ENGINEERING LYSINE DEMETHYLASES TO ORTHOGONALLY PROBE CELLULAR FUNCTIONS

by

Megan Breski

BS, Shippensburg University, 2013

MS, University of Pittsburgh, 2015

Submitted to the Graduate Faculty of

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2017

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This dissertation was presented

by

Megan Breski

It was defended on

November 17, 2017

and approved by

Candace M Kammerer, PhD
Associate Professor
Human Genetics
Graduate School of Public Health
University of Pittsburgh

Beth Roman, PhD
Associate Professor
Human Genetics
Graduate School of Public Health
University of Pittsburgh

Jeremy J Martinson, DPhil
Assistant Professor
Infectious Diseases and Microbiology
Graduate School of Public Health
University of Pittsburgh

Committee chair
Zsolt Urban, PhD
Associate Professor
Human Genetics
Graduate School of Public Health
University of Pittsburgh

Zsolt Urban, PhD

ENGINEERING LYSINE DEMETHYLASES TO ORTHOGONALLY PROBE CELLULAR FUNCTIONS

Megan Breski, PhD

University of Pittsburgh, 2017

ABSTRACT

Post-translational modifications (PTMs) are an essential mechanism that increases the complexity of cellular processes. Protein methylation is a common PTM that can result in structural variation altering protein-protein/protein-DNA interactions. Several enzymes have been identified that have the ability to remove methyl groups from proteins. Many of these enzymes are histone demethylases, which are vital epigenetic modifiers. Lysine demethylases (KDMs) remove methyl groups from proteins, such as histones, resulting in a transition of chromatin accessibility for transcription factors. Jumonji C (JmjC) domain containing KDMs are a family of approximately 30 2-ketoglutarate (2-KG/2-OG) dependent enzymes known to demethylate histones within the chromatin nucleosomal core structure for the regulation of gene expression. Histone demethylases have tissue specific expression patterns and are active towards particular lysine residues, which suggest they are involved in specific cellular pathways. Irregular expression of KDM4s is associated with several types of cancer, but the roles of each KDM4 member in oncogenesis or cancer progression is convoluted. It is a difficult task to study an individual epigenetic enzyme and elucidate non-histone substrates due to the redundant cofactor utilization and their functional rapidity. To orthogonally study KDM4s, a bump-and-hole activation and inhibition system was developed. To elucidate non-histone catalytic

substrates of KDM4s, a photo-crosslinkable unnatural amino acid (UAA) was incorporated in the catalytic domain to induce covalent binding to substrates for subsequent identification. Elucidating the roles of individual epigenetic modifiers will greatly impact public health because it will allow for improved therapeutic interventions and diagnostic biomarkers.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to thank the chemistry department and the Dr. Kabirul Islam group for allowing this work to be performed. I have learned several techniques and concepts that will benefit me in my future career. I would also like to thank the post-doctoral and graduate students for teaching me various techniques that I was not familiar with. I give tremendous thanks to my friends and family for their lasting support, encouragement and prayers throughout my graduate school journey. I am blessed to have met so many inspiring people to keep me determined and focused to reach my goals, so I give many thanks to all of them.

1.0 INTRODUCTION

1.1 POST-TRANSLATION MODIFICATIONS

Protein post-translational modifications (PTMs) refer to reversible or irreversible chemical alterations to proteins in the form of particular chemical groups or even larger macromolecule additions, such as entire proteins like ubiquitin. Over 500,000 proteins have been identified to be post-translationally modified, with majority of the PTMs being glycosylation, phosphorylation, acetylation, and methylation (1). Ultimately, the role of PTMs is to enable signaling and regulatory mechanisms by altering protein function. PTMs are effective by changing protein dynamics, structure and binding competencies, requiring PTMs to be installed and removed systematically for optimal cellular health. There are over 200 types of PTMs identified with a large portion of them catalyzed by modifying enzymes (2). Mutations to the site of PTM as well as disrupted modifying enzyme activity can result in a plethora of diseases. Lysine mutations in the androgen receptor (AR) have been implicated in Kennedy's disease, an inherited neurodegenerative disorder due to a loss of acetylation. AR amino acid substitution K630A or both K632A and K633A result in a significant reduced rate of ligand-dependent nuclear translocation and disrupted proteosomal degradation due to the non-acetylated AR mutants misfolding and forming aggregates with several other proteins, including ubiquitin ligase E3 (3-6). Another example of PTM site mutation being implicated in disease involves

glycosylation. The etiology of autosomal dominant spongiform encephalopathy involves the loss of N-linked glycosylation in the prion protein (PRNP), due to the amino acid substitution T183A (7-10). One of the most well known PTMs is phosphorylation, which is added and removed by kinases and phosphatases, respectively. Protein phosphorylation predominantly occurs on serine residues, but is also seen on threonine and tyrosine (11). Kinases are involved in a plethora of cellular pathways, one of them being the cell cycle. In fact, there is an entire class of cyclin-dependent kinases (CDKs), which regulates cell cycle. There is abundant literature on CDKs having altered activity, resulting in a misregulated cell cycle leading to various cancers, making them great therapeutic targets (12-13). Clearly, PTMs and their related modifying enzymes are critical to proper cellular functioning. One area of specific interest is epigenetic PTMs and associated proteins.

1.2 EPIGENETICS AND POST TRANSLATIONAL MODIFICATIONS

The human genome contains 3 billion base pairs compacted into cells in the form of chromatin. Chromatin is made up of nucleosomal subunits consisting of -147 bp of negatively charged DNA wrapped around a positively charged histone protein octamer core (14-16). Gene expression patterns are dictated by the organization of chromatin, which can be highly dynamic. Euchromatin results in gene expression due to the accessibility of the DNA to various transcription factors, whereas, heterochromatin is highly compact, hindering transcription factor binding, resulting in gene silencing. Chemical modifications occur on both DNA and histones, known as epigenetic marks, to regulate transcription and are regulated by specific enzymes for

proper cell functioning. Epigenetics is the study of gene regulation, chromatin organization and phenotypic changes that do not involve changes in the DNA sequence (17-20).

DNA gets methylated by DNA methyltransferases (DNMTs) on cytosine bases adjacent to guanine (CpG), becoming 5-methylcytosine (5mC) and resulting in silenced chromatin. CpG sequences allow for the opposing DNA strands to harbor diagonal 5mCs. DNMTs either act as de novo DNMTs, which bind non-methylated DNA and generate the initial methylation pattern or act as maintenance DNMTs, which use an existing DNA strand as a template to copy the methylation pattern after DNA replication occurs. Promoters and regulatory components are the most common regions of the genome that become methylated, and it is hypothesized that methylated promoters will prevent transcription factors from binding and initiating gene expression. In contrast, various DNA binding proteins including transcription factors bind unmethylated CpGs to keep regions from becoming methylated, acting as protection for promoters or as regulation for genome interacting loci. DNA methylation is generally a stable modification, but the methylation patterns change dramatically during epigenetic reprogramming that occurs during fertilization. Epigenetic reprogramming is critical for imprinting mechanisms necessary for proper human development. DNA methylation patterns are lost when fertilization occurs to regain the pluripotency status, but the imprinting-related DNA methylation is re-established in the primordial germ cell stage (21-25).

The process of 5mC erasure involves three members of the ten-eleven translocation (TET) family, which are 2-ketoglutarate (2-KG), Fe(II) and oxygen dependent dioxygenases. These enzymes have been shown to successively oxidize 5mC to 5-hydroxymethyl cytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxycytosine (5CaC). The unmodified cytosine is

completely restored through base-excision repair mechanisms and potentially a decarboxylase that has yet to be identified (26-28).

1.3 HISTONE POST-TRANSLATIONAL MODIFICATIONS

Epigenetic modifications not only occur on DNA, but also on histones. The nucleosomal histone octamer core is composed two of each histone H2A, H2B, H3 and H4. It is important to note that nucleosomes are highly dynamic to allow for proper gene expression. As state earlier, the organization of chromatin can be dictated by DNA methylation status, but it is also highly dependent on what is called the 'histone-code'. The N-terminal tails of histones extend from the nucleosomal core, exposing them for posttranslational modifications (PTMs). With histones being an integral part of chromatin, it is important to understand how they are being modified epigenetically and what it results in phenotypically. The most common epigenetic marks found on histones are acetylation and methylation on arginine and lysine residues. The addition and removal of these marks result in various phenotypes depending on the site of the modification and often coordinate with the DNA methylation status for efficient chromatin dynamics. Different families of chromatin modifying proteins are involved in the installation, recognition and removal of the epigenetic mark depending on the chemical group and the amino acid on which the modification resides (29-33).

PTMs can also be in the form of acetylation, methylation, phosphorylation, ubiquitination, sumoylation and proteolytic processing and vary depending on histone (H2A, H2B, H3, H4), cell type, developmental stage and disease. Together the PTMs make up the

‘histone code.’ Each PTM has an independent function, but it is important to recognize that the downstream effects are due to the crosstalk and combination of specific PTMs.

Negatively charged acetyl groups are added to lysine residues on proteins, changing the electronic properties of the substrate and, subsequently, binding partner affinity. The antagonistic relationship between histone acetyltransferases (HATs) and histone deacetylases (HDACs) regulates this epigenetic mark. HATs catalyze the addition of acetyl groups to histone lysine residues, whereas, HDACs catalyze their removal. When histones are acetylated, the positively charged histone tails are neutralized, decreasing the affinity for negatively charged DNA (34-37). There are various methylation states that histone residues exhibit: mono- (me1), di- (me2) and tri- (me3) and are most abundant on histone H3 and H4 lysine or arginine residues. Methylated lysine residues have also been observed on nucleosome linker histone H1. These modification states dictate which effector proteins recognize the modification and vary throughout development, tissue type and pathologies. Methylation modification state can increase (via methyltransferases activity) and decrease (via demethylase activity), providing a dynamic chemical process essential for various biological pathways (38-42).

Unlike histone acetylation, methylation does not affect the electronic properties of the histone. Instead, there are several effector proteins that recognize and read the methylation status to induce successive events, such as protein recruitment for chromatin silencing or activation. Because of this phenomenon, chromatin state is not dictated solely by the presence of a methyl group, but is dependent on the location of the methylation and the degree of methylation (Figure 1). Generally, H3K9 and H3K27 methylation are associated with transcriptional silencing, but both have been observed in some regions of transcriptionally active chromatin. Also, methylation of H3K4 and H3K36 are associated with transcriptional

activation, but can be seen in chromatin regions of transcriptional silencing. The variation is dictated by the location of methylation mark in relation to gene elements – promoter region vs. gene body - and other PTMs that are present (42-60).

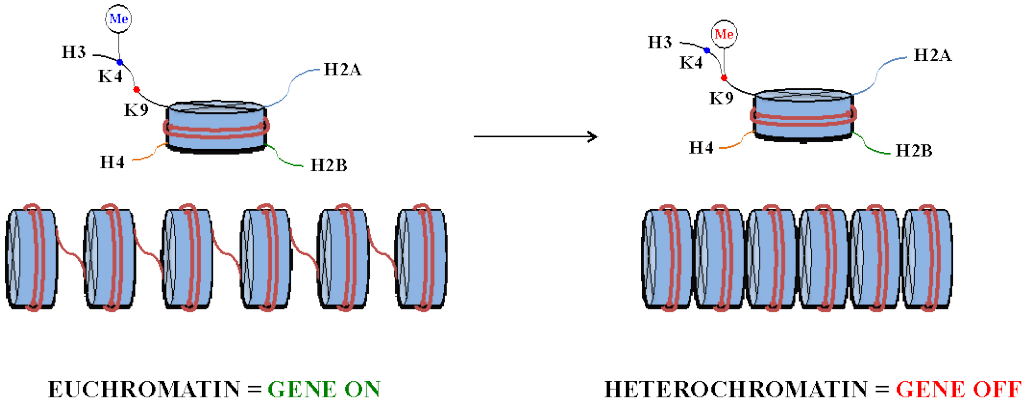


Figure 1: Chromatin organization and PTMs.

Post translational modifications occur on different sites within nucleosomal histone tails. Methylated lysine 4 of histone H3 generally results in a euchromatic chromatin state allowing for transcription to occur. This methyl group can be removed and different lysine residues can be methylated to result in a different chromatin organization. Methylated histone H3 lysine 9 results in a compact chromatin organization, keeping transcription from occuring.

1.3.1 Histone Lysine Methylation dynamics

Epigenetic modification patterns dictate differentiation and vary depending on developmental stage, cell type and tissue type. Several proteins are involved in the dynamic processes that regulate the addition and removal of the covalent epigenetic modifications and these proteins interact with each other to influence downstream pathways such as transcription, replication, and expression at the gene and protein level. Specific proteins that ‘write’, ‘read’ and ‘erase’ the chemical mark tightly regulate each PTM. In relation to histone lysine methylation, the ‘writers’ are lysine methyltransferases (KMTs) and the ‘erasers’ are lysine demethylases

(KDMs) (Figure 2). Lysine methyl readers can be found as domains within KMTs/KDMs or could be proteins without enzymatic function.

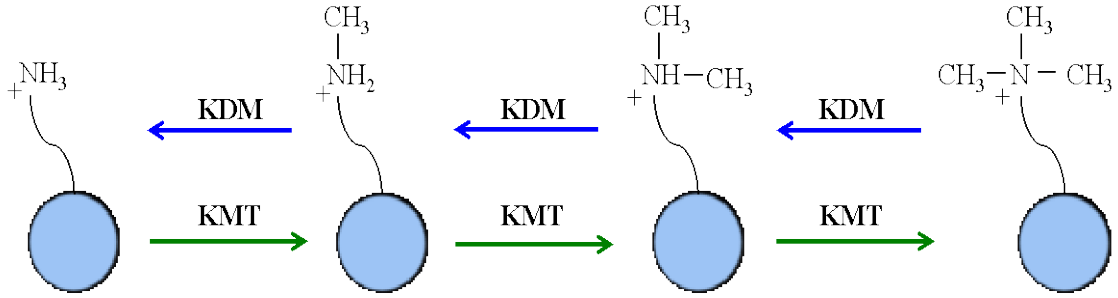


Figure 2: Lysine Methylation Dynamics.

Lysine residues of proteins get methylated and demethylated by writers (KMTs) and erasers (KDMs), respectively. KMTs can add 1, 2 or 3 methyl groups making the lysine mono-, di-, or tri- methylated. Similarly, KDMs can remove these methyl marks successively. Different KMTs and KDMs add and remove methyl groups depending on protein substrate, site of lysine, and degree of methylation.

More than 50% of the chromatin-modifying enzymes in humans are Protein methyltransferases (PMTs), which transfer methyl groups onto various protein substrates acting as 'writers'. The enzymatic activity of these enzymes is dependent on the evolutionarily conserved catalytic SET [Su(var)3-9, Enhancer-of-zest, Trithorax] domain. There are seven main families of lysine methyltransferases (KMTs) including EZ, SET1, SET2, SMYD, SUV39, SUV4-20, RIZ. SET domain containing enzymes aid in the generation of bivalent chromatin for genes to be poised for both activation and repression. Each KMT acts site specifically adding a specific number of methyl groups to mono-, di- or tri- methylate (61-68).

Methyltransferase G9a, also known as EHMT2, transfers two methyl groups from universal methyl donor S-adenosylmethionine (SAM) onto lysine residues through its SET (Su (var), enhancer of zeste, trithorax domain) domain. Specific amino acid residues within the

histone substrates are essential for G9a activity. For instance, H3K9 methylation by G9a is abolished when the neighboring arginine (H3R8) is substituted for any other amino acid. This shows that methyltransferases have site specificity for a specific sequence motif similar to DNA methyltransferases such as DNMT1 and DNMT3 (69-72).

The motif recognized by G9a is found in several non-histone proteins and they have been observed to be methylated. In fact, G9a is automethylated at its N-terminus, resembling tri-methylated H3K9 (H3K9me3) (73). Other non-histone protein substrates and interacting partners for G9a include imprinting associated proteins DNMT1, HDAC1 (Histone deacetylase 1), and p53 (74-75).

The interaction of G9a with other proteins clearly demonstrates the crosstalk of chromatin marks. As DNA get replicated, DNMT1 is responsible for maintaining the DNA methylation pattern on daughter strand DNA, while G9a methylates H3K9. G9a and DNMT1 localize and bind to the chromatin simultaneously during replication. Additionally, these proteins were shown to form a ternary complex with loading factor proliferating cell nuclear antigen (PCNA) and the knockdown of DNMT1 perturbs the complex binding and function resulting in insufficient replication (76). Tumor suppressor p53 has been shown to induce expression of the well characterized imprinted gene PEG3. PEG3 is highly expressed in the brain and is important for specific neuronal development in the hypothalamus14. G9a and Euchromatin Histone methyltransferase 1 protein (EuHMTase1, also known as G9a like protein - GLP) dimethylate p53 at the C-terminal lysine residue 373 resulting in an inactive p53 protein, reduced apoptosis and reduced Peg3 (77-78).

G9a targets euchromatic regions and methylates H3K9 and H3K27 resulting in gene silencing. G9a mediated H3K9 methylation negatively regulates H3K4 methylation supported

by G9a knockout (G9a-/-) resulting in a significant increase in H3K4 methylation. G9a-/- mice are non-viable and die at an early embryonic age of E9.5-E12.5 making it likely that G9a mediated methylation is an essential mechanism regulating gene expression in early development. G9a and GLP are frequently found as homo- or heterodimers or in transcriptional regulatory complexes such as E2F6, CtBP-1, and CDP/cut. The E2F6 complex is related to de novo CpG DNA methylation by recruiting DNMT3 to proximal promoters where the complex is bound. Loss of G9a or GLP mediated H3K9 methylation result in lethality potentially due to a gross relocalization of heterochromatin protein 1 (HP1), an important protein for position-effect variegation and x-chromosome inactivation; significant reduction of H3K9 mono- and dimethylation and DNA methylation; and induction of Mage-a genes, which are clustered on the X-chromosome and are associated with HDAC chromatin binding and the Notch signaling pathway. Interestingly, point mutations in the SET domain of G9a that render the enzyme catalytically inactive have been expressed in G9a(-/-) embryonic stem cells (ESCs) and the mutant can moderately restore DNA methylation levels, but not H3K9me2 levels suggesting that the enzymatic activity of G9a is partially dispensable for DNA methylation. G9a also has an Ankyrin repeat domain (ANK) that plays a key role in recognition and adaption to interacting partners. Within the ANK domain, there is a hydrophobic cage that recognizes and strongly binds to methylated N-terminal H3K9 peptides. G9a has also been shown to recruit and directly interact with de novo DNA methyltransferases DNMT3a and DNMT3b via its ANK domain and when ANK is deleted, de novo DNA methylation is impaired, but not completely lost. The deletion of the ANK domain does not seem to have a dramatic effect on H3K9me3 levels and because HP1 recognizes this histone methyl mark and can recruit DNMTs, this may explain why the DNA methylation is not completely missing. There seems to be two unique

epigenetic mechanisms G9a uses for gene repression. G9a methylates H3K9 through its SET domain, which will eventually lead to HP1 binding chromatin and recruiting DNMTs for CpG methylation and, secondly, by directly recruiting and interacting with DNMTs for de novo DNA methylation. Histone methylation can regulate gene expression by antagonizing the effects that other histone modifications (e.g. histone acetylation) have on gene expression through the binding of chromodomains (79-90).

Histone KDMs act as 'erasers' by removing methyl groups in a site and state specific manner similar to the KMTs (Table 1). Many histone demethylases harbor 'reader' domains, such as the Tudor domain, to recognize the modifications and act as effector proteins. Histone

Table 1: Histone Lysine Writers and Erasers.

Histone Lysine Writers and Erasers		
Site of Modification	**Writers**	**Erasers**
H1K186me1	EHMT1, EHMT2	
H1K25me1	EZH2, EHMT2, EHMT1	KDM4D
H3K27me1	EZH1, EZH2,EHMT1, EHMT2	
H3K27me2	EZH1, EZH2, NSD3	PHF8, KDM6B, KDM7
H3K27me3	EZH2, NSD3	KDM6A, KDM6B
H3K36me1	ASH1L	
H3K36me2	SETMAR, NSD1, SMYD2, ASH1L	KDM2A, KDM2B, KDM4A, KDM4B, KDM4C, KDM8
H3K36me3	NSD2	KDM4A,KDM4B, KDM4C, NO66
H3K4me1	SETD7	KDM1A
H3K4me2	NSD3	KDM5A, KDM5D, KDM1A
H3K4me3	MLL, MLL3, MLL4, PRDM9, SETD1A, SETD1B, SMYD3	PHF8, KDM2B, KDM5A, KDM5B, KDM5C, KDM5D, NO66
H3K79me1	DOT1L	
H3K79me2	DOT1L	
H3K79me3	DOT1L	
H3K9me1	EHMT2	KDM3A, KDM3B,KDM1A
H3K9me2	EHMT1, EHMT2, PRDM2	PHF8, KDM3A, KDM3B, KDM4A, KDM4B, KDM4C, KDM4D, KDM7, KDM1A, KDM1B
H3K9me3	SETDB1, SETDB2, SUV39H1, SUV39H2	KDM4A, KDM4B, KDM4C, KDM4D, KDM4E
H4K20me1	SETDB, NSD2	PHF8
H4K20me2	SETDB, SUV420H1, SUV420H2, NSD1	
H4K20me3	SUV420H1, SUV420H2, NSD2	

Table showing different histone methylation sites and degrees and their respective writers and erasers

demethylation is an important biological phenomenon and will be the focus of the document and expanded upon in the upcoming chapters.

1.4 LYSINE DEMETHYLASES

Lysine methylation was originally thought to be an irreversible modification, but the discovery of Lysine Specific Demethylase 1 (KDM1A or LSD1) in 2004 allowed for the understanding that this is a regulated mark that can be added and removed. KDM1A catalyzes the removal of methyl groups from H3K4me1 and H3K4me2 as well as some non-histone proteins such as p53 and DNMT1. KDM1A is a flavin-dependent amine oxidase and consists of an amine oxidase catalytic domain, an N-terminal SWIRM domain involved in recognition of the specific target and an N-terminal extension. KDM1A can be thought of as an inducer of transcriptional silencing because it specifically targets H3K4me1/2, which is known to be associated with transcriptional activation, and has been shown to be involved in protein-protein interactions with transcriptional repressive complexes (e.g. CoREST and HDAC1/2) (91).

Following the discovery of KDM1A, several more lysine demethylases were identified. There are over 30 enzymes discovered that have the ability to remove methyl marks from lysine residues, but the mechanism of action is completely different compared to KDM1A and KDM2B. Almost all of the KDMs identified harbor a catalytic Jumonji C (JmjC) domain and are dependent on molecular oxygen, iron and metabolite 2-ketoglutarate (2-KG). The human genome encodes 32 JmjC-domain containing genes with almost all of them showing demethylase activity. There are various combinations of domains (e.g. PHD, Tudor, ARID,

JmjN) incorporated in JmjC-domain containing proteins, which define their function. The lysine demethylases are categorized into seven subfamilies based on structure and sequence homology and each of them target specific methylated lysine residues based on sequence and methylation state (Figure 3) (92).

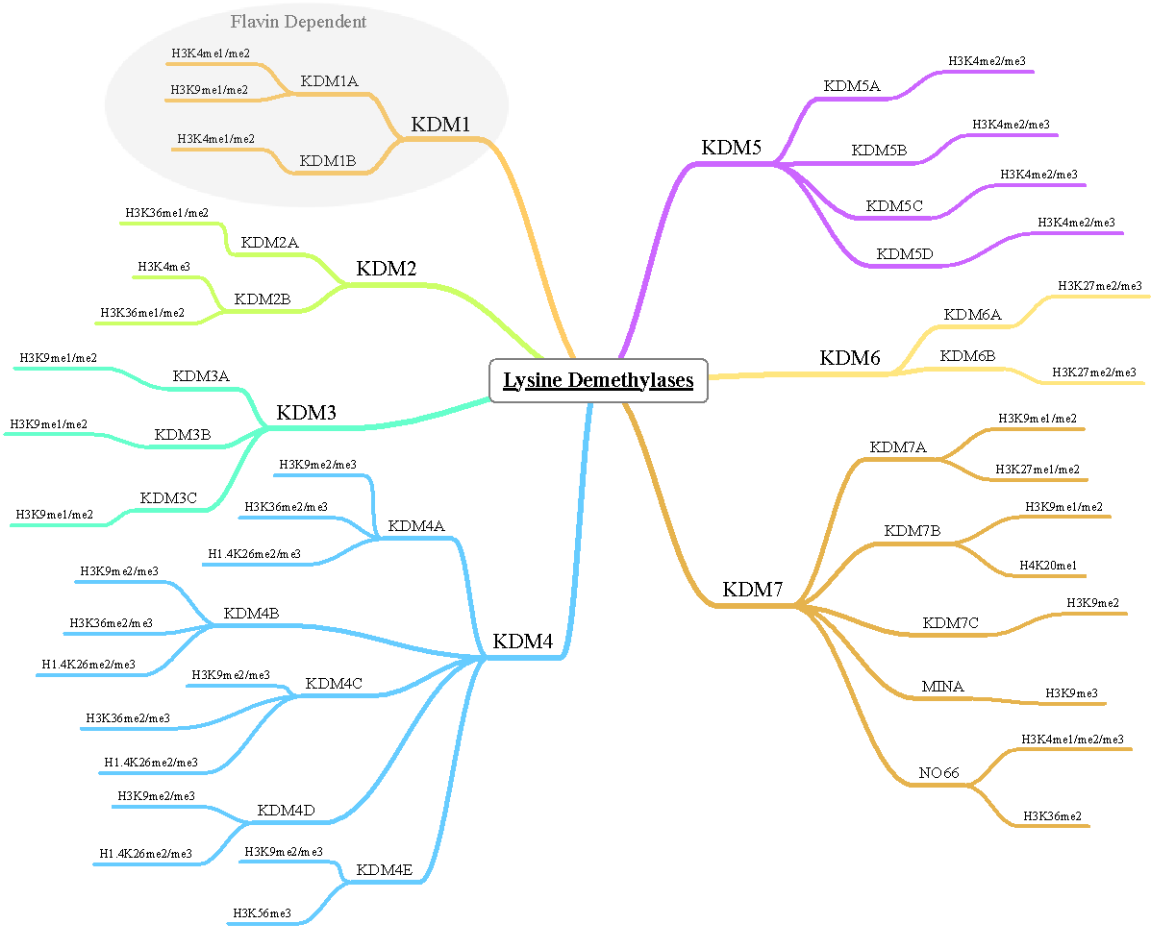


Figure 3: Lysine Demethylase Subfamilies and Substrates.

The nodal map shows the subfamiles of demethylases with their corresponding members and substrates. The flavin dependent lysine demethylases are shown in the grey circle. All other lysine demethylases are 2-ketoglutarate dependent.

Majority of the KDM proteins are JmjC domain containing demethylases requiring Fe(II), 2-KG and molecular oxygen (O_2) for the catalysis of a dioxygenase reaction that results in demethylated lysine residues on histones. Within the catalytic domain, a His-Glu/Asp-His triad positions the Fe(II) allowing 2-KG and oxygen to bind forming highly reactive oxoferryl (IV). 2-KG is decarboxylated, releasing succinate and carbon dioxide. The methyllysine target is hydroxylated, resulting in an unstable intermediate (hydroxymethyl), which spontaneously decomposes to formaldehyde, producing the demethylated product in either the di-, mono- or non-methylated state. To prevent the accumulation of formaldehyde, Formate Dehydrogenase (FDH) utilizes NAD+ to oxidize formaldehyde and ultimately produce additional carbon dioxide (Figure 4) (93-94).

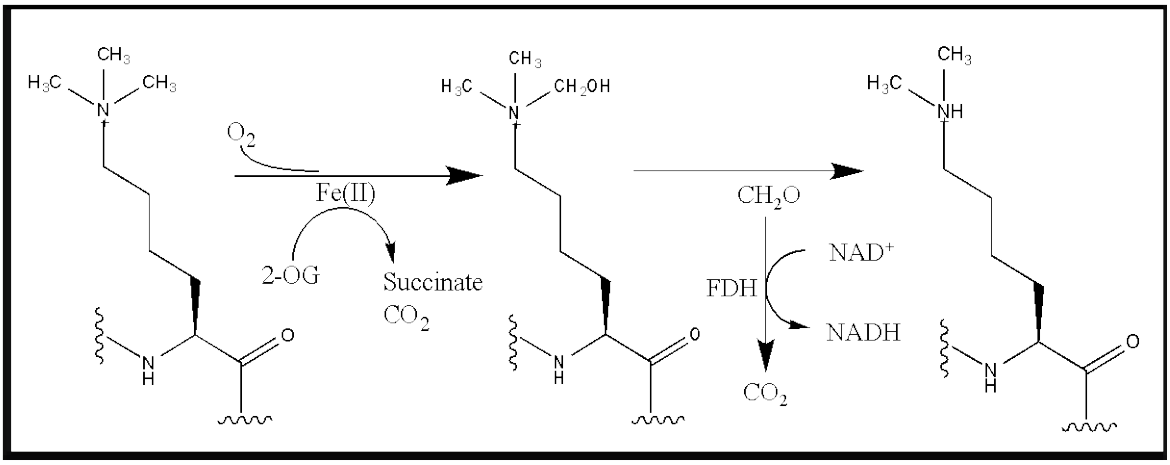


Figure 4: JmjC Domain Containing Enzymatic Mechanism.

1.4.1 Lysine Demethylase Family 4

There are five functional KDM4 genes: KDM4A-E. Orthologs of KDM4A-C are seen in all vertebrates whereas orthologs of KDM4D are seen only in placental mammals providing

support that the five membered KDM4 subfamily is highly conserved. KDM4s are JmjC domain containing proteins with several unique features to each of the five members (93-95).

The domains within KDMs play essential roles for correct protein function. KDM4A-C contain N-terminal JmjN and JmjC domains and on the C-terminus contain a double PHD and double Tudor domain. KDM4D and KDM4E lack the PHD and Tudor domains (Figure 5). The JmjC domain is the catalytic domain and is therefore required for demethylase activity. Structural integrity of the protein is provided by the JmjN domain, which also interacts with the catalytic domain for proper substrate binding positioning. The PHD and Tudor domains are important for localization and protein-protein interactions by acting as reader domains and recognizing specific methylated lysine residues. The KDM4A Tudor domains recognize and bind repressive histone marks H3K4me2/me3 and H4K20me2/me3, whereas the PHD domains bind methylated and unmodified residues on many histone tails, providing a more flexible recognition pathway of histone modifications (93-95).

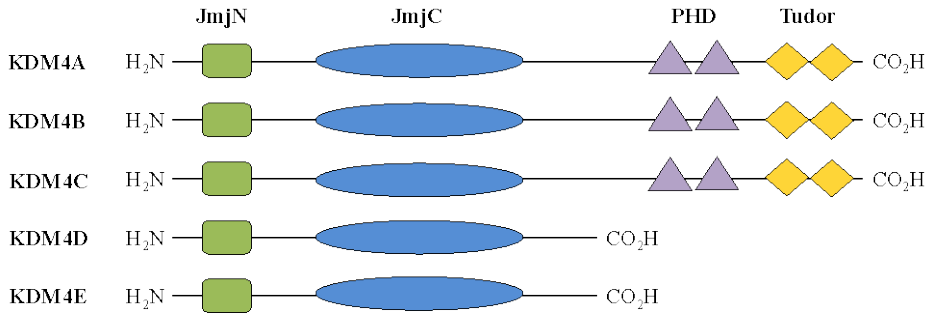


Figure 5: KDM4 family of enzymes and linear protein domain map.

KDM4s are considered transcriptional repressors because of their target substrates. KDM4A-C demethylate H3K9me2/me3 and H3K36me2/me3, where KDM4D and –E demethylate H3K9me2/me3. Crystallography has shown differences in the surface residues of

the histone binding pocket between KDM4A-C and KDM4D-E. The structural variation within the substrate binding pocket of KDM4 members explains the electrostatic differences between the KDM4s causing interactions with substrates to be different. The variation of substrate-enzyme interactions may explain the difference in activity toward H3K36. KDM4A demethylates trimethylated lysine residues more efficiently than dimethylated residues and has approximately 5-fold higher activity towards H3K9me3 than H3K36me3. X-ray crystallography analysis of KDM4A shows that the electrostatic environment in the methylated histone binding pocket does not allow productive positioning toward the iron-containing catalytic domain of monomethylated lysine residues. This further supports the methylation state specificity of these KDM4 proteins (92-95).

1.4.1.1 Expression and Localization Patterns of KDM4s

KDM4s have both redundant and unique expression patterns at the tissue, cellular, and genomic level. KDM4s predominantly localize to the nucleus to regulate chromatin dynamics, but they are also found in the cytosol where they regulate important cellular mechanisms such as protein translation, making their functions diverse and widespread. The normal expression patterns of the KDM4 genes suggest that the associated proteins have non-overlapping biological functions. KDM4A-C are ubiquitously expressed in human tissues with the highest expression being in the spleen, ovaries and colon. KDM4D and KDM4E are mainly expressed in the testes, but KDM4E expression is low relative to the other KDM4 genes. Knockout and transgenic studies have been performed in various biological models such as *D. melanogaster, C. elegans,* and mice to study the physiological functions KDM4s. Double homozygous mutants of both *Drosophila KDM4* orthologs, *dKDM4A* and *B,* are not viable and die in the second larval stage. *C. elegans* contain a single *KDM4* gene and depletion of this gene results in germ line

apoptosis and slows DNA replication. Conditional heart-specific *KDM4A* knockout in mice as well as transgenic mice have demonstrated that KDM4A promotes cardiac hypertrophy (96-103).

Vertebrate KDM4A–C proteins contain a conserved double tudor domain and a potential zinc-finger domain at the carboxy terminus, which mediate binding to specific histone modification states. Deletion of the C-terminal domain in KDM4 proteins result in a change of sub-cellular localization, changed demethylase activity and disruption of other KDM4 functions (104-107). KDM4A, KDM4B and KDM4C double tudor domains display different histone-binding preferences. KDM4B has been shown to co-localize with heterochromatin mark H3K23me3. H3K23me3 binding by KDM4B stimulates the demethylation of H3K36 and serves as a link to mammalian germ cell development. The double tudor domain of KDM4C has shown highly specific binding to H3K4me3 and has been associated with mitotic chromatin at transcription start sites (108).

1.4.1.2 Protein Complexes Containing KDM4s

The KDM4s have also been shown to interact with a diverse array of proteins including those associated with transcription or DNA repair, which can direct their genomic localization or regulate cellular pathways outside of gene expression. Kim et al. recently showed that KDM4A promotes prostate tumorigenesis by interacting with transcription factor, ERG (erythroblast transformation-specific related gene) through the up-regulation of Yes-associated protein 1 (YAP1) by decreased H3K9me3 at the YAP1 promoter (109).

KDM4A has been shown to be an E2F1 coactivator and the KDM4A-E2F1 complex controls tumor metabolism. KDM4A and E2F1 associate together on target gene promoters and

enhance E2F1 chromatin binding and transcriptional activity. Pyruvate dehydrogenase kinases (PDKs) PDK1 and PDK3 are direct targets of KDM4A and E2F1 and control the switch between glycolytic metabolism and mitochondrial oxidation. Downregulation of KDM4A leads to elevated activity of pyruvate dehydrogenase and mitochondrial oxidation, resulting in excessive accumulation of reactive oxygen species providing mechanistic insights to the role of KDM4A in prostate cancer (110).

KDM4s can act as both a transcriptional activator and repressor depending on interacting proteins. When forming complexes with androgen receptor (AR) or estrogen receptor (ER), KDM4A stimulates their transcriptional activity and induces the expression of target genes that are important for proliferation in prostate and breast cancer. KDM4A acts as a transcriptional repressor when binding with repressive factors such as nuclear receptor corepressor (N-CoR) and histone deacetylases (111-113). KDM4 catalytic activity is not always required for transcriptional regulation. The KDM4 family can act as signaling proteins to recruit various effector proteins to carry out different functions.

Chromatin immunoprecipitation (ChIP) experiments suggest that KDM4B regulates mitotic clonal expansion by forming a complex with DNA binding transcription factor, CCAAT/enhancer-binding protein (C/EBPβ) to regulate cell cycle genes *Cdc45l* and *Cdc25c* (114). Additionally, molecular chaperone heat shock protein 90 (Hsp90) interacts with KDM4B as for protein stabilization. KDM4B is ubiquitinated on lysines 337 and 562 for proteasomal degradation when Hsp90 is inhibited, which alters histone methylation patterns and provides evidence for a regulatory pathway that could be targeted therapeutically (115).

Self-renewal and pluripotency of embryonic stem cells (ESCs) are mechanisms regulated at several levels including histone modification regulation. A functional RNAi screen

for histone demethylases identified five potential demethylases essential for mouse ESC identity. KDM4B and KDM4C are necessary for self-renewal of ESCs and induced pluripotent stem cell generation. Genome-wide occupancy studies reveal target sites unique to KDM4B and KDM4C as well as common to both KDM4B and KDM4C belong to functionally separable Core, Polycomb repressive complex (PRC), and Myc regulatory modules, respectively. KDM4B and Nanog act through an interconnected regulatory loop, whereas KDM4C interacts with PRC2 to result in transcriptional repression. The Polycomb repressive complex 2 (PRC2) catalyzes di-methylation and tri-methylation of lysine 27 of histone H3 and plays an essential role in the maintenance of repressive chromatin states. The core components of the PRC2 complex are the H3K27 methyltransferase EZH2, SUZ12, EED and RBAP46/48. Other accessory proteins of PRC2 include PHF1, JARID2 and AEBP2. This demonstrates that members of the KDM4 subclass exhibit distinct and combinatorial functions as well as highlights the mechanisms of functional crosstalk between the different enzymes and reader domains (116-117).

KDM4A-C have also been shown to have non-histone substrates. Ponnaluri, et al. demonstrated KDM4A-C catalytic activity towards methylated non-histone peptides from WIZ, CDYL1, CSB and G9A42, which suggests remarkable flexibility in substrate recognition (118). Homo- and hetero-dimers of the KDM4 protein members can form resulting in another regulatory mechanism. KDM4A-C can become incorporated into multiprotein complexes with the SWI/SNF chromatin-remodeling complex and can interact with inhibitory complexes such as histone deacetylase complexes (HDACs). The formation of dimers and incorporation into functionally diverse multiprotein complexes supports that KDM4 proteins have a significant role in directing gene expression in development, homeostasis and pathology (119-124).

1.4.1.3 KDM4 Inhibition

Many histone demethylase inhibitors have been designed and are classified into five groups: 2-ketoglutarate analogs, iron chelators, zinc chelators, catalytic domain inhibitors, and prodrugs. Synthetic N-oxalylglycine (NOG) is an amide analog of 2-KG making it a competitive inhibitor of the 2-KG dependent oxygenases. NOG has been identified as a natural product in rhubarb and spinach, but is not present in human HEK293T cells (125). Because NOG is a 2-KG competitive inhibitor, it is not selective for particular families of oxygenases. Several groups have developed NOG derivatives to increase the selectivity of NOG towards KDMs. Hamada et. al developed four NOG derivatives providing increased selectivity towards KDM4s. Dimethylester derivatives of the NOG analogues increased cellular permeability and demethylase inhibition in cellular assays observed by the accumulation of H3K9me3 (126). This work supports that 2-KG competitive inhibitors can be altered to specifically inhibit protein families and suggests that upon increased engineering, specifically inhibit individual proteins.

Natural metabolites induce specific cellular responses and can act as KDM inhibitors. Another 2-KG competitive inhibitor is 2-hydroxyglutarate (2-HG), which is an oncometabolite. Isocitrate dehydrogenase 1 and 2 (IDH1/2) are $NADP^+$-dependent enzymes that catalyze oxidative decarboxylation of isocitrate to 2-KG. Somatic mutations in IDH1/2 (R132H) have been identified in several cancers, which result in a gain-of-function allowing for the conversion of 2-KG to D-2-HG. The accumulation of D-2-HG becomes so overwhelming in the cell where the capacity of D-2-HG dehydrogenase is flooded and the enzyme cannot effectively perform to minimize 2-HG to safe levels (127-129). Hypoxia can induce the production of L-2-HG in an IDH1/2 independent fashion. Instead, hypoxia induces promiscuous substrate usage by lactate

dehydrogenase A (LDHA) and LDHA-mediated 2-KG reduction (130). Both mechanisms to increase 2-HG levels in the cell result in reduced 2-KG levels and 2-KG utilization inhibition causing disrupted chromatin organization.

One of the most potent broad range 2-KG dependent oxygenase inhibitors identified is known as 5-carboxy-8-hydroxyquinoline (IOX1). Cellular assays show that IOX1 is active against cytosolic and nuclear 2-KG oxygenases without esterification. IOX1 can translocate the active site Fe(II), which is a rare induced metal movement, to disrupt 2-KG and substrate binding (131). IOX1 is highly potent towards KDM4 proteins *in vitro*, but the potency is 100-fold lower in cellular assays, likely due to the poor cellular permeability characteristics. Schiller et al developed n-octyl ester derivative of IOX1 making it more hydrophobic and increase the cellular permeability. With this modification, the cellular EC_{50} value improved from 100 μM to 4 μM (132). The reported therapeutic inhibitors lack selectivity and specificity towards histone demethylases, which becomes problematic regarding nondesirable targets. Therapeutic inhibitors for histone demethylases remain in the preclinical phase for various types of cancer to be included in therapeutic cocktails.

1.5 KDMS AND PUBLIC HEALTH

Gene and protein expression dictate cellular phenotype and cellular health. Chromatin dynamics and post-translational modifications are critical complex mechanisms responsible for the expression of genes and specific protein functions. KDMs are post-translational modifying enzymes involved in several cellular processes including chromatin dynamics, protein translation and protein-protein interactions. KDMs are critical for ideal cellular health and

human development and have been implicated in diverse pathological phenotypes. KDM4s have been associated with mental, cardiac, neurological, and viral disorders as well as with several cancers.

1.5.1 KDM4s and Cancer

The overexpression or underexpression on the gene and protein level of the KDM4 subfamily has been observed in several pathological phenotypes. There is extensive literature on KDM4 involvement in cancer development and progression. KDM4A overexpression and amplification is seen in several cancers such as ovarian, breast, squamous cell carcinomas, prostate, lung and colon cancer (133-139). KDM4A overexpression triggers localized chromosomal instability, specifically 1q12, 1q21 and Xq13.1 copy gains, which are rich with oncogenes. These localized copy gains can occur due to reduced HP1γ binding to chromatin and an extended S phase. When HP1γ cannot bind to chromatin, it results in an open, accessible chromatin state allowing for the re-replication of oncogenic regions. HP1γ bound to chromatin hinders KDM4A from being able to modify histones. The antagonistic behavior between KDM4A and HP1γ may contribute to the explanation as to how KDM4A encourages DNA replication (100, 105).

Overexpression of KDM4 proteins also disrupts the DNA mismatch repair (MMR) pathway resulting in poor chromatin integrity. Mismatch recognition proteins bind H3K36me3 to ensure proper DNA MMR occurs. When KDM4C is overexpressed, H3K36me3 is less abundant hindering mismatch recognition proteins from binding appropriate regions. The dysregulated pathway allows for mutations to go overlooked and promote genomic instability. Proper DNA MMR can be restored through the downregulation of KDM4C. Because KDM4A

and KDM4B demethylate H3K36me3, similar results are expected with the overexpression of these proteins (133, 140-142).

KDM4A overexpression can also allow cells to evade apoptosis or decreases apoptosis all together by extending the G2/M phase. The activating protein 1(AP1) transcription factor is a heterodimer of the JUN and FOS proteins, which controls apoptosis, cell proliferation and differentiation. AP1 is induced by a positive feedback loop where KDM4A is a priming factor. KDM4A can indirectly induce the expression of AP1 genes because histone demethylation facilitates AP1 binding to JUN and FOSL1 promoters, which stimulates the feedback loop. The overexpression of KDM4A can result in the overexpression of AP1 genes, which will promote cell growth and metastasis (134).

KDM4A can repress the expression of tumor suppressors, such as CHD5, by impeding senescence. KDM4A can directly bind with tumor suppressors (e.g. p53) rendering them inactive and can form complexes with the androgen and estrogen receptors making them more sensitive resulting in the overexpression of oncogenes. ER-positive breast cancer cells with depleted KDM4A showed decreased expression of oncogenes c-Jun and cyclin D1. Proliferation inhibition of ER-negative and ER-positive breast cancer cells was observed with KDM4A knockdown. These results show that KDM4A is necessary for ER-positive and –negative breast cancer and supports the idea that there are multiple pathways that KDM4A is involved in, which can influence breast cancer development and/or progression. Contrasting with the results for breast cancer, the overexpression of KDM4A in HeLa cervical carcinoma did not result in tumor growth. With opposing results for the overexpression of KDM4A in different cancer cell lines, general conclusions cannot be made based on the expression of this protein, but rather it is cancer and tissue specific (135, 143).

KDM4B plays a significant role in nuclear receptor responsiveness in breast and prostate cancer. KDM4B is highly expressed in estrogen receptor (ER)-positive breast cancer subtypes, can bind to the ER, act on the repressive mark of H3K9me3, and recruit chromatin remodeling complexes. The activity of KDM4B when bound with the ER can induce oncogenic expression (e.g. MYB and MYC) resulting in proliferation. KDM4B expression level is positively correlated with prostate cancer severity, potentially due to it interactions with the androgen receptor (AR). AR transcription is upregulated and the AR is stabilized when cooperating with KDM4B via degradation inhibition. KDM4B is involved in metastasis and hypoxia and can promote cell survival (144-147).

The overexpression of KDM4C has been observed in breast and prostate cancer, esophageal squamous cell carcinoma, metastatic lung sarcomatoid carcinoma, non-solid tumors and medulloblastoma. KDM4C can promote tumorigenesis by activating oncogenes (e.g. MDM2), binding the AR nuclear receptor and establishing stem cell-like phenotypic characteristics. Overexpression of KDM4C resulted in hypomethylation of H3K9 and knockdown studies in breast cancer cells results in the inhibition of proliferation and metastasis (148-152).

Cancer cells are known to have an uncanny ability to evade apoptosis. Tumor Necrosis Factor Related Apoptosis Inducing Ligand (TRAIL) binds to cell surface receptors (DR4/DR5) to induce the caspase cascade that ultimately triggers apoptosis. Decoy receptors (DcRT) are also present on the cell surface as an internal regulatory mechanism to avoid excessive apoptosis. TRAIL expression is decreased in several cancer types including acute myeloid leukemia and prostate cancer. Targeting TRAIL as a cancer therapeutic strategy is not a new concept, but there are major limitations with this path such as compound stability, small

molecule specificity and toxicity. When AML cells were treated with DNA methyltransferase (DNMT) inhibitors and Histone Deacetylase (HDAC) inhibitors, TRAIL expression increases leading to increased apoptosis. Additionally, histone demethylases (KDM4A/B) are upregulated in several cancers including prostate cancer. Targeting KDM4A/B in prostate cancer cells results in an inhibition of proliferation. Because TRAIL seems to be regulated epigenetically, groups have investigated if and how KDM4s regulate TRAIL induced apoptosis. Previously, a KDM4A/B inhibitor termed C-4 was developed. TRAIL and DR5 mRNA and protein levels increase in a C-4 concentration dependent manner in prostate cancer cell lines, whereas C-4 has no effect on the levels of the monitored genes/proteins in normal cells. Additionally, IP experiments showed that C-4 increases cell surface DR5 abundance. Silencing KDM4A in prostate cancer cells shows a significant increase in TRAIL and DR5 mRNA and protein expression, whereas, silencing KDM4B does not result in an increase in these apoptosis related factors. Both knockdown and inhibition of KDM4A result in induced caspase-8 activation and apoptosis in cancer cells with no effect on normal cells. Apoptosis and caspase cleavage is only observed when both TRAIL and KDM4A/B inhibitor was applied in prostate cancer cells. Apoptosis, TRAIL expression, and de-phosphorylation of apoptosis related proteins dramatically increases when both TRAIL is expression is induced with ONC201 compound and KDM4A/B are inhibited with C-4. ChIP experiments show that KDM4A binds to the TRAIL and CHOP promoter. CHOP transcriptionally upregulates DR5. Inhibiting KDM4A (siRNA or C-4) in cancer cells results in significantly higher CHOP mRNA and protein levels. KDM4A inhibition by C-4 also results in an increase in histone epigenetic modifications in the promoter regions of TRAIL and CHOP. This shows that KDM4A is a key epigenetic silencer of TRAIL

in cancer cells, KDM4A regulates DR5 via CHOP epigenetic modifications and targeting KDM4A increases tumor sensitivity to TRAIL-initiated apoptosis therapies (153-155).

It has become clear that the KDM4 subfamily members play various and diverse roles in cancer development and progression making them potential therapeutic targets or biomarkers. When considering the therapeutic strategies involving KDM4s, it is important to keep in mind that the activity and interacting proteins are context dependent and vary depending on cancer type.

1.6 BUMP HOLE APPROACH

Approximately 20,000 protein-coding genes give rise to massive diversity among proteins, which are responsible for the complex biochemical processes that occur within the human body. Many of these proteins share similar structures and domains allowing them to use the same cofactor while having a different function. This idea is true for the proteins that require 2-ketoglutarate as a cofactor. Families of proteins often contain the same domains, which allow the human body to be more efficient, but make studying the function, localization and cellular pathway roles of a specific protein increasingly difficult. Various techniques and approaches have been applied to understand the role of specific proteins such as introducing single nucleotide mutations (for gain- or loss-of function) and knockout/knockdown and observing the phenotypic effects. It is known that many protein-protein complexes form to carry out many processes, there are compensatory mechanisms in place for adaptability and biological processes often occur on the millisecond to minute time scale (e.g. PTMs). With these characteristics and biological mechanisms in place, it is challenging to determine if the

phenotypic effects are due to the under- or overexpression of the targeted protein or if there is involvement from other proteins not being controlled for. Therefore, it is necessary to use approaches that can increase specificity and have the ability of temporal control.

Classical approaches in pharmacology involve optimization of small molecule derivatives to explore the functions of specific protein targets. Recently, an allele-specific chemical genetics (ASCG) approach has been developed and tested by several groups. This approach involves engineering both the protein of interest and the small molecule to be structurally complementary using recombinant DNA technology and organic synthesis. Site-directed mutagenesis is performed on the protein rendering it activated or inhibited by the synthetically modified ligand. The affinity for each other would essentially make the optimized system orthogonal and completely independent of the native pathway (156-161).

The druggability of proteins, increase in enzyme structural knowledge and accessibility to cofactor activators and inhibitors make the bump-hole approach increasingly more applicable and feasible for a diverse array of protein families. A bulky group is introduced onto the ligand (Bump) and is accommodated by the space (Hole) created through site directed mutagenesis on the enzyme. The native enzyme due to steric hindrance will not accept the bulky ligand and the engineered enzyme will not accept the native ligand because of lack of shape/ size complementarity. This creates an allele specific activator or inhibitor for the modified enzyme that can be used to monitor isoforms within the same family in a completely independent fashion.

This approach has been demonstrated on several PTM related protein families such as kinases, acetyltransferases, and methyltransferases. There are approximately 520 human protein kinases and ~20,000 phosphorylation sites resulting in a complex interacting network. De-

convoluting the biological functions of these proteins involves identifying substrates for specific kinases. Analog-sensitive alleles have been developed for >80 kinases. Shokat and colleagues developed [γ -32 P]-labeled orthogonal ATP analogues containing sterically bulky groups ATP (162). These ATP analogues were utilized by mutant kinase alleles as alternative cofactors to label their substrates with [32 P]-phosphate. Only the mutant allele was active on these bulky ATP analogues and not the wild type, allowing orthogonal identification of substrates for only the engineered kinase. This strategy resulted in novel candidate substrates of v-Src such as coflin, Dok-1, and calumenin to be identified. Additionally, this approach led to greater than 200 substrates identified for the cell-cycle regulator CDK1 in budding yeast and with further engineering, the Shokat group developed another set of ATP analogues resulting in the identification of >70 substrates for CDK1 (163-164).

Mutant alleles carrying unnatural amino acids (UAAs) with biologically orthogonal chemical properties are powerful systems for function elucidation. Complementary chemical or physical stimuli can be applied to reveal the roles of proteins bearing UAAs in complex in vivo environments. Using nonsense suppressor mutagenesis, the Schultz group successfully incorporated UAAs embellished with distinct reactivity into mutant alleles in a site-specific manner. A highly evolved orthogonal pair of tRNA and its cognate synthetase accomplished such allele-specific incorporation of UAAs into proteins. Remarkably, steric complementarity was employed to expand the hydrophobic pocket of the synthetase to accommodate the “ bump” in UAAs. Introduction of space in the active site also made the engineered synthetase orthogonal to the canonical amino acids. This powerful allele-specific chemical genetic approach has thus far been employed to incorporate more than 70 UAAs in a wide range of proteins having regulatory, structural, and/or enzymatic roles. Incorporation of photoactivable

UAAs into proteins, for example, allowed rapid activation of an allele in high spatiotemporal resolution using light to study its role in fast dynamic biological events or to identify and characterize transient protein− protein interactions (165-166).

1.7 PHOTO-CROSSLINKING APPROACH

There are 20 natural amino acid building blocks that the genetic code utilizes for the generation of proteins in all organisms. The functional groups of the natural amino acids include carboxylic acids and amides, a thiol and thiol ether, alcohols, basic amines, and alkyl and aryl groups. Post-translational modifications occur on these building blocks to define structure and function of proteins allowing for an enormous amount of functional outcomes to occur. Because of the chemical nature of functional manipulation of proteins, synthetic chemists and biologists have been able to expand the genetic code of several organisms to allow for the incorporation of unnatural amino acids into proteins.

Protein expression requires two major steps: transcription of DNA to mRNA and translation of mRNA to protein. The mature mRNA harbors triplet base pairs, codons, that correspond to one of the 20 amino acid building blocks or acts as a stop codon where the protein translation is complete. Translation occurs in the cytoplasm with ribosomes, which are composed of two subunits: large 50S and the small 30S each containing translation related RNA molecules. Transfer RNA (tRNA) is an adaptor molecule that recognizes both the codon via complementary base-pairing using an anti-codon and the corresponding amino acid. The association of specific tRNAs and amino acids occurs through aminoacyl tRNA synthetases, which places the correct amino acid on tRNAs by an esterification process. Translation occurs

via three main steps including initiation, elongation and termination. Initiation occurs when the mRNA is bound by a complex of initiation factors and then by the ribosome and necessary RNAs. Elongation refers to the amino acid chain becoming longer through the ribosome moving along the mRNA in the 5'-3' direction using elongation factors and corresponding tRNAs associating with the complex. Termination occurs when one of three codons, UAA, UAG, or UGA, is reached by the ribosome. There are not tRNAs that recognize the termination codons. Instead, release factors bind facilitating the release of the mRNA and dissociation of the ribosome. The amino acid chain generated undergoes folding by chaperone proteins and may get post-translationally modified for proper protein functioning.

There are three key features for successful unnatural amino acid incorporation including:

1. UAA incorporation should be site specific
2. tRNA-UAA association should be orthogonal
3. tRNA-UAA binding should out-compete release factors

To incorporate unnatural amino acids, the genetic code must be expanded to utilize codons that are not naturally recognized by tRNAs. To generate this system, several groups have evolved tRNAs and tRNA synthetases from various organisms to orthogonally generate tRNA-UAA complexes with anti-codons to infrequent termination codons. A plethora of unnatural amino acids have been introduced into proteins through the engineering of tRNA-tRNA synthetase pairs and the utilization of underutilized codons. One of the most infrequently used codons is the termination codon UAG, also known as the amber suppressor codon. Unnatural amino acids can be selectively incorporated at UAG sites using tRNAs that have the correct anti-codon and tRNA synthetases that can bind the unnatural amino acid and catalyze the attachment to the proper tRNA. For instance, para-azidophenylalanine (pAzF) can be recognized by archaea

evolved tRNA-tRNA synthetase (pEVOL-AzF) and site specifically incorporated at amber codons. UAA incorporation has been utilized for higher resolution interactome mapping, localization patterns and optical control for pathway activation (165-170).

Current methods for determining localization patterns and potential substrates for KDMs involve nonspecific crosslinking. For instance ChIP uses formaldehyde to crosslink amide bonds, which is strictly proximity based. The crosslinking can occur within any region of the protein, is not temporally controllable and is nonspecific. The incorporation of a photo-induced crosslinking unnatural amino acid allows for crosslinking to occur within specific protein domains (i.e. catalytic domain vs. reader domain) using site directed mutagenesis, temporal control through dictating when the protein is exposed to the activating light, and is highly specific through the orthogonal engineering.

1.8 OBJECTIVES

The main objective of this work is to develop three biochemical systems for KDM4 proteins that are orthogonal, can be applied in mammalian cells and provide a template to develop similar systems for other KDMs. The development of these systems will allow for the elucidation of genes and proteins regulated by individual KDMs and is necessary due to the limitations of current technologies used to study this family of enzymes. The combination of small molecules and genetic techniques results in a chemical biology tool that is both highly specific and rapidly employable, which addresses the issues of compensatory mechanisms, off target effects and low resolution of current methods such as knock-out studies, non-specific small molecules and chromatin immunoprecipitation. The goal is to develop three tools that can

eventually be employed in live models (i.e. mammalian cells, zebrafish, mice) to elucidate genes and proteins that are regulated by the catalytic activity of individual KDMs to better understand their roles in different contexts (healthy vs. diseased). This will expand the knowledge of KDMs functions in various cellular pathways, which will contribute to the development of improved therapeutic strategies and interventions and potentially unmask diagnostic and prognostic biomarkers. This work is being performed in the hopes that the tools developed can serve as a basic framework to generate similar systems for other 2-KG dependent enzymes.

1.8.1 Aim 1 – Develop an Orthogonal KDM4 Bump-Hole Activation System

The system will involve engineering the catalytic domain of KDM4A and the fourth carbon of cofactor, 2-KG, which will act independently of and not influence the native KDM4A-2-KG pair. Additionally, the engineered pair should be chromatin site and state specific, similar to the native pair. This engineered system should be able to mimic the native KDM4A system and will ideally be applicable to other KDM4 members and potentially other JmjC domain containing KDMs (Figure 6).

The KDM4A JmjC domain has been mutated to expand the 2-KG binding site for the accommodation of bulky 2-KG analogues. The mutant enzyme and a select few 2-KG analogues have been screened for enzymatic activity in the presence of trimethylated histone H3K9 peptides. An *in vitro* engineered KDM4A-2-KG has been identified, but should be improved for higher orthogonality, proof of concept in live mammalian cells and generality to other JmjC domain containing enzymes. The objectives of this aim are as follows:

1. Screen and biochemically characterize additional 2-KG analogues for H3K9 demethylase activity by KDM4 mutants.
2. Test the engineered KDM4-2-KG pairs in physiologically relevant environments including mammalian cell lysates and live cells.
3. Apply this bump-hole activation approach to other JmjC domain containing KDMs.

This will allow for the engineered KDM4-2-KG pair to be employed in live cells as a gain of function approach to elucidate genes that are regulated by the catalytic activity of individual KDM4s in various contexts.

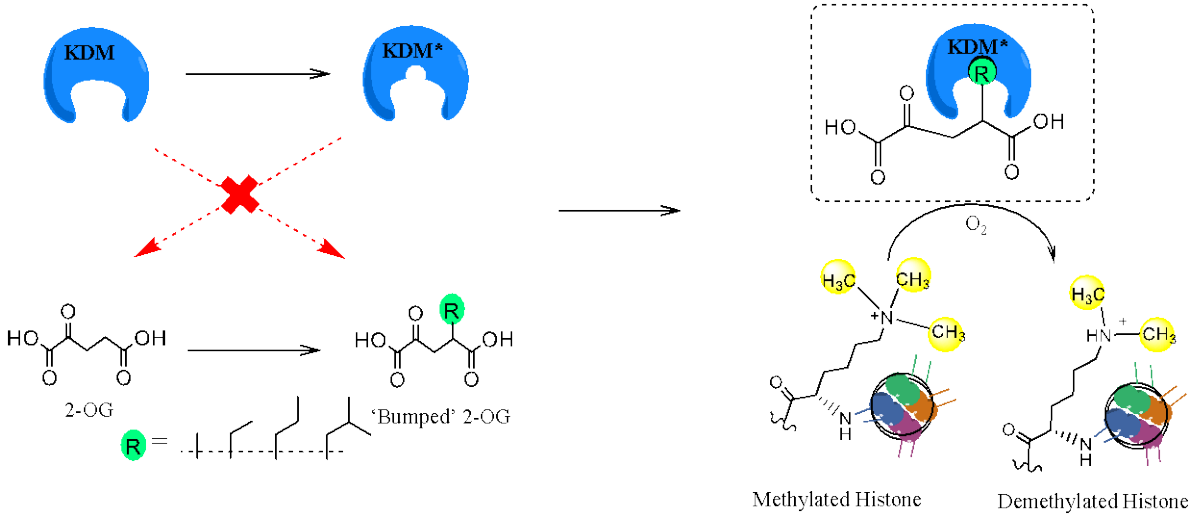


Figure 6: Bump Hole activation concept

The KDM of interest is site specifically mutated to generate a 'hole' which will 1) generate low affinity binding to the native cofactor (2-OG/2-KG) and 2) accommodate a bulky (bump) cofactor and bind at a higher affinity. The native KDM, as well as other 2-OG dependent enzymes, will not have the structural capacity to bind the 'bumped' cofactor making the system orthogonal. Applying this system to KDMs will result in orthogonal and controllable histone demethylation.

1.8.2 Aim 2 – Develop an Orthogonal KDM4 Bump-Hole Inhibition System

This system will involve screening active KDM4 mutants for inhibition by 2-KG competitive inhibitor analogues such as 2-HG and NOG. Similar to the activation system, the inhibitors will be modified into bulky derivatives that will orthogonally inhibit KDM4 demethylation (Figure 7). Ideally, the system generated for KDM4 will be generalizable to additional KDMs. To accomplish this aim, the following objectives include:

1. Screen and biochemically characterize 2-HG analogues for KDM demethylase inhibition.
2. Test the engineered KDM4-2-HG pairs in physiologically relevant environments including mammalian cell lysates.
3. Screen and biochemically characterize potent 2-KG competitive inhibitor, NOG, analogues for KDM4A demethylase inhibition.

Following the completion of the aims listed above, the orthogonal pairs identified need to be functionally validated in live cells. This will allow for the engineered KDM4-inhibitor pair to be employed in live cells as a loss of function approach to elucidate genes that are regulated by the catalytic activity of individual KDM4s in various contexts.

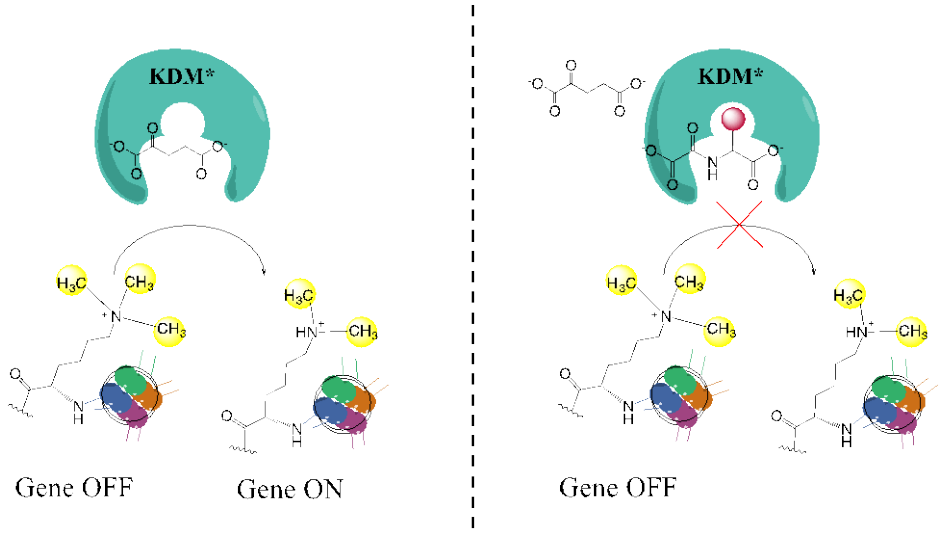


Figure 7: Bump Hole inhibition concept.

The KDM of interest is site specifically mutated to generate a 'hole' which will 1) generate low affinity binding to the native inhibitor (2-HG or NOG) and 2) accommodate a bulky (bump) inhibitor and bind at a higher affinity. The native KDM, as well as other 2-OG dependent enzymes, will not have the structural capacity to bind the 'bumped' inhibitor making the system orthogonal. Applying this system to KDMs will result in orthogonal and controllable inhibition of histone demethylation.

1.8.3 Aim 3 – Develop Photo-induced Crosslinkable KDM4

The catalytic domain of KDM4s were analyzed to target residues for the substitution with a photo-crosslinkable UAA. Successful site-specific incorporation and photo-induction of the UAA will allow for the capture of catalytic substrates of KDM4s. This system allows for high specificity and temporal control for the identification of member specific substrates. The development of this system for each KDM4 member will allow for the identification of non-histone protein substrates as well as the identification of genomic localization with the

resolution of where each KDM4 is catalytically active (Figure 8). The following objectives are necessary to be accomplished for the development of this system:

1. Generate and biochemically characterize photo-crosslinkable KDM4A-UAA mutants.
2. Generalize the approach in other KDM4 members.
3. Test the KDM4-UAA mutants in mammalian cell environments.

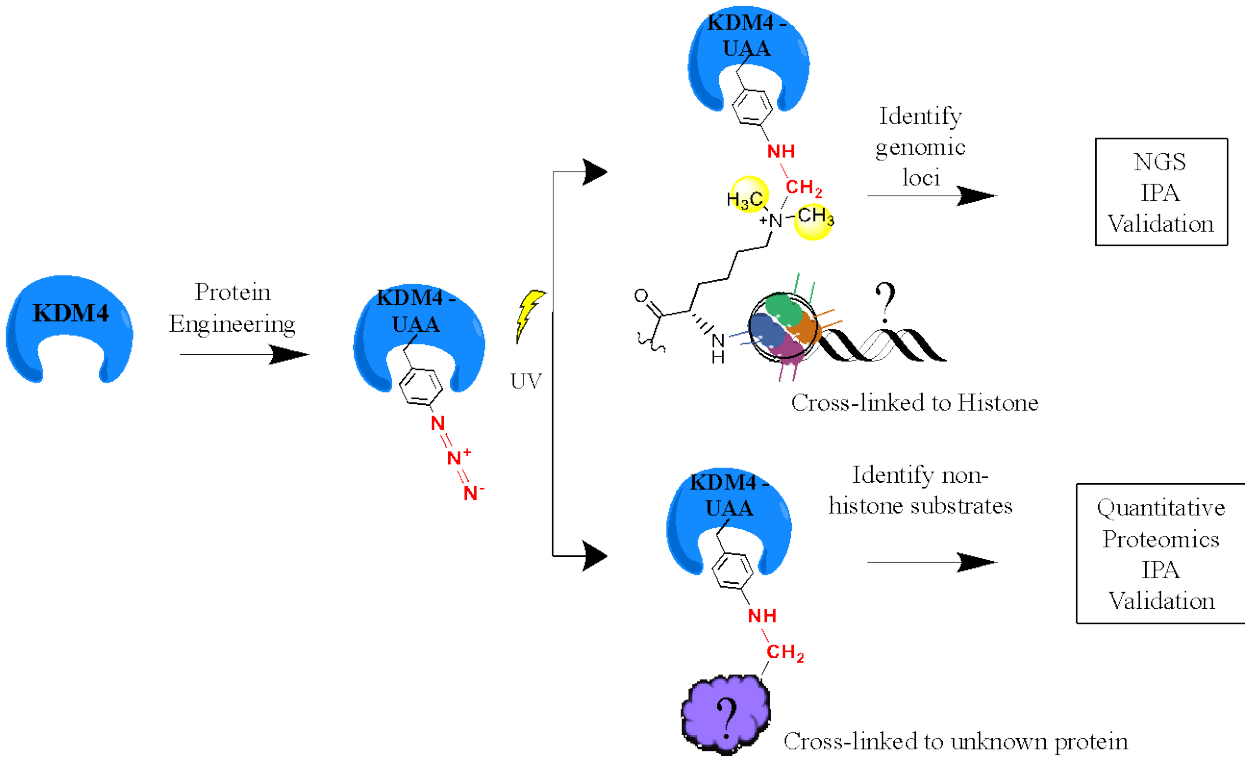


Figure 8: Photo-induced Crosslinking Concept

The KDM of interest is site specifically mutated to incorporate a photo-inducible UAA. Upon exposure to UV light, the UAA will be activated and crosslinking will be induced. Crosslinking will take place between the KDM4 and the substate bound within close proximity to the UAA. KDM4 crosslinked to nucleosomal histones can be captured to isolate the surrounding nucleosomal DNA to determine catalytic genomic localization. Additionally, non-histone proteins that crosslink to the KDM4-UAA can be determined via proteomics analysis. Together these data will unmask the catalytic role of KDM4s in various cellular pathways.

The development of a photo-crosslinkable KDM4 will allow for KDM4 non-histone protein substrates to be identified. Following the identification of these substrates, pathway analysis should be conducted to determine how the demethylation of specific proteins effects different cellular mechanisms. Additionally, through next generation sequencing (NGS) technologies, genomic loci where histones are directly demethylated by KDM4s will be unmasked. The demethylation at these sites may result in transcriptional changes that can be monitored by (quantitative polymerase chain reaction) qPCR. This work will result in genes and proteins that can be monitored using the systems developed in aim 1 and aim 2 of this document. For instance, if the quantitative proteomics data shows that the engineered KDM4A crosslinks to a transcription factor, pathway analysis can be performed to determine which genes are regulated by the transcription factor identified. These genes can then be monitored by qPCR following the employment of the engineered KDM4A-2-KG or KDM4A-inhibitor. All of the tools developed through this work will be complementary to each other and will aid in the improvement, validation and progression of these systems.

Future investigations using these tools will allow for the catalytic functions of individual KDM4s to be elucidated in different contexts. The functions of each KDM can then be compared to further understand functional redundancies and uniqueness. Additionally, the behavior and roles of an individual KDM can be compared in different contexts. For instance, the systems developed for KDM4A can be applied to healthy cells, cancer cells, and neuropathic cells. The results for each model can then be compared to identify how the role of KDM4 changes in each phenotype. Theoretically, this can be performed for each KDM to allow for the comparison of functions across enzyme and phenotype to better understand the complex

network of this epigenetic family for the development of therapeutic strategies and identification of biomarkers.

2.0 MATERIALS AND METHODS

2.1 PLASMIDS, MUTAGENIC PRIMERS, CELL LINES AND ANTIBODIES

All the plasmids are for bacterial expression and obtained as gifts from individual laboratories or purchased from Addgene. Mutagenic primers were obtained from Integrated DNA Technologies (Table 7). Human embryonic kidney 293T (HEK293T) cells were obtained from the American Type Culture Collection (ATCC) and used in the current study following manufacturer's protocol. All the antibodies (Table 2) used in the current study are purchased from established vendors and used following manufacturer's protocol.

Table 2: Antibody Information

Antibody	Catalog Number	Vendor
Histidine Tag	sc-8036	Santa Cruz Biotech
Histone H3	701517	Thermo Scientific
Histone H3K27me3	GTX50901	GeneTex
Histone H3K27me3	61017	Active Motif
Histone H3K36me2	61019	Active Motif
Histone H3K36me3	61021	Active Motif
HIstone H3K4me3	61379	Active Motif
Histone H3K9me1	39681	Active Motif
Histone H3K9me2	39683	Active Motif
Histone H3K9me3	61013	Active Motif
Histone H4	61521	Active Motif
JMJD2A/KDM4A	39815	Active Motif
JMJD2B/KDM4B	PA5-24611	Thermo Scientific
JMJD2C/KDM4C	NB110-38884	Novus bio
JMJD2D/KDM4D	PA5-41590	Thermo Scientific
Strep-tag	2-1509-001	IBA
Goat Anti-Rabbit IgG-HRP	5/1/4030	Southern Biotech
HRP Goat anti-Mouse IgG	15014	Active Motif
Rabbit anti-sheep IgG-HRP	5/1/6150	Southern Biotech
Donkey anti-goat IgG-HRP	NBP1-74815	Novus bio

2.2 SYNTHESIS AND PURIFICATION OF PEPTIDES

All the peptides were synthesized by the University of Pittsburgh Peptide Synthesis Facility (Table 3). Analytical-scale separation was performed using ZORBAX reversed-phase C18 (5μm, 4.6 × 250 mm) column with UV detection at 280 nm. The column was equilibrated with 0.1% aqueous trifluoroacetic acid solution prior to each injection. Analytical separation was performed with a linear gradient of acetonitrile to 10% in 15 min and then to 70% in 5 min in 0.1% aqueous trifluoroacetic acid with a flow rate of 1 mL/min. The crude peptides were purified using preparative reversed-phase HPLC (XBridge C18, 5 μm, 10 x 250 mm column) eluting with a flow rate of 5 mL/min and a gradient of acetonitrile starting from 0% to 90% in 15 min and then to 100% in 18 min in aqueous trifluoroacetic acid (0.01%). The purified peptides were first concentrated by SpeedVac concentrator followed by lyophilization. The dried peptides were resuspended in water containing 0.01% TFA and stored at -80 before use. Concentrations of the peptides were determined based on the observation that 1 mg/ml peptide generates an absorbance value (A205) of 30 at 205 nm. The integrity of the purified peptide was confirmed by MALDI mass spectrometry.

Table 3: Histone H3 Peptides.

Histone H3 Peptides		
Peptide	**Sequence**	**Molecular Weight (Da)**
H3K4me3	H2N-ARTK(Me3)QTARKSTGGKA-CONH2	1601
H3K9me3	H2N-ARTKQTARK(Me3)STGGKAPRKQLK(Biotin)-CONH2	2692
6CF-H3K9me3	6CF-HN-ARTKQTARK(Me3)STGGKA-CONH2	1959
TAMRA-H3K9me3	TAMRA-HN-ARTKQTARK(Me3)STGGKA-CONH2	2382
H3K27me3	H2N-APRKQLATKAARK(Me3)SAPATGGVK-CONH2	2474.61
H3K36me3	H2N-KSAPSTGGVK(Me3)KPHRYKPGTGK(Biotin)-CONH2	2562
H3R2me2	H2N-AR(Me2)TKQTARKSTGGKAPRK(Biotin)-CONH2	2195.26

2.3 BACTERIAL EXPRESSION AND PURIFICATION OF PROTEINS

2.3.1 Lysine Demethylase Family 4 (KDM4)

The N-terminal 6xHis-tagged human KDM4A-jmjC domain (catalytic domain of KDM4A) bacterial expression construct pNIC28-Bsa4 (Addgene ID: 38846) was obtained from Addgene. The wild type KDM4A plasmid was transformed into E. Coli BL21 (DE3) competent cells (Invitrogen) using pNIC28-Bsa4 kanamycin-resistant vector. A single colony was picked up and grown overnight at 37 in 10 mL of Luria-Bertani (LB) broth in the presence of 50 μg/mL kanamycin and 35 μg/mL chloramphenicol. The culture was diluted 100-fold and allowed to grow at 37 ℃ to an optical density (OD600) of 0.8, and protein expression was induced overnight at 17 °C with 0.6 mM IPTG in an Innova 44® Incubator shaker (New Brunswick Scientific). Proteins were purified as follows: harvested cells were resuspended in 15 mL lysis buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM β-mercaptoethanol, 10% glycerol, 25 mM imidazole, Lysozyme, DNase, and Roche protease inhibitor cocktail). The cells were lysed by pulsed sonication (Qsonica-Q700), and centrifuged at 13000 rpm for 40 min at 4 °C. The soluble extracts were subject to Ni-NTA agarose resin (Thermo) according to manufacturer's instructions. After passing 20 volumes of washing buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM β-mercaptoethanol, 10% glycerol, and 25 mM imidazole), proteins were eluted with a buffer containing 50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM β-mercaptoethanol, 10% glycerol, and 400 mM imidazole. Proteins were further purified by size exclusion chromatography (Superdex-200) using AKTA pure FPLC system (GE healthcare) with buffer containing 50 mM Tris-HCl pH 8.0, 200 mM NaCl, and 10% glycerol. Purified proteins were concentrated using Amicon Ultra-10k centrifugal filter device (Merck Millipore Ltd.). The

protein concentration was determined using Bradford assay kit (BioRad Laboratories) with BSA as a standard. The concentrated proteins were stored at -80°C before use. KDM4A variants were generated using the QuikChange Lightning site-directed mutagenesis kit (Agilent Technologies). The resulting mutant plasmids were confirmed by DNA sequencing. KDM4A variants were expressed and purified as stated for KDM4A-WT.

The N-terminal strep-tagged human KDM4A, KDM4B and KDM4D JmjC catalytic domain for bacterial expression were kindly provided by Prof. Raymond Trievel, University of Michigan.The native constructs were transformed into E. coli Rosetta [DE3] competent cells. A single colony was picked up and grown overnight at 37 °C in 10 mL of Luria-Bertani (LB) broth in the presence of 100 μg/mL ampicillin and 35 μg/mL chloramphenicol. The culture was diluted 100- fold and allowed to grow at 37 °C to an optical density (OD600) of 0.8, and protein expression was induced overnight at 18 °C with 1 mM IPTG in an Innova 44® Incubator shaker (New Brunswick Scientific). Proteins were purified as follows: harvested cells were resuspended in 15 mL lysis buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM β-mercaptoethanol, 10% glycerol, Lysozyme, DNase, and Roche protease inhibitor cocktail). The cells were lysed by pulsed sonication (Qsonica-Q700), and centrifuged at 13000 rpm for 40 min at 4 °C. The soluble extracts were subject to Strep-tactin resin (Qiagen, cat# 30004) according to manufacturer's instructions. After passing 20 volumes of washing buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM β-mercaptoethanol, 10% glycerol), proteins were eluted with a buffer containing 50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM β-mercaptoethanol, 10% glycerol, and 5 mM D-desthiobiotin. Proteins were further purified by size exclusion chromatography (Superdex-200) using AKTA pure FPLC system (GE healthcare) with buffer containing 50 mM Tris-HCl pH 8.0, 200 mM NaCl, and 10% glycerol. Purified proteins were

concentrated using Amicon Ultra-10k centrifugal filter device (Merck Millipore Ltd.). The protein concentration was determined using Bradford assay kit (BioRad Laboratories) with BSA as a standard. The concentrated proteins were stored at -80°C before use.

The N-terminal 6xHis-tagged human KDM4C catalytic domain construct for bacterial expression was a kind gift from Prof. Danica Fujimori, UCSF. Plasmids were transformed into BL21 [DE3] star cells and grown at 37 °C to an optical density of 0.6. Protein expression was induced with 0.3 mM IPTG overnight at 17 °C. Following bacterial cell lysis, expressed protein was purified first by nickel affinity chromatography and secondly by size exclusion chromatography using a superdex 200 FPLC column. The proteins were concentrated via centrifugation. Protein concentration was determined using Bradford assay.

2.3.2 Lysine Demethylase Family 6 (KDM6)

The plasmid constructs for the catalytic domain of KDM6B was kindly provided by Prof. Christopher Schofield of the University of Oxford. Purified plasmid was transformed into Rosetta [DE3] competent cells. Propagation of plasmid-containing cells was performed in Luria Bertani (LB) broth at 37 °C. Once the optical density reached ~0.8, protein expression was induced with 0.3 mM IPTG shaking at 17 °C overnight. The cells were collected by centrifugation at 4000 rpm for 30 minutes. Following bacterial cell lysis, expressed protein was purified first by nickel affinity chromatography and secondly by size exclusion chromatography using a superdex 200 FPLC column.

2.3.3 Formaldehyde Dehydrogenase (FDH)

The N-terminal 6xHis-tagged *P. putida* FDH bacterial expression construct in pET28 vector was obtained from the Bhagwat laboratory at the Wayne State University The plasmid was transformed into E. coli Rosetta [DE3] competent cells. A single colony was picked up and grown overnight at 37 °C in 10 mL of Luria-Bertani (LB) broth in the presence of 50 μg/mL kanamycin and 35 μg/mL chloramphenicol. The culture was diluted 100-fold and allowed to grow at 37 °C to an optical density (OD600) of 0.8, and protein expression was induced overnight at 18 °C with 1 mM IPTG in an Innova 44® Incubator shaker (New Brunswick Scientific). Proteins were purified as follows: harvested cells were resuspended in 15 mL lysis buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM β-mercaptoethanol, 10% glycerol, 25 mM imidazole, Lysozyme, DNase, and Roche protease inhibitor cocktail). The cells were lysed by pulsed sonication (Qsonica-Q700), and centrifuged at 13000 rpm for 40 min at 4 °C. The soluble extracts were subject to Ni-NTA agarose resin (Thermo) according to manufacturer's instructions. After passing 20 volumes of washing buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM β-mercaptoethanol, 10% glycerol, and 25 mM imidazole), proteins were eluted with a buffer containing 50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM β-mercaptoethanol, 10% glycerol, and 400 mM imidazole. Proteins were further purified by size exclusion chromatography (Superdex-200) using AKTA pure FPLC system (GE healthcare) with buffer containing 50 mM Tris-HCl pH 8.0, 200 mM NaCl, and 10% glycerol. Purified proteins were concentrated using Amicon Ultra-10k centrifugal filter device (Merck Millipore Ltd.). The protein concentration was determined using Bradford assay kit (BioRad Laboratories) with BSA as a standard. The concentrated proteins were stored at -80 °C before use.

2.3.4 Histone H3

Gene sequence encoding wild type *Xenopus laevis* histone H3 was a kind gift from Prof. Minkui Luo at the Memorial Sloan-Kettering Cancer Center. The plasmid containing histone H3-C110A was transformed into BL21 codon plus (DE3) RIPL competent cells. A single colony was picked up and grown overnight at 37 °C in 10 mL of LB broth with 100 μg/mL ampicillin and 35 μg/mL chloramphenicol. The inoculation culture was diluted 1:100 fold in fresh LB medium and cells were grown at 37 °C until OD600 reached to ~0.7. Protein expression was induced by the addition of 0.3 mM IPTG followed by growing for an additional 3 h at 37 °C. Cells were harvested by centrifugation at 5000 rpm for 30 min, and then resuspension of the pellet in 5 mL of lysis buffer (10 mM Tris-HCl pH 7.5, 2 M guanidinium hydrochloride (GdnHCl), 5 mM β-mercaptoethanol, 10% glycerol, DNase, Lysozyme and Roche protease inhibitor cocktail). The cells were lysed by pulsed sonication and centrifuged at 20,000g for 40 min at 4 °C. Insoluble histone was recovered from inclusion bodies by dissolving in 6 M GdnHCl and 10 mM Tris-HCl pH 7.5, and incubated for 10 min at room temperature followed by centrifugation at 20,000g for 40 min at 4 °C. The soluble histone supernatant was purified by size exclusion chromatography on a Superdex-200 using AKTA pure FPLC system. Fractions were concentrated using Amicon Ultra-4 centrifugal 3K filter and further purified with preparative reverse-phase HPLC (XBridge C18, 5 μm, 10 x 250 mm column) eluting with a flow rate of 4 mL/min starting from 10% acetonitrile to 70 % in 15 min and then to 100 % over 5 min in aqueous trifluoroacetic acid (0.01%). The purified protein was concentrated by SpeedVac followed by lyophilization. The protein was stored at -80°C before use.

2.3.5 G9a

The N-terminal 6xhis tagged human G9a construct for bacterial expression was sequenced and purified. Purified plasmid was transformed into BL21 [DE3] competent cells. Propagation of plasmid-containing cells was performed in Luria Bertani (LB) broth at 37 °C with associated antibiotics. Once the optical density reached ~0.8, protein expression was induced with 0.5 mM IPTG shaking and 200 mL of 100 mM zinc chloride per 1 L flask was added to aid protein folding. Cultures incubated at 17 °C overnight with agitation. The cells were collected by centrifugation at 4000 rpm for 30 minutes. Following bacterial cell lysis, expressed protein was purified first by nickel affinity chromatography and secondly by size exclusion chromatography using a superdex 200 FPLC column. SDS-PAGE and coomassie staining verified isolated protein. Proteins were concentrated using Amicon Ultra-10k centrifugal filter device (Merck Millipore Ltd.) Protein concentration was determined using Bradford solution and BSA standard curve. The protein was stored at -80°C until use in biochemical assays.

2.3.6 Unnatural Amino Acid (UAA) Incorporation

KDM4 catalytic domains were mutated with an amber suppressor codon (TAG) substitution at a designated site using the mutagenesis protocol. The confirmed mutant KDM4 and tRNA-tRNA synthetase plasmids (pEVOL-AzF or pEVOL-BpF) were co-transformed into competent BL21 [DE3] star cells. Following transformation, single colonies were propagated in 10 mL LB broth with associated antibiotics. The cultures were then centrifuged and the pelleted cells were resuspended in 3 mL LB broth and transferred to 1 liter of GMML media (Table 4-5). The cells were grown to an OD_{600} of 0.8 by shaking at 37 °C. Once the desired optical density was

reached, the unnatural amino acid (Azidophenylalanine, AzF, or Benzophenone, BpF) was added to the flasks and the cultures incubated at 17 °C for 30 minutes. L-Arabinose was then and the cells incubated for an additional 30 minutes at 17 °C. KDM4 expression was induced at 17 °C with 0.25 mM IPTG. Proteins were purified as follows: harvested cells were resuspended in 15 mL lysis buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM β-mercaptoethanol, 10% glycerol, Lysozyme, DNase, and Roche protease inhibitor cocktail). The cells were lysed by pulsed sonication (Qsonica-Q700), and centrifuged at 13000 rpm for 40 min at 4 °C. The soluble extracts were subject to Strep-tactin resin (Qiagen, cat# 30004) according to manufacturer's instructions. After passing 20 volumes of washing buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM β-mercaptoethanol, 10% glycerol), proteins were eluted with a buffer containing 50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM β-mercaptoethanol, 10% glycerol, and 5 mM D-desthiobiotin. Proteins were further purified by size exclusion chromatography (Superdex-200) using AKTA pure FPLC system (GE healthcare) with buffer containing 50 mM Tris-HCl pH 8.0, 200 mM NaCl, and 10% glycerol. Purified proteins were concentrated using Amicon Ultra-10k centrifugal filter device (Merck Millipore Ltd.). The protein concentration was determined using Bradford assay kit (BioRad Laboratories) with BSA as a standard. The concentrated proteins were stored at -80°C before use. Protein concentrations were validated by SDS-PAGE and LC-MS. The AzF preservation was determined by a copper catalyzed click reaction followed by SDS-PAGE and in-gel fluorescence.

Table 4: GMML Media for KDM4-UAA Expression.

GMML Media Composition (for 1 Liter)		
Reagent	**Concentration**	**Volume**
M9 media powder	10.5 g/L	950 mL
Glycerol	20%	50 mL
Leucine	4 mg/mL	10 mL
Calcium Chloride	20 mg/mL	1 mL
Magnesium Sulfate	120 mg/mL	1 mL
D-biotin	5 mg/mL	0.2 mL
Thiamine-HCl	5 mg/mL	0.2 mL
Glucose	0.2 mg/mL	10 mL
Heavy Metal Stock Solution		1 mL

Table 5: Heavy Metal Stock Solution for GMML media.

Heavy Metal Stock Solution (1 L of 1000X)	
Reagent	**Quantity**
Na_2 $MoSO_4$	500 mg
$CoCl_2$	250 mg
$CuSO_4$	175 mg
$MnSO_4$	1 g
$MgSO_4$	8.75 g
$ZnSO_4$	1.25 g
$FeCl_2$	1.25 g
$CaCl_2$	2.5 g
H_3BO_3	1 g
1 M HCl	to 1 Liter

2.4 MUTAGENESIS

Mutant plasmids were generated using the QuikChange Lightning site-directed mutagenesis kit (Agilent Technologies) and confirmed via Sanger sequencing by GeneWiz and primers from IDT (Table 6). The purified plasmids were transformed into BL21 [DE3] or Rosetta [De3] cells and expressed and purified similar to their native counterparts.

Table 6: Mutagenesis Primers.

Bump-Hole Mutant Primers	
Mutant	**Fwd. Primer**
KDM4A-Y132A	5'-CATTCAATCCTCCAATC**GCC**GGTGCAGATGTGAATGG-3'
KDM4A-Y132G	5'-CATTCAATCCTCCAATC**GGC**GGTGCAGATGTGAATGG-3'
KDM4A-Y175A	5'-GAGGGTGTGAACACCCCA**GCC**CTGTACTTTGGCATG-3'
KDM4A-Y175F	5'-GGTGTGAACACCCCA**TTC**CTGTACTTTGGCATG-3'
KDM4A-Y175G	5'-GGTGTGAACACCCCA**GGC**CTGTACTTTGGCATG-3'
KDM4A-Y175V	5'-GGTGTGAACACCCCA**GTG**CTGTACTTTGGCATG-3'
KDM4A-Y177A	5'-GTGAACACCCATACCTG**GCC**TTTGGCATGTGGAAGAC-3'
KDM4A-Y177F	5'-GTGAACACCCATACCTG**TTC**TTTGGCATGTGGAAGAC-3'
KDM4A-Y177G	5'-GTGAACACCCCATACCTG**GGC**TTTGGCATGTGGAAGAC-3'
KDM4A-Y177V	5'-GTGAACACCCCATACCTG**GTG**TTTGGCATGTGGAAGAC-3'
KDM4A-F185A	5'-GGCATGTGGAAGACATCC**GCC**GCTTGGCACACTGAAGA-3'
KDM4A-F185G	5'-GGCATGTGGAAGACATCC**GGC**GCTTGGCACACTGAAG-3'
KDM4A-F185I	5'-GGCATGTGGAAGACATCC**ATC**GCTTGGCACACTGAAG-3'
KDM4A-F185T	5'-GGCATGTGGAAGACATCC**ACC**GCTTGGCACACTGAAG-3'
KDM4A F185V	5'-GGCATGTGGAAGACATCC**GTG**GCTTGGCACACTGAAG-3'
KDM4A-N198A	5'-GACCTCTACAGCATC**GCC**TACCTGCACTTTGG-3'
KDM4A-N198G	5'-GGACCTCTACAGCATC**GGC**TACCTGCACTTTGGAG-3'
KDM4A-K206A	5'-CTGCACTTTGGAGAACCA**GCC**TCCTGGTACTCTGTTCC-3'
KDM4A-K206G	5'-CTGCACTTTGGAGAACCA**GGC**TCCTGGTACTCTGTTCC-3'
KDM4A-W208A	5'-CTTTGGAGAACCAAAGTCC**GCC**TACTCTGTTCCACCTGAG-3'
KDM4A-W208F	5'-GGAGAACCAAAGTCC**TTC**TACTCTGTTCCACCTGAG-3'
KDM4A-W208G	5'TGGAGAACCAAAGTCC**GGC**TACTCTGTTCCACC-3'
KDM4A-W208V	5'GGAGAACCAAAGTCC**GTG**TACTCTGTTCCACCTG-3'
KDM4A-K241A	5'- GCA TTT CTC CGC CAC **GCC** ATG ACC CTG ATT TCC -3'
KDM4A-K241G	5'- GCA TTT CTC CGC CAC **GGC** ATG ACC CTG ATT TCC -3'
KDM4A-A286G	5'-CCATGGTTTTAACTGT**GGC**GAGTCTACCAATTTTGC-3'
KDM4A-S288A	5'-GGTTTTAACTGTGCGGAG**GCC**ACCAATTTTGCTACCC-3'
KDM4A-S288G	5-GGTTTTAACTGTGCGGAG**GGC**ACCAATTTTGCTACCCG-3'
KDM4A -N290A	5'- CTG TGC GGA GTC TAC C**GCC**TT TGC TAC CCG TCG G -3'
KDM4A -N290G	5'- CTG TGC GGA GTC TAC C**GGC**TT TGC TAC CCG TCG G -3'
KDM4B-F186G	5'- CATGTGGAAGACCACC**GGC**GCCTGGCACAC -3'
KDM4C-F187G	5'- GCATGTGGAAGACCACG**GGC**GCATGGCACATCGAAG -3'
KDM4D-F189G	5'- GCATGTGGAAAACCACG**GGC**GCTTGGCATACAGAG -3'
KDM6B-V1472A	5'- CTG TGC ACT GG**G CC**C AGG CCA CCG GC -3'
KDM6B-V1472G	5'- CTG TGC ACT GG**G GC**C AGG CCA CCG GC -3'
Histone H3 Mutagenesis Primers	
Mutant	**Fwd. Primer**
H3K4C	5'- GCG CGT ACT **TGC** CAG ACG GCT CG -3'
H3K9C	5'- GCA GAC GGC TCG G**TG C**TC CAC CGG CGG -3'
H3K27C	5'- CTA CCA AGG CTG CTC GC**T GC**A GCG CGC CGG CTA C -3'
H3K36C	5'- CTA CCG GCG GCG TG**T GC**A AGC CTC ACC GTT AC -3'
Amber Codon Mutants	
Mutant	**Fwd. Primer**
I71TAG	5'- GGT CAT TCC TGC CCC C**TA G**CA ACA GCT GGT GAC G -3'
V171TAG	5'- GGA TCA CCA TTG AGG GT**T AG**A ACA CCC CAT ACC TG -3'
Y175TAG	5'- GGT GTG AAC ACC CCA **TAG** CTG TAC TTT GGC ATG -3'
Y177TAG	5'- GAA CAC CCC ATA CCT G**TA G**TT TGG CAT GTG GAA G -3'
D191TAG	5'- GCT TGG CAC ACT GAA **TAG** ATG GAC CTC TAC AGC -3'
K241TAG	5'- GCA TTT CTC CGC CAC **TAG** ATG ACC CTG ATT TCC -3'
S288TAG	5'- GGT TTT AAC TGT GCG GAG **TAG** ACC AAT TTT GCT ACC CG -3'
N290TAG	5'- CTG TGC GGA GTC TAC C**TA G**TT TGC TAC CCG TCG G -3'
V313TAG	5'- GCT CCT GTA GAA AGG ACA TG**T AG**A AGA TCT CCA TGG -3'
Y175TAG + Y177G	5'- GTG TGA ACA CCC CA**T AG**C TGG GCT TTG GCA TGT G -3'
Y175G + Y177TAG	5'- GTG TGA ACA CCC CA**G GC**C TG**T AG**T TTG GCA TG -3'
KDM4B-I72TAG	5'- GAT CCC GGC GCC C**TA G**CA GCA GGT G -3'
KDM4B-Y178TAG	5'- CAC GCC CTA CCT G**TA G**TT CGG CAT GTG GAA GAC C -3'
KDM4B-K242TAG	5'- CTT CCT GCG GCA T**TA G**AT GAC CCT CAT CTC -3'
KDM4C-I73TAG	5'- GTC TAA AGG AGC CC**T AG**G TGC GGG TCT TGC AAA G -3'
KDM4D-L75TAG	5'- CTT AAT AGC CAC TCC C**TA G**CA GCA GGT GGC -3'
KDM4D-Y181TAG	5'- CAC CCT ACT TG**T AG**T TTG GCA TGT GGA AAA CC -3'
KDM4D-K245TAG	5'- CTT CCT GCG GCA C**TA G**GT GGC CCT CAT C -3'

2.5 GENERATION OF HISTONE METHYL LYSINE ANALOGUES (MLAS)

For the generation of trimethylated H3KC9Me3, 1 mg of lyophilized histone was dissolved in 98 µl of alkylation buffer (4 M Guanidine-Hydrochloride, 1 M HEPES pH 7.8, 10 mM D/Lmethionine). Once histone was fully dissolved, 2 µl of 1 M DTT (prepared fresh) was added to the histone solution followed by 1 h incubation at 37 °C. The solution was then added to 10 mg of (2-bromoethyl) trimethyl ammonium bromide (Sigma, cat# 117196) and protected from light. After incubating for 2.5 h at 50 °C (gently agitating every 30 min), 1 µl of 1 M DTT was added to quench the reaction. The solution incubated at 50 °C for an additional 2.5 h. The reaction was quenched by adding 5 µl β-mercaptoethanol (β-ME). A PD10 (GE, cat# 17-0851-01) column was used to desalt the solution and the histones were eluted with 2 mM β-ME. SDS-PAGE was performed to determine presence of the histone protein. The samples were lyophilized and then resuspended in 50 mM Tris pH 8.0. Protein concentration was determined by Bradford assay and the molecular weight was verified by LC-MS. Protein was stored in -80 °C until used in demethylase activity assay.

2.6 HEK293T CELL CULTURE AND CELL LYSIS

Human embryonic kidney (HEK) 293T cells were grown in Dulbecco modified Eagle medium (DMEM) (Gibco) supplemented with 10% fetal calf serum in a humidified atmosphere

containing 5% CO2 in a T75 flask. At ~80% confluence, cells were treated with 20 µM of KDM4 inhibitor n-octyl-IOX1 (EMD Millipore 5.30537.0001)13,14 dissolved in DMSO to generate hypermethylated histones. 24 hours post treatment, cells were harvested and lysed with 300 µL of cold RIPA buffer (Sigma) supplemented with 1X Roche protease inhibitor cocktail and 5 mM TCEP by sonicating for 15 min at amplitude of 100 with a repeating 1 min pulse cycle at 4 °C. Cell lysates were centrifuged at 12,000 rpm for 30 min at 4 °C to remove cell debris. The supernatant was then passed through the detergent removal spin column (Pierce # 87778) and eluted with Tris buffer (50 mM Tris-HCl pH 8.0, 10% glycerol, 2 mM TCEP, 1X Roche protease inhibitor cocktail) following manufacturer's protocol. Protein concentration was determined by Bradford assay (Bio-Rad Laboratories). This stock solution was used for subsequent demethylase activity assays as described below.

2.7 WESTERN BLOTTING

Equal volumes of the assay sample were separated on SDS-PAGE and transferred onto a 0.2 µm supported nitrocellulose membrane (Bio-Rad Laboratories) at a constant voltage of 40 for 2.5 h at 4 °C. Membranes were blocked with 20 mL of TBST buffer (50 mM Tris HCl pH 7.4, 200 mM NaCl, 0.01% Tween-20) with 5% nonfat dry milk for 1 hr. at room temperature with gentle shaking. The blocking buffer was then removed and membranes were washed with 20 mL of TBST buffer. Immunoblotting was performed with 1:500 diluted primary antibody (Histone-H3 mAb, cat# 61475, Active Motif; Histone H3K9Me3 mAb, cat# 61013, Active Motif; Histone H3K9Me1, cat# 39681, Active Motif) for 12 h at 4 °C. The antibody solutions were removed and membranes were washed three times with TBST buffer. The blots were then incubated with

HRP-conjugated secondary antibody Goat anti-Mouse IgG (cat# 15014, Active motif) with 5% nonfat dry milk (1:10000 dilution) in TBST for 1.5 h at room temperature. After similar washing, protein bands were visualized by chemiluminescence using VISIGLO HRP Chemiluminescent substrates A and B (cat# N252-120ML and N253- 120ML, aMReSCO) following manufacturer's protocol.

2.8 LC-MS ANALYSIS

Protein samples were loaded onto a PLRP-S column (Higgins Analytical, 5 μm, 1000A, 300 μm i.d.× 100 mm) of the LC system (Ultimate 3000, Dionex, Sunnyvale, CA) online coupled to an electrospray ionization (ESI) time-of-flight mass spectrometer (micrOTOF, BrukerDaltonics, Billerica, MA). Chromatographic separation was performed at a constant flow rate of 3.5 μL/ min using a binary solvent system (solvent A: water with 2.5% acetonitrile and 0.1% formic acid; solvent B: acetonitrile with 0.1% formic acid) and a linear gradient program (0-5 min, 5% B; 5- 10 min, 5-30% B; 10-35 min, 30-70% B; 35-44 min, 70% B; 44-45 min, 70-5% B; 45-60 min, 5% B). Mass spectra were acquired in positive ion mode over the mass range m/z 50 to 3000. ESI spectra were deconvoluted to obtain molecular ion masses with Data Analysis 3.3 (Bruker Daltonics, Billerica, MA) using the MaxEnt algorithm resulting in a mass accuracy of 0.01%.

2.9 IN VITRO ASSAYS

2.9.1 MALDI-TOF Demethylase Assay

To monitor enzymatic activity, a demethylase activity assay was optimized and then observed through MALDI-TOF. Each histone demethylase assay sample included 10 μM enzyme, 10 μM peptide, 1 mM 2-KG/analogues, 50 mM Tris pH 8, 50 μM $(NH4)_2Fe(SO_4)_2$, and 2 mM L-ascorbic acid with a total assay volume of 20 μl. Fe(II) supplement and L-ascorbic acid were prepared freshly. The 2-KG was added last to the assay sample and briefly centrifuged. The samples were incubated at 37 °C for three hours. To observe demethylase activity, 1 μL of assay sample was applied to the MALDI plate followed by 1 μL of CHCA matrix. The sample was analyzed using Voyager or Bruker on the Reflector Positive mode. The negative control included all components of the assay except for the demethylating enzyme.

For KDM inhibition studies, the demethylation protocol remains the same as for the activity, but the inhibitor was added (1-4 mM for 2-HG analogues; 100 μM for NOG analogues) to the samples prior to the addition of 2-KG analog (80 μM).

2.9.2 Coupled Fluorescence Assay

The coupled fluorescence assay was used for enzyme kinetics experiments.

2.9.2.1 FDH Fluorescence Activity Assay

Following FDH expression and purification, an activity assay was performed prior to the initiation of the coupled fluorescence assay used for kinetic characterization of bump-hole pairs. The assay was composed of the following reagents: 50 mM HEPES pH 8.0, 50 μM $(NH4)_2Fe(SO_4)_2$, 1 mM L-ascorbic acid, 1mM NAD+, 1 mM formaldehyde, and 100 nM FDH. The control sample did not contain FDH. The samples incubated at 37 °C for 10 minutes. The accumulation of NADH was measured by fluorescence intensity (excitation= 340 nm, emission= 490 nm) on a TECAN Infinite M1000Pro. The data were analyzed using Graphpad Prism software and plotted using RFU following normalization (Figure 9). All samples were performed in duplicates (172).

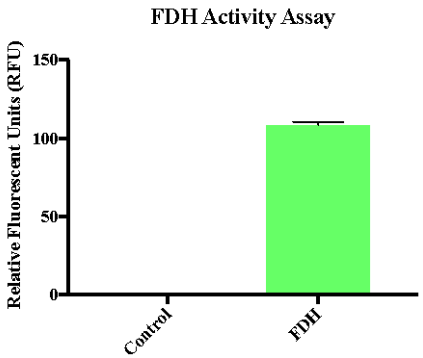


Figure 9: FDH Activity Assay.

Fluorescence activity assay for FDH. Relative fluorescence units (RFU) plotted for both control sample and Experimental sample.

2.9.2.2 NADH calibration curve

To convert the relative fluorescence unit (RFU) values collected in the coupled fluorescence intensity assay, an NADH calibration curve was generated by measuring fluorescence intensity (Figure 10). NADH (ACROS, cat #124530050) was added to assay buffer (50 mM HEPES pH 8, 50 μM $(NH4)_2Fe(SO_4)_2$, 1 mM L-ascorbic acid, 1 mM NAD+, 200 nM FDH, and 1 μM KDM4A) in varying concentrations (0-10 μM). The experiments were performed in 384-well white Corning plates. The fluorescence intensity was measured with excitation = 340 nm and emission = 490 nm on a TECAN Infinite M1000Pro. Triplicate values were taken for each concentration (172).

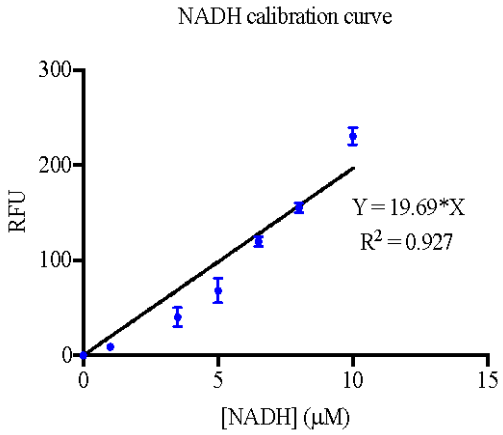


Figure 10: NADH calibration curve.

2.9.2.3 KDM4 Kinetics Characterization

To determine the catalytic efficiency of the wild type and engineered KDM4A-2-KG pairs, the demethylase activity was measure by fluorescence intensity of accumulated NADH by employing a coupled fluorescence intensity assay. The experiments were performed in 384-well white Corning plates. The assay was composed of an enzyme cocktail (50 mM HEPES pH

8.0, 50 μM $(NH4)_2Fe(SO_4)_2$, 1 mM L-ascorbic acid, 1 mM NAD+, 200 nM FDH, and 1 μM KDM4A or KDM4A-F185G and a substrate cocktail (0-175 μM 2-KG/analogue (color key in Figure 11) and 300 μM H3K9Me3 peptide. The substrate cocktail was first applied to the microplate followed by the enzyme cocktail. The accumulation of NADH was measured by fluorescence intensity (excitation= 340 nm, emission= 490 nm) every 30 seconds over a period of 15 minutes on a TECAN Infinite M1000Pro. The data were analyzed using Graphpad Prism software and converted to μM of H3K9Me3 demethylated using the NADH calibration curve. Only points within the linear range (Figure 11) were used to calculate the slope for each 2-KG concentration and the values were fitted to the Michaelis-Menten equation to get the K_m. Experiments were performed in duplicates (172).

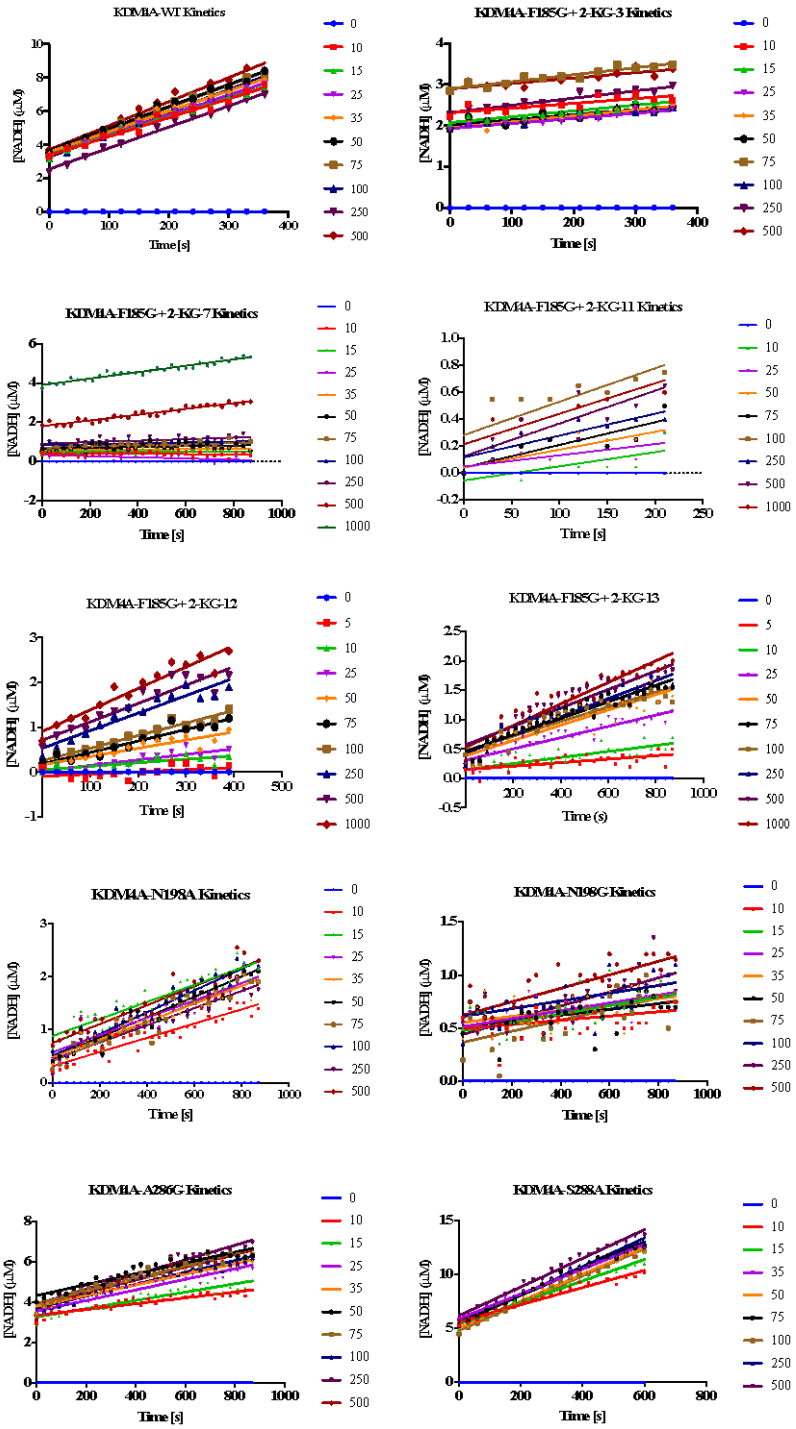


Figure 11: Coupled Fluorescence Assay linear curves for Kinetic Characterization.

Native and mutant KDM4 enzymes were kinetically evaluated. RFU values were converted to concentration of NADH using the NADH calibration curve. This was plotted against time (seconds) to generate the slope for the Michaelis-Menten equation. Each linear line of best fit corresponds to the concentration of 2-KG (number key on right of each graph).

2.9.3 G9a -KDM4 coupled assay

Bacterially expressed histone H3 was resuspended in 0.1% Trifluoroacetic acid (TFA) and then subjected to a two-step assay where the histone was first methylated by G9a and then demethylated by KDM4A-F185G. 20 µg of histone was incubated with 4 µM N-6xhis-G9a, 125 µM SAM and 50 mM Tris pH 8 at room temperature for 10 minutes. Following incubation, N-6xhis-G9a was pulled down using 50 µl Nickel beads by shaking the sample with Nickel resin for 40 minutes at room temperature. The cleared sample was then transferred to a fresh microcentrifuge tube and the histones were subjected to demethylation by F185G. The demethylation assay was composed of the following: 20 µg methylated histone, 50 µM F185G, 1 mM 2-KG analog, 50 mM Tris pH 8.0, 100 µM $(NH4)_2Fe(SO_4)_2$, 2 mM L-ascorbic acid, and 10% glycerol. After 1.5 h of incubation at 37ºC, the sample was subjected to western blotting. The negative control for the G9a methylation assay included a sample without G9a. The negative control for the demethylation assay included a sample without F185G. All samples were performed in duplicates.

2.9.4 Demethylation in HEK293T cell lysates

For demethylase investigation, 100 µg of HEK293T cell lysate was incubated with 50 µM of KDM4-WT or KDM4-mutant in buffer containing 25 mM Tris pH 8.0, 100 µM $(NH4)_2Fe(SO_4)_2$, 2 mM L-ascorbic acid, 10% glycerol and 1 mM 2-KG/analogues. After 2.5 h of incubation at 37 ºC the samples were analyzed for demethylation by western blotting.

For KDM inhibition studies, the demethylation protocol remains the same as for the activity, but the inhibitor was added (1-4 mM for 2-HG analogues; 100 μM for NOG analogues) to the samples prior to the addition of 2-KG analog (80 μM).

2.9.5 Copper Catalyzed Click Assay

Bacterially expressed and purified KDM4 protein (WT or AzF mutant) was either exposed to 365 nM light (for varying amounts of time) or kept out of UV range. Following exposure, 10 μg enzyme was incubated with 100 μM Tris(benzyltriazolylmethyl)amine (TBTA), 100 μM TAMRA-alkyne, 1 mM copper sulfate ($CuSO_4$), and 2 mM freshly dissolved Ascorbic acid for 2 hours at 37 C shaking at 350 rpm. Following the incubation, in-gel fluorescence was performed using 8% SDS-PAGE to visualize clicked band. Negative controls included samples without enzyme as well as samples with wild type enzyme.

2.9.6 In-gel Fluorescence Crosslinking

The optimized in gel fluorescence conditions are as follows: 30 uM enzyme, 5 uM TAMRA-H3K9me3 peptide, 100 uM $(NH4)_2Fe(SO_4)_2$, 2 mM Ascorbate, 50 mM Tris-HCl and 200 mM 2-KG, 1mM NOG. The samples incubated at room temperature for approximately 2 minutes and were then exposed to 365 nM UV light for 40 minutes in an ice bath. The samples were then mixed with SDS-PAGE loading dye and analyzed via 8% SDS-PAGE fluorescence. The gels were coomassie stained to validate that the enzyme was in each sample. Negative controls included the lack of enzyme, the addition of wild type enzyme that will not undergo

crosslinking and the lack of UV exposure for each sample. All samples were performed in duplicates.

Optimization of this system included varying enzyme concentration, type of fluorephore utilized, varying peptide concentration, varying incubation and UV exposure time, varying 2-KG concentration and the addition of different 2-KG competitive inhibitors.

2.10 DEMETHYLATION IN LIVE HEK293T CELLS

HEK293T cells were transfected with 15 ug full length native or mutant (F185G) KDM4A plasmid using Lipofectamine 2000. The cells were treated with 1 mM cell permeable dimethylester-2-KG analog 24 hours post transfection. The cells were harvested 24 hours after treatment with compound. Histone extraction was performed and Histone H3 was quantified using 15% SDS-PAGE and coomassie staining. Western blotting was then performed on equal amounts of histone using antibodies for H3K9me1, H3K9me3 and Histone H3. All experiments were performed in duplicates. Control samples included the transfection with empty PCMV vector and no treatment.

3.0 RESULTS AND DISCUSSION

3.1 KDM BUMP HOLE ACTIVATION

Studies have shown that orthogonal protein:ligand pairs can be generated through a ’bump-and-hole’ approach where site directed mutagenesis is performed on a protein of interest to generate a hole, which can be allele-specifically activated or inhibited by a compensatory bulky ligand. Because of cofactor 2-KG redundancy, we generated a bump-and-hole system to study the KDM4 family of epigenetic modifiers. Previously, we have shown that KDM4A can be specifically activated by bulky 2-ketoglutarate cofactors. Upon mutagenesis, substrate specificity is conserved and the approach in other KDM4 members (KDM4B, KDM4C, KDM4D) showed similar orthogonal results. Although the 2-KG analogues from our previous study did not activate other native 2-KG dependent oxygenases, TET2, FTO, and KDM6B, our specificity screening was not exhaustive and the catalytic efficiency of the engineered system is moderately reduced compared to the native KDM4A-2-KG pair (171-172). Thus, we expanded the repertoire of 2-KG analogues to generate an improved allele-specific activation system for KDM4s.

3.1.1 Generation of 2-Ketogluarate Library

Previously, we identified a phenylalanine residue within the JmjC domain that is conserved in KDM4A-D and is perfectly poised towards the 2-KG cofactor (Figure 12A-B). We found that this residue can accommodate a bulky 2-KG cofactor analog upon mutagenesis to a significantly smaller glycine residue. Specifically, residue F185 of KDM4A and its corresponding residues in KDM4C and KDM4D (F187, F189, respectively) efficiently utilize 2-KG-ethyl (3) and 2-KG-isobutyl (7) for demethylase activity when mutated to a glycine. We aimed to improve our orthogonal KDM4-2-KG engineered pair by identifying a bulkier 2-KG analog with similar kinetic efficiency to reduce potential small molecule promiscuity during in-cell experiments. First, the 2-KG analog library was expanded for demethylase activation screening.

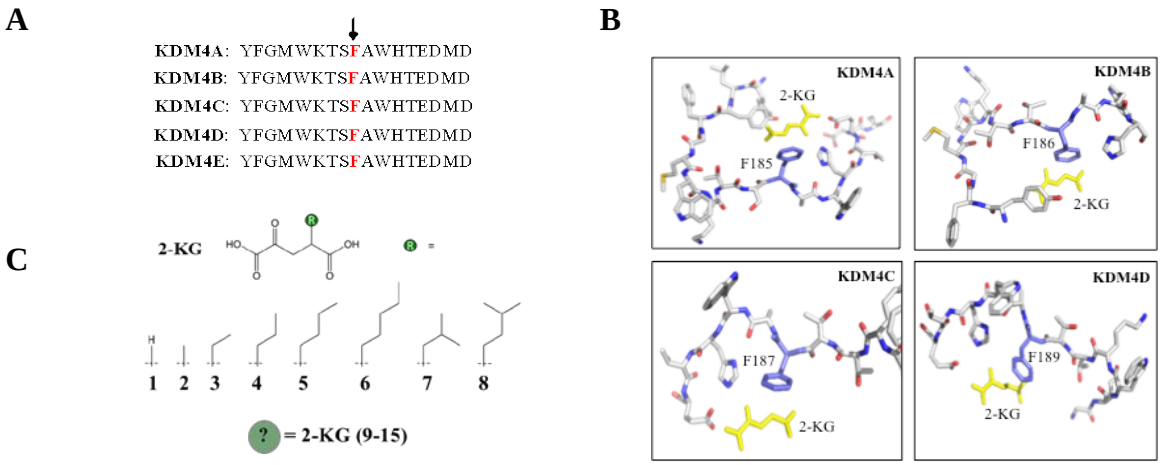


Figure 12: Development of KDM4 Hole and Bumped 2-KG Analogues

A) Amino acid sequence alignment of KDM4 subfamily A-E. The red highlighted phenylalanine (F) residue is the site of interest for the hole being generated in each enzyme. B) Crystal structure of KDM4A-D (respective PDB: 2YBS, PDB: 4LXL, PDB: 4XDO, PDB:4HON). The 2-KG (yellow) binding pocket is shown with surrounding amino acids in white and the phenylalanine residue of interest in blue. C) 2-Ketoglutarate analog library. Analogues with will be referred to by the bold number below each functional group.

A total of eight 2-ketoglutarate analogues were previously screened for KDM4A demethylase activation. Seven additional 2-KG bulky analogues were synthesized, purified and validated for screening for demethylase activation of KDM4 native and mutant constructs (Figure 12C).

3.1.2 Biochemical characterization of 2-KG analogues for KDM4 activation

The JmjC catalytic domain of native and mutant KDM4A, KDM4C and KDM4D were bacterially expressed and purified for demethylase activity screening with the new generation of 2-KG analogues (**9-15**). The KDM4 family is known to demethylate histone H3K9me2/me3 (Figure 13A). To measure demethylase activity, a simple mass spectrometry MALDI-TOF assay was utilized. Demethylation is a mass change of 14 Daltons resulting in a resolvable mass peak in the MALDI-TOF spectrum (Figure 13B). The proportion product (H3K9me1/me2) for each screening is depicted in the heat maps of Figure 13C-E. The demethylase assay screening for all KDM4 members resulted in similar activation patterns previously observed when using the first generation of 2-KG analogues (**1-8**). The native constructs were completely activated by 1 and were not activated by **9-15**. The mutant KDM4 constructs were not activated by 1 and were activated to some extent by almost all of the 2-KG analogues screened. KDM4A-F185G was activtated by all of the 2-KG analogues, with the greatest extent of activation observed when paired with **3, 7, 11, 12**, and **13**. KDM4C-F187G was activated most strongly by **3, 7,** and **11** and KDM4D-F189G was strongly activated by **3, 7,** and moderately activated by **12** and **13**. The cyclical 2-KG derivatives – **9, 14**, and **15**, failed to induce demethylase activity, demonstrating that the JmjC engineering cannot accommodate these extremely bulky residues for the enzymes hydroxylation mechanism. Together these results further demonstrate the

orthogonality of the engineered system and provide a larger library of compounds to investigate further for proof of concept validation.

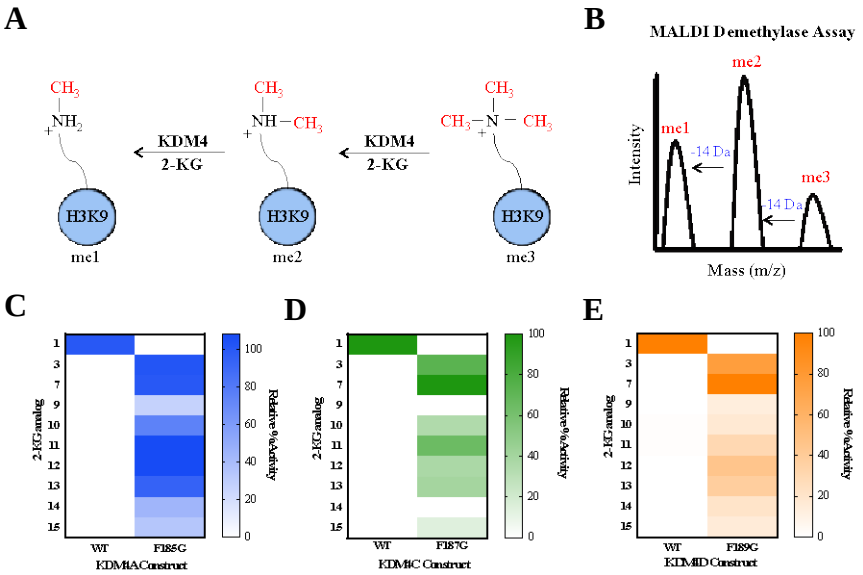
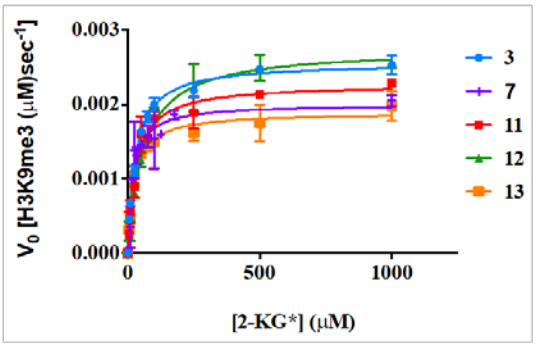


Figure 13: Screening Engineered KDM4-2-KG Pairs for Demethylase Activity.

A) Schematic of successive lysine demethylation by KDM4 through the utilization of 2-KG. B) MALDI-TOF spectrum cartoon depicting mass change corresponding to methyl status. The removal of one methyl group is a 14 Dalton loss. C-E) Heat map diagram for KDM4A/C/D bump-hole activation screening with **1, 3, 7, 9-15**.

The MALDI-TOF demethylase screening results show that **3**, **7**, **11**, **12** and **13** can activate the KDM4 mutants, but the assay was developed to reduce the potential false negatives. Functionally, enzymes rapidly perform their mechanism over and over again, meaning one KDM4 molecule can demethylate several molecules of histone H3K9 successively. This mechanism allows for one individual KDM4 molecule to remove two methyl groups from the same H3K9 substrate. The MALDI-demethylase assay uses a 1:1 enzyme to substrate ratio to eliminate the 'rate of demethylation' variable. Additionally, there are over 60 enzymes in humans that have been identified that utilize 2-KG as a cofactor. The average cellular

concentration of 2-KG in human cells is approximately 40 μM. The concentration of 2-KG applied in the MALDI demethylase assay is 1 mM to eliminate the 'effective concentration' variable. Epigenetic enzymes are kinetically efficient allowing a plethora of proteins to be functional with only 40 μM of their necessary cofactor. To determine if the KDM4-2-KG engineered pairs identified are kinetically competent, a coupled-fluorescence assay was implemented. Previously, we identified KDM4A-WT needs approximately 12 μM 2-KG (**1)** tobe 50% active. The Km for KDM4A-F185G was within 5-fold for derivatives **3, 7, 11, 12** and **13**. The Km for KDM4A-F185G was the lowest for **7** and **13**, but they had the lowest turnover rate depicted by the Michaelis-Menten curve in Figure 14. Interestingly, **12** had a turnover rate similar to **3**, but the Km is the highest of the compounds screened (F185G + **3** = 29.49 ± 1.8 μM; F185G + **7** = 19.72 ± 4.2 μM; F185G + **11** = 30.81 ± 4.8 μM; F185G + **12** = 58.07 ± 5.4 μM; F185G + **13** = 23.9 ± 3.6 μM). The kinetic assessment mildly differentiates superior compounds for the KDM4A-F185G pair, but because they are all within a similar range and close to the kinetic characteristics of the native system, all compounds were investigated further. In addition to kinetic characterization, these compounds have been screened with other 2-KG dependent enzymes to determine potential cofactor promiscuity. Native FTO, TET and KDM6B constructs did not show activity with **3**, **7**, **11**, **12**, or **13**.



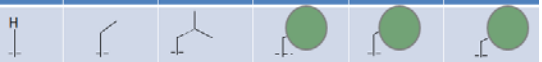
	1	3	7	11	12	13
						
Km (μM)	–	29.49 ± 1.8	19.72 ± 4.2	30.81 ± 4.8	58.07 ± 5.4	23.9 ± 3.6

Figure 14: Michaelis-Menten Plot of KDM4A-F185G-2-KG Pairs

Kinetic characterization of KDM4A-F185G-2-KG analog pairs. Top chart shows the Michaelis-Menten plot with each curve corresponding to the corresponding 2-KG analog (see key). Bottom table provides values for the Km identified for each 2-KG analog when paired with KDM4A-F185G.

3.1.3 Demethylation towards Full Length Histones

The engineered KDM4A-2-KG pairs were subjected to a coupled G9a-KDM4A assay to determine if the systems are active towards full length histone substrates and to demonstrate the dynamic nature of histone methylation modifications. Bacterially expressed full length histones were exposed to lysine methyltransferase G9a and universal methyl donor, S-adenosylmethionine (SAM), to generate dimethylated H3K9, a substrate of KDM4A. The newly modified H3K9 was then incubated with the KDM4A-F185G-2-KG pair of choice to generate monomethylated H3K9 product. All engineered pairs resulted in abundant H3K9me1 compared to the H3 product produced by G9a and SAM (Figure 15A-B). The bump hole pairs for KDM4A/C/D were then tested in HEK293T cell lysates to determine if they are applicable in a physiologically relevant environment. HEK293T cells were treated with potent KDM4

inhibitor, n-octyl-IOX1 to enrich the cells with KDM4 substrate, H3K9me3 (Figure 15C). Upon exposure of lysates to the engineered pairs, there was a dramatic reduction in H3K9me3 levels for all systems (Figure 15D-F). In cell lysates, F185G showed greatest activity when paired with **11, 12,** and **13**, depicted by the loss of H3K9me3 signal. KDM4A-F185G paired with **11-13** seems to improve demethylation activity suggesting that the bump-hole system has been improved compared to our previous results. F187G and F189G showed similar levels of activation when implemented in live cells.

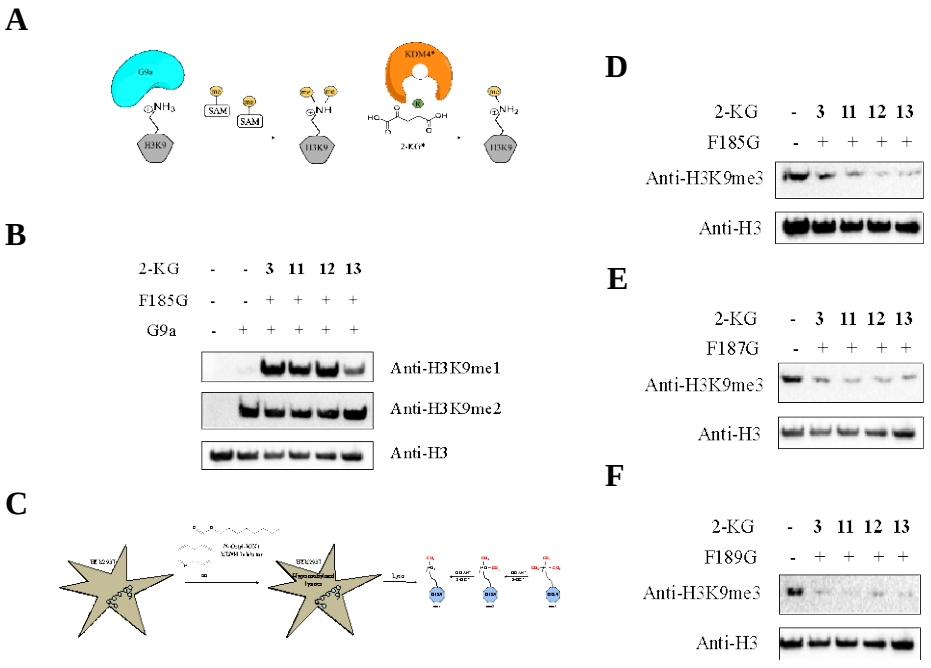


Figure 15: Bump Hole Activation towards Full Length Histones.

A) Cartoon schematic of the G9a-KDM4 coupled assay. G9a dimethylates Histone H3K9. The dimethylated histone is then incubated with the KDM4-2-KG pair to remove the methyl groups. B) Western blot of G9a-KDM4 coupled assay. KDM4A-F185G was paired with **3, 11-13** following histone methylation by G9a. Western blotting was performed to evaluate levels of H3K9me1 and H3K9me2. C) Cartoon schematic of bump hole activation in HEK293T lysates. Western blot of bump hole activation in HEK293T cells for D) KDM4A-F185G, E) KDM4C-F187G, F) KDM4D-F189G. Western blotting was performed to evaluate levels of H3K9me3.

3.1.4 Demethylase Bump-Hole Activation in HEK293T cells

The *in vitro* investigation of the KDM4 bump-hole systems suggests complete orthogonality and kinetic competence providing confidence to implement this approach in live mammalian cells. The cell membrane is composed of a lipid bilayer for regulated trafficking of molecules between the cell's internal and external environment. Because of the membrane's hydrophobic characteristics, 2-KG and the 2-KG analogues are not cell permeable. To permeablize the cofactors of interest, **1**, **3**, and **7** were esterified to become dimethyl-ester derivatives (**1**$_{DM}$, **3**$_{DM}$, **7**$_{DM}$, respectively, Figure 16A). This process increases the hydrophobicity of the cofactors, making them more permeable. Cellular esterases will cleave the ester bonds to allow for the compounds to resume the structure screened *in vitro* and be accommodated in the KDM4 mutant catalytic domain, as well as, reduce their permeability effectively sequestering the compound in the cell.

Full length constructs of native KDM4A and F185G mutant KDM4A were transiently transfected into HEK293T cells treated with KDM4 inhibitor, IOX1. HEK293T cells were then treated with **1**$_{DM}$, **3**$_{DM}$, or **7**$_{DM}$. Differences in levels of H3K9me3 were compared between samples via western blotting which showed modest decreases in cells treated compared to empty control cells (Figure 16B). There is a significant decrease in H3K9me3 upon cellular treatment with F185G + **7** compared to the treatment with the native system and the empty control cells. Additional experiments were performed to conclude that the decrease in trimethylated H3K9 observed is not due to compound treatment alone or transfection treatment alone. The implementation of the developed KDM4A-2-KG bump hole system in HEK293T cells is promising, but needs further validation and can be improved. Additional esterified 2-KG

derivatives need to be tested including **11-13$_{DM}$** and potentially even longer ester derivatives. For instance, the addition of an octyl-ester chain has shown to greatly improve permeability of compounds, including IOX1, so this may be a strategy to implement onto these 2-KG analogues to improve potency.

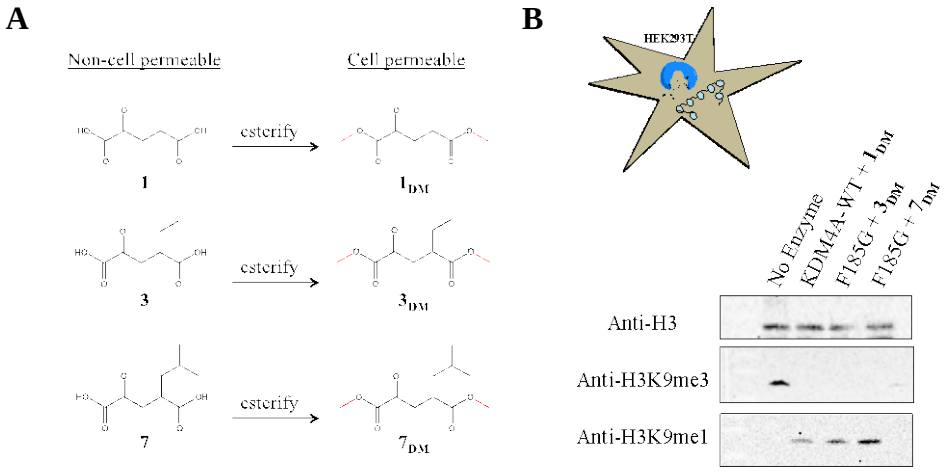


Figure 16: Bump Hole Activation of KDM4A in HEK293T cells.

A) Cell permeable 2-KG compound structures. B) Western blot of bump hole activation in HEK293T cells. Levels of H3K9me3 and H3K9me1 were evaluated for cells treated with enzyme and compound and compared to the control sample (No enzyme).

3.1.5 Generality of Method to other KDMs members

KDM6B is another JmjC domain containing enzyme that is known to demethylate H3K27me1/me2/me3 to transcriptionally activate target genes (Figure 17A). H3K27me3 is a crucial histone modification for the repression of various genes involved in cell cycle regulation and development. Additionally, hypermethylation of H3K27 contributes to cell proliferation in cancers such as pancreatic ductal adenocarcinoma. KDM6B is downregulated in several cancers potentially keeping tumor suppressor genes inactivated. To assess the generality of the bump-

hole approach developed for KDM4, KDM6B was engineered in a similar fashion and screened for demethylase activity when paired with the 2-KG analogues. The crystal structure of KDM6B was used to target residues for mutagenesis for catalytic domain expansion. There are several bulky residues residing in the 2-KG binding pocket of KDM6B. Previous engineering performed on KDM4A shows that altering some of these corresponding residues completely abolishes enzymatic activity, so these residues were not targeted for mutagenesis in our initial experiments. Valine-1472 is positioned directly above the C4 position of 2-KG, similar to F185 in KDM4A, making it an ideal target for mutagenesis to accommodate bulky 2-KG analogues (Figure 17B, PDB 4EZH). Following mutagenesis of KDM6B to V1472A or V1472G, demethylase activation was screened using H3K27me3 peptide. The native KDM6B was not activated by any 2-KG analogues, which further demonstrates the orthogonality of the F185G engineered system. Interestingly, V1472A was not activated by any 2-KG analogues, but V1472G mutant showed strong demethylase activity when paired with **1-6,** and **11** (Figure 17C). V1472G, however, does not provide a fully orthogonal system, as it accepts native 2-KG, but a concentration dependent assay suggests that V1472G will only be weakly activated by 2-KG under more physiological conditions. Demethylation of H3K27me3 peptide by native and mutant KDM6B enzymes was monitored upon activation by varying concentrations of 2-KG (Figure 17D). KDM6B almost completely converted H3K27me3 to non-methylated peptide with 20 μM 2-KG (75% H3K27me0), whereas, V1472A only provided ~50% non-methylated peptide and some trimethylated product remained when incubated with 20 μM 2-KG. At 20 μM 2-KG, V1472G did not produce any non-methylated product and majority of the peptide was di- and tri-methylated. V1472G did not completely demethylate H3K27 even at 1 mM 2-KG. With a cellular concentration of 2-KG being approximately 40 μM, V1472G may not be strongly

activated in cells, allowing for the system to be highly, but not completely orthogonal. Importantly, the mutagenesis does not disrupt substrate specificity, as the mutants did not show demethylase activity towards H3K9me3 or H3K36me3 peptide, keeping the functional integrity intact (Figure 17E). Lastly, KDM6B-V1472G was tested for activation in HEK293T cell lysates (Figure 17F). Because H3K9me3 and H3K27me3 are complementary marks, hypermethylation of H3K9 often results in hypermethylation of H3K27. HEK293T cells were treated with KDM4 inhibitory, n-octyl-IOX1, lysed, and incubated with the engineered KDM6B system. H3K27me3 signals were diminished in the lysates incubated with KDM6B-V1472G + **3/11** compared to controls. These results strongly suggest that the corresponding KDM4-F185 residue in other JmjC containing proteins may be a 'gate-keeper' residue for generating demethylase bump-hole systems. Because the orthogonality for the KDM6B engineering is not perfect, further structural alterations may need to be done for other KDMs to provide completely orthogonal pairs. Potential residues to mutate include the tryptophan, tyrosine and phenylalanine residues residing in the 2-KG binding pocket.

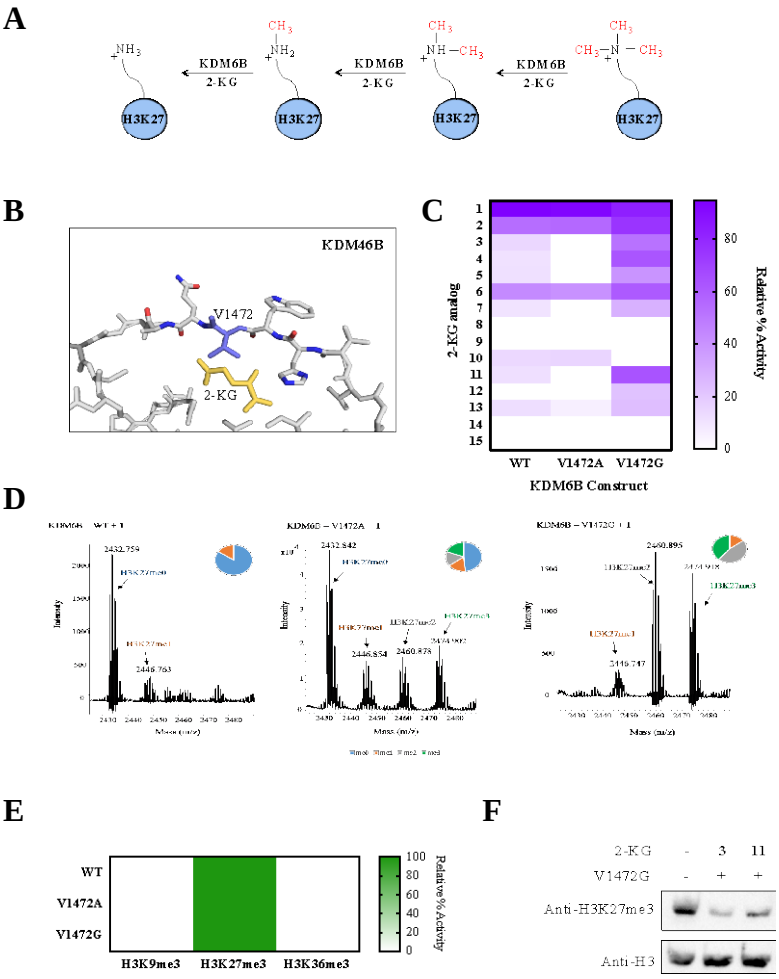


Figure 17: Bump Hole Activation of KDM6B.

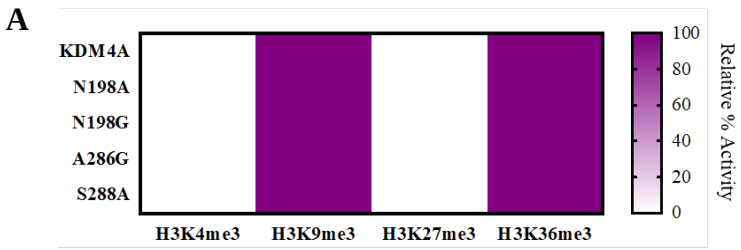
A) Cartoon schematic of histone demethylation by KDM6B. KDM6B can completely remove methyl groups from histone 3 lysine 27 (H3K27). B) Crystal structure of KDM6B (PDB: 4EZH). 2-KG (yellow) binding pocket is shown with the surrounding amino acid residues in white and the valine of interest in blue. C) Heat map diagram of demethylase activity towards H3K27me3 for the KDM6B bump hole pairs. D) MALDI spectrum showing activity of KDM6B-WT/V1472A/V1472G with 20 μM 2-KG towards H3K27me3. Pie charts in the upper right corner of each spectrum depict the methyl status distribution for each spectrum (me0: blue, me1: orange, me2: grey, me3: green). E) Heat map diagram of substrate specificity for KDM6B mutants. F) Western blot of KDM6B-V1472G + **3/11** bump hole activation in HEK293T cell lysates. Levels of H3K27me3 were compared for each sample.

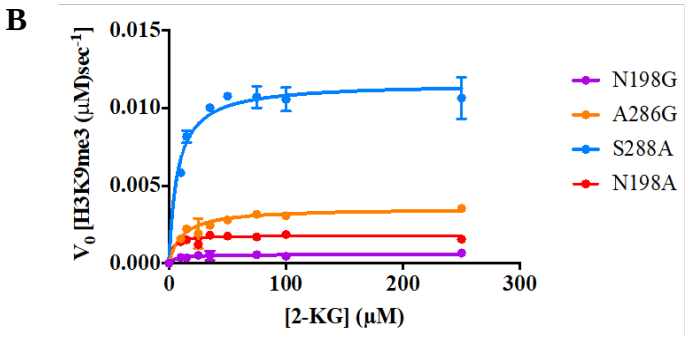
3.2 KDM BUMP HOLE INHIBITION

Several JmjC domain containing enzymes are overexpressed in disease, providing the rationale to develop an orthogonal inhibition system in tandem with the bump-hole activation. Having both an activator and inhibitor for the same protein will provide an 'on and off' switch for gene regulation by a specific JmjC domain containing enzyme of choice.

3.2.1 Generating and Biochemically Characterizing KDM4A Mutants

Additional KDM4A mutants were generated for potential inhibition. Again, the crystal structure of KDM4A was analyzed to target specific amino acids for mutagenesis. Four mutants were generated harboring an expanded catalytic domain: N198A, N198G, A286G and S288A. Similar to the bump-hole activation system, these mutants were screened for demethylase activity when paired with **1-15**. All four of the mutants only showed activity with **1** towards H3K9me3 and H3K36me3. Additionally, the mutagenesis did not disrupt the substrate specificity of these enzymes depicted by the lack of demethylation towards H3K4me3 and H3K27me (Figure 18A). These mutants were then kinetically characterized. Interestingly, the mutants had lower Km values compared to the Km identified for the native KDM4A (WT; N198A; N198G; A286G; S288A) (Figure 18B). Also, KDM4A-S288A most efficiently used **1** compared to the other mutants based on the Michaelis-Menten curve in Figure 18.





	N198A	N198G	A286G	S288A
Km (μM)	2.325 ± 1.0	6.3 ± 4.6	11.45 ± 1.6	7.4 ± 1.3

Figure 18: Biochemical Characterization of Bump Hole Inhibition Pairs.

A) Heat map diagram of demethylase activity of KDM4A mutants towards H3K4me3, H3K9me3, H3K27me3 and H3K36me3. B) Michaelis-Menten curves for KDM4A mutants when paired with 2-KG-**1**. The respective Km values are in the table below the graph.

3.2.2 Screening KDM4A Inhibition with 2-HG Analogues

To initiate the development of this system we used known 2-KG competitive inhibitor, 2-Hydroxyglutarate (2-HG), as a skeleton to develop 2-HG bulky analogues (Figure 19A). Following the synthesis and purification of 2-HG analogues, KDM4A/C/D, KDM6B and their respective mutants (F185G, F187G, F189G, V1472A, V1472G) were screened for demethylase inhibition via MALDI. All native constructs were strongly inhibited by **16** and were not inhibited by **17-20**. F185G, F187G and F189G were strongly inhibited by **17-20** clearly demonstrating that the mutagenesis accommodates the bulky group of the 2-HG analogues.

Additionally, the approach is further validated for generality among JmjC domain containing enzymes observed by the orthogonal inhibition pairs generated for V1472A and V1472G, specifically when paired with **20** (Figure 19B). The bump hole inhibition pairs were then tested in HEK293T cell lysates to determine if they are applicable in a physiologically relevant environment. HEK293T cells were treated with potent KDM4 inhibitor, n-octyl-IOX1 to enrich the cells with KDM4 substrate, H3K9me3. Upon exposure of lysates to the engineered pairs, there was a dramatic reduction in H3K9me3 levels for systems incubated with 2-KG analog. The bump hole inhibition systems were demonstrated in cell lysates when the 2-HG analog was added to the sample (Figure 19C). Compounds **17** and **18** strongly inhibited F185G, but failed to inhibit KDM4A-WT. Through this demethylase inhibition screening, orthogonal systems have now been generated for KDM4A, KDM4C, KDM4D and KDM6B, but the effective inhibitory concentration of **1** towards KDM4A is in the millimolar range. This screening provides a robust template for the development of a bump-hole inhibition system for JmjC domain containing enzymes, but needs to be improved for the implementation in cells as toxicity may be a limiting factor for the 2-HG analogues.

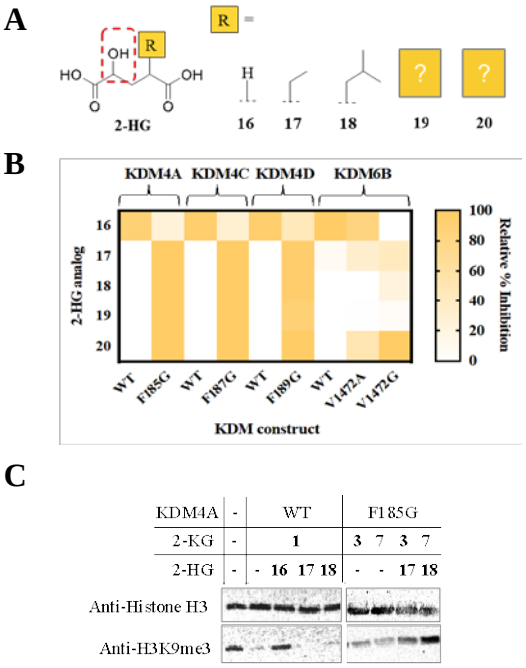


Figure 19: Screening Engineered KDM-2-HG pairs for Inhibition.

A) 2-hydroxyglutarate (2-HG) analogues. The bold number below each functional group will be used to refer to compounds. B). Heat map diagram of demethylase inhibition of KDM bump hole pairs towards H3K9me3 or H3K27me3 for KDM4 or KDM6B, respectively. C) Western blot of bump hole inhibition in cell lysates for KDM4A-F185G paired with **3** or **7** activators and **17** or **18** inhibitors.

3.2.3 Screening KDM4A Inhibition with NOG Analogues

To further develop the bump-hole inhibition system, a more potent 2-KG competitive inhibitor was chosen as the template. N-oxalylglycine (NOG, **21**) is similar in structure to both 2-KG and 2-HG, but has an IC_{50} value of ~17 μM for KDM4A. A total of 32 bulky NOG analogues were screened against the KDM4A mutants (F185G, N198A, N198G, A286G, S288A) for H3K9me3 demethylase inhibition via MALDI (Figure 20A-B). H3K9me3 demethylation by native KDM4A (WT) was completely inhibited by **21** and was weakly inhibited 10 of the analogues.

Unsurprisingly, F185G was inhibited to some extent by almost all of the NOG analogues. F185G was only weakly inhibited by **21** and was strongly inhibited by 14 of the 32 bulky analogues. Interestingly, the KDM4A mutants that readily accept the native 2-KG (**1**) are strongly inhibited by **32**, which is one of the largest NOG compounds in the library. Utilizing NOG as a template has proven to be successful based on this preliminary screening. When screening 2-HG analogues, 1 mM of 2-HG inhibitor was used in the assay, whereas, the NOG screening was performed with only 100 μM of NOG analog inhibitor.

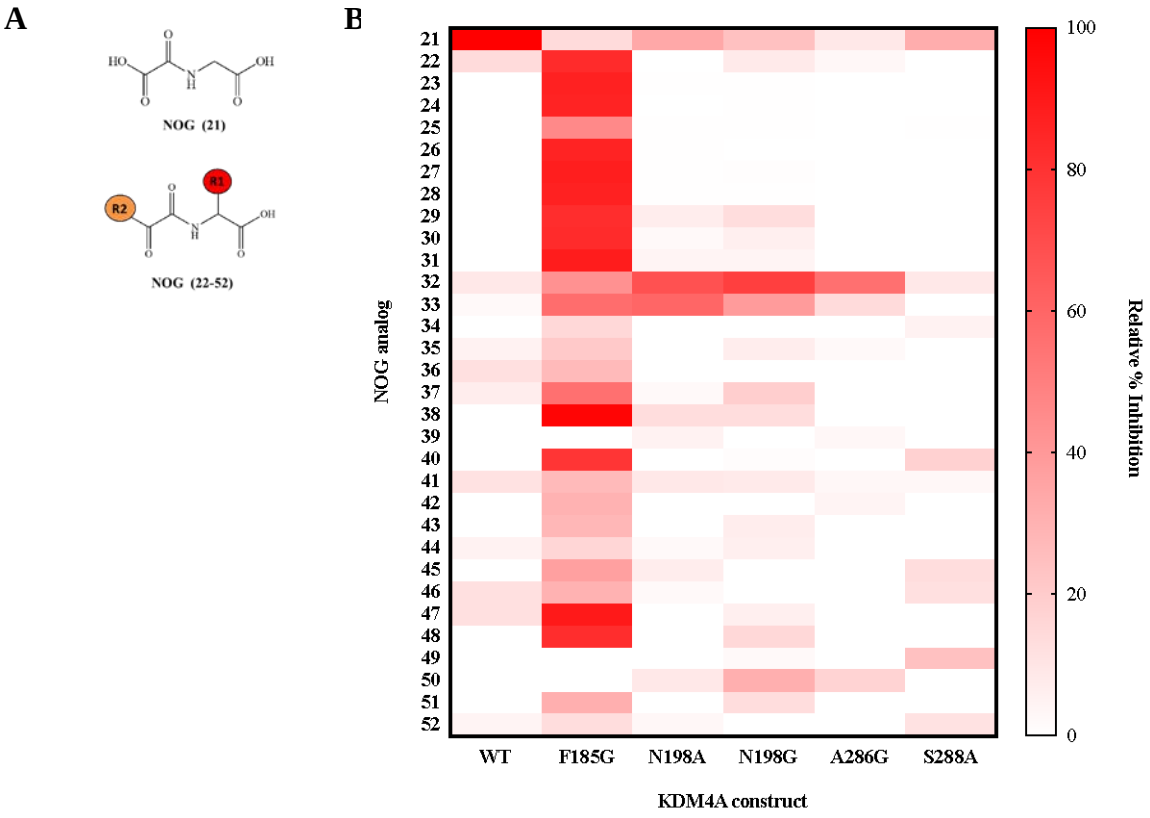


Figure 20: Screening Engineered KDM4A-NOG Pairs for Inhibition.

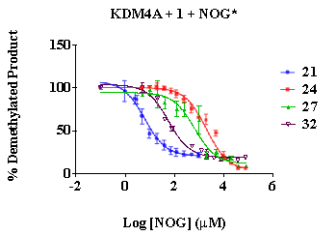
A) N-oxalylglycine (NOG) native compound (**21**) and modification sites (R1 and R2). Compounds **22-52** refer to different NOG analogues that are modified at either R1 or R2 or at both modification sites. B) Heat map diagram depicting percent demethylase inhibition towards H3K9me3. KDM4A construct is shown at the bottom and the compound number is on the left.

3.2.4 IC_{50} Determination

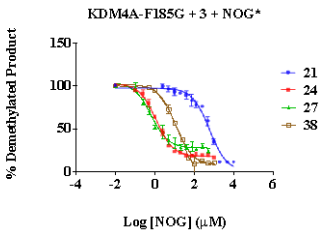
Because NOG is a potent and broad range inhibitor, the IC_{50} of select pairs was measured for the WT and F185G KDM4A (Figure 21A-C). NOG **32** was of immediate curiosity because the use of a mutant that accepts the native 2-KG in live cells could decrease the amount of engineering done and allow a simpler model to be utilized. Because **32** weakly inhibits KDM4A-WT, the IC_{50} needed to be identified. Unfortunately, the IC_{50} of **32** towards KDM4A-WT is low, making it a rather potent inhibitor. Because of this, orthogonality will likely not be achieved once applied in cells. Further engineering on this molecule should be performed to develop the bump-hole system desired. This molecule may be of use as a broad range KDM4 inhibitor once made cell permeable.

KDM4A-F185G is potently inhibited by **24**, **25**, and **28** with micromolar or submicromolar IC_{50} values (F185G + **24** = 1.225; F185G + **27** = 0.6263; F185G + **28** = 10.77). The structural complementarity developed by the system is clearly demonstrated by the difference in IC_{50} value for **21** towards the native and mutant enzyme (WT + **21** = 6.42; F185G + **21** = 605.9). Additionally, **24**, **27**, and **38** did not effectively inhibit KDM4A-WT. The NOG analogues that inhibit KDM4A-F185G have been screened for inhibition of additional 2-KG dependent enzymes and this has further supported their specificity and the orthogonality of this system. Together, these data provide a promising KDM4A-NOG bump hole system that could potentially be employed in concert with the bump hole activation system.

A



B



C

KDM4A-NOG* IC_{50}					
NOG	**21**	**24**	**27**	**32**	**38**
WT (µM)	6.425	2073	645.5	57.58	--
F185G (µM)	605.9	1.225	0.6263	--	10.77

Figure 21: IC50 Determination.

IC_{50} curves for A) KDM4A-WT and B) KDM4A-F185G when paired with **21, 24, 27, 32**. C) Table showing functional group of NOG analog and IC_{50} value (µM).

3.3 PHOTO-INDUCIBLE KDM4 CROSSLINKERS

The specific substrates for each KDM4 member are not well characterized due to the lack of appropriate tools. Chromatin immunoprecipitation (ChIP) is beneficial because it shows where proteins are localizing across the genome, but it is incapable of reporting catalytic activity of enzymes. I plan to incorporate a photo-crosslinkable unnatural amino acid (pUAA) into the catalytic domain of KDM4A without disrupting its enzymatic properties (demethylase activity and substrate specificity) to identify non-histone substrates and genomic loci specifically regulated by this enzyme.

The incorporation of the pUAA into the catalytic domain of KDM4A requires the expansion of the genetic code.[46] Twenty canonical amino acids are used as the building blocks

for protein construction within living systems. Adding the pUAA as a building block involves a unique codon that is not assigned to a native amino acid. For incorporation of the pUAA at specific locations, an evolved orthogonal tRNA/aminoacyl tRNA-synthetase pair will be employed into the host cell (Figure 22).

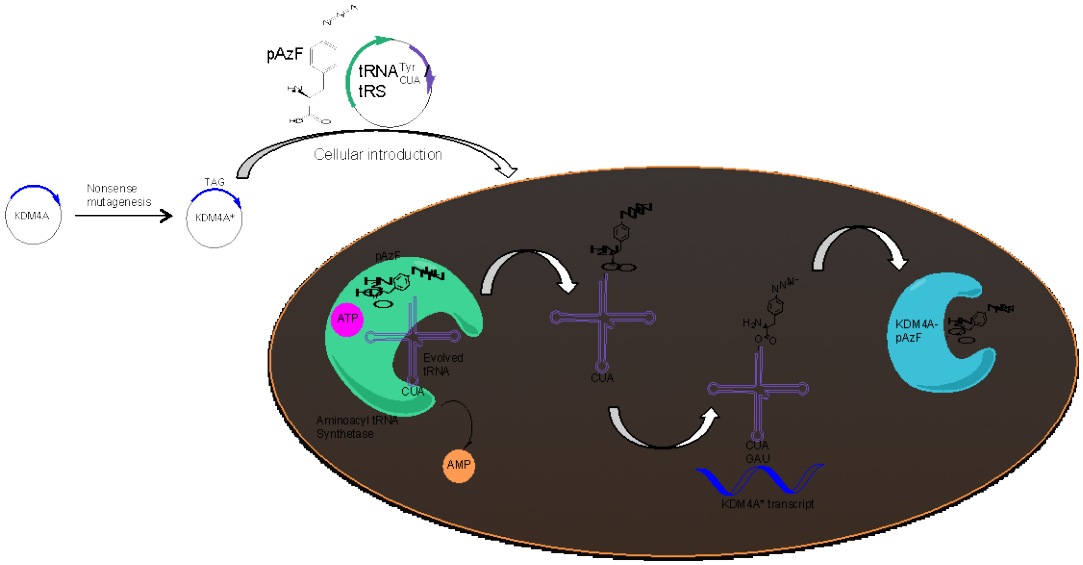


Figure 22: Unnatural Amino Acid Incorporation.

The gene of interest is site specifically mutated with the amber suppressor codon (TAG). Cells are treated with the mutated plasmid, unnatural amino acid, and evolved tRNA synthetase/tRNA pair. The evolved system will incorporate the unnatural amino acid into the protein at the amber codon site.

3.3.1 Generating KDM4A Mutants Containing Unnatural Amino Acids

To express a photo-crosslinkable KDM4A in bacterial cells, the catalytic domain (JmjC) of KDM4A was assessed using the crystal structure (PDB: 2YBS) to identify target residues for the incorporation of para-azidophenylalanine (pAzF), the pUAA of choice based on proximity to the histone substrate (Figure 23A). Residues within 5 angstroms to the histone substrate were chosen for mutagenesis rationalized by the average nitrogen-carbon bond length being

approximately 1.5 angstroms and mutagenesis subtly altering amino acid positions. Following mutagenesis, successful incorporation of pAzF or benzophenone (BpF) (Figure 23B) was confirmed via SDS-PAGE and LCMS-MS (Figure 23C-D). For proper cross-linking to occur, the azide group of the pAzF needs to be intact prior to UV light exposure. To determine if the pAzF was compromised prior to cross-linking assays, a copper catalyzed click assay was performed. Fluorescent Tetramethylrhodamine (TAMRA) labeled alkyne was incubated with the purified protein of interest (KDM4A-I71AzF) in azide-alkyne copper catalyzed click conditions, both before and after exposure to 365 nm UV light. The samples were run on SDS-PAGE for fluorescent analysis. The fluorescent band is only visible prior to UV exposure confirming that the azide group is intact following expression and purification and the protein is UV photo-inducible demonstrated by the lack of a signal (Figure 23E-F).

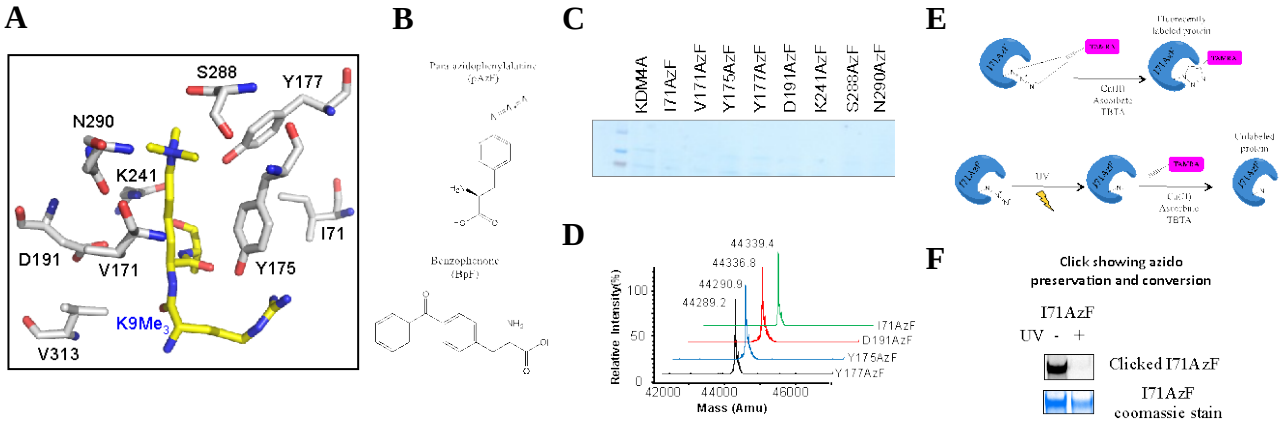


Figure 23: Generating KDM4A Photo-Crosslinkable Mutants.

A) Crystal structure of KDM4A substrate binding pocket (PDB:). Residues targeted for mutagenesis (white) are within close proximity to the bound H3K9me3 (yellow). B) Photo-inducible unnatural amino acids with crosslinking ability. Top is para-azidophenylalanine (pAzF) and bottom is benzophenone (BpF). C) Coomassie stained SDS-PAGE gel of bacterially expressed and purified KDM4A-UAA catalytic domains. D) LCMS-MS results for expressed and purified KDM4A-UAA catalytic domains showing the successful incorporation of pAzF. E) Copper catalyzed click schematic for incorporation and preservation of pAzF into KDM4A. F) In-gel fluorescence and Coomassie stained SDS-PAGE gel of copper catalyzed click reaction on KDM4A-I71AzF with and without exposure to 365 nm UV light.

3.3.2 Assessing the Enzymatic Activity of KDM4A-UAA Mutants

A total of 15 KDM4A-pUAA mutants were successfully expressed and purified, with 2 harboring the bulkier UAA known as BpF (Figure 23B). All mutants were screened for demethylase activity towards H3K9me3 peptide using the e optimized MALDI-TOF demethylase assay. Three of the 15 mutants, I71AzF, Y177AzF, and K241AzF, maintained demethylase activity (Figure 24A) and were further analyzed for substrate specificity remaining intact. To determine if substrate specificity remained intact for the active mutants, demethylase assays were performed on H3K4me3, H3K27me3 and H3R2me2 peptides and measured by

MALDI-TOF. The native and mutant enzymes only showed demethylase activity towards H3K9me3 peptide indicating that substrate specificity is conserved (Figure 24B-C). This screening confirms that 3 of the 15 mutants bind KDM4A histone substrates and maintain the specificity which will reduce false positives in specificity which will reduce false positives in future crosslinking experiments.

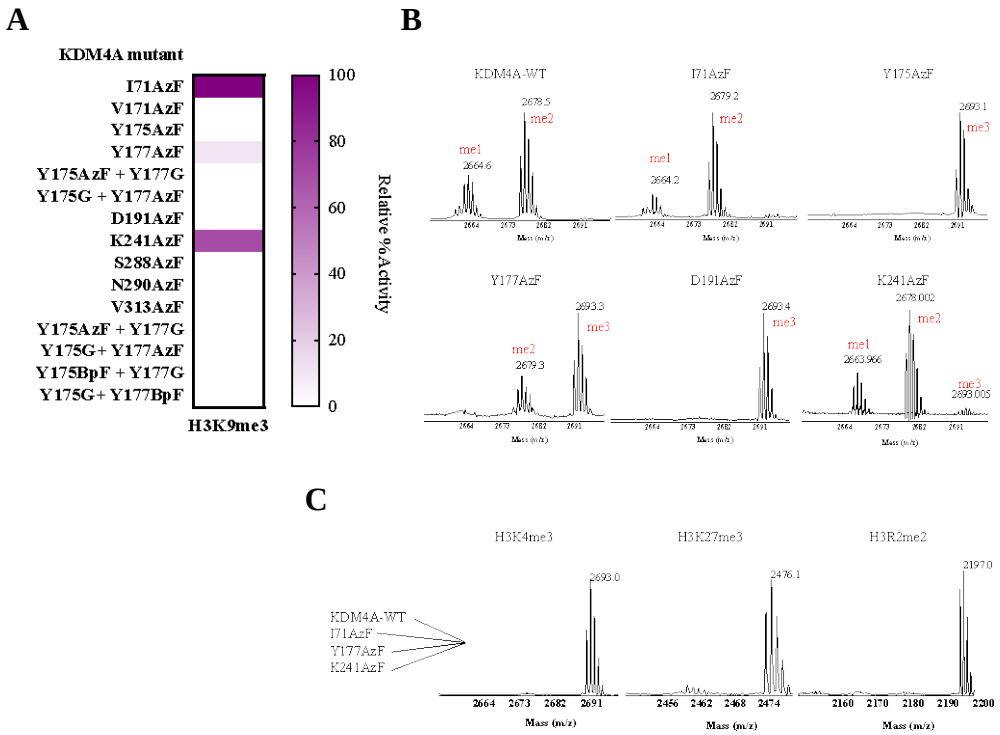


Figure 24: Assessing the Enzymatic Activity of KDM4A-UAAs.

A) Heat map diagram showing catalytic activity of KDM4A-UAAs towards H3K9me3. B) MALDI-TOF spectra showing demethylase activity towards H3K9me3. C) MALDI-TOF spectra showing demethylase activity of active KDM4A-UAAs (I71AzF, Y177AzF, K241AzF) towards H3K4me3, H3K27me3, H3R2me2 to demonstrate substrate specificity.

3.3.3 Screening KDM4A Mutants for Crosslinking Abilities

All mutants were then screened *in vitro* for crosslinking capabilities towards fluorescently labeled carboxyfluorescein (6CF) H3K9me3 peptide by exposing the samples to 365 nm light for 20 minutes and performing in-gel fluorescence (Figure 25A). KDM4A-I71AzF and KDM4A-K241AzF successfully crosslinked to 6CF-H3K9me3, which provides a proof of concept as well as insights that maintained enzymatic activity might be an appropriate filter for determining potential crosslinkable mutants. KDM4A-I71AzF more efficiently crosslinks the peptide substrate compared to the K241AzF mutant based on the higher intensity signal observed via in-gel fluorescence (Figure 25B). Additionally, I71 is further from the methylated lysine residue within the substrate compared to other target residues (Figure 23A). The activity screening and crosslinking assays suggest that I71 does not have a role in substrate binding or orienting the histone for demethylation, but is in close enough proximity to a histone residue to crosslink.

The I71AzF-6CF-H3K9me3 signal validates that the enzyme is photo-inducible, but the signal is not robust. To optimize the crosslinking system to increase efficiency, a time-dependent copper catalyzed click assay was performed. Purified KDM4A-I71AzF was exposed to 365 nm UV light for varying amounts of time (0-60 minutes) followed by the initiation of the click reaction. In-gel fluorescence was performed to determine the optimal UV exposure time for complete nitrene formation. The azide species remained intact with up to 30 minutes of UV exposure and there was a complete loss of azide signal at 40 minutes (Figure 25C). All crosslinking experiments were performed with 40 minutes of UV exposure time following this experiment. In addition to the time-dependent click assay, the fluorphore-labeled histone

H3K9me3 peptide substrate was changed from an unstable carboxyfluorescein (6CF) label to a more robust and sensitive trimethylrhodamine derivative – TAMRA label. KDM4A-I71AzF was then tested for more thorough crosslinking towards TAMRA-H3K9me3 peptide and 40 minutes of 365 nm exposure time. The crosslinked species signal was clearly more intense under these conditions providing confidence that the protein is an effective cross-linker (Figure 25D). This is important because there are likely non-histone substrates that are less abundant in cells compared to histone H3 and the crosslinkable enzyme must be efficient to covalently bind to highly transient interacting proteins.

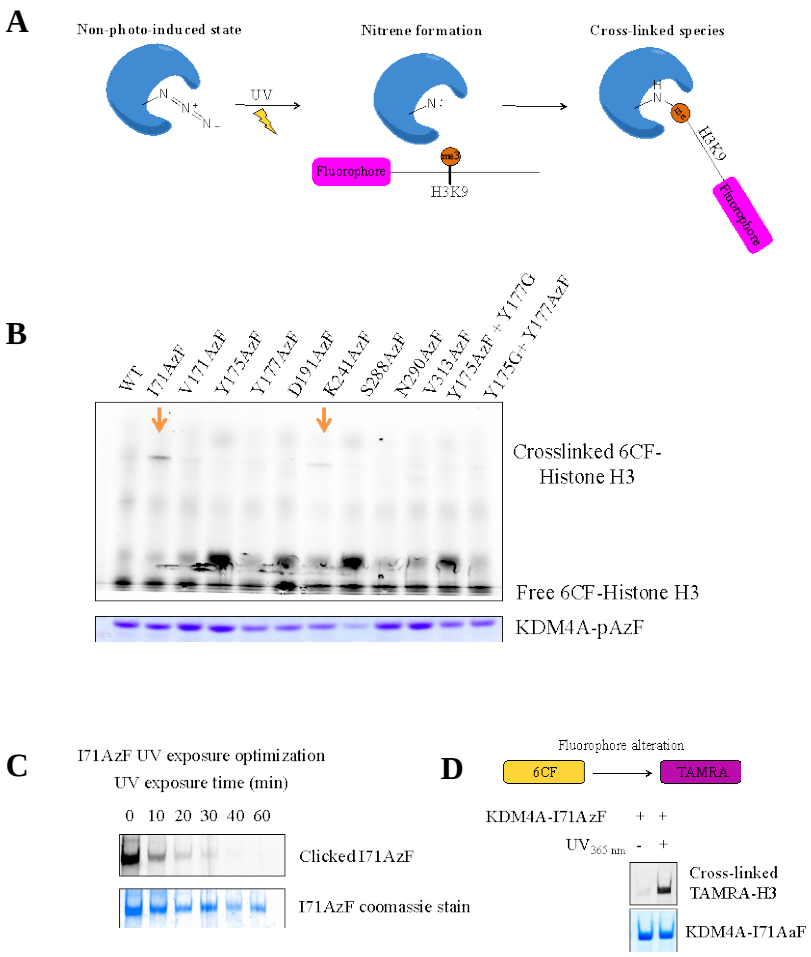


Figure 25: Screening KDM4A Mutants for Crosslinking Ability.

A) Photo-induced crosslinking conceptual cartoon. B) In-gel fluorescence of KDM4A-UAA crosslinked to 6CF-H3K9me3. Orange arrow indicates mutants with an evident crosslinking band. Coomassie stained SDS-PAGE shows KDM4A protein is present in each sample. C) In-gel fluorescence of copper catalyzed click assay for UV exposure time optimization. Coomassie stained SDS-PAGE of KDM4A-I71AzF below. D) In-gel fluorescence of KDM4A-I71AzF crosslinked to TAMRA-H3K9me3 with coomassie stained SDS-PAGE showing KDM4A-I71AzF in each sample below.

3.3.4 Photo-crosslinkable KDM4C and KDM4D

Because the KDM4 members are sequentially and structurally homologous (Figure 26A), the KDM4A mutants showing activity (enzymatic and/or crosslinking) were generated in KDM4B-D. These mutants were expressed and purified as stated for KDM4A and similar activity and crosslinking experiments were conducted. KDM4B does not express suggesting that the protein folding may be compromised with the incorporation of the unnatural amino acid at the I72 site or the catalytic domain vector needs to be optimized for increased protein yield. KDM4C-I73AzF is catalytically active towards known native substrates H3K9me3 and H3K36me3. KDM4D-L75AzF is catalytically active towards H3K9me3 and not towards H3K36me3, demonstrating substrate specificity similar to the native KDM4D. The demethylase activity of KDM4C-I73AzF and KDM4D-L75AzF was also tested towards non-native substrates, H3K4me3 and H3K27me3 (Figure 26B). Fortunately, the substrate specificity remains intact. KDM4C-D active mutants were then subjected to the click assay to validate azide competence. Both I73AzF and L75AzF showed intense signals prior to UV exposure and a loss of signal with 30 minutes of 365 nm light treatment confirming the preservation and photoinducibility of the azidophenylalanine residue (Figure 26C).

KDM4C and KDM4D AzF mutants were then screened for photo-induced crosslinking to TAMRA-H3K9me3 peptide. The conditions optimized for KDM4A-I71AzF were implemented when performing the crosslinking experiments for I73AzF and L75AzF. Both KDM4C-I73AzF and KDM4D-L75AzF provided intense crosslinking signals when exposed to 40 minutes of 365 nm UV light (Figure 26D). The biochemical investigation of the KDM4A,-C,

and -D photo-crosslinkable mutants has been performed using known KDM4 substrate Histone H3 peptide.

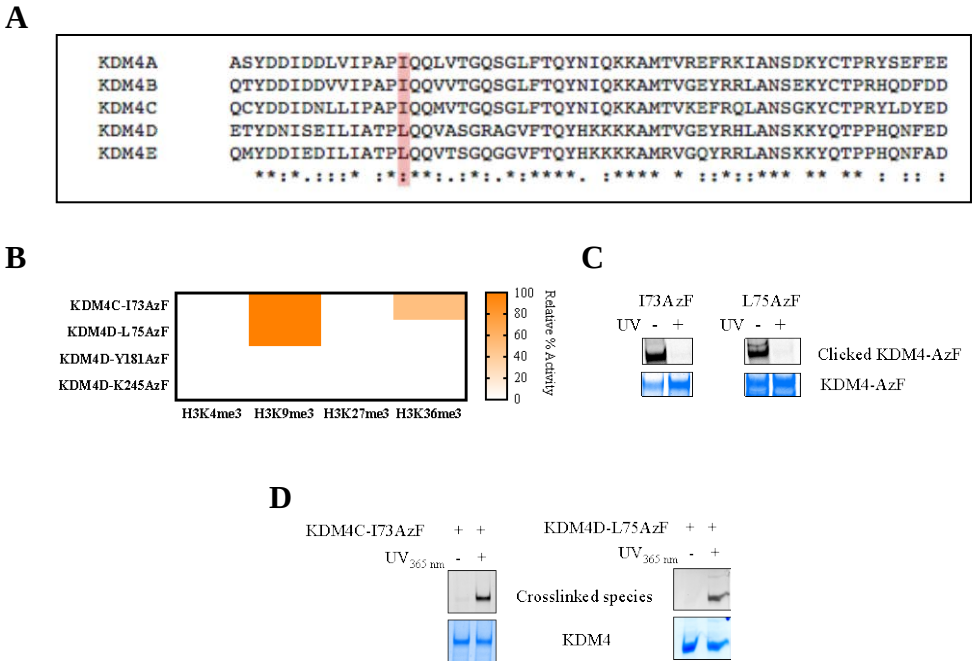


Figure 26: Generating Photo-Crosslinkable KDM4C and KDM4D.

A) Amino Acid sequence alignment showing the conservation of the corresponding isoleucing 71 (highlighted in red) in across KDM4 members. B) Heat map diagram showing the demethylase activity of the KDM4C and KDM4D UAA mutants towards H3K4me3, H3K9me3, H3K27me3, H3K36me3. C) Copper catalyzed click in-gel fluorescence image for KDM4C-I73AzF and KDM4D-L75AzF. Coomassie image below shows that the KDM4 protein is present in all samples. D) In-gel fluorescence of KDM4C-I73AzF and KDM4D-L75AzF crosslinked to TAMRA-H3K9me3. Coomassie stained enzyme below.

3.3.5 Crosslinking Full Length Histones

Because the native substrate is full length histones, HEK293T histones were extracted and used in the crosslinking assay. KDM4A-I71AzF successfully crosslinks to mammalian histone H3 (Figure 27). Ultimately, the goal is to identify KDM4 enzymatic substrates. An important consideration when developing this system, is that the KDM4-AzF mutants identified as photo-

inducible are catalytically active and will effectively reduce the substrate concentration, subsequently decreasing the probability of capturing non-histone proteins through this method. The affinity of non-methylated and mono-methylated H3K9 is dramatically reduced compared to H3K9me2/3 for KDM4s. This fact makes it essential for the crosslinking to occur before the enzyme can remove the methyl groups. To facilitate this strategic method, 2-KG competitive inhibitors – IOX1 and NOG, were incubated with KDM4A-I71AzF and then the enzyme was subjected to crosslinking towards HEK293T extracted histones (Figure 27A). Upon the addition of the inhibitors, the crosslinking signal increased providing additional insights and precautions for future live cell photo-crosslinking experiments. Crosslinking to the histones extracted from HEK293T cells shows that I71AzF can identify and crosslink to full length histone H3 in a mixture, but the experiment did not resolve which methylated species it crosslinked to. The activity assays previously performed show that the enzymatic specificity is conserved in the mutants, but this does not necessarily mean that binding is not occurring and if the non-enzymatic histone peptide substrate is bound within the catalytic domain, there is a chance of crosslinking. To determine if the KDM4A,-C, and –D AzF mutants crosslink non-specifically, histone H3 methyl lysine analogues (H3 MLAs) were generated and screened for crosslinking. H3 MLAs are histone derivatives where cysteine residues replace specific amino acid residues of choice allowing for alkylation chemistry to be performed so that modifications of interest can be site specifically placed. To generate H3 MLAs, full length histone H3 was mutated at K4, K9, K27, or K36 to incorporate a cysteine residue (K4C, K9C, K27C, K36C). Following mutagenesis, the histones were bacterially expressed, purified, modified to mimic a trimethylated lysine residue through established alkylation chemistry, and verified via LCMS (Figure 27B). These histones were subsequently utilized for photo-crosslinking experiments

with KDM4A-I71AzF, KDM4C-I73AzF, and KDM4D-L75AzF. All mutants crosslinked to H3K9Cme3 and H3K36Cme3 following UV exposure and did not crosslink H3K4Cme3. Interestingly, All mutants crosslinking H3K27Cme3, which has recently been shown to bind KDM4s and potentially be an enzymatic substrate of this protein family, but that has yet to be validated (Figure 27C). This experiment does show that there may be some level of promiscuity for the crosslinking events, suggesting that there may be potential false positives in the non-histone substrate investigation. Another reason for the nonspecific crosslinking occurring in this experiment could be that these are isolated systems where there is only one histone species available for the enzyme to bind to. Crosslinking may be more specific in full cell lysates or in live cells because the higher affinity proteins will bind and outcompete the non-enzymatic substrates with lower affinity for KDM4 proteins.

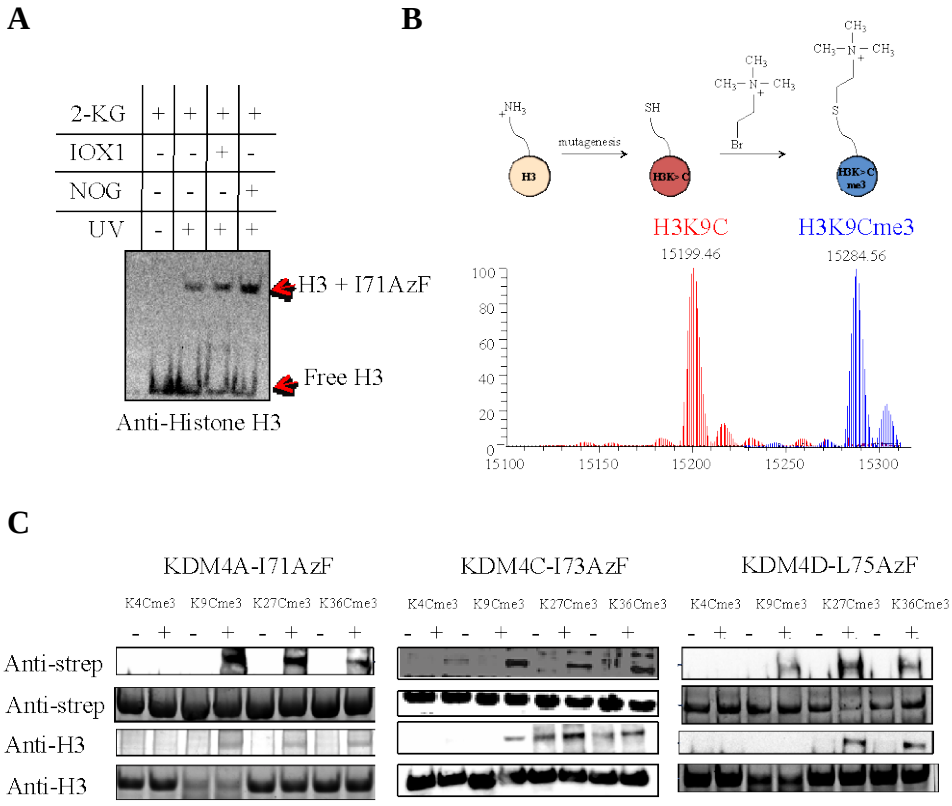


Figure 27: Crosslinking Full Length Histones.

A) Western blot of I71AzF crosslinked to histone H3 extracted from HEK293T cells using anti-Histone H3 antibody. Crosslinking experiments were performed under different conditions including with and without inhibitor. The bottom red arrow indicated free-non-crosslinked histone H3. The top red arrow is designating where the crosslinked species is shown. B) Top - cartoon depicting the generation of Histone methyl lysine analogues. Bottom – LCMS spectrum showing the conversion from H3K9C to H3K9Cme3. C) Western blot of KDM4A-I71AzF, KDM4C-I73AzF, and KDM4D-L75AzF crosslinked to Histone H3 MLAs – K4Cme3, K9Cme3, K27Cme3 and K36Cme3. Western blotting was performed with anti-strep tag antibody to detect the N-terminal strep tag (WSHPQFEK) located on the enzyme and with anti-H3 antibody to detect high molecular weight bands for H3. Each sample had a negative control where the sample was not exposed to UV light.

3.3.6 Crosslinking in Mammalian Cell lysates

With robust data supporting photo-induced crosslinking for three KDM4 members, the system was tested in HEK293T whole cell lysates. HEK293T cells were hypermethylated on KDM4 substrates by treatment with n-octyl-IOX1 to increase the abundance of substrates to be captured. Whole cell lysates were incubated with KDM4-AzF catalytic domain which contains an N-terminal unique peptide tag known as the Strep tag (WSHPQFEK). The strep tag allows for high affinity pull down using strep-tactin resin. This system has been shown to dramatically increase protein capture yield and reduce non-specific binding compared to other affinity pull down systems. Following exposure to 365 nm UV light to induce crosslinking, the KDM4-AzF free protein and crosslinked complexes were pulled down using streptactin resin and analyzed via western blotting and coomassie staining (Figure 28A). To confirm that the crosslinking is actually taking place, immunoblotting was performed with Histone H3 antibody (Anti-H3) because it is a known substrate for KDM4 proteins. KDM4A-I71AzF and KDM4D-L75AzF samples showed high molecular weight bands using the anti-H3 antibody in samples exposed to UV light, but not in samples that did not have UV light treatment, providing evidence that crosslinking occurred (Figure 28B). The samples were immunoblotted with Strep-tag antibodies (Anti-strep) to observe any signals above the molecular weight of the KDM4-AzF protein (>44 kDa), which would suggest crosslinking to several proteins. Several unique signals were observed in the UV light exposed sample for KDM4A-I71AzF compared to the corresponding sample that was not exposed to 365 nm light. KDM4D-L75AzF samples do show unique signals in the UV exposed lysates, but there are not as many high molecular weight bands compared to that in the KDM4A-I75AzF samples (Figure 28B middle). Following the detection of crosslinked band via western blotting, the samples were analyzed by coomassie staining.

Both KDM4A-I71AzF and KDM4D-L75AzF stained gels recapitulated the western blot analysis. With further analysis of the stained gel images, additional high molecular weight bands become apparent in each of the mutant samples exposed to UV light (Figure 28B right). Although there is apparent evidence of non-histone substrate crosslinking for KDM4A-I71AzF, it is less obvious for the KDM4D-L75AzF sample. This may be due to several reasons. KDM4D may have fewer non-histone protein substrates in general, the substrates for KDM4D may not be as abundant in these lysates or may be highly context dependent and require specific scaffolding proteins, and/or the crosslinking conditions need to be optimized in an enzyme unique fashion. For instance the strep-tag may become inaccessible for pull down or detection when KDM4D is crosslinked to certain protein substrates. Likely, all of these possibilities play a role in the reduced number of signals observed.

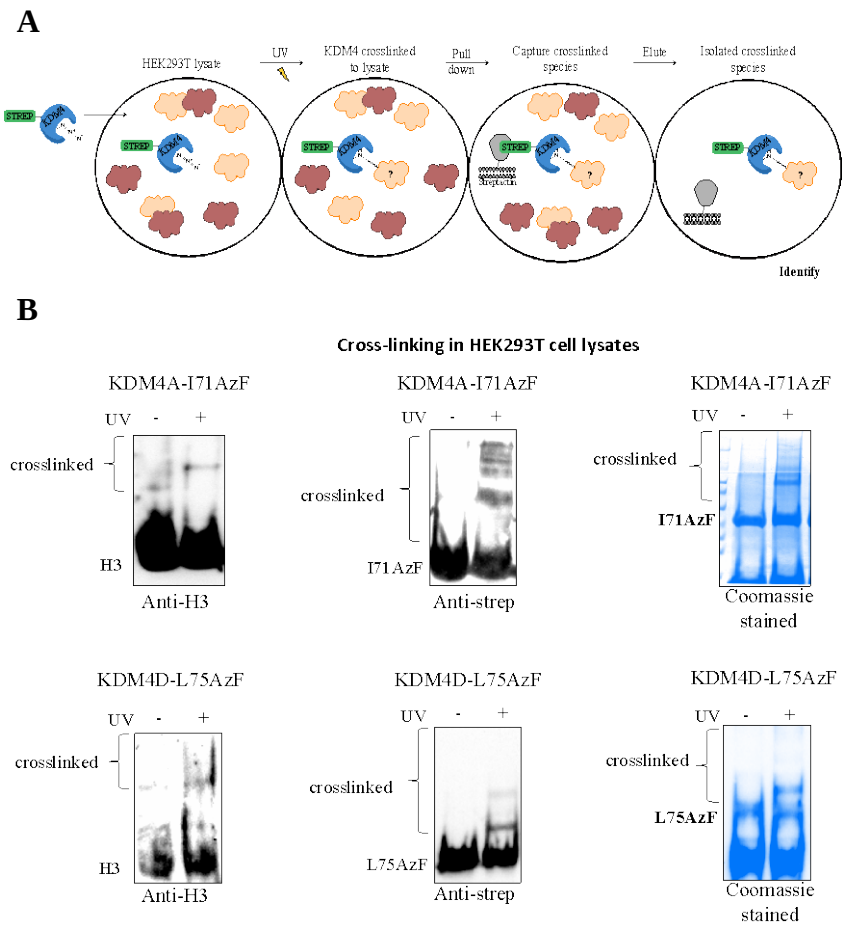


Figure 28: Crosslinking Mammalian Cell Lysates.

A) Cartoon depicting photo-induced crosslinking in HEK293T cell lysates. B) Western blots showing crosslinked bands in HEK293T lysates for KDM4A-I71AzF (top) and KDM4D-L75AzF (bottom) using anti-H3 antibody (left) and anti-strep tag antibody (middle). Coomassie stained SDS-PAGE for crosslinking detection (right) for KDM4A-I71AzF (top) and KDM4D-L75 (bottom).

4.0 CONCLUSIONS

There are over 30 JmjC domain-containing enzymes that are involved in regulating cellular function and identity through their post-translational modification functions. All JmjC domain-containing enzymes utilize metabolite 2-KG to demethylate lysine residues from proteins. This cofactor redundancy makes it difficult to selectively activate or inhibit these enzymes independently with small molecules. Because there are so many members with the same mechanism, functional redundancy prohibits unbiased and clear functional elucidation through classical genetic techniques. The number of proteins that get methylated on their lysine residues is massive and it is well understood that lysine methylation is a transient chemical mark regulated by KMTs and KDMs. Nucleosomal histone proteins are of great importance in chromatin organization and are methylated to induce changes in chromatin state and therefore gene expression. It is critical to understand each individual KDMs catalytic role in healthy and diseased cells for improved cancer therapeutic intervention, but the current tools are not fully competent to achieve this.

This body of work is the first to provide an orthogonal demethylase activation system for KDM4. Additionally, a gatekeeper residue may have been identified, providing a framework to develop this system for additional JmjC domain-containing enzymes. Future directions will include analytical assessment of cell permeability of the 2-KG analogues, observing gene expression changes upon the implementation of the engineered pairs for KDM4A-D in both

healthy and diseased cells, and generating similar systems for additional KDMs to compare gene expression regulatory roles.

Another approach to elucidating the roles of each KDM is through inhibition. An orthogonal KDM4-NOG system has successfully been developed with high potency *in vitro*. This engineered system will act as a highly complementary tool to the bump hole activation system generated by providing an 'ON and OFF' switch. This tool needs to be further developed and generalized among additional KDMs. Cell permeable derivatives of the inhibitors need to be designed and tested. In addition, this system needs to be tested in live mammalian cells. To maximize the potential of this system, the inhibitors could be photo-caged to increase temporal and regional control. Nonetheless, the bump-hole activation and inhibition show great promise for future experiments to investigate the independent roles of KDMs.

KDMs likely have non-histone protein substrates, but none have been identified in live cells. Several ChIP studies provide data on localization and potential interacting partners, but this method does not indicate if the associated protein partner is necessarily an enzyme substrate. To improve this resolution to get information on which proteins are catalytic targets of KDM4s, the incorporation of an unnatural amino acid with covalent crosslinking capabilities is an attractive approach. Here, photo-inducible crosslinking KDM4A, KDM4C, and KDM4D mutants have been generated, biochemically characterized, screened for crosslinking and tested in physiologically relevant environments. The data generated here suggest that several high molecular weight proteins crosslink to KDM4s following UV exposure, encouraging the pursuit of proteomics analysis to identify these proteins. Future experiments with this system include implementing this in live cells and organisms to elucidate substrates in various cellular environments and during different stages of development. The engineered system can also be

improved by utilizing a different unnatural amino acid for increased crosslinking efficiency and higher stability. Additionally, the crosslinking system for KDM4B needs to be optimized further. KDM4B-UAA mutants did not successfully express which may be due to protein instability or poor vector choice. Lastly, crosslinking to nucleosomal histones and full nucleosome pull down should be pursued to identify genomic catalytic sites.

In summary, three chemical biology tools have been developed for the KDM4 subfamily and the potential for expansion and application to additional KDMs is high. This will allow for the roles of epigenetic modifiers to be identified and compared to one another in various environments. These tools have potentially high public health impact as there are an abundance of KDMs implicated in diverse pathological phenotypes where their mechanistic nature and cellular pathway involvement has not been fully unmasked. The three tools developed here may aid in the strategic design of therapeutic approaches for neurological and cardiac diseases as well as numerous cancers. Additionally, these tools may be able to identify diagnostic and prognostic biomarkers for various disorders depending on the results of each protein.

APPENDIX: ABBREVIATIONS

2-HG: 2-Hydroxyglutarate

2-KG: (2)-Ketoglutarate (also referred to as 2-OG = 2-oxoglutarate)

5CaC: 5-Carboxycytosine

5fC: 5-formylcytosine

5hmC: 5-hemimethylcytosine

5mC: 5-methylcytosine

6CF: 6-Carboxyfluorescein

ANK: Ankyrin repeat domain

AP1: Activating Protein 1

AR: Androgen receptor

BpF: Benzophenone

CDK: Cyclin dependent kinase

ChIP: Chromatin Immunoprecipitation

CpG: Cytosine-phosphate-Guanine base sequence

DNMT: DNA methyltransferase

DTT: dithiothreitol

ER: Estrogen receptor

ERG: erythroblast transformation-specific related gene

ESC: Embryonic stem cells

EuHMTase1: Euchromatin Histone methyltransferase 1 protein

FDH: 10-formyltetrahydrofolate dehydrogenase

FPLC: Fast protein liquid chromatography

GLP: G9a like protein

H2: Histone H2

H3: Histone H3

H4: Histone H4

HAT: Histone acetyltransferases

HDAC: Histone deacetylase

HEK293T: Human Embryonic Kidney Cells – 293 Transfection

HP1: heterochromatin protein 1

HPLC: High performance liquid chromatography

IDH: Isocitrate dehydrogenase

IOX1: 5-carboxy-8-hydroxyquinoline

IP: Immunoprecipitation

IPA: Ingenuity Pathway Analysis

JMJC: Jumonji C domain

JMJN: Jumonji N domain

K4, K9, K27, K36: Lysine 4, Lysine 9, Lysine 27, Lysine 36

kDa: kilodalton

KDM: Lysine demethylase

KMT: Lysine methyltransferase

LSD: flavin-dependent lysine specific demethylases

m^{6}A: N_6-methyladenosine

MBD2: Methylated DNA binding domain-containing protein 62

METTL3: methyltransferase like 3 domain

MMR: Mismatch repair

MT: methyltransferase

NADH: Nicotinamide adenine dinucleotide

N-CoR: nuclear receptor corepressor

NGS: Next Generation Sequencing

NOG: N-oxalylglycine

PAD: peptidylarginine deiminases

pAzF or AzF: Azidophenylalanine

PCNA: proliferating cell nuclear antigen

PDK: Pyruvate dehydrogenase kinases

PKMT: Protein lysine methyltransferases

PMT: Protein methyltransferases

PRMT: Protein arginine methyltransferases

PRNP: Prion protein

PTM: Post translational modification

qPCR: Quantitative polymerase chain reaction

RNF: Ring finger protein

SAM: S-adenosylmethionine- binding domain

SCF: Skp, Cullin, F-box containing complex

SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis

SET: Su (var), enhancer of zeste, trithorax domain

TAMRA: Tetramethylrhodamine

TET: human ten-eleven-translocation domain

TRAIL: Tumor Necrosis Factor Related Apoptosis Inducing Ligand

UAA: Unnatural amino acid

YAP1: Yes-associated protein 1

BIBLIOGRAPHY

1. Khoury, George A., Richard C. Baliban, and Christodoulos A. Floudas. "Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database." *Scientific reports* 1 (2011): 90.

2. Li, Shuyan, et al. "Loss of post-translational modification sites in disease." *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*. NIH Public Access, 2010.

3. Gioeli, Daniel, and Bryce M. Paschal. "Post-translational modification of the androgen receptor." *Molecular and cellular endocrinology* 352.1 (2012): 70-78.

4. Ko, Soyoung, et al. "Lysine methylation and functional modulation of androgen receptor by Set9 methyltransferase." *Molecular endocrinology* 25.3 (2011): 433-444.

5. Thomas, Monzy, et al. "Androgen receptor acetylation site mutations cause trafficking defects, misfolding, and aggregation similar to expanded glutamine tracts." *Journal of Biological Chemistry* 279.9 (2004): 8389-8395.

6. Fu, Maofu, et al. "Acetylation of androgen receptor enhances coactivator binding and promotes prostate cancer cell growth." *Molecular and cellular biology* 23.23 (2003): 8563-8575.

7. Helenius, Ari, and Markus Aebi. "Intracellular functions of N-linked glycans." *Science* 291.5512 (2001): 2364-2369.

8. Stimson, Elaine, et al. "Site-specific characterization of the N-linked glycans of murine prion protein by high-performance liquid chromatography/electrospray mass spectrometry and exoglycosidase digestions." *Biochemistry* 38.15 (1999): 4885-4895.

9. Mitra, Nivedita, et al. "N-linked oligosaccharides as outfitters for glycoprotein folding, form and function." *Trends in biochemical sciences* 31.3 (2006): 156-163.

10. Rudd, Pauline M., et al. "Glycosylation differences between the normal and pathogenic prion protein isoforms." *Proceedings of the National Academy of Sciences* 96.23 (1999): 13044-13049.

11. Nestler, E. J., and P. Greengard. "Protein phosphorylation is of fundamental importance in biological regulation." *Basic Neurochemistry: Molecular, Cellular and Medical Aspects, 6th edition. Lippincott-Raven, Philadelphia, PA* (1999).

12. Van den Heuvel, Sander, and Ed Harlow. “Distinct Roles for Cyclin-Dependent Kinases in Cell Cycle Control.” Science, vol. 262, no. 5142, 1993, pp. 2050–2054. JSTOR, JSTOR, www.jstor.org/stable/2882984.

13. Finn, Richard S., Alexey Aleshin, and Dennis J. Slamon. "Targeting the cyclin-dependent kinases (CDK) 4/6 in estrogen receptor-positive breast cancers." *Breast Cancer Research* 18.1 (2016): 17.

14. Sawicki, Mark P., et al. "Human genome project." *The American journal of surgery* 165.2 (1993): 258-264.

15. Luger, Karolin, et al. "Crystal structure of the nucleosome core particle at 2.8 Å resolution." Nature 389.6648 (1997): 251-260.

16. Richmond, Timothy J., and Curt A. Davey. "The structure of DNA in the nucleosome core.” *Nature* 423.6936 (2003): 145-150.

17. Jaenisch, Rudolf, and Adrian Bird. "Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals." *Nature genetics* 33 (2003): 245-254.

18. Bannister, Andrew J., and Tony Kouzarides. "Regulation of chromatin by histone modifications." *Cell research* 21.3 (2011): 381-395.

19. Vaissière, Thomas, Carla Sawan, and Zdenko Herceg. "Epigenetic interplay between histone modifications and DNA methylation in gene silencing." *Mutation Research/Reviews in Mutation Research* 659.1 (2008): 40-48.

20. Egger, Gerda, et al. "Epigenetics in human disease and prospects for epigenetic therapy." *Nature* 429.6990 (2004): 457-463.

21. Medina-Franco, José L., et al. "Discovery and development of DNA methyltransferase inhibitors using in silico approaches." *Drug Discovery Today* 20.5 (2015): 569-577.

22. Robert, Marie-France, et al. "DNMT1 is required to maintain CpG methylation and aberrant gene silencing in human cancer cells." *Nature genetics* 33.1 (2003): 61-65.

23. Fuks, François, et al. "DNA methyltransferase Dnmt1 associates with histone deacetylase activity." *Nature genetics* 24.1 (2000): 88-91.

24. Bird, Adrian P. "CpG-rich islands and the function of DNA methylation." *Nature* 321.6067 (1986): 209-213.

25. Razin, Aharon, and Arthur D. Riggs. "DNA methylation and gene function." *Science* 210.4470 (1980): 604-610.

26. Okano, Masaki, et al. "DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development." *Cell* 99.3 (1999): 247-257.

27. Pfaffeneder, Toni, et al. "The discovery of 5-formylcytosine in embryonic stem cell DNA." *Angewandte Chemie* 123.31 (2011): 7146-7150.

28. Ito, Shinsuke, et al. "Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine." *Science* 333.6047 (2011): 1300-1303.

29. He, Housheng Hansen, et al. "Nucleosome dynamics define transcriptional enhancers." *Nature genetics* 42.4 (2010): 343-347.

30. Fleming, Alastair B., et al. "H2B ubiquitylation plays a role in nucleosome dynamics during transcription elongation." *Molecular cell* 31.1 (2008): 57-66.

31. Ahmad, Kami, and Steven Henikoff. "Epigenetic consequences of nucleosome dynamics." *Cell* 111.3 (2002): 281-284.

32. Morales, Violette, and Hélène Richard-Foy. "Role of histone N-terminal tails and their acetylation in nucleosome dynamics." *Molecular and cellular biology* 20.19 (2000): 7230-7237.

33. Zentner, Gabriel E., and Steven Henikoff. "Regulation of nucleosome dynamics by histone modifications." *Nature structural & molecular biology* 20.3 (2013): 259-266.

34. Yang, X. J., and E. Y. Seto. "HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention." (2007): 5310-5318.

35. Wang, Zhibin, et al. "Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes." *Cell* 138.5 (2009): 1019-1031.

36. Brown, Christine E., et al. "The many HATs of transcription coactivators." *Trends in biochemical sciences* 25.1 (2000): 15-19.

37. Hassig, Christian A., and Stuart L. Schreiber. "Nuclear histone acetylases and deacetylases and transcriptional regulation: HATs off to HDACs." *Current opinion in chemical biology* 1.3 (1997): 300-308.

38. Wang, Yanming, et al. "Human PAD4 regulates histone arginine methylation levels via demethylimination." *Science* 306.5694 (2004): 279-283.

39. Cuthbert, Graeme L., et al. "Histone deimination antagonizes arginine methylation." *Cell* 118.5 (2004): 545-553.

40. Strahl, Brian D., et al. "Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1." *Current Biology* 11.12 (2001): 996-1000.

41. Torres-Padilla, Maria-Elena, et al. "Histone arginine methylation regulates pluripotency in the early mouse embryo." *Nature*445.7124 (2007): 214-218.

42. Martin, Cyrus, and Yi Zhang. "The diverse functions of histone lysine methylation." *Nature reviews Molecular cell biology* 6.11 (2005): 838-849.

43. Guccione, Ernesto, et al. "Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive." *Nature*449.7164 (2007): 933-937.

44. Shilatifard, Ali. "Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation." *Current opinion in cell biology* 20.3 (2008): 341-348.

45. Sierra, Jose, et al. "The APC tumor suppressor counteracts β-catenin activation and H3K4 methylation at Wnt target genes." *Genes & development* 20.5 (2006): 586-600.

46. Nightingale, Karl P., et al. "Cross-talk between histone modifications in response to histone deacetylase inhibitors MLL4 links histone H3 acetylation and histone H3K4 methylation." *Journal of Biological Chemistry* 282.7 (2007): 4408-4416.

47. Tachibana, Makoto, et al. "G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription." *The EMBO journal* 27.20 (2008): 2681-2690.

48. Snowden, Andrew W., et al. "Gene-specific targeting of H3K9 methylation is sufficient for initiating repression in vivo." *Current biology* 12.24 (2002): 2159-2166.

49. Tachibana, Makoto, et al. "Functional dynamics of H3K9 methylation during meiotic prophase progression." *The EMBO journal* 26.14 (2007): 3346-3359.

50. Wu, Li-Peng, et al. "Histone deacetylase inhibitor depsipeptide activates silenced genes through decreasing both CpG and H3K9 methylation on the promoter." *Molecular and cellular biology* 28.10 (2008): 3219-3235.

51. Yuan, Wen, et al. "H3K36 methylation antagonizes PRC2-mediated H3K27 methylation." *Journal of Biological Chemistry*286.10 (2011): 7983-7989.

52. Lee, Min Gyu, et al. "Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination." *Science* 318.5849 (2007): 447-450.

53. Wu, Hui, et al. "Histone methyltransferase G9a contributes to H3K27 methylation in vivo." *Cell research* 21.2 (2011): 365.

54. Yoo, Kyung Hyun, and Lothar Hennighausen. "EZH2 methyltransferase and H3K27 methylation in breast cancer." *International journal of biological sciences* 8.1 (2012): 59.

55. Lee, Jung-Shin, and Ali Shilatifard. "A site to remember: H3K36 methylation a mark for histone deacetylation." *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 618.1 (2007): 130-134.

56. Klose, Robert J., et al. "Demethylation of histone H3K36 and H3K9 by Rph1: a vestige of an H3K9 methylation system in Saccharomyces cerevisiae?." *Molecular and cellular biology* 27.11 (2007): 3951-3961.

57. Klose, Robert J., et al. "The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36." *Nature*442.7100 (2006): 312-316.

58. Sen, Payel, et al. "H3K36 methylation promotes longevity by enhancing transcriptional fidelity." *Genes & Development* 29.13 (2015): 1362-1376.

59. Lu, Xu, et al. "The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure." *Nature structural & molecular biology* 15.10 (2008): 1122-1124.

60. Yu, Yongxin, et al. "Histone H3 lysine 56 methylation regulates DNA replication through its interaction with PCNA." *Molecular cell*46.1 (2012): 7-17.

61. Chen, Dagang, et al."Regulation of transcription by a protein methyltransferase." *Science* 284.5423 (1999): 2174-2177.

62. Trievel, Raymond C., et al. "Structure and catalytic mechanism of a SET domain protein methyltransferase." *Cell* 111.1 (2002): 91-103.

63. Jelinic, Petar, Jean-Christophe Stehle, and Phillip Shaw. "The testis-specific factor CTCFL cooperates with the protein methyltransferase PRMT7 in H19 imprinting control region methylation." *PLoS biology* 4.11 (2006): e355.

64. Dillon, Shane C., et al. "The SET-domain protein superfamily: protein lysine methyltransferases." *Genome biology* 6.8 (2005): 227.

65. Huang, L., et al. "Correlation of chromosomes 1p and 19q status and expressions of O6-methylguanine DNA methyltransferase (MGMT), p53 and Ki-67 in diffuse gliomas of World Health Organization (WHO) grades II and III: a clinicopathological study." *Neuropathology and applied neurobiology* 35.4 (2009): 367-379.

66. Liu, Feng, et al. "Protein lysine methyltransferase G9a inhibitors: design, synthesis, and structure activity relationships of 2, 4-diamino-7-aminoalkoxy-quinazolines." *Journal of medicinal chemistry* 53.15 (2010): 5844-5857.

67. Tang, Jie, et al. "PRMT 3, a type I protein arginine N-methyltransferase that differs from PRMT1 in its oligomerization, subcellular localization, substrate specificity, and regulation." *Journal of Biological Chemistry* 273.27 (1998): 16935-16945.

68. Frankel, Adam, et al. "The novel human protein arginine N-methyltransferase PRMT6 is a nuclear enzyme displaying unique substrate specificity." *Journal of Biological Chemistry* 277.5 (2002): 3537-3543.

69. Yevstafiev, Dmytro. *Deciphering the mechanism of G9a spreading genome-wide*. Diss. Université d'Ottawa/University of Ottawa, 2015.

70. Pal, Sharmistha, et al. "Human SWI/SNF-associated PRMT5 methylates histone H3 arginine 8 and negatively regulates expression of ST7 and NM23 tumor suppressor genes." *Molecular and cellular biology* 24.21 (2004): 9630-9645.

71. Patnaik, Debasis, et al. "Substrate specificity and kinetic mechanism of mammalian G9a histone H3 methyltransferase." *Journal of Biological Chemistry* 279.51 (2004): 53248-53258.

72. Collins, Robert E., et al. "The ankyrin repeats of G9a and GLP histone methyltransferases are mono-and dimethyllysine binding modules." *Nature structural & molecular biology* 15.3 (2008): 245-250.

73. Chin, Hang Gyeong, et al. "Automethylation of G9a and its implication in wider substrate specificity and HP1 binding." *Nucleic acids research* 35.21 (2007): 7313-7323.

74. Chuikov, Sergei, et al. "Regulation of p53 activity through lysine methylation." *Nature* 432.7015 (2004): 353-360.

75. Rathert, Philipp, et al. "Protein lysine methyltransferase G9a acts on non-histone targets." *Nature chemical biology* 4.6 (2008): 344-346.

76. Estève, Pierre-Olivier, et al. "Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication." Genes & development 20.22 (2006): 3089-3103.

77. Wilkinson, Lawrence S., William Davies, and Anthony R. Isles. "Genomic imprinting effects on brain development and function." Nature Reviews Neuroscience 8.11 (2007): 832-843.

78. Huang, Jing, et al. "G9a and Glp methylate lysine 373 in the tumor suppressor p53." Journal of Biological Chemistry 285.13 (2010): 9636-9641.

79. Bharathy, N., et al., G9a, a multipotent regulator of gene expression. Epigenetics, 2013. 8(1): p. 16-22. 13.

80. Lehnertz B, Northrop JP, Antignano F, Burrows K, Hadidi S, Mullaly SC, et al., Activating and inhibitory functions for the histone lysine methyltransferase G9a in T helper cell differentiation and function. J Exp Med, 2010. 207: p. 915–22.

81. Tachibana, Makoto, et al. "G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis." *Genes & development* 16.14 (2002): 1779-1791.

82. Velasco, Guillaume, et al. "Dnmt3b recruitment through E2F6 transcriptional repressor mediates germ-line gene silencing in murine somatic tissues." Proceedings of the National Academy of Sciences 107.20 (2010): 9281-9286.

83. Ogawa, Hidesato, et al. "A complex with chromatin modifiers that occupies E2F-and Myc-responsive genes in G0 cells." Science 296.5570 (2002): 1132-1136.

84. Tachibana, Makoto, et al. "Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9." Genes & development 19.7 (2005): 815-826.

85. Heard, Edith, et al. "Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation." Cell 107.6 (2001): 727-738.

86. Stevenson, Brian J., et al. "Rapid evolution of cancer/testis genes on the X chromosome." Bmc Genomics 8.1 (2007): 129.

87. Dong, Kevin B., et al. "DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity." The EMBO journal 27.20 (2008): 2691-2701.

88. Collins, Robert E., et al. "The ankyrin repeats of G9a and GLP histone methyltransferases are mono-and dimethyllysine binding modules." Nature structural & molecular biology 15.3 (2008): 245-250.

89. Epsztejn-Litman, Silvina, et al. "De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes." Nature structural & molecular biology 15.11 (2008): 1176-1183.

90. Smallwood, Andrea, et al. "Functional cooperation between HP1 and DNMT1 mediates gene silencing." Genes & development 21.10 (2007): 1169-1178.

91. Culhane, J. and Cole, P., LSD1 and the Chemistry of Histone Demethylation. Current Opinion in Chemical Biology, 2007. 11(5): p. 561-568.

92. Labbe, R.M., Holowatyj, A., Yang, Z., Histone lysine demethylase (KDM) subfamily 4: structures, functions and therapeutic potential. American Journal of Translational Research, 2014. 6(1): p. 1-15

93. Klose, R.J., Kallin, E.M., and Zhang, Y., JmjC-domain containing proteins and histone demethylation. Nature Reviews Genetics, 2006. 7: p. 715-727.

94. Berry, W. and Janknecht, R., KDM4/JMJD2 Histone Demethylases: Epigenetic Regulators in Cancer Cells. The Journal of Cancer Research, 2013. 73: p. 2936.

95. Hillringhaus, L., et al., Structural and Evolutionary Basis for the Dual Substrate Selectivity of Human KDM4 Histone Demethylase Family.The Journal of Biological Chemistry, 2011. 286: p. 41616–41625.

96. Iwamori N, Zhao M, Meistrich ML, Matzuk MM. The testis-enriched histone demethylase, KDM4D, regulates methylation of histone H3 lysine 9 during spermatogenesis in the mouse but is dispensable for fertility. Biol Reprod. 2011;84:1225–1234.

97. Krupp M. RNA-Seq Atlas--a reference database for gene expression profiling in normal tissue by next-generation sequencing. Bioinformatics. 2012;28:1184–1185.

98. Shin S, Janknecht R. Diversity within the JMJD2 histone demethylase family. Biochem Biophys Res Commun. 2007;353:973–977.

99. Tsurumi A. Drosophila Kdm4 demethylases in histone H3 lysine 9 demethylation and ecdysteroid signaling. Sci Rep. 2013;3:2894.

100. Black JC, Allen A, Van Rechem C, Forbes E, Longworth M, Tschöp K, Rinehart C, Quiton J, Walsh R, Smallwood A, Dyson NJ, Whetstine JR. Conserved antagonism between JMJD2A/KDM4A and HP1γ during cell cycle progression. Mol Cell. 2010;40:736–748.

101. Skarnes WC. A conditional knockout resource for the genome-wide study of mouse gene function. Nature. 2011;474:337–342.

102. Kawazu M, Saso K, Tong KI, McQuire T, Goto K, Son DO, Wakeham A, Miyagishi M, Mak TW, Okada H. Histone demethylase JMJD2B functions as a co-factor of estrogen receptor in breast cancer proliferation and mammary gland development. PLoS One. 2011;6:e17830.

103. Eppig JT. The Mouse Genome Database (MGD): comprehensive resource for genetics and genomics of the laboratory mouse. Nucleic Acids Res. 2012;40:D881–886.

104. Fodor, Barna D., et al. "Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells." *Genes & development* 20.12 (2006): 1557-1562.

105. Black, Joshua C., et al. "KDM4A lysine demethylase induces site-specific copy gain and rereplication of regions amplified in tumors." *Cell* 154.3 (2013): 541-555.

106. Kupershmit, Ilana, et al. "KDM4C (GASC1) lysine demethylase is associated with mitotic chromatin and regulates chromosome segregation during mitosis." *Nucleic acids research* 42.10 (2014): 6168-6182.

107. Lohse, Brian, et al. "Posttranslational modifications of the histone 3 tail and their impact on the activity of histone lysine demethylases in vitro." *PLoS One* 8.7 (2013): e67653.

108. Pedersen, Marianne Terndrup, et al. "The demethylase JMJD2C localizes to H3K4me3-positive transcription start sites and is dispensable for embryonic development." *Molecular and cellular biology* 34.6 (2014): 1031-1045.

109. Kim, Tae-Dong, et al. "Histone demethylase JMJD2A drives prostate tumorigenesis through transcription factor ETV1." *The Journal of clinical investigation* 126.2 (2016): 706.

110. Wang, Ling-Yu, et al. "KDM4A coactivates E2F1 to regulate the PDK-dependent metabolic switch between mitochondrial oxidation and glycolysis." *Cell reports* 16.11 (2016): 3016-3027.

111. Shin, Sook, and Ralf Janknecht. "Activation of androgen receptor by histone demethylases JMJD2A and JMJD2D." *Biochemical and biophysical research communications* 359.3 (2007): 742-746.

112. Gaughan, Luke, et al. "KDM4B is a master regulator of the estrogen receptor signalling cascade." *Nucleic acids research*41.14 (2013): 6892-6904.

113. Wong, Rebecca Lee Yean, and Cheryl Lyn Walker. "Molecular pathways: environmental estrogens activate nongenomic signaling to developmentally reprogram the epigenome." *Clinical Cancer Research* 19.14 (2013): 3732-3737.

114. Guo, L., et al. "Histone demethylase Kdm4b functions as a co-factor of C/EBPβ to promote mitotic clonal expansion during differentiation of 3T3-L1 preadipocytes." *Cell Death & Differentiation* 19.12 (2012): 1917-1927.

115. Ipenberg, Inbal, et al. "Heat shock protein 90 (Hsp90) selectively regulates the stability of KDM4B/JMJD2B histone demethylase." *Journal of Biological Chemistry* 288.21 (2013): 14681-14687.

116. Das, Partha Pratim, et al. "Distinct and combinatorial functions of Jmjd2b/Kdm4b and Jmjd2c/Kdm4c in mouse embryonic stem cell identity." *Molecular cell* 53.1 (2014): 32-48.

117. Torres, Idelisse Ortiz, and Danica Galonić Fujimori. "Functional coupling between writers, erasers and readers of histone and DNA methylation." *Current opinion in structural biology* 35 (2015): 68-75.

118. Ponnaluri, VK Chaithanya, et al. "Identification of non-histone substrates for JMJD2A–C histone demethylases." *Biochemical and biophysical research communications* 390.2 (2009): 280-284.

119. Kumar, Sunil, Prasad Venkata, and Yonggyun Kim. "Suppressive activity of a viral histone H4 against two host chromatin remodelling factors: lysine demethylase and SWI/SNF." *Journal of General Virology* 97.10 (2016): 2780-2796.

120. Curtis, Brenda J., et al. "Histone lysine demethylases function as co-repressors of SWI/SNF remodeling activities during Drosophila wing development." *Developmental biology* 350.2 (2011): 534-547.

121. Hayakawa, Tomohiro, and Jun-ichi Nakayama. "Physiological roles of class I HDAC complex and histone demethylase." *BioMed Research International* 2011 (2010).

122. Brauchle, Michael, et al. "Protein complex interactor analysis and differential activity of KDM3 subfamily members towards H3K9 methylation." *PloS one* 8.4 (2013): e60549.

123. Ogawa, Hidesato, et al. "A complex with chromatin modifiers that occupies E2F-and Myc-responsive genes in G0 cells." *Science*296.5570 (2002): 1132-1136.

124. Van Rechem, Capucine, et al. "Lysine demethylase KDM4A associates with translation machinery and regulates protein synthesis." *Cancer discovery* 5.3 (2015): 255-263.

125. Al-Qahtani, Khalid, et al. "The broad spectrum 2-oxoglutarate oxygenase inhibitor N-oxalylglycine is present in rhubarb and spinach leaves." *Phytochemistry* 117 (2015): 456-461.

126. Hamada, Shohei, et al. "Synthesis and activity of N-oxalylglycine and its derivatives as Jumonji C-domain-containing histone lysine demethylase inhibitors." *Bioorganic & medicinal chemistry letters*19.10 (2009): 2852-2855.

127. Linster, Carole L., Emile Van Schaftingen, and Andrew D. Hanson. "Metabolite damage and its repair or pre-emption." *Nature chemical biology* 9.2 (2013): 72-80.

128. Dang, Lenny, et al. "Cancer-associated IDH1 mutations produce 2-hydroxyglutarate." *Nature* 462.7274 (2009): 739-744.

129. Fantin, Valeria R., et al. "Cancer-associated IDH1 mutations produce 2-hydroxyglutarate." (2010): 33-33.

130. Intlekofer, Andrew M., et al. "Hypoxia induces production of L-2-hydroxyglutarate." *Cell metabolism* 22.2 (2015): 304-311.

131. Hopkinson, Richard J., et al. "5-Carboxy-8-hydroxyquinoline is a broad spectrum 2-oxoglutarate oxygenase inhibitor which causes iron translocation." *Chemical science* 4.8 (2013): 3110-3117.

132. Schiller, Rachel, et al. "A cell-permeable ester derivative of the JmjC histone demethylase inhibitor IOX1." *ChemMedChem* 9.3 (2014): 566-571.

133. Awwad, S., Ayoub, N., Overexpression of KDM4 lysine demethylases disrupts the integrity of the DNA mismatch repair pathway. Biology Open, 2015. 4(3).

134. Ding, X., Pan, H., Li, J., Zhong, Q., Chen, X., Dry, S.M., Wang, C.Y., Epigenetic activation of AP1 promotes squamous cell carcinoma metastasis. Sci Signal, 2013. 6:ra28, 1–13, S0–15.

135. Mallette, F.A., Richard, S., JMJD2A promotes cellular transformation by blocking cellular senescence through transcriptional repression of the tumor suppressor CHD5.Cell Rep, 2012. 2: p. 1233–1243.

136. Guerra-Calderas, L., Gonzalez-Barrios, R., Herrera, L., Cantu de Leon, D., Soto-Reyes, E., The role of the histone demethylase KDM4A in cancer. Cancer Genetics, 2014.

137. Liu, Y., An epigenetic role for PRL-3 as a regulator of H3K9 methylation in colorectal cancer. Gut, 2013. 62: p. 571–581.

138. Kawazu, M., Saso, K., Tong, K.I., McQuire, T., Goto, K., Son, D.O., Wakeham, A., Miyagishi, M., Mak, T.W., Okada, H., Histone demethylase JMJD2B functions as a co-factor of estrogen receptor in breast cancer proliferation and mammary gland development. PLoS One, 2011. 6:e17830.

139. Liu, G., Bollig-Fischer, A., Kreike, B., van de Vijver, M.J., Abrams, J., Ethier, S.P., Yang, Z.Q., Genomic amplification and oncogenic properties of the GASC1 histone demethylase gene in breast cancer. Oncogene, 2009. 28: p. 4491–4500.

140. Schmidt, Christine K., and Stephen P. Jackson. "On your mark, get SET (D2), go! H3K36me3 primes DNA mismatch repair." *Cell*153.3 (2013): 513-515.

141. Li, Feng, et al. "Regulation of mismatch repair by histone code and posttranslational modifications in eukaryotic cells." *DNA repair*38 (2016): 68-74.

142. Liu, Dekang, Guido Keijzers, and Lene Juel Rasmussen. "DNA mismatch repair and its many roles in eukaryotic cells." *Mutation Research/Reviews in Mutation Research* 773 (2017): 174-187.

143. Zhao, Rui, et al. "CHD5 a tumour suppressor is epigenetically silenced in hepatocellular carcinoma." *Liver International* 34.6 (2014).

144. Chu, Chia-Han, et al. "KDM4B as a target for prostate cancer: structural analysis and selective inhibition by a novel inhibitor." *Journal of medicinal chemistry* 57.14 (2014): 5975-5985.

145. Coffey, Kelly, et al. "The lysine demethylase, KDM4B, is a key molecule in androgen receptor signalling and turnover." *Nucleic acids research* 41.8 (2013): 4433-4446.

146. Yang, Jun, et al. "The role of histone demethylase KDM4B in Myc signaling in neuroblastoma." *JNCI: Journal of the National Cancer Institute* 107.6 (2015).

147. Qiu, Mei-Ting, et al. "KDM4B and KDM4A promote endometrial cancer progression by regulating androgen receptor, c-myc, and p27kip1." *Oncotarget* 6.31 (2015): 31702.

148. Yamamoto, Shinzo, et al. "Histone demethylase KDM4C regulates sphere formation by mediating the cross talk between Wnt and Notch pathways in colonic cancer cells." *Carcinogenesis* 34.10 (2013): 2380-2388.

149. Kim, Tae-Dong, et al. "Pro-growth role of the JMJD2C histone demethylase in HCT-116 colon cancer cells and identification of curcuminoids as JMJD2 inhibitors." *American journal of translational research* 6.3 (2014): 236.

150. Cheung, Ngai, et al. "Targeting aberrant epigenetic networks mediated by PRMT1 and KDM4C in acute myeloid leukemia." *Cancer Cell* 29.1 (2016): 32-48.

151. Liu, Gang, et al. "Genomic amplification and oncogenic properties of the GASC1 histone demethylase gene in breast cancer." *Oncogene* 28.50 (2009): 4491-4500.

152. Luo, Weibo, et al. "Histone demethylase JMJD2C is a coactivator for hypoxia-inducible factor 1 that is required for breast cancer progression." *Proceedings of the National Academy of Sciences*109.49 (2012): E3367-E3376.

153. Wang, Junjian, et al. "Silencing the epigenetic silencer KDM4A for TRAIL and DR5 simultaneous induction and antitumor therapy." *Cell Death & Differentiation* 23.11 (2016): 1886-1896.

154. Sobczak-Kupiec, Agnieszka, et al. "Role of TRAIL and miR-34a as Therapeutic Agents in Prostate Cancer: Increasing the Armory of Micro-Musketeers." *Molecular Oncology: Underlying Mechanisms and Translational Advancements*. Springer International Publishing, 2017. 237-245.

155. Backe, Marie Balslev, et al. "Lysine demethylase inhibition protects pancreatic β cells from apoptosis and improves β-cell function." *Molecular and Cellular Endocrinology* (2017).

156. Islam, Kabirul. "Allele-specific chemical genetics: concept, strategies, and applications." *ACS chemical biology* 10.2 (2014): 343-363.

157. Baud, Matthias GJ, et al. "A bump-and-hole approach to engineer controlled selectivity of BET bromodomain chemical probes." *Science* 346.6209 (2014): 638-641.

158. Kapoor, Tarun M., and Timothy J. Mitchison. "Allele-specific activators and inhibitors for kinesin." *Proceedings of the National Academy of Sciences* 96.16 (1999): 9106-9111.

159. Chapman, Eli, et al. "A small molecule inhibitor selective for a variant ATP-binding site of the chaperonin GroEL." *Bioorganic & medicinal chemistry letters* 19.3 (2009): 811-813.

160. Islam, Kabirul, et al. "Defining efficient enzyme–cofactor pairs for bioorthogonal profiling of protein methylation." *Proceedings of the National Academy of Sciences* 110.42 (2013): 16778-16783.

161. Islam, Kabirul, et al. "Bioorthogonal profiling of protein methylation using azido derivative of S-adenosyl-L-methionine." *Journal of the American Chemical Society* 134.13 (2012): 5909-5915.

162. Garske, A. L., Peters, U., Cortesi, A. T., Perez, J. L., Shokat, K. M., Chemical genetic strategy for targeting protein kinases based on covalent complementarity. Proc. Natl. Acad. Sci. U.S.A. 108, 15046−15052.

163. Ubersax, Jeffrey A., et al. "Targets of the cyclin-dependent kinase Cdk1." *Nature* 425.6960 (2003): 859-864.

164. Blethrow, Justin D., et al. "Covalent capture of kinase-specific phosphopeptides reveals Cdk1-cyclin B substrates." *Proceedings of the National Academy of Sciences* 105.5 (2008): 1442-1447.

165. Liu, Chang C., and Peter G. Schultz. "Adding new chemistries to the genetic code." *Annual review of biochemistry* 79 (2010): 413-444.

166. Wang, Lei, et al. "Expanding the genetic code of Escherichia coli." *Science* 292.5516 (2001): 498-500.

167. Xie, Jianming, and Peter G. Schultz. "A chemical toolkit for proteins—an expanded genetic code." *Nature Reviews Molecular Cell Biology* 7.10 (2006): 775-782.

168. Baker, Austin S., and Alexander Deiters. "Optical control of protein function through unnatural amino acid mutagenesis and other optogenetic approaches." *ACS chemical biology* 9.7 (2014): 1398-1407.

169. Pham, Nam D., Randy B. Parker, and Jennifer J. Kohler. "Photocrosslinking approaches to Interactome mapping." *Current opinion in chemical biology* 17.1 (2013): 90-101.

170. Sudhamalla, Babu, et al. "Site-specific azide-acetyllysine photochemistry on epigenetic readers for interactome profiling." *Chemical Science* (2017).

171. Breski, Megan. *Allele specific approach to study histone demethylation using engineered KDM4A-(2)-Ketoglutarate pairs*. Diss. University of Pittsburgh, 2015.

172. Reprinted with permission from Breski, Megan, et al. "Engineering Biological C–H Functionalization Leads to Allele-Specific Regulation of Histone Demethylases." *Journal of the American Chemical Society* 138.41 (2016): 13505-13508. http://pubs.acs.org/doi/10.1021/jacs.6b08653. Copyright 2017 American Chemical Society.