Inhibiting Multiple Myeloma SET protein, a Histone Methyltransferase, via High Throughput Screens

Honors Thesis

March 2011

Thomas J Hartley

Department of Pathology
The University of Michigan

Contents

Acknowledgements

I want to thank graduate student Matt Pauly for beginning the MMSET project and providing us with the tools to succeed.

I would like to thank Marcelo Murai for over the past year teaching me everything I would need to know about protein expression, purification and storage.

I would also like to thank the other lab members, George Lund, Shihan He, Trupta Purohit, Chris Jones, Gireesh Reddy, and Aibin Shi for their reliability and expertise within their own respective areas.

Special thanks also goes to the Yali Dou Lab for generously giving us access their liquid scintillation counter, and Elizabeth Townsend for her assistance with this assay.

I would most like to thank Dr. Jolanta Grembecka and Dr. Tomasz Cierpicki.
Dr. Grembecka has taught me everything about laboratory procedures from how to properly aspirate a pipette to how to run fluorescence polarization, isothermal titration calorimetry, and time-resolved fluorescence resonance energy transfer assays. She is also the one who made me realize the importance of our work and how it fits into the bigger picture. Dr. Cierpicki taught me about the usefulness of analyzinging protein structure, and how to obtain this information with x-ray crystallography or nuclear magnetic resonance spectroscopy. He also made me excited about making discoveries, and taught me that a great scientist must be creative, and makes use of all resources available to him.

Without these individuals, this research and my thesis would not have been possible.

Figures

Tables

Abstract

The overexpression of MMSET, due to a chromosomal translocation, has been observed in numerous cancers including brain, epithelial, lymphoid and most commonly, in multiple myeloma. As many as 20% of newly diagnosed cases of multiple myeloma are due to this type of translocation. MMSET functions by methylating lysine residue 36 on the C-terminal tail of histone H3, which alters the regulation of local genes that are believed to be involved in tumor emergence and progression. Our research focuses on developing small molecule inhibitors of MMSET via high throughput screens. Our results indicate that AlphaLISA is the most suitable assay for our objective. Unfortunately our current constructs of MMSET are not active enough to be detected by this method. Future plans of research aim to use site-directed mutagenesis to improve the activity of MMSET *in vitro*.

Introduction

Importance of Research

Multiple myeloma (MM) is a universally fatal malignancy of mature plasma cells, the antibody-producing leukocyte of the human body [8]. The disease is characterized by the accumulation of abnormal plasma cells at multiple sites in bone marrow tissue [4]. Symptoms of this disease vary greatly because they are dependent on what area of the body is affected, which makes its diagnosis a difficult task [9]. Ordinary complications include renal failure, bone problems, anemia, and various infections due to the fact that MM comprises the immune system [9]. In 2007, over 60,000 people in the United States had a history of MM, and 20,180 people were estimated to be diagnosed with MM in 2010 [1]. Multiple myeloma remains an incurable disease with a median life expectancy of approximately four years [13].

Physiology of the SET Domain

Nucleosomes, the DNA-packaging unit in eukaryotes, are composed of eight proteins called histones [15]. Post-translational, covalent modifications of the NH_2-terminal tails of these histones change their binding characteristics, and thus alter chromatin structure [5, 11]. Modified histones will recruit or dispel protein complexes, and in doing so, regulate gene expression [6, 15]. Methylation of lysine residues is an example of these modifications and is facilitated by SET domain-containing proteins [5]. The ~130 amino acid, catalytic SET domain is highly conserved in nature [11, 15], which suggests that it plays a crucial role in normal cellular functions.

ROLE OF MMSET IN MULTIPLE MYELOMA

The multiple myeloma SET domain-containing protein (MMSET) is named as such because of its SET domain that is characteristic to histone methyltransferases, and because its overexpression is associated with multiple myeloma due to a chromosomal translocation [6]. Recent studies show that 15-20 % of newly diagnosed cases of MM are be caused by a t(4;14)(p16; q32) chromosomal translocation [13], which is the second most common translocation associated with MM and has the worst prognosis [6]. While the t(4;14) translocation dysregualates both the MMSET gene and the fibroblast growth factor receptor 3 (FGFR3) gene [4], overexpression of MMSET is universal to MM tumors with this type of translocation whereas FGFR3 is not overexpressed in one-third of these cases [6]. This indicates that the overexpression of MMSET is the underlying factor that causes MM in this type of translocation. Furthermore, MMSET overexpression has been shown to promote cellular adhesion, clonogenic growth and tumorigenicity in MM [6]. siRNA-mediated knockdown of MSMET in MM cell lines have demonstrated that this protein is advantageous to the survival of MM cells [13]. Complete knockout of MMSET affected expression of genes involved in key survival processes, such as cell cycle, apoptosis and adhesion [13].

MMSET is believed to function by methylating lysine residue 27 and 36 on the terminal tail of histone H3 [4, 15]; thus an overexpression of this protein would cause hyper-methylation of histone H3, causing chromatin to be in a more open structural state [8]. Of MM patients that lack the t(4;14) translocation, 10% of these individuals have null mutations that inactivate the production of UTX, a histone demethylase [13]. Whether by histone methylation via MMSET or by the lack of demethylation due to a

deficiency of KDM6A, it seems highly plausible that an increased methylation of histone tails contributes to the emergence and progression of MM.

VARIOUS OTHER DISEASES

MMSET expression has also been shown to vary directly with tumor proliferation activity in glioblastoma multiforme [6], the most common and malignant primary brain tumor [14]. In fact, blocking MMSET expression via RNA interference actually suppresses the growth of glioma cells [6]. Additionally, comparative studies of other tumor types and their normal tissue counterparts have shown overexpression of MMSET in 15 cancers, including brain, epithelial, lymphoid and leukemia [5, 6].

MMSET is also known as Wolf-Hirschhorn syndrome candidate 1 protein (WHSC1) due to its malfunction being the causative agent in this neurological disorder. WHS is most often caused by a partial deletion of a region on the short arm of chromosome 4 (4p16), and is characterized by severe mental retardation [9]. Normally-functioning MMSET plays a restorative role within the genome, recruiting proteins like p53-binding protein 1 to repair breaks that occur in DNA, whereas malfunctioning MMSET prevents DNA from undergoing proper repair [9].

These data demonstrate that MMSET is likely involved in tumor emergence and progression, and that an understanding of this protein and its mechanisms may lead to promising drug discoveries. Our research is aimed to characterize the interaction between MMSET and histone H3, and develop small molecule inhibitors of its methylation function by implementing high throughput techniques.

Protein Purification

INTRODUCTION

The MMSET gene encodes three isoforms as a result of alternative splicing. Our research focuses on the SET and PWWP domains of this protein, which are conserved in two of these isoforms. Because full length MMSET would be difficult to manage and is unnecessary for our objective, various constructs were made that contain different elements of the entire protein. The protein used for the majority of our experiments is made up of amino acids 952 – 1216 (265aa) of full length MMSET, contains an N-terminus alpha helix and the SET domain; this will be hence forth referred to as MMSET-helix. Another construct we became interested in is made up of amino acids 877 – 1216 (340aa) of full length MMSET, contains an N-terminal PWWP domain, alpha helix and SET domain; this will be hence forth referred to as MMSET-pwwp. All protein sequences are listed in Appendix. Expression vectors encoding NSD1 and NSD3, homologs to MMSET, were synthesized because they also contain the alpha helix and SET domain, and are believed to have similar structures.

METHODS

cDNAs encoding the MMSET-helix, MMSET-pwwp, NSD1 and NSD3 were ordered from GenScript. Each target gene was cloned into pET-Mocr expression vector. The pET-Mocr expression vector is pET21 that has been modified with an N-terminal monomeric Ocr (Mocr)-His6 tag. Unlabeled proteins were obtained from *Escherichia coli* cells cultured in LB medium, while ^{15}N-labeled protein were cultured in minimal

medium supplemented with ^{15}N ammonium sulfate (Cambridge Isotope Labs). MMSET-helix and MMSET-pwwp, unlabeled and ^{15}N-labeled, were expressed in RosettaTM (DE3) cells (EMD) and were grown to an OD_{600} of 0.600, at which point they were induced at 20°C for 6 hours with 0.5mM IPTG. Maximum yield of protein during cell lysis was obtained with the French press. To keep our target proteins in the soluble fraction of the cell lysate, specific extraction buffers were used for each protein. MMSET-helix is located in the soluble fraction of the cell lysate when using an extraction buffer of 50mM Tris buffer (pH 8.5), 250mM NaCl, and 1mM TCEP. MMSET-pwwp requires an extraction buffer of 50mM Tris buffer (pH 9.0), 500mM NaCl, 1mM TCEP, and 5% glycerol. The proteins were purified using affinity chromatography employing nickel-agarose (GE Healthcare) (Figure 4). To remove the Mocr fusion tag, the proteins were cleaved by tobacco etch virus (TEV) protease. Completion of cleavage was identified by viewing the cleavage products via SDS-PAGE using 4% to 20% Ready Gels (Bio-Rad).

The proteins were then dialyzed to 50mM Tris buffer (pH 9.0), 50mM NaCl and 1mM TCEP and then applied to a Q sepharose column (GE Healthcare) for ion exchange chromatography (Figure 5). The elution buffer used for both proteins was 50mM Tris buffer (pH 9.0), 1M NaCl and 1mM TCEP. Protein intended for crystallization was further purified by Sephacryl S-100 size exclusion chromatography. As a final step, proteins were again dialyzed to 50mM Tris buffer (pH 9.0), 50mM NaCl, 1mM TCEP and frozen at -80 °C for further experiments. Protein concentration was determined by absorbance measurements at 280nm. The extinction coefficients at 280nm for MMSET-helix and MMSET-pwwp are of 22,920 M^{-1} cm^{-1} and 52,370 M^{-1} cm^{-1}, respectively [3].

Although expression vectors containing a Mocr-His6 tag were synthesized for NSD1 and NSD3, these proteins have yet to be expressed.

RESULTS

MMSET-helix is located in the soluble fraction of the cell lysate when using an extraction buffer of 50mM Tris buffer (pH 8.5), 250mM NaCl, 1mM TCEP. However, the extended MMSET-pwwp protein was not stable using the same extraction buffer as MMSET-helix, which caused aggregation, degradation and the protein to not be able to be cut with TEV protease (Figure 1). We tested several buffers for extraction of MMSET-pwwp, varying the concentrations of NaCl, Triton X-100 and glycerol, to determine which yielded the most of our target protein in the soluble fraction. The concentration of total protein in each extraction buffer was measured using a NanoDrop Spectrophotometer (Thermo Scientific) (Table 1), and samples were run on an SDS-PAGE gel to determine the location of our target protein (Figure 2), whether it was in a the soluble or the insoluble fraction. Although the buffers containing Triton X-100 yielded more total protein in the soluble fraction, the SDS-PAGE gel revealed that our target protein was at a higher concentration in the buffers containing glycerol. The optimal extraction buffer was 50mM Tris buffer (pH 9.0), 500mM NaCl, 1mM TCEP, and 5% glycerol. This buffer provided cleaner protein after being passed on the affinity column (Figure 3) and a successful removal of the MOCR tag via TEV protease.

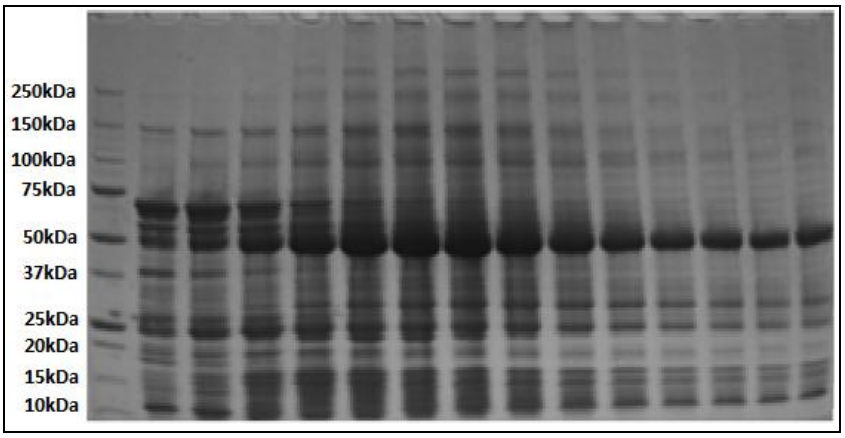

Figure 1. MMSET-pwwp using initial extraction buffer after affinity column. Target protein ~50 kDa due to additional Mocr fusion tag.

Extraction Buffer	Fraction	Concentration [mg/mL]
A	Soluble	25.3
	Insoluble	23.4
B	Soluble	20.0
	Insoluble	15.4
C	Soluble	26.7
	Insoluble	13.0
D	Soluble	28.7
	Insoluble	13.1
E	Soluble	23.2
	Insoluble	16.1
F	Soluble	26.7
	Insoluble	17.9

Table 1. The concentration of total protein in each extraction buffer measured using NanoDrop Spectrophotometer.

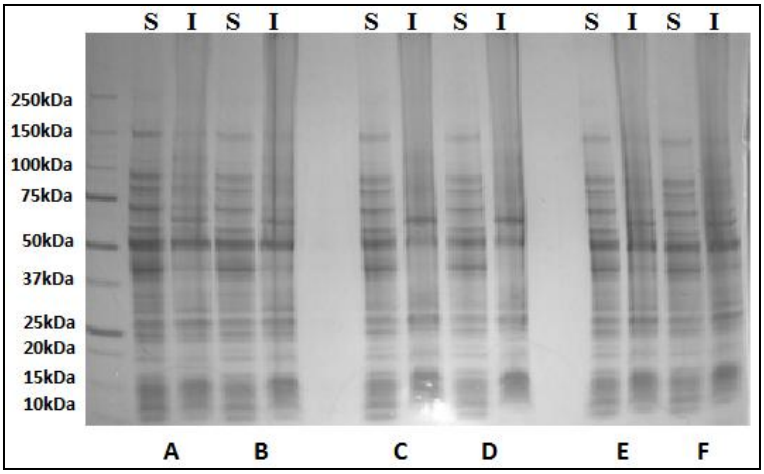

Figure 2. An SDS-PAGE gel showing the soluble (S) and insoluble (I) fractions of MMSET-pwwp after being extracted using buffers with various conditions.

For Table 1 and Figure 2, the standard extraction buffer is 50mM Tris buffer (pH 9.0), 250mM NaCl, and 1mM TCEP. Experimental buffer conditions: A) 300mM NaCl; B) 500mM NaCl; C) 300mM NaCl and 0.1% Triton X-100; D) 500mM NaCl and 0.1% Triton X-100; E) 300mM NaCl and 5% glycerol; F) 500mM NaCl and 5% glycerol.

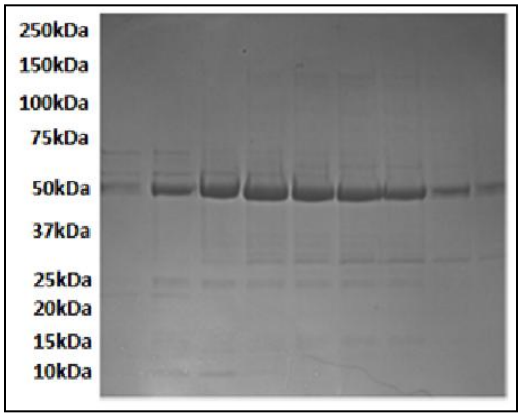

Figure 3. MMSET-pwwp after affinity column using optimized extraction buffer (50mM Tris buffer (pH 9.0), 500mM NaCl, 1mM TCEP, and 5% glycerol).

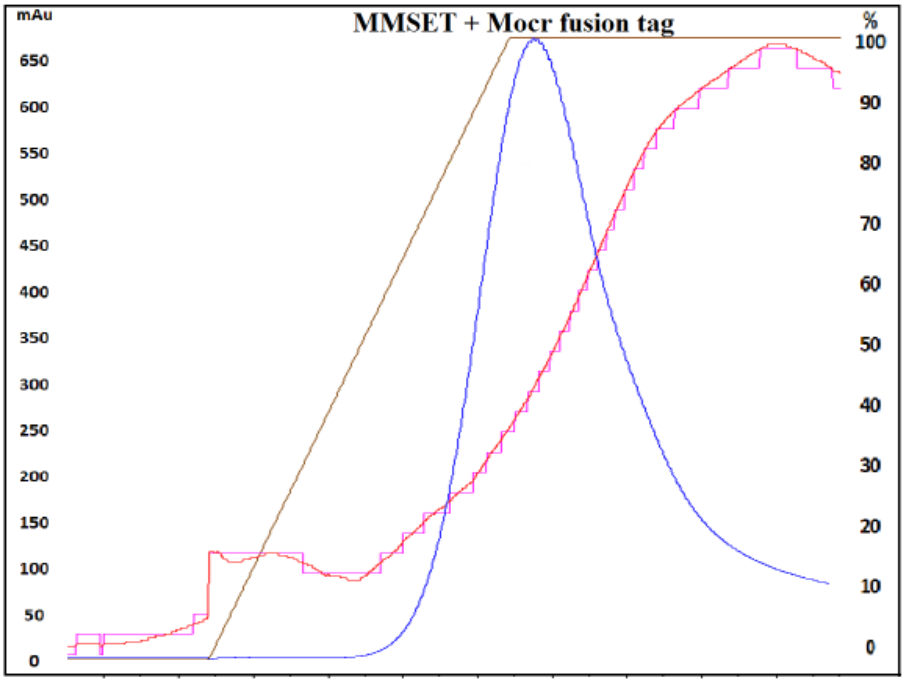

Figure 4. Elution of MMSET on NiNTA column. The blue line is absorbance (mAu). The brown line is percentage of elution buffer (50mM Tris buffer (pH 8.5), 250mM NaCl, 1mM TCEP, and 200mM imidazole).

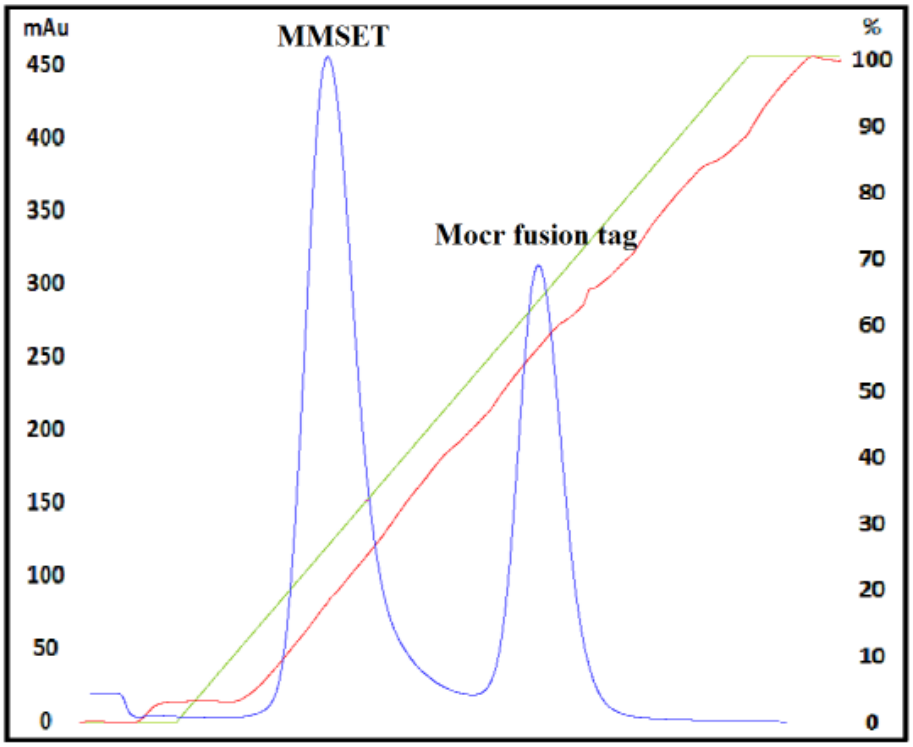

Figure 5. Elution of MMSET and Mocr fusion tag on Q Sepharose column. The blue line is absorbance (mAu). The green line is percentage of elution buffer (50mM Tris buffer (pH 9.0), 1M NaCl, and 1mM TCEP).

Immunofluorescence

INTRODUCTION

Our purpose of immunoblotting was to demonstrate the specific methylation activity of MMSET at lysine 36 (K36) on histone H3, and to optimize conditions for its enzymatic function to be used in our high throughput screen.

METHODS

5µM of MMSET-helix was incubated with 50ng/µL histone H3 and 100µM of S-adenosyl-methionine (SAM) in 50mM Tris buffer (pH 8.5), 5mM $MgCl_2$, and 0.2mM DTT for 1 hour at 30°C. The reaction was stopped with the addition of Laemmli buffer. The reaction products were resolved by SDS-PAGE using a 4% to 20% Ready Gel (Bio-Rad), and transferred to a PVDF membrane (Invitrogen). The membrane was blocked for 1 hour with 5% BSA in TBST, then probed with primary antibody, monoclonal Di-Methyl-Histone H3(Lsy36) Rabbit mAb (Cell Signaling Technology), overnight at 4°C. The following day, the membrane was washed three times for 10 minutes with TBST. Secondary anti-rabbit antibodies were then added to the membrane and allowed to incubate for 1 hour at room temperature. The membrane was washed again three times for 10 minutes with TBST. Chemiluminescent reagent (Bio-Rad) was added to the membrane and examined by western blot analysis using autoradiography film (Denville Scientific).

RESULTS

The films from our western blot were conclusive to indicate that histone methylation increases with increasing concentrations of MMSET-helix (figure 6). Unfortunately the primary antibody was not as accurate as advertised and we observed nonspecific binding. In Figure 7, we see the antibody binding to histone H3 even in the absence of MMSET. Even with this issue, the assay was still able to indicate that MMSET is more active at a higher temperature (in this case 30°C).

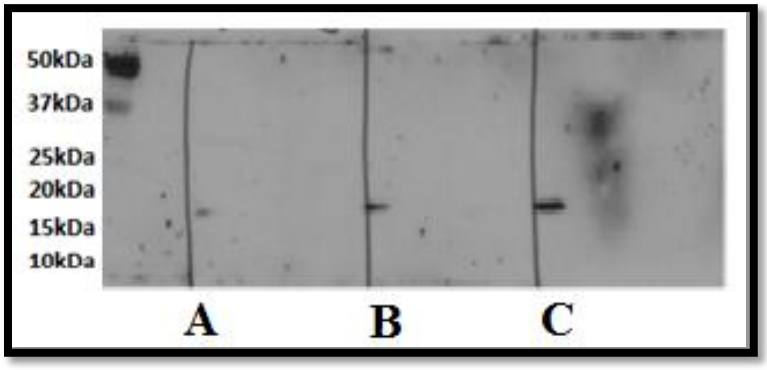

Figure 6. Western blot detecting dimethylated histone H3 with MMSET-helix at 0µM (A), 1µM (B) and 5µM (C). Signal intensity increases with increasing enzyme concentration as we would expect. Histone H3 ~15kDa.

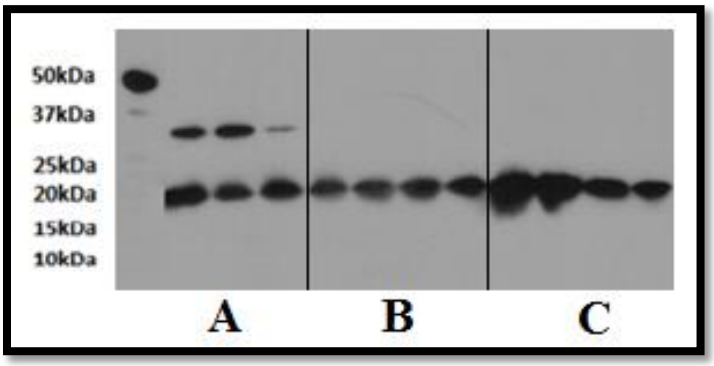

Figure 7. Western blot detecting dimethylated histone H3 with MMSET-helix.
A) 0µM MMSET-helix, incubated at room temperature; B) 5µM MMSET-helix, incubated at room temperature; C) 5µM MMSET-helix, incubated at 30°C.
Column A shows signal in the absence of enzyme which indicates non-specific binding between recombinant histone H3 and our primary antibody. A more intense signal in column C leads us to believe MMSET-helix is more active at a higher temperature.

DISCUSSION

This immunoblotting assay was unique from our other methyltransferase assays in that the substrate was recombinant histone H3 instead of the shorter H3 peptide. Using recombinant histone H3 as a substrate made this reaction more representative of the reaction that occurs *in vivo*, although our histone H3 is in an unfolded state. Since the interaction between MMSET and histone H3 has yet to be characterized, it is possible that the PWWP domain of full length MMSET assists in its enzymatic activity by coordinating with the histone peptide near lysine 36. Once a reliable antibody is obtained, further experiments can be done to test this hypothesis with histone H3 or entire nucleosomes. Since this method was not reliable for our purposes, we resorted to other means of testing the enzymatic activity of MMSET.

Liquid Scintillation Spectrometry

INTRODUCTION

We used liquid scintillation spectrometry, a radioactive technique, to optimize conditions for the methyltransferase activity of MMSET-helix *in vitro*. Radioactive isotopes are unstable due to being composed of too few or too many neutrons and are known to spontaneously rearrange their nuclei, which causes an emission of particles or electromagnetic radiation. This aspect of radioactive isotopes allows for them to be detectable but still react in the same chemical manner as stable isotopes. The radioactive emissions can be measured indirectly by exposing radioactive-labeled material to a photographic emulsion, such as x-ray film. This method can be used to determine the location of a radioactive substance within a tissue. However, autoradiography is difficult to quantify. To accurately measure the intensity of radiation emitted, a liquid scintillation counter can be used to measure the emissions directly. By mixing a radioactive component with a scintillation cocktail, radioactive decay will cause the fluid to fluoresce. The scintillation cocktail contains two components: an aromatic organic solvent, such as toluene, which absorbs the energy from radioactive decay; and molecules of phosphors, which convert the absorbed energy into light. A scintillation counter quantifies the amount of radioactivity by counting the number of flashes given off. Typically this data is expressed in average counts per minute (CPMA) [2, 10].

METHODS

We used tritium-labeled methyl groups on S-adenosyl-methionine (^{3}H-SAM) for the radioactive component because SAM is a cofactor of MMSET and acts as a methyl

group donor. The high sensitivity of the scintillation method was useful because tritium has a relatively low energy of emission. We carried out the reactions of MMSET-helix with H3 peptide 21-44 (AnaSpec) and ^{3}H-SAM at various conditions. 20µL of the each reaction was then spotted onto a gridded P81 filter paper (Whatman) and put to a heater until dry. The filter paper was then washed three times for 10 minutes with freshly-made 50mM $NaHCO_3$ (pH 9.0) and again put to dry. Lastly, each reaction on the filter paper was cut into separate squares, added to plastic translucent vials, vortexed with 10mL of Ultima Gold scintillation cocktail (Perkin Elmer) and analyzed using a liquid scintillation counter. The samples were measured for five minutes each. Our negative controls included a sample without MMSET-helix, a sample without H3 peptide and a sample without any reaction spotted onto filter paper.

RESULTS

Whatman Grade P81 filter paper is composed of cellulose phosphate and is strong cation exchanger that will bind medium sized proteins. Whether a given protein will bind depends on the isoelectric point (pI) of the protein and the buffer in which it is solubilized. The proteins in this reaction are MMSET-helix and H3 peptide 21-44 (AnaSpec) which have theoretical pI values of 6.67 and 11.25, respectively. All of our reactions were carried out with buffer at pH 8.5 - 11.5. These conditions allow for the methylated H3 peptide to bind to the filter paper, and MMSET-helix and ^{3}H-SAM to be removed during the washing step. MMSET-helix did not bind to the paper because it is an anion at the pH we used, and ^{3}H-SAM did not bind because it is not a protein. Using

this same procedure, we could also test the activity of MMSET-pwwp (theoretical pI of 7.86) with either H3 peptide or recombinant histone H3 (theoretical pI of 11.13).

The optimal conditions for this reaction were 5µM MMSET-helix, 500µM H3 peptide 21-44 (AnaSpec) and 0.5µL of ^{3}H-SAM at 0.55µCi/µL (Perkin Elmer) in 20mM glycine buffer (pH 11), 10% glycerol, 1mM TCEP in a volume of 40µL and incubated for 1 hour at 37°C. These conditions provided a CPMA value that was 85-fold greater than a reaction that lacked the enzyme, and 35-fold greater than a reaction that lacked the substrate. Longer incubation time did give greater results, but after 1 hour of incubation the reaction was at ~80% its maximum value. Samples carried out under the same conditions differed in CPMA by 11%, and samples measured from the same reaction differed in CPMA by 3%.

DISCUSSION

Liquid scintillation spectrometry has proven to be a reliable method to measure the methylation activity of MMSET-helix. This method was useful to our research because it allowed us to use the same substrate (H3 peptide 21-44) that we will use in our high throughput screen, which means that we could implement the same conditions for optimal activity of the enzyme. However, this assay cannot determine level of methylation per substrate. We do not know if our enzyme is attaching one, two or three methyl groups to the substrate. We only know that the transfer of methyl groups from SAM to H3 peptide is being catalyzed by MMSET-helix. In the future, this method will be used to test activity of other MMSET constructs, and to test the inhibition activity of compounds on other histone methyltransferases to ensure specificity of our compounds.

AlphaLISA

INTRODUCTION

Amplified Luminescent Proximity Homogeneous Assay Screen (AlphaScreen) is another highly sensitive assay that uses beads to detect molecular interactions, such as protein:protein and protein:small molecule interactions. This assay, which is a development from Luminescent Oxygen Channeling Assay technology, uses singlet oxygen molecules as a means for short-distance energy transfer. Donor beads contain a photosensitizing agent that excites ambient oxygen to a singlet state when irradiated at 680nm. When the singlet oxygen reaches the Acceptor bead, a cascade of internal energy transfer results in the emission of light at a specified wavelength. This specific wavelength can then be detected by a microplate reader. A chemiluminescent signal will only be generated if the beads are within the diffusional range of the singlet oxygen (approximately 200nm). Close proximity of the beads occurs when they are bound, via an antigen-antibody interaction, to the same substrate or conjugated substrates. Donor beads are coated with streptavidin that binds to biotinylated proteins. Acceptor beads are coated with antibodies specific for each experiment [12].

AlphaLISA is a newer development that is named as such due to its use of AlphaScreen detection mechanisms and its use of an antibody pair to detect a substrate, a technique characteristic to the classic Enzyme-Linked Immunosorbent Assay (ELISA). The Acceptor beads for AlphaScreen emit light at a wavelength between 520nm and 620nm, whereas the Acceptor beads in the more advanced AlphaLISA screen emit light with greater intensity and at a well-defined wavelength of 615nm. By emitting a defined wavelength of light, the AlphaLISA screen has better accuracy because it is less

susceptible to interferences. AlphaLISA has a simple protocol, does not require a wash step to remove non-specifically adsorbed reactants, and is analyzed on a convenient microtiter plate. We decided on AlphaLISA because of its high sensitivity and suitability for high throughput screening.

METHODS

We used MMSET-helix for our enzyme and biotinylated H3 peptide 21-44 (AnaSpec) as a substrate. Donor beads (Perkin Elmer) were coated in streptavidin which would bind to the biotin on our substrate, and Acceptor beads (Perkin Elmer) were coated with an antibody that was specific to the dimethylated lysine 36 of H3 peptide (H3 peptide K36me2). We incubated 5µM MMSET-helix, 5µM biotinylated H3 peptide 21-44 and 100µM SAM in 20mM glycine buffer (pH 11), 10% glycerol, and 1mM TCEP at 37°C for 1 hour. Incubation was carried out inside 1.5mL tubes to prevent evaporation. After incubation, pH was adjusted to 8.0, and 10µL of the reaction was transferred to a 384-well OptiPlate (Perkin Elmer). 5µL of Acceptor beads were added and incubated at room temperature for 1 hour. 10µL of Donor beads were then added and incubated at room temperature for 30 minutes. Final volume was 25µL, and each bead was at a final concentration of 20µg/mL. The plate was read using a Pherastar Plus FS microplate reader (BMG) with an AlphaLISA module.

RESULTS

Positive controls for this assay (H3 peptide K36me2) provided an Alpha Signal that was over 100-fold greater than the negative control (non-methylated H3 peptide or H3 peptide K36me3) (Figure 8). Although we do not expect our enzyme to provide an Alpha Signal as great as the positive control, we have yet to even reach a signal that is more than 2-fold greater than the negative control.

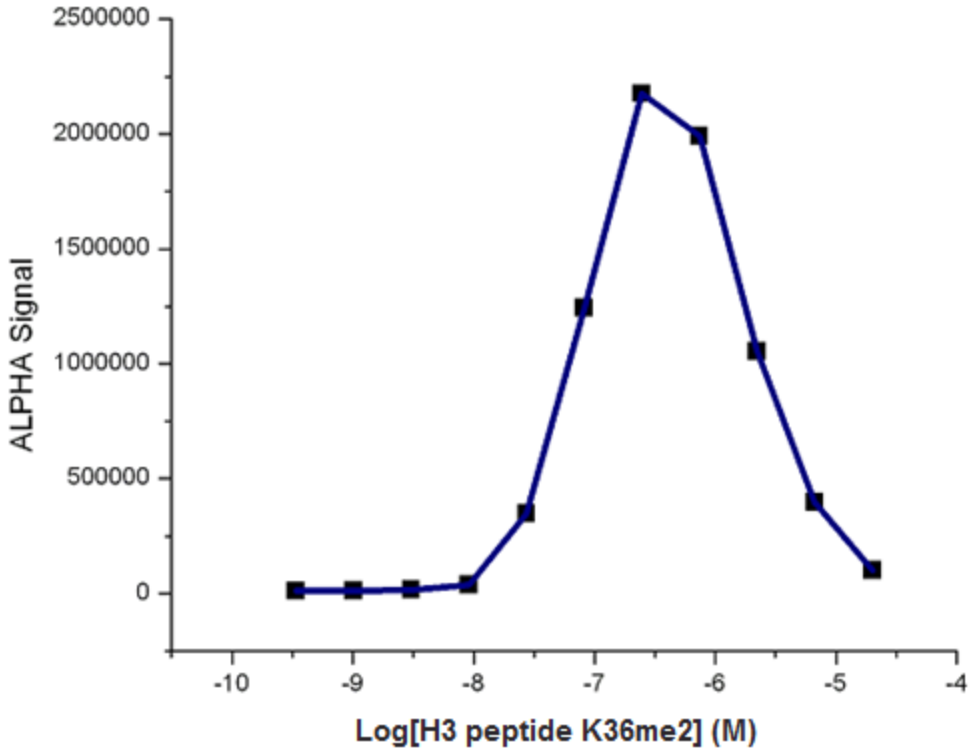

Figure 8. This graph displays the Alpha Signal emitted from a serial dilution of H3 peptide K36me2. This indicates that for maximum signal output (~2,200,000), dimethylated substrate should be at a concentration of 0.3µM. In this same assay, non-methylated H3 peptide and H3 peptide K36me3 gave signals at the background level (~20,000).

DISCUSSION

The controls for AlphaLISA reveal that it gives a strong signal and would be an excellent candidate for high throughput screens. The current issue is making our enzyme active enough to be detectable with this method. Recent results have shown a decrease in signal when the positive control (H3 peptide K36me2) is incubated with our enzyme. This could mean that our enzyme is trimethylating the already dimethylated substrate. Trimethylation has been shown to occur in NSD1, a close homolog to our enzyme [11]. If this is the case, it would cause decrease in signal because our Acceptor beads do not bind trimethylated lysine residues. This would not have affected our results from liquid scintillation spectrometry because that assay did not give evidence to the methylation state of individual substrates; it only showed the total methylation present in the sample. It is also a possibility that it our enzyme is methylating a different amino acid residue on our substrate, such as H3K27 [5]. This would show no signal with AlphaLISA because our Acceptor beads are specific to lysine 36. Again, this situation would not have affected the signal from liquid scintillation spectrometry.

Analyzing Protein Structure

INTRODUCTION

Our objective was to gain insight about the function of MMSET by studying its structure using x-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy. MMSET-helix had consistently been a weak enzyme *in vitro*, and solving its structure could assist in finding the reason for this. X-ray crystallography data would also assist in modifying inhibitory compounds to improve their affinity. NMR would allow us to analyze binding characteristics by examining conformation changes due to labeled enzyme binding its substrate. Overall, any information about the normal function of MMSET would help to create ways to disrupt this process.

METHODS

Crystallization

We prepared standard crystallization plates of MMSET-helix using the Index (Hampton Research) and Wizard I & II (Emerald BioSystems) screens on 2-drop chamber, 96-well crystallization plates (Hampton Research) at 4°C and 18°C with various protein concentrations. See Table 2 below for all conditions tested.

Experimental Conditions for the Crystallization of MMSET-helix

Date Began	Temperature	Concentrations of MMSET-helix	Screen Type
17-Nov-10	18°C	5mg/mL, 10mg/mL	Hampton Index
16-Dec-10	4°C	2mg/mL, 5mg/mL	Hampton Index
20-Jan-11	4°C	4mg/mL, 8mg/mL	BioSystems Emerald Wizard I&II
25-Jan-11	4°C	10mg/mL, 12mg/mL	BioSystems Emerald Wizard I&II
26-Jan-11	4°C	10mg/mL, 15mg/mL	Hampton Index
24-Feb-11	18°C	8.5mg/mL, 12.5mg/mL	Hampton Index
25-Feb-11	4°C	8.5mg/mL, 12.5mg/mL	Hampton Index

Table 2. This chart shows all of the conditions used in effort to crystallize MMSET-helix.

Nuclear Magnetic Resonance

To examine interaction of MMSET and histone H3, we incubated ^{15}N-MMSET-helix at ~120μM with H3 peptide 21-44 at concentrations of 0μM, 300μM and 600μM and measured ^{1}H-^{15}N TROSY-HSQC spectra. Protein samples were diluted with 50mM Tris buffer (pH 7.5), 100mM NaCl and 1mM TCEP to a final volume of 140μL. Samples were then supplemented with 10μL of deuterium oxide, briefly centrifuged and transferred to 3mL NMR tubes. The experiments were recorded at 30°C employing Bruker Avance III 600-MHz spectrometer equipped with cryogenic probe.

RESULTS

We did not obtain crystals during our crystallization screens. Our NMR results revealed that in the binding between MMSET-helix and H3 peptide is very weak, due to hardly any shift being observed on the spectra (data not shown). However, the spectrum of ^{15}N-MMSET-helix alone did exhibit very strong signals in the center of the spectrum, which indicate an unstructured region of amino acids on the C-terminal tail (Figure 9).

DISCUSSION

The evidence of an unstructured C-terminal tail is consistent with the theoretical structure, which indicates that the last 12 amino acid residues are disordered [7]. This is likely the reason for the protein failing to crystallize. We plan to introduce an early stop codon mutation that will eliminate the final 12 residues of the C-terminal tail. From the sequence alone, we know that an alpha helix lies between the SET and PWWP domains.

It has been hypothesized that this helix is shared by both domains and may assist in stabilizing the SET domain. Further crystallizations should be performed using a construct that contains the PWWP domain.

Since NSD1 and NSD3 are known homologs of MMSET, knowing their structure would be helpful in predicting the structure of MMSET. Just this month, a paper was published indicating that scientists had successfully crystallized NSD1. Upon examining their construct, we noticed that they did not include many residues past the post-SET domain. In fact, they end their protein at the same location that we were planning to insert an early stop codon. This gives further evidence to our hypothesis and will likely lead to successful crystallization.

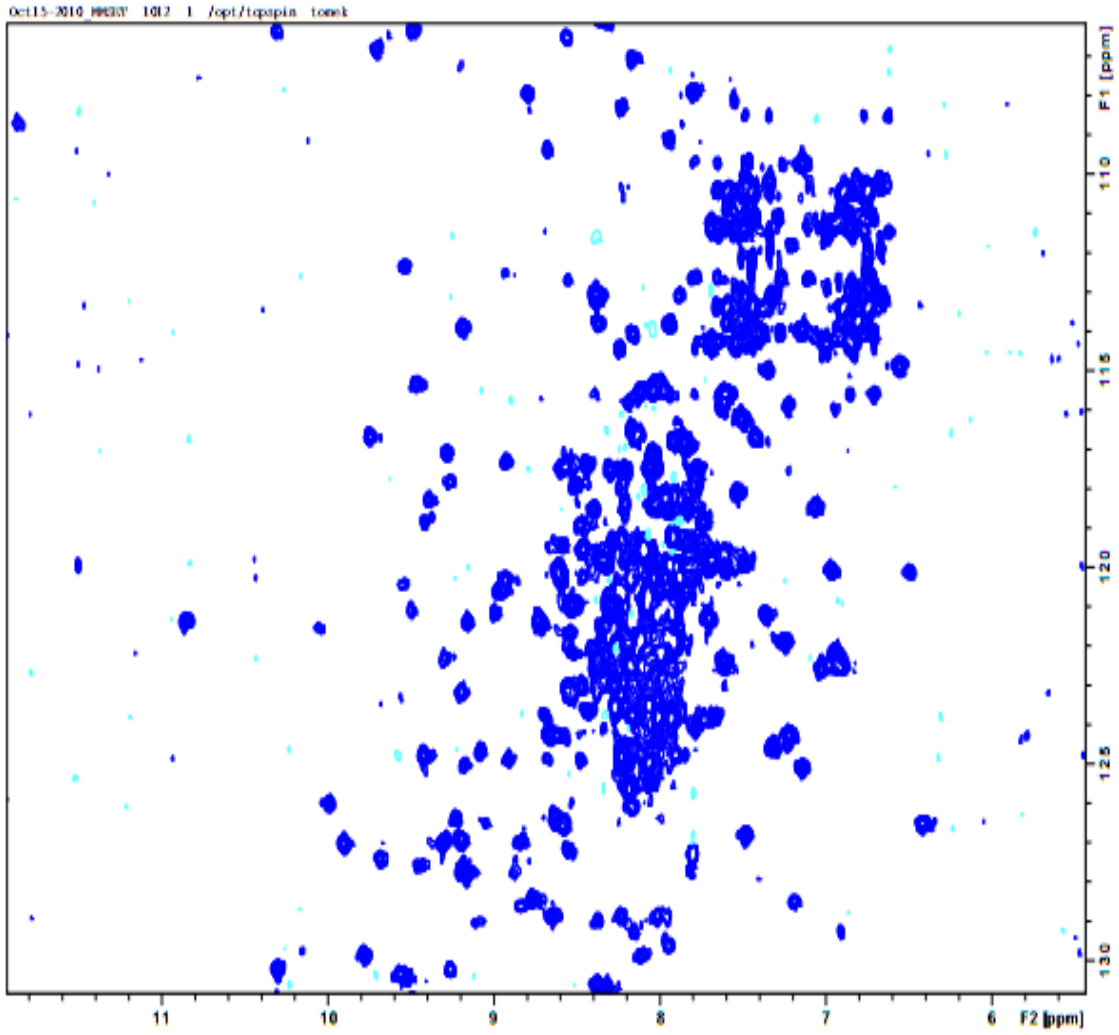

Figure 9. NMR TROSY-HSQC spectrum of ^{15}N-MMSET-helix. Strong signals in the center of the spectrum indicate a region of unstructured amino acids on the C-terminal tail.

Conclusions and Directions for Future Research

Since developing small molecule inhibitors is our objective, we must first improve our AlphaLISA to prepare it for high throughput screening. There are several ways to do this.

One approach we will take is to improve the potency of MMSET. As previously stated, the PWWP domain is in close association with the SET domain, and these domains may even share an alpha helix. Given this fact, we hypothesize that the PWWP domain will improve the enzyme's activity by binding to the peptide near lysine 36.

We could also improve the activity of MMSET by mutating the amino acid residues of the post-SET loop that is believed to inhibit the activity of this enzyme [11] (Figure 12). Data from the crystallization of NSD1 indicate that these homologs contain a post-SET loop that naturally undergoes conformational changes, which opens or closes the lysine binding channel for methylation [11]. This discovery is consistent with the theoretical model of MMSET-helix that shows conformational changes of the post-SET loop, indicated by a gap in the amino acid chain (Figure 11). This newly developed theory states that nucleosome contact stabilizes the loop in a conformation that is non-inhibitory. By disabling the autoregulatory activity of the post-SET loop, we may be able to drastically increase the activity of the enzyme by exposing the active site (Figure 13). We have two proposals to test this hypothesis: we will change the vital residues involved in the loop to glycine and alanine residues, and we will create another construct that deletes the loop entirely (Figure 10).

Whether by including the PWWP domain or by disabling the post-SET loop, development of a small molecule inhibitor will progress much faster with a more active enzyme.

We will test the activity of these new constructs with liquid scintillation spectrometry because that has been the most accurate for our research. Once a more active enzyme is obtained, we can move forward with the optimization of AlphaLISA. The only remaining condition that would likely need optimization is the time of incubation. Ideally for this assay, the enzyme must progress enough to dimethylate H3 peptide to the level of 0.3μM in the sample, as indicated in Figure 8. Then, the activity of this enzyme will be screened with a 50,000 compound library.

Once several inhibitory compounds are identified, we can work to increase the potency of these compounds. Information regarding the enzyme:substrate interaction would help during this process, and may be obtained through NMR experiments. If MMSET-pwwp does have greater activity than MMSET-helix, is likely due to the fact that MMSET-pwwp has a greater binding affinity to the substrate, which would result in a greater shift on NMR spectra. We would then be able to identify which amino acids are involved in the enzyme:substrate interaction by running NMR experiments with ^{13}C, ^{15}N-MMSET-pwwp. Characterizing this interaction will give us a better understanding of how to inhibit the enzyme.

By crystallizing our enzyme modified with a shortened C-terminal tail, it will open up the option of crystallizing our enzyme while in solution with compounds. We can then identify the location of its binding and predict ways to improve its binding affinity. Lastly, it will be necessary to ensure the specificity of our inhibitory compound

by testing its effects with similar histone methyltransferases. Non-specific inhibitory effects of methyltransferases would be toxic to cells. NSD1 and NSD3 will be perfect candidates for this test.

The small molecule inhibitors that we can identify will be most useful in development of drug therapies for individuals suffering from the associated diseases. While this project is far from being complete, it is headed in the right direction and appears to be on the precipice of making rapid progress. These next few months of research will be very exciting and may lead to promising breakthroughs with numerous implications.

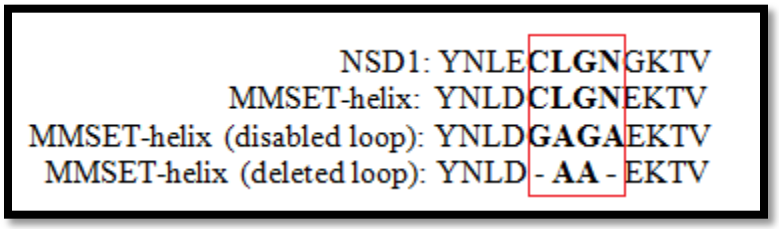

Figure 10. Amino acid sequences comparing NSD1, MMSET-helix and two future constructs of MMSET-helix. The vital four vital residues in the post-SET loop are conserved in both proteins, indicating that they behave in a similar fashion. Removing or disabling this loop should increase the activity of our enzyme.

All models were generated using PyMol Molecular Graphics System Version 1.4.

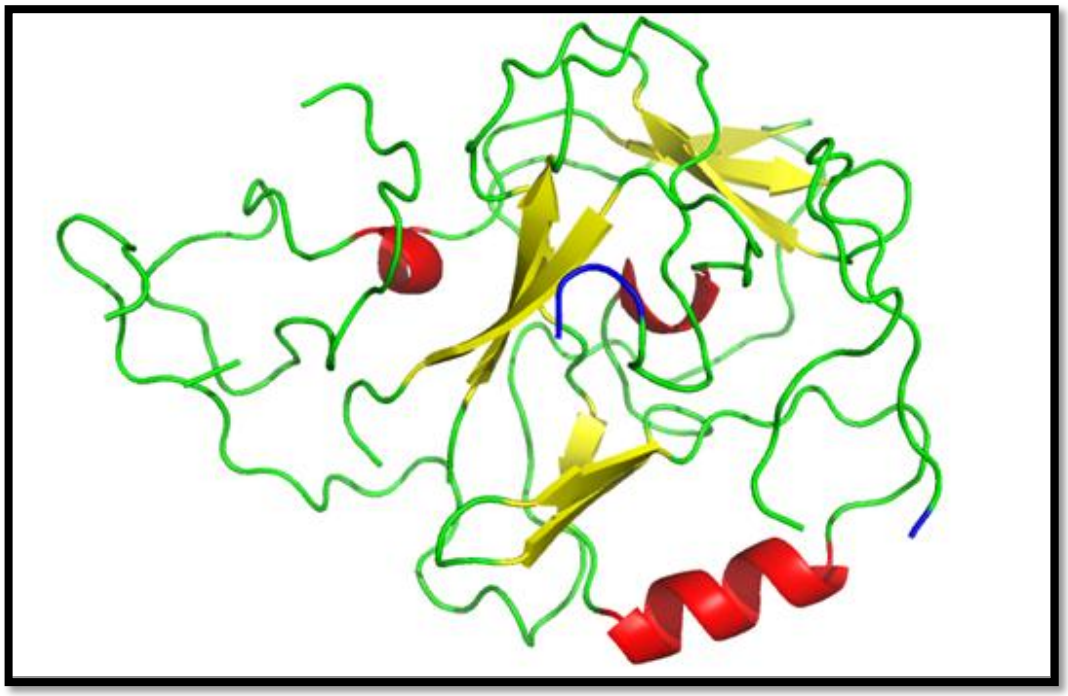

Figure 11. Cartoon theoretical model of MMSET-helix. Gaps in amino acid sequence indicate regions of conformational change. Blue regions represent the location of four sequential amino acid residues involved in the post-SET loop. This figure shows that our protein also undergoes conformational changes in the post-SET loop, which further implies that it behaves in a similar fashion as the post-SET loop of NSD1.

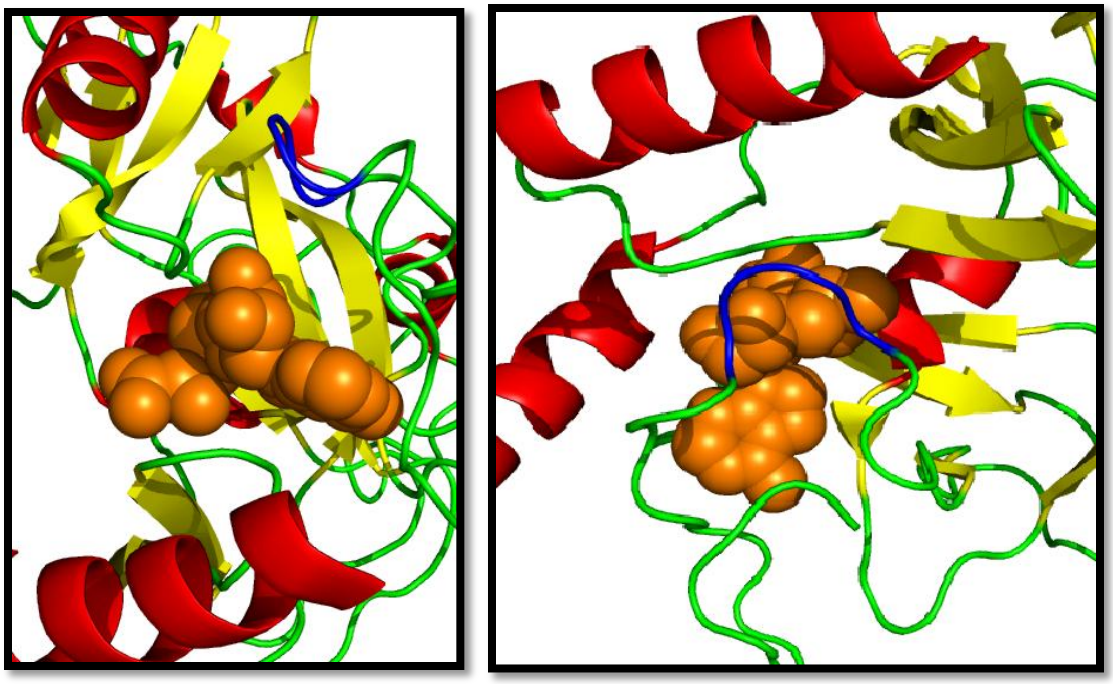

Figure 12. Cartoon representation of NSD1. Orange compound is SAM located in the active site of the enzyme. Blue region is the post-SET loop. These images clearly show that the post-SET loop could inhibit access to SAM in the active site.

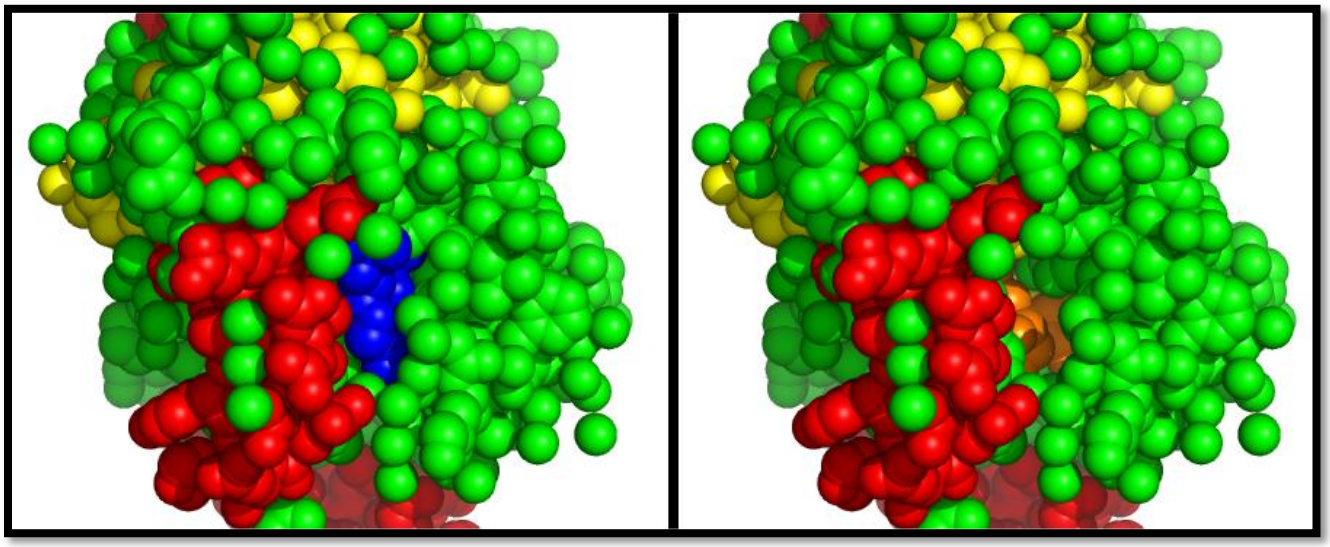

Figure 13. Spherical representation of NSD1. On the left is NSD1 fully intact. Blue region is the post-SET loop. On the right is NSD1 with the post-SET loop removed. Orange compound is SAM. These images support our hypothesis that removing the post-SET loop on our enzyme will increase its activity by making the SAM in the active site more accessible.

Appendix

H3 peptide (21-44) (24 aa)
ATKAARKSAPSTGGVKKPHRYRPG
pI = 11.75
M.W. = 2,521.9 Da
Extinction coefficient = 1,490 M^{-1} cm^{-1}

Biotinylated H3 peptide (21-44) (26aa)
ATKAARKSAPSTGGVKKPHRYRPG-GK(Biotin)-NH2
pI = 11.76
M.W. = 2,724.1 Da
Extinction coefficient = 1,490 M^{-1} cm^{-1}

Recombinant Histone H3 (136 aa)
MARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIR
RYQKSTELLIRKLPFQRLVREIAQDFKTDLRFQSSAVMALQEACEAYLVGLFEDT
NLCAIHAKRVTIMPKDIQLARRIRGERA
pI = 11.13
M.W. = 15,404.0 Da
Extinction coefficient = 4,470 M^{-1} cm^{-1}

MMSET-helix (952-1216) (265 aa)
RGIGRVFKNALQEAEARFREIKLQREARETQESERKPPPYKHIKVNKPYGKVQIY
TADISEIPKCNCKPTDENPCGFDSECLNRMLMFECHPQVCPAGEFCQNQCFTKRQ
YPETKIIKTDGKGWGLVAKRDIRKGEFVNEYVGELIDEEECMARIKHAHENDITH
FYMLTIDKDRIIDAGPKGNYSRFMNHSCQPNCETLKWTVNGDTRVGLFAVCDIP
AGTELTFNYNLDCLGNEKTVCRCGASNCSGFLGDRPKTSTTLSSEE
pI = 6.67
M.W. = 30,198.26 Da
Extinction coefficient = 22,920 M^{-1} cm^{-1}

PWWP-SET (340aa)
KLHFQDIIWVKLGNYRWWPAEVCHPKNVPPNIQKMKHEIGEFPVFFFGSKDYY
WTHQARVFPYMEGDRGSRYQGVRGIGRVFKNALQEAEARFREIKLQREARETQ
ESERKPPPYKHIKVNKPYGKVQIYTADISEIPKCNCKPTDENPCGFDSECLNRML
MFECHPQVCPAGEFCQNQCFTKRQYPETKIIKTDGKGWGLVAKRDIRKGEFVNE
YVGELIDEEECMARIKHAHENDITHFYMLTIDKDRIIDAGPKGNYSRFMNHSCQP
NCETLKWTVNGDTRVGLFAVCDIPAGTELTFNYNLDCLGNEKTVCRCGASNCSG
FLGDRPKTSTTLSSEE
pI = 7.86
M.W. = 39,282.68 Da
Extinction coefficient = 52,370 M^{-1} cm^{-1}

Full Length NSD2 (1365aa)

MEFSIKQSPLSVQSVVKCIKMKQAPEILGSANGKTPSCEVNRECSVFLSKAQLSSS
LQEGVMQKFNGHDALPFIPADKLKDLTSRVFNGEPGAHDAKLRFESQEMKGIGT
PPNTTPIKNGSPEIKLKITKTYMNGKPLFESSICGDSAADVSQSEENGQKPENKAR
RNRKRSIKYDSLLEQGLVEAALVSKISSPSDKKIPAKKESCPNTGRDKDHLLKYN
VGDLVWSKVSGYPWWPCMVSADPLLHSYTKLKGQKKSARQYHVQFFGDAPER
AWIFEKSLVAFEGEGQFEKLCQESAKQAPTKAEKIKLLKPISGKLRAQWEMGIVQ
AEEAASMSVEERKAKFTFLYVGDQLHLNPQVAKEAGIAAESLGEMAESSGVSEE
AAENPKSVREECIPMKRRRRAKLCSSAETLESHPDIGKSTPQKTAEADPRRGVGS
PPGRKKTTVSMPRSRKGDAASQFLVFCQKHRDEVVAEHPDASGEEIEELLRSQW
SLLSEKQRARYNTKFALVAPVQAEEDSGNVNGKKRNHTKRIQDPTEDAEAEDTP
RKRLRTDKHSLRKRDTITDKTARTSSYKAMEAASSLKSQAATKNLSDACKPLKK
RNRASTAASSALGFSKSSSPSASLTENEVSDSPGDEPSESPYESADETQTEVSVSSK
KSERGVTAKKEYVCQLCEKPGSLLLCEGPCCGAFHLACLGLSRRPEGRFTCSECA
SGIHSCFVCKESKTDVKRCVVTQCGKFYHEACVKKYPLTVFESRGFRCPLHSCVS
CHASNPSNPRPSKGKMMRCVRCPVAYHSGDACLAAGCSVIASNSIICTAHFTAR
KGKRHHAHVNVSWCFVCSKGGSLLCCESCPAAFHPDCLNIEMPDGSWFCNDCR
AGKKLHFQDIIWVKLGNYRWWPAEVCHPKNVPPNIQKMKHEIGEFPVFFFGSKD
YYWTHQARVFPYMEGDRGSRYQGVRGIGRVFKNALQEAEARFREIKLQREARE
TQESERKPPPYKHIKVNKPYGKVQIYTADISEIPKCNCKPTDENPCGFDSECLNRM
LMFECHPQVCPAGEFCQNQCFTKRQYPETKIIKTDGKGWGLVAKRDIRKGEFVN
EYVGELIDEEECMARIKHAHENDITHFYMLTIDKDRIIDAGPKGNYSRFMNHSCQ
PNCETLKWTVNGDTRVGLFAVCDIPAGTELTFNYNLDCLGNEKTVCRCGASNCS
GFLGDRPKTSTTLSSEEKGKKTKKKTRRRRAKGEGKRQSEDECFRCGDGGQLVL
CDRKFCTKAYHLSCLGLGKRPFGKWECPWHHCDVCGKPSTSFCHLCPNSFCKEH
QDGTAFSCTPDGRSYCCEHDLGAASVRSTKTEKPPPEPGKPKGKRRRRRGWRRV
TEGK

pI = 9.00

M.W. = 152,258.2 Da

Extinction coefficient = 136,710 M^{-1} cm^{-1}

NSD1 (231aa)

SKELRQLQEDRKNDKKPPPYKHIKVNRPIGRVQIFTADLSEIPRCNCKATDENPCG
IDSECINRMLLYECHPTVCPAGGRCQNQCFSKRQYPEVEIFRTLQRGWGLRTKTD
IKKGEFVNEYVGELIDEEECRARIRYAQEHDITNFYMLTLDKDRIIDAGPKGNYA
RFMNHCCQPNCETQKWSVNGDTRVGLFALSDIKAGTELTFNYNLE**CLGN**GKTV
CKCGAPNCSGFL

pI = 8.10

M.W. = 26,431.1 Da

Extinction coefficient = 22,920 M^{-1} cm^{-1}

References

1. Altekruse, S F., C L. Kosary, M Krapcho, N Neyman, and R Aminou. "SEER Stat Fact Sheets: Myeloma." *Surveillance, Epidemiology and End Results*. National Cancer Institure, Nov. 2009. Web. 29 Mar. 2011. <http://seer.cancer.gov/statfacts/html/mulmy.html>.

2. Brown, M J., P W. Ind, P J. Barnes, D A. Jenner, and C T. Dollery. "A sensitive and specific radiometric method for the measurement of plasma histamine in normal individuals." *Analytical Biochemistry* 109.119 Mar. (1980): 142-46. *Google Scholar*. Web. 29 Mar. 2011.

3. [Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A.;*Protein Identification and Analysis Tools on the ExPASy Server;* (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005). pp. 571-607

4. Hideshima, Teru, P L. Bergsagel, W M. Kuehl, and Kenneth C. Anderson. "Advances in biology of multiple myelomsa: clinical applications." *Blood Journal* 104.3 Aug. (2004): 607-18. *Google Scholar*. Web. 29 Mar. 2011. <http://bloodjournal.hematologylibrary.org.proxy.lib.umich.edu/content/104/3/607.full.pdf+html>.

5. Ji-Young, Kim, Kee Hae Jin, Choe Nak-Won, Kim Sung-Mi, and Eom Gwang-Hyeon. "Multiple-myeloma-related WHSC1/MMSET isoform RE-IIBP is a histone methyltransferase with transcriptional repression activity." *Molecular and Cellular Biology* 28.6 Mar. (2008): 2023-34. *PubMed*. Web. 29 Mar. 2011. <http://mcb.asm.org.proxy.lib.umich.edu/cgi/content/full/28/6/2023?view=long&pmid=18172012>.

6. Kassambara, Alboukadel, Bernard Klein, and Jerome Moreaux. "MMSET is overexpressed in cancers: link with tumor aggressiveness."*Biochemical and Biophysical Research Communications* 379.420 Feb. (2009): 840-45. *PubMed*. Web. 29 Mar. 2011.

7. Kelley, Lawrence A., and Michael J. Sternberg. "Protein structure prediction on the Web: a case study using the Phyre server."*Nature Protocols* 426 Feb. (2009): 363-71. *Google Scholar*. Web. 27 Mar. 2011.

8. Martinez-Garcis, E, R Popovic, D J. Min, S M. Sweet, and P M. Thomas. "The MMSET histone methyl transferase switches global histone methylation and alters gene expression in t(4;14) multiple myeloma cells." *Blood Journal* 117.16 Jan. (2011): 211-20. *PubMed*. Web. 29 Mar. 2011. <http://bloodjournal.hematologylibrary.org.proxy.lib.umich.edu/content/117/1/211.long>.

9. Nellia, Robert. "Mayo Researchers Pinpoint How One Cancer Gene Functions." Mayo Clinic. National Institutes of Health, 2 Feb. 2011. Web. 29 Mar. 2011. <http://www.mayoclinic.org/news2011-rst/6158.html>.

10. "Principles and Applications of Liquid Scintillation Counting." *National Diagnostics*, 2010. Web. 29 Mar. 2011. <http://www.ehs.psu.edu/radprot/LSC_Theory2.pdf>.

11. Qi, Qiao, Li Yan, Wang Mingzhu, Danny Reinberg, and Xu Rui-Ming. "The Structure of NSD1 Reveals an Autoregulatory Mechanism Underlying Histone H3K36 Methylation." *The Journal of Biological Chemistry* 286.1011 Mar. (2011): 8361-68. *PubMed*. Web. 29 Mar. 2011. <http://www.jbc.org.proxy.lib.umich.edu/content/286/10/8361.full.pdf+html>.

12. Reisine, Terry, Richard M. Eglen, Philippe Roby, and Nathalie Rouleau. "The Use of AlphaScreen Technology in HTS: Current Status." *Current Chemical Genomics* 1 Feb. (2008): 2-10. *PubMed*. Web. 29 Mar. 2011. <http://www.ncbi.nlm.nih.gov.proxy.lib.umich.edu/pmc/articles/PMC2775125/>.

13. Smith, Emma M., Kevin Boyd, and Faith E. Davies. "The potential role of epigenetic therapy in multiple myeloma." *British Journal of Haematology* 148.5 Nov. (2009): 702-13. *PubMed*. Web. 29 Mar. 2011. <http://onlinelibrary.wiley.com.proxy.lib.umich.edu/doi/10.1111/j.1365-2141.2009.07976.x/pdf>.

14. Steck, Peter A., Azra H. Ligon, and Paul Cheong. "Two tumor suppressive loci on chromosome 10 involved in human." *Genes, Chromosomes and Cancer* 4.12 Apr. (1995): 255-61. *Google Scholar*. Web. 29 Mar. 2011. <http://onlinelibrary.wiley.com.proxy.lib.umich.edu/doi/10.1002/gcc.2870120404/abstract>.

15. Xiao-Jian, Sun, Wei Ju, Wu Xin-Yan, Hu Ming, and Wang Lann. "Identification and Characterization of a Novel Human Histone H3 Lysine 36-specific Methyltransferase." *The Journal of Biological Chemistry* 280.4222 Aug. (2005): 35261-71. *PubMed*. Web. 28 Mar. 2011. <http://www.jbc.org.proxy.lib.umich.edu/content/280/42/35261.full.pdf+html>.