection reagent, according to manufacturer's instructions, using 2.5µl DharmaFECT reagent and a final siRNA concentration of 250nM per 1mL reaction.

To obtain LSD1 knockdowns in K562 cells, K562 cells were seeded in IMDM medium, supplemented with 20% BIT and 8µg/ml polybrene and infected with virus packaged with plasmids encoding LSD1-targeting shRNA TRCN0000046071 or TRCN0000046072, followed by selection with 2µg/ml puromycin.

**UNC1215 Treatment of K562 Cells**

K562 cells were seeded at a concentration of 200,000 cells/ml and incubated with 3µM UNC1215. Fresh UNC1215 was added every 24 hours. Samples for RNA extraction were collected every 24 hours.

**Erythroid and Megakaryocytic Differentiation**

7 million inducible 3xFlag-L3MBTL3 or empty vector K562 cells were seeded in 25ml media and cultured overnight in the presence of 1µg/ml doxycycline, with or without 3µM UNC1215 (stock concentration was 10mM). Next day, 10nM PMA (or ethanol as the vehicle control) was added to stimulate megakaryocytic differentiation, or 50µM hemin (or NH4OH as the vehicle control) to stimulate erythroid differentiation. Samples were taken every 24 hours post addition of the differentiation agent for a total of 96 hours. Doxycycline and UNC1215 were
supplemented every 24hrs. Megakaryocytic differentiation experiments were repeated in triplicates. A paired t-test was used to analyze significance.

**RNA Isolation, Real-Time-Quantitative PCR, and Primer Sequences**

RNA was extracted using RNeasy Mini Kit or QIAzol Lysis Reagent (Qiagen) and reverse transcribed using qScript cDNA Synthesis Kit (Quanta Biosciences) with random primers. Expression levels of CD41 were determined using TaqMan primer/probe set (Invitrogen, Hs0116228_m1) and normalized against TBP (Invitrogen, Hs00427620_m1). All remaining RT-PCR reactions were performed using SYBR Green SuperMix (Bio-Rad). Real-time PCR reactions were conducted on a 7900HT real-time PCR instrument (Applied Biosystems). Data were normalized against TBP. The following primer sequences were used (written in 5' to 3'):

- **TBP-fwd**: TAGCAGCACGGTATGAGCAACT
- **TBP-rvs**: GGGCATATTGTTGCACTGAGA

- **L3MBTL3-fwd**: ACCAAGGAGGATGGAGAAGAGA
- **L3MBTL3-rev**: TAGGAGGCAAACCTGCTTTAG

- **Hes1-fwd**: AGCACAGAAAGTCATCAAAGCC
- **Hes1-rvs**: CATTTCAGAATGTCCGCCTTC

- **Notch3-fwd**: GTCAGTGGACTCAACACCATCT
- **Notch3-rev**: CCTCATCCTCTTCAGTGCTCGCTC

- **Gfi1b-fwd**: CTAGCTTCTCCTGGGACACCTT
- **Gfi1b-rvs**: TGGAGAAGACCTTGTTGCACT

- **C-myb-fwd**: CAGGTGCTACCAACACAGAACC
- **C-myb-rvs**: GATGGAGTGAGTGGTGTTCTC

- **Spi1-fwd**: CAGAAGACCTGGTGCCCTATGA
- **Spi1-rvs**: CAG TAA TGG TCG CTA TGG CTC TC

- **LSD1-fwd**: GCCGAAGGTAGAGTACAGAGAGA
LSD1-rvs: CATACCCCTCCAGAACTGTCGTC
Meis1-fwd: CTTTGCCTCAGGTCCCGTGA
Meis1-rvs: CCGTAATGGGTAGATCGTCG
CD10-fwd: GGCCCTTTATGGTACAACCTCA
CD10-rvs: TGCAATCAAATCCTCGACCACA
CD61-fwd: TGTAAGAAGTTTGACCGGGAG
CD61-rvs: TGCCAGTGTCCTTAAGCTCTTT
HBG-fwd: CCTGAGAACTTCAAGCTCCTGGAA
HBG-rvs: TGGTATCTGGAGGACAGGGCAC
HBA-fwd: GGTCAACTTCAAGCTCCTAAGC
HBA-rvs: GCTCACAGAAGCCAGGAACTTG
CD71-fwd: GGTCAAGACAGCGCTCAAAAAC
CD71-rvs: CATAACCCCGAGATTCTCCAC

3.6 References

Chapter 4

Closing Remarks and Future Directions
4.1 Thesis Summary

This study has investigated various aspects of L3MBTL3 biology, including its chromatin localization, its interaction partners, and its role in megakaryopoiesis. Future studies investigating L3MBTL3 should aim to pinpoint its role as a transcriptional activator or repressor (which may differ depending on the context), its genomic localization, and its mechanistic role in the BHC complex. An attempt should also be made to confirm the findings regarding its role in megakaryopoiesis in a non-cancer cell line. Other interesting avenues of research would be to further investigate the role of L3MBTL3 in erythropoiesis, in Notch signaling, and in neuronal development - processes L3MBTL3 may be involved in as a result of its association with LSD1 and CoREST. Future research should also focus on attempting to decipher the function of as of yet uncharacterized MBT domain proteins.

4.2 A Growing Understanding of the Mammalian MBT Family of Proteins

While the MBT family of proteins consists of only 3 members in Drosophila, in mammals it has expanded to 9 members. In recent years, a deeper understanding of the biological function of the MBT family of proteins has been gained. Most of the earlier studies used various in vitro biochemical approaches including X-ray crystallography, fluorescence polarization, isothermal calorimetry, peptide pull-downs, and more recently, peptide arrays (1-5). These studies were crucial for the understanding of the various characteristics of the MBT family as a whole, including the fold and structure of the MBT domains, the mechanism of methyl-lysine binding and substrate selectivity. Other studies utilized cell biology techniques and knockout mice to decipher the biological function of the MBT proteins. Interestingly, the knockout mice of different MBT proteins demonstrated distinct phenotypes, with some displaying embryonic (L3MBTL2 and L3MBTL3) or early post-natal (MBTD1) lethality due to various abnormalities, while others (L3MBTL1 and SFMBT1) surviving with no detectable physiological irregularities (6-9). However, various cell biology approaches have shown that L3MBTL1 and SFMBT1 do have important biological functions, such as a role in hematopoiesis in the case of L3MBTL1 (10), and roles in spermatogenesis (11), myogenesis (12), and epithelial to mesenchymal transition (13) in the case of SFMBT1. Therefore, MBT proteins have developed diverse biological roles during evolution, although some redundancy or functional compensation from
other MBT proteins may exist, particularly during embryonic development. The biological functions and mechanistic roles of a number of MBT proteins still remain unknown. The current study attempted to address this gap by studying the biological and mechanistic function of L3MBTL3 in megakaryocytic development. Future studies should continue investigating the biological roles of more poorly characterized MBT family members.

4.3 Novel Approaches in the Study of the Histone Code

The current study used chemical biology as a tool to probe the function of L3MBTL3. While chemical biology is not a new field, its application to the study of epigenetics is emerging as an invaluable investigative approach. A number of studies have reported the use of chemical probes against various methyltransferases (14-18) and against BET bromodomains (19-20) and obtained significant biological insight into the target's biological function. Similarly, the current study demonstrated how the use of UNC1215 can increase our biological understanding of L3MBTL3. UNC1215 turned out to be an invaluable tool in the study of L3MBTL3, particularly due to the as of yet unexplained failure of achieving a successful L3MBTL3 knockdown in certain cell lines. Future studies should continue the challenging task of designing specific, potent, and cell permeable chemical probes against various MBT proteins with the goal of further understanding their biological function and the specific contribution of the methyl-lysine MBT domains. While not necessarily of therapeutic benefit, such chemical probes are instrumental in enriching our molecular and mechanistic knowledge and understanding of the target's biological function.

Another emerging, novel technique that will very likely prove to be immeasurably useful is the clustered regularly interspaced short palindromic repeats (CRISPR)-CRISPR-associated (Cas) system. CRISPR and Cas9 play a role in the bacterial immune response system against bacteriophages by directing degradation of complementary foreign nucleotide sequences. In eukaryotic cells, the CRISPR-Cas system can be engineered to cut out genes, introduce mutations into the genome, target proteins to particular genomic loci, and a number of other applications (21). This may be a particularly useful approach in the case of L3MBTL3 due to the challenging nature of knocking it down, but can also significantly benefit the study of the MBT family of proteins as a whole.
4.4 References


