Biochemical and Functional Studies of the Histone Demethylases

A thesis submitted to the Board of the Faculty of Physical Sciences of the University of Oxford in partial fulfilment of the requirements for the degree of Master of Science (By Research)



Sophie Therese Williams
Wadham College, Oxford
September 2013

Abstract

Histone lysine methylation is an important post-translational modification (PTM), directing the regulation of gene expression in eukaryotic genomes. Methylation states are dynamically altered by the activity of methyl-transferases and demethylases, including the JmjC 2OG oxygenase Lysine Demethylases (KDMs). KDMs have been associated with perturbed chromatin methylation profiles seen in several pathologies, including cancer and developmental disorders. There is much interest in elucidating both the substrate profile of the KDMs and the full biological functions of KDMs. For this reason, "chemical probes" for KDMs are of particular interest. Selective and potent inhibitors of these enzymes will help to identify the extent to which KDMs drive disease states, and whether they can be targeted as drug candidates in specific contexts.

This thesis presents a systematic study on the sequence and methylation state selectivity specific demethylation of the isolated catalytic domains of the JmjC-domain containing oxygenases, all assigned as KDMs in the literature, using histone H3 synthetic peptide arrays. Demethylation activity was monitored at H3K4, K9, K27, K36 and K79. Results clarify the lysine position and methylation specific demethylation activities of the catalytic domain of the histone demethylases *in vitro*. Unexpected demethylation profiles were observed for the KDM4 subfamily, showing both demethylation at K4 and K27, as well as K9 and K36 (for KDM4A-C).

Several novel substrates were identified and further biochemical characterisation was performed using kinetic and structural analyses. Demethylation activities were quantified for KDM4A and E at tri-methyl K27, while K4 demethylation activity was repeatedly too low to analyse, suggesting it is not a catalytically productive substrate. Structural analyses confirm H3K27me3 binds in the active site of KDM4A in a similar conformation to characterised substrates. Experiments exploring the potential for KDM4A to demethylate *in vivo* showed full length KDM4A to demethylate at K27, suggesting that the novel activity is not just an artefact of the catalytic domain.

With an interest to design specific chemical probes for the KDMs and 2OG oxygenases, I developed an mRNA-display peptide screen and identified potential cyclic peptides inhibitors for the prolyl hydroxylase PHD2. The results from this work are preliminary, but point to wider cellular functions of PHD2 in metabolism and oxygen sensing.

Acknowledgements

Many thanks to Prof. Christopher Schofield and Dr Akane Kawamura, for their support and guidance this year. Thank you also to my colleagues, Louise, Sarah, Richard, Olly and the CJS group, and to Nick for being so encouraging, even when I thought it was impossible!

Contents

Figures vii

Tables ix

Abbreviations x

Chapter 1. Introduction 1

1.1 Epigenetics and the Histone Protein Code 1

1.1.1. The Nucleosome 1

1.1.2. Histone Protein Post Translational Modifications 1

1.1.3. Lysine Demethylation 3

1.2. 2OG Oxygenases 4

1.2.1. Prolyl Hydroxylases 6

1.2.2. DNA Demethylases 8

1.2.3. Histone Demethylases 8

1.2.4. Hydroxylases 9

1.3. Structures of 2OG Oxygenases 9

1.4. Small Molecule Inhibitors for 2OG oxygenases 10

1.5. Project objectives 11

Chapter 2: Materials and Methods 13

2.1 Mass spectrometry and Kinetics methods 13

2.1.1 MALDI end-time point 13

2.1.2 MALDI Kinetics Determination 14

2.1.3 MALDI competition 14

2.1.4 FDH Assay 14

2.2 Microbiological Techniques: Cloning and Protein Purification 15

2.2.1 Plasmid Cloning 15

2.2.2 Protein Constructs and Purification Details 15

2.2.3 Purification Buffers and Protocols 17

2.3 Cell Culture Techniques 19

2.4 Solid Phase Peptide Synthesis 20

2.5 In Vitro Translation System 22

Chapter 3: Substrate Selectivity of the Histone Lysine Demethylases 24

3.1 Introduction **24**

3.2 Developing a Systematic Screen for Demethylase Activity **25**

3.2.1 Phylogenetic Analysis of the KDMs 25

3.2.2 Choosing Peptides for the Demethylation Screen 27

3.3 Systematic Screen of Histone Lysine Demethylase activity 29

3. 4 Novel histone demethylation activity for the KDM4 subfamily **39**

3.4.1 Michaelis-Menten Enzyme Kinetics 43

3.4.2 KDM4A and KDM4E Kinetic Analysis of Demethylation at K27 and K4 48

3.4.3 Competition Experiments of KDM4A Demethylation at K4, K9, K27 and K36 50

3.5 Investigating the Biological Relevance of KDM4A Demethylation at H3K4 and K27 52

3.6 Crystallographic Studies of KDM4A Bound to H3K27 ..... **56**

3.7 Conclusions **57**

Chapter 4. Cyclic Peptide Inhibitors as Biochemical Tools **58**

4.1. Development of Cyclic Peptide Chemical Probes **58**

4.2 *In vitro* Inhibitor Screening: mRNA display technology **62**

Conclusions **69**

References **72**

Appendix 1: Peptide Library **81**

Appendix 2: Mass Spectrometry Figures of Demethylation Activity **84**

Figures

1.1A The structure of DNA packaging, from chromosome to nucleosome
1.1B Histone methylation patterns upon histone proteins 3 & 4
1.1C The histone lysine demethylase mechanism
1.1D View of non-haem Fe(II) coordination site of KDM4A
1.2 Outline mechanism for the 2OG oxygenases
1.3 Examples of modifications catalysed by 2OG oxygenases
1.4 JmjC-2OG oxygenase catalysed demethylation and hydroxylation reactions
1.5 Structures of KDM inhibitors

2.1 SDS PAGE gels to illustrate the purity of enzymes used in the substrate screening
2.2 Chromatogram and SDS PAGE gel of KDM4A after gel filtration
2.3 Chromatogram and SDS PAGE gel of KDM6B

3.2.1 Sequence alignment of catalytic domains for the JmjC domain containing 20G oxygenase enzymes
3.2.2 Kinetic parameters for KDM4E and different length K9 tri-methyl peptides
3.2.3 Time course demethylation profile for KDM4A with 21-mer H3K9 tri-methylated peptide
3.3.1 Summary of H3 lysine demethylation activity of the JmjC domain containing proteins
3.3.2 Comparison of crystal structures of KDM7B with KDM2A with a sequence alignment of the JmjC domains for both enzymes
3.3.3 Overlay of KDM4A with KDM4D
3.3.4 Alignment of JMJD5 and FIH
3.4.1 End point MALDI for all members of the KDM4 subfamily with the mono-, di- and tri-methylated H3K27 peptide.
3.4.2 Time courses of KDM4A and KDM4E K27 demethylation
3.4.3 The reactions underlying FDH monitoring of demethylase activity
3.4.4 Calibration curve for formaldehyde and AFU
3.4.5 Concentration dependent formaldehyde production of KDM4A with pepF as measured by FDH
3.4.6 Kinetic analysis of KDM4A with 21-mer H3K4me3 peptide
3.4.7 Kinetic analysis of KDM4E with 21-mer H3K4me3 peptide
3.4.8 Competition between demethylation at trimethyl H3K27 and HK4 versus K9 and K36
3.4.9 Domain structures for FL-KDM4 enzymes
3.4.10 Demethylation of FL-KDM4A of Alta Bioscience peptides at H3K4, K9, K27 and K36
3.4.11 IF staining of HeLA nuclei after expression of i) wild type or ii) H188A mutant KDM4A

4.1.1 Examples of clinically used cyclic peptides
4.1.2 Schematics to illustrate elements of the *in vitro* and RaPID expression systems
4.1.3 In vitro cyclic peptide screening as used in the RaPID system.
4.2.1 OctetRED binding of His_6-Biotin PHD2 and KDM2A to biotin-coated sensors
4.2.2 Purification of His_6-PHD2 and His_6-Biotin-PHD2
4.2.3 Assessing purity of puromycin-linked mRNA
4.2.4 Determination of cDNA concentration bound to immobilised His_6-Bitoin-PHD2
4.2.5 Agarose gel of PCR amplification of cDNA from the bound cyclic peptides

Tables

2.1 Demethylation assay conditions for MALDI analysis
2.2 Fmoc protected amino acids used in solid phase peptide synthesis

3.4.1 Calibration concentration for formaldehyde and AFU
3.4.2 Reaction conditions tested for FDH assays for KDM4A with pepF
3.4.3 Kinetic parameters determined using FDH for KDM4A and 4E with 21-mer, tri-methylated K4, K9, K27 and K36

4.2.1 Constructs used in development of mRNA-display screen
4.2.2 Peptide sequences identified by using the RaPID system
4.2.3 NICB BLAST for non-redundant protein sequences

Abbreviations

2HG	2-hydroxyglutarate
2OG	2-oxyglutarate
2TY	2 times Tryptone Yeast media
ABHs 1-8	Alkylation repair protein B homologues 1-8
AFU	Arbitrary Fluorescence Units
Ala	Alanine
AlkB	Alkylation repair protein B
Arg	Arginine
Asn	Asparagine
bp	base pair
CAD	C-terminal transactivation domain
CAS	Clavaminic acid synthase
CAW	Chloroacetyltyrosine
CBP/p300	Crebb Binding Protein/protein 300
cDNA	complementary Deoxyribonucleic acid
CIAc-OSu	Chloromethylcarbonyl-oxysuccinamide
CODD	C-terminal oxygen-dependent degradation domain
C-P3H	Collagen prolyl hydroxylase 3
C-P4H	Collagen prolyl hydroxylase 4
Cys	Cysteine
Da	Dalton
DAPI	4'-6-diamindo-2-phenylindole
DMOG	Dimethyl- N-oxalylglycine
dmPDCA	Dimethyl pyridine dicarboxylic acid
DNA	Deoxyribonucleic acid
DSBH	Double-stranded beta-helix
FBXL11	Fbox, Leucine-rich protein 11
FDH	Formaldehyde dehydrogenase
FIH	Factor Inhibiting HIF (Hypoxia-inducable factor asparaginyl hydroxylase)
FIT	Flexible *in vitro* Translation system
FL-KDM4A	Full-length lysine demethylase 4A
FTO	Fat mass and obesity associated protein
Gly	Glycine
H2A/B	Histone protein 2A/B
H3	Histone protein 3
H3K14	Histone protein 3 lysine 14
H3K27	Histone protein 3 lysine 27
H3K4	Histone protein 3 lysine 4
H3K79	Histone protein 3 lysine 79

H3K9	Histone protein 3 lysine 9
H3S10	Histone protein 3 serine 10
H4	Histone protein 4
H4R2me3	Histone protein 4 trimethyl-arginine 2
H4R3	Histone protein 4 arginine 3
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HEK	Human embryonic kidney cells
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIF	Hypoxia Inducible Factor
HOX	Homeobox
IPNS	Isopenicillin N-Synthase
JARID1C	Jumonji acetyltransferase-rich interactive domain 1
JmjC	Jumonji-C
JMJD1A	Jumonji-C Domain Containing 1A
JMJD2A-E	Jumonji-C Domain Containing 2A-E
JMJD3A	Jumonji-C Domain Containing 3A
JMJD4A	Jumonji-C Domain Containing 4A
JMJD5	Jumonji-C Domain Containing 5
JMJD6	Jumonji-C Domain Containing 6
KDM	Lysine Demethylase
KDM1A	Lysine Demethylase 1A
KDM2-7	Lysine Demethylases 2-7
KDM2A	Lysine Demethylase 2A
KDM4A-E	Lysine Demethylase 4A-E
LSD1	Lysine Specific Demethylase 1
Lys	Lysine
MALDI-TOF	Matrix-assisted laser desorption/Ionisation time of flight
Mina53	MYC-induced nuclear antigen
mRNA	messenger Ribonucleic acid
MS	Mass spectrometry
NAD^+	Nicotamine adenine dinucleotide
NADH	Nicotamine adenine dinucleotide, reduced
NFATc1	Nuclear factor of activated T-cells, cytoplasmic 1
NO66	Nucleolar protein 66
NODD	N-terminal oxygen-dependent degradation domain
NOG	N-oxalylglycine
P450	Protein 450
PDC	Pyruvate Dehydrogenase Complex
PDCA	Pyridine dicarboxylic acid
PDK3	Pyruvate Dehydrogenase Kinase-3

PHD finger	Plant Homeobox domain finger
PHD	Prolyl hydroxylase
PHF8	Plant Homeobox Finger Protein 8
PTM	Post-translational modification
RaPID	Random non-standard Peptide Integration Discovery
RED	Rapid Binding Detection
RNA	Ribonucleic acid
RT-PCR	Real time polymerase chain reaction
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electophoresis
Ser	Serine
SPPS	Solid Phase Peptide Synthesis
SPR	Surface Plasmon Resonance
TauD	Taurine dioxygenase
TCA cycle	Tricarboxylic acid cycle
TET	Ten eleven translocation enzymes
tRNA	transfer Ribonucleic acid
Tyr	Tyrosine
UTX	Ubiquitously transcribed tetratricopeptide repeat, X chromosome
UTY	Ubiquitously transcribed tetratricopeptide repeat, Y chromosome
Val	Valine

Chapter 1. Introduction

1.1 Epigenetics and the Histone Protein Code

1.1.1. The Nucleosome

The nucleosome is a DNA/protein complex which compacts DNA into 146 base pair (bp) units.[2] The complex is made up of an octamer of core histone proteins, with two copies of H2A, H2B, H3 and H4 (Fig. 1.1 A, B). The core proteins are well conserved in eukaryotes from yeast to mammals.[3] The tails of histone proteins H3 and H4, which protrude from the nucleosome, are more conserved than the equivalent regions of the less accessible H2A and 2B proteins. Despite this, the *N*-terminus of histone H3 is one of the most dynamically modified regions of all proteins.[4]

1.1.2. Histone Protein Post Translational Modifications

The nucleosome complex, and not free DNA, is accessed by regulatory enzymes. In order for the DNA bound to the histone core to be transcribed, it must be unwound from the nucleosome core to become accessible to the transcriptional machinery. This "open" or "closed" state of DNA is referred to as euchromatin or heterochromatin, and can be controlled in by many different mechanisms.[5] Chemical modifications to DNA and histone proteins play important roles in transcriptional regulation (Fig.1.1 A, B).[6] Patterns of post-translational modifications (PTMs), including acetylation,[7,8] methylation,[9,10] phosphorylation,[11] ubiquitination[12] and biotinylation[13,14] to the tails of the histone proteins are important factors in regulating the sets of genes that are expressed in different contexts (Fig.1.1 B). The presence of PTMs upon histone proteins is controlled by dynamic processes of addition and removal ("writing and erasing"), which change according to the wider cellular environment. Such chemical modifications can act through directing recruitment and recognition of transcriptional machinery and by changing the biophysical properties of DNA-histone protein interactions.[15,16]

N^{ε}-acetylation of lysine residues is regulated by histone acetyltransferases (HATs) and histone deacetylases (HDACs). The presence of acetylated residues reduces the binding affinity for negatively charged DNA. The acetyl groups neutralise the otherwise basic and

positively charged histone proteins, inducing a euchromatin state.[8] Several domain structures in chromatin remodelling enzymes and transcriptional activators are known to bind selectively to acetylated lysine residues. This illustrates that lysine N^{ε}-acetylation can both direct general changes in chemical characteristics and specific enzyme recruitment.

Phosphorylation of serine and threonine residues is commonly associated with transcriptional activation.[17] Like acetylation, the presence of negatively charged phosphate groups is thought to neutralise the positive charge of the histone tail, weakening interactions between the histone proteins and DNA.

Biotinylation of histone proteins has been identified on the histone proteins H2A, H3 and H4. Biotin addition is catalysed by two enzymes, biotinase and holocarboxylase synthase.[13] While some reports suggest biotinylation of histones is associated with heterochromatin,[18] the lack of appropriate analytical techniques available has limited further evaluation of biotinylation in epigenetic regulation.[18]

Ubiquitination occurs on the histone proteins H2A and H2B via the E1-3 ubiquitination pathway. Ubiquitination takes place on the "core" protein sequences, as opposed to the exposed tail regions. Ubiquitination of H2A is associated with heterochromatin states and DNA repair mechanisms,[12,19] while ubiquitination of H2B is associated with euchromatin states by inducing methylation at H3K4 residues.[12]

Methylation at N^{ε}- and N^{ω}- positions of lysine- and arginine-residues can inhibit or activate transcription, depending upon the location. DNA is also subject to methylation, in particular cytosine bases.[20] Methylated lysines in histone proteins were first identified in 1964.[21] Since then, a great deal of information has come to light regarding the associations between histone protein lysine methylation and transcriptional functions,[22] and with methylated DNA.[23] Methylation at H3K9 and 27 is generally considered repressive, while methylation at K4 and K36 is thought to be generally activating.[6,22] Removal of these methyl groups reverses the transcriptional state.

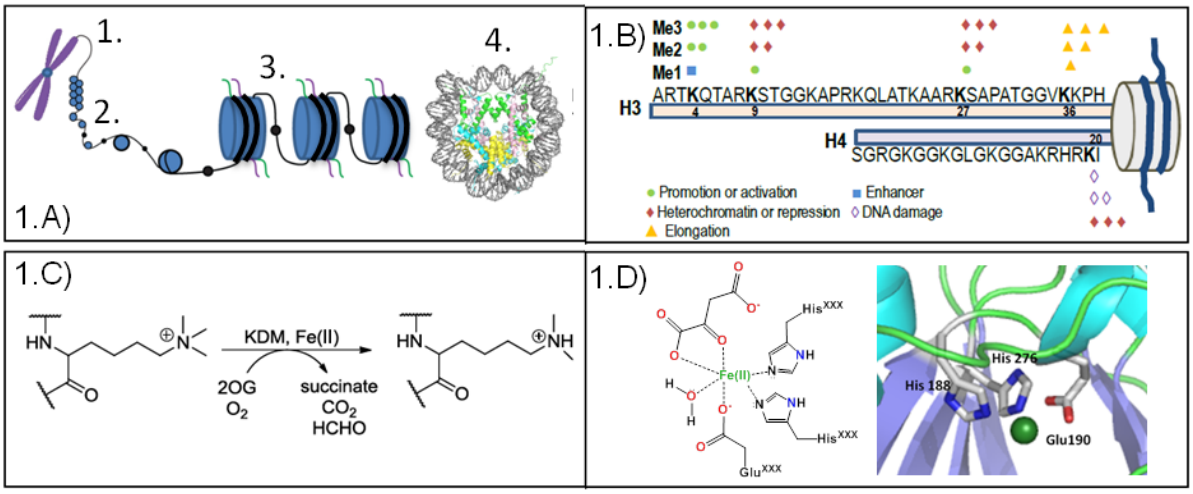


Fig. 1 A-D: 1A: The structure of DNA packaging, from chromosome (1) to chromatin (2) to nucleosome (3). The DNA is coiled on the outside of the nucleosome particle, with the *N*-terminal tails reaching out from the complex core (4, PDB 4GJF). 1B: Histone methylation patterns upon histone proteins 3 and 4. 1C: The histone lysine demethylase mechanism, hydroxylation on an epsilon methyl group results in collapse of the hemi-acetyl, and formation of formaldehyde. Hydroxylation is a stable mark if present on the carbon chain (for example C5 hydroxylation catalysed by JMJD6). 1D: View of non-haem Fe(II) coordination site of KDM4A (PDB 2OX0).

Modifications in close proximity to methylated residues are thought to "tune" demethylation and methylation by modulating the binding of histone methyltransferases/demethylases to their substrates. However, the specific biochemical roles that histone modifications play in transcriptional regulation are largely unclear, and the available evidence implies that histone PTMs do not act as simple molecular "switches". PTMs on the same residue can be associated with increased transcriptional activity, while in a different environment they can inhibit transcription.[24,25,26]

Histone modifications are known to regulate genetic processes in combinatorial ways, both within one protein, and also across proteins within the same nucleosome core. For example, in *S. cerevisiae* phosphorylation of the H3 serine 10 (H3S10) promotes acetylation at H3 lysine 14 (H3K14), whereas the ubiquitination of H2B triggers H3 lysine 4 (H3K4) methylation by the Set1 methyl-transferase, and H3 lysine 79 (H3K79) methylation by the Dot1 methyltransferase.[26,12]

1.1.3. Lysine Demethylation

The 2OG oxygenase lysine demethylases are defined by the Lysine (K) Demethylase (KDM) nomenclature. Until 2004, histone lysine methylation was thought to be irreversible. The

discovery of the Lysine- Specific Demethylase (LSD1, KDM1A), a flavin-dependent monoamine oxidase which catalyses the removal of methyl groups from both H3K4 di- and mono- methylated lysine residues changed this view.[27] LSD1 does not demethylate at tri-methyl lysine residues. The discovery of a second lysine demethylase, FBXL11 (or KDM2A), which catalyses the removal of methyl groups from H3K36me2 led to the identification of a larger family of KDMs.[28] JmjC domain-containing enzymes are Fe(II) and 2-oxoglutarate (2OG) dependent oxygenases, which utilise a haem-free, HxD/ExxH Fe(II) coordination site (Fig. 1.1 C, D). This triad is conserved in most 2OG oxygenases (see Sections 1.2 and 1.3). In contrast to LSD1, the JmjC domain containing KDMs can act on all three tri-, di- and mono-lysine N^{ϵ}-methylation states of lysine residues.

1.2. 2OG Oxygenases

The 2OG dependent oxygenases are the largest known family of non-haem oxygenases.[29,30] 2OG oxygenases are involved in a wide range of biological roles and are conserved in the genomes of most aerobic organisms.[29] 2OG dependent oxygenases regulate a variety of biological processes including the hypoxic response,[31] nucleic acid repair,[32,33] and modification,[34] fatty acid metabolism[35] and histone modification,[36] and many are of interest as potential drug targets.

The catalytic cycle of the 2OG oxygenases has been studied in details for enzymes including clavaminic acid synthase (CAS)[37,38] and taurine dioxygenase (TauD).[39,40] The hydroxylation step follows a radical-based oxygen atom insertion, similar to that of cytochrome P450 mono-oxygenases (Fig 1.2, and accompanying mechanism annotation).[41]

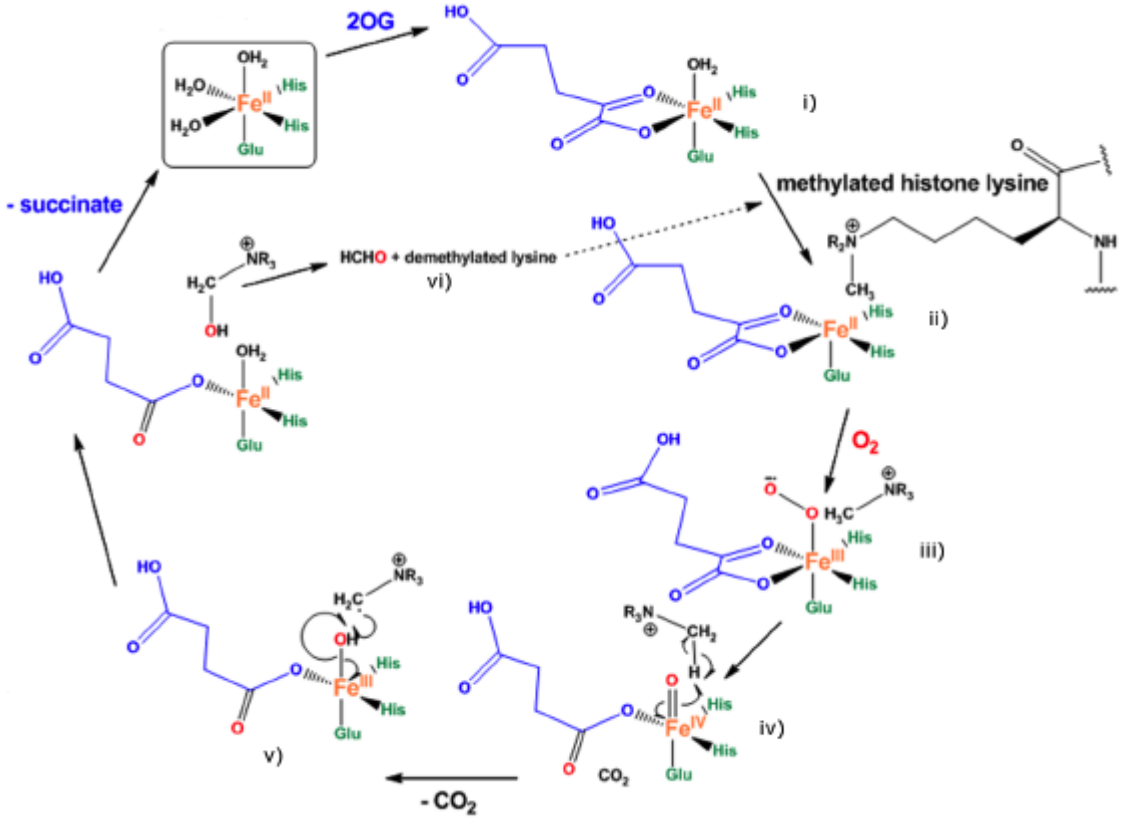


Fig 1.2: Outline mechanism for the 2OG oxygenase demethylation reaction, adapted from Rose *et al.* (i) 2OG binds to the iron centre, coordinating via its 1- and 5- carboxylate groups. (ii) Binding of the alkyl bond containing "prime" substrate to the Fe-2OG complex is thought to facilitate dioxygen binding. Crystallographic data suggests dioxygen binds in a site previously occupied by water. (iii) Oxidative decarboxylation of 2OG generates a Fe(IV) intermediate, succinate and CO_2. (iv) This reactive Fe(IV) species abstracts a hydrogen from the substrate alkyl bond. (v) Hydroxylation of the radical alkyl intermediate occurs. Both oxygen atoms of the dioxygen cosubstrate have been transferred into products. This is followed by a release of succinate and the hydroxylated product. The active site is then regenerated. (vi) In the case of hydroxylated, 2-methyl lysine groups, the product released from the active site undergoes elimination of formaldehyde, reducing the methylation state of the lysine residue.

The 2OG oxygenases are distinct from the P450 enzymes by the presence of an enzyme-bound Fe(II) without a prosthetic haem group, and requirement for 2OG (which undergoes oxidative decarboxylation to form CO_2 and succinate). Many 2OG oxygenases require additional reducing agents in order to catalyse their reactions, including ascorbate (vitamin C); although, it is not clear to what extent other reducing agents, such as glutathione, are used by 2OG oxygenases. The exact role of reducing agents in the catalytic scheme is also uncertain. Recent studies suggest that ascorbate may perform several functions, including maintaining levels of Fe(II) and influencing enzyme stability.[42] Examples of reactions catalysed by 2OG oxygenases are summarised in Fig. 1.3.

1.2.1. Prolyl Hydroxylases

Prolyl -4 and -3 hydroxylation modifications are abundant in biology, present both in collagen (the most abundant protein in humans) and the transcription factor Hypoxia Inducible Factor (HIF)-1α, which is central to cellular oxygen-concentration sensing.

Collagen Prolyl Hydroxylase 4 (C-P4H) was the first identified human 2OG oxygenase and catalyses prolyl-4-hydroxylation with *trans*-stereochemistry.[43,44,45] This post translational modification (PTM) has been shown to enhance collagen stability. A related protein, Prolyl Hydroxylase 3 was found to catalyse prolyl-3-hydroxylation, although this modification was shown to destabilise collagen secondary structure.[45] Collagen hydroxylations have also been recorded at the 5-carbon position of lysine residues, which are thought to stabilise collagen in conjunction with glycosylation modifications.[46]

Oxygen sensing in humans is regulated by the HIF-1/2α isoforms and prolyl hydroxylases (PHD1-3). These enzymes catalyse the hydroxylation of proyl residues in the 4-position within *N*- and *C*-terminal Oxygen-dependent Degradation Domains (CODD and NODD) of HIF-1/2α. Hydroxylation in these regions recruits the von Hippel-Lindau tumour suppressor protein, resulting in ubiquitin-mediated proteosomal degredation of HIF-1/2α proteins. In low oxygen concentrations, the PHDs are unable to catalyse efficient hydroxylation and the HIF-1/2α isoforms are stabilised. The HIF-1/2α proteins are then able to migrate into the nucleus, dimerise with HIF-1β and activate the transcription of genes involved in the hypoxic response. A second mechanism of oxygen dependent HIF-1/2α regulation has also been identified; hydroxylation of Asn803 by FIH in in the *C*-terminal Trans-activation Domain (CAD) prevents interaction of HIF-α with transcriptional co-activators, including CBP/p300.

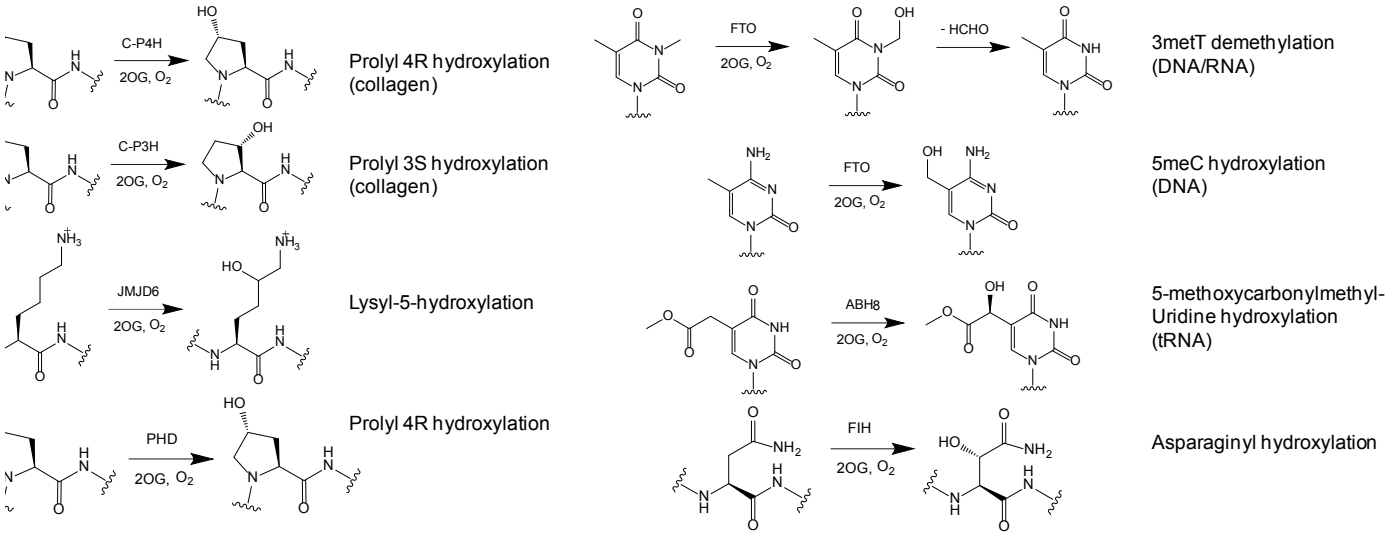


Figure 1.3: Examples of modifications catalysed by 2OG oxygenases. Abbreviations used: collagen prolyl hydroxylase 3 (C-P3H), collagen prolyl hydroxylase 4 (C-P4H), prolyl hydroxylase (PHD), Factor Inhibiting HIF 1alpha (hypoxia-inducible factor asparaginyl hydroxylase, FIH), fat mass and obesity associated protein (FTO),

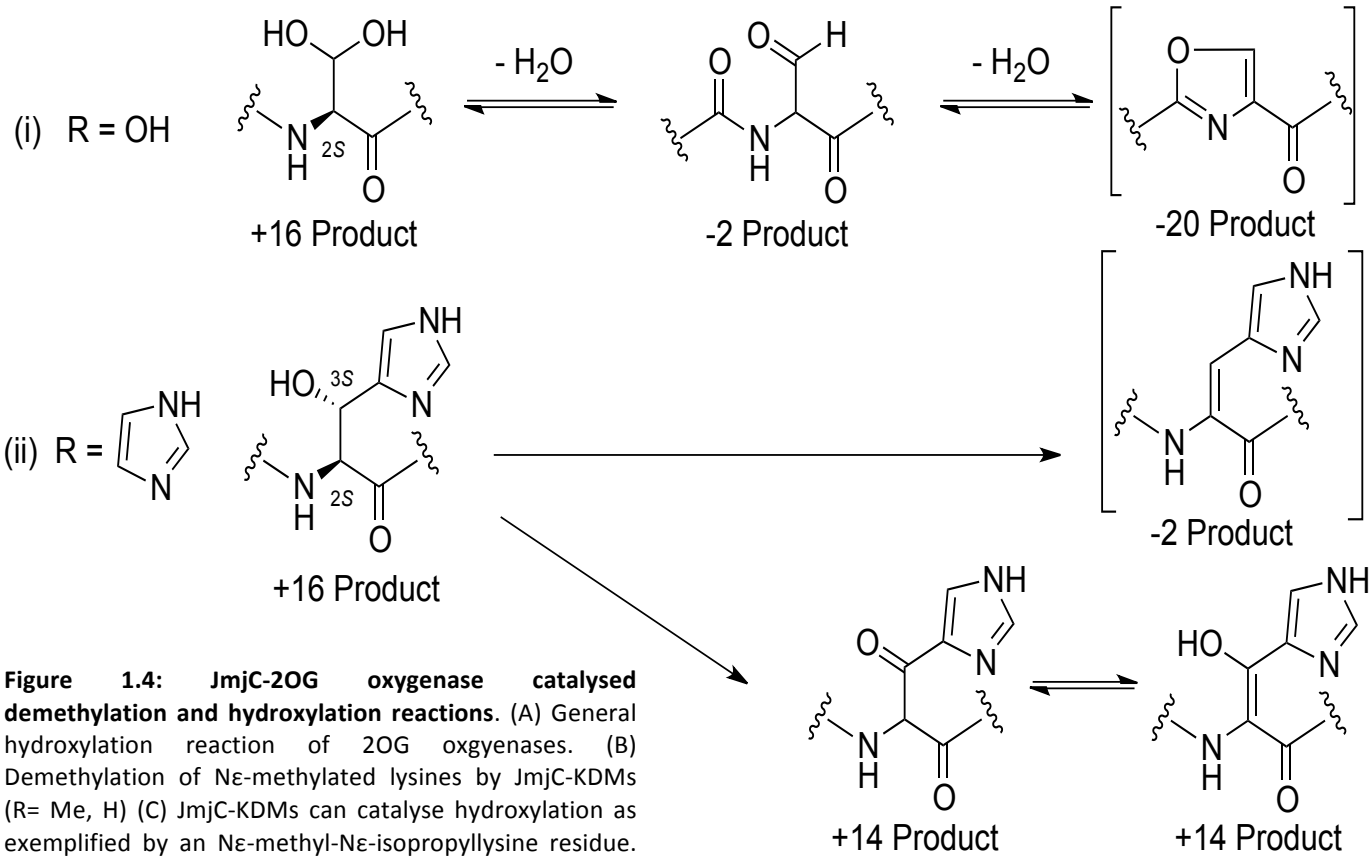


Figure 1.4: JmjC-2OG oxygenase catalysed demethylation and hydroxylation reactions. (A) General hydroxylation reaction of 2OG oxgyenases. (B) Demethylation of Nε-methylated lysines by JmjC-KDMs (R= Me, H) (C) JmjC-KDMs can catalyse hydroxylation as exemplified by an Nε-methyl-Nε-isopropyllysine residue. (D) asymmetric and symmetric dimethylarginine demethylation is catalysed by JmjC-KDMs (unpublished). (E) We have identified JmjC hydroxylase that (separately) catalyse hydroxylation at C-2, C-3, and C-4 positions of lysine-residues. (F) The promiscuity of the JmjC-hydroxylase FIH. Hydroxylations of various residues of FIH are shown in the box (i)-(iii).

1.2.2. DNA Demethylases

DNA demethylation was first identified in *E. coli* as catalysed by the 2OG oxygenase AlkB.[47] AlkB catalyses the repair of *N*-alkylated single-stranded DNA damage caused by alkylating reagents. AlkB can tolerate 1-methyladenine, 3-methylcytosine and, although less effectively, 1-methylguanine and 3-methylthymidine.[32] Eight Human homologues of AlkB have also been identified (ABHs 1-8), although only two, ABH2 and 3 are thought to be active DNA demethylases. ABH8 has been shown to hydroxylate 5-methylcarbonylmethyluridine on tRNA, indicating a breadth of substrate possibilities for these enzymes. Other human 2OG oxygenases are also known to chemically modify DNA. For the example, Fat and Obesity associated enzyme (FTO) catalyses demethylation of 3-methylthyamine and 3-methyluracil,[48] while the Ten Eleven Translocation (TET) enzymes catalyse hydroxylation of 5-methyl cytosine.[20]

1.2.3. Histone Demethylases

To date, there are 31 JmjC domain-containing 2OG oxygenases annotated as histone demethylases.[23] All of the validated JmjC KDMs include defined “accessory” domains in addition to the JmjC domain, which in most cases direct substrate specificity: for example the Tudor and Plant Homeobox Domains (PHDs). These are thought to be epigenomic “readers”, with Tudor domains associating with methylated lysine residues and arginines,[24,25] as well as directing protein-protein and protein-nucleotide interactions, while the PHDs are involved in chromatin regulation and direct affinity to lysines with different methylation states.[26,27,28]

The KDMs are of particular interest as therapeutic targets. They have been implicated in many pathologies, including tumorogenesis,[49] infertility[50] and cognitive-development disorders[51], to name a few. Alterations in methylation status are particularly prevalent in cancers.[52] Several studies have implicated KDMs to drive or modulate patterns of chromatin modification associated with cancerous phenotypes. For example, KDM4A expression is associated with site specific copy-gain mutations, a cancer-state hallmark.[53] In other studies, KDM4A has been associated with cancer-cell proliferation.[54] The absence of a comprehensive survey of the roles of KDMs in disease reflects a more general lack of

information regarding the biological roles of these enzymes. There is a need in the scientific community for more chemical and biological tools to elucidate the numerous biological roles epigenetic modifying enzymes perform.

1.2.4. Hydroxylases

In addition to the JmjC- KDMs, a set of JmjC enzymes do not contain domains associated with histone or chromatin binding (referred to here as the JmjC hydroxylases. See Results for justification of this nomenclature). This subfamily includes JMJD5, JMJD6, Mina53 and NO66, which have also been reported as histone KDMs.[55,56,57,58] However, more recent work has shown JMJD6 to catalyse lysyl-5(S)-hydroxylation on splicing related proteins [59,60] while Mina53 and NO66 have recently been shown to be histidyl- hydroxylases of ribosomal proteins.[61] In addition, it has been proposed that JMJD5 hydroxylates the transcription factor NFATc1.[62] Whilst not displayed in all family members, 2OG oxygenase hydroxylases can be remarkably promiscuous with respect to their sequence selectivity. For instance, Factor Inhibiting HIF (FIH), the first JmjC-domain containing protein shown to be active as an oxygenase,[63,64] was also demonstrated to be highly promiscuous, capable not only of hydroxylating an asparaginyl residue in HIF-1/2α but also multiple asparaginyl, aspartate and histidine residues in ankyrin repeat-domain proteins (Fig. 1.4).[65] Studies in our research group have since confirmed that FIH is able to hydroxylate a number of amino acids, including non-natural amino acids. The work with FIH suggests that other currently assigned JmjC enzymes may catalyse modifications upon other substrates than those identified.

1.3. Structures of 20G Oxygenases

Structural studies of the 2OG oxygenases Isopenicillin N-Synthase (IPNS), involved in the biosynthesis of the penicillin and cephalosporin antibiotics, identified the “signature” double-stranded beta-helix (DSBH) core fold, or “jelly-roll” domain.[66,67] The eight beta-strand fold is also present in many metalloenzymes which bind divalent metals. The highly conserved metal binding site contains a facial triad of two histidine residues and one carboxylic group (either an aspartate, glutamate or, more rarely, the non-carboxylate side chain of a histidine residue). Loops between the fourth and fifth anti-parallel beta strands,

and structural elements spanning the jelly-roll motif have been shown to vary between 2OG oxygenases, and are thought to direct substrate specificity.

1.4. Small Molecule Inhibitors for 2OG oxygenases

A number of inhibitors have been identified for the 2OG oxygenases which include i) "generic" molecules specific to the 2OG oxygenase family, ii) and small molecules specific to 2OG oxygenase subfamilies.[68,69] There is much interest in designing molecules selective for individual enzymes in this family.

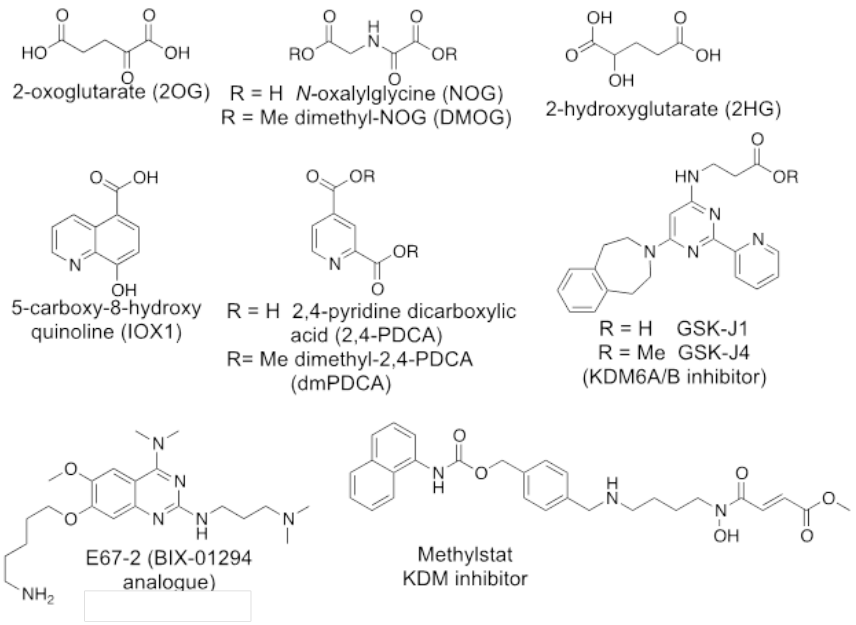


Fig.1.5: Structures of KDM inhibitors (and pro-drugs). Active site targeting generic inhibitors (NOG, IOX1, 2,4-PDCA, methylstat), KDM6 selective probe (GSK-J1) and histone substrate mimetic inhibitor (E67-2) are shown. 2HG is an oncometabolite inhibiting KDM4.

The generic 2OG oxygenase inhibitors include molecules that i) chelate transition metals, sequestering Fe(II) from the active site and ii) mimic the 2OG substrate, which compete for binding at the active site, limiting catalytic turnover (for example 2,4 PDCA, a metal chelator and NOG, a 2OG mimetic compound, Fig. 1.5). More tailored molecules, which target subfamilies of the 2OG oxygenases, reflect common chemical properties of the subfamily's substrates. For example, E67-2 and Methylstat, which are KDM specific inhibitors.

Many rational-design approaches are available for improving small molecule potency against a specific target. Molecule design can utilise i) the presence of allosteric binding sites, ii) changes in active site and protein-surface charge, iii) disruption of protein-protein binding, iv) structural rearrangements, v) the displacement of structurally required water molecules from the enzyme active site and vi) molecules which combine structural motifs from both the enzyme substrate and cofactor ("linked inhibitors"). The generic KDM inhibitor Methylstat adopts this last approach. The molecule maximises interactions within the enzyme active site by including both a cofactor and substrate mimic in one molecule. Novel design methods may enable enzyme-level selectivity within the KDMs, and 2OG oxygenases more specifically. In Chapter 4, I elaborate on the potential for peptide-based molecules to provide potent and selective molecular targets for the 2OG oxygenases.

1.5. Project objectives

Many of the 2OG oxygenases are currently pursued as therapeutic targets both by academic and industrial groups. Despite growing therapeutic interest in these enzymes, the biological function and relevant substrates of many 2OG oxygenases, especially that of KDMs, are far from clear. Systematic profiling is therefore required to clearly define the substrates of KDMs and 2OG oxygenases. These results will be useful in future assay designs and interpretation of biological data in both cellular and clinical contexts.

One approach to gain greater understanding of the cellular functions of these enzymes involves the design of small molecules to specifically modulate enzyme function. Termed "chemical probes", these molecules are useful tools for validating the link between a specific enzyme or set of enzymes in cells, and a phenotype. This allows further exploration of identified targets as a clinical potential outside of a clinical setting.

While there have been two small molecule chemical probes declared for 2OG oxygenases to date (Structural Genomics Consortium (SGC) website; KDM6 subfamily (Kruidenier *et al.* 2012) and PHD (Chowdhry)), development of these probes remains a challenge due to homologies between the 2OG oxygenase subfamilies. Therefore more efficient strategies of probe development are required for targeting other 2OG oxygenases.

The aims of this project are as follows:

i) To conduct a histone lysine methylation profile for 15 JmjC-2OG oxygenases (Chapter 3).

ii) To develop an *in vitro* screen for cyclic and linear peptide chemical probes for KDMs (Chapter 4).

Chapter 2: Materials and Methods

All materials were supplied by Sigma-Aldrich unless otherwise stated. Water (dH_2O) was used from Millipore Elix 10 system.

2.1 Mass spectrometry and Kinetics methods

2.1.1 MALDI end-time point

Enzyme	[2OG] / µM	[Ascorbate] / µM	[Fe^{II}] /µM	2OG K_m /µM	Buffer Conditions
KDM2A	200	100	10	12.45 ± 1.4 ǂ	50 mM HEPES pH 7.5
KDM3A	200	100	10	-	50 mM HEPES pH 7.5, 150mM NaCl
KDM4A/E	200	100	10	12-14.0 ǂ	50 mM HEPES pH 7.5
KDM5C	200	100	10	5.4 ± 0.4 [24]	50 mM HEPES pH 7.5
KDM6A	200	100	10	10 ± 1.3 ǂ	50 mM HEPES pH 7.5, 150mM NaCl, 5% glycerol
KDM6B	200	100	10	8.2 ± 1.0 ǂ	50 mM HEPES pH 7.5, 150mM NaCl, 5% glycerol
KDM7A	200	100	10	16.0 ± 1.0 ǂ	50 mM HEPES pH 7.5, 150mM NaCl
JMJD5	200	100	10	-	50 mM HEPES pH 7.5, 150mM NaCl 5% glycerol, 10 mM DTT
JMJD6	500	100	50	10.1 ± 1.4 ǂ	50 mM HEPES pH 7.5, 150mM NaCl 5% glycerol, 10 mM DTT
MINA53	200	100	10	-	50 mM HEPES pH 7.5, 150mM NaCl 5% glycerol, 10 mM DTT
NO66	200	100	10	-	50 mM HEPES pH 7.5, 150mM NaCl 5% glycerol, 10 mM DTT
FIH	200	100	10	10.0 ± 1.0 [23]	50 mM TRIS pH 7.5, 150mM NaCl

Table 2.1 Demethylation assay conditions for MALDI analysis. Kinetic parameters as reported, ǂ indicates unpublished data.

Demethylation activities on peptides were measured by Matrix-assisted laser-desorption/ionization (MALDI) time-of-flight (TOF) mass spectrometry (MS), using a BRUKER UltraFlex TOF/TOF instrument. Reaction conditions varied between enzymes as optimised for each reaction (see table). The concentrations of 2OG varied depending upon the K_m values for each enzyme. In the case of JMJD6, the low IC50 for succinate, the biproduct of 2OG coupled oxidative decarboxylation required high levels of 2OG to outcompete any succinate produced by the reaction. All reactions used 1 µM enzyme, 10 µM peptide, 2-oxyglutarate (2OG) $FeSO_4.7H_2O$, sodium ascorbate. Reactions were incubated for 60-120 min at 37 ^oC, quenched with 1:3 MeOH by volume and spotted using α-Cyano-4-hydroxycinnamic acid (CHCA) matrix. Relative intensities of demethylation/hydroxylation

observed in the mass spectra were used to calculate the percentage demethylation. For demethylation were more than one product was observed (i.e. the removal of more than one methyl group in the tri- and di-methylated substrates) the percentage demethylation is reported by reduction of the highest methylation state.

2.1.2 MALDI Kinetics Determination

In some instances, Michaelis-Menten enzyme kinetics assays were performed using MALDI endpoint assays. Reactions were quenched within the linear range, as determined by an initial time course. Peptide concentration was varied from 10 µM to 1000 µM in 10 µL reactions and incubated at room temperature. Reaction conditions: 50 mM HEPES pH7.5, Fe(II) ammonium sulphate 10 µM, sodium ascorbate 100 µM, 2OG 200 µM, enzyme 2 µM. Reactions were undertaken in duplicate. Demethylation was determined by MALDI, as stated above. Kinetic parameters were fitted using available equations in GraphPad Prism 5.

2.1.3 MALDI competition

Reaction conditions were used as in 2.1.1, although scaled to a final volume of 50 µL, and included a second peptide at 10 µM. Samples were mixed and incubated in 0.5 ml Eppendorf tubes at room temperature. At each time point (0, 30 s, 2, 5, 10, 15, 20, 30 min etc.) 5 µL of the reaction mixture was removed, quenched with a 1:2 ratio of sample to MeOH and snap frozen in liquid nitrogen. The zero time point was quenches without the addition of enzyme. Samples were stored at -80 ^{o}C until analysed.

2.1.4 FDH Assay

Production of formaldehyde as a by-product of demethylation was observed using the FDH/NAD^+ coupled assay.[70] Buffer: 50 mM HEPES pH7.5 with 0.01% Tween20, Fe(II) ammonium sulphate 10 µM, sodium ascorbate 100 µM, 2OG 200 µM, NAD^+ 500 µM, enzyme 2 µM with 0.025 U FDH enzyme per assay (total volume 25 µL), conducted at room temperature. Peptides at various concentrations were diluted into a clear-bottom black Greiner 384-well plate, a mix of 2OG, ascorbate, Fe(II) and NAD^+ were added in advance of the enzyme mix, which was added and mixed within 30 sec of the first plate measurement. Subsequent measurements were made every 30 sec after for periods between 30 min and 1 hr. The k_m and V_{max} values were calculated during the linear phase of formaldehyde

production, as determined by the slope of fluorescence signal recorded. Measurements were made using a PHERAstar FS (BMG Labtech) plate reader with 355 nm extinction and 460 nm emission).

2.2 Microbiological Techniques: Cloning and Protein Purification

Standard sterile techniques were employed throughout. All instruments and pipette tips were sterilised in a TouchClave® II autoclave (LTE Scientific) at 121°C for 20 min.

2.2.1 Plasmid Cloning

Inserts for KDM4B, C, D, KDM6A, B, PHF8, KDM2A, PHD2 and FIH were amplified from expression vectors (pNIC28-BsaI or PET28(+)) and cloned into the pNIC-Bio3 plasmid (SGC). Constructs were expressed with biotin supplement in BL21(DE3)Rosetta co-transfected with pCDF-BirA.

2.2.2 Protein Constructs and Purification Details

Recombinant enzyme purification was conducted either by myself or another member of the CJS group (as annotated). Other materials, such as vectors or recombinant proteins are annotated if donated from outside the group or purchased.

KDM2A (T. Smart): 1-517aa, His_6 *N*-terminal fusion pNIC28-BsaI vector (SGC), BL21 *E. coli* expression. Purified by 5 ml His-Trap FPLC chromatography. TEV-catalysed His_6 cleavage prior to size exclusion chromatography (Superdex S75).[17]

KDM3A (construct unknown): His_6 *C*-terminal fusion, pFB-CTHF10-LIC vector transformed into DH10bac and transfected in Sf9 cells. Protein and vector produced by the SGC.[18]

KDM4A (S. Williams): 1-359aa His_6 *N*-terminal fusion pNIC28-Bsa4 (SGC), expressed in the phage resistant derivative of *E. coli* BL21(DE3). Purified by Ni(II)-NTA resin using a gravity column and then S75 Superdex size exclusion chromatography.[19,20] KDM4B-E were expressed and purified using the methods outlined in Ng *et al.* (2007). TEV-catalysed His_6

cleavage was undertaken for KDM4B-E.[19] KDM4B (S. Williams): 1-365aa His_6 *N*-terminal fusion pNIC28-Bsa4. Purified by 5 ml His-Trap FPLC and Superdex S75 size exclusion chromatography. KDM4C (CJS): 1-366aa His_6 *N*-terminal fusion pNIC28-Bsa4. Purified by 5 ml His-Trap FPLC and Superdex S75 size exclusion chromatography. KDM4D (CJS): 1-358aa His_6 *N*-terminal fusion pNIC28-Bsa4. Purified by 5 ml His-Trap FPLC and Superdex S75 size exclusion chromatography. KDM4E (S. Williams): 1-337aa His_6 *N*-terminal fusion pNIC28-Bsa4. Purified by 5 ml His-Trap FPLC and Superdex S75 size exclusion chromatography.

KDM5C: 1-765aa His_6 *C*-terminal fusion, pFB-CTHF10-LIC transformed into DH10bac and transfected in Sf9 cells. Protein and vector produced by the SGC.18 KDM6A: 880-1401aa pNIC-NHTrxT (SGC) BL21 *E. coli* expression.[21]

KDM6B (S. Williams): 1141-1590aa His_6 *N*-terminal fusion pNIC28-Bsa4, BL21 *E. coli* expression.[18] Purified by 5 ml His-Trap FPLC and Superdex S75 size exclusion chromatography.

KDM7A (L. Walport): 79–447aa His_6 *N*-terminal fusion pNIC28-Bsa4, BL21 *E. coli* expression (construct includes the PHD Finger domain and JmjC domain). Purified by 5 ml His-Trap FPLC and Superdex S75 size exclusion chromatography.

JMJD5 (S. Williams, Dr S. Wilkins): 1-416aa Rossetta DE3 *E. coli* expression, pET28a(+) vector (Novagen). JMJD6: 1-403aa His_6 *N*-terminal fusion pET28(+) vector, BL21 *E. coli* expression.[10] Purified by 5 ml His-Trap FPLC and Superdex S75 size exclusion chromatography.

NO66 (R. Sekirnik): 116-641aa, His_6 *N*-terminal fusion pET28a(+) vector, BL21 *E. coli* expression.[11] Purified by 5 ml His-Trap FPLC and Superdex S75 size exclusion chromatography.

MINA53 (R. Sekirnik): 26-465aa, His_6 *N*-terminal fusion pET28a(+) vector, BL21 *E. coli* expression.[11] Purified by 5 ml His-Trap FPLC and Superdex S75 size exclusion chromatography

FIH (Dr A. Hardy): 1–348aa, His_6 *N*-terminal fusion pET28a(+), BL21 *E. coli* expression.[23] Purified by 5 ml His-Trap FPLC and Superdex S75 size exclusion chromatography.

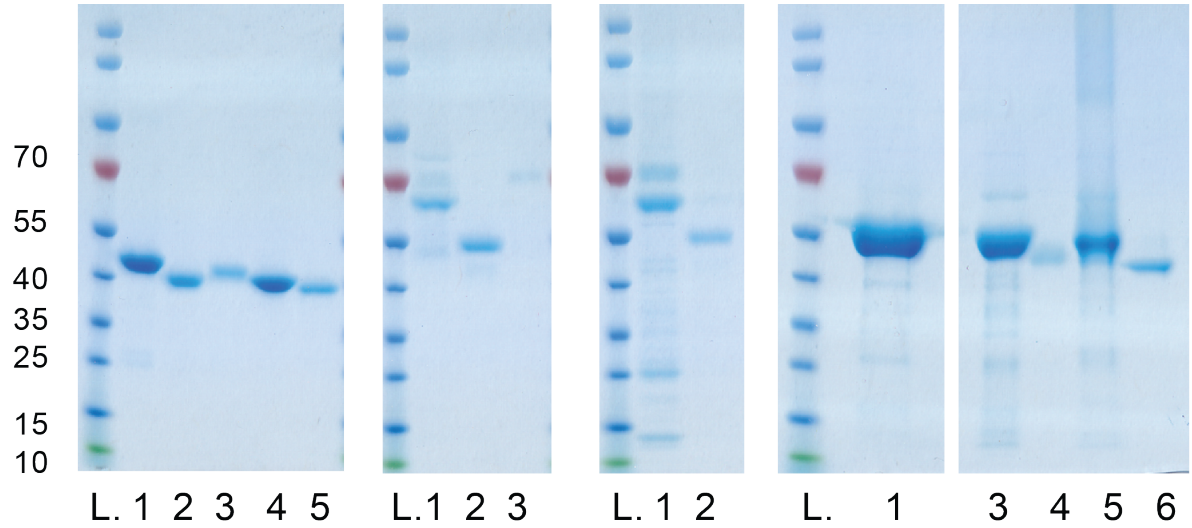


Fig 2.1: SDS PAGE gels to illustrate purity of enzymes used in the substrate screening (left to right, 1-4). L, Ladder. 1. KDM4A (44.2 kDa), KDM4B (42.0 kDa), KDM4C (42.6 kDa), KDM4D (41.3 kDa), KDM4E (38.6 kDa). 2 FBXL11 (62.5 kDa), KDM3A (40.1 kDa), KDM6B (62.4 kDa). 3 UTX (64.2 kDa), PHF8 (44.8 kDa). 4 Mina53 (50.1 kDa), NO66 (58.9 kDa), JMJD5 (49.5 kDa), JMJD6 (48.9 kDa) and FIH (42.5 kDa).

2.2.3 Purification Buffers and Protocols

Example Ni(II)-(NTA) Resin Purification – KDM4A/JMJD2A Bacterial pellet was thawed in 100ml glass beaker with 20ml binding buffer. Pellet was re-suspended over 30 min, 4 ^{o}C. Cells were sonicated over ice (5s ON 5s OFF, 30sec – 30sec OFF cycle x 12) before DNase (1 mg) was added. Cell lysate was then incubated for 20 min over ice. Cell debris was pelleted at 20K rpm, 4 ^{o}C, 15 min. Supernatent was then filtered (0.44 µm) and passed over Biorad 25ml column with 3ml Ni-NTA beads (equilibrated in binding buffer, 2x 10 ml).

Elution from Ni-NTA beads: 60 ml of lysate was passed over the Ni-NTA beads. Column washing: 2 x 10 ml of Binding buffer, 3 x 10 ml wash buffer. Elution: 1 ml x 6, 2 ml x 6, 10ml. Fraction protein content was assessed by Nanodrop (A280). Fractions from the Elution steps were concentrated and injected onto a Superdex 75 gel filtration column (pre-equilibrated into Gel Filtration buffer). SDS Page analysis shows expression of protein (44.3 kDa, ext. 75.3 x 10^3 Abs) and purification to ~90%. Fractions 8,9 and 16-21 were pooled and concentrated to a final concentration of 12.9 mg/ml (293 µM).

Buffer	Materials
Binding buffer	500mM NaCl, 5% glycerol, 50mM HEPES (pH7.5), 10mM Imidazole
Wash buffer	500mM NaCl, 5% glycerol, 50mM HEPES (pH7.5), 20mM Imidazole
Elution buffer	500mM NaCl, 5% glycerol, 50mM HEPES (pH7.5), 250mM Imidazole
Filtration buffer:	500mM NaCl, 5% glycerol, 10mM HEPES (pH7.5),

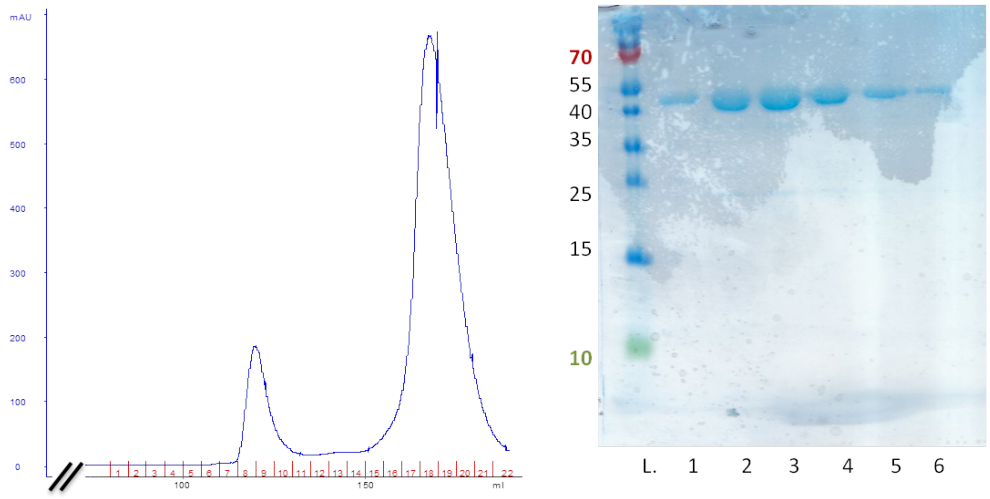


Fig. 2.2 – Chromatogram and SDS PAGE gel of KDM4A after gel filtration (Superdex 75). Gel lanes correspond to fractions 9, 16, 18, 19, 20, 21. All fractions were pooled and concentrated for use in demethylation assays.

Example Ni(II) 5ml His-Trap Purification – KDM6B Bacterial pellet was thawed in a 100ml glass beaker with 20ml binding buffer. Pellet was re-suspended over 30 min, 4 ^{o}C. Cells were sonicated over ice (5s ON 5s OFF, 30sec – 30sec OFF cycle x 12) before DNase (1 mg) was added. Cell lysate was then incubated for 20min over ice. Cell debris was pelleted at 20K rpm, 4 ^{o}C, 15min. Supernatent was then filtered (0.44 µm) and passed over a 5ml His-Trap© (GE Healthcare) column (equilibrated with 15 ml binding buffer). Elution from 5ml His-Trap©: 60 ml of lysate was passed over the column. Washing: 25 ml Binding Buffer, 25 ml Wash Buffer, 100% gradient to Elution Buffer over 30 ml, and continued for 25 ml (all at 1 ml / min). Fractions from chromatogram were run on an 8% acrylamide SDS-PAGE gel (shown below). The fractions corresponding to lanes 7-11 were pooled, concentrated <1.5ml and loaded onto a Superdex 75 gel filtration column (pre-equilibrated into Gel Filtration buffer). SDS-PAGE analysis after gel filtration shows expression of protein (62.4 kDa, ext. 62.5 x 10^3

Abs) and purification to ~90%. Fractions 12-16 were pooled and concentrated to a final concentration of 34.93 mg/ml 524.6 µM.

Buffer	Materials
Binding buffer	500mM NaCl, 5% glycerol, 50mM HEPES (pH7.5), 10mM Imidazole
Wash buffer	500mM NaCl, 5% glycerol, 50mM HEPES (pH7.5), 20mM Imidazole
Elution buffer	500mM NaCl, 5% glycerol, 50mM HEPES (pH7.5), 250mM Imidazole
Gel filtration buffer	500mM NaCl, 5% glycerol, 10mM HEPES (pH7.5), 0.5mM TCEP

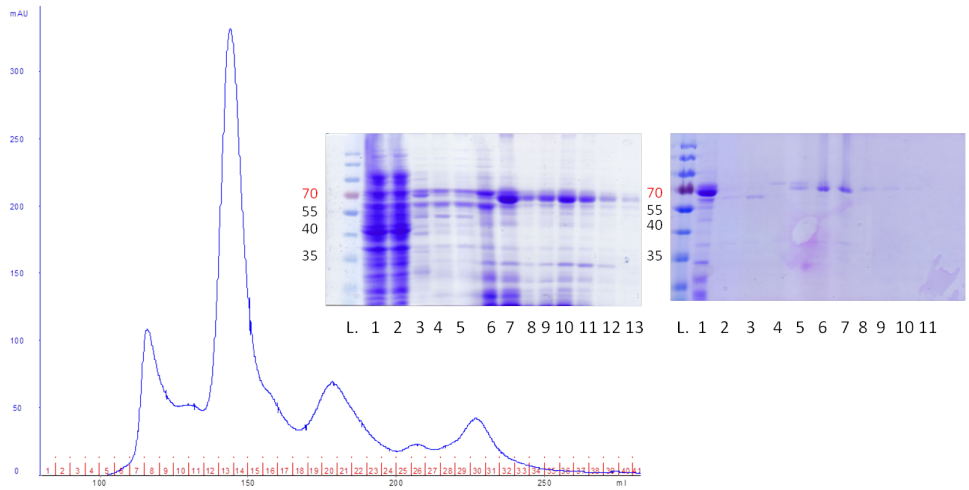


Fig.2.3 – Chromatogram and SDS PAGE gel of KDM6B using i) 5 ml His-Trap Ni(II) purification and ii) gel filtration (Superdex 75). Gel lanes correspond to fractions 8, 10, 11, 12, 13, 14, 15, 17, 23, 30. Lane one is the input control. Fractions 12-16 were pooled and concentrated for use in demethylation assays.

2.3 Cell Culture Techniques

Expression and Purification of Full length KDM4A Full length KDM4A was transiently transfected in HEK293T (50 µg DNA per 15 cm^2 diameter plate, 80% confluency, grown in DMEM, PAA Laboratories) and harvested at 24 hrs. Cells were washed in PBS, lysed in 50mM HEPES, 150 mM NaCl, 0.01% NP40, 1% protease inhibitor cocktail (SIGMA), 1% DNAseI by volume (10 mg/ml stock) and incubated with 60 µL FLAG-Magnetic Beads (SIGMA) for 4 hrs at 4 ^{o}C. Beads were washed with lysis buffer for four 600 µL washes, and then only 50 mM HEPES for two 300 µL washes (to remove NP40 and NaCl before MS analysis). The beads were then made to the initial volume of 60 µL and used in further assays. Demethylation reactions were conducted with 1 µL of the immobilised anti-FLAG beads with 10 µM peptide, 200 µM 2-oxyglutarate (2OG), 10 µM $FeSO_4.7H_2O$, 100 µM sodium ascorbate.

Reactions were incubated for 120 min at 37 ^{o}C, quenched with 1:3 MeOH by volume. The demethylation products were observed by MALDI.

Immunofluorescence of Global Histone Marks HeLa cells were maintained in OptiMEM reduced serum medium (Invitrogen) supplemented with 0.5% foetal calf serum (FCS) (Invitrogen) and 1% penicilin-streptomycin (PAA laboratories). Cells were seeded (1,500 per well) into a 96-well optical grade plate (Becton Dickinson) and left overnight to adhere. Cells were transfected with pcDNA3 FLAGx3-KDM4A (1-359) wild type and H188A mutant for 24 hrs. Cells were washed with PBS, fixed in 4% paraformaldehyde (20 min), and permeabilised with 0.5% TritonX-100 (10 min) in PBS. After 30 min blocking (3% FCS in PBS), cells were incubated overnight with primary antibody (1:500) anti-histone H3 antibodies (K9me3 (Abcam, ab8988), K4me3 (Diagenode, pAB-003-050), K27me3 (Millipore, 07-449), K36me3 (Abcam, ab9050) and further incubated with the secondary antibody (1:500) goat anti-rabbit Alexafluor 488 (Invitrogen) (1 hr). Cell nucleus was stained with DAPI (1:1000), (Invitrogen). **Image acquisition and analysis:** To analyse the relative intensity of tri-methyl lysine states at H3K4, K9, K27 and 36 between the mutant and wild type samples an automated imaging analysis system (Beckton Dickinson Pathway) was used. For each well in a 96-well plate, the system acquired a 3-by-2 tile-scanned image for Alexafluor 488 and DAPI. The Pathway's software used the DAPI staining to identify nuclei as regions-of-interests. It is also able to identify and isolate transfected cells from non-transfected cells. For each nucleus, the software extracted the average intensity of the histone staining, followed by the average intensity of all the nuclei in that particular well. The non-transfected cells are used as a background fluorescence control. The average intensity for each well, with the background subtracted, was plotted for the wild type and mutant using GraphPad Prism 5.

2.4 Solid Phase Peptide Synthesis

All peptides synthesised in this work used a CS-Bio automated solid phase peptide synthesiser (CS-Bio 336X machine) or, for the in house histone-screen, the Intarvis MultiPep machine. Lists of peptides used and synthesised is included in Appendix 1. All amino acids

were supplied as α-amine Fmoc protected amino acids. Standard amino acids were supplied by CSBio. Methylated lysine was purchased from Novabiochem.

Amino acid	Fmoc protected amino acid for SPPS
Alanine	Fmoc-Ala-OH
Arginine	Fmoc-Arg(Pbf)-OH
Asparagine	Fmoc-Asn(Trt)-OH
Aspartic acid	Fmoc-Asp(OtBu)-OH
Cysteine	Fmoc-Cys(Trt)-OH
Glutamic acid	Fmoc-Glu(OtBu)-OH
Glutamine	Fmoc-Gln(Trt)-OH
Glycine	Fmoc-Gly-OH
Histidine	Fmoc-His(Trt)-OH
Isoleucine	Fmoc-Ile-OH
Leucine	Fmoc-Leu-OH
Lysine	Fmoc-Lys(Boc)-OH
Methionine	Fmoc-Met-OH
Phenylalanine	Fmoc-Phe-OH
Proline	Fmoc-Pro-OH
Serine	Fmoc-Ser(tBu)-OH
Threonine	Fmoc-Thr(tBu)-OH
Tryptophan	Fmoc-Trp(Boc)-OH
Tyrosine	Fmoc-Tyr(tBu)-OH
Valine	Fmoc-Val-OH
Monomethyl lysine	Fmoc-Lys(me, Boc)-OH
Dimethyl lysine	Fmoc-Lys(me2)-OH
Trimethyl lysine	Fmoc-Lys(me3)-OH

Table 2.2: Fmoc protected amino acids used in SPPS

All synthesis was undertaken in dimethylformamide (DMF, 100 µmol scale for CB-Bio, 10 µmol scale for the Intarvis- Multipep machine). Activating agents 1-hydroxybenzotriazole (HOBt) and diisopropylcarbodiimide (DIC) were used in combination with piperidine, as the Fmoc deprotection reagent. Poly(aminomethylstyrene) resin (Polymer Labs) was used for synthesis, together with the Rink Amide linker (CS-Bio). Amino acids were used at 10x excess (1 mmol), and activating reagents (DIC, HOBt) were also used at 1 mmol. More expensive methylated amino acids were used at 3-5x excess. Each amino acid was activated by addition of DIC/HOBt for 30 min prior to addition to the resin. After addition to the resin, the activated amino acid was mixed with the resin/growing peptide chain for 2 h, prior to removal of excess reagent and washing of the resin. Piperidine (20 % in DMF) was used in

two treatments (8 min and 2 min) to deprotect the *N*-terminal of the peptide, followed by extensive washing of the resin, and addition of the next amino acid.

2.5 In Vitro Translation System

FIT system Flexizyme RNA and $tRNA^{fMet}_{CAU}$ were prepared as previously reported.[71] Translation factors, enzymes and ribosome were purchased from GE Healthcare. Reactions were undertaken as previously described.[72-74]

Preparation of puromycin-fused mRNA library RNAs consisting of 8−12 repeated NNK random sequences (5'-GGGUU AACUU UAAGA AGGAG AUAUA CAU AUG $(NNK)_m$ UGC GGC AGC GGC AGC GGC AGC UAG GACGG GGGGC GGAAA-3', m = 8−12) were prepared by *in vitro* transcription according to the previously described method[75]. The mRNA library was ligated with a puromycin linker (5'-CTCCC GCCCC CCGTC C-$(SPC18)_5$-CC-puromycin-3') by T4 RNA ligase. The ligated product was purified by phenol-chloroform extraction and ethanol precipitation.

RT PCR cDNA Quantification 1x Real-Time PCR mix (Invitrogen)

PCR Reaction	
Water	835
10x PCR buffer	100
250mM MgCl2	10
5mM dNTPs	50
100uM T7g10M.F48	2.5
100uM G5San13.R39	2.5
Total	**1000ul**

Real-Time PCR	
1xPCR mix	100
1/10,000 SYBR green	1
1xTaqDNA polymerase	1.5
Total	200ul

The Reverse Transcription Product was incubated in PCR mix at 95°C for 5 min. The supernatant was removed while the sample was still hot using pre-heated magnet, to ensure that peptides were not bound to the immobilized target. For the RT-PCR, 1ul of positive sample was to 19 µL real-time PCR mix, and 1 µL of negative control sample to 19

µL real-time PCR mix. For the standards, 1 µL of 2 x 10^4, 10^6, 10^8 cDNA samples were added to 19ul RT PCR mix. The conditions for RT PCR were as follows: 94 ^oC 40 sec, 61 ^oC 40 sec, 72 ^oC 40 sec, repeated for 30 cycles; 61>72 ^oC (0.5C/ sec).

Phenol/Chloroform Extraction: An equivalent volume of phenol/chloroform/isoamyl alcohol (25:1:1) was added to the DNA or RNA and vortexed. The samples were centrifuged (13,000 rpm, for 10 min). The supernatant was extracted, and the PCI non-aqueous layer disposed in organic halogen waste (standard PPE required). Equivalent volumes of chloroform/isoamyl alcohol (25:1) and were then added to the recovered supernatant, vortexed and centrifuge at 13,000 rpm, 10 min. The supernatant was then extracted again. The volume of the recovered product was doubled with 100% ethanol, vortexed, and then centrifuged at 13,000 rpm for 15 min. The supernatant was removed, air dried for 5 min and then dissolved in 30 µL water.

Chapter 3: Substrate Selectivity of the Histone Lysine Demethylases

3.1 Introduction

This chapter describes a systematic study on the sequence and methylation state selectivity of the isolated catalytic domains of human JmjC-domain containing oxygenases, all assigned as KDMs in the literature, using histone H3 synthetic peptide arrays. The KDM nomenclature will be used (KDM#), expect in the cases of enzymes which show no demethylation activity. These are annotated by the 2OG oxygenase nomenclature.

No systematic study of histone mark catalytic activity for the lysine demethylases has been previously conducted. Prior to the discovery of KDM1A/LSD1[27] and KDM2A/FBXL11[28], the presence of methyl-lysines on histone proteins was well documented, although they were considered stable and irreversible marks[76]. Over the last decade, more than 20 KDMs have been reported in the literature, the majority of which are JmjC-domain containing proteins that belong to the larger family of 2OG oxygenases.[77]

The first identified KDM, KDM2A was tested against each of the known marks (H3K4, K9, K27, K36, K79 and H4R3 and K20) and shown to be a K36-di- and mono-methyl specific demethylase.[28] Other KDMs were quickly identified, including JMJD1A (KDM3A/JHD2A)[78] and KDM4A/JMJD2A[79]. The latter was the first report to investigate the activities of several KDM enzymes (KDM4A-E/JMJD2A-E) in a single study. A substrate profile was conducted as part of the biochemical characterisation of KDM4A-C, showing specific demethylation at K9 and K36, tri- and di-methylated lysines, although the *in vitro* demethylation activity reported for these marks is relatively low (<25 %) after 2-5 hr incubations at 37 ^{o}C.

Reports of novel histone demethylases often include a screen for demethylase activity against each of the known histone marks using either mass spectrometry methods or antibody based detection methods. In some cases, lack of specificity of antibody detection, particularly in the presence of post-translational modifications, has caused controversy over the claimed function. For example, JMJD6 was initially reported as a H4R3me2 demethylase using antibody detection methods[56], although later publications could not reproduce this

result.[80,81,60] Webby *et al.* (Science, 2009) subsequently demonstrated JMJD6 to act as a lysyl hydroxylase (as discussed later). It is possible that in the work of Chang *et al.* hydroxylation of the substrate protein blocks antibody binding, giving a reduced signal, rather than indicating a loss in signal due to demethylation of H4R3me2. More generally, antibodies can show non-specific associations in heterogeneous samples. Since antibody detection is qualitative, rather than quantitative, high background signal can result in false positive results. Therefore, mass spectrometry methods, which can directly measure changes in chemical modifications and isolate the site of modification (by MS/MS fragmentation), may in general be more robust detection techniques for demethylation / hydroxylation.

A mass spectrometry based detection method using Matrix-Assisted Laser Desorption/Ionization MS (MALDI-TOF MS) was used to conduct the demethylation screen. MALDI is a soft-ionization technique, which detects masses most commonly in a [M+1] ionisation state and has a detection limit (in the reflectron mode) of 8000 Da, within the range necessary for the demethylation screen (peptide masses range from 800 to 3000 Da). MALDI enables processing of samples, without additional chromatography elution time and use of a 96 “spot-plate”, which can be used to queue samples for analysis.

The results clarify the lysine position and methylation specific demethylation activities of the catalytic domain of the histone demethylases *in vitro*, as well as indicating greater promiscuity, both of demethylase and hydroxylase function for several of the enzymes tested. For the novel substrates, further biochemical characterisation was undertaken using kinetic and structural analyses (Sections 3.4 and 3.6).

3.2 Developing a Systematic Screen for Demethylase Activity

3.2.1 Phylogenetic Analysis of the KDMs

A set of enzymes was chosen to represent different subfamilies of the JmjC containing 2OG oxygenases (studied enzymes are underlined in the phylogenetic tree in Fig. 3.2.1). The figure illustrates a multiple sequence alignment (Clustal Omega, EBI) of the catalytic domains of each of the human JmjC-domain containing 2OG oxygenases (as annotated in

Uniprot). As discussed in the Introduction, there is a large number of human 2OG oxygenases, which although contain a JmjC catalytic domain, catalyse other oxidative modifications distinct from lysine demethylation. From available sequence-based and structural evidence, non-JmjC 2OG oxygenases show marked differences in structural motifs within the substrate binding and catalytic sites to those of the proposed histone lysine demethylases.[66] Thus, they have not been included in this phylogenetic tree alignment.

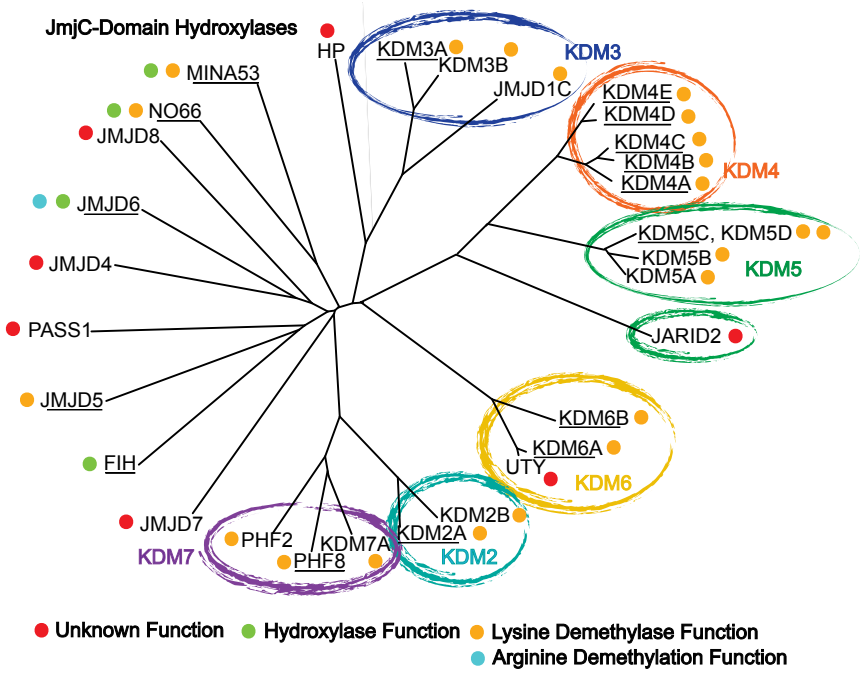


Figure 3.2.1: Sequence alignment of catalytic domains for the JmjC domain-containing 2OG oxygenase enzymes (CLUSTL Omega, PHYLIP3.6). Subfamilies are represented in branches. Catalytic function and proposed demethylase substrate is listed next to enzymes tested in this screen (underlined).

Topographical distances between branches present on the phylogenetic tree (Fig. 3.2.1) represent degrees of difference in sequence, and allow a systematic comparison of sequence between enzymes. While sequence-based phylogenetic analysis cannot offer direct information about structural or substrate similarity, the sub-groupings of enzymes according to sequence act as a surrogate for functional, and to a degree enzyme-evolution information (the more similar enzymes are to each other, the more likely their substrate profile will be similar, and more closely related enzyme sequences are likely to have evolved from a single origin in lower organisms).

The phylogenetic tree identifies seven subfamilies of potential histone lysine demethylases: groups KDM2-7 and a subset of loosely associated JmjC-domain containing enzymes clustering close to FIH – a known protein-hydroxylase.[64] At least one member of each of these subfamilies was chosen to represent members from the subgroup, although it should be noted that substrate selectivity is not necessarily conserved between members of the same subfamily (for example, the substrate selectivity of the KDM4 subfamily).

The enzymes were also chosen for the availability of expression and purification protocols. Proteins used only included catalytic domains. The proteins used were purified to ≥85% purity, as assessed by SDS-PAGE gel (Materials and Methods). All proteins were obtained from within the CJS group, with the exception of KDM3A and KDM5C, which were expressed using baculovirus expression systems in insect cells and provided for use in this work by the Structural Genomics Consortium, SGC).

3.2.2 Choosing Peptides for the Demethylation Screen

Histone peptides used in this screen included the Alta Bioscience H3 library of 21-mer peptides and a collection of in-house peptides, which vary in length and methylation site (sequences are listed in Appendix 1, Peptide library). The K4 and K9 comprise of the same histone sequence (A1-A21), only differing in position of methylated lysine. All the Alta Bioscience peptides included an *N*-terminal biotin group, with an aminohexanoic-acid linker region (mass 339.8 Da, not present on the in-house peptides).

In some instances, modifications and affinity tags present on the termini of peptides have been known to affect catalytic activity of histone lysine demethylases (for example PHF8). In this screen, no difference in catalytic efficiency was found between peptides with or without the linker region for demethylases other than PHF8. Rather, peptide length was seen to be a more significant factor in demethylation turnover (as discussed below).

Substrate	K_M (μM)	k_{cat} (s^{-1})
2OG	14.0 ± 1.8	0.087 ± 0.003
AR**Kme3**STGGK	21.6 ± 1.3	0.063 ± 0.00
AR**Kme2**STGGK	28.5 ± 3.9	0.065 ± 0.002
ARTKQTAR**Kme3**STGGKA	17.4 ± 2.5	0.073 ± 0.002
ARTKQTAR **Kme2**STGGKA	32.3 ± 7.4	0.070 ± 0.004
TAR**Kme3**STGGKAPRKQLA	14.8 ± 3.9	0.067 ± 0.003
ARTKQTAR**Kme3**STGGKAPRKQLATKVA	0.8 ± 0.4	0.075 ± 0.001

Table 3.2.2: Kinetic parameters for KDM4E and different length K9 tri-methyl peptides. KDM4E shows greater affinity for longer peptides. Determined by Dr N. Rose using MALDI-TOF detection

The effect of length of peptide on catalytic efficiency has been explored for several members of the 2OG oxygenases, including the KDM4s (Table 3.2.2, Dr N. Rose[82]). In most cases, longer peptides allow for greater binding between the substrate and enzyme, enhancing turnover, justifying the use of longer peptide in this study.

The peptides used in demethylation assays contain mono-, di- and tri-methylated lysines at each of the K4, K9, K27 and K36 positions. After assigning the methylation state specificity for the demethylases, peptides including either tri- or di-methylated K79 sites were tested for demethylation.

The assay conditions for demethylation used cofactor concentration above K_m values for known optimal substrate for each enzyme (Materials and Methods), to reduce the number of variables potentially limiting catalytic activity. The buffer compositions were matched to optimized conditions from the literature. In all cases (except JMJD5), a positive control peptide was included in the screen to ensure enzyme activity. For JMJD5, no published peptide substrate has been identified, and so the observation of 2OG turnover to form succinate was used as a surrogate for enzyme activity.

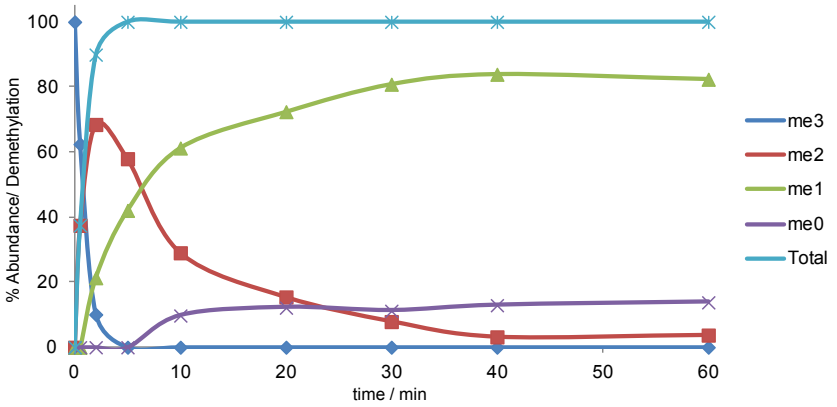


Fig. 3.2.3: Time course demethylation profile for KDM4A with 21-mer H3K9 tri-methylated peptide (as determined by MALDI). Each mark (me3-me0) is reported as a relative abundance. Total demethylation follows the conversion of tri-methyl state to other products.

Time courses were run for run for each enzyme substrate pair to ensure that the activity was in the linear range, and the assay conditions were optimal. Overall, enzyme and peptide reactions of 1-1.5 hrs at 37 ^{o}C, was sufficient. Reactions were quenched by methanol and analysed by MALDI-TOF. An example of demethylation profile for KDM4A over a time course of 60 min (Fig. 3.2.3) shows the conversion of tri-methyl to di-methyl, mono-methyl and zero-methylation states.

3.3 Systematic Screen of Histone Lysine Demethylase activity

The majority of my findings were consistent with already published results; however, in certain cases, unexpected substrate promiscuity was observed. A heat map of demethylation activities (Fig.3.3.1) summarises the N^{ε}-methyl lysine specific activity of the JmjC 2OG oxygenases on the H3K4 K9 K27 and K36 peptides. No activity was seen for any demethylase or hydroxylase on K79 (data not shown). Examples of MALDI data for each enzyme are given in Appendix 2 (labelled S.1-13). The results of the substrate profiling of representatives from each of the human KDM subfamilies are summarised below:

KDM2A. The peptide screening results reveal KDM2A/FBXL11 catalyses demethylation at H3K36 (Fig. 3.3.1, S.1). It does not demethylate tri-methyl lysine, but does act on di- and mono-methylated states. These results are consistent with published studies.[83,84]

KDM3A. The peptide screening results reveal KDM3A catalyses demethylation at di- and mono-specific demethylation at H3K9 (Fig.3.3.1, S.2). These results are consistent with previous reports. [78]

% Demethylation by catalytic domains of JmjC 2OG oxygenases

Subfamily	Enzyme	H3K4 Me3	H3K4 Me2	H3K4 Me1	H3K9 Me3	H3K9 Me2	H3K9 Me1	H3K27 Me3	H3K27 Me2	H3K27 Me1	H3K36 Me3	H3K36 Me2	H3K36 Me1	+16
KDM2	KDM2A	×	×	×	×	×	×	×	×	×	×	44.35 ±0.31	67.20 ±0.00	
KDM3	KDM3A	×	×	×	×	90.27 ±1.72	85.67 ±0.52	×	×	×	×	×	×	
KDM4	KDM4A	39.05 ±0.12	✓*	×	94.71 ±0.53	80.07 ±5.96	×	74.69 ±0.30	33.52 ±0.15	✓*	84.50 ±0.02	39.42 ±5.74	×	
	4B	×	×	×	91.35 ±0.06	71.59 ±0.35	×	24.36 ±6.63	23.08 ±6.06	×	54.36 ±1.09	24.99 ±0.01	×	
	4C	×	×	×	94.90 ±0.35	56.56 ±3.70	×	34.51 ±10.64	21.99 ±3.74	×	57.12 ±0.76	23.6 ±11.80	×	
	4D	×	×	×	78.24 ±1.05	66.95 ±8.39	×	73.44 ±0.40	36.63 ±6.20	×	×	×	×	
	4E	9.83	✓*	×	53.17 ±4.52	80.58 ±0.53	29.80 ±3.15	73.44 ±0.40	36.63 ±6.20	✓*	×	×	×	
KDM5	KDM5C	94.19 ±0.12	72.52 ±0.50	×	×	×	×	×	×	×	×	×	×	
KDM6	KDM6A	×	×	×	×	×	×	10.64 ±0.42	86.49 ±0.10	68.10 ±9.34	×	×	×	
	KDM6B	×	×	×	×	×	×	34.32 ±9.33	63.36 ±16.15	72.20 ±7.82	×	×	×	
Hydrox.	MINA53	×	×	×	×	×	×	×	×	×	×	×	×	75.46 ±2.10
	NO66	×	×	×	×	×	×	×	×	×	×	×	×	71.75 ±2.66
	JMJD5	×	×	×	×	×	×	×	×	×	×	×	×	NA
	JMJD6	×	×	×	×	×	×	×	×	×	×	×	×	95.50 ±6.43
	FIH	×	×	×	×	×	×	×	×	×	×	×	×	89.92 ±0.65

% Hydroxylation

% Activity 100 0

Fig 3.3.1: Summary of H3 lysine demethylation activity of the JmjC domain containing proteins. Only catalytic domains used. Peptide sequences used were as follows (bold residue modified): H3K4me3, Biotin-spacer-ART**K**QTARKSTGGKAPRKQLA; K9, Biotin-spacer-ARTKQTARKSTGG**K**APRKQLA; K27, Biotin-spacer-KAPRKQLATKAAR**K**SAPATGG, K36 Biotin-spacer-SAPATGGV**K**KPHRYRPGTVAL. Asterisk* indicates demethylation observed from tri-methyl to mono-methyl, but not from di-methyl to mono-methyl. PHF8 data is not included in this table since i) a full length construct was used, including domains other than the JmjC domain, and ii) a free *N*-terminus is required for activity, and so these peptides were screened for activity from an in-house library without *N*-terminal modifications. The substrates used for the hydroxylases are taken from the literature (MINA53 and NO66, Ge *et al.* 2012, JMJD6 Webby *et al.* 2009, FIH Cockman *et al.* 2006).

KDM4A-E. KDM4A-C showed demethylation at K9 and K36 tri- and di-methylated lysine, with some residual activity at the mono-methyl state leading to complete demethylation (Fig. 3.3.1, S.3-5). KDM4D and E showed demethylation at K9 tri- and di-methyl lysine, again with limited demethylation of the mono-methyl state (S.6-7). No demethylation was

observed at K36 tri- or di-methyl lysine for KDM4D and E. Alternative activities not previously reported for KDM4 subfamily were also observed.

Unexpectedly, the catalytic domain of all members of the KDM4 family demethylated the H3K27 methylated peptides at tri- and di-methylated sequences (S.3, Fig. 3.4.1). In addition KDM4A and KDM4E, but not other members of the family, were able to demethylate at H3K4 although only after longer incubation periods (>3 hr incubation at 37 ^{o}C). This difference in alternative function across the subfamily may arise from increased "overall" enzyme activity (KDM4A and E showed consistently higher levels of demethylation over the other members of the subfamily), or from more distinct differences in the substrate-binding pocket. The biochemical characterisation of these novel demethylation activities by KDM4 enzymes is discussed in more detail below (Section 3.3)

KDM5C. KDMD5C demethylated K4 tri- and di-methyl lysine, but did not show mono-methyl activity under the conditions tested (Fig.3.3.1, S.9). This is consistent with published work.[85]

KDM6A and B. KDM6A and B demethylate tri-, di- and mono-methylated H3K27 (Fig.3.3.1, S.10), and show no activity at other H3 modified lysine residues. This is in agreement with the literature.[86]

KDM7B The KDM7B/PHF8 catalytic domain has been reported to show a degree of promiscuity between di- and mono-methylated lysine, catalysing demethylation at K9, K27 and K36 di- and mono-demethylation.[51] PHF8 expressed with only the catalytic domain was not active. Instead, a construct containing both the catalytic domain and PHD finger domain was tested for demethylation. The same demethylation profile was observed (H3K9, K27 and K36 di-methyl and mono-methyl demethylation). This multi-domain construct showed greater demethylation activity on di-methyl and mono-methyl K9 when a H3K4me3 modified residue was also present. This reflects the ability of the PHD finger in PHF8 to bind at this site and direct demethylase activity at K9 in the presence of tri-methylated K4.

JmjC-Domain Containing Hydroxylases. The hydroxylases JMJD6, Mina53, NO66, JMJD5 and FIH were tested for demethylation and/or hydroxylation with the Altabioscience, H3 library and the in-house library of untagged histone peptides. No demethylation was observed on any of the tested histone peptides. As a positive control for enzyme activity, all enzymes were also tested against reported substrates for hydroxylation, where available.

JMJD6 was tested against the H3 library, as well as a selection of modified peptides derived from H4, the proposed histone substrate for JMJD6. It has been a reported that JMJD6 catalyses arginine demethylation on H4R3(symmetric)Me2.[56] However, several groups have disputed this report. [80,87,88] In my screen, no lysine demethylation or arginine demethylation of H3 or H4 peptides was observed. However, under the same conditions, JMJD6 was able to hydroxylate a control peptide from LUC7-like2 (LUC7L2), a protein associated with RNA splicing (S.11).[59]

Histone peptide hydroxylation has been observed for JMJD6 on histone proteins (Dr M. Mantri [81] and in the literature [80]). This was not clearly repeated in my screen, although this is likely to be due to a reduction in enzyme activity compared to previous studies. Further studies should be conducted to elucidate the full histone peptide hydroxylation activity of JMJD6.

For NO66 and MINA53, hydroxylation on histidine residues in ribosomal protein peptides was observed (S.12). These peptides are derived from the 60A ribosomal proteins L7 and L8. The presence of irreversible hydroxylation modifications to these proteins in the ribosome is thought to influence transcription efficiency.[61] This is consistent with the proposed role of Mina53 and NO66 in cell proliferation,[89,90,57,91] their interaction with the Myc onogene pathway [89,57] and their up-regulation in several cancers and association with poor prognosis.[89,90,91,92-95,

FIH showed no demethylase activity on histone peptides, but did catalyse hydroxylation of peptides from HIF1α and ankyrin repeat domains[96] (S.13).

For JMJD5, no peptide substrate for hydroxyation has been reported to date, although increased succinate turnover has been shown in the presence of NFATc1.[62] In this screen no demethylation was observed on any histone peptide, including H3K36me2, the reported substrate of JMJD5.[55]

Discussion of Results

Systematic screening using mass spectrometry detection methods has confirmed some of the peptides substrates and methylation-states specificities for the JmjC KDMs.

KDM2A/FBXL11. KDM2A has been shown to have both histone and non-histone substrates (p65).[97] In this screen of histone peptides, KDM2A showed fidelity towards di- and mono-methylated H3K36, suggesting the enzyme requires specific chemical motifs in order to recognise its substrates. However, the reported p65 peptide substrates (Lys218, NSGSCLGGDEIFLICD-**Kme1** and Lys221, VQ-**Kme2**-EDIEVYFTGPGWEAR)[97] show little similarity to the H3K36 sequence (SAPATGGV**-Kme2-**KPHRYRPGTVAL). One general similarity is a double-glycine motif in proximity to the demethylated lysines. However, with the modified lysines positioned towards the end of the peptides, it is hard to determine any common structural motifs. KDM2A may only able to demethylate this peptide substrate because of the lack of surrounding residues. Attempts to synthesise these peptides were not successful, and so could not be included in the general screen. To assess the residues involved in directing KDM2A activity at these sites, peptides which include varying lengths of p65 sequence around these modification sites should be synthesised and tested for demethylation. The absence of a clear rationale for specificity towards these two peptides suggests KDM2A may also demethylate at other methylated lysines.

The methylation state specificity of KDM2A towards di- and mono-methylated lysine is conserved between the two substrates (Fig.3.3.2). It is of interest to compare KDM2A active site residues to those of other KDMs which i) accept the substrate (even if the demethylation state is different), and ii) accept the same methylation-state, but shows a difference substrate preference.

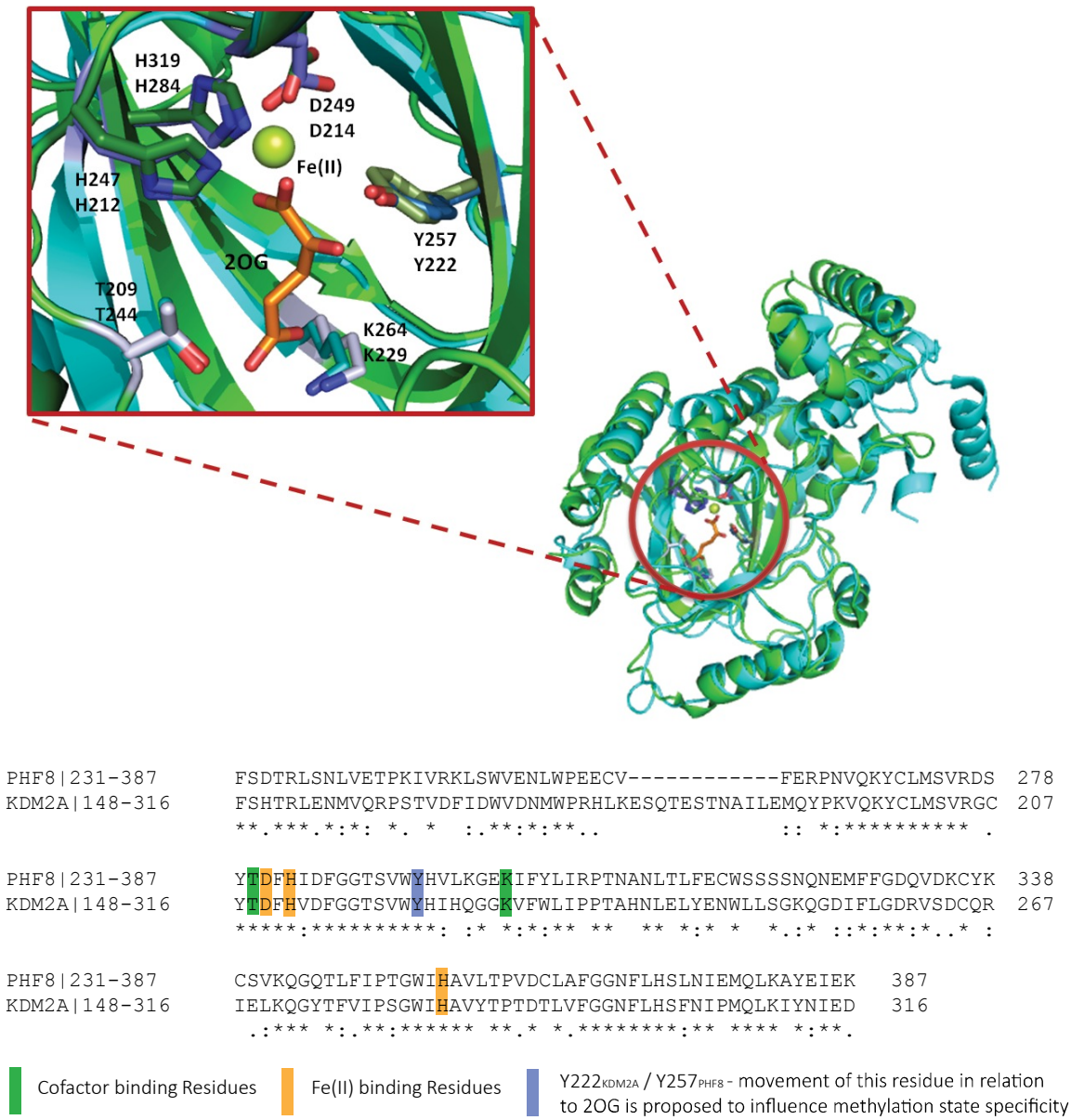


Fig. 3.3.2: Comparison of crystal structures of KDM7B/PHF8 (PDB 2WWU, green) with KDM2A/FBXL11 (PDB 2YU1, cyan), with a sequence alignment of the JmjC domains for both enzymes. Key residues involved in substrate binding and/or catalysis are highlighted. $Y222_{KDM2A}$ (blue) is conserved between KDM2A and PHF8, but not with KDM4A, which utilises an asparagine (N198). The larger tyrosine is proposed to sit closer to the metal in KDM2A and PHF8, while KDM4A utilises N198 to interact with the C-5-carboxylate of bound inhibitor NOG (Yue *et al.* 2010).

Crystallographic studies of the catalytic domain of KDM2A as compared to KDM7B (another di- and mono-methyl specific KDM) suggest this activity is directed by specific residues in the active site. The tyrosine Y222 of KDM2A, conserved in both KDM2A and KDM7B, but not in KDM4A (a tri-methyl specific demethylase) appears to direct this specificity.[1] This could be one rationale for methylation specific activity of KDM2A; however, it should also be

noted that much structural similarity to KDM4A present in the active site, close to the catalytic metal centre. This suggests that other residues, besides Y222, will be involved in directing methylation state specificity.

KDM3A. Studies using antibody detection techniques *in vivo* suggest both KDM3A and B act as di- and mono-methyl K9 demethylases involved in androgen receptor mediated gene transcription activation.[98] No catalytic activity has been observed for KDM3C.[99]

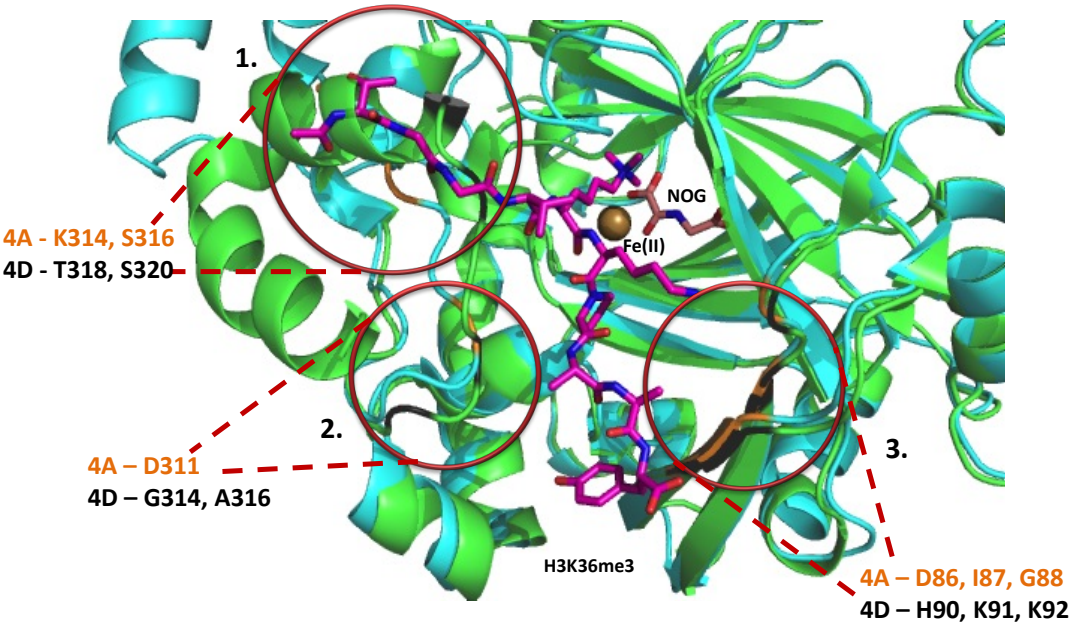


Fig.3.3.3: Overlay of KDM4A (2OQ7, cyan, K36 peptide, magenta) with KDM4D (3DXU, green). Regions 1-3 correspond to key differences in protein sequence at substrate binding site. Region 1: Orange residues, Lys314, Ser316 (4A); black resiudes, Thr318, Ser320 (4D). Region 2: Orange residues, Asp311 (4A), balck residues Gly314, Ala316 (4D). Region 3: Orange residues Asp86, Ile87, Gln88, (4A) black residues, His90, Lys91, Lys92 (4D).

KDM4. The KDM4 subfamily has been studied extensively and the demethylation profile of KDM4A-C and KDM4D/E towards H3K9 and H3K36 has been reported.[79,100] Both branches of the KDM4 subgroup, KDM4 A-C and D/E, showed greatest activity for the tri-methylated K9 site, but different demethylation activity towards K9 (all KDM4 members) and K36 (only KDM4A-C).[79] This substrate selectivity between K9 and K36 is thought to arise from differences in substrate binding residues located between this pocket and the Zn(II) binding

site. KDM4A-C enzymes use these residues to interact with the histone-substrate backbone at Arg40 and Tyr41. However, these are not conserved in KDM4D or E.[101]

An overlay of the KDM4A and KDM4D crystal structures (as depicted in Fig.3.3.3) illustrates how the residues in KDM4A and E differ around this site.[100] A loop region extends from the active site in KDM4A, creating space for a "kink" at Pro30, enabling it to contact residues on the external surface of the enzyme. However, the equivalent region in KDM4D is positioned closer to the active site, limiting space for the peptide to bind which suggests why the affinity for this peptide is too poor to allow demethylation.

The methylation state-specific activity for the KDM4s has been rationalised by the relative proximity of methyl groups to the catalytic iron centre and, possibly, the rate of rotation of the presented methyl group.[102] However, results in this chapter show demethylation can occur to the zero-methyl state. This suggests that methylation-state specific activity may be related to rate of catalytic turnover, rather than simple structural differences in the catalytic core. This proposal is supported by findings that the more active construct and enzyme purification batches for KDM4A and E, as compared to B-D, show greater demethylation to the me0 state.

The methylation state is important for maximising interactions between the substrate and KDM4 active site. A tri-methyl lysine may act as a better "anchor" than the di-methylated state, with di- methyl peptides showing lower kinetic-turnover efficiency than for the tri-methyl peptide analogues.[103] This is consistent with results presented here, with the KDM4s showing demethylation of the tri-methyl state to the zero methylation state, while no mono-methyl state is seen for the K9me2 peptide, and very little demethylation to K9me0 observed at K9me1. The ability of Ng *et al.* (2009) to crystallize KDM4A in the presence of K9 and K36 mono-methylated peptides suggests a degree of affinity consistent with demethylation of the mono-methyl state, even if demethylation was not observed.[102]

Recent work has suggested KDM4A-C enzymes can catalyse demethylation of lysines in non-histone proteins.[104] There is no direct similarity in sequence between the identified non-

histone substrates and histone substrates. To some extent, this supports our hypothesis that the lysine demethylase enzymes are more promiscuous than thought. KDM4A has previously been seen to demethylate at a mutant H3K27me3 sequence when a di-glycine motif is incorporated at the positions +3 and +4 to the modified lysine (mimicking the K9 sequence);[105] however, work presented here suggests this is not essential for activity at K27.

In light of the interesting demethylation profile for the KDM4 enzymes, further biochemical investigations were carried out (Section 3.4).

KDM5C. The subgroup of KDM5A-D shows specific demethylation at H3K4, as identified *in vivo* using antibody techniques, which suggests KDM5A plays a key role in gene up-regulation during development.[106] A binding screen of modified histone peptides identified only H3K4 di- and mono-methyl lysine as significant hits, both *in vitro* and *in vivo*.[107]

KDM6A/B. The enzymes UTX and JMJD3 are specific H3K27me3 demethylase enzymes. They have been implicated in gene repression during development, and somatic mutations in these enzymes are associated with several pathologies, from cancer to congenital heart disease. The third enzyme, UTY, is thought to be catalytically inactive, although unpublished studies from our group suggest this not to be case. Evidence from the literature also supports a role for UTY in gene expression; endogenous levels of this enzyme in male foetuses has been shown to partially recover the function of UTX in its absence.[108]

KDM7B The KDM7 enzymes comprise of PHF8, PHF2 and KIA1718. All three are thought to act as di-methyl demethylases at K9 and K27.[51,109,110,110] For KDM7B, demethylation activity at H3K36 on peptides was also observed in this screen, as reported.[51] However, it is disputed as to whether demethylation at this site occurs *in vivo*.[111,112] Histone proteins contain a repeated "RKS" motif, present at K9 and K27. This similarity between the sequences surrounding the methyl-lysine may enable PHF8 to demethylate at both sites. However, little sequence similarity exists between these sites and K36. Close structural homology between KDM7B and the K36 di-methyl specific lysine demethylase KDM2A supports evidence for demethylase activity of PHF8 at this site.[1] *In vivo,* the presence of

extra non-catalytic domains may direct demethylation activity towards K9, K27 and not K36. The PHD finger in KDM7B is known to bind to H3K4 trim-methyl lysine, directing activity towards residues closer to the histone *N*-terminus.[112]

JmjC Hydroxylases No members of this subgroup of the JmjC domain-containing enzymes showed any histone demethylation activity in this study. Instead, for several enzymes hydroxylation was observed. This observation is consistent with the crystal structures of these enzymes which show more similarity to FIH in catalytic residue positions within the active site than to the lysine demethylase enzymes (MINA53, 2XDV; NO66, 4DIQ; JMJD5, 4GJZ (as shown below, Fig. 3.3.4), JMJD6 ,3K2O[113]; FIH 2CGN[114]). The structure of JMJD5, for instance, correlates most closely to FIH rather than to KDM2A or the KDM4A-C enzymes, which are known di-methyl H3K36 demethylases.[115] For JMJD6, lysine hydroxylation on histone peptides incorporating dimethyl arginine has been reported, showing that the enzyme has a preferential activity for hydroxylation of lysine over demethylation of arginine. No hydroxylation of peptides by JMJD6 was observed in this screen, but either combinations of modifications or longer incubation periods may be required to see histone peptide activity.

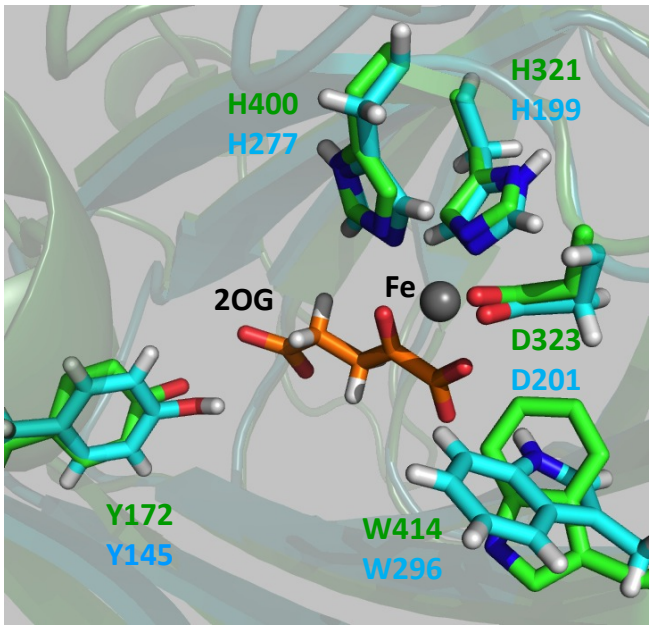


Fig. 3.3.4: Alignment of JMJD5 (4GJZ, green) and FIH (1MZF, cyan). Fe(II) and 2OG binding residues are conserved.

The domain structure of the JmjC domain-only proteins also suggests alternative roles to histone protein demethylation. Their lack of annotated binding accessory domains, for example like the PHD and Tudor domains identified in lysine histone demethylases, limits these enzymes' involvement in epigenetic regulation involving the complex combination of modifications present along histone protein tails.

3. 4 Novel histone demethylation activity for the KDM4 subfamily

As described above, unexpected demethylation activity at H3K27 was observed for each member of the KDM4 family, and in the case of KDM4A/JMJD2A, activity was seen at K4, K9, K27 and K36 (although not at K79, Fig. 3.4.1).

Although screens of these enzymes have been published previously, no evidence for activity at K27 has been reported. There is some variation between the expression construct length (1-350 v.s.1-359 in this study) and demethylase buffers (Tris-HCL pH 7.5, rather than HEPES pH 7.5). This may have resulted in less active recombinant protein in other studies, and therefore weakly observable demethylation as compared to activities reported here.

Demethylation was only observed for the tri-methyl and di-methyl states of the H3K27 peptide. For the mono-methyl state, the combination of low affinity of the enzyme to the K27 backbone sequence and the low methylation state may reduce the overall catalytic efficiency to such an extent that demethylation is no longer observable under these conditions.

Experiments were undertaken to determine the best substrate for demethylation at K27 (Fig. 3.4.2). KDM4A and E were tested for demethylation against several length K27 peptides over a time course of 120 min at room temperature. A small amount of background demethylation (5%) was observed for Pep49 (AARKme3SAPATG) in the zero time point.

As with previously identified substrates (Section 3.2), longer peptides showed greater catalytic turnover. This observation is consistent with the literature (Hillringhaus *et al.* 2011). Both 8-mer and 12-mer peptides showed less than 20% demethylation after 2 hrs

with KDM4A and E. The 21- and 25-mer peptides showed markedly higher levels of demethylation (between 55-75% for both KDM4A and E). The slight increase in demethylation for the 21-mer may suggest the presence of the biotin tag activates demethylation, although this trend has not been observed in other assays, suggesting the result here to be an artefact.

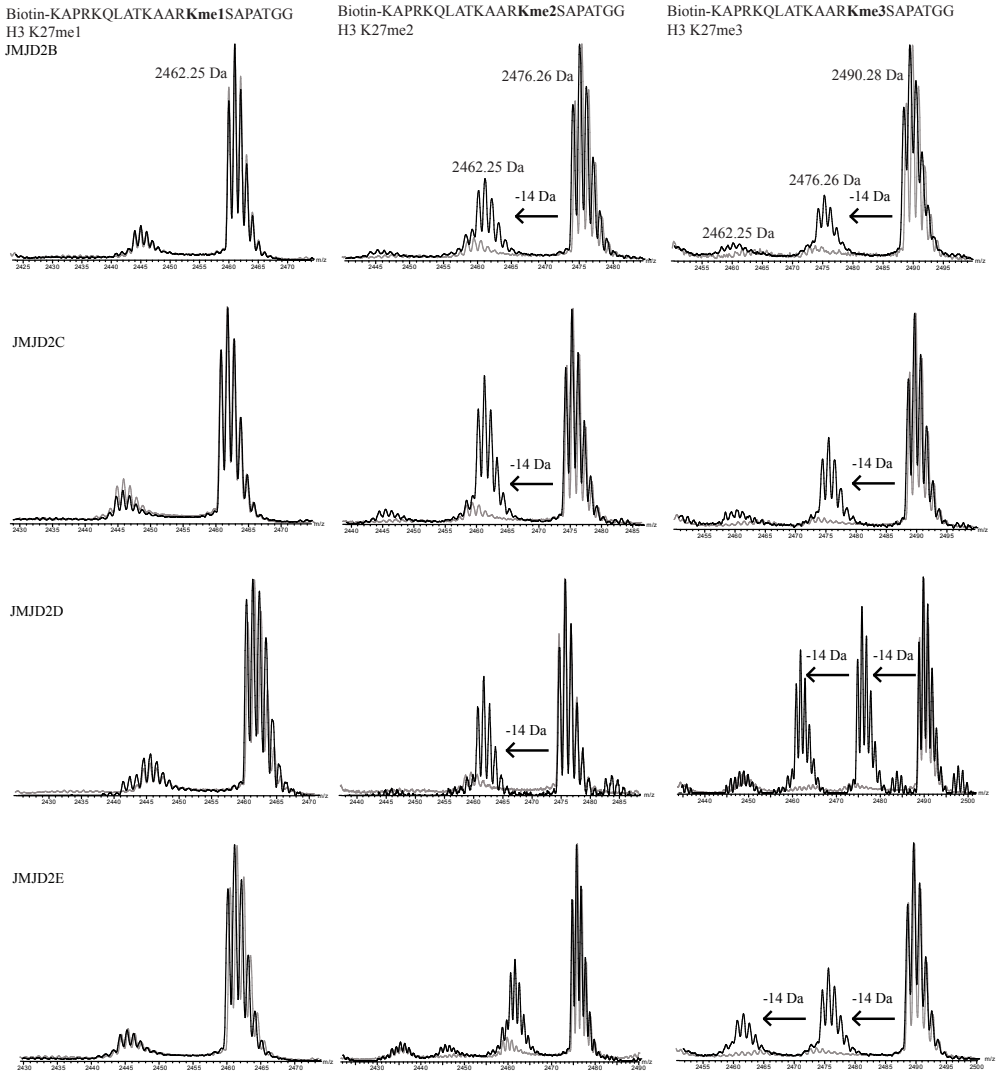


Fig. 3.4.1: End point MALDI for all members of the KDM4 subfamily with the mono-, di- and tri-methylated H3K27 peptide. Interestingly, KDM4D and E, which do not demethylate on K36 are active at K27.

Interestingly, there is a marked difference in catalytic turnover for the KDM4A and E. KDM4A shows greater overall demethylation with the shorter peptides, while KDM4E shows greater initial rates of demethylation on longer peptides, although the total demethylation

of the two enzymes is similar in the presence of longer peptides. Since the demethylation of K27 appears to be weaker than for K9 and K36, and that enzyme turnover is relatively inefficient, maximum levels of catalysis will be required to characterise kinetic parameters. Comparing activity between K9 (Fig. 3.2.4E) and K27 shows K27 to require longer incubation periods to reach comparative levels of demethylation. Maximum demethylation occurs at K9 within 10 minutes, while demethylation at K27 is still not saturated at comparable length peptides after 2 hrs, suggesting either i) the K27 peptide binding is low, or ii) although the peptide binds in the active site, but in a non-catalytically active confirmation. This is discussed further in Sections 3.4.2 and 3.6.

As longer peptides showed greatest demethylation activity, the 21-mer amino acids peptides were employed in further studies of K27 demethylation (Ala14-Ala35, as used in the screen from the Alta Bioscience library, although synthesised without the *N*-terminal biotin tag).

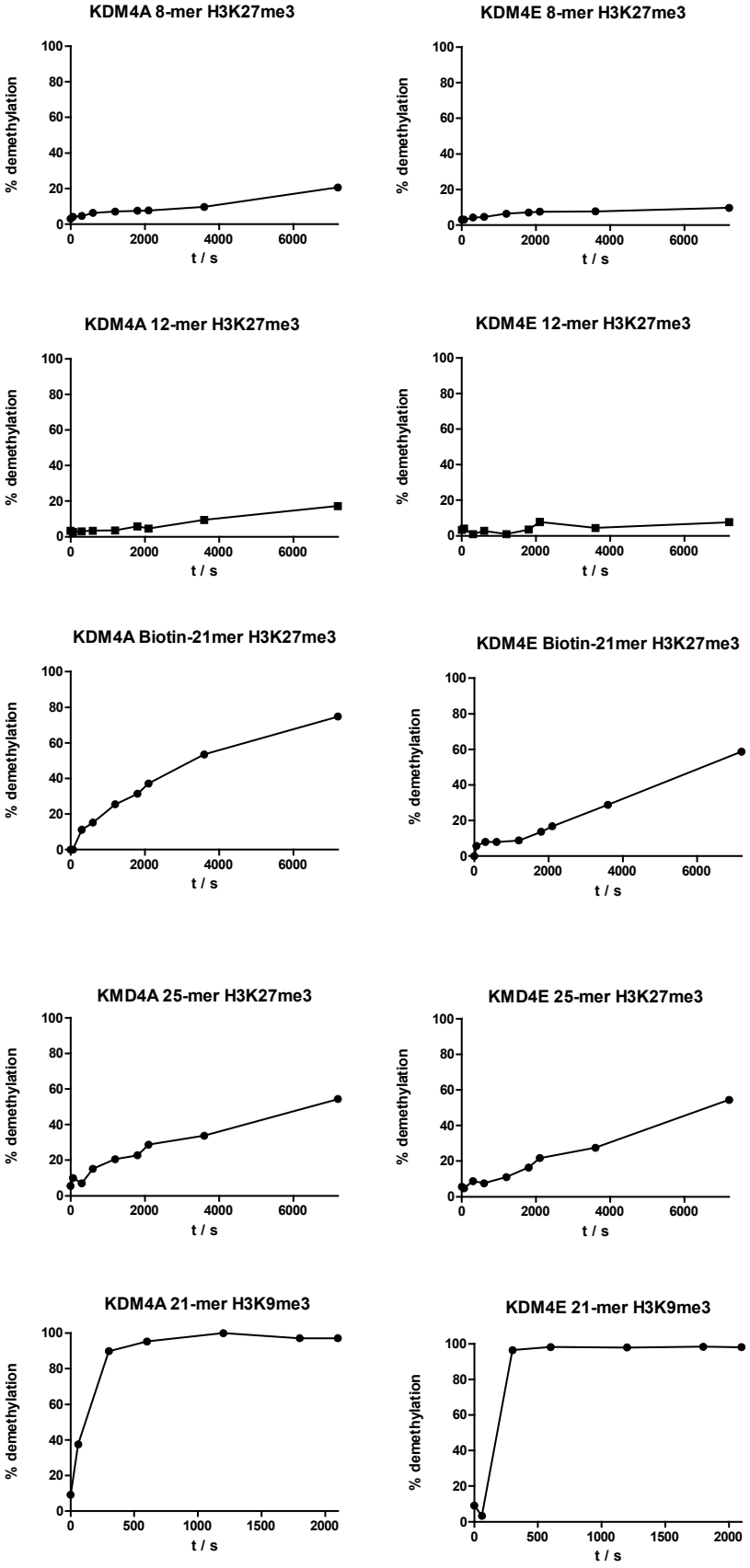


Fig. 3.4.2: Demethylation time courses of KDM4A and 4E on H3K27me3 peptides.

Peptide sequences:
8-mer AAR-Kme3-SAPATG,
12-mer ATKAAR-Kme3-SAPAT,
Biotin-21mer APRKQLATKAAR-Kme3-SAPATGGV,
25-mer STGGKAPRKQLATKAAR-Kme3-SAPATGGV,
K9me3, ARTKQYAR-Kme3-STGGKAPRKQLA.

Time course (0- 120 min) of KDM4A and KDM4E incubated in a ratio of 1:10 enzyme to peptide concentration as a 50 µL reaction (demethylation conditions as listed in Materials and Methods). For each time point, 5 µL was remove, quenched with 10 µL MeOH, and snap frozen in liquid nitrogen. Samples were stored at -80 ^{o}C until analysed by MALDI-TOF. Demethylation was calculated as a ratio of the sum of -14 and -28 Da peak intensity to the starting material intensity and plotted in GraphPad Prism 5.

3.4.1 Michaelis-Menten Enzyme Kinetics

In order to assess enzyme-substrate affinity and catalytic turnover for members of the KDM4 subfamily with the H3K27 sequence using Michaelis-Menten variables, demethylation assays were conducted with varying concentrations of peptide using either continuous reactions (using formaldehyde dehydrogenase, FDH coupled fluorescence-based monitoring)[116] or single-point reactions (stopped in the linear range of enzyme activity and monitored by MALDI).

The FDH assay utilises formaldehyde dehydrogenase from *Pseudomonas putida* to oxidise formaldehyde, produced as a bi-product of the demethylase reaction, to formic acid. This reaction is coupled to the reduction of nicotinamide adenine dicucleotide (NAD^+) to its reduced form, NADH.

$$E + S \underset{k_{-1}}{\overset{k_1}{\leftrightharpoons}} ES \xrightarrow{k_{cat}} E + P$$

$$v = \frac{d[P]}{dt} = \frac{v_{max}[S]}{k_m + [S]}$$

Equations 3.1, 3.2: 1) General reaction scheme for enzyme kinetics. The Michaelis-Menten constant k_m is related to these parameters by [$k_m = (k_{-1} + k_{cat})/k_1$], and is a measure for the affinity between enzyme and substrate, while the turnover number k_{cat} is generally representative of the rate of the rate-determining step. 2) The Michaelis-Menten equation relating substrate concentration to initial velocity under steady-state conditions. k_m = Michaelis-Menten constant, V_{max} = maximal velocity.

The production of NADH is monitored by fluorescence spectroscopy, with excitation and emission wavelengths at 355 nm and 460 nm respectively. Since the production of NADH is directly proportional to formaldehyde production by demethylation, demethylase activity can be monitored indirectly through the production of NADH (Fig. 3.4.3).

The benefit of using FDH as compared to end-point reactions is that real time data can be processed without further analysis. Furthermore, the number of total reactions in each experiment can be greatly increased over individual, end-point reactions.

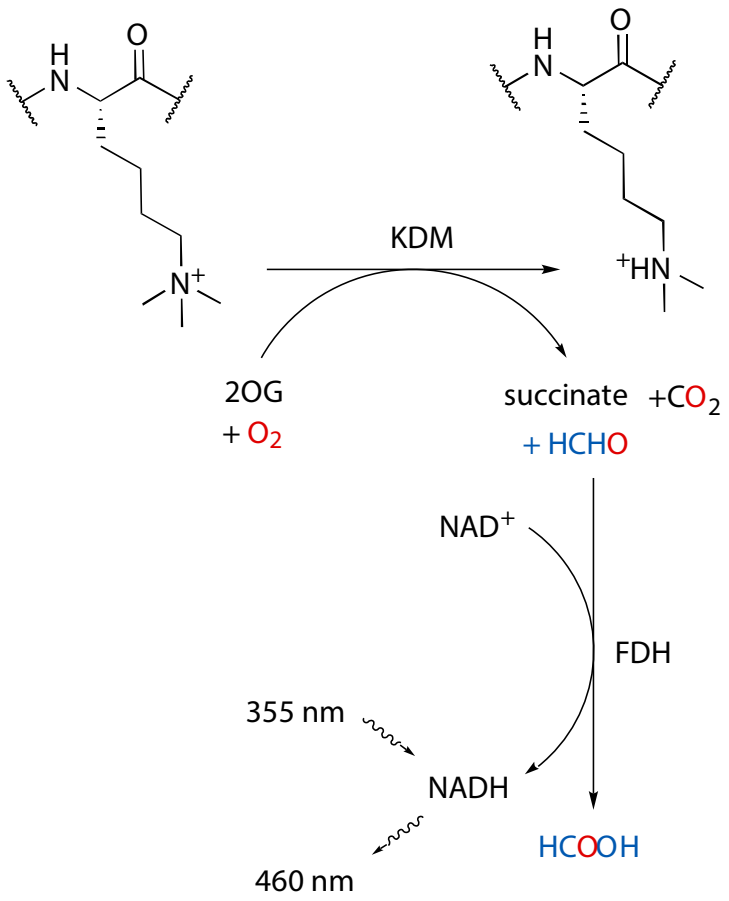


Fig. 3.4.3: The reactions underlying FDH monitoring of demethylase activity. Formaldehyde is oxidised to formic acid, which is coupled to the reduction of NAD+ to NADH, which is fluorescent, and is monitored as a surrogate of demethylation.

FDH Assay Optimisation

The assay conditions for FDH assays originally described by Couture *et al.* have been adapted for the KDM4 enzymes by Dr A. Kawamura.[116] KDM4A and E were purified using the constructs and purification methods described in Materials and Methods. FDH was purchased from Sigma as a lyophilized powder (3 U/mg) and used without further purification. Stock concentrations of FDH were frozen at 1 U/ 100 µL in FDH buffer (50mM HEPES, 0.01% TWEEN, total volume 50 µL) and thawed on use. NAD^+ was used at 500 µM final concentration. The Fe(II), 2OG and ascorbate concentrations were used as stated in the demethylation assays, although the enzyme concentration was increased to 2 µM final to maximise the signal of formaldehyde production above background (as utilised in publications reporting inhibition data of KDM4 enzymes using FDH[117]). The total volume of each reaction was 25 µL.

In this study, a new BMG Labtech© Pherastar plate reader was used for FDH assay kinetics, for increased sensitivity and accuracy. The instrument includes two needle injection pumps, allowing the dispensation of reagents directly into the plate while measurements are recorded. This plate reader has a larger signal threshold and intensity window than previously used instruments (Perkin Elmer© Envison plate reader), and so the protocol was optimized to reflect this.

Parameters relating to instrument settings and reaction protocol were tested to: i) reduce background noise, ii) increase enzyme signal above background and iii) limit air bubbles and other disturbances within the reaction well (since these interfered with the fluorescence readings).

Formaldehyde /µM	Log10 [F]	AFU
0	-	38771.67
0.076	-0.16	42765.00
0.23	0.31	55447.33
0.69	0.79	72694.67
2.06	1.27	94949.00
6.18	1.74	111239.67
18.519	2.22	117519.67
55.56	2.70	122798.33
166.67	-0.50	38771.67
500	-0.16	42765.00

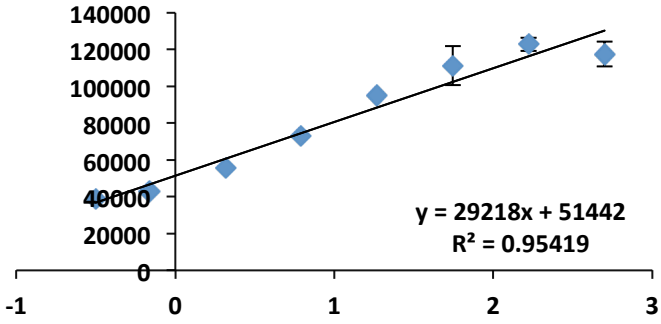


Table 3.4.1: Calibration concentrations for formaldehyde and AFU measured on the BMG Labtech© Pherastar plate reader. **Figure 3.4.4:** Calibration curve and y = mx + c equation fit (Excel), y = 29218x + 51442.

A calibration curve was required to correlate the levels of fluorescence with formaldehyde production, in order to convert the Arbitrary Fluorescence Units (AFU), which varies between machine and filter manufacturer, to formaldehyde concentration in µM. FDH and NAD^+ were mixed prior to the addition of formaldehyde.

Formaldehyde was titrated across the plate (concentrations listed in Table 3.4.1) and the fluorescence signal measured every 30 seconds across a period of 30 minutes. The difference between baseline fluorescence (no-formaldehyde control) and maximum fluorescence (highest-formaldehyde concentration) was plotted against formaldehyde concentration to generate a calibration curve (Fig. 3.4.4).

To ensure the signal of fluorescence remained within the detection limits of the instrument, test assays were run to optimise the gain setting (the window of detection) for KDM4 FDH. A setting of 1834 nm was found to give both the greatest window of detection without reducing the fluorescence signal.

Initial test reactions were run to optimise the injection protocol, ensuring that the reagents were mixed within the space of 30 sec prior to detection, and with as few disturbances to the reaction volume (bubbles/menisci etc.). The protocols tested included i) manual pipetting using multi-channel pipette, ii) enzyme-mix auto injection and iii) enzyme-mix and substrate-mix auto injection (as summarised in Table 3.4.2).

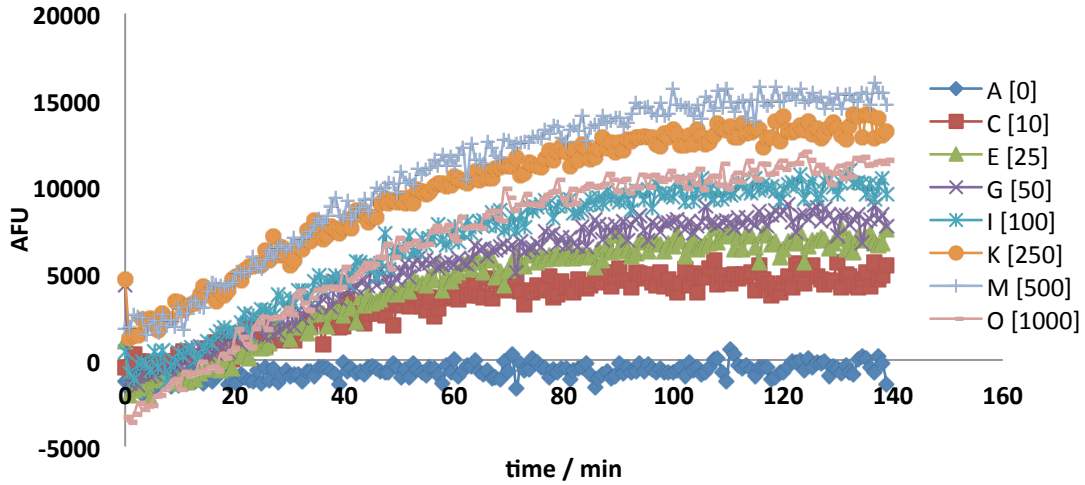


Fig. 3.4.5: Concentration dependent formaldehyde production of KDM4A with pepF (8-mer K9me3), as measured by FDH. No-enzyme control signal was averaged and deducted from the averaged reaction signal. The linear range extends to around 30 min.

Different peptide concentrations were all performed as triplicate repeats within the same plate. The peptide dilutions were made in a 96-well plate and transferred to a fluorescence assay plate (black, clear bottom 384 well Greiner© square-welled plate). A substrate mix of Fe(II), 2OG, ascorbate and NAD+ and enzyme mix of FDH and demethylase were mixed individually and aliquoted into separate lanes of a 96 well plate. Prior to mixing the reagents, the plate reader program was adjusted to either mix and read fluorescence or inject, mix and read fluorescence. These settings were optimised, as discussed below.

The K9me3 8-mer peptide (AR**Kme3**STGGK) was used in initial reactions. This is a well-characterised substrate for both KDM4A and E. It is easily synthesised and is more cost-effective to use in assay optimisation than longer, more expensive peptides. Fig. 3.4.5 illustrates how the fluorescence in AFU varies with time by peptide concentration of 8-mer H3K9me3 for KDM4A. Triplicate values are averaged and the average background for each concentration subtracted.

The injectors require a dead-volume of at least 1 mL and the lines needed to be primed with mixtures prior to dispensing liquid. Speed of injection was also varied in assays 2,3 and 4. Liquid dispensing speed did not greatly affect accuracy or replicate reproducibility, and so the higher speed was chosen to allow for faster injection times, close to the speed of reading fluorescence between samples.

Reagent mix	Manual pipetting	Injector 1	Injector 2	Shake time/ s	Read interval / s	Assay Identifier
Enzyme	✓	×	×	15	30	1
Substrate	✓	×	×	15	30	
Enzyme	×	✓	×	15	30	2
Substrate	✓	×	×	15	30	
Enzyme	×	×	✓	15	30	3
Substrate	✓	×	×	15	30	
Enzyme	×	×	✓	15	30	4
Substrate	×	✓	×	15	30	

Table 3.4.2: Reaction conditions tested for FDH assays for KDM4A with 8-mer, K9me3 peptide (PepF). Injection sample speed was also varied between 150 and 300 µL / s in assays 2,3 and 4. No difference was seen in efficiency of mixing between slower and faster speeds. 300 µL / s was used in further reactions.

It was found that manual pipetting (assay condition 1) was optimal for ensuring reproducibility between triplicates. While using the auto injectors reduced the time between enzyme-substrate mixing and the initial readings, the liquids were either too viscous (with high enzyme concentrations and detergent) to be dispensed accurately or mixed in the reaction wells sufficiently, or bubbles were introduced in some wells in a non-predictable manner.

MALDI-TOF MS Kinetic Analysis

The MALDI-TOF MS kinetics reactions utilised the same demethylation conditions as for the FDH assays, although replacing the reaction buffer with only 50 mM HEPES (pH 7.5) and the detection reagents with buffer. Reaction volumes were also scaled down to 10 µL (from 25 µL) since MALDI-TOF MS detection is more sensitive at lower peptide concentrations, relative to the FDH assay. This also reduced the overall amount of peptide and enzyme required.

The linear range of demethylation (the initial rate phase) was determined for the highest concentration of peptide (1000 µM) and applied to the other concentrations. This was performed for each peptide substrate.

The peptide dilutions were made in a 96-well plate and transferred to a second plate for the demethylation reaction. The substrate mix of 2OG and buffer (4 µL) was mixed with the peptide (1 µL). The enzyme mix of ascorbate, Fe(II), 2 µM enzyme and buffer was added to start the reaction.

3.4.2 KDM4A and KDM4E Kinetic Analysis of Demethylation at K27 and K4

Using the protocol described above, the kinetic parameters for demethylation on longer peptides were tested. Histone peptides of 21 amino acids, corresponding to sequences around tri-methyl H3K4, K9, K27 and K36 were synthesised for kinetics studies (CSBio using Solid Phase Peptide Synthesis (SPPS), Materials and Methods and peptide library, A.3). All amino acids were supplied as α-amine Fmoc protected amino acids. Standard amino acids

Enzyme(construct)	Substrate	k_m µM	V_{max} s^{-1}	k_{cat}/k_m $(\mu M.s)^{-1}$
KDM4A (1-359)	H3K4 (1-21aa)	-	-	-
	H3K9 (1-21aa)	15.89±7.7	1.31±0.12	0.041
	H3K27 (10-31aa)	127.9±37.4	0.97±0.09	0.00035
	H3K36 (23-44 aa)	17.94±3.14	1.99±0.07	0.012
KDM4E (1-337)	H3K4 (1-21aa)	-	-	-
	H3K9 (1-16aa)	14.8±3.9 *	-	0.0045 *
	H3K9 (1-25aa)	15.89±7.7 *	-	0.0937 *
	H3K27 (10-31aa)	75.15±31.24		0.0054

Table 3.4.3: Kinetic parameters determined using FDH for KDM4A and 4E with 21-mer, tri-methylated K4, K9, K27 and K36. Asterisk * identifies values calculated by Dr N. Rose. Since no demethylation was observed at K79, this mark was not included. GraphPad Prism 5 software was sued for the k_m determination.

(the 21-naturally occurring amino acids) were supplied from CSBio, while the methylated lysines were purchased from Novabiochem.

The catalytic efficiency of each substrate was compared, in the first instance using FDH and KDM4A. The H3K9 21-mer peptide was the best substrate (k_m 15.89 ± 7.7 µM, k_{cat}/k_m = 0.0414 $(\mu M.s)^{-1}$ see kinetics table, Table 3.4.3), followed by H3K36 (k_m 17.94 ± 3.14 µM, k_{cat}/k_m = 0.012 $(\mu M.s)^{-1}$), which is consistent with published results.[116,79] The H3K9 and K36 peptides are both better substrates compared to K27 (K_m 127.9 ± 37 µM, k_{cat}/k_m = 0.00035 $(\mu M.s)^{-1}$). This indicates that KDM4A has a stronger affinity and higher catalytic efficiency at the H3K9 and K36 tri-methyl peptides in comparison to K27.

For H3K4, the demethylation signal was not clearly discernible above background levels of fluorescence from the peptide and KDM4A and FDH enzymes. While enzyme dependent formaldehyde production was observed (above background levels) for higher peptide concentrations, a K_m value could not be determined.

In order to maximise demethylation signal, single end-point reactions analysed by MALDI-TOF MS detection were used for kinetic characterisation of KDM4A demethylation at H3K4me3 (Fig. 3.4.6). Both enzyme and peptide concentration were varied, to maximise potential demethylation activity.

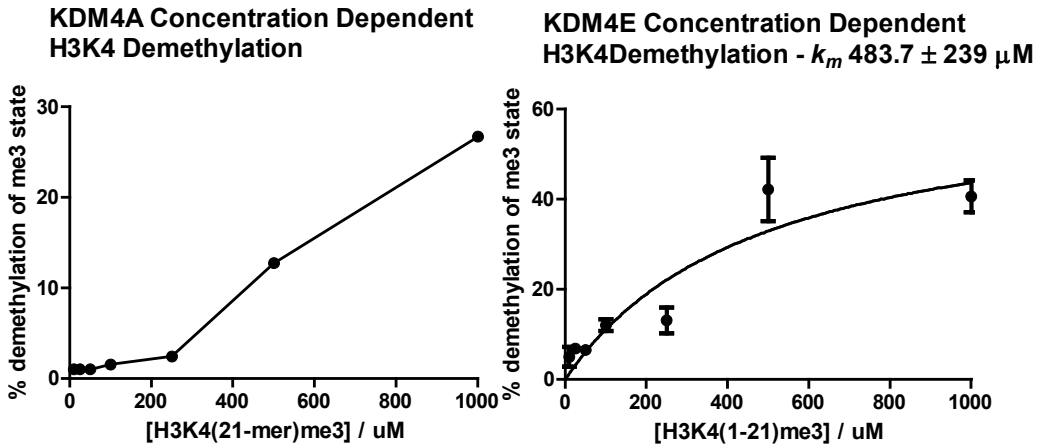


Fig. 3.4.6, 7: Kinetic analyses of KDM4A and 4E with 21-mer H3K4me3 peptide.

For KDM4E, previous work by Dr N. Rose[82] identified kinetic parameters for several H3K9 tri-methyl peptide lengths (Table 3.1.1). In FDH assays conducted with KDM4E using the 21-mer peptide and conditions outlined above, the time span of the linear range was difficult to determine since formaldehyde production showed fast initial rates, with a long-tail off to a low saturation point. This gave large error bars for low peptide concentrations, and indeterminate kinetic parameters. It was decided to continue kinetic evaluation of KDM4E demethylation at H3K4 and K27 by endpoint kinetics. The enzyme activity for lower peptide concentrations was still markedly low, under linear conditions, when assayed as individual time points.

Such poor kinetic data suggests the H3K4 mark is, if at all, a very poor substrate for KDM4A and E, and most likely the other KDM4 demethylases.

3.4.3 Competition Experiments of KDM4A Demethylation at K4, K9, K27 and K36

To test the preferential substrate activity of KDM4A towards K4, K9 and K36 over K27, competition experiments were undertaken, monitoring demethylation of two peptides simultaneously in the same reaction mix.

Equal concentrations of H3K4, K27, K36 21-mer peptides, and the Alta bioscience H3K9 (to distinguish between the masses of the K4 and K9 peptide) were used in a 1:10 ratio of enzyme to peptide and monitored over a course of 0-30 and 0-120 min (Fig. 3.4.8).

The competition experiments confirm results from the kinetic evaluation, regarding substrate preference. The K9 and K36 marks were demethylated to the mono-methyl lysine state before demethylation of K27 occurs (Fig. 3.4.8 A, B). Interestingly, in the presence of K9, K27 and K36 peptides, no K4 demethylation was observed (Fig. 3.4.8 C, D, E). The H3K4 peptide is likely outcompeted for binding at the active site when other modified lysines are available, even the mono-methyl state at K9, which is too weak a binder to undergo further demethylation under standard incubation
conditions.

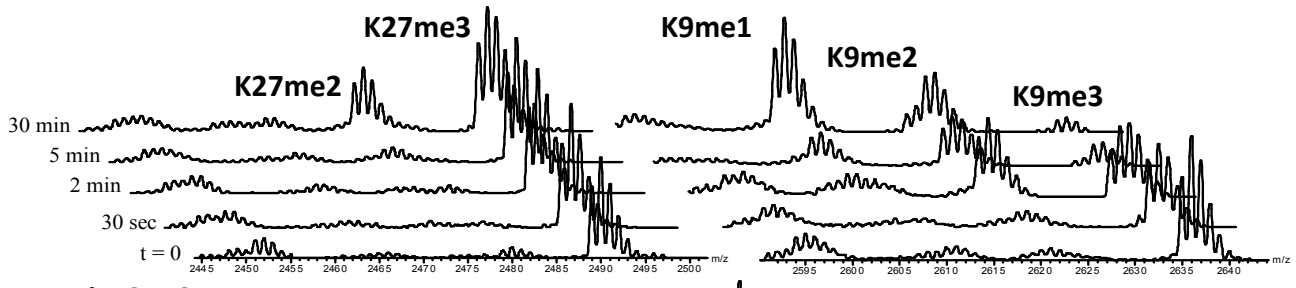


Fig. 3.4.8 A

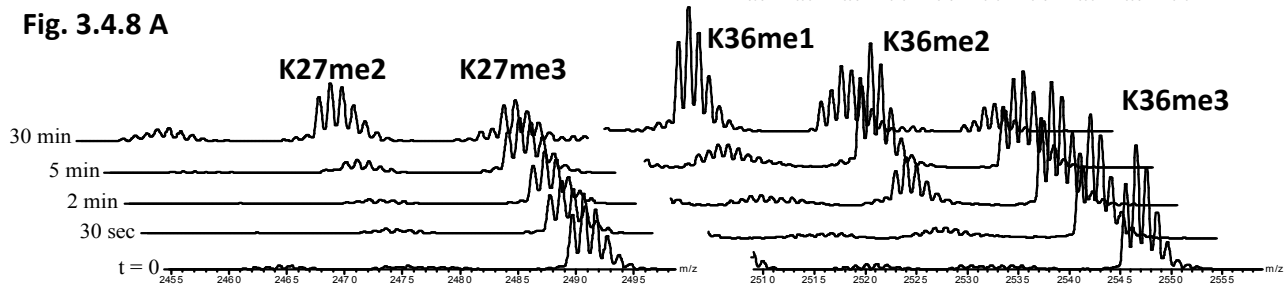


Fig. 3.4.8 B

Fig.3.4.8 A, B: Competition between demethylation at tri-methyl H3K27 versus K9 and K36. Peptides used: K9me3 Biotin-spacer- ARTKQTAR-Kme3-STGGKAPRKQLA, K27me3 Biotin-spacer- KAPRKQLATKAAR-Kme3-SAPATGG, K36me3 Biotin-spacer- SAPATGGV-Kme3-KPHRYRPGTVAL.

Fig.3.4.8 C,D,E: Competition between demethylation at tri-methyl HK4 versus K9 and K36. No demethylation at K4 was observed. Peptides used: K4me3 ARTKQTAR-Kme3-STGGKAPRKQLA, K9me3 Biotin-spacer-ARTKQTAR-Kme3-STGGKAPRKQLA, K27me3 Biotin-spacer-KAPRKQLATKAAR-Kme3-SAPATGG, K36me3 Biotin-spacer-SAPATGGV-Kme3-KPHRYRPGTVAL.

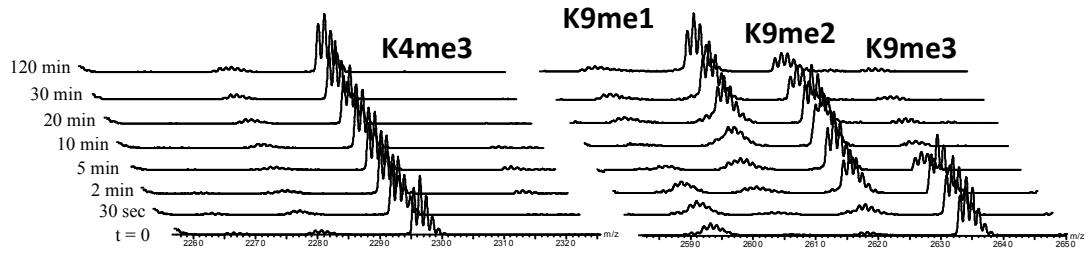


Fig. 3.4.8 C K4me3 21-mer K9me3 *N*-terminal biotin 21-mer

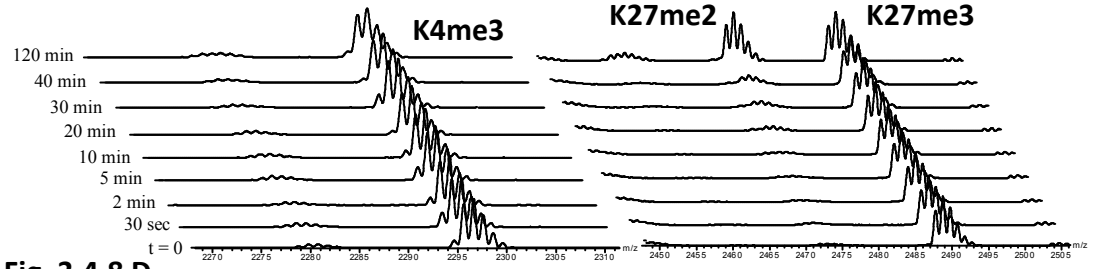


Fig. 3.4.8 D K4me3 21-mer K27me3 *N*-terminal biotin 21-mer

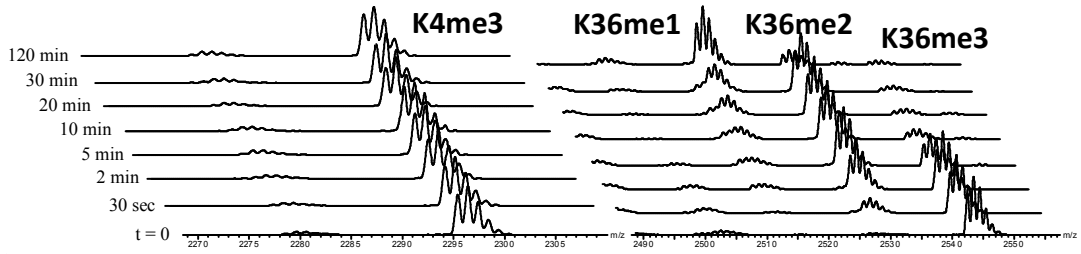


Fig. 3.4.8 E K4me3 21-mer K36me3 *N*-terminal biotin 21-mer

3.5 Investigating the Biological Relevance of KDM4A Demethylation at H3K4 and K27

The calculated kinetic parameters for KDM4A activity at H3K27 indicate it to be a less active substrate or weaker binder than previously characterised substrates, H3K9 and K36.

Much information is available about the alteration of global methylation states at K9me3 and K36me3 by KDM4A (Introduction), although little is known about the phenotype associated with alternative substrates.

Experiments were devised to test what effect KDM4A has on tri-methylation levels at H3K4 and K27 in more biologically relevant scenarios. The first experiments looked at whether the presence of alternative domains (JmjN. PHD and double Tudor domains, see Fig.3.4.9) within KDM4A narrowed the substrate selectivity of the JmjC domain *in vitro*, since the isolated catalytic domain, as used in this screen could, show potentially greater catalytic promiscuity in the absence of these regulatory domains.

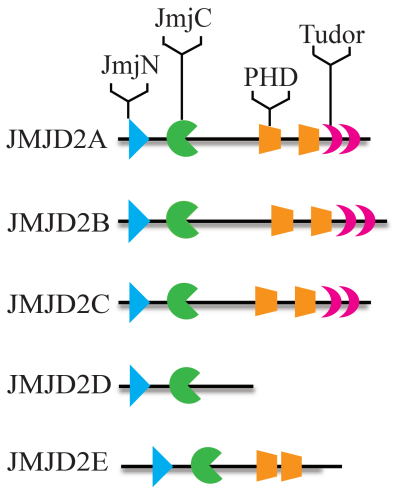


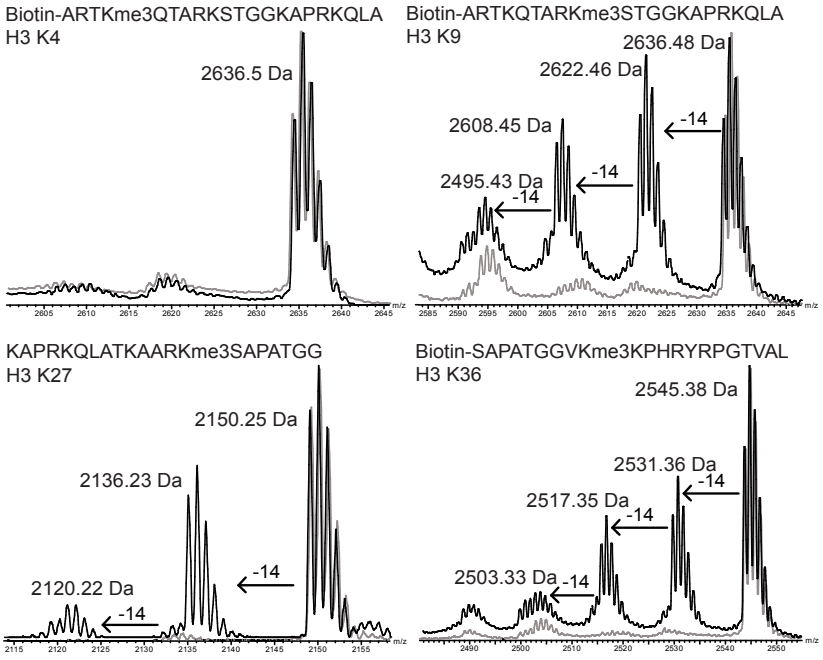
Fig.3.4.9: Domain structures for full-length KDM4 enzymes, as illustrated by Uniprot,

The second experiments looked at the effect of transiently over-expressing KDM4A in mammalian cells (HeLa) on global methylation levels at H3K4, K9, K27 and K36. This would test whether, in the presence of interaction partners or other biological apparatus, high, localised concentrations of KDM4A would be able to alter the global methylation status of chromatin at K27 and K4, as seen in previous experiments for K9 and K36.

Demethylation Activity by Full Length KDM4A

Full length (FL) KDM4A was cloned into a pcDNA3-FLAGx3 vector (construct from Prof. Yi Zhang). To transiently express FL, FLAGx3-fusion KDM4A, Human Embryonic Kidney cells 293T (HEK293T) were grown to 80% confluency in 15 cm^2 cell-culture dishes and transfected with FL-KDM4A DNA for 24 hrs, washed, scraped and lysed for further purification (Materials and Methods). The presence of the FLAG affinity tag allowed for purification of

FL-KDM4A from the cell lysate using anti-FLAG antibody coated magnetic beads, and samples of purified FL-KDM4A were stored at -80 ^{o}C until required for demethylation assays.



KDM4A	H3K4 %			H3K9 %			H3K27 %			H3K36 %		
	Me3	Me2	Me1	Me3	Me2	Me1	Me3	Me2	Me1	Me3	Me2	Me1
WT FL	×	×	×	74.27 ±10.00	51.22 ±3.05	×	45.62 ±3.60	28.41 ±3.20	×	58.67 ±4.67	31.18 ±8.34	×

Fig.3.4.10: Demethylation of FL KDM4A of Alta Bioscience peptides at H3K4, K9, K27 and K36 as analysed by MALDI.

To test demethylation of FL-KDM4A, demethylation assays were scaled-up to a total volume of 20 µL, maintaining all cofactor concentrations apart from the use of 2 µL of magnetic beads (with FL-KDM4A immobilised) rather than recombinant enzyme and the addition of protease inhibitor cocktail (at a 1 in 100 diltuion) to the 50 mM HEPES buffer. Reactions were incubated at 37 ^{o}C for 2 hrs, quenched with MeOH and analysed by MALDI-TOF MS.

Demethylation of K9, K27 and K36 by full length KDM4A was observed, but no demethylation at K4 (Fig. 3.4.10). The presence of alternative domains does not apparently restrict activity at K27, but does so at K4.

Demethylation Activity by KDM4A in HeLa Cells

To assess whether KDM4A demethylates at K27 tri-methyl lysine in cells, wild type and H188A catalytically inactive mutant KDM4A was transiently over-expressed in HeLa cells. Cells were plated in a 96 well plate and transfected with pcDNA3-FLAGx3 FL-KDM4A over a 24 hr period. Changes in global methylation levels where determined by antibody-detection based immuno-fluorescence as show in Fig.3.4.11 (undertaken with Dr Clarence Yapp, SGC, Materials and Methods). Antibodies for H3K4, K9, K27 and K36 tri-methylation states were used, along with an anti-FLAG antibody for identifying cells expressing wild type or mutant KDM4A and AlexaDAPI staining to identify nuclei. Secondary antibodies were used to assay changes in fluorescence.

Image acquisition and analysis of the fixed cells utilised the Beckton Dickinson pathway, an automated high-content imaging platform. The software isolates the DAPI stained nuclei, and correlates the nuclear region of the cell with the intensity of staining for the histone mark of interest. An average of all nuclear, histone staining intensity is then determined for all cells in a single well (a single assay point), and averaged again between replicates.

While the over-expressed wild type KDM4A did alter the methylation profile at K9 and K36, no comparable difference was observed at K27 or K4 for the wild type under the conditions tested (P values all indicated significant changes, reflecting the small error in each sample).

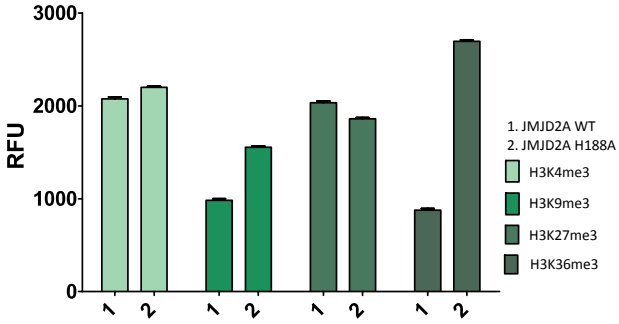


Fig.3.4.11 – IF staining of HeLa nuclei after expression of i) wild type or ii) H188A mutant KDM4A. Marks probed for included H3K4, K9, K27 and K36. Data was processed using the BD Pathway & GraphPad Prism 5.

This may indicate that KDM4A is directed away from K27 demethylation in cells, but since the change in signal for global demethylation is dependent upon the relative abundance of the mark and histone-modifying enzyme, the absence of a change of K27 methylation levels may only reflect the greater abundance of K27 tri-methyl lysine across the genome.

In addition, as demonstrated by *in vitro* competition experiments (Fig. 3.4.8), the demethylation of H3 sites may be sequential in cells. H3K27me3 demethylation may only take place after both H3K9 / H3K36 sites are demethylated, and time course over longer periods (2-7 days) may demonstrate altered H3K27 global methylation levels in these conditions.

3.6 Crystallographic Studies of KDM4A Bound to H3K27

Crystallographic studies (conducted by Dr Rashed Chowdury) suggest that KDM4A is able to bind with the K27 peptide (Fig 3.6.1) in a catalytically productive mode. The co-crystal structure with KDM2A and 25-mer H3K27me3 peptide shows a well defined tri-methyl lysine within the active site.

Poor resolution of the peptide backbone beyond the modified lysine and neighbouring residues is probably due to flexibility of the peptide, reflecting slow/weak binding as supported by low catalytic parameters for this substrate. The electron density for Arg26 and Ser28 side chains is not clear, although the Ser28 side chain seems to be pointing into the

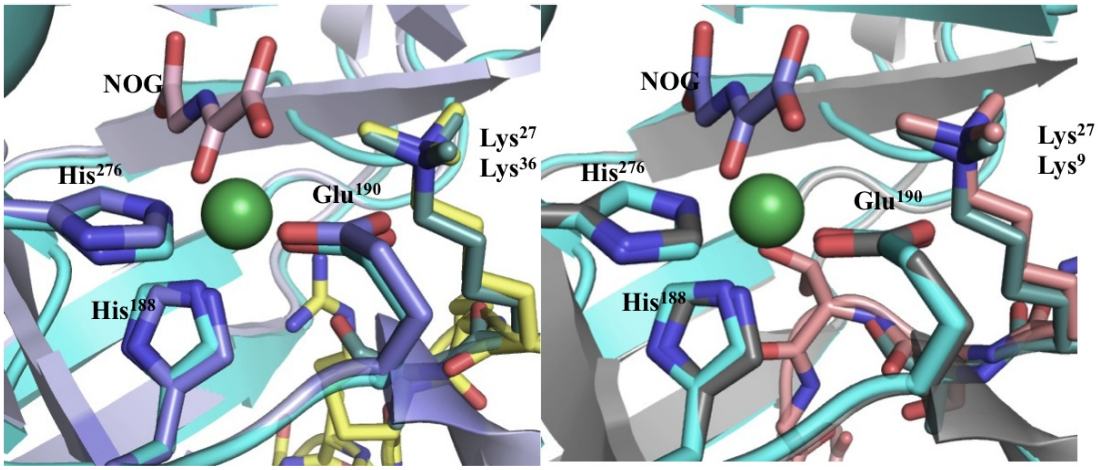


Fig. 3.6.1: KDM4A (2YBS, purple with K36me3 peptide, yellow) with KDM4A (cyan, with K27me3, teal); KDM4A (2OQ6, grey with K9me3 peptide, pink) with KDM4A (cyan, with K27me3, teal).

region where the K9 peptide extends, indicating the K27 peptide may extend into the open face of the KDM4A substrate binding pocket. However, at the methylated lysine, the K27 shows strong similarity to the positioning of K9 and K36 tri-methyl lysines (PDB 2OQ6, 2YBS).

3.7 Conclusions

The implications of the results are potentially wide ranging. A systematic screen of 2OG dependent JmjC-domain activity has, (i) confirmed site and state specific demethylation for 16 JmjC histone lysine demethylase enzymes on peptide substrates (ii) showed that some of these enzymes are more promiscuous than previously thought, and (iii) showed some enzymes with a JmjC domain to have no demethylase activity. These studies are necessary to validate the biochemical activities of JmjC-domain enzymes on peptide targets by a robust method, and to correlate this with information available in the literature.

Although the majority of the findings presented here have been undertaken *in vitro*, and the cell based assay reported here does not show any changes in global methylation states at either K4 or K27 under the conditions tested, it is at least possible that in specific biological contexts these alternative activities could function to regulate histone tail methylation. If KDM4 enzymes are involved in removing repressive K27 and K4 methylation states *in vivo*, it would be interesting to consider how these functions would relate to the gene-activating K9 and K36 demethylation by the KDM4 enzymes. Further biological studies should be undertaken to show whether KDM4 enzymes are involved in more complex feedback mechanisms between active or repressive methylation states. Thus, although activities in cells can differ from those of the isolated enzymes, the available evidence suggests that if an activity is observed with an isolated protein, it is likely to occur in cells (as seen with FIH and ankyrin hydroxylation). Therefore, it is important that enzymatic functions identified in cellular contexts using antibody detection should be supported by biochemical studies.

Chapter 4. Cyclic Peptide Inhibitors as Biochemical Tools

4.1. Development of Cyclic Peptide Chemical Probes

One key chemical approach for elucidating biological functions of enzymes is the design of synthetic molecules to modulate enzyme-activity on a specific target. These molecules, termed “chemical probes”, are designed to inhibit enzyme function with a high degree of potency and selectivity over related enzymes. In recently published studies of chemical probes, structure-based rational design has been important for ensuring these parameters.

In the field of epigenetics, chemical probes have been already utilised to identify the biological roles of histone deacetylases,[118,119,120] siturins,[121] histone acetyl-transferases,[122] histone methyl-transferases[123,124,125] and histone demethylases [126,127,128]

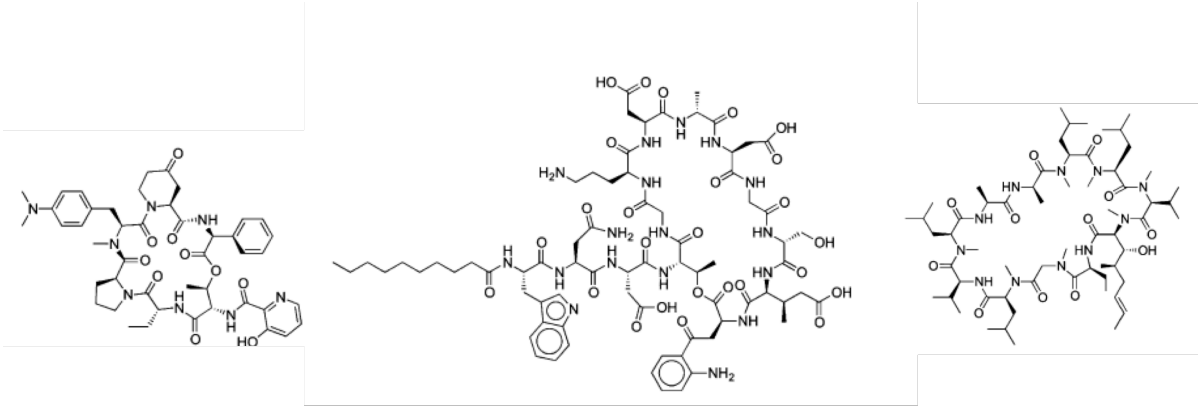
Chemical probes have also been designed for chromatin “reader” enzymes, including the BET-bromodomain enzymes (JQ1[129], PFI-1[130], PREBBP and EP300 Bromodomains (I-CBP112), and methylated-lysine reader domains (UNC1215[131]).

While some of these compounds have continued into clinical trials (HDACs), the primary role of chemical probes is not to direct phenotype in a predictable manner, but to elucidate the functions played by epigenetic enzymes in gene regulation. Chemical probes can play a crucial role in first validating potential drug targets, before investing in expensive clinical trials. The usefulness of chemical probes to the pharmaceutical industry has been demonstrated by their financial investment into collaborative projects involving chemical tool development.

One inherent issue with probe design for the KDMs is discriminating between enzymes within subfamilies. As discussed in the introduction, there are multiple approaches for rational drug design. However, in cases where enzymes share the same cofactor and substrate, their substrate binding site and catalytic centre will show strong homology, and structure-based design to discriminate between them is limited. For example, GSK-J1 shows

selectivity for KDM6 H3K27me3 demethylases over other subfamilies, but does not discriminate as effectively between JMJD3 and UTX.[127]

A novel class of chemical probes based on cyclic peptide have been designed to overcome this limitation of KDM subfamily redundancy (Kawamura 2013, under review). Peptide sequences were identified using the integration of the Flexible I*n vitro* Translation system (FIT) with Random non-standard Peptide Integration Discovery (RaPID).[132,133] Larger, macro-cyclic molecules are able to exploit more interactions with both the active site and substrate-binding site, compared to smaller molecules, and their chemical characteristic can be altered by including novel amino acids or epitope markers (for example cell permeability sequences).



Pristinamycin **Daptomycin** "LY146032" **Cyclosporin**

Fig. 4.1.1: Examples of clinically used cyclic peptides. Pristinamycin is an antibiotic which binds to the 50 S ribosomal subunit and inhibits peptide enlongation. Daptomycin disrupts bacterial cell membrane function in Gram-positive organisms. Cyclosporin is an immunosuppressant drug, used widely after organ transplantation to reduce host rejection.

Peptides have been used in a large variety of clinical applications, from antibiotics to cancer treatments (Fig. 4.1.1).[134,135] Many of these mimic cyclic peptide sequences present in nature, for example anti-bacterial cyclic peptides being derived from pathogen defence mechanisms. Peptide-based drugs are often limited in their efficacy because of peptidase-mediated degradation. However, the cyclised analogues, with no free terminus for endopeptidase activity and with potential stabilising secondary structure interactions between strands, can be more resistant.[136]

Cyclic peptides can be synthesised using Solid Phase Peptide Synthesis (SPPS) in conjunction with a number of cyclisation procedures.[137] However, this approach requires either a sequence to be known in advance. Cyclisation of substrate sequences is one route for inhibitor design, although limits the potential for maximising weakly interaction/distance amino acids in the active site not normally associated with substrate selectivity. Novel technologies are available for the screening of random peptide sequences against a desired protein target. The binding partner sequences generated can be chemically modified to tune potency and selectivity.

The RaPID methodology uses a codon-replaced mRNA template library, in which Met is exchanged for a modified amino acid (*N*-chloroacetyl-tyrosine) able to undergo auto peptide cyclisation (Fig. 4.1.2, 3). The mRNA fragments include an AUG start codon, random internal sequence and terminal UGC, to code for cysteine. The *in vitro* initiation tRNA is charged with an amino acid including a chloro-acetyl moiety and the *C*-terminal Cys acts as a nucleophile towards the chloro-acetyl group, forming a thioether-macrocyle.

In the selection process, as illustrated in Fig. 4.1.2, the mRNA-puromycin linked fragments can be translated to form peptide-mRNA conjugates and retro-transcribed to give the cDNA template (4, 5). The puromycin acts as a linker between the mRNA and growing peptide chain. The peptide fragments auto-cyclise in solution (4) and bind to the specific “bait” target (6). Non-specific sequences are washed off (7), and the mRNA used to retro-translate the cDNA sequences of binding partners (8). Repeated cycles of selection allow for the enrichment of peptide sequences selective for the target.

Once sufficient levels of cDNA have been recovered (compared to a no-input control), the cDNA sequences of the interaction partners are determined. The linear peptide analogues can then be synthesised on large scale using sold phase peptide synthesis and cyclised using choromethylcarbonyl-oxysuccinamide (ClAc-OSu).

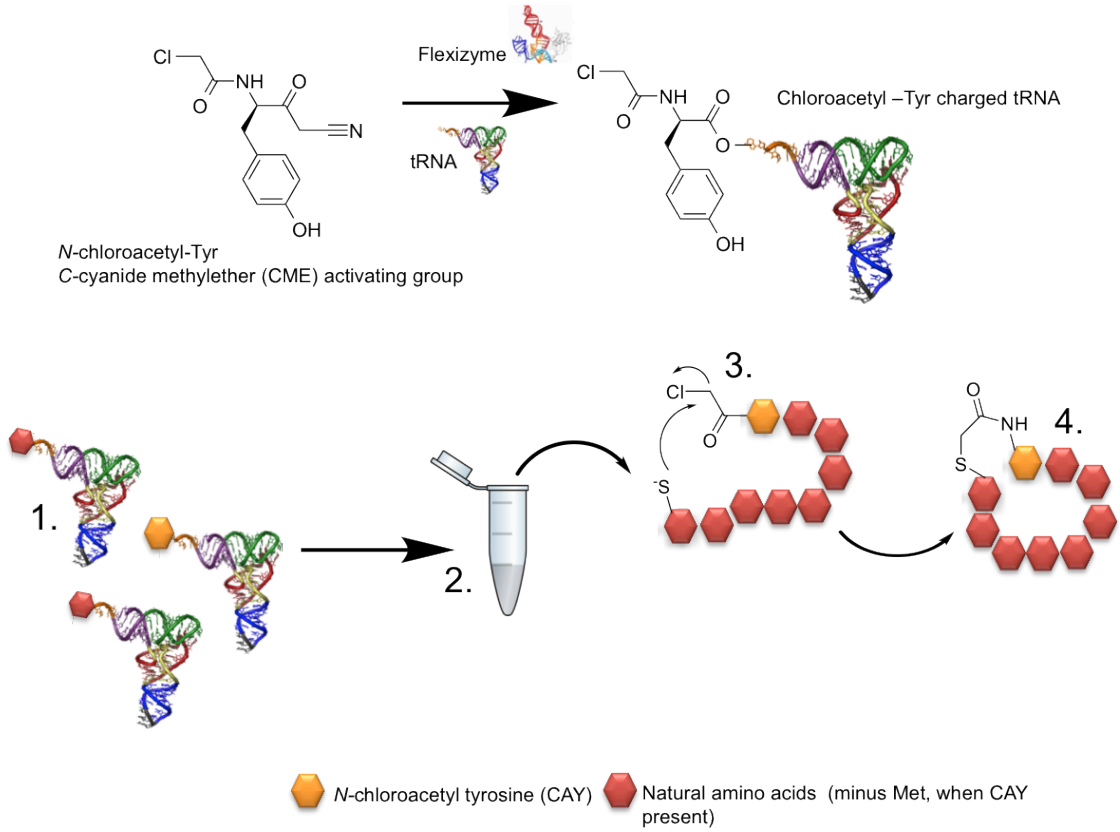


Fig. 4.1.2: Schematics to illustrate elements of the *in vitro* and RaPID expression systems. Top) FIT methodology uses a codon-replaced mRNA template library, in which Met is exchanged for a modified amino acid (N-chloroacetyl-tyrosine). The Flexizyme is used to catalyse the esterification of CAY to its tRNA conjugate. Below) A schematic to show *in vitro* translation of peptides including the chloro-acetyl Tyr (along with other natural amino acids) and the *C*-terminal Cys nucleophilic ring closure (3), forming a thioether-macrocyle (4).

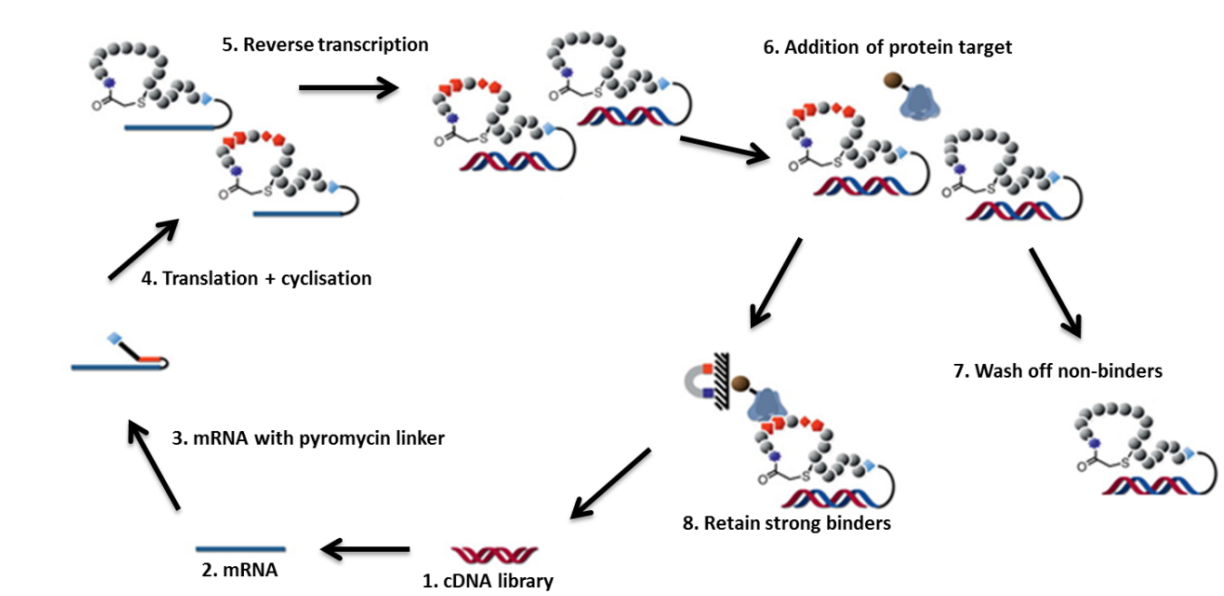


Fig. 4.1.3: *In vitro* cyclic peptide screening as used in the RaPID system. Adapted from Hipolito *et al.* 2012. (2) A cDNA library is translated to mRNA, (3) mRNA is ligated to a puromycin linker, (4, 5) mRNA-puromycin linked fragments are translated to form peptide-mRNA conjugates and retro-transcribed to give the cDNA template, (6) peptides are bound to a bound target protein, (7) non-specific sequences are washed off (8), strong binders are used as a cDNA library for the next selection round. Repeated cycles of selection allow for the enrichment of peptide sequences which bind strongly to the target protein.

4.2 *In vitro* Inhibitor Screening: mRNA display technology

The biological function and relevant substrates of many 2OG oxygenases are far from validated in cells. To complement the panel of available chemical probes for KDMs with peptide-based molecules, development of materials for an *in vitro* peptide-based inhibitor selection screen was undertaken.

The RaPID mRNA display system generates peptides that bind a chosen target protein. The target protein must be immobilised beads to "fish" interaction partners from a mixture of both binding and non-binding peptides. The non-specific interaction partners are removed by repeated rounds of stringent washing conditions. This technique requires target proteins to be expressed and purified to a sufficiently high level, and to be capable of binding the target protein under harsh conditions.

To express KDMs and 2OG oxygenases for the mRNA-display screen, a dual affinity-tag construct was used (pNIC-Bio3), including both an *N*-terminal His_6 and *C*-terminal biotinylation-site. Most of the enzymes included as potential targets have been cloned into and expressed adequately from Pet28(+)/PNIC28-Bsa vectors, which includes only an *N*-terminal His_6-tag. The second fusion sequence (Avi-tag) is recognised by Biotin-[acetyl-CoA-carboxylase] synthetase (BirA, co-expressed from pCDF-BIRA alongside the recombinant protein). On the addition of biotin to the growth media, BirA catalyses the transfer of an activated biotin-5'-adenylate to a lysine residue in the Avi-tag fusion sequence. This reaction is very efficient, and allows for complete biotinylation of recombinant protein. This His6-tag allows purification of the protein by Ni(II)-affinity chromatography, whereas the biotin group allows for strong binding of the recombinant protein to streptavidin-coated beads, considered more effective and specific than the equivalent His_6-tag, Ni(II) interaction, which is optimal for the mRNA-display peptide selection. The biotin-streptavadin interaction is stronger and more specific than the His_6-tag-Ni(II) interaction, and is optimal for mRNA selection.

In this study, six KDMs and the prolyl-hydroxylase PHD2 were cloned into PNIC-Bio3 (Table 4.2.1). All constructs were transformed into BL21(DE3) Rossetta codon-optimised *E. coli*.

Expression trials were conducted for selected enzymes (Table 4.2.1, Materials and Methods) to test for optimal biotinylation. Soluble protein expression was optimised for KDM4C, KDM4D, KDM2A and PHD2 using the following conditions: 2TY media, growth at 37 ^{o}C until OD_{600} = 6.5-0.8 and overnight expression at 18 ^{o}C while shaking at 180 rpm. ESI-MS was conducted on these four proteins to determine percentage of biotinylation. While KDM4C and D showed limited biotinylation and degradation (*), KDM2A and PHD2 were confirmed as 100% biotinylated.

Enzyme	pNIC-Bio3 Clone	BL21(DE3) Rosetta / pCDF-BirA Expression	Expression Trial Conducted	Mass Spectrometry Confirmation of Biotinylation
KDM4C	✓	✓	✓	✓
KDM4D	✓	✓	✓	✓
KDM6B	✓	✓	✓	
KDM6A	✓	✓		
PHF8	✓	✓	✓	
KDM2A	✓	✓		*
PHD2	✓	✓		*

Table 4.2.1: Constructs used in development of mRNA-display screen.

KDM2A and PHD2 were tested for whether the biotin-tag was exposed on the protein surface, and therefore able to bind to the streptavidin-immobilised beads. PHD2 and KDM2A were tested for their ability to bind to streptavadin-coated sensors, as a test for the proteins' affinity to streptavadin. Binding was monitored by Octet Rapid Binding Detection (RED) binding over a time course of 45 minutes, at varying protein concentrations (Materials and Methods). OctetRED is a modified Surface Plasmon Resonance (SPR) technique that allows for the quantification of binding to the surface of boated receptors. The association curves indicate that PHD2 and KDM2A bind to the sensors in a concentration dependent manner: Fig. 4.2.1.

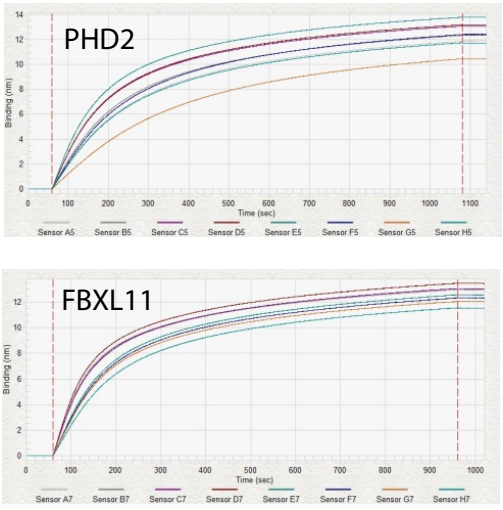


Fig. 4.2.1: OctetRED binding of His6-Biotin PHD2 and FBXL11 to biotin-coated sensors. Concentration dependent binding illustrates the biotin tag is exposed

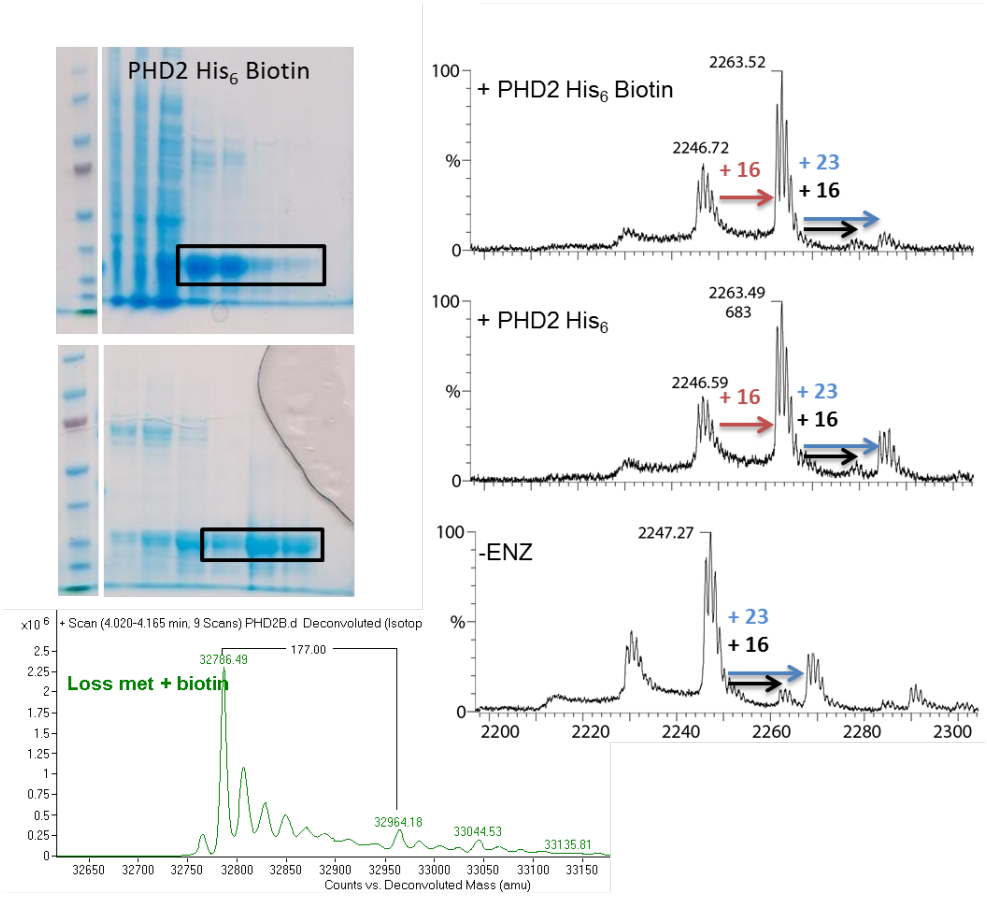


Fig. 4.2.2: Purification of His_6-PHD2 and His_6-Biotin-PHD2 Top: SDS PAGE gel of PHD2 His-Biotin using His-Trap Ni(II) purification (top) and gel filtration (Superdex 75, bottom). Fractions corresponding to lanes 4-6 were pooled and concentrated to 28.1 mg/ml (901 µM) were pooled and concentrated for use in in vitro mRNA display of cyclic peptide binding partners. LHS Bottom: ESI-MS confirmation of 100% biotinylation of purified His6-Biotin-PHD2. RHS: MALDI spectra indicate that the His6-Biotin-PHD2 construct is as active a hydroxylase as the standard His6-PHD2 construct.

To verify the correct folding of PHD2, despite the inclusion of the biotinylated tag, its enzyme activity was tested and compared to that of the His_6-tagged construct. Both His_6-PHD2 and His_6-Biotin-PHD2 were able to hydroxylate a HIF-1α CODD-substrate peptide over 30 min (Fig. 4.2.2), showing the presence of the biotin group does not affect substrate binding or catalytic activity.

The His_6-Biotin-PHD2 was then used by Dr A. Kawamura in a screen for mRNA-conjugate cyclic peptide-binders using the RaPID system. As described previously, the *N*-terminal Met was replaced with the non-natural, modified *N*-chloroacetyl-*L*-Tyr or *N*-chloroacetyl-*D*-Tyr ($ClAc^{L}Tyr$ and $ClAc^{D}Tyr$) amino acid. Each peptide included at least one Trp residue, which together with the modified $ClAc^{L/D}Tyr$ and terminal Cys, increased the UV auto-absorbance

of the peptides. Once isolated and purified, the peptides were dissolved in water and their concentrations verified using a spectrometer at 280 nm.

After 5 rounds of selection, an increase in percentage of recovery of cDNA from 0.0003 to 5.01% (as determined by RT-PCR quantification). The increased recovery reflects the convergence of peptides binding to the target protein.

A number of cyclic peptide hits were identified from this initial screen (Table 4.2.2). One motif common to sequences PHD2-001-005 is a serine residue 4 positions from the *C*-terminus, preceded by either a valine or leucine. Three of the hits include WDxxT and TFYLSR motifs. Since these sequences were all selected using PHD2, it is interesting to note that while 3 share 50% of the TFYLSR sequence, the other sequences are markedly different. This could reflect differences in binding location (catalytic site vs. allosteric site), or binding conformation, or that the motif may have been selected for by another bias in the screen.

ID	Sequence	Frequency
PHD2_001	^{D}YIWDTGTFYLSRTC	2
PHD2_002	^{D}YIWDTRTFYLSRTC	1
PHD2_003	^{D}YVWLTDTWVLSRTC	1
PHD2_004	^{D}YVWDPRTFYLSRIC	13
PHD2_005	^{D}YWNPLTIYVSRVC	1
PHD2_006	^{D}YGWPPIVSRTWPYC	1

Table 4.2.2 Peptide sequences identified by Dr Akane Kawamura.

It is also interesting to note that only three of the sequences (PHD2_004-6) included proline residues, given that PHD2 is a prolyl-hydroxylase.

If these peptides reflect peptide sequences which will bind to PHD2 in *in vitro*, then they may also reflect sequences able to interact with PHD2 *in vivo*. To investigate whether these sequences reflected the biological roles of PHD2, the sequences were searched against the human proteome (NBLAST, human database). Interesting hits are summarised in Table 4.2.3.

Sequence searched	Protein hit	Residues/sequence	E value	Function
TFYLSR	TRIM65	7-11	0.53	Ubiquitin protein ligase E3A
	Pyruvate Dehydrogenase Kinase	158-162	0.53	Antagonist to pyruvate dehydrogenase
	Protein unc-79 homolog	2079-83	0.56	NALCN Na^+ channel component
IWDTGTFYLSRT	ANKMY1	138-144	4.1	Ankyrin repeat MYND P1

Table 4.2.3: NICB BLAST for non-redundant protein sequences. Values with E < 5 reported for sequences listed in Table 4.2.2. Sequence searches removed *N*- and *C*- residues, since these were biased in the mRNA library design.

The "TFYLSR" motif in these peptides is also found in proteins that function in regulatory pathways involving PHD2. For example, Pyruvate Dehydrogenase Kinase-3 (PDK3) regulates the activity of Pyruvate Dehydrogensae Complex (PDC), a key metabolic complex responsible for catalysing decarboxylation of pyruvate to acetyl-CoA and NADH prior to entry into the TCA cycle. Four isoforms of PDKs have been identified to catalyse the phosphorylation of PDC, with PDK3 showing the strongest affinity and highest enzyme activity.[138] Recent work has identified elevated HIF-1α levels to induce PDK3 expression, although it is not known whether it is a direct HIF-1α target gene. Since PHD2 activity under normoxia reduces HIF-1α levels, a potential interaction with PDK3 may indicate a role for PHD2 in metabolism oxygen sensing.[139]

Cyclic peptides are considered a closer mimic of protein binding partners than the linear analogues. Cyclisation allows a more fixed structure, with potentially more interactions to a binding. It was of interest to determine whether performing the same selection process using linear peptides would yield similar binding-sequences.

To generate a linear peptide specific for PHD2, a peptide mRNA screen was conducted using His_6-Biotin-PHD2. The screen was identical to the cyclic peptide screen performed, except that it did not include the $ClAc^{L/D}$Tyr modified amino acid.

First, an mRNA library coding for peptide sequences between 8 and 12 residues long (libraries NNK8-NNK12) was incubated with T4RNA ligase and a puramycin-CCA-linker to

generate mRNA-puramycin conjugates. The mRNA was purified by chloroform extraction and assessed for purity (Fig. 4.2.3).

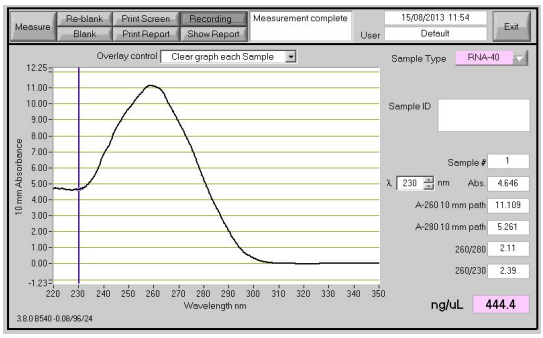


Fig. 4.2.3 – Assessing purity of puromycin-linked mRNA

The puromycin-mRNA was translated using an *in vitro* translation system (GE Healthcare) and incubated with His_6-Biotin-PHD2 immobilised streptavidin magnetic beads (Invitrogen). After washing (to remove non-specific binders) the immobilised beads were used as a template for reverse transcription of the bound linear peptide-puromycin-mRNA conjugates (with the His_6-Biotin-PHD2 and cyclic peptide binding partners still attached). The resulting cDNA was then quantified by Real Time PCR (RT-PCR, Fig. 4.2.4). The standard curve (3,4,5,6) shows the bound cyclic peptide concentration (7) to above 10^8, although lower than the input (1), illustrating a degree of selection from the mRNA pool.

The RT-PCR cDNA product was amplified by standard PCR for 15 cycles (to reduce polymerase introduced errors). An agarose gel shows a dingle band in each of the replicate PCR reactions, suggesting a degree of fidelity of the DNA polymerase. This cDNA was purified (Fig. 4.2.5) and used for mRNA transcription. After this stage, the mRNA has to be re-ligated with puromycin before the cycle can begin again.

Due to time constraints, the screen could not be completed, and was ended here. Nevertheless, after repeated rounds of selection have been completed, it will be interesting to compare the recovered sequences to those generated in the cyclic peptide screen.

Sample ID	Sample Name	Cт
1	No template (negative control)	27.78
2	Input (negative control)	-
3	10^5	22.50
4	10^6	18.45
5	10^7	14.93
6	10^8	11.31
7	Bound (sample)	9.97

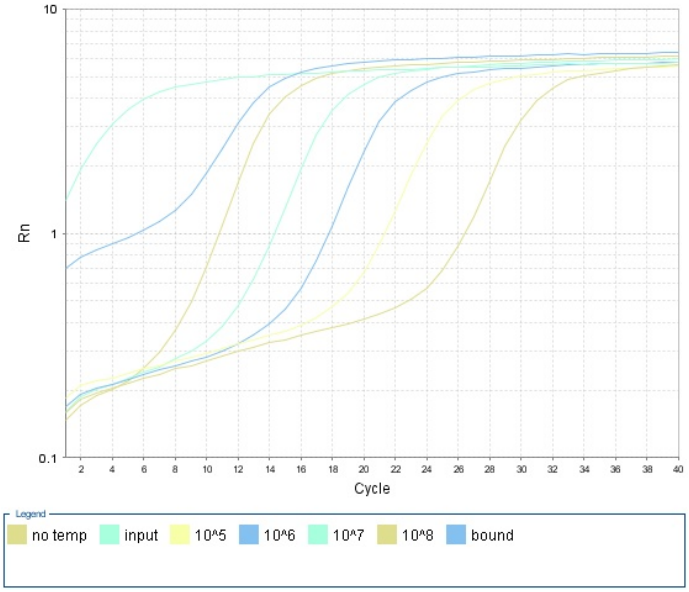


Fig. 4.2.: Determination of cDNA concentration bound to immobilised His_6-Biotin-PHD2 using RT-PCR. The melting curve mid-points are listed in the above table.

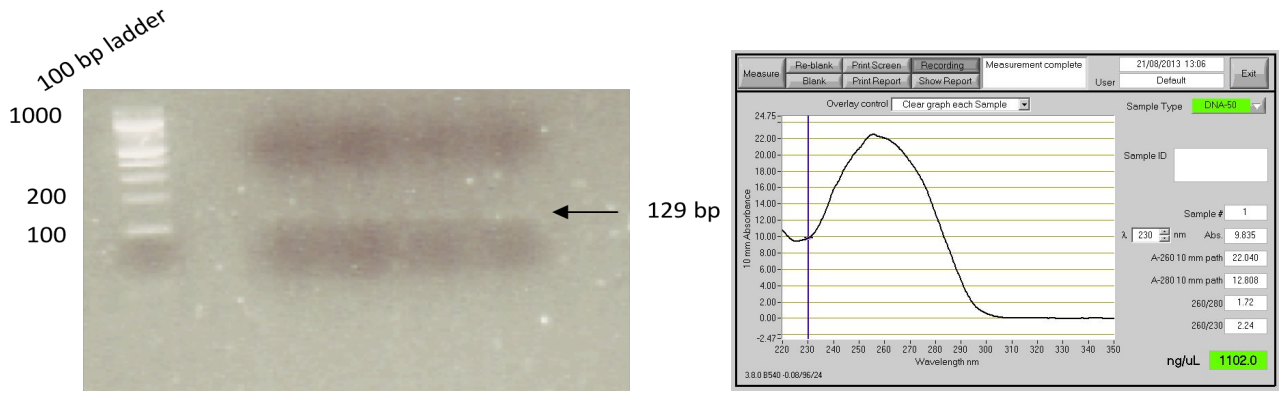


Fig. 4.2.5: i) Agarose gel of PCR amplification of cDNA from the bound cyclic peptides (100 bp ladder, band >100 bp, as indicated by the arrow) ii) DNA was purified and used for a second round of transcription.

Conclusions

The project objectives of this dissertation, as outlined in the Introduction, were to:

1. Conduct a broad screen to determine the histone methylation profile of 15 2OG oxygenases.
2. Work towards the development of an *in vitro* screen for cyclic and linear peptide chemical probes for KDMs.

In summary, I have achieved:

1. i) A screen of JmjC 2OG-oxygenase demethylase function, which led to the identification of unexpectedly greater substrate tolerance for KDM4s.

   ii) Detailed investigations of KDM4A demethylation at K4 and K27 tri-methyl lysines. This required kinetic analyses by FDH and MALDI based assays. The demethylation activity at methylated K4 by both enzymes was shown to be very limited, although activity at K27 was more substantial, suggesting this may be a previously unidentified substrate for these enzymes.

   iii) Expression of full length KDM4A and determination of the enzyme's demethylation profile (which was active at K9, K27 and K36, but not K4). Over expression of full length KMD4A in mammalian cells demonstrated KDM4A demethylation of H3K27 was not observed on a global level.

2. Design of recombinant materials for chemical probe design and initial screening for linear peptide selection.

In light of my results from Chapter 3:

i) A clear distinction in substrate preference should be recognised between the JmjC KDMs and hydroxylases – which included MINA53, NO66, JMJD6, JMJD5 and FIH. No demethylation activity was observed for any of these enzymes, while members of the KDMs showed specificity for both lysine position and methylation status.

ii) The substrate profile for KDM4A-C and D/E should be considered to be more promiscuous, with respect to the sequence surrounding demethylated lysines. Interestingly, even though KDM4A-C and D/E show differences in methylation at K36, both accept tri-methyl K27. Therefore, demethylation activity at K27 and K4 is potentially directed by sequence motifs common to those at K9 (the RK(S) motif). The greater similarity between K9 and K27, over K4, may explain the low levels of demethylation observed for the latter. The rate of demethylation at K27 was much lower than for other confirmed substrates (k_{cat}/k_m H3K9me3 0.041, K36me3 0.012, K27me3 0.00035 $(\mu M.s)^{-1}$). The KDMs are not very efficient enzymes, with k_{cat}/k_m values in the order of $(\mu M.s)^{-1}$. Kinetic parameters have indicated the K_c9me3 modified nucleosome[140] to be a better substrate than the equivalent, 8-mer peptide by only 37% (330 v.s. 240 μM) which is surprisingly low.[103] These activities may be limited by *in vitro* reaction conditions, e.g. the absence of interaction partners, but suggest only small relative increase in catalytic efficiency in the presence of an intact substrate. This may correlate with the biological functions of epigenetic modifying enzymes, which have to respond in a dynamic manner to the presence of chemical modifications and recruited complexes to histone protein tails. It may be proposed that if demethylation reactions were very efficient, the window to "tune" activity could be small.

iii) The cellular specific demethylation activity of the KDMs should be investigated further. Immunofluorescence data suggests over expression of KDM4A to not affect global H3K4 or H3K27 methylation after 24 hrs. However, as discussed in Chapter 3, this may relate to increased abundance of H3K27me3 vs. K9me3, or indicate KDM4A is only active towards K27 after total demethylation of K9 and K36. Although activities in cells can differ from those of the isolated enzymes, the available evidence suggests that if an activity is observed with an isolated protein, is it likely to occur in cells (as seen with FIH and Ankyrin hydroxylation).

The work carried out in Chapter 4 demonstrated that peptide sequences could be used to design chemical probes. The use of RaPID translation systems and standard *in vitro* translation in combination with mRNA-display generates target specific protein sequences which may modulate enzyme activity. These methods also give a (relatively) non-biased means of identifying interacting sequences, which can then be tested against sequences in databases of the human proteome for novel biological roles.

Several further lines of investigation should be followed on from this work.

1. Assessing whether K27/K4 demethylation by KDM4s occurs in cells using methods which monitor gene expression, such as RNA quantification and luciferase assays, focusing on genes associated with these specific marks (such as HOX genes, which are predominantly K27-trimethylated).
2. Determination of mode of binding to PHD2 by the cyclic peptide hits. Further work is being conducted to elucidate if these are sufficiently potent and selective PHD2 inhibitors, or whether they block binding of PHD2 with interaction partners.
3. Continuation of linear peptide screen against PHD2, to determine whether similar sequence motifs are generated.

References

1 Yue, W. W. *et al.* Crystal structure of the PHF8 Jumonji domain, an Nε-methyl lysine demethylase. *FEBS Letters* **584**, 825-830, doi:http://dx.doi.org/10.1016/j.febslet.2009.12.055 (2010).

2 Luger, K. Structure and dynamic behavior of nucleosomes. *Current opinion in genetics & development* **13**, 127-135 (2003).

3 Thatcher, T. H. & Gorovsky, M. A. Phylogenetic analysis of the core histones H2A, H2B, H3, and H4. *Nucleic Acids Research* **22**, 174-179, doi:10.1093/nar/22.2.174 (1994).

4 Hyland, E. M. *et al.* Insights into the Role of Histone H3 and Histone H4 Core Modifiable Residues in Saccharomyces cerevisiae. *Molecular and Cellular Biology* **25**, 10060-10070, doi:10.1128/mcb.25.22.10060-10070.2005 (2005).

5 Bannister, A. J., Schneider, R. & Kouzarides, T. Histone Methylation: Dynamic or Static? *Cell* **109**, 801-806, doi:http://dx.doi.org/10.1016/S0092-8674(02)00798-5 (2002).

6 Allis, C. D., Jenuwein, T., Reinberg, D. & Caparros, M.-L. Epigenetics. *Cold Spring Harbor Press* (2007).

7 Mitra, D., Parnell, E. J., Landon, J. W., Yu, Y. & Stillman, D. J. SWI/SNF binding to the HO promotor requires histone acetylation and stimulates TATA-binding protein recruitment. *Molecular and Cellular Biology* **26**, 4095-4110 (2006).

8 Sterner, D. E. & Berger, S. L. Acetylation of histones and transcription-related factors. *Microbiol Mol Biol Rev* **64**, 435-459 (2000).

9 Zhang, Y. & Reinberg, D. Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes & Development* **15**, 2343-2360 (2001).

10 Di Lorenzo, A. & Bedford, M. T. Histone arginine methylation. *FEBS Letters* **585**, 2024-2031 (2011).

11 Baba, A. *et al.* PKA-dependent regulation of the histone lysine demethylase complex PHF2-ARID5B. *Nat Cell Biol* **13**, 668-675.

12 Weake, V. M. & Workman, J. L. Histone Ubiquitination: Triggering Gene Activity. *Molecular Cell* **29**, 653-663, doi:http://dx.doi.org/10.1016/j.molcel.2008.02.014 (2008).

13 Ballard, T. D. *et al.* Biotinidase catalyzes debiotinylation of histones. *European Journal of Nutrition* **41**, 78-84 (2002).

14 Hassan, Y. I. & Zempleni, J. A novel, enigmatic histone modification: biotinylation of histones by holocarboxylase synthetase. *Nutrition Reviews* **66**, 721-725 (2008).

15 Berger, S., Kouzarides, T., Shiekhattar, R. & Shilatifard, A. An operational definition of epigenetics. *Genes & Development* **23**, 781-783 (2009).

16 Fischle, W., Wang, Y. & Allis, C. D. Histone and chromatin cross-talk. *Current Opinion in Cell Biology* **15**, 172-183, doi:http://dx.doi.org/10.1016/S0955-0674(03)00013-9 (2003).

17 Fischle, W. *et al.* Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* **438**, 1116-1122 (2005).

18 Stanley, J. S., Griffin, J. B. & Zempleni, J. Biotinylation of histones in human cells. Effects of cell proliferation. *European Journal of Biochemistry* **268**, 5424-5429 (2001).

19 Srivenugopal, K. S., Yuan, X. H., Friedman, H. S. & Ali-Osman, F. Ubiquitination-dependent proteolysis of O6-methylguanine-DNA methyltransferase in human and murine tumor cells following inactivation with O6-benzylguanine or 1,3-bis(2-chloroethyl)-1-nitrosourea. *Biochemistry* **35**, 1328-1334 (1996).
20 Tahiliani, M. *et al.* Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science* **324**, 930-935 (2009).
21 Murray, K. The Occurrence of iε-N-Methyl Lysine in Histones. *Biochemistry* **3**, 10-15 (1964).
22 Greer, E. L. & Shi, Y. Histone methylation: a dynamic mark in health, disease and inheritance. *Nature Reviews Genetics* **13**, 343-357 (2012).
23 Cedar, H. & Bergman, Y. Linking DNA methylation and histone modification: patterns and paradigms. *Nature Reviews Genetics* **10**, 295-304 (2009).
24 Kirmizis, A. *et al.* Arginine methylation at H3R2 controls deposition of H3K4 trimethylation. *Nature Letters* **449**, 928-934 (2007).
25 Li, X. *et al.* H4R3 methylation facilitates β-globin transcription by regulating histone acetyltransferase binding and H3 acetylation. *Blood* **115**, 2028-2037 (2010).
26 Walter, W. *et al.* 14-3-3 Interaction with Histone H3 Involves a Dual Modification Pattern of Phosphoacetylation. *Molecular and Cellular Biology* **28**, 2840-2849, doi:10.1128/mcb.01457-07 (2008).
27 Shi, Y. *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941-953 (2004).
28 Tsukada, Y. *et al.* Histone Demethylation by A Family of JmjC Domain-containing Proteins. *Nature* **439**, 811-816 (2006).
29 Loenarz, C. & Schofield, C. J. Physiological and biochemical aspects of hydroxylations and demethylations catalyzed by human 2-oxoglutarate oxygenases. *Trends in Biochemical Sciences* **36**, 7-18 (2011).
30 Ryle, M. J. & Hausinger, R. P. Non-heme iron oxygenases. *Current Opinion in Chemical Biology* **6**, 193-201, doi:http://dx.doi.org/10.1016/S1367-5931(02)00302-2 (2002).
31 Loenarz, C., Chowdhury, R., Schofield, C. J. & Flashman, E. Oxygenases for oxygen sensing. *Pure and Applied Chemistry* **80**, 1837-1847 (2008).
32 Bleijlevens, B. *et al.* Dynamic states of the DNA repair enzyme AlkB regulate product release. *Embo Reports* **9**, 872-877 (2008).
33 Delaney, J. C. *et al.* AlkB reverses etheno DNA lesions caused by lipid oxidation *in vitro* and *in vivo*. *Nature Structural and Molecular Biology* **12**, 855-860 (2005).
34 Loenarz, C. & Schofield, C. J. Oxygenase Catalyzed 5-Methylcytosine Hydroxylation. *Chemistry & Biology* **16**, 580-583 (2009).
35 Mukherji, M. *et al.* The chemical biology of branched-chain lipid metabolism. *Progress in Lipid Research* **42**, 359-376 (2003).
36 Walport, L. J., Hopkinson, R. J. & Schofield, C. J. Mechanisms of human histone and nucleic acid demethylases. *Curr Opin Chem Biol* **16**, 1-10 (2012).
37 Zhou, J., Gunsior, M., Bachmann, B. O., Townsend, C. A. & Solomon, E. I. Substrate Binding to the α-Ketoglutarate-Dependent Non-Heme Iron Enzyme Clavaminate Synthase 2: Coupling Mechanism of Oxidative Decarboxylation and Hydroxylation. *Journal of the American Chemical Society* **120**, 13539-13540, doi:10.1021/ja983534x (1998).

38 Zhou, J. *et al.* Spectroscopic Studies of Substrate Interactions with Clavaminate Synthase 2, a Multifunctional α-KG-Dependent Non-Heme Iron Enzyme: Correlation with Mechanisms and Reactivities. *Journal of the American Chemical Society* **123**, 7388-7398, doi:10.1021/ja004025+ (2001).

39 Ryle, M. J., Padmakumar, R. & Hausinger, R. P. Stopped-Flow Kinetic Analysis of *Escherichia coli* Taurine/a-Ketoglutarate Dioxygenase: Interactions with a-Ketoglutarate, Taurine and Oxygen. *Biochemistry* **38**, 15278-15286 (1999).

40 Ho, R. Y. N. *et al.* Resonance Raman Studies of the Iron(II)−α-Keto Acid Chromophore in Model and Enzyme Complexes. *Journal of the American Chemical Society* **123**, 5022-5029, doi:10.1021/ja0041775 (2001).

41 Wong, L.-L. Cytochrome P450 monooxygenases. *Current Opinion in Chemical Biology* **2**, 263-268, doi:http://dx.doi.org/10.1016/S1367-5931(98)80068-9 (1998).

42 Salnikow, K. *et al.* Depletion of Intracellular Ascorbate by the Carcinogenic Metals Nickel and Cobalt Results in the Induction of Hypoxic Stress. *Journal of Biological Chemistry* **279**, 40337-40344, doi:10.1074/jbc.M403057200 (2004).

43 Di Lullo, G. A., Sweeney, S. M., Korkko, J., Ala-Kokko, L. & San Antonio, J. D. Mapping the Ligand-binding Sites and Disease-associated Mutations on the Most Abundant Protein in the Human, Type 1 Collagen. *Journal of Biological Chemistry* **277**, 4223-4231 (2002).

44 Myllyharju, J. & Kivirikko, K. I. Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends in Genetics* **20**, 33-43 (2004).

45 Jenkins, C. L. & Raines, R. T. Insights on the conformational stability of collagen. *Natural Product Reports* **19**, 49-59 (2002).

46 Ruotsalainen, H. *et al.* Glycosylation catalyzed by lysyl hydroxylase 3 is essential for basement membranes. *Journal of Cell Science* **119**, 625-635 (2006).

47 Welford, R. W. D., Schlemminger, I., McNeill, L. A., Hewitson, K. S. & Schofield, C. J. The Selectivity and Inhibition of AlkB. *Journal of Biological Chemistry* **278**, 10157-10161, doi:10.1074/jbc.M211058200 (2003).

48 Gerken, T. *et al.* The Obesity-Associated FTO Gene Encodes a 2-Oxoglutarate-Dependent Nucleic Acid Demethylase. *Science* **318**, 1469-1472 (2007).

49 Kampranis, S. C. & Tsichlis, P. N. Histone demethylases and cancer. *Adv Cancer Res* **102**, 103-169 (2009).

50 Liu, Z. *et al.* Jmjd1a Demethylase-regulated Histone Modification Is Essential for cAMP-response Element Modulator-regulated Gene Expression and Spermatogenesis. *Journal of Biological Chemistry* **285**, 2758-2770, doi:10.1074/jbc.M109.066845 (2010).

51 Loenarz, C. *et al.* PHF8, a gene associated with cleft lip/palate and mental retardation, encodes for an Nε-dimethyl lysine demethylase. *Human Molecular Genetics*, doi:10.1093/hmg/ddp480 (2009).

52 Berdasco, M. & Esteller, M. Aberrant epigenetic landscape in cancer: how cellular identity goes awry. *Dev Cell* **19**, 698-711 (2010).

53 Black, Joshua C. *et al.* KDM4A Lysine Demethylase Induces Site-Specific Copy Gain and Rereplication of Regions Amplified in Tumors. *Cell* **154**, 541-555, doi:http://dx.doi.org/10.1016/j.cell.2013.06.051 (2013).

54 Kogure, M. *et al.* Deregulation of the histone demethylase JMJD2A is involved in human carcinogenesis through regulation of the G1/S transition. *Cancer Letters* **336**, 76-84, doi:http://dx.doi.org/10.1016/j.canlet.2013.04.009 (2013).

55 Hsia, D. A. *et al.* KDM8, a H3K36me2 histone demethylase that acts in the cyclin A1 coding region to regulate cancer cell proliferation. *Proc Natl Acad Sci U S A* **107**, 9671-9676 (2010).

56 Chang, B., Chen, Y., Zhao, Y. & Bruick, R. K. JMJD6 Is a Histone Arginine Demethylase. *Science* **318**, 444-447, doi:10.1126/science.1145801 (2007).

57 Tsuneoka, M., Koda, Y., Soejima, M., Teye, K. & Kimura, H. A Novel Myc Target Gene, mina53, That Is Involved in Cell Proliferation. *Journal of Biological Chemistry* **277**, 35450-35459, doi:10.1074/jbc.M204458200 (2002).

58 Sinha, K. M., Yasuda, H., Coombes, M. M., Dent, S. Y. & de Crombrugghe, B. Regulation of the osteoblast-specific transcription factor Osterix by NO66, a Jumonji family histone demethylase. *Embo J* **29**, 68-79.

59 Webby, C. J. *et al.* Jmjd6 Catalyses Lysyl-Hydroxylation of U2AF65, a Protein Associated with RNA Splicing. *Science* **325**, 90-93, doi:10.1126/science.1175865 (2009).

60 Mantri, M. *et al.* The 2-Oxoglutarate-Dependent Oxygenase JMJD6 Catalyses Oxidation of Lysine Residues to give 5S-Hydroxylysine Residues. *ChemBioChem* **12**, 531-534 (2011).

61 Ge, W. *et al.* Oxygenase-catalyzed ribosome hydroxylation occurs in prokaryotes and humans. *Nat Chem Biol* **8**, 960-962, doi:http://www.nature.com/nchembio/journal/v8/n12/abs/nchembio.1093.html#supplementary-information (2012).

62 Youn, M.-Y. *et al.* JMJD5, a Jumonji C (JmjC) Domain-containing Protein, Negatively Regulates Osteoclastogenesis by Facilitating NFATc1 Protein Degradation. *Journal of Biological Chemistry* **287**, 12994-13004, doi:10.1074/jbc.M111.323105 (2012).

63 Lando, D., Peet, D. J., Whelan, D. A., Gorman, J. J. & Whitelaw, M. L. Asparagine Hydroxylation of the HIF Transactivation Domain: A Hypoxic Switch. *Science* **295**, 858-861, doi:10.1126/science.1068592 (2002).

64 Hewitson, K. S. *et al.* Hypoxia-inducible factor (HIF) asparagine hydroxylase is identical to factor inhibiting HIF (FIH) and is related to the cupin structural family. *Journal of Biological Chemistry* **277**, 26351-26355 (2002).

65 Yang, M. *et al.* Asparagine and Aspartate Hydroxylation of the Cytoskeletal Ankyrin Family Is Catalyzed by Factor-inhibiting Hypoxia-inducible Factor. *Journal of Biological Chemistry* **286**, 7648-7660, doi:10.1074/jbc.M110.193540 (2011).

66 Aik, W., McDonough, M. A., Thalhammer, A., Chowdhury, R. & Schofield, C. J. Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases. *Curr Opin Struct Biol* (2012).

67 Clifton, I. J. *et al.* Crystal structure of carbapenem synthase (CarC). *Journal of Biological Chemistry* **278**, 20843-20850 (2003).

68 Rose, N. R., McDonough, M. A., King, O. N. F., Kawamura, A. & Schofield, C. J. Inhibition of 2-oxoglutarate dependent oxygenases. *Chemical Society Reviews* **40**, 4364-4397 (2011).

69 Rose, N. R. *et al.* Inhibitor scaffolds for 2-oxoglutarate-dependent histone lysine demethylases. *J Med Chem* **51**, 7053-7056 (2008).

70 Lizcano, J. M., Unzeta, M. & Tipton, K. F. A Spectrophotometric Method for Determining the Oxidative Deamination of Methylamine by the Amine Oxidases. *Analytical Biochemistry* **286**, 75-79, doi:http://dx.doi.org/10.1006/abio.2000.4782 (2000).

71 Yamagishi, Y. *et al.* Natural product-like macrocyclic N-methyl-peptide inhibitors against a ubiquitin ligase uncovered from a ribosome-expressed de novo library. *Chemistry & biology* **18**, 1562-1570, doi:10.1016/j.chembiol.2011.09.013 (2011).
72 Clemons, W. M., Jr. *et al.* Crystal structure of the 30 S ribosomal subunit from Thermus thermophilus: purification, crystallization and structure determination. *Journal of molecular biology* **310**, 827-843, doi:10.1006/jmbi.2001.4778 S0022-2836(01)94778-3 [pii] (2001).
73 Shimizu, Y. *et al.* Cell-free translation reconstituted with purified components. *Nature biotechnology* **19**, 751-755 (2001).
74 Josephson, K., Hartman, M. C. & Szostak, J. W. Ribosomal synthesis of unnatural peptides. *J Am Chem Soc* **127**, 11727-11735, doi:10.1021/ja0515809 (2005).
75 Hayashi, Y., Morimoto, J. & Suga, H. In vitro selection of anti-Akt2 thioether-macrocyclic peptides leading to isoform-selective inhibitors. *ACS chemical biology* **7**, 607-613, doi:10.1021/cb200388k (2012).
76 Bannister, A. J., Schneider, R. & Kouzarides, T. Histone Methylation: Dynamic or Static? *Cell* **109**, 801-806 (2002).
77 Klose, R. J., Kallin, E. M. & Zhang, Y. JmjC-domain-containing proteins and histone demethylation. *Nature Reviews Genetics* **7**, 715-727 (2006).
78 Klose, R. J. *et al.* The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. *Nature* **442**, 312-316 (2006).
79 Whetstine, J. R. *et al.* Reversal of Histone Lysine Trimethylation by the JMJD2 Family of Histone Demethylases. *Cell* **125**, 467-481 (2006).
80 Webby, C. J. *et al.* Jmjd6 catalyses lysyl-hydroxylation of U2AF65, a protein associated with RNA splicing. *Science* **325**, 90-93, doi:10.1126/science.1175865 (2009).
81 Mantri, M. *Mechanistic Studies of JMJD6, a Fe(II) and 2OG Dependent Lysyl Hydroxylase* Doctor of Philosophy thesis, University of Oxford, (2011).
82 Rose, N. R. *Structural, Mechanistic and Inhibition Studies on the histone Lysine Demethylases* Doctor of Philosophy thesis, University of Oxford, (2009).
83 Tsukada, Y. & Zhang, Y. Purification of histone demethylases from HeLa cells. *Methods* **40**, 318-326 (2006).
84 Yamane, K. *et al.* JHDM2A, a JmjC-Containing H3K9 Demethylase, FacilitatesTranscription Activation by Androgen Receptor. *Cell* **125**, 483-495 (2006).
85 Klose, R. J. *et al.* The Retinoblastoma Binding Protein RBP2 Is an H3K4 Demethylase. *Cell* **128**, 889-900, doi:http://dx.doi.org/10.1016/j.cell.2007.02.013 (2007).
86 Hong, S. *et al.* Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases. *Proceedings of the National Academy of Sciences* **104**, 18439-18444, doi:10.1073/pnas.0707292104 (2007).
87 Hahn, P. *et al.* Analysis of Jmjd6 Cellular Localization and Testing for Its Involvement in Histone Demethylation. *PLoS One* **5**, e13769 (2010).
88 Han, G. *et al.* The Hydroxylation Activity of Jmjd6 Is Required for Its Homo-oligomerization. *Journal of Cellular Biochemistry*.
89 Suzuki, C. *et al.* Identification of Myc-associated protein with JmjC domain as a novel therapeutic target oncogene for lung cancer. *Molecular Cancer Therapeutics* **6**, 542-551, doi:10.1158/1535-7163.mct-06-0659 (2007).

90 Teye, K. *et al.* Increased Expression of a Myc Target Gene Mina53 in Human Colon Cancer. *The American Journal of Pathology* **164**, 205-216, doi:http://dx.doi.org/10.1016/S0002-9440(10)63111-2 (2004).
91 Zhang, Y. *et al.* The Human mineral dust-induced gene, mdig, is a cell growth regulating gene associated with lung cancer. *Oncogene* **24**, 4873-4882 (2005).
92 Fukahori, S. *et al.* Immunohistochemical expressions of Cap43 and Mina53 proteins in neuroblastoma. *Journal of Pediatric Surgery* **42**, 1831-1840, doi:http://dx.doi.org/10.1016/j.jpedsurg.2007.07.008 (2007).
93 Ishizaki, H. *et al.* Overexpression of the myc target gene Mina53 in advanced renal cell carcinoma. *Pathology International* **57**, 672-680, doi:10.1111/j.1440-1827.2007.02156.x (2007).
94 Ogasawara, S. *et al.* Accelerated expression of a Myc target gene Mina53 in aggressive hepatocellular carcinoma. *Hepatology Research* **40**, 330-336, doi:10.1111/j.1872-034X.2009.00604.x (2010).
95 Tsuneoka, M. *et al.* Mina53 as a Potential Prognostic Factor for Esophageal Squamous Cell Carcinoma. *Clinical Cancer Research* **10**, 7347-7356, doi:10.1158/1078-0432.ccr-03-0543 (2004).
96 Coleman, M. L. *et al.* Asparaginyl hydroxylation of the notch ankyrin repeat domain by factor inhibiting hypoxia-inducible factor. *Journal of Biological Chemistry* **282**, 24027-24038 (2007).
97 Lu, T. *et al.* Regulation of NF-κB by NSD1/FBXL11-dependent reversible lysine methylation of p65. *Proceedings of the National Academy of Sciences* **107**, 46-51, doi:10.1073/pnas.0912493107 (2010).
98 Okada, Y., Scott, G., Ray, M. K., Mishina, Y. & Zhang, Y. Histone demethylase JHDM2A is critical for Tnp1 and Prm1 transcription and spermatogenesis. *Nature* **450**, 119-123, doi:http://www.nature.com/nature/journal/v450/n7166/suppinfo/nature06236_S1.html (2007).
99 Brauchle, M. *et al.* Protein Complex Interactor Analysis and Differential Activity of KDM3 Subfamily Members Towards H3K9 Methylation. *PLoS ONE* **8**, e60549, doi:10.1371/journal.pone.0060549 (2013).
100 Hillringhaus, L. *et al.* Structural and evolutionary basis for the dual substrate electivity of the human KDM4 histone demethylase family. *The Journal of Biological Chemistry* **286**, 41616-41625 (2011).
101 Hillringhaus, L. *et al.* Structural and Evolutionary Basis for the Dual Substrate Selectivity of Human KDM4 Histone Demethylase Family. *Journal of Biological Chemistry* **286**, 41616-41625, doi:10.1074/jbc.M111.283689 (2011).
102 Ng, S. S. *et al.* Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity. *Nature* **448**, 87-91, doi:http://www.nature.com/nature/journal/v448/n7149/suppinfo/nature05971_S1.html (2007).
103 Shiau, C. *et al.* Reconstitution of Nucleosome Demethylation and Catalytic Properties of a Jumonji Histone Demethylase. *Chemistry & Biology* **20**, 494-499, doi:http://dx.doi.org/10.1016/j.chembiol.2013.03.008 (2013).
104 Ponnaluri, V. K. C., Vavilala, D. T., Putty, S., Gutheil, W. G. & Mukherji, M. Identification of non-histone substrates for JMJD2A–C histone demethylases.

Biochemical and Biophysical Research Communications **390**, 280-284, doi:http://dx.doi.org/10.1016/j.bbrc.2009.09.107 (2009).

105 Chen, Z. *et al.* Structural basis of the recognition of a methylated histone tail by JMJD2A. *Proceedings of the National Academy of Sciences* **104**, 10818-10823, doi:10.1073/pnas.0704525104 (2007).

106 Christensen, J. *et al.* RBP2 belongs to a family of demethylases, specific for tri- and dimethylated lysine 4 on histone 3. *Cell* **128**, 1063-1076, doi:10.1016/j.cell.2007.02.003 (2007).

107 Wang, G. G. *et al.* Haematopoietic malignancies caused by dysregulation of a chromatin-binding PHD finger. *Nature* **459**, 847-851, doi:http://www.nature.com/nature/journal/v459/n7248/suppinfo/nature08036_S1.html (2009).

108 Shpargel, K. B., Sengoku, T., Yokoyama, S. & Magnuson, T. UTX and UTY Demonstrate Histone Demethylase-Independent Function in Mouse Embryonic Development. *PLoS Genet* **8**, e1002964, doi:10.1371/journal.pgen.1002964 (2012).

109 Yokoyama, A. *et al.* KIAA1718 is a histone demethylase that erases repressive histone methyl marks. *Genes to Cells* **15**, 867-873, doi:10.1111/j.1365-2443.2010.01424.x (2010).

110 Baba, A. *et al.* PKA-dependent regulation of the histone lysine demethylase complex PHF2-ARID5B. *Nat Cell Biol* **13**, 668-675, doi:http://www.nature.com/ncb/journal/v13/n6/abs/ncb2228.html#supplementary-information (2011).

111 Zhu, Z. *et al.* PHF8 is a histone H3K9me2 demethylase regulating rRNA synthesis. *Cell Res* **20**, 794-801 (2010).

112 Feng, W., Yonezawa, M., Ye, J., Jenuwein, T. & Grummt, I. PHF8 activates transcription of rRNA genes through H3K4me3 binding and H3K9me1/2 demethylation. *Nat Struct Mol Biol* **17**, 445-450, doi:http://www.nature.com/nsmb/journal/v17/n4/suppinfo/nsmb.1778_S1.html (2010).

113 Mantri, M. *et al.* Crystal Structure of the 2-Oxoglutarate- and Fe(II)-Dependent Lysyl Hydroxylase JMJD6. *Journal of Molecular Biology* **401**, 211-222 (2010).

114 Hewitson, K. S. *et al.* Hypoxia-inducible Factor (HIF) Asparagine Hydroxylase Is Identical to Factor Inhibiting HIF (FIH) and Is Related to the Cupin Structural Family. *Journal of Biological Chemistry* **277**, 26351-26355, doi:10.1074/jbc.C200273200 (2002).

115 Del Rizzo, P. A., Krishnan, S. & Trievel, R. C. Crystal Structure and Functional Analysis of JMJD5 Indicate an Alternate Specificity and Function. *Molecular and Cellular Biology* **32**, 4044-4052, doi:10.1128/mcb.00513-12 (2012).

116 Couture, J. F., Collazo, E., Ortiz-Tello, P. A., Brunzelle, J. S. & Trievel, R. C. Specificity and mechanism of JMJD2A, a trimethyllysine-specific histone demethylase. *Nature Structural and Molecular Biology* **14**, 689-695 (2007).

117 Rose, N. R. *et al.* Plant growth regulator daminozide is a selective inhibitor of human KDM2/7 histone demethylases. *J Med Chem* **55**, 6639-6643 (2012).

118 Suzuki, T., Itoh, Y. & Miyata, N. Isoform-Selective Histone Deacetylase Inhibitors. *Current Pharmaceutical Design* **14**, 529-544, doi:10.2174/138161208783885335 (2008).

119 New patented histone deacetylase inhibitors. *Expert Opinion on Therapeutic Patents* **19**, 1727-1757, doi:doi:10.1517/13543770903393789 (2009).
120 Paris, M., Porcelloni, M., Binaschi, M. & Fattori, D. Histone Deacetylase Inhibitors: From Bench to Clinic. *Journal of Medicinal Chemistry* **51**, 1505-1529, doi:10.1021/jm7011408 (2008).
121 Cole, P. A. Chemical probes for histone-modifying enzymes. *Nat Chem Biol* **4**, 590-597 (2008).
122 Cockman, M. E. *et al.* Posttranslational hydroxylation of ankyrin repeats in I kappa B proteins by the hypoxia-inducible factor (HIF) asparaginyl hydroxylase, factor inhibiting HIF (FIH). *Proceedings of the National Academy of Sciences of the United States of America* **103**, 14767-14772 (2006).
123 Greiner, D., Bonaldi, T., Eskeland, R., Roemer, E. & Imhof, A. Identification of a specific inhibitor of the histone methyltransferase SU(VAR)3-9. *Nature chemical biology* **1**, 143-145 (2005).
124 Spannhoff, A., Hauser, A. T., Heinke, R., Sippl, W. & Jung, M. The emerging therapeutic potential of histone methyltransferase and demethylase inhibitors. *ChemMedChem* **4**, 1568-1582 (2009).
125 Spannhoff, A. *et al.* Target-Based Approach to Inhibitors of Histone Arginine Methyltransferases. *Journal of Medicinal Chemistry* **50**, 2319-2325, doi:10.1021/jm061250e (2007).
126 Schmidt, D. M. Z. & McCafferty, D. G. trans-2-Phenylcyclopropylamine Is a Mechanism-Based Inactivator of the Histone Demethylase LSD1†. *Biochemistry* **46**, 4408-4416, doi:10.1021/bi0618621 (2007).
127 Kruidenier, L. *et al.* A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. *Nature* **488**, 404-408 (2012).
128 Luo, X. *et al.* A Selective Inhibitor and Probe of the Cellular Functions of Jumonji C Domain-Containing Histone Demethylases. *Journal of the American Chemical Society* **133**, 9451-9456, doi:10.1021/ja201597b (2011).
129 Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468**, 1067-1073, doi:http://www.nature.com/nature/journal/v468/n7327/abs/nature09504.html#supplementary-information (2010).
130 Fish, P. V. *et al.* Identification of a Chemical Probe for Bromo and Extra C-Terminal Bromodomain Inhibition through Optimization of a Fragment-Derived Hit. *Journal of Medicinal Chemistry* **55**, 9831-9837, doi:10.1021/jm3010515 (2012).
131 James, L. I. *et al.* Discovery of a chemical probe for the L3MBTL3 methyllysine reader domain. *Nat Chem Biol* **9**, 184-191, doi:http://www.nature.com/nchembio/journal/v9/n3/abs/nchembio.1157.html#supplementary-information (2013).
132 Hipolito, C. J. & Suga, H. Ribosomal production and in vitro selection of natural product-like peptidomimetics: The FIT and RaPID systems. *Current Opinion in Chemical Biology* **16**, 196-203, doi:http://dx.doi.org/10.1016/j.cbpa.2012.02.014 (2012).
133 Reid, P., Goto, Y., Katoh, T. & Suga, H. in *Ribosome Display and Related Technologies* Vol. 805 *Methods in Molecular Biology* (eds Julie A. Douthwaite & Ronald H. Jackson) Ch. 19, 335-348 (Springer New York, 2012).
134 Sewald N, J. H. *Peptides: Chemistry and Biology*. (Wiley-VCH, 2002).

135 Kastin, A. *Handbook of biologically active peptides*. (Access Online via Elsevier, 2013).
136 Tyndall, J. D., Nall, T. & Fairlie, D. P. Proteases universally recognize beta strands in their active sites. *Chemical reviews* **105**, 973-1000 (2005).
137 White, C. J. & Yudin, A. K. Contemporary strategies for peptide macrocyclization. *Nature Chemistry* **3**, 509-524 (2011).
138 Bowker-Kinley, M., Davis, W., Wu, P., HARRIS, R. & POPOV, K. Evidence for existence of tissue-specific regulation of the mammalian pyruvate dehydrogenase complex. *Biochem. j* **329**, 191-196 (1998).
139 Lu, C.-W., Lin, S.-C., Chen, K.-F., Lai, Y.-Y. & Tsai, S.-J. Induction of pyruvate dehydrogenase kinase-3 by hypoxia-inducible factor-1 promotes metabolic switch and drug resistance. *Journal of Biological Chemistry* **283**, 28106-28114 (2008).
140 Simon, M. D. *et al.* The Site-Specific Installation of Methyl-Lysine Analogs into Recombinant Histones. *Cell* **128**, 1003-1012, doi:http://dx.doi.org/10.1016/j.cell.2006.12.041 (2007).

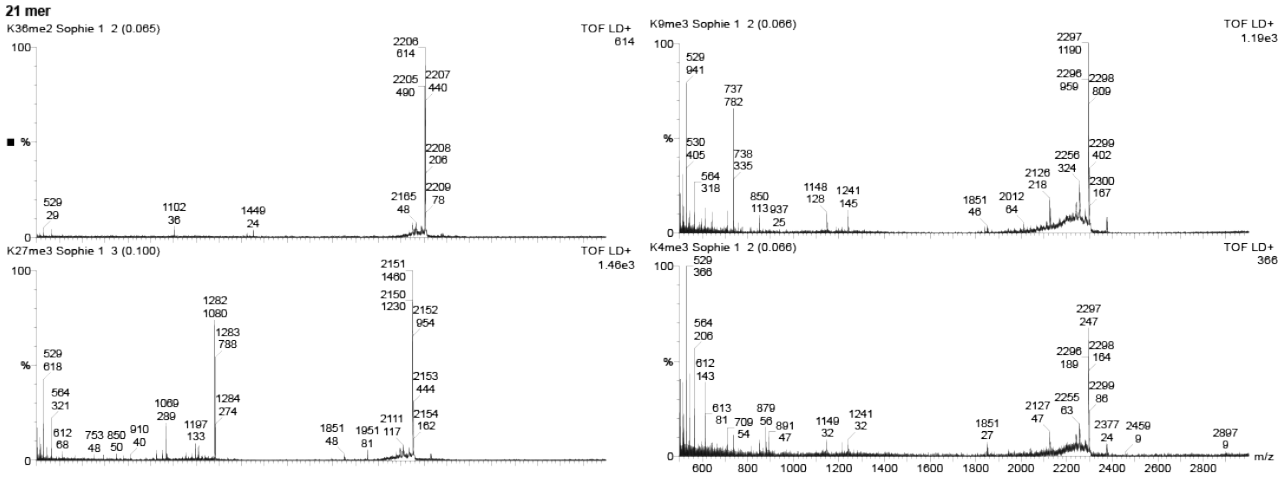
Appendix 1: Peptide Library

The library of *N*-terminal biotinylated peptides was purchased from Alta Bioscience. Other peptides were either purchased from GLBiochem (combination peptides) or synthesised using a CSBio 336X machine (FDH peptides *) or the Intarvis MultiPep (histone screen, synthesised by Dr. W. Ge).

Library	Reference / histone mark	Amino Acid Sequence	Mass /D a
Alta Bioscience	A10 / K4	Biotin-spacer- ART Kme1 QTARKSTGGKAPRKQLA	2608.5
MALDI assays	A11 / K4	Biotin-spacer- ART Kme2 QTARKSTGGKAPRKQLA	2622.5
	A12 / K4	Biotin-spacer- ART Kme3 QTARKSTGGKAPRKQLA	2636.5
	B7 / K9	Biotin-spacer- ARTKQTAR Kme1 STGGKAPRKQLA	2608.5
	B8 / K9	Biotin-spacer- ARTKQTAR Kme2 STGGKAPRKQLA	2622.5
	B9 / K9	Biotin-spacer- ARTKQTAR Kme3 STGGKAPRKQLA	2636.5
	F1 / K27	Biotin-spacer- KAPRKQLATKAAR Kme1SAPATGG	2462.3
	F2 / K27	Biotin-spacer- KAPRKQLATKAAR Kme2 SAPATGG	2476.3
	F3 / K27	Biotin-spacer- KAPRKQLATKAAR Kme3 SAPATGG	2490.3
	F11 / K36	Biotin-spacer- SAPATGGV Kme1 KPHRYRPGTVAL	2517.4
	F12 / K36	Biotin-spacer- SAPATGGV Kme2 KPHRYRPGTVAL	2531.4
	G1 / K36	Biotin-spacer- SAPATGGV Kme3 KPHRYRPGTVAL	2545.4
	G11 / K79	Biotin-spacer- VREIAQDF Kme1 TDLRFQSSAVMA	2766.5
	G12 / K79	Biotin-spacer- VREIAQDF Kme2 TDLRFQSSAVMA	2794.6
CSBio SPPS	K4	ART Kme3 QTARKSTGGKAPRKQLA *	2297.6
FDH assays	K9	ARTKQTAR Kme3 STGGKAPRKQLA *	2297.6
	K27	KAPRKQLATKAAR Kme3 SAPATGG *	2151.1
	K36	SAPATGGV Kme3 KPHRYRPGTVAL *	2206.5
GLBiochem	K4	AR-Tphos-Kme3-QTARKSTGGKAPRKQLA	2377.6
MALDI assays	K4	A-Rme2(s)-T-Kme3-QTARKSTGGKAPRKQLA	2325.6
	K9	ARTKQTAR-Kme3-S-Tphos-GGKAPRKQLA	2377.6
	K9	ARTKQTAR-Kme3-Sphos-TGGKAPRKQLA	2377.6
	K9	ARTKQTA-Rme2(s)-Kme3-STGGKAPRKQLA	2325.6
	K9	ARTKQTARKSTGGKAPRKQLA	2254.6
Intarvis MultiPep	H1 / K9	ARK(Me)3SAPAT	841
MALDI assays	H6	AR-Kme3-STAAK	915
K9	H7	ARK-Kme3-PSTGGK	1008
	H8	ARK-Kme3-ATGGK	871
	H22	QTAR-Kme3-STGG-C	1050
	H23	QTAR-Kme2-STGG-C	1036
	H15	QTAR-Kme1-STGG	1022

	H78	TKQTA-R2me(a)-Kme2-STGGK	1319.4
	H79	TKQTA-R2Me(a)-Kme3-STGGK	1305.4
	H80	TKQTA-R2Me(s)-Kme2-STGGK	1319.4
	H81	TKQTA-R2Me(s)-Kme3-STGGK	1305.4
	H12	ARTKQTAR-K3me-STGGKA	1603.8
	08_16	ARTKQTAR-Kme1-STGGKAPRKQLATKA	2569.3
	08_2	ARTKQTAR-Kme1-STGGKAPRKQ	2084.6
K4	H14	ART-Kme2-QTARKS	1391
	H 16	ART-Kme1-QTARGGC	1162
	H 17	ART-Kme3-QTARGGC	1191
	H 68	A-R2me(a)-T-Kme3-QTARKSTG	1375.4
	H 69	A-R2me(s)-T-Kme3-QTARKSTG	1375.4
	08_13	ART-Kme1-QTARKSTGGKAPRKQLATKA	2569.3
	08_14	ART-Kme2-QTARKSTGGKAPRKQLATKA	2583.3
	08_15	ART-Kme3-QTARKSTGGKAPRKQLATKA	2597.3
K27	H 5	AR-Kme3-STGGK	844
	H 25	KAAR-Kme2-SAPATGGC	1245
	H 26	KAAR-Kme3-SAPATGGC	1259
	H 27	KAAR-Kme1-SAPATGGC	1231
	H 50	ATKAAR-Kme3-SAPAT	1215.3
	H 82	ATKAA-R2me(a)-Kme2-SAPAT	1229.3
	H 83	ATKAA-R2me(a)-Kme1-SAPAT	1215.3
	H 84	ATKAA-R2me(s)-Kme2-SAPAT	1229.3
	H 85	ATKAA-R2me(s)-Kme1-SAPAT	1215.3
	08_3	LATKAAR-Kme2-SAPSTGGVKK	1799.4
	08_4	LATKAAR-Kme1-SAPSTGGVKK	1785.4
	23-mer	GGKAPRKQLATKAAR-Kme3-SAPATGG	2265.5
	24-mer	TGGKAPRKQLATKAAR-Kme3-SAPATGG	2366.6
	25-mer	STGGKAPRKQLATKAAR-Kme3-SAPATGG	2453.4
K36	08_19	KSAPSTGGV-Kme1-KPHRYRPGTVALRE	2607.2
	08-20	KSAPSTGGV-Kme2-KPHRYRPGTVALRE	2621.2
	08_21	KSAPSTGGV-Kme3-KPHRYRPGTVALRE	2635.2
	08_31	TGGV-Kme2-KPHR	1007.2
	08-32	TGGV-Kme3-KPHR	1021.2
H4R2		Biotin-spacer- SG Rme1 GKGGKGLGKGGAKRHRKV	2288.2
		Biotin-spacer- SG Rme2 GKGGKGLGKGGAKRHRKV	2302.2

MALDI to confirm HPLC fractions and masses of synthesised peptides:



Appendix 2: Mass Spectrometry Figures of Demethylation Activity

Summary of supplementary figures

S.1	KDM2A/FBXL11 demethylation at K36me2/1
S.2	KDM3A/JMJD1A demethylation at K9me2/1
S.3-5	KDM4A-C/JMJD2A-C demethylation at K9me3/2/1 and K36me3/2
S.6-7	KDM4D-E/JMJD2D/E – demethylation at K9me3/2/1
S.8	Demethylation at K27 by KDM4B-E
S.9	KDM5C/JARID1C demethylation at K4me3/2
S.10	KDM6A, B/ UTX, JMJD3 demethylation at K27me3/2/1
S.11	Hydroxylation by JMJD6
S.12	Hydroxylation by NO66 and MINA53
S.13	Hydroxylation by FIH

S.1 – KDM2A/FBXL11

Fig. S.1-A
Biotin-SAPATGGV**Kme1**KPHRYRPGTVAL
H3 K36me1

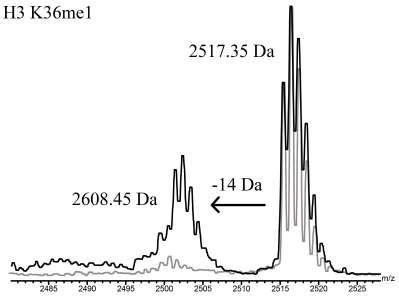


Fig. S.1-B
Biotin-SAPATGGV**Kme2**KPHRYRPGTVAL
H3 K36me2

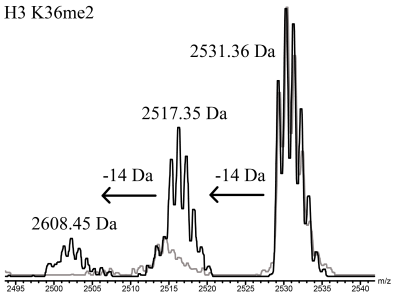
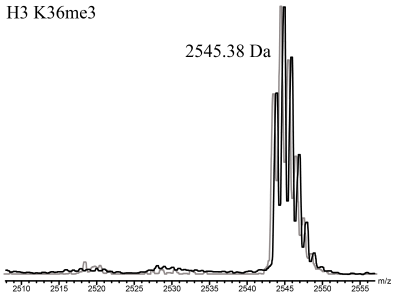


Fig. S.1-C
Biotin-SAPATGGV**Kme2**KPHRYRPGTVAL
H3 K36me3



S.2 – KDM3A/JMJD1A

Fig. S.2A
Biotin-ARTKQTAR**Kme1**STGGKAPRKQLA
H3 K9-me1

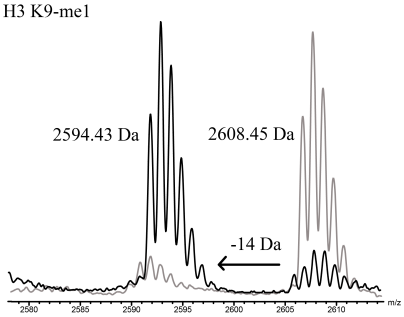


Fig. S.2B
Biotin-ARTKQTAR**Kme2**STGGKAPRKQLA
H3 K9-me2

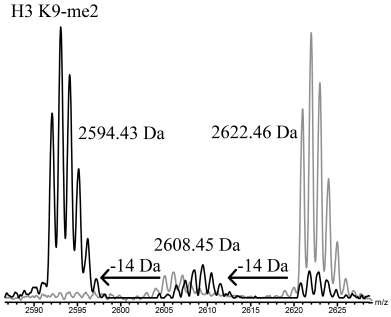
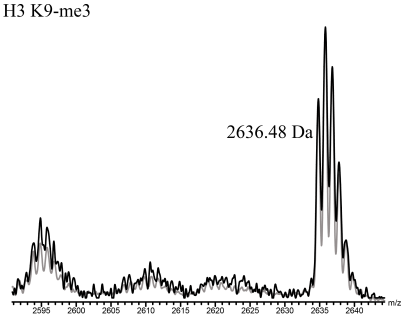
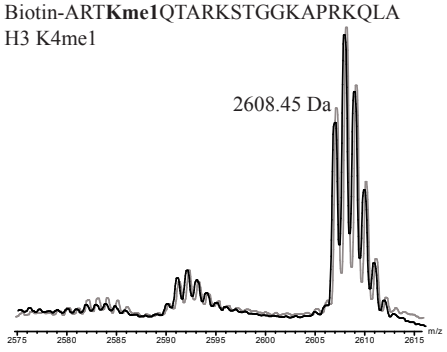


Fig. S.2C
Biotin-ARTKQTAR**Kme3**STGGKAPRKQLA
H3 K9-me3

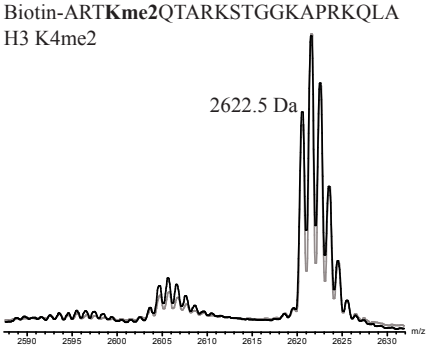


S.3 – KDM4A/JMJD2A

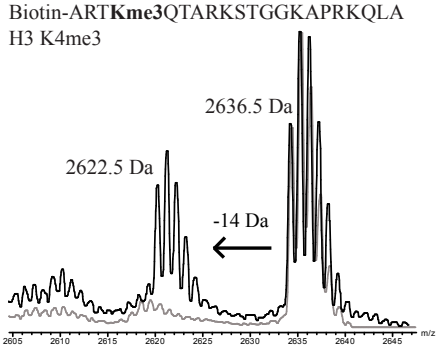
Biotin-ART**Kme1**QTARKSTGGKAPRKQLA
H3 K4me1



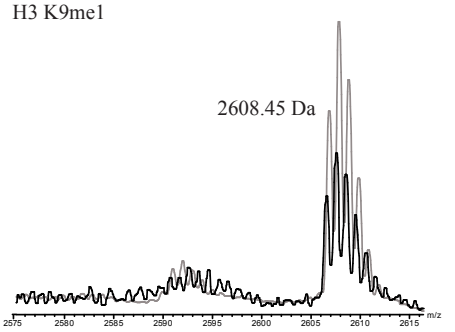
Biotin-ART**Kme2**QTARKSTGGKAPRKQLA
H3 K4me2



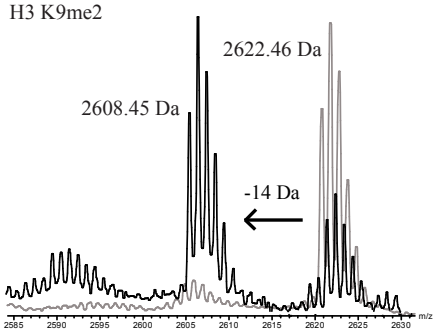
Biotin-ART**Kme3**QTARKSTGGKAPRKQLA
H3 K4me3



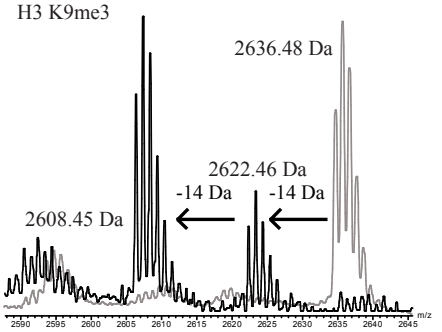
Biotin-ARTKQTAR**Kme1**STGGKAPRKQLA
H3 K9me1



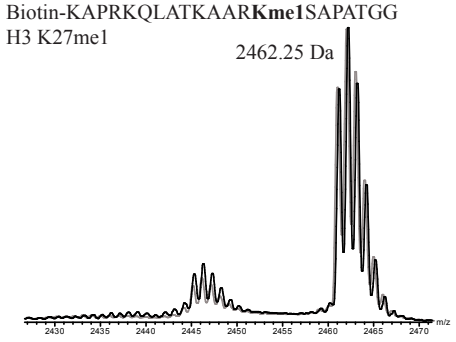
Biotin-ARTKQTAR**Kme2**STGGKAPRKQLA
H3 K9me2



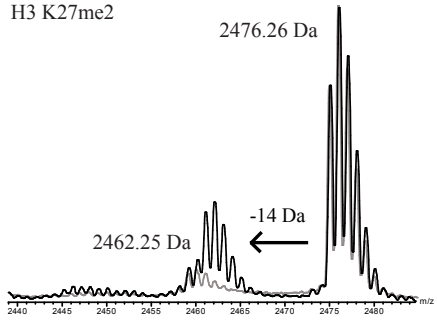
Biotin-ARTKQTAR**Kme3**STGGKAPRKQLA
H3 K9me3



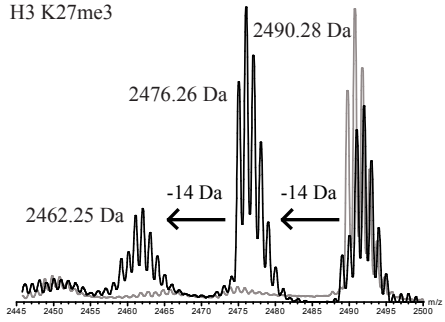
Biotin-KAPRKQLATKAAR**Kme1**SAPATGG
H3 K27me1



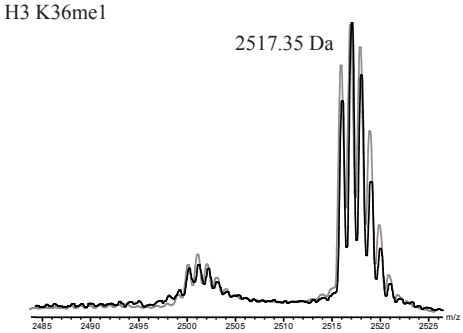
Biotin-KAPRKQLATKAAR**Kme2**SAPATGG
H3 K27me2



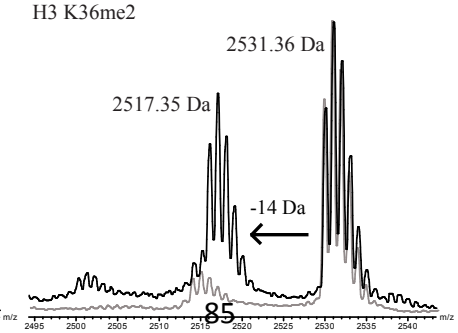
Biotin-KAPRKQLATKAAR**Kme3**SAPATGG
H3 K27me3



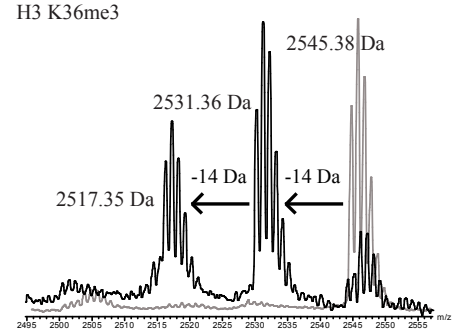
Biotin-SAPATGGV**Kme1**KPHRYRPGTVAL
H3 K36me1



Biotin-SAPATGGV**Kme2**KPHRYRPGTVAL
H3 K36me2



Biotin-SAPATGGV**Kme3**KPHRYRPGTVAL
H3 K36me3



S. 4 – KDM2B/JMJD2B

Fig. S.3A
Biotin-ARTKQTARKSTGG**Kme1**APRKQLA
H3 K9-me1

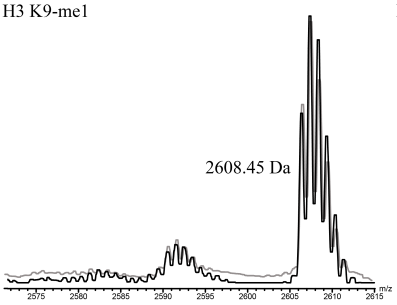


Fig. S.3B
Biotin-ARTKQTARKSTGG**Kme2**APRKQLA
H3 K9-me2

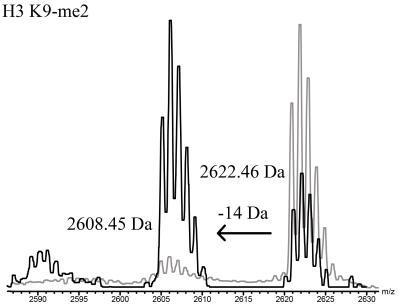


Fig. S.3C
Biotin-ARTKQTARKSTGG**Kme3**APRKQLA
H3 K9-me3

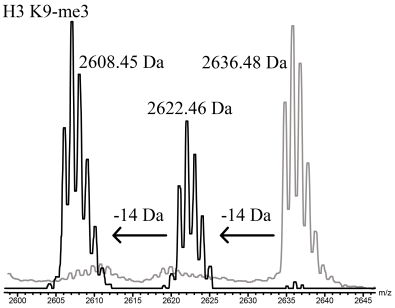


Fig. S.3D
Biotin-SAPATGGV**Kme1**KPHRYRPGTVAL
H3 K36-me1

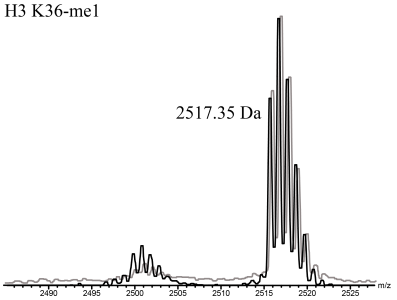


Fig. S.3E
Biotin-SAPATGGV**Kme2**KPHRYRPGTVAL
H3 K36-me2

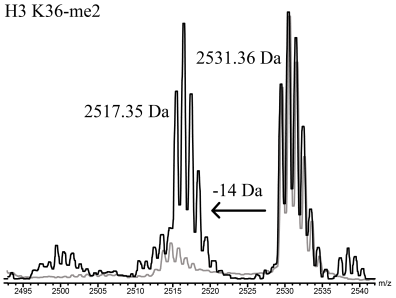
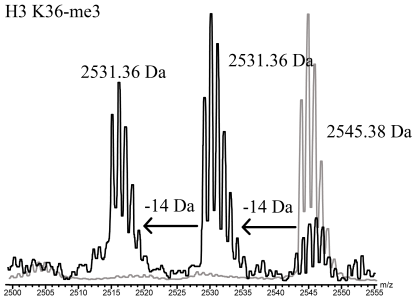


Fig. S.3F
Biotin-SAPATGGV**Kme3**KPHRYRPGTVAL
H3 K36-me3



S. 5 – KDM4C/JMJD2C

Fig. S.4A
Biotin-ARTKQTARKSTGG**Kme1**APRKQLA
H3 K9-me1

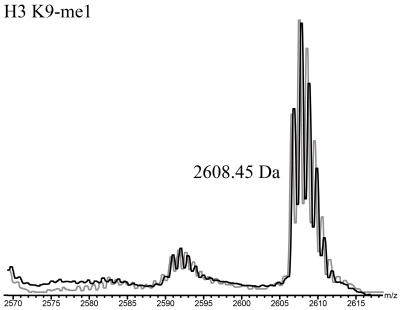


Fig. S.4B
Biotin-ARTKQTARKSTGG**Kme2**APRKQLA
H3 K9-me2

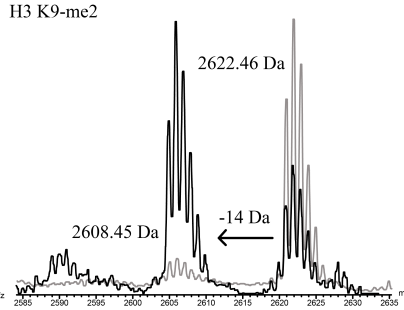


Fig. S.4C
Biotin-ARTKQTARKSTGG**Kme3**APRKQLA
H3 K9-me3

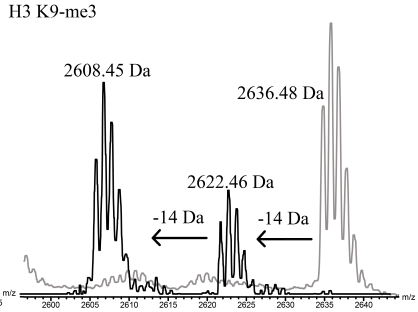


Fig. S.4D
Biotin-SAPATGGV**Kme1**KPHRYRPGTVAL
H3 K36-me1

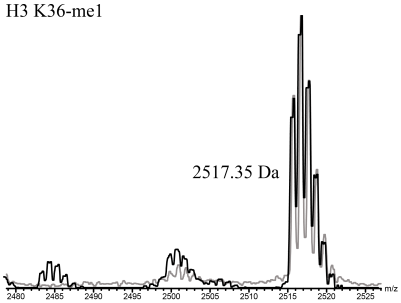


Fig. S.4E
Biotin-SAPATGGV**Kme2**KPHRYRPGTVAL
H3 K36-me2

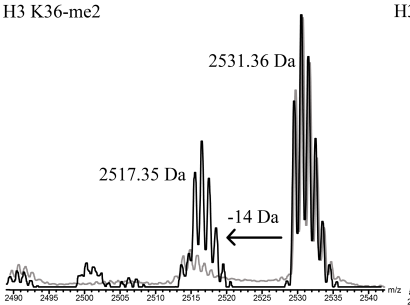
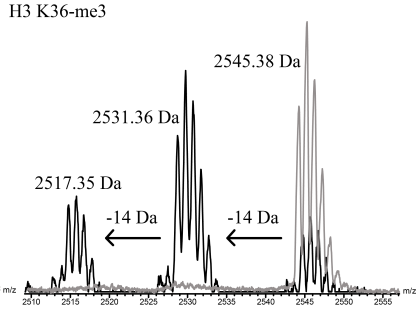


Fig. S.4F
Biotin-SAPATGGV**Kme3**KPHRYRPGTVAL
H3 K36-me3



S. 6 – KDM4D/JMJD2D

Fig. S.5A
Biotin-ARTKQTARKSTGG**Kme1**APRKQLA
H3 K9-me1

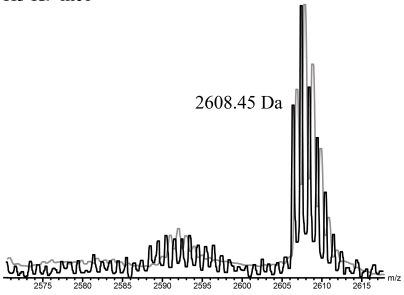


Fig. S.5B
Biotin-ARTKQTARKSTGG**Kme2**APRKQLA
H3 K9-me2

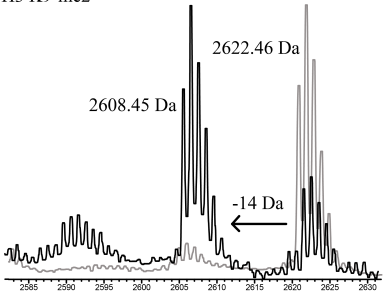
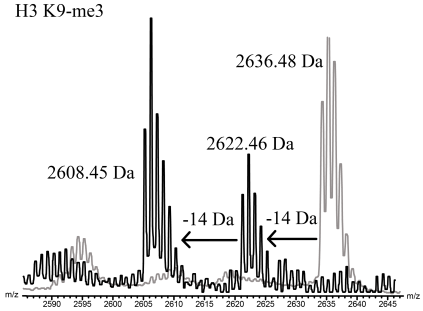


Fig. S.5C
Biotin-ARTKQTARKSTGG**Kme3**APRKQLA
H3 K9-me3



S. 7 – KDM4E/JMJD2E

Fig. S.6A
Biotin-ARTKQTARKSTGG**Kme1**APRKQLA
H3 K9-me1

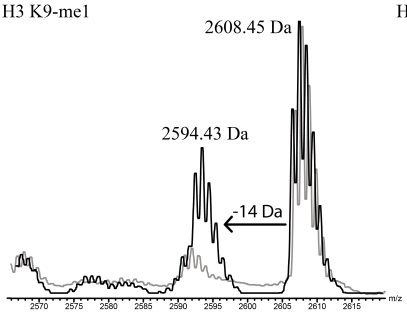


Fig. S.6B
Biotin-ARTKQTARKSTGG**Kme2**APRKQLA
H3 K9-me2

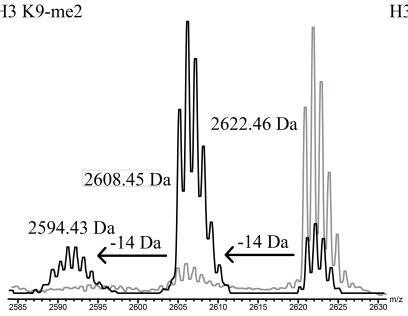
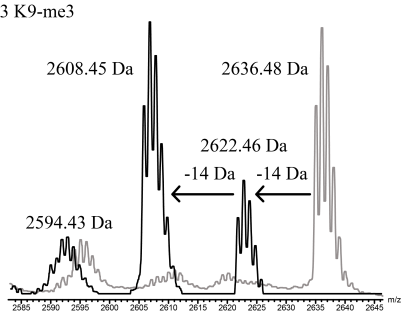
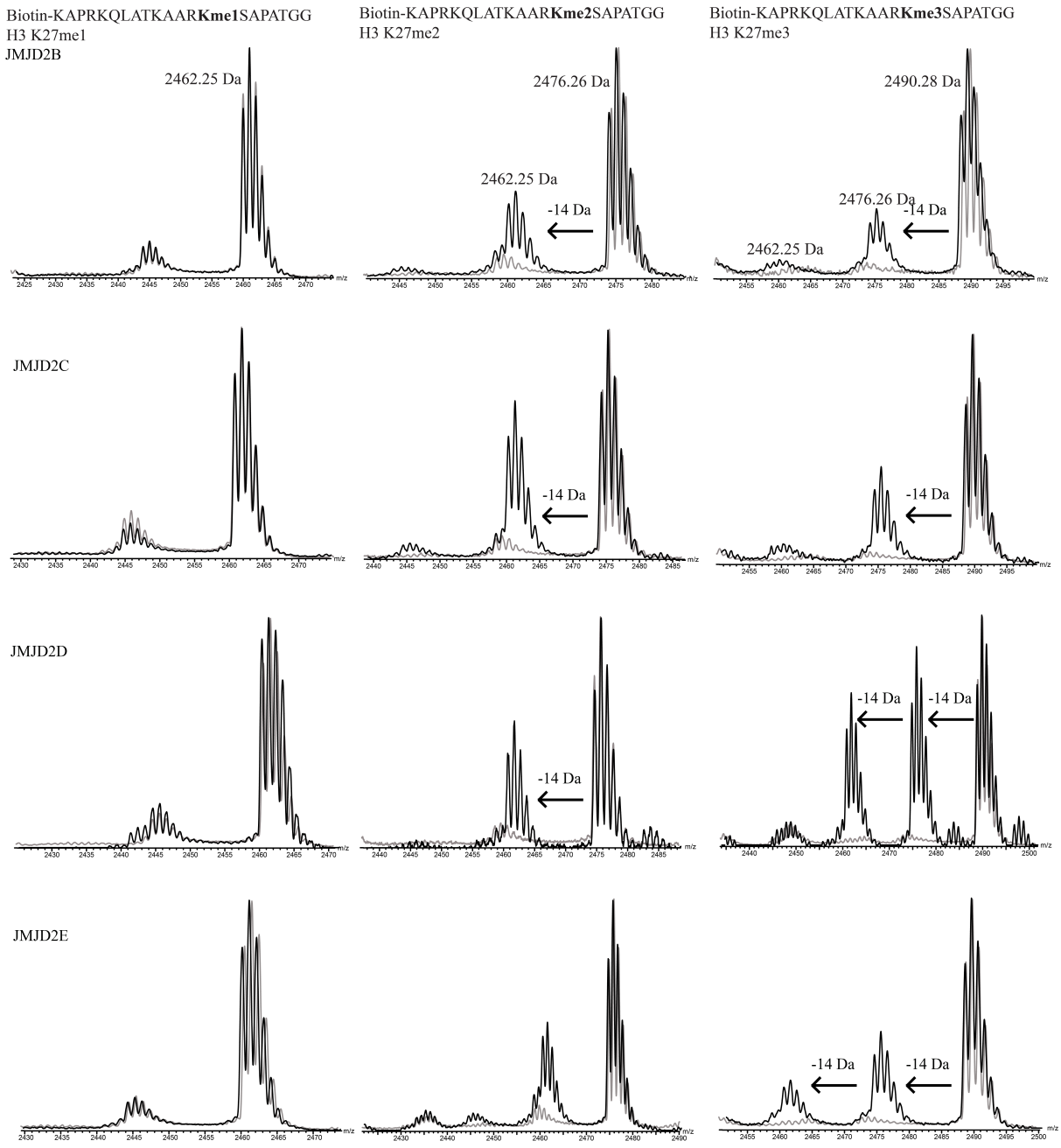


Fig. S.6C
Biotin-ARTKQTARKSTGG**Kme3**APRKQLA
H3 K9-me3



S.8 – Demethylation at K27 by KDM4 B-E



S.9 - KDM5C/JARID

Fig. S.9A
Biotin-ART**Kme1**QTARKSTGGKAPRKQLA
H3 K4-me1

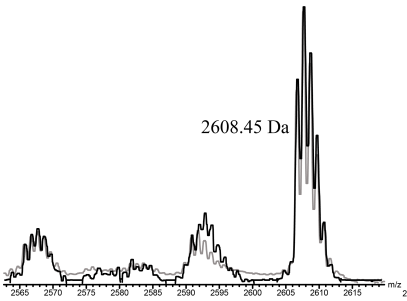


Fig. S.9B
Biotin-ART**Kme2**QTARKSTGGKAPRKQLA
H3 K4-me2

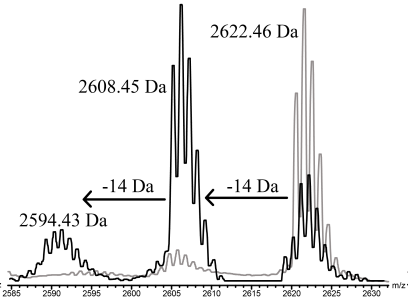
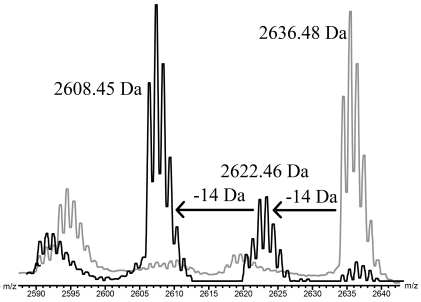


Fig. S.9C
Biotin-ART**Kme3**QTARKSTGGKAPRKQLA
H3 K4-me3



S.10 - KDM6A, B/UTX, JMJD3

Fig. S.10A
Biotin-KAPRKQLATKAAR**Kme1**SAPATGG
H3 K27me1

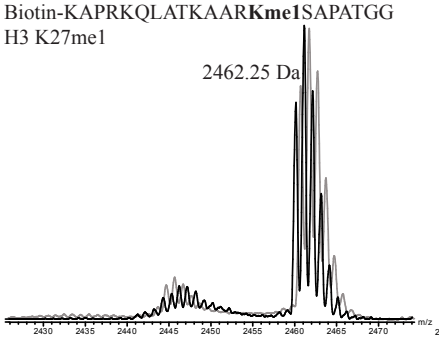


Fig. S.10B
Biotin-KAPRKQLATKAAR**Kme2**SAPATGG
H3 K27me2

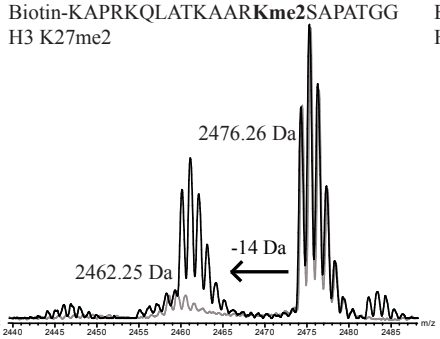


Fig. S.10C
Biotin-KAPRKQLATKAAR**Kme3**SAPATGG
H3 K27me3

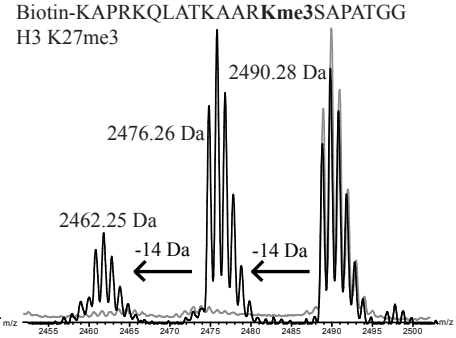


Fig. S.10D
Biotin-KAPRKQLATKAAR**Kme1**SAPATGG
H3 K27me1

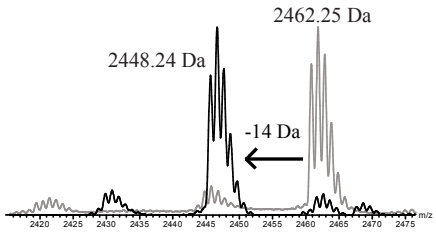


Fig. S.10E
Biotin-KAPRKQLATKAAR**Kme2**SAPATGG
H3 K27me2

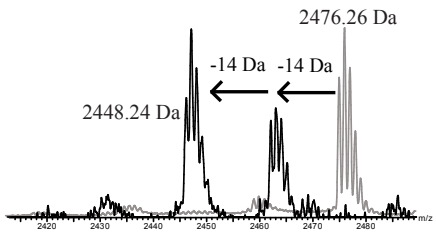
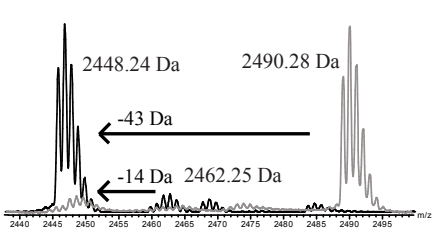
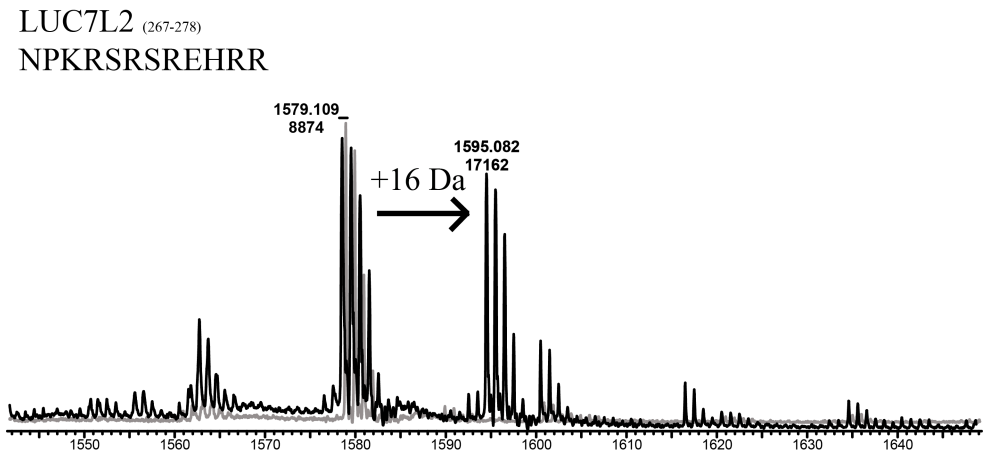


Fig. S.10F
Biotin-KAPRKQLATKAAR**Kme3**SAPATGG
H3 K27me3

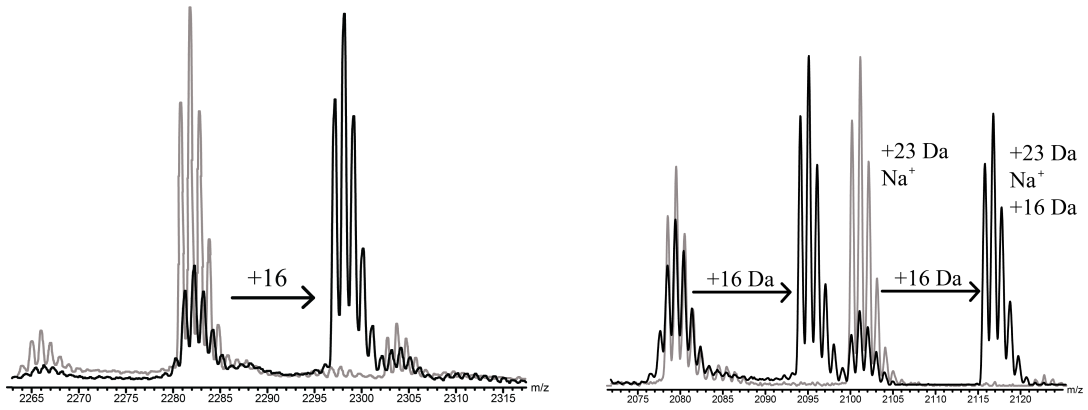


S.11 – Hydroxylation by JMJD6



S.12– Hydroxylation by MINA53 and NO66

LHS - MINA53 histidine hydroxylation of: GRGNAGGL**H**HHRINFDKYHP
RHS - NO66 histidine hydroxylation of CNPVEHPFGGGN**H**QHIGKPST
The bold residue indicates the hydroxylated histidine (as elucidated in Ge et al. 2012)



S. 13 – Hydroxylation by FIH

Synthetic Ankyrin HLEVVKLLLEAGADVNAQDK
2162.2 Da

