Spring 5-1-2014

Molecular Mechanisms for Regulating the Assembly of the Mixed Lineage Leukemia-1 (MLL1) Core Complex

Michael Hu

Follow this and additional works at: https://surface.syr.edu/honors_capstone

Part of the Biochemistry Commons

Recommended Citation
https://surface.syr.edu/honors_capstone/758

This Honors Capstone Project is brought to you for free and open access by the Syracuse University Honors Program Capstone Projects at SURFACE. It has been accepted for inclusion in Syracuse University Honors Program Capstone Projects by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.
Molecular Mechanisms for Regulating the Assembly of the Mixed Lineage Leukemia-1 (MLL1) Core Complex

A Capstone Project Submitted in Partial Fulfillment of the Requirements of the Renée Crown University Honors Program at Syracuse University

Michael Hu
Candidate for Bachelor of Science Degree and Renée Crown University Honors
May 2014

Honors Capstone Project in Biochemistry

Capstone Project Advisor: Dr. Michael Cosgrove, Professor

Capstone Project Reader: Dr. Sam Chan, Professor

Honors Director: Stephen Kuusisto, Director

Date: May 6, 2014
Abstract

Epigenetics is the study of heritable traits that are not caused by alterations to DNA. An example of such a mechanism is histone methylation, specifically histone H3 lysine 4 (H3K4). Histones are proteins that wrap and package DNA into nucleosomes. Modifications to a histone’s chemistry can up or down regulate specific gene expression programs through changes in nucleosome positioning. Human mixed lineage leukemia protein-1 (MLL1) is a member of the SET1 family of H3K4 methyltransferases. Previous studies in the Cosgrove lab have shown that an interaction between two protein sub-complexes, MLL1 and WRAD (WDR5, RbBP5, Ash2L, DPY30), leads to dimethylation. H3K4 dimethylation plays an important role in normal development and hematopoiesis in humans. Disruption of core complex formation from amino acid point mutations can lead to a decrease, or even complete loss, of dimethylation.

Kabuki Syndrome (KS) is a multiple malformation disorder that was recently discovered to be associated with mutations in mixed lineage leukemia protein-2 (MLL2), a paralog of MLL1 and member of the SET1 family of proteins. Using the experimental model established with MLL1, tests were conducted that determined the importance of a highly conserved arginine residue (R). In MLL1, this R3765 (MLL2 R5340) was mutated into leucine, which revealed a severe loss of H3K4 dimethylation resulting from failure of MLL1 to interact with WDR5. Additional KS and non-KS (control) MLL1 mutations from MLL2 screens of patients were tested and suggested a correlation between changes in H3K4 methylation and KS.

Protein arginine deiminases (PADs) are enzymes that post translationally modify arginine into citrulline, a non-standard amino acid. Previous work with KS led us to hypothesize a possible regulation system of PAD and SET1 family proteins. While PAD is able to disrupt complex formation by changing R3765 into citrulline, thereby down-regulating dimethylation, mutations in SET1 family proteins can elevate dimethylation when in complex with WRAD. Thereby creating an “on” and “off” system of protein interaction. Previous work in the Cosgrove lab showed preferential interaction between PAD2 and MLL1. Initial assays were conducted and optimized that supported this hypothesis, as well as laying the ground work for future experimentation.
Executive Summary

Mixed lineage leukemia protein-1 (MLL1) and WRAD (WDR5, RbBP5, Ash2L, DPY30) are two parts of a protein core complex called MWRAD. Separately, MLL1 and WRAD can only monomethylate, which means they add one methyl (CH$_3$) group onto a histone called histone H3 lysine 4 (H3K4). When in complex, MWRAD is able to dimethylate or add two methyl groups. Histones are proteins that consolidate DNA by wrapping them up. Without histones the DNA would be very long and each human cell would contain 1.8 meters of DNA, as opposed to 90 micrometers, a ten-thousand fold difference. Usually genes and traits are thought to be passed down through each individual’s genetic DNA code. However, there exists a new field of study called epigenetics, whereby inherited traits are passed down through modifications outside of DNA, an example being histone methylation.

The Cosgrove lab has worked extensively with this model system and have discovered a highly conserved six amino acid motif in MLL1 called the Win motif that acts as a bridge with WDR5. Mutations in this region lead to a disruption of complex formation between MLL1 and WRAD, thereby decreasing the amount of H3K4 dimethylation. This can be thought as a lock and key mechanism where the Win motif is inserted into a WRAD binding site.

My work has primarily revolved around that region of MLL1 and understanding the mechanisms behind the regulation of this MLL1 core complex. During my three years, I was tasked with two separate projects. The first focused on Kabuki Syndrome (KS), a congenital disorder that has been linked with MLL2. MLL1 and MLL2 are part of a family of proteins, so they are very similar in nature. Therefore, it was hypothesized that we could test KS and non-KS (controls) mutations found in MLL2 using MLL1. Results showed a correlation between mutations in KS patients and a loss of dimethylation.

After successfully publishing a paper with members of the Cosgrove lab, I was assigned to a second project dealing with protein arginine deiminase (PAD), an enzyme that modifies an amino acid called arginine. Prior work showed a relation between PAD2 and MLL1 leading us to believe that there could exist a regulation mechanism for H3K4 dimethylation. Essentially PAD2 would hypothetically lower dimethylation in the case that there is an overexpression of dimethylation, as the human body maintains a delicate homeostasis by keeping gene expression in a specific range. Furthermore, detrimental mutations to PAD2 could be a cause for KS, as it would change one of the highly conserved amino acids in the Win motif. Unfortunately I was only able to lay the ground works for this project due to time constraints. Future work will hopefully be able to verify the data that was obtained.
Acknowledgements

First, I want to thank Professor Michael Cosgrove for giving me the unique opportunity of working in his research lab. My experiences over the past three years have helped me develop my understanding of the scientific method and greatly improved my research skills. The knowledge I have learned has already helped me get a job out of college at a research company, Ortho Clinical Diagnostics, in Rochester.

Also, thanks to all the members of the Cosgrove lab, both past and present, who have supported me during experiments, answered my multitude of questions, and helped me write this thesis. I would not have made it through all the data collection without you. Especially my graduate student, Venkat, for taking me under his wing, despite being stranded in India during the first few weeks.

In addition, thank you Syracuse University Renée Crown Honors Program for providing me the funding during the summer of 2013 through the Crown and Wise-Marcus 50 Year Friendship Award. The resources you offer students allowed me to not only pursue my research interests, but also produce a final product through the capstone project.

Lastly, I wanted to thank the SU and SUNY Upstate biochemistry faculty for sharing their advice and insight regarding the science field. Their passion and energy was quite contagious. Four years ago, I chose Syracuse University without knowing where it was located. The professors and advisors here have truly given me an amazing college experience. I’m extremely lucky to have attended such a great intuition that cares about their undergraduate students.
## Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>i</td>
</tr>
<tr>
<td>Executive Summary</td>
<td>ii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>iii</td>
</tr>
<tr>
<td>Table of Content</td>
<td>1</td>
</tr>
<tr>
<td>I. Introduction</td>
<td>2</td>
</tr>
<tr>
<td>II. Materials and Method</td>
<td>8</td>
</tr>
<tr>
<td>III. Results</td>
<td>12</td>
</tr>
<tr>
<td>IV. Discussion</td>
<td>26</td>
</tr>
<tr>
<td>References</td>
<td>28</td>
</tr>
</tbody>
</table>
Chapter I

Introduction

*The Human Mixed Lineage Leukemia Protein-1 (MLL1) Core Complex*

Nucleosomes are fundamental units in genetic regulation and they consist of DNA wrapped around eight histone proteins. Changes in the position of a nucleosome can cause varying levels of over-expression or suppression of genes. Chromatin, containing multiple nucleosomes, can be transcriptionally activated or deactivated with histone lysine methylation, a crucial epigenetic system\(^1\). Epigenetics is the study of gene expression programs that does not originate from alterations in the DNA. Post-translational modifications of histones can reveal or cover binding sites for proteins involved in gene expression. Mixed lineage leukemia protein -1, MLL1, named as such due to mutations frequently leading to acute leukemia, is one of six members in the SET1 family of H3K4 methyltransferases with the other five being: MLL2, MLL3, MLL4, SET1a, and SET1b\(^1\). The SET1 family is one of many SET domain proteins that exist based on similarities in sequence and structural features. Alone, the SET1 family can only participate in the slow monomethylation of H3K4\(^1\). However, dimethylation is possible in the presence
of the WDR5, RbBP5, ASH2L, and DPY-30 (WRAD) complex\(^1\). WRAD acts similarly to MLL1 and can monomethylate H3K4\(^1\). MWRAD, in contrast, shows a ~600 fold increase in H3K4 dimethylation\(^1\). Interestingly, mono- and dimethylation occur sequentially at two unique sites in the MWRAD complex\(^1\). Previous and current work in Dr. Michael Cosgrove’s lab focuses on understanding the structure and function of MWRAD. H3K4 dimethylation is important in the normal development and hematopoiesis of humans\(^1\). Under wild-type conditions, the MLL1 core complex is capable of H3K4 dimethylation, but point mutations in the amino acid sequence altering the amount of dimethylation have been linked with a variety of human diseases\(^1\). If the core complex is unable to form, then there will only be monomethylation\(^1\).

**Post Translational Modifications in Kabuki Syndrome**

Kabuki Syndrome (OMIM 147920) is a human multiple malformation disorder categorized by intellectual disability, facial abnormalities, heart problems, hearing loss, growth deficiency, and short stature\(^2\). Due to the rare nature of this disease (1 in 32,000 with only 400 reported cases), it was not until 1981 that two separate scientists, Norio Niikawa and Yoshikazu Kuroki, identified Kabuki Syndrome (KS)\(^2\). Named as such due to the similarities with Kabuki makeup, a type of Japanese theatre\(^2\). Since most reported incidences arose *de novo*, KS is suggested to be an autosomal dominant disorder or X-linked recessive\(^2\). Classification of KS is still not completely understood, as many of the conditions are shared with other diseases like Down Syndrome\(^2\). Currently, four of the five cardinal features must be met in order to make a KS diagnosis, including: facial appearance, skeletal
abnormalities, fetal finger pads, intellectual disability, and short stature\(^3\). Exome sequencing in the early-2000s, as well as large spectrum analysis of KS patients that followed were unable to pinpoint one shared gene at fault, however the majority of individuals tested had a mutation in MLL2\(^2,3\). Nine of ten patients in the initial screen had an altered MLL2 gene, while two thirds of a larger (\(n = 43\)) screen had mutations in MLL2\(^2\). Seventy two percent of a 110 KS family screen had adversely affected MLL2\(^3\). Though not all those diagnosed with KS had the MLL2 mutant, a possible explanation for those without MLL2 mutations could be incorrect identification, the existence of another related developmental disease, or mutations existing in the noncoding region\(^2\). Loss-of-function mutations in MLL2 likely alters H3K4 methylation in KS and, as such, is believed to be linked to the problems in early human development for Kabuki Syndrome patients\(^2\). In addition, tests done on MLL2 knock out mouse models were lethal\(^3\). KS is one of many syndromes classified by developmental delay and disrupted histone methylation\(^3\).

**Significance of the WDR5 Interaction (Win) Motif**

The WD repeat protein-5 (WDR5) is a conserved component in SET1 family complexes that is shared across many different species\(^4\). It was discovered in 2008 that there exists a bridge between MLL1 and WDR5 that assists in core complex formation\(^4\). The *Win* motif, a six amino acid residue sequence (GSARAE) is located in the N-SET region of MLL1 and contains a highly conserved arginine (R3765) that was shown to be bind in the pocket of WDR5\(^4\). R3765 exists in all members of the SET1 family and is important in the enzymatic activity of the
MWRAD complex. Mutations in the Win motif could lead to a significant disruption in the dimethylation activity attributed to normal hematopoiesis and human growth.

**Mutations Affecting Histone H3 Lysine 4 (H3K4) Methylation**

During the KS mutation screen, MLL2 arginine 5340 was found to be mutated into a leucine. This amino acid position is homologous to MLL1 R3765 in the Win motif. Due to the correlation between two previous research papers, it was hypothesized that by utilizing the established in vitro model system of MLL1, the effects of a mutation in MLL2 could be determined given the great amount of similarity between the two proteins. MLL1 is a paralog of MLL2, thereby presenting the opportunity to understand the effects of the KS MLL2 R5340 mutation. This change in amino acid to leucine leads to a loss of charge, shorter atomic structure, and increased hydrophobicity. In addition, a non-KS related residue R3746 was mutated into a histidine as a control mutation that should not impact H3K4 methylation. Other screened MLL2 mutations were tested by a graduate student in the research lab. The identification and understanding of Kabuki Syndrome point mutations will hopefully open doors for possible treatments or earlier diagnosis.

**Peptidylarginine Deiminase (PAD) and the Impact of Citrullination**

PAD enzymes catalyze the citrullination of arginine and have been implicated in diseases such as rheumatoid arthritis, breast cancer, and multiple sclerosis. The conversion of the NH$_2^+$ group on arginine to a keto group on citrulline leads to a
loss of the positive charge. The mammalian PAD family of enzymes consist of PAD1, PAD2, PAD3, and PAD4\(^8\). Of the four, PAD2 is expressed in the skeletal muscle, brain, spleen, and secretory glands\(^8\). In order for these enzymes to function properly, they require calcium ions (Ca\(^{2+}\))\(^8\). It has been hypothesized that PAD could be used as a possible inhibitor to histone methyltransferases\(^9\). For example, while MLL1 and WRAD interaction leads to dimethylation, PAD could possibly disrupt the interaction if there was an overexpression of dimethylation. In this way, citrullination can be used to counter a gain-of-function genetic mutation. Given that histone methylation is an epigenetic modification that has been linked with cancer, enzymes that directly or indirectly inhibit methylation could act as drugs to counter uncontrolled cell growth\(^9\). Furthermore, this field of study has only just begun to scratch the surface, if an inhibitor with specificity towards a histone methyltransferase were created, then it could aid in the fight against cancer, as well as other diseases\(^9\). In the past, the belief was that histone methylation was a permanent genetic modification\(^10\). Prior work with MLL1 R3765 indicated that a mutation on this amino acid could lead to a complete loss of dimethylation. It is hypothesized that an overexpression of PAD2 on this arginine could lead to a patient with Kabuki Syndrome. It was shown in previous experiments that PAD2 catalyzes citrullination of R3765 and R3771 in MLL1. While based on previous work, citrullination of R3765 will likely disrupt complex formation, it is unclear what functional consequences result from R3771 citrullination.
Research Project Overview

The goal of work in my thesis was to understand how MLL1 core complex (Fig. 1) formation is regulated by R3765 and R3771 in MLL1.

In order to better understand the impact of MLL1 missense mutations, an experimental model (Fig. 2) was used to compare complex formation and H3K4 methylation with wild-type MLL1.
Chapter II

Materials and Methods

*Mutagenesis.* Point mutations were made using a site directed mutagenesis kit (QuickChange II, Stratagene) on truncated wild-type human MLL1 containing amino acid residues 3745-3969 in a pGST vector to introduce R3746H, R3765L, and R3771T mutations. Plasmids were sequenced by the DNA Sequencing Core Facility at SUNY Upstate Medical University to verify the presence of the desired mutation and the absence of additional unintended mutations. Primers were purchased from Integrated DNA Technologies.

*Protein Expression/Purification.* Wild-type and mutant MLL1, as well as full-length wild-type human WDR5, RbBP5, ASH2L, and DPY-30 were individually expressed in *Escherichia coli* (Rosetta II, Novagen) and purified using affinity chromatography, as previously described⁴. In addition, proteins were dialyzed for buffer exchange and further purified using a gel filtration column (Superdex, 200, GE Healthcare) with filtered 20mM Tris (pH 7.5), 300mM NaCl, 1mM Tris(2-carboxyethyl)phosphine,
and 1µM ZnCl$_2$. Fractions from each step were selectively chosen after using SDS-PAGE gels to confirm the protein by molecular weight and maximize purity. Lastly, purified proteins were concentrated, if necessary, using a centrifuge concentration tubes (Millipore, Minicon).

**SDS-PAGE.** Polyacrylamide gel electrophoresis (PAGE) was utilized to determine concentrations and purity of protein samples, as well as separating experimental time course reactions into individual protein components. Sodium dodecyl sulfate (SDS), a detergent, denatures proteins by unfolding them to become linear, thereby assisting in sample analysis. Bands on the gel (BioRad, Mini Protein TGX) were compared to a Precision Plus Protein Standard (BioRad). Coomassie stained and destained images were captured using a Gel Doc (BioRad) and bands were quantified using Image Lab.

**Nanodrop Spectroscopy.** Concentrations (mg/mL) of gel filtration purified proteins were determined with a Nandrop 2000c (Thermo Scientific) at 280nm. Relative values were obtained using 2µL of protein.

**Bradford Assays.** Accurate concentrations (mg/mL) of proteins were obtained through Bradford Assays compared with a standard curve plot. Absorbance at 280nm were used to standardize concentrations of proteins prior to experiments.
Methyltransferase Assays. In vitro experiments using the MLL1 core complex at 7µM were done over a time course to test for mono-, di-, and tri-methylation activity as previously described. 250µM of S-Adenosyl Methionine (AdoMet), as well as 10µM of H3 (residues 1-20) peptide, were added to the MWRAD complex and incubated at 15°C for 24 hours. Aliquots were withdrawn at different time points and quenched using 0.5% Triflouroacetic acid (TFA). Each sample was then diluted 1:5 using α-cyano-4-hydroxycinnamic acid, dried, and analyzed in reflectron mode on the SUNY ESF MALDI-TOF mass spectrometer (Bruker Autoflex III). Data obtained was the average of 100 shots/position at 5 different positions chosen at random for each sample. Relative methylation values were determined using mMass and plotted on a Microsoft Excel spread sheet to compare amounts of methylation over time.

Analytical Ultracentrifugation. Wild-type and mutant MLL1 were mixed with various stoichiometric concentrations of WDR5 to determine if binding was effected by a point mutation. 3mm two-sector charcoal filled Epon centerpieces with quartz windows were assembled and samples were loaded. The Beckman Coulter ProteomLab XL-A analytical ultracentrifuge with absorbance optics was run at 10°C and 50,000RPM using a 4 hole An-50 Ti rotor. The program scanned 0-min time intervals for 300 scans and the data was analyzed using the continuous distribution method c(s) in SEDFIT.
**GST Pulldown Assays.** Wild-type and mutant GST-MLL1 proteins were mixed with the WRAD complex to compare binding affinity of the MLL1 core complex after citrullination with PAD2. The proteins, assay buffer (50mM Tris (pH 8.5), 0.2M NaCl, 5% Glycerol, and 10mM CaCl$_2$), and PAD2 were incubated at room temperature with glutathione beads over several time points. Reactions were stagger started, so as to end at the same time. Buffer washes were conducted to eliminate non-GST-bound proteins. SDS-PAGE was used to separate proteins on a gel and determine if WRAD successfully bound to GST-MLL1 in the presence of PAD2 over time.

**Western Blotting.** Novex NuPAGE gels were used to prepare electrophoresis samples. However, instead of coomassie staining, bands were transferred to a nitrocellulose membrane using electroblotting and 20x NuPAGE transfer buffer. The membrane was then verified using Poncea S stain and the antibody was added. Lastly, the membrane was probed using an antibody that detects the citrullinated MLL1 protein.
Chapter III

Results

Expression and Purification of MLL1 Mutants

Human MLL1 3745-3969 mutants R3765L, R3746H, and R3771T were expressed by inducing at OD600 1.0 using Isopropyl β-D-1 thiogalactopyranoside (IPTG), which is a molecule that promotes transcription of the lac operon. Verification of induction can be seen on the left (Fig. 3) with an SDS-PAGE gel.

Figure 3 shows:
Lane 1 = BioRad Standard Ladder (50kDa)
Lane 2 = Uninduced MLL1-GST Protein
Lane 3 = Induced MLL1-GST Protein

The induced lane shows a significant increase in band intensity of protein between lane 2 and lane 3, which confirms induction.
Harvested pellets were lysed through as microfluidizer with an EDTA Free Protease Inhibitor Cocktail Tablet to avoid protein degradation. Purification of MLL1 mutants proceeded through a GST affinity chromatography, dialysis for buffer exchange, second GST affinity chromatography run, and finally gel filtration chromatography. At the end of each column run, sample tubes were chosen based off UV 280nm elution peaks when MLL1 came off the beads. These tubes, as well as the input and flow through, were displayed on a SDS-PAGE gel to analyze intensity of bands correlating to amount of protein and presence of contaminants. Sample aliquots of protein were saved throughout the process of purification for different experiments and stored at -80°C. GST fusion proteins after GST run 1 were used for GST pulldown assays, cleaved GST run 2 proteins were used for methylation assays, and gel filtration MLL1 was used for binding assays using sedimentation velocity analytical ultracentrifugation. Band intensity shown after gel filtration, though high in purity, were noticeably weaker and a loss of protein was seen each time for all mutants and wild-type MLL1 proteins. Destained coomassie gels can be seen below (Fig. 4, Fig. 5, Fig. 6) correlating to GST1, GST2, and gel filtration.

Figure 4 (missing ladder) shows:
Lane 1 = Input
Lane 2 = Flowthrough
Lane 3 – 9 = Select MLL1 Protein Tubes
Figure 5 shows:
Lane 1 = BioRad Standard Ladder (25kDa)
Lane 2 = Dialysis Input
Lane 3 = Dialysis Output
Lane 4 – 15 = Select MLL1 Protein Tubes

Figure 6 shows:
Lane 1 = BioRad Standard Ladder (25kDa)
Lane 2 = Non-Concentrated Gel Filtration MLL1 Protein
Lane 3 = Concentrated Gel Filtration MLL Protein

**MLL1 R3765L – Disruption of Complex Formation**

Hypothesis: If the R3765L missense mutation in MLL1 causes loss-of-function due to changes in physical properties of the amino acid from an arginine to a
leucine, then this would be indicative of R5340L doing the same in MLL2 due to MLL1 and MLL2 being paralogs to each other in humans.

In order to analyze methylation activity (mono-, di-, and trimethylation activity), a time course experiment and mass spectrometry were utilized to determine if MLL1 and WRAD were binding based on methylation data compared to wild-type MLL1. Each methylation (CH$_3$) results in a 15Da increase. Plotted data did indicate a severe loss of dimethylation between M$_{WT}$WRAD (Fig. 7a, Fig. 7b) and M$_{R3765L}$WRAD (Fig. 8a, Fig. 8b), while there was not a noticeable difference in the level of monomethylation. This assay was repeated to validate the correlation between mutation and disruption of binding.

Below are the results from the experiment conducted on August 2, 2012. Figures 7a (WT) and 8a (R3765L) are the plots from mMass (mono-, di-, tri- from left to right), while Figures 7b (WT) and 8b (R3765L) are the same values plotted by Microsoft Excel over the time period of 24 hours.

**Key:**
- **Unmodified (0)**
- **Monomethylation (1)**
- **Dimethylation (2)**
- **Trimethylation (3)**
Figure 8b
Once the loss of dimethylation was established when mutating R3765, analytical ultracentrifugation was applied to better analyze complex formation. This real time optical detection system using UV light absorption is able to distinguish molecules based on shape and molecular mass. In wild-type conditions, the MLL1 and WDR5 complex sediments around 2.80, while separately MLL1 is ~1.80 and WDR5 is ~2.37. Data obtained in July 2013 further supported the original hypothesis that the MLL1 R3765L mutation negatively effects core complex formation and, in doing so, greatly reduces, if not completely eliminates, dimethylation. Fig. 9 (M\textsubscript{WT}WDR5) and 10 (M\textsubscript{R3765L}WDR5) show the plot from SEDFIT (left) and values graphed on Microsoft Excel (right). A minor peak on the left exists for both experiments and was attributed to either protein degradation or contamination.

Figure 9
Hypothesis: If the R3746H mutation was found in patient screens, but was not correlated with KS. Then we therefore hypothesized that it would behave like wild-type MLL1.

To test this hypothesis a similar MALDI-TOF and AUC experimentation set up was used to test the similarity of wild-type and R3756H MLL1 proteins. Both of which supported the initial hypothesis that was posited.

Figures 11a (WT) and 12a (R3746H) are peaks from mMass (mono-, di-, trimethylation activity from left to right). Figures 11b (WT) and 12b (R3746H) are plotted values on Microsoft Excel. Though a small variance exists, the methylation trend is very similar.

**Key:**
- Unmodified (0)
- Monomethylation (1)
- Dimethylation (2)
- Trimethylation (3)
Sedimentation experiments agreed with the MALDI-TOF results in that complex was forming normally, despite this mutation. Figures 13 (WT) and 14 (R3746H) show MLL1-WDR5 complex formation, however experimental data from MLL1 R3746H was determined to have concentrations that were not stoichiometric. An excess of wild-type WDR5 led to the additional peak around 2.24.
Hypothesis: If the arginines at positions MLL1 3765 and 3771 f MLL1 are citrullinated by PAD2, then either the R3771 is exposing R3765 to PAD2 or there is no significance in this citrullination. Prior experimentation has already determined the importance of both the Win motif and MLL1 R3765 in MWRAD complex formation\(^7\).

Tests had been done before by Valerie Vought with PAD2 after MALDI-TOF results showed an arginine 6 residues down from the highly conserved Win motif arginine had been citrullinated in both MLL1. In addition, PAD2 displayed a high affinity to MLL1 in comparison to the other members of the SET1 family and PAD4.

To determine if PAD2 is a possible model for regulation of MLL1 core complex assembly, MLL1 R3771T was tested for disruption of methylation. Figure 15 lends evidence towards a similarity with wild-type MLL1 methylation behavior.
A time course glutathione pulldown assay was created to determine if PAD2
citrullination negatively impacted core complex formation. Fusion GST-MLL1
WT and R3771T protein were used to bind with the glutathione beads.
Optimization tests were conducted to hone down the original conditions. Figure
19 shows a dramatic decrease in band intensity for WDR5 (top circle) and DPY30
(bottom circle) over the time period of 90 minutes. The left side being MLL1 WT
and right side being MLL1 R3771T in complex with WRAD.
Repeat experiments showed similar results, which led to the goal of determining if a difference existed between citrullination with and without the arginine at position 3771 of MLL1. Due to the faint intensity of bands, analysis was unable to draw a conclusive answer.

Initial steps were taken to utilize western blotting towards offering a better method of quantification. Figure 20 shows a successful dot blot of an antibody that selectively binds with MWRAD that has been citrullinated by PAD2. Faint bands were seen with MWRAD and No CaCl$_2$, which was attributed to background binding. Controls of histone H3 (residues 1-20) + PAD2 showed slight antibody binding, while the absence of CaCl$_2$ showed no intensity, and the enzyme, PAD2, alone did not bind with the antibody.
Unfortunately western blotting results for MLL1 WT and MLL1 R3771T were deemed inconclusive. Only two experimental attempts were made during the academic school year. The first of which is seen below in Figure 20 (MLL1 WT, coomassie) and Figure 21 (MLL1 WT, western blot). The data did show that there was an inverse relationship between MLL1 core complex formation from the glutathione pulldown and citrulline antibody binding. This supported the hypothesis that PAD2 severely decreases complex formation through the citrullination of R3765, and R3771 in the process. Both Figure 21 and 22 have the same corresponding lanes. Lane 1 is the BioRad Standard Ladder, lane 2 is the input, lane 3 through 10 are time points from zero to 1 hour, lane 11 is the WRAD control, and lane 12 is the No CaCl₂ control. As less complex shows up over time on Figure 21, more citrulline antibody binds over time on Figure 22.
Chapter IV

Discussion and Conclusion

Positive Correlation of MLL1 Mutations and Dimethylation

MLL1 R3765 was determined to have a significant impact on MLL1 core complex formation and normal dimethylation\(^7\). Results obtained from MALDI-TOF and AUC were repeated by Stephen Shinsky, a graduate student in Dr. Cosgrove’s lab. Unlike other Kabuki Syndrome mutations in MLL2, this arginine located in the Win motif led to an almost complete loss of dimethylation. Hence, using a paralog of MLL2 R5340, the results supported the initial hypothesis that the arginine located at 3765 is crucial in the interaction of MLL1 with WDR5.

PeptidylArginine Deiminase 2

Initial results were not conclusive, but support the hypothesis that MLL1 R3771 may play a role in citrullination of MLL1 R3765. A flaw exists when the SET1 family peptides were sent for MALDI-TOF analysis of the 1Da difference of arginine and citrulline. MLL1 and MLL4, unlike MLL2, MLL3, SET1a, and SET1b included both 3 extra amino acids before and after the central residues. This could have aided in the stability of interaction with PAD2 and PAD4.
The arginine located at 3771 is not needed for the interaction of MLL1 with WDR5 and is supported by crystal structures that show movement of R3771 lending further evidence that this amino acid does not interact with the surface of WDR5.

Further research is needed to standardize the testing of SET1 family proteins and PAD2/4 either by sending equal length peptide sequences or by using truncated constructs in the GST pulldown assay. Shorter intervals of time may help in quantifying MLL1 core complex formation. Repeating western blot assays under optimized conditions will allow for more accurate quantification of PAD2 enzymatic activity over time and determine whether the R3771 plays a role in the speed of R3765 citrullination.

*Research Overview and Techniques learned from BCM460*

These past three years of research have been eye opening in teaching me both fundamental research techniques and developing a scientific mindset. Much of the biochemistry that I learned while in Dr. Cosgrove’s lab has already aided me in my undergraduate studies and has helped me obtain interviews for research positions after graduation.

Many thanks go out to the members, both past and present, of the lab. I could not have done this without your assistance. Though I wish more data could have been collected and better experimental techniques could have been used, I believe BCM460 has given me invaluable skills that will continue to be utilized in the future.
References