Characterization of the interaction between MIER1β and lysine 27 of Histone 3

By

Julia Ferguson

A thesis submitted to the

School of Graduate Studies

in partial fulfillment of the requirements for the degree of

Master of Science.

Division of BioMedical Sciences

Faculty of Medicine

Memorial University

October 2015

St. John's Newfoundland and Labrador

ABSTRACT

Acetylation (ac) and tri-methylation (me3) of lysine number 27 on histone 3 (H3K27) play a very important role for early development and differentiation in embryonic stem cells, and often determine which cell type- specific processes are silenced and activated. H3K27 marked genes are often aberrantly expressed in adult cancers. When H3K27 is tri-methylated (H3K27me3), it is an indicator for a transcriptionally inactive gene. When H3K27 is acetylated (H3K27ac), it is an indicator for a transcriptionally active gene. This study focused on determining the binding site for H3K27ac within the mesoderm induction early response 1 (MIER1) protein, a fibroblast growth factor (FGF) inducible early response gene that is known to recruit chromatin modifiers and transcription factors for gene repression. The first 83 amino acids of MIER1 had been previously shown to bind H3K27ac. Alignment verification of the first 83 amino acids of MIER1 between different species revealed highly conserved sequences that represent the potential binding site. Site directed mutagenesis was employed to mutate highly conserved hydrophobic aromatic amino acid sequences in the full-length protein, as well as create deletion constructs. The mutated MIER1 proteins were then tested against wild type MIER1 protein for their ability to bind H3K27ac in peptide pull down assays. The results of this study have localized the C-terminal end of the 83 amino acid binding site to amino acids 36-50, and have shown that the amino acid sequence ^{38}TLE^{40} is required in part for binding.

ACKNOWLEDGEMENT

First and foremost I would like to thank my supervisor Dr. Laura Gillespie who gave me such an amazing opportunity to study in such a fast paced and dynamic field. I will always be grateful for the guidance and direction you have shown to help me better understand how remarkable the world of science can be. I would also like thank Dr. Gary Paterno and Dr. Jon Church for their time and expertise in reviewing my work.

I am very grateful for the irreplaceable technical help from Corinne Mercer and thank you for answering all my many questions. I would also like to thank all of the members of the Terry Fox Lab. Your friendship and laughter will always be treasured and greatly appreciated, especially on those days when nothing seemed to work.

To my parents, siblings, friends, and all of my family that supported me when I needed it the most. I will always be grateful to know how much I was loved and how encouraged I felt even though you all could not be here in St. John's with me. I would like to especially thank my big sisty who was there to listen, hug and keep me going whenever I needed it.

I could not have made it through this degree without my phenomenal husband Josh. Thank you for all the love and support you showed me and always being there to help me try harder and persevere.

Lastly, I would like to dedicate this work to some of the most amazing men I have ever known and will be forever missed. To Mr. James Gates and Mr. Allister Ferguson, thank you for reminding me where I am from and where I can go. You both taught me that life is hard work, and to always apply yourself and never give up.

TABLE OF CONTENTS

List of Figures

List of Tables

List of Abbreviations

µg	microgram
µl	microlitre
^{o}C	degrees Celsius
aa	amino acid
Amp	ampicillin
AD	activation domain
BAH	bromo-adjacent homology
BPTF	bromodomain and PHD finger transcription factor
BRD	bromodomain
BX-C	bithorax complex
CBP	CREB-binding protein
CDYL	chromodomain protein, Y like
ChIP on chip	chromatin immunoprecipitation on chip
CIAP	calf intestinal alkaline phosphatase
CoREST	co-repressor for element-1-silencing transcription factor
DNA	deoxyribonucleic acid
EDTA	ethylenediamine tetra-acetic acid
EED	embryonic ectoderm development
ELM2	EGL27 and MTA1 homology 2
ERα	estrogen receptor alpha
EERα	estrogen related receptor alpha
EZH1/2	enhancer of zeste homolog 1/2
FGF	fibroblast growth factor
g	gram
GNAT	gcn5 N-acetyltransferases
GST	glutathione S-transferase
GST-MIER1	glutathione S-transferase tagged MIER1
h	hour(s)
HAT	histone acetyltransferase
HDAC	histone deacetylase
HDM	histone lysine demethyltransferases
HMT	histone methyl transferase/lysine methyltransferases
HIV-1	human immunodeficiency virus type 1
HRP	horseradish peroxidase
H3K27	histone 3 lysine 27
H3K27ac	histone 3 lysine 27 acetylated
H3K27me1	histone 3 lysine 27 mono-methylated
H3K27me2	histone 3 lysine 27 di-methylated
H3K27me3	histone 3 lysine 27 tri-methylated
IPTG	isopropyl- β-D- thiogalactopyranoside
JMJD	Jumonji domain containing proteins
KAT	lysine acetyl transferase
kDa	kilodalton
KMT	lysine methyl transferase/histone methyl transferase

LB	Luria-Bertani broth
LSD	lysine-specific demethylase
mier1	mesoderm induction early response 1 gene (DNA/RNA)
MIER1	mesoderm induction early response 1 protein
M	molar
MBT	malignant brain tumor domain
Min	minute(s)
Ml	milliliter
MLL	myeloid/lymphoid or mixed-lineage Leukemia
mM	millimolar
MTA-1	metastasis-associated protein
NCoR	nuclear receptor corepressor
NES	nuclear export signal
NLS	nuclear localization signal
NMR	nuclear magnetic resonance
NURF	nucleosome remodeling factor
NuRD	nucleosome remodeling deacetylase
Nt	nucleotide
PCR	polymerase chain reaction
PBS	phosphate buffered saline
PCAF	p300 coactivator associated factor
PcG	Polycomb group
PHD	plant homeodomain
PMSF	phenylmethylsulfonyl fluoride
PRC2	polycomb recessive complex 2
PTM	posttranslational modification
RA	retinoic acid
RbAp46/48	retinoblastoma associated protein 46/68
REST	repressor for element-1-silencing transcription factor
RNA	ribonucleic acid
Rpm	revolutions per minute
SANT	domain first identified in **S**WI3, **A**DA1, **N**CoR, **T**FIIIB
SAM	s-adenosylmethionine
SET	SU(VAR)3–9, enhancer-of-Zeste, Trithorax
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate- polyacrylamide gel electrophoresis
SMRT	silencing mediator of retinoic acid and thyroid hormone receptor
Sp1	promoter specific transcription factor 1
Suz12	suppressor of zeste 12
TRX	trithorax protein

Chapter 1- Introduction

1.1 General Introduction

1.1.1 Structure and Function of Chromatin

Chromatin in a eukaryotic cell's nucleus consists of DNA and proteins.[1,2] The DNA is tightly wrapped around histone proteins which form the nucleosome structure.[1] The nucleosome consist of 146 base pairs of DNA wrapped around the core histone proteins. There are four core histones: histone 2A (H2A), histone 2B (H2B), histone 3 (H3) and histone 4 (H4). The histones are organized into an octamer with two copies of each, as shown in Figure 1.1. All the core histones are organized in a fold like manner of three helices separated by two short strap loops, which are then formed into heterodimers by folding the helices into a "handshake motif" or "helix-strand-helix" motif. These helices are held by the linker histone 1 (H1). This alignment of the histones and DNA in the nucleosome is a major factor in the DNA accessibility and other histone-histone interactions.[1–4] Another important feature on the core histones is the unstructured N-terminal histone tails that protrude from the nucleosome. These histone tails are made up of varying sequences of amino acids for histones H1, H2A, and H2B. However, the histone tail amino acid sequences from H3 and H4 are nearly identical in all eukaryotes, indicating that sequence conservation is essential for function. [2] The histone tails are often subjected to post translational modifications (PTMs). There are many different PTMs, and examples of the most common modifications include: methylation, acetylation, phosphorylation, ubiquitination, sumoylation, and citrullination and is shown in Figure 1.1. [5, reviewed in 6] These PTMs serve as markers for transcriptional regulation as they dictate whether the DNA is accessible or not for undergoing transcription

from the histone proteins, and also serve as markers in various histone-histone interactions. The PTMs ultimately act as mechanisms by which chromatin modification occurs within the cell. The different PTMs will be discussed in the next section.

1.1.2 The Histone Code

For a long time the functions of histone proteins were not known, nor was the significance of the PTM marks on the histone tails. Historically it was believed that the nucleosome core was unchanging, and the positively charged acetyl groups were responsible for loosening the DNA from the histone by neutralizing the charge on the negatively charged DNA, thus allowing the DNA to be transcribed. [7,8] It is now known that within the cell, the nucleosome core is perpetually adjusting in response the cells environment. An example that changes the nucleosome structure is the different PTMs on histone tails that are a signal for gene activation, gene inactivation, recruiting histone modifying complexes, and various other complexes and signals, thus acting as a "code" for what is to happen in that particular site. Drs. Strahl and Allis named this "code" the "histone code" in 2000.[8] They proposed that there is a certain "language" of multiple modifications that is encoded onto the histone tail domains that is read by proteins to cause specific downstream functions.[8,9] The "histone code" that is printed on the N-terminal tails, is crucial since the enhancer DNA itself is not responsible for the initiation of gene expression in the cell. Instead, multiple transcription factors binding to the enhancer DNA mediates the recruitment of chromatin modifiers and other transcription complexes that generate additional regulatory information from the "histone code".

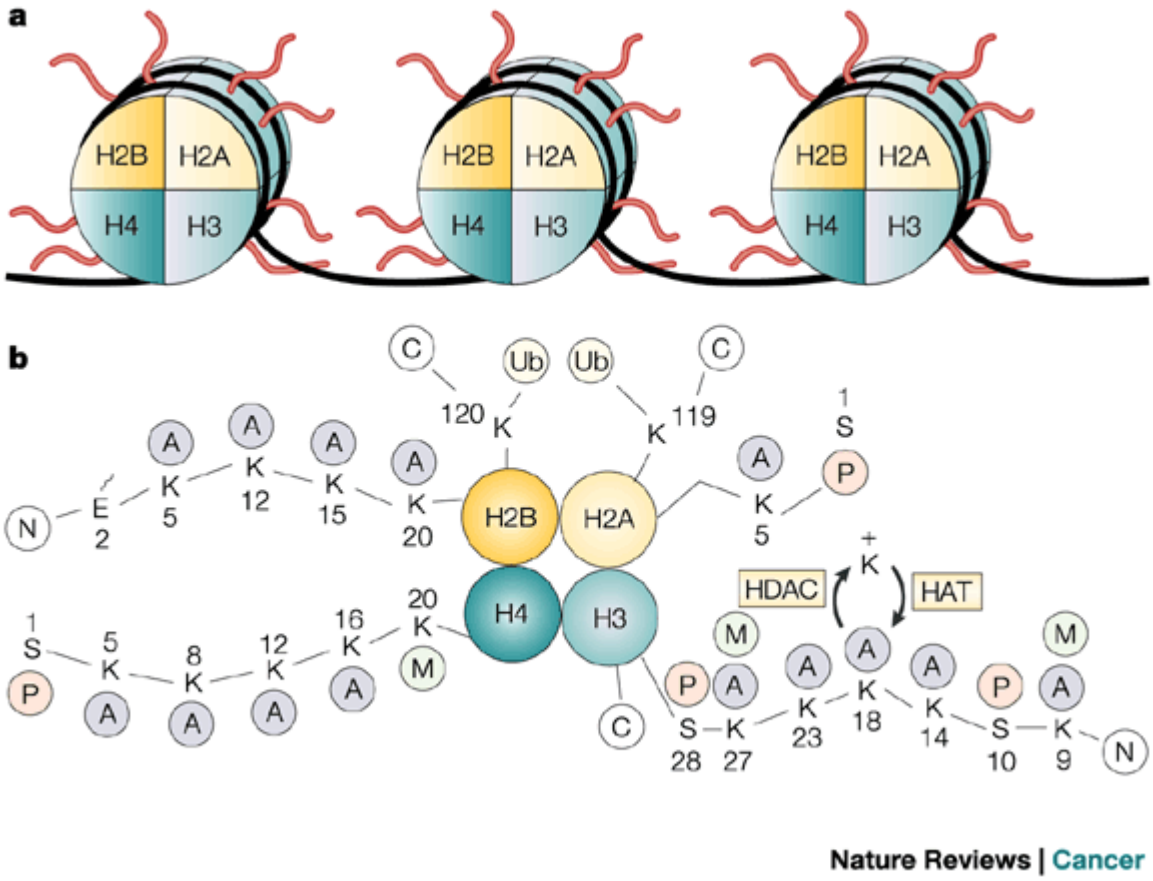

Figure 1.1: Histone structure

a. There are two copies of each of the four core proteins: H2A , H2B, H3, and H4 forming an octamer. The DNA is wrapped around the histone proteins (shown in black) to form the nucleosomes. The N-terminal tails of core histones are shown in red protruding from the histone proteins. B. The N-terminal tails protruding from the core histones are broken down into the known amino acids that can be post transitionally modified with the position number of the amino acid (aa) beside it. Lysines (K) are potentially acetylated/deacetylated by histone acetyltransferases (HATs) and histone deacetylases (HDACs). The lysine 18 on histone 3 is shown to be modified by HATs or HDACs in this schematic. Reprinted from Nature Reviews Cancer, Paul A. Marks, Richard A. Rifkind, Victoria M. Richon, Ronald Breslow, Thomas Miller & William K. Kelly. Histone deacetylases and cancer: causes and therapies, pp. 194-202. Copyright 2001, with permission from RightsLink / Nature Publishing Group)[10]

Since the environment within the cell is constantly changing, and DNA transcription does not need to continuously occur, the histone code is only temporal and each modification occurs only when needed. The "DNA code", however, is permanent and contains all the information for printing and manufacturing the "histone code". [11]

To access the information preserved on the DNA certain histone modifying proteins or protein complexes are recruited to each particular modification on the histone tail to "read" what the modification is and then begin downstream effects. Other modifying complexes could be recruited to "write" or "erase" modification by adding or removing a PTM, respectively. These three modifying complexes work together in regulating transcription of each respective target gene. [12,13]

1.1.3 Writers and Erasers of Methylation

There are different readers, writers and erasers for the various types of PTMs on histone tails. But there are only a few amino acids that are covalently modified. These are lysine (K), arginine (R), serine (S), threonine (T), tyrosine (Y), histidine (H) and glutamic acid (E).[14] The most common covalent modification that occurs is methylation and acetylation on lysines.[15] The acetylation modification on histones was first identified by Alfrey in 1964 as a potential regulator of RNA synthesis. Scrutinization of this process since then has led to the discovery that lysine acetylation plays a key role in transcriptional activity: essentially genes being turned "on". [14,16] The methylation modification on histones was first identified by Murry in 1964 but the significance of the modification was not known until decades later.[17] More recent discoveries have found that lysine methylation is often correlated with transcriptional

repression and condensation of the chromatin: essentially the genes being turned "off". However, depending on the location of the methylated lysines, this modification can be an indicator for gene activation, which will be discussed further on in section 1.1.6.[17,18]

Histone lysine methylation plays an essential role in chromatin organization and function. This modification is connected to many processes including genomic imprinting, chromosome X- inactivation, heterochromatin formation and transcriptional regulation.[19] The writers of histone methylation are histone lysine methyltransferases (HMTs or KMT's). HMTs add methyl groups to lysine residue from the co-factor S-adenosylmethionine (SAM) onto the amine of the specific lysine residue, as shown in Figure 1.2. The ε-amino group can be mono-,di-, or tri-methylated, and each methyl group added can change the function.[19] Methylation causes a condensed DNA chromatin structure that is referred to as heterochromatin. Other than the disruptor of telomeric silencing-1- like (Dot1L), all HMTs use an evolutionally conserved SET (SU(VAR)3–9, enhancer-of-Zeste, Trithorax) domain for lysine methylation.[20] Compared to other histone modifying enzymes, HMTs often methylate specific lysine residues. There are around 10 human methyltransferases, and such examples of these are: SET9, SET1, MLL, and SMYD3 methylate H3K4; G9a, EHMT1/2, and SUV39H1/2 methylate H3K9; G9a and EZH2 methylate H3K27; NSD1; SETMAR methylate H3K36; DOT1L methylates H3K79; SET8 and NSD1 methylate H4K20. [19–21]

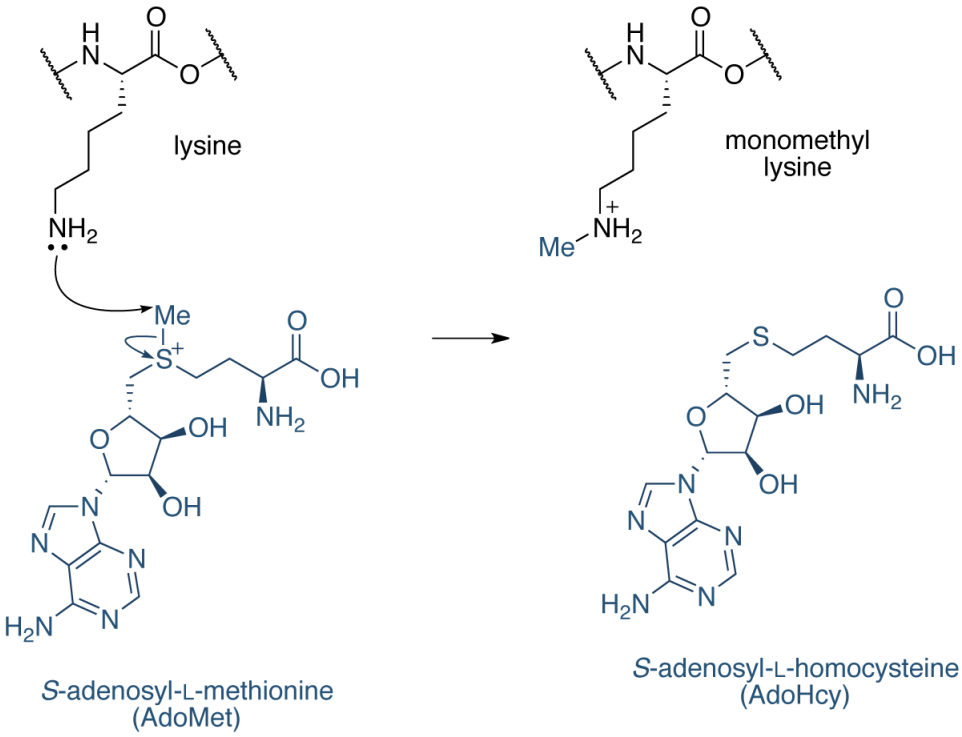

Figure 1.2: Lysine Methylation

The writers of histone methylation are histone lysine methyltransferases (HMTs or KMT's). They add a methyl group to lysine residue from the co-factor S-adenosylmethionine (SAM) onto the amine of the specific lysine residue. (Reprinted from ATDbio, Epigenetics. Copyright. 2005-2015, with permission from ATDbio Ltd)[22]

Histone lysine demethyltransferases (HDMs) are responsible for removing the methyl group from the lysine residue. The HDMs are divided into two families: lysine-specific demethylase (LSD) proteins and Jumonji domain (JMJD) containing proteins. The JMJD family includes: FBXL, JMJC1-3, PHF, JARID, and other JMJD subfamily proteins.[19,20]

LSD proteins are linked to different modifying complexes involved in transcription, heterochromatin spreading, and stress-induced responses. LSD1 demethylates H3K4me2/me1, H3K9me1/2, and other non-histone substrates, such as p53. FBXL proteins are involved in regulating transcriptional elongation by demethylating H3K36me1/2. The JMJD1 proteins demethylate genes related to tumor suppression, androgen receptor targeting, and thyroid receptor responding by demethylating H3K9me2. [20] JMJD2 proteins remove methyl groups from H3K9me2/3, which is associated with transcriptional elongation and transcriptional repression. JMJD3 proteins are specific to demethylating H3K27me2/3, which is essential for normal development, cell fate decision, and halting pluripotency. PHF proteins demethylate H3K9me2 and H3K27me2 repressing marks. JARID proteins demethylate the coding regions of transcribed genes on H3K4me2/3 and have also been known to work as a transcriptional co-repressor. Different point mutations found in X-linked mental retardation are linked to JARID activity. [23] JARID2, also known as Jumonji, is phylogenetically related to the JARID1 family, but no related enzymatic activity has been shown. [19,20]

1.1.4 Writers and Erasers of Acetylation

Histone acetyltransferases (HAT's or KAT's) are the writers of acetyl modifications on lysines. They neutralize the positive charge on the ε-amino group by transferring an acetyl group from acetyl-CoA onto the lysine residue. This leads to accessibility of the chromatin for transcriptional activation, as shown in Figure 1.3. This accessible state of chromatin is also referred to as euchromatin.[21] The different types of HATs that modify lysines are classified by their catalytic activity. These include: Gcn5 N-acetyltransferases (GNATs) (which contains Gcn5, PCAF, Elp3, Hat1, Hpa2 and Nut1), the Morf, Ybf2 (Sas3), Sas2 and Tip60 (MYST) HATs, and other proteins including p300/CBP (CREB-binding protein). Taf1 and other nuclear receptor co-activators, have been reported to show HAT activity, but do not contain true consensus HAT domain. [24,25] While acetylation of histone tails is mostly linked to transcription, acetylation of non-histone substrates has also been linked to DNA repair and replication.[6]

Histone deacetylases (HDACs) are the erasers of the acetylation modification by removing the acetyl motif from the ε-amino group of the lysine. The removal of the acetyl group which is known as N^{ε}-deacetylation[26] is, shown in Figure 1.3. This causes a decrease in the space between the nucleosome and DNA that is wrapped around it. This diminishes accessibility for transcription factors, and leads to transcriptional repression. There are 18 different HDACs classified into two families: the classical and silent information regulator 2 (Sir2)-related protein (sirtuin), and four classes depending on their similarity of yeast proteins. Class I includes: HDAC 1, HDAC 2, HDAC3, HDAC8 and are similar to the deacetylase yeast protein RPD3. Class II includes: HDAC 4, HDAC 5, HDAC6, HDAC 7, HDAC 9,

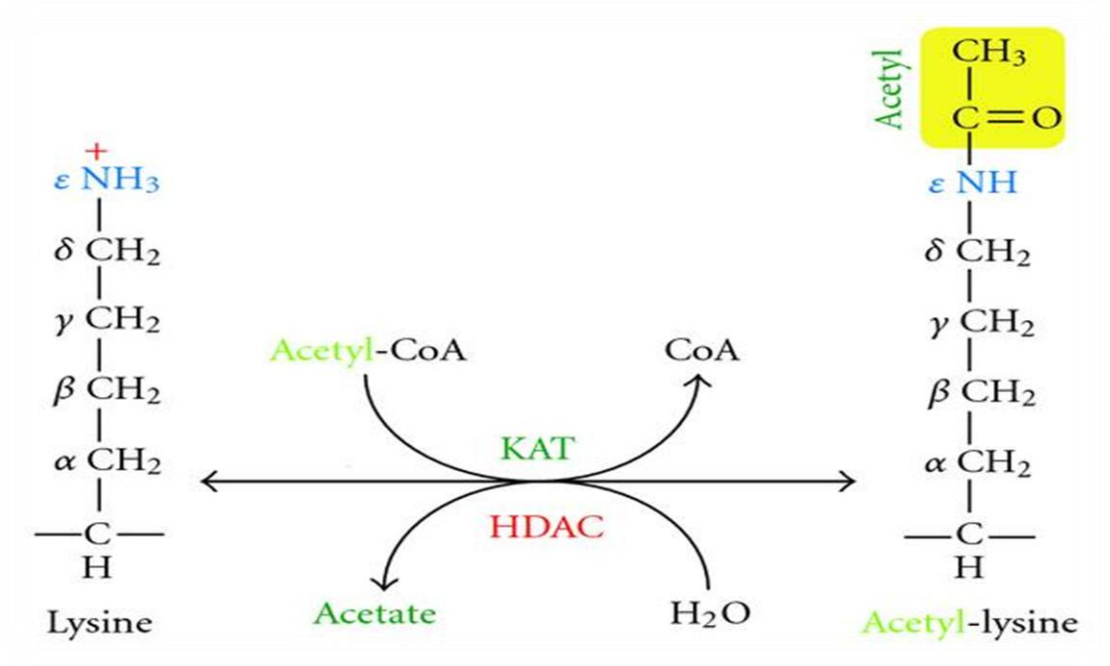

Figure 1.3: Lysine Acetylation and Deacetylation

HATs or KATs write the acetyl modification on specific lysines by transferring an acetyl group from acetyl-CoA to the lysine residue on the ε-amino group, neutralizing the positive charge. HDACs remove the ε-amino group from the lysine and release acetate. (Reprinted from International journal of cell biology, Kim, Go-Woon Gocevski, Goran, Wu, Chao-Jung Xiang-Jiao Yang, Dietary, Metabolic, and Potentially Environmental Modulation of the Lysine Acetylation Machinery, pp. 351-365. Copyright 2010, with permission from Elsevier)[27]

HDAC 10 and are similar to the yeast protein HDCA1. [28] HDAC6 and HDAC10 have two active sites and therefore often classified as class IIA. In the catalytic center of HDAC11 are conserved residues found in both class I and class II of HDACs. Therefore, it is sometimes placed in class IV. These deacetylases are known to be inhibited by vorinostat [suberoylanilide hydroxamic acid (SAHA)] and trichostatin A (TSA). Class III in mammals contain deacetylating sitrunis proteins (SIRT1, SIRT2, SIRT3, SIRT4, SIRT5, SIRT6, SIRT7) and Sir2 in yeast proteins.[26,28,29]

HDACs need to be able to enter the nucleus because often their main target protein or protein complexes are found in the nucleus. HDACs do this through their nuclear localization signal (NLS) or by co- localization in conjunction with other proteins HDACs. Most HDACs contain a NLS, and are found in the nucleus. However, some HDACs are found in the cytoplasm, depending on their different catalytic domains.[29] Class I HDACs are found almost exclusively in the nucleus.

HDAC1 and HDAC2 are exclusively nuclear, since they do not have a nuclear export signal (NES). HDAC3, however, contains NLS and a NES. This suggests that HDAC3 can be recruited to the cytoplasm. Class II HDACs are able to exit and enter in and out of the nucleus in response to certain cellular signals, however, HDAC6 is predominantly found in the cytoplasm. Class IV is generally found in the nucleus; but HDAC11 has been shown to be co-localized to the cytoplasm with HDAC6. [29,30]

Within the past decade, HDAC inhibitors have become widely produced and advanced into clinical drugs for the treatment of certain cancers, Alzheimer's, arthritis, and even anti-latency therapy for persistent human immunodeficiency virus type 1 (HIV-1) infection.[31–34] Because HDACs play an important role in regulating cellular gene expression, differentiation,

development, and the maintenance of cellular homeostasis, this altered expression of HDACs often plays an active role in tumor growth and progression. Aberrant expression of HDACs activity has been found in tumor samples. Different studies have found that there is an increase in HDAC1 expression in gastric, prostate, colon, and breast carcinomas. Other studies have found that the over-expression of HDAC2 has been noted in cervical, gastric cancers, and in colorectal carcinoma. Also, high levels of HDAC3 and HDAC6 have been reported to be expressed in colon and breast cancer specimens.[35–41]

Applying HDAC inhibitors towards classes of HDACs that have known involvement in modifying complexes could be an important application for clinical treatments. Class I HDACs have been reported to be associated with at least four major chromatin modifying complexes that control gene silencing. HDAC3 is recruited uniquely to the silencing mediator of retinoic acid and thyroid hormone receptor (SMRT) and nuclear receptor co-repressor (NCoR) repression complexes.[42,43] HDAC1 and HDAC2 are activated through recruitment into several co-repressor complexes including the Sin3A[44], co-repressor for element-1-silencing transcription factor (CoREST)[45], and nucleosome remodeling deacetylase (NuRD)[46,47] complexes.

One study found the need for the compound inositol phosphate: D-myo-inositol-1,4,5,6-tetrakisphosphate [Ins(1,4,5,6)P4] in the binding pocket of HDACs to act as a conserved regulator of Class 1 HDACs and their recruited repressor complex.[31] They found that HDAC3 forms a complex with the extended SANT domain from the SMRT co-repressor.[48] Also, the interaction between HDAC3 and the SMRT-SANT domain has been shown to require the presence of Ins(1,4,5,6)P4. By using sequence conservation, other class I HDAC complexes are thought to bind inositol phosphates. The interaction between HDAC1 and metastasis-

associated protein 1 (MTA1) from the NuRD complex uses the MTA1's EGL27 and MTA1 homology 2(ELM2) and SWI3, ADA1, NCoR, TFIIIB (SANT) domains. The structure reveals that this complex also has an inositol phosphate binding pocket at the interface between the MTA1- SANT domain and HDAC1. The HDAC1:MTA1 structure provides insight into how the combination of the ELM2-SANT domains are able to recruit histone deacetylase enzymes. Together, the two domains make very extensive interactions with HDAC1 by wrapping completely around the catalytic domain and covering both the carboxy- and amino-terminal ends. This suggests that the two domains might be a common feature in class I HDAC: co-repressor complexes and could be used in the future for developing specific inhibitors to mimic the inositol phosphate interaction with the protein and the ELM2-SANT domains.

1.1.5 Reader Domains and Their Targets

Readers of modified amino acid residues have conserved effector domains within proteins that recognize and bind to specific modifications and cause downstream effects that will either stabilize the chromatin or recruit other complexes to further modify the chromatin. Reader proteins typically use a cavity or surface groove to hold the modified histone residue, and they determine what the modification is, and/ or determine the number of modifications. For example a reader protein may determine whether the amino acid is mono, di- or tri-methylated.[14,49] These domains include chromodomains, bromodomains (BRD), WD40 repeats, tudors, plant homeodomain (PHD), and malignant brain tumor (MBT).[9,50] BRDs recognize the acetyl modification on the ε-N-acetylated lysine residues and some PHD fingers

can recognize various acetylated lysines.[15,51] Chromodomains, WD40 repeats, tudors, certain PHD fingers, and MBT domains recognize methyl modification on specific lysine residues.[14,49] The downstream effects caused by reader proteins and their respective domains, along with their associated complexes will be discussed further on in this section.

BRDs selectively target acetylated lysines. In humans, there are over 40 known proteins that contain more than 60 different BRDs. BRDs fold into 120 amino acid, a four α helix structure (αZ, αA, αB and αC), which is linked by loop regions of varying length (ZA and BC loops) where the acetyl binding site is located. BRD containing proteins often cooperate in chromatin remodeling complexes, transcriptional control, and methyl or acetyltransferase activity. The organization of BRD-containing proteins is evolutionarily conserved, and the BRD domain is often found with other epigenetic reader domains. Observed combinations include the PHD, as well as multiple BRDs, and various other domains that are known to moderate protein interactions, an example is the bromo-adjacent homology (BAH) domain.[52–55]

The recognition of acetylated histones through the BRD uses a small selection of conserved amino acids. One study used site-directed mutagenesis to determine which of these amino acids specifically bind acetylated lysine residues found in a cluster of hydrophobic amino acids on the inner structure (often called a hydrophobic pocket) of a BRD.[56] They mutated amino acids ^{809}Y, ^{802}Y, ^{760}Y and ^{752}V in the BRD of the HAT co-activator PCAF (p300/CBP-associated factor). These are highly conserved hydrophobic and aromatic amino acids throughout the BRD family, and indicate that they are an important feature of the BRDs binding ability. They found that substitution of alanine for ^{809}Y completely obliterated

binding to the lysine-acetylated H4 peptide, whereas the Y802A, Y760A and V752A mutants had significantly reduced ligand-binding affinity. In order to determine if the mutations disrupted the overall BRD fold, they compared the 3D structure using 15N-HSQC (Heteronuclear Single Quantum Coherence) spectra of the mutants to that of the wild-type protein. The Y809A mutant was only affected for a few residues near the mutation site. However, mutations of the other residues in the hydrophobic binding pocket caused greater changes in the protein conformation, particularly in the ZA loop. Nuclear magnetic resonance (NMR) structural analysis and mutagenesis have shown that certain amino acids (^{809}Y), are essential for the BRD interaction with the acetyl group of acetylated lysine, and other amino acids (^{802}Y, ^{760}Y and ^{752}V) imply both structural and functional roles in the recognition.[56]

The PHD fingers are readers of PTMs and were first discovered by Schindler in 1993: who, noticed a stretch of cystines that were always between two plant homeodomains in the protein HAT3 in *Arabidopsis thaliana*.[57] The PHD finger motif is about 60 amino acids in length, and resembles the metal binding RING domain, which binds four cysteines, one histidine, three cysteines (C4-H-C3) and binds two Zn^{+2}. In yeast and fly the PHD proteins are nuclear and bind DNA or chromatin, to mediate control over transcription, and the chromatin structure. As with other writers or erasers, reader containing proteins often target specific sites. For example, PHD containing proteins can be sorted into two categories depending on what they have been shown to target. The first category of PHD fingers target H3K4me3 and unmodified H3. The second category of PHD fingers target H3K9me3, and various acetylated lysine residues on H3 and H4[reviewed in 51,57–60]

Chromodomains are readers of methylated lysines; they are folded into three anti-parallel beta strands against a C-terminal alpha helix that mediate protein–protein and/or protein–

nucleic acid interactions. The chromodomain is evolutionally conserved and is a member of the OB-fold class found in DNA binding proteins of archaea, bacteria and eukaryotes. Chromodomain proteins are divided into three classes: proteins with a single chromodomain, proteins with an N-terminal chromodomain with a structurally similar C-terminal chromo shadow domain, and proteins with paired tandem chromodomains. The chromodomain binds to the methylated residue as a beta strand in a conserved groove on the surface of the chromodomain, which forms a beta sandwich. Different proteins containing chromodomains target certain lysine residues such as: CHD1 on H3K4, HP1 on H3K9 and Polycomb proteins on H3K27. [58,61]

Tudor and MBT domains have also been shown to recognize the methyl group on lysine residues, and are known as the 'royal family' along with chromodomains for their similar structure and target residues. Tudor domains have been shown to bind to H3K4, H3K9 and H4K20. MBT domains have been shown to target H3K4 and H4K20. [reviewed in 6,14,50]

Often reader effector motifs are found in a complex of proteins that control epigenetic regulation along with writer and eraser proteins; this allows for a modification to be analyzed, written on or removed, thus controlling the gene activity surrounding that histone modification. Proteins with BRDs, chromodomains, PHD and MBT domains have been shown to act together with other chromatin modifiers in complexes to remodel the chromatin. There are many well-known examples of these complexes. The NuRD complex which is involved in ATP-dependent chromatin remodeling and histone deacetylase activities contains: the histone deacetylase proteins HDAC 1/2 , the histone-binding proteins RbAp46 and RbAp48, the MTA1 (or MTA2 / MTA3), the methyl-CpG-binding domain protein MBD3/2, and the chromodomain binding protein CHD3/CHD4/CHD5. [62,63] The NURF

complex is an ISWI-containing ATP-dependent chromatin-remodeling complex. The NURF subunit BPTF (BRD and PHD finger transcription factor), associates with chromatin and recruits the associated ATPase, SNF2L (also known as ISWI and SMARCA1), to the HOXC8 promoter. This complex is recruited to H3K4me3 to maintain Hox gene expression patterns during development.[60,64] The co-repressor NCoR1 remodeling complex (contains TIFβ protein with a BRD and PHD finger) promotes heterochromatin formation by interacting with the heterochromatin chromodomain containing protein HP1.[65] There is a known REST/CDYL/G9a complex where CDYL(chromodomain protein, Y like) bridges REST(repressor for element-1-silencing transcription factor), a neuronal gene repressor and histone methyltransferases for gene repression and suppression of cellular transformation.[66]

1.1.6 Known Methylated and Acetylated Lysine Residues and their Functions

Acetylated lysine residues distinguish transcriptional activation. However, acetylation of specific lysines on a particular histone tail does not seem to be as important compared to the overall acetylation level of histone proteins. Although, HATs are not residue specific, HMTs are often lysine specific to what they methylate. HATs are able to acetylate many different residues on histone proteins, and are limited to which histone they can acetylate.[25,67] HAT enzymes have tissue specific functions, and different HAT enzymes will target different lysine residues at stage specific developmental times.[24] Acetylated lysines can also act as active enhancers elements in ES cells for starting tissue-specific patterning.[68]

Lysine residues that are methylated can be an indicator for transcriptional activation or repression. Genes that are methylated at lysine 4 (H3K4), lysine 36 (H3K36) and lysine 79

(H3K79) of histone H3 are indicators of active chromatin. Whereas methylation at lysine 9 (H3K9) and lysine 27 (H3K27) of histone H3 and methylation at lysine 20 of histone H4 (H4K20) are indicators of repressed chromatin. Di-methylation at H3K4 can indicate both inactive and active genes, whereas tri-methylation is only found on active genes.[19] The HMT that adds the methyl modifications are also often lysine specific as previously stated. The levels of methylated lysines can also vary depending on the stage of development and the tissue specificity in which they are located.[69,70]

1.1.7 H3K27 Acetylation and Methylation

Histone 3 lysine 27 can be acetylated, mono-, di-, and tri-methylated. However, the acetyl or methyl modifications are placed separately, and do not occur at the same time. The acetylated modification on lysine 27 was first discovered in yeast and has since been studied in mice and humans.[71,72]

Acetylation of H3K27 (H3K27ac) has been shown to be associated with many active genes. One study found that the trithorax proteins (TRX), is required for H3K27 acetylation. The TRX proteins and their founding member trithorax belong to a collection of proteins that are responsible for the maintenance of gene expression. They are typically found in activating gene expression.[73] Myeloid/lymphoid or mixed-lineage Leukemia (MLL), is the human homolog of trithorax. One study found that trithorax associates with CBP and was required for maintenance of transcriptionally active states and thus repressing Polycomb silencing.[74] This will be discussed further on in the section. H3K27ac is found in high levels in the early stages of *Drosophila* embryos and begins to decline after four hours as the opposite

methylation modification occurs and H3K27me3 is increased.[74] The same study found that knockdown of the methyltransferse E(Z), which is responsible for methylating H3K27, decreases H3K27me3 and increases H3K27ac in bulk histones. Moderate over-expression of CBP *in vivo* caused a global increase in H3K27ac and a decrease in H3K27me3. This decrease in methylation levels increased Polycomb mutants due to lack of methyl dependent silencing. The TRX over-expression also led to an increase in H3K27ac and then a decrease in H3K27me3, resulting in defects in Polycomb silencing.[74]

H3K27ac acts as an important enhancer mark that can determine active from poised enhancer elements. When enhancer proximal genes lack H3K27ac enrichment, they display lower expression levels compared to the average enhancer proximal gene. This modification can be written by the CREB binding protein (p300/CBP), and is associated with active promoters in mammalian cells.[68]

The other PTM methylation, particularly tri-methylation on H3K27 is a repressive identifier of transcription. H3K27me3 is important in the differentiation of cells. Tri-methylation is catalyzed by the Polycomb recessive complex 2 (PRC2) which is part of the Polycomb group (PcG) proteins. The repression of most developmental gene regulation through histone modifications requires the PcG proteins. They are highly conserved transcriptional repressors and are required for correct body patterning since they control Hox gene expression.[75,76] Components of PRC2 are essential in the earliest stages of vertebrate development. The Polycomb proteins were first studied in *Drosophila* and through gene deletion studies. In *Drosophila* the core subunits are: enhancer of zeste [E(Z)], suppressor of zeste (Su[z]12), extra sex combs (ESC), and multicopy suppressor of IRA1 (MSI1).[77] In humans the PRC2 core components are: enhancer of zeste homolog 1/2 (EZH2), suppressor of

zeste 12 (SUZ12), embryonic ectoderm development (EED), and retinoblastoma associated protein 46/68 (RbAp46/48). EZH2 is a H3K27 methyltransferse, and SUZ12 is required for the methyl activity of EZH2.[78] The EED subunit is a reader of the methyl modification on H3K27, and RbAp46/48 has acts as a histone chaperone with H3.[49,78]

It has been shown that PRC2 and Polycomb proteins are essential for development, (especially in homeotic conversions) and PRC2 is an important regulator in many cell types. Some cell types regulated by PRC2 include human stem cells; adult and embryonic (ES), neural, epidermal and haematopoietic. In human ES cells PRC2 works with pluripotency factors Oct4, SOX2 and NANOG to stop lineage-specific genes and keep the pluripotent state.[79] Most H3K27me3 targets of PRC2 are specified in stem cells, and early embryogenesis. They often target developmental processes such as gastrulation, mesoderm formation and neurulation, and either the methylation modifications are maintained on H3K27me3 or they are lost upon transcriptional activation during development.[81]

In mouse ES cells it was found that 512 genes bound by H3K27me3 are occupied by components of the PRC2 complex.[81] These regions show a notable enrichment in genes related to transcription and development. The transcription and development processes include: organogenesis, morphogenesis, pattern specification, neurogenesis, cell differentiation, embryonic development, and cell-fate commitment. Also, the target genes within development and transcription functional groups overlap, which indicates that most of the PcG target genes are for transcription factors and have crucial roles in a variety of developmental processes.[81]

In different species it has been shown loss of PRC2 results in non functioning embryo or death. Using temperature sensitive *Drosophila* mutants of E(Z), that loss of functional E(Z)

caused a loss of PcG protein binding to the chromosomes, and deprived histones of K27 methylation. It has also been shown that an HMTs negative SET domain mutant *Drosophila* E(Z) does not repress HOX genes *in vivo*, which implicates that H3K27 is the essential substrate that PRC2 modifies for Polycomb repression in differentiation.[77,82] In mice removal of PRC2 and other PcG proteins is lethal after implantation of the embryo.[83]

The role of tri-methylated and acetylated H3K27 is also important in the developmental the bithorax complex (BX-C). The BX-C in *Drosophila* is a group of homeotic genes that give rise to body segmentation plan.[84] The expression of each gene is governed by cis- regulatory domains with one for each parasegment. Repression of the domains is dependent on the PcG proteins which controls H3K27me3 of H3. Parasegment-specific indication of PcG function on H3K27me3 locations across the BX-C in successive parasegments shows a "stairs" pattern that revealed the boundaries of the BX-C regulatory domains function. Acetylated H3K27 is broadly enriched across active domains, in a pattern harmonious to that of H3K27me3.[84]

H3K27me3 and H3K27ac are mutually exclusive, but either can occur together with the modification H3K4me3 in regions referred to as bivalent domains.[74] Bivalent domains occur in nucleosomes with two modifications simultaneously occurring. Bivalent domains often occur in the promoter of specific transcription factors, and are thought to keep the genes poised in order to respond to developmental cues. These cues dictate whether or not the targeted genes will then be activated or repressed depending on the received signals.[85] These domains can also occur in adult tissue as well.[86] H3K27me3 can occur in large domains spanning several hundred of kilobases or found individually on genes. It has been found that these domains can be associated with gene families, such as the Hox gene clusters.[79]

1.2 Mesoderm Induction Early Response 1: MIER1

1.2.1 MIER1

Mesoderm Induction Early Response 1 (MIER1) was first identified in *Xenopus laevis* (*Xmier1*) as a fibroblast growth factor (FGF) activated gene. Its expression levels were found to increase after the induction of mesoderm differentiation of FGFs in *Xmier1*.[87] It is highly evolutionarily conserved between species, and a later study found the amino acid sequence of *Xenopus laevis* and human ortholog shared 91% similarity. There is also a 95% amino acid sequence similarity between the human and mouse orthologs. The SANT domain of MIER1 between the human, mouse and *Xenopus* orthologs are idenitical.[88,89] The human *MIER1* gene is located on chromosome 1p31.2, spans 63 kb, and consists of 17 exons. It encodes six protein isoforms. Through different splicing of a facultative intron, the human *MIER1* gene produces distinct carboxyl termini isoforms (α and β), and will be discussed in the next section.[90] During early embryo development, *Xmier1* mRNA levels are evenly expressed in early cleavage, then slightly increase at blastula stage and decrease by 6-fold during gastrula, neurula and tailbud stages. It is then undetectable in the following developing stages.[87] Further studies also determined that the subcellular location of MIER1α changes with breast cancer progression in ductal epithelial cells. There is a dramatic shift of MIER1α from the nucleus to the cytoplasm during the progression of invasive carcinoma.[91] Investigation of the expression pattern in human tissue revealed that the human MIER1α isoform is found in specific cell types for almost every reproductive and endocrine tissue except the thyroid gland, which suggests it has particular function in endocrine tissue.[92]

The common internal region in MIER1 contains several domains used for its role in regulating transcription (shown in Figure 1.4). These domains include an N-terminal acidic region that functions as a transcriptional activator.[87] An ELM2 domain that has been shown to recruit HDAC1 and repress transcription.[93] A SANT domain shown to interact with the promoter specific transcription factor (Sp1) to interfering with transcription.[94] The possible histone binding region in MIER1 has been shown to bind to H3K27ac. (Paterno et al, unpublished) In addition, MIER1 has been shown to bind CBP in its N-terminal domain (the histone binding region as well as acidic and ELM2 domains) and inhibit its HAT activity.[95] MIER1 has been shown to interact with HMT G9a [96], and its role with this chromatin modifying enzyme has yet to be determined.

1.2.2 MIER1 Structure

A number of studies have been conducted to characterise the structure and function of MIER1. The *mier1* gene encodes for six different protein isoforms with the same common internal region, through exon skipping, a facultative intron, three different polyadenylation signals, and two alternate promoters (P1 and P2).[90] In humans there are three different N- and two different C- termini. The different N-terminal isoforms (named N1, N2 and N3 in GenBank) result from the use of the two different promoters and alternative splicing. Use of the P1 promoter can produce the exon 3A with the inclusion of the cassette and transcribes the isoform N1. Inclusion of 3A encodes for an additional sequence containing a *bona fide* nuclear export signal (NES).[97] The P1 promoter also encodes for N2 isoform and the use of P2 promoter encodes the N3 isoform. The N2 and N3 isoforms (named N2 and N3 in

GenBank) have different start sequences, (ML in N2 and MAE in N3) but encode the same protein sequence after the start sequence.[97]

The two C-terminal isoforms known as MIER1α and MIER1β are created by the use of a facultative intron. The α isoform results from the removal of the facultative intron and encodes a much smaller protein with the C-terminus consisting of 23 amino acids. MIER1 α contains a LXXLL motif which is important in the interaction with nuclear hormone receptors.[90] MIER1α has been shown to interact with estrogen receptor (ERα) in MCF7 cells and MIER1α over expression has been shown to inhibit estrogen-stimulated anchorage-independent growth. [91] MIER1α is found in the cytoplasm because it does not have an NLS but it can be shuttled into the nucleus via interaction with nuclear proteins HDAC1/2.[98] The β isoform results from the inclusion of the facultative intron and encodes a much larger protein with 102 amino acids in the C-terminus. The β isoform has a NLS and is more widely expressed than MIER1α.[90] There are two different distinct amino acid sequences for MIER1α and MIER1 β isoforms, which suggests that they both have distinctly different functions. For my experiments, I will be using the MIER1β isoform since it is the more commonly expressed isoform.

Studies of MIER1, and its different isoforms have shown it shares similar domains to other known epigenetic regulators, such as the: acidic region, a possible histone binding region, ELM2 domain, and the SANT domain,[87] as shown in Figure 1.4. Thus, MIER1 recruits chromatin modifiers through its conserved domains and transcription factors to suppress transcriptional activity. The acidic region has an activation domain (amino acids 1-175) containing four stretches of acidic amino acids at the N-terminus that have been shown to

activate transcription.[87] This function is often found in proteins that bind transcription factors.[99,100] The histone binding region has been shown to bind H3K27ac and has been narrowed to amino acids 1-83 at the N-terminus (Paterno et al., unpublished data). Currently it is known that the transcriptional activation function occurs in amino acids 1-109, and the histone binding region in amino acids 1-83. Neither has been characterized to the smallest possible region, which is why they overlap.

The SANT domain, named for the transcription factors : SW13, ADA2, N-CoR and TF1113 (giving the acronym SANT) was originally found based on its homology to the DNA binding domain c-myb.[101] The SANT domain was discovered as a small motif around 50 amino acids in length, and is present in many nuclear receptor co-repressors.[102] Proteins that contain one or more SANT domains are often involved in protein-protein interactions. The SANT domain has been shown to interact with complexes that contain HMTs[75], HATs[103] and HDACs.[104,105] SANT domains are found in proteins regulating development including: Egl-27 and CoREST[106], N-CoR[101] and SMRT[102]. SANT domains have also been shown to be a part of the chromatin remodelling complexes MTA-1 and MTA-2.[45,75]

The ELM2 domain (EGL-27 and MTA homology 2) was first identified in Egl-27, a *Caenorhabditis elegans* protein and encodes a protein similar to MTA1.[107] This similar protein is a component of a complex with ATP-dependent nucleosome remodelling, and histone deacetylation activities. It is also involved in protein complexes that are known to regulate the activity of transcription factors involved in embryonic patterning.[107] The ELM2 domain is evolutionarily conserved. MIER1 contains both an ELM2 domain and a SANT

domain. There are 13 known human co-repressor complexes that contain both domains.[31] This implies a structural, and/or functional relationship between the domains.

1.2.3 MIER1 as an Epigenetic Regulator

MIER1 acts as an epigenetic regulator enforcing transcriptional repression through two distinct mechanisms. The first mechanism is the interaction with transcription factors estrogen receptor α (ER) and Sp1. The MIER1α isoform has been shown to interact with estrogen receptor α (ER). Over expression of MIER1α dramatically decreases estrogen-stimulated growth of breast carcinoma cells illustrating that MIER1α functions as an ER co-repressor.[91] The SANT domain has been found to interact with Sp1. This interaction causes the removal of Sp1 from the promoter of the target genes, including repressing transcription of MIER1's own promoter which interferes with transcription.[94]

The second mechanism is the recruitment/regulation of chromatin modifying enzymes through its N-terminal region, and conserved domains ELM2 and SANT. MIER1 has been shown to bind the CREB binding protein (CBP) through its N-terminal region, including the histone binding and ELM2 domains. CBP contains a BRD, and is a well-known HAT. However, interaction with MIER1 inhibits its HAT activity.[95] The ELM2 domain of MIER1 has been shown to recruit HDAC1, and repress transcription.[93] In recent studies MIER1 has been shown to increase HDAC activity to H3K27ac, and remove the acetyl group. (Paterno et al., unpublished) MIER1 has also been shown to interact with HMT G9a through its SANT domain, although MIER1's exact function with the chromatin modifier G9a is unknown[96](Paterno et al., unpublished data).

Given the information that MIER1 recruits the above modifying enzymes it is most likely that MIER1 works within a complex of these enzymes to repress transcription from H3K27 target genes. The methylation of H3K27 indicates repressed transcription, so the active transcription modification of acetylation must be removed from the lysine to begin repression. HDAC could be working in complex with MIER1 to remove the acetyl group from K27. Also with MIER1 blocking the acetylase activity of CBP the lysine cannot be re-acetylated. It has been shown that in *Drosophila* CBP specifically acetylates H3K27ac over other HATs.[74] Therefore it would be interesting to study if other HATs in humans would come to acetyl K27 or not. However in order for this process to occur MIER1 must read that there is an acetyl group on K27. As mentioned previously, reader effector motifs are often found in complexes controlling gene repression. Since MIER1 does not contain any recognizable reader domains (such as bromo-, chromo-, PHD or MBT), it likely uses a similar mechanism as other readers domain to recognize the acetyl modification and then recruit and bind the other modifiers CBP, HDAC, and G9a.

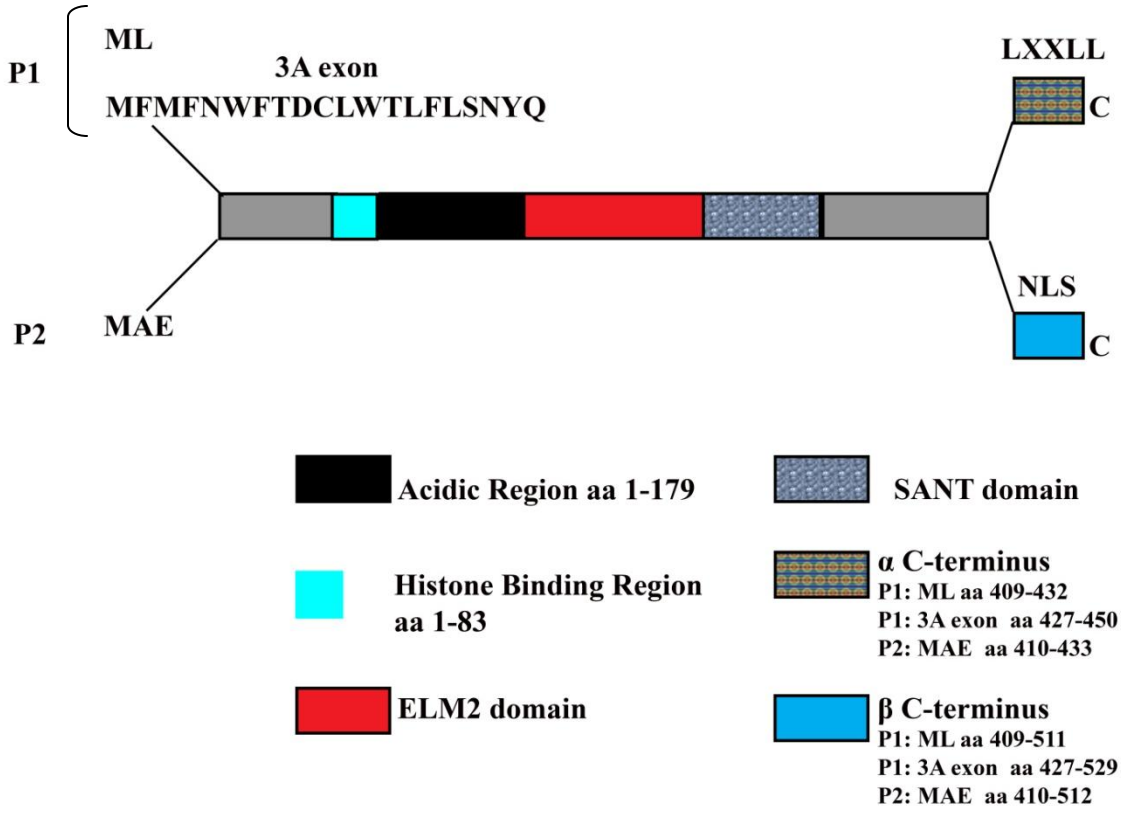

Figure 1.4: Domain Structure of MIER1

P1 and P2 give rise to different MIER1 transcripts through different promoter use. Transcription from the P1 promoter produces either a start site of ML- or MFMFNWFTDCLWTLFLSNYQ- , the sequence encoded by exon 3A. Transcription from the P2 promoter produces an isoform with a start site that begins with MAE-. Through different use of the facultative intron 16 at the C-terminus creates two isoform MIER1α and MIER1β. MIER1β results from the inclusion of the facultative intron. The β isoform contains a nuclear localization signal (NLS). The MIER1α isoform is produced when the intron is not included and results in a smaller protein. The α isoform contains a classic LXXLL motif that is used for interaction with nuclear receptors. All of the MIER1 isoforms share the same central domains including the acidic domain, histone interaction domain, SANT domain, and the ELM2 domain, as depicted above.

1.3 Purpose of this Study and Objectives

MIER1 has been found to interact, and recruit chromatin modifiers, and mostly likely work within a complex to repress transcription. In recent data from a histone peptide array and peptide pull down arrays, MIER1 has been shown to only interact when H3K27 is acetylated from among many different single or combinations of modifications (acetylated, methylation, phosphorylation, citrullination, and unmodified) on the four core histones H2A, H2B, H3 and H4. MIER1 is known to interact with HDAC, and in recent evidence, MIER1 has been shown to increase HDAC activity towards the H3K27ac. It then removes the acetyl modification from lysine 27, and this represses transcription. (Paterno et al., unpublished data) Recent data has demonstrated through the use of a deletion construct of MIER1, that the binding site of H3K27ac is in the first 83 amino acids in MIER1 (Paterno et al., unpublished data). **Given the above information, the purpose of this study was to determine the crucial binding sequence in MIER1 of H3K27ac. We hypothesized that the binding site is in a conserved sequence that has been maintained through evolution, and is essential for the interaction of MIER1 on acetylated lysine 27 to repress transcription.**

Objective 1: Identification of Null-Binding H3K27ac full Length Mutant MIER1.
Three conserved hydrophobic aromatic amino acids were mutated to all alanines. A Not1 restriction site was also introduced at the mutation site, for easy verification of the mutation. The mutant DNA was then purified. The plasmid DNA was sent away for sequencing. Full length mutant MIER1 would later be tested to see if it would still bind to acetylated H3K27 or not. This would determine if the chosen amino acids were the binding location in MIER1 for H3K27ac.

Objective 2: Create a series of deletion constructs from amino acids 1-83 of MIER1.

Creation of several different deletion constructs from the first 83 amino acids in MIER1 were completed to be tested if they would still bind to acetylated H3K27 and to narrow down the area of amino acids required for binding.

Objective 3: Determination of a possible binding site in MIER1 for H3K27ac.

A number of peptide pull down assays were performed with all of the full length mutant MIER1 proteins, deletion construct proteins, and wild type MIER1 to determine the binding ability of the MIER1 proteins to acetylated H3K27 and possible binding site.

Chapter 2- Materials and Methods

2. Materials and Methods

2.1 Plasmids and Constructs

A. GST-*hMIER1* β

*hMIER1*β – (accession number AY124191.1) GST-*himer1 β* fusion constructs were produced by Zhihu Ding (Terry Fox Lab, Division BioMedical Science, Faculty of Medicine, Memorial) B, PCR amplified products were cloned into TOPO PCR2.1 vector and digested with *EcoR1*, then inserted into the *EcoR1* sites in the pGEX 4T.1 (GE Health Care). This plasmid contains a GST-tag for isolating GST-fusion protein in experiments.

B. GST-*hMIER1*β Deletion Constructs

Deletion constructs of MIER1β (amino acids 1-35, 1-50 and 1-63) were PCR amplified (primers found in Table 2.4) from pcDNA3-*MIER1β,* cloned into TOPO PCR2.1 vector then digested with *EcoR1* and ligated into pGEX 4T.1 at the *EcoR1* site. Deletion construct amino acids 1-179 was PCR amplified using primers her34k and mi-her1 T100 (listed in Table 2.4 pg 44) from CS3+MT-*MIER1β*, cloned into pCR3.1, then digested with *EcoR1* and ligated into pGEX 4T.1 at the *EcoR1* site by Tina Blackmore (Terry Fox Lab, Division BioMedical Science, Faculty of Medicine, Memorial). This plasmid contains a GST-tag for isolating GST-fusion protein in experiments.

C. pcDNA3-*hMIER1*β

This construct was used as a template to amplify MIER1 sequence by PCR for cloning deletion constructs. MIER1β was cloned into this vector (Invitrogen) at *EcoR1* site of pcDNA3 by Yuan Lew (Terry Fox Lab, Division BioMedical Science, Faculty of Medicine, Memorial).

2.2 Site-Directed Mutagenesis

2.2.1 Thermal Cycling for Producing Full Length MIER1 Mutants

Site Directed Mutagenesis was performed using mutagenesis primers (found in Table 2.1) this process resulted in mutating three amino acids to three alanine amino acids, and creating a Not1 restriction site.(Figure 3.2 depicting amino acid mutation pg. 56) Primers for mutations were designed using the Agilent Technologies QuikChange Primer Design https://www.genomics.agilent.com/primerDesignProgram.jsp. Site Directed Mutagenesis was performed using two complimentary oligonucleotides containing the desired mutation, flanked by unmodified nucleotide sequences using thermal cycling, as per the manufacturers protocol (QuikChange ® Lightning Site-Directed Mutagenesis Kit, cat # 210518). Briefly, mutant strand synthesis was created by mixing in a, thin walled PCR tube : 2.5 µl of 10× reaction buffer, 50 ng (10 ng/µl) of *mier1*β plasmid (wild type), 125 ng (100ng/µl) of oligonucleotide forward primer, 125 ng (100ng/µl) of oligonucleotide reverse primer (specific primer sequences found in Table 2.1), 1 µl of dNTP mix, 1 µl of QuikSolution reagent and dH2O to a final volume of 50 µl. Lastly, 1µl of QuikChange® Lightning Enzyme was added. Thermal cycling conditions were set using the parameters from Stratagene for the QuikChange Lightning Site-Directed Mutagenesis Kit, found in Table 2.2. After thermal cycling, the parental DNA was digested using 1 µl of the *Dpn1* enzyme with an incubation of 20 minutes at 37 ºC. *Dpn1* treated DNA was transformed in XL Blue Cells as follows: 50 µl of cells were aliquoted in to 1.7 ml eppendorf tube, and 3 µl of *Dpn1* treated DNA was mixed gently with a pipette tip and left to incubate for 30 minutes on ice. Cells were then heat shocked for 30 seconds at 42 ºC and placed on ice for two minutes. 250 µl of pre-warmed Luria-Bertani

broth (LB) [10 g tryptone (Fisher, cat # BP1421-500), 5 g yeast (Fisher, cat # BP1422-500), 10 g NaCl (Fisher, cat # S640-10), 1 L dH2O, autoclaved] was aseptically added to each transformation and left to shake at 250-300 revolutions per minute (rpm) at 37 ºC for one hour. 100 µl was spread on pre-warmed agar plates containing (50 µg/ml) ampicillin (Amp; Sigma, cat # A-9518) and left inverted overnight at 37 ºC to grow. The next day, one colony was re-streaked on a new LB Amp plate to confirm mutant plasmid transformation, and left to grown inverted overnight at 37 ºC. One colony was picked the next day then the DNA was purified, as follows: single colonies were grown overnight with 5 mls of LB and 5 µl of Amp (50µl/mg), and left to shake overnight at 37 ºC, 250-300 rpm. The following morning, cells were pelleted at 1232 xg for 15 to 20 minutes. The pellet was re-suspended, and plasmid DNA was collected as per manufacturer's protocol (Promega Wizard Plus SV Mini Preps DNA Purification System, cat # A1460). Briefly, after the cells had been pelleted, the supernatant was drained off and the pellet re-suspended in 250 µl of re-suspension solution. To break open the cells; 250 µl of a lysis buffer was used the tube was inverted to mix, and then it was left to incubate at room temperature until the solution was clear. Furthermore 10 µl of an Alkaline Protease solution was added to inactivate endonucleases, and other proteins released during the lysis. The protease was added to the tube, inverted to mix the sample and left at room temperature for five minutes. After incubation, 350 µl of a neutralization solution was added, and the tube was inverted to mix and then centrifuged at 14,000 xg for 10 minutes. The cleared lysate was decanted to the spin columns, and centrifuged at 14,000xg for one minute, and flow through discarded. The column was washed 2x with a wash solution first with750 µl and then 250 µl. The spin columns were centrifuging at 14,000 xg for one-two minutes and flow through discarded. 50 µl of nuclease free water was added to the spin

column and allowed to sit for a few minutes. Then the plasmid DNA was eluted by centrifuging at 14,000 xg for one minute. The eluted DNA was stored at -20°C. The concentration (ng/µl) of the eluted DNA was determined by Nanodrop Spectrophotometer 2000. The machine was blanked with 2µl of nuclease free water and then 2 µl of plasmid DNA was measured.

2.2.2 Restriction Digest to Verify Mutation Site

Purified DNA of full length MIER1 mutants were treated with Not1 restriction enzyme (15 U/µl; Invitrogen, cat # 15441-017) as a preliminary check to determine if the desired mutation was created. 5µl DNA (500 ng) was digested with 2-3x excess restriction enzyme, 4 µl 10x REact® 3 buffer (Invitrogen, cat# 16303-018) and dH_2O up to 40 µl. Samples were incubated in a 37 ^oC water bath for one hour. 5 µl from the digested DNA was mixed with 2 µl 6x loading buffer (Ficoll®-400 in dH_20, Sigma cat #F4375) and 5 µl of dH_20, then loaded to a 1% agarose gel. The samples were run around 100V for 45 minutes to confirm appropriate size of digestion. Agarose gels were prepared by melting UltraPure™ agarose (Invitrogen, cat # 16500-500) in 1X TBE Buffer [Tris (Fisher, cat # BP152-1)/Borate (Fisher, cat # BP168-500) /EDTA (Fisher, BP120-500)]. SYBR® Safe DNA gel stain(1 in 10,000 dilution) (Invitrogen, cat # S33102) was added to slightly cooled melted agarose in 1X TBE before liquid was poured into gel casting, and left to solidify at room temperature for about 30 minutes. Electrophoresis was carried out in 1X TBE.

Table 2.1: PCR primer pairs used for preparing MIER1 Full Length Mutants

Full Length Mutant	Forward Primer	Reverse Primer
YGY	5'GGC GAC ATG CCA ATT CAT GAA CTT CTC AGC CTT GCG GCC GCT GGT AGT ACT GTT CGA CTA CCT GAA GAA G 3'	5' CTT CTT CAG GTA GTC GAA CAG TAC TAC CAG CGG CCG CAA GGC TGA GAA GTT CAT GAA TTG GCA TGT CGC C 3'
FDP	5' GGT TCA GCA ACA TCA GAT GAC CAT GAA GCG GCC GCA TCA GCT GAC ATG CTG GTT CAT GAT T 3'	5'AAT CAT GAA CCA GCA TGT CAG CTG ATG CGG CCG CTT CAT GGT CAT CTG ATG TTG CTG AAC C 3'
MPI	5' GAA ATA GAA GAT CTT GCA AGG GAA GGC GAC GCG GCC GCT CAT GAA CTT CTC AGC CTT TAT GGT TAT GG 3'	5'CCA TAA CCA TAA AGG CTG AGA AGT TCA TGA GCG GCC GCG TCG CCT TCC CTT GCA AGA TCT TCT ATT TC 3'
TLE	5' GCT GAC ATG CTG GTT CAT GAT TTT GAT GAT GAA CGA GCG GCC GCA GAG GAA GAA ATG ATG GAA GGA GAA 3'	5' TTC TCC TTC CAT CAT TTC TTC CTC TGC GGC CGC TCG TTC ATC ATC AAA ATC ATG AAC CAG CAT GTC AGC 3'

Table 2.2: Thermal Cycling Conditions for Full Length MIER1 Mutant Synthesis

Segment	Cycles	Temperature	Time
1	1	95 °C	2 minutes
2	18	95 °C	20 seconds
		60 °C	10 seconds
		68 °C	30 seconds/kb of plasmid length =30 sec*FL β 1500 bp+ 4900 bp pGEX=195 secs=3.20 mins
3	1	68 °C	5 minutes

2.3 Cloning of Deletion Constructs

2.3.1 Amplification of Deletion Constructs

Deletion constructs were amplified from MIER1β (wild type) pcDNA3 using Platinum® *Taq* HiFi DNA Polymerase High Fidelity (5 U/µl; Invitrogen, cat # 11304-011). PCR was performed in thin walled PCR tubes in an Eppendorf Mastercycler Gradient Thermocycler (specific conditions found in Table 2.3 pg 43). Each reaction contained: 10 µl of pcDNA3 (10 ng), 5 µl 10X High Fidelity PCR buffer (Invitrogen, P/N 52045), 2 µl 50mM $MgSO_4$ (Invitrogen P/N 52044), 1 µl 10mM dNTPs(2.5 mM of each dNTP), 100ng/µl her47 (forward primer), 100 ng/µl reverse primer (specific sequence found in Table 2.4, found on pg 44) , 0.2 µl Platinum® *Taq* DNA High Fidelity(Invitrogen, cat# 11304-011), and dH_2O up to 50 µl.

After amplification samples were put on ice, additional 3' A-overhangs were added to the reactions by mixing with the pipette tip 1 unit of *Taq* polymerase (5U/µl; Invitrogen, cat # 18038-042). The tube was then incubated at 72 ºC for 10 minutes without cycling. The tube was then immediately put on ice and used in the TOPO® cloning reaction.

2.3.2 TOPO Cloning

The amplified adenylated PCR product was cloned into pCR™ 2.1- TOPO ® vector as per manufacturers protocol (TOPO ®TA® Cloning Kit Invitrogen, cat # 45-0641). Briefly, 4 µl of PCR product was mixed with, 1µl of salt solution (1.2 M NaCl and 0.06 M $MgCl_2$), and 1 µl of TOPO® vector. It was allowed to incubate at room temperature for five minutes. 2 µl of TOPO® cloned reaction was transformed into 50 µl XL Blue cells, as in section 2.2.1. Colony PCR was performed as in section 2.3.3, on multiple colonies to check for inserts.

Positive colonies that contained the appropriate insert were purified by bacterial mini-prep, as follows: single colonies were grown overnight with five mls of LB and 5 µl of Amp (50µg/ml), and left to shake overnight 37 ºC at 250-300 rpm. The following morning cells were pelleted at 1232 xg for 15 to 20 minutes. The pellets were re-suspended and plasmid DNA collected as per manufacturers protocol (Promega Wizard Plus SV Mini Preps DNA Purification System, cat # A1460) as described in section 2.2.1. The concentration (ng/µl) of the eluted DNA was determined by Nanodrop Spectrophotometer 2000. The machine was blanked with 2 µl of nuclease free water and then 2 µl of plasmid DNA was measured. The TOPO® cloned deletion constructs were sent for sequencing to TCAG DNA Sequencing Facility (Toronto, Canada) to confirm proper sequence orientation, as described in section 2.4.5

2.3.3 Colony PCR

Following bacterial transformation as in section 2.2.1, individual colonies were picked up with a sterile disposable pipette tip and re-suspended in a master mix. PCR was performed in thin walled PCR tubes in an Eppendorf Mastercycler Gradient Thermocycler conditions are found in Table 2.5. Each reaction contained 1 µl of template DNA (re-suspended colony as the template DNA) with master mix made containing: 5 µl 10X PCR buffer (Invitrogen, P/N Y02028) , 1.5 µl 50mM $MgCl_2$ (Invitrogen P/N Y02016), 4 µl 10mM dNTPs(2.5mM of each dNTP), 2 µl of 100µg/ml $pCRS_P$/2.1 and pCRT/2.1 respectively (specific primer sequences for different vectors found in Table 2.6 pg 46), 0.2µl *Taq* DNA polymerase, and dH_2O up to 50 µl. The pipette tip from each colony after re-suspension in the master mix was then

scratched on a LB Amp plate (50 µg/ml) and left to grow overnight at 37ºC. The colonies were checked again in the same manner to confirm the insert or ligation reaction. Positive colonies were purified as in section 2.2.1. PCR products were visualized by agarose gel electrophoresis, by mixing 10 µl of the PCR products with 2 µl of 6x DNA loading dye, loaded and run on a 1.2% agarose gel for 25 minutes around 115V, to confirm appropriate size of insert. Agarose gels were made by melting UltraPure™ agarose with 1X TBE Buffer (Tris /Borate/EDTA). SYBR® Safe DNA gel stain(1 in 10,000 dilution) was added to slightly cooled melted agarose, in 1X TBE before liquid was poured into gel casting, and left to solidify at room temperature for about 30 minutes. Electrophoresis was carried out in 1X TBE.

2.4 Sub-cloning

2.4.1 Restriction Digest of Deletion Construct Inserts and Vector

The purified TOPO® cloned deletion construct inserts vector (pGEX 4T.1; from section 2.3.2) were both digested with EcoR1 (15U/µl) (Invitrogen, cat # 152-02013). The vector and inserts 5µl (500ng), respectively were digested with 2-3x excess appropriate restriction enzyme, 4µl 10X buffer H (Invitrogen, cat # A1501A), and mixed with dH_2O up to 40µl. Samples were incubated in a 37 ^oC water bath for 1 hour. 5 µl of the digested DNA was mixed with 2 µl of 6x loading buffer and 5 µl of dH_20, then loaded to a 1.2% agarose gel to confirm digestion of inserts and vector had occurred. Agarose gels were made by melting UltraPure™ agarose in 1X TBE Buffer (Tris/Borate/EDTA). SYBR® Safe DNA gel stain (1

in 10,000 dilution) was added to slightly cooled melted agarose in 1X TBE before liquid was poured into gel casting, and left to solidify at room temperature for about 30 minutes. Electrophoresis was carried out in 1XTBE.

2.4.2 Alkaline Phosphatase Treatment of DNA

The digested vector DNA was treated with Calf Intestinal Alkaline Phosphastase (CIAP; 20 U/µl; Invitrogen, cat # 18009-019) to remove the 5'-phosphate groups, in order prevent re-circularization and re-ligation of the vector. Up to 10 pmol of digested DNA was used with 0.01U/µl CIAP/pmol of ends in CIAP buffer. Samples were incubated in a 37 ^{o}C water bath for 30 minutes. Another aliquot of the same amount of CIAP used before was added, and left to incubate at 37 ^{o}C for another 30 minutes. CIAP treated DNA was then mixed with 6x DNA loading dye and loaded to a 1.2% agarose gel for gel purification. Agarose gels were made by melting UltraPure™ agarose in 1X TBE Buffer (Tris/Borate/EDTA). SYBR® Safe DNA gel stain (1 in 10000 dilution) was added to slightly cooled melted agarose in 1X TBE, before liquid was poured into gel casting, and left to solidify at room temperature for about 30 minutes. Electrophoresis was carried out in 1XTBE.

2.4.3 Agarose Gel Purification

Agarose gels were visualized with UV light, and the appropriate sized DNA fragment was cut from the gel using a disposable razor blade. The gel slice was then placed in a pre-weighted 1.7 microcentrifuge tube. Purification of DNA from gel pieces was performed as per manufactures protocol using the Qiagen QIAquick Gel Extraction Kit, using all the kits

reagents (Qiagen, cat # 28704). Briefly, 3 volumes of Buffer QG to 1 volume of gel were added to the gel slice and left to incubate at 50 °C for 10 minutes. After the gel slice was dissolved completely, the color of the mixture was checked for the same yellow color of Buffer QG. If the mixture was not the appropriate yellow color 10 µl of 3M sodium acetate was added to readjust the pH to pH≤7.5. 1 gel volume of isopropanol (Sigma, cat # I9516) was added to the sample, mixed and then transferred to a QIAquick spin column, in a provided 2 ml collection tube. The spin column was centrifuge at 10,000xg for one minute and flow-through was discarded. Next, 500 µl of Buffer QG was added to the QIAquick column, then centrifuge for one minute at 10,000 xg. To wash the column, 750 µl of Buffer PE was added to QIAquick column, and centrifuge for one minute at 10,000xg, the flow-through was discarded and the QIAquick column was centrifuged again for one minute at 10,000 xg. The QIAquick column was put into a clean 1.7 ml microcentrifuge tube and DNA was eluted with 30 µl of dH_2O. The dH_2O was allowed to sit on the column for around one minute, and then centrifuged for one minute at 10,000xg. From the extracted DNA 2 µl was mixed with 10 µl of 6 x loading buffer and run on a 1.2 % agarose gel to confirm the extracted DNA was the right size. An agarose gel was made by melting UltraPure™ agarose with 1X TBE Buffer. SYBR® Safe DNA gel stain (1 in 10,000 dilution) was added to slightly cooled melted agarose, in 1X TBE before the liquid was poured into the gel casting, and left to solidify at room temperature for about 30 minutes. Electrophoresis was carried out in 1X TBE with a voltage around 120V for 20 minutes.

2.4.4 Ligation

A 20 µl ligation reaction was prepared using 1:3 molar ratio of vector: insert gel purified samples from section 2.4.3. The concentrations needed for the inserts were calculated using the following formula basing the vector concentration on the band intensity similarity to the 1000 bp marker in the 1 kb plus DNA ladder.

$(ng\ vector\ \times kb\ of\ insert \div kb\ of\ vector) \times$ insert to insert ratio 3/ 1

(roughly 11.5 ng of construct aa 1-63, 9.1 ng of construct aa 1-50, 6.4 ng of construct aa 1-35, 3.5 ng of construct aa 1-19, and 100 ng of vector) mixed with 5x ligase reaction buffer (Invitrogen, cat P/N Y90001), 1µl of T4 DNA ligase (5U/µl) (Invitrogen, cat # 15224-041), 0.5 µl of 10 mM dATP (Promega, E601D) and dH_2O to 20 µl. The reaction was gently mixed with a pipette tip, and left to incubate at room temperature for two hours. 2 µl from each ligation reaction was transformed as in section 2.2.1 in 50 µl of XL Blue cells. Colonies were PCR checked (as in section 2.3.3) for ligated insert in the pGEX 4T.1 vector with appropriate vector primer (found in Table 2.6, pg 46). One colony was re-streaked on a new LB Amp plate to confirm ligated insert and left to grown inverted overnight at 37 ºC. Positive colonies that contained the appropriate ligated insert were purified by bacterial mini-prep, as follows: single colonies were grown overnight with 5 mls of LB, 5µl of Amp (50µg/ml), and left to shake overnight 37 ºC at 250-300 rpm. The following morning cells were pelleted at 1232 xg rpm for 15 to 20 minutes. The pellets were re-suspended and plasmid DNA collected as per manufacturers protocol (Promega Wizard Plus SV Mini Preps DNA Purification System, cat # A1460) as described in section 2.2.1. The concentration (ng/µl) of the eluted DNA was determined by Nanodrop Spectrophotometer 2000. The machine was blanked with 2 µl of

nuclease free water and then 2 µl of plasmid DNA was measured. The deletion constructs were sent for sequencing to confirm proper sequence orientation, as described in section 2.4.5.

2.4.5 Sequencing Analysis

The full length mutants and deletion constructs correct sequences were confirmed by the TCAG DNA Sequencing Facility (Toronto, Canada). The samples were sent to be processed, in the required concentrations by the facility of 200-300 ng in 7.0 µl of purified template DNA. Also included was 50 ng in 0.7 µl of primers for each forward and reverse sequencing reaction, in two separate tubes for a total volume of 7.7 µl in each tube.

Table 2.3: PCR Thermal cycling conditions for creating MIER1 Deletion Constructs

Cycles	Temperature	Time
1	94 °C	2 minutes
30	94 °C	30 seconds
	55 °C	30 seconds
	68 °C	30 seconds
1	68 °C	5 minutes
1	30 °C	1 second

Table 2.4: PCR primer pairs used for preparing MIER1 Deletion Constructs

Deletion Construct	Forward Primer *	Reverse Primer**
aa 1-179	5’ CAC CAT GGC GGA GCC ATC TGT TGA ATC 3’	5’ GCC CTA CCA GTC TTC TGA TGG AAT 3’
aa 1-63	5’ CAC CAT GGC GAC ATC TGT TGA ATC 3’	5’ CG GAA TTC CTA GTC GCC TTC CCT TGC AAG 3’
aa 1-50	5’ CAC CAT GGC GAC ATC TGT TGA ATC 3’	5’ CG GAA TTC CTA GTT TGT TTC TCC TTC CAT CAT 3’
aa 1-35	5’ CAC CAT GGC GAC ATC TGT TGA ATC 3’	5’ CG GAA TTC CTA ATC AAA ATC ATG AAC CAG 3’
aa 1-19	5’ CAC CAT GGC GAC ATC TGT TGA ATC 3’	5’ CG GAA TTC CTA ATC GTC TGA TGT TGC TGA 3’

Primers for deletions constructs were checked with IDT Oliogo Analyzer for the correct Tm, GC% content (around 40%), hairpin structures and self- dimer.
*(her 47 for all except 1-179) (her 34K for 1-179)
**(mi-her1 T-100 for 1-179)

Table 2.5: PCR Thermal cycling conditions for checking Deletion Constructs inserts

Cycles	Temperature	Time
1	94 °C	5 minutes
30	60 °C	1 minute
	72 °C	1 minute
	94 °C	1minute
1	60 °C	1 minute
1	72 °C	7 minutes
1	30 °C	1 second

Table 2.6: PCR primer pairs used for checking for inserts

Primer	Forward Primer	Reverse Primer
pCRSp/2.1	5’ GCT CAC TCA TTA GGC ACC C 3’	
pCRT/2.1		5’ GGC GAT TAA GTT GGG TAA CG 3’
pGEX 1.1	5’ GGG CTG GCA AGC CAC GTT TGG T 3’	
pGEX 1.2		5’ CCG GGA GCT GCA TGT GTC AGA GG 3’

2.5 Peptide Pull Down Assay with GST-fusion Proteins

2.5.1 GST-fusion protein production

GST fusion plasmids were transformed into competent Rosetta Cells™ (Millipore, cat #70953-4). First 100 µl of competent cells were aliquoited to 1.7 ml eppendorf tube, and 10 ng of purified DNA was mixed gently with a pipette tip, and left to incubate for 30 minutes on ice. Cells were then heat shocked for 45 seconds at 42 ºC, and placed on ice for two minutes. 500 µl of pre-warmed LB was aseptically added to each transformation, and left to shake at 250-300 rpm at 37 ºC for one hour. 100 µl was spread on pre-warmed agar plates containing (50 µg/ml) ampicillin, and plates were left inverted overnight at 37º. The next day, one colony was re-streaked on a new LB Amp plate to confirm mutant plasmid transformation, and left to grown inverted overnight at 37ºC. The following day, single colonies were grown overnight with 5 mls of LB plus ampicillin (50µg/ml). The morning of the third day 1 ml was added to 250ml of LB and Amp (50 µg/ml), and left shaking at 37 ºC to grow to an OD_{600} of 0.80-1.0. To check optical density, 1 ml of LB medium was added to a cuvette, and used as a blank to calibrate the machine 1 ml of inoculated culture was added to another cuvette and the measurement taken. OD was read on a spectrophotometer at a wavelength of 595. Once the culture reached the desired OD it was then induced with 25 µl of 1M IPTG (isopropyl- β-D-thiogalactopyranoside; Fisher cat #BP1755-1), and left shaking overnight at room temperature. In the morning cells were pelleted at 3000 rpm for 10 minutes at 4 ºC in 250 ml Nalgene polypropylene bottles. The liquid was drained off and the pellet frozen until it was solid. The pellet was thawed at room temperature, and re- suspended in 5 mls of ice-cold 1X

PBS (phosphate buffered saline; Sigma, cat # P3813). The re-suspended pellet was transferred to a new sterile 50 ml falcon tube and 50 µl of 0.2 M PMSF (phenylmethylsulfonyl fluoride; Sigma, cat # P7626) was added. The cells were then lysed by sonication (70% amplitude, 2 x 15 second pulses with 30 second break in-between). Next, 500 µl of 10% NP-40 (Nonidet P40'; Sigma, cat # 74385) was added, and samples were left on ice for 20 minutes. The tubes were then centrifuged at 10,000 x g for 15 minutes. The soluble faction was removed, and stored in 1 ml amounts in labelled 1.7 ml eppendorf tubes at -70 ºC. SDS-PAGE (sodium dodecyl sulfate – polyacrylamide gel electrophoresis) was used to determine the concentration and purity of GST-Fusion proteins.

2.5.2 Isolation and detection of GST-fused proteins

Glutathione Hi-Cap Matrix beads (Qiagen, cat# 30900) were aliquoited into 1.7 ml eppendorf tubes and washed by adding 1 ml of GST-Pull Down Buffer [20mM Tris HCL pH7.5, 150mM NaCl, 0.1% NP40 in PBS, 10% Glycerol (Fisher, cat #56-81-5; added right before use), dH_20], inverting until the beads were suspended in the buffer, then allowing rotation for one minute at room temperature. The beads were centrifuged at 5000 x g for 30 seconds and the supernatant was aspirated off by vacuum. This was repeated two more times. Then the beads were re-suspended in a 50% slurry with GST- Pull Down Buffer. 50 µl of the 50% slurry of Glutathione Hi-Cap Matrix beads were aliquoted into new 1.7 ml tubes. The beads were washed twice with 1 ml of 1 x PBS in the same manner as with GST Pull-Down Buffer. After washing, varying amounts of soluble GST-fusion protein (25 µl, 50 µl, 100µl and 200 µl) and enough GST Pull Down Buffer to equal 1 ml + 5 µl of 0.2 M PMSF was

added to the sample and rotated at 4 ^{o}C for 1 hour. After incubation, the beads were washed with 1 ml 2x with GST Pull Down Buffer with 2% NP-40 solution added to wash away unbound proteins. A final wash with 1 ml of 150mM NaCl was done using the same method as above. To the washed beads, 30 µl of 2 x SSB (Sodium dodecyl sulfate [SDS] Sample Buffer (Sigma, cat # L5750) + Bromophenol blue dye [BDH Chemicals, cat # 44305]) was added and the samples were vortexed lightly then boiled for four minutes to denature the proteins. The samples were allowed to cool on ice for three minutes, before lightly vortexing, then centrifugation at 5000 x g for one minute to bring the beads to the bottom of the tube. The supernatants were then run on a 7% [for full length MIER1(wild type) and mutant full length MIER1] or 10% SDS-PAGE gel (deletion constructs) and stained with Coomassie Blue Stain (Biorad, cat # 161-0400) for one hour shaking at room temperature. The gel was then de-stained overnight in 1 L of destain (740ml dH2O, 200 ml methanol [ACP, cat# M-3640], and 60 ml glacial acetic acid [ACP, cat # A-0302]) and dried under vacuum for one hour.

2.5.3 Peptide Pull Down Assay

Glutathione Hi-Cap Matrix beads, 100 µl per sample were aliquoted in a 1.7 ml eppendorf tube and washed by adding 1 ml of GST-Pull Down Buffer [20 mM Tris HCL pH7.5, 150mM NaCl, 0.1 % NP40 in PBS, 10 % Glycerol (added right before use), 1 mM DTT [DL-Dithiothreitol (added right before use; Sigma, cat # D5545-1)], dH_2O], inverting until the beads were suspended in the buffer, and then allowing them to be gently vortexed. The beads were then centrifuged at 5000 xg for 30 seconds. This process was repeated two more times. Then the beads were re-suspended to a 50% slurry in GST- Pull Down Buffer and for each

pull down reaction, 100 µl of 50 % slurry was used. Each sample reaction contained a total volume of 500 µl in a 1.7 ml eppendorf tube: the GST protein (pre-determined amount), GST-Pull Down Buffer and 2.5 µl of 0.2 M PMSF. To make up for the different volumes of proteins being used, any of the protein samples with less than the highest volume amount of protein being used per experiment had a 1 % NP-40 solution added so that the total volume equalled 500 µl. The tubes were then rotated for one hour at 4ºC. After an hour the tubes were centrifuged at 5000 xg for 30 seconds. The supernatant was aspirated off and the beads were washed 4x in the same manner as above. The beads were then washed with 1 ml of 150 mM NaCl as per GST-Pull Down Buffer. Then, 100 µl of Elution Buffer [10mM Glutathione (Sigma, cat # G4251), 100 mM Tris HCL pH8] was added to the beads, and allowed to rotate at half rotation for 20 minutes. After 20 minutes the supernatant was removed and stored; 100 µl more of Elution Buffer was added and allowed to rotate at half rotation for 20 minutes more, then added later. Eluted protein was then run over buffer exchanged de-salting columns (Zeba® Spin Desalting Columns 0.5 mls cat # 89883), as per manufacturer's protocol using 300 µl of interaction buffer [GST PD Buffer, 1mg/mL BSA cat # A7888 Sigma, 1mM PMSF]. In a 1.7 ml eppendorf tube 150µl of Interaction Buffer, 40 µl of buffer exchanged protein and 10µl of a 1/50 diluted, 1 mg/ml stock of Histone peptide (AnaSpec Histone Peptides, GK Biotin) H3 (21-44; cat# 64440-25)-unmodified [only used for full length mutant TLE] and H3K27ac (21-43; cat # 64637-025) was mixed and left to incubate overnight at 4 ºC on half rotation. An input sample was also prepared by mixing 2 µl of buffer exchanged protein, and 28 µl of 1xSSB + Bromophenol blue dye and stored at -20ºC. In the morning, 50 µl of a 50 % NeutraAvidin ® Agarose Resin (Thermo Scientific cat # 29200) beads were prepared in the same manner as Glutathione Hi-Cap Matrix beads, and added to each overnight reaction plus

200 µl of interaction buffer. The tubes were allowed to rotate for one hour at 4ºC. After 1 hour, tubes were centrifuged at 5000 xg for 30 seconds, and the supernatant was aspirated off. The beads were then washed 4x with GST-Pull Down Buffer and then once with 150 mM NaCl as before. The beads were re-suspended in 30 µl of 1.5 X SSB+ Bromophenol blue dye, boiled for four minutes, allowed to cool on ice, vortexed lightly and then centrifuged for 30 sec at 5000 x g. The supernatant loaded on a 7 % or 10 % SDS-polyacrylamide gel, depending on the size of the protein being tested. The gels were transferred to a PVDF membrane with the Biorad Trans-Blot® Turbo™ Transfer System at a voltage of 25 V, a constant current of 2.5 Amps, and 12 minutes for full length MIER1 mutants and 25 V, constant current of 2.5 Amps, and 10 minutes for deletion constructs. The membranes were blocked in SuperBlock® T20 (TBS) Blocking Buffer (ThermoScientific, cat #37356). Western Blot analysis was performed with an Anti- GST (0.25 ug/ml, mouse monoclonal (GenScript, cat# A00865) antibody in SuperBlock® for one hour, then washed 5 times in one hour, shaking at room temperature in 1 X TBS-T [20 mM Tris pH 7.6, 150 g NaCl, Tween 20 (Sigma, cat # P2287), dH_2O]. Secondary Antibody 1:5000 dilution of sheep-anti mouse- horseradish peroxidase (HRP; GE Health Care UK Limited cat# NA931V) in SuperBlock® was added, and was left to shake at room temperature for one hour. The membrane was washed 5 times in 1hour again by shaking in 1 X TBST at room temperature and analyzed with chemiluminescent detection using ECL Clarity Western Blotting system and Hyperfilm HCL (BioRad, cat # 170-5061).

Chapter 3: Results

3.1 Creating Full Length MIER1 Mutants

3.1.1 Mutating Conserved Amino Acids within the first 83 Amino Acids of MIER1

In previous research pull down assays were performed in order to determine which region in MIER1 bound H3K27ac. The amino acids 1-83 were determined to be the smallest region capable of binding to H3K27ac. This was found using different size constructs of the MIER1 protein, until binding was still observed to the H3K27ac peptide. (Paterno et al., unpublished data)

The purpose of this project was to more precisely define which of the first 83 amino acids were required to specifically bind H3K27ac. It has been demonstrated that the binding site of proteins to acetylated lysines, uses a bromodomain (BRD) with highly conserved hydrophobic and aromatic amino acids within a hydrophobic pocket. One study found specific tyrosines and a valine were the essential amino acids, in the hydrophobic pocket formation required for binding to the acetylated lysine.[56] These aromatic side chains form a hydrophobic core, and are involved in the stability, and formation of the secondary, and tertiary structure along with the function of the protein.[52,56,108,109] However, MIER1 does not contain a BRD but might still use the same method of utilizing a hydrophobic pocket formed by aromatic side chains to bind H3K27ac.

The way in which reader domains bind to hydrophobic amino acids in acetylated histones was used to find a novel binding domain for acetylated lysines in the MIER1 protein.

Knowing that the sequence of amino acids 1-83 contained the binding site this area was examined for a possible hydrophobic pocket used for binding. Since there is no available 3-D model of MIER1 it was not possible to determine if there is a hydrophobic pocket in the tertiary structure of the folded protein. Since the amino acids used in BRDs to bind acetylated lysine residues are conserved amino acids, it was assumed that the amino acids H3K27ac is binding of MIER1 are likely conserved through different species. It was also assumed that the amino acids used in binding are likely to be hydrophobic as well. Also given that other binding sites have been identified within the conserved domains of MIER1 by mutating conserved amino acids, we used this knowledge in conjunction to determine the binding site.[93]

An alignment was used with the first 83 amino acids of MIER1 from fifteen different species to identify conserved amino acids in, human, rabbit, horse, elephant, dog, cat, whale, walrus, chicken, *Xenopus*, mouse, zebrafish, daphnia, *Drosophila*, and *C.elegans.* Figure 3.1 showed which amino acids in the sequence were conserved between the many species. The alignment figure most importantly compared which hydrophobic, and aromatic amino acids were conserved between the different species.

The first set of experiments used site-directed mutagenesis to study the binding site since other researchers have used this method to mutate conserved amino acids to elicit a change in function of the protein, or determine a binding location.[56,74] Sites were chosen to be mutated from the alignment in Figure 3.1, and these were the amino acids: [^{73}YGY^{75}], [^{64}MPI^{66}], [^{22}FDP^{24}]. The first full length mutant produced was the YGY to AAA. Alanines are most commonly used for site directed mutagenesis, because they are small and aliphatic in nature which causes the biggest change in a binding conformation. Glutamine has also been used as a substitution in site directed mutagenesis.[56,74,110] A Not1 restriction site was introduced at

location of the mutagenesis in the wild type DNA sequence. Each mutation consisted of three amino acids, which meant the sequence was long enough to change the wild type DNA sequence into the Not1site: GCGGCCGC. The wild type codons were being mutated to all, alanines which can be expressed as: GCT, GCC, GCA or GCG. Therefore, wild type codons were arranged to solely alanines, and in the process created the sequence to also be a Not1 restriction site. The introduction of the restriction site where the mutation was located was used to initially verify the mutagenesis, and it can also distinguish the mutated MIER1β protein from the wild type MIER1β protein. If ever in the event there was ever mix of the two a restriction digest could be performed, and the mutated full length MIER1β should show two bands from the digest. One from the cut at the mutation site, and the other at the multiple cloning site in the pGEX 4T.1 vector. The wild type MIER1β should show only one band from a cut in the multiple cloning site of the pGEX 4T.1 vector. A Not1 restriction site was introduced at the mutation site in all full length MIER1 mutants. This can be seen in Figure 3.2 which is a schematic of mutation change, and introduced restriction site. Three other full length MIER1β mutants were created based on the same criteria as above where the hydrophobic amino acids, were mutated with aromatic side chains in order to identify the H3K27ac binding site. The two other full length mutations were MPI-AAA, at amino acids 64-66, and FDP-AAA at amino acids 22-24. All of the full length mutants were sent away for sequencing to confirm mutation, as described in section 2.4.5.

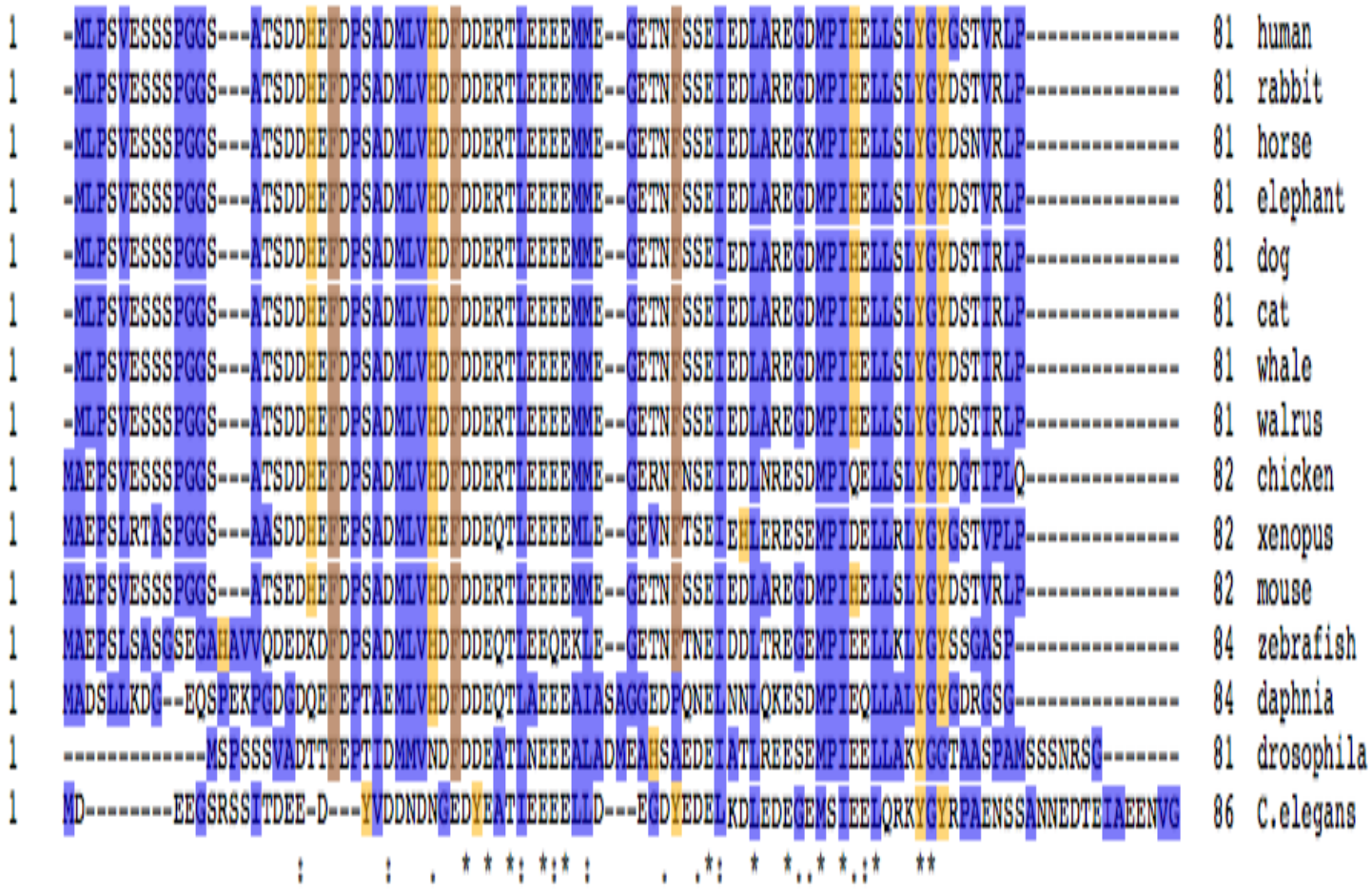

Figure 3.1: Conserved 1-83 amino acid Sequence of MIER1 in Various Species

Around the first 83 amino acids of MIER1 aligned in 15 different species to determine which amino acids were logical candidates for mutate, and locate the binding site of H3K27ac. The highlighted blue amino acids are hydrophobic, and the highlighted brown amino acids are aromatic, or have aromatic side chains. The amino acids are in sequence starting at amino acid one as seen on the left hand side of the diagram in each species, and numbered progressively to around 83 amino acids (ranging from 81-86) as seen beside the species name on the right hand side of the diagram. At the bottom of the diagram are colons indicating conservation between groups of strongly similar properties (scoring > 0.5 in the Gonnet PAM 250 matrix). A period indicates groups of weakly similar properties (scoring =< 0.5 in the Gonnet PAM 250 matrix). An asterisk indicates a single conserved residue between the species.

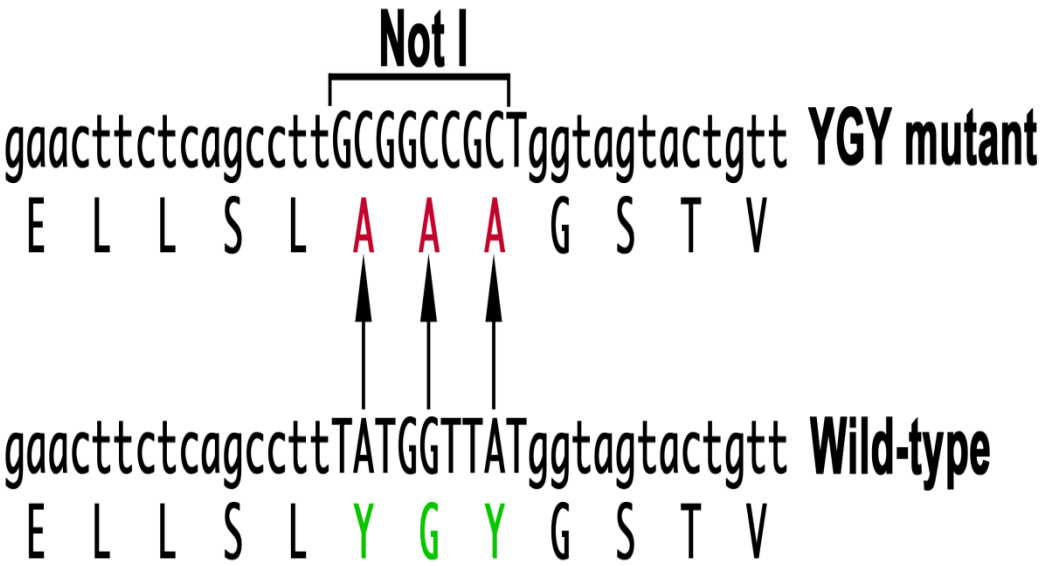

Figure 3.2: MIER1 sequence showing mutation and restriction site

This diagram shows the wild type MIER1β amino acid sequence with ^{73}Y, ^{74}G, ^{75}Y highlighted in green on the bottom. Above the mutated amino acid sequence of all alanines is highlighted in red. The nucleotide sequence of the wild type and mutant sequences are in capital letters, and in the mutant sequence a Not1 restriction site can be seen.

3.1.2 Expression of Full Length MIER1 Mutants

Restriction digests were performed with the full length MIER1β mutants, and wild type MIER1β plasmid DNA sequences to quickly determine if the mutagenesis created the mutations. If the mutagenesis was successful, and the wild type sequence was altered to the sequence for Not1, then the digest would show two bands. The first band would be the Not1 restriction site in the pGEX 4T.1 vector, and the second band would be the size distance from the first restriction site to the second Not1 site in MIER1β at the mutation site. There is a Not1 restriction site in the GST-vector near the multiple cloning sites of wild type MIER1β. This was used as a positive control, and shows up as a linearized plasmid band equal in size to the GST-vector plus MIER1β (approximately 6550bp). The digests were run on an agarose gel, and visualized under UV light to determine if there was an extra band in the full length mutant DNA from the restriction enzyme.

The restriction digest showed that there was one band in the wild type MIER1β lane around, 6550bp which is the expected size of the GST pGEX 4T.1 vector (approximately 5000bp), and MIER1β (approximately 1550 bp) combined. In the full length mutant lane there, are two bands of the appropriate size fragment of the vector, and the mutation site in MIER1β demonstrating the desired mutation was created. The first band in the full length mutant lane is from the restriction cut at Not1 site in the multiple cloning site of the GST-vector resulting in, a linearized GST-vector, which is around 5000 bp. The second band is from the Not1 restriction cut at the mutated site around 1550 bp. The T38A, L39A, E40A mutant versus wild type digest, can be seen in Figure 3.3. Restriction digests of the plasmid DNA from all full length mutants showed an extra band from the introduced mutation. All

full length mutants were sent away for DNA sequencing to confirm that the wild type sequence was mutated sequence, as described in Section 2.4.5.

The GST-fusions proteins (wild type MIER1β, full length MIER1β mutants and GST-vector proteins) were prepared, purified and run on a SDS-PAGE gel with varying amounts of protein. This was followed by staining with Coomassie Blue, as in sections 2.5.1 and 2.5.2 at different time points whenever the protein was needed for an experiment. However, since all of the GST-fusion proteins were not produced at the same time they therefore were not accurately measured for concentration, until compared on the same SDS-PAGE gel. By running all of the proteins together allowed us to determine what amount for the GST-fusion proteins was needed to use for a consistent 1 µg of protein per sample. To determine how much protein was needed in a 1 µg sample for each of the mutant, and wild type full length MIER1β proteins for use in future experiments a known volume with 1 µg of just GST protein was run alongside the full length wild type, and mutant MIER1β proteins for comparison.

Once the determined amounts for each full length MIER1β 1 mutant, wild type MIER1β, and GST proteins were decided a sample of each was then run on the same gel for comparison, and can be seen in Figure 3.4.

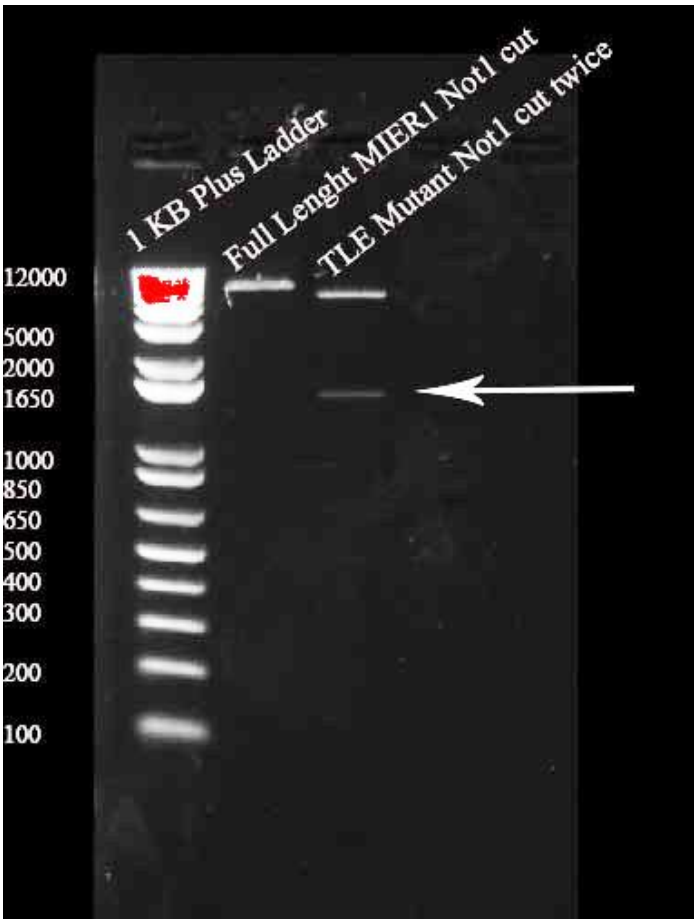

Figure 3.3: Restriction Digest of TLE Mutant

Restriction Digests were performed as described in the Material and Methods. Wild type MIER1 and full length mutant MIER1 DNA was incubated with a Not1 restriction enzyme in a 37 °C water bath for one hour, then run on 1% agarose gel to separate DNA by appropriate size fragments and visualized under UV light. A single band was expected in the full length wild type MIER1 lane representing linearized DNA from Not1 restriction site cut in the GST-vector of about 6550bp. This band can be seen in the second lane. In the third lane which contains full length MIER1 mutant a band can be seen with a fragment cut, for the Not1 site in the GST vector around the expected 5000bp. The expected second cut fragment is around 1550 bp from the introduced Not1 restriction site and can be seen in the third lane with an arrow indicating the second fragment cut.

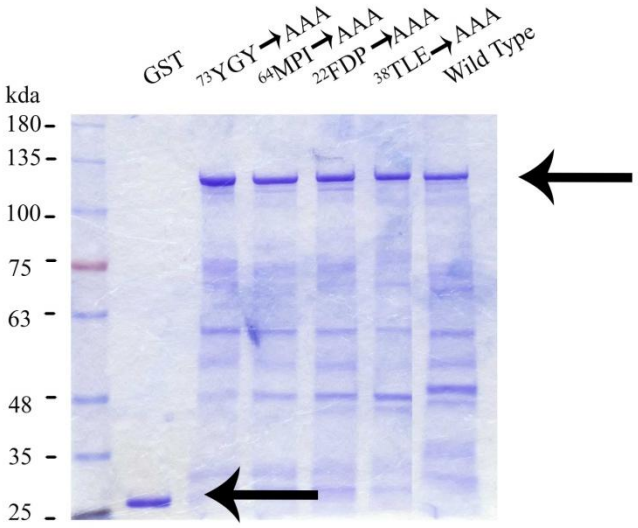

Figure 3.4: GST-Fusion Protein of Wild Type, Full Length Mutant MIER1β and GST

GST fusion proteins were prepared as described in Material and Methods. Around 1µg of full length MIER1β GST-fusion proteins were purified through binding to Glutathione Hi-Cap Matrix beads (Qiagen, cat# 30900) and then eluted and run on a SDS-PAGE gel. The gel was stained with Coomassie Blue and dried down. The top arrow indicates the full length mutants and wild type MIER1β GST-fusion proteins running around 125 kDa. The lane-labeled GST does not contain a full length GST-fusion MIER1β protein, but the GST vector protein alone. The bottom arrow indicates the GST vector protein running around 26 kDa.

3.2 Full Length MIER1β Mutants YGY, MPI and FDP and Wild Type MIER1 Binding to H3K27ac

To determine whether the full length MIER1β mutants would bind H3K27ac, peptide pull down assays were completed using purified full length MIER1β mutant, and wild type GST-fusion proteins (as described in section 3.1.2). Peptide pull down assay have been described as a simple, and unbiased assay that allows for the identification of histone effector proteins, and various biotinylated histone peptides modified at particular residues.[111] Utilizing this assay for our experiments allowed the results to be determined if there was a difference in binding between the full length wild type MIER1β protein, and the full length mutant MIER1β protein in an unbiased manner.

As a positive control for these assays wild type MIER1β protein was used since it is known to bind H3K27ac. (Paterno et al., unpublished) The mutant and wild type proteins were independently incubated with, a biotin tagged H3K27ac peptide. Then the protein-peptide complex was incubated with neutravidin conjugated beads eluted, and run on a SDS-PAGE gel. The gels were then transferred to a PVDF membrane, and followed by Western Blot analysis with an Anti-GST antibody. 1μg for each GST-fusion protein was used in this assay, as shown on the Coomassie Blue stained SDS-page gel in Figure 3.4.

The initial peptide pull-down tested the full length MIER1β mutant Y73A, G74A, Y75A, and whether it could bind H3K27ac compared to the wild type protein. The assays showed the full length mutant protein to have the same binding capability as wild type MIER1β. Therefore the chosen mutated amino acids were not the critical amino acids needed for binding (Figure 3.5A&3.5B). The peptide pull down assay in Figure 3.5A&3.5B with the full

length Y73A, G74A, Y75A mutant, and wild type MIER1β was performed by Corinne Mercer (Terry Fox Lab, Division BioMedical Science, Faculty of Medicine, Memorial).

Next, another set of peptide pull down assays were completed to test the binding ability for the full length mutants M64A, P65A, I66A and F22A, D23A, P24A. The alanine substitution for these conserved amino acids ^{64}MPI^{66} and ^{22}FDP^{24} also demonstrated the same binding capability as the wild type protein, and therefore they did not contain critical amino acids needed for binding despite being highly conserved against species (Figure 3.5C&3.5D).

These data show the three full length MIER1β mutants (YGY, MPI, and FDP) which have the same binding ability as the wild type MIER1β (Figure 3.5). Therefore, the critical amino acids needed for binding H3K27ac. were not found in these amino acid sequences. The peptide pull down assays were repeated three times in total for each full length mutant to, ensure reproducible results.

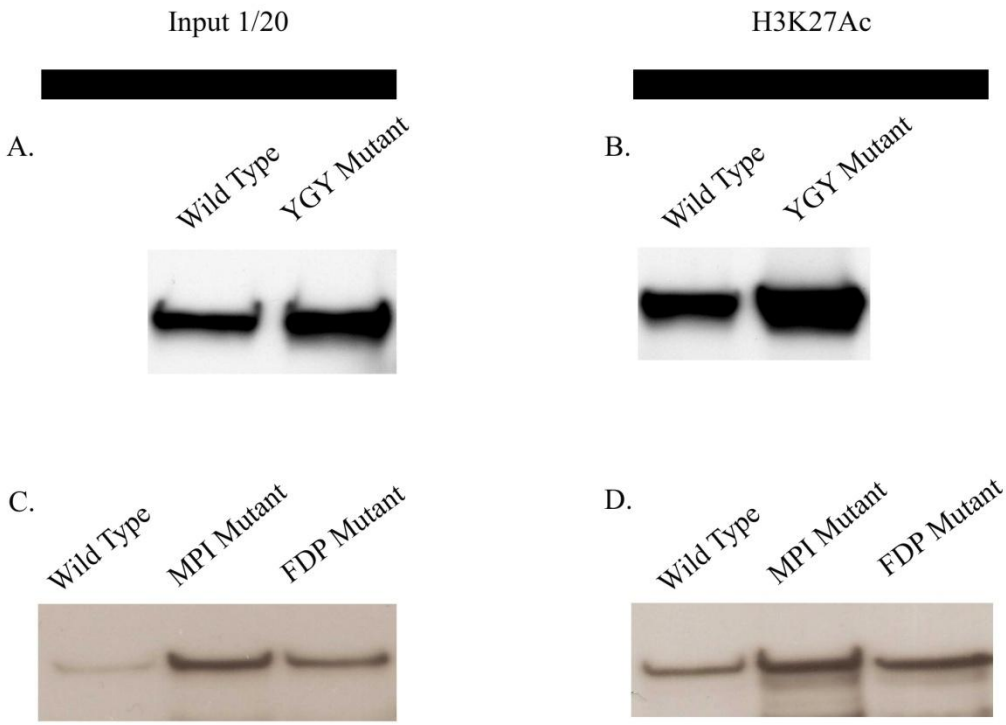

Figure 3.5: Full Length MIER1β Mutants YGY, MPI and FDP do not contain critical amino acids for binding

Eluted GST fusion proteins were purified, bound to the biotinylated peptide and eluted according to the protocol in Material and Methods. Eluted purified GST-fusion proteins were incubated separately overnight with a biotinylated H3K27ac peptide. Then the protein-peptide complex was incubated with avidin conjugated beads. The protein-peptide bound complex was eluted by boiling with SSB loaded, and run on a SDS-PAGE gel. It was then transferred to a PVDF membrane and a Western Blot analysis was conducted with an anti-GST antibody to detect if any GST-tagged proteins had been pulled down. Panel A depicts the ($1/20^{th}$) input of the wild type and full length mutant YGY used to show the relative amount of eluted purified protein used, compared to the amount in the protein-peptide interaction as seen in panel B. Panel C depicts the ($1/20^{th}$) input of the wild type and full length mutants MPI and FDP used to show the relative amount of eluted purified protein compared to the amount used in the protein-peptide interaction as seen in panel D.

3.3 Creating MIER1 Deletion Constructs amino acids 1-63, 1-50 and 1-35

After site-directed mutagenesis in three different locations of conserved residues within MIER1β did not find the amino acids needed for binding, another approach was used to find the region that binds H3K27ac. Instead of looking at the alignment of the first 83 amino acids (Figure 3.1), and selecting conserved amino acids that were hydrophobic with aromatic side chains, four smaller deletion constructs were made from amino acids 1-83. This allowed us to try and narrow down, the area of known binding from 83 amino acids to a smaller area. For each construct approximately every 15 amino acids from the C-terminal end were deleted in succession. Deleting 15 amino acids from the C-terminal end to progressively make smaller construct, allowed me to examine smaller areas at time to determine if binding is occurring. If binding is occurring in a larger construct but not the progressively smaller construct then it would be likely that the binding sequence is located in between the two. Care was taken not to delete part of a conserved sequence of aromatic, and hydrophobic amino acids. For example, deletion construct aa 1-63 was not created as aa 1-65 (which would have been roughly 15 amino acids from amino acid 83) since amino acid 65 is in the middle of three conserved hydrophobic amino acids (amino acid sequence 64M,65P, 66I).

The creation of four deletion constructs was attempted, starting with removing approximately 15 amino acids to create amino acids constructs 1-63, the 1-50, 1-35 and 1-19. We could not recover deletion construct 1-19 into the GST-vector, and so only constructs aa1-63, 1-50 and 1-35 were prepared into GST-fusion proteins. After they were PCR amplified (primers found in Table 2.4) from pcDNA3-*MIER1β*, the deletions were cloned into TOPO PCR2.1 vector. The deletions were digested with *EcoR1* and ligated into a GST tagged pGEX 4T.1 vector at the *EcoR1* site.

The GST-fusions proteins (deletion construct and GST-vector proteins) were prepared, purified and run on a SDS-PAGE gel with varying amounts of protein as in sections 2.5.1, and 2.5.2. Once each amount was determined to be around 1 µg, the chosen volume amount of GST-fusion protein for each deletion construct, and GST-vector alone was loaded onto a SDS-PAGE gel stained with Coomassie Blue. A diagram of the deletion constructs and GST-vector used can be seen in Figure 3.6A, and the Coomassie stained gel can be seen in Figure 3.6B.

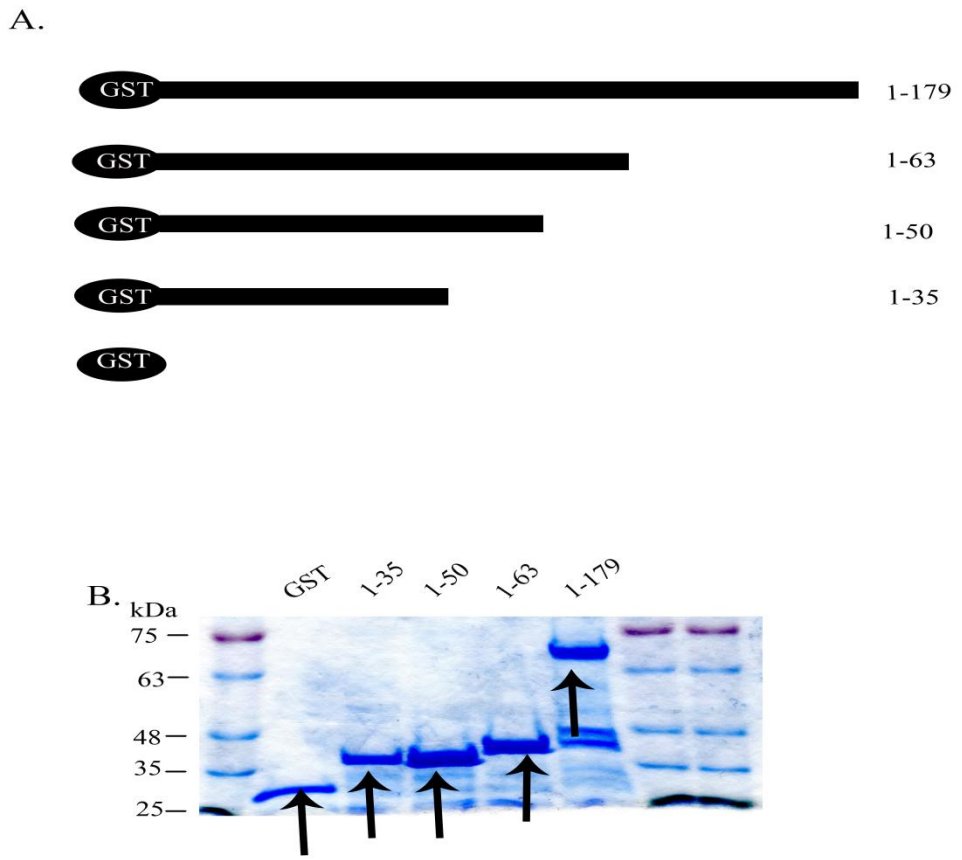

Figure 3.6: Deletion Constructs of MIER1

Panel A. This diagram represents GST deletion constructs. The GST tag is depicted on each deletion construct and then the GST tagged vector by itself. The numbers of amino acids in each respective construct are indicated on the right of each construct. Panel B. GST fusion proteins were prepared as described in Material and Methods. Around 1µg of deletion construct GST-fusion proteins were purified by binding to Glutathione Hi-Cap Matrix beads (Qiagen, cat# 30900), then eluted from the beads and run on a SDS-PAGE gel. The gel was then stained with Coomassie Blue and dried down. The lane labeled GST does not contain a full length GST-fusion deletion construct protein but the GST vector protein alone. Arrows indicate the band of interest for each of the deletion construct and GST-vector proteins.

3.4 Deletion Constructs 1-63 and 1-50 Bind H3K27ac but not Deletion Construct 1-35

Previous studies conducted by other individuals in the lab indicated that the first 83 amino acids of MIER1 bind H3K27ac. Peptide pull-down assays were performed with deletion constructs to further narrow down the area where the binding site is located in MIER1β. Purified deletion constructs and GST-vector proteins were incubated, with a biotin-tagged H3K27ac peptide. The protein-peptide complexes were pulled down with avidin conjugated beads, eluted and run onto a SDS-PAGE gel. To analyze the pull-down samples the gels were transferred to a PVDF membrane, and subject to Western Blot analysis with an Anti-GST antibody. Another deletion construct with amino acids 1-179 was used for a positive control in place of the construct 1-83 in each experiment, since it includes the known binding area to H3K27ac. As a negative control GST-vector protein was also incubated with a biotin tagged H3K27ac peptide since, GST protein is known to not bind H3K27ac or unmodified H3 (Paterno et al., unpublished). The relative amounts of GST-fusion proteins were used in this assay and are shown on the Coomassie Blue stained SDS-PAGE gel in Figure 3.6. These results were repeated twice more.

The peptide pull-down assay showed that the deletion amino acids constructs 1-63 and 1-50 of MIER1 bound to H3K27ac. The deletion construct amino acids 1-35, however, did not bind, as seen in Figure 3.7. Since the construct 1-63, and 1-50 bound this suggests that amino acids 51-63 are not critically required for binding since the construct 1-50 bound to the peptide. This suggests binding requires some or all of the amino acids 1-50. The exact area needed for binding cannot be confirmed from these experiments since the deletions were only from the C-terminus, and deletion constructs that start at the N-terminus, and delete from the other end were not created. To confirm if the 35-50 amino acids is critical sequence required

for binding, or if there is an amino acid located in 1-35 sequence that is also required for binding along with the amino acids 35-50, deletion constructs starting from the N-terminal end, similar to removing approximately 15 amino acids from the C-terminal of the first 83 amino acids: 16-83, 30-83, 45-83 and 60-83 amino acid constructs should be made. Then test the constructs with peptide pull down assays to determine which sequence between the C-terminal, and N-terminal deletions is necessary for binding the H3K27ac peptide.

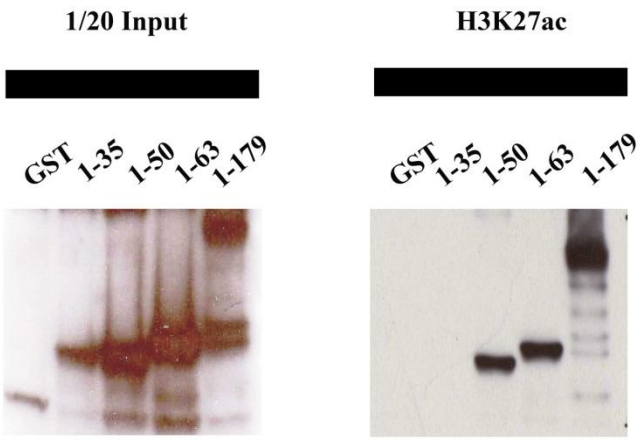

Figure 3.7: Deletion Constructs Amino Acids 1-63 and 1-50 Bind to H3K27ac but not 1-35

Eluted deletion construct GST fusion proteins were purified, incubated with biotinylated peptide and eluted according to the protocol in Material and Methods. Eluted purified GST-fusion proteins were incubated separately overnight with a biotinylated H3K27ac peptide. Later they were incubated with neutravidin conjugated beads, and any protein-peptide bound complex was eluted by boiling with SSB, loaded, and run on a SDS-PAGE gel. They were then transferred to a PVDF membrane and a Western Blot analysis was conducted with an Anti-GST antibody to detect if any GST-tagged proteins had been pulled down. The ($1/20^{th}$) input lane shows the relative amount of eluted purified protein compared to the amount used in the protein-peptide interaction. The GST vector lane acts as a control, since it has been shown to not bind with the peptide.

3.5 Full Length MIER1 Mutant TLE has Reduced Binding Compared to Wild Type MIER1 to H3K27ac

The results from the last section, testing binding of the amino acid deletion constructs 1-63, 1-50 and 1-35, revealed that the binding requires some or all of the amino acids 35-50. The next approach was to utilize the direct and fast method of site-directed mutagenesis, with three amino acids in MIER1β for H3K27ac. The amino acid sequence ^{38}TLE^{40} was chosen to mutate to all alanine amino acids by consulting the alignment diagram in Figure 3.1 because it was a conserved sequence between species, and the ^{39}L is a hydrophobic acid which might be important in binding.

The same approach for creating all the full length mutants in T38A, L39A, E40A was implemented as that used in sections 3.1.1 and 3.1.2. A restriction digest of the full length mutant T38A, L39A, E40A versus the wild type, shows an extra band in the mutant plasmid DNA for the Not1 site in the mutated DNA, along with the band from the cut at the Not1 site in the GST-vector (Figure 3.3). The relative amount of GST-fusion protein concentration for the full length mutant T38A, L39A, E40A compared to the other full length mutants [Y73A, G74A, and Y75A], [M64, P65A, I66A] and [F22A, D23A, P24A], along with the wild type, and the GST protein by itself can be seen in Figure 3.4.

In order to determine if the full length mutant TLE contained the critical amino acids in MIER1β for binding of H3K27ac, peptide pull assays were performed. As a positive control for this experiment, wild type MIER1β. A negative control used for this experiment was the GST-vector. Purified full length mutant ^{38}TLE^{40}, wild type, and GST-vector proteins were incubated with a biotin-tagged H3K27ac peptide. The protein-peptide complex was then incubated with neutravidin conjugated beads, eluted and run on a SDS-PAGE gel. To analyze

the pull downs, the gel was transferred to a PVDF membrane, followed by a Western Blot analysis with an Anti-GST antibody. To verify binding capacity the TLE full length mutant, wild type MIER1β, and GST-vector proteins were also incubated with a biotin tagged H3 (21-44) peptide. The protein-peptide complexes were incubated with neutravidin conjugated beads, eluted, run on the same SDS-page gel as the acetylated peptide bound protein complexes, and subjected to the same Western Blot conditions of the other samples in each experiment. These were chosen because the wild type MIER1β protein is known to bind H3K27ac but does not bind to the unmodified H3 peptide. The GST protein was also known not to bind H3K27ac or unmodified H3 (Paterno et al. unpublished). Each pull down was repeated three times with the same controls each time.

As evident in Figure 3.8, the peptide pull down assays illustrated that the full length mutant ^{38}TLE^{40} had a reduced binding capability when compared to the binding of the wild type protein to H3K27ac. This information indicates that the amino acids T38A, L39A, E40A, are important in the binding of H3K27ac, but, the T38A, L39A, E40A amino acid sequence is not the entire binding sequence.

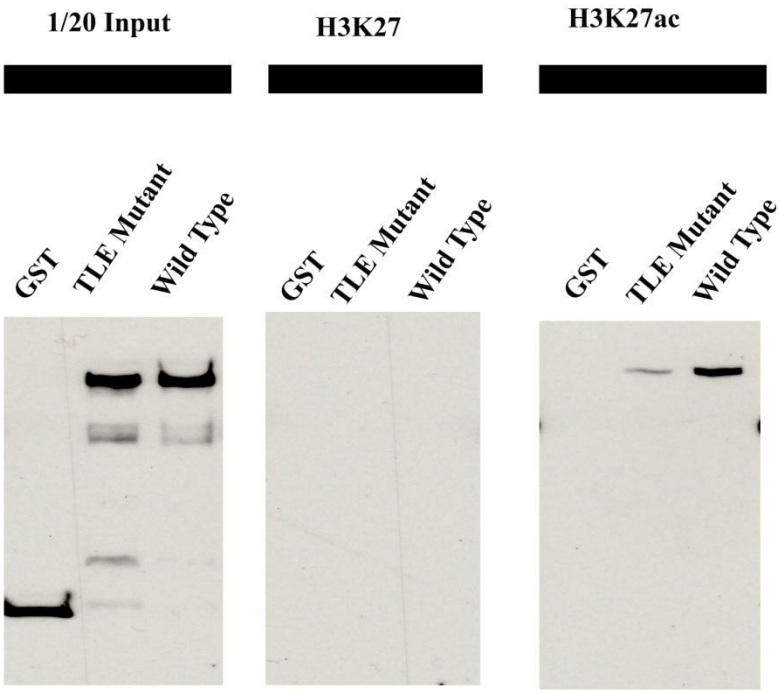

Figure 3.8: Full Length MIER1β Mutant TLE Contains Critical Amino Acids for Binding H3K27ac

Eluted GST fusion proteins were purified, bound to the biotinylated peptide and eluted according to the protocol in Material and Methods. Eluted purified GST-fusion proteins were separately incubated overnight with a biotinylated H3K27ac peptide and H3 (21-44) peptide. Then the protein-peptide complex was incubated with neutravidin conjugated beads, and any protein-peptide bound complex was eluted by boiling with SSB, loaded and run on a SDS-PAGE gel. It was then transferred to a PVDF membrane and a Western Blot analysis was conducted with an anti-GST antibody to detect if any GST-tagged proteins had been pulled down. Panel A depicts the ($1/20^{th}$) input of the wild type. Full length mutant TLE showed the relative amount of eluted purified protein, used compared to the amount in the protein-peptide interaction as seen in panel C. Panel B depicts no binding of the GST-fusion proteins to the unmodified peptide H3 (21-44). Panel C depicts the reduced binding capability of the full length TLE mutant compared to the wild type, and no binding of the GST-vector to the acetylated peptide.

Chapter 4- Discussion

4.1 General Discussion

Lysine 27 on histone 3 (H3K27) plays a crucial role in the differentiation of ES cells. Depending on whether the residue is acetylated or methylated it allows for active or inactive gene transcription.[74] In a recent study it was discovered that MIER1 recognizes, and binds to the acetylated residue of lysine 27 on histone 3 (Paterno et al., unpublished data). Additional research has found that H3K27ac binds to MIER1 in the first 83 amino acids of the protein (Paterno et al., unpublished data). The purpose of the present study was to determine the critical binding sequence in MIER1 for H3K27ac. It is important to find and characterize the site at which H3K27ac, is binding to MIER1 since both of these modifications exert important functions in the cell.

Given the known role of MIER1 as a chromatin modifier, and its ability to bind an acetylated lysine, it can be assumed that the binding site in MIER1 for H3K27ac, would function in the same manner as other chromatin modifying proteins that bind acetylated lysines. The purpose of this study was to determine the essential amino acids, in MIER1 that bind H3K27ac.

The initial full length mutant $^{73\text{-}75}$YGY (Y73A, G74A Y75A) was chosen because it had a strong sequence conservation of hydrophobic, and aromatic amino acids that have been shown previously to be directly involved in binding to a substrate.[56] Importantly, in other studies of proteins binding acetylated histones, these conserved hydrophobic and aromatic amino acids were shown to be essential to the structure, and thus the function, of the protein.[56] Perhaps,

one might assume that the mutation of these specific amino acids should change the relation of the secondary and tertiary structures formed by the protein binding to the acetylated lysine. However, this full length mutant had the same binding ability to H3K27ac as the full length wild type protein, as shown in peptide pull-down assays (Figure 3.5). The other two full length mutants studied MPI (M64A, P65A, I66A), and FDP (F22A, D23A, P24A) had the same binding to H3K27ac as the full length wild type MIER1 GST-fusion protein. This demonstrates that these conserved hydrophobic (P,I, and F), and aromatic (Y,F) amino acids do not play a role in the binding of H3K27ac to MIER1, nor the potential folded tertiary structure that the MIER1 protein forms to H3K27ac.

The next series of experiments used deletion constructs of the first 83 amino acids in MIER1 to examine a smaller region, which could contain the binding sequence to H3K27ac. Peptide pull-down assays with the three deletion constructs aa1-63, 1-50 and 1-35 were completed. These studies revealed that amino acids from 1 to 50 were sufficient for the interaction with the H3K27ac peptide. However, the deletion construct aa 1-35 did not bind to the H3K27ac peptide (Figure 3.7).

Approximately the same amount of protein for each of the deletion constructs (about 1µg) was used in the peptide pull down assays as shown in Figure 3.6. However the equimolar amounts of the protein were not taken into account. The differences in the molecular weights of each of the fusion proteins, and the amount of protein, should have been adjusted according to the largest molecular weight fusion construct (which would have been aa 1-179) so the same number of moles was used in each sample. In future experiments I would digest all the deletion construct protein bands from an SDS-page gel then extract the peptides and run mass spectrometry on each sample to determine what amount was actually in each sample.

The results from the last set of peptide pull-downs between the full length TLE mutant and the H3K27ac peptide revealed the full length mutant^{38}TLE^{40} (T38A, L39A, and E40A) reduced the binding to H3K27ac compared to the full length wild type MIER1β GST-fusion protein (Figure 3.8).

MIER1 recognizes and binds to the acetylated modification on H3K27 (Paterno et al, unpublished data). Thus we believe binding of H3K27ac to MIER1 is involved in its ability to act as a transcriptional repressor. Initially, the hypothesis was that the binding site in MIER1 was a conserved sequence of hydrophobic and aromatic amino acids. This study demonstrated that the conserved amino acids in the ^{38}TLE^{40} sequence is involved in the binding, but are not the only amino acids responsible for binding.

There are a couple of ways H3K27ac could be binding to MIER1. One way maybe similar to that of the BRD protein which uses hydrophobic, and aromatic amino acids to form a binding pocket. However, in the literature it has been shown there are other possible amino acids, and amino acid sequences sufficient for binding acetylated lysines. An arginine, asparagine, or acidic stretches of amino acids, including those in the SANT domain, could be involved in the docking to an acetylated histone. [54,108,112–115]

I found that while mutating the amino acids sequence ^{38}TLE^{40} in the wild type protein to all alanines, and particularly mutating a conserved hydrophobic amino acid (^{39}L) in the full length mutant MIER1 protein these three amino acids cannot account for total binding. However, the amino acids ^{38}TLE^{40} are essential in contributing to binding. It is likely that the binding site consists of several critical amino acids before the first 50 amino acids. When the wild type MIER1 protein is folded into the proper binding conformation it could possibly use other amino acids, from down or upstream of the ^{38}TLE^{40} sequence.

According to the literature, an arginine[108] or asparagine[54,115] might be required in the binding of proteins to acetylated lysines. An arginine or asparagine could be involved in MIER1 binding to H3K27ac along with other hydrophobic aromatic amino acids, and/or the ^{38}TLE^{40} amino acid sequence. The method used in this study to determine the binding site was modeled after similar experiments with BRDs binding to acetylated substrates. Some studies have found that along with the hydrophobic amino acids, BRDs require an arginine[108] or asparagine[54,115] for completing the folded structure of the protein and the binding pocket. Therefore, it is possible that along with hydrophobic amino acids an arginine or asparagine is required for binding H3K27ac. Interestingly, there is an arginine at ^{37}R, and an asparagine at ^{50}N. Both of these are within the first 50 amino acids in MIER1, which I showed were required for H3K27ac binding in MIER1. Either amino acid might be involved in binding and should be examined.

Another acetylated lysine binding mechanism found in the literature, involves conserved acidic amino acid stretches. A number of studies have shown that other acidic regions, or acidic activation domains, are involved in the binding of the acetylated substrates. These acidic areas additionally have an affinity towards readers, and writers of acetylated lysines in various proteins suggesting a role in transcriptional regulation.[110-113] For example, the tumor suppressor protein p53 that is known to bind acetylated lysines, contains an activation domain with several acidic amino acid residues that interact with different proteins such as Mouse double minute 2 homolog (Mdm2), and the CREB binding protein. [116,117] The ligand binding area in the secondary structure of p53's acidic transactivation domain has conserved hydrophobic amino acids in conserved positions that are critical to function of the protein.[116] Other examples of proteins that use acidic residues in transcriptional activity, include the

Epstein–Barr virus nuclear protein 2 (EBNA2), and herpes simplex virion protein 16 (VP16). They have also been shown to recruit, and bind the writer proteins p300/CBP using their acidic activation domains, and use the writers HAT domain in activating transcription.[112,116] The SANT domain which is often involved in protein-protein interactions, and can bind to histone tails to remodel nucleosomes, is covered in large areas of acidic residues. These acidic patches have been proposed to act as binding locations for histone tails or highly basic proteins.[113] While the SANT domain in MIER1, does not act in the binding of H3K27ac, the possibility that it serves another purpose in binding amino acid residues or other basic proteins is conceivable. (Paterno et al., unpublished)

Given that these examples of different proteins (p53, EBNA2 and VP16) use acidic amino acid stretches to bind acetylated substrates, along with their affinity towards different reader, and writer proteins and MIER1's affinity towards similar readers and writers MIER1 may use one of the acidic stretches of amino acids starting at the ^{40}E to bind H3K27ac. Further characterization of amino acids upstream of 50 must be explored to determine if an arginine, asparagine, or a stretch of acidic amino acids is required in binding H3K27ac.

It is possible that the amino acids required for binding MIER1 to H3K27ac are in multiple locations, up or downstream, of the ^{38}TLE^{40} sequence. In 2006, Li *et al.*,[118] determined that a a PHD finger of the human BPTF (bromodomain and PHD domain transcription factor) is required to bind to H3(1-15)K4me3. This is the largest subunit of the ATP-dependent chromatin-remodeling complex, NURF. They studied the binding site of the H3(1–15)K4me3-peptide which was localized in the PHD domain using nuclear magnetic resonance (NMR). The trimethylated lysine of K4 is positioned within a cage of four conserved

aromatic amino acids at different positions in the sequence of the domain. The amino acid ^{3}Y forms the base, while ^{10}Y, ^{17}Y and ^{32}W, form three walls.[118]

Furthermore, in 2003 Min *et al.*,[119] used X-ray crystallography to generate a 1.4-Å-resolution structure of the chromodomain in the Polycomb complex that was bound to a H3K27me3 peptide, in order to understand how the chromodomain in the protein was recognizing and binding to trimethyl modification. The structure revealed a conserved model of methylated lysines in a crystal lattice, that is made up of conserved amino acids found in the chromodomains of the Polycomb protein family. They looked at the similarities between the *Drosophila* Polycomb protein (dPC) with the H3K27me3 peptide compared to the HP1 protein, and the H3K9me3 peptide. Many features of the interactions between the dPC, and the H3K27me3 peptide were similar between the HP1 and H3K9m3 peptides. Most importantly, they found that K27me3 is bound in a hydrophobic pocket made up of three aromatic residues, ^{26}Y, ^{47}W, and ^{50}W similar to the K9me3 peptide.[119] These findings further suggest that the binding site in MIER1 of H3K27ac might be at multiple highly conserved hydrophobic aromatic amino acids in different locations in the MIER1 amino acid sequence. When folded into the protein structure it creates a binding pocket.

4.2 Implications

The implications of determining the binding site of H3K27ac in MIER1 is significant and perhaps linked to the epigenetic regulation of gene expression. Recent studies have found that loss of the acetylated H3K27 and particularly increased rates of trimethylated H3K27 by EZH2 has led to silencing of tumor suppressor genes and metastasis of different cancers.[120–122] An example of this is with ZEB1 induction, global H3K27ac levels are found to decrease while levels of methylation were found to be increased in ZEB1 genes located at H3K27. ZEB1 is an E-box binding transcription factor which is a major suppressor of epithelial genes in lung cancer. It is activated in cancer cells, and involved in invasion, metastasis, stem-like properties and it is a major factor in the chain of events leading to changes in gene expression during the epithelial to mesenchymal transition (EMT) in non-small cell lung cancer. HDAC inhibitors were found to increase the expression of ZEB1 target genes. Targeting these epigenetic modifications (ac and me3) would be expected to reduce metastasis.[120]

MIER1 binds to H3K27ac (Paterno et al., unpublished), and MIER1 has been shown to recruit HDAC1[93] to enable deacetylation of H3K27 in target genes and which could lead to increased methylation these genes. Thus, MIER1 could possibly be a marker for cancers with observed acetylated lysine loss with tri-methylated lysine gain. This process could be enabled since MIER1 has been shown to block the HAT CBP from acetylating lysines, and recruit HDAC1 to remove the acetyl group from the lysine residue.[93] (Paterno et al., unpublished data). This then creates the opportunity for HMTs to add methyl groups to the lysine, marking it for transcriptional repression. The only HMT to trimethylate H3K27 is the PRC2 complex via its EZH2 component.[78] High levels of EZH2 in adults cells have been shown to lead to

poor prognosis in many cancers including prostate cancer, breast cancer, lymphoma, myeloma, colorectal cancer, endometrial cancer, bladder cancer and melanoma.[12,123] Currently, there are many therapeutic treatments being developed, and used in the clinic, that either target HDACs, or certain HMTs, such as EZH2.

MIER1 has been shown to interact with ERα, and over expression of MIER1 in invasive ductal carcinoma causes restriction of estrogen-stimulated anchorage-independent growth. [91] The ERα receptor, along with similar hormone receptors such as estrogen-related receptor alpha (ERRα), are PTM to control gene regulation. Specific lysine residues in DNA binding domain of ERRα can be acetylated by p300 coactivator associated factor (PCAF), and deacetylated by HDAC8.[124] Given that MIER1 has been demonstrated to recruit, and block the HAT activity of p300/CBP, it is possible that MIER1 can exert the same affinity towards a similar reader protein, PCAF to control acetyl activity.[95] Also given the knowledge that MIER1 interacts with ERα, and other nuclear receptors, it should be investigated if MIER1 associates with ERRα to possibly control breast cancer progression.[91]

In view of the fact that MIER1 is a known chromatin modifier and repressor of transcription[93,95](Paterno et al., unpublished), and since H3K27 PTMs have been found in many aggressive cancers, it would be logical to investigate treatments involving MIER1 bound complexes. This could include MIER1 and HDAC in a complex together. Recently, Millard *et al.,* (2013) studied HDACs in complexes with chromatin modifiers for the purpose developing inhibtors.[31] They found HDACs are often found in repressor complexes, so it is reasonable to develop inhibitors for this complex. In the same paper, the ELM2 and SANT domains are shown to act together by completely wrapping around HDAC, covering the active sites and allowing other domains to recruit nucleosomes to the complex. Often ELM2 and

SANT domains are juxtaposed in co-repressor proteins suggesting a conserved function, and most of these proteins have been shown to recruit HDAC1 and/or HDAC2.[31] MIER1 recruits HDAC1/2 through its ELM2 domain. While there is no known function yet involving both the ELM2 and SANT domain in MIER1, given conservation, a function is likely for both domains. The SANT domain in MIER1 is known to repress transcription of Sp1 at MIER1's promoter and has been shown to interact with the HMT G9a.[93,94] Once again this suggests there is a functional relationship between the two domains, potentially when MIER1 is bound to H3K27ac and must recruit other chromatin modifiers in order to alter the acetylated active state of the chromatin to trimethylate and inactive H3K27 bound genes. Thus, developing inhibitors of HDAC1s ability to form in a complex with MIER1 might be a viable drug treatment. If MIER1 could not bind HDAC then this would lessen the capacity of HDAC to remove the acetyl group. This would then lead to less unmodified H3K27 available. Also there would be less amounts of inactive chromatin that might be targeted by aberrant EZH2 activity, thus making the chromatin available for active transcription.

The role of MIER1 in epigenetic regulation has only been studied with MIER1 GST-fusion proteins and cell lines. Also MIER1 has only recently been found to bind to H3K27ac *in vitro* using protein samples.(Paterno et al., unpublished) Originally however, the function of MIER1 was first investigated in the developing embryo, and was found to be upregulated after the induction of mesoderm differentiation by FGFs during early embryo develop in *Xmier1*, then undetectable in the following developing stages.[87] It should be investigated in embryonic stem cells (ES) if there is a link to genes associated with modified H3K27 that are involved in recruitment of tissue-specific transcription factors, which lead to the establishment of tissue-specific gene expression pattern in embryos. Especially since both PTM acetylation and

methylation have been shown to have important impacts on H3K27 bound genes. Particularly, acetylation of H3K27 has been shown to be associated with the trithorax protein which is responsible for the maintenance of gene expression.[73] Also the trimethylation of lysine 27 is a target for the PRC2 complex at specific areas that are marked for early embryogenesis. Given this information, it is possible that MIER1 exerts a different function as a transcriptional regulator in those genes during development depending on whether they are acetylated or methylated, and should be investigated.

4.3 Future Studies

While much progress has been made in narrowing down the critical binding site in MIER1 for H3K27ac, more work is necessary to determine the exact binding sequence. Once the binding site has been located in MIER1, the potential to uncover the precise function of MIER1 when bound to H3K27ac associated genes can be further characterized. When this site has been extensively studied, the potential for MIER1 as a chromatin modifier may serve as a therapeutic aid for certain cancers. Below are some of the experiments that can be completed in order to further understand the mechanism of MIER1 activity in the cell.

4.3.1 Determine Binding site in MIER1 of H3K27ac

As mentioned in the results, it would be worthwhile to further narrow down the sequence in MIER1 that binds H3K27ac. The optimal way to determine this interaction site would be to create deletion constructs from amino acids 1-83. Using the strategy outlined here, four deletion constructs could be produced that progressively remove 15 amino acids. However, instead of deleting from the C-terminal end, the process could start by deleting from the N-terminal end of MIER1. Deleting about 15 amino acids at a time (amino acids: 16-83, 30-83, 45-83 and 60-83) allows for a broader region to be tested for the binding location of H3K27ac. Then, depending on the results of those experiments compared to the sequence know from this study (aa 36-50) it can be defined to an exact sequence of amino acids. A peptide pull down array could be performed to observe which amino acids in the deletion constructs are interacting with the acetylated H3K27 peptide. Deletion constructs containing amino acids 16-83, 30-83 and 45-83 should be particularly useful.

4.3.2 Determine exact amino acids required for binding of H3K27ac in MIER1

Once the essential region in MIER1 has been determined through the use of deletion constructs, site directed mutagenesis of conserved amino acids in the that region could be the best way to create a full length MIER1 that is not capable of binding H3K27ac. Once the full length MIER1 mutant protein is produced, the mutant should be tested against the wild type full length MIER1 in a peptide pull down assay to confirm that it does not bind the H3K27ac peptide.

4.3.3 Characterize the binding site in MIER1 of H3K27ac

To fully understand what is happening when MIER1 is binding H3K27ac, the MIER1 site needs to be characterized. Once the binding amino acid sequence has been determined in MIER1 for H3K27ac, the structural basis of how the binding is occurring can be studied. Using X-Ray crystallography, the sequence in MIER1 that binds H3K27ac can be visualized, and how the protein folds and where the modification interacts with the protein can be further studied. The purified MIER1 protein deletion construct of amino acids 1-83 (with the GST-tag digested) should be used to crystallize with the H3K27ac (21-43) peptide since it contains the binding site and encompasses the entire sequence known to bind the peptide.

4.3.4 Demonstrate an *in vivo* effect on adult cells with Mutant MIER1

The physiological effect of having a full length mutant MIER1 that cannot bind to acetylated H3K27 in healthy breast epithelial cells could be studied. A method to test this would be to transfect a full length mutant MIER1 plasmid, into the epithelial breast cells. Then the cells should be studied to see if there are any physiological changes. Ultimately we want to determine what effects MIER1 has in regulating H3K27 associated genes, and in turn, cancer progression. So it should also be determined whether there is an increase in growth or any other physiological changes to the normal breast cells that would be indicative of changes in cancerous growth.

4.3.5 Determine any role MIER1 has on cells during development associated to H3K27 regulated genes

MIER1 was first discovered as a fibroblast growth factor activated early response gene. The expression levels of *Xmier1* were up-regulated after the induction of mesoderm differentiation by FGFs.[90] It has been demonstrated in the literature that early developmental cues are epigenetically pre-marked in human ES cells, and indicate the different roles of H3K27me3 and H3K27ac as regulatory elements in the early stages of human embryogenesis.[69] As the GST-protein of MIER1 has only been shown to bind H3K27ac, it would be interesting to determine if it might be linked to genes that are active in ES cells. These would include the trithorax protein (MLL, the human homolog), which is responsible for the maintenance of gene expression of acetylated H3K27 target genes.[73] This is

especially interesting because trithorax associates with CBP, and is required for the maintenance of transcriptionally active states that repress Polycomb silencing. Because MIER1 is known to recruit chromatin modifiers and transcription factors in adult cells[93–96](Paterno et al., unpublished), it should be studied if MIER1 is exerting a vital role development and differentiation of ES cells. A chromatin immunoprecipitation on ChIP (ChIP on chIP) could be performed to determine if in MIER1 there are H3K27 associated genes binding in the promoters, enhancers or repressor regions of assocaited genes. Then, to determine if MIER1 is being recruited to modify tissue-specific transcription factors, a transgenic reporter assay in mice embryos could be performed to determine cell-type and stage-specific expression characteristic of developmental genes.

Bibliography

1. Kornberg RD TJ. Chromatin structure; oligomers of the histones. *Science (80-).* 1974;184(4139):865-868. doi:10.1126/science.184.4139.865 The digital archives of Science, 1880-1996.

2. Arents G, Burlingame RW, Wang BC, Love WE, Moudrianakis EN. The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. *Proc Natl Acad Sci U S A.* 1991;88(22):10148-10152.

3. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature*. 1997;389(6648):251-260.

4. Arents G, Moudrianakis EN. Topography of the histone octamer surface: repeating structural motifs utilized in the docking of nucleosomal DNA. *Proc Natl Acad Sci U S A.* 1993;90(22):10489-10493.

5. Yang X-J, Seto E. Lysine acetylation: codified crosstalk with other posttranslational modifications. *Mol Cell*. 2008;31(4):449-461. doi:10.1016/j.molcel.2008.07.002.

6. Kouzarides T. Chromatin modifications and their function. *Cell*. 2007;128(4):693-705. doi:10.1016/j.cell.2007.02.005.

7. Svotelis A, Bianco S, Madore J, et al. H3K27 demethylation by JMJD3 at a poised enhancer of anti-apoptotic gene BCL2 determines ERα ligand dependency. *EMBO J*. 2011;30(19):3947-3961. doi:10.1038/emboj.2011.284.

8. Strahl BD, Allis CD. The language of covalent histone modi ® cations. 2000;403(January):41-45.

9. Jenuwein T, Allis CD. Translating the histone code. *Science*. 2001;293(5532):1074-1080. doi:10.1126/science.1063127.

10. Paul A. Marks, Richard A. Rifkind, Victoria M. Richon, Ronald Breslow TM& WKK. Histone deacetylases and cancer: causes and therapies. *Nat Rev Cancer 1*. 2001;1:194-202. doi:doi:10.1038/35106079.

11. Agalioti T, Chen G, Thanos D. Deciphering the transcriptional histone acetylation code for a human gene. *Cell*. 2002;111(3):381-392.

12. Wang GG, Allis CD, Chi P. Chromatin remodeling and cancer, Part I: Covalent histone modifications. *Trends Mol Med*. 2007;13(9):363-372. doi:10.1016/j.molmed.2007.07.003.

13. Zhang Y, Yang X, Gui B, et al. Corepressor protein CDYL functions as a molecular bridge between polycomb repressor complex 2 and repressive chromatin mark trimethylated histone lysine 27. *J Biol Chem*. 2011;286(49):42414-42425. doi:10.1074/jbc.M111.271064.

14. Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol*. 2007;14(11):1025-1040. doi:10.1038/nsmb1338.

15. Filippakopoulos P, Picaud S, Mangos M, et al. Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell*. 2012;149(1):214-231. doi:10.1016/j.cell.2012.02.013.

16. Allfrey V, Faulkner R, Mirsky A. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *... Sci United States ...*. 1964;315(1938):786-794. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC300163/.

17. MURRAY K. THE OCCURRENCE OF EPSILON-N-METHYL LYSINE IN HISTONES. *Biochemistry*. 1964;3:10-15. doi:10.1021/bi00889a003.

18. Cosgrove MS. Writers and readers: deconvoluting the harmonic complexity of the histone code. *Nat Struct Mol Biol*. 2012;19(8):739-740. doi:10.1038/nsmb.2350.

19. Völkel P, Angrand P-O. The control of histone lysine methylation in epigenetic regulation. *Biochimie*. 2007;89(1):1-20. doi:10.1016/j.biochi.2006.07.009.

20. Li KK, Luo C, Wang D, Jiang H, Zheng YG. Chemical and Biochemical Approaches in the Study of Histone Methylation and Demethylation. *Med Res Rev*. 2012;(4):815-867. doi:10.1002/med.

21. Jenuwein T. Re-SET-ting heterochromatin by histone methyltransferases. *Trends Cell Biol*. 2001;11(6):266-273. http://www.ncbi.nlm.nih.gov/pubmed/11356363.

22. atdbio. Epigenetics. 2014. http://www.atdbio.com/content/56/Epigenetics.

23. Laumonnier F, Holbert S, Ronce N, et al. Mutations in PHF8 are associated with X linked mental retardation and cleft lip/cleft palate. *J Med Genet*. 2005;42(10):780-786. doi:10.1136/jmg.2004.029439.

24. Lee KK, Workman JL. Histone acetyltransferase complexes: one size doesn't fit all. *Nat Rev Mol Cell Biol*. 2007;8(4):284-295. doi:10.1038/nrm2145.

25. Kimura A, Matsubara K, Horikoshi M. A decade of histone acetylation: marking eukaryotic chromosomes with specific codes. *J Biochem*. 2005;138(6):647-662. doi:10.1093/jb/mvi184.

26. Yang X-J, Seto E. HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. *Oncogene*. 2007;26(37):5310-5318. doi:10.1038/sj.onc.1210599.

27. Kim G-W, Gocevski G, Wu C-J, Yang X-J. Dietary, metabolic, and potentially environmental modulation of the lysine acetylation machinery. *Int J Cell Biol*. 2010;2010(Figure 1):632739. doi:10.1155/2010/632739.

28. Dokmanovic M, Clarke C, Marks P a. Histone deacetylase inhibitors: overview and perspectives. *Mol Cancer Res*. 2007;5(10):981-989. doi:10.1158/1541-7786.MCR-07-0324.

29. De Ruijter AJM, van Gennip AH, Caron HN, Kemp S, van Kuilenburg ABP. Histone deacetylases (HDACs): characterization of the classical HDAC family. *Biochem J*. 2003;370(Pt 3):737-749. doi:10.1042/BJ20021321.

30. Longworth MS, Laimins L a. Histone deacetylase 3 localizes to the plasma membrane and is a substrate of Src. *Oncogene*. 2006;25(32):4495-4500. doi:10.1038/sj.onc.1209473.

31. Millard CJ, Watson PJ, Celardo I, et al. Class I HDACs share a common mechanism of regulation by inositol phosphates. *Mol Cell*. 2013;51(1):57-67. doi:10.1016/j.molcel.2013.05.020.

32. Arrowsmith CH, Bountra C, Fish P V, Lee K, Schapira M. Epigenetic protein families: a new frontier for drug discovery. *Nat Rev Drug Discov*. 2012;11(5):384-400. doi:10.1038/nrd3674.

33. Vojinovic J, Damjanov N. HDAC inhibition in rheumatoid arthritis and juvenile idiopathic arthritis. *Mol Med*. 2011;17(5-6):397-403. doi:10.2119/molmed.2011.00030.

34. Barton KM, Archin NM, Keedy KS, et al. Selective HDAC Inhibition for the Disruption of Latent HIV-1 Infection. *PLoS One*. 2014;9(8):e102684. doi:10.1371/journal.pone.0102684.

35. JH, Choi, Kwon HJ, Yoon BI, Kim JH, Han SU, Joo HJ KD. Expression profile of histone deacetylase 1 in gastric cancer tissues. *Jpn J Cancer Res*. 2001;92(12):1300-1304.

36. Halkidou K, Gaughan L, Cook S, Leung HY, Neal DE, Robson CN. Upregulation and nuclear recruitment of HDAC1 in hormone refractory prostate cancer. *Prostate*. 2004;59(2):177-189. doi:10.1002/pros.20022.

37. Zhang Z, Yamashita H, Toyama T, et al. HDAC6 Expression Is Correlated with Better Survival in Breast Cancer HDAC6 Expression Is Correlated with Better Survival in Breast Cancer. 2004:6962-6968.

38. Zhang Z, Yamashita H, Toyama T, et al. Quantitation of HDAC1 mRNA expression in invasive carcinoma of the breast. *Breast Cancer Res Treat*. 2005;94(1):11-16. doi:10.1007/s10549-005-6001-1.

39. Song J, Noh JH, Lee JH, et al. Increased expression of histone deacetylase 2 is found in human gastric cancer. *APMIS*. 2005;113(4):264-268. doi:10.1111/j.1600-0463.2005.apm_04.x.

40. Huang BH, Laban M, Leung CH-W, et al. Inhibition of histone deacetylase 2 increases apoptosis and p21Cip1/WAF1 expression, independent of histone deacetylase 1. *Cell Death Differ*. 2005;12(4):395-404. doi:10.1038/sj.cdd.4401567.

41. Wilson AJ, Byun D-S, Popova N, et al. Histone deacetylase 3 (HDAC3) and other class I HDACs regulate colon cell maturation and p21 expression and are deregulated in human colon cancer. *J Biol Chem*. 2006;281(19):13548-13558. doi:10.1074/jbc.M510023200.

42. Wen YD, Perissi V, Staszewski LM, et al. The histone deacetylase-3 complex contains nuclear receptor corepressors. *Proc Natl Acad Sci U S A*. 2000;97(13):7202-7207.

43. Guenther MG, Lane WS, Fischle W, Verdin E, Lazar MA, Shiekhattar R. WD40-repeat protein linked to deafness A core SMRT corepressor complex containing HDAC3 and TBL1 , a WD40-repeat protein linked to deafness. 2000:1048-1057. doi:10.1101/gad.14.9.1048.

44. Laherty CD, Yang W, Sun J, Davie JR, Seto E, Eisenman RN. Histone deacetylases associated with the mSin3 Corepressor Mediate Mad Transcriptional Repression. 1997;89:349-356.

45. Humphrey GW, Wang Y, Russanova VR, et al. Stable histone deacetylase complexes distinguished by the presence of SANT domain proteins CoREST/kiaa0071 and Mta-L1. *J Biol Chem*. 2001;276(9):6817-6824. doi:10.1074/jbc.M007372200.

46. Xue Y, Wong J, Moreno GT, Young MK, Co J, Wang W. NURD , a Novel Complex with Both ATP-dependent chromatin-remodeling and Histone Deacetylase Activities. *Mol Cell*. 1998;2:851-861.

47. Zhang Y, Ng H, Erdjument-bromage H, Tempst P, Bird A, Reinberg D. Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev*. 1999;13:1924-1935.

48. Watson PJ, Fairall L, Santos GM, Schwabe JWR. Structure of HDAC3 bound to co-repressor and inositol tetraphosphate. *Nature*. 2012;481(7381):335-340. doi:10.1038/nature10728.

49. Yun M, Wu J, Workman JL, Li B. Readers of histone modifications. *Cell Res*. 2011;21(4):564-578. doi:10.1038/cr.2011.42.

50. Kim J, Daniel J, Espejo A, et al. Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Rep*. 2006;7(4):397-403. doi:10.1038/sj.embor.7400625.

51. Musselman C a, Kutateladze TG. PHD fingers: epigenetic effectors and potential drug targets. *Mol Interv*. 2009;9:314-323. doi:10.1124/mi.9.6.7.

52. Mujtaba S, Zeng L, Zhou M-M. Structure and acetyl-lysine recognition of the bromodomain. *Oncogene*. 2007;26(37):5521-5527. doi:10.1038/sj.onc.1210618.

53. Filippakopoulos P, Knapp S. The bromodomain interaction module. *FEBS Lett*. 2012;586(17):2692-2704. doi:10.1016/j.febslet.2012.04.045.

54. Zeng L, Zhang Q, Gerona-Navarro G, Moshkina N, Zhou M-M. Structural basis of site-specific histone recognition by the bromodomains of human coactivators PCAF and CBP/p300. *Structure*. 2008;16(4):643-652. doi:10.1016/j.str.2008.01.010.

55. Josling G a, Selvarajah S a, Petter M, Duffy MF. The role of bromodomain proteins in regulating gene expression. *Genes (Basel)*. 2012;3(2):320-343. doi:10.3390/genes3020320.

56. Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, Zhou M. Structure and ligand of a histone acetyltransferase bromodomain bromodomain is similar to that of acetyl-CoA by histone acetyl-. *Nature*. 1999;399:491-496. www.nature.com.

57. Schindler U, Beckmann H, Cashmore a R. HAT3.1, a novel Arabidopsis homeodomain protein containing a conserved cysteine-rich region. *Plant J*. 1993;4(1):137-150.

58. Mellor J. It takes a PHD to read the histone code. *Cell*. 2006;126(1):22-24. doi:10.1016/j.cell.2006.06.028.

59. Bienz M. The PHD finger, a nuclear protein-interaction domain. *Trends Biochem Sci*. 2006;31(1):35-40. doi:10.1016/j.tibs.2005.11.001.

60. Wysocka J, Swigut T, Xiao H, et al. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature*. 2006;442(7098):86-90. doi:10.1038/nature04815.

61. Eissenberg JC. Structural biology of the chromodomain: form and function. *Gene*. 2012;496(2):69-78. doi:10.1016/j.gene.2012.01.003.

62. Urvalek AM, Gudas LJ. Retinoic Acid and Histone Deacetylases Regulate Epigenetic Changes in Embryonic Stem Cells. *J Biol Chem*. 2014. doi:10.1074/jbc.M114.556555.

63. Eberl HC, Spruijt CG, Kelstrup CD, Vermeulen M, Mann M. A map of general and specialized chromatin readers in mouse tissues generated by label-free interaction proteomics. *Mol Cell*. 2013;49(2):368-378. doi:10.1016/j.molcel.2012.10.026.

64. Alkhatib SG, Landry JW. The nucleosome remodeling factor. *FEBS Lett*. 2011;585(20):3197-3207. doi:10.1016/j.febslet.2011.09.003.

65. Agricola E, Randall R a, Gaarenstroom T, Dupont S, Hill CS. Recruitment of TIF1γ to chromatin via its PHD finger-bromodomain activates its ubiquitin ligase and transcriptional repressor activities. *Mol Cell*. 2011;43(1):85-96. doi:10.1016/j.molcel.2011.05.020.

66. Mulligan P, Westbrook TF, Ottinger M, et al. CDYL bridges REST and histone methyltransferases for gene repression and suppression of cellular transformation. *Mol Cell*. 2008;32(5):718-726. doi:10.1016/j.molcel.2008.10.025.

67. Shia W-J, Pattenden SG, Workman JL. Histone H4 lysine 16 acetylation breaks the genome's silence. *Genome Biol*. 2006;7(5):217. doi:10.1186/gb-2006-7-5-217.

68. Creyghton MP, Cheng AW, Welstead GG, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A*. 2010;107(50):21931-21936. doi:10.1073/pnas.1016071107.

69. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann S a, Flynn R a, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*. 2011;470(7333):279-283. doi:10.1038/nature09692.

70. Rintisch C, Heinig M, Bauerfeind A, Schafer S, Mieth C, Patone G, Hummel O, Chen W, Cook S, Cuppen E, Colomé-Tatché M, Johannes F, Jansen RC NH, Werner M, Pravenec M, Vingron M HN. Natural variation of histone modification and its impact on gene expression in the rat genome. *Genome Res*. 2014;24(6):942-953. doi:10.1101/gr.169029.113.

71. Suka N, Suka Y, Carmen AA, et al. Technique Histone Acetylation Site Usage in Yeast Heterochromatin and Euchromatin. *Mol Cell*. 2001;8(2):473-479. doi:10.1016/S1097-2765(01)00301-X.

72. Garcia B a, Hake SB, Diaz RL, et al. Organismal differences in post-translational modifications in histones H3 and H4. *J Biol Chem*. 2007;282(10):7641-7655. doi:10.1074/jbc.M607900200.

73. Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G. Genome regulation by polycomb and trithorax proteins. *Cell*. 2007;128(4):735-745. doi:10.1016/j.cell.2007.02.009.

74. Tie F, Banerjee R, Stratton C a, et al. CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. *Development*. 2009;136(18):3131-3141. doi:10.1242/dev.037127.

75. Hansen KH, Helin K. Epigenetic inheritance through self-recruitment of the polycomb repressive complex 2. *Epigenetics*. 2009;4(3):133-138. http://www.ncbi.nlm.nih.gov/pubmed/19377285.

76. Lee TI, Jenner RG, Boyer L a, et al. Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell*. 2006;125(2):301-313. doi:10.1016/j.cell.2006.02.043.

77. Lund AH, van Lohuizen M. Polycomb complexes and silencing mechanisms. *Curr Opin Cell Biol*. 2004;16(3):239-246. doi:10.1016/j.ceb.2004.03.010.

78. Margueron R, Reinberg D. The Polycomb complex PRC2 and its mark in life. *Nature*. 2011;469(7330):343-349. doi:10.1038/nature09784.

79. Young MD, Willson T a, Wakefield MJ, et al. ChIP-seq analysis reveals distinct H3K27me3 profiles that correlate with transcriptional activity. *Nucleic Acids Res*. 2011;39(17):7415-7427. doi:10.1093/nar/gkr416.

80. Mohn F, Weber M, Rebhan M, et al. Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol Cell*. 2008;30(6):755-766. doi:10.1016/j.molcel.2008.05.007.

81. Boyer L a, Plath K, Zeitlinger J, et al. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature*. 2006;441(7091):349-353. doi:10.1038/nature04733.

82. Pengelly AR, Copur Ö, Jäckle H, Herzig A, Müller J. A histone mutant reproduces the phenotype caused by loss of histone-modifying factor Polycomb. *Science*. 2013;339(6120):698-699. doi:10.1126/science.1231382.

83. Leeb M, Wutz A. Establishment of epigenetic patterns in development. *Chromosoma*. 2012;121(3):251-262. doi:10.1007/s00412-012-0365-x.

84. Bowman SK, Deaton AM, Domingues H, et al. H3K27 modifications define segmental regulatory domains in the Drosophila bithorax complex. *Elife*. 2014:e02833. doi:10.7554/eLife.02833.

85. Pietersen AM, van Lohuizen M. Stem cell regulation by polycomb repressors: postponing commitment. *Curr Opin Cell Biol*. 2008;20(2):201-207. doi:10.1016/j.ceb.2008.01.004.

86. Voigt P, Tee WW, Reinberg D. A double take on bivalent promoters. *Genes Dev*. 2013;27(12):1318-1338. doi:10.1101/gad.219626.113.

87. Paterno GD, Li Y, Luchman H a., Ryan PJ, Gillespie LL. cDNA Cloning of a Novel, Developmentally Regulated Immediate Early Gene Activated by Fibroblast Growth Factor and Encoding a Nuclear Protein. *J Biol Chem*. 1997;272(41):25591-25595. doi:10.1074/jbc.272.41.25591.

88. Paterno GD, Mercer FC, Chayter JJ, Yang X, Robb JD, Gillespie LL. Molecular cloning of human er1 cDNA and its differential expression in breast tumours and tumour-derived cell lines. *Gene*. 1998;222(1):77-82.

89. Thorne LB, Grant AL, Paterno GD, Gillespie LL. Cloning and characterization of the mouse ortholog of mi-er1. *DNA Seq*. 2005;16(3):237-240. doi:10.1080/10425170500069783.

90. Paterno GD, Ding Z, Lew Y-Y, Nash GW, Mercer FC, Gillespie LL. Genomic organization of the human mi-er1 gene and characterization of alternatively spliced isoforms: regulated use of a facultative intron determines subcellular localization. *Gene*. 2002;295(1):79-88.

91. McCarthy PL, Mercer FC, Savicky MWJ, Carter B a, Paterno GD, Gillespie LL. Changes in subcellular localisation of MI-ER1 alpha, a novel oestrogen receptor-alpha interacting protein, is associated with breast cancer progression. *Br J Cancer*. 2008;99(4):639-646. doi:10.1038/sj.bjc.6604518.

92. McCarthy PL, Paterno GD, Gillespie LL. Protein expression pattern of human MIER1 alpha, a novel estrogen receptor binding protein. *J Mol Histol*. 2013;44(4):469-479. doi:10.1007/s10735-012-9478-z.

93. Ding Z, Gillespie LL, Paterno GD. Human MI-ER1 Alpha and Beta Function as Transcriptional Repressors by Recruitment of Histone Deacetylase 1 to Their Conserved ELM2 Domain Human MI-ER1 Alpha and Beta Function as Transcriptional Repressors by Recruitment of Histone Deacetylase 1 to Their C. *Mol Cell Biol*. 2003;23(1):250-258. doi:10.1128/MCB.23.1.250.

94. Ding Z, Gillespie LL, Mercer FC, Paterno GD. The SANT domain of human MI-ER1 interacts with Sp1 to interfere with GC box recognition and repress transcription from its own promoter. *J Biol Chem*. 2004;279(27):28009-28016. doi:10.1074/jbc.M403793200.

95. Blackmore TM, Mercer CF, Paterno GD, Gillespie LL. The transcriptional cofactor MIER1-beta negatively regulates histone acetyltransferase activity of the CREB-binding protein. *BMC Res Notes*. 2008;1:68. doi:10.1186/1756-0500-1-68.

96. Wang L, Charroux B, Kerridge S, Tsai C-C. Atrophin recruits HDAC1/2 and G9a to modify histone H3K9 and to determine cell fates. *EMBO Rep*. 2008;9(6):555-562. doi:10.1038/embor.2008.67.

97. Clements J a, Mercer FC, Paterno GD, Gillespie LL. Differential splicing alters subcellular localization of the alpha but not beta isoform of the MIER1 transcriptional regulator in breast cancer cells. *PLoS One*. 2012;7(2):e32499. doi:10.1371/journal.pone.0032499.

98. Li S, Paterno GD, Gillespie LL. Nuclear localization of the transcriptional regulator MIER1α requires interaction with HDAC1/2 in breast cancer cells. *PLoS One*. 2013;8(12):e84046. doi:10.1371/journal.pone.0084046.

99. Massagué J, Seoane J, Wotton D. Smad transcription factors. *Genes Dev*. 2005;19(23):2783-2810. doi:10.1101/gad.1350705.

100. Bryant GO, Prabhu V, Floer M, et al. Activator control of nucleosome occupancy in activation and repression of transcription. *PLoS Biol*. 2008;6(12):2928-2939. doi:10.1371/journal.pbio.0060317.

101. Aasland R, Stewart AF GT. The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and , TFIIIB. *Trends Biochem Sci*. 21(3):87-89.

102. De la Cruz X, Lois S, Sánchez-Molina S, Martínez-Balbás M a. Do protein motifs read the histone code? *Bioessays*. 2005;27(2):164-175. doi:10.1002/bies.20176.

103. Boyer LA, Langer MR, Crowley KA, Tan S, Denu JM, Peterson CL. Essential Role for the SANT Domain Chromatin Remodeling Enzymes. 2002;10:935-942. doi:10.1016/S1097-2765(02)00634-2.

104. Yu J, Li Y, Ishizuka T, Guenther MG, Lazar M a. A SANT motif in the SMRT corepressor interprets the histone code and promotes histone deacetylation. *EMBO J*. 2003;22(13):3403-3410. doi:10.1093/emboj/cdg326.

105. Barrios AP, Gómez A V, Sáez JE, et al. Differential properties of transcriptional complexes formed by the CoREST family. *Mol Cell Biol*. 2014. doi:10.1128/MCB.00083-14.

106. You a, Tong JK, Grozinger CM, Schreiber SL. CoREST is an integral component of the CoREST- human histone deacetylase complex. *Proc Natl Acad Sci U S A*. 2001;98(4):1454-1458. doi:10.1073/pnas.98.4.1454.

107. Solari F, Bateman a, Ahringer J. The Caenorhabditis elegans genes egl-27 and egr-1 are similar to MTA1, a member of a chromatin regulatory complex, and are redundantly required for embryonic patterning. *Development*. 1999;126(11):2483-2494. http://www.ncbi.nlm.nih.gov/pubmed/10226007.

108. Plotnikov AN, Yang S, Zhou TJ, Rusinova E, Frasca A, Zhou M-M. Structural insights into acetylated-histone H4 recognition by the bromodomain-PHD finger module of human transcriptional coactivator CBP. *Structure*. 2014;22(2):353-360. doi:10.1016/j.str.2013.10.021.

109. Steiner S, Magno A, Huang D, Caflisch A. Does bromodomain flexibility influence histone recognition? *FEBS Lett*. 2013;587(14):2158-2163. doi:10.1016/j.febslet.2013.05.032.

110. Gray IC, Barnes MR. Amino Acid Properties and Consequences of Substitutions. Gray MRB and IC, ed. *Bioinforma Genet*. 2003;4:289-304.

111. Wysocka J. Identifying novel proteins recognizing histone modifications using peptide pull-down assay. *Methods*. 2006;40(4):339-343. doi:10.1016/j.ymeth.2006.05.028.

112. Chabot PR, Raiola L, Lussier-Price M, et al. Structural and functional characterization of a complex between the acidic transactivation domain of EBNA2 and the Tfb1/p62 subunit of TFIIH. *PLoS Pathog*. 2014;10(3):e1004042. doi:10.1371/journal.ppat.1004042.

113. Boyer LA, Latek RR, Peterson CL. The SANT domain: a unique histone-tail-binding module? *Nat Rev Mol Cell Biol*. 2004;5(February):158-163.

114. Wang L, Grossman SR, Kieff E. Epstein-Barr virus nuclear protein 2 interacts with p300, CBP, and PCAF histone acetyltransferases in activation of the LMP1 promoter. *Proc Natl Acad Sci U S A*. 2000;97(1):430-435.

115. Ember SWJ, Zhu J-Y, Olesen SH, et al. Acetyl-lysine binding site of bromodomain-containing protein 4 (BRD4) interacts with diverse kinase inhibitors. *ACS Chem Biol*. 2014;9(5):1160-1171. doi:10.1021/cb500072z.

116. Lee H, Mok KH, Muhandiram R, et al. Local structural elements in the mostly unstructured transcriptional activation domain of human p53. *J Biol Chem*. 2000;275(38):29426-29432. doi:10.1074/jbc.M003107200.

117. Lee CW, Martinez-yamout M a, Dyson HJ, Wright PE. NIH Public Access. *Biochemistry*. 2011;49(46):9964-9971. doi:10.1021/bi1012996.Structure.

118. Li H, Ilin S, Wang W, et al. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature*. 2006;442(7098):91-95. doi:10.1038/nature04802.

119. Min J, Zhang Y, Xu R-M. Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genes Dev*. 2003;17(15):1823-1828. doi:10.1101/gad.269603.

120. Roche J, Nasarre P, Gemmill R, et al. Global Decrease of Histone H3K27 Acetylation in ZEB1-Induced Epithelial to Mesenchymal Transition in Lung Cancer Cells. *Cancers (Basel)*. 2013;5(2):334-356. doi:10.3390/cancers5020334.

121. Javaid S, Zhang J, Anderssen E, et al. Dynamic chromatin modification sustains epithelial-mesenchymal transition following inducible expression of snail-1. *Cell Rep*. 2013;5(6):1679-1689. doi:10.1016/j.celrep.2013.11.034.

122. Leroy G, Dimaggio P a, Chan EY, et al. A quantitative atlas of histone modification signatures from human cancer cells. *Epigenetics Chromatin*. 2013;6(1):20. doi:10.1186/1756-8935-6-20.

123. Holm K, Grabau D, Lövgren K, et al. Global H3K27 trimethylation and EZH2 abundance in breast tumor subtypes. *Mol Oncol*. 2012;6(5):494-506. doi:10.1016/j.molonc.2012.06.002.

124. Wilson BJ, Tremblay AM, Deblois G, Sylvain-Drolet G, Giguère V. An acetylation switch modulates the transcriptional activity of estrogen-related receptor alpha. *Mol Endocrinol*. 2010;24(7):1349-1358. doi:10.1210/me.2009-0441.