Determining the Kinetic Parameters of the Protein Lysine Methyltransferase SET7/9 using Liquid Chromatography and Tandem Mass Spectrometry

by Nicholas James Dupuis Stesco

A Thesis submitted to the Faculty of Graduate Studies in The University of Manitoba in Partial Fulfilment of the Requirements of the Degree of

MASTER OF SCIENCE

College of Pharmacy

Rady Faculty of Health Science

University of Manitoba

Winnipeg

Table of Contents

Materials and Methods

Results

Discussion

Acknowledgements

I would like to thank my supervisor, Dr. Ted Lakowski, for all of his help, guidance, and teaching. I would like to thank my advisory committee members, Dr. Michael Namaka and Dr. Mazdak Khajehpour for agreeing to be on my committee and for their input on my work. I would also like to thank Ryan Lillico and Sabina Ozog for all of their help. Funding was provided by NSERC.

For my friends and family who have supported me in many ways. And especially for my partner, Silka Weil, who has always been there for me.

Abstract

Histone lysine methylation is an epigenetic post-translational modification which can modulate gene expression and has been implicated in various forms of cancer. The aim of this research is to determine the kinetic parameters of the lysine methyltransferase SET7/9 using liquid-chromatography and tandem mass spectrometry (LC-MS/MS) techniques.

Reactions are performed *in vitro* using SET7/9 enzyme, S-adenosyl-L-methionine (SAM), and recombinant histone H3 or a peptide of the N-terminal tail of histone H3. Analysis is performed using two assays, one to quantify histone modifications (especially mono-, di-, and trimethyl lysine) and one to quantify S-adenosyl-L-homocysteine (SAH), a co-product of all SAM-dependent methyltransferase reactions. The data collected are then used to calculate the apparent kinetic parameters of the reaction, K_m^{app} and V_{max}^{app}. For SET7/9 the K_m^{app} and V_{max}^{app} for SAM were 2.24 ± 0.97 µM and 0.047 ± 0.0057 pmol/min when full-length histone H3 was used, and 0.22 ± 0.03 µM and 0.19 ± 0.004 pmol/min when H3 peptide was used; the Km^{app} and Vm_{ax}^{app} for histone H3 were 1.21 ± 0.53 µM and 0.16 ± 0.018 pmol/min.

Tables:

List of Figures:

List of Abbreviations:

BME: β-mercaptoethanol

BSA: bovine serum albumin

CpG: cytosine-phosphate-guanidine

ELISA: enzyme-linked immunosorbent assay

FPLC: fast protein liquid chromatography

H3K4: lysine 4 of histone H3

H3K4Me: monomethylated lysine 4 on histone H3

H3K4Me2: dimethylated lysine 4 on histone H3

H3K4Me3: trimethylated lysine 4 on histone H3

H3T11P: phosphorylated threonine 11 on histone H3

HP1α: heterochromatin protein 1 α

K_m: Michaelis-Menten constant

K_m^{app}: apparent K_m

IMAC: immobilized metal ion affinity chromatography

K: lysine

KMe: monomethyl lysine

KMe2: dimethyl lysine

KMe3: trimethyl lysine

L3MBTL1: lethal (3) malignant brain tumour-like protein 1

LC-MS/MS: liquid chromatography and tandem mass spectrometry

MBT: malignant brain tumour

miRNA: micro RNA

MRM: multiple reaction monitoring

MYND: myeloid translocation protein 8, nervy, and deaf-1

N-CoR: nuclear receptor co-repressor 1

NMR: nuclear magnetic resonance

PHD: plant homeodomain

pI: isoelectric point

PKMT: protein lysine methyltransferase

PRMT: protein arginine methyltransferase

PTM: post-translational modification

PWWP: proline-tryptophan-tryptophan-proline

RMe: monomethyl arginine

aRMe2: asymmetric dimethylarginine

sRMe2: symmetric dimethylarginine

SAH: S-adenosyl-L-homocysteine

SAM: S-adenosyl methionine

SET: suppressor of variegation, enhancer of zeste, and trithorax

TCA: trichloroacetic acid

V_{max}: maximum reaction velocity

V_{max}^{app}: apparent V_{max}

1.0 Introduction

1.1 Epigenetics

Epigenetics is the study of inheritable changes in gene expression which are not directly coded into the genomic DNA sequence. Rather, epigenetic changes in gene expression are mediated by methylation of DNA methylation, micro RNA (miRNA), and post-translational modifications to histones. DNA methylation is one of the most studied examples of epigenetic phenomena. DNA methylation typically occurs on cytosine-phosphate-guanine (CpG) sites (CpG islands), producing 5-methylcytosine. When such CpG methylation happens in the promoter of a gene it typically results in a decrease in gene expression due to formation of heterochromatin (*1*). This mechanism will be explained below. By contrast, unmethylated CpG islands have been shown to promote a transcriptionally active state in DNA (*2*). DNA methylation at these sites decreases gene expression and has also been shown to increase deposition of other epigenetic marks. The post-transcriptional regulation of gene expression is regulated by RNA silencing, which is regulated by miRNA. These miRNA target mRNA, which are then digested in a target mRNA specific manner. In contrast, histones are involved in packaging DNA. Post-translational modifications to histones regulate how densely or loosely DNA is packed. On a localized level, gene specific histone PTMs can induce gene specific relaxation of DNA packaging which allows better access to gene transcription machinery.

The protein effectors that mediate epigenetic changes in gene expression can be categorized as ‘readers, writers, or erasers’. Epigenetics writers are enzymes which form an epigenetic mark, such as DNA methylation. A reader is any protein which recognizes and interacts with a specific epigenetic mark. Finally, an eraser is an enzyme which removes epigenetic marks. The interplay

between members of these three categories gives us a set of epigenetic mechanisms which are collectively referred to as the histone code (*3*). However more and more evidence shows that these events are also influenced, or even caused, by processes which are not historically considered "epigenetic" (*2*). Most notable is the post-translational modification (PTM) of non-histones by enzymes which were originally characterized as epigenetic writers (*1*). The term epigenetics also refers to chromatin biology and how DNA structure is affected by interactions between chromatin and histones (*1*, *4*, *5*).

Histones operate as parts of an octameric complex, which is composed of two tetramers of the histones H2A, H2B, H3, and H4. The full complex is positively-charged, allowing it to interact electrostatically with the negatively-charged phosphate backbone of DNA (*6*). When the full complement of histones is bound to a segment of DNA, the unit is called the nucleosome. A nucleosome is defined as 147 DNA base pairs wound around a histone octamer (**Figure 1A**) (*1*). This complex is responsible for packaging DNA so that it can fit into the nucleus of a cell. However, this packaging aspect is intimately related to the role that histones play in the control of gene expression. Through a process of condensing or relaxing, the packaging of chromatin can form heterochromatin or euchromatin respectively (**Figure 1**). Genes within condensed, tightly packed heterochromatin are not as accessible to transcriptional machinery, while genes within relaxed, loosely packed euchromatin are available to the transcriptional machinery. The condensation of chromatin is dictated in part by a variety of PTMs to histones that modulate interactions between DNA and histones and recruit transcription factors, co-activators, and co-repressors (*7*). The pattern of PTMs to histones is a heritable trait passing gene expression information to daughter cells after cell division.

Figure 1. A nucleosome is made up of a histone octamer with 147 bp of DNA wound around it. The nucleosome is the base unit of chromatin. There are two ways in which nucleosomes hold their segments of DNA: (A) loosely wound euchromatin, which allows for transcription of genes, and (B) tightly wound heterochromatin, which limits transcription of genes. The dynamic between hetero- and euchromatin is regulated, in part, by histone modifications.

1.2 Post-translational modifications (PTMs) of histones

There are many PTMs which are commonly found on histones (**Figure 2**). Epigenetic marks on histones include, but are not limited to, methylation of lysine and arginine; acetylation of lysine; phosphorylation of serine, tyrosine, and threonine; ubiquitylation of lysine; sumoylation of lysine; glycosylation of serine and threonine; deimination of arginine; ADP ribosylation of

glutamic acid and arginine; and proline isomerization among many others (*6*). These PTMs are placed predominantly on the N-terminal tail of histones. However, another common PTM that has been observed is the removal of 6 to 21 amino acids from histone tails. Researchers have shown relationships between chromatin states and the presence of these marks on specific histone residues. These marks can increase transcription through recruitment of transcription factors, or prevent gene transcription by obstructing the binding of factors (*6*). For example, phosphorylation of threonine 11 in histone H3 (H3T11P) has been shown to recruit WDR5, which is part of a methyltransferase complex which targets lysine residues. This increases trimethylation of H3K4, which is generally shown to be a positive mark, associated with transcriptionally-active euchromatin (*8*, *9*).

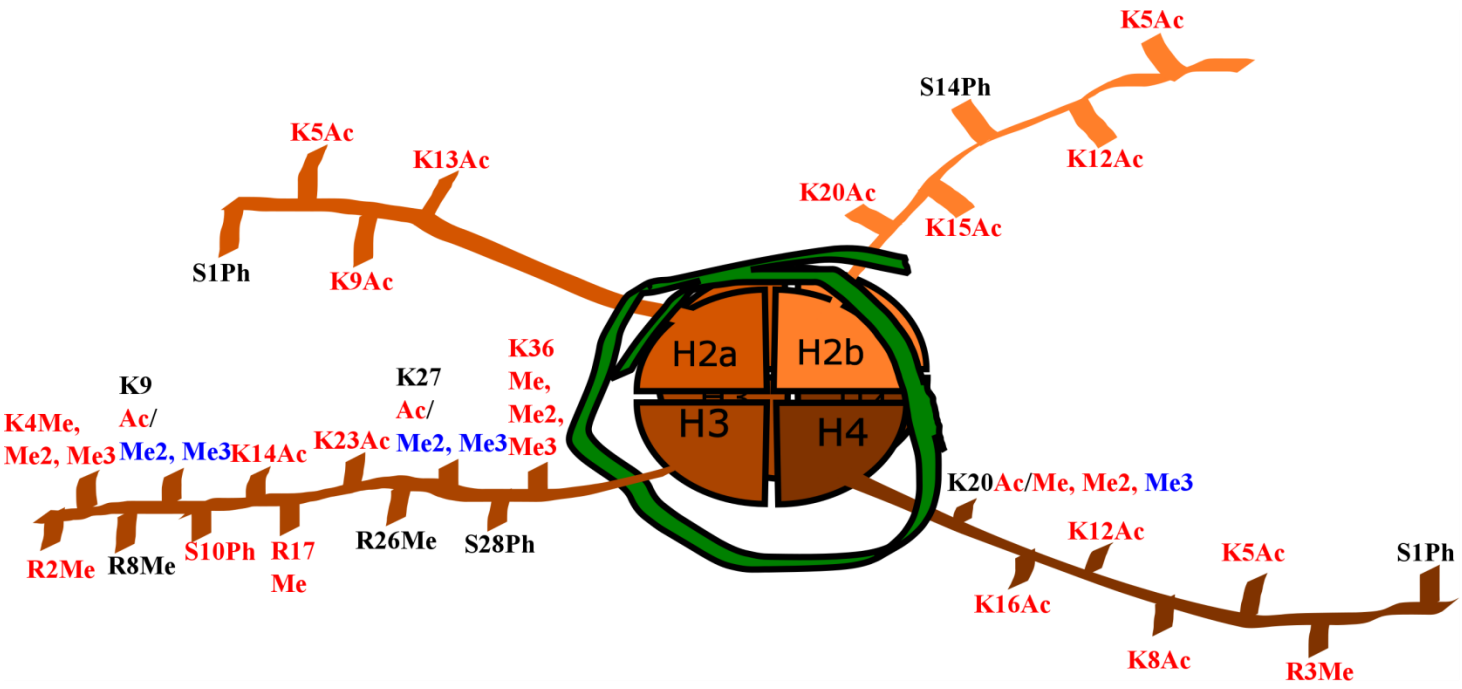


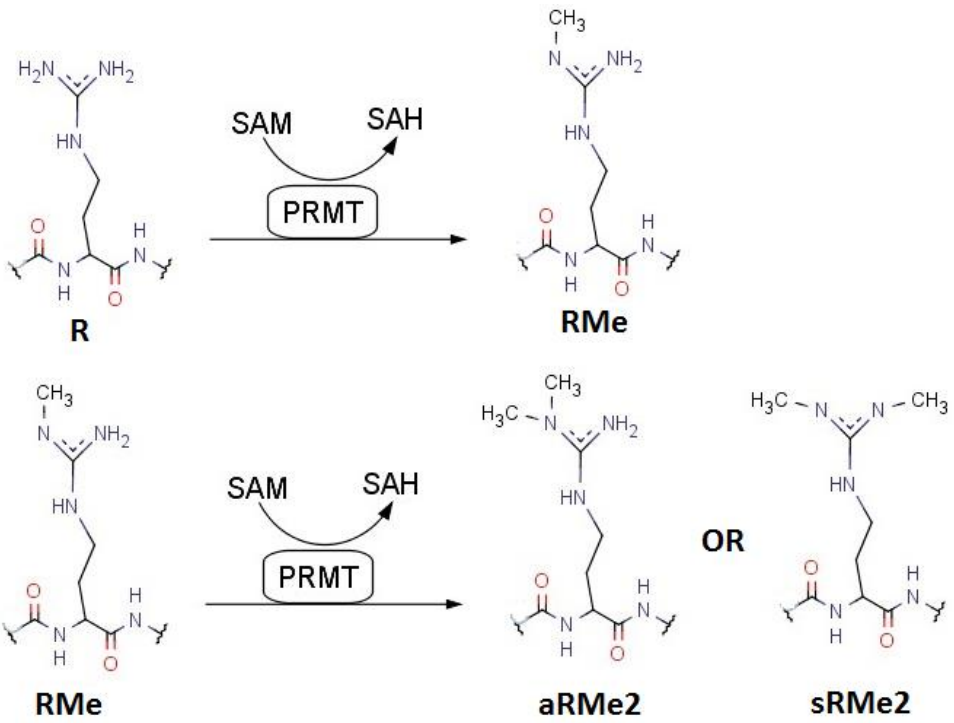
Figure 2. Various PTMs found on the N-terminal tails of histones. PTMs which are associated with increased gene expression are shown in red, while PTMs associated with decreased gene

expression are shown in blue. The combination of histone PTMs leading to increased or decreased transcription is referred to as the histone code (*3*).

The combination of PTMs found on histones and how specific combinations increase or decrease transcription is sometimes referred to as the histone code (*8*, *3*). However, generalized definitions in this field are still being debated and it is important to understand that any one piece of information regarding the histone code must ultimately fit into the larger context of many other epigenetic marks. However, it is also important to focus on specific aspects of the code so as to come to broader conclusions which can eventually be more easily fit together. As such, we consider the term epigenetics to refer to changes in gene expression as a result of both chromatin structural changes and protein recruitment by histone modifications, and will focus on histone lysine methylation.

Methylation of histones is known to take place on lysine and arginine residues, and particularly on H3 and H4. A wide variety of these methyl marks are regulated by enzymes known as methyltransferases. Arginine is known to be either mono- or di-methylated. Dimethylated arginine exists in either a symmetric or asymmetric form. aRMe2 and sRMe2 are produced by type 1 and type 2 protein arginine methyltransferases (PRMT) respectively (*10*). Methyllysine can be found in mono-, di-, or tri-methylated forms; different protein lysine methyltransferases (PKMTs) are able to methylate lysine to different degrees. The methylation reaction uses S-adenosyl methionine (SAM) as a co-substrate, transferring the methyl group attached to the sulfur atom in the compound to the protein substrate, thereby producing S-adenosyl L-homocysteine (SAH) as a co-product (**Figure 3**) (*4*).

A


SAM
SAH
PRMT
R
RMe
RMe
aRME2
OR
sRME2

B

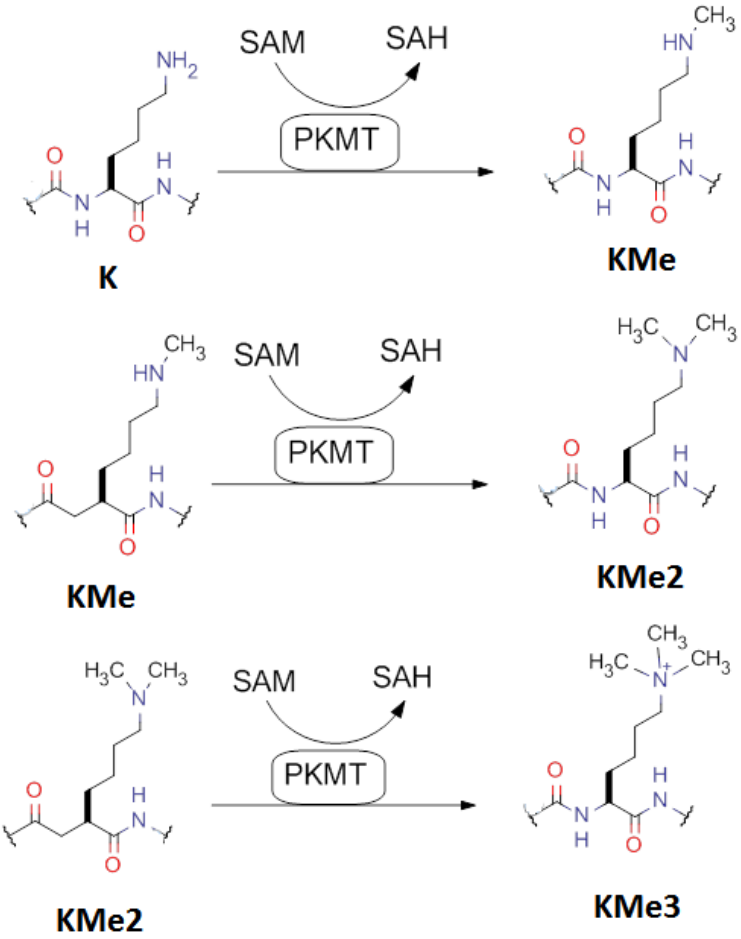

SAM
SAH
PKMT
K
KMe
KMe
KMe2
KMe2
KMe3

Figure 3. (A) Methylation reactions catalysed by PRMTs. PRMTs can catalyse the formation of RMe, aRMe2, and sRMe2 (*11*). (B) Methylation reactions catalysed by PKMTs. PKMTs can catalyse the formation of KMe, KMe2, or KMe3 depending on their ability to bind to the substrate (*12*).

Histone methylation typically influences gene expression by altering chromatin structure, usually by recruiting or impeding recruitment of transcription factors to promoter regions through their reader domains. This is thought to occur primarily by the methylation mark serving as a binding site or preventing binding because methylation does not affect the charge state of histones (*6*). The transcription factors being affected can be gene activators or repressors. Specific examples will be discussed further below. Protein reader domains which are known to interact with various methyl lysine marks include chromodomains, plant homeodomain (PHD) finger, Tudor, proline-tryptophan-tryptophan-proline (PWWP) motif, and malignant brain tumour (MBT) domains (*6*). For example, trimethylation of histone H3 at lysine 9 or lysine 27 (H3K9me3 or H3K27me3) represses promoters governed by RNA polymerase II (*13*). Another methyl lysine-recognizing protein which is important for chromatin remodeling is heterochromatin protein 1α (HP1α). This protein binds to H3K9me3 using a chromodomain (*6*). HP1α proteins then recognize and bind to histone H3 via a chromoshadow domain to induce condensation of chromatin (*4*). It is of note that numerous bacteria and archaea have been found to express lysine methyltransferases. For example, the hyperthermophilic archaea *Sulfolobus solfataricus* has a β-glycosidase enzyme which is methylated on up to five residues in order to protect the enzyme from high temperatures (*14*). While the purpose of methylation in these

prokaryotes in still largely unknown, this shows that methylation and the enzymes responsible for it are found in very diverse contexts.

1.3 Protein lysine methyltransferase activity and disease

There are two major conserved protein domains which are commonly found in histone lysine methyltransferases (PKMTs). These are MYND and SET domains. The SET domain is responsible for methylation of a substrate using SAM as a co-factor; as such, most PKMTs with the exception of DOT1L contain a SET domain (*6*, *15*). The SET domain is 130 amino acids long and it is named after several *Drosophila* proteins: Suppressor of variegation, Enhancer of zeste, and Trithorax. The SET domain contains several different regions – the C-SET, N-SET, I-SET, and post-SET regions – and is the part of SET domain PKMTs which gives them their catalytic activity (*15, 16*). MYND stands for Myeloid translocation protein 8, Nervy, and Deaf-1, and this domain is also found within those proteins (*13*). It is a zinc-finger domain comprised of numerous histidine and cysteine residues, and is involved in protein-protein interactions and possibly DNA binding (*17, 18*). The MYND domain is 40 to 60 amino acids long, with two zinc-binding motifs. These motifs show a conserved spacing of histidine and cysteine residues, following either a C-X-X-C or C/H-X-X-X-C pattern (*19*). These zinc-binding motifs have been shown, *via* NMR Nuclear Overhauser Effect Spectroscopy, to interact with one another, which allows for the folding of the MYND domain within the rest of the protein, stabilizing the disulfide bonds present while in the reducing environment of the intracellular compartment (*19*). The MYND domain has been shown to mediate protein-protein interactions, however, mutation studies have shown that loss of the MYND domain does not affect methylation activity (*18*, *20*).

The SET domain is responsible for binding both of these enzymes' substrates, SAM and the protein to be modified. Structural studies of PRMTs have shown that binding of each substrate proceeds sequentially, with SAM binding first. SAM binding causes a conformational change in the enzyme which opens a second binding site, allowing the enzyme to bind its peptide substrate (*21*). Computational quantum mechanics, molecular mechanics and molecular dynamics simulations have predicted that this is true for PKMTs as well (*22*). Methyltransferases bind the peptide and SAM in different sites and bring them together to facilitate the transfer of the methyl group from the bound SAM to the bound peptide (*23*). The binding sites are within 3 Å of each other within the enzyme structure and are connected by a channel in the SET domain which is only just large enough for a methyl group to move through (*22*, *24*). Further structural studies have shown that the peptide-binding site in PKMTs recognizes its specific substrate via a cage of aromatic amino acids, especially conserved tyrosine residues (*8*, *24*).

PKMTs have been implicated in a variety of disease states, particularly in many different forms of cancer (*25*). Perhaps this is because PKMTs operate throughout the genome, so it is not surprising when they are revealed to have an effect on genes which regulate the cell cycle (*8*, *18*, *26*). Abnormal expression of many different PKMTs has been linked to either onset of disease states or poor patient outcomes, or both, to the point that some researchers have proposed using these expression patterns as biomarkers (*25*–*29*). While the proposed link between PKMTs and disease is usually their effect on either the expression or activity of proteins involved in DNA damage response and cell cycle regulation, the mechanisms in specific cases are still largely unknown (*30*).

SET7/9 is a SET domain PKMT which is regarded as H3K4-specific and is only capable of mono-methylating this residue (*6*). It has been suggested that this is due to a tyrosine in the

active site of this enzyme which decreases the active site volume so that it cannot accommodate di- or tri-methyl lysine (*31*). In PKMTs capable of catalysing formation of lysine di- and tri-methylation this tyrosine residue is replaced by a phenylalanine residue, meaning that the amino acid present at this position is what determines the methylation activity of the enzyme (*32*, *33*). The deposition of H3K4Me by SET7/9 is a transcriptionally positive epigenetic mark, meaning that it is associated with a transcriptionally active chromatin state. Methylation at this position also inhibits the binding of a histone deacetylase complex, meaning this mark is also positively associated with increased histone acetylation (*34*). SET7/9 has also been shown to methylate non-histone proteins such as the tumour suppressor protein p53, estrogen receptor α, androgen receptor, and TAF10 (*34–36*). Due to its effects on cell cycle regulators, increased SET7/9 activity has been linked to poor prognosis in hepatocellular carcinoma through its association with metastasis, larger tumour size, and recurrence (*26*). However, other studies have shown that increased SET7/9 expression can decrease cancer cell proliferation (*37*, *38*). For example SET7/9 was shown to methylate β-catenin, a mediator of the Wnt/β-catenin signaling pathway which promotes cell proliferation and tumorigenesis. Methylation of β-catenin by SET7/9 decreased its half-life within the cells, reducing proliferation of cervical cancer cells (*38*). Low SET7/9 expression has also been associated with onset of breast cancer and poor patient outcomes in cases of gastric cancer (*37*, *39*). A possible reason for this is that methylation of p53 by SET7/9 stabilizes the tumour suppressor, allowing it to better initiate transcription (*36*, *40*).

A sub-family of methyltransferases is characterized by the presence of both a SET and a MYND domain. These are referred to as SMYD which stands for SET and MYND domain containing enzymes. These enzymes possess a SET domain which is interrupted with an MYND domain. The N-terminal portion of the SET domain (N-SET) is followed by a MYND domain,

which is followed by the I-SET domain, and the C-terminal portion of the SET domain (C-SET) (*13*). It is then followed by a cysteine-rich zinc binding fold, referred to as the post-SET domain (*13*). The SMYD SET core contains three highly conserved protein motifs; a GxG motif, a NHxCxPN motif, and a GEExxxxY motif (*15*). In this notation, x is a stand-in for any amino acid. These motifs enable SMYD enzymes to bind SAM and mutation of either the NHxCxPN or GEExxxxY will abolish enzymatic activity (*15*). SMYD3 is known as an H3K4me3-specific PKMT and is often cited as a proto-oncogene. It has been shown that over-expression of SMYD3 in cell culture increases cell proliferation; 80 genes have been shown to be targets of SMYD3 activity, many of which are cell cycle regulators (*13*). Van Aller *et al* showed that SMYD3 is also capable of methylating H4K5, and that suppression of SMYD3 activity in human breast carcinoma and hematoma cells resulted in the loss of the ability of the cancer cells to form colonies in an anchorage-independent environment (*41*). This provides a link between the activity of SMYD3 and cancer cell phenotype. SMYD3 has also been shown to bind a transcription repressor protein, nuclear receptor co-repressor 1 (N-CoR), by way of an interaction between the MYND domain of SMYD3 and a proline-rich PXLXP motif present in N-CoR (*13*, *19*).

SMYD2 differs from other SMYD PKMTs by the fact that it dimethylates its histone H3 substrate, H3K36 (*42*). It can also methylate histone H2A and H4 and is an oncogene because of its ability to methylate the tumour suppressors p53 and retinoblastoma protein (Rb). As such, the function of p53 is impaired by methylation at K370 by SMYD2 (*15*). Since the activity of p53 as a transcriptional activator is highly involved in cell cycle control, regulation of apoptosis, and the cellular response to DNA damage, impairment of p53 has high potential to contribute to oncogenesis (*36*, *43*). SMYD2 also methylates Rb at K860 and K810, giving the transcriptional

repressor L3MBTL1 a binding site on this protein (*42*, *44*). L3MBTL1 has been shown to bind to RbK860Me via a 3xMBT domain, acting as a co-repressor of Rb (*43*, *44*). The ability of SMYD2 to methylate these non-histone proteins has been associated with onset and/or poor prognosis of several types of cancers such as esophageal squamous cell carcinoma, gastric cancer, breast cancer, and chronic lymphocytic leukemia (*45*, *29*, *30*, *46*). SMYD2 contains two functionally important, conserved tyrosine residues: Y240 and Y258. Y240 is necessary for methylransferase activity, since a Y240F mutation has been shown to prevent the function of the enzyme (*47*). Y258 has been shown to facilitate the proper orientation of the ε-amino group of the peptide substrate within the binding pocket of the enzyme (*15*). SMYD2 is highly expressed within the human heart and brain, and over-expression results in a general up-regulation of genes (*15*, *20*).

Enhancer of Zeste homolog 2, or EZH2, is a SET domain PKMT which is part of the polycomb repressor complex 2 and produces the repressive histone modifications H3K9Me, H3K27Me and Me2 (*48*, *27*). Many types of lymphomas contain aberrant EZH2 activity, and over-expression of EZH2 has been shown to occur in basal-like breast cancer and oral cancers (*23*, *46*, *27*). The methyl marks deposited on H3K27 by EZH2 serve as a binding site for the polycomb repressor complex 1, which ubiquitinates histone H2A at K119, causing chromatin compaction and decreasing gene expression (*48*). However H3K27 methylation is also associated with H3K4 methylation, which is known to increase gene expression (*48*). As has been noted with SET7/9, a mutation replacing a key tyrosine residue in the active site of EZH2 has been shown to expand its methylating abilities. A Y641F mutation in this enzyme has been shown to trimethylate H3K27 (*23*). As with other SET domain PKMTs, EZH2 has been shown to act on non-histone proteins, such as the transcription factor STAT3 (*48*). Although this still suggests a

role in regulation of expression, it is not expressly epigenetic in nature because methylated STAT3 may not be inherited.

G9a is a PKMT which methylates H3K9 as well as H3K27, producing repressive marks (*24*). It produces predominantly KMe and KMe2, but will produce KMe3 during long incubations *in vitro* (*49, 50*). The H3K9 methylation mark is a binding site for HP1α, which induces a transition to heterochromatin upon binding. This makes the PTM catalyzed by G9a a repressive mark (*49, 50*). Like other PKMTs, G9a has been shown to methylate non-histone proteins. It has been shown to methylate proteins with domains which bind to methylated lysines such as widely interspaced zinc finger motifs protein (WIZ) (*50*). G9a has also been shown to methylate certain epigenetic writers, showing some methylation of the histone deacetylase 1 and DNA methyltransferase-1 as well as the N-terminal domain of other G9a molecules, in a process called automethylation where a methyltransferase can methylate either itself or other molecules of the same enzyme nearby (*50, 51*). The methylation of these non-histones by G9a has been shown to give HP1α a binding site, particularly on WIZ and G9a (*50*). The activity of G9a for non-histones could therefore play a part in the regulation of heterochromatin formation. G9a activity has been associated with lung cancer progression, in particular with an increase in metastasis of lung cancer cells. Chen *et al* have shown that an increase in H3K9Me2 as a result of G9a activity in lung cancer cells decreases expression of a cell adhesion molecule, Ep-Cam, leading to increased metastasis and a poor prognosis (*28*).

The enzymes discussed so far have all been SET domain enzymes. DOT1L is one of very few PKMTs which do not contain a SET domain, along with the yeast PKMT DOT1 (*52*). The first enzyme in this family was discovered in yeast and was called disruptor of telomeric silencing, or DOT1 (*52*). The homologue in humans is called DOT1-like, or DOT1L. Aside from

the absence of a SET domain, what distinguishes DOT1L from other PKMTs is the fact that it methylates histone H3 within the globular domain rather than on the N-terminal tail. DOT1L is known to mono- and dimethylate H3K79, an epigenetic mark which is associated with active genes (*53, 54*). Abnormal DOT1L activity is involved in some types of mixed-lineage leukemia

Mixed lineage leukemia is a form of acute myeloid leukemia which represents 10% of all acute leukemia and is caused by chromosomal translocations of a SET domain methyltransferase, mixed lineage leukemia 1 (MLL1). MLL1 has both DNA binding and histone lysine methyltransferase domains. The histone lysine methyltransferase domain catalyzes the formation of the permissive modification H3K4 methylation. Mixed lineage leukemia results in abnormal hematopoiesis characterized by self-renewing hematopoietic precursor cells and thus leukemia (*55*). Normal hematopoiesis is controlled by the Homeobox protein A9 (HOXA9) and its expression is regulated by the gene MLL (*56*). MLL binds to the DNA of the *HOXA9* promote, methylating neighbouring histone H3 lysine 4 (H3K4) residues. MLL also binds MOF, which acetylates histone H4 lysine 16 (H4K16) (**Figure 4A**) (*57*). Together these epigenetic processes prepare chromatin for transcription and regulate HOXA9 expression (*55*). Mixed lineage leukemia is caused by chromosomal translocations of *MLL* and more than 60 forms are known that result in fusions between MLL and a variety of proteins that eliminate the histone lysine methylation activity of MLL but not its DNA binding capacity (*58, 59*). Acute lymphocytic leukemia fused 9 (AF9) produces one of these fusions MLL-AF9 which eliminates MOF binding while recruiting the lysine methyltransferase DOT1L. This eliminates H4K16 acetylation and H3K4 methylation while introducing ectopic H3K79 methylation, rewriting the histone code at the HOXA9 promoter, resulting in HOXA9 overexpression and leukemia (**Figure 4B**) (*56, 60*). HOXA9 plays a central role in mixed lineage leukemia. Transplantation of

cells overexpressing HOXA9 into mice rapidly induces leukemia (*61*, *55*). HOXA9 is overexpressed in mixed lineage leukemia cell lines and this is correlated with poor prognosis in humans (*62*). Moreover, cells expressing the MLL-AF9 mutant fusion protein become adapted to the higher expression levels of HOXA9 and become stalled with respect to differentiation in the precursor stage. Reduction in the expression of HOXA9 to normal levels has been shown to result in rapid death of cells expressing the MLL-AF9 fusion. Through its recruitment by MLL-AF9, DOT1L inhibition has a specific effect on mixed-lineage leukemic cells, making it an attractive target for therapies (*63*).

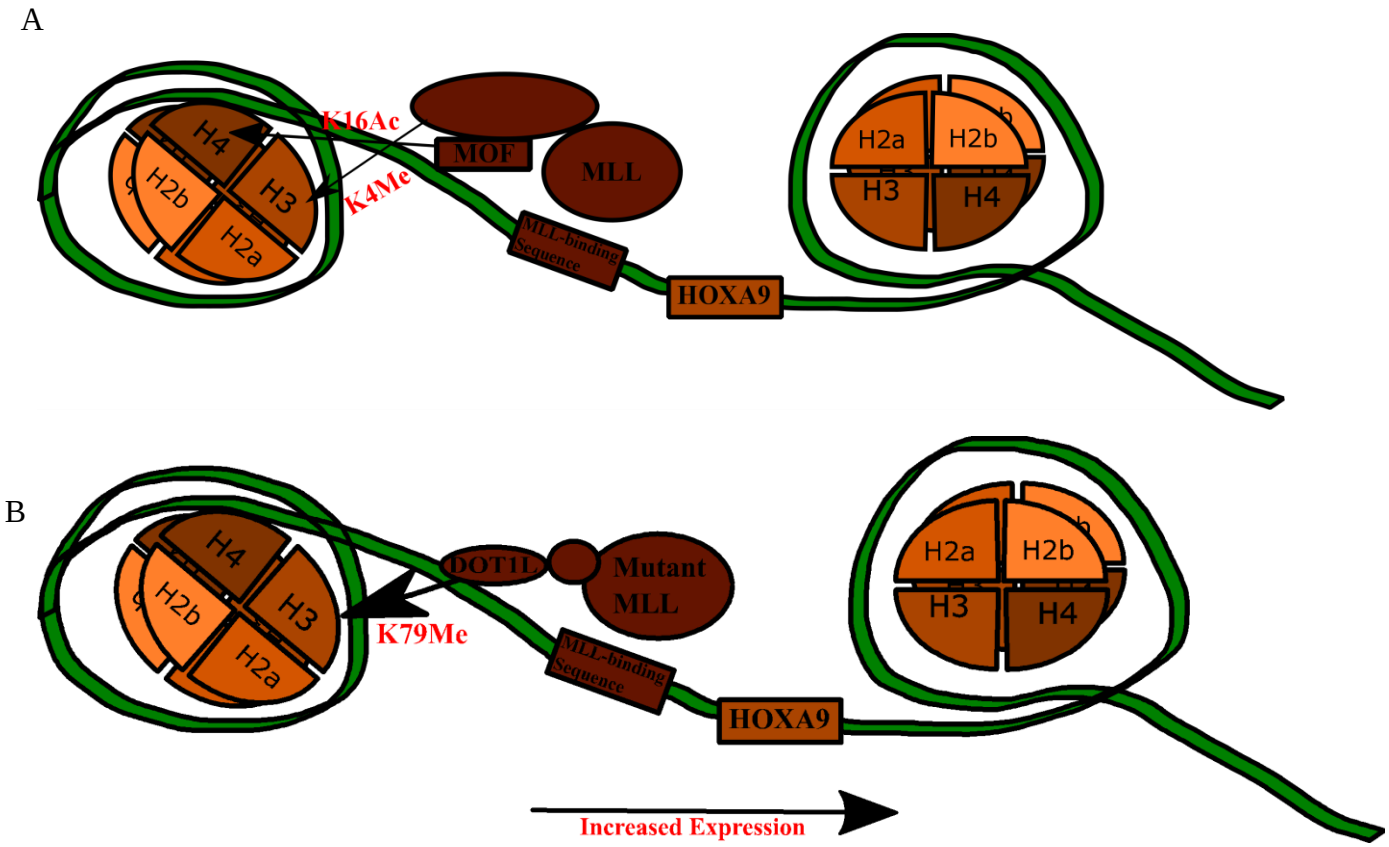


Figure 4. Recruitment of epigenetic enzymes by (A) Wild type MLL recruits MOF and binds to the *HOXA9* promoter, resulting in methylation of H3K4 and acetylation of H4K16 by MLL and

MOF respectively. Together these modifications regulate normal *HOXA9* expression, resulting in normal hematopoiesis. (B) Mutant MLL-AF9 fusion protein recruits DOT1L and binds to the *HOXA9* promoter, resulting in the loss of the H3K4 methylation and H4K16 acetylation marks and introducing methylation of H3K79 by DOT1L. This leads to overexpression of *HOXA9* and onset of leukemia (*56*, *64*).

SET7/9 has been shown to be overexpressed in tumour cells compared to non-tumour cells, and knockdown of SET7/9 in cultured cells has been shown to result in cell cycle arrest prior to the G1 phase (*26*). Increased G9a expression has been associated with various aggressive forms of lung cancer (*63*). Due to this variety of mechanisms through which PKMTs can be involved in the onset of disease, PKMTs have become popular targets for drug development and aberrant expression of PKMTs has been indicated as potential biomarkers indicating poor prognosis (*27*). To this end, kinetic analyses of these enzymes are important because that is how we can determine the mechanism of action for these enzymes and the conditions they require. By increasing our understanding of how epigenetic enzymes work, we can more reliably produce drugs to target them.

The enzymes responsible for modifying these residues are thought to be very specific with regard to their substrates; the enzyme SET7/9 has been shown to methylate lysine 4 on the H3 protein (*34*). However, epigenetic enzymes can have substantially less substrate specificity, especially *in vivo*. For example, some methyltransferases have shown the ability to transfer bulky alkynes to peptides instead of simple methyl groups using modified analogues of SAM containing transferable alkyne groups (*65*, *66*). Others have been shown to methylate residues which have been chemically modified (*21*). The fact that these enzymes can accept substrates and co-factors which have been radically altered from their usual forms suggests that there is

some flexibility in the active sites of methyltransferases. If this is the case, these enzymatic reactions may be much less specific than previously thought. A limiting factor in studying these marks is the difficulty of obtaining useful data. Epigenetic enzymes may behave differently *in vitro* compared with *in vivo*, or assays may have difficulty detecting the placement of specific marks.

1.4 Lysine methyltransferase enzyme kinetics

All PKMTs are bisubstrate enzymes, meaning that they bind two substrates, SAM and a lysine within a protein, to catalyze the methylation of this lysine. There are different ways in which bisubstrate enzyme reactions can proceed. This depends on whether the reaction is random or ordered. In a random reaction the enzyme can bind its substrates in either order, and the products can leave in any order (**Figure 5**). In an ordered reaction, there is a specific order in which the substrates must be bound and in which the products leave. There is also the possibility of a double displacement reaction, where one substrate binds to the enzyme, covalently modifies it, and then leaves the enzyme. The second substrate is then bound by the enzyme and the modification is transferred to the second substrate, typically *via* nucleophilic attack (*67*). In any of these situations the binding of one substrate may alter the enzyme's affinity for the other substrate. Most methyltransferase enzymes have been determined to proceed *via* either a random or ordered bisubstrate mechanism; none have been observed to have a double displacement mechanism. PRMT6 has been shown to proceed *via* an ordered reaction where the binding of SAM to the enzyme causes a disordered region in the N-terminus of the enzyme to take on a shape which allows for binding of the protein substrate (*11*, *68*). Most PKMTs, however, seem to follow a random bisubstrate mechanism (*49*, *69*, *70*). We will therefore assume that SET7/9 follows a random bisubstrate mechanism such as that depicted in **Figure 5**. The next

consideration is whether these enzymes follow a processive or distributive mechanism. A processive reaction is one in which the enzyme can catalyze multiple rounds of the reaction without dissociating from its substrate (*71*). A common example of this is found in DNA polymerase enzymes, which catalyze the addition of many nucleotides to a growing chain without releasing the chain. By contrast, a distributive mechanism is one in which the enzyme must release all substrates between rounds of catalysis (*70*). Most methyltransferases have been determined to follow a distributive mechanism (*11, 68*). However, some research has argued that certain methyltransferases may be processive. For example, *Patnaik et al* argued that G9a is a processive enzyme because it behaves similarly and has sequence similarity to DIM-5, another processive PKMT (*49*).

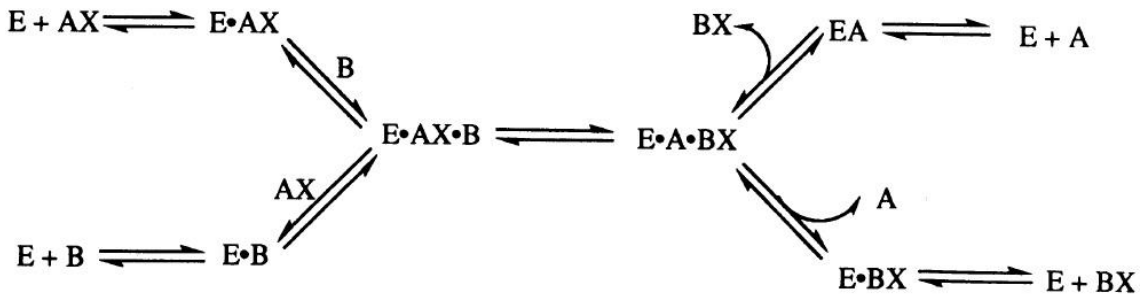


Figure 5. Diagram of the possible mechanisms of a random ordered bi-bi enzymatic reaction. In the context of PKMT reactions E refers to the PKMT, while AX refers to SAM and B refers to the protein substrate. In this mechanism the enzyme can bind to each substrate in either order, but the reaction must proceed through the E·AX·B complex (*67*).

1.5 Methyltransferase activity assays

The choice of technique when studying epigenetics is therefore a very important consideration. Many assays have been developed for the detection of epigenetic marks and measuring the activity of epigenetic enzymes, including radioactive based assays, mass spectrometry based assays including proteomic methods and the use of antibodies which bind to specific PTMs. However many of these methods are qualitative rather than quantitative, giving data that can be somewhat limited in scope.

1.5.1 Radioactive methods

Radiography assays have been widely used to detect product formation in *in vitro* methylation reactions. The goal of these assays is usually to show that enzymes of interest have methyltransferase activity. These involve carrying out the reaction in the presence of radiolabeled SAM, typically either [^{14}C methyl]SAM or [^{3}H methyl]SAM such that a radioactive methyl-group is transferred to the substrate protein producing a radio-labeled methylated substrate (*31*, *72*). Therefore, once the methylation reaction has taken place, there should be detectable radioactivity on the substrate provided that the protein is an efficient substrate for the PKMT in question. The proteins can then be separated *via* SDS-PAGE and any radioactivity can be detected by film, storage phosphor imaging, or by cutting out the product bands and performing scintillation counting (*72*).

Scintillation proximity-based assays are another radioactive type assay that use radiolableled SAM (typically [^{14}C-methyl]SAM or [^{3}H-methyl]SAM). These assays have the advantage of being faster to develop than storage phosphor screens. The assay uses small beads containing a solid core of scintillant. The beads are coated with coupling molecules such as

antibodies or streptavidin if the peptides being used are biotinylated. The peptides stick to the beads and any which have been methylated with radiolabeled SAM will cause the scintillant within the bead to give off light. This is due to the transfer of energy to the scintillant by the β particles given off by the radiolabeled transferred methyl group. Any β particles from unbound molecules containing ^{3}H will dissipate into the solution, and are not within a sufficient proximity to the scintillation beads to produce any light. This prevents the assay from producing high background signal from unreacted radiolabeled SAM. The scintillation signal can be measured using a scintillation counter (*72*). One disadvantage of this is that there is no step for removal of residual ^{3}H-methyl SAM which may non-specifically bind to the scintillant beads. This can result in some background signal intensity, making quantification difficult. Such background may be eliminated by an additional wash step but this may wash away analyte and should not normally be needed (*73*).

Another method that is commonly used for detection of radioactivity incorporated into histones is a filter-binding assay. This consists of preparing reactions with a PKMT, histones, and radiolabeled SAM. After the reaction is finished, it is spotted onto nitrocellulose filter paper. The filter binds proteins such as histones and the rest of the reaction components, particularly unreacted radiolabeled SAM, can be washed away, allowing for quantification of radioactivity from methylated histones *via* phosphor screen imaging or scintillation counting (*74*). This has been used in a scaled-up form for medium throughput screening of PKMT inhibitors (*72*). These assays are quantitative because the radiolabeled SAM can be used to produce a standard curve (*73*).

Another consideration with radioactive assays is the use of ^{3}H-methylSAM versus ^{14}C-methylSAM. While either results in the transfer of radioactivity to substrate peptides, ^{14}C-

methylSAM is considered a more 'stable' interaction. This is because ^{3}H could passively interact with the substrate peptide by displacing hydrogen atoms, rather than directly through methylation. It has been shown that hydrogen ions from water in the reaction solution can exchange with the methylated-lysine methyl protons (*23*). This phenomenon is called solvent kinetic isotope effect (SKIE) and it has been studied using a proton inventory technique. In this technique, the enzymatic reaction is conducted in a solution that consists of 90% deuterated water (D_2O). The reaction is allowed to proceed and the SKIEs are determined by looking at the ratio of reaction rates in D_2O compared to those in H_2O. The proton inventory is similar, but works by measuring reaction rates at 0-90% D_2O solvent (*23*). This technique has yielded interesting results in regard to the deprotonation of lysine during methylation. This deprotonation is a key step in the reaction because it makes the N_ε on lysine into a nucleophile. Computational studies have suggested that this might be the function of a conserved tyrosine residue within the enzyme, but a proton inventory experiment showed that this may be a function of bulk solvent instead (*22*). Moreover, the relative energies of beta decay for ^{3}H-methylSAM *versus* ^{14}C-methylSAM and their respective half-lives are also considerations. While ^{3}H has a half-life of more than 12 years, ^{14}C has a half-life of 5000 years and nearly 10-fold higher energy for beta-decay. Finally it must be noted that the additional regulatory burden and the potential health risk of using radio-nuclides must be taken into consideration when choosing to use such assays.

1.5.2 Mass spectrometry methods

Mass spectrometry-based assays present an exciting alternative to those methods discussed above. Depending on the methodology mass spectrometry can be quantitative and very sensitive. Mass spectrometry is able to directly detect both substrates and products of reactions based on their mass and ionization in solution, allowing researchers to monitor every aspect of a reaction

of interest at once. Once a mass spectrometry assay has been validated, it can be run for a long time with minimal supervision. This means that many samples can easily be tested at once.

Many liquid chromatography and tandem mass spectrometry (LC-MS/MS) assays used in the field of epigenetics fall into the category of proteomics analyses. These generally involve the detection of PTMs on short peptides by measuring an increase in the mass of the peptide corresponding to the PTM of interest. In methylation studies, this corresponds to an increase of 14 Da for every methyl group added (*72*). Some difficulties inherent in these analyses include the difficulty discerning analyte identity of geometric isomers, for example symmetric and asymmetric dimethyl arginine. However, greater selectivity can be achieved by LC-MS/MS. This first separates analytes according to their elution time by chromatography. The samples are then differentiated *via* mass-to-charge ratio (*m/z*) as the precursor ion in the first mass spectrometer, then by the *m/z* ratio of the characteristic product ions produced upon fragmentation in a collision chamber. Such techniques potentially allow unambiguous differentiation of compounds that have similar retention times and even identical molecular weights since these compounds often fragment into product ions with differing (*m/z*) (*11*, *75*). A further technique used to simplify these analyses is chemical derivatization of the analytes. An example of this technique is using propionate anhydride to propionylate unmodified lysine residues within sample peptides. This has the advantages of increasing analyte hydrophobicity, allowing for easier peptide separation using reverse phase HPLC (*76*). It also allows for more uniform digestion of sample peptides using the protease trypsin. This is because with all of the lysine residues in a given peptide chemically derivatized trypsin can only cleave at arginine residues, usually resulting in fragments which are suitable for LC-MS/MS (*77*). However this practice may have undesirable consequences, such as displacement of the epigenetic marks which are of interest. Derivatization

could also add the same mass as an epigenetic PTM to analytes, which make it impossible to differentiate a naturally-occurring PTM from derivatization. For example, a propionylated lysine residue has the same mass as a trimethyllysine residue. A technique which some have used in place of derivatization is to treat histone samples with a protease capable of hydrolyzing proteins to completion, leaving only individual amino acids (*78*, *79*). This allows for easier identification of parent ions, and the detected amount of modified residues can be normalized to the amount of unmodified residue. This was used by *Edrissi et al* in an attempt to quantify formyl lysine and other PTMs within TK6 cells (*78*). After complete enzymatic hydrolysis of histones by pronase (a mixture of proteases from *Streptomyces griseus*), total formyl lysine and other PTM were quantified via LC-MS/MS using deuterated internal standards (*78*). However, the assay as it was performed was not validated. Lillico *et al* developed and validated a similar technique to quantify many more modifications to histones simultaneously, each as an individual modified amino acid (*79*). This study showed that the method by *Edrissi et al* was not quantitative under the conditions they used and that acid hydrolysis was superior to enzymatic hydrolysis in all cases except acetyl lysine. A disadvantage inherent in this technique is that it is then impossible to determine the exact location of the PTMs that were detected within the protein sequence. The usefulness of mass spectrometry techniques in the context of studying lysine methyltransferases, along with limitations inherent to these techniques, will be discussed further in other sections of this work.

Many mass spectrometry techniques from the field of proteomics have been used to study epigenetic enzymes. In particular, these techniques have been used to isolate modified peptides within cell lysate samples and to determine the precise location of modifications within proteins. Acuto *et al* observed that tryptic peptides containing at least one RMe usually have a basic

isoeletric point (pI), so they can be separated from other peptides via isoelectric focusing (*80*). Therefore, 2D-PAGE can be used to separate out methylated tryptic peptides (*81, 80*). The bands corresponding to these peptides of interest can then be isolated from the gel and analysed with LC-MS/MS. However, proteomic techniques were originally designed for detection of target proteins in samples. As such, while these techniques can be quite sensitive for detection of target proteins they are not designed to be quantitative. A major impediment to quantitation using proteomic MS techniques is their inability to cover the entire sequence of many proteins. This means that some sites of PTMs cannot be detected, much less quantified.

There are two general approaches to the use of MS in proteomics, the bottom-up approach and the top-down approach (*82*). In the bottom-up approach, a protein of interest is isolated and fragmented into peptides, usually *via* enzymatic proteolysis with trypsin. The mass spectra of these peptides are then used as a 'fingerprint' to identify the protein of interest in samples (*82*). With mass spectrometers capable of fragmentation, these peptides can be further fragmented in order to produce diagnostic ions for detection of the protein of interest. This reduces the number of peptides required for protein identification (*82*). However due to the initial fragmentation of the target protein into peptides required by bottom-up proteomics, it is impossible to cover 100% of the protein sequence (*83*). In the context of epigenetics research, this means that studies conducted using this technique may only present a partial description of PTMs found on target peptides, potentially missing PTMs on segments of a protein. In the top-down approach, intact protein ions are detected and fragmented by the mass spectrometer (*84*). This approach requires a mass spectrometer with high resolving power to accurately detect ions with masses as large as intact proteins, and is therefore limited to specialized laboratories (*85*). The top-down approach has the advantage of being able to cover most of a protein's sequence if

the protein is below 30 kDa in mass (*83*). As the mass of the target proteins increases, sequence coverage decreases. Tsybin *et al* reported 20% sequence coverage of a 150 kDa antibody using a top-down approach with a quadrupole time-of-flight mass spectrometer (*86*). While top-down mass spectrometry techniques will likely become able to achieve greater sequence coverage as mass spectrometer technology advances, it currently leaves these techniques unsuitable for quantitation of PTMs in epigenetic research.

1.5.3 Methods using antibodies

Antibody-based assays are also commonly used in the field of epigenetics, such as western blots, ELISA, and antibody-based pulldown of peptides followed by identification *via* LC-MS (*87–89*). The binding of an antibody to a modified residue on the protein substrate allows for the PTM to be detected. Detection can be achieved either by the production of a signal, such as fluorescence or light in an enzyme-coupled assay, or by subsequent LC-MS/MS detection of peptides to which the antibodies have bound (*72*, *90*). The antibodies used in these assays are typically raised against specific PTMs or SAH (*72*, *91*). These antibodies have been produced by introducing peptide libraries which contain epigenetic marks in various contexts. For example, Guo *et al* produced antibodies against RMe using peptides containing RMe in a central position and within an RGG motif (*88*). This is done to increase the specificity of the antibodies. It is hoped that antibodies produced in this way will strengthen antibody-based assays, as previous work with antibodies against epigenetic marks has shown that they are often limited in their selectivity (*81*). These antibodies can also be used for immunoprecipitation, which enriches proteins of interest from total cell protein samples. The antibodies are conjugated to a solid substrate, often beads coated with protein A, and the proteins of interest are bound by the antibodies to remove them from solution (*88*). Co-immunoprecipitation is very similar, but

has the goal of pulling down a protein of interest as well as other proteins with which the protein of interest associates (*81*).

These types of assays have been used to great effect for many years in many different sections of biological research. However, in the context of epigenetics there are problems inherent in these assays which stem from the antibodies themselves. Antibodies are prone to being highly variable in their efficacy from one batch to another, resulting in the need to validate antibodies whenever a new batch is to be used, and whenever an antibody is going to be used in a new assay (*87*, *92*). Furthermore, in the context of histone PTMs the differences that these antibodies are being designed to detect are quite small. For example, an antibody which is specific for H3K9 trimethylation would have to be able to bind to the trimethyl mark while also being unable to bind the mono- or dimethyl lysine which may be present at H3K9 . In practise, studies have shown that cross-reactivity occurs with these antibodies. Bock *et al* used peptide arrays containing peptide sequences corresponding to the N-terminal tail sequences of histones H3, H4, H2A, and H2B with various common epigenetic marks in order to test the binding specificity of a series of antibodies which were designed to bind specific PTMs (*92*). They observed several false positives, showing that these antibodies were binding non-specifically in some cases. They also showed several false negatives, particularly in cases where the histone peptide contained more than one epigenetic modification. Their conclusion was that the presence of a modification may block an antibody from binding where it is supposed to. The propensity of antibodies to bind non-specifically to epigenetic PTMs, along with the potential for PTMs to block an antibody from binding, decrease the usefulness of antibody-based assays in epigenetic studies.

1.6 Future considerations

One finding that appears consistently in the literature is that PKMTs can methylate more than just one or two residues within histones and possibly on multiple non-histone proteins. SET7/9 has been shown to methylate residues within transcription factors such as TATA box-binding protein associated factor 10, leading to a higher affinity of the protein for RNA polymerase II (*34*). Non-histone targets have been found for many other PKMTs as well (*16*). More and more relationships between different epigenetic marks are being found. For example, research has shown that methylation of H3R2 can inhibit trimethylation of H3K4 (*93*). Future studies will need to identify such interactions between marks so as to clarify context-based effects of these additions. Enzyme specificity is another field that warrants further research. Previous work with PRMTs has shown some flexibility at the active sites of these enzymes (*66*). Further exploration of such flexibility in substrate binding by PKMTs could help us to better understand methylation and its biological function.

1.7 Hypothesis and aims

We hypothesize that the activity of the methyltransferase SET7/9 can be quantified using LC-MS/MS to measure the total methylated lysines liberated from methylated products and SAH which is the co-product of the reaction. We expect to only see KMe formation as a result of SET7/9 activity, and we expect that the amount of SAH produced equals the amount of KMe produced. Such LC-MS/MS techniques will allow determination of enzymatic kinetic parameters.

Based on the hypothesis, the primary aim of this research is to determine the apparent kinetic parameters K_m^{app} and V_{max}^{app} of SET7/9 for both of its substrates, SAM and histone H3.

Once those values are found, the secondary aim is to determine substrate specificity and potential cross-talk between different PTMs that occurs on histones to determine if substrate histones with prior modifications affect SET7/9 K_m and V_{max}.

2.0 Materials and methods

2.1 SET7/9 expression and purification

A pET28a-LIC plasmid containing a gene coding for the full-length SET7/9 enzyme with a N-terminal 6xHis tag was obtained from Dr. Masoud Vedadi, of the Structural Genomics Consortium in Toronto, Ontario (**Figure 6**). This plasmid was transformed into chemically competent *E. coli* BL21(DE3) pLysS cells (Invitrogen). 106 ng of the plasmid sample was added to 50 µL of cells, incubated on ice for 30 min, then heatshocked for 30 s in a 42°C water bath. After heatshock, the cells were put back on ice and 250 µL of room temperature SOC medium were added. The mixture was then incubated at 37°C while being shaken at 225 rpm for 1 hour. The mixture was then split in two and plated on two different LB solid medium plates containing 34 µg/mL chloramphenicol and 50 µg/mL kanamycin in order to isolate individual colonies of transformants. The plates were incubated overnight at 37°C. An individual colony was used to inoculate 100 mL of 2X LB nutrient broth containing 34 µg/mL chloramphenicol, 50 µg/mL kanamycin, and 1% glucose. This culture was incubated at 37°C while being shaken at 200 rpm overnight, and centrifuged at 15000x g for 20 min the following morning to inoculate two 1 L cultures with cell pellet. The 1L cultures were grown in 2X LB nutrient broth containing 34 µg/mL chloramphenicol, 50 µg/mL kanamycin, and 1% glucose. These cultures were incubated at 37°C while being shaken at 200 rpm for approximately 6 hours, or until an OD600 between 1 and 1.5 was reached. The cultures were then kept on ice for approximately 1.5 hours. Induction of plasmid expression was then achieved by addition of 1 M IPTG to a final concentration of 1 mM. The cultures were then incubated overnight while being shaken at 200 rpm, approximately 16 hours. During this incubation stage, the heater in the incubator was turned off and 2 flasks filled with ice were placed in the incubator with the cultures. After ~1 hour of incubation, any

melted ice was replaced. This way the temperature remained between 20 – 23°Cfor the first hour post induction of expression. The temperature read 28°C at the end of the incubation period.

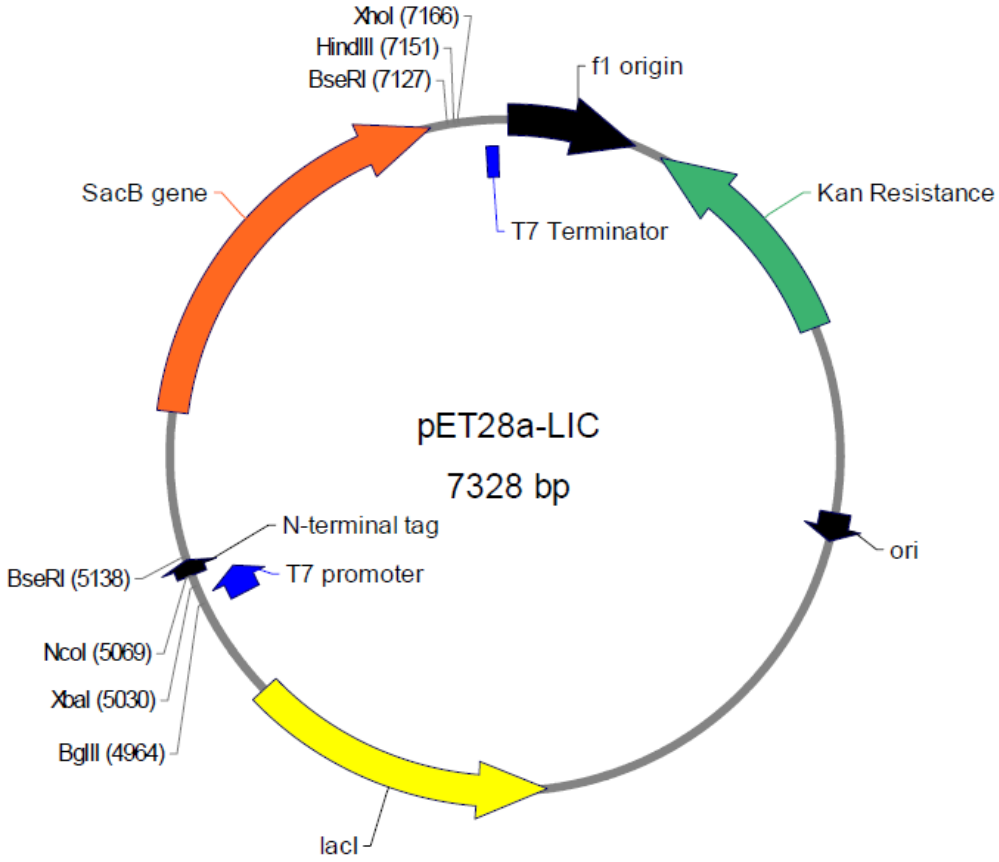


Figure 6. Plasmid map of pET28a-LIC showing selectable markers, expression system, and multiple cloning site (*94*).

Bacterial cultures were centrifuged at 15000 x g for 20 min. at 4°C and the growth medium was poured off of the cell pellet. The pellet was resuspended in lysis buffer (20 mM $NaHPO_4$, 100 mM NaCl, 1 mM PMSF, 7 mM β-mercaptoethanol, 0.1% Triton X-100, 0.5 mg/mL lysozyme) with an EDTA-free protease inhibitor tablet (Life Technologies) and 10 U DNase, and sonicated on ice for seven 30 s pulses of sonication with a 30 s pause on ice between each pulse. The sample was then centrifuged at 15000 x g for 15 min at 4°C. The supernatant

was removed and passed through a 0.45 µm filter (Whatman). The sample was then loaded onto a 5 mL HisTrap FF immobilized metal ion affinity chromatography column (GE Healthcare) on an AKTÄ Purifier FPLC at a flow rate of 1 mL/min. The column was then washed with buffer for 10 column volumes at 3 mL/min. The protein of interest was eluted at 3 mL/min in a step-wise imidazole gradient starting at 5 mM imidazole and ending at 400 mM imidazole. Each step in the gradient increased the concentration of elution buffer (B) by 2.5%, giving an increase of 10 mM imidazole per step. The first step was from 0% elution buffer B to 2.5%, followed by stepping up to 5%, 7.5%, 10%, and 12.5%. After these steps, a final fraction was collected at 100% B (400 mM imidazole) to ensure that no protein remained on the column. Six 10 mL fractions were collected during the gradient. The fraction with the highest UV absorbance measured during elution was then loaded onto a HiPrep 26/10 desalting column (GE Healthcare) to exchange the buffer and remove salt and imidazole from the protein of interest. After being loaded onto the column at a flow rate of 1 mL/min, the sample was eluted in 50 mM Tris buffer (pH 7.5). The resulting 15 mL fraction was then concentrated down to ~2 mL using a Spin-X UF20 concentrator with a 10 kDa molecular weight cutoff (Corning).

The concentrated fraction was then run on a 10% SDS-PAGE gel to both verify that the protein, SET7/9, had been produced and purified and to determine the concentration of the protein via densitometry. This was done using 10 mm mini Tris-glycine gels. These were prepared according to a standard protocol (*95*). Samples were diluted in sample dilution buffer (4xTris-Cl pH 6.8; 30% glycerol, 1% SDS, 0.28 M BME, 0.012% bromphenol blue, and H_2O), then loaded into wells. Gels were run at 100V for ~45 min (*95*).

Presence of SET7/9 was determined by comparison to a PageRuler standard protein weight ladder (Invitrogen). Densitometry was performed by running the fraction containing

SET7/9 on a gel alongside varying known concentrations of BSA (Sigma-Aldrich). SDS-PAGE gel imaging was done using an Alpha-Innotech FluorChem FC2 imaging system, and ImageJ image processing and analysis software was used to plot a BSA standard curve for determining protein concentration by densitometry (*96*). Glycerol was then added to the fraction to a final concentration of 5% and the fraction was split into 200 µL aliquots for storage at -80°C.

2.2 DOT1L expression and purification

A pET28-MHL plasmid containing a gene coding for residues 1-420 of the DOT1L enzyme with a N-terminal 6xHis tag was also obtained from Dr. Masoud Vedadi, of the Structural Genomics Consortium in Toronto, Ontario. This plasmid was transformed into chemically competent *E. coli* BL21(DE3) pLysS cells (Invitrogen). Bacterial transformation, growth of cultures, induction of protein expression, lysis of bacteria, and purification were performed as described above for SET7/9 purification, with the exception that the fraction collected from the HisTrap column was desalted into HEPES buffer (50 mM, pH 8). After the desalting step, the collected fraction was incubated on ice for 1 hr with 10 U of DNase. This was done to hydrolyse any DNA which was co-purified bound to DOT1L (*53*). The fraction was then loaded onto a HiTrap SP strong cation exchange column. The protein was eluted with HEPES buffer containing 1M NaCl over a 50 min gradient. Any fractions collected were precipitated by addition of TCA to a final concentration of 12%. After the TCA was added, the fractions were incubated on ice for 15 min at which point they were centrifuged at 12000 x g for 15 min. The supernatant was removed using a Pasteur pipette and the pellet was washed with ice cold acetone. This step was repeated once, then the precipitated proteins were resuspended in Tris buffer and run on a 12% SDS-PAGE gel along with BSA in 0.625 µM, 1.25 µM, 2.5 µM, and 5

μM concentrations SDS-PAGE was then performed as described above to confirm the presence of the protein and to measure its concentration *via* densitometry.

2.3 *in vitro* methylation reactions

Methylation reactions have a final volume of 100 μL and consist of enzyme, buffer, protein substrate, and SAM (Cayman Chemical Co.). The enzymes used were SET7/9 (NEB), G9a (NEB), expressed SET7/9, and expressed DOT1L. The reactions were prepared in 0.300 mL HPLC inserts and incubated for 60 min, an incubation time which has been shown to be within the linear activity range for SET7/9 as well as many other SET domain methyltransferases and has been widely used in *in vitro* assays with SET7/9 (*38*, *97*, *90*, *98*, *99*). Experiments using SET7/9 (NEB), G9a (NEB), and DOT1L were incubated at 37°C, while experiments using expressed SET7/9 were incubated at room temperature (*53*, *100*). Different buffers were used in different experiments. For the experiments with G9a a buffer consisting of 0.5 mM DTT, 100 mM NaCl, and 50 mM Tris-HCl, pH 9 was used. For the experiments with SET7/9 (NEB), a buffer with 0.25 mM BME, 2 mM NaCl, and 100 mM ammonium acetate, pH 9 was used. For the experiments with expressed SET7/9 and DOT1L, the buffer consisted of 50 mM ammonium acetate, 0.25 mM DTT, 0.01% TritonX-100, pH 9. Protein substrates used were histone H3 peptides (Epigentek), full-length histone H3 (Cayman Chemical Co.), and chicken nucleosomes (Epicypher). The concentrations of histone and SAM stocks were determined by UV absorbance and calculated from the extinction coefficients (*101*). The components of the methylation reaction were all combined directly in HPLC inserts to limit mechanical loss of analytes. Two compounds, 2-chloroadenosine and 2-aminobutyric acid, are also added as internal standards in concentrations of 0.25 μM and 0.17 μM respectively. In all experiments, the enzyme was added to a final concentration of 100 nM. MilliQ water is added to bring the final reaction volume up to

100 µL. The samples are incubated at room temperature for 1 hr, at which point the methylation reaction is stopped by incubation at 75°C for 5 min. The samples are then split into 2 portions for analysis by two different LC-MS/MS assays; one for quantification of post-translational modifications to proteins and one for quantification of SAH. All samples are then dried completely in a Savant SPD1010 SpeedVac Concentrator for 50 min with no heat (Thermo Scientific).

2.4 Post-translational modifications assay

The PTM assay was performed as previously described, using a Shimadzu LC-MS 8040 with the addition of 2-aminobutyric acid as an internal standard that was added directly to the reaction buffer (*79*). Samples for use in this assay must first undergo acid hydrolysis in order to break any proteins into their component amino acids. This is achieved by placing all HPLC inserts containing the dried samples taken from the methylation experiments into ELDEX vacuum hydrolysis vessels with 250 µL of 6N HCl, evacuating the air from the vessels with a vacuum pump, and incubating them at 110°C for 24 hrs. The samples are then removed from the hydrolysis vessels, dried again in a Savant SPD1010 SpeedVac concentrator for 50 min with no heat, and reconstituted in 60 µL of 0.05% formic $acid_{(aq)}$ for LC-MS/MS analysis. For separation of analytes by LC, a Primesep 200 (Sielc) mixed function cation exchange column was used with a pH gradient for elution. The mobile phases were (A) 0.05% formic $acid_{(aq)}$ and (B) 1% formic acid in 50% $acetonitrile_{(aq)}$. The assay has a 10 minute per sample runtime and the MRM transitions for the analytes are listed in **Table 1**. Of the analytes included in the assay, only K, KMe, KMe2, and KMe3 were quantified because they were of particular interest for the enzymes being studied. The other MRM transitions were kept in the assay for qualitative detection. Quantification was done with an 11-point standard curve using known concentrations of KMe,

KMe2, and KMe3, and K (Epigentek). The curve was prepared by making a sample with 5 µM KMe, KMe2, and KMe3, and 50 µM K and then using that sample as the top level of a series of 1 in 2 dilutions. The internal standard, 2-aminibutyric acid, was then added to a concentration of 1.7 µM.

Table 1. MRM transitions and collision energies for PTM assay analytes

Analytes	**Ionization**	**Retention Time (min.)**	**MRM Transition (m/z)**	**Collision Energy (eV)**
Monomethyl lysine	+	3.54	161.1 →84.1	-17.0
Dimethyl lysine	+	3.68	174.7 →83.9	-22.0
Trimethyl lysine	+	3.81	189.0 →84.1	-22.0
Lysine	+	3.45	146.9 → 130.0	-14.0
Acetyl lysine	+	2.22	189.2 → 126.0	-14.0
Arginine	+	3.47	175.2 → 70.1	-25.0
Monomethyl arginine	+	3.61	188.8 → 70.0	-24.0
Asymmetric dimethyl arginine	+	3.95	203.0 → 45.9	-22.0
Symmetric dimethyl arginine	+	3.32	202.8 → 171.9	-14
2-aminobutyric acid	+	1.52	104.1 →58.1	-26.0

2.5 SAH assay

The rationale behind this assay is that for every methylation event which occurs during the reaction period, one molecule of SAH is produced. Therefore, measuring SAH production should give a value which is representative of the methylation that has taken place. An advantage of this assay is that it will detect any SAH produced regardless of where the methylation is taking place, meaning that it does not have to simultaneously detect and quantify multiple analytes. This assay was performed as described previously with the addition of 2-chloroadenosine as an internal standard (*102*). Samples for use in this assay are reconstituted in 40 µL of 0.1% formic $acid_{(aq)}$ for LC-MS/MS analysis. A Waters C18 UPLC™ column was used with mobile phase (A) 0.1% formic $acid_{(aq)}$ and (B) 0.1% formic acid in 50% $acetonitrile_{(aq)}$. This assay has a 9 minute runtime, and the MRM transitions for the analytes are listed in Table 2.

Table 2. MRM transitions and collision energies for SAH assay analytes

Analytes	**Ionization**	**Retention Time (min.)**	**MRM Transition (m/z)**	**Collision Energy (eV)**
S-adenosyl-L-homocysteine	+	1.72	385.1 → 136.1 385.1 → 134.1	-25.0 -20.0
2-chloroadenosine	+	5.32	302.0 → 170.0	-17.0

2.6 Treatment of data

Data from both LC-MS/MS assays are plotted as reaction velocity *vs* the concentration of varied substrate. Values from the negative control samples were taken as background and subtracted from experimental values. The plots were fit to a rectangular hyperbola *via* non-linear regression in SigmaPlot to determine the apparent K_m and V_{max} (K_m^{app} and V_{max}^{app}) of each enzyme with regard to substrates, SAM and histone H3. The curve to which the data were fit is described by an equation in the form of the Michaelis-Menten equation (Equation 1),

$$v = \frac{V_{max}[S]}{K_m + [S]} \qquad \textbf{Equation 1}$$

where [S] is the concentration of the substrate which was varied during the experiment. This is an acceptable method of analysis because the equation describing the velocity of a bisubstrate enzymatic reaction simplifies to an equation taking the same form as the Michaelis-Menten equation when one substrate is at a fixed, saturating concentration (see below). The reaction velocity, given the reaction diagram in **Figure 5**, must take into account both the forward and reverse reactions for both the binding of substrates and for the products. This means taking into account 9 separate equilibria, their constants, and the V_{max} for both the forward and reverse reactions, V_{max1} and V_{max2} respectively. One also needs to take into account the overall equilibrium of the reaction and its equilibrium constant, K_e. For all dissociation constants, we consider if the dissociation is from the enzyme with the two products or two substrates bound, in which case the designation α comes before the constant (i.e. αK_{AX}). Dissociation of one product or substrate to form free enzyme and product or substrate have no such designation. With all of these considerations in mind, the complete reaction velocity can be described by **Equation 2**.

$$v = \frac{V_{max1}V_{max2}([B][AX]-\frac{[BX][A]}{K_e})}{\alpha K^{AX}K^{B}V_{max2}+\alpha K^{B}V_{max2}[AX]+\alpha K^{AX}V_{max2}[B]+\frac{\alpha K^{A}V_{max1}[BX]}{K_e}+\frac{\alpha K^{BX}V_{max1}[A]}{K_e}+V_{max2}[B][AX]+\frac{V_{max1}[BX][A]}{K_e}}$$

Equation 2

This equation is far too complex and difficult to deal with experimentally and equally difficult to fit to a curve, even using sophisticated software capable of fitting multiple parameters simultaneously. In order to simplify **Equation 2** we will assume that the concentrations of the products BX and A and the initial rate of the reverse reaction are zero. We will further reduce the reaction diagram in **Figure 5** to consider the reaction to form E·A·BX to be slow with respect to substrate binding (rapid equilibrium) as others have invoked (*103*). With these assumptions, the initial rate of the forward direction reduces to **Equation 3**

$$v = \frac{V_{max}[AX][B]}{\alpha K^{AX}K^{B}+\alpha K^{B}[AX]+\alpha K^{AX}[B]+[AX][B]}$$ **Equation 3**

where V_{max} is the maximum velocity of the forward direction. αK^{AX} is a dissociation constant describing the binding of AX to the enzyme+B complex and αK^{B} is a dissociation constant for dissociation of B from the enzyme+B complex.In order to observe what happens to **Equation 3** at saturating concentrations of B we can simplify **Equation 3** by dividing it by the concentration of B, yielding **Equation 4**

$$v = \frac{V_{max}[AX]}{\alpha K^{AX}\left(1+\frac{K^{B}}{[B]}\right)+[AX](1+\frac{\alpha K^{B}}{[B]})}$$ **Equation 4**

At high saturating concentrations of B, $K^B/[B]$ and $\alpha K^B/[B]$ both approach zero, moreover as B becomes large and saturating $\alpha K^{AX}=K_m^{A\ ap,}$ which is the corresponding apparent Michaelis constant.

$$v = \frac{V_{max}^{app}[AX]}{K_m^{AX.app}+[AX]} \qquad \textbf{Equation 5}$$

This also holds true for saturating concentrations of B. **Equation 3** simplifies differently when the concentration of AX is saturating. Once again $\alpha K^B=K_m^{B\ app}$ which is the corresponding apparent Michaelis constant. To derive, this one divides equation 2 by [AX] resulting in

Equation 6

$$v = \frac{V_{max}[B]}{\alpha K^B\left(1+\frac{K^{AX}}{[AX]}\right)+[B]\left(1+\frac{\alpha K^{AX}}{[AX]}\right)} \qquad \textbf{Equation 6}$$

Equation 6 simplifies further at fixed, high saturating concentrations of AX, where the terms $K^{AX}/[AX]$ and $\alpha K^{AX}/[AX]$ both approach zero. This gives **Equation 7.**

$$v = \frac{V_{max}^{app}[B]}{K_m^{B.app}+[B]} \qquad \textbf{Equation 7}$$

This method of data analysis has been widely employed in this field and is recommended over methods such as the double reciprocal plot, which can introduce further error into calculations (*67*, *100*). In these calculations, SAM can be thought of as AX and histone H3 can be thought of as B, although in principle it does not matter as the mechanism is random. Treating the data in this manner involves making a few assumptions about the reaction taking place. Firstly, we

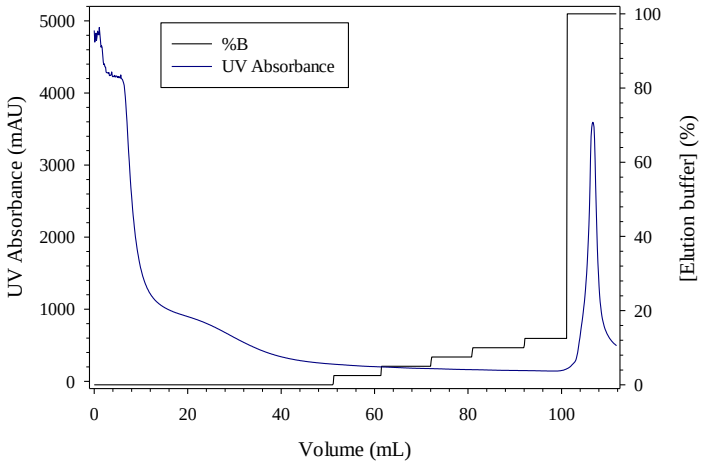
assume that SET7/9 is following a random ordered mechanism as shown in **Figure 5**. Next, we are assuming that one of the substrates in each experiment is fixed at a saturating concentration, so that the steps leading to **Equation 5** and **Equation 7** are valid. Assuming that no product is being formed, our equations do not have to take the dissociation of A and BX from the enzyme into account.

3.0 Results

3.1 SET7/9 expression and purification

The bacterial yield (wet weight) was 8.0 g/L of *E. coli* when centrifuged. After lysis of the cells, IMAC was performed and the His-tagged protein of interest was eluted on a step-wise gradient and during each step a fraction was collected. The absorbance peak observed during the elution at 100% B suggests that the protein of interest was collected in elution volume 102-112 mL (**Figure 7**). This fraction was loaded onto a HiPrep 26/10 desalting column (GE Healthcare) which was equilibrated with 50 mM Tris buffer (**Figure 8**). As shown in **Figure 8**, the separation of the peaks in absorbance and conductivity shows that the protein has been separated from any salt and imidazole which was left over from the His-tag affinity column purification. The collected desalted fraction was then run on a 10% SDS-PAGE gel along with BSA in 0.625 µM, 1.25 µM, 2.5 µM, and 5 µM concentrations (**Figure 8B**). The molecular weight of the protein of interest was determined using ProtParam to be 41.6 kDa. The BSA bands' intensities were plotted in ImageJ image processing software and used to generate a standard curve to which the intensity of the SET7/9 bands were compared in order to determine the concentration of SET7/9 in the purified sample. The concentration of the purified SET7/9 was determined to be 9.05 µM.

A



B

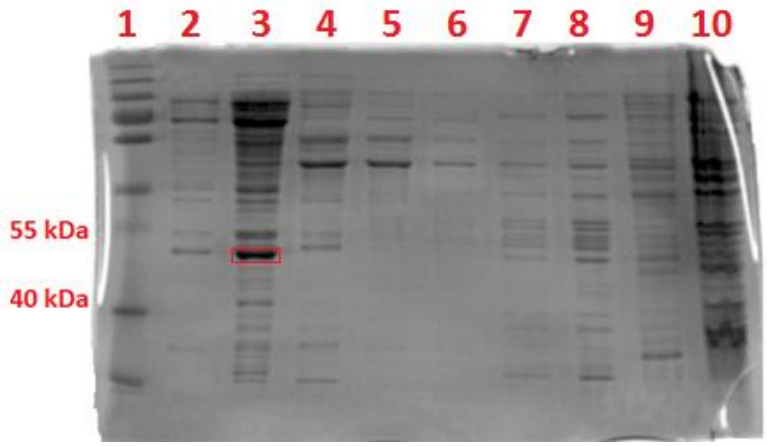


Figure 7. (A) FPLC chromatogram depicting the purification of the His-tagged SET7/9 from bacterial lysate using a HisTrap FPLC column (B) 10% SDS-PAGE gel showing fractions collected from the HisTrap column during FPLC of SET7/9. PageRuler protein standard is in lane 1, elution volume 92-102 mL in lane 2, elution volume 102-112 in lane 3, elution volume 52-62 mL in lane 4, elution volume 62-72 mL in lane5, elution volume 72-82 mL in lane 6, elution volume 82-92 mL in lane 7, what was collected during sample loading in lane 8, column wash in lane 9, and unpurified sample in lane 10. The red box in lane 3 is outlining the band corresponding to SET7/9.

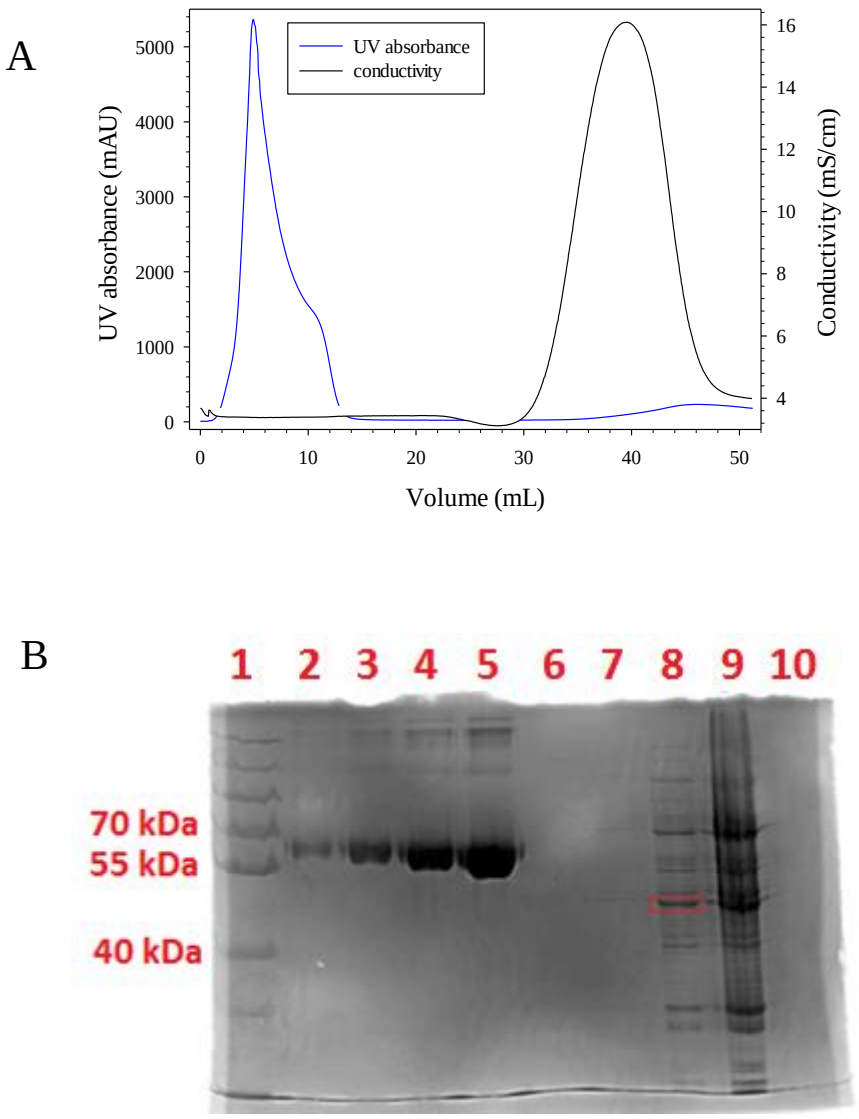


Figure 8. (A) FPLC chromatogram showing the elution of SET7/9, shown in the peak in absorbance, from a desalting column, followed by the salt and imidazole, shown in the peak in conductivity. (B) 10% SDS-PAGE gel showing PageRuler protein standard in lane 1, BSA standards in lanes 2-5, and purified protein sample in lanes 7-9. Different amounts of protein were run in these lanes. Lanes 6 and 10 were left blank. The band containing the protein of interest has been highlighted in lane 8.

3.2 DOT1L expression and purification

As with SET7/9, the absorbance peak observed during the elution at 100% B suggests that the protein of interest was collected in elution volume 112-125 mL (**Figure 9**). **Figure 10** shows that the protein has been separated from any salt and imidazole which was left over from the His column purification. **Figure 11** shows that the protein of interest does not appear to bind to the cation exchange column. The molecular weight of the protein of interest was determined using ProtParam to be 47.8 kDa. The BSA band intensities were plotted in ImageJ image processing software and used to generate a standard curve to which the intensity of the DOT1L bands were compared in order to determine the concentration of DOT1L in the purified sample. The concentration of the purified DOT1L was determined to be 3.3 µM. The purified enzyme was then used for an *in vitro* methylation experiment to show the activity of the resulting protein. This experiment used 100 nM DOT1L and varied the concentration of SAM while holding histone H3 at 10 µM. The results showed no DOT1L activity, with no apparent difference between control and experimental samples (**Figure 12)**.

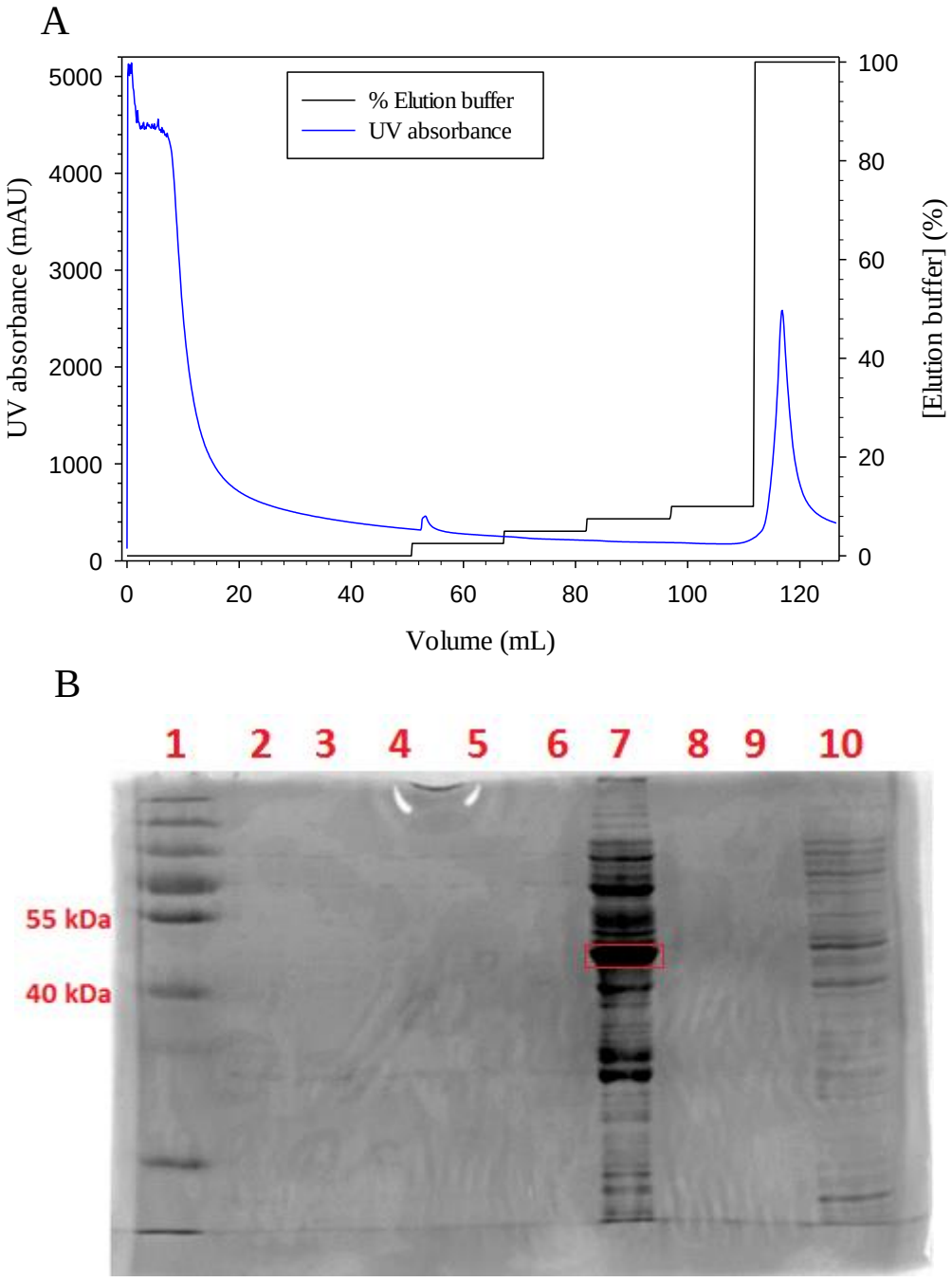


Figure 9. (A) FPLC chromatogram showing the purification of His-tagged protein of interest from bacterial lysate using a HisTrap FPLC column. The bacteria had been transformed with a plasmid containing the gene for the first 420 amino acids of DOT1L. (B) 12% SDS-PAGE gel showing the fractions collected from purification of DOT1L-containing lysate with a HisTrap

column; PageRuler protein standard is in lane 1, lane 2 is blank, elution volume 49-68 mL is in lane 3, elution volume 68-82 mL is in lane 4, elution volume 82-96 mL is in lane 5, elution volume 96-112 mL is in lane 6, elution volume 112- 125 is in lane 7, lanes 8 and 9 are blank, and the column wash is in lane 10. The outlined band in lane 6 corresponds to the expressed DOT1L (1-420).

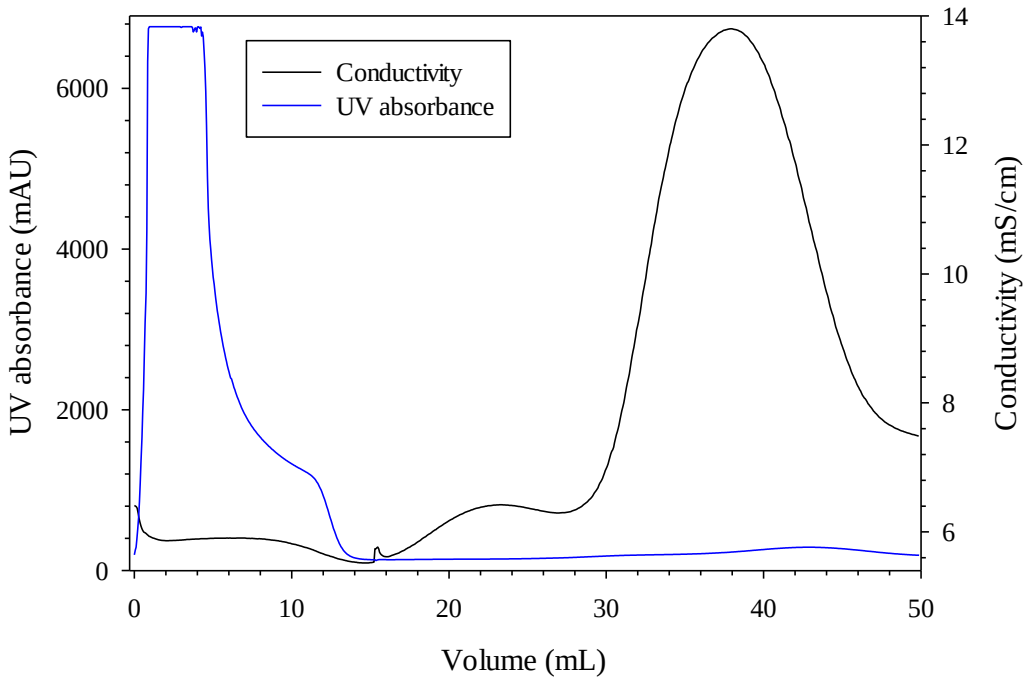


Figure 10. FPLC chromatogram showing the elution of DOT1L, shown in the peak in UV absorbance, from a HiPrep 26/10 desalting column, followed by the salt and imidazole, shown in the peak in conductivity.

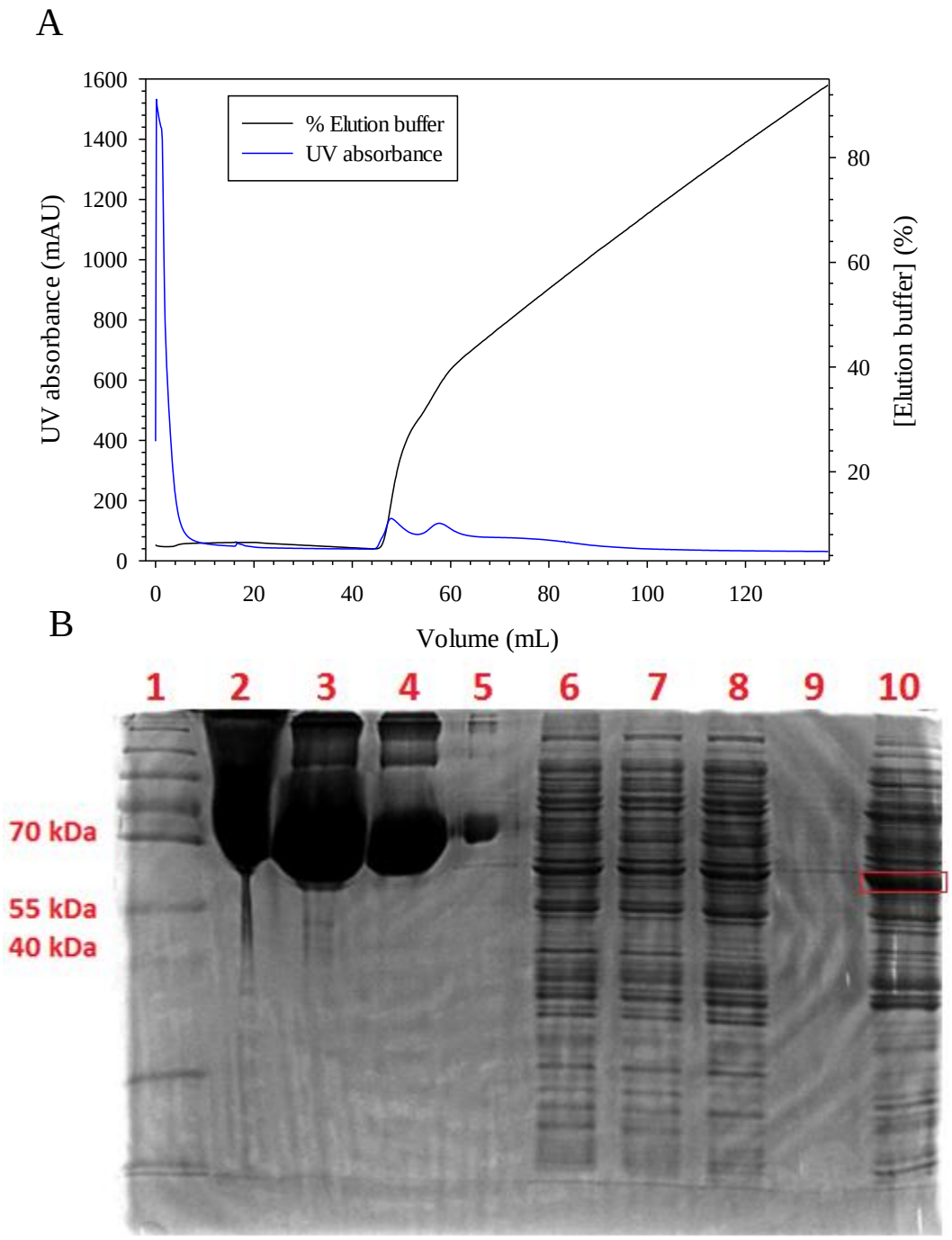


Figure 11. (A) FPLC chromatogram showing the elution of DOT1L, shown in the peak in absorbance, from a HiTrap SP strong cation exchange column. (B) A 12% SDS-PAGE gel containing PageRuler protein standard in lane 1, varying BSA concentrations in lanes 2-5, and varying concentrations of eluent from the SP column in lanes 6-10.

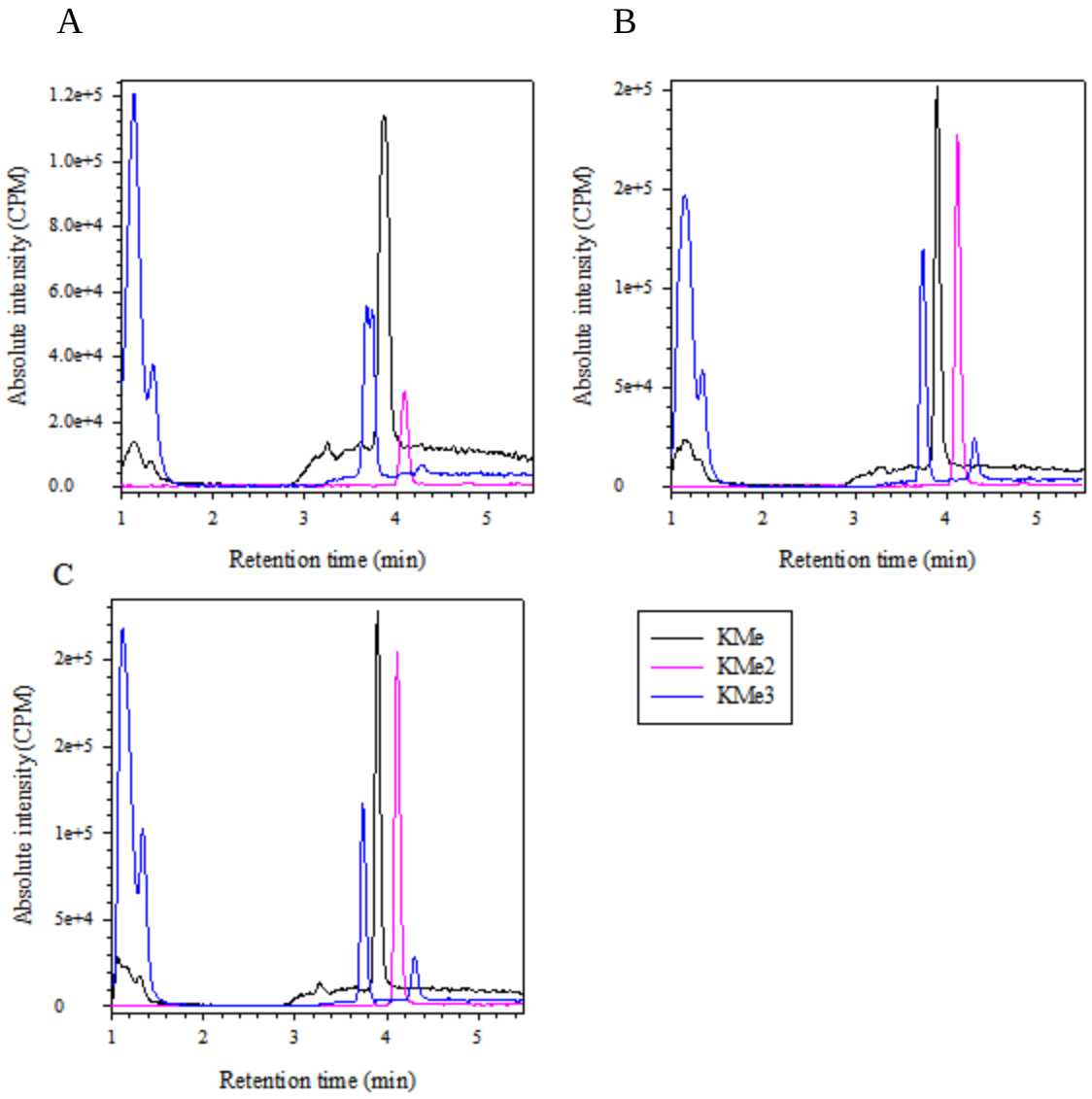


Figure 12. DOT1L PTM chromatograms showing (A) no histone control, sample contained only buffer, 0.1 µM DOT1L, and 10 µM SAM (B) no SAM control, sample contained only buffer, 0.1 µM DOT1L, and 10 µM histone (C) DOT1L activity test, sample contained buffer, DOT1L, histone and SAM. 0.1 µM DOT1L, 10 µM histone H3, and 10 µM SAM were used.

3.3 G9a methylation experiment

In vitro methylation reactions were performed using G9a purchased from New England Biolabs. The concentration of histone H3 was held at 5 μM and the concentration of SAM was varied. The results from this experiment were inconclusive, with no difference between the experimental samples and the negative controls (**Figure 13**). This experiment used a Tris-based buffer, and we suspected that the buffer may be affecting the background signal in the MS. We tested Tris, HEPES, and ammonium acetate buffers to determine their effect on background signal in the MS (**Figure 14).** These results did not explain the lack of methylation observed with G9a, however they did show that certain buffers contribute to background signal in these assays.

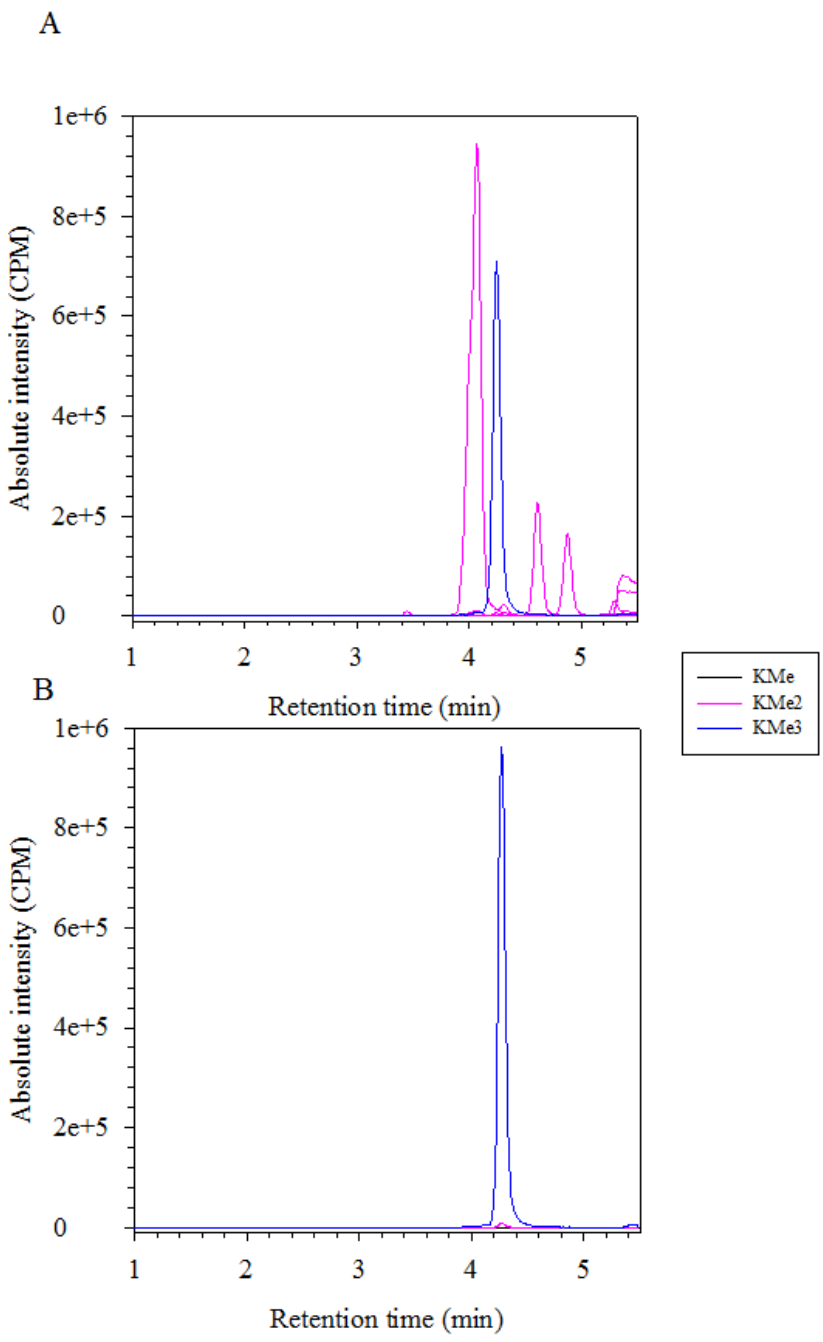


Figure 13. LC-MS/MS chromatograms showing the PTMs KMe, KMe2, and KMe3 produced by G9a under the following experimental conditions (A) no enzyme control, sample contained only buffer, 25 µM SAM, and 5 µM histone H3 (B) 0.1 µM G9a with 5 µM SAM and 5 µM histone H3.

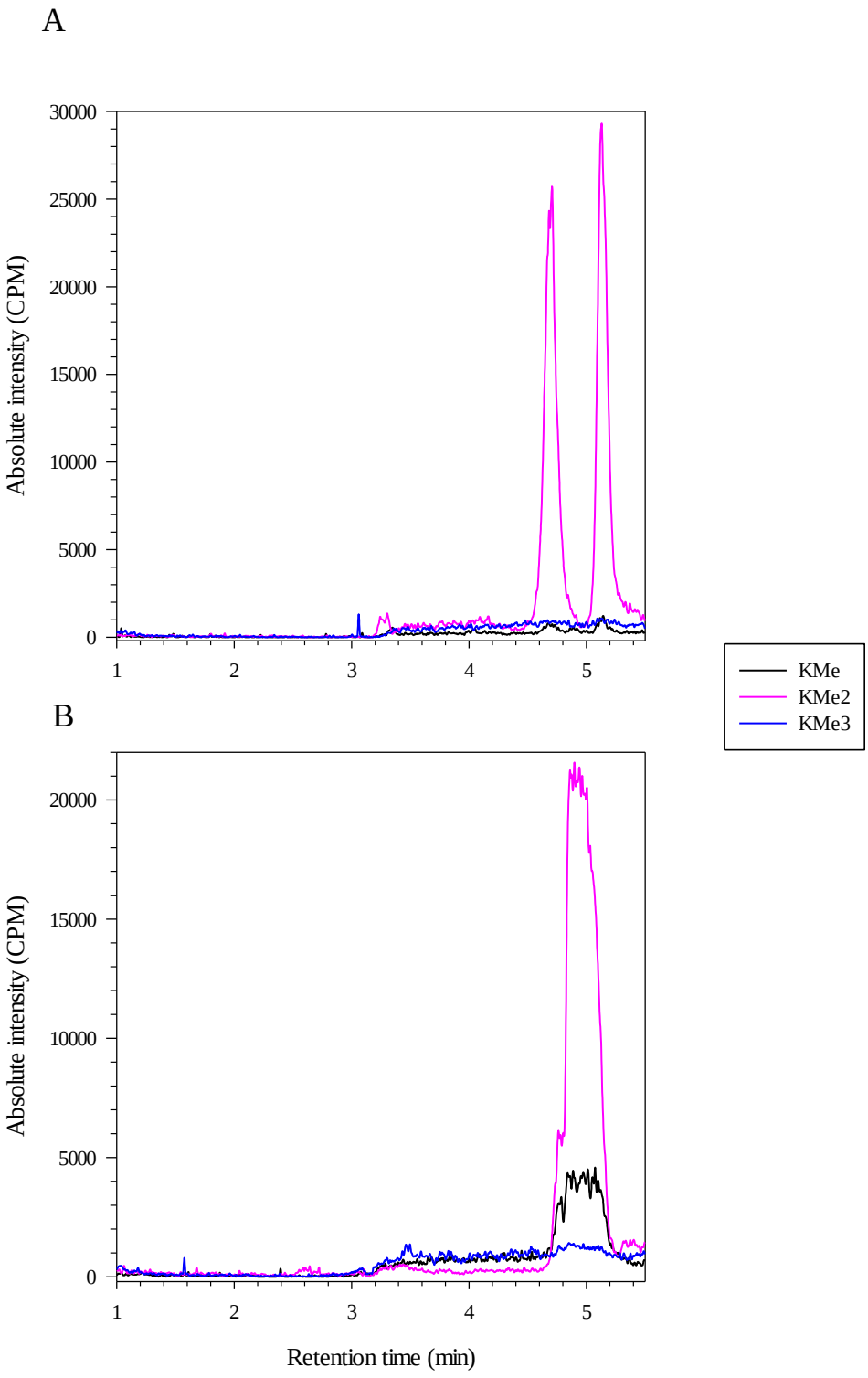


Figure 14. PTM assay chromatograms showing the KMe, KMe2, and KMe3 channels from LC-MS/MS for (A) blank Tris buffer, and (B) blank HEPES buffer.

3.4 Preliminary work with H3 peptides and SET7/9

Work with the PKMT SET7/9 began with recombinant enzyme purchased from New England Biolabs (NEB). This enzyme was used in *in vitro* methylation reactions with varying SAM concentrations and a peptide substrate consisting of the first 21 amino acids in the sequence of histone H3 (**Figure 15**). Methylation reactions were performed by varying the concentration of SAM and holding the concentration of H3 peptide constant. The amount of KMe and SAH produced in these reactions was measured via LC-MS/MS. **Figure 16** and **Figure 17** show chromatograms comparing chemical standards to experimental samples for the PTM assay and the SAH assay respectively. **Figure 18** shows chromatograms from the control samples for this experiment. The amount of each analyte found in the no SAM control was taken as background signal and subtracted from the amount found in each of the other samples, and the resulting values were plotted (**Figure 19**).

ART**K**QTARKSTGGKAPRKQLA

Figure 15. The sequence of the histone H3 peptides containing the first 21 amino acids of histone H3 (Epigentek). The location at which SET7/9 is supposed to methylate this substrate, K4, is highlighted in red.

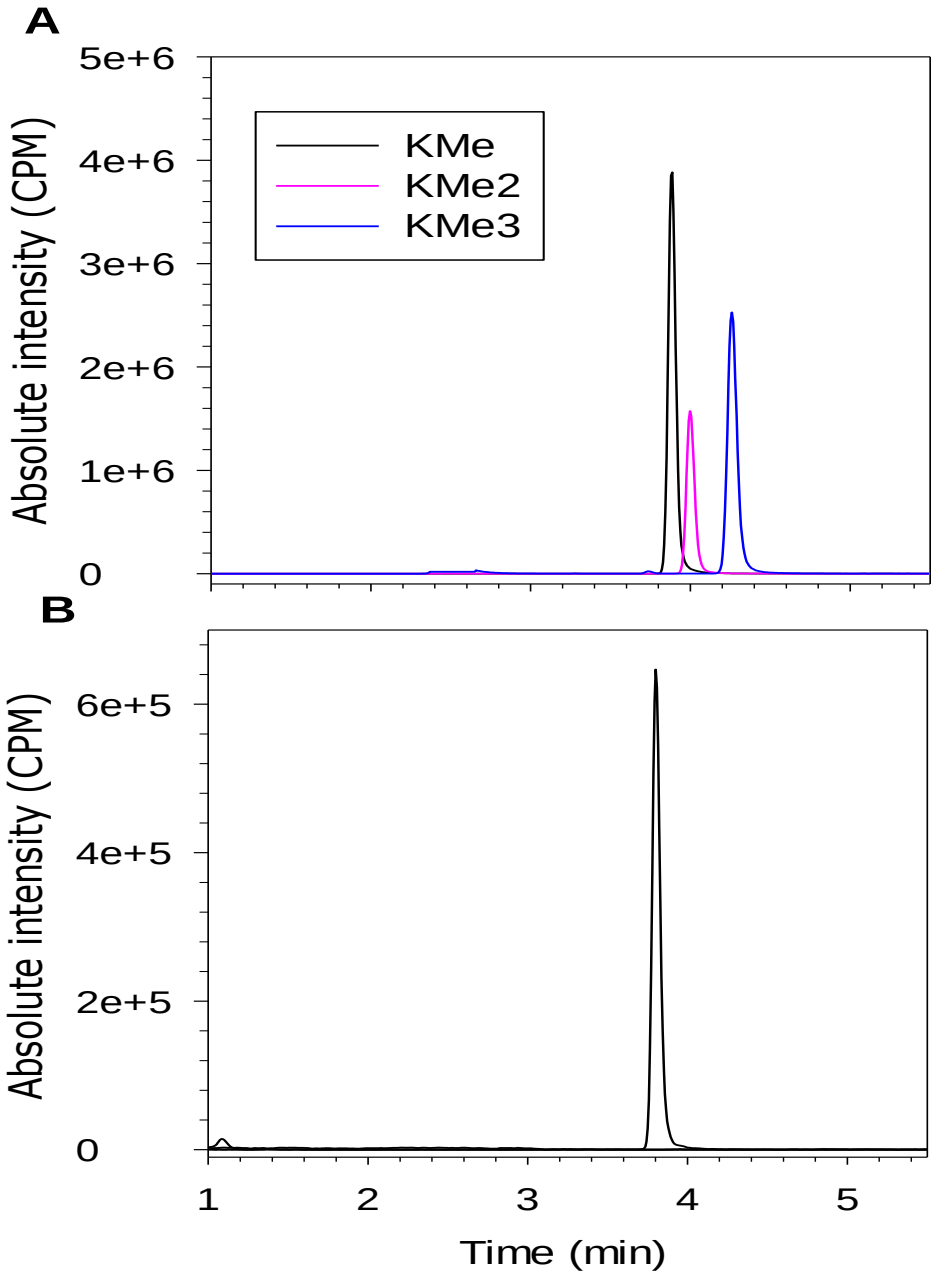


Figure 16. LC-MS/MS chromatogram showing (A) 5 µM KMe, KMe2, and KMe3 standards used for quantification of analytes, and (B) all methyllysines produced by SET7/9 in a reaction containing 5 µM SAM, 10 µM histone H3 peptide, and 0.1 µM SET7/9 (NEB).

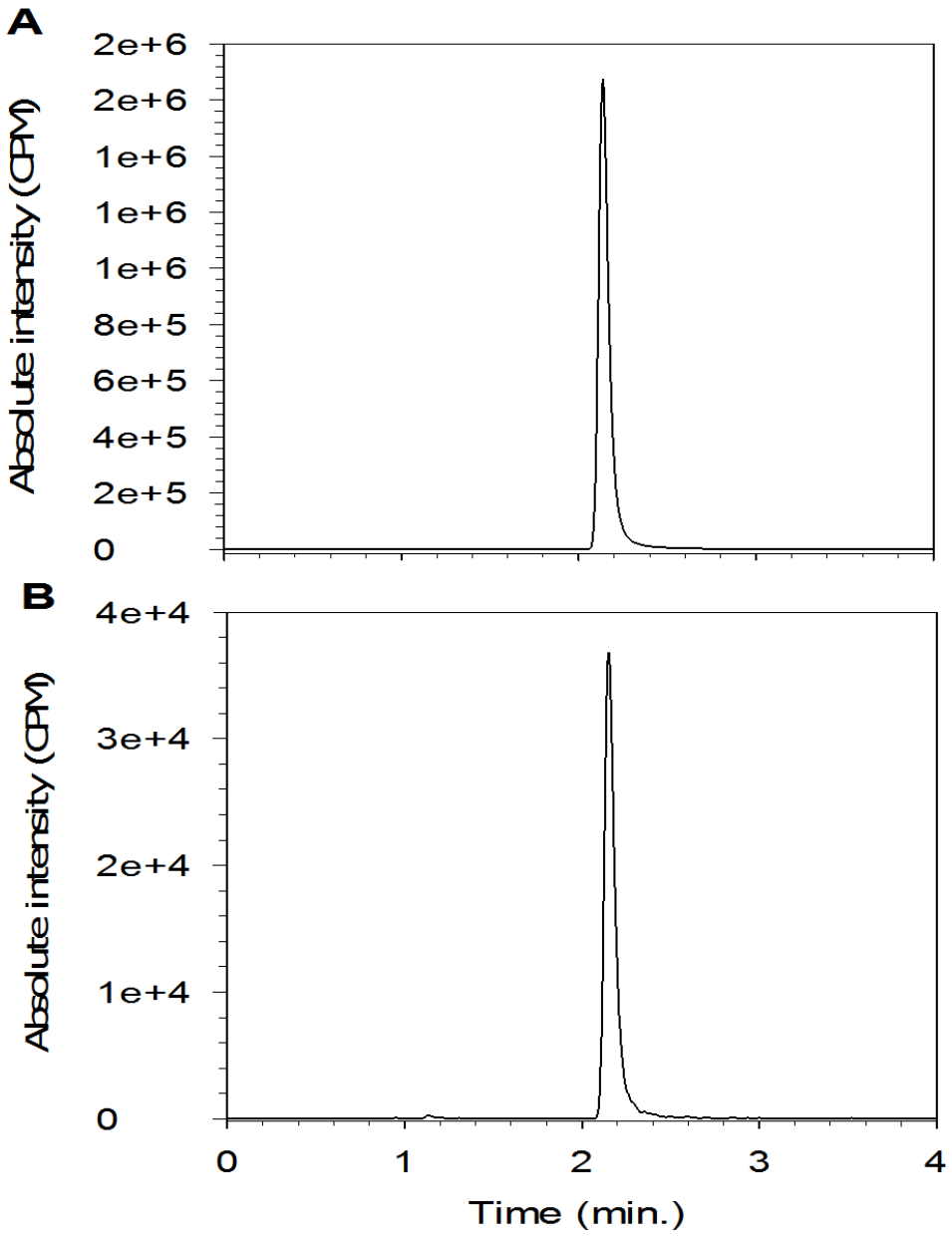


Figure 17. LC-MS/MS assay detecting SAH. The chromatograms depict (A) 10 µM standard SAH used for quantification of analytes, and (B) SAH produced in an experimental SET7/9 sample containing 5 µM SAM, 10 µM histone H3 peptide, and 0.1 µM SET7/9 (NEB).

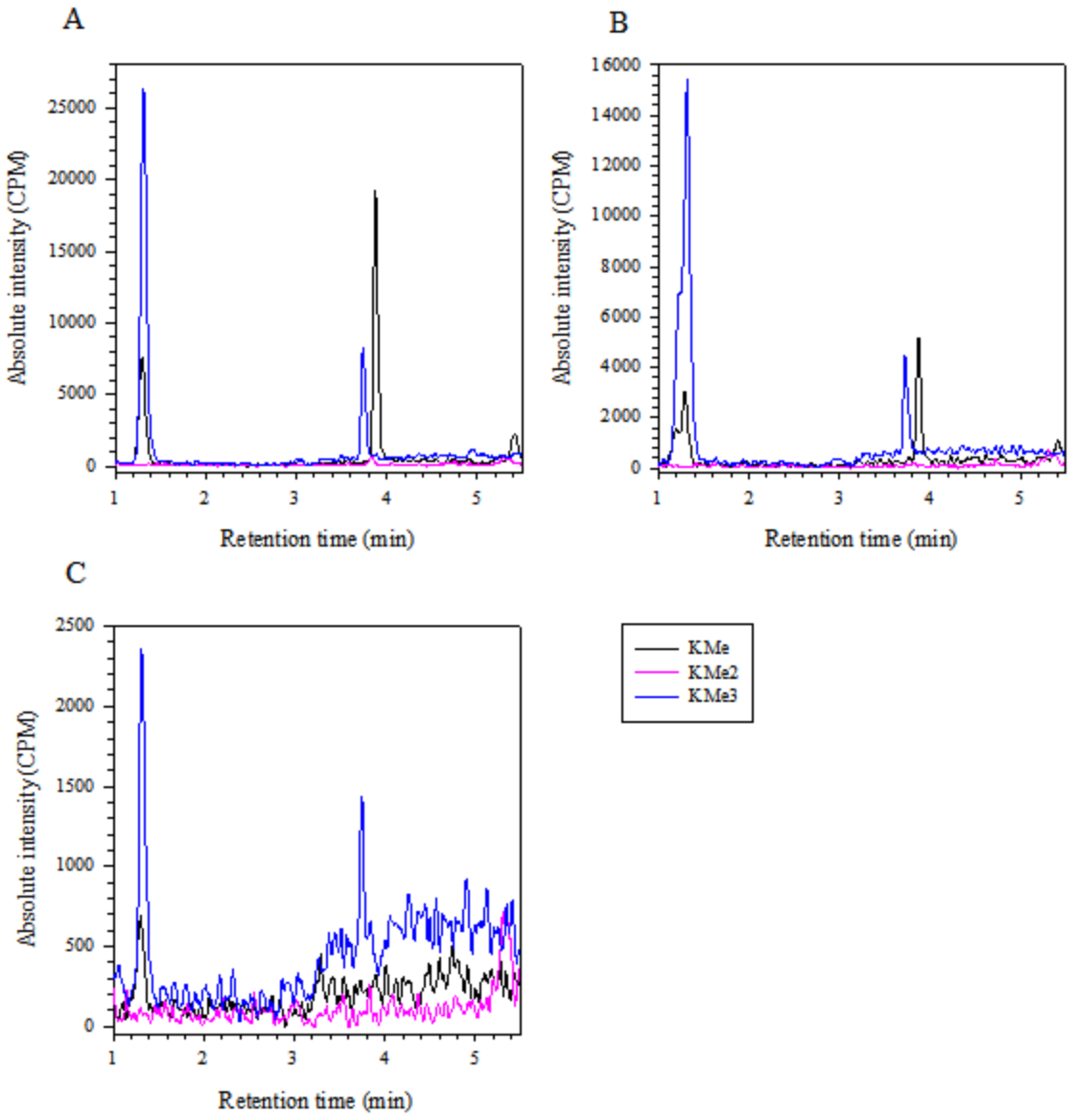


Figure 18. Control chromatograms from experiments with SET7/9 (NEB) for KMe, KMe2, and KMe3 (A) 0.1 µM enzyme, 10 µM histone H3, no SAM, (B) 0.1 µM enzyme, 5 µM SAM, no histone H3, and (C) 10 µM histone H3, 5 µM SAM, no enzyme.

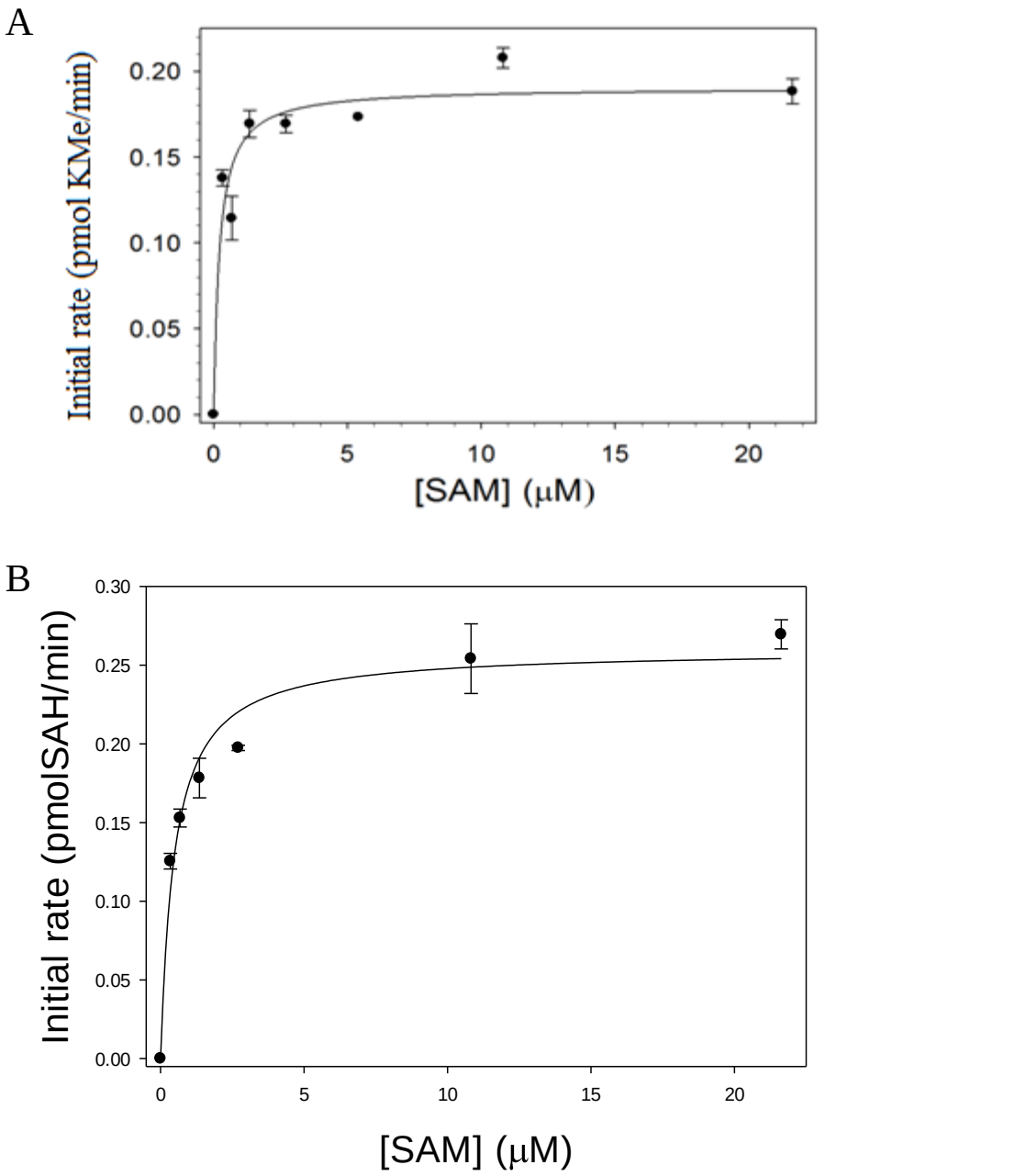


Figure 19. (A) Initial rate of KMe formation by SET7/9 (NEB) as the concentration of SAM varies, adjusted R^2=0.93 (B) Initial rate of SAH formation by SET7/9 (NEB) as the concentration of SAM varies, adjusted R^2=0.97. 0.1 μM SET7/9 was used and the concentration of H3 peptide was fixed at 10 μM. The data in both plots are fit to **Equation 5** using non-linear regression in SigmaPlot in order to determine the K_m^{app} and V_{max}^{app} of SET7/9 (NEB) for SAM. The circles are means of 3 samples ±SD.

3.5 SET7/9 PTM results:

Methylation reactions were performed with the expressed SET7/9 while varying the concentration of SAM and holding the concentration of histone H3 constant. Four negative controls were used: a no SAM control, a no histone control, a no enzyme control, and a control containing only enzyme in buffer. The amount of monomethyl lysine produced in these reactions was measured via LC-MS/MS. The amount of monomethyl lysine found in the no SAM control was taken as background signal and subtracted from the amount found in each of the other samples, and the resulting values were plotted (**Figure 20**). The K_m^{app} and V_{max}^{app} given by fitting experimental data to **Equation 5** or **Equation 7** are shown in **Table 3**. This experiment was then repeated, this time varying the concentration of histone H3 and holding the concentration of SAM constant. The data were treated the same as described above, except the no histone control was subtracted as background. The K_m^{app} and V_{max}^{app} of SET7/9 for histone H3 are also shown in **Table 3**. **Figure 21** shows chromatograms of the control samples from these experiments.

Table 3. Kinetic parameters of SET7/9 as calculated from experimental data.

Assay	**Varied substrate**	**Fixed substrate**	**K_m^{app} (μM)**	**V_{max}^{app} (pmol/min)**
PTM[1]	SAM	Histone	2.24 ± 0.97	0.047 ± 0.0057
PTM[1]	Histone	SAM	1.21 ± 0.53	0.16 ± 0.018
SAH[1]	SAM	Histone	8.54 ± 0.92	0.079 ± 0.0036
PTM[2]	SAM	H3 peptide	0.22 ± 0.03	0.19 ± 0.004
SAH[2]	SAM	H3 peptide	0.48 ± 0.07	0.26 ± 0.008

[1]experiments performed with expressed SET7/9

[2]experiments performed with purchased SET7/9 and histone H3 peptide

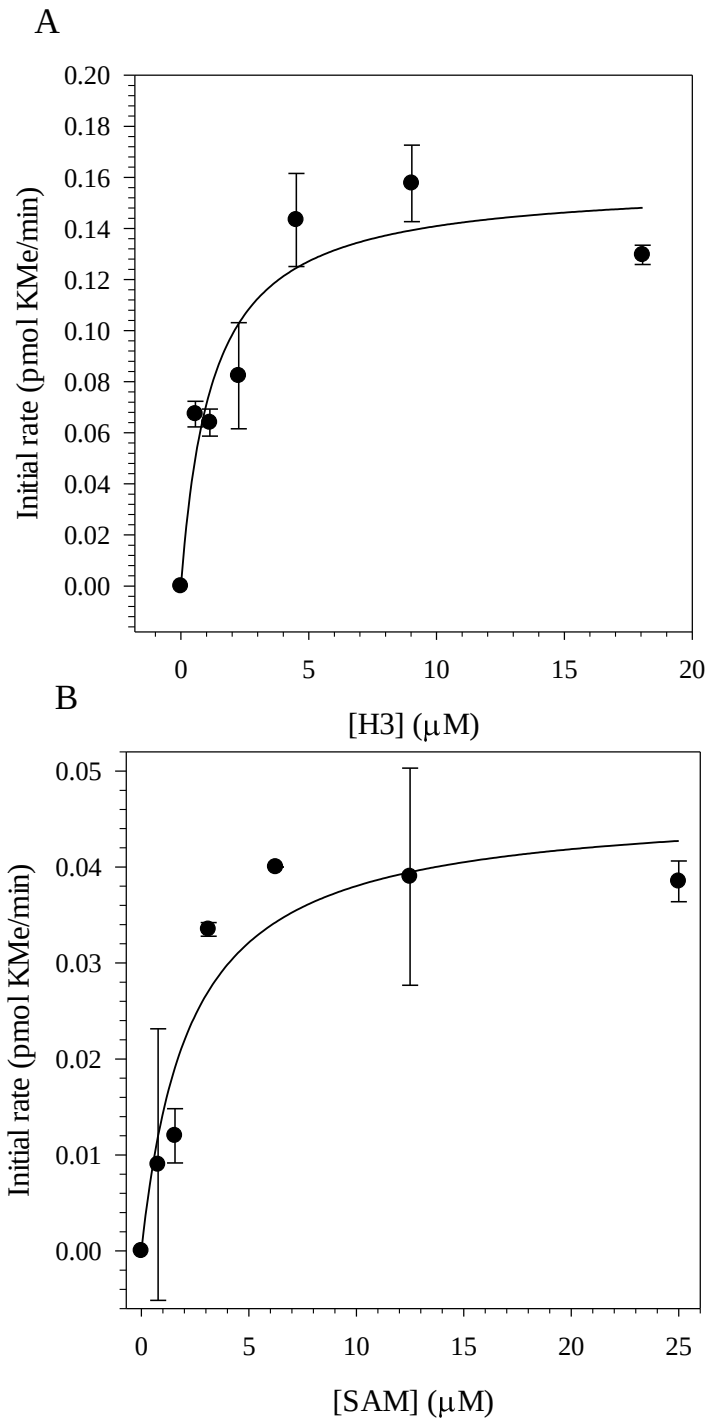


Figure 20. (A) Initial rate of KMe formation as the concentration of histone H3 increases, adjusted R^2=0.88. Expressed SET7/9 was used at a concentration of 0.1 μM and the concentration of SAM was fixed at 10 μM. The data were fit to **Equation 7** using non-linear

regression in SigmaPlot in order to determine the K_m^{app} and V_{max}^{app} of SET7/9 for histone H3 (**Table 3**). Each point is a mean of 2 samples ± SD. (B) Initial rate of KMe formation as the concentration of SAM increases, adjusted R^2=0.89. Expressed SET7/9 was used at a concentration of 0.1 µM and the concentration of histone H3 was fixed at 10 µM. The data are fit to **Equation 5** using non-linear regression in SigmaPlot in order to determine the K_m^{app} and V_{max}^{app} of SET7/9 for SAM (**Table 3**). Each point is a mean of 2 samples ± SD.

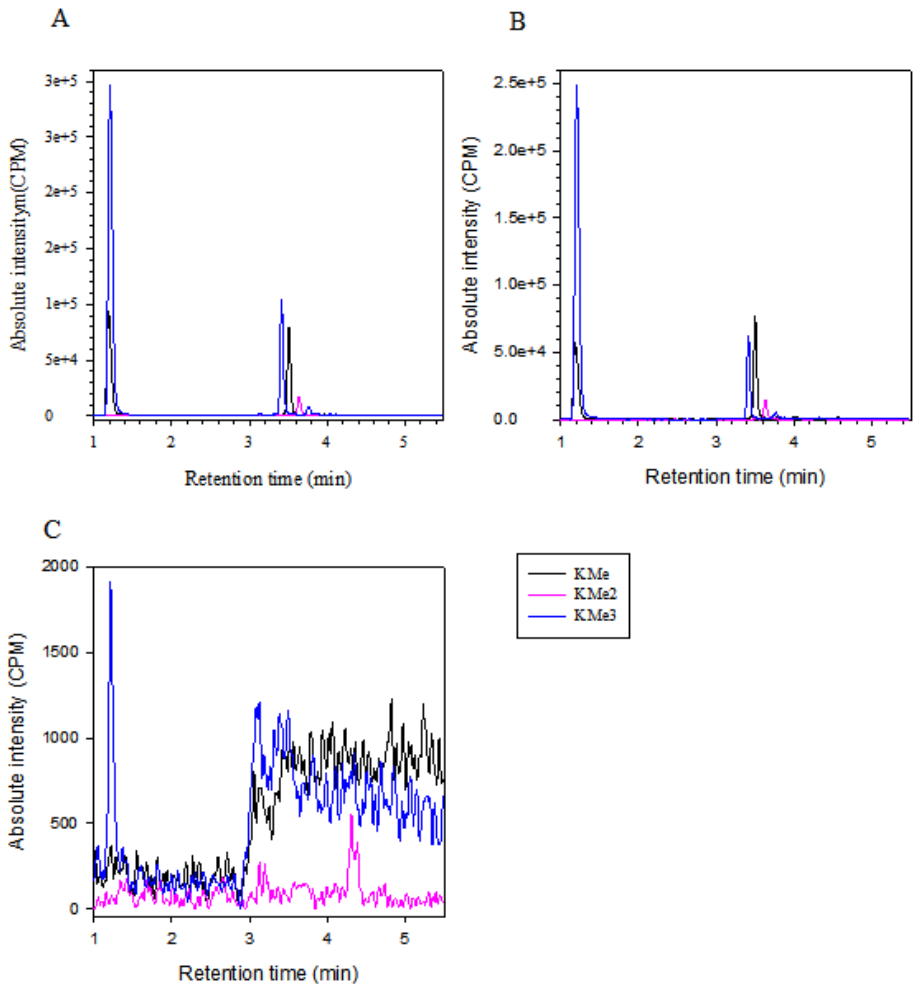


Figure 21. Chromatograms from experiments with expressed SET7/9 for KMe, KMe2, and KMe3 from control samples. (A) 0.1 µM enzyme, 10 µM histone, no SAM, (B) 0.1 µM enzyme, no histone, 5 µM SAM and (C) no enzyme, 10 µM histone, 5 µM SAM.

3.6 SET7/9 SAH assay results:

Samples were also taken from the methylation reactions discussed above in order to measure the SAH produced during these reactions by LC-MS/MS. Since each methylation event produces 1 molecule of SAH, SAH production is directly proportional to methylation rate. The data from the SAH assay were treated the same as the PTM data; measurements from control samples were subtracted from experimental samples in order to eliminate background signal. These data were then plotted as described above, and the K_m^{app} and V_{max}^{app} of the enzyme for SAM were determined (**Figure 22**). These values are shown in **Table 3**. The experiment was repeated to determine these values for histone H3 but the SAH assay was unable to detect SAH in the majority of the samples.

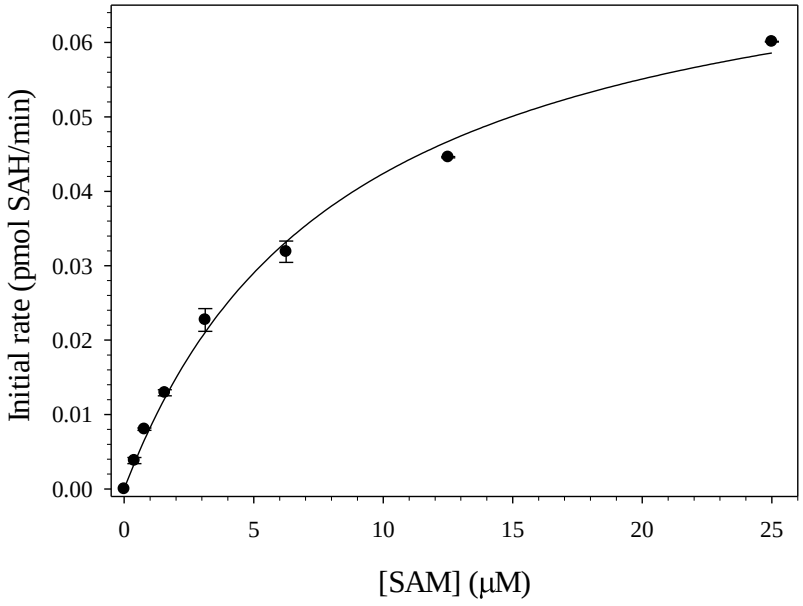


Figure 22. Initial rate of SAH formation by expressed SET7/9 with increasing concentration of SAM. The data are fit to **Equation 5** using non-linear regression in SigmaPlot in order to determine the K_m^{app} and V_{max}^{app} of SET7/9 for SAM (**Table 3)**, adjusted R^2=0.99. Each point is a mean of 2 samples ± SD. Expressed SET7/9 was used at a concentration of 0.1 μM and the concentration of histone H3 was fixed at 10 μM.

4.0 Discussion

4.1 Preliminary SET7/9 experiments

The work with SET7/9 purchased from NEB and the H3 peptides confirmed that SET7/9 was only producing monomethyl lysine as was found previously (*34*).These results were also a proof-of-principle that the *in vitro* methylation reaction format with LC-MS/MS as detection works. The K_m^{app} of SET7/9 with regard to SAM determined from the PTM data was 0.22 ± 0.03 µM, while the K_m^{app} determined from the SAH data was 0.48 ± 0.07 µM. The difference between these values stands out, as the K_m^{app} from the PTM assay is less than half that from the SAH assay. This was thought to be due to automethylation of the enzyme, and data from some of the controls looked like they might support this hypothesis. Methylation reactions with H3 peptide were also performed with varying concentrations of H3 peptide, but the highest concentration of the histone H3 peptide tested did not yield maximal enzymatic activity. This resulted in an almost linear curve when plotted, such that fitting the data to the Michaelis-Menten equation via non-linear regression would not yield reliable kinetic parameters. Some SAH data from these experiments were suggestive of automethylation by the enzyme as samples containing no H3 peptide contained SAH while the no enzyme control contained no detectable SAH, but further experiments were unable to corroborate this. We therefore concluded that SET7/9 does not automethylate.

4.2 Experiments with expressed SET7/9

The protein expression and purification protocol used here produced active protein and only required a 5 hour purification process. **Figure 6** shows that the protein of interest is co-purified with other proteins. A further purification step using a HiTrap Q strong anion exchange

column (GE Lifesciences) was attempted with other batches, but this resulted in a purer protein which had no biological activity. This may be explained by loss of activity during the extended purification process. Alternatively, the wrong protein may have been purified during subsequent purification steps, resulting in the loss of SET7/9 and hence its activity.

The PTM assay used here can detect and quantify KMe, KMe2, and KMe3, as well as many other PTMs, but the data from this assay which are shown above only relate to KMe. This is because no KMe2 or KMe3 was found to be produced by SET7/9. This confirms the identity of SET7/9 as a monomethyl transferase. **Figure 21** shows that the signal for KMe2 and KMe3 found in these samples is background signal coming from the enzyme preparation. When compared to **Figure 18**, it is clear that the background signal is similar to that found with purchased enzyme. This could indicate that the recombinant protein is highly methylated by the bacterial expression host, as bacteria do have their own lysine methyltransferase enzymes (*14*). Such background activity may also be caused by *in vitro* methylation of the contaminating proteins. The values reported here for the K_m^{app} of SET7/9 for H3 are similar to those previously reported, although many differing K_m values can be found. The value we found for the K_m^{app} for SAM is higher than the value reported by the Structural Genomics Consortium (SGC) group from whom we received the SET7/9 plasmid, while even higher values for the K_m^{app} of SET7/9 for SAM have also been reported (*12*, *100*). The study reported by the SGC which used this SET7/9 expression construct determined the kinetic parameters the same way as we have here, by plotting methylation data and calculating the parameters by fitting the data to **Equation 5** or **Equation 7** in SigmaPlot (*100*).

The decrease in enzyme activity observed at higher substrate concentrations observed in **Figure 20** may suggest substrate inhibition. Substrate inhibition is a phenomenon that is

observed in a wide variety of enzymes, particularly during *in vitro* experiments. It can be a result of having given an enzyme much more substrate than it would encounter in a physiological setting, or it can be a feature of the enzyme's mechanism (*104*). In the latter case, substrate inhibition generally serves to control the amount of a reaction's end-products which are being made. In the context of methyltransferases, this may constitute another type of control governing the placement of methyl lysine marks. Certain other methyltransferases have been shown to exhibit substrate inhibition, such as SUV39H1 (*49*, *69*).

An issue with the data obtained from these assays is that they do not yield the same values (**Table 3**). As stated above, one molecule of SAH is produced during each methylation event. Therefore, there should be an equal amount of SAH and KMe being produced. However this is not the case, leading to different values for K_m^{app} and V_{max}^{app} depending on which data are being plotted. This could be caused by a few phenomena such as enzymatic automethylation, error in the assays due to high background signal, or errors in detection from analytes being below the lower limit of detection or contamination of the samples with endogenous bacterial methylation or SAH. Other methyltransferases have been shown to automethylate – such as G9a, MLL1, and all PRMTs – but this has not yet been shown with SET7/9 (*31*, *50*, *51*). Automethylation has been shown in some cases to affect the activity of the automodified enzyme, meaning that this could be another avenue for control of these epigenetic marks. Such is the case for CARM1 and PRMT 8 (*31*, *105*). Automethylation of G9a has not been reported to have an effect on enzymatic activity (*51*). Methylation occurring on the enzyme itself could also explain the discrepancies between the KMe and SAH data. If the enzyme is being methylated while it is still in the bacterium in which it is expressed, then that could falsely increase the signal for KMe in the LC-MS/MS assay. In fact, the controls which contained only the enzyme

preparation in buffer did show a background signal for KMe, KMe2, and KMe3 (**Figure 21**). While some data from experiments with purchased SET7/9 and H3 peptide may have indicated the possibility of automethylation, no data from experiments with expressed SET7/9 have corroborated that. However, if automethylation was occurring at very low levels, then it may have been hidden by the background signal. If the enzyme was methylated prior to purification, either through its own activity or by bacterial methyltransferases, then any site which could have been automethylated may have been occupied prior to experimentation. However, such complete pre-methylation would have prevented the discovery of automethylation in G9a, MLL1, and PRMTs.

The differences between values could also be due to the enzyme preparation considering the purchased SET7/9 and the SET7/9 expressed and purified in this work. The estimation of enzyme concentration in the purified fraction could have been incorrect or measured differently, which would lead to differences in the values obtained. The purified fraction also contained impurities, such as other proteins which were co-purified. The presence of these impurities could be responsible, as the activity of an enzyme may be affected by other proteins. Furthermore, any of the protein impurities could have contained methylated lysines which could have contributed to the signal detected in these experiments. There were also issues seen in the assays. The SAH assay in particular had trouble with sensitivity, such that no SAH was detectable in some samples at low substrate concentrations. The SAH assay could therefore be somewhat less reliable than the PTM assay. A difficulty with the PTM assay, however, was that the controls gave back higher KMe values than some samples with low substrate concentrations. This meant that subtracting the control from these low concentration samples resulted in a negative value, which then had to be excluded from further calculations. However it is more likely that the issue lies

with the enzyme preparation or with the *in vitro* methylation assays, as both LC-MS/MS assays have been previously validated (*79*, *102*). The differences between the values obtained with purchased SET7/9 and those obtained with SET7/9 which we expressed are expected, as activity of enzymes can vary from one batch to another.

4.3 DOT1L expression

The purification of DOT1L was performed according to the protocol outlined by Yu *et al* (*53*). **Figures 9-11** show that the protein of interest was produced and is present in samples taken throughout the purification process. However activity tests performed with purified DOT1L all returned inconclusive, with no observable methyltransferase activity (**Figure 12**). The reason for this may be that the purification process was too long, resulting in inactive enzyme. To confirm this we could attempt expression of DOT1L, but omit the cation exchange column step. This would lower the purity of the final product, but may yield active enzyme by decreasing the likelihood of product degradation or denaturation.

4.4 G9a methylation experiment

In vitro methylation reactions were performed using G9a purchased from New England Biolabs. The results from this experiment were also inconclusive, with no difference between the experimental samples and the negative controls (**Figure 12)**. This was likely caused by the use of a Tris-based buffer, which caused a high level of background (**Figure 13**). It was the failure of this experiment that prompted the testing of different buffer components as sources of erroneous signal, leading to the use of ammonium acetate as a main buffer component for the experiments with SET7/9. The lower contribution of ammonium acetate to the background signal could be due to the fact that, under the vacuum conditions in the SpeedVac while the samples are being

dried, the ammonium and acetate in the samples are converted into ammonia and acetic acid, which are in a gas form. This was much of the buffer is likely removed from the samples, leaving fewer ions which can cause background signal.

4.5 Conclusions

This study has shown that the activity of PKMTs can be detected and measured using LC-MS/MS assays which quantify the methyllysines produced by the enzyme and the SAH produced as a by-product. This is novel because our methods are the first to be developed which can measure the entirety of a PKMT's activity. Other techniques used in the field of epigenetics may not detect some crucial information due to shortcomings inherent in other analytical techniques. For example proteomics assays seldom achieve full protein sequence coverage, so there is a possibility that they will not detect all of the modifications present on the protein substrate. Importantly, if those modifications cannot be detected they cannot be quantified (*72*). Other techniques, such as radiography and scintillation based assays, are unable to distinguish between different modifications, such as KMe, KMe2, and KMe3. Our PTM assay has confirmed that SET7/9 only catalyzes the formation of KMe, definitively showing that no KMe2 or KMe3 was present on the substrate proteins *in vitro*. The fact that the LC-MS/MS assays employed here can detect the full extent of the activity of a PKMT, and that the PTM assay can detect which modifications are being formed, make these assays uniquely well-suited for the study of epigenetic enzymes. With these techniques, future research in epigenetics will be able to give a more complete view of the activity of these enzymes.

4.6 Future Experiments

The next steps in this research program will be to perform the same *in vitro* methylation experiments while varying the concentrations of both substrates at the same time, with the middle of the concentration range for each being the K_m^{app} which we have determined for SET7/9. Having 6 concentration levels for each substrate will yield 36 samples which, when the data are plotted as outlined above, will give a family of 6 curves. The shape of these curves will help to elucidate the mechanism of action of this enzyme, specifically the order in which the substrates bind to the enzyme (*67*). This can be further confirmed by performing product inhibitor experiments with SET7/9. This methylation reaction and analysis protocol will also be used in the future to test new compounds for their ability to inhibit SET7/9. Epigenetic enzymes are known to have roles in many different cellular processes, from regulating transcription to DNA strand break repair (*8*, *40*, *63*, *106*, *107*). Understanding how these enzymes will inform the drug development process.

Another aspect of epigenetic methylation that is of interest is whether or not the presence of other epigenetic PTMs on the protein substrate influences the deposition of the methyl lysine mark on the enzymatic level. This will be tested using this reaction platform by performing the methylation reactions described here using pre-modified histones. These histones are available for purchase and are offered with modifications such as acetylation, all types of methylation, and phosphorylation at various spots on histone H3 (Epigentek). Peptides with modifications at the H3K9 position would be of particular interest since methylation there has previously been shown to inhibit the placement of methyl marks at H3K4 (*50*). Determining the effect of these other PTMs on the kinetic parameters of SET7/9 would elucidate a previously unrecognized relationship between the modifications at the enzymatic level. For example, inhibition of

methylation at H3K4 by prior modification at H3K9 – such as mono-, di-, or trimethylation, or acetylation – would likely show itself as an increase in the K_m of SET7/9 for the peptide substrate and/or a decrease in the V_{max}. Modifications at other locations on histone H3 may increase deposition of KMe on H3K4 by SET7/9. This could manifest as a decrease in K_m of SET7/9 for the pre-modified substrate and/or an increase in V_{max}. Of interest in this category are peptides with acetylation at H3K14 or K18 since lysine acetylation is generally a transcriptionally active marker (*106*). If use of pre-modified proteins affects the activity of SET7/9, then this would indicate that there is some form of control exerted over deposition of epigenetic PTMs at the enzymatic level by other PTMs. Relationships between the deposition of some PTMs on histones have been shown before, so the discovery of more such cases seems likely (*48, 106*). Study of this epigenetic cross-talk is essential to further our understanding of the larger framework of epigenetics.

5.0 References

1. M. a Dawson, T. Kouzarides, B. J. P. Huntly, Targeting epigenetic readers in cancer. *N. Engl. J. Med.* **367**, 647–57 (2012).

2. N. R. Rose, R. J. Klose, Understanding the relationship between DNA methylation and histone lysine methylation. *Biochim. Biophys. Acta*. **1839**, 1362–1372 (2014).

3. B. D. Strahl, C. D. Allis, The language of covalent histone modifications. *Nature*. **403**, 41–45 (2000).

4. A. N. Richart, C. I. W. Brunner, K. Stott, N. V Murzina, J. O. Thomas, Characterization of chromoshadow domain-mediated binding of heterochromatin protein 1α (HP1α) to histone H3. *J. Biol. Chem.* **287**, 18730–7 (2012).

5. R. Collepardo-Guevara *et al.*, Chromatin unfolding by epigenetic modifications explained by dramatic impairment of internucleosome interactions: A multiscale computational study. *J. Am. Chem. Soc.* **137**, 10205–10215 (2015).

6. A. J. Bannister, T. Kouzarides, Regulation of chromatin by histone modifications. *Cell Res.* **21**, 381–95 (2011).

7. C. Rivera, Z. A. Gurard-Levin, G. Almouzni, A. Loyola, Histone lysine methylation and chromatin replication. *Biochim. Biophys. Acta*. **1839**, 1433–1439 (2014).

8. M. A. Dawson, T. Kouzarides, Cancer epigenetics: from mechanism to therapy. *Cell*. **150**, 12–27 (2012).

9. J.-Y. Kim *et al.*, A role for WDR5 in integrating threonine 11 phosphorylation to lysine 4 methylation on histone H3 during androgen signaling and in prostate cancer. *Mol. Cell*. **54**, 613–25 (2014).

10. P. ’t Hart, T. M. Lakowski, D. Thomas, A. Frankel, N. I. Martin, Peptidic partial bisubstrates as inhibitors of the protein arginine N-methyltransferases. *Chembiochem*. **12**, 1427–32 (2011).

11. T. M. Lakowski, A. Frankel, Kinetic analysis of human protein arginine N-methyltransferase 2: formation of monomethyl- and asymmetric dimethyl-arginine residues on histone H4. *Biochem. J.* **421**, 253–61 (2009).

12. K. Guitot *et al.*, Label-free measurement of histone lysine methyltransferases activity by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Anal. Biochem.* **456**, 25–31 (2014).

13. K. W. Foreman *et al.*, Structural and functional profiling of the human histone methyltransferase SMYD3. *PLoS One*. **6**, e22290 (2011).

14. S. Lanouette, V. Mongeon, D. Figeys, J. F. Couture, The functional diversity of protein lysine methylation. *Mol. Syst. Biol.* **10**, 1–26 (2014).

15. A. D. Ferguson *et al.*, Structural basis of substrate methylation and inhibition of SMYD2. *Structure*. **19**, 1262–73 (2011).

16. H.-M. Herz, A. Garruss, A. Shilatifard, SET for life: biochemical activities and biological functions of SET domain-containing proteins. *Trends Biochem. Sci.* **38**, 621–39 (2013).

17. EMBL-EBI, Zinc finger, MYND-type. *Wellcome Trust Genome Campus* (2014), (available at http://www.ebi.ac.uk/interpro/entry/IPR002893).

18. N. Spellmon, J. Holcomb, L. Trescott, N. Sirinupong, Z. Yang, Structure and Function of SET and MYND Domain-Containing Proteins. *Int. J. Mol. Sci.* **16**, 1406–1428 (2015).

19. F. Kateb *et al.*, Structural and functional analysis of the DEAF-1 and BS69 MYND

domains. *PLoS One*. **8**, e54715 (2013).

20. M. Abu-Farha *et al.*, The Tale of Two Domains: Proteomics and Genomics Analysis of SMYD2, A New Histone Methyltransferase. *Mol. Cell. Proteomics*. **7**, 560–572 (2007).

21. T. M. Lakowski, P. 't Hart, C. A. Ahern, N. I. Martin, A. Frenkel, Nη-substituted arginyl peptide inhibitors of protein arginine N-methyltransferases. *ACS Chem. Biol.* **5**, 1053–1063 (2010).

22. H.-B. Guo, H. Guo, Mechanism of histone methylation catalyzed by protein lysine methyltransferase SET7/9 and origin of product specificity. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 8797–802 (2007).

23. D. R. Kipp, C. M. Quinn, P. D. Fortin, Enzyme-dependent lysine deprotonation in EZH2 catalysis. *Biochemistry*. **52**, 6866–78 (2013).

24. R. C. Trievel, B. M. Beach, L. M. a. Dirk, R. L. Houtz, J. H. Hurley, Structure and Catalytic Mechanism of a SET Domain Protein Methyltransferase. *Cell*. **111**, 91–103 (2002).

25. A. Benard *et al.*, Histone trimethylation at H3K4, H3K9 and H4K20 correlates with patient survival and tumor recurrence in early-stage colon cancer. *BMC Cancer*. **14**, 531 (2014).

26. Y. Chen *et al.*, Increased Expression of SETD7 Promotes Cell Proliferation by Regulating Cell Cycle and Indicates Poor Prognosis in Hepatocellular Carcinoma. *PLoS One*. **11**, e0154939 (2016).

27. Z. Li *et al.*, The polycomb group protein EZH2 is a novel therapeutic target in tongue cancer. *Oncotarget*. **4**, 2532–2549 (2013).

28. M. W. Chen *et al.*, H3K9 histone methyltransferase G9a promotes lung cancer invasion and metastasis by silencing the cell adhesion molecule Ep-CAM. *Cancer Res.* **70**, 7830–7840 (2010).

29. S. Komatsu *et al.*, Overexpression of SMYD2 contributes to malignant outcome in gastric carcinoma. *Br. J. Cancer*. **112**, 357–364 (2015).

30. W. Oliveira-Santos *et al.*, Residual expression of SMYD2 and SMYD3 is associated with the acquisition of complex karyotype in chronic lymphocytic leukemia. *Tumor Biol.* **37**, 9473–9481 (2016).

31. A. Patel *et al.*, Automethylation activities within the mixed lineage leukemia-1 (MLL1) core complex reveal evidence supporting a "two-active site" model for multiple histone H3 lysine 4 methylation. *J. Biol. Chem.* **289**, 868–84 (2014).

32. P. A. Del Rizzo, R. C. Trievel, Molecular basis for substrate recognition by lysine methyltransferases and demethylases. *Biochim. Biophys. Acta*. **1839**, 1404–1415 (2014).

33. P. A. Del Rizzo *et al.*, SET7/9 catalytic mutants reveal the role of active site water molecules in lysine multiple methylation. *J. Biol. Chem.* **285**, 31849–31858 (2010).

34. J.-F. Couture, E. Collazo, G. Hauk, R. C. Trievel, Structural basis for the methylation site specificity of SET7/9. *Nat. Struct. Mol. Biol.* **13**, 140–6 (2006).

35. Y. Jiang *et al.*, Structural Insights into Estrogen Receptor α Methylation by Histone Methyltransferase SMYD2, a Cellular Event Implicated in Estrogen Signaling Regulation. *J. Mol. Biol.* **426**, 3413–3425 (2014).

36. S. Chuikov *et al.*, Regulation of p53 activity through lysine methylation. *Nature*. **432**, 353–360 (2004).

37. Y. Akiyama, Y. Koda, S. Byeon, S. Shimada, T. Nishikawaji, Reduced expression of SET7/9, a histone mono-methyltransferase, is associated with gastric cancer progression. *Oncotarget*. **7**, 3966–3983 (2015).

38. C. Shen *et al.*, SET7/9 regulates cancer cell proliferation by influencing B-catenin stability. *FASEB J.* **29**, 4313–4323 (2015).

39. Y. Song *et al.*, SET7/9 inhibits oncogenic activities through regulation of Gli-1 expression in breast cancer. *Tumor Biol.* **37**, 9311–9322 (2016).

40. J. K. Kurash *et al.*, Methylation of p53 by Set7/9 Mediates p53 Acetylation and Activity In Vivo. *Mol. Cell*. **29**, 392–400 (2008).

41. G. S. van Aller *et al.*, Smyd3 regulates cancer cell phenotypes and catalyzes histone H4 lysine 5 methylation. *Epigenetics*. **7**, 340–343 (2012).

42. S. Xu, C. Zhong, T. Zhang, J. Ding, Structure of human lysine methyltransferase Smyd2 reveals insights into the substrate divergence in Smyd proteins. *J. Mol. Cell Biol.* **3**, 293–300 (2011).

43. S. M. Carr, A. Poppy Roworth, C. Chan, N. B. La Thangue, Post-translational control of transcription factors: Methylation ranks highly. *FEBS J.* **282**, 4450–4465 (2015).

44. L. A. Saddic *et al.*, Methylation of the retinoblastoma tumor suppressor by SMYD2. *J. Biol. Chem.* **285**, 37733–37740 (2010).

45. Q. Liu, M. Wang, Histone lysine methyltransferases as anti-cancer targets for drug discovery. *Acta Pharmacol. Sin.* **37**, 1–8 (2016).

46. L. Liu, S. Kimball, H. Liu, A. Holowatyj, Genetic alterations of histone lysine methyltransferases and their significance in breast cancer. *Oncotarget*. **6**, 2466–2482

(2014).

47. M. A. Brown, R. J. Sims, P. D. Gottlieb, P. W. Tucker, Identification and characterization of Smyd2: a split SET/MYND domain-containing histone H3 lysine 36-specific methyltransferase that interacts with the Sin3 histone deacetylase complex. *Mol. Cancer*. **5** (2006) (available at http://www.ncbi.nlm.nih.gov/pubmed/16805913\nhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1524980).

48. T. Ezponda, J. D. Licht, Molecular pathways: deregulation of histone h3 lysine 27 methylation in cancer-different paths, same destination. *Clin. Cancer Res.* **20**, 5001–8 (2014).

49. D. Patnaik *et al.*, Substrate specificity and kinetic mechanism of mammalian G9a histone H3 methyltransferase. *J. Biol. Chem.* **279**, 53248–53258 (2004).

50. P. Rathert *et al.*, Protein lysine methyltransferase G9a acts on non-histone targets. *Nat. Chem. Biol.* **4**, 344–6 (2008).

51. H. G. Chin *et al.*, Automethylation of G9a and its implication in wider substrate specificity and HP1 binding. *Nucleic Acids Res.* **35**, 7313–7323 (2007).

52. A. T. Nguyen, Y. Zhang, The diverse functions of Dot1 and H3K79 methylation. *Genes Dev.* **25**, 1345–1358 (2011).

53. W. Yu *et al.*, Catalytic site remodelling of the DOT1L methyltransferase by selective inhibitors. *Nat. Commun.* **3**, 1–11 (2012).

54. B. J. Venters, B. F. Pugh, Chromatin meets RNA polymerase II. *Genome Biol.* **8**, 319 (2007).

55. E. Kroon *et al.*, Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. *EMBO J.* **17**, 3714–3725 (1998).

56. A. V. Krivtsov, S. A. Armstrong, MLL translocations, histone modifications and leukaemia stem-cell development. *Nat. Rev. Cancer*. **7**, 823–833 (2007).

57. Y. Dou *et al.*, Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell*. **121**, 873–885 (2005).

58. H. G. Drexler, H. Quentmeier, R. A. F. MacLeod, Malignant hematopoietic cell lines: in vitro models for the study of MLL gene alterations. *Leukemia*. **18**, 227–32 (2004).

59. S. a Shinsky *et al.*, A non-active-site SET domain surface crucial for the interaction of MLL1 and the RbBP5/Ash2L heterodimer within MLL family core complexes. *J. Mol. Biol.* **426**, 2283–99 (2014).

60. Y. Okada *et al.*, hDOT1L links histone methylation to leukemogenesis. *Cell*. **121**, 167–178 (2005).

61. U. Thorsteinsdottir, E. Kroon, L. Jerome, F. Blasi, G. Sauvageau, Defining Roles for HOX and MEIS1 Genes in Induction of Acute Myeloid Leukemia. *Mol. Cell. Biol.* **21**, 224–234 (2001).

62. M. He *et al.*, Two isoforms of HOXA9 function differently but work synergistically in human MLL-rearranged leukemia. *Blood Cells, Mol. Dis.* **49**, 102–106 (2012).

63. P. Agarwal, S. P. Jackson, G9a inhibition potentiates the anti-tumour activity of DNA double-strand break inducing agents by impairing DNA repair independent of p53 status. *Cancer Lett.* **380**, 467–475 (2016).

64. R. K. Slany, The molecular biology of mixed lineage leukemia. *Haematologica*. **94**, 984–993 (2009).

65. W. Peters *et al.*, Enzymatic site-specific functionalization of protein methyltransferase substrates with alkynes for click labeling. *Angew. Chem. Int. Ed. Engl.* **49**, 5170–3 (2010).

66. O. Binda *et al.*, A chemical method for labeling lysine methyltransferase substrates. *Chembiochem*. **12**, 330–4 (2011).

67. R. A. Copeland, *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis* (John Wiley & Sons Inc., New York, NY, 2000).

68. T. M. Lakowski, A. Frankel, A kinetic study of human protein arginine N-methyltransferase 6 reveals a distributive mechanism. *J. Biol. Chem.* **283**, 10015–10025 (2008).

69. H. G. Chin, D. Patnaik, P. O. Estève, S. E. Jacobsen, S. Pradhan, Catalytic properties and kinetic mechanism of human recombinant Lys-9 histone H3 methyltransferase SUV39H1: Participation of the chromodomain in enzymatic catalysis. *Biochemistry*. **45**, 3272–3284 (2006).

70. L. M. a Dirk *et al.*, Kinetic manifestation of processivity during multiple methylations catalyzed by SET domain protein methyltransferases. *Biochemistry*. **46**, 3905–3915 (2007).

71. S. F. M. Van Dongen, J. A. A. W. Elemans, A. E. Rowan, R. J. M. Nolte, Processive Catalysis. *Angew. Chemie - Int. Ed.* **53**, 11420–11428 (2014).

72. M. Luo, Current chemical biology approaches to interrogate protein methyltransferases. *ACS Chem. Biol.* **7**, 443–463 (2012).

73. L. Xia, H. de Vries, A. P. IJzerman, L. H. Heitman, Scintillation proximity assay (SPA) as a new approach to determine a ligand's kinetic profile. A case in point for the adenosine A1 receptor. *Purinergic Signal.* **12**, 115–126 (2016).

74. D. Greiner, T. Bonaldi, R. Eskeland, E. Roemer, A. Imhof, Identification of a specific inhibitor of the histone methyltransferase SU(VAR)3-9. *Nat. Chem. Biol.* **1**, 143–5 (2005).

75. T. M. Lakowski, C. Zurita-Lopez, S. G. Clarke, A. Frankel, Approaches to measuring the activities of protein arginine N-methyltransferases. *Anal. Biochem.* **397**, 1–11 (2010).

76. S. Lin *et al.*, Stable-isotope-labeled histone peptide library for histone post-translational modification and variant quantification by mass spectrometry. *Mol. Cell. Proteomics*. **13**, 2450–66 (2014).

77. Y.-M. Kuo, R. a Henry, A. J. Andrews, A quantitative multiplexed mass spectrometry assay for studying the kinetic of residue-specific histone acetylation. *Methods*. **70**, 127–133 (2014).

78. B. Edrissi, K. Taghizadeh, P. C. Dedon, Quantitative analysis of histone modifications: formaldehyde is a source of pathological n(6)-formyllysine that is refractory to histone deacetylases. *PLoS Genet.* **9**, e1003328 (2013).

79. R. Lillico, M. G. Sobral, N. Stesco, T. M. Lakowski, HDAC inhibitors induce global changes in histone lysine and arginine methylation and alter expression of lysine demethylases. *J. Proteomics*. **133**, 125–133 (2016).

80. T. Uhlmann *et al.*, A method for large-scale identification of protein arginine methylation. *Mol. Cell. Proteomics*. **11**, 1489–99 (2012).

81. S. M. Carlson, O. Gozani, Emerging Technologies to Map the Protein Methylome. *J. Mol.*

Biol. **426**, 3350–3362 (2014).

82. B. Bogdanov, R. D. Smith, Proteomics by fticr mass spectrometry: TOP down and bottom up. *Mass Spectrom. Rev.* **24**, 168–200 (2005).

83. B. A. Garcia, What Does the Future Hold for Top Down Mass Spectrometry? *J. Am. Soc. Mass Spectrom.* **21**, 193–202 (2010).

84. B. Macek, L. F. Waanders, J. V Olsen, M. Mann, Top-down Protein Sequencing and MS3 on a Hybrid Linear Quadrupole Ion Trap-Orbitrap Mass Spectrometer. *Mol Cell Proteomics*. **5**, 949–958 (2006).

85. Y. Mao, S. G. Valeja, J. C. Rouse, C. L. Hendrickson, A. G. Marshall, Top-down structural analysis of an intact monoclonal antibody by electron capture dissociation-fourier transform ion cyclotron resonance-mass spectrometry. *Anal. Chem.* **85**, 4239–4246 (2013).

86. Y. O. Tsybin *et al.*, Structural analysis of intact monoclonal antibodies by electron transfer dissociation mass spectrometry. *Anal. Chem.* **83**, 8919–8927 (2011).

87. T. Hattori *et al.*, Renewable, recombinant antibodies to histone post-translational modifications. **10**, 992–995 (2013).

88. A. Guo *et al.*, Immunoaffinity enrichment and mass spectrometry analysis of protein methylation. *Mol. Cell. Proteomics*. **13**, 372–87 (2014).

89. T. L. Graves, Y. Zhang, J. E. Scott, A universal competitive fluorescence polarization activity assay for S-adenosylmethionine utilizing methyltransferases. *Anal. Biochem.* **373**, 296–306 (2008).

90. N. Gauthier *et al.*, Development of Homogeneous Nonradioactive Methyltransferase and

Demethylase Assays Targeting Histone H3 Lysine 4. *J. Biomol. Screen.* **17**, 49–58 (2012).

91. T. A. Klink *et al.*, Development and validation of a generic fluorescent methyltransferase activity assay based on the Transcreener® AMP/GMP Assay. *J. Biomol. Screen.* **17**, 59–70 (2013).

92. I. Bock *et al.*, Detailed specificity analysis of antibodies binding to modified histone tails with peptide arrays. *Epigenetics*. **6**, 256–263 (2011).

93. D. Hyllus *et al.*, PRMT6-mediated methylation of R2 in histone H3 antagonizes H3 K4 trimethylation. *Genes Dev.* **21**, 3369–80 (2007).

94. pET28a-LIC (Plasmid #26094). *Addgene*, (available at https://www.addgene.org/26094/).

95. F. M. Ausubel *et al.*, *Current Protocols in Molecular Biology Current Protocols in Molecular Biology* (John Wiley & Sons Inc., 2003).

96. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*. **9**, 671–675 (2012).

97. B. Xiao *et al.*, Structure and catalytic mechanism of the human histone methyltransferase SET7/9. *Nature*. **421**, 652–656 (2003).

98. E. Diaz *et al.*, Development and Validation of Reagents and Assays for EZH2 Peptide and Nucleosome High-Throughput Screens. *J. Biomol. Screen.* **17**, 1279–92 (2012).

99. K. Y. Horiuchi *et al.*, Assay development for histone methyltransferases. *Assay Drug Dev. Technol.* **11**, 227–36 (2013).

100. D. Barsyte-lovejoy *et al.*, (R) -PFI-2 is a potent and selective inhibitor of SETD7 methyltransferase activity in cells. *Proc. Natl. Acad. Sci.* **111**, 12853–12858 (2014).

101. K. Luger, T. J. Rechsteiner, a J. Flaus, M. M. Waye, T. J. Richmond, Characterization of nucleosome core particles containing histone proteins made in bacteria. *J. Mol. Biol.* **272**, 301–311 (1997).

102. T. M. Lakowski, A. Frankel, Sources of S-adenosyl-l-homocysteine background in measuring protein arginine N-methyltransferase activity using tandem mass spectrometry. *Anal. Biochem.* **396**, 158–160 (2010).

103. R. Alberty, The relationship between Michaelis constants, maximum velocities and the equilibrium constant for an enzyme-catalyzed reaction. *J. Am. Chem. Soc.* **75**, 1928–1932 (1953).

104. M. C. Reed, A. Lieb, H. F. Nijhout, The biological significance of substrate inhibition: A mechanism with diverse functions. *BioEssays*. **32**, 422–429 (2010).

105. M. B. C. Dillon, H. L. Rust, P. R. Thompson, K. A. Mowen, Automethylation of protein arginine methyltransferase 8 (PRMT8) regulates activity by impeding S-adenosylmethionine sensitivity. *J. Biol. Chem.* **288**, 27872–27880 (2013).

106. T. Kouzarides, Chromatin Modifications and Their Function. *Cell*. **128**, 693–705 (2007).

107. P. K. Mazur, O. Gozani, J. Sage, N. Reynoird, Novel insights into the oncogenic function of the SMYD3 lysine methyltransferase. *Transl. Cancer Res.* **5**, 330–333 (2016).