

Investigating the substrate selectivity and the inhibitor of the human KDM7 sub-family of JmjC histone lysine demethylases

Yijia Zhang
St Cross College, University of Oxford

June 2019

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 Doctor of Philosophy

Abstract

The regulation of gene expression by the post-translational modifications (PTMs) of histone tails is important in eukaryotic cell biology and of current medical interest. Specific PTMs to histone tails, including acetylation, phosphorylation and methylation, can either promote or inhibit transcription of specific sets of genes. Histone methylation was once thought to be irreversible but is no known to be dynamic. The extent of N^{ϵ}-methyl lysine- and arginine-methylation is regulated by sets of methyl transferases and demethylases.

The work described in this Thesis focused on the substrate selectivity and inhibition of the human KDM7 subfamily of JmjC histone demethylases (KDMs). KIAA1718 (KDM7A), PHF8 (KDM7B) have been demonstrated to be KDMs while PHF2 (KDM7C) is a proposed pseudo-enzyme. The KDM7 enzymes are part of the large family of JmjC KDMs which are 2-oxoglutarate (2OG) dependent oxygenases and which use Fe(II) as a cofactor. Most JmjC KDMs are multidomain proteins with the catalytic activity lying in their JmjC domain. However, other domains may regulate demethylation. The KDM7 enzymes are of particular interest because of their interesting demethylation activities and relatively simple domain architectures. Their plant homeodomains (PHDs) are proposed to bind to the histone H3K4me3 mark and in doing so either promote (PHF8), or inhibit (KIAA1718), demethylation by the JmjC domain of H3K9me2.

Following an introduction to the epigenetic roles of KDMs (Chapter 1), Chapter 2 describes work on the kinetics of the KDM7 enzymes. Consistent with prior reports. PHF8 was found to demethylate H3K9me2 and its activity is enhanced by the presence of H3K4me3. KIAA1718 is reported to demethylate H3K9me2 and H3K27me2, the demethylation of which was either inhibited (H3K9me2) or enhanced (H3K27me2) by the binding of H3K4me3 mark. However, my results raise questions about the effect of H3K4me3 on demethylation of H3K27me2 by KIAA1718. New potential histone substrates were identified and characterised for both PHF8 and KIAA1718.

Mechanistic studies were performed in Chapter 3 to investigate factors regulating KDM7 substrate binding. Mutagenesis studies and metal ion binding studies of wild-type and variant PHF8/PHF2 proteins revealed that although the HxD…H iron binding motif is crucial, other factors are involved in regulate the KDM activities of the KDM7 subfamily. The overall results reveal the PHD fingers are of primary importance in substrate binding.

Using AlphaScreen and fluorescence polarisation methods, novel inhibitors for the KDM7 enzymes were investigated along with the mode of inhibition. In one line of investigation, a cyclic peptide inhibitor was identified to selectively inhibit the KDM activity of PHF8 and promote the activity of KIAA1718 at the H3K9me2 mark.

Overall, the work described in this Thesis is supportive of the proposal that the PHD fingers of the KDM7 enzymes regulate their demethylation activities. Indeed, binding to the H3K4me3 mark by their PHD fingers appears to dominates substrate binding by the KDM7 enzymes. Blocking the binding of the PHDs of KDM7s can inhibit (PHF8) or promote (KIAA1718) demethylation activity.

Acknowledgments

Firstly, I would like to thank my supervisors, Professor Chris Schofield and Dr Akane Kawamura. Chris, for providing me the opportunity to join the group and for his enthusiasm and ideas. Akane, for being endless patient and supportive, and for always keeping me on track. I feel grateful to have them as my supervisors.

A number of people have contributed directly to the work in this thesis. I am particularly indebted to Dr Louise Walport, Dr Richard Hopkinson, Dr Anthony Tumber, Dr Bhaskar Bhushan, Dr Alexandre Erdmann, Dr Philip Gerken, Oliver Coleman, Roman Belle and Xian Deng for materials, protocols and experimental assistance. All specific contributions are acknowledged in the text. Thanks also to H. Manso Jubier, Dr Adam Hardy and Wendy Sobey for keeping the lab running smoothly.

An enormous number of people in the research groups have been both excellent colleagues and friends during my D.Phil. Particular thanks for technical advice and useful discussions go to Dr Louise Walport, Dr Kate Dunne, Joanna Bonnici, Dr Richard Hopkinson, Leah Taylor-Kearney, Gareth Langley, Dr Hanna Tarhonskaya. Special thanks for proofreading go to Dr Christopher Lohans, Dr Louise Walport, Dr Kate Dunne, Grace Roper and Dr Bhaskar Bhushan. Many others have provided a lovely working environment: thank you Joan, Rachel, Kerstin, Kristina, Saiful, Ilaria, Jurgen, Lennart, Chiara, Dong, Shifali, Cyrille, David, Hilal, Anka and everyone I am bound to have missed.

Most importantly, my heartfelt thanks go to my parents for their support over the past years; and my fiancé Bowen for his love, his patience and his companionship during my highs and lows.

Declaration

The work described in this report is entirely my own, except where I have either acknowledged help from a named person or given reference to a published source.



Yijia Zhang

Contents

	Abstract	i
	Acknowledgements	ii
	Declaration	iii
	Contents	iv
	Abbreviations	vii

Chapter 1: Introduction

1.1	Epigenetics	1
1.2	DNA packaging, nucleosomes and histones	2
1.3	Post-translational modifications of histone tails	4
1.3.1	Acetylation and deacetylation	6
1.3.2	Methylation	7
1.3.3	Demethylation	9
1.3.4	Other post-translational modifications	10
1.4	The 2-oxoglutarate dependent oxygenases	11
1.5	The JmjC histone demethylases	15
1.6	Histone readers	18
1.6.1	Plant homeodomains	18
1.6.2	Bromodomains	19
1.6.3	Tudor domains	19
1.6.4	Chromodomains	20
1.7	The KDM7 subfamily	20
1.8	Thesis objectives	22
1.9	References	23

Chapter 2: Biochemical studies on the KDM7 subfamily

2.1	Chapter overview	30
2.2	Introduction	31
2.3	Results and discussion	35
2.3.1	Protein expression and purification	35
2.3.2	Oprimisation of *in vitro* assay conditions for demethylation of dimethyllysine by KDM7 family enzymes	41

2.3.3 Investigations on the *in vitro* demethylation activities of PHF8 48
2.3.4 Investigations on KIAA1718 catalysed demethylation reactions 52
2.3.5 Potential alternatives substrates for KIAA1718 and PHF8 59
2.3.6 Histone substrate screening of PHF2 69
2.4 Summary 71
2.5 References 73

Chapter 3: Substrate and metal ion binding studies on the KDM7 subfamily

3.1 Chapter overview 77
3.2 Introduction 78
3.2.1 Previous studies on KDM7 subfamily 78
3.2.2 Techniques of interaction studies 80
3.2.3 The principle of fluorescence polarisation (FP) assays 82
3.3 Results and discussion 84
3.3.1 Purification of PHF8H319Y and PHF2Y321H 84
3.3.2 Activity of variant proteins 85
3.3.3 Buffer optimisations for KDM7s and their variants for substrate binding 87
3.3.4 Substrate binding affinities of KDM7 enzymes and their variants 88
3.3.5 Buffer optimisations for KDM7s and their variants for methylstat binding 97
3.3.6 Methylstat binding studies on KDM7 enzymes and their variants 100
3.3.7 Metal coordination studies on KDM7 JmjC domain active site 106
3.4 Summary 112
3.5 References 114

Chapter 4: Inhibition studies on the KDM7 subfamily

4.1 Chapter overview 117
4.2 Introduction 118
4.2.1 Inhibitors targeting the KDM7 and KDM2 subfamilies 118
4.2.2 Principles of the AlphaScreen assay 119
4.3 Results and discussion 122
4.3.1 Establishing AlphaScreen assay for inhibitor screening 122
4.3.2 Establishing FP displacement assay for inhibitor screening 135

4.3.3 Investigating the mode of inhibition of a specific inhibitor of the KDM2/7 subfamilies 139
4.3.4 Identifying novel cyclic peptide inhibitors of the KDM7 subfamily 145
4.3.5 Amiodarone derivatives inhibit KDM7s in a PHD finger independent manner 151
4.4 Summary 156
4.5 References 157

Chapter 5: Materials and Methods

5.1 General procedures 160
5.2 DNA constructs generation 164
5.3 Recombinant enzyme gene expression in *E*.coli 177
5.4 Recombinant enzyme purification 181
5.5 Activity assays 183
5.6 References 188

Chapter 6: Summary and conclusions

6.1 Summary and conclusions 191
6.2 References 196

Appendix 1 Sequence alignment of the KDM7 subfamily 198

Appendix 2 Recombinant KDM7 constructs 199

Appendix 3 Publications arising from work described in this Thesis 200

Abbreviations

2OG	2-oxoglutarate
2TY	2 × Tryptone-Yeast Extract
ATP	Adenosine Triphosphate
APS	Ammonium Persulfate
BSA	Bovine Serum Albumin
CHCA	α-cyano-4-hydroxycinnamic Acid
CTP	Cytidine Triphosphate
Da	Dalton (mass unit)
DNA	Deoxyribonucleic Acid
DSBH	Double Stranded Beta Helix
DSF	Differential Scanning Fluorimetry
DTT	Dithiothreitol
E. coli	Escherichia coli
EDTA	Ethylenediaminetetracetic Acid
Et	Ethyl
FAD	Flavin-adenine Dinucleotide
FDH	formaldehyde dehydrogenase
FIH	Factor Inhibiting HIF
Fmoc	Fluorenylmethyloxycarbonyl
FPLC	Fast Protein Liquid Chromatography
GTP	Guanosine Triphosphate
H3	Histone 3
HAT	Histone Acetyltransferase

HDAC	Histone Deacetylase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIF	Hypoxia Inducible Factor
HMT	Histone methyltransferases
HPLC	High Pressure Liquid Chromatography
IEX	Ion Exchange Chromatography
IMAC	Immobilised Metal Affinity Chromatography
IOX1	5-Carboxy-8-hydroxyquinoline
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal Titration Calorimetry
JmjC-KDM	JumonjiC- Domain Containing KDM
KDM	Histone Lysine Demethylase
KMT	Lysine methyltransferases
LIC	Ligation Independent Cloning
LSD	Lysine Specific Demethylase
MALDI-TOF	Matrix-Assisted Laser Desorption Ionisation Time of Flight
Me	Methyl
MS	Mass Spectrometry
MW	Molecular Weight
NAD	Nicotinamide Adenine Dinucleotide
NF- κβ	nuclear factor-κβ
NMR	Nuclear Magnetic Resonance
NOG	N-oxalylglycine
NTP	Nucleoside Triphosphate

OD	Optical Density
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
PHD	Plant Homeodomain
PHF	Prolyl Hydroxylase Domain containing protein
PKA	Protein Kinase A
PMSF	Phenylmethylsulfonyl fluoride
PRMT	Protein Arginine Methyltransferase
PTM	Post-translational Modification
RMT	Protein Arginine Methyltransferases
RNA	Ribonucleic Acid
RT	Room Temperature
SAM	S-adenosyl Methionine
sd	Standard Deviation
SDM	Site Directed Mutagenesis
SDS	Sodium Dodecyl Sulfate
SEC	Size Exclusion Chromatography
sem	Standard Error of the Mean
SP-Seph	Sulphopropyl Sepharose
TB	Terrific Broth
TCA	Tricarboxylic Acid
TEMED	N,N,N',N'-Tetramethyl-Ethane-1,2-Diamine
TET	Ten-Eleven Translocation Enzyme
TEV	Tobacco Etch Virus

TFA	Trifluoroacetic Acid
TIPS	Triisopropylsilane
T_m	Melting Temperature
Tris	Tris(hydroxymethyl)aminomethane
UV	Ultraviolet
WT	Wild Type

Common one- and three-letter amino acid abbreviations are used throughout the text.

Chapter 1 Introduction

1.1 Epigenetics

Deoxyribonucleic acid (DNA) represents the code based on which all proteins are synthesised, with messenger ribonucleic acid (mRNA) acting as an intermediate between the two (i.e. 'DNA makes mRNA makes proteins').[1] It is often said that the diploid cells in the human body possess the same genotype, derived from embryonic stem cells, but can differentiate to phenotypically distinct cell types. The gene expression profile of each cell is specialised, depending on their location, function, and environmental factors, suggesting that a complex regulatory mechanism is employed during cellular differentiation and in the normal life cycle of a cell.[2]

'Epigenetics' was first defined by Waddington as the study of the mechanisms involved in determining how a single genotype (i.e. DNA sequence) gives rise to different phenotypes.[2] This definition pictured cellular differentiation during development as a ball (the undifferentiated cell) rolling downhill into inescapable valleys, with its ultimate fate determined by the valley into which it settles.[3] Each valley represents a particular cell type, and once settled, a cell cannot change types.

More recently, one definition of epigenetics concerns heritable changes in gene expression that are not attributed to changes in their underlying DNA sequence.[4] The mechanisms of epigenetic regulation, including non-coding RNA, and modifications to DNA and histones, are also considered as contribution to epigenetic regulation.[5,6] However, questions remain as to the heritability of epigenetic marks.[7] Two models have been proposed: the reprogramming model (which largely occurs in gametes, in which the state of chromatin is

reset to achieve totipotency), and the inheritance model (which suggests that the state of chromatin in embryonic cells is inherited from mature germ cells, and directs the cellular development).[8] However, the extent to which epigenetic marks can be inherited between generations of cells in higher organisms remains unclear.[8]

1.2 DNA packaging, nucleosomes and histones

Each animal cell contains around two metres of 'linear' DNA,[9] which is extensively compacted into chromatin by several levels of packing.[10] In the first level of packing, ~146 base pairs of DNA wrap around a histone octamer complex, stabilised by electrostatic interactions between positively charged histones and negatively charged DNA, to form a nucleosome (Figure 1.1 A). Each histone octamer complex contains two copies of each of the four histone proteins, histone 2A (H2A), histone 2B (H2B), histone 3 (H3) and histone 4 (H4).[11,12] These histone proteins initially form heterodimers (H2A with H2B, and H3 with H4), and two H2A/H2B dimers combine with two H3/H4 dimers to produce the histone octamer. Individual nucleosomes are joined together by linker regions of DNA, forming the 'beads on a string' motif (Figure 1.1 B). Addition of linker histone 1 (H1) stabilises the DNA/histone complex, and leads further compaction into a 30 nm fibre, from which DNA can further condense to form a chromatin fibre. The density of chromatin fibre varies according to the phase of the cell cycle.[13] in general, chromatin presents in two forms: the open and transcriptionally permissive euchromatin, and the tightly compacted heterochromatin.[13] During metaphase, chromatin adopts a highly-condensed structure (size decreased more than 1000-fold) to form mitotic chromosomes.[14]

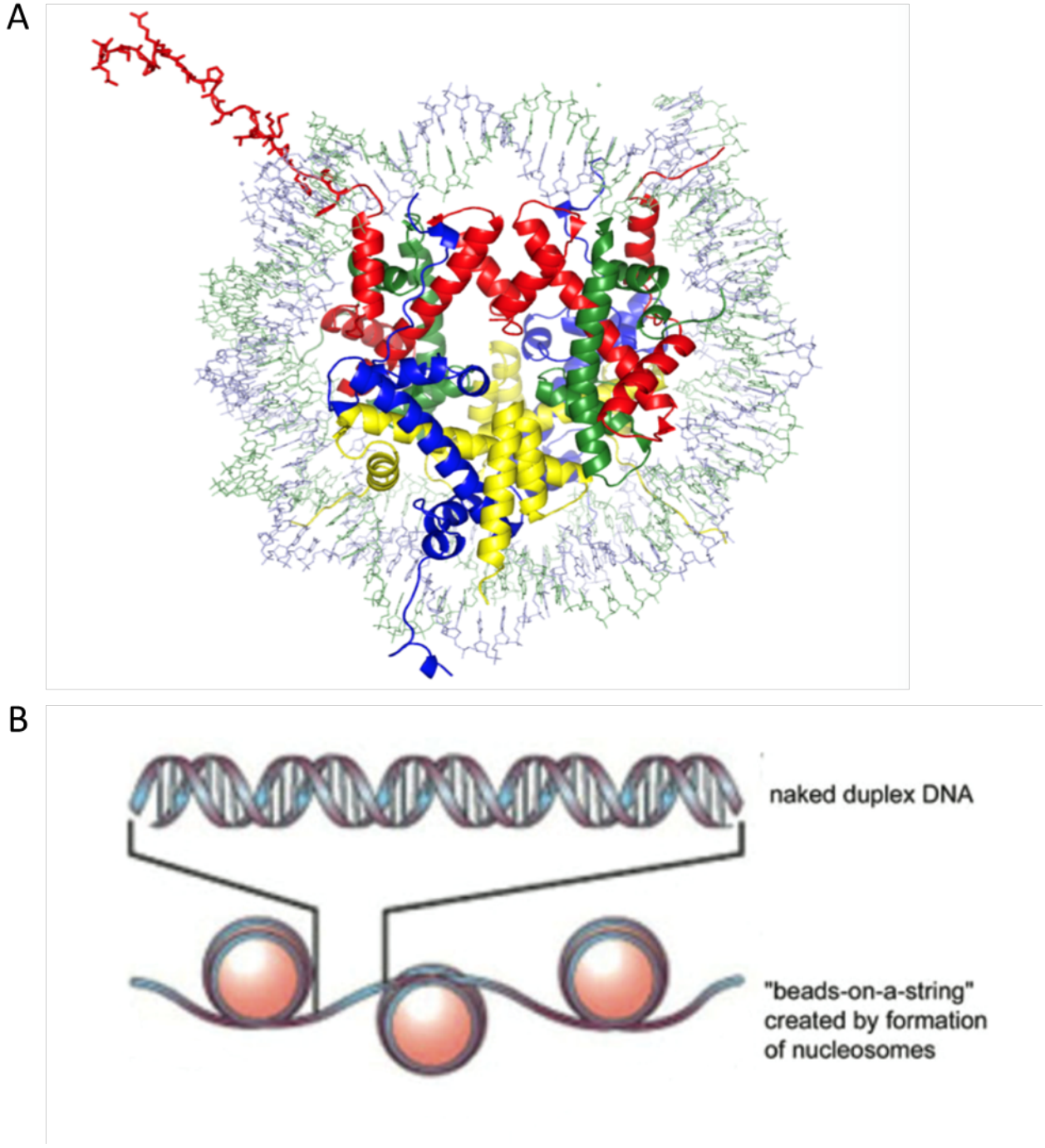


Figure 1.1 A) View from a crystal structure of a nucleosome (PDB ID: 1AOI).[11] DNA is shown in pale blue and green. Histone 3 (H3) proteins are shown in red, histone 4 (H4) in green, histone 2A (H2A) in yellow and histone 2B (H2B) in blue. **B)** Scheme showing the 'beads on a string' form of the nucleosome.[13] Figure reproduced from reference 13.

Histones are relatively small basic proteins, usually containing 100 to 150 residues. Their structures consist of a *C*-terminal globular core domain containing three α-helices separated

by two loops, and an unstructured *N*-terminal tail that protrudes out from the nucleosome.[15] The *N*-terminal tails of histones are thought to play a significant role in maintaining the structural stability of nucleosomes.[15] NMR and ion mobility mass spectrometric studies suggest that the *N*-terminus of histone proteins is highly disordered in solution.[16,17] Evidence from computational studies shows that the H2B, H3 and H4 tails have α-helical regions, suggesting that the precise structures of the histone tails may be subject to biological context.[18–21]

In some contexts, histone proteins H2A, H2B and H3 may be replaced by variants which differ from the canonical histones in size, sequence and chromosomal location.[22] As an example, H2A.X is involved in response to DNA damage repair.[23] Variants of H3 include H3.1 and H3.2, which together represent the canonical H3 protein. H3.3 marks active chromatin, while H3t is specific to testes and CENP-A and is found in centromeric nucleosomes.[22,24–27] Variation in the sequence and structure of these histone variants allows for the regulation of gene expression by altering the stability and structure of the nucleosome and changing the types of post-translational histone modifications that are possible.[22] Mutation or dysregulation of these histone variants is associated with cancers and other diseases.[28]

1.3 Post-translational modifications of histone tails

Epigenetic regulation can be achieved by the production of noncoding RNA, ATP-dependent remodelling of chromatin, and the covalent modification of DNA and histone proteins.[29,30] These modifications can impact on the structure and charge of chromatin, ultimately modulating the interaction of DNA with transcriptional machinery, thereby affecting transcription.

Proteins are synthesised *via* ribosomal translation of mRNA.[31] Following synthesis, proteins can be modified in multiple ways, including by proteolytic cleavage, the addition of cofactors, and modification with small chemical groups (e.g. acetyl and methyl groups).[32–34] These modifying processes, which normally occur following translation, constitute post-translational modifications (PTMs).[35] Histones are subject to extensive PTMs, most of which occur on the unstructured *N*-terminal histone tail (Figure 1.4).[36] Modification of lysine, arginine, threonine, tyrosine and serine residues has been found to impact many epigenetic phenomena (Figure 1.4).[35,37] PTMs can alter the charge and conformation of histone tails, impacting on the association between histones and DNA,[36] thereby modulating transcription of the corresponding genes.[35,38–40] Different combinations of PTMs on a histone tail can lead to different outcomes, and thus the set of PTMs on a histone are often referred to as the 'histone code'.[41] PTMs serve as 'marks', and are regulated by epigenetic enzymes. Histone marks are installed by proteins known as 'writers', removed by proteins known as 'erasers', and recognised by proteins known as 'readers', providing a highly dynamic method of epigenetic regulation.[42] The cross talk between different marks is also thought to be of importance; for example, modification at one position may facilitate the binding of an epigenetic enzyme and direct its action to a nearby residue.[40,43] Well-studied histone PTMs include acetylation, methylation and phosphorylation, as briefly discussed in the following sections (section 1.4.1 – section 1.4.3).

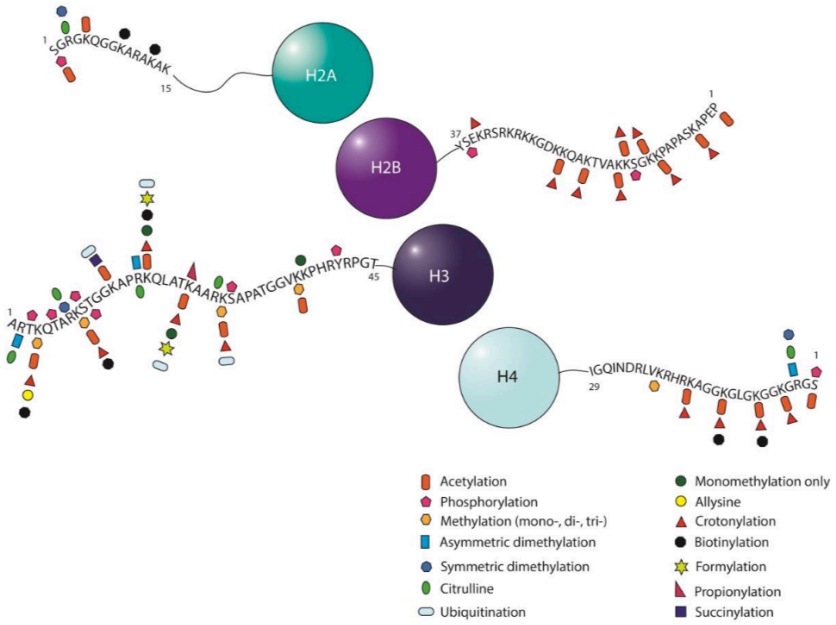


Figure 1.4 Post-translational modifications (PTMs) known to occur on the *N*-termini of the four canonical core histones.[44] Figure reproduced with permission.[45]

1.3.1 Acetylation and deacetylation

Histone acetylation may occur on a lysine residue, neutralising the positive charge of its side chain, thereby weakening the interaction between the histone and negatively charged DNA (Figure 1.5).[46–49] As a result, DNA compaction can be relaxed by histone acetylation, increasing the accessibility of the DNA to other proteins and transcription factors, and thus modulating the transcription of particular genes.[46,50] On the other hand, the removal of an acetyl group from a lysine residue, known as deacetylation, promotes more condensed chromatin packaging, and thus represses gene transcription.[51]

Acetyl groups can be added by histone acetyltransferases (HATs) and removed by histone deacetylases (HDACs). HATs utilise acetyl-coenzyme A (acetyl-CoA) as the donor of an acetyl group and are directly involved in energy homeostasis since acetyl-CoA is a key intermediate in the tricarboxylic acid (TCA) cycle. There are four main subfamilies of HATS: Gcn5-related *N*-acetyltransferases (GNATs), MYST (MOZ, Ybf2, Sas2 and Tip60), p300/CBP, and Rtt109,

which vary in selectivity for different histone positions.[52] HDACs are divided into four classes based on sequence homology: Class I (HDAC1, 2 and 8), Class II (HDAC4, 5, 7, 9 and 10), Class III (SIRT1-7) and Class IV (HDAC11).[53,54] Like HATs, HDACs also differ in selectivity, and have also been reported to accept non-histone substrates.[53,55,56] HDACs are involved in many biological processes including cell migration and cell-cell interactions, and dynamically regulate gene expression together with HATs.[56]

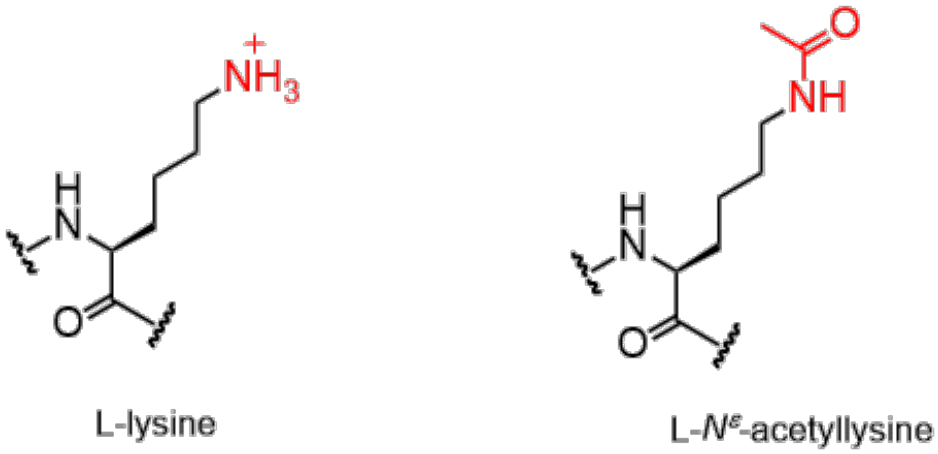


Figure 1.5 Acetylation of lysine residues neutralises its positive charge. The terminal group and the post-translational modification (PTM) are shown in red.

1.3.2 Methylation

Histone methylation occurs on lysine- and arginine- residues. Lysine can be mono-, di- and tri-methylated (Figure 1.6 A), while arginine can be mono- or di-methylated (Figure 1.6 B). The dimethylation of arginine can be symmetric (Rme_{2s}) with one methyl group on each terminal nitrogen of the guanidino group, or asymmetric (Rme_{2a}) with both methyl groups on the same terminal nitrogen (Figure 1.6). Unlike lysine acetylation, methylation does not impact the charge of lysine- or arginine- residues, and thus does not directly change the structure of chromatin.[57] Instead, histone methylation can signal for transcription, depending on the position of the mark on histone and the methylation state.[58]

Lysine and arginine methylation is catalysed by histone methyltransferases (HMTs) which use *S*-adenosyl methionine (SAM) as a methyl group donor.[59] Lysine methylation is catalysed by a large family of lysine methyltransferases (KMTs), most of which possess a SET domain (acronym for Su(Var)3-9, enhancer-of-zeste, trithorax) that catalyses the methylation reaction.[60–63] As with the KMTs, the enzymes responsible for histone arginine methylation (protein arginine methyltransferases (PRMTs)) are SAM-dependent.[64]

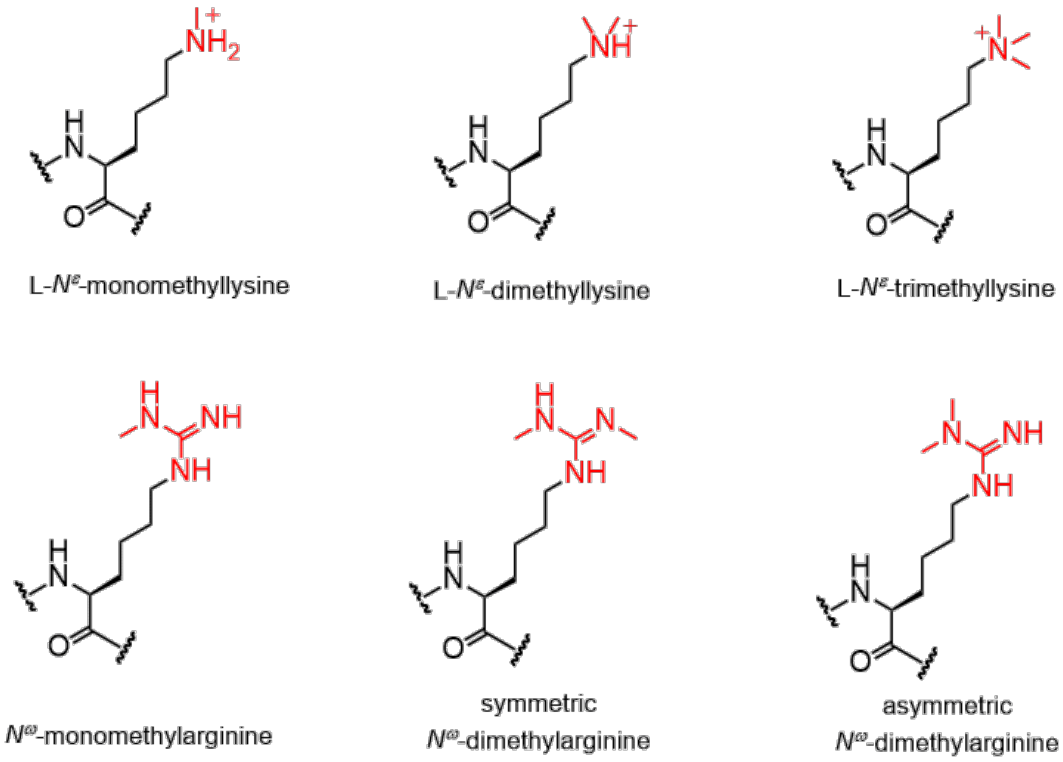


Figure 1.6 Structures of methylated lysine- and arginine- residues. The post-translational modifications (PTMs) are shown in red.

Methylation of lysines on the tails of the four canonical histones has been identified, and has been ascribed multiple functions. For histone 3 (H3), methylation at positions 4, 36 and 79 is related to gene activation, while methylation at positions 9 and 27 usually represses genes.[65] The methylated lysine marks are localised to particular gene regions; for example, histone 3 trimethylated lysine at position 4 (H3K4me3) is often found in the promoter

regions of active genes,[66,67] while H3K9me3 and H3K27me3 are enriched in the promotors of inactive genes.[68,69] In addition to histone tails, methylation of a core residue on histone 3 (H3K79) has also been identified, as catalysed by DOT1L.[65,70,71]

1.3.3 Demethylation

Histone methylation was thought to be irreversible until 2004, when lysine-specific demethylase 1 (LSD1) was identified.[72] Also known as histone lysine demethylase 1 A (KDM1A), LSD1 is a flavin-adenine dinucleotide (FAD) dependent oxidase which catalyses the removal of mono- and dimethyl groups on H3K4 and H3K9 *via* the formation of an imine intermediate (Figure 1.7 A).[72,73] LSD1 has also been reported to demethylate a number of non-histone substrates, including p53 and DNMT1. A second member of the LSD family, LSD2 (also known as KDM1B) was identified as an H3K4me1/me2 demethylase in 2009.[74] More recently, LSD2 has been reported to ubiquitylate O-GlcNAc transferase (OGT), thereby promoting the proteasome-dependent degradation of OGT.[75]

Another family of histone lysine demethylases (KDMs) is the Jumonji-C histone lysine demethylases (JmjC KDMs), which belongs to the 2-oxoglutarate dependent dioxygenase (2OG oxygenase) superfamily. The JmjC KDMs employ an oxidative demethylation mechanism which does not require the formation of an imine intermediate, and thus the JmjC KDMs can demethylate all three states of lysine methylation (Figure 1.7 B).[76] The detailed catalytic mechanism of the JmjC KDMs is discussed in the following section (section 1.5), as it shares the same mechanism as other 2OG oxygenases.[77]

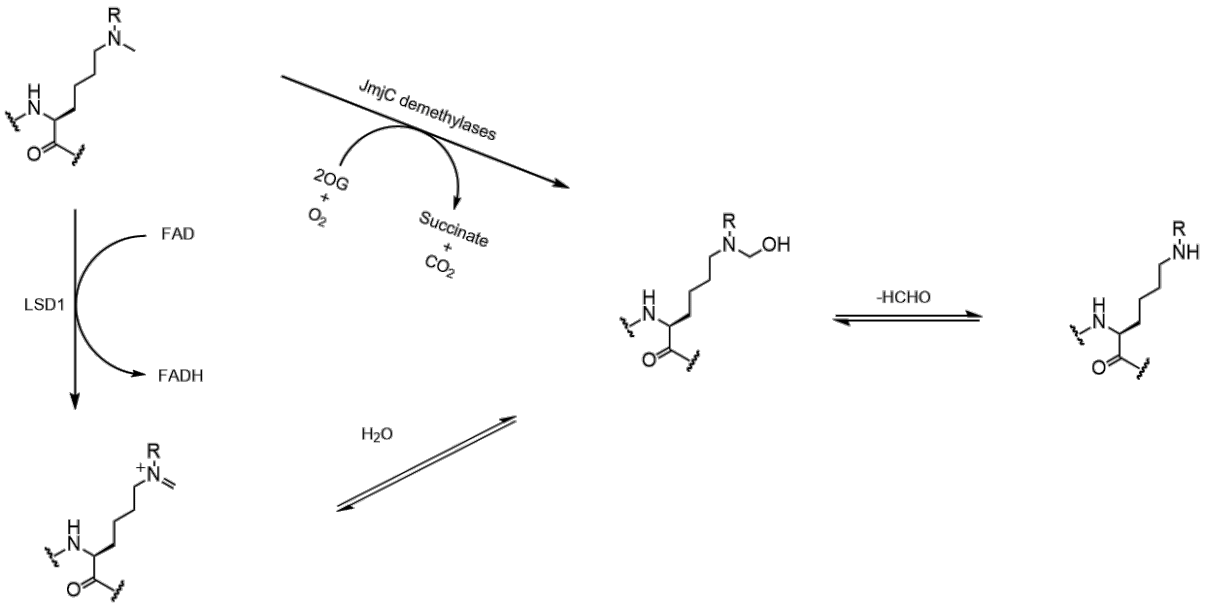


Figure 1.7 Scheme showing the different mechanisms of the LSD1 and JmjC lysine demethylases.[78,79]

1.3.4 Other post-translational modifications

Other PTMs on histone tails include phosphorylation, ubiquitylation, crotonylation, malonylation, succinylation and formylation.[80–85] Phosphorylation can occur on histone serine, threonine and tyrosine residues (Figure 1.8). Histone phosphorylation is catalysed by protein kinases, and removed by histone phosphatases.[86] The phosphorylation of histones was found to be involved in chromatin compaction during meiosis and mitosis, as well as chromatin relaxation when associated with H3 acetylation.[84,87–89] The phosphorylation of histone variant H2AX may be present in the nucleosomes around the break points of damaged DNA, and facilitates the recruitment of the necessary DNA repair enzymes.[86]

Ubiquitylation is usually considered to be the process of marking the histone protein for degradation.[90] Ubiquitin, which is a 76-residue protein, is attached to the lysine residue of histone *via* an enzymatic cascade reaction, and is itself further ubiquitylated to form the

mark for degradation.[85] Other functions of ubiquitylation include signalling for protein recruitment and structural changes to the DNA.[91]

Crotonylation, malonylation, succinylation and formylation are all examples of different types of lysine acylation modifications.[80–83] Unlike the well-studies acetylation mark, these acylation marks were only recently discovered, and little is known about their impact.

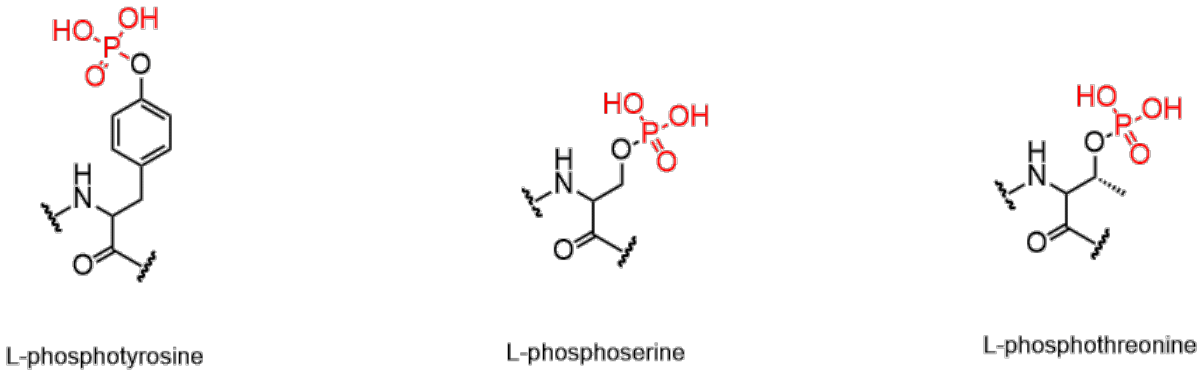


Figure 1.8 Phosphorylation of tyrosine, serine and threonine residues. The post-translational modifications are shown in red.

1.4 The 2-oxoglutarate dependent oxygenases

The 2-oxoglutarate (2OG) dependent oxygenases represent a superfamily of enzymes present in almost all organisms, with more than 60 members present in humans.[92] The 2OG oxygenases are able to catalyse a wide variety of oxidative reactions, including epoxidation, desaturation, ring closure and hydroxylation (Figure 1.9).[92,93] The diverse reactivity, substrate selectivity and functionality of the 2OG oxygenases has been extensively studied.

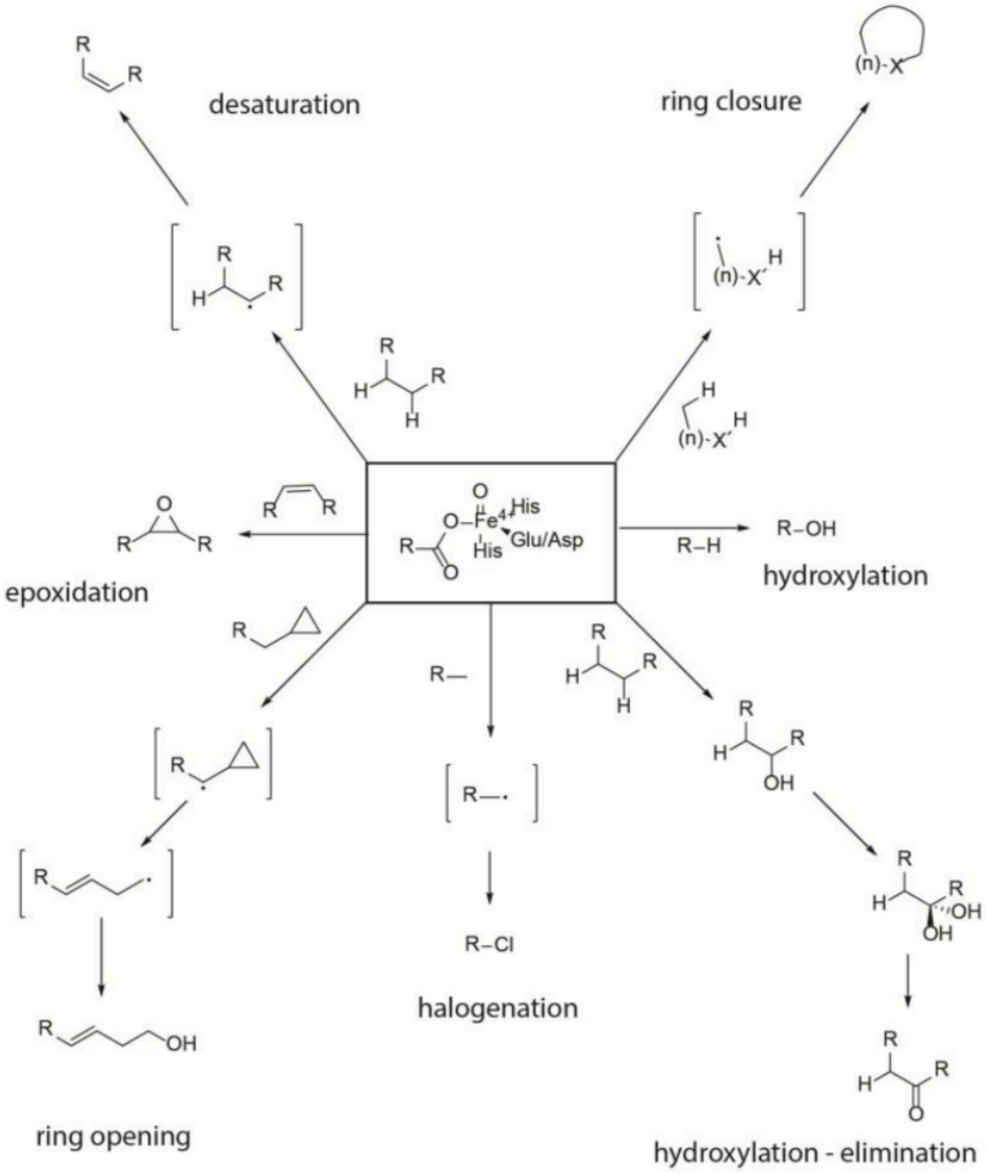


Figure 1.9 Range of reactions catalysed by members of the 2OG oxygenase superfamily of enzymes.[93]

In 1995, it was first observed by Roach *et al*. that isopenicillin *N*-synthase (IPNS), a 2OG oxygenase related oxidase enzyme involved in the synthesis of penicillin,[94] possesses a double-stranded β-helix (DSBH, commonly known as the 'jelly-roll' motif) and an iron-binding motif formed by three protein residues (Figure 1.10 A). Subsequent studies showed that most structurally characterised 2OG oxygenases contain this DSBH fold, and a conserved HxD/E…H iron-binding motif.[95,96] As observed in the crystal structure of deacetoxycephalosporin-C synthase (DAOCS), a 2OG oxygenase involved in cephalosporin

biosynthesis, the catalytic Fe(II) is octahedrally coordinated by the three amino acids and three water molecules in its resting state (Figure 1.10 B).[97] The addition of 2OG displaces two of these water molecules, allowing one of the 2OG carboxylates to coordinate to the Fe(II), and the other 2OG carboxylate to be bound by a basic residue (arginine or lysine) near the iron-binding site.[97–101] The third water molecule that binds to the Fe(II) centre is displaced upon oxygen binding, which is thought to be facilitated by the binding of the substrate to the active site.[97,99,102]

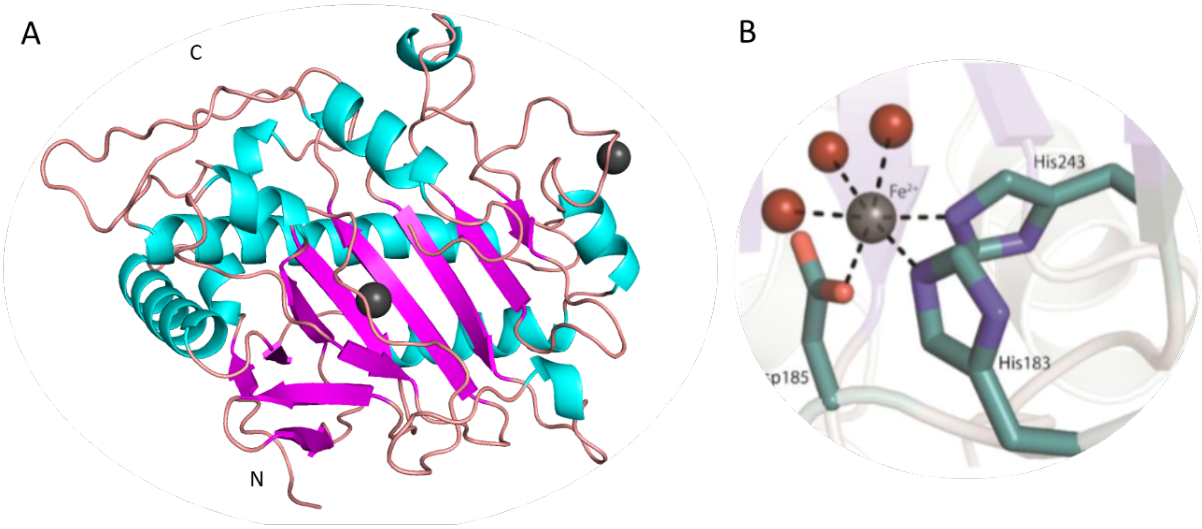


Figure 1.10 View from the crystal structure of **A)** IPNS, in which the DSBH was first observed (PDB ID: 1IPS). **B)** The resting state of the iron binding site of DAOCS (PDB ID: 1RXF).

The study of DAOCS, along with kinetic and spectroscopic analyses of other 2OG oxygenases, led to the proposal of a conserved reaction mechanism for the 2OG oxygenase superfamily (Figure 1.11).[97–104] The individual steps are as follows:

1) 2OG binds to the Fe(II) in the active site in a bidentate manner, displacing two water molecules.

2) The prime substrate binds to the active site, weakening the binding of the remaining water molecule to the Fe(II). Displacement of this final water molecule allows the Fe(II) to react with a molecule of oxygen.

3) Fe(II) reacts with the oxygen molecule, and a Fe(III)-superoxo species is generated.

4) A ferryl intermediate is formed.

5) Decarboxylation of 2OG produces a reactive Fe(IV)-oxo species and releases CO_2.

6) The Fe(IV)-oxo species abstracts a hydrogen atom from the substrate C-H bond, forming Fe(III)-OH and a substrate carbon radical.

7) The hydroxyl group of Fe(III)-OH is transferred to the substrate carbon radical, returning the enzyme to its Fe(II) state. If this hydroxyl group is part of a hemiaminal, as in the case of JmjC demethylases (see Figure 1.7), it may collapse to form formaldehyde.

8) The hydroxylated substrate and the succinate by-product leave the active site, and two water molecules bind to Fe(II). The enzyme is returned to its resting state and is prepared for the next catalytic cycle.

Interestingly, while ascorbate is not directly involved in the catalytic cycle of 2OG oxygenases, studies have suggested that ascorbate is beneficial for the activity of a number of these enzymes, including the collagen hydroxylases and the hypoxia-inducible factor (HIF) hydroxylases.[105,106] The presence of ascorbate is usually thought to prevent Fe(II) oxidation and thus to preserve the enzymatic activity of the 2OG oxygenases.[105]

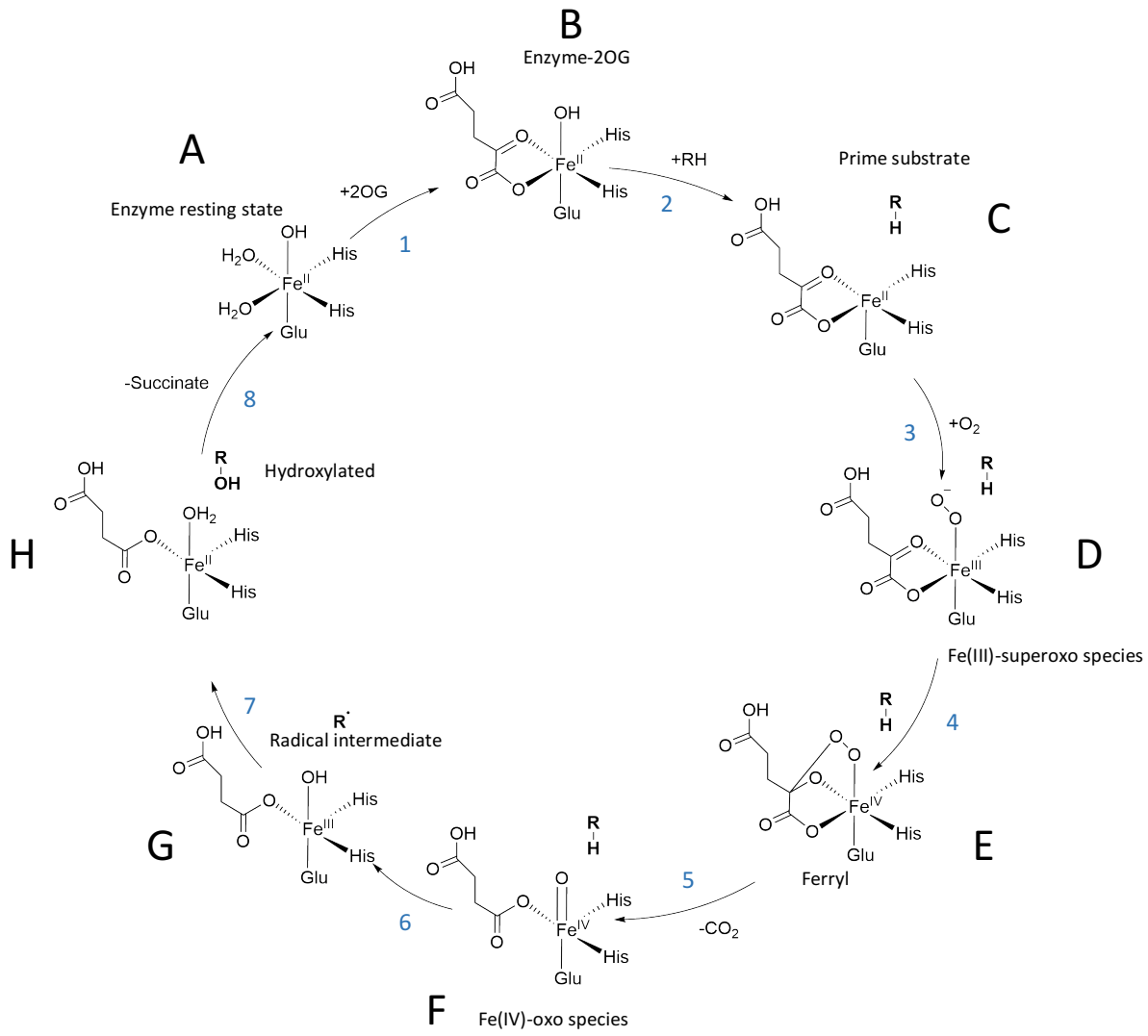


Figure 1.11 Consensus mechanism of the hydroxylation catalysed by the 2OG oxygenase superfamily.[107] Complex A-C have been directly reveal by crystal structures; complex F has been evidenced by spectroscopy studies; complex H has been identified using stopped-flow UV-visible studies; complex D,E and G have not been observed.

1.5 The JmjC histone demethylases

The Jumonji C (JmjC) domain was first identified in a gene trap experiment of JARID2 in mice.[108] Mutation of the gene for JARID2 was found to result in aberrant neural plate and tube formation, producing abnormal neural grooves. Hence, the gene responsible for this was named Jumonji, based on the Japanese word for 'cruciform'.[108] The JmjC domain was later found to be indispensable for the demethylation activity of FBXL11, which was subsequently named JmjC domain-containing histone demethylase 1A (JHDM1A), and whch

is also known as KDM2A.[109] In humans, over 30 JmjC domain-containing proteins have been identified to date, the majority of which have been validated as demethylases, with the remainders being hydroxylases or catalytically unidentified proteins.[92,110]

The JmjC demethylases primarily act on histones, and are known as epigenetic regulators.[111] Recent studies have also identified non-histone substrates for some JmjC demethylases *in vitro* and in cells.[112,113] As members of the 2OG oxygenase superfamily, JmjC demethylases have the typical 'jelly-roll' motif and iron-binding motif of the other enzymes in this group (Section 1.5). The mechanism of the demethylation reaction catalysed by the JmjC demethylases follows the general mechanism of the 2OG oxygenases (Figure 1.11), with the further reaction of a hydroxymethyl group to form a molecule of formaldehyde and the demethylated product (Figure 1.7). Interestingly, the JmjC demethylases differ from the classical 2OG oxygenases in that the co-factor 2OG binds on the 4th β-strand of the DSBH core (except KDM6 on the 1st β-strand and KDM3B on the 8th β-strand), rather than the 8th β-strand as is seen with most other 2OG oxygenases.[98] In addition, JmjC demethylases often have a lysine residue binding to the co-factor 2OG, while most other 2OG oxygenases use an arginine residue.

The JmjC histone demethylase family is the largest family of histone demethylases (KDMs), with members that remove most of the known lysine methylation marks.[79] Phylogenetic analyses divide the human JmjC KDMs into six subfamilies (KDM2-7) based on their substrate specificities and sequence similarities (Table 1.1).[114] In addition to the catalytic domain, most KDMs also contain reader domains which recognise and/or bind to substrates, thereby facilitating substrate specificity and protein-protein interactions.[115–117] For example,

the plant homeodomains (PHDs) of the KDM7 subfamily are known to bind to H3K4me3, and the CXXC domains of FBXL11 are reported to bind to non-methylated CpG DNA.[115,118,119] A detailed description of the histone readers can be found in the next Section (Section 1.7).

Table 1.1 Substrate specificity and domains of the characterised JmjC-KDMs.[114] Blue shading indicates demethylation activity which has been observed both in cells and *in vitro*, green shading *in vitro* only, and orange shading in cells only. 1, 2, 3 indicates the methylation level, i.e. me1, me2, me3.

	H3K4			H3K9			H3K27			H3K36			H3K79			H4K20			H1.4K26		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
KDM2A																					
KDM2B																					
KDM3A																					
KDM3B																					
KDM3C																					
KDM4A																					
KDM4B																					
KDM4C																					
KDM4D																					
KDM4E																					
KDM5A																					
KDM5B																					
KDM5C																					
KDM5D																					
KDM6A																					
KDM6B																					
KDM6C																					
KIAA1718																					
PHF8																					
PHF2																					

1.6 Histone Readers

Most chromatin-associated proteins contain one or more binding domains specific for histone tails.[120] The binding of these reader domains localises the protein in the correct positions in the chromatin environment and helps to modulate their activities.[120] Descriptions of some of the histone readers are given below.

1.6.1 Plant homeodomains

In 1993, the plant homeodomain (PHD) was first identified in the homeodomain protein HAT3.1 in *Arabidopsis thaliana*. However, it was not until 2006 that it was identified as a reader of trimethylated lysine 4 of histone H3 (H3K4me3). Since then, multiple binding substrates have been reported for PHD, including unmodified H3K4 and H3K9me3.[121–124] The human genome encodes over 200 PHD-containing proteins. Most of the identified PHD fingers are involved in the control of gene transcription and chromatin dynamics.[125]

Plant homeodomains are also known as PHD fingers, as they contain two zinc finger motifs (Figure 1.12).[126] PHD fingers bind their substrates *via* an aromatic cage which confers substrate selectivity.[125] For H3K4me3/2 reading PHD fingers, the composition and positioning of key cage-forming residues are very diverse, as revealed by a structure-based sequence alignment.[125] The N-terminal sequence around the first zinc finger is thought to be of critical importance in histone peptide recognition. For example, mutagenesis studies suggested that alanine substitution of the *N*-terminal residues N295/D297, E298, D304 and E307 of the first PHD finger of the autoimmune regulator 1 (AIRE1) significantly reduces its binding to H3 substrates.[122,127,128]

PTMs on histone *N*-terminal tails can influence the binding of PHD fingers. For example, the presence of H3R2me2a has been observed to negatively correlate with the H3K4me3 binding of bromodomain and PHD finger transcription factor (BPTF).[129] The H3K4 readout by the CHD4 PHD finger enhanced its H3K9 methylation activity.[130,131] Therefore, the reading of a particular mark on a histone tail by PHD fingers depends on its local PTM composition (PTM cross talk).

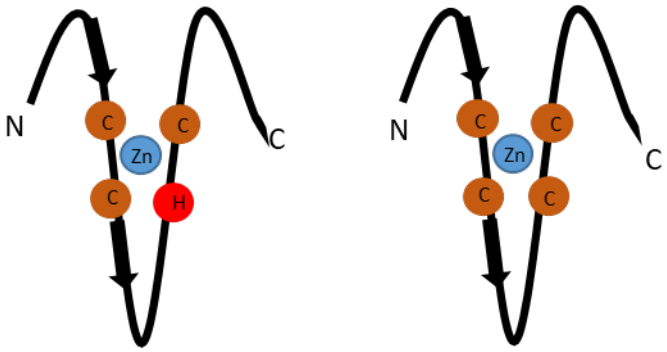


Figure 1.12 Zinc binding motifs of the PHD fingers of PHF8.

1.6.2 Bromodomains

There are over 40 bromodomain-containing proteins that have been identified in the human proteome.[132] Bromodomains recognise and bind to the acetylated lysine residue on histone tails. In addition to histone substrates, many non-histone substrates have been found for bromodomains; for example, the bromodomain of CREB binding protein binds to an acetylated lysine residue of p53.[133] Bromodomains consist of a left-handed 4 α-helix bundle and the loops between the helices contribute to the substrate specificity of the bromodomain-containing proteins.[134]

1.6.3 Tudor domains

Tudor domains are involved in many cellular processes, including RNA metabolism, histone modification and DNA damage response.[135] Tudor domains bind to methylated arginine or

lysine residues, and are probably the most well-characterised family of methylarginine-binding domains.[136] The Tudor domain itself consists of 5-stranded bent anti-parallel β-sheet which forms an aromatic cage to bind the methylated residues. The openness of the aromatic cage is thought to define its substrate specificity.[135]

1.6.4 Chromodomains

Chromodomains were originally identified in two *Drosophila* proteins, Polycomb and HP1, which were subsequently found in most plants and animals.[137] The chromodomains found in chromatin-associated proteins bind to methylated lysine residues.[138] For some proteins, DNA binding has also been reported.[139] Chromodomains contain three anti-parallel β-sheets and a α-helix, made up of around 50 amino acids.[140]

1.7 The KDM7 subfamily

The subsequent Chapters in this Thesis are primarily focused on the human KDM7 subfamily of JmjC histone demethylases. The KDM7 subfamily consists of three members: KIAA1718 (KDM7A), PHF8 (KDM7B) and PHF2 (KDM7C).[77] Each of these proteins contains a PHD finger and an *N*-terminal JmjC domain. The PHD fingers of PHF8 and KIAA1718 recognise and bind the trimethylated lysine of H3K4me3, while the JmjC domain removes the methyl group from the lysine residue at position 9 (PHF8) or 27 (KIAA1718).[115] Although the two enzymes share ~82% sequence identity between their PHD fingers and 64% identity between the JmjC domains (Figure 1.13, Appendix 1), their substrates differ; PHF8 catalyses the demethylation of H3K9me2, whereas the reported substrate of KIAA1718 is H3K27me2.[115,141] The substrate specificity is thought to be determined (at least partially) by the different length of the linker region between the PHD finger and the JmjC domain. However, the basis of the substrate specificity of PHF8 and KIAA1718 is not fully understood.

PHF8 and KIAA1718 are of biological importance; mutations in the JmjC domain of human PHF8 (deletions or a missense mutation F279S) cause inherited X-linked mental retardation which is often accompanied by cleft palate and/or cleft lip,[142,143] while KIAA1718 is associated with neural induction and brain inflammation.[144–146]

PHF2 is involved in chondrogenesis and adipogenesis.[147,148] Despite sharing significant sequence identity with PHF8 and KIAA1718, PHF2 was found to be inactive *in vitro*. Some cellular studies suggest that H3K9 is a 'substrate' of PHF2, and western blot analysis indicated that the PHD finger of PHF2 also binds H3K4me2/3.[119,149] PHF2 was found by Baba *et al.* to be activated by Protein Kinase A (PKA)-mediated phosphorylation on four possible serine residues.[150] However, without structural information, the impact and mechanism of these phosphorylation events are uncertain.

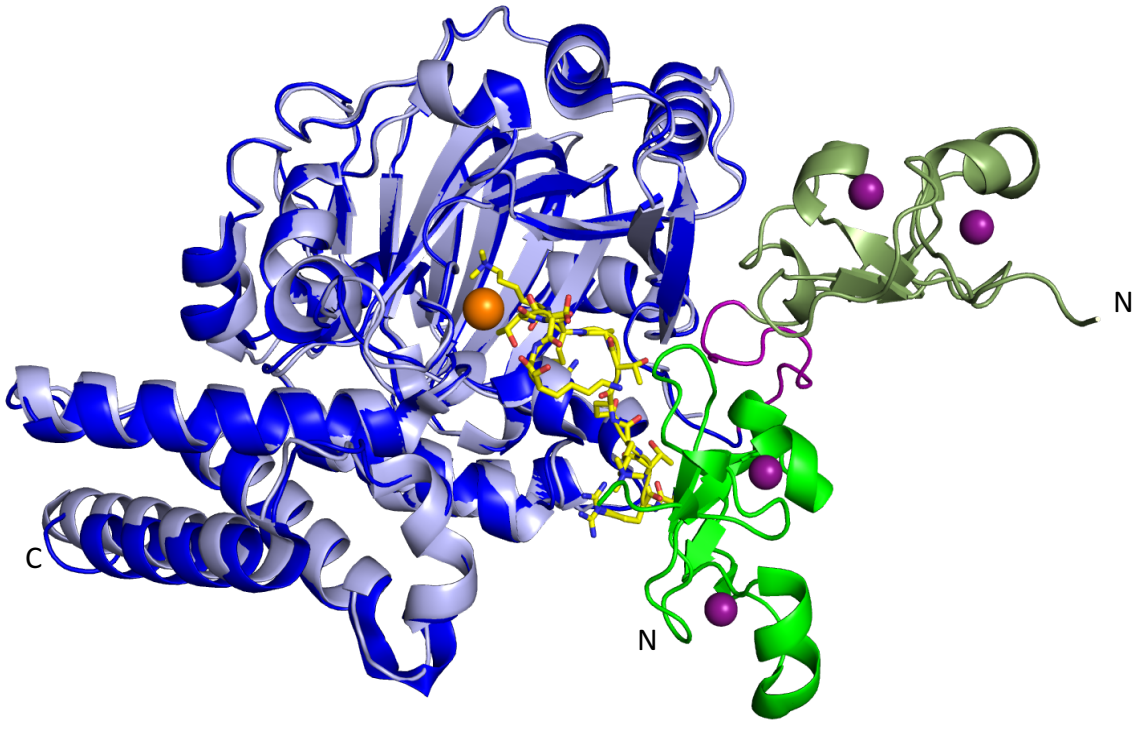


Figure 1.13 View from a crystal structure of PHF8 (light blue for JmjC domain and green for PHD finger, PDB ID: 3kv4) and KIAA1718 (dark blue for JmjC domain and forest green for PHD finger, PDB ID: 3kv5). H3K4me3K9me2 peptide is shown in yellow sticks, zinc ion in purple spheres and iron in orange sphere.

1.8 Thesis objectives

The existing experimental data on the KDM7 subfamily of histone lysine demethylases defined two functional domains and several H3 substrates *in vitro* and in cells. However, the causes of the different demethylation activities across the family were unclear. Although the KDM7 enzymes are of biological importance, no selective inhibitors have been discovered for this family. The work described in the subsequent Chapters of this Thesis concerns the biochemical study of the KDM7 subfamily of enzymes, the underlying mechanism of substrate specificity and diversity, and the identification of novel inhibitors using two screening platforms that are amenable to high-throughput screening.

The specific aims of the research described in this Thesis are to:

1) Characterise the *in vitro* demethylation activity of each member of the KDM7 subfamily;

2) Identify potential new substrates for PHF8 and KIAA1718;

3) Provide insight into the basis of the substrate selectivity and diversity of PHF8 and KIAA1718;

4) Explore the effects of the different iron-binding motifs of the KDM7 enzymes

5) Study the catalytically inactive PHF2 from the perspectives of substrate binding and inhibition;

6) Establish screening platforms to evaluate inhibitors concerning the KDM7 enzymes.

The biochemical characterisation work is described in Chapter 2, while mechanistic studies and inhibition studies are described in Chapter 3 and Chapter 4, respectively. Chapter 5 describes the experimental techniques used in this work.

1.12 References

1 F. Crick, *Nature*, 1970, **227**, 561–3.
2 C. H. Waddington, *The Strategy of The Genes*, 1957.
3 C. H. Waddington, *Int. J. Epidemiol.*, 2012, **41**, 10–13.
4 A. D. Goldberg, C. D. Allis and E. Bernstein, *Cell*, 2007, **128**, 635–638.
5 A. Bird, *Nature*, 2007, **447**, 396–398.
6 S. L. Berger, T. Kouzarides, R. Shiekhattar and A. Shilatifard, *Cold Spring Harb. Lab. Press*, 2009, 781–783.
7 P. Bheda and R. Schneider, *Trends Cell Biol.*, 2014, **24**, 712–723.
8 M. E. Gill, S. Erkek and A. H. F. M. Peters, *Curr. Opin. Cell Biol.*, 2012, **24**, 387–396.
9 R. T. Kamakaka and S. Biggins, *Genes Dev.*, 2005, **19**, 295–310.
10 R. Kornberg, *Science (80-.).*, 1974, **184**, 868–871.
11 K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent and T. J. Richmond, *Nature*, 1997, **389**, 251–260.
12 G. Arents, R. W. Burlingame, B. C. Wang, W. E. Love and E. N. Moudrianakis, *Proc. Natl. Acad. Sci.*, 1991, **88**, 10148–10152.
13 N. Gilbert, S. Boyle, H. Fiegler, K. Woodfine, N. P. Carter and W. A. Bickmore, *Cell*, 2004, **118**, 555–566.
14 C. L. Woodcock and R. P. Ghosh, *Cold Spring Harb. Perspect. Biol.*, 2010, **2**, 1–25.
15 W. Iwasaki, Y. Miya, N. Horikoshi, A. Osakabe, H. Taguchi, H. Tachiwana, T. Shibata, W. Kagawa and H. Kurumizaka, *FEBS Open Bio*, 2013, **3**, 363–369.
16 M. Gao, P. S. Nadaud, M. W. Bernier, J. A. North, P. C. Hammel, M. G. Poirier and C. P. Jaroniec, *J. Am. Chem. Soc.*, 2013, **135**, 15278–15281.
17 K. Saikusa, S. Fuchigami, K. Takahashi, Y. Asano, A. Nagadoi, H. Tachiwana, H. Kurumizaka, M. Ikeguchi, Y. Nishimura and S. Akashi, *Anal. Chem.*, 2013, **85**, 4165–4171.
18 D. T. Jones, *J. Mol. Biol.*, 1999, **292**, 195–202.
19 N. A. Vellore and R. Baron, *ChemMedChem*, 2014, **9**, 484–494.
20 H. Liu and Y. Duan, *Biophys. J.*, 2008, **94**, 4579–4585.
21 D. A. Potoyan and G. A. Papoian, *Proc. Natl. Acad. Sci.*, 2012, **109**, 17857–17862.
22 P. B. Talbert and S. Henikoff, *Nat. Rev. Mol. Cell Biol.*, 2010, **11**, 264–275.
23 I. H. Ismail and M. J. Hendzel, *Environ. Mol. Mutagen.*, 2008, **49**, 73–82.
24 O. Witt, W. Albig and D. Doenecke, *Exp. Cell Res.*, 1996, **306**, 301–306.
25 D. K. Palmer, K. O'Day, H. L. Trong, H. Charbonneau and R. L. Margolis, *Proc. Natl. Acad. Sci. U. S. A.*, 1991, **88**, 3734–8.
26 W. C. Earnshaw and N. Rothfield, *Chromosoma*, 1985, **91**, 313–321.
27 K. Ahmad and S. Henikoff, *Mol. Cell*, 2002, **9**, 1191–1200.
28 P. J. Skene and S. Henikoff, *Development*, 2013, **140**, 2513–2524.
29 J. C. Chuang and P. A. Jones, *Pediatr. Res.*, 2007, **61**, 24R.
30 P. B. Becker and W. Hörz, *Annu. Rev. Biochem.*, 2002, **71**, 247–273.
31 R. Haselkorn and L. B. Rothman-Denes, *Protein Synthesis*, 1973, vol. 1.
32 C. A. Bailey, L. Gerber, A. D. Howard and S. Udenfriend, *Proc. Natl. Acad. Sci.*, 1989, **86**, 22 LP-26.
33 D. A. I. Mavridou, J. M. Stevens, L. Mönkemeyer, O. Daltrop, K. Di Gleria, B. M. Kessler, S. J. Ferguson and J. W. A. Allen, *J. Biol. Chem.*, 2012, **287**, 2342–2352.
34 A. M. Arnaudo and B. A. Garcia, *Epigenetics Chromatin*, 2013, **6**, 24.

35 A. J. Bannister and T. Kouzarides, *Cell Res.*, 2011, **21**, 381.
36 J. A. Law and S. E. Jacobsen, *Nat. Rev. Genet.*, 2010, **11**, 204–220.
37 T. Kouzarides, *Cell*, 2007, **128**, 693–705.
38 K. E. Gardner, C. D. Allis and B. D. Strahl, *J. Mol. Biol.*, 2011, **409**, 36–46.
39 T. Jenuwein and C. D. Allis, *Science (80-.).*, 2001, **293**, 1074 LP-1080.
40 J.-S. Lee, E. Smith and A. Shilatifard, *Cell*, 2010, **142**, 682–685.
41 B. D. Strahl and C. D. Allis, *Nature*, 2000, **403**, 41.
42 T. G. Gillette and J. A. Hill, *Circ. Res.*, 2015, **116**, 1245–1253.
43 T. Zhang, S. Cooper and N. Brockdorff, *EMBO Rep.*, 2015, **16**, 1467 LP-1481.
44 H. Huang, B. R. Sabari, B. A. Garcia, C. D. Allis and Y. Zhao, *Cell*, 2014, **159**, 458–458.e1.
45 Rachel Shamo-Schiller, 2016.
46 T. R. Hebbes, A. W. Thorne and C. Crane-Robinson, *EMBO J.*, 1988, **7**, 1395–1402.
47 T. R. Hebbes, A. L. Clayton, A. W. Thorne and C. Crane-Robinson, *EMBO J.*, 1994, **13**, 1823–1830.
48 W. A. Krajewski and P. B. Becker, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 1540–1545.
49 N. Suka, Y. Suka, A. A. Carmen, J. Wu and M. Grunstein, *Mol. Cell*, 2001, **8**, 473–479.
50 K. Struhl, *Genes Dev.*, 1998, **617**, 599–606.
51 E. Seto and M. Yoshida, *Cold Spring Harb. Perspect. Biol.*, 2014, **6**, a018713.
52 S. Y. Roth, J. M. Denu and C. D. Allis, *Annu. Rev. Biochem.*, 2001, **70**, 81–120.
53 A. J. M. de RUIJTER, A. H. van GENNIP, H. N. CARON, S. KEMP and A. B. P. van KUILENBURG, *Biochem. J.*, 2003, **370**, 737 LP-749.
54 X.-J. Yang and E. Seto, *Oncogene*, 2007, **26**, 5310.
55 M. S. Longworth and L. A. Laimins, *Oncogene*, 2006, **25**, 4495.
56 A. Valenzuela-Fernández, J. R. Cabrero, J. M. Serrador and F. Sánchez-Madrid, *Trends Cell Biol.*, 2008, **18**, 291–297.
57 M. Lachner, R. J. O'Sullivan and T. Jenuwein, *J. Cell Sci.*, 2003, **116**, 2117 LP-2124.
58 M. Lachner, R. J. O\textquoterightSullivan and T. Jenuwein, *J. Cell Sci.*, 2003, **116**, 2117–2124.
59 D. D. Le and D. G. Fujimori, *Curr. Opin. Chem. Biol.*, 2012, **16**, 507–515.
60 F. van Leeuwen, P. R. Gafken and D. E. Gottschling, *Cell*, 2002, **109**, 745–756.
61 J. Min, Q. Feng, Z. Li, Y. Zhang and R.-M. Xu, *Cell*, 2003, **112**, 711–723.
62 A. Patel, V. E. Vought, V. Dharmarajan and M. S. Cosgrove, *J. Biol. Chem.*, 2011, **286**, 3359–3369.
63 S. Rea, F. Eisenhaber, D. O'Carroll, B. D. Strahl, Z.-W. Sun, M. Schmid, S. Opravil, K. Mechtler, C. P. Ponting, C. D. Allis and T. Jenuwein, *Nature*, 2000, **406**, 593.
64 A. Di Lorenzo and M. T. Bedford, *FEBS Lett.*, 2010, **585**, 2024–2031.
65 A. Barski, S. Cuddapah, K. Cui, T.-Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev and K. Zhao, *Cell*, 2007, **129**, 823–837.
66 H. Santos-Rosa, R. Schneider, A. J. Bannister, J. Sherriff, B. E. Bernstein, N. C. T. Emre, S. L. Schreiber, J. Mellor and T. Kouzarides, *Nature*, 2002, **419**, 407.
67 R. Schneider, A. J. Bannister, F. A. Myers, A. W. Thorne, C. Crane-Robinson and T. Kouzarides, *Nat. Cell Biol.*, 2003, **6**, 73.
68 R. Cao, L. Wang, H. Wang, L. Xia, H. Erdjument-Bromage, P. Tempst, R. S. Jones and Y. Zhang, *Science (80-.).*, 2002, **298**, 1039–1043.
69 K. H. Hansen, A. P. Bracken, D. Pasini, N. Dietrich, S. S. Gehani, A. Monrad, J. Rappsilber, M. Lerdrup and K. Helin, *Nat. Cell Biol.*, 2008, **10**, 1291.

70 J. Mellor, *Mol. Cell*, 2009, **35**, 729–730.
71 A. T. Nguyen and Y. Zhang, *Genes Dev.*, 2011, **3**, 1345–1358.
72 Y. Shi, F. Lan, C. Matson, P. Mulligan, J. R. Whetstine, P. A. Cole, R. A. Casero and Y. Shi, *Cell*, 2004, **119**, 941–953.
73 N. S. Scrutton, *Nat. Prod. Rep.*, 2004, **21**, 722–730.
74 A. Karytinos, F. Forneris, A. Profumo, G. Ciossani, E. Battaglioli, C. Binda and A. Mattevi, *J. Biol. Chem.* , 2009, **284**, 17775–17782.
75 Y. Yang, X. Yin, H. Yang and Y. Xu, *Mol. Cell*, 2015, **58**, 47–59.
76 R. Schiller, G. Scozzafava, A. Tumber, J. R. Wickens, J. T. Bush, G. Rai, C. Lejeune, H. Choi, T.-L. Yeh, M. C. Chan, B. T. Mott, J. S. O. McCullagh, D. J. Maloney, C. J. Schofield and A. Kawamura, *ChemMedChem*, 2014, **9**, 566–571.
77 R. J. Klose, E. M. Kallin and Y. Zhang, *Nat. Rev. Genet.*, 2006, **7**, 715–727.
78 F. Forneris, C. Binda, M. A. Vanoni, A. Mattevi and E. Battaglioli, *FEBS Lett.*, 2005, **579**, 2203–2207.
79 S. M. Kooistra and K. Helin, *Nat. Rev. Mol. Cell Biol.*, 2012, **13**, 297.
80 H. Huang, S. Lin, B. A. Garcia and Y. Zhao, *Chem. Rev.*, 2015, **115**, 2376–2418.
81 M. Tan, H. Luo, S. Lee, F. Jin, J. S. Yang, E. Montellier, T. Buchou, Z. Cheng, S. Rousseaux, N. Rajagopal, Z. Lu, Z. Ye, Q. Zhu, J. Wysocka, Y. Ye, S. Khochbin, B. Ren and Y. Zhao, *Cell*, 2011, **146**, 1016–1028.
82 L. Dai, C. Peng, E. Montellier, Z. Lu, Y. Chen, H. Ishii, A. Debernardi, T. Buchou, S. Rousseaux, F. Jin, B. R. Sabari, Z. Deng, C. D. Allis, B. Ren, S. Khochbin and Y. Zhao, *Nat. Chem. Biol.*, 2014, **10**, 365.
83 T. Jiang, X. Zhou, K. Taghizadeh, M. Dong and P. C. Dedon, *Proc. Natl. Acad. Sci.*, 2007, **104**, 60 LP-65.
84 P. Cheung, K. G. Tanner, W. L. Cheung, P. Sassone-Corsi, J. M. Denu and C. D. Allis, *Mol. Cell*, 2000, **5**, 905–915.
85 S. Braun and H. D. Madhani, *EMBO Rep.*, 2012, **13**, 619–630.
86 E. P. Rogakou, D. R. Pilch, A. H. Orr, V. S. Ivanova and W. M. Bonner, *J. Biol. Chem.*, 1998, **273**, 5858–5868.
87 P. N. I. Lau and P. Cheung, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 2801–2806.
88 Y. Wei, L. Yu, J. Bowen, M. A. Gorovsky and C. D. Allis, *Cell*, 1999, **97**, 99–109.
89 D. Rossetto, N. Avvakumov and J. Côté, *Epigenetics*, 2012, **7**, 1098–1108.
90 M. Hochstrasser, *Annu. Rev. Genet.*, 1996, **30**, 405–439.
91 A. Wyce, T. Xiao, K. A. Whelan, C. Kosman, W. Walter, D. Eick, T. R. Hughes, N. J. Krogan, B. D. Strahl and S. L. Berger, *Mol. Cell*, 2007, **27**, 275–288.
92 C. Loenarz and C. J. Schofield, *Nat. Chem. Biol.*, 2008, **4**, 152.
93 I. J. Clifton, M. A. McDonough, D. Ehrismann, N. J. Kershaw, N. Granatino and C. J. Schofield, *J. Inorg. Biochem.*, 2006, **100**, 644–669.
94 P. L. Roach, I. J. Clifton, V. Fülöp, K. Harlos, G. J. Barton, J. Hajdu, I. Andersson, C. J. Schofield and J. E. Baldwin, *Nature*, 1995, **375**, 700.
95 E. L. Hegg and L. Q. Jr, *Eur. J. Biochem.*, 2004, **250**, 625–629.
96 W. Aik, M. A. McDonough, A. Thalhammer, R. Chowdhury and C. J. Schofield, *Curr. Opin. Struct. Biol.*, 2012, **22**, 691–700.
97 K. Valegård, A. C. T. van Scheltinga, M. D. Lloyd, T. Hara, S. Ramaswamy, A. Perrakis, A. Thompson, H.-J. Lee, J. E. Baldwin, C. J. Schofield, J. Hajdu and I. Andersson, *Nature*, 1998, **394**, 805.
98 M. A. McDonough, C. Loenarz, R. Chowdhury, I. J. Clifton and C. J. Schofield, *Curr.*

Opin. Struct. Biol., 2010, **20**, 659–672.
99 J. Zhou, M. Gunsior, B. O. Bachmann, C. A. Townsend and E. I. Solomon, *J. Am. Chem. Soc.*, 1998, **120**, 13539–13540.
100 Z. Zhang, J. Ren, D. K. Stammers, J. E. Baldwin, K. Harlos and C. J. Schofield, *Nat. Struct. Biol.*, 2000, **7**, 127.
101 J. C. Price, E. W. Barr, B. Tirupati, Bollinger J. Martin and C. Krebs, *Biochemistry*, 2003, **42**, 7497–7508.
102 E. Flashman, L. M. Hoffart, R. B. Hamed, J. M. Bollinger Jr, C. Krebs and C. J. Schofield, *FEBS J.*, 2010, **277**, 4089–4099.
103 R. W. D. Welford, J. M. Kirkpatrick, L. A. McNeill, M. Puri, N. J. Oldham and C. J. Schofield, *FEBS Lett.*, 2005, **579**, 5170–5174.
104 L. M. Hoffart, E. W. Barr, R. B. Guyer, J. M. Bollinger and C. Krebs, *Proc. Natl. Acad. Sci.*, 2006, **103**, 14738 LP-14743.
105 H. J. Knowles, R. R. Raval, A. L. Harris and P. J. Ratcliffe, *Cancer Res.*, 2003, **63**, 1764 LP-1768.
106 E. Flashman, S. L. Davies, K. K. Yeoh and C. J. Schofield, *Biochem. J.*, 2010, **427**, 135 LP-142.
107 R. P. Hausinger, *Crit. Rev. Biochem. Mol. Biol.*, 2004, **39**, 21–68.
108 T. Takeuchi, Y. Yamazaki, Y. Katoh-Fukui, R. Tsuchiya, S. Kondo, J. Motoyama and T. Higashinakagawa, *Genes Dev.* , 1995, **9**, 1211–1222.
109 K. Yamane, C. Toumazou, Y. Tsukada, H. Erdjument-Bromage, P. Tempst, J. Wong and Y. Zhang, *Cell*, 2006, **125**, 483–495.
110 W. Ge, A. Wolf, T. Feng, C. H. Ho, R. Sekirnik, A. Zayer, N. Granatino, M. E. Cockman, C. Loenarz, N. D. Loik, A. P. Hardy, T. D. W. Claridge, R. B. Hamed, R. Chowdhury, L. Gong, C. V. Robinson, D. C. Trudgian, M. Jiang, M. M. MacKeen, J. S. McCullagh, Y. Gordiyenko, A. Thalhammer, A. Yamamoto, M. Yang, P. Liu-Yi, Z. Zhang, M. Schmidt-Zachmann, B. M. Kessler, P. J. Ratcliffe, G. M. Preston, M. L. Coleman and C. J. Schofield, *Nat. Chem. Biol.*, 2012, **8**, 960–962.
111 S. Markolovic, T. M. Leissing, R. Chowdhury, S. E. Wilkins, X. Lu and C. J. Schofield, *Curr. Opin. Struct. Biol.*, 2016, **41**, 62–72.
112 T. Lu, M. W. Jackson, B. Wang, M. Yang, M. R. Chance, M. Miyagi, A. V Gudkov and G. R. Stark, *Proc. Natl. Acad. Sci.*, 2010, **107**, 46 LP-51.
113 V. K. C. Ponnaluri, D. T. Vavilala, S. Putty, W. G. Gutheil and M. Mukherji, *Biochem. Biophys. Res. Commun.*, 2009, **390**, 280–284.
114 R. L. Hancock, K. Dunne, L. J. Walport, E. Flashman and A. Kawamura, *Epigenomics*, 2015, **7**, 791–811.
115 J. R. Horton, A. K. Upadhyay, H. H. Qi, X. Zhang, Y. Shi and X. Cheng, *Nat. Struct. Mol. Biol.*, 2010, **17**, 38–44.
116 S. Tu, Y.-C. Teng, C. Yuan, Y.-T. Wu, M.-Y. Chan, A.-N. Cheng, P.-H. Lin, L.-J. Juan and M.-D. Tsai, *Nat. Struct. &Amp; Mol. Biol.*, 2008, **15**, 419.
117 W. Yao, Y. Peng and D. Lin, *Biochem. Biophys. Res. Commun.*, 2010, **396**, 323–328.
118 N. P. Blackledge, J. C. Zhou, M. Y. Tolstorukov, A. M. Farcas, P. J. Park and R. J. Klose, *Mol. Cell*, 2010, **38**, 179–190.
119 J. R. Horton, A. K. Upadhyay, H. Hashimoto, X. Zhang and X. Cheng, *J. Mol. Biol.*, 2011, **406**, 1–8.
120 X. de la Cruz, S. Lois, S. Sánchez-Molina and M. A. Martínez-Balbás, *BioEssays*, 2005, **27**, 164–175.

121 S. Chakravarty, L. Zeng and M.-M. Zhou, *Structure*, 2009, **17**, 670–679.
122 F. Chignola, M. Gaetani, A. Rebane, T. Org, L. Mollica, C. Zucchelli, A. Spitaleri, V. Mannella, P. Peterson and G. Musco, *Nucleic Acids Res.*, 2009, **37**, 2951–2961.
123 S. Iwase, B. Xiang, S. Ghosh, T. Ren, P. W. Lewis, J. C. Cochrane, C. D. Allis, D. J. Picketts, D. J. Patel, H. Li and Y. Shi, *Nat. Struct. &Amp; Mol. Biol.*, 2011, **18**, 769.
124 S. Eustermann, J.-C. Yang, M. J. Law, R. Amos, L. M. Chapman, C. Jelinska, D. Garrick, D. Clynes, R. J. Gibbons, D. Rhodes, D. R. Higgs and D. Neuhaus, *Nat. Struct. &Amp; Mol. Biol.*, 2011, **18**, 777.
125 Y. Li and H. Li, *Acta Biochim. Biophys. Sin. (Shanghai).*, 2012, **44**, 28–39.
126 M. Bienz, *Trends Biochem. Sci.*, 2006, **31**, 35–40.
127 G. G. Wang, J. Song, Z. Wang, H. L. Dormann, F. Casadio, H. Li, J.-L. Luo, D. J. Patel and C. D. Allis, *Nature*, 2009, **459**, 847.
128 T. Org, F. Chignola, C. Hetényi, M. Gaetani, A. Rebane, I. Liiv, U. Maran, L. Mollica, M. J. Bottomley, G. Musco and P. Peterson, *EMBO Rep.*, 2008, **9**, 370 LP-376.
129 H. van Ingen, F. M. A. van Schaik, H. Wienk, J. Ballering, H. Rehmann, A. C. Dechesne, J. A. W. Kruijzer, R. M. J. Liskamp, H. T. M. Timmers and R. Boelens, *Structure*, 2008, **16**, 1245–1256.
130 R. E. Mansfield, C. A. Musselman, A. H. Kwan, S. S. Oliver, A. L. Garske, F. Davrazou, J. M. Denu, T. G. Kutateladze and J. P. Mackay, *J. Biol. Chem.* , 2011, **286**, 11779–11791.
131 C. A. Musselman, R. E. Mansfield, A. L. Garske, F. Davrazou, A. H. Kwan, S. S. Oliver, H. O'Leary, J. M. Denu, J. P. Mackay and T. G. Kutateladze, *Biochem. J.*, 2009, **423**, 179 LP-187.
132 P. Filippakopoulos and S. Knapp, *FEBS Lett.*, 2012, **586**, 2692–2704.
133 S. Mujtaba, Y. He, L. Zeng, S. Yan, O. Plotnikova, Sachchidanand, R. Sanchez, N. J. Zeleznik-Le, Z. Ronai and M.-M. Zhou, *Mol. Cell*, 2004, **13**, 251–263.
134 S. Mujtaba, L. Zeng and M.-M. Zhou, *Oncogene*, 2007, **26**, 5521.
135 C. Chen, T. J. Nott, J. Jin and T. Pawson, *Nat. Rev. Mol. Cell Biol.*, 2011, **12**, 629.
136 Y. Yang and M. T. Bedford, *Nat. Rev. Cancer*, 2012, **13**, 37.
137 R. Paro and D. S. Hogness, *Proc. Natl. Acad. Sci. U. S. A.*, 1991, **88**, 263–267.
138 J. C. Eissenberg, *Gene*, 2001, **275**, 19–29.
139 K. Bouazoune, A. Mitterweger, G. Längst, A. Imhof, A. Akhtar, P. B. Becker and A. Brehm, *EMBO J.*, 2002, **21**, 2430–2440.
140 J. Li, Z. Li, J. Ruan, C. Xu, Y. Tong, P. W. Pan, W. Tempel, L. Crombet, J. Min and J. Zang, *PLoS One*, 2011, **6**, e25104.
141 C. Loenarz, W. Ge, M. L. Coleman, N. R. Rose, C. D. O. Cooper, R. J. Klose, P. J. Ratcliffe and C. J. Schofield, *Hum. Mol. Genet.*, 2009, **19**, 217–222.
142 F. E. Abidi, M. G. Miano, J. C. Murray and C. E. Schwartz, *Clin. Genet.*, 2007, **72**, 19–22.
143 D. Kleine-Kohlbrecher, J. Christensen, J. Vandamme, I. Abarrategui, M. Bak, N. Tommerup, X. Shi, O. Gozani, J. Rappsilber, A. E. Salcini and K. Helin, *Mol. Cell*, 2010, **38**, 165–178.
144 A. Yokoyama, Y. Okuno, T. Chikanishi, W. Hashiba, H. Sekine, R. Fujiki and S. Kato, *Genes to Cells*, 2010, **15**, 867–873.
145 C. Huang, Y. Xiang, Y. Wang, X. Li, L. Xu, Z. Zhu, T. Zhang, Q. Zhu, K. Zhang, N. Jing and C. D. Chen, *Cell Res.*, 2010, **20**, 154–165.
146 T. Suganuma and J. L. Workman, *Cell Res.*, 2010, **20**, 861–862.
147 Y. Okuno, F. Ohtake, K. Igarashi, J. Kanno, T. Matsumoto, I. Takada, S. Kato and Y. Imai, *Diabetes*, 2013, **62**, 1426–1434.

148 J. H. Park, M. Jung and K. C. Moon, *Oncotarget*, 2018, **9**, 142–151.
149 K. Hata, R. Takashima, K. Amano, K. Ono, M. Nakanishi, M. Yoshida, M. Wakabayashi, A. Matsuda, Y. Maeda, Y. Suzuki, S. Sugano, R. H. Whitson, R. Nishimura and T. Yoneda, *Nat. Commun.*, 2013, **4**, 1–11.
150 A. Baba, F. Ohtake, Y. Okuno, K. Yokota, M. Okada, Y. Imai, M. Ni, C. A. Meyer, K. Igarashi, J. Kanno, M. Brown and S. Kato, *Nat. Cell Biol.*, 2011, **13**, 668–675.

Chapter 2 Biochemical Studies on the KDM7 subfamily

2.1 Chapter overview

The work described in this Chapter concerns the biochemical characterisation of three members of human histone lysine demethylase (KDM) 7 subfamily, i.e. KIAA1718 (KDM7A), PHF8 (KDM7B) and PHF2 (KDM7C). Initially, the *in vitro* demethylation activity of each KDM7 was investigated using MALDI-TOF MS-based assays. Further work investigated the substrate selectivity of PHF8 and KIAA1718, resulting in identification of potential new substrates for these enzymes. Finally, the potential interaction of modified lysine peptides with KDM7s was investigated; the results provide insight into the cross talk between the PHD finger and JmjC catalytic domains of each protein and the contributions of the PHD fingers to substrate specificity.

The modified H1.4 peptides used in this Chapter were synthesised by Dr Richard Hopkinson.

2.2 Introduction

The KDM7 subfamily of human KDMs consists of three members: KIAA1718 (KDM7A), PHF8 (KDM7B), and PHF2 (KDM7C).[1] The three enzymes share significant sequence identity and similarity (Appendix 1&2, Sequence alignment of KDM7 enzymes). All the members of the KDM7 subfamily contain a PHD finger and a JmjC catalytic domain (Figure 2.1). The PHD fingers bind the histone tail at H3K4me3, an interaction which is proposed to tune the substrate selectivity of the JmjC catalytic domains of KIAA1718 and PHF8.[2] The JmjC domains contain an iron-binding motif (HxD…H/Y) at their active sites, which is the site of demethylation of dimethyl- or monomethyl-lysine residues, at H3K9me1/2 and maybe other residues (see below).

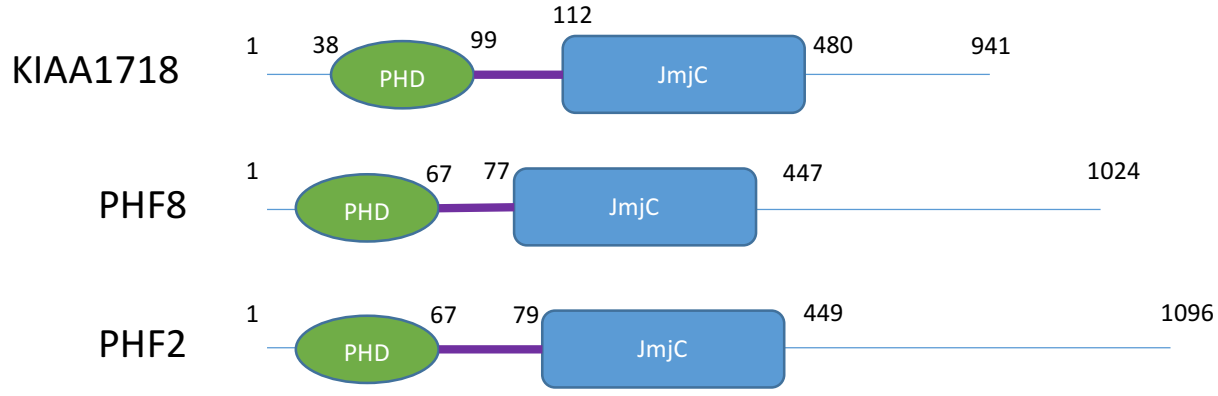


Figure 2.1 Schematic of the domains of the KDM7 enzymes. The structures of each construct are depicted based on the full-length protein. Residue numbers are shown above each schematic. The linker regions of each enzyme are highlighted in bold purple.[2,3]

Members of the KDM7 subfamily are implicated in animal development and cell growth. For example, mutations in the PHD finger and JmjC domain of the human PHF8 gene are linked to midline defects such as autism, X-linked mental retardation, and cleft lip and palate.[4–7] Additionally, knockdown of the zebrafish PHF8 homologue causes delayed brain development

and correlates with abnormal facial growth.[8] PHF8 has also been reported to regulate HIF1α hypoxia signalling by binding to H3K4me3, and to contribute to the progression of castration-resistant prostate cancer (CRPC).[9,10] Overexpression studies with KIAA1718 using a mouse embryonic stem cell (ESC) differentiation system, and in early chick embryos, have provided evidence that KIAA1718 promotes neural induction.[11,12] KIAA1718 is also reported to mediate the upregulation of intercellular adhesion molecule 1 (ICAM1) which is involved in the adhesion and transmigration of leukocytes across the endothelium, contributing to the progression of brain inflammation.[13] PHF2 plays an important role in adipogenesis. [14,15] Thus, knockout studies in mice and knock-down studies in 3T3-L1 cells suggested that PHF2 promotes fat cell differentiation by binding to the gene promotors of adipogenic transcription factors.[14,15] Cell-based assays demonstrated that PHF2 regulates the expression of p53 by demethylating H3K9me2 and is thus associatied with p53-mediated cancers.[14,16] PHF2 is also reported to inhibit rDNA transcription by competing with PHF8 for H3K4me2/3 binding in cells.[17] However, no published data has shown *in vitro* evidence of lysine demethylation catalysed by PHF2.

The aforementioned biological studies mean it is important to define the biochemical roles of the human KDM7 proteins. PHF8 is reported to be an H3K9me1/2 and H3K36me2 demethylase, while KIAA1718 accepts H3K27me1/2 and H3K9me1/2 as substrates.[2,12,18–21] PHF8 is also reported to demethylate H4K20me1 in cells and on nucleosomal substrates *in vitro*, but the cellular antibody based studies have not been supported by *in vitro* work with peptidic substrates.[8,22]

On substrate binding to PHF8, crystallographic studies reveal that the linker region between the PHD finger and JmjC domain of PHF8 adopts a bent conformation (Figure 2.2 A), such that the binding to H3K4me3 facilitates the catalytic domain binding to the K9 position.[2] The linker region of KIAA1718 is three amino acid residues longer than that of PHF8, and has been suggested to result in a more rigid and extended conformation (Figure 2.2 B), allowing the active site of the JmjC domain to bind at the K27 position when binding to H3K4me3.[2] The JmjC domain of KIAA1718 is therefore positioned further far away from the PHD finger, preventing the demethylation activity at K9 by binding of the PHD finger to H3K4me3. Biochemical studies on the *in vitro* demethylation activity of PHF8 and KIAA1718 have been carried out using formaldehyde dehydrogenase coupled (FDH) assay (Table 2.1); however, the factors contributing to their substrate specificity have not been fully investigated.

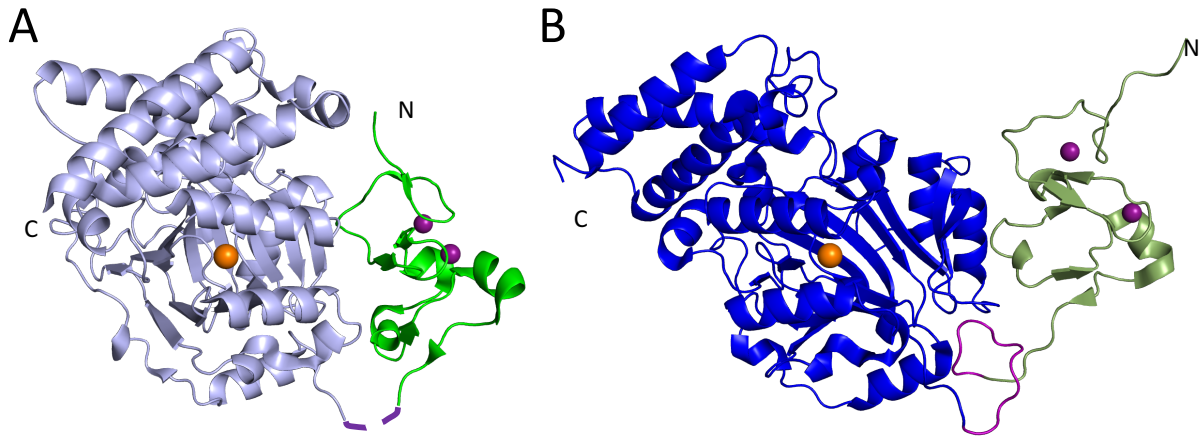


Figure 2.2 View from a crystal structure of **A)** PHF8 (PDB ID: 3kv4) JmjC domain (light blue), PHD finger (green) and disordered linker region (purple dashed line). **B)** KIAA1718 (PDB ID: 3kv5) JmjC domain (blue), PHD finger (forest green) and extended linker region (purple loop). Zinc ion shown in purple sphere and iron in orange.

Unlike PHF8 and KIAA1718, the catalytic activity of PHF2 remains questionable. The unusual active site binding motif of PHF2 suggests it might be catalytically inactive. PHF2 contains a tyrosine (Y321) in place of the second histidine residue (H319 for PHF8 and H354 for KIAA1718) within the canonical 2OG oxygenase iron binding motif, although PHF2 has been

demonstrated to bind Ni(II) with a K_D of 58 ± 4 μM.[3] PHF2 was reported to demethylate H3K9me1 *in vivo* by western analysis, and the PHD finger of PHF2 to bind H3K4me3.[23] A study by Baba *et al.* suggests that PHF2 becomes active upon the phosphorylation of a serine (S1056) by protein kinase A (PKA) allowing it to demethylate H3K9me2 on native histones *in vitro*.[24] However, no demethylation on peptidic substrates by PHF2 has been discovered to date, and thus no further biochemical studies have been reported. To help address the substrate selectivity of KDM7 enzymes, the *in vitro* activities were characterised using MALDI-TOF MS assays.

Table 2.1 The kinetic parameters from the previous reported biochemical work on KDM7 subfamily.[2]

Name	Substrate(s)	K_M (substrate)/μM	Method
PHF8 $_{PHD\text{-}JmjC}$	$H3_{1\text{-}24}$K4me3K9me2	8.0 ± 1.2	FDH
	$H3_{1\text{-}24}$K9me2	134 ± 18	
KIAA1718 $_{PHD\text{-}JmjC}$	$H3_{1\text{-}24}$K4me3K9me2	--	FDH
	$H3_{1\text{-}24}$K9me2	< 1.2	

2.3 Results and discussion

The KDM7 enzymes appear to represent an excellent model system for studying how non-catalytic domains regulate activities at the catalytic JmjC domain thus were selected for detailed biochemical analyses. The expression constructs of the recombinant KDM7 enzymes used and their respective sequences in this thesis are given in Table 2.2. The constructs included those for dual domain proteins (i.e. proteins containing both PHD finger and JmjC domain), as well as the JmjC domains and PHD fingers separately (except for KIAA1718 where the JmjC domain only construct was not generated). Purification tags were cleaved for all recombinant KDM7 proteins used in this Chapter, unless otherwise stated.

Table 2.2 The KDM7 constructs used in this Thesis. Constructs used in previous studies are also shown (residue number*). The protein tags used for protein production are defined.

Name	Residue number	Residue number*	Tag
PHF8 $_{PHD-JmjC}$	1 - 447	1 – 447[2,25]	TrxT
PHF8 $_{JmjC}$	78 - 447	62 – 447[2]	TrxT
PHF8 $_{PHD}$	1 - 65	2 - 66[26]	TrxT
KIAA1718 $_{PHD-JmjC}$	38 - 480	1 - 488 / 38 - 488[2]	TrxT
KIAA1718 $_{PHD}$	38 - 96	--	TrxT
PHF2 $_{PHD-JmjC}$	1 - 449	1 - 451[3]	Z-Basic
PHF2 $_{JmjC}$	59 - 449	60 - 451[3]	Z-Basic
PHF2 $_{PHD}$	1 - 67	1 - 70[3]	Z-Basic

2.3.1 Protein expression and purification

The expression methods for the production of all the KDM7 subfamily derived proteins were similar. The targeted enzyme sequences were cloned into the appropriate T7-promoter containing vectors (pET-28-a backbone) using the ligand independent cloning (LIC-cloning) protocol (Chapter 5, Materials and Methods).[27] The plasmids of interest were transformed and expressed into DE3 lysogen containing *Escherichia coli* (*E. coli*) cells , and the harvested

cell pellets were stored at -80 °C before purification. Detailed protocols can be found in Chapter 5 Materials and Methods.

The purification methods for the KDM7 derived proteins were based on the purification method reported for KIAA1718 by Rose *et al*., and were optimised for each protein construct.[28] All the KDM7 enzymes used in this Thesis are shown on two SDS-PAGE gels (Figure 2.3). This section describes the purification of wild-type KDM7 enzymes; the expression trials and purification procedures for variants of KDM7 proteins can be found in Chapter 3. All proteins where purification tags were cleaved were purified using three sequential columns (a Ni-affinity or an ion-exchange column, a reverse Ni-affinity column after tag-cleavage and a gel filtration column). Recombinant wild-type KDM7 proteins were at least 90% pure as judged by SDS-PAGE analysis (Figure 2.3 A, lane 1-7), except for $PHF2_{JmjC}$ which had some apparent degradation products (or other impurities) present (Figure 2.3 A, lane 8). The tagged proteins were subjected to two column chromatography purifying procedures (a Ni-affinity column or an ion-exchange column, followed by a gel filtration column) and were at least 80% pure by SDS-PAGE analysis (Figure 2.3 B, lane 1-8). $PHF2Y321H_{JmjC}$ appeared to show bands with larger molecular weights than that calculated (Figure 2.3 A lane 12 & Figure 2.3 B lane 11), possibly due to binding of $PHF2Y321H_{JmjC}$ with contaminants or oligomerisation.

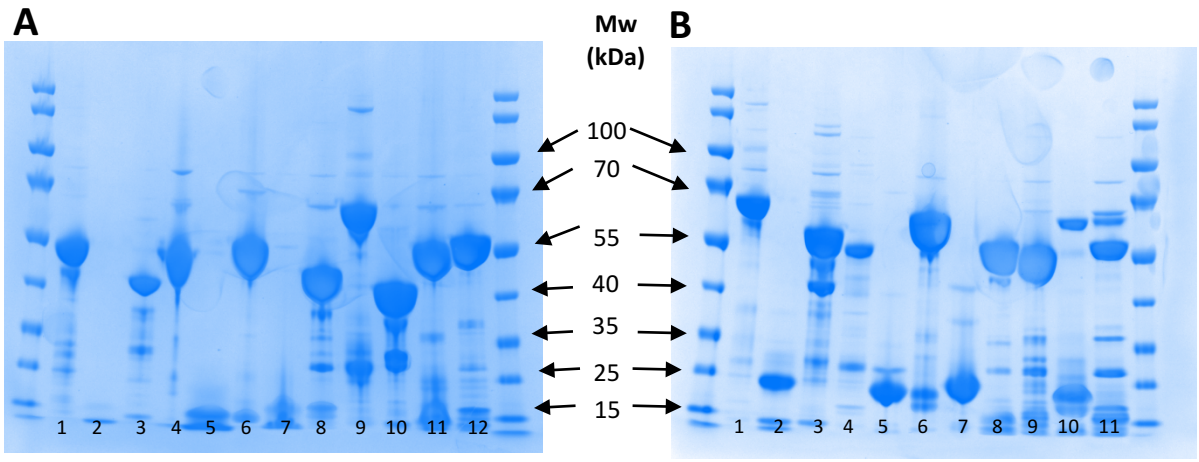


Figure 2.3 SDS-PAGE gel images analyses of purified **A)** non-tagged and **B)** tagged KDM7 enzymes. Calculated masses of the untagged and tagged proteins, respectively, are given. **1:** $PHF8_{PHD\text{-}JmjC}$, 51.7K and 65.5K; **2:** $PHF8_{PHD}$, 7.5K and 21.5K; **3:** $PHF8_{JmjC}$, 42.9K and 56.7K; **4:** $KIAA1718_{PHD\text{-}JmjC}$, 51.4K and 65.5K; **5:** $KIAA1718_{PHD}$, 6.8K and 21.0K; **6:** $PHF2_{PHD\text{-}JmjC}$, 51.3K and 60.3K; **7:** $PHF2_{PHD}$, 7.7K and 16.7K; **8:** $PHF2_{JmjC}$, 44.6K and 53.6K; **9:** $PHF8H319Y_{PHD\text{-}JmjC}$, 51.7K and 65.5K (note that for this protein only, the non-tagged protein is in image B and tagged protein is in image A); **10:** $PHF8H319Y_{JmjC}$, 42.9K (Figure A), $PHF2Y31H_{PHD\text{-}JmjC}$ 60.3K (Figure B); **11:** $PHF2Y31H_{PHD\text{-}JmjC}$, 51.3K (Figure A), $PHF2Y321H_{JmjC}$ 53.6K (Figure B); **12:** $PHF2Y321H_{JmjC}$, 44.6K (Figure A).

For PHF8 and KIAA1718, the proteins with a TrxT tag (using pNH-TrxT vector) showed high solubility and stability during the chromatographic purification. The TrxT tags also improved the expression levels of proteins (typically 10 mg purified protein was obtained from 1 L cell culture) when compared to proteins constructs in the pNIC28-Bsa4 vector, which encodes for proteins with a hexa-histidine (His_6) tag,). Therefore, the pNH-TrxT vector was selected as the plasmid of choice for the large-scale expression of PHF8 and KIAA1718. To obtain >90% puifiued protein (by SDS-PAGE), the target protein was purified from cell lysates using at least two chromatography steps, i.e. a Ni-affinity column followed by a size exclusion column. The cleavage of tagged protein was carried out using the Tobacco Etch Virus (TEV) protease[29], which was usually carried out *via* overnight incubation after the first Histrap® column.

All three members of KDM7 subfamily were found to precipitate and degrade during the purification procedure (Figure 2.4 showing PHF8 as an example), indicating that minimising the duration of the purification process would render a more active batch of protein. After

optimisation, a 24-hour purification protocol (over two days) was established, which enabled apparently full cleavage of the tags and avoided visible precipitation. In summary the optimised procedure was as follows: the stored cell pellet was thawed, sonicated and then purified using a HisTrap™ column (except for PHF2 that used a SP Seph column), all within a 4 hour period. TEV catalysed cleavage was then immediately initiated (over 12 hours). The resultant protein solution (including PHF2) was next quickly (< 5min) passed through another HisTrap™ column to remove the his-tagged TEV protease and uncleaved protein (if still present) before running through a gel filtration column over 2 hours. In the final step, protein was concentrated using Amicon® ultra centrifugal filters (MilliPore), then aliquoted (2 hours). Therefore, the overall purification process lasted no more than 24 hours (see detailed protocol in Chapter 5 Materials and Methods). Importantly, it was found that freshly prepared TEV protease (i.e. the material used had to have been purified within 3 months) was needed to achieve nearly 100% cleavage within 12 hours and to avoid any wastage *via* uncleaved protein.

Different columns were tested to optimise the purification protocol. The use of Histrap FF® (fast flow) and Histrap HP® (high performance) 5 mL columns resulted in a similar, if not the same, protein purity and purification speed. Gel filtration columns Superdex™ 75 10/300 GL (S75) and Superdex™ 200 10/300 GL (S200) are technically capable of separating all the JmjC domain-containing proteins, since their molecular weights were within the separation range of both columns (Table 2.3).[30] However, in practice, use of the S75 column yielded higher purity than the S200 column, especially when the protein loading was high. Further, the PHD finger proteins could only be purified by the S75 column given their relatively low molecular

weight (typically 7k Da without a tag). Therefore, the S75 column was used throughout the purification of all proteins described in this Thesis.

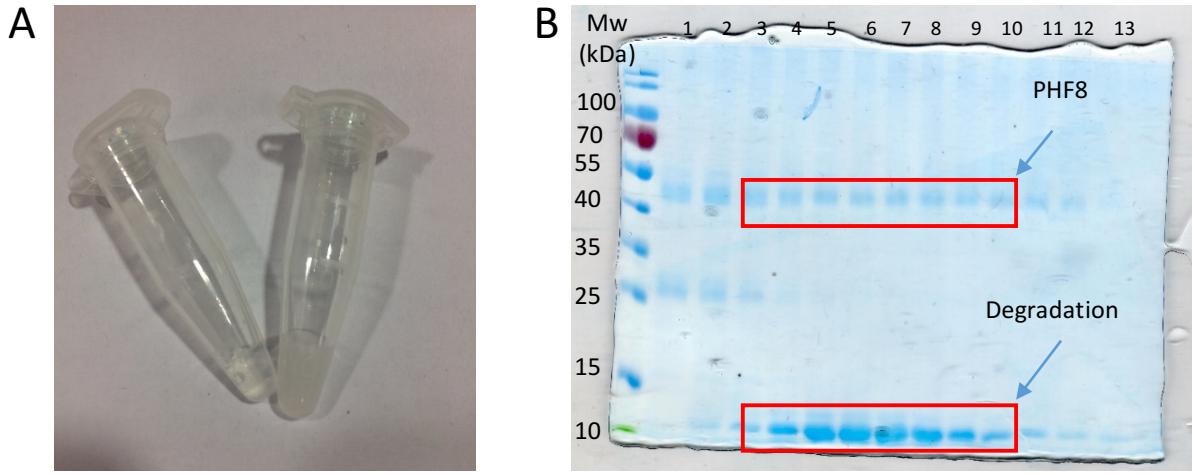


Figure 2.4 PHF8 can precipitate during the purification process. **A)** PHF8 protein solution after Histrap column (left) and after overnight TEV cleavage (right). **B)** SDS-PAGE gel image showed degradation of pNIC28-Bsa4-PHF8$_{PHD\text{-}JmjC}$ after a S75 gel filtration column chromatography. Lane 1-13 shows 13 fractions from gel filtration column. Lane 3-10 were exampled from elution peak. Flow rate: 1 ml/min; temperature :4 °C; buffer: 50 mM HEPES (pH 7.5), 500 mM NaCl, 1 mM TCEP.

Table 2.3 Column data of S75 10/300 GL column and S200 10/300 GL column from GE Healthcare®.[30]

	S75	***S200***
Matrix	Cross-linked agarose and dextran	
Average Particle Size	13 µm	
pH Stability Range	3 to 12	
Exclusion Limit	Approx. 1×10^5	Approx. 1.3×10^6
Optimum Separation Range	**3,000 – 70,000**	**10,000 – 600,000**
Maximum Flow Rate r.t.	1.5 ml/min	1.0 ml/min

For PHF2, Ni-affinity chromatography was avoided since previous studies by Xian Deng (a former Master student in Schofield group) indicated that nickel ions could destabilise PHF2 protein (further studies of metal binding can be found in Chapter 3 Section 3.2.6). To obtain sufficiently stable and pure PHF2 protein, optimisation of the purification procedure was carried out based on that established for PHF8 and KIAA1718. An ion-exchange (IEX) column was used to replace the HisTrap™ column in the first step of the PHF2 purification protocol in order to prevent the protein from contacting nickel ions. Based on Xian Deng's study, IEX +

size-exclusion column was chosen to purify PHF2. A considerable amount of degradation during purification procedure was observed for PHF8 when using the pNIC28-Bsa4 vector (Figure 2.4B). In an attempt to minimise degradation, the PHF2 gene was subcloned into the pNIC-ZB vector (containing a Z-Basic tag). The highly charged Z-Basic tag is an engineered tag based on the sequence of protein Z and is used to achieve selective capture of target proteins.[31] The strong ion binding affinity of Z-Basic tag renders the tagged protein to bind more tightly to the IEX column over other protein contaminants[31]. This enables the tagged protein to be eluted using a high-salt buffer (500 mM NaCl) and most of the impurities to be removed using a low-salt wash buffer. Although degradation of PHF2 protein was still observed during IEX column purification (Figure 2.5 A), the size-exclusion column after TEV cleavage removed most of the apparent degradation products (Figure 2.5 B). The overall PHF2 purification process was carried out in less than 24 hours, as optimised for PHF8 and KIAA1718, to minimise degradation. Moreover, because the Z-Basic tag significantly improved protein solubility, the overall yield was also improved. Finally, to keep proteins stable at -80 °C, 5% glycerol (v/v final) was added into the protein solution before snap-freezing using liquid nitrogen.

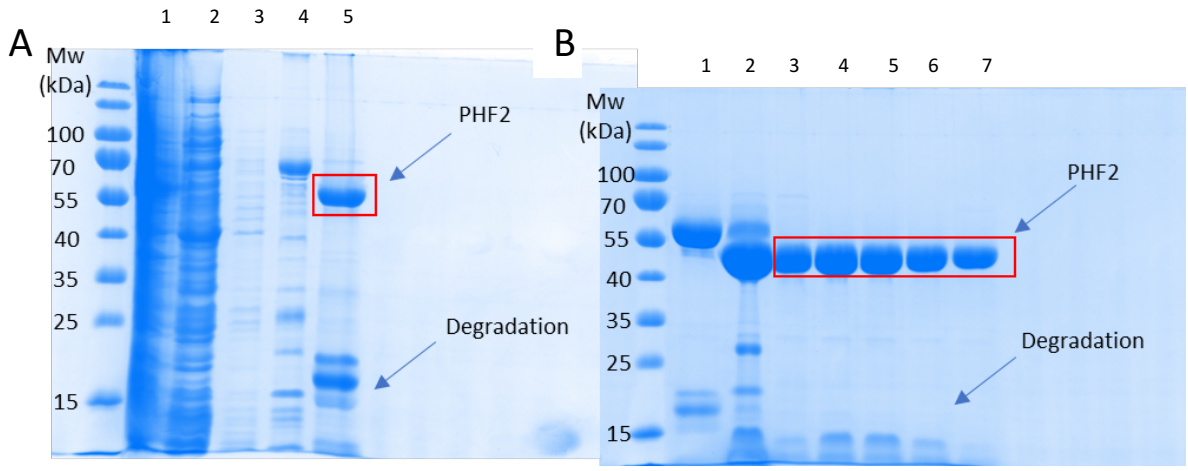


Figure 2.5 Purification of PHF2 using the pNIC-ZB vector. **A)** SDS-PAGE gel image shows the purity of PHF2 fractions from SP seph column. Degradation occurred at low molecular weight area. Lane 1: Cell lysate (filtered); Lane 2: Flow-through; Lane 3: 40 mM NaCl wash; Lane 4: 200 mM NaCl wash; Lane 5: 500 mM NaCl elution. **B)** SDS-PAGE gel image is showing the cleaved PHF2 protein with most of the degradation removed by size-exclusion column. Lane 1: PHF2 with tag; Lane 2: PHF2 with tag cleaved (column load); Lane 3-7: gel filtration fractions. Final PHF2 protein could achieve over 90% pure.

2.3.2 Optimisation of *in vitro* assay conditions for demethylation of dimethyllysine by KDM7 family enzymes

Several publications have reported *in vitro* assays to measure the activity of recombinant KDM7 enzymes under a variety of conditions.[2,3,18,21] Loenarz *et al.* have reported the demethylation of a $H3_{1-19}K9me1$ peptide by PHF8 in 50 mM Tris-HCl (2-amino-2-(hydroxymethyl)-1,3-propanediol hydrochloride) buffer (pH 7.5).[18] Horton *et al*. measured the demethylation activities of both PHF8 and KIAA1718 against $H3_{1-24}K9me2$ and $H3_{1-35}K27me2$ in 50 mM HEPES (4-(2-hydroxyethyl) piperazine-1-ethanesulfonic acid) buffer (pH 7.0).[2] ceKIAA1718 (*Caenorhabditis elegans* ortholog of KIAA1718) was reported to catalyse the removal of the methyl marks at H3K9 and H3K27 in 20 mM Tris-HCl buffer (pH 7.5) by Hanqing Lin *et al*.[21] Thus, to select optimal assay conditions for the KDM7s in enzymatic assays, demethylation reactions catalysed by the KDM7s in both HEPES and Tris-HCl buffers at pH 7.5 were initially examined.

In the buffer optimisation assays, Fe(II) (200 µM), ascorbate (2 mM) and 2OG (2 mM) were used in excess in an attempt to ensure maximum activity. 100 mM NaCl was also included in the buffers to prevent the enzymes from precipitating. The KDM7 enzyme (5 µM) and histone substrate (10 µM) concentrations were kept constant throughout the buffer optimisation assays. The demethylation reaction mixture was incubated at 37 °C for 1 hour before being quenched by 2% (v/v) aqueous formic acid. All end-point reaction mixtures were analysed by MALDI-TOF MS. For PHF8 and PHF2, the tested buffers did not substantially influence their demethylation activities, as the extent of demethylation observed was almost the same in HEPES or Tris-HCl (Figure 2.6 A & B). For KIAA1718, however, its demethylation activity against $H3_{1-15}K9me2$ was doubled in HEPES buffer compared to Tris-HCl; interestingly the extent of $H3_{1-15}K4me3K9me2$ demethylation remained similar in the two buffers (Figure 2.6 C). Based on these results (and to remain consistent with protein purification (section 2.3.1)), HEPES was used for all subsequent KDM7 assays.

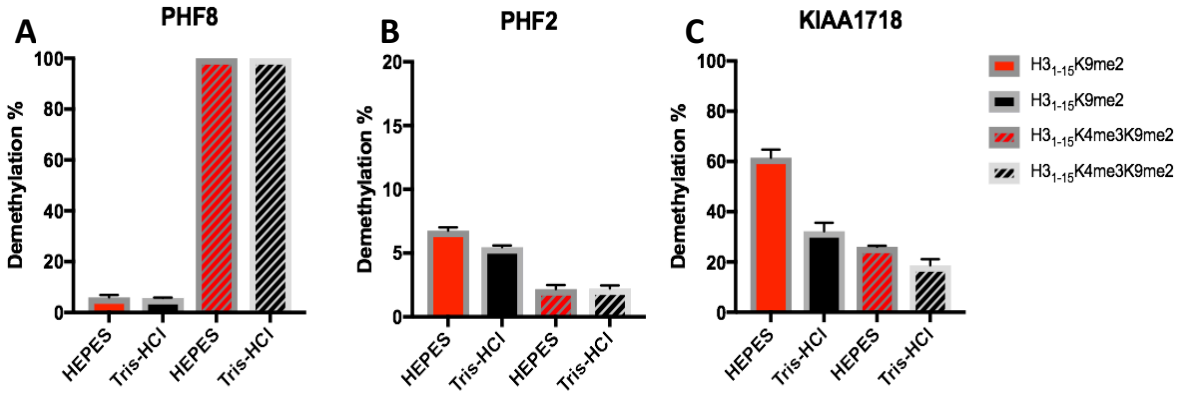


Figure 2.6 KDM7 enzyme catalsyed demethylation of $H3_{1-15}K4me3K9me2$ or $H3_{1-15}K9me2$ in either 50 mM HEPES buffer or 50 mM Tris-HCl buffer (containing 100 mM NaCl). Red: HEPES; Black: Tris-HCl. Solid colour: $H3_{1-15}K9me2$; Slash pattern: $H3_{1-15}K4me3K9me2$. **A)** PHF8. **B)** PHF2. **C)** KIAA1718. Reactions were quenched by 2% (v/v) aqueous formic acid after 1-hour incubation at 37 °C. The extent of demethylation (percentage) was analysed by MALDI-TOF MS. Figures were generated using Prism® software. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical duplicates, and error bars give ± one s.d.

Next, the effects of different pH values on demethylation were tested (pH 6.0, 7.0, 8.0 and 9.0). Demethylation reactions were carried out in the same way as for buffer optimisation, but this time in all cases using 50 mM HEPES as the buffer. For isolated PHF8, demethylation activities against both peptides were the highest in pH 7.0 buffer (Figure 2.7 A & B). Interestingly, demethylation of $H3_{1-15}$K9me2 by PHF8 was only observed at pH 7.0 (Figure 2.7 A), suggesting that the interaction between H3K9me2 and the JmjC domain of PHF8 might be more susceptible to pH changes than the H3K4me3 interaction. The relationship between pH and KIAA1718 activity on $H3_{1-15}$K4me3K9me2 followed the same trend as observed for PHF8 (Figure 2.7 D). Unlike PHF8, KIAA1718 demethylated $H3_{1-15}$K9me2 at pH 7.0, 8.0 and 9.0 (Figure 2.7 C), although the greatest extent of demethylation of $H3_{1-15}$K9me2 by KIAA17718 was still observed at pH 7.0. PHF2 only showed at most 5% (if any) possible demethylation activity under the conditions tested, which made it difficult to carry out further characterisations (Figure 2.7 E & F). These results suggest that pH had a significant influence on KDM7 enzymatic activity and that of the tested conditions pH 7.0 was optimal for PHF8 and KIAA1718. Therefore, pH 7.0 was selected for the *in vitro* demethylation assays for the KDM7s.

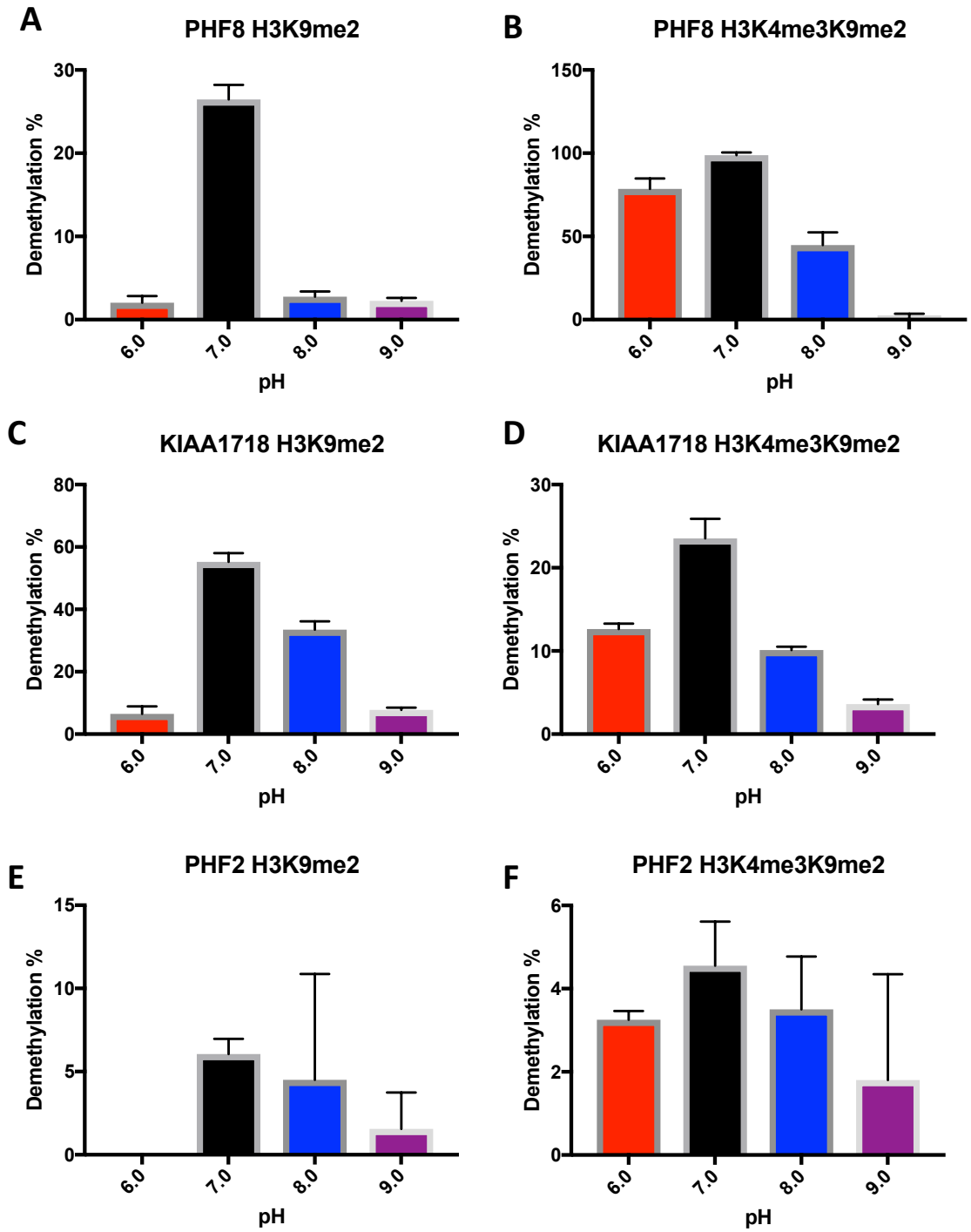


Figure 2.7 KDM7 enzyme catalsyed demethylation of $H3_{1-15}$K4me3K9me2 and $H3_{1-15}$K9me2 in 50 mM HEPES buffer, 100 mM NaCl, various pH. Red: pH 6.0; Black: pH 7.0; Blue: pH 8.0; Purple: pH 9.0. **A)** PHF8 against $H3_{1-15}$K9me2. **B)** PHF8 against $H3_{1-15}$K4me2K9me2. **C)** KIAA1718 against $H3_{1-15}$K9me2. **D)** KIAA1718 against $H3_{1-15}$K4me3K9me2. **E)** PHF2 against $H3_{1-15}$K9me2. **F)** PHF2 against $H3_{1-15}$K4me3K9me2. Reactions were quenched with 2% (v/v) aqueous formic acid after 1-hour incubation at 37 °C. Demethylation percentage was analysed by MALDI-TOF MS. Figures were generated using Prism® software. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

The last factor examined was the concentration of NaCl, which is known to have a significant effect on some KDM catalysed demethylase activities.[32] The demethylation reactions of KDM7 enzymes against both $H3_{1-15}$K4me3K9me2 and $H3_{1-15}$K9me2 were investigated. The buffers used in the screening assays were 50 mM HEPES buffer (pH7.0) with 0, 50, 100, and 500 mM NaCl. The experimental procedure was the same as for the investigations on the effects of the buffering system and pH. The extent of demethylation after 1-hour incubation was measured under each condition. For PHF8, with $H3_{1-15}$K4me3K9me2, 100% of demethylation was achieved after 1-hour incubation under all the conditions tested (Figure 2.8 B), indicating that PHF8 is very active with $H3_{1-15}$K4me3Kme2, and so may not be influenced substantially by the NaCl concentration, under the tested conditions. The final demethylation conversions of $H3_{1-15}$K9me2 were less than 40% for all tested conditions; the highest percentage demethylation was achieved with 50 mM NaCl (Figure 2.8 A).

To further investigate the effect of salt on PHF8 and $H3_{1-15}$K4me3K9me2, time course analyses were recorded where demethylation reactions were quenched at various time points. The results reveal that the concentration of NaCl negatively affect the initial rate of the demethylation of PHF8 using $H3_{1-15}$K4me3K9me2 (Figure 2.9). Reactions under all conditions reached a plateau after 15 minutes, but the reaction in 50 mM NaCl buffer manifested the highest reaction rate. Notably, some white precipitation was observed upon diluting enzyme stock into the buffer without NaCl. Therefore, a 50 mM HEPES, pH 7.0, 50 mM NaCl buffer was chosen for PHF8 assays based on a combination of the enzymatic activity and stability considerations.

With KIAA1718, low salt conditions (0, 50 mM NaCl) were found to be preferred for $H3_{1-15}$K9me2 demethylation (Figure 2.8 C); the highest extent of $H3_{1-15}$K4me3K9me2

demethylation was achieved in 50 and 100 mM NaCl buffer (Figure 2.8 D). PHF2 did not show obvious demethylation under any of the condition examined. Therefore, the same buffer as PHF8 was selected for subsequent KIAA1718 demethylation studies.

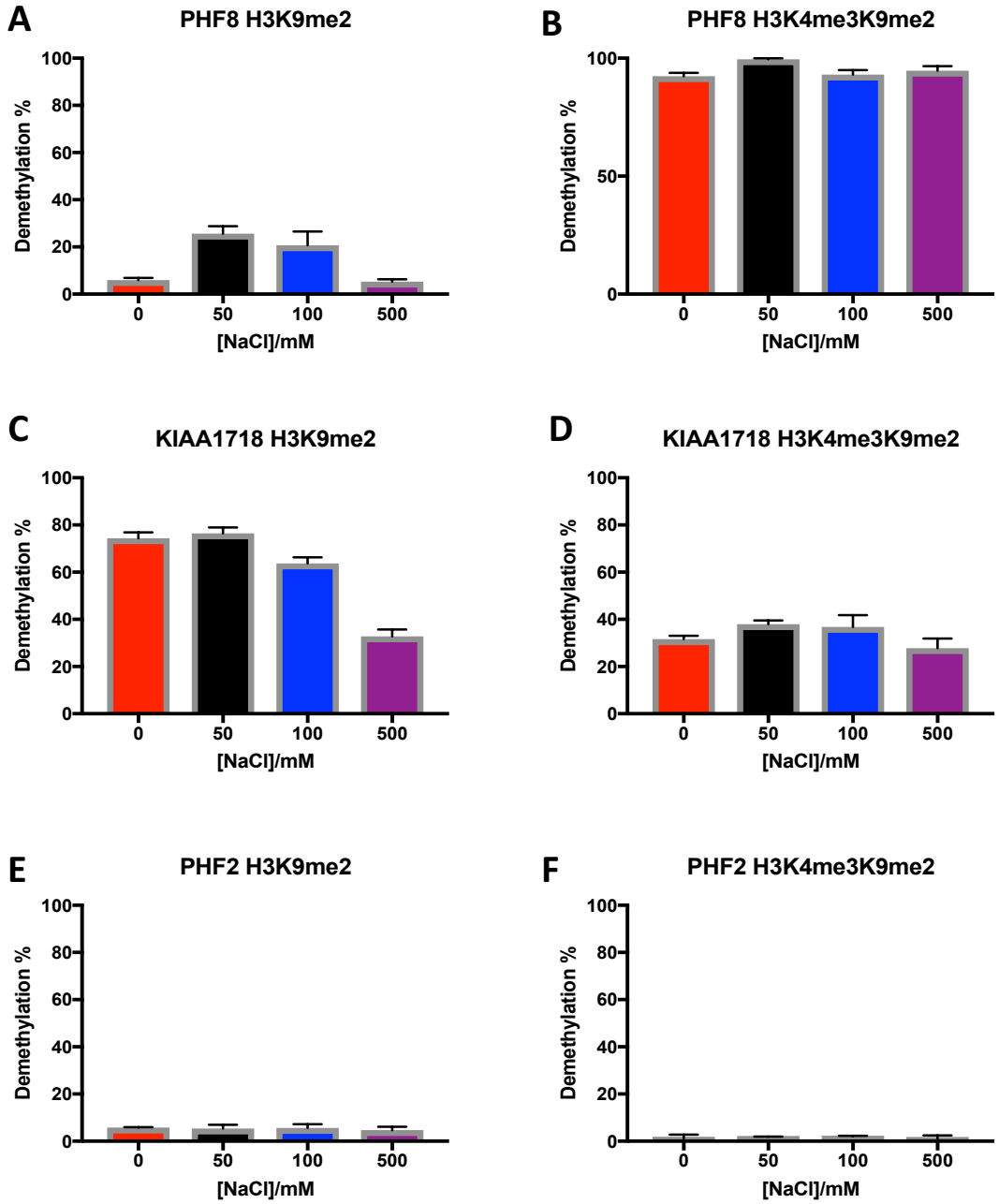


Figure 2.8 KDM7 enzymes catalysed demethylation of $H3_{1-15}K4me3K9me2$ and $H3_{1-15}K9me2$ in 50 mM HEPES buffer (pH 7.0) with various concentration of NaCl. Red: 0 mM; Black: 50 mM; Blue: 100 mM; Purple: 500 mM. **A)** PHF8 against $H3_{1-15}K9me2$. **B)** PHF8 against $H3_{1-15}K4me3K9me2$. **C)** KIAA1718 against $H3_{1-15}K9me2$. **D)** KIAA1718 against $H3_{1-15}K4me2K9me2$. **E)** PHF2 against $H3_{1-15}K9me2$. **F)** PHF2 against $H3_{1-15}K4me3K9me2$. Reactions were quenched by 2% (v/v) aqueous formic acid after 1-hour incubation at 37 °C. The extent of demethylation was analysed by MALDI-TOF MS. Figures were generated using Prism® software. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

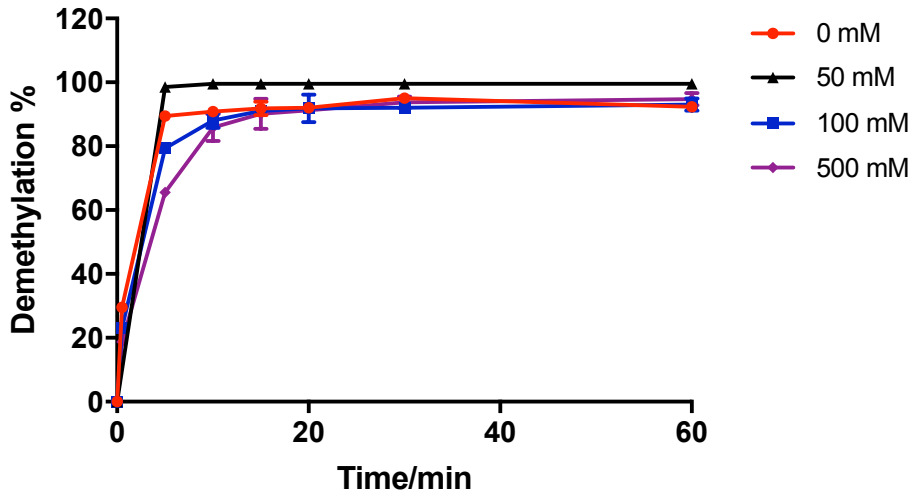


Figure 2.9 Time courses of PHF8 catalysed demethylation of $H3_{1-15}$K4me3K9me2. Reactions were quenched by 2% (v/v) aqueous formic acid after certain time period of incubation at 37 °C. The extent of demethylation was analysed by MALDI-TOF MS. Figures were generated using Prism® software. A pre-quenched sample was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

To summarise the results described in this section, a graphic of the conditions tested is shown in Figure 2.10, with the optimal condition highlighted in bold blue.

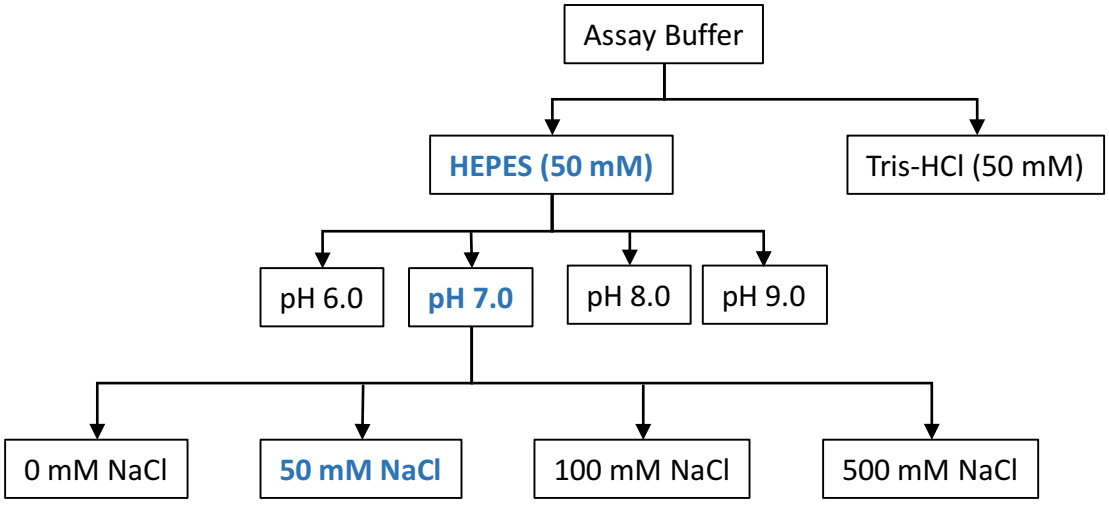


Figure 2.10 *In vitro* demethylation assay buffer optimisation for KDM7s.

2.3.3 Investigations on the *in vitro* demethylation activities of PHF8

Previous work reported the demethylation activities of PHF8 against H3K9me1/me2, H3K36me1/me2 and H4K20me1 substrates, both with isolated proteins and in cells.[2,8,18] Notably, Horton *et al.* demonstrated that the demethylation activity of PHF8 on H3K9me2 is promoted by the presence of H3K4me3.[2] To further investigate the substrate preference of PHF8, and to validate the activity of isolated PHF8, time courses of $PHF8_{PHD\text{-}JmjC}$ with $H3_{1\text{-}15}$K4me3K9me2 and $H3_{1\text{-}15}$K9me2 peptides were recorded using the optimised buffer conditions (section 2.3.2). In these assays, 50 µM of Fe (II), 1 mM of 2OG and 2 mM of ascorbate were used to match the literature conditions.[2]

The time course experiments revealed that $PHF8_{PHD\text{-}JmjC}$ has relatively weak / slow activity on the $H3_{1\text{-}15}$K9me2 peptide, with a $t_{1/2}$ of more than 1 hour (Figure 2.11 A). By contrast, for the peptide modified by methylation at both H3K4me3 and H3K9me2, the activity of PHF8 was enhanced by more than 12-fold, with a $t_{1/2}$ of less than 5 minutes (Figure 2.11 B), implying that the presence of H3K4me3 significantly increases PHF8 activity on H3K9me2. The individually plotted figure of $H3_{1\text{-}15}$K9me0/1/2 also revealed an interesting pattern of demethylation. The curve of $H3_{1\text{-}15}$K9me1 started to flatten and reached its steady state after 60 min (Figure 2.11 C, blue), indicating that H3K9me1 accumulates and is demethylated into $H3_{1\text{-}15}$K9me0 at a similar rate. Since the increment of $H3_{1\text{-}15}$K9me1 equals to the demethylation of $H3_{1\text{-}15}$K9me2, the plateau of $H3_{1\text{-}15}$K9me1 curve suggests that PHF8 also has a similarly low activity for $H3_{1\text{-}15}$K9me1 as compared to $H3_{1\text{-}15}$K9me2. On the other hand, for the $H3_{1\text{-}15}$K4me3K9me2 peptide, the relative abundance of $H3_{1\text{-}15}$K4me3K9me2 dropped to about 10% within 5 minutes while the curve for $H3_{1\text{-}15}$K4me3K9me1 mostly stayed flat,

indicating that the monomethyl group of $H3_{1-15}K4me3K9me1$ was removed efficiently by PHF8 (Figure 2.11 D). This discrimination in reaction rates indicates that PHF8 has a higher activity for $H3_{1-15}K4me3K9me1$ compared to $H3_{1-15}K9me1$.

In conclusion, the presence of H3K4me3 enhances PHF8 activity on both H3K9me2 and H3K9me1, consistent with the work by Horton *et al.*[2] The $PHF8_{PHD-JmjC}$ construct used in this Thesis was the same as their work, but 15 mer peptides replaced 24 mer due to their availability. To understand the possible effects of the length of substrate peptide, time courses against 21 mer H3K4me3K9me2 and H3K9me2 were then measured using the same conditions.

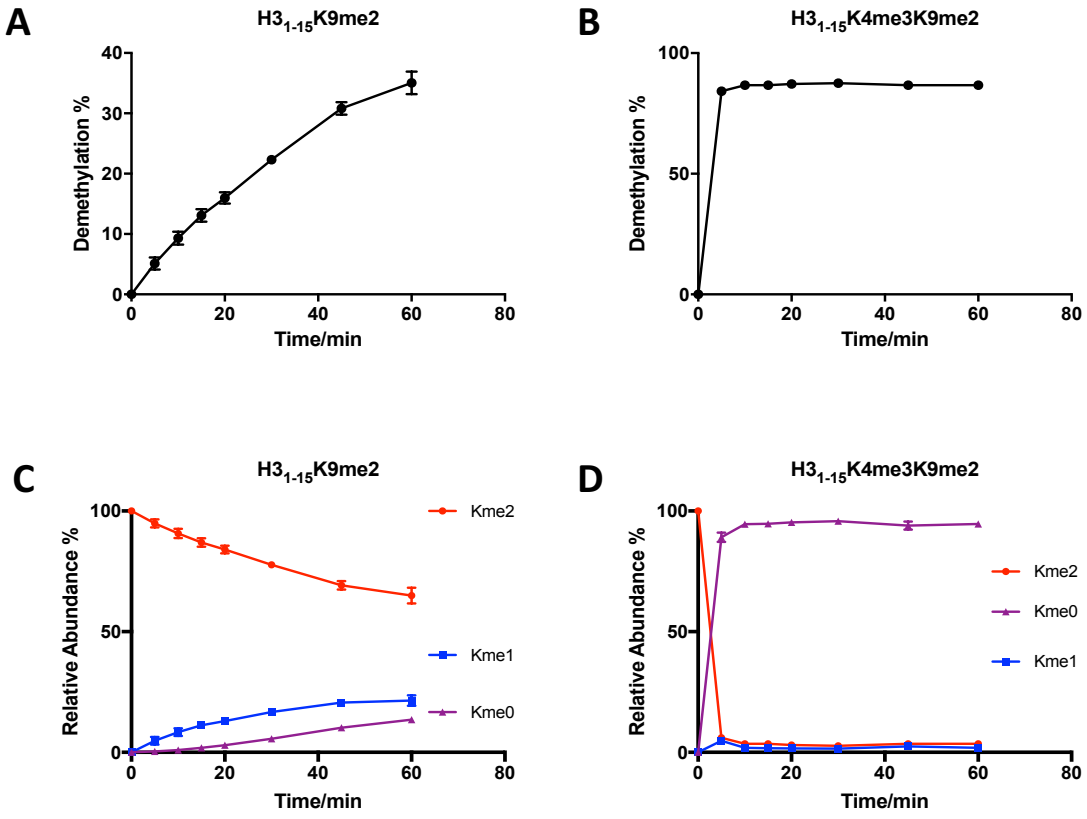


Figure 2.11 $PHF8_{PHD-JmjC}$ demethylation assays using MALDI-TOF MS. **A)** Overall demethylation time course of $H3_{1-15}K9me2$. **B)** Overall demethylation time course of $H3_{1-15}K4me3K9me2$. **C)** Individual plots of Kme0, Kme1 and Kme2 of $H3_{1-15}K9me2$. **D)** Individual plots of Kme0, Kme1 and Kme2 of $H3_{1-15}K4me3K9me2$. A pre-quenched sample was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

The addition of 6 extra amino acids to the *C*-terminal of the substrate boosted PHF8 activity against H3K9me2 by at least 30-fold, with a $t_{1/2}$ of 2 min (Figure 2.12 A); the overall demethylation of $H3_{1-21}$K4me3K9me2 was similar to that of the 15 mer peptide (Figure 2.12 B), with a $t_{1/2}$ of less than 2 min. Two minutes was the earliest time point measured due to constraints of the experimental set up. Therefore, it is difficult to precisely compare the PHF8 activity against 21 mer and 15 mer H3K4me3K9me2 peptides since their $t_{1/2}$ values cannot be measured accurately using this chosen method. However, close inspection of the individual plots of H3K9me0/1/2 indicates that the increase of H3K9me0 (21 mer) was slower than that of H3K9me0 (15 mer), (Figure 2.11 D, purple and Figure 2.12 D, purple), with a much larger accumulation of the monomethylated peptide. Moreover, the time for 90% conversion to the fully demethylated peptide was 5 min for the 15 mer peptide and 1 hour for the 21 mer. The slower accumulation of H3K9me0 indicated slower turnover of H3K9me1 suggests that the activity of PHF8 for H3K4me3K9me1 is attenuated by elongation of the peptide, though activity on the H3K4me3K9me2 is not significantly influenced.

To summarise the results described in this section, $PHF8_{PHD\text{-}JmjC}$ demethylated dimethyl- and monomethyl- K9 on histone 3 tail peptides. Demethylation of H3K9me2 and H3K9me1 by PHF8 was boosted by the presence of H3K4me3, and the length of peptidic substrates was found to influence PHF8 activity (Table 2.4).

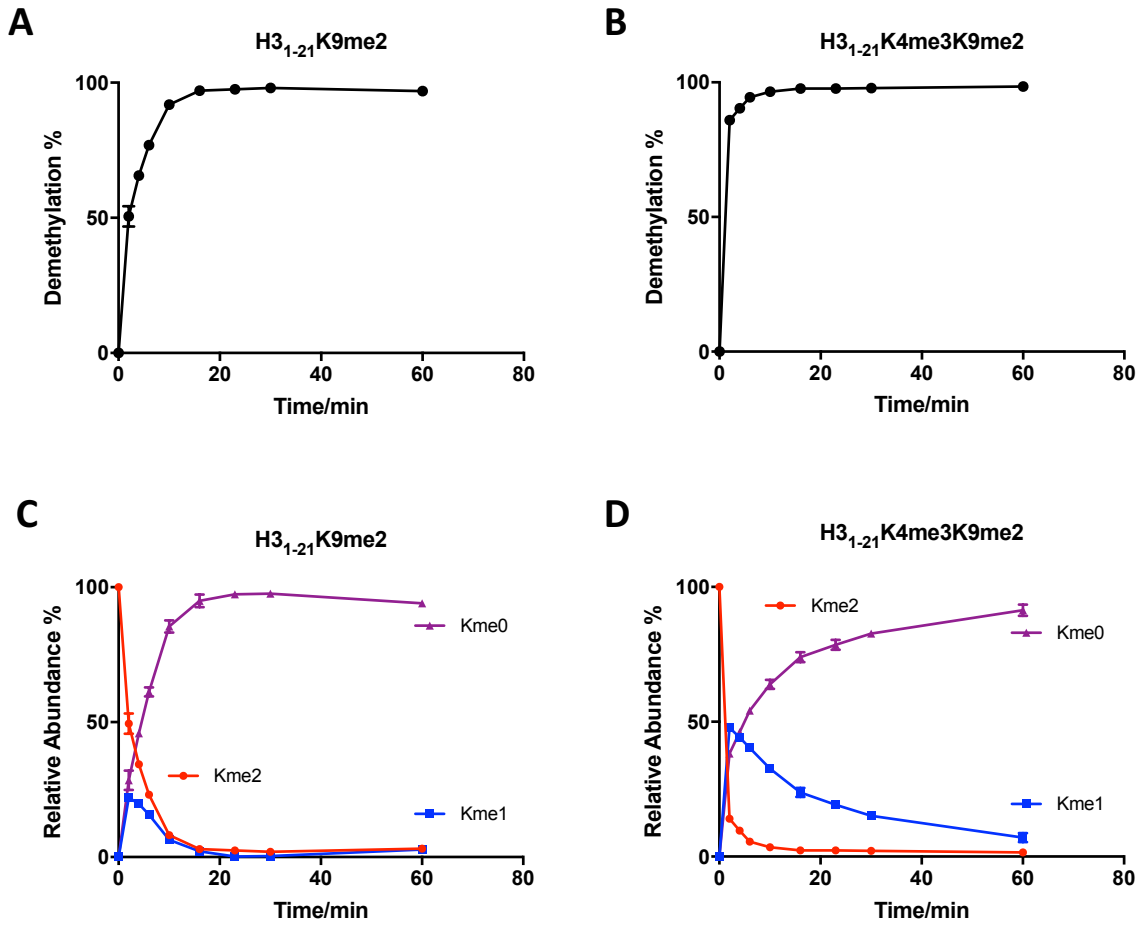


Figure 2.12 Demethylation of histone peptides by PHF8$_{PHD-JmjC}$. **A)** Overall demethylation time course of H3$_{1-21}$K9me2. **B)** Overall demethylation time course of H3$_{1-21}$K4me3K9me2. **C)** Individual plots of Kme0, Kme1 and Kme2 of H3$_{1-21}$K9me2. **D)** Individual plots of Kme0, Kme1 and Kme2 of H3$_{1-21}$K4me3K9me2. Figures were generated using Prism® software. A pre-quenched sample was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

Table 2.4 Summary of the K9me2 demethylation reactions catalysed by PHF8. The values marked with asterisks are reported in the literature.[2]

	H3K4me3K9me2			H3K9me2		
	15 mer	**21 mer**	**24 mer**	**15 mer**	**21 mer**	**24 mer**
$t_{1/2}$	< 5 min	< 2 min	5 min*	> 60 min	2 min	60 min*
90% conversion	5 min	4 min	20 min*	> 60 min	10 min	> 60 min*

2.3.4 Investigations on KIAA1718 catalysed demethylation reactions

KIAA1718 is reported to demethylate $H3_{1-24}K9me2$ and $H3_{1-35}K27me2$.[2,21] According to studies by Horton *et al.*, the binding of KIAA1718 to H3K4me3 reduces demethylation of H3K9me2 while facilitating that of H3K27me2. To investigate this substrate preference, demethylation assays with a series of peptides were carried out with $KIAA1718_{PHD\text{-}JmjC}$. Time courses were undertaken for each peptide using the same *in vitro* demethylation assay conditions as for $PHF8_{PHD\text{-}JmjC}$ (Section 2.3.3). $t_{1/2}$ of histone substrates and the time for 90% conversion were used to compare the activities with different peptides. Time-courses were carried out with H3K9me2 (15 mer and 21 mer), H3K4me3K9me2 (15 mer and 21 mer), H3K27me2 (21 mer) and H3K4me3K27me2 (35 mer).

Compared to $PHF8_{PHD\text{-}JmjC}$, $KIAA1718_{PHD\text{-}JmjC}$ showed relatively high demethylation activity with $H3_{1-15}K9me2$, with a $t_{1/2}$ of 20 min (Figure 2.13 A). In contrast to $PHF8_{PHD\text{-}JmjC}$, the presence of H3K4me3 decreased KIAA1718 activity on H3K9me2, with the $t_{1/2}$ increasing to more than 60 min (Figure 2.13 B), though not as significantly as that observed by Horton et al, possibly due to the differences in peptide substrate length. The accumulation rate of $H3_{1-15}K9me0$ was also faster than $H3_{1-15}K4me3K9me0$, indicating that demethylation of $H3_{1-15}K9me1$ is also attenuated by the presence of H3K4me3 (Figure 2.13 C & D).

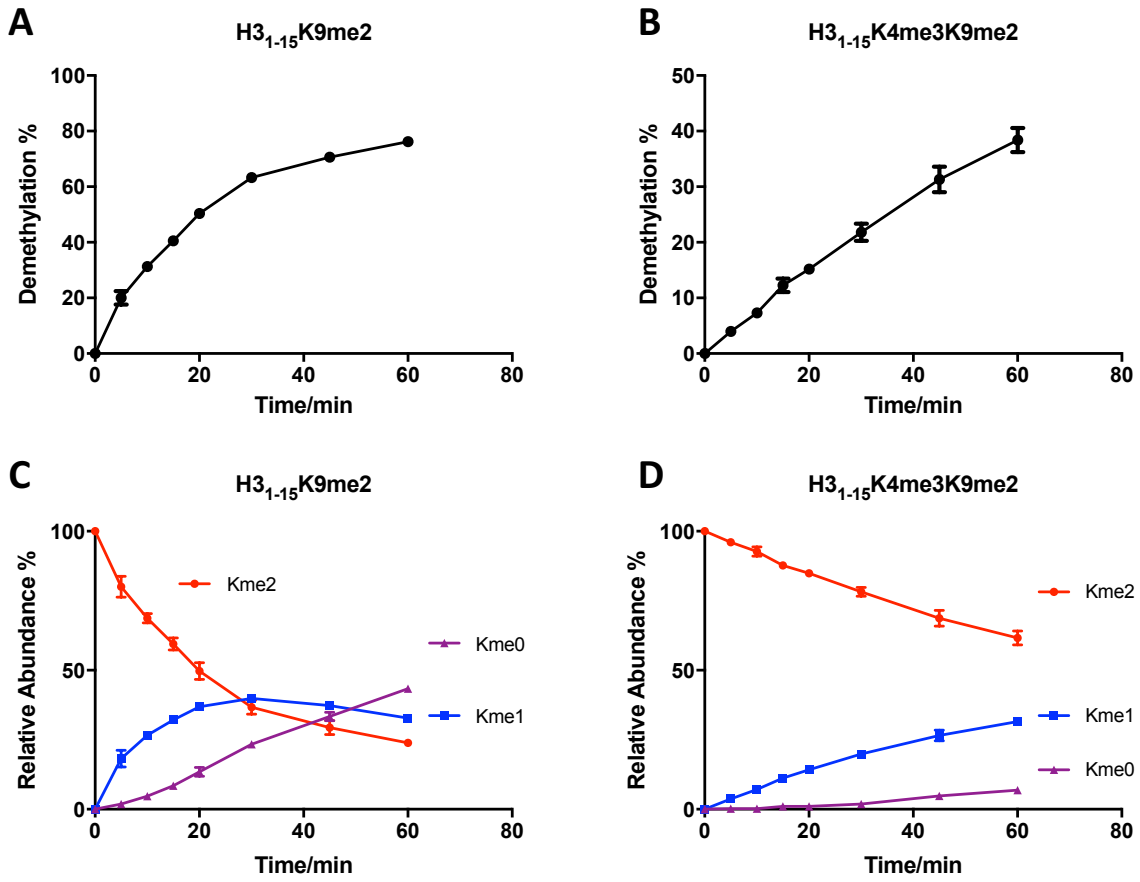


Figure 2.13 KIAA1718$_{PHD-JmjC}$ demethylation assays by MALDI-TOF MS. **A)** Overall demethylation time course of H3$_{1-15}$K9me2. **B)** Overall demethylation time course of H3$_{1-15}$K4me3K9me2. **C)** Individual plots of Kme0, Kme1 and Kme2 of H3$_{1-15}$K9me2. **D)** Individual plots of Kme0, Kme1 and Kme2 of H3$_{1-15}$K4me3K9me2. Figures were generated using Prism® software. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

To further investigate the substrate specificity of KIAA1718 and the effect of peptide length, other demethylation assays were carried out using the 21 mer histone H3K4me3K9me2 and H3K9me2 peptides. KIAA1718 activity on 21 mer H3K9me2, which gave a $t_{1/2}$ of 3 min, was increased compared to the 15 mer peptide (Figure 2.14 A). By contrast, demethylation of H3$_{1-21}$K4me3K9me2 was nearly abolished by the addition of the six amino acids to the *C*-terminal (Figure 2.14 B). These results are consistent with the published data using 24 mer peptides, where the $t_{1/2}$ of H3$_{1-24}$K9me2 was 4 min, supporting the assumption that the shorter peptide length was responsible for the differences observed with the 15 mer peptides.[2]

The opposite trend was observed with PHF8, where the presence of H3K4me3 substantially enhanced PHF8 activity against H3K9me2. One explanation for this difference is the different length linker (Figure 2.1). According to Horton *et al*., when H3K4me3 and H3K9me2 are both present, the binding of PHD finger to H3K4me3 acts like an anchor, and the demethylation of H3K9me2 by JmjC domain follows.[2] The linker KIAA1718 is longer than PHF8 and adopts a rigid conformation. It is likely that when the PHD finger is bound to H3K4me3, the distance between H3K4 and H3K9 becomes too short for the JmjC domain of KIAA1718 to reach the dimethyl-lysine at position 9. On the other hand, since the linker of PHF8 is shorter and more flexible, position 9 is in an optimal position for the JmjC domain to capture H3K9me2.

It is interesting that as the length of peptide is increased from 15 mer to 21 mer the demethylation of H3K9me2 catalysed by KIAA1718 becomes more efficient (Figure 2.13 & 2.14, also see Table 2.5). For PHF8, the highest demethylation activity observed was with the 21 mer peptide and was dramatically decreased with the 15 and 24 mer peptides (Figure 2.11 & 2.12, also see Table 2.4). However, this difference on the effect of peptide length remains to be investigated further.

It was also proposed that H3K4me3K27me2 is a better substrate than H3K4me3K9me2 for KIAA1718 as the distance between H3K4 and H3K27 is preferable.[2] To confirm this theory, the demethylation of H3K27 peptides was next investigated.

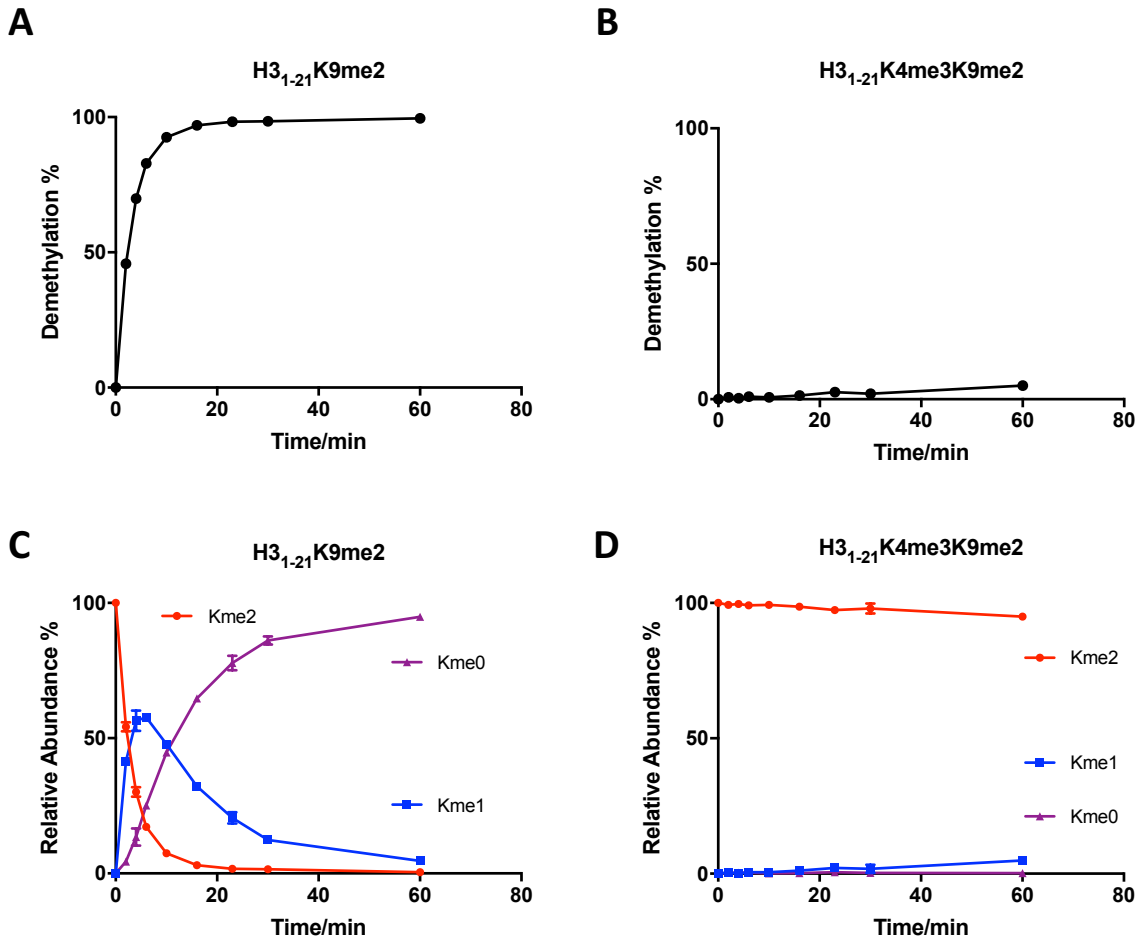


Figure 2.14 Demethylation of histone peptides by KIAA1718$_{PHD-JmjC}$. **A)** Overall demethylation time course of H3$_{1-21}$K9me2. **B)** Overall demethylation time course of H3$_{1-21}$K4me3K9me2. **C)** Individual plots of Kme0, Kme1 and Kme2 of H3$_{1-21}$K9me2. **D)** Individual plots of Kme0, Kme1 and Kme2 of H3$_{1-21}$K4me3K9me2. Figures were generated using Prism® software. A pre-quenched sample was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

To further analyse the theory of linker-aided KDM7 substrate selectivity, demethylation assays were carried out with H3K27me2 peptides. The aim of work this was to understand the demethylation activities of PHF8 and KIAA1718 on H3K27me2 (with and without H3K4me3), to compare them to H3K9me2 and to validate whether H3Kme3 has an inhibitory effect on the demethylation of H3K27me2. A 35 mer H3K4me3K27me2 peptide (synthesised by CEM) was used to be consistent with literature experiments.[2] Due to the difficulty in 35

mer H3K27me2 synthesis, a 21 mer peptide was used instead (residues 21-44), which had also been examined in the literature.[2]

A time course using the $H3_{21-44}K27me2$ as the KIAA1718 substrate manifested a relatively high activity, with a $t_{1/2}$ of 6 min (Figure 2.15 A). Interestingly, however, when H3K4me3 was presented in the dual methylated peptide $H3_{1-35}K4me3K27me2$, the activity dropped by 5-fold, with a $t_{1/2}$ of about 30 min (Figure 2.15 B). This experiment also showed that the rate of $H3_{1-35}K4me3K27me0$ formation (Figure 2.15 C purple) was slower than that of $H3_{21-44}K27me0$ (Figure 2.15 D purple). This observation is consistent with the results using H3K9me2, but not supportive to the suggestion that H3K4me3 facilitates H3K27me2 demethylation *in vitro*.[2]

Interestingly, the $t_{1/2}$ of $H3_{1-35}K27me2$ reported by Horton *et al.* is about 19 min, 2.4-fold longer than for the $H3_{21-44}K27me2$ peptide.[2] Two possible explanations for this are that: either the extra 9 amino acids from 36 to 44 are crucial to (or at least facilitate) the demethylation of H3K27me2, or that binding of the PHD finger to the unmodified H3K4 position (or amino acids in this region) has a negative effect on H3K27me2 demethylation. To test the first explanation, KIAA1718 demethylation assays of 44 mer and 35 mer H3K4me3K27me2/H3K27me2 peptides should be carried out side by side. The substrate preference of KIAA1718 could be investigated by comparing the $t_{1/2}$, time for 90% conversion and K_M values.

As for the second explanation, ITC binding assays of KIAA1718 with an $H3_{1-21}K4me0$ peptide carried out by Mr Oliver Coleman and Mr Roman Belle showed that the K_D of this peptide was within the single digital μM range (data not shown). This suggests that the binding of

KIAA1718 on H3K4me0 cannot be neglected. Thus, a second possible explanation is that the binding of the PHD finger of KIAA1718 has a negative effect on H3K27me2 demethylation. Given the observed binding of KIAA1718 to the fully unmethylated peptide, it may be that the distance between H3K4 and H3K27 is still not the best fit for KIAA1718. Due to the difficulty of 44 mer peptide synthesis, the first explanation was not further investigated in the work described this Thesis, but should be considered for future work. Histone substrates with dimethyl-lysine between H3K9 and H3K27 (i.e. H3K14/18/23/36) and modified H1.4 peptides were screened to identify if there were a more optimal Kme3-Kme2 distance for KIAA1718; the results are described in the next section (Chapter 2, Section 2.3.5).

The parameters of *in vitro* demethylation assays of KIAA1718 against all the peptidic substrates described in this section, i.e. H3K4me3K9me2, H3K9me2, H3K4me3K27me2 and H3K27me2, are summarised in Table 2.5.

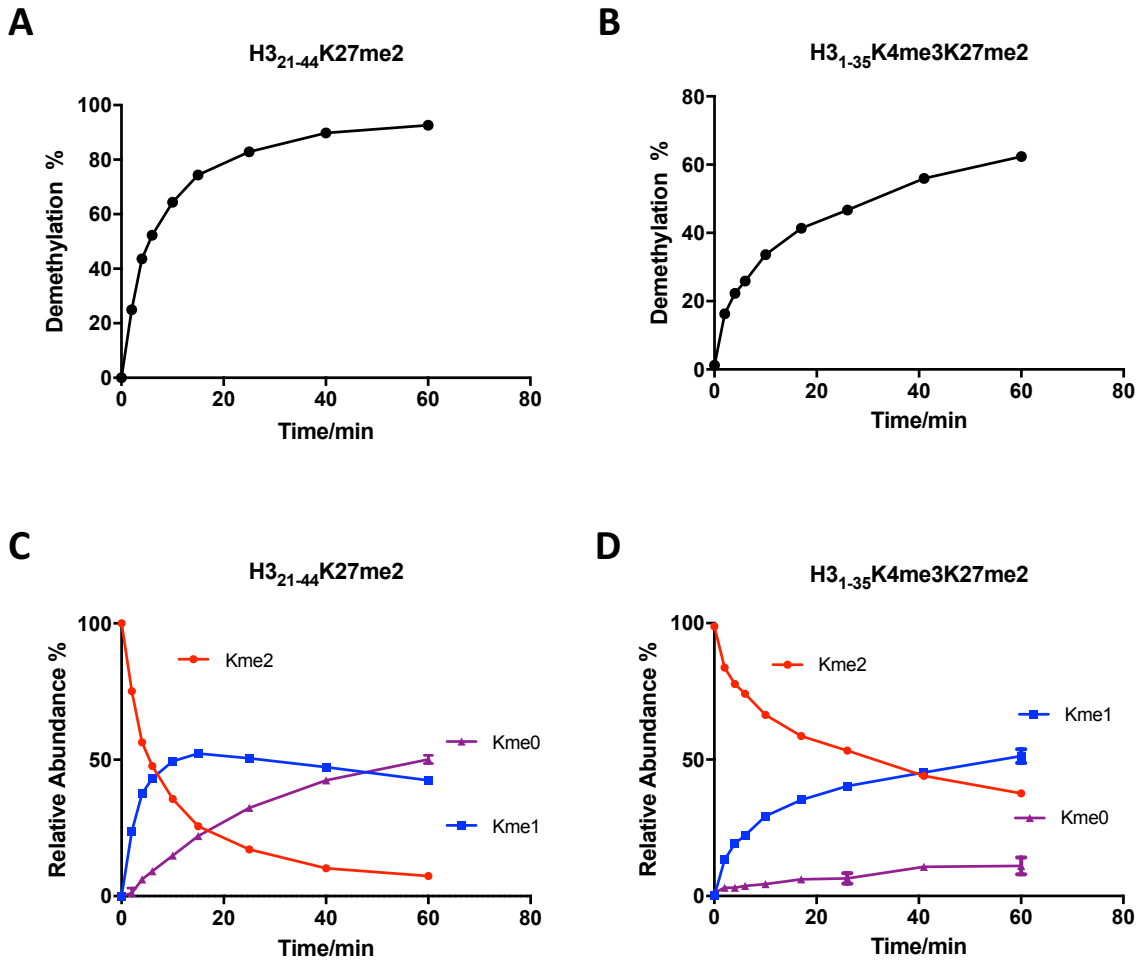


Figure 2.15 Demethylation of histone peptides by KIAA1718$_{PHD-JmjC}$. **A)** Overall demethylation time course of H3$_{21-44}$K27me2. **B)** Overall demethylation time course of H3$_{1-35}$K4me3K27me2. **C)** Individual plots of Kme0, Kme1 and Kme2 of H3$_{21-44}$K27me2. **D)** Individual plots of Kme0, Kme1 and Kme2 of H3$_{1-35}$K4me3K27me2. Figures were generated using Prism® software. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

Table 2.5 Summary of the demethylation reactions catalysed by KIAA1718. The values marked with asterisks are reported in the literature.[2] The '-' symbol indicates no apparent evidence for demethylation was observed.

	H3K4me3K9me2			H3K9me2		
	15 mer	**21 mer**	**24 mer**	**15 mer**	**21 mer**	**24 mer**
$t_{1/2}$	>60 min	-	-*	20 min	3 min	4 min*
90% conversion	>60 min	-	-*	>60 min	10 min	15 min*
	H3K4me3K27me2			**H3K27me2**		
	35 mer		**35 mer**	**21 mer**	**21 mer**	**35 mer**
$t_{1/2}$	30 min		10 min*	6 min	8 min*	19 min*
90% conversion	>60 min		60 min*	40 min	30 min*	>60 min*

2.3.5 Potential alternative substrates for KIAA1718 and PHF8

The possibility of potential alternative substrates for KIAA1718 and PHF8 was then investigated. First, a series of dimethylated histone 3 (with and without H3K4me3) peptides were incubated with $KIAA1718_{PHD\text{-}JmjC}$ and $PHF8_{PHD\text{-}JmjC}$ using the standard demethylation assay described in the previous sections. This study aimed to investigate whether KIAA1718 and PHF8 possessed any demethylation activity on other dimethyl-lysine positions (14, 18, 23 and 36), and to compare them with the activities on H3K9me2 and H3K27me2. End-point demethylation percentages after 1 hour were determined using MALDI-TOF MS. All the peptides that had demonstrated over 50% demethylation were further investigated by time course studies, if purified peptides (i.e. not crude peptides) were available in the research group. Second, a histone peptide library of H3 and H4 peptides (AltaBioscience, Figure 2.16) was used for wider alternative substrate screening to potentially identify new substrates. The peptide library screening results are shown below (Table 2.6). Finally, four 15 mer modified H1.4 peptides, H1.4K26me1, H1.4K26me2, H1.4K26me3 and H1.4K21me3K26me2, were subject to time course studies with PHF8 and KIAA1718.

Histone Set 3 H3,H4 N-terminal library

A1	Control 1	E1	Ac-KQLATΦAARKSAPATGGVKKP -spacer-Biotin
A2	ARTKQTARKSTGGKAPRKQLA -spacer-Biotin	E2	Ac-KQLATΠAARKSAPATGGVKKP -spacer-Biotin
A3	AΞTKQTARKSTGGKAPRKQLA -spacer-Biotin	E3	Ac-KQLATΘAARKSAPATGGVKKP -spacer-Biotin
A4	AΨTKQTARKSTGGKAPRKQLA -spacer-Biotin	E4	Ac-KQLATΘAAΨKSAPATGGVKKP -spacer-Biotin
A5	AΨΩKQTARKSTGGKAPRKQLA -spacer-Biotin	E5	Ac-KQLATΘAAΨΔSAPATGGVKKP -spacer-Biotin
A6	AΨΩΘQTARKSTGGKAPRKQLA -spacer-Biotin	E6	Ac-LATKAARKSAPATGGVKKPHR -spacer-Biotin
A7	AΨTΘQTARKSTGGKAPRKQLA -spacer-Biotin	E7	Ac-LATKAAΞKSAPATGGVKKPHR -spacer-Biotin
A8	ARΩKQTARKSTGGKAPRKQLA -spacer-Biotin	E8	Ac-LATKAAΨKSAPATGGVKKPHR -spacer-Biotin
A9	ARΩΘQTARKSTGGKAPRKQLA -spacer-Biotin	E9	Ac-LATKAAΨΔSAPATGGVKKPHR -spacer-Biotin
A10	ARTΦQTARKSTGGKAPRKQLA -spacer-Biotin	E10	Ac-LATKAAΨΔΣAPATGGVKKPHR -spacer-Biotin
A11	ARTΠQTARKSTGGKAPRKQLA -spacer-Biotin	E11	Ac-LATKAAΨKΣAPATGGVKKPHR -spacer-Biotin
A12	ARTΘQTARKSTGGKAPRKQLA -spacer-Biotin	E12	Ac-LATKAARΔSAPATGGVKKPHR -spacer-Biotin
B1	ARTΘQTARΔSTGGKAPRKQLA -spacer-Biotin	F1	Ac-LATKAARΔΣAPATGGVKKPHR -spacer-Biotin
B2	ARTΘQTARΘSTGGKAPRKQLA -spacer-Biotin	F2	Ac-LATKAARΦSAPATGGVKKPHR -spacer-Biotin
B3	ARTKQTARΔSTGGKAPRKQLA -spacer-Biotin	F3	Ac-LATKAARΠSAPATGGVKKPHR -spacer-Biotin
B4	ARTKQTARΔΣTGGKAPRKQLA -spacer-Biotin	F4	Ac-LATKAARΘSAPATGGVKKPHR -spacer-Biotin
B5	ARTKQTARΔSΩGGKAPRKQLA -spacer-Biotin	F5	Ac-LATKAAΨΘSAPATGGVKKPHR -spacer-Biotin
B6	ARTKQTARΔΣΩGGKAPRKQLA -spacer-Biotin	F6	Ac-LATKAARΘΣAPATGGVKKPHR -spacer-Biotin
B7	ARTKQTARΦSTGGKAPRKQLA -spacer-Biotin	F7	Ac-LATKAAΨΘΣAPATGGVKKPHR -spacer-Biotin
B8	ARTKQTARΠSTGGKAPRKQLA -spacer-Biotin	F8	Ac-LATKAARKΣAPATGGVKKPHR -spacer-Biotin
B9	ARTKQTARΘSTGGKAPRKQLA -spacer-Biotin	F9	SGRGKGGKGLGKGGAKRHRKV -spacer-Biotin
B10	ARTKQTARΘSΩGGKAPRKQLA -spacer-Biotin	F10	ΣGRGKGGKGLGKGGAKRHRKV -spacer-Biotin
B11	ARTKQTARΘΣTGGKAPRKQLA -spacer-Biotin	F11	SGΞGKGGKGLGKGGAKRHRKV -spacer-Biotin
B12	ARTKQTARΘΣΩGGKAPRKQLA -spacer-Biotin	F12	SGΨGKGGKGLGKGGAKRHRKV -spacer-Biotin
C1	ARTKQTARKΣTGGKAPRKQLA -spacer-Biotin	G1	SGΨGΔGGKGLGKGGAKRHRKV -spacer-Biotin
C2	ARTKQTARKSΩGGKAPRKQLA -spacer-Biotin	G2	ΣGΨGΔGGKGLGKGGAKRHRKV -spacer-Biotin
C3	ARTKQTARKΣΩGGKAPRKQLA -spacer-Biotin	G3	SGRGΔGGKGLGKGGAKRHRKV -spacer-Biotin
C4	Ac-KQTARKΣΩGGΔAPRKQLATKA -spacer-Biotin	G4	SGRGΔGGΔGLGΔGGAKRHRKV -spacer-Biotin
C5	Ac-KQTARKΣΩGGΘAPRKQLATKA -spacer-Biotin	G5	SGRGΔGGΔGLGKGGAKRHRKV -spacer-Biotin
C6	Ac-KQTARKSΩGGΔAPRKQLATKA -spacer-Biotin	G6	SGRGKGGΔGLGKGGAKRHRKV -spacer-Biotin
C7	Ac-KQTARKSΩGGΘAPRKQLATKA -spacer-Biotin	G7	SGRGKGGΔGLGΔGGAKRHRKV -spacer-Biotin
C8	Ac-KQTARΔSTGGΔAPRKQLATKA -spacer-Biotin	G8	SGRGKGGKGLGΦGGAKRHRKV -spacer-Biotin
C9	Ac-RKSTGGΔAPRKQLATKAARKS -spacer-Biotin	G9	SGRGKGGKGLGΠGGAKRHRKV -spacer-Biotin
C10	Ac-RKSTGGΔAPΨKQLATKAARKS -spacer-Biotin	G10	SGRGKGGKGLGΘGGAKRHRKV -spacer-Biotin
C11	Ac-RKSTGGΦAPRKQLATKAARKS -spacer-Biotin	G11	SGRGKGGΔGLGΦGGAKRHRKV -spacer-Biotin
C12	Ac-RKSTGGΠAPRKQLATKAARKS -spacer-Biotin	G12	SGRGKGGΔGLGΘGGAKRHRKV -spacer-Biotin
D1	Ac-RKSTGGΘAPRKQLATKAARKS -spacer-Biotin	H1	Ac-GGKGLGΦGGAΔRHRKVLRDNI -spacer-Biotin
D2	Ac-RKSTGGΘAPΨKQLATKAARKS -spacer-Biotin	H2	Ac-GGKGLGΘGGAΔRHRKVLRDNI -spacer-Biotin
D3	Ac-RKSTGGKAPΞKQLATKAARKS -spacer-Biotin	H3	Ac-GGKGLGKGGAΔRHRKVLRDNI -spacer-Biotin
D4	Ac-RKSTGGKAPΨKQLATKAARKS -spacer-Biotin	H4	Ac-GKGGAKRHRKVLRDNIQGITK -spacer-Biotin
D5	Ac-RKSTGGΔAPRΔQLATKAARKS -spacer-Biotin	H5	Ac-GKGGAKRHRΔVLRDNIQGITK -spacer-Biotin
D6	Ac-GGKAPΨΔQLATKAARKSAPAT -spacer-Biotin	H6	Ac-GKGGAΔRHRΔVLRDNIQGITK -spacer-Biotin
D7	Ac-GGKAPRΔQLATKAARKSAPAT -spacer-Biotin	H7	Ac-GKGGAΔRHRΦVLRDNIQGITK -spacer-Biotin
D8	Ac-GGKAPRΔQLATΘAARKSAPAT -spacer-Biotin	H8	Ac-GKGGAΔRHRΘVLRDNIQGITK -spacer-Biotin
D9	Ac-GGKAPRΔQLATΔAARKSAPAT -spacer-Biotin	H9	Ac-GKGGAKRHRΦVLRDNIQGITK -spacer-Biotin
D10	Ac-KQLATΔAARKSAPATGGVKKP -spacer-Biotin	H10	Ac-GKGGAKRHRΠVLRDNIQGITK -spacer-Biotin
D11	Ac-KQLATΔAAΨKSAPATGGVKKP -spacer-Biotin	H11	Ac-GKGGAKRHRΘVLRDNIQGITK -spacer-Biotin
D12	Ac-KQLATΔAAΨΔSAPATGGVKKP -spacer-Biotin	H12	Control 2

Δ = acetyl-Lysine	Φ = monomethyl-Lys	Π = dimethyl-Lys	Θ = trimethyl-Lys
Σ = phospho-Ser	Ω = phospho-Thr	Ξ = monomethyl-Arg	Ψ = asym dimethyl-Arg
Spacer = aminohexanoic acid, Ahx		Ac- = N-terminal acetylation	

Figure 2.16 Histone H3, H4 N-terminal library peptide sequences.

Table 2.6 Histone H3, H4 N-terminal library peptide screening results. Peptide numbers match Figure 2.16. Red cross indicates no demethylation product (-14 Da) was observed. Green tick indicates at least 20% demethylation was observed.

	KIAA1718	PHF8	PHF2		KIAA1718	PHF8	PHF2
A1	✗	✗	✗	E1	✗	✗	✗
A2	✗	✗	✗	E2	✓	✗	✗
A3	✗	✗	✗	E3	✗	✗	✗
A4	✗	✗	✗	E4	✗	✗	✗
A5	✗	✗	✗	E5	✗	✗	✗
A6	✗	✗	✗	E6	✗	✗	✗
A7	✗	✗	✗	E7	✗	✗	✗
A8	✗	✗	✗	E8	✗	✗	✗
A9	✗	✗	✗	E9	✗	✗	✗
A10	✗	✗	✗	E10	✗	✗	✗
A11	✗	✗	✗	E11	✗	✗	✗
A12	✗	✗	✗	E12	✗	✗	✗
B1	✗	✗	✗	F1	✗	✗	✗
B2	✗	✗	✗	F2	✗	✗	✗
B3	✗	✗	✗	F3	✓	✗	✗
B4	✗	✗	✗	F4	✗	✗	✗
B5	✗	✗	✗	F5	✗	✗	✗
B6	✗	✗	✗	F6	✗	✗	✗
B7	✗	✓	✗	F7	✗	✗	✗
B8	✓	✓	✗	F8	✗	✗	✗
B9	✗	✗	✗	F9	✗	✗	✗
B10	✗	✗	✗	F10	✗	✗	✗
B11	✗	✗	✗	F11	✗	✗	✗
B12	✗	✗	✗	F12	✗	✗	✗
C1	✗	✗	✗	G1	✗	✗	✗
C2	✗	✗	✗	G2	✗	✗	✗
C3	✗	✗	✗	G3	✗	✗	✗
C4	✗	✗	✗	G4	✗	✗	✗
C5	✗	✗	✗	G5	✗	✗	✗
C6	✗	✗	✗	G6	✗	✗	✗
C7	✗	✗	✗	G7	✗	✗	✗
C8	✗	✗	✗	G8	✗	✗	✗
C9	✗	✗	✗	G9	✗	✗	✗
C10	✗	✗	✗	G10	✗	✗	✗
C11	✗	✗	✗	G11	✗	✗	✗
C12	✓	✗	✗	G12	✗	✗	✗
D1	✗	✗	✗	H1	✗	✗	✗
D2	✗	✗	✗	H2	✗	✗	✗
D3	✗	✗	✗	H3	✗	✗	✗
D4	✗	✗	✗	H4	✗	✗	✗

D5	✗	✗	✗	H5	✗	✗	✗
D6	✗	✗	✗	H6	✗	✗	✗
D7	✗	✗	✗	H7	✗	✗	✗
D8	✗	✗	✗	H8	✗	✗	✗
D9	✗	✗	✗	H9	✗	✗	✗
D10	✗	✗	✗	H10	✗	✗	✗
D11	✗	✗	✗	H11	✗	✗	✗
D12	✗	✗	✗	H12	✗	✗	✗

Interestingly, in addition to demethylation of H3K9me2 and H3K27me2, consistent with the time course studies described in the previous section, KIAA1718 also catalysed demethylation of $H3_{1-24}$K14me2, $H3_{1-24}$K18me2 and $H3_{1-24}$K23me2 to nearly 100% after 1-hour incubation and demethylated H3K36me2 (21 mer) to about 70% (Figure 2.17). When in combination with H3K4me3, the demethylation of H3K14me2, H3K18me2 and H3K23me2 also reached to nearly 100% (Figure 2.17), none of which have previously been reported as KIAA1718 substrates.

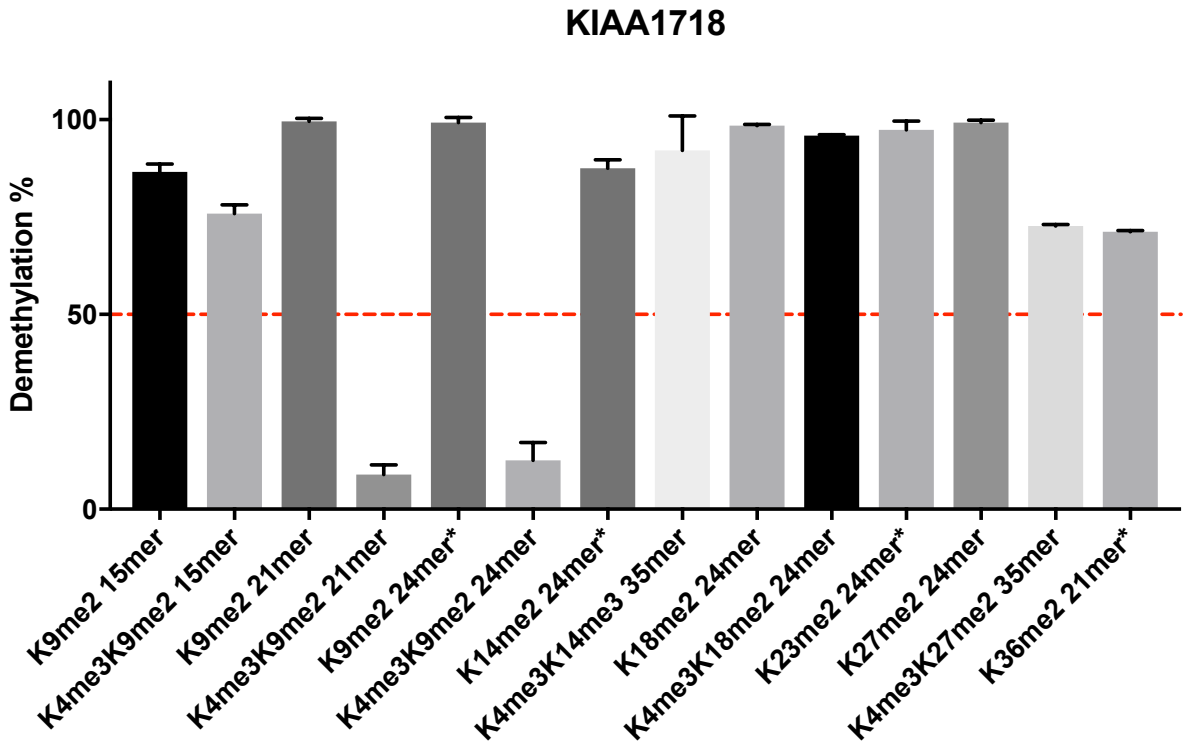


Figure 2.17 KIAA1718$_{PHD-JmjC}$ substrate single concentration demethylation assays. Peptides with asterisks (*) are crude peptides, i.e. not purified by HPLC chromatography. Reactions were carried out in 50 mM HEPES (pH 7.0), 50 mM NaCl and 5 mM TCEP. Each reaction contained 5 μM KIAA1718, 10 μM peptide, 50 μM Fe(II), 1 mM 2OG and 2 mM ascorbate. The extent of demethylation was analysed by MALDI-TOF MS after 1-hour incubation at 37 °C. Figures were generated using Prism® software. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

To further investigate the substrate demethylation by KIAA1718, all the purified peptides that were demethylated to over 50% by KIAA1718 in the screening assay (i.e. H3$_{1-35}$K4me3K14me2, H3$_{1-24}$K18me2 and H3$_{1-24}$K4me3K18m2) were then subjected to time course studies. The overall plots of time course of H3K18me2 and H3K4me3K18me2 were similar (Figure 2.18 A & C), which suggested that H3K4me3 was not important for KIAA1718 catalysed demethylation at H3K18me2. The increase in unmethylated peptide was faster with H3K18me2 than H3K4me3K18me2 (Figure 2.18 B & D), suggesting the presence of H3K4me3 may have a slightly adverse effect on H3K18me1 demethylation. In the case of H3$_{1-35}$K4me3K14me2, barely any demethylation of K4me3K14me1 was observed over the time course (Figure 2.18 F). Unfortunately, with no pure H3K14me2 peptide, the effect of H3K4me3 on H3K14me2 could not be determined. Future work should include the synthesis of an H3K14me2 peptide as well as H3K23me2 peptides to verify their demethylation by KIAA1718.

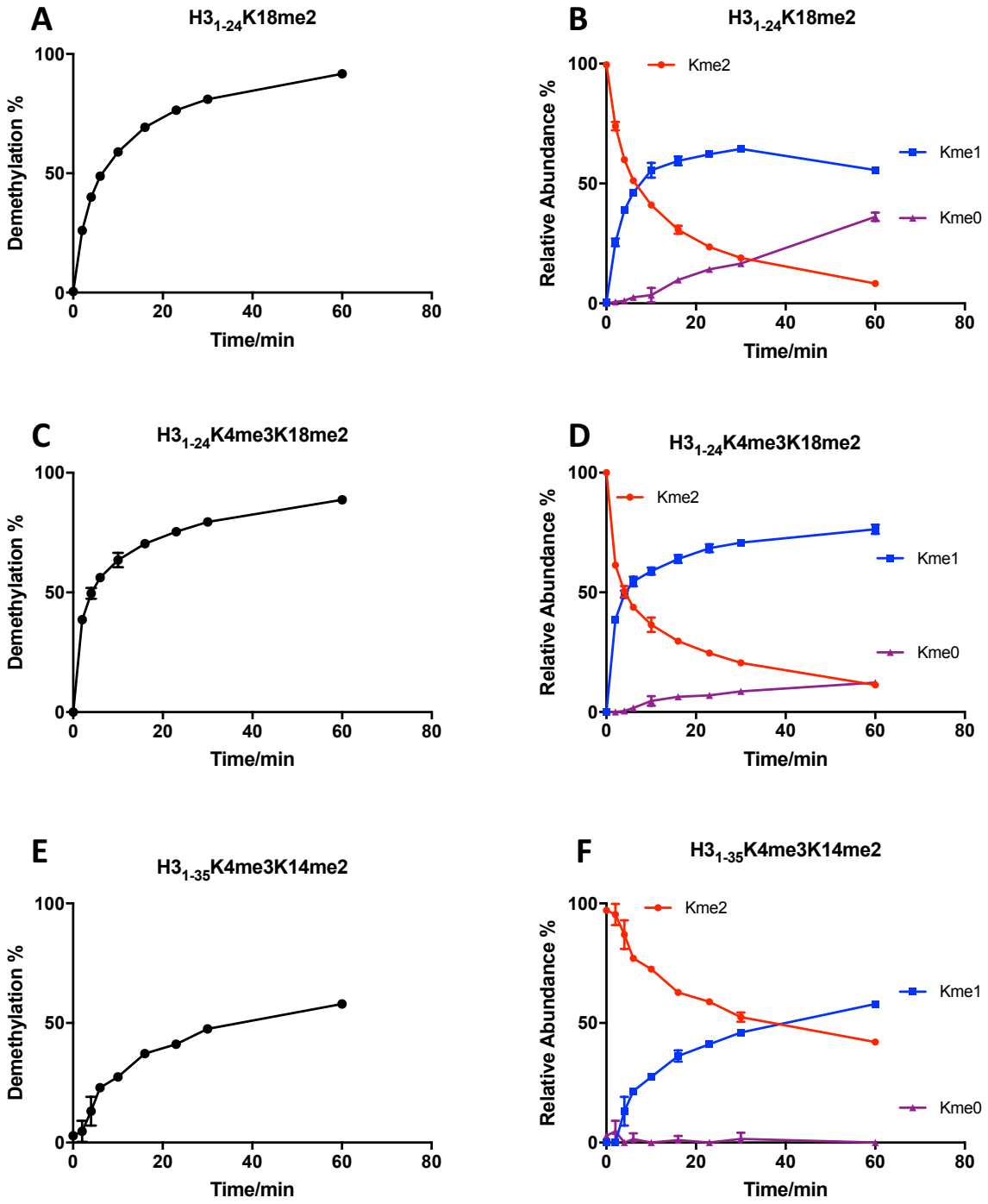


Figure 2.18 KIAA1718$_{PHD-JmjC}$ demethylation assays by MALDI-TOF MS. **A)** Overall demethylation time course of H3$_{1-24}$K18me2. **B)** Individual plots of Kme0, Kme1 and Kme2 of H3$_{1-24}$K18me2. **C)** Overall demethylation time course of H3$_{1-24}$K4me3K18me2. **D)** Individual plots of Kme0, Kme1 and Kme2 of H3$_{1-24}$K4me3K18me2. **E)** Overall demethylation time course of H3$_{1-35}$K4me3K14me2. **F)** Individual plots of Kme0, Kme1 and Kme2 of H3$_{1-35}$K4me3K14me2. Figures were generated using Prism® software. A pre-quenched sample was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

Following KIAA1718$_{PHD-JmjC}$, the same set of assays were conducted with PHF8$_{PHD-JmjC}$. The results of demethylation assays of histone 3 dimethylated substrates showed that PHF8$_{PHD-JmjC}$ primarily demethylated K9me2 with and without the presence of H3K4me3 (Figure 2.19 red box). Interestingly, demethylation of H3K14me2, H3K23me2 and H3K27me2 by PHF8$_{PHD-JmjC}$ was also observed, to a lower extent than that of H3K9me2 (Figure 2.19). Demethylation of H3K14me2 by PHF8$_{PHD-JmjC}$ seemed to be enhanced by the binding to H3K4me3, with the end-point demethylation increasing from 45% to 87% (Figure 2.19). Time course analysis of H3$_{1-35}$K4me3K14me2 confirmed that it was a substrate for PHF8, although no demethylation of H3K14me1 was observed (Figure 2.20). PHF8 also demethylated H3K27me2 with about 40% demethylation observed after an 1-hour incubation; however, PHF8 had no activity against H3$_{1-35}$K4me3K27me2 (Figure 2.19), potentially in line with the binding of H3K4me3 positioning the peptide for K9 demethylation. H3$_{1-24}$K23me2 (crude peptide) could also be demethylated by PHF8 but to a lesser extent than with KIAA1718. By contrast to previously reported results, PHF8$_{PHD-JmjC}$ did not show activity against H3K36me2 21 mer peptide under the condition tested in this Thesis.[18]

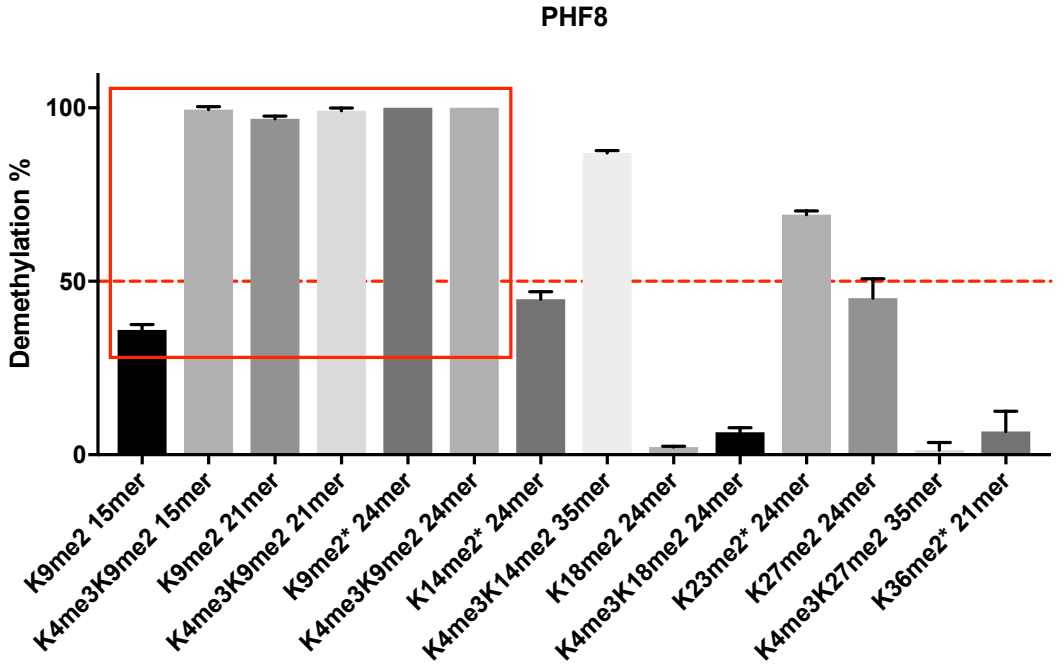


Figure 2.19 $PHF8_{PHD\text{-}JmjC}$ substrate single concentration demethylation assays. Peptides with asterisks (*) are crude peptides, i.e. not purified by HPLC chromatography. Reactions were carried out in 50 mM HEPES (pH 7.0), 50 mM NaCl and 5 mM TCEP. Each reaction contained 5 µM PHF8, 10 µM peptide, 50 µM Fe(II), 1 mM 2OG and 2 mM ascorbate. The extent of demethylation was analysed by MALDI-TOF MS after 1-hour incubation at 37 °C. Figures were generated using Prism® software. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

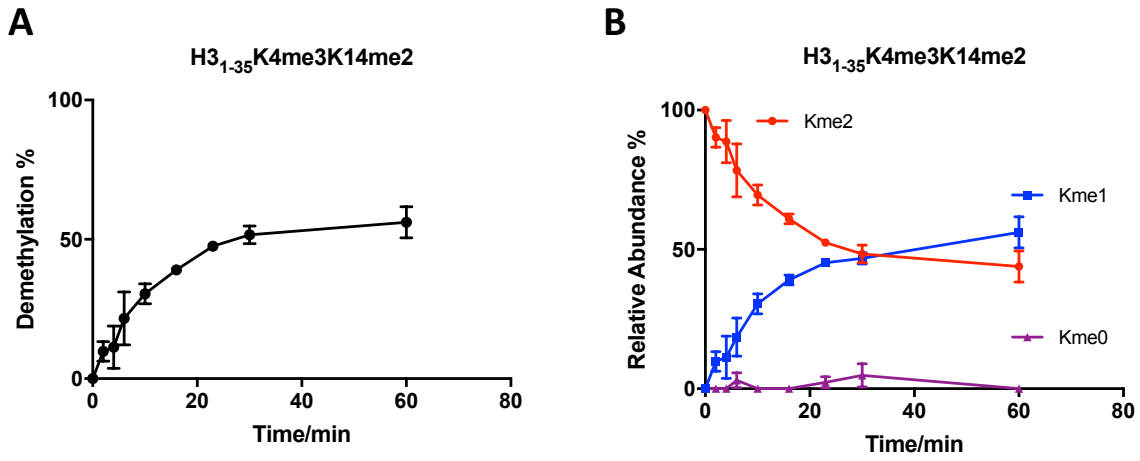


Figure 2.20 $PHF8_{PHD\text{-}JmjC}$ demethylation assays by MALDI-TOF MS. **A)** Overall demethylation time course of $H3_{1\text{-}35}$K4me3K14me2. **B)** Individual plots of Kme0, Kme1 and Kme2 of $H3_{1\text{-}35}$K4me3K14me2. Figures were generated using Prism® software. A pre-quenched sample was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

Finally, modified H1.4 peptides were tested with PHF8 and KIAA1718 to further investigate the preference for a specific Kme3-Kme2 distance. Histone isotype 4 (H1.4) is a 'linker' histone that binds to nucleosomal DNA but do not form part of the histone octamer.[33] Recent studies suggest that histone 1 isotypes and their post-translational modifications have functional roles in the transcriptional regulation of cells.[34,35] H1.4K26 is the most abundantly methylated site on H1.4 among multiple methylated lysine sites.[36] Given the similarity of the sequence of H1.4 to the first 15 amino acid residues of H3 (Figure 2.21), H1.4K21me3K26me2 was selected as a mimic of H3K4me3K9me2.

```
H3                              ARTKQTARKSTGGKA
H1.4     MSETAPAAPAAPAPAEKTPVKKKARKSAGAAKRKASGPPV
```

Figure 2.21 Sequence alignment of H1.4 residues 1-40 and H3 residues 1-15.

The H1.4 peptides (15 mer), including H1.4K21me2K26me2, H1.4K26me1, H1.4K26me2 and H1.4K26me3, were synthesised by Dr Richard Hopkinson. Time courses were recorded for each peptide with both KIAA1718$_{PHD\text{-}JmjC}$ and PHF8$_{PHD\text{-}JmjC}$ using the same MALDI-TOF MS assay method as for the H3 alternative substrates that was previously described in this section. The time courses of H1.4 peptides were taken with H3$_{1\text{-}15}$K4me3K9me2 and H3$_{1\text{-}15}$K9me2 side by side to compare the extent of demethylation. H3K9me2 was the still the best observed substrate for KIAA1718 of all the peptides tested (Figure 2.22 A). By contrast, KIAA1718 activity with H1.4K26me2, which was designed to mimic the H3K9me2, was much lower and similar to that with H3K4me3K9me2 (Figure 2.22 A). Interestingly, the activity of KIAA1718 with H1.4K21me3K26me3 was enhanced compared to its H3 counterpart, indicating that the Kme3-Kme2 distance is not the only factor affecting the demethylation activity of KIAA1718. For PHF8, none of the H1.4 peptides were demethylated to more than 10% after 1-hour incubation (Figure 2.22 B), suggesting that PHF8 is more specific in selecting its substrates. The work of H1.4 peptides has been reported in Walport *et al*. FEBS Letters, 2018.[37]

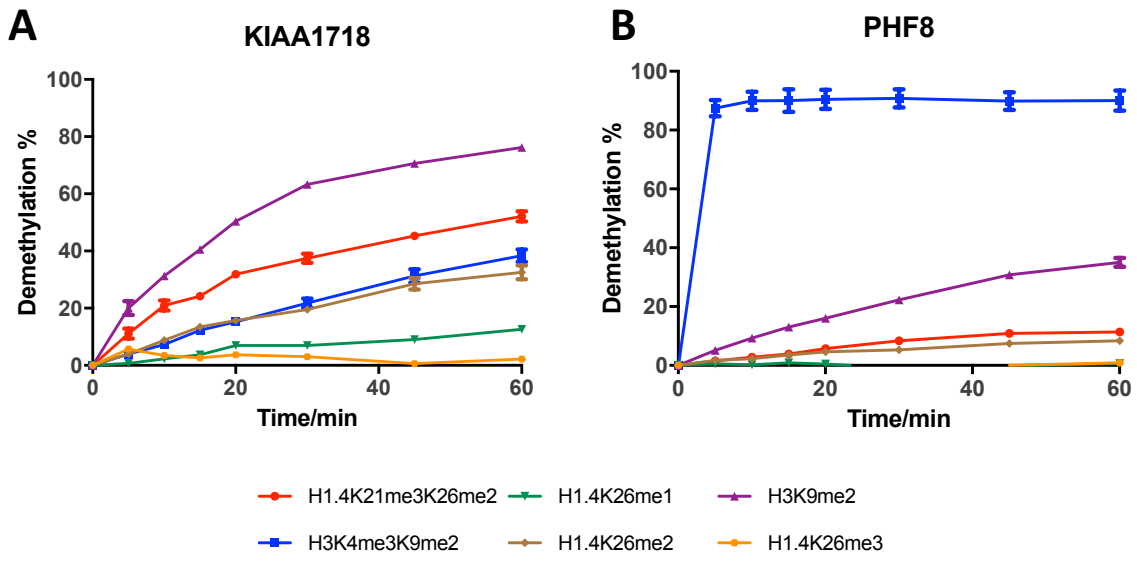


Figure 2.22 Overall demethylation time course of 15 mer H1.4 peptides (10 µM) comparing with $H3_{1-15}$K4me3K9me2 (10 µM) and $H3_{1-15}$K9me2 (10 µM) catalysed by **A)** $KIAA1718_{PHF-JmjC}$ (5 µM); **B)** $PHF8_{PHD-JmjC}$ (5 µM). Figures were generated using Prism® software. A pre-quenched sample was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

To conclude, the kinetic results with KIAA1718 and PHF8 described in this section demonstrate that, (a) combinatorial PTMs can alter the substrate preference of the KDM7s, and (b) the histone substrate scope may be wider than that reported. Although KIAA1718 and PHF8 share significant similarity in their JmjC domains and possess the same iron-binding motif, the demethylation activities of KIAA1718 and PHF8 are not the same (Table 2.7). Additionally, KIAA1718 and PHF8 demethylation assays against the H3, H4 peptide library showed that $KIAA1718_{PHD-JmjC}$/$PHF8_{PHD-JmjC}$ did not demethylate other peptides than the ones that had already been characterised (negative data not shown). These new results inspired further substrate screening of PHF2, which was found to be inactive on H3K9me2 *in vitro*.

Table 2.7 Summary of the demethylation reactions catalysed by KIAA1718 and PHF8 against alternative substrates.

	KIAA1718		PHF8	
	$t_{1/2}$	90% conversion	$t_{1/2}$	90% conversion
$H3_{1-35}$K4me3K14me2	30 min	>60 min	30 min	>60 min
$H3_{1-24}$K18me2	6 min	60 min	-	-
$H3_{1-24}$K4me3K18me2	5 min	60 min	-	-

2.3.6 Histone substrate screening of PHF2

In an attempt to profile the *in vitro* demethylation activity of PHF2, some histone H3 and H4 peptides and ARID5B peptide were tested by Xian Deng (former Master's student in the CJS group) (Table 2.8). ARID5B has been reported to interact with and recruit PHF2 , promoting H3K9me2 demethylation at target gene.[38] However, no demethylation was obsevred .

Table 2.8 Previous *in vitro* demethylation assays of PHF2.

Peptide substrate tested in the MALDI assay		
Peptide	**Sequence**	**Demethylation Observed?**
H3K4me1	ARTK(me1)QTARKSTGGKAPRKQLATKA	No
H3K4me2	ARTK(me2)QTARKSTGGKAPRKQLATKA	No
H3K4me3	ARTK(me3)QTARKSTGGKAPRKQLATKA	No
H3K9me1	ARTKQTARK(me1)STGGKAPRKQ	No
H3K9me2	ARTKQTARK(me2)STGGKAPRKQ	No
H3K9me3	ARTKQTARK(me3)STGGKAPRKQ	No
H3K27me1	LATKAARK(me1)SAPSTGGVKK	No
H3K27me2	LATKAARK(me2)SAPSTGGVKK	No
H3K27me3	LATKAARK(me3)SAPSTGGVKK	No
H3K36me1	KSAPSTGGVK(me1)KPHRYRPGTVALRE	No
H3K36me2	KSAPSTGGVK(me2)KPHRYRPGTVALRE	No
H3K36me3	KSAPSTGGVK(me3)KPHRYRPGTVALRE	No
H4K20me1	KGGAKRHRK(me1)VLRDNIQ	No
H4K20me2	KGGAKRHRK(me2)VLRDNIQ	No
H4K20me3	KGGAKRHRK(me3)VLRDNIQ	No
H3K4me3K9me2	ARTK(me3)QTARK(me2)STG	No
ARID5B	ALYKYMKERK(me2)TPIERIP	No

Demethylation assays were therefore undertaken using peptides methylated at H3K9/14/18/23/27/36 using the same conditions as had been optimised for PHF8 and KIAA1718. However, no demethylation of any these substrates by PHF2 was observed (Figure 2.23). Screening against the full H3, H4 peptide library also offered no hits for PHF2 catalysed demethylation. To investigate the mechanism of the demethylation catalysed by the KDM7 subfamily and to discover the potential cause of inactive PHF2 *in vitro*, mechanistic studies were carried out as described in the next Chapter (3).

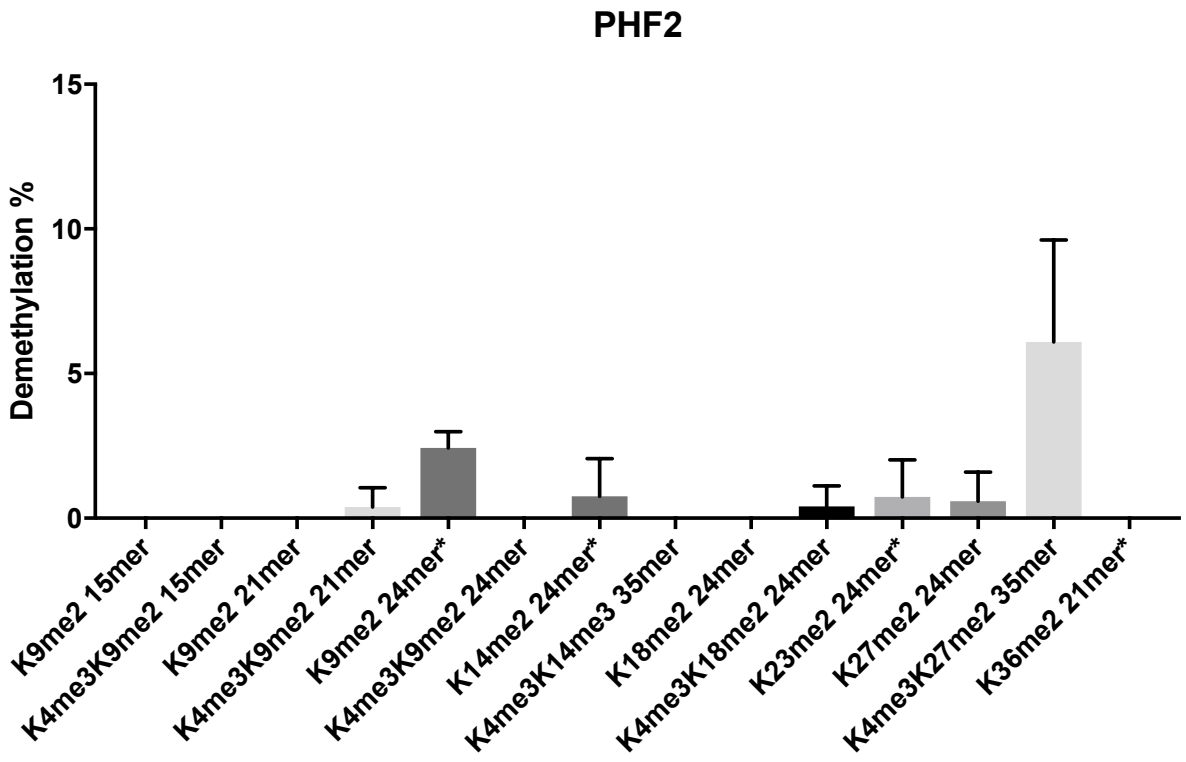


Figure 2.23 $PHF2_{PHD\text{-}JmjC}$ substrate single concentration demethylation assays. Peptides with asterisks (*) are crude peptides, i.e. not purified by HPLC chromatography. Reactions were carried out in 50 mM HEPES (pH 7.0), 50 mM NaCl and 5 mM TCEP. Each reaction contained 5 µM PHF2, 10 µM peptide, 50 µM Fe(II), 1 mM 2OG and 2 mM ascorbate. Demethylation percentages were analysed by MALDI-TOF MS after 1-hour incubation at 37 °C. Figures were generated using Prism® software. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

2.4 Summary

This Chapter has described *in vitro* biochemical studies on the three human KDM7 subfamily members. The demethylation activities of PHF8 and KIAA1718 were verified using *in vitro* MALDI-TOF MS demethylation assays. PHF8 preferentially demethylates H3K9me2, and this demethylation activity can be enhanced by the presence of H3K4me3. PHF8 was also found to accept H3K14me2, H3K23me2 and H3K27me2. The presence of H3K4me3 increased the demethylation of H3K14me2, whilst inhibiting that of H3K27me2. Unlike previous reports, PHF8 did not demethylate H3K36me2 under the conditions tested. The results support the proposal that KIAA1718 has a broader substrate spectrum than PHF8.[18]

KIAA1718 accepts H3K9me2, H3K14me2, H3K18me2, H3K23me2, H3K27me2 and H3K36me2 as its substrates. The demethylation activity of KIAA1718 on H3K9me2 was attenuated by the presence of H3K4me3, while demethylation on H3K18me2 by KIAA1718 was not significantly affected. Kinetics studies also suggest that KIAA1718 has a relatively low activity on the monomethyllysine at all tested positions, compared to PHF8. However, the effect of H3K4me3 on H3K27me2 demethylation remains a question mark. No *in vitro* demethylation activity has been discovered for PHF2.

The new histone substrates of KIAA1718 may be of biological importance. H3K14me2 is involved in the in the regulation of the transcription of bone marrow stem cells trans-differentiation.[39] H3K23me2 is enriched in the heterochromatic regions and is suggested to have a role in organising the genome in distinct heterochromatic regions during embryogenesis.[40] By contrast to the extensive studies of H3K9me2 and H3K27me2 marks, H3K18me2 is poorly understood. Future work should also consider overexpression of

KIAA1718 in cells and examine the change in the level of H3K14me2, H3K18me2 and H3K23me2 using immunoprecipitation. If the overexpression of KIAA1718 results in the decrease of one or some of the marks, it may indicate that these post-translational marks are substrates of KIAA1718 in cells. The results thus suggest that KIAA1718 may also associate with cell differentiation and be relevant studies of the poorly understood post-translational marks on histone3. When developing inhibitors for KIAA1718, scientists should consider the danger of increasing H3K14/18/23me2 levels in addition to H3K9me2.

The results in this Chapter suggest that the substrate binding by the PHD finger appears to exert a significant influence on the demethylation activity of the JmjC domains of both PHF8 and KIAA1718. The length and conformation of linker region are also proposed to affect the substrate selectivity of the KDM7s. KIAA1718 and PHF8 have a different preference of Kme3-Kme2 combinations. Future work should consider the synthesis of purified relatively long peptides including H3K14me2, H3K23me2, H3K4me3K23me2 to further understand the effect of H3K4me3 on the KDM7s demethylation of these positions.

2.5 References

1. Klose, R. J., Kallin, E. M. & Zhang, Y. JmjC-domain-containing proteins and histone demethylation. *Nat. Rev. Genet.* **7,** 715–727 (2006).
2. Horton, J. R. *et al.* Enzymatic and structural insights for substrate specificity of a family of jumonji histone lysine demethylases. *Nat. Struct. Mol. Biol.* **17,** 38–44 (2010).
3. Horton, J. R., Upadhyay, A. K., Hashimoto, H., Zhang, X. & Cheng, X. Structural basis for human PHF2 Jumonji domain interaction with metal ions. *J. Mol. Biol.* **406,** 1–8 (2011).
4. Abidi, F. E., Miano, M. G., Murray, J. C. & Schwartz, C. E. A novel mutation in the PHF8 gene is associated with X-linked mental retardation with cleft lip/cleft palate. *Clin. Genet.* **72,** 19–22 (2007).
5. Koivisto, A. M. *et al.* Screening of mutations in the PHF8 gene and identification of a novel mutation in a Finnish family with XLMR and cleft lip/cleft palate. *Clin. Genet.* **72,** 145–149 (2007).
6. Kleine-Kohlbrecher, D. *et al.* A Functional Link between the Histone Demethylase PHF8 and the Transcription Factor ZNF711 in X-Linked Mental Retardation. *Mol. Cell* **38,** 165–178 (2010).
7. Qiao, Y. *et al.* Autism-associated familial microdeletion of Xp11.22. *Clin. Genet.* **74,** 134–144 (2008).
8. Qi, H. H. *et al.* Histone H4K20/H3K9 demethylase PHF8 regulates zebrafish brain and craniofacial development. *Nature* **466,** 503–507 (2010).
9. Maina, P. K. *et al.* c-MYC drives histone demethylase PHF8 during neuroendocrine differentiation and in castration-resistant prostate cancer. *Oncotarget* **7,** 75585–75602 (2016).
10. Maina, P. K. *et al.* Histone demethylase PHF8 regulates hypoxia signaling through HIF1α and H3K4me3. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1860,** 1002–1012 (2017).
11. Huang, C. *et al.* The dual histone demethylase KDM7A promotes neural induction in early chick embryos. *Dev. Dyn.* **239,** 3350–3357 (2010).
12. Huang, C. *et al.* Dual-specificity histone demethylase KIAA1718 (KDM7A) regulates neural differentiation through FGF4. *Cell Res.* **20,** 154–165 (2010).
13. Choi, J. Y. & Jo, S. A. KDM7A histone demethylase mediates TNF-α-induced ICAM1 protein upregulation by modulating lysosomal activity. *Biochem. Biophys. Res. Commun.* **478,** 1355–1362 (2016).
14. Lee, K. H. *et al.* PHF2 histone demethylase acts as a tumor suppressor in association with p53 in cancer. *Oncogene* **34,** 2897–2909 (2015).
15. Okuno, Y. *et al.* Epigenetic regulation of adipogenesis by PHF2 histone demethylase. *Diabetes* **62,** 1426–1434 (2013).
16. Yang, J. *et al.* Epigenetic regulation of megakaryocytic and erythroid differentiation by PHF2 histone demethylase. *J. Cell. Physiol.* **233,** 6841–6852 (2018).
17. Shi, G. *et al.* PHD finger protein 2 (PHF2) represses ribosomal RNA gene transcription by antagonizing PHF finger protein 8 (PHF8) and recruiting methyltransferase SUV39H1. *J. Biol. Chem.* **289,** 29691–29700 (2014).
18. Loenarz, C. *et al.* PHF8, a gene associated with cleft lip/palate and mental retardation, encodes for an Nε-dimethyl lysine demethylase. *Hum. Mol. Genet.* **19,** 217–222

(2009).
19. Qiu, J. *et al.* The X-linked mental retardation gene PHF8 is a histone demethylase involved in neuronal differentiation. *Cell Res.* **20,** 908–918 (2010).
20. Yokoyama, A. *et al.* KIAA1718 is a histone demethylase that erases repressive histone methyl marks. *Genes to Cells* **15,** 867–873 (2010).
21. Lin, H. *et al.* Coordinated regulation of active and repressive histone methylations by a dual-specificity histone demethylase ceKDM7A from Caenorhabditis elegans. *Cell Res.* **20,** 899–907 (2010).
22. Liu, W. *et al.* PHF8 mediates histone H4 lysine 20 demethylation events involved in cell cycle progression. *Nature* **466,** 508–512 (2010).
23. Wen, H. *et al.* Recognition of histone H3K4 trimethylation by the plant homeodomain of PHF2 modulates histone demethylation. *J. Biol. Chem.* **285,** 9322–9326 (2010).
24. Baba, A. *et al.* PKA-dependent regulation of the histone lysine demethylase complex PHF2-ARID5B. *Nat. Cell Biol.* **13,** 668–675 (2011).
25. Yue, W. W. *et al.* Crystal structure of the PHF8 Jumonji domain, an Nε-methyl lysine demethylase. *FEBS Lett.* **584,** 825–830 (2010).
26. Fortschegger, K. *et al.* PHF8 targets histone methylation and RNA polymerase II to activate transcription. *Mol. Cell. Biol.* **30,** 3286–3298 (2010).
27. Arie Geerlof. Ligase independent cloning (LIC) Protocol. *EMBL - Hambg.* (2006).
28. Rose, N. R. *et al.* Inhibitor scaffolds for 2-oxoglutarate-dependent histone lysine demethylases. *J. Med. Chem.* **51,** 7053–7056 (2008).
29. Raran-Kurussi, S., Cherry, S., Zhang, D. & Waugh, D. S. Removal of Affinity Tags with TEV Protease. in *Heterologous Gene Expression in E.coli: Methods and Protocols* (ed. Burgess-Brown, N. A.) 221–230 (Springer New York, 2017). doi:10.1007/978-1-4939-6887-9_14
30. GE Healthcare. Superdex 200 10 / 300 GL Buffers and solvent resistance Optimization In-depth information. *Instr. 71-5017-96 AK* 1–2 (2014).
31. Hedhammar, M. & Hober, S. Zbasic-A novel purification tag for efficient protein recovery. *J. Chromatogr. A* **1161,** 22–28 (2007).
32. Wigle, T. J. *et al.* A high-throughput mass spectrometry assay coupled with redox activity testing reduces artifacts and false positives in lysine demethylase screening. *J. Biomol. Screen.* **20,** 810–820 (2015).
33. Izzo, A. & Schneider, R. The role of linker histone H1 modifications in the regulation of gene expression and chromatin dynamics. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1859,** 486–495 (2016).
34. Fyodorov, D. V., Zhou, B. R., Skoultchi, A. I. & Bai, Y. Emerging roles of linker histones in regulating chromatin structure and function. *Nat. Rev. Mol. Cell Biol.* **19,** 192–206 (2018).
35. Hergeth, S. P. & Schneider, R. The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. *EMBO Rep.* **16,** 1439–1453 (2015).
36. Lu, A. *et al.* Mapping of Lysine Monomethylation of Linker Histones in Human Breast Cancer Mapping of Lysine Monomethylation of Linker Histones in Human Breast and Its Cancer Department of Proteomics and Signal Transduction , Max-Planck Institute for 2nd Department of S. *J. Proteome Res.* 4207–4215 (2009).
37. Walport, L. J. *et al.* Mechanistic and structural studies of KDM-catalysed demethylation of histone 1 isotype 4 at lysine 26. *FEBS Lett.* (2018). doi:10.1002/1873-3468.13231

38. Hata, K. *et al.* Arid5b facilitates chondrogenesis by recruiting the histone demethylase Phf2 to Sox9-regulated genes. *Nat. Commun.* **4,** 1–11 (2013).
39. Liao, X., Liao, Y., Zou, Y., Li, G. & Liao, C. Epigenetic modifications of histone H3 during the transdifferentiation of Thy-1(+) Lin(-) bone marrow cells into hepatocytes. *Mol. Med. Rep.* **12,** 7561–7567 (2015).
40. Vandamme, J. *et al.* H3K23me2 is a new heterochromatic mark in Caenorhabditis elegans. *Nucleic Acids Res.* **43,** 9694–9710 (2015).

Chapter 3 Substrate and Metal Ion Binding Studies on the KDM7 Subfamily

3.1 Chapter overview

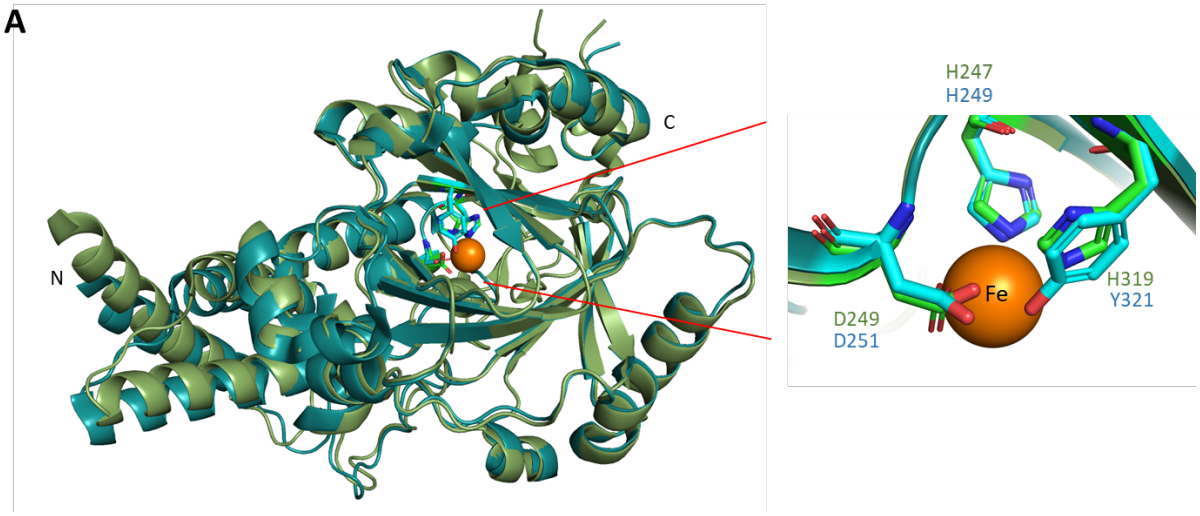
The work described in the previous Chapter (Chapter 2) described studies relating to the kinetics and the substrate selectivity of KDM7 enzymes in relation to N^{ε}-methylated lysine marks at various positions on the human histone 3 (H3) tail. The work described in this Chapter aimed to investigate the substrate binding affinities of different KDM7 enzymes and domains, and the relationships of these with catalytic activities. Both the PHD finger and JmjC domains of KDM7s were characterised using fluorescent ligands mimicking histone peptide or 2OG binding. Mutagenesis studies were carried out on PHF2 and PHF8 to study the importance of the Y321 (PHF2) and H319 (PHF8) residues in coordinating the active-site iron and how these may influence enzymatic activity. Finally, studies on the metal ion binding of KDM7 enzymes were carried out using *apo* forms of the PHF8 and PHF2 enzymes.

The modified histone peptides used in this Chapter were synthesised by Mr Roman Belle.

3.2 Introduction

3.2.1 Previous studies on KDM7 subfamily

All three members of the KDM7 subfamily posess one PHD finger and one JmjC domain (Chapter 2, Figure 2.1). The sequences of the PHD fingers are highly conserved (~ 70% identity) across the KDM7 subfamily.[1,2] The PHD fingers of the KDM7s are reported to bind H3K4me3 most tightly, with binding gradually weakening from the tri-, to the di- to the mono-methyl derivatives of H3K4.[1] This selective and potent binding of the KDM7 PHD fingers to H3K4me3 contributes to their substrate selectivity of the KDM activity.[2] On the other hand, the JmjC domains and the PHD-JmjC linker regions are more varied across the KDM7 subfamily, with ~ 50% identity in the JmjC domains across the subfamily.[1,2] The JmjC domains of KIAA1718 and PHF8, the closer of the three members, share ~65% identity (~75% similarity) and they contain the same core iron binding motif.[2] Interestingly, as described in Chapter 2 (Section 2.3.3 & 2.3.4), the JmjC domains of KDM7 enzymes have different substrate preferences. PHF8 shows the highest demethylation activity against H3K9me2 (over H3Kme2 at other positions) and shows no demethylation activity against H3K18me2. KIAA1718 is more promiscuous, and can demethylate H3K9me2, H3K14me2, H3K18me2, H3K23me2, H3K27me2 and H3K36me2 (Chapter 2, Section 2.3.5). The presence of H3K4me3 has an detrimental effect on the demethylation of H3K9me2 catalysed by KIAA1718 while enhances that by PHF8. PHF2 is thought to be catalytically inactive *in vitro*[1] and the iron binding motif different from that of PHF8 and KIAA1718 (Figure 3.1).[1] However, the molecular determinants of catalysis and substrate selectivity / binding of the KDM7 subfamily are not fully understood, so it is possible that PHF2 has an as yet unidentified catalytic activity.



B

```
•  PHF2          CVKDSYTDFHIDSGGASAWYHVLKGEKTFYLIRPASANISLYERWRSASNHSEMFFADQV 299
•  PHF8          SVRDSYTDFHIDFGGTSVWYHVLKGEKIFYLIRPTNANLTLFECWSSSSNQNEMFFGDQV 297
•  KIAA1718      GVQDSYTDFHIDFGGTSVWYHVLWGEKIFYLIKPTDENLARYESWSSSVTQSEVFFGDKV 281
•                 * ********* ** * ***** *** **** *   *    * * *     * ** * *

•  PHF2          DKCYKCIVKQGQTLFIPSGWIYATLTPVDCLAFAGHFLHSLSVEMQMRAYEVERRLKLGS 359
•  PHF8          DKCYKCSVKQGQTLFIPTGWIHAVLTPVDCLAFGGNFLHSLNIEMQLKAYEIEKRLSTAD 357
•  KIAA1718      DKCYKCVVKQGHTLFVPTGWIHAVLTSQDCMAFGGNFLHNLNIGMQLRCYEMEKRLKTPD 341
•                 ****** **** *** * *** * **  ** ** * *** *   **   ** * **
```

Figure 3.1 Metal binding by the KDM7 histone demethylases. **A)** View from a crystal structure of PHF2 (Blue, PDB ID: 3PU8) in alignment with PHF8 (Green, PDB ID: 3KV4). Residues constituting the iron binding motif are shown as sticks; iron is shown as an orange sphere. The iron-binding residues are H249, D251, Y321 for PHF2 (in green) and H247, D249, H319 for PHF8 (in blue). **B)** partial sequence alignment of part of the JmjC region of the KDM7 family. In PHF8 and KIAA1718, the iron binding motif is HxD…H (highlighted in orange). The second His residue is altered to a Tyr residue in PHF2.

3.2.2 Techniques of interaction studies

Techniques including isothermal calorimetry (ITC), intrinsic fluorescence (FI), NMR spectroscopy, and antibody based methods (e.g. AlphaScreen technology) can be used to quantitatively measure protein-substrate interactions. Each of these has advantages and disadvantages. ITC has been widely used for quantitative studies of protein-protein and protein-ligand interactions, and its use is often considered a gold standard in biomolecular interaction studies.[3–5] ITC can provide an accurate measurement of binding affinities based on the changes in the heat absorbed or released during the interaction.[5] This label-free and antibody free method can be automated, thus reducing manual effort. However, ITC requires a significant amount of protein. For a typical protein-substrate binding reaction with K_D of 10 μM (e.g. using PEAQ AutoITC (Malvern) available in the group), a single measurement consumes 300 μL of protein (50 μM), while, an FP assay in a 384-well plate only uses 40 μL of protein (50 μM), i.e. less than 1/7 of the protein that ITC requires.[6] Other disadvantages of ITC are that a single assay can take one or two hours (including pre- and after- washes); this is a problem for proteins prone to precipitation or degradation such as the KDM7s. Further, current ITC protocols cannot measure triplicate repeats at the same time, increasing the technical error due to the asynchronous operations.

While protein NMR spectroscopy can accurately determine the amino acid residues that are involved in binding, it requires a significant amount of protein, and it is labour-intensive and some types of protein NMR is difficult for high mass proteins such as the KDM7s.

Intrinsic fluorescence (FI) can be used to detect protein-ligand binding and relies on changes relating to fluorescent amino acid residues in the protein of interest.[7] Upon excitation with light (subject to the type of amino acid residue and its context), light is emitted in different

directions. If a fluorescent residue (e.g. tryptophan) is within or adjacent to a binding site, the signal may be modulated upon binding to that binding site. FI does not require fluorescent labelled ligand/protein, but the signal window for Fl sometimes can be too low for analysis due to the limited number of appropriate fluorescent amino acid residues.

The 'AlphaScreen' method is a well-established high-throughput inhibitor screening antibody based method that uses the amplified luminescent proximity homogeneous assay (ALPHA) technology, which requires excitation and emission beads for signal detection.[8] A considerable advantage of AlphaScreen against other methods is that it uses relatively small amounts of protein (down to picogram levels), but the complicated operation and light-sensitive nature of the detection beads can limit its usage in binding affinity measurements. The AlphaScreen method was used for inhibitor screening for KDM7 enzymes, as described in the next Chapter (Chapter 4). Based on the advantages and disadvantages of each of the aforementioned technique, the fluorescence polarisation (FP) method was selected as an appropriate technique for the use in the work described in this Chapter, due to its robustness and throughput (see below).

3.2.3 The principles of fluorescence polarisation (FP) assays

Fluorescence polarisation (FP) is a solution-based technique that is used to directly measure molecular interactions. Since the 1990s, FP assays have been used to quantitate ligand-protein interactions in drug discovery; its application has been extended to studying receptor-ligand interactions, protein-DNA interactions and proteolysis.[9–16] FP measures the change of the degree of polarisation of fluorophores.[17] When a fluorophore that is tumbling in solution is excited with plane-polarised light, it emits light in various directions (Figure 3.2).

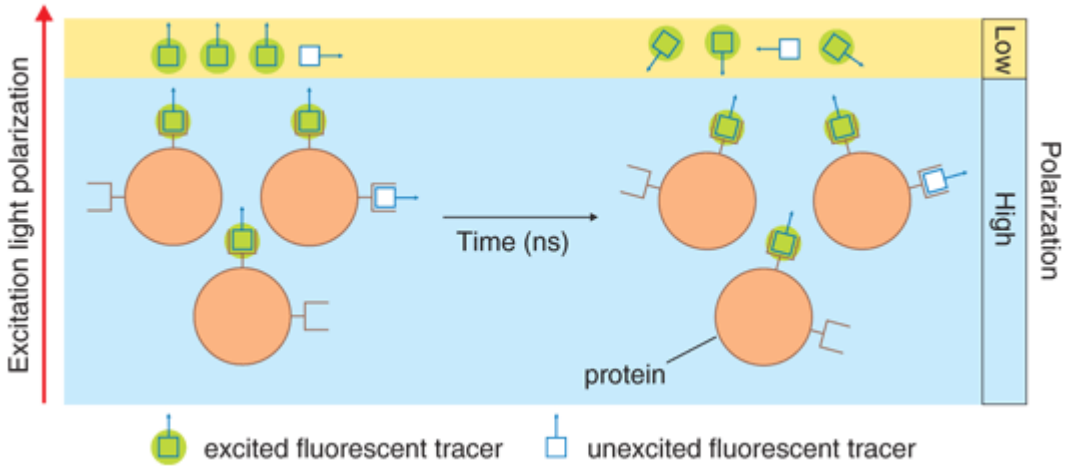


Figure 3.2 The principle of FP. The extent of polarisation increases as the fluorophores binding to macromolecules.[18]

The faster the fluorophore rotates, the more even the distribution of emitted light, and thus the weaker the light intensity is in a particular direction. If the fluorophore binds to a macromolecule (e.g. protein), the rotational speed of the fluorophore is reduced, resulting in stronger emission in one direction and higher polarisation. The ratio of free to bound fluorophore decreases as binding affinity rises, resulting in an increase of the FP overall.[9] Quantitatively, the value of FP is defined as the differences in the intensity of emission light

parallel ($I \parallel$) and perpendicular ($I \perp$) to the emission light plane, divided by the total emission light intensity (Equation 1).[19] Measurement of the differences in FP before and after the binding event can provide insight into the binding affinity.

$$FP = \frac{I\parallel - I\perp}{I\parallel + I\perp} \qquad \text{(Equation 1)}$$

The fluorophores used in this thesis were FI-methylstat (Figure 3.3 A) and fluorescein isothiocyanate (FITC) labelled histone peptides ($H3_{1\text{-}15}$K4me3K9me2-FITC and $H3_{1\text{-}15}$K9me2-FITC). Methylstat (Figure 3.3 B) is reported to be a selective inhibitor of Jumonji C domain-containing histone demethylases (JmjC-KDMs).[20] Its fluorescent analogue FI-methylstat has been used in FP assays to evaluate JmjC-KDMs inhibitors targeting the JmjC domain.[21] In this Chapter, FI-methylstat was used, not only to investigate the binding activity of the JmjC domains of the KDM7 subfamily, but also to study non-catalytic JmjC domain-containing proteins such as PHF2.

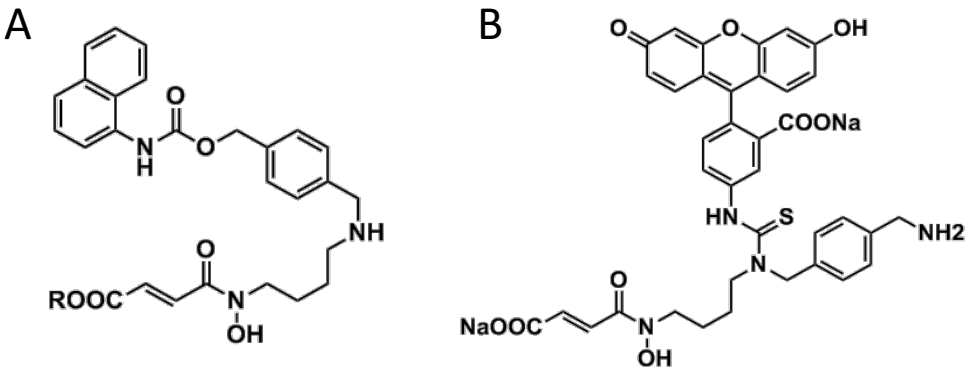


Figure 3.3 Structures of methylstat and its fluorescent analogue. **A)** Methylstat. **B)** FI-methylstat.[21]

3.3 Results and discussion

3.3.1 Purification of PHF8H319Y and PHF2Y321H

To investigate the importance of the iron coordinating residues in the metal binding affinity and catalysis of KDM7 enzymes, mutational analysis was carried out at the active sites of both PHF2 and PHF8. A PHF2 variant containing the canonical HxD…H (Y321H) Fe(II) binding triad, and variant PHF8 containing a PHF2-like HxD…Y (H319Y) motif were generated by site-directed mutagenesis (abbreviated as PHF2Y321H and PHF8H319Y, respectively) using standard methods as described in Chapter 5 (Materials and Methods). It was proposed that if the H to Y alternation present in PHF2 did not affect its iron binding affinity, the PHF8H319Y variant might preserve enzymatic activity. Alternatively, if this substitution was the (sole) cause of PHF2 inactivity *in vitro*, the PHF2Y321H variant may rescue the KDM activity as the binding motif would be the same as in the active PHF8 and KIAA1718 enzymes.

Expression trials for a pNH-TrxT plasmid bearing gene encoding for PHF8H319Y$_{PHD\text{-}JmjC}$ point mutant were carried out in BL21 (DE3) cells under six expression conditions varying the temperature and concentrations of the inducing reagent IPTG (isopropyl β-D-1-thiogalactopyranoside). The best expression condition found was the same as that for wild-type PHF8$_{PHD\text{-}JmjC}$, as judged by SDS-PAGE analysis (Figure 3.4 A). A pilot purification trial of PHF8H319Y$_{PHD\text{-}JmjC}$ suggested that the purification protocol for PHF8$_{PHD\text{-}JmjC}$ was suitable, yielding >90% purity by SDS-PAGE analysis (Figure 3.4 B). Therefore, the wild-type expression and purification procedures were used for the variant PHF8 proteins (Chapter 2, Section 2.3.1). The expression protocols for variant PHF2 proteins were adopted from that used by Xian Deng and the purification procedures was the same as that described for the

wild-type proteins (Chapter 2, Section 2.3.1). All of the variant proteins of PHF8 and PHF2 were at least 80% pure as judged by SDS-PAGE analysis (Chapter 2, Figure 2.3).

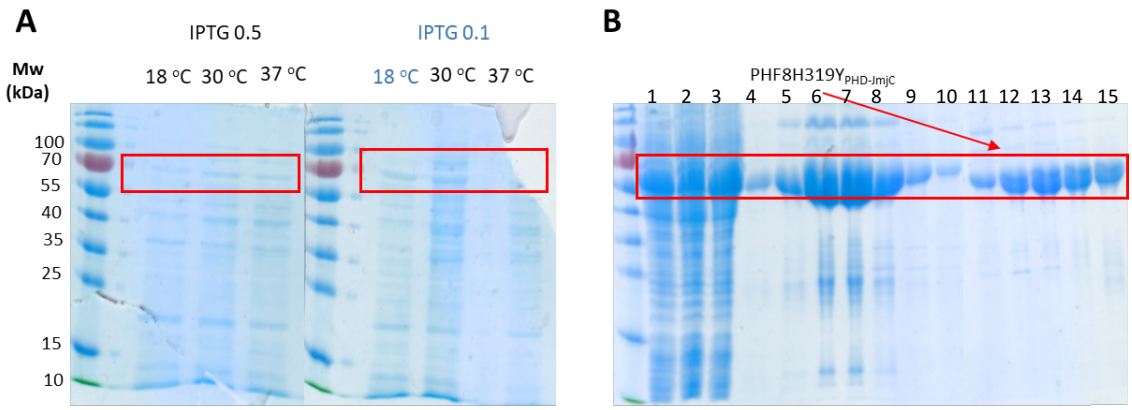


Figure 3.4 SDS-PAGE gel images of **A)** Expression trials PHF8H319Y$_{PHD\text{-}JmjC}$ with 0.5 mM and 0.1 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG) inducing at 18 °C, 30 °C, 37 °C. Condition selected is coloured in blue. **B)** Purification trial of PHF8H319Y$_{PHD\text{-}JmjC}$ using the same purification protocol of PHF8$_{PHD\text{-}JmjC}$. Lane 1: cell lysate; lane 2: HisTrap® column load; lane 3: pass-through; lane 4: wash; lane 5-10: HisTrap® column fractions; lane 11-15: gel filtration column fractions. The selected expressing condition is labeled in blue. Bands on SDS-PAGE gel indicating PHF8H319Y$_{PHD\text{-}JmjC}$ are highlighted in red boxes.

3.3.2 Activity of variant proteins

To compare the activity of the PHF8 and PHF2 variant proteins, MALDI-TOF MS-based activity assays were carried out for each isolated protein using the H3$_{1\text{-}15}$K4me3K9me2 peptide substrate (Chapter 5, Materials and Methods). The reaction mixtures were quenched (1:1) with 2% (v/v) aqueous formic acid solution after one hour. Interestingly, PHF8H319Y$_{PHD\text{-}JmjC}$ lost its demethylation activity on H3$_{1\text{-}15}$K4me3K9me2 (Figure 3.5 A). It was proposed that this loss of activity due to the PHF8 H to Y substitution might be caused by the reduction in iron binding capacity, since the substitution was within the central iron binding motif. Assays were thus further carried out using higher concentrations of Fe(II) (1 mM and 10 mM for every 5 µM of PHF8H319Y$_{PHD\text{-}JmjC}$), corresponding to 200-fold and 2000-

fold excess with respect to the enzyme concentration, in order to compensate for potential reduction in Fe(II) binding. However, as before no demethylation activity was detected under these conditions (Figure 3.5 B). The inactivation of PHF8 by the H to Y substitution is consistent with the proposal that HxD…H 'triad' coordination is indispensable for PHF8 KDM catalysis.

Interestingly, PHF2Y321H$_{PHD\text{-}JmjC}$, which restored the canonical HxD…H Fe(II) chelating motif of PHF8 to PHF2, also did not show any demethylation activity against H3$_{1\text{-}15}$K4me3K9me2 (Figure 3.5 C) under the tested conditions. Activity assays were also carried out using H3$_{1\text{-}15}$K9me2 peptide, but again no demethylation activity was observed (data not shown). These results with PHF8H319Y$_{PHD\text{-}JmjC}$ and PHF2Y321H$_{PHD\text{-}JmjC}$ reveal that whilst the HxD…H motif is required, it is not sufficient for JmjC KDM activity.

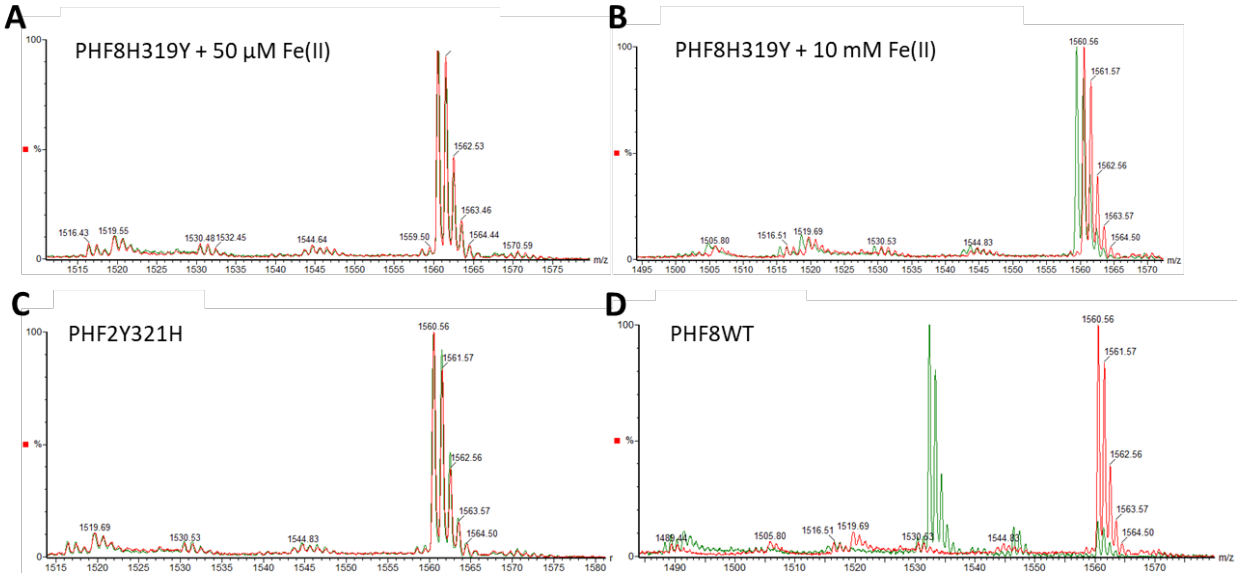


Figure 3.5 Representative MALDI-TOF MS spectra of **A)** PHF8H319Y$_{PHD\text{-}JmjC}$, **B)** PHF8H319Y$_{PHD\text{-}JmjC}$ (higher Fe(II) concentration), (the 1 Da shift between the red and green spectra is likely due to the calculation error of the MALDI-TOF analyzing software), **C)** PHF2Y321H$_{PHD\text{-}JmjC}$ and **D)** PHF8$_{PHD\text{-}JmjC}$ with H3$_{1\text{-}15}$ K4me3K9me2 peptide. Each reaction contained 5 µM enzyme, 10 µM peptide, 200 µM 2-oxoglutarate (2OG), 1 mM oL-ascorbate, 50 µM (for **A**) or 10 mM (for **B**) $(NH_4)_2Fe(SO_4)_2$. PHF8$_{PHD\text{-}JmjC}$ was used as a positive control and no enzyme reaction (in red) as a negative control. Reactions were quenched with 2% (v/v) aqeuos formic acid after 1 hour at 37 °C.

3.3.3 Buffer optimisations for KDM7s and their variants for substrate binding

The pH dependence and the buffer salt effects on KDM7 binding affinity has not been previously described. The results from activity assays indicated that HEPES pH 7.0 buffer with 50 mM NaCl is preferred for the demethylation of histone peptides by PHF8 and KIAA1718 (Chapter 2, section 2.3.2). However, as no KDM activity was observed for isolated PHF2, the same buffer may not be most suitable for PHF2. To investigate whether the same trend of pH and NaCl preference occurs for PHF2 as for PHF8 and KIAA1718, FP binding assays of $PHF2_{PHD\text{-}JmjC}$ were carried out. The binding assays for PHF8 and KIAA1718 were carried out using the optimal buffer tested for activity assays, i.e. 50 mM HEPES (pH 7.0) and 50 mM NaCl as described in Chapter 2.

Initially, FP binding assays with $H3_{1\text{-}15}$K4me3K9me2-FITC were conducted to investigate the buffer pH preference of PHF2 for substrate binding. The results at pH 7.0 and 8.0 suggested that the substrate binding affinity of PHF2 is higher at neutral pH 7.0 compared to pH 8.0, when other buffer components remained the same (Figure 3.6 A & B). The buffers in each pH value also varied in NaCl concentrations (50 mM, 250 mM and 500 mM). The K_D values of $PHF2_{PHD\text{-}JmjC}$ binding with $H3_{1\text{-}15}$K4me3K9me2-FITC in 50 mM, 250 mM and 500 mM NaCl buffer (pH 8.0) were 0.77 $\pm$ 0.08 µM, 0.91 $\pm$ 0.11 µM and 1.36 $\pm$ 0.14 µM respectively (Figure 3.6 A). The data points for pH 7.0 buffer follow the same trend although they could not be interpreted to give a reliable K_D value (Figure 3.6 B), suggesting that the presence of NaCl in the assay buffer has a negative effect on the substrate binding of $PHF2_{PHD\text{-}JmjC}$. Therefore, a buffer containing 50 mM HEPES (pH 7.0), 50 mM NaCl was selected to maintain the optimal binding affinity while preventing precipitation, with the addition of 0.01% (v/v) of Tween-20 in order to hinder non-specific binding of the protein to the plate.

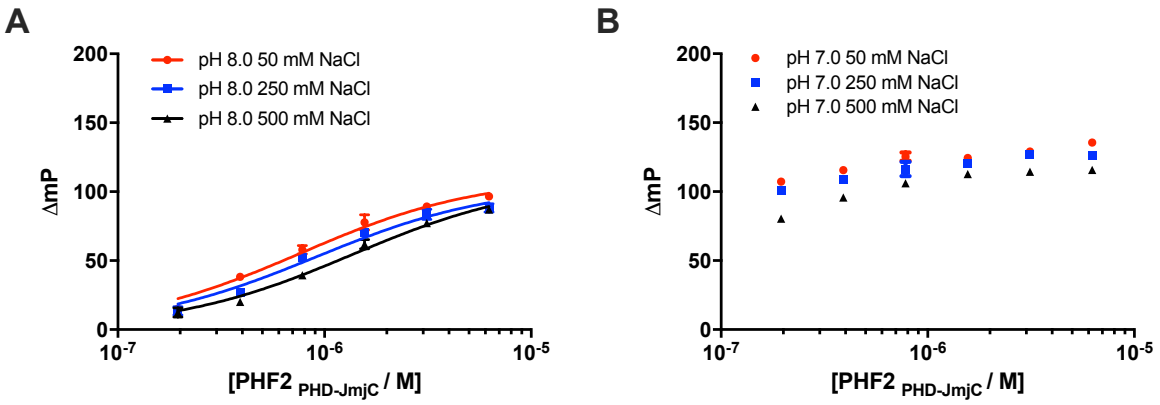


Figure 3.6 FP assays measuring the binding of $PHF2_{PHD\text{-}JmjC}$ to $H3_{1\text{-}15}K4me3K9me2$-FITC in **A)** 50 mM HEPES pH 8.0 and 0.01% (v/v) Tween-20 and NaCl. Data were fitted to a one-site specific binding model using Prism® software. R square values indicating the goodness of fitting are 0.99, 0.98 and 0.99 for red, blue and black curves respectively. **B)** 50 mM HEPES pH 7.0 and 0.01% (v/v) Tween-20 and NaCl. Red, black and blue triangles/circles represent fluorescence polarisation at each protein concentration in buffer with 50, 250 mM and 500 Mm NaCl. PHF2 (various concentrations) was incubated with $H3_{1\text{-}15}K4me3K9me2$-FITC (0.05 µM) and $NiSO_4$ (100 µM) in a clear bottom 96-well plate for 20 minutes in the dark at room temperature before analysis using a PHERA Star FS® plate reader. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

3.3.4 Substrate binding affinities of KDM7 enzymes and their variants

To better understand the substrate binding affinity of wild-type and variant KDM7s, a series of substrate binding assays were conducted using FP. First, $KIAA1718_{PHD\text{-}JmjC}$ was tested with two fluorescent labelled H3 peptides, $H3_{1\text{-}15}K4me3K9me2$-FITC and $H3_{1\text{-}15}K9me2$-FITC. According to the kinetic studies described in Chapter 2, and as reported,[2] KIAA1718 catalyses demethylation of H3K9me2; the presence of H3K4me3 reduced demethylation activity at K9me2 (Chapter 2, Section 2.3.4). Interestingly, the binding assay results revealed significant differences in the binding affinities of KIAA1718 for $H3_{1\text{-}15}K4me3K9me2$-FITC and $H3_{1\text{-}15}K9me2$-FITC, but in the opposite direction to the activity results. $KIAA1718_{PHD\text{-}JmjC}$ bound to $H3_{1\text{-}15}K4me3K9me2$-FITC with a K_D of 0.15 $\pm$ 0.01 µM (Figure 3.7 A), while the binding to $H3_{1\text{-}15}K9me2$-FITC was too weak to give a K_D value (Figure 3.7 B). This result is consistent with the previously reported K_D of 0.29 $\pm$ 0.04 µM for $H3_{1\text{-}24}K4me3K9me2$ as

determined by ITC (K_D of $H3_{1-24}K9me2$ was 60 ± 19 μM).[1,2] Notably, the binding of KIAA1718-$_{PHD-JmjC}$ in the presence and absence of 2OG overlapped at each enzyme concentration (Figure 3.7), suggesting that the presence of 2OG does not influence the binding of KIAA1718 to the tested substrates, $H3_{1-15}K4me3K9me2$-FITC and $H3_{1-15}K9me2$-FITC. This is contrary to the current knowledge of many 2OG oxygenases, where 2OG binding precedes substrate binding by stabilising the enzyme-substrate complex,[22] and thus prompted a hypothesis that the substrate binding of KIAA1718 is principally determined by binding to the PHD finger.

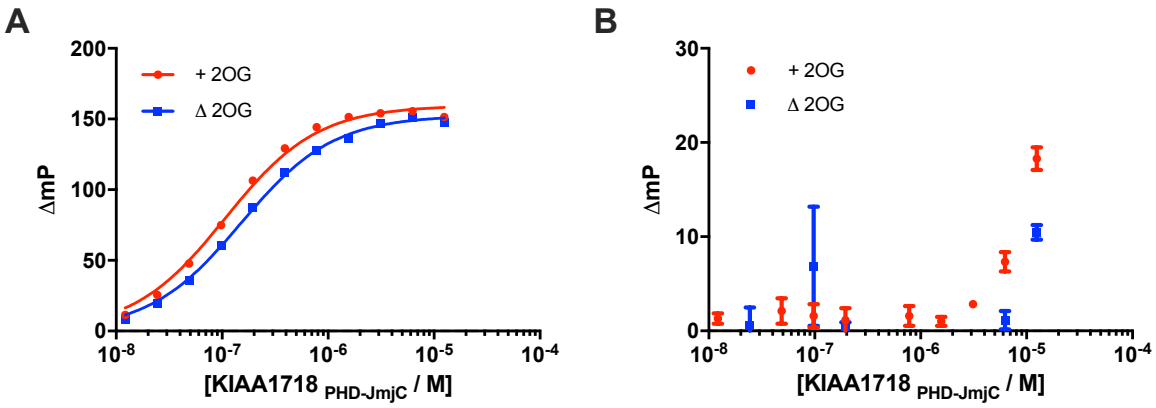


Figure 3.7 FP binding assays of $KIAA1718_{PHD-JmjC}$ to **A)** $H3_{1-15}K4me3K9me2$-FITC and **B)** $H3_{1-15}K9me2$-FITC with and without 2OG (100 μM). $KIAA1718_{PHD-JmjC}$ were incubated with $H3_{1-15}K4me3K9me2$-FITC (0.05 μM) or $H3_{1-15}K9me2$-FITC (0.05 μM) and NiSO4 (100 μM) in a clear bottom 96-well plate for 20 minutes in the dark at room temperature before analysing by a PHERA Star FS® plate reader. Data set A were fitted to a one-site specific binding model using Prism® software. R square values indicating the goodness of fitting are both 1.00 for red and blue curves respectively. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

To further investigate the contribution of the PHD finger to the substrate binding of KIAA1718, FP binding assays on $H3_{1-15}K4me3K9me2$-FITC and $H3_{1-15}K9me2$-FITC were carried out with $KIAA1718_{PHD}$. $KIAA1718_{PHD}$ bound $H3_{1-15}K4me3K9me2$-FITC with a K_D of 0.42 ± 0.05 μM (Figure 3.8), 2.8-fold higher than that of $KIAA1718_{PHD-JmjC}$ (Figure 3.7). Despite the 2.8-fold difference between $KIAA1718_{PHD}$ and $KIAA1718_{PHD-JmjC}$, both proteins bound to $H3_{1-}$

$_{15}$K4me3K9me2-FITC with K_D values in the nanomolar range (Table 3.1), suggesting that the PHD finger is important in KIAA1718 substrate binding. As for the binding with $H3_{1-15}$K9me2-FITC, the K_D value could not be measured since hardly any binding was observed even with the highest concentration of $KIAA1718_{PHD}$ tested (Figure 3.8).

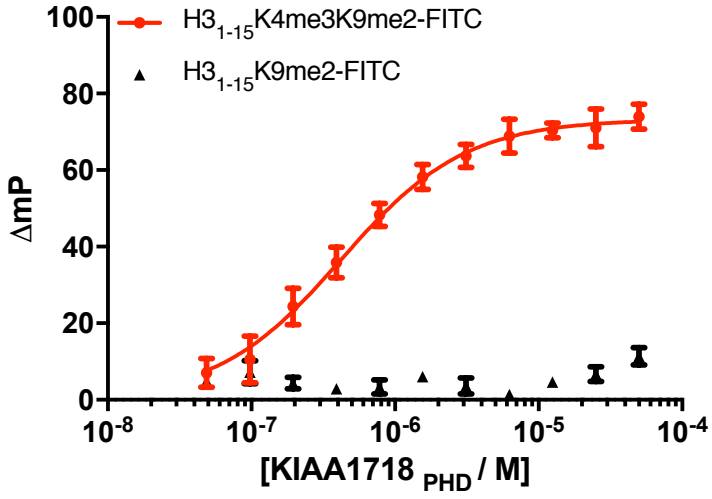


Figure 3.8 FP binding assays of $KIAA1718_{PHD}$ to $H3_{1-15}$K4me3K9me2-FITC and $H3_{1-15}$K9me2-FITC. $KIAA1718_{PHD}$ were incubated with $H3_{1-15}$K4me3K9me2-FITC (0.05 µM) or $H3_{1-15}$K9me2-FITC (0.05 µM) and NiSO4 (100 µM) in a clear bottom 96-well plate for 20 minutes in the dark at room temperature before analysing by a PHERA Star FS® plate reader. Data were fitted to a one-site specific binding model using Prism® software. R square value indicating the goodness of fitting is 0.96 for red curve. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

Next, it was asked whether the same contribution from the PHD finger to substrate binding for KIAA1718 manifests in the case of PHF8, which demethylates H3K9me2 more efficiently when combined with H3K4me3 (Chapter 2, Section 2.3.3). FP binding assays on $H3_{1-15}$K4me3K9me2-FITC and $H3_{1-15}$K9me2-FITC were carried out with $PHF8_{PHD-JmjC}$, $PHF8_{PHD}$ and $PHF8_{JmjC}$. $PHF8_{PHD-JmjC}$ bound to $H3_{1-15}$K4me3K9me2-FITC tightly, with an K_D of 0.83 $\pm$ 0.08 µM (Figure 3.9 A), while binding with $H3_{1-15}$K9me2-FITC gave an K_D of 7.48 $\pm$ 1.62 µM (Figure 3.9 B), indicating that substrate binding to $PHF8_{PHD-JmjC}$ is dominated by binding of the H3K4me3 mark. In addition, as for KIAA1718, the presence of 2OG did not influence the

substrate binding to $PHF8_{PHD\text{-}JmjC}$ (Figure 3.9). To investigate which domain facilitates the binding of H3K4me3, FP binding assays with $PHF8_{PHD}$ and $PHF8_{JmjC}$ were further conducted. $PHF8_{PHD}$ was observed to bind to $H3_{1\text{-}15}$K4me3K9me2-FITC with a K_D of 0.16 $\pm$ 0.01 μM (Figure 3.9 C) and did not efficiently bind to $H3_{1\text{-}15}$K9me2-FITC (> 50 μM, Figure 3.9 D). $PHF8_{JmjC}$ did not show any binding to $H3_{1\text{-}15}$K4me3K9me2-FITC (Figure 3.9 E), in line with the enzymatic activity studies (Chapter 2, Section 2.3.3). It was therefore concluded that the interaction with H3K4me3 by the PHD finger is the main contributor to PHF8 substrate binding. FP binding assays with $PHF8H319Y_{PHD\text{-}JmjC}$ further supported this proposal. $PHF8H319Y_{PHD\text{-}JmjC}$ bound to $H3_{1\text{-}15}$K4me3K9me2-FITC with an K_D of 1.10 $\pm$ 0.10 μM (Figure 3.9 F), similar to $PHF8_{PHD\text{-}JmjC}$ (Figure 3.9 A). The H to Y alteration which inactive the PHF8 enzyme did not affect the substrate binding of PHF8.

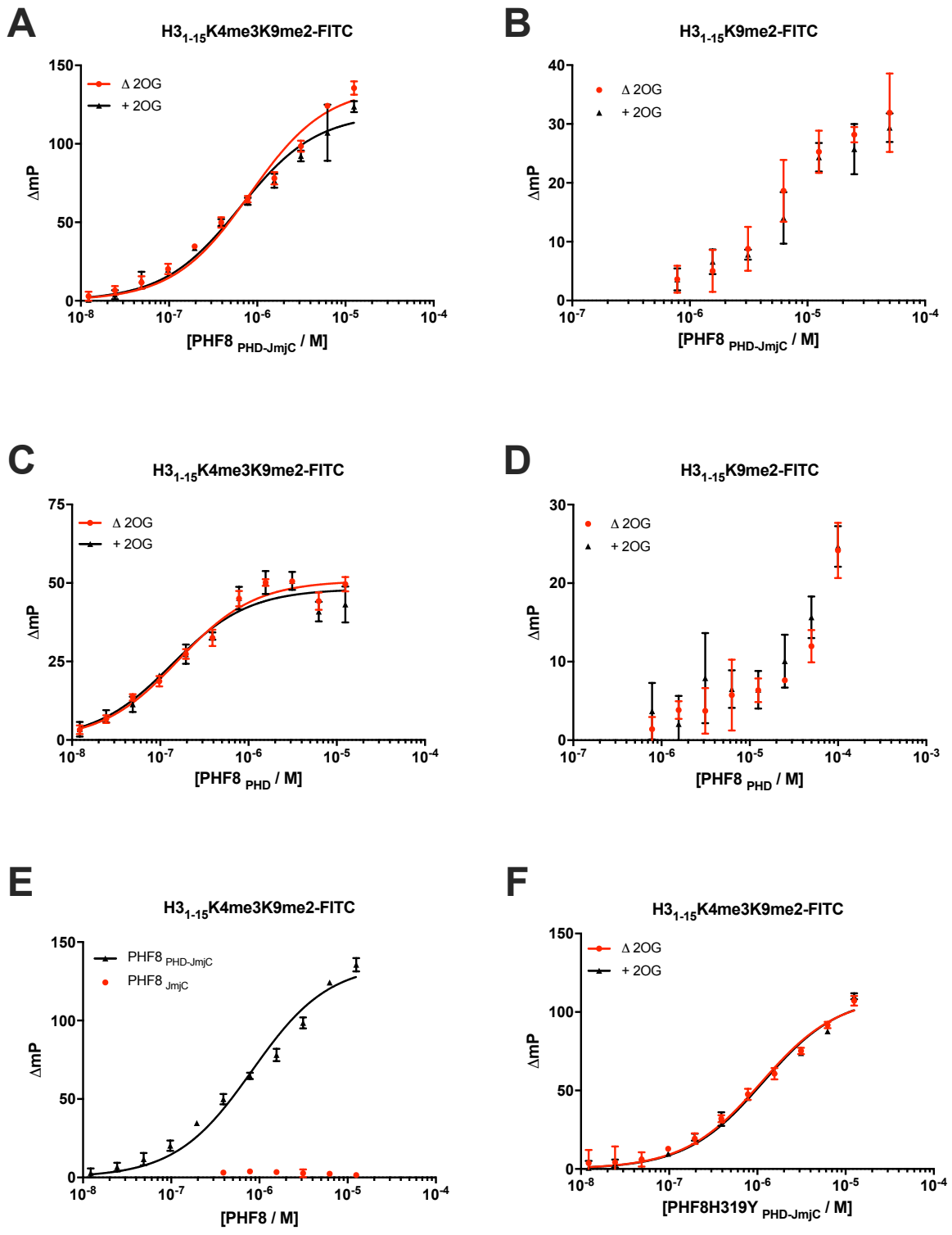


Figure 3.9 FP measurement of PHF8 proteins binding to histone substrates with and without 2OG (100 µM). **A)** $PHF8_{PHD\text{-}JmjC}$ binding to $H3_{1\text{-}15}K4me3K9me2$-FITC. R square values indicating the goodness of fitting are 0.98 and 0.97 for red and black curves respectively. **B)** $PHF8_{PHD\text{-}JmjC}$ binding to $H3_{1\text{-}15}K9me2$-FITC. **C)** $PHF8_{PHD}$ binding to $H3_{1\text{-}15}K4me3K9me2$-FITC. R square values indicating the goodness of fitting are 0.98 and 0.95 for red and black curves respectively. **D)** $PHF8_{PHD}$ binding to $H3_{1\text{-}15}K9me2$-FITC. **E)** $PHF8_{JmjC}$ binding to $H3_{1\text{-}15}K4me3K9me2$-FITC. R square values indicating the goodness of fitting is 0.98 for black curve. **F)** $PHF8H319Y_{PHD\text{-}JmjC}$ binding to $H3_{1\text{-}15}K4me3K9me2$-FITC. R square values indicating the goodness of fitting are both 0.98 for red and black curves.

PHF8 proteins were incubated with $H3_{1-15}$K4me3K9me2-FITC (0.05 µM) or $H3_{1-15}$K9me2-FITC (0.05 µM) and NiSO4 (100 µM) in a clear bottom 96-well plate for 20 minutes in the dark at room temperature before analysing using a PHERA Star FS® plate reader. Data were fitted to a one-site specific binding model using Prism® software. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

PHF2, which has no reported catalytic activity in isolated form, was subsequently tested for binding by FP binding assays. $PHF2_{PHD\text{-}JmjC}$ and $PHF2_{PHD}$ gave binding K_D values less than 1 µM, while $PHF2_{JmjC}$ showed no detectable binding (Figure 3.10 A). The presence of 2OG did not influence the binding of $PHF2_{PHD\text{-}JmjC}$ on $H3_{1-15}$K4me3K9me2-FITC (Figure 3.10 B). The overall results with PHF2 proteins clearly imply that the interaction between the H3K4me3 mark and the PHD finger of PHF2 dominates substrate binding.

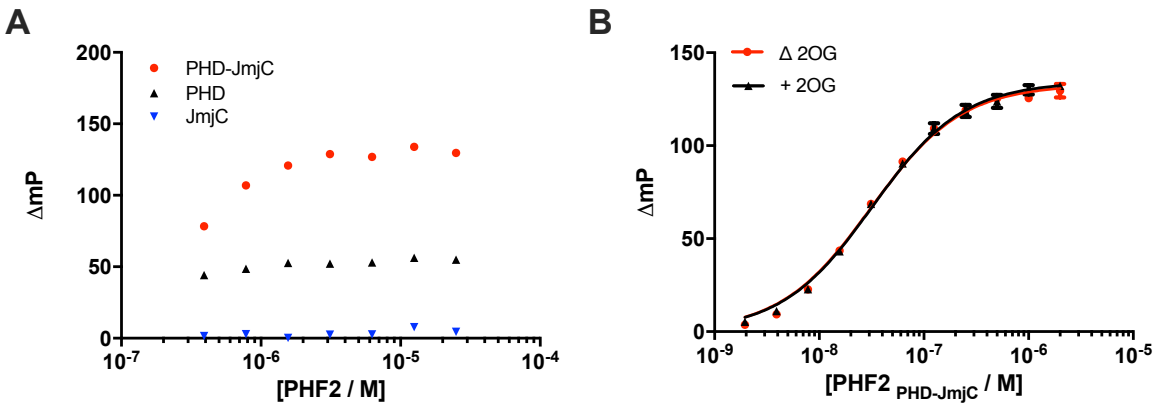


Figure 3.10 FP measurements of: **A)** Various domains of PHF2 **B)** $PHF2_{PHD\text{-}JmjC}$ binding to $H3_{1-15}$K4me3K9me2-FITC with and without 2OG (100 µM). PHF2 proteins were incubated with $H3_{1-15}$K4me3K9me2-FITC (0.05 µM) and NiSO4 (100 µM) in a clear bottom 96-well plate for 20 minutes in the dark at room temperature before analysing using a PHERA Star FS® plate reader. Data set B were fitted to a one-site specific binding model using Prism® software. R square values indicating the goodness of fitting are both 0.99 for red and black curves. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

Finally, FP binding assays with $PHF8H319Y_{PHD\text{-}JmjC}$ and $PHF2Y321H_{PHD\text{-}JmjC}$ to $H3_{1-15}$K4me3K9me2-FITC gave similar K_D values to their wild-type counterparts. The $PHF8H319Y_{PHD\text{-}JmjC}$ gave a K_D of 1.10 $\pm$ 0.10 µM, similar to the K_D of 0.83 $\pm$ 0.08 µM for

PHF8$_{PHD-JmjC}$ (Figure 3.11 A). PHF2Y321H$_{PHD-JmjC}$ and PHF2$_{PHD-JmjC}$ bound to H3$_{1-15}$K4me3K9me2-FITC in a similar manner despite the lack of measurable K_D values (Figure 3.11 B).

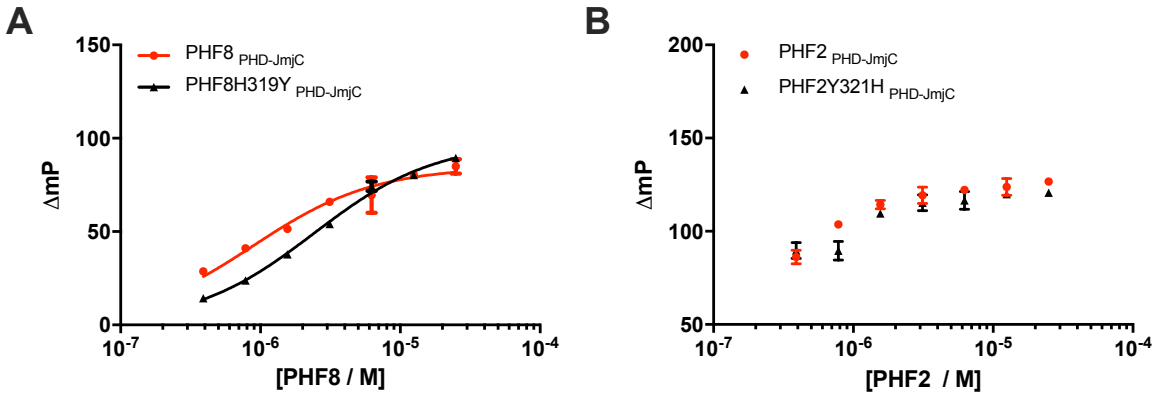


Figure 3.11 FP measurements of: **A)** PHF8 and **B)** PHF2 proteins binding to H3$_{1-15}$K4me3K9me2-FITC. PHF2 or PHF8 proteins were incubated with H3$_{1-15}$K4me3K9me2-FITC (0.05 µM) and NiSO4 (100 µM) in a clear bottom 96-well plate for 20 minutes in the dark at room temperature before analysing using a PHERA Star FS® plate reader. Data set A were fitted to a one-site specific binding model using Prism® software. R square values indicating the goodness of fitting are 0.95 and 0.99 for red and black curves respectively. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

In summary the K_D values of KDM7s binding to the two peptides H3$_{1-15}$K4me3K9me2-FITC and H3$_{1-15}$K9me2-FITC is shown in Table 3.1. Based on the results described above, interactions with the PHD fingers appears to drive the binding of KDM7s to their histone substrates, and hence may regulate their enzymatic activity. To investigate the influence of substrate binding by the PHD fingers on the enzymatic activities of KDM7s, kinetic assays with a tri-ethyl H3K4 peptide (H3$_{1-15}$K4et3K9me2, synthesised by Mr Roman Belle) with PHF8$_{PHD-JmjC}$ were conducted using MALDI MS-based demethylation assays (Chapter 5, section 5.5), as well as with the H3$_{1-15}$K4me3K9me2 peptide (Figure 3.12 A). H3$_{1-15}$K4et3K9me2 was demonstrated to bind to PHF8 more tightly than H3$_{1-15}$K4me3K9me2 by Roman Belle using ITC assays (unpublished data). The results of kinetic assays were attempted to be fitted by the GraphPad Prism7 software using a substrate inhibition fitting

model. However, no reliable parameters could be generated, likely due the effects of complicated combined substrate/product inhibition. Nonetheless, the results of the kinetic assays indicate that the apparent initial rate of demethylation catalysed as by $PHF8_{PHD\text{-}JmjC}$ is substantially increased with $H3_{1\text{-}15}K4et3K9me2$ compared to $H3_{1\text{-}15}K4me3K9me2$ (Figure 3.12 A). This observation further supports the role of the PHD in substrate binding and hence catalysis.

The results of a competition experiments using $H3_{1\text{-}15}K4et3K9me2$ and $H3_{1\text{-}15}K4me3K9me2$ against $H3_{1\text{-}15}K4me3K9me2$-FITC with $PHF8_{PHD\text{-}JmjC}$ suggested that $H3_{1\text{-}15}K4et3K9me2$ (IC_{50} = 0.69 µM) binds to PHF8 approximately four times tighter than $H3_{1\text{-}15}K4me3K9me2$ (IC_{50} = 2.6 µM) (Figure 3.12 B).

Overall, the results described in this Section demonstrate that the PHD fingers of the KDM7 enzymes appear to dominate substrate binding and influence the enzymatic activity of PHF8.

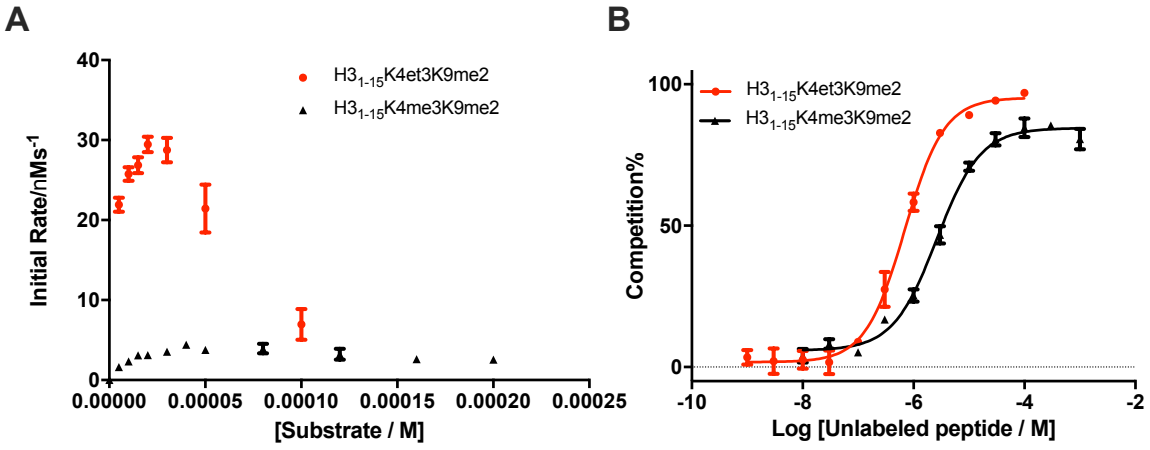


Figure 3.12 A) Measurement of K_M values of demethylation reactions of $H3_{1\text{-}15}K4et3K9me2$ and $H3_{1\text{-}15}K4me3K9me2$ catalysed by $PHF8_{PHD\text{-}JmjC}$. **B)** FP measurement of displacement IC_{50} values of $H3_{1\text{-}15}K4et3K9me2$ and $H3_{1\text{-}15}K4me3K9me2$ against $H3_{1\text{-}15}K4me3K9me2$-FITC with $PHF8_{PHD\text{-}JmjC}$. Data set B were fitted to a one-site specific binding model using Prism® software. R square values indicating the goodness of fitting are 0.99 and 0.98 for red and black curves respectively. For both data set A and B, a sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

Table 3.1 Summary of the K_D of KDM7s binding to $H3_{1-15}$K4me3K9me2-FITC and $H3_{1-15}$K9me2-FITC. NM indicates that the K_D values are not measurable under the tested conditions.

		K_D H3K4me3K9me2/nM no 2OG	K_D H3K4me3K9me2/nM with 2OG	K_D H3K9me2/μM no 2OG	K_D H3K9me2/μM with 2OG
KIAA1718	PHD	419.2 ± 54.0	-	-	-
	PHD-JmjC	150.4 ± 4.3	106.9 ± 3.7	> 50	> 50
	JmjC	NM	NM	NM	NM
PHF2	PHD	< 1 μM	-	-	-
	PHD-JmjC	31.3 ± 1.4	32.1 ± 1.3	-	-
	JmjC	NM	NM	NM	NM
PHF8	PHD	159.5 ± 14.1	135.4 ± 18.2	> 50	> 50
	PHD-JmjC	826.2 ± 75.8	638.4 ± 68.1	NM	NM

3.3.5 Buffer optimisations for KMD7s and their variants for methylstat binding

Methylstat is reported to be a selective JmjC domain-containing histone demethylase (JmjC-KDM) inhibitor.[21] FI-methylstat is a fluorescent analogue of methylstat.[21] Methylstat is predicted to target the JmjC domain, at both substrate and 2OG binding pockets. The binding of $PHF2_{PHD\text{-}JmjC}$ to FI-methylstat was evaluated in the set of buffers used for the binding of $H3_{1\text{-}15}$K4me3k9me2-FITC. Thus, FI-methylstat will likely to bind (if at all) to the JmjC domain of PHF2. Under all the conditions tested, $PHF2_{PHD\text{-}JmjC}$ was observed to bind to FI-methylstat with a K_D >10 µM. as observed for $H3_{1\text{-}15}$K4me3k9me2-FITC binding to PHF2 (Figure 3.6), increased salt concentrations negatively affect FI-methylstat binding to PHF2 (Figure 3.13). It was difficult to compare the influence of pH since the weak binding affinity precluded determining reliable K_D measurements at either of the tested pH values. However, use of the pH 8.0 buffer gave a more intense signal for the same protein concentration compared to that at pH 7.0 (Figure 3,13). Therefore, a buffer containing 50 mM HEPES (pH 8.0), 50 mM NaCl was selected for FI-methylstat binding studies. As for $H3_{1\text{-}15}$K4me3k9me2-FITC binding, 0.01% (v/v) Tween-20 was also added in the binding buffer.

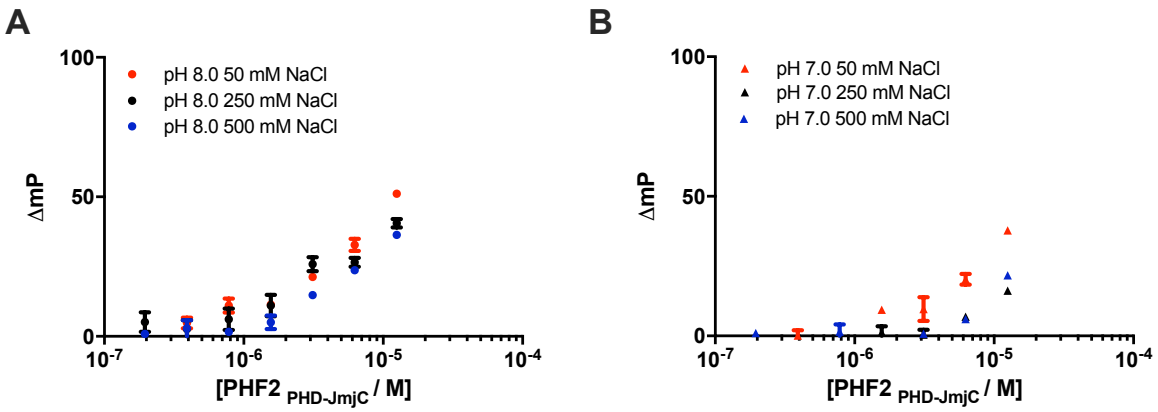


Figure 3.13 FP assays measuring the binding of $PHF2_{PHD\text{-}JmjC}$ to FI-methylstat in **A)** 50 mM HEPES pH 8.0 and 0.01% (v/v) Tween-20 and NaCl. **B)** 50 mM HEPES pH 7.0 and 0.01% (v/v) Tween-20 and NaCl. Red, black and blue triangles/circles represented the fluorescence polarisation at each protein concentration in buffer with 50, 250 mM and 500 Mm NaCl. PHF2 (various concentrations) was incubated with FI-methylstat (0.0125 µM) and $NiSO_4$ (100 µM) in a clear bottom 96-well plate for 20 minutes in the dark at room temperature before analysing by a PHERA Star FS® plate reader. Data were fitted to a one-site specific binding model using Prism® software. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

Finally, FP binding assays of FI-methylstat by $PHF8_{PHD\text{-}JmjC}$ and $PHF8H319Y_{PHD\text{-}JmjC}$ were carried out using pH 7.0 and pH 8.0 buffers, in order to investigate whether the H to Y substitution changes the pH preference of PHF8 binding to FI-methylstat. $PHF8_{PHD\text{-}JmjC}$ preferred the buffer in pH 7.0 over pH 8.0 (Figure 3.14 A), as did $PHF8H319Y_{PHD\text{-}JmjC}$ (Figure 3.14 B). Interestingly, although $PHF8H319Y_{PHD\text{-}JmjC}$ showed the same pH preference as $PHF8_{PHD\text{-}JmjC}$, it seemed to be more affected by the buffer pH than the wild-type protein, with more than 13-fold difference in the binding K_D values at pH 7.0 compared to pH 8.0 (5-fold for $PHF8_{PHD\text{-}JmjC}$, Table 3.2).

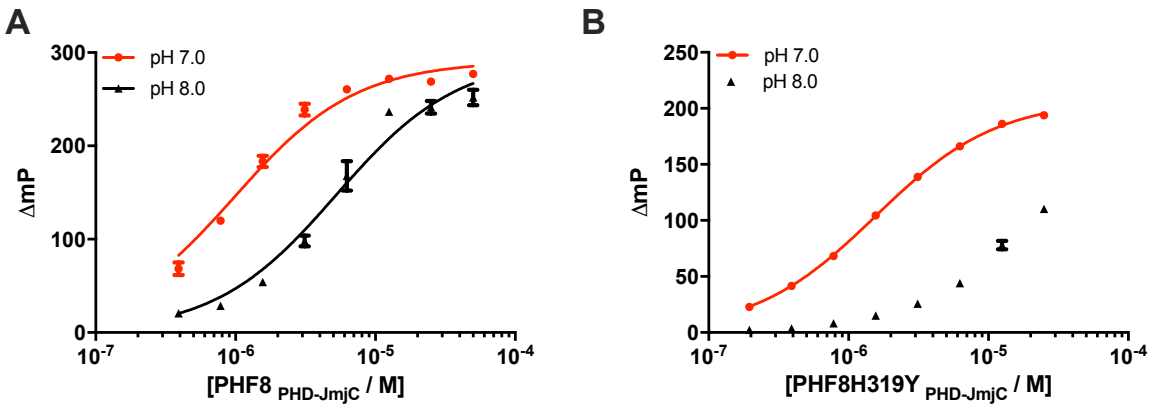


Figure 3.14 FP assays measuring the binding of **A)** $PHF8_{PHD\text{-}JmjC}$ to FI-methylstat in 50 mM HEPES, 50 mM NaCl and 0.01% (v/v) Tween-20. R square values indicating the goodness of fitting are 0.99 and 0.97 for the red and black curves respectively. **B)** $PHF8H319Y_{PHD\text{-}JmjC}$ to FI-methylstat in 50 mM HEPES and 150 mM and 0.01% (v/v) Tween-20. R square value indicating the goodness of fitting is 1.00 for the red curve. PHF8 proteins (various concentrations) were incubated with FI-methylstat (0.0125 µM) and NiSO4 (100 µM) in a clear bottom 96-well plate for 20 minutes in the dark at room temperature before analysing by a PHERA Star FS® plate reader. Data were fitted to a one-site specific binding model using Prism® software. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

The binding K_D values resulting from the binding assays described in this Section are summarised in Table 3.2. 50 mM HEPES (pH 7.0), 50 mM NaCl and 0.01% (v/v) Tween-20 was used as the buffer for KDM7 FP substrate binding assays and KIAA1718/PHF8 FI-methylstat binding assays. 50 mM HEPES (pH 8.0), 50 mM NaCl and 0.01% (v/v) Tween-20 was used for PHF2 FI-methylstat binding assays. In scenarios where PHF2 was apparently destabilised (as evidenced by precipitation observed during the operation of the FP assays), 5 mM TCEP was added to improve the enzyme stability.

Table 3.2 K_D values of FI-methylstat measured by FP binding assays of KDM7 enzymes.

			$PHF2_{PHD\text{-}JmjC}$			
pH	7.0			8.0		
NaCl / mM	50	250	500	50	250	500
K_D / µM	<1	<1	<1	0.77 ± 0.09	0.91 ± 0.01	1.36 ± 0.14
	$PHF8_{PHD\text{-}JmjC}$			$PHF8H319Y_{PHD\text{-}JmjC}$		
pH	7.0	8.0		7.0	8.0	
NaCl / mM	50	50		150	150	
K_D / µM	0.98 ± 0.07	5.25 ± 0.61		1.56 ± 0.03	>20	

3.3.6 Methylstat binding studies of KDM7 proteins and their variants

This Section describes work using FI-methylstat as a probe to investigate the mode of binding of the JmjC domains of KDM7s. It was aimed at further understanding the relationships between enzymatic activity and binding through the JmjC domain of KDM7s. It was envisaged that, especially for the catalytically inactive PHF2, such fluorescent binding studies may be useful for studying co-factor/substrate binding.

For each member of the KDM7 subfamily, the binding assays of FI-methylstat were carried out in the presence and absence of 100 μM 2OG. Protein constructs containing JmjC domains that were the subject of FP methylstat binding assays were: $KIAA1718_{PHD\text{-}JmjC}$, $PHF8_{PHD\text{-}JmjC}$, $PHF8_{JmjC}$, $PHF8H319Y_{PHD\text{-}JmjC}$, $PHF8H319Y_{JmjC}$, $PHF2_{PHD\text{-}JmjC}$, $PHF2_{JmjC}$, $PHF2Y321H_{PHD\text{-}JmjC}$ and $PHF2Y321H_{JmjC}$. In the work described in this Section, only binding of KDM7 proteins to FI-methylstat was investigated. More detailed experiments and results about FP displacement assays are described in the next Chapter (Chapter 4, Section 4.3.2).

For $KIAA1718_{PHD\text{-}JmjC}$, binding to methylstat was relativelytight with a K_D of 0.13 $\pm$ 0.01 μM (without 2OG, Figure 3.15). When 100 μM 2OG was incubated together with the binding mixture, the K_D was elevated (K_D > 10 μM). This observation is consistent with the fact that methylstat binds at the 2OG binding site in the JmjC-domain, supporting the proposal that FI-methylstat is competitive with respect to 2OG binding.

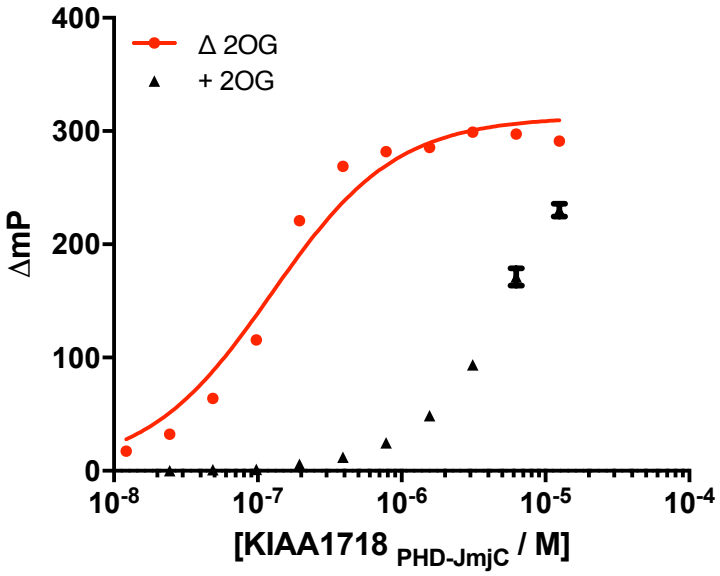


Figure 3.15 2OG disturbed FI-methylstat binding of KIAA1718. Red circles represented the fluorescence polarisation at each enzyme concentration in the absence of 2OG. Data were fitted in a one-site specific binding model using Prism® software, where the bottom of fitting curves were fixed at 0 ΔmP. R square value indicating the goodness of fitting is 0.99 for the red curve. A no enzyme reaction was used as negative control for normalisation. Points are plotted as the mean of technical triplicates, and error bars give ± one s.d.

PHF8 showed a different level of binding affinity towards FI-methylstat compared to KIAA1718. $PHF8_{PHD\text{-}JmjC}$ bound the FI-methylstat with a K_D of 0.98 ± 0.07 μM (Figure 3.16 A) and $PHF8_{JmjC}$ bound with a similar magnitude, giving a K_D of 1.33 ± 0.04 μM (Figure 3.16 B). Moreover, as for KIAA1718, addition of 2OG weakened the binding of FI-methylstat to both $PHF8_{PHD\text{-}JmjC}$ and $PHF8_{JmjC}$ by more than tenfold (Figure 3.16 A and B). Although PHF8 and KIAA1718 shared ~ 65% identity between their JmjC domains and their iron-binding motif are the same (Appendix 1&2), the disparity of methylstat binding affinity implies that the interactions involved in methylstat binding at the active site are not the same for the two KDM7 enzymes. This observation may reflect observations on the different selectivities of the JmjC domains of PHF8 and KIAA1718 towards dimethylated lysine at various positions on histone H3 peptides (Chapter 2 Section 2.3.5). However, attempts on crystallographic study using the KDM7 protein crystal structures available in the Protein Data Bank (PDB) did not find other interactions which may cause the different substrate preference of PHF8 and

KIAA1718. Detailed biophysical analyses, such as protein NMR studies, are required so as to clarify which interactions are involved in cofactor and substrate binding.

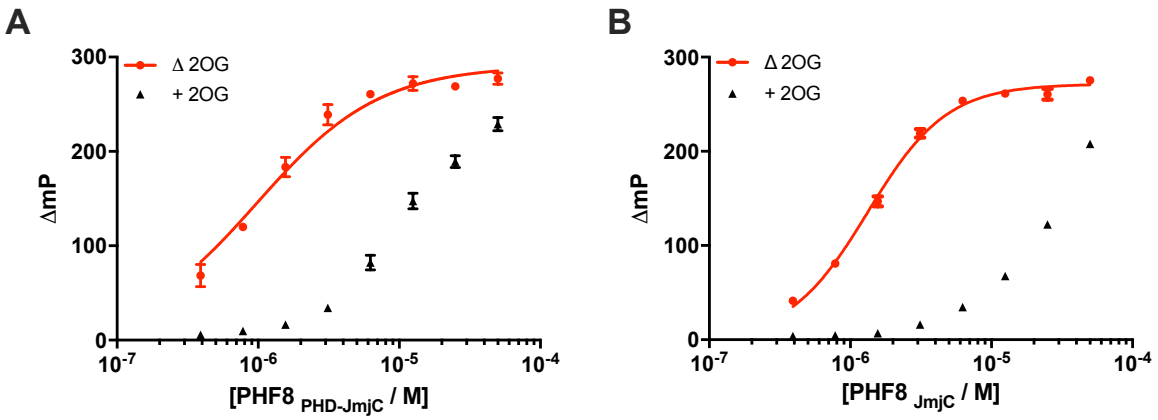


Figure 3.16 FP measurement of **A)** $PHF8_{PHD\text{-}JmjC,}$ R square value indicating the goodness of fitting is 0.99 for the red curve. **B)** $PHF8_{JmjC}$ binding to FI-methylstat +/- 2OG (100 µM). R square value indicating the goodness of fitting is 1.00 for the red curve. Red circles represented the fluorescence polarisation at each enzyme concentration in the absence of 2OG. Data were fitted in a one-site specific binding model using Prism® software, where the bottom of fitting curves were fixed at 0 ΔmP. A no enzyme reaction was used as negative control for normalisation. Points are plotted as the mean of technical triplicates, and error bars give ± one s.d.

$PHF8H319Y_{PHD\text{-}JmjC}$ gave a binding K_D of 1.56 $\pm$ 0.03 µM (Figure 3.17 A) with FI-methylstat, and $PHF8H319Y_{JmjC}$ gave a K_D of <1 µM (estimated from the incomplete curve, Figure 3.17 B), which were the same as for the $PHF8_{PHD\text{-}JmjC}$ and $PHF8_{JmjC}$ respectively within the error range, suggesting that although the His to Tyr substitution causes PHF8 to be non-catalytic and likely disturbs the iron binding motif, the alteration appears not to have influenced its binding to FI-methylstat within detection limits. The combination of these results suggests that there may be different amino acid residues / active site interactions responsible for cofactor and substrate binding in different KDM7s.

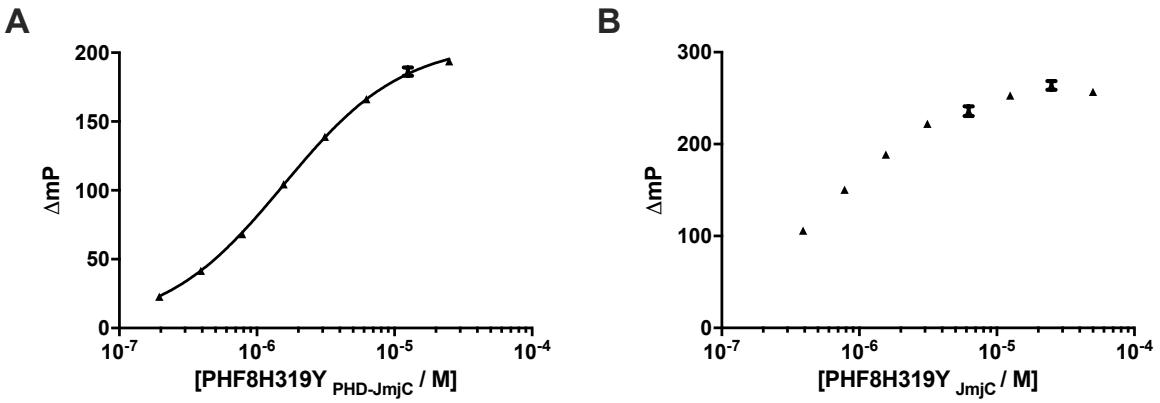


Figure 3.17 FP measurement of **A)** $PHF8_{PHD\text{-}JmjC}$ **B)** $PHF8_{JmjC}$ binding to FI-methylstat in the absence of 2OG. Data set A was fitted in a one-site specific binding model using Prism® software, where the bottom of fitting curves were fixed at 0 ΔmP. R square value indicating the goodness of fitting is 1.00 for the black curve. A no enzyme reaction was used as negative control for normalisation. Points are plotted as the mean of technical triplicates, and error bars give ± one s.d.

$PHF2_{PHD\text{-}JmjC}$ binds FI-methylstat with a K_D of 2.5 $\pm$ 0.21 μM (Figure 3.18 A), while $PHF2_{JmjC}$ with K_D of 6.37 $\pm$ 1.03 μM (Figure 3.18 C), about three to four-fold higher than the corresponding PHF8 constructs (Figure 3.16). $PHF2Y321H_{PHD\text{-}JmjC}$ and $PHF2Y321H_{JmjC}$ did not gave accurate K_D values due to the incomplete curves observed over the tested concentration ranges (Figure 3.18 B & D). By contrast with PHF8 and KIAA1718, the binding of FI-methylstat to $PHF2_{PHD\text{-}JmjC}$ and $PHF2Y321H_{PHD\text{-}JmjC}$ was not affected by the addition of 2OG.

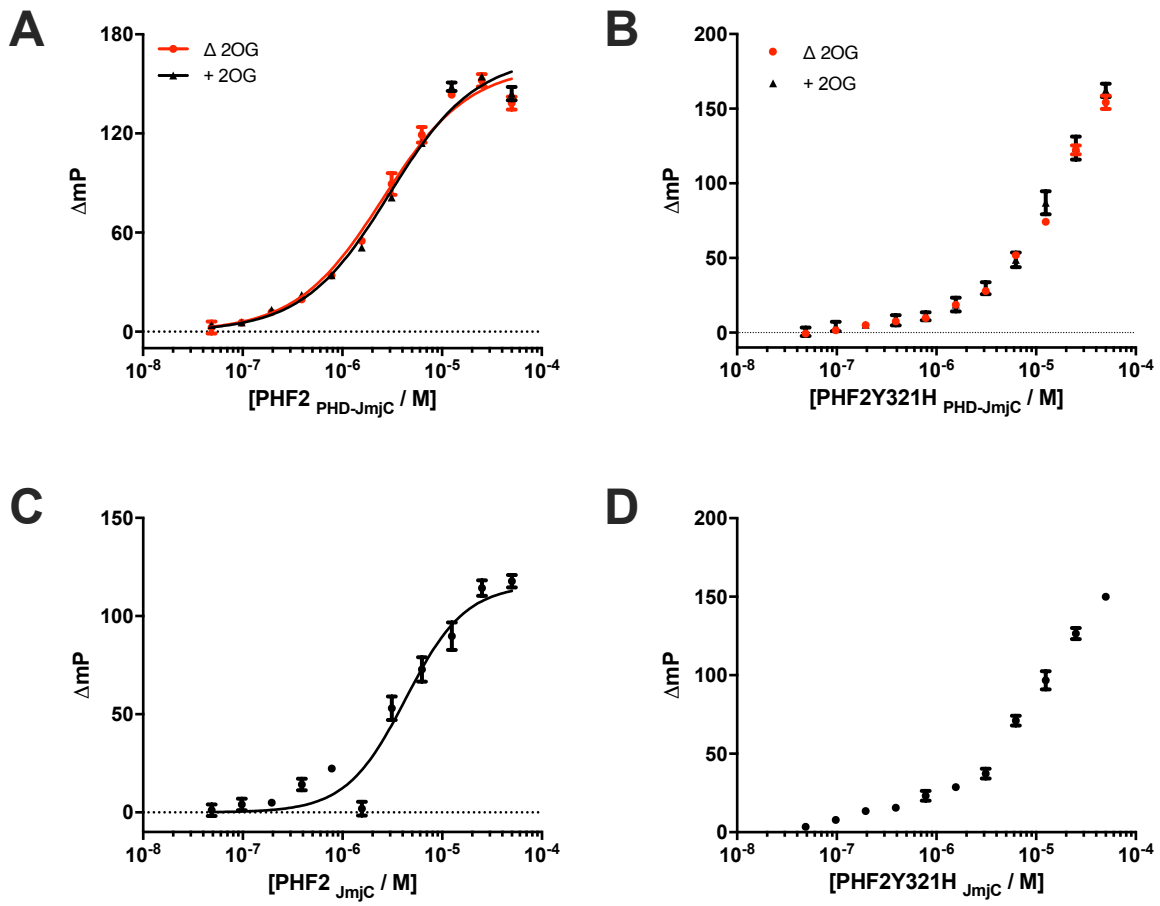


Figure 3.18 FP measurement of **A)** $PHF2_{PHD\text{-}JmjC}$ and **B)** $PHF2Y321H_{PHD\text{-}JmjC}$ binding to FI-methylstat +/- 2OG. Red circles represent fluorescence polarisation at each enzyme concentration in the absence of 2OG. **C)** $PHF2_{JmjC}$ and **D)** $PHF2Y321H_{JmjC}$ binding to FI-methylstat without 2OG. Data were fitted in a one-site specific binding model using Prism® software. A no enzyme reaction was used as negative control for normalisation. Points are plotted as the mean of technical triplicates, and error bars give ± one s.d.

Together with the activity studies described in Chapter 2, the observation of 2OG not affecting PHF2 FI-methylstat binding raises the question as to whether 2OG binds to PHF2 (potentially as the cofactor for demethylation or other reaction). NMR observed 2OG turnover assays were thereby carried out for $PHF2_{PHD\text{-}JmjC}$. Comparing with $PHF8_{PHD\text{-}JmjC}$ which turned over 2OG into succinate over 1-hour incubation, no 2OG turnover product was detected for $PHF2_{PHD\text{-}JmjC}$ (Figure 3.19), further supporting the proposal that 2OG may be not utilized by PHF2.

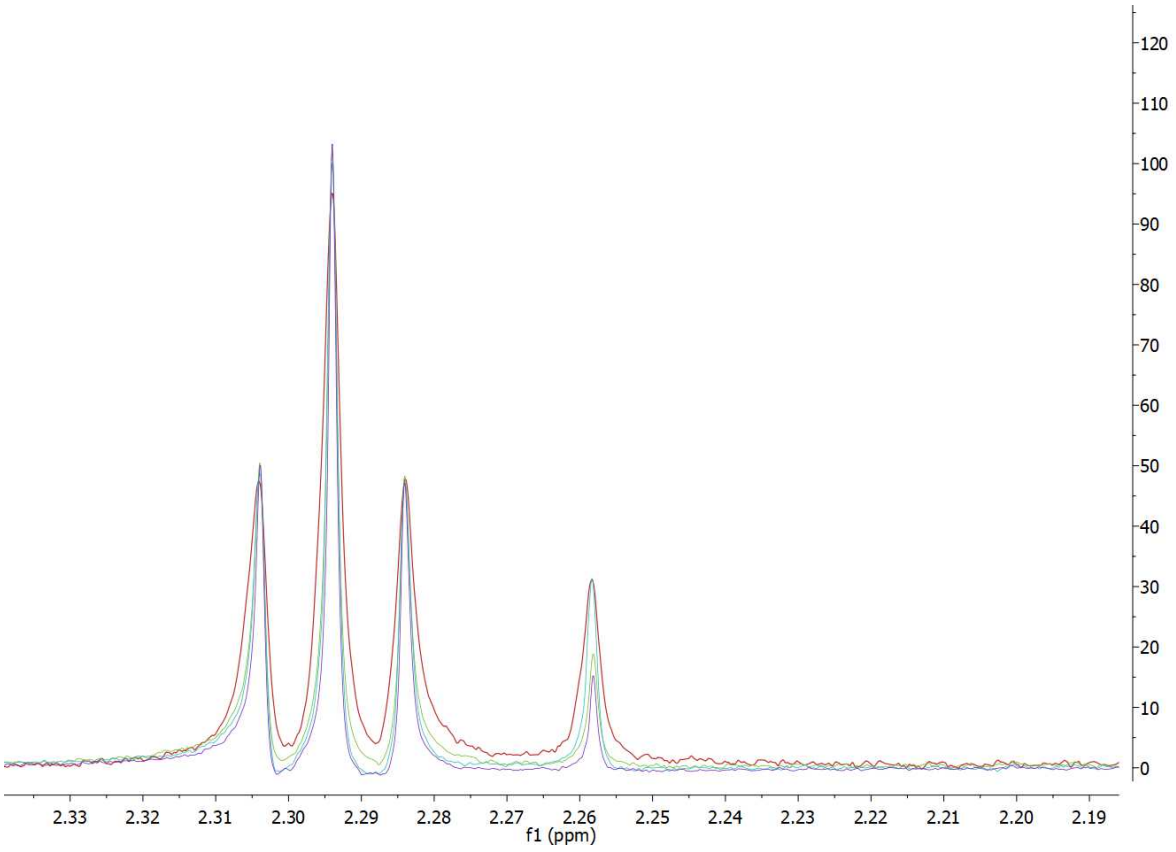


Figure 3.19 NMR ^{1}H studies of 2OG turnover by $PHF2_{PHD\text{-}JmjC}$ and $PHF8_{PHD\text{-}JmjC}$. Peaks around 2.26 ppm represent the peaks for 2OG turnover product – succinate. Red: $PHF8_{PHD\text{-}JmjC}$ turnover assay. Blue: $PHF8_{PHD\text{-}JmjC}$ negative control. Green: $PHF2_{PHD\text{-}JmjC}$ turnover assay. Purple: $PHF2_{PHD\text{-}JmjC}$ negative control.

A summary of the binding parameters of all the FP binding assays using FI-methylstat as described in this Section is shown in Table 3.3. In conclusion, these results show that FI-methylstat is a useful tool for investigating ligand binding to JmjC domains. Importantly, since PHF2 is catalytic inactive, the FP assay provides new opportunities to investigate PHF2 (and its inhibitors), opening up possibilities for other JmjC domains without catalytic activity (e.g. JARID2, HR).[23,24]

Table 3.3 FI-methylstat binding K_D values of KDM7 subfamily.

	PHF8		PHF8H319Y		KIAA1718
Construct	PHD-JmjC	JmjC	PHD-JmjC	JmjC	PHD-JmjC
K_D/ µM	0.98 ± 0.07	1.33 ± 0.04	1.56 ± 0.03	NM	0.13 ± 0.01
	PHF2		**PHF2Y321H**		
Construct	PHD-JmjC	JmjC	PHD-JmjC	JmjC	
K_D/ µM	2.5 ± 0.21	6.37 ± 1.03	NM	NM	

3.3.7 Metal coordination studies in KDM7 JmjC-domain active site

To determine the significance of the tyrosine residue in the iron binding motif of PHF2, iron binding assays were conducted on apo $PHF2_{JmjC}$ and $PHF8_{JmjC}$ enzymes (i.e. non-metalled forms). Apo $PHF8_{JmjC}$ and apo $PHF2_{JmjC}$ (wild-type and variants) were obtained by treating the holoenzymes (i.e. fully metallated) with 3 mM EDTA at 4 °C overnight. The fluorescence intensity (FI) assay was chosen initially for Fe(II) binding measurement of apo PHF8 because of its efficient nature and lack of requirement for fluorescent ligands. However, attempts to obtain decent Fe(II) binding curves for apo $PHF8_{JmjC}$ and $PHF2_{JmjC}$ both failed due to the low signal window (data not shown).

Next, non-denaturing mass spectrometry was used to investigate the apo status of enzymes and to detect binding. This work was carried out in collaboration with Dr Joseph Gault (Professor Dame Carol Robinson group, Oxford). The apo enzymes ($PHF8_{JmjC}$, $PHF8H319Y_{JmjC}$, $PHF2_{JmjC}$ and $PHF2Y321H_{JmjC}$) were analysed using a Q-Exactive™ Hybrid Quadrupole-Orbitrap™ Mass Spectrometer (ThermoFisher) to verify their apo status. All enzymes were observed to be fully apo within detection limits.

The apoenzymes were then incubated with various concentration of $(NH_4)_2Fe(SO_4)_2$ on ice for 5min before subsequent analysis by MS. The concentration of $(NH_4)_2Fe(SO_4)_2$ used was initially the same as the apoenzyme tested. If the starting sample was not. detected well in the MS machine, the concentration of $(NH_4)_2Fe(SO_4)_2$ was raised gradually up to 40 μM in a stepwise manner. Overall, the metal ion concentration was kept low to prevent damage to the mass spectrometer. Apo $PHF2_{JmjC}$ and apo $PHF2Y321H_{JmjC}$ both were observed to bind iron under the condition tested (enzyme: $(NH_4)_2Fe(SO_4)_2$ = 1: 10) (Figure 3.20 A). Unfortunately, the spectra were not sufficiently well resolved to conduct quantitative calculation of binding affinities due to the interference of SO_4^{2-} ions. The spectrum of apo $PHF8H319Y_{JmjC}$ did not show a +Fe peak with the most saturating metal iron concentration, whereas the apo $PHF8_{JmjC}$ showed about 50% binding at 2 μM $(NH_4)_2Fe(SO_4)_2$ (enzyme: $(NH_4)_2Fe(SO_4)_2$= 1: 1, Figure 3.20 B). These results indicate that the PHF8H319Y variant had lost its enzymatic activity most likely due to loss of Fe(II) binding affinity. The MS method is in principle amenable for K_D determinations. However, the apoenzyme forms of PHF2 and PHF8 tended to precipitate even at low concentrations, and it was challenging to gain clean spectra to determine the percentage of the bound enzyme. Therefore, ITC assays were investigated for quantitative measurement of metal binding.

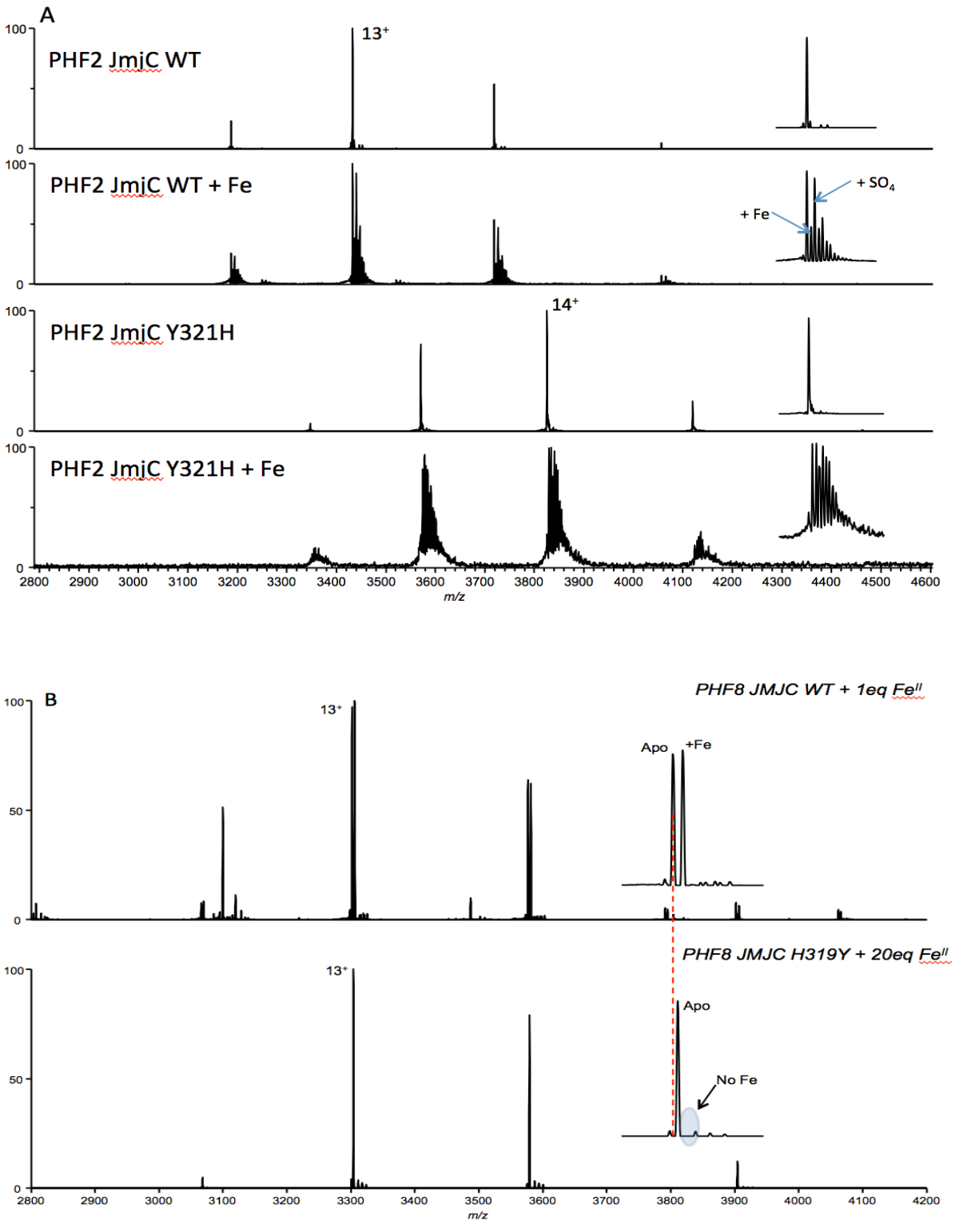


Figure 3.20 Non-denaturing mass spectrometry detecting iron binding for: **A)** PHF2 JmjC wild-type and the Y321H variant. The $(NH_4)_2Fe(SO_4)_2$ (20 µM) was incubated with an enzyme (2 µM) 5 minutes on ice before spray. **B)** PHF8 JmjC wild-type and H319Y variant. $(NH_4)_2Fe(SO_4)_2$ (2 µM for wild-type, 40 µM for H319Y variant) was incubated with an enzyme (2 µM) 5 minutes on ice before analysis.

Since Fe(II) readily oxidises under aerobic conditions, ITC assays to investigate metal ion binding were carried out under standard atmospheric condition using Ni(II), which has been used previously as a replacement for Fe(II) in work on JmjC KDMs.[1] The K_D Ni(II), values determined for apo $PHF2_{JmjC}$ and apo $PHF2Y321H_{JmjC}$ were 43.5 μM and 52.0 μM, respectively (Figure 3.21 A), consistent with the reported K_D values of apo PHF2 protein for Ni(II) measured using ITC by Horton *et al*.[1] These results suggest that PHF2 preserves some metal binding affinity of Ni(II), although not as well as apo $PHF8_{JmjC}$ (Figure 3.21 B). The K_D of $PHF8H319Y_{JmjC}$ for Ni(II) was about 2 μM (Figure 3.21 C), i.e. an increase of at least fifty-fold compared to $PHF8_{JmjC}$. Note, the ITC results should be regarded as preliminary and have not as yet been repeated to exclude artefacts such as enzyme batch difference and injection method which could have an influence on the K_D value. It will also be important to carry out studies in parallel on full length proteins. Also, the shape of the binding curve for apo $PHF8_{JmjC}$ was different from the other three enzymes, as it did not appear to manifest a standard binding curve as typically observed for apoenzyme binding.[6] Typically, when an apoenzyme binds a metal ion, the enzyme will become more stable and thus release heat which changes the temperature of the reaction mixture.[5] This change of temperature is detected by the sensor of ITC and forms each peak on the figure. As the binding progresses, the heat released decreases and reaches a plateau. However, in the case of $PHF8_{JmjC}$, the curve first decreased then increased after ten injections, suggesting an initial absorbance of heat and release of heat after the tenth injection. The cause of this abnormal shape of the curve for $PHF8_{JmjC}$ is unclear.

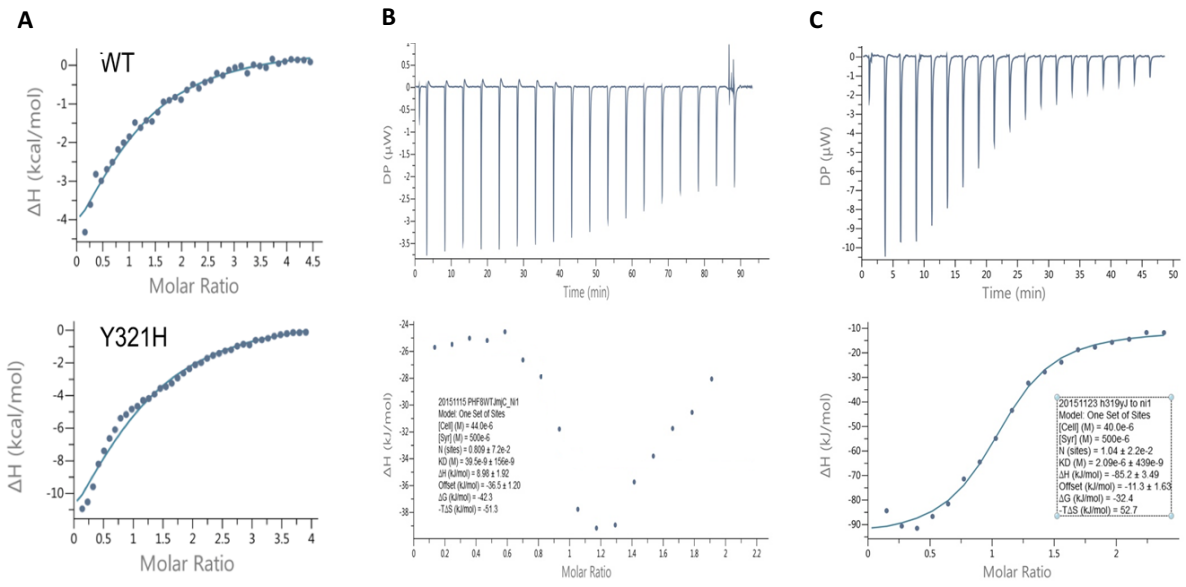


Figure 3.21 ITC titration for nickel binding affinity of **A)** $PHF2_{JmjC}$ (WT) and $PHF2Y321H_{JmjC}$ (Y321H). **B)** $PHF8_{JmjC}$. **C)** $PHF8H319Y_{JmjC}$.

The ITC method is limited by its time-consuming and protein-consuming nature. To achieve efficient screening of metal binding for PHF2 and PHF8, T_m shift assays were carried out. This assay measures the change of denaturation temperature under various conditions. Herein, the metal binding affinity was the primary investigation. Four buffers with pH from 6.0 to 9.0 were tested in the T_m shift assays, with different metal ion including (Cu(II), Mg(II), Co(II), Zn(II) and Mn(II)). Solutions containing metal ion (50 μM) were mixed with the apoenzyme (5 μM) and a heating program (20 °C 30min, heating from 20 °C to 95 °C over 60min, back and stayed at 20 °C) was started immediately. A melting curve for each condition was generated and T_m values were calculated by GraphPad Prism7 software using the Boltzmann Sigmoidal non-linear fitting.

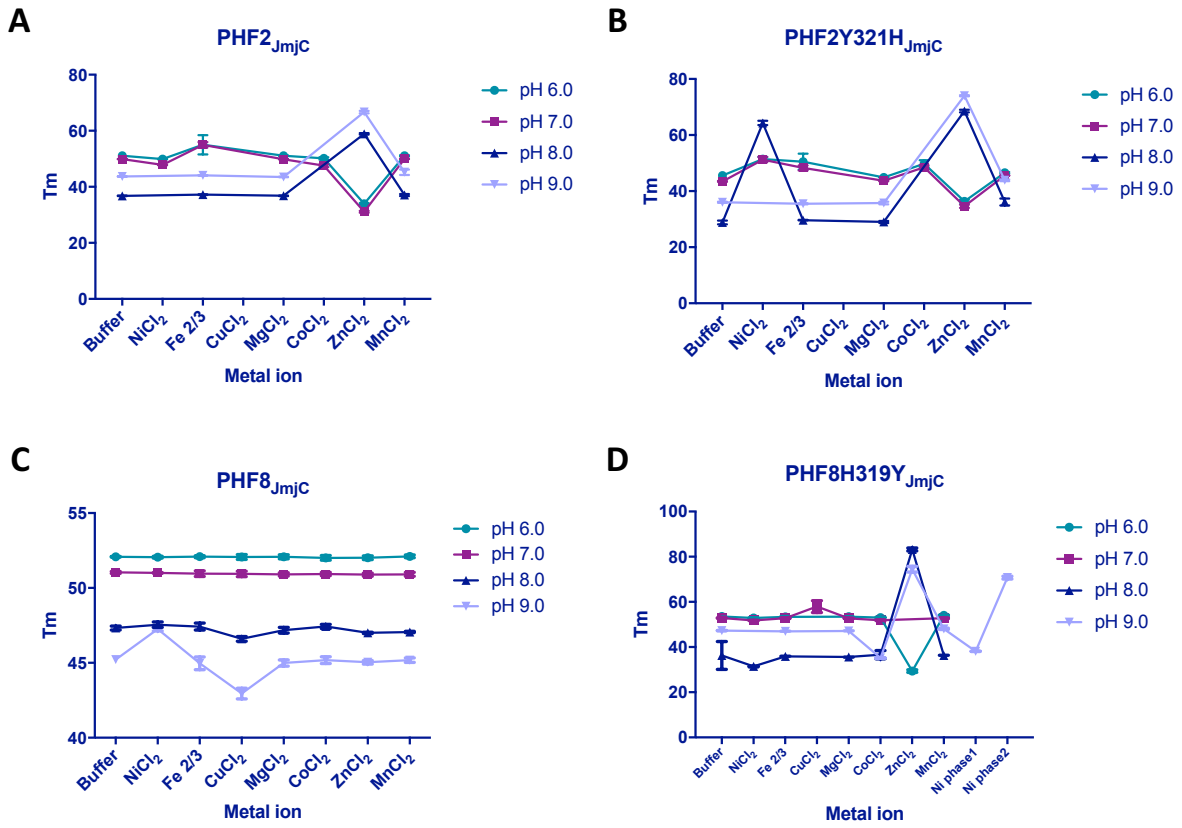


Figure 3.22 T_m shift assays of $PHF2_{JmjC}$, $PHF2Y321H_{JmjC}$, $PHF8_{JmjC}$ and $PHF8H319Y_{JmjC}$.

The T_m values of apo $PHF8_{JmjC}$ decreased from 54 °C to 45 °C as the pH was increased from 6.0 to 9.0, indicating the stability of apo $PHF8_{JmjC}$ decreases with an increase in buffer pH (Figure 3.22 C). $PHF2_{JmjC}$ gave similar T_m values at pH 6.0 and 7.0 (without added metal ions present in the buffer), and was apparently slightly more unstable at pH 9.0 (Figure 3.22 A). This phenomenon of pH affecting stability was also observed with $PHF2Y321H_{JmjC}$ and $PHF8H319Y_{JmjC}$ (Figure 3.22 B & D). Interestingly, while none of the metals substantially increased the stability of apo $PHF8_{JmjC}$, the other apoenzymes tested had different responses to Ni(II) compared to Zn(II), the reason for which is unclear. With Fe(II), the reactions were oxidised to Fe(III), as indicated by a colour change. Future work should thus consider conducting the assays in a glove box or under the protection of nitrogen gas to evaluate the effects of Fe(II) on stability.

3.4 Summary

This Chapter describes mechanistic studies on the KDM7 subfamily, focusing on the characterisation of the binding affinity of PHD finger and JmjC domain. The PHD fingers of all three KDM7 proteins tightly bind to the H3K4me3 with K_D values ranging from nanomolar to single digit micromolar level (Table 3.1). The binding to the H3K4me3 mark by the PHD finger thus appears to dominate substrate binding to PHF8 so regulating its enzymatic activity. By improving binding to the PHD finger (by using a substrate analogue), the apparent V_{max} of demethylation was also increased (Figure 3.12). The results thus imply that targeting the PHD fingers of PHF8 and KIA1718B may be a viable way of modulating their activity, including (if desired) in a therapeutic sense. By implication, this may also be the case for other JmjC KDMs. PHD fingers can recognise and bind to H3K4me3 with high affinity using via interactions involving an aromatic cage,[25] which may be amenable to drug targeting, though as yet no potent PHD inhibitors are reported.

For the investigation of the JmjC domains, FI-methylstat was used as a tool for to mimic substrate binding. FP binding assays employing FI-methylstat were used to investigate the binding activity at the the JmjC domains of PHF8 and KIAA1718 and provided an opportunity for studying the apparently non-catalytic JmjC domain of PHF2. The $PHF2_{JmjC}$ construct used in this Thesis was found to be inactive in the activity assay described in Chapter 2. However, $PHF2_{JmjC}$ showed a strong binding affinity for FI-methylstat, suggesting that other factors than the presence of the typical iron-binding triad can affect the enzymatic activity of KDM7 emzymes. Future work could consider studies on the pH dependence on KDM7s and link it to the adjustment of the activities of these enzymes in cancer and normal cells, which can manifest different pH values. PHF8 is overexpressed in prostate cancer cells and is of

potential medicinal chemistry interest.[26] Studying JmjC enzyme activity as a function of changes in cellular pH is of interest and may open up a new avenue for cancer drug development.

3.4 References

1 J. R. Horton, A. K. Upadhyay, H. Hashimoto, X. Zhang and X. Cheng, *J. Mol. Biol.*, 2011, **406**, 1–8.

2 J. R. Horton, A. K. Upadhyay, H. H. Qi, X. Zhang, Y. Shi and X. Cheng, *Nat. Struct. Mol. Biol.*, 2010, **17**, 38–44.

3 A. K. Dutta, J. Rosgen and K. Rajarathnam, *Methods Mol Biol*, 2015, **1229**, 315–324.

4 G. Pochetti and R. Montanari, *Protoc. Exch.*, 2012, 1–13.

5 C. F. Quinn, M. C. Carpenter, M. L. Croteau and D. E. Wilcox, *Isothermal Titration Calorimetry Measurements of Metal Ions Binding to Proteins*, Elsevier Inc., 1st edn., 2016, vol. 567.

6 Malvern, ITC protocol, https://www.malvernpanalytical.com/en/products/technology/microcalorimetry/isothermal-titration-calorimetry.

7 B. Quantum, *Luminescence and the nature of light*, 2000.

8 N. R. Rose, E. C. Y. Woon, A. Tumber, L. J. Walport, R. Chowdhury, X. S. Li, O. N. F. King, C. Lejeune, S. Ng, T. Krojer, M. C. Chan, A. M. Rydzik, R. J. Hopkinson, K. Hing, M. Daniel, C. Strain-damerell, C. Gileadi, G. Kochan, K. K. Yeoh, P. J. Ratcliffe, N. Burgess-brown, F. Von Delft, S. Muller-knapp, B. Marsden, P. E. Brennan, M. A. Mcdonough, R. J. Klose, C. J. Schofield and A. Kawamura, *J. Med. Chem.*, 2012, **55**, 6639–6643.

9 W. A. Lea and A. Simeonov, *Expert Opin Drug Discov.*, 2012, **6**, 17–32.

10 M. Kecskés, T. S. Kumar, L. Yoo, Z. G. Gao and K. A. Jacobson, *Biochem. Pharmacol.*, 2010, **80**, 506–511.

11 H. Wei, Z. Yan and S. J. Richard, *J. Biomol. Screen.*, 2002, **7**, 215–222.

12 D. M. Jameson, *Burns*, 2011, **110**, 2685–2708.

13 P. Banks, M. Gosselin and L. Prystay, *J. Biomol. Screen.*, 2000, **5**, 239–334.

14 T. Jia, C. Fu, C. Huang, H. Yang and N. Jia, *ACS Appl. Mater. Interfaces*, 2015, **7**, 10013–10021.

15 J. Du and E. R. Strieter, *Anal. Biochem.*, 2018, **550**, 84–89.

16 and D. K. Sang Hyeon Lee, Sujin Ham, Seungsoo Nam, Naoki Aratani, Astuhiro Osuka, Eunji Sim, *J. Phys. Chem.*, , DOI:10.1021/acs.jpcb.8b00213.

17 L. M. Levine, M. L. Michener, M. V. Toth and B. C. Holwerda, *Anal. Biochem.*, 1997, **247**, 83–88.

18 ThermoFisher, Fluorescence Polarization (FP) - Note 1.4, https://www.thermofisher.com/uk/en/home/references/molecular-probes-the-handbook/technical-notes-and-product-highlights/fluorescence-polarization-fp.html.

19 F. Perrin, *J Phys Radium*, 1926, **7**, 390–401.

20 X. Luo, Y. Liu, S. Kubicek, J. Myllyharju, A. Tumber, S. Ng, K. H. Che, J. Podoll, T. D. Heightman, U. Oppermann, S. L. Schreiber and X. Wang, *J. Am. Chem. Soc.*, 2011, **133**, 9451–9456.

21 W. Xu, J. D. Podoll, X. Dong, A. Tumber, U. Oppermann and X. Wang, *J Med Chem*, 2013, **56**, 5198–5202.

22 M. Li, S. Martinez, R. P. Hausinger and J. P. Emerson, *Biochemistry*, 2018, acs.biochem.8b00730.

23 C. C. Thompson and G. M. J. Beaudoin, *Adv. Dev. Biol.*, 2006, **16**, 357–387.

24 D. Landeira and A. G. Fisher, *Trends Cell Biol.*, 2011, **21**, 74–80.

25 K. Fortschegger and R. Shiekhattar, *Epigenetics*, 2011, **6**, 4–8.

26 P. K. Maina, P. Shao, Q. Liu, L. Fazli, S. Tyler, M. Nasir, X. Dong, H. H. Qi, P. K. Maina, P. Shao, Q. Liu, L. Fazli, S. Tyler, M. Nasir, X. Dong and H. H. Qi, *Oncotarget*, 2016, **7**, 75585–75602.

Chapter 4 Inhibition Studies on KDM7 Subfamily

4.1 Chapter overview

The work described in this Chapter concerned the establishment of two inhibitor screening platforms. An AlphaScreen based method and a fluorescence polarisation (FP) based method. These methods not only enable high-throughput screening of KDM inhibitors, but also provide a way of studying catalytically inactive enzymes *in vitro*. Three examples of inhibitor characterisation using AlphaScreen and FP assays are described as demonstration of the utility of the screening platforms.

The cyclic peptides used in this Chapter were synthesised by Mr Oliver D. Coleman. (S,S)-6 KDM2A inhibitor was synthesised by Dr Philip A. Gerken. AMI derivatives were synthesised by Dr Alexandre Erdmann.

4.2 Introduction

4.2.1 Inhibitors targeting the KDM7 and KDM2 subfamilies

Many histone lysine demethylases (KDMs) are involved in epigenetic regulation, cellular development and other diseases, and thus become potent targets for drug discovery.[1] However, although there are variations in detail, the available evidence is that all JmjC KDMs share the same general catalytic mechanism and 2OG-binding mode, making it difficult to design selective inhibitors of a particular enzyme. Inhibitors of the KDM7 subfamily with some degree of selectivity have been reported (Figure 4.1).[2–10] However, since the KDM7 subfamily is closely related to the KDM2 subfamily, it has been difficult to achieve selectivity between the two subfamilies.[3–5,11] Peptide mimicking inhibitors and inhibitors that bind to KDMs allosterically have also been reported.[12,13] However, these inhibitors are often limited by their cellular permeability. Discovering novel highly selective and cell permeable inhibitors is important for biological investigations and chemical probes development concerning the KDM7 subfamily.

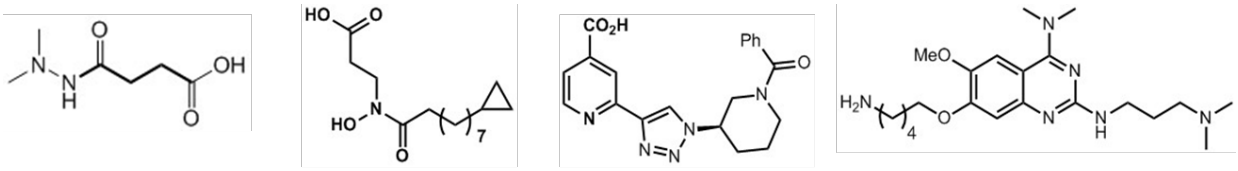


Figure 4.1 Structures of the reported KDM2/7 inhibitors.[3–6]

4.2.2 Principles of the AlphaScreen assay

MALDI-TOF MS-based assay and formaldehyde hydrogenase (FDH) coupled assays can both be used to determine inhibitor IC_{50} values for KDM7s.[14] However, both assays require relatively high concentrations of enzyme (typically micromolar concentrations of enzyme for each reaction are required). Further, the MALDI-TOF MS-based assay is laborious, preventing it use I high-throughput screening applications. In addition, the formaldehyde dehydrogenase (FDH) enzyme used for FDH coupled assay may be inhibited by the tested inhibitors, leading to false positive results. Therefore, in order to carry out screening of large compound libraries, an AlphaScreen assay was developed to achieve high-throughput screening.

The amplified luminescent proximity homogeneous assay (ALPHA) based technology uses a pair of beads (donor bead and acceptor bead) which can emit light when brought into close proximity.[15] Laser excitation of the donor bead at 680 nm causes a photosensitiser to convert triplet oxygen to the excited singlet state $^1O_2$. The acceptor bead, when sufficiently close to the donor bead, interacts with $^1O_2$ and emits a luminescence between 520 and 620 nm, transferring the excited singlet oxygen back to its ground state.[16] For the screening assays (called as AlphaScreen in this Thesis) with KDM7s, the donor bead is coated with streptavidin, which binds to the biotin group at the *C*-terminus of biotinylated histone peptides.[17] The acceptor bead varies depending on the type of AlphaScreen assay being carried out.

For application of an AlphaScreen based activity assay involving detection of the demethylated product, the acceptor bead is coated with protein-A, which binds to an

antibody of choice that recognises the product peptide (Figure 4.2). When the antibody binds selectively to the 'chemical mark' on the product peptide and the streptavidin binds to the biotin group conjugated onto the same peptide, the donor bead and acceptor bead are brought into close proximity, enabling a luminescent signal to be observed upon donor bead excitation (Figure 4.2). The assay for KDM7 employs an antibody recognising the H3K9me1 mark on a histone peptide, the concentration of which increases as the demethylation reaction proceeds. The resultant change in luminescence emitted can be used to measure the demethylation of the substrate peptide.

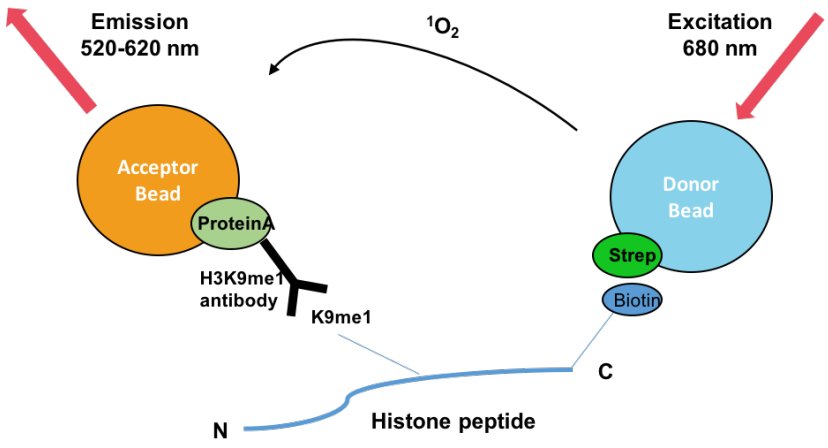


Figure 4.2 Setup of an AlphaScreen peptide turnover assay for JmjC-KDMs. The interaction of the antibody with biotinylated peptide can be detected using an AlphaScreen IgG detection kit (PerkinElmer).

The AlphaScreen binding assay can enable the detection of protein-protein (or protein-peptide) interactions, in this case, the acceptor bead is a nickel chelating bead which binds to the poly-His-tagged protein of interest.[17] When the poly-His-tagged protein binds to the biotinylated peptide, which binds to streptavidin coated acceptor beads, the donor and acceptor beads are brought into close proximity, giving a luminescent signal upon donor bead excitation (Figure 4.3). As the concentration of the poly-His-tagged protein or the

biotinylated peptide increases, the fraction of peptide bound to the protein changes. The resultant change in luminescence emitted can be used to measure the binding affinity of a protein to a peptide.

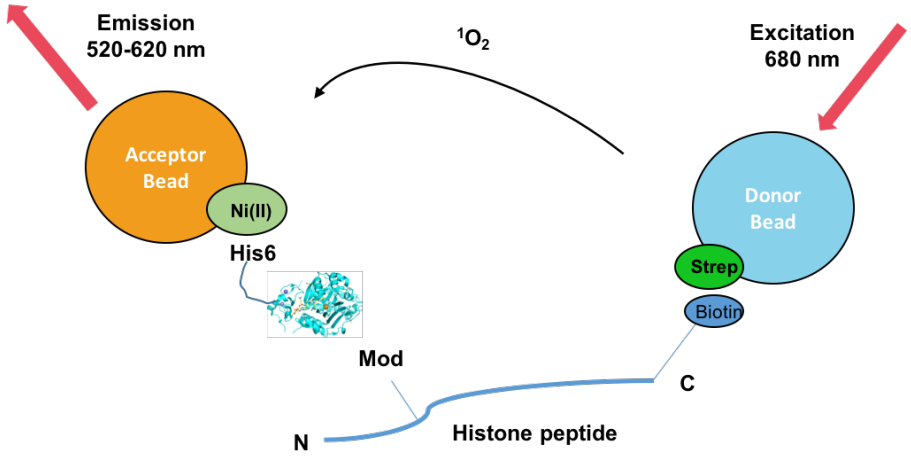


Figure 4.3 Setup of an AlphaScreen binding assay.

4.3 Results and discussions

4.3.1 Establishing AlphaScreen assay for inhibitor screening

AlphaScreen activity assay for $PHF8_{PHD\text{-}JmjC}$

As mentioned in the introduction to this chapter, AlphaScreen assays require biotinylated peptides. Initially, a MALDI-TOF MS activity assay was conducted (by Dr Louise Walport) to investigate whether *N*- or *C*- terminally biotinylated peptide was most suitable for the desired assay.[18] Only the *C*-terminally biotinylated peptide was observed to be demethylated by $PHF8_{PHD\text{-}JmjC}$. This is unsurprising given that PHF8 is known to interact with the *N*-terminal region of H3.[14] Thus, a $H3_{1\text{-}21}$K4me3K9me2 GGK fusion peptide biotinylated at the *C*-terminal lysine side-chain (described as $H3_{1\text{-}21}$K4me3K9me2-biotin) was selected as the substrate peptide of $PHF8_{PHD\text{-}JmjC}$ for use in AlphaScreen assays (GLS Biochem > 95% pure). The *C*-terminal GGK residues were attached as a linker to the biotin group.

It was envisaged that AlphaScreen assays could be set up in two ways, i.e. signal-gain and signal-loss, using an antibody that either recognises the enzymatic product H3K9me1, or one that recognises the substrate H3K9me2, respectively. Antibodies against both peptides are commercially available from Abcam. Since PHF8 and KIAA1718 catalysed the removal of the methyl groups at H3K9 not H3K4 (Chapter 2, section 2.3.4), the anti-H3K9 antibodies were proposed to be suitable for detecting the demethylation of both H3K9me2 and H3K4me3K9me2 as catalysed by the KDM7s. The anti-H3K9me2 antibody (ab1220, Rabbit, Abcam) had been tested by Dr. Akane Kawamura and found to be effective at a 1:2000 dilution.[17] The anti-H3K9me1 antibody (ab8896, Rabbit, Abcam) had been tested and found to be effective at a 1:2000 dilution by Dr. Louise Walport.[18] The comparison of the two set-ups of the assay showed that the signal-gain assay generated an AlphaScreen signal that

was more sensitive to monitor the progression of the demethylation reaction than the signal-loss assay (Figure 4.4).[18] Thus, it was decided to pursue the signal-gain assay for the KDM7 enzymes.

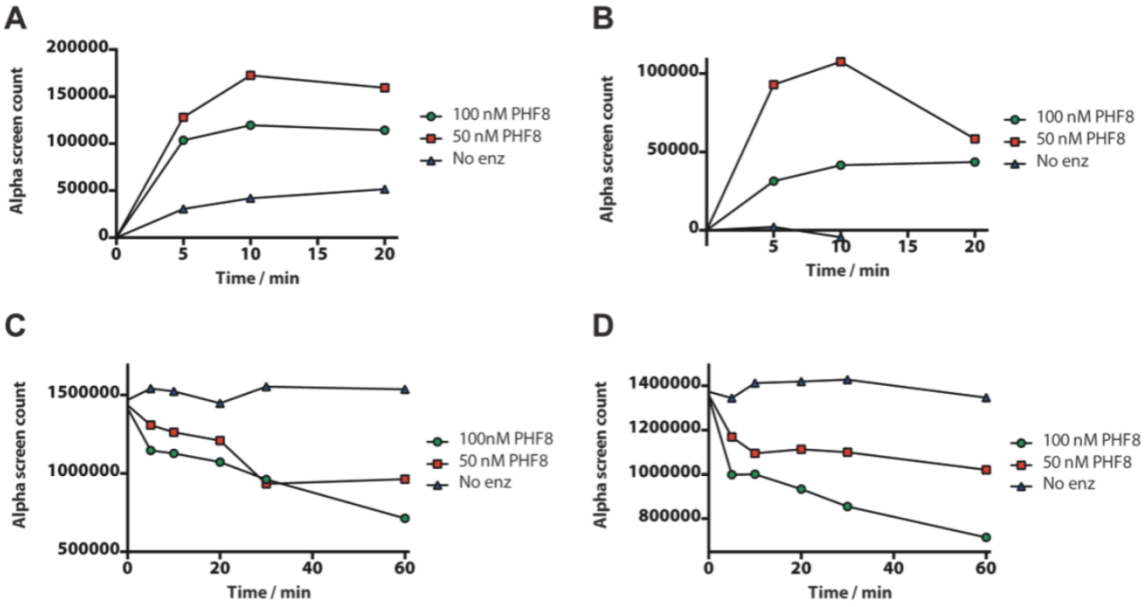


Figure 4.4 AlphaScreen assays comparing signal-gain and signal-loss set-ups. Signal-gain assays (ab8896) with **A)** $H3_{1\text{-}21}$K4me3K9me2-biotin (300 nM) or **B)** $H3_{1\text{-}21}$K4me3K9me2-biotin (100 nM). Signal-loss assays (ab1220) with **C)** $H3_{1\text{-}21}$K4me3K9me2-biotin (25 nM) or **D)** $H3_{1\text{-}21}$K4me3K9me2-biotin (25 nM) and $H3_{1\text{-}21}$K4me3K9me2 peptide (75 nM). Figure reproduced with permission from Walport, DPhil Thesis, 2013.[18]

Several elements of the AlphaScreen activity assay were then investigated for $PHF8_{PHD\text{-}JmjC}$. Despite the well-established assays by Dr Kawamura and Dr Walport, the batches of enzyme, antibody and detection beads had all changed from their work, which requires a re-development of the signal-gain assays for KDM7s. Several conditions for the AlphaScreen activity assay with $PHF8_{PHD\text{-}JmjC}$ were examined, included the linear range of the AlphaScreen detection beads (i.e. the luminescence signal generated in relation to the product concentration), antibody dilution, peptide concentration, enzyme concentration, the kinetic linear range of demethylation reaction, and bead incubation time. As high salt concentration had previously been shown to suppress the enzyme activity of KDM7s

(Chapter 2, section 2.3.1), a 50 mM HEPES buffer (pH 7.0) with no NaCl was used as the AlphaScreen assay buffer. To help maintain enzyme stability and to reduce non-specific binding to plasticware, 0.01% Tween-20 and 0.1% (w/v) BSA were added to the buffer immediately before use.

The linear range of the AlphaScreen detection beads was then investigated. For activity assays, an IgG Detection Kit containing Streptavidin Donor beads and Anti-rabbit IgG conjugated Acceptor beads was used (PerkinElmer). As described in Section 4.2, the signal increases as the demethylated product peptide is generated. However, the increase in the AlphaScreen signal counts does not always increase linearly with the accumulation of the demethylation product, but follows a trend shown in Figure 4.5.[19] At low product peptide concentrations, increases in peptide initially give rise to increased AlphaScreen signal counts. At high peptide concentrations, however, further increases in peptide concentrations leads to a decrease in the signal due to the full occupation of the binding sites of antibody or beads, which is known as the “Hook effect” or “Prozone effect”.[19] Therefore, it is necessary to determine the linear part of the curve where the signal increases proportionally with peptide concentration.

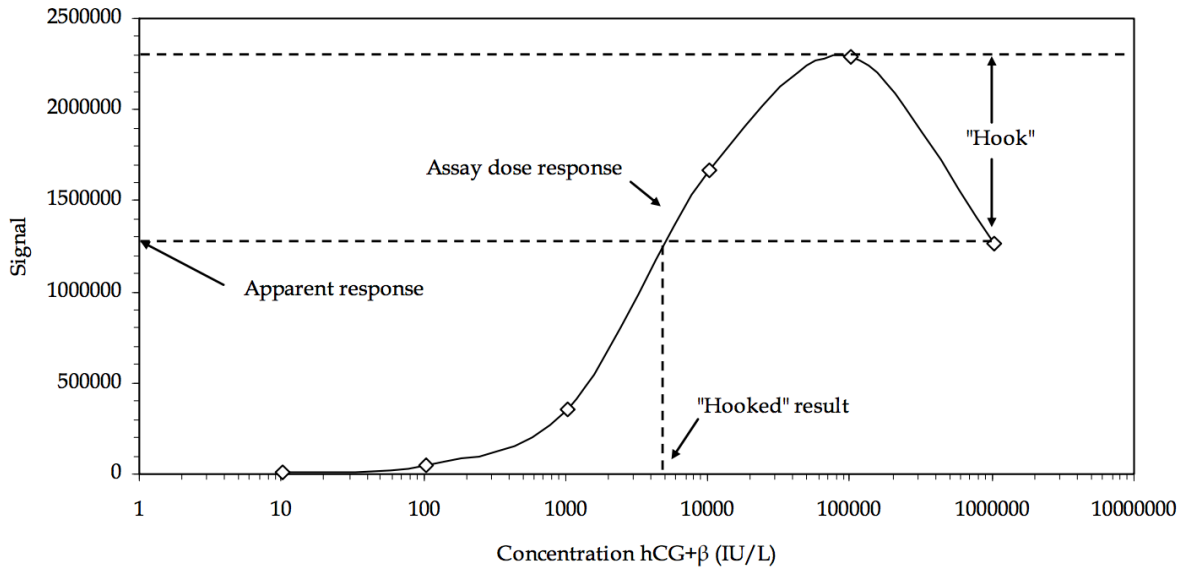


Figure 4.5 Schematic of assay dose response curve and the Hook effect.[19]

In the assay with $PHF8_{PHD\text{-}JmjC}$, H3K9me1 is generated as the demethylation product and therefore an appropriate antibody was required to detect this monomethyl-lysine residue. However, since the optimal concentrations of enzyme, peptide and antibody had not been determined at this stage, a linear range assay investigation was carried out using the biotin conjugated IgG control (Bn-IgG) provided by the manufacturer (PerkinElmer). The Bn-IgG control solution was subjected to a 2-fold dilution and mixed with the Donor/Acceptor beads (the concentrations used were according to the manufacturer's instructions). The mixture was then allowed to equilibrate for 1 hour before being read by an EnVision plate reader (PerkinElmer). The linear range was found to be from 1.6 nM to 6.4 nM of the Bn-IgG control, with AlphaScreen count from 100,000 to 1,000,000 (Figure 4.6), suggesting that the readout of activity assays should not exceed 1,000,000.

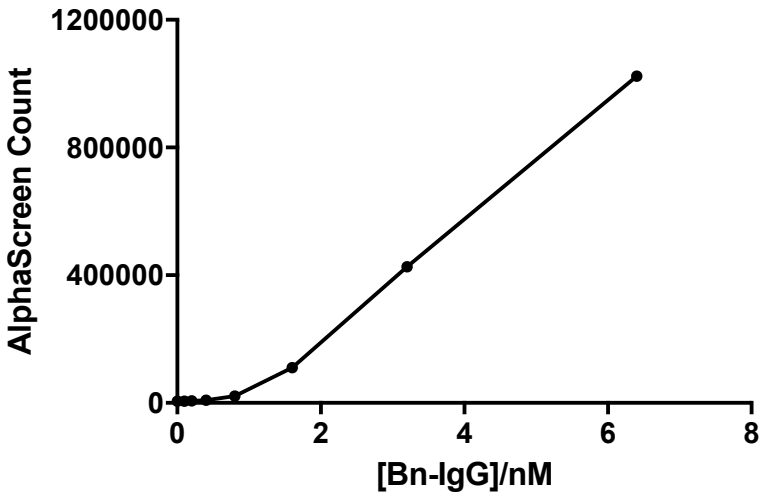


Figure 4.6 Evaluation of the linear range of IgG beads. Data are plotted as the mean of technical duplicates, and error bars give ± one s.d.

Next, the use of the antibody was optimised. The dilution of the H3K9me1 antibody (ab8896) was 1:2000 as tested by Dr. Louise Walport. However, since the batch of antibody had changed, it was necessary to verify the amount used again. The new batch of antibody ab8896 (stock concentration 1 mg/mL) was compared with the old batch (stock concentration 0.4 mg/mL) at the same concentration. The previous concentration used was 0.2 µg/mL (0.4 mg/mL divided by 2000) which in equal to a 1:5000 dilution of the new batch. Both batches of antibody were tested against a titration of $H3_{1\text{-}21}$K9me1-biotin and $H3_{1\text{-}21}$K4me3K9me2-biotin to investigate the selectivity of antibody binding at various peptide concentrations. The results revealed that the old batch selectively bound to H3K9me1-biotin at a 1:2000 dilution, whereas the new batch gave high counts for both peptides (Figure 4.7 A), indicating that the concentration of the new batch was too high to gain selectivity. The high signal count could potentially obscure changes in signal generated by slight changes in demethylation. The absolute count of signal was also out of the linear range of the beads (Figure 4.6), meaning that the signal count may not be proportional with the peptide

concentration. Thus, the calculation of demethylation using a high signal count would not be accurate.

To bring the signal obtained with the new ab8896 antibody batch into the linear range of the assay and to gain more selectivity, two further concentrations of the new batch of ab8896 were tested (i.e. 1:20000 and 1:80000 dilutions). At a 1:20000 dilution, the antibody was found to be selective for H3K9me1-biotin and the signal differences between H3K9me1-biotin and H3K4me3K9me2-biotin were within the beads' linear range (Figure 4.7 B). From the curve of 1:20000 dilution of ab8896, 300 nM of peptide was selected as an optimal concentration for the use in the AlphaScreen assay, in order achieve a to balance between antibody selectivity and maximising signal count.

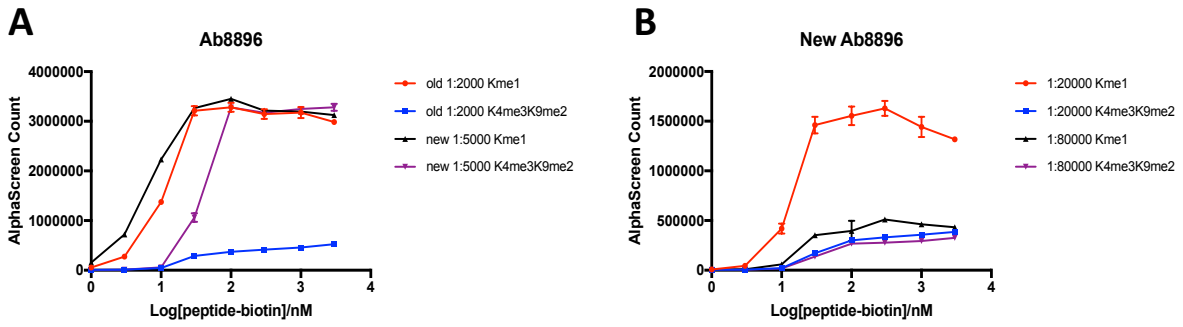


Figure 4.7 Evaluation of the selectivity of antibody ab8896 (Abcam) for H3K9me1 over H3K4me3K9me2 at different concentrations by AlphaScreen. **A)** Comparison of new and old batch of ab8896 at 0.2 µg/mL. **B)** Two dilutions of the new batch ab8896. Figures were generated using GraphPad Prism7 software. A sample without biotinylated peptide was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

To ensure the binding of Donor/Acceptor beads to biotinylated peptide and antibody reached equilibrium before analysing the plate, the AlphaScreen signal was monitored after various incubation times. The AlphaScreen count remained steady after 45 to 50 minutes' incubation at room temperature (Figure 4.8). It was thereby decided that 1 hour was a suitable incubation time for the assay.

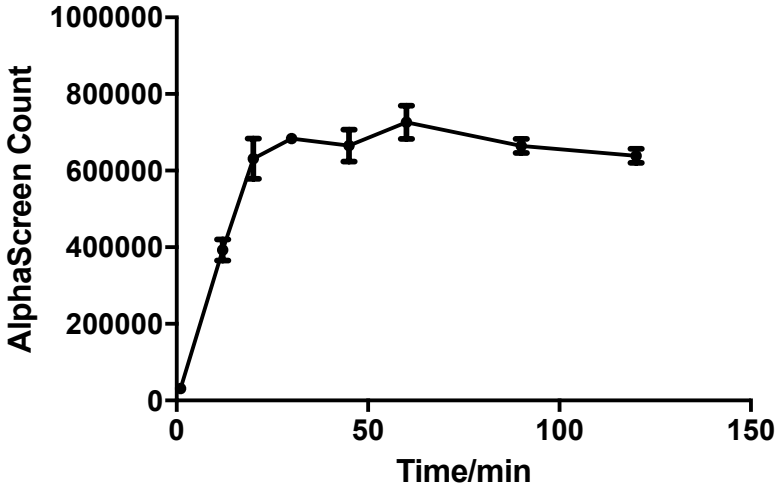


Figure 4.8 Evaluation of AlphaScreen incubation time with Donor and Acceptor beads. Figures were generated using GraphPad Prism7 software. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

The final factors investigated were the concentration of $PHF8_{PHD\text{-}JmjC}$ and the kinetic linear range of the demethylation reaction catalysed by it. Time courses of the demethylation reaction catalysed by various concentrations of $PHF8_{PHD\text{-}JmjC}$ (from 50 pM to 15 nM) were analysed (Figure 4.9 A). To gain a reasonable signal count, enzyme concentrations lower than 2 nM were initially excluded.

For inhibitor IC_{50} evaluations, the incubation time was also considered based on the linear range of the reaction. In the KDM7 demethylation reaction, the monomethyl-lysine at position 9 is an intermediate which accumulates at the early stage of demethylatio; this undergoes further demethylation to give unmodified lysine as the reaction progresses. The antibody ab8896 selectively recognises the monomethyl-lysine at position 9 and thus the AlphaScreen signal represents the amount of K9me1. To ensure the amount of K9me1 accurately reflected the demethylation of K9me2, the quenching point should be within the linear range of the demethylation reaction. From the standpoint of the assay operation, it was also required that the linear range of the reaction was not so long such that PHF8

degrades during the incubation, or too short such that quenching the reaction in sufficient time becomes difficult. Taking these factors into consideration, use of 10 nM of PHF8 with a quenching point at 10 min was selected to be the most suitable assay condition for inhibitor screening (Figure 4.9 B).

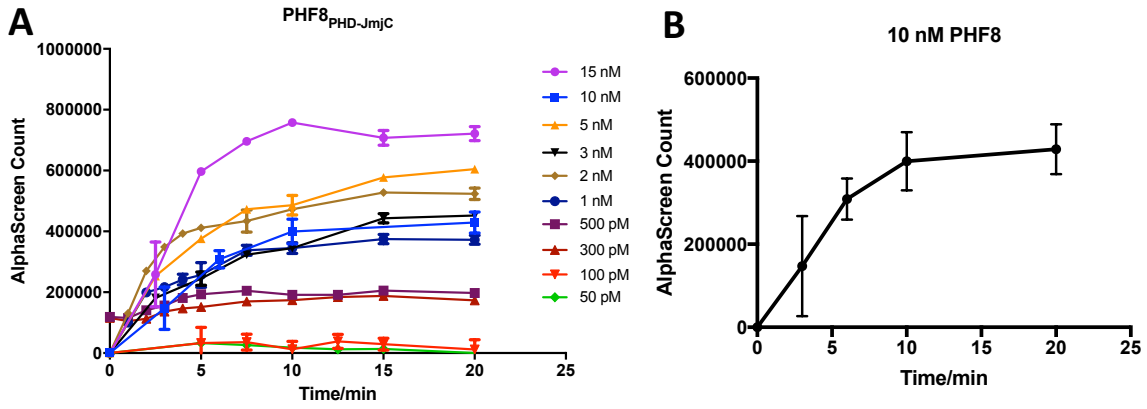


Figure 4.9 A) AlphaScreen assays of $PHF8_{PHD\text{-}JmjC}$ (various concentrations) on H3K4me3K9me2-biotin peptide (300 nM). **B)** The selected assay condition for the assay: 10 nM of PHF8. Figures were generated using GraphPad Prism7 software. A sample of pre-quenched reaction was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

AlphaScreen activity assay for $KIAA1718_{PHD\text{-}JmjC}$

Several optimised conditions from that of $PHF8_{PHD\text{-}JmjC}$ were applied to the AlphaScreen activity assay for $KIAA1718_{PHD\text{-}JmjC}$, including the use of detection beads and antibody ab8896. As for the biotinylated peptide, $H3_{1\text{-}21}K9me2$-biotin was used for the observation of best $KIAA1718_{PHD\text{-}JmjC}$ activity (Chapter 2, section 2.3.4). The concentration of $H3_{1\text{-}21}K9me2$-biotin was optimised to be 30 nM based on the titration assay with ab8896 (Figure 4.10).

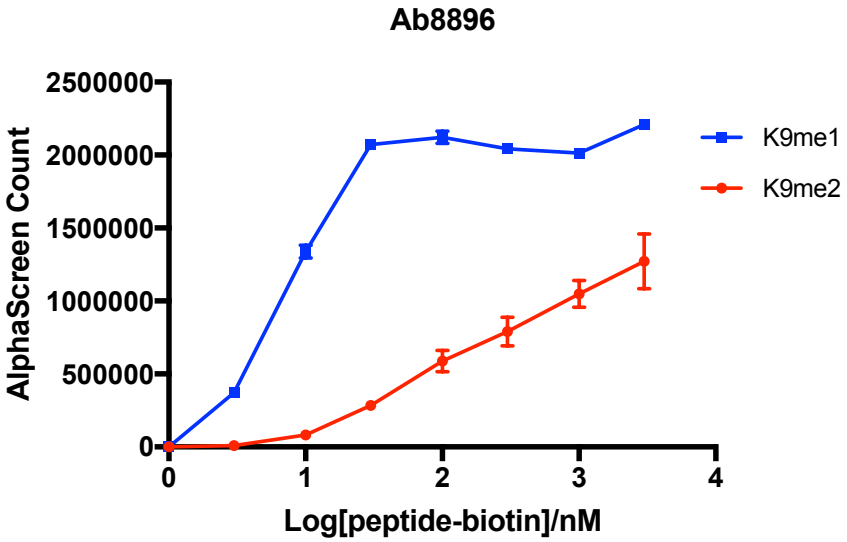


Figure 4.10 Evaluation of the selectivity of Ab8896 of H3K9me1 over H3K9me2. Figures were generated using GraphPad Prism7 software what version. A sample without biotinylated peptide was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

Four concentrations of KIAA1718$_{PHD\text{-}JmjC}$ were reacted with 30 nM of H3$_{1\text{-}21}$K9me2-biotin and time courses were recorded (Figure 4.11). Based on the linear range of the demethylation assay and the signal strength (applying the sample principle described for PHF8), 10 nM of KIAA1718 with quenching point at 10 min was selected to be the suitable assay conditions for KIAA1718 inhibitor screening.

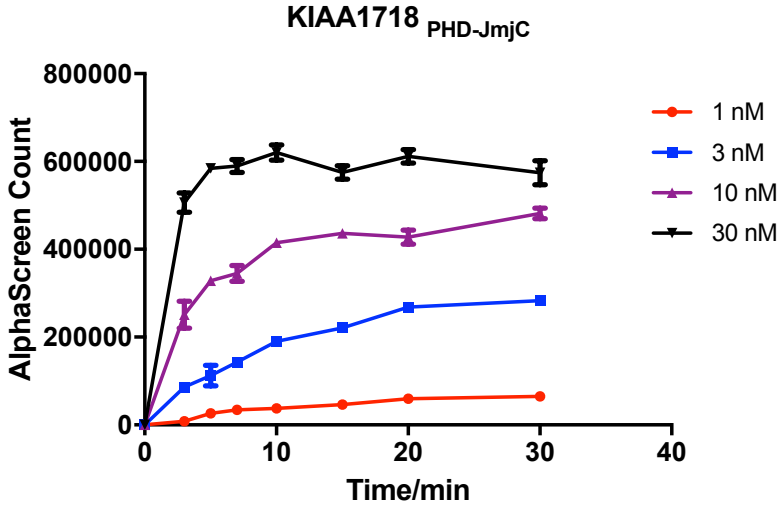


Figure 4.11 AlphaScreen assays of KIAA1718$_{PHD\text{-}JmjC}$ (various concentrations) on H3$_{1\text{-}21}$K9me2-biotin peptide (30 nM). The selected assay condition for the assay: 10 nM of KIAA1718. Figures were generated using GraphPad Prism7 software. A sample of pre-quenched reaction was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

AlphaScreen inhibition assay for PHF8 and KIAA1718

To evaluate whether the optimised AlphaScreen activity assay conditions for PHF8 and KIAA1718 described above were useful for inhibitor screening, pilot IC_{50} assays with three known KDM inhibitors were carried out on $PHF8_{PHD\text{-}JmjC}$ and $KIAA1718_{PHD\text{-}JmjC}$. The compounds used were known 2OG competive inhibitors, Daminozide, IOX1 and 2,4-pyridinedicarboxylic acid (2,4-PDCA) (Figure 4.12).[2,6]

Commonly known as Alar, Daminozide is a plant growth retardant that is used to control the growth of ornamental plants.[20,21] Daminozide is found to inhibit the plant 2OG oxygenases and a screen of human KDMs showed that it specifically inhibits the KDM2/7 subfamilies.[6,22,23] 2,4-PDCA and IOX1 are well known broad-spectrum inhibitors of 2OG oxygenases.[2,24,25]

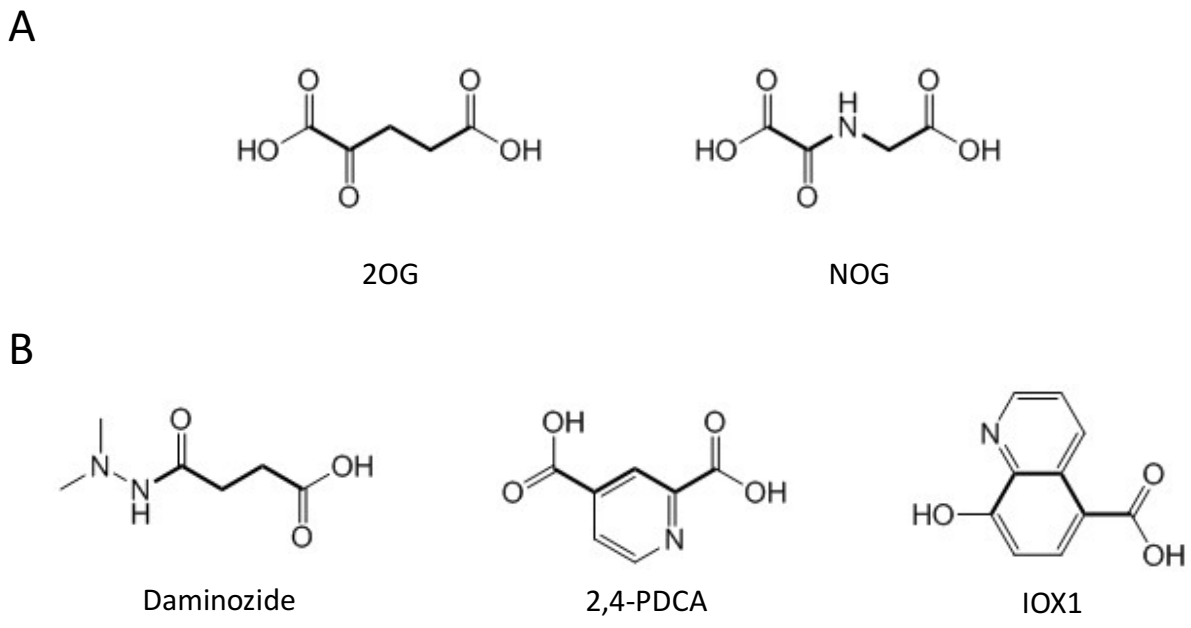


Figure 4.12 Structure of 2-Oxoglutarate (2OG) and its analogues reported as JmjC histone lysine demethylase (KDM) inhibitors. **A)** 2OG and *N*-Oxalylglycine (NOG). **B)** 2,4-PDCA, IOX1 and Daminozide.[2,6,24]

The IC_{50} assays were carried out with the optimised assay conditions for $PHF8_{PHD\text{-}JmjC}$ and $KIAA1718_{PHD\text{-}JmjC}$ separately. Thus, $PHF8_{PHD\text{-}JmjC}$ (10 nM) or $KIAA1718_{PHD\text{-}JmjC}$ (10 nM) was pre-incubated with increasing concentrations of the inhibitors (30 nM to 100 μM) for 15 min before the addition of peptide and cofactors. The reaction was quenched after 10 min, at which point the mixture of antibody and detection beads were added. The plate was measured after 1 hour using an Envision plate reader (Perkin Elmer). The extent of inhibition was determined relative to a control reaction containing no inhibitor. By plotting this value against the log of the inhibitor concentration, an IC_{50} value was determined. Daminozide was found to inhibit PHF8 with an IC_{50} of 0.47 μM (Figure 4.13 A), which is similar to the reported IC_{50} value of 0.55 μM.[6] The IC_{50} of Daminozide for KIAA1718 was 0.15 μM (Figure 4.13 B), which was 15-fold lower than the IC_{50} value measured by MALDI-TOF MS (2.2 μM tested by Dr. Louise Walport). The difference is possibly due to the difference in the type of assay used and differences in the enzyme concentrations employed (200 fold more in MALDI relative to AlphaScreen). IOX1 and 2,4-PDCA did not show obvious inhibition of PHF8 or KIAA1718 under the conditions tested, consistent with the results using FP assay (Chapter 4, section 4.3.2). These results demonstrated that the setup of AlphaScreen assay is appropriate for inhibitor screening.

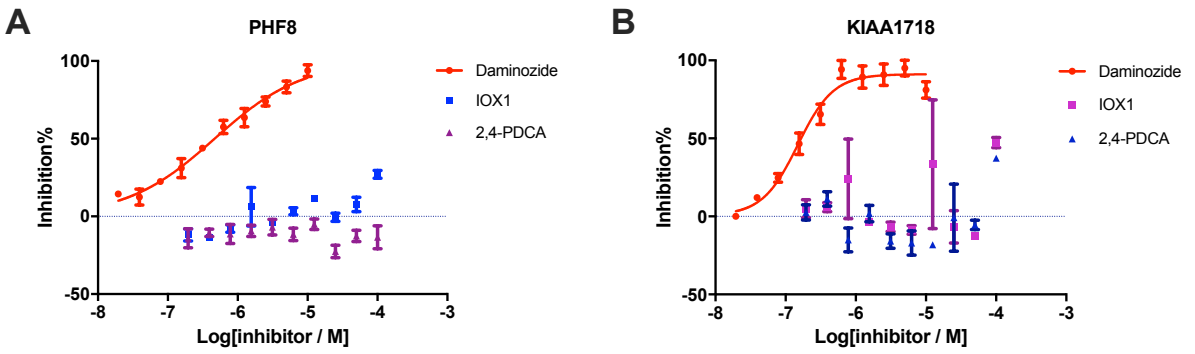


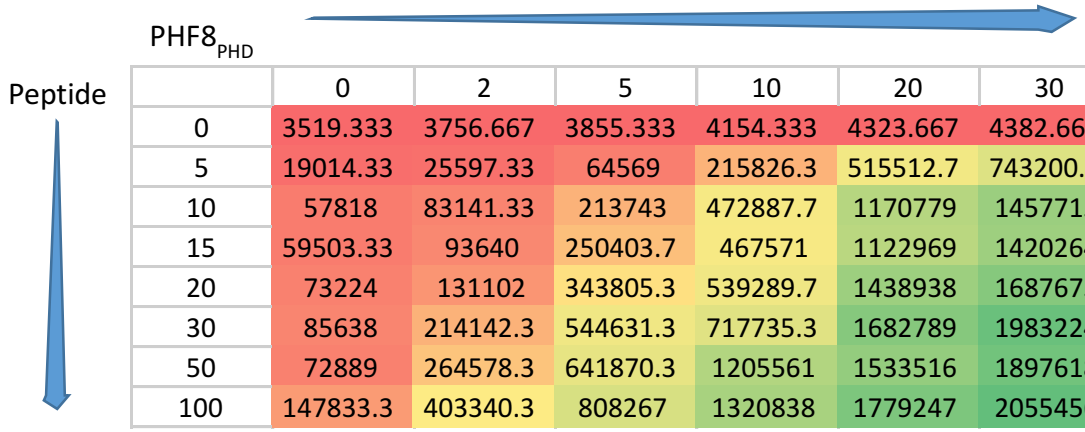
Figure 4.13 IC_{50} determination for Daminozide, IOX1 and 2,4-PDCA against **A)** $PHF8_{PHD\text{-}JmjC}$ (10 nM) and **B)** $KIAA1718_{PHD\text{-}JmjC}$ (10 nM) measured by AlphaScreen. Data were fitted to a log(inhibitor) vs. response – Variable slope (four parameters) model using GraphPad Prism7 software. R square values indicating the goodness of fitting for Daminozide curves were 0.97 (Figure 4.13 A) and 0.95 (Figure 4.13 B). A sample without enzyme was used as a negative control for normalisation. A sample without inhibitor was used as a 'no inhibition' control. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

AlphaScreen binding assay for $PHF8_{PHD}$

Unlike 2OG competitors, some JmjC-KDM inhibitors are designed to interrupt the binding of PHD fingers to their substrates. As the effect of this disruption on catalytic activity is uncertain, activity assays may not be suitable to evaluate these inhibitors. To test the inhibitors that targeted the PHD finger of PHF8, an AlphaScreen binding assay was set up to perform PHD-inhibitor screening which was also amenable for a high-throughput manner. The His-tagged $PHF8_{PHD}$ was pre-incubated with the tested inhibitors (if present) and then mixed with the $H3_{1\text{-}21}$K4me3K9me2-biotin peptide and the detection beads to equilibrium. No antibody was required as there were no demethylation products.

Three factors were optimised for the setup of this inhibitor screening platform, the concentration of $H3_{1\text{-}21}$K4me3K9me2-biotin, the concentration of $PHF8_{PHD}$ and the incubation time with the detection beads. For the first two concentrations, a heat map of binding was generated using six concentrations of protein (0 to 30 nM) and eight concentrations of peptide (0 to 0.1 µM). The average signal count of the technical triplicates

for each concentration combination manifested a map where the signal gradually increased from the top-left to the bottom-right (Figure 4.14). 10 nM of $H3_{1-21}K4me3K9me2$-biotin and 10 nM of $PHF8_{PHD}$ were determined to be suitable concentrations for the binding assay as these were within the linear range of the binding signal.



Peptide \ $PHF8_{PHD}$	0	2	5	10	20	30
0	3519.333	3756.667	3855.333	4154.333	4323.667	4382.667
5	19014.33	25597.33	64569	215826.3	515512.7	743200.3
10	57818	83141.33	213743	472887.7	1170779	1457715
15	59503.33	93640	250403.7	467571	1122969	1420264
20	73224	131102	343805.3	539289.7	1438938	1687673
30	85638	214142.3	544631.3	717735.3	1682789	1983224
50	72889	264578.3	641870.3	1205561	1533516	1897618
100	147833.3	403340.3	808267	1320838	1779247	2055456

Figure 4.14 AlphaScreen heat map of $PHF8_{PHD}$ (0, 2, 5, 10, 20 and 30 nM) binding assay with $H3_{1-21}K4me3K9me2$-biotin (0, 5, 10, 15, 20, 30, 50, 100 nM). Intensity counts are averaged from technical triplicates.

For the acceptor beads, a histidine (nickel chelate) detection kit (PerkinElmer) containing Streptavidin Donor beads and nickel chelating (Ni-NTA) Acceptor beads was used. To determine the influence of bead incubation time on the IC_{50} values of the inhibitors, a non-biotinylated $H3_{1-15}K4me3K9me2$ peptide was used to compete with the biotinylated peptide with incubation time periods of 1, 2 and 3 hours. The IC_{50} values obtained for the three conditions were 22.6 µM, 28.4 µM and 30.7 µM respectively (Figure 4.15). The inhibition curves of 1-hour and 2-hour incubation overlapped with each other, while the curve of 3-hour incubation shifted and contained higher error bars at lower concentrations. Therefore, for the binding assay and inhibitor screening assay, the reactions were incubated with the detection beads for 1 hour before analysis.

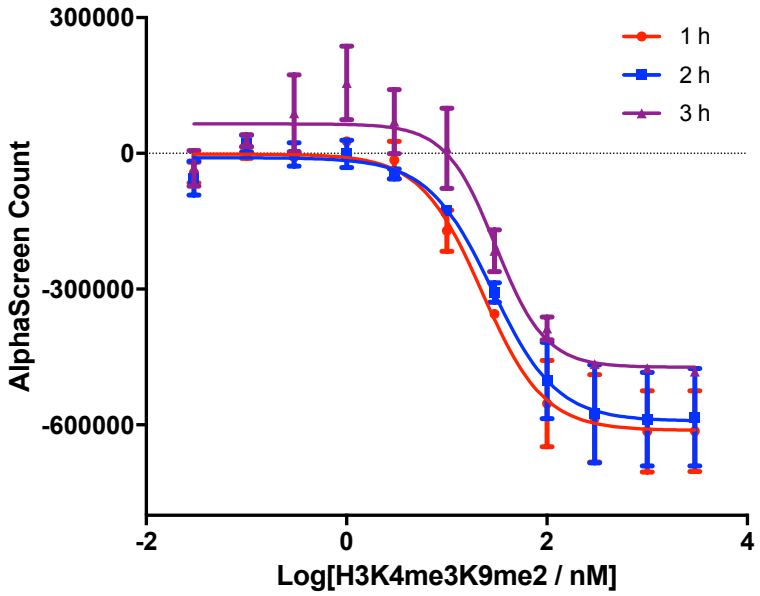


Figure 4.15 Evaluation of the incubation time with detection beads. Red: 1 hour. Blue: 2 hours. Purple: 3 hours. Data were fitted to a log(inhibitor) vs. response – Variable slope (four parameters) model using GraphPad Prism7 software. R square values indicating the goodness of fitting were 0.95, 0.93 and 0.92 for red, blue and purple curves respectively. A sample without enzyme was used as a negative control for normalisation. A sample without inhibitor was used as a 'no inhibition' control. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

4.3.2 Establishing FP displacement assay for KDM7 inhibitor screening

An FP displacement assay was developed for inhibitor screening, based on the FP binding assay of KDM7s that was established and described in the previous Chapter (Chapter 3, section 3.2.4 and 3.2.5). JmjC-domain inhibitors, like 2OG mimics and iron-chelators, were tested for displacement activity against the binding of KDM7 enzymes to FI-methylstat (structure in Chapter 3, section 3.1.4).[26] PHD finger inhibitors were tested against the binding of KDM7 enzymes to H3K4me3K9me2-FITC. For both types of FP displacement assays, the KDM7 enzymes were pre-incubated with a concentration range of the tested inhibitors (from 3 nM to 10 mM) for 15 min before the addition of FI-methylstat or H3K4me3K9me2-FITC. The reaction was incubated in dark to equilibrium (typically 20 min) before analysis. The concentration of each KDM7 enzymes to be used was the same as the K_D values generated from the binding assay (Chapter 3, Table 3.1 & Table 3.3).

The assay conditions for the JmjC-domain inhibitors was first evaluated, using 2OG and several known 2OG mimicking inhibitors (i.e. the compounds in Figure 4.12). In the FP displacement assays with $PHF8_{PHD\text{-}JmjC}$, Daminozide gave an IC_{50} value of 0.49 μM, consistent with the reported data.[6] Interestingly, despite structurally being similar to 2OG (IC_{50} = 55.6 μM, Figure 4.16), NOG only weakly displaced FI-methylstat at high concentrations (Figure 4.16). The commonly known broad-spectrum inhibitors IOX1 and 2,4-PDCA did not show any displacement effect, consistent with the AlphaScreen results (Figure 4.13).

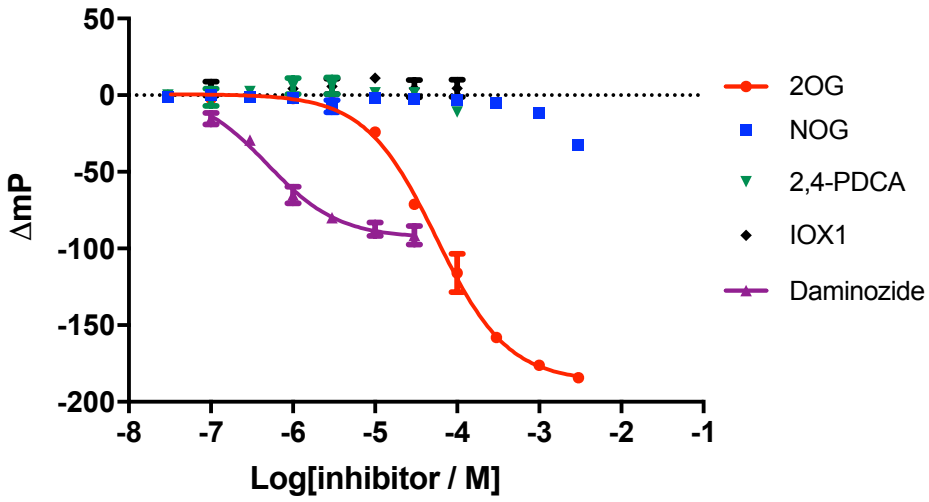


Figure 4.16 $PHF8_{PHD\text{-}JmjC}$ FP displacement assay measuring the IC_{50} values of 2OG, NOG, 2,4-PDCA, IOX1 and Daminozide. Data were fitted to a log(inhibitor) vs. response – Variable slope (four parameters) model using GraphPad Prism7 software. The fitting was fixed at 0 ΔmP at the top. R square values indicating the goodness of fitting were 0.99 and 0.95 for red and purple curves respectively. A sample without enzyme was used as a negative control for normalisation. A sample without inhibitor was used as a 'no inhibition' control. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

The same assays were carried out for $KIAA1718_{PHD\text{-}JmjC}$. Similarly to PHF8, inhibitors other than 2OG and Daminozide were not observed to compete with FI-methylstat efficiently (Figure 4.17). Daminozide inhibited KIAA1718 with an IC_{50} value less than 50 nM (the exact value could not be measured due to an incomplete dose-response curve). The IC_{50} value for 2OG against KIAA1718 was 1.39 μM, 40-fold lower than that of PHF8. This could indicate that KIAA1718 has a higher affinity for the cosubstrate 2OG compared to PHF8. However,

the IC_{50} value represents the concentration of inhibitor when half of the enzyme had been inhibited. Due to the different concentration used for PHF8 (1 µM) and KIAA1718 (0.15 µM) selected by their binding K_D of FI-methylstat, the IC_{50} values are not directly comparable.

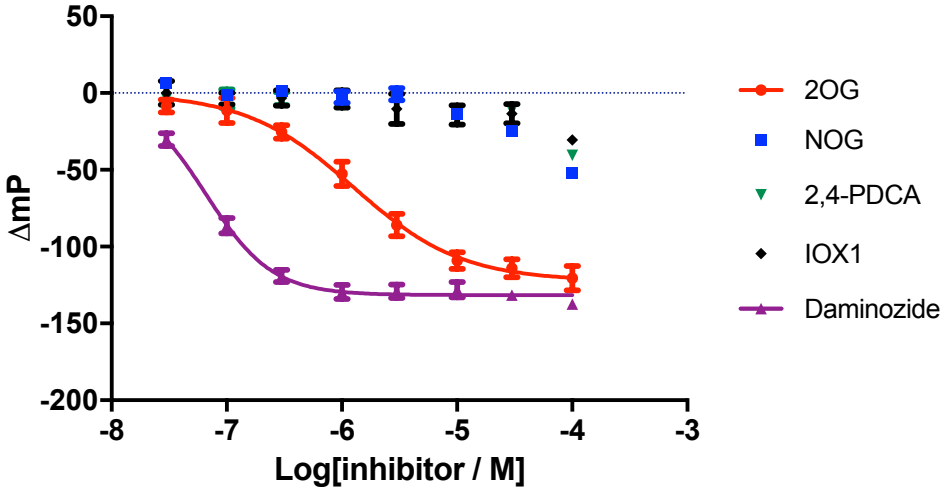


Figure 4.17 $KIAA1718_{PHD\text{-}JmjC}$ FP displacement assay measuring the IC_{50} values of 2OG, NOG, 2,4-PDCA, IOX1 and Daminozide. Data were fitted to a log(inhibitor) vs. response – Variable slope (four parameters) model using GraphPad Prism7 software. The fitting was fixed at 0 ΔmP at the top. R square values indicating the goodness of fitting were 0.95 and 0.96 for red and purple curves respectively. A sample without enzyme was used as a negative control for normalisation. A sample without inhibitor was used as a 'no inhibition' control. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

Although the *in vitro* activity assay for PHF2 had not identified any substrates (Chapter 2, section 2.3.6), the FP binding assay showed that $PHF2_{PHD\text{-}JmjC}$ and $PHF2_{JmjC}$ both bound to FI-methylstat with a K_D of single digital µM (Chapter 3, section 3.2.5). These FP assays provided a method to investigate the non-catalytic PHF2 and its interaction with small molecules. Like for $PHF8_{PHD\text{-}JmjC}$ and $KIAA1718_{PHD\text{-}JmjC}$, $PHF2_{PHD\text{-}JmjC}$ was subjected to the same set of inhibitor screening. Daminozide was used as a positive control for PHF8 to ensure negative binding results of compounds were not from the operational mistakes. Unfortunately, no potent inhibitor was found for PHF2. Marginal displacement of FI-methylstat was observed for NOG at 1 mM, however, the IC_{50} was too high to measure accurately (Figure 4.18).

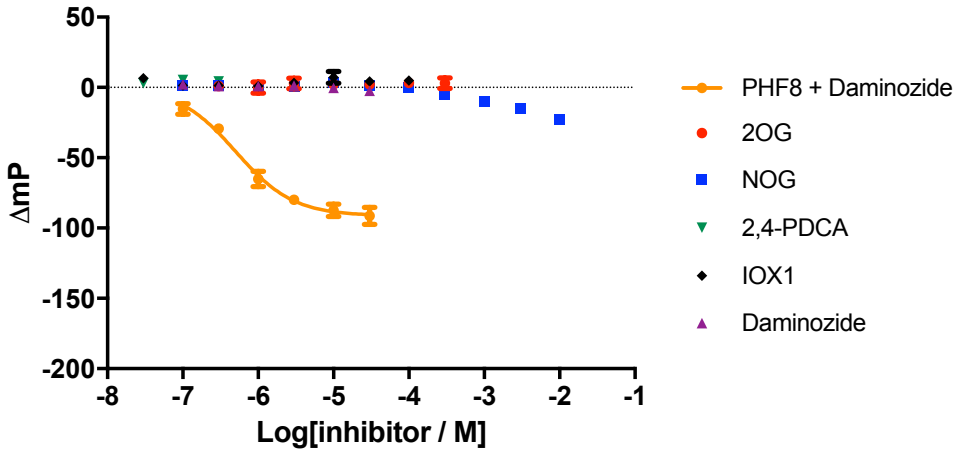


Figure 4.18 $PHF2_{PHD\text{-}JmjC}$ FP displacement assay measuring the IC_{50} values of 2OG, NOG, 2,4-PDCA, IOX1 and Daminozide. Data were fitted to a log(inhibitor) vs. response – Variable slope (four parameters) model using GraphPad Prism7 software. The fitting was fixed at 0 ΔmP at the top. R square value indicating the goodness of fitting was 0.95 for yellow curve. A sample without enzyme was used as a negative control for normalisation. A sample without inhibitor was used as a 'no inhibition' control. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

Next, the FP displacement assay for the inhibitors targeting the PHD finger was applied to the unlabelled $H3_{1\text{-}15}$K4me3K9me2 peptide with the KDM7 proteins containing the PHD finger only. As expected, the unlabelled peptide was able to compete with its FITC labelled analogue for all three enzymes. The IC_{50} values of KIAA1718, PHF8 and PHF2 were 2.3 µM, 5.2 µM, 1.1 µM respectively (Figure 4.19).

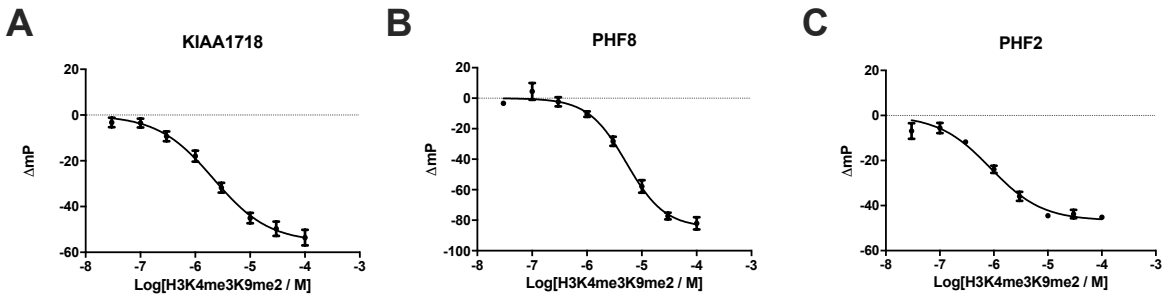


Figure 4.19 FP displacement assay of **A)** KIAA1718; **B)** PHF8; **C)** PHF2 using unlabelled H3 peptide (0.03 to 100 µM) and H3K4me3K9me2-FITC (50 nM). Data were fitted to a log(inhibitor) vs. response – Variable slope (four parameters) model using GraphPad Prism7 software. The fitting was fixed at 0 ΔmP at the top. R square value indicating the goodness of fitting were 0.97, 0.97 and 0.97 for A, B and C. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

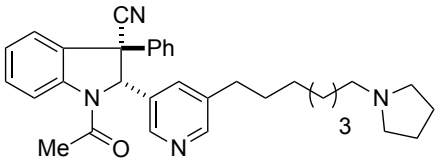
The inhibitors tested by FP displacement assay for KDM7s and their IC_{50} values are summarised in Table 4.1. Using a 384-well assay plate, IC_{50} values of 16 inhibitors were measured (8 concentrations, and in triplicate) within 25 min (20 min of incubation and 5 min of plate reading). Collectively, the results demonstrate that the FP displacement assay is a suitable assay for KDM7 inhibitor screening, as well as for non-catalytic enzyme studies.

Table 4.1 Summary of the IC_{50} values tested using FP displacement assay for the KDM7 subfamily. "NM"indicated that the IC_{50} values are not able to be measured precisely.

	PHF8	PHF2	KIAA	Probe
2OG	55.6 µM	NM	1.39 µM	FI-methylstat
NOG	> 100 µM	>10 mM	>100 µM	FI-methylstat
2,4-PDCA	NM	NM	NM	FI-methylstat
IOX1	NM	NM	NM	FI-methylstat
Daminozide	0.49 µM	NM	<50 nM	FI-methylstat
$H3_{1-15}$K4me3K9me2	5.2 µM	1.1 µM	2.3 µM	$H3_{1-15}$ K4me3K9me2-FITC

4.3.3 Investigating the mode of inhibition of a specific inhibitor of the KDM2/7 subfamilies

A screen of a library of known binders to methyllysine reader domains and histone methyl transferases carried out by the Structural Genomics Consoritum, Oxford has identified small molecule series containing an indole scaffold that inhibit KDM2A. Further medicinal chemistry efforts (by Dr Philip Gerkin) yielded a selective and cell permeable KDM2A inhibitor, (*S*,*S*)-6, with an IC_{50} of 0.16 µM against recombinant KDM2A (Figure 4.20). Since KDM2 and KDM7 subfamilies are closely related[27], cross-screening against PHF8 and KIAA1718 was carried out to see whether they are also catalytically inhibited by (*S*,*S*)-6.



(*S*,*S*)-**6**. KDM2A IC_{50} 0.16 μM

Figure 4.20 Structure of (*S*,*S*)-6.[11]

IC_{50} values were determined using the AlphaScreen method, employing the conditions described in the previous section (Section 4.3.1). A pilot assay showed that (*S*,*S*)-6 was more potent against KIAA1718 than PHF8 (figure not shown), thus a dilution of compound from 10 μM was tested with KIAA1718 (dilution from 100 μM was tested with PHF8). The inhibition assay with each enzyme was repeated two or three times and the average of duplicates or triplicates was taken as the final IC_{50} value (Table 4.2). Pleasingly, (S,S)-6 had > 50-fold selectivity for KIAA1718 (IC_{50} of 0.19 μM, Figure 4.21 A) over PHF8 (IC_{50} of >10 μM, Figure 4.21 B), suggesting this was also a potent and selective inhibitor for KIAA1718 over PHF8.

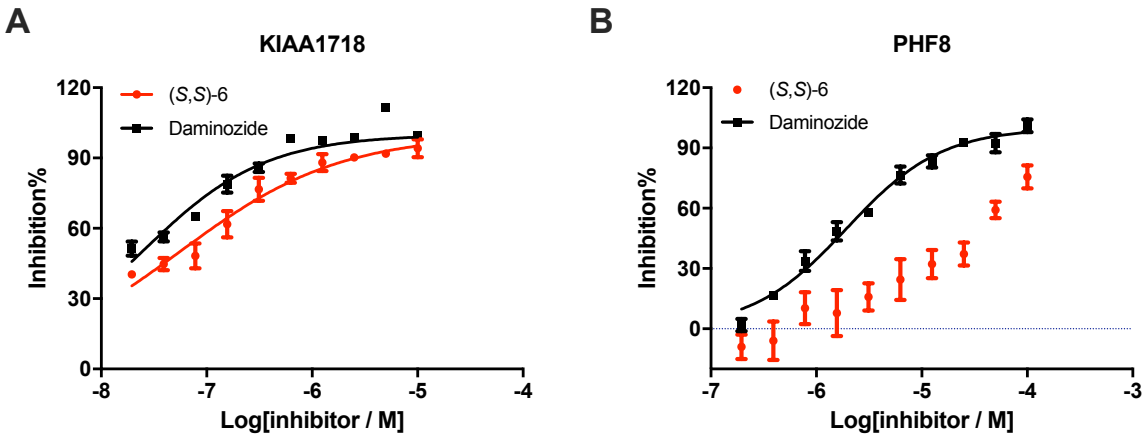


Figure 4.21 IC_{50} determination for (*S*,*S*)-6 and Daminozide against A) KIAA1718 (10 nM) B) PHF8 (10 nM) measured by AlphaScreen. Data were fitted to a log(inhibitor) vs. response – Variable slope (four parameters) model using GraphPad Prism7 software. The fitting was fixed at 0% inhibition at the bottom and 100% at the top. R square values indicating the goodness of fitting were 0.95 (red), 0.92 (black) for A and 0.97 (black) for B. A sample without enzyme was used as a negative control for normalisation. A sample without inhibitor was used as a 'no inhibition' control. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

Table 4.2 Summary of the IC_{50} values of (*S*,*S*)-6 against KIAA1718 and PHF8 calculated in repetitive assays.

	Repeat 1	N	Repeat 2	N	Repeat 3	N	Average	S.D.
KIAA1718	0.18 μM	3	0.14 μM	2	0.26 μM	2	0.19 μM	0.06
PHF8	>10 μM	3	>10 μM	3	-	-	>10 μM	-

(*S*,*S*)-6 was further tested against representative members from each of the other KDM subfamilies, i.e. KDM3A from KDM3 subfamily; KDM4A, C and D from KDM4 subfamily; KDM5B, C from KDM5 subfamily and KDM6B from KDM6 subfamily. IC_{50} values of these enzymes from KDM subfamilies other than KDM2/7 are more than 75-fold of KIAA1718 and of KDM2A (Table 4.3). The compound (*S*,*S*)-6 therefore appeared to be selective for the closely related KDM2 and KDM7 subfamilies.

Table 4.3 (*S*,*S*)-6 Selectivity vs other KDMs (IC_{50} μM). AlphaScreens for JmjC-KDMs (except KDM7s) were carried out by Dr Anthony Tumber.

KDM2A	KDM3A	KDM4A	KDM4C	KDM4D	KDM5B	KDM5C	KDM6B	KIAA1718	PHF8
0.16	39	639	430	321	17	16	20	0.19	> 10

To investigate the mode of inhibition of (*S*,*S*)-6 and its selectivity between the two KDM7 enzymes tested, FP displacement assays were carried out with KDM2A, KIAA1718 and PHF8. Since (*S*,*S*)-6 was designed to target the JmjC domain, the FI-methylstat competition assay was first carried out with 0.12 μM of KDM2A. The concentration of KDM2A was equal to the K_D value of its binding to FI-methylstat (Figure 4.22 A). By contrast to the strong inhibition against the enzymatic activity of KDM2A, (*S*,*S*)-6 displaced FI-methylstat with an IC_{50} of more than 100 μM, while the known broad 2OG oxygenase inhibitor 2,4-PDCA had an IC_{50} of 13.2 μM (Figure 4.22 B). FP assay results for the other inhibitors carried out by Dr Anthony Tumber also confirmed the relatively high IC_{50} value of (*S*,*S*)-6 against KDM2A (data not shown).

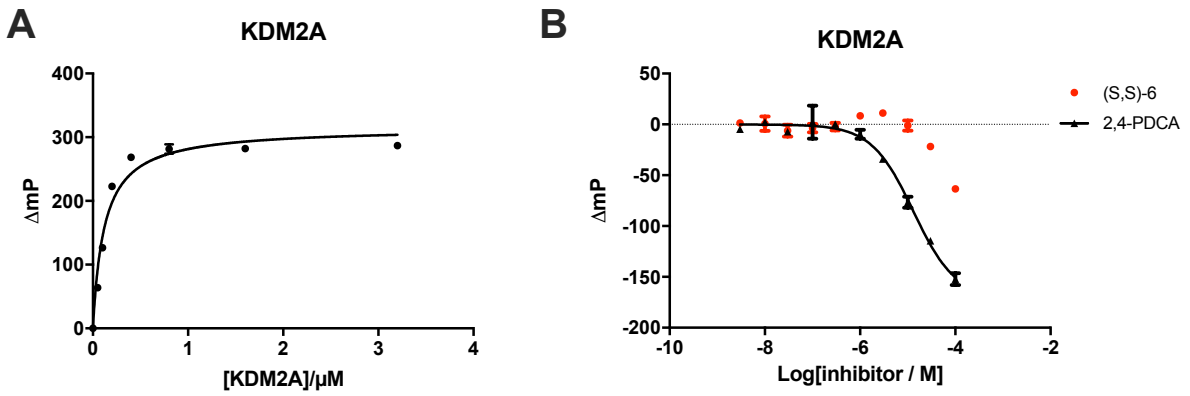


Figure 4.22 FP assays measuring **A)** binding of KDM2A to 12.5 nM FI-methylstat (K_D value of KDM2A to FI-methylstat was 012 μM as determined by fitting data into a one-site specific binding model using GraphPad Prism7 software). **B)** IC_{50} values of (*S*,*S*)-6 and 2,4-PDCA against KDM2A. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

Displacement assays with (*S*,*S*)-6 were also carried out for KIAA1718 and PHF8. Surprisingly, no strong displacement was found against PHF8 or KIAA1718 (Figure 4.23), indicating that (*S*,*S*)-6 may not directly bind to the JmjC domains of these enzymes.

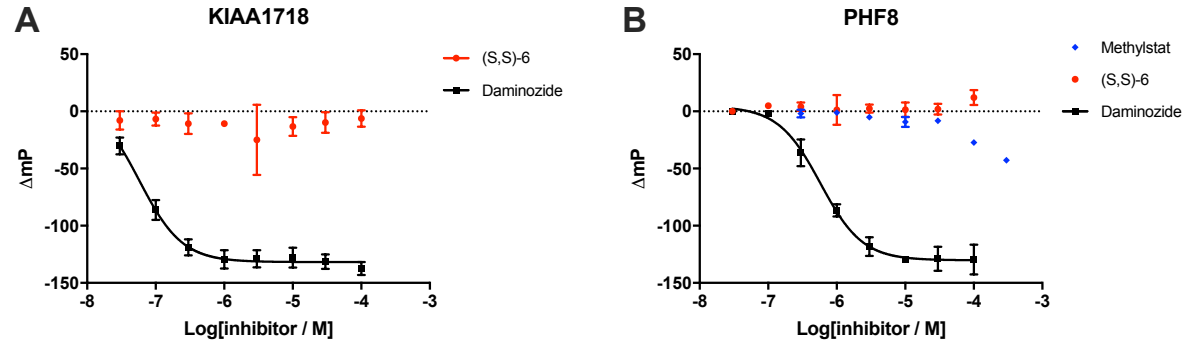


Figure 4.23 IC_{50} determination for (*S*,*S*)-6, Daminozide and Methylstat acid against **A)** KIAA1718 (0.13 μM) **B)** PHF8 (1 μM) measured by FP displacement assay with FI-methylstat (12.5 nM). Data were fitted to a log(inhibitor) vs. response – Variable slope (four parameters) model using GraphPad Prism7 software. The fitting was fixed at 0 ΔmP at the top. R square values indicating the goodness of fitting were 0.96 for A (black curve) and 0.99 for B (black curve). A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

The IC_{50} values of (*S*,*S*)-6 and other inhibitors against KDM2A, KIAA1718 and PHF8 are summarised in Table 4.4.

Table 4.4 IC_{50} values for Methylstat acid, Daminozide and (*S*,*S*)-6 against KDM2A, KIAA1718 and PHF8 measuring by FP displacement assay.

FI-methylstat	IC_{50} / µM (% inh at 100 µM)		
	KDM2A	KIAA1718	PHF8
(S,S)-6	> 100 (21%)	> 100	> 100
Daminozide	0.23	0.10	0.59
Methylstat acid	0.14	-	131.83

Next, the FP displacement assay was carried out using $H3_{1-15}$K4me3K9me2-FITC to investigate whether (*S*,*S*)-6 competes with the substrate. Although more promising than the competition with FI-methylstat, (*S*,*S*)-6 was unable to give a full dose-response curve at the concentrations tested and the IC_{50} values against all three enzymes were more than 100 µM (Figure 4.24). Therefore, the FP displacement assays suggest that the competition of (*S*,*S*)-6 against $H3_{1-15}$K4me3K9me2-FITC binding to the KDM7 dual domain enzymes may, in part, contribute to the inhibition of KDM7 catalysis and/or the mechanism of inhibition by (*S*,*S*)-6 is complex.

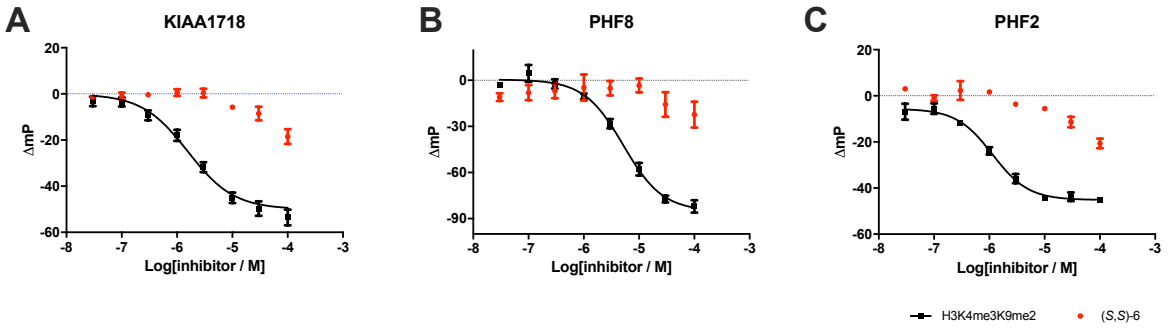


Figure 4.24 IC_{50} determination for (*S*,*S*)-6 against **A)** KIAA1718 (0.2 µM enzyme) **B)** PHF8 (1 µM enzyme) **C)** PHF2 (0.03 µM enzyme) measured by FP displacement assay with $H3_{1-15}$K4me3K9me2-FITC (50 nM). Data were fitted to a log(inhibitor) vs. response – Variable slope (four parameters) model using GraphPad Prism7 software. The fitting was fixed at 0 ΔmP at the top. R square values indicating the goodness of fitting were 0.96, 0.97 and 0.98 for A, B and C. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

As demonstrated in the previous Chapter (Chapter 3, section 3), the initial binding of KDM7 enzymes to $H3_{1-15}$K4me3K9me2-FITC is dominated by interactions with their PHD fingers. To assess whether (*S*,*S*)-6 bound to the PHD finger domain, an FP displacement assay with $KIAA1718_{PHD}$ was carried out. Interestingly, but perhaps not surprisingly, given that the KDM2A construct used for these assays contained a JmjC domain, the inhibitor did not displace $H3_{1-15}$K4me3K9me2-FITC even at 100 µM (Figure 4.25). Together with the results against FI-methylstat, the FP displacement assays revealed that the (*S*,*S*)-6 does not directly compete for binding with active-site targeting methylstat to the JmjC domains or H3K4me3 peptide binding to PHD fingers of KDM2A and KIAA1718 alone. Given the construct of KIAA1718 used in this assay contained a PHD finger, a JmjC domain and a linker between them, it may suggest that the (*S*,*S*)-6 binds to the linker region of KIAA1718, or the dual domain interface. This could explain the different inhibitory effects of (*S*,*S*)-6 on KIAA1718 and PHF8, since the sequences of their linker regions are significantly different.[14]

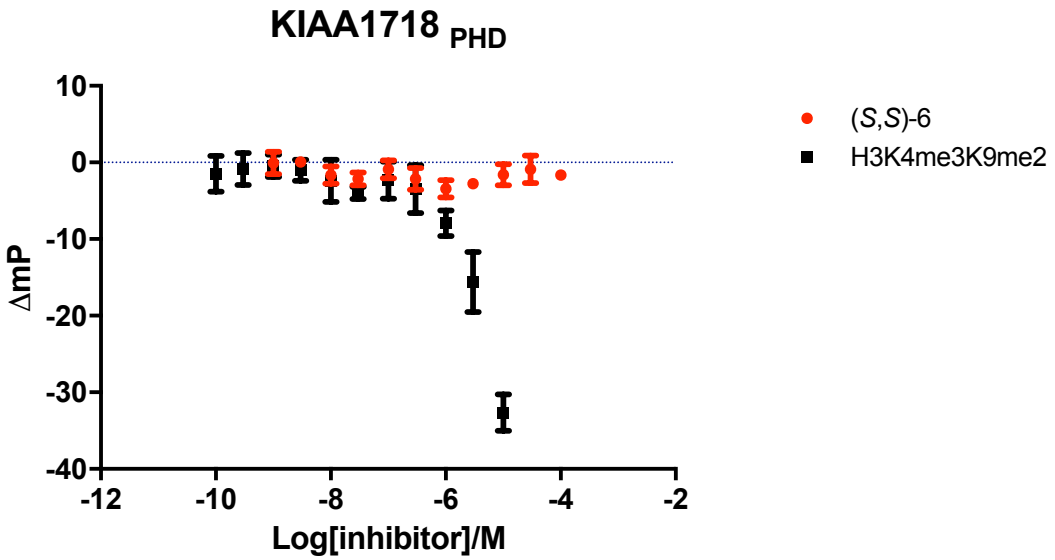


Figure 4.25 FP displacement assay of (*S*,*S*)-6 competing with $H3_{1-15}$K4me3K9me2-FITC (50 nM) on binding to $KIAA1718_{PHD}$ (0.5 µM). A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

The sequence alignment of the KDM2A construct used in this study with the linker region of KIAA1718 reveals ~ 40% identity and ~ 55% similarity (Figure 4.26). These residues appear in the catalytic centre of KDM2A and within the linker of KIAA1718, but are not in the linker of PHF8. The inhibitor may thus recognise these similar residues and interrupts the demethylation activity of KIAA1718 non-competitively, by which the different IC_{50} values against KIAA1718 and PHF8 can also be explained. In conclusion, the KDM2A inhibitor (*S*,*S*)-6 does not directly bind to the JmjC domain or the PHD finger of KIAA1718 and may be a non-competitive inhibitor of KIAA1718. Part of the work described in this section has been published in Gerken *et al.*, *Angew*. *Chem*. 2017.

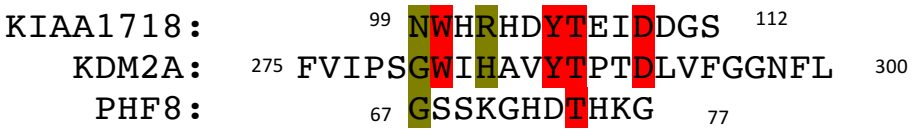


Figure 4.26 Sequence alignment of KIAA1718 (residue 99-112), KDM2A (residue 275-300) and PHF8 (residue 67-77). Red boxes label identical residues and orange boxes label similar residues.

4.3.4 Identifying novel cyclic peptide inhibitors of the KDM7 subfamily

The cyclic peptide inhibitors of the KDM7 subfamily have several potential advantages over iron-chelating inhibitors and peptide mimics, such as target selectivity and low toxicity that make them an attractive strategy for drug development.[28] From 2006 to 2015, nine cyclic peptide drugs have been approved by the Food and Drug Administration (FDA) and European Medicines Agency (EMA), which are applied in bacterial and fungal infection and gastrointestinal disorders.[29] As most of the reported inhibitors of KDM7 subfamily have not achieved true selectivity (they usually also inhibit the KDM2 subfamily)[3–6], cyclic peptide inhibitors were investigated for KDM7 subfamily to achieve selectivity.

Three potential cyclic peptide inhibitors OC3, OC6 and OC9 were identified by Oliver Coleman using an mRNA-display platform (since the work is in preparation of publication, the sequences are not disclosed here). Initially, a pilot MALDI-TOF MS-based activity assay with OC3, OC6 and OC9 (10 μM and 100 μM) demonstrated their inhibition of the *in vitro* demethylation activity of KIAA1718$_{PHD\text{-}JmjC}$ and PHF8$_{PHD\text{-}JmjC}$ (data not shown). Time course studies of the two enzymes with the cyclic peptides were subsequently carried out. OC3 and OC6 both inhibited the demethylation activity of KIAA1718 on H3$_{1\text{-}15}$K9me2, bringing the $t_{1/2}$ from 25 min to 71 min and 30 min respectively (Figure 4.27 A). Interestingly, OC9 slightly enhanced KIAA1718 activity with a $t_{1/2}$ of 19 min (Figure 4.27 A). This observation could reflect release of PHD finger from binding to the non-methylated lysine at position H3K4. An enhancement effect of OC9 on KIAA1718 activity was further supported by the demethylation reaction of KIAA1718$_{PHD\text{-}JmjC}$ with H3$_{1\text{-}15}$K4me3K9me2 (Figure 4.27 B). OC9 enhanced the activity by more than 4.5-fold, whilst OC3 and OC6 both appeared to slightly inhibit the demethylation of H3$_{1\text{-}15}$K4me3K9me2 by KIAA1718$_{PHD\text{-}JmjC}$ (Table 4.5).

For PHF8$_{PHD\text{-}JmjC}$, OC3, OC6 and OC9 all inhibited activity on H3$_{1\text{-}15}$K4me3K9me2 and H3$_{1\text{-}15}$K9me2. The extent of demethylation of H3$_{1\text{-}15}$K9me2 at 60 min suggested that OC9 inhibited the demethylation by ~ 30% (Figure 4.27 C). The reaction of PHF8$_{PHD\text{-}JmjC}$ with H3$_{1\text{-}15}$K4me3K9me2 was too fast to show any significant difference in $t_{1/2}$. However, a slower turnover rate at the first 10 min was observed with OC9, changing the time required for 90% conversion from less than 5 min to 10 min (Figure 4.27 D). Overall, the time course studies suggested that for both PHF8 and KIAA1718, OC9 had the most potent effect on both enzymes as either as an inhibitor or an 'enhancer'. The parameters of the time course study are summarised in Table 4.5.

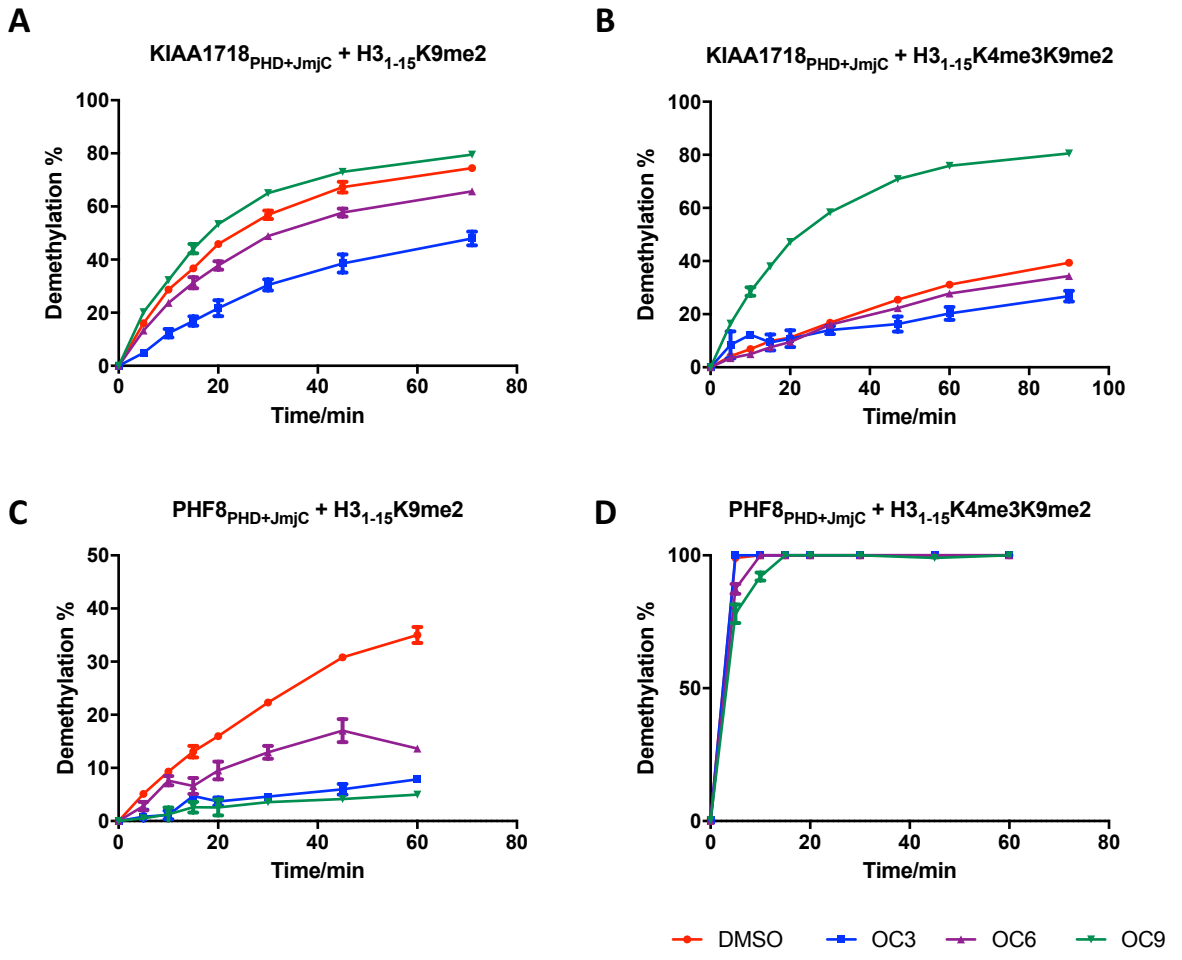


Figure 4.27 MALDI-TOF MS assays of KDM7 enzymes with OC3, OC6 and OC9 (10 μM in 0.1% DMSO). **A)** KIAA1718$_{PHD-JmjC}$ (5 μM) with H3$_{1-15}$K9me2 (10 μM). **B)** KIAA1718$_{PHD-JmjC}$ (5 μM) with H3$_{1-15}$K4me3K9me2 (10 μM). **C)** PHF8$_{PHD-JmjC}$ (5 μM) with H3$_{1-15}$K9me2 (10 μM). **D)** PHF8$_{PHD-JmjC}$ (5 μM) with H3$_{1-15}$K4me3K9me2 (10 μM). An assay with 0.1% DMSO was used as the standard time course for comparison. A pre-quenched sample was used for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

Table 4.5 $t_{1/2}$ of the demethylation reactions catalysed by KIAA1718 and PHF8 with cyclic peptide inhibitors.

	KIAA1718		PHF8 (% at 60 min)	
	H3K9me2	H3K4me3K9me2	H3K9me2	H3K4me3K9me2
OC3	71 min	> 90 min	> 60 min (8%)	< 5 min
OC6	30 min	> 90 min	> 60 min (14%)	< 5 min
OC9	19 min	20 min	> 60 min (5%)	< 5 min
DMSO	25 min	> 90 min	> 60 min (35%)	< 5 min

To measure the IC_{50} values of the cyclic peptides, AlphaScreen inhibition assays were carried out with KIAA1718$_{PHD-JmjC}$ and PHF8$_{PHD-JmjC}$, which did not succeed in obtaining IC_{50} values of the cyclic peptides (Figure 4.28 A & B). Two reasons were proposed regarding the

unsuccessful AlphaScreen assays: the cyclic peptides may emit or alter the luminescence signal and so interfere with the signal generated from the detection beads, or they may interact with the antibody and caused a loss of antibody selectivity. To evaluate the potential interference from cyclic peptide inhibitors, an assay with H3K9me1 was carried out. In this assay, OC3, OC6 and OC9 (0, 10, 100 µM in 0.1% DMSO final) were incubated with $H3_{1\text{-}21}K9me1$ (10 nM), antibody Ab8896, and the detection beads for 1 hour. Analysis of the end-point AlphaScreen ouput suggested that the cyclic peptides do indeed influence the assay read-out (Figure 4.28 C).

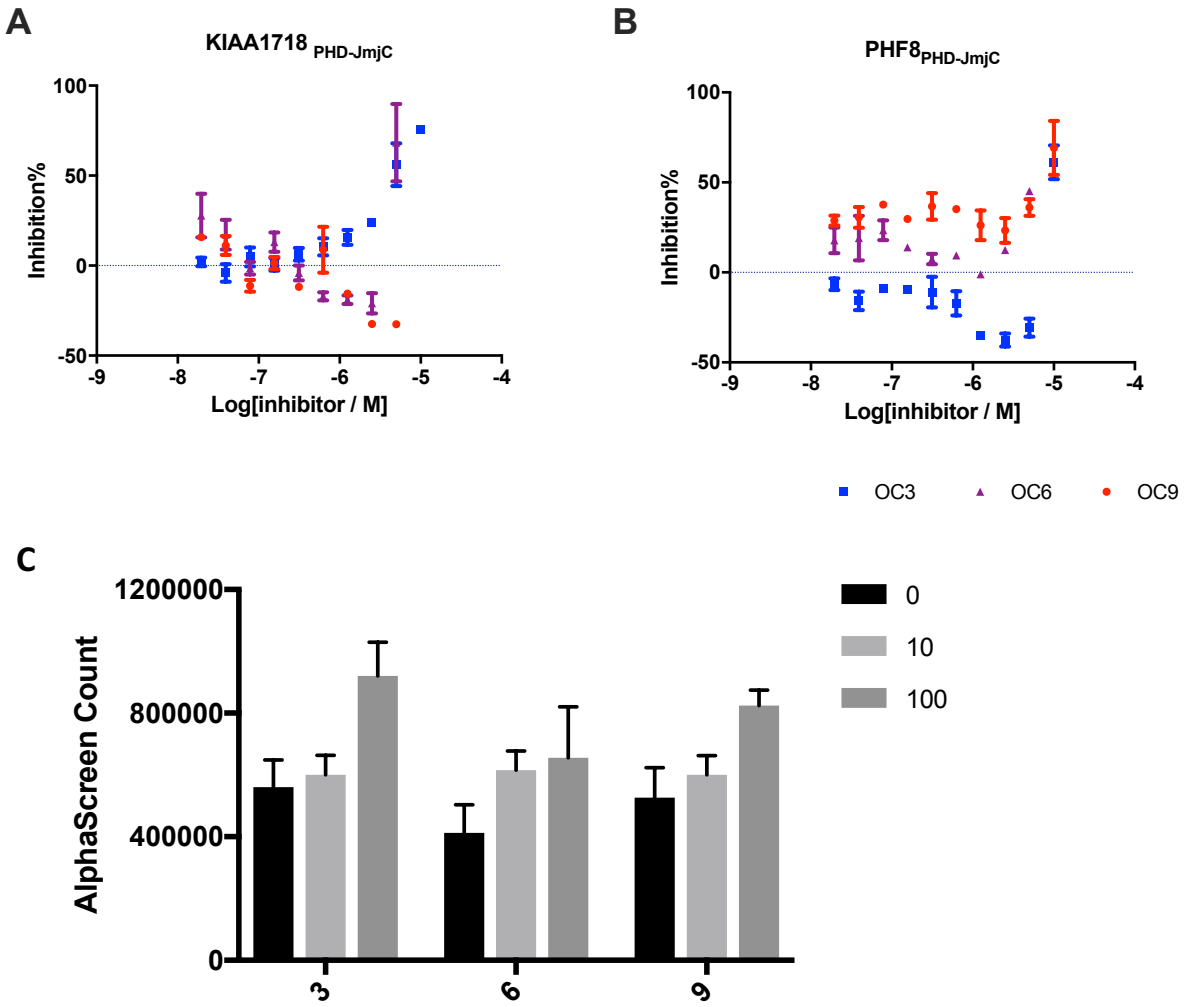


Figure 4.28 AlphaScreen assays of KDM7 enzymes with OC3, OC6 and OC9 (in 0.1% DMSO). **A)** IC_{50} values measurement of $KIAA1718_{PHD\text{-}JmjC}$ (10 nM) on $H3_{1\text{-}15}K9me2$ (30 nM). **B)** IC_{50} values measurement of $PHF8_{PHD\text{-}JmjC}$ (10 nM) on $H3_{1\text{-}15}K4me3K9me2$ (0.3 µM). **C)** Signal interference test. A sample with 0.1% DMSO was used for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

Another possible reason contributing to the unsuccessful AlphaScreen assays may be the choice of substrate. For both KIAA1718 and PHF8, the assay was conducted using their best known substrate (i.e. KIAA1718 with H3K9me2 and PHF8 with H3K4me3K9me2); the reactions of these two substrates were only slightly influenced by the cyclic pepitides as suggested by the MALDI-TOF MS assays (Figure 4.27). Therefore, the IC_{50} values may be too high to measure by the current AlphaScreen assay set-up. Screening across representative enzymes from other KDM subfamilies using MALDI-TOF MS demonstrated that OC9 iss a selective inhibitor of PHF8 with an IC_{50} value of 1.4 μM and an 'enhancer' of KIAA1718 with an EC_{50} of 4.0 μM (assays conducted by Oliver Coleman).

In order to test the proposal of the mode of inhibition of the cyclic peptides, FP displacement assays of OC3, OC6 and OC9 were carried out on $KIAA1718_{PHD}$ (binding to $H3_{1-15}$K4me3K9me2-FITC) and $PHF8_{JmjC}$ (binding to FI-methylstat). OC9 competed against $H3_{1-15}$K4me3K9me2-FITC on $KIAA1718_{PHD}$ with an IC_{50} of 0.5 μM while OC3 and OC6 did not (Figure 4.29 A). These observations suggest that only OC9 selectively targets the PHD finger. The assay with $PHF8_{JmjC}$ showed that OC3 and OC9 did not displace FI-methystat on the JmjC domain and OC6 was slightly competitive at high concentration (Figure 4.29 B). An assay using Protein NMR (conducted by Dr Martine Abboud and Oliver Coleman) also confirmed the binding of OC9 to the PHD finger of PHF8. Together with MALDI-TOF MS and RapidFire data, it was concluded that OC9 was a selective PHD finger-targeting inhibitor of KIAA1718 and PHF8.

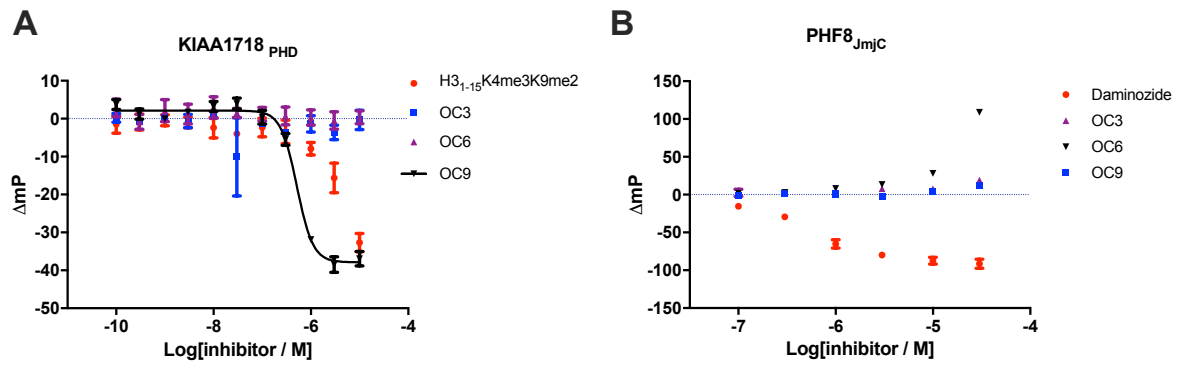


Figure 4.29 FP displacement assay measuring the IC_{50} of OC3, OC6 and OC9 (in 0.1% DMSO) against **A)** $KIAA1718_{PHD}$ (0.5 µM). **B)** $PHF8_{JmjC}$ (1 µM). Data set A were fitted to a log(inhibitor) vs. response – Variable slope (four parameters) model using GraphPad Prism7 software. R square value indicating the goodness of fitting was 0.98 for black curve. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

4.3.5 Amiodarone derivatives inhibit KDM7s in a PHD finger independent manner

In 2012, Wagner *et al.* identified amiodarone (AMI), an antiarrhythmic drug, as an inhibitor of the PHD finger 3 of KDM5A ($KDM5A_{PHD3}$). Inspired by the amiodarone derivatives, a series of compounds were synthesised by Dr Alexandre Erdmann (Figure 4.30).

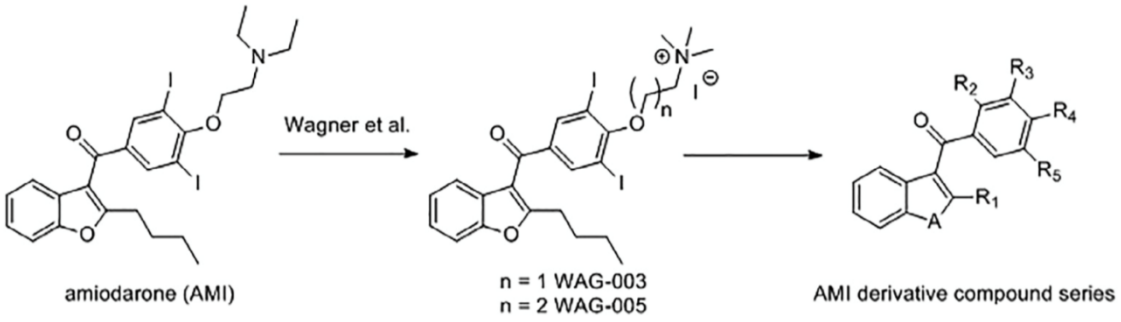
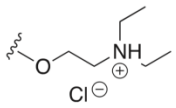
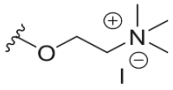
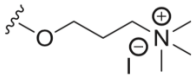
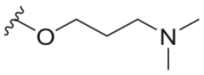
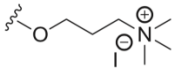
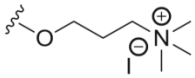
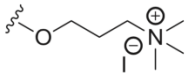
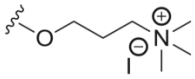
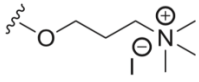


Figure 4.30 Structures of amiodarone (AMI) and its derivatives reported by Wagner *et al.* and synthesised by Dr Alexandre Erdmann.

IC_{50} values of the compounds against the binding of $H3_{1-21}K4me3$ to $KDM5A_{PHD3}$ were measured using AlphaScreen activity assays (conducted by Dr Bhaskar Bhushan, Table 4.6). Although selected compounds inhibited the *in vitro* activity of $KDM5A_{PHD3}$, none was found to have a lower IC_{50} value than the parent compound AMI (Table 4.6). The IC_{50} values of the AMI derivatives inhibiting the *in vitro* activity of $KDM5A_{PHD3}$ are shown in Table 4.6. Since the PHD finger of KDM7 subfamily of enzymes plays a significant role in substrate binding and selectivity, the effect of the AMI derivatives on the catalytic activity of $KIAA1718_{PHD\text{-}JmjC}$ and $PHF8_{PHD\text{-}JmjC}$ was investigated. A pilot test with $PHF8_{PHD\text{-}JmjC}$ was carried out using a MALDI-TOF MS-based assay. WAG-003, WAG-005 and compound 2a (10 µM and 100 µM) were incubated with $PHF8_{PHD\text{-}JmjC}$ (5 µM) for 15 min before the addition of $H3_{1\text{-}15}K4me3K9me2$ and cofactors to start the demethylation reaction. The percentage demethylation was measured after 1-hour incubation at room temperature. WAG-003

showed about 30% inhibition (from 30% to 20%) at 100 µM and WAG-005 showed 50% (from 30% to 15%) inhibition at 100 µM (Figure 4.31).

Table 4.6 Structures and IC_{50} values of Amiodarone (AMI) and its derivatives against $KDM5A_{PHD3}$. NM: not measurable within the solubility range of the compound. Average ± StdDev (n ≥ 3 independent replicates) shown. $H3_{1-21}K4me3$ IC_{50}: 0.22 ± 0.02 µM. Final concentrations of protein and peptide were 25 nM.

Compounds	IC_{50} (µM)	A	R_1	R_2	$R_3 = R_5$	R_4
AMI	99 ± 23	O	butyl	H	I	
WAG-003	1658 ± 490	O	butyl	H	I	
WAG-005	562 ± 130	O	butyl	H	I	
1a	166 ± 20	O	butyl	H	H	
2a	340 ± 44	O	butyl	H	H	
2b	844 ± 23	O	methyl	H	H	
2c	533 ± 109	S	methyl	H	H	
2d	1310 ± 39	NMe	H	H	H	
7b	2305 ± 82	O	methyl		H	**H**

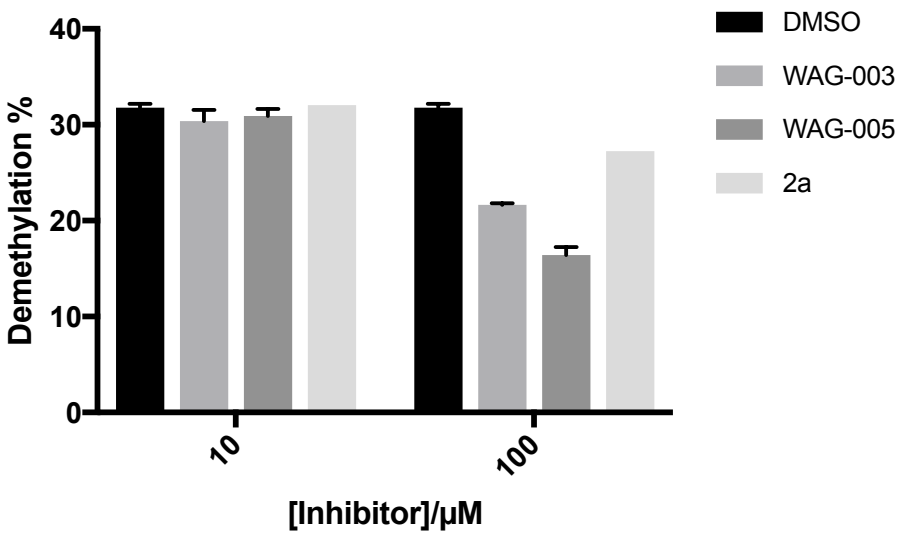


Figure 4.31 Percentage demethylation of $PHF8_{PHD\text{-}JmjC}$ on $H3_{1\text{-}15}K4me3K9me2$ after 1-hour incubation with WAG-003, WAG-005 and compound 2a (10 μM and 100 μM in 1% DMSO) and 1% DMSO. Reactions were quenched by 2% (v/v) aqueous formic acid after 1-hour incubation at room temperature. The extent of demethylation (percentage) was analysed by MALDI-TOF MS. Figures were generated using GraphPad Prism7 software. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical duplicates (except for 2a where a single assay was carried out), and error bars give ± one s.d.

The IC_{50} values of the inhibitors were measured by the AlphaScreen assays. Since the compounds were designed to target the PHD finger, AlphaScreen assays were carried out with $PHF8_{PHD}$ on $H3_{1\text{-}21}K4me3K9me2$-biotin using the unlabelled histone peptide for comparison. Interestingly, as was observed for $KDM5A_{PHD3}$,[30] none of the compounds tested showed a strong inhibitory effect against the binding activity of $PHF8_{PHD}$ (Figure 4.32). For example, compound 2a at 100 μM inhibited the *in vitro* catalytic activity of PHF8 on $H3_{1\text{-}15}K9me2$ by ~30% (Figure 4.31), however, it did not displace the $H3_{1\text{-}21}K4me3K9me2$-biotin peptide from $PHF8_{PHD}$ at the same concentration (Figure 4.32), suggesting that compound 2a likely inhibits PHF8 in a PHD finger independent manner.

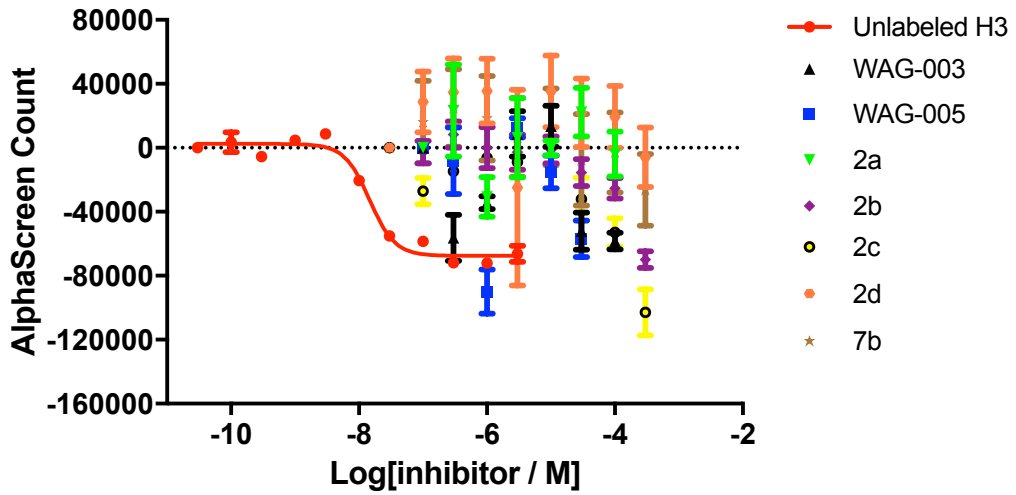


Figure 4.32 IC_{50} values measurement of the representative AMI derivatives against $PHF8_{PHD}$ (10 nM) using AlphaScreen displacement assay. Data were fitted to a log(inhibitor) vs. response – Variable slope (four parameters) model using GraphPad Prism7 software. The fitting was fixed at 0 at the top. R square value indicating the goodness of fitting was 0.97 for blue curve. A sample without enzyme was used as a negative control for normalisation. Data are plotted as the mean of technical triplicates, and error bars give ± one s.d.

A cross-screening of the compound library against $PHF8_{PHD\text{-}JmjC}$ and $KIAA1718_{PHD\text{-}JmjC}$ compared the percentage inhibition by co-incubation of compounds (100 µM) with 10 µM of $PHF8_{PHD\text{-}JmjC}$ or 100 µM of $KIAA1718_{PHD\text{-}JmjC}$ using MALDI-TOF MS assays (Chapter 5, section 5.5). The IC_{50} values of activity inhibition against $PHF8_{PHD\text{-}JmjC}$ and $KIAA1718_{PHD\text{-}JmjC}$ were measured by Dr Bhaskar Bhushan using AlphaScreen activity assays. The results suggest that the AMI derivatives 1a, 2a and 2d were significantly more potent inhibitors of KIAA1718 catalytic activity than inhibitors of PHD finger peptide binding. Promisingly, 2b was found to displace the biotinylated peptide on $KIAA1718_{PHD\text{-}JmjC}$ with an IC_{50} of 5.6 $\pm$ 1.5 µM (Table 4.7). However, 2b also completely inhibited the activity of $KIAA1718_{PHD\text{-}JmjC}$ at 100 µM suggesting that 2b can also act at the JmjC domain. Although none of the tested AMI derivatives was selective to PHD finger, they still provided a direction of further design of the AMI derivatives. The described work has potential applications for the future

development of PHD-finger targeting inhibitors for KDMs and was published in Bhushan *et al.* 2017.[30]

Table 4.7 IC_{50} values of the AMI derivatives for displacement of $H3_{1-21}K4me3$-biotin against $PHF8_{PHD-JmjC}$ and $KIAA1718_{PHD-JmjC}$ by AlphaScreen (conducted by Dr Bhaskar Bhushan), and their percentage inhibition of demethylation activity at 100 µM by MALDI-TOF MS.

Compounds	$PHF8_{PHD-JmjC}$		$KIAA1718_{PHD-JmjC}$	
	Inhibition% at 100 µM (MALDI)	IC_{50} (AS)	Inhibition% at 100 µM (MALDI)	IC_{50} (AS)
AMI	22%	-	83%	80 ± 26
WAG-003	22%	-	17%	-
WAG-005	33%	-	100%	16 ± 4
1a	95%	-	100%	295 ± 48
2a	89%	120 ± 32	100%	149 ± 32
2b	32%	126 ± 34	100%	5.6 ± 1.5
2d	24%	>300	100%	88 ± 17

4.4 Summary

This work described in this Chapter concerns inhibition studies on the KDM7 subfamily. Two inhibitor screening platforms, the AlphaScreen based platform and the FP based platform were established for the inhibitor screening of the KDM7 enzymes. Using these two platforms, novel inhibitors for the KDM7 enzymes were investigated along with their modes of inhibition.

The mode of inhibition of a selective KDM2/7 inhibitor (*S,S*)-6 was investigated. Although (*S,S*)-6 has been found to inhibit the enzymatic activity of KDM2A and KIAA1718 with IC_{50} values of 0.16 μM and 0.19 μM, respectively (using the AlphaScreen activity assays), FP competition assays of KIAA1718 proteins imply that the compound does not target either the PHD finger or the JmjC domain. Sequence alignment of KDM2A and KIAA1718/PHF8 suggested that (*S,S*)-6 is likely to bind to the centre of KDM2A and the linker region of KIAA1718, which share significant sequence similarity between each other.

The FP screening method also helped the identification of a novel cyclic peptide inhibitor of PHF8 and an 'enhancer' of KIAA1718 that binds selectively to their PHD fingers. Another example of the utility of the FP and AlphaScreen screening platforms was in the measurement of the IC_{50} values of a series of amiodarone (AMI) derivatives, which have been reported to inhibit the PHD finger 3 of KDM5A ($KDM5A_{PHD3}$). The potent AMI derivatives that inhibit the KDM activity of PHF8 and KIAA1718 have been found not directly bind to the PHD finger which they were designed to target. Overall, the results show that the AlphaScreen based method and the FP based method are appropriate for screening KDM7 inhibitors as well as investigating the mode of inhibition.

4.5 References

1 C. H. Arrowsmith, C. Bountra, P. V. Fish, K. Lee and M. Schapira, *Nat. Rev. Drug Discov.*, 2012, **11**, 384–400.

2 R. J. Hopkinson, A. Tumber and C. Yapp, *Chem Sci*, 2015, **4**, 3110–3117.

3 K. S. England, A. Tumber, T. Krojer, G. Scozzafava, S. S. Ng, M. Daniel, A. Szykowska, K. Che, F. Von Delft, N. A. Burgess-Brown, A. Kawamura, C. J. Schofield and P. E. Brennan, *Medchemcomm*, 2014, **5**, 1879–1886.

4 T. Suzuki, H. Ozasa, Y. Itoh, P. Zhan, H. Sawada, K. Mino, L. Walport, R. Ohkubo, A. Kawamura, M. Yonezawa, Y. Tsukada, A. Tumber, H. Nakagawa, M. Hasegawa, R. Sasaki, T. Mizukami, C. J. Schofield and N. Miyata, *J. Med. Chem.*, 2013, **56**, 7222–7231.

5 A. K. Upadhyay, D. Rotili, J. W. Han, R. Hu, Y. Chang, D. Labella, X. Zhang, Y. S. Yoon, A. Mai and X. Cheng, *J. Mol. Biol.*, 2012, **416**, 319–327.

6 N. R. Rose, E. C. Y. Woon, A. Tumber, L. J. Walport, R. Chowdhury, X. S. Li, O. N. F. King, C. Lejeune, S. Ng, T. Krojer, M. C. Chan, A. M. Rydzik, R. J. Hopkinson, K. Hing, M. Daniel, C. Strain-damerell, C. Gileadi, G. Kochan, K. K. Yeoh, P. J. Ratcliffe, N. Burgess-brown, F. Von Delft, S. Muller-knapp, B. Marsden, P. E. Brennan, M. A. Mcdonough, R. J. Klose, C. J. Schofield and A. Kawamura, *J. Med. Chem.*, 2012, **55**, 6639–6643.

7 L. Kruidenier, C. W. Chung, Z. Cheng, J. Liddle, K. Che, G. Joberty, M. Bantscheff, C. Bountra, A. Bridges, H. Diallo, D. Eberhard, S. Hutchinson, E. Jones, R. Katso, M. Leveridge, P. K. Mander, J. Mosley, C. Ramirez-Molina, P. Rowland, C. J. Schofield, R. J. Sheppard, J. E. Smith, C. Swales, R. Tanner, P. Thomas, A. Tumber, G. Drewes, U. Oppermann, D. J. Patel, K. Lee and D. M. Wilson, *Nature*, 2012, **488**, 404–408.

8 J. Liang, B. Zhang, S. Labadie, D. F. Ortwine, M. Vinogradova, J. R. Kiefer, V. S. Gehling, J. C. Harmange, R. Cummings, T. Lai, J. Liao, X. Zheng, Y. Liu, A. Gustafson, E. Van der Porten, W. Mao, B. M. Liederer, G. Deshmukh, M. Classon, P. Trojer, P. S. Dragovich and L. Murray, *Bioorganic Med. Chem. Lett.*, 2016, **26**, 4036–4041.

9 J. R. Horton, X. Liu, M. Gale, L. Wu, J. R. Shanks, X. Zhang, P. J. Webber, J. S. K. Bell, S. C. Kales, B. T. Mott, G. Rai, D. J. Jansen, M. J. Henderson, D. J. Urban, M. D. Hall, A. Simeonov, D. J. Maloney, M. A. Johns, H. Fu, A. Jadhav, P. M. Vertino, Q. Yan and X. Cheng, *Cell Chem. Biol.*, 2016, **23**, 769–781.

10 X. Luo, Y. Liu, S. Kubicek, J. Myllyharju, A. Tumber, S. Ng, K. H. Che, J. Podoll, T. D. Heightman, U. Oppermann, S. L. Schreiber and X. Wang, *J. Am. Chem. Soc.*, 2011, **133**, 9451–9456.

11 P. A. Gerken, J. R. Wolstenhulme, A. Tumber, S. B. Hatch, Y. Zhang, S. Müller, S. A. Chandler, B. Mair, F. Li, S. M. B. Nijman, R. Konietzny, T. Szommer, C. Yapp, O. Fedorov, J. L. P. Benesch, M. Vedadi, B. M. Kessler, A. Kawamura, P. E. Brennan and M. D. Smith, *Angew. Chemie - Int. Ed.*, 2017, **56**, 15555–15559.

12 U. Leurs, B. Lohse, K. D. Rand, S. Ming, E. S. Riise, P. A. Cole, J. L. Kristensen and R. P. Clausen, *ACS Chem. Biol.*, 2014, **9**, 2131–2138.

13 A. Kawamura, M. Münzel, T. Kojima, C. Yapp, B. Bhushan, Y. Goto, A. Tumber, T. Katoh, O. N. F. King, T. Passioura, L. J. Walport, S. B. Hatch, S. Madden, S. Müller, P. E. Brennan, R. Chowdhury, R. J. Hopkinson, H. Suga and C. J. Schofield, *Nat. Commun.*, , DOI:10.1038/ncomms14773.

14 J. R. Horton, A. K. Upadhyay, H. H. Qi, X. Zhang, Y. Shi and X. Cheng, *Nat. Struct. Mol.*

Biol., 2010, **17**, 38–44.
15 R. M. Eglen, T. Reisine, P. Roby, N. Rouleau, C. Illy, R. Bossé and M. Bielefeld, *Curr. Chem. Genomics*, 2008, **1**, 2–10.
16 E. F. Ullman, H. Kirakossian, A. C. Switchenko, J. Ishkanian, M. Ericson, C. A. Wartchow, M. Pirio, J. Pease, B. R. Irvin, S. Singh, R. Singh, R. Patel, A. Dafforn, D. Davalian, C. Skold, N. Kurn and D. B. Wagner, *Clin. Chem.*, 1996, **42**, 1518–1526.
17 A. Kawamura, A. Tumber, N. R. Rose, O. N. F. King, M. Daniel, U. Oppermann, T. D. Heightman and C. Schofield, *Anal. Biochem.*, 2010, **404**, 86–93.
18 Walport L. J., Structural and Functional Studies of Chromatin Modifying Enzymes, D.Phil Thesis, University of Oxford, 2013.
19 J. Schiettecatte, E. Anckaert and J. Smitz, *Adv. Immunoass. Technol.*, , DOI:10.5772/35797.
20 R. Hasegawa, R. Cabral, T. Hoshiya, K. Hakoi, T. Ogiso, P. Boonyaphiphat, T. Shirai and N. Ito, *Int. J. Cancer*, 1993, **54**, 489–493.
21 B. Toth, L. Wailcave, K. Patii, I. Schmeltz and D. Hoffmann, 1977, 3497–3500.
22 R. G. S. Brown, H. Kawaide, Y. Y. Yang, W. Rademacher and Y. Kamiya, *Physiol. Plant.*, 1997, **101**, 309–313.
23 W. Rademacher, *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 2000, **51**, 501–531.
24 C. C. Thinnes, K. S. England, A. Kawamura, R. Chowdhury, C. J. Schofield and R. J. Hopkinson, *Biochim. Biophys. Acta - Gene Regul. Mech.*, 2014, **1839**, 1416–1432.
25 R. P. Nowak, A. Tumber, C. Johansson, K. H. Che, P. Brennan, D. Owen and U. Oppermann, *Curr. Opin. Chem. Biol.*, 2016, **33**, 151–159.
26 W. Xu, J. D. Podoll, X. Dong, A. Tumber, U. Oppermann and X. Wang, *J Med Chem*, 2013, **56**, 5198–5202.
27 R. J. Klose, E. M. Kallin and Y. Zhang, *Nat. Rev. Genet.*, 2006, **7**, 715–727.
28 A. Zorzi, K. Deyle and C. Heinis, *Curr. Opin. Chem. Biol.*, 2017, **38**, 24–29.
29 J. Owens, B. Hughes and A. Mullard, *Nat. Rev. Drug Discov.*, , DOI:10.1038/nrd.2016.15.
30 B. Bhushan, A. Erdmann, Y. Zhang, R. Belle, C. Johannson, U. Oppermann, R. J. Hopkinson, C. J. Schofield and A. Kawamura, *Bioorganic Med. Chem.*, 2018, **26**, 2984–2991.

Chapter 5 Materials and Methods

All Chemicals, unless otherwise stated, were obtained from Sigma-Aldrich. Tryptone, yeast extract and agar were obtained from Oxoid. Sodium chloride for culture media was from Fisher BioReagents. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was from Apollo Scientific.

Standard molecular biology and biochemical procedures were applied thoughout the work. Water was purified using Millipore Elix® 10 followed by filtration through a 0.22 µm filter using a Millipore Milli-Q system. Cell growing flasks, media and plastic tubes were sterilised by autoclaving for 20 minutes at 121 ^{o}C where relevant. For water used in DNA applications and heat-senstive solutions, a further sterilisation step was performed manually using a 0.22 µm filter in addition to using sterile MilliQ water.

5.1 General Procedures

Competent cells

Table 5.1 defines the *E*. coli Strains used in this Thesis.

Table 5.1 *Escherichia coli* (*E.* coli) strains used in this Thesis.

Name	Genotype	Usage	Sourcing
XL10 Gold	Tet^{r} Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F′ proAB lacI^{q}ZΔM15 Tn10 (Tet^{r}) Amy Cam^{r}]	Plasmid transformation	Prepared in-house from commercial stock (Stratagene)*
BL21 (DE3)	*E. coli* B F^{-} dcm ompT hsdS (r_B^- m_B^-) gal λ(DE3)	Protein expression	Prepared in-house from commercial stock (Stratagene)*

*Protocol stated elsewhere.

Some of the developed experimental procedures described in this Chapter have been reported (see Chapters 2, Section 2.3.1; Chapter 3, Section 3.3.1); Other reported procedures, as developed previously and used in the experimental work described in this thesis, are referenced throughout.

Peptide sequences

The peptides used in this thesis were made in-house or obtained from GL BioChem Ltd, (Shanghai China) (Table 5.2). All peptides were prepared as C-terminal amides (Table 5.2).

Table 5.2 Sequences of the peptides used in this thesis.

Peptide names	Sequences
H3 $_{1-15}$ K4me3K9me2	ARTK(me3)QTARK(me2)STGGKA
H3 $_{1-15}$ K9me2	ARTKQTARK(me2)STGGKA
H3 $_{1-24}$ K4me3K14me2	ARTK(me3)QTARKSTGGK(me2)APRKQLATKA
H3 $_{1-24}$ K14me2	ARTKQTARKSTGGK(me2)APRKQLATKA
H3 $_{1-24}$ K4me3K18me2	ARTK(me3)QTARKSTGGKAPRK(me2)QLATKA
H3 $_{1-24}$ K18me2	ARTKQTARKSTGGKAPRK(me2)QLATKA
H3 $_{1-29}$ K4me3K23me2	ARTK(me3)QTARKSTGGKAPRKQLATK(me2)AARKSA
H3 $_{1-29}$ K23me2	ARTKQTARKSTGGKAPRKQLATK(me2)AARKSA
H3 $_{1-35}$ K4me3K27me2	ARTK(me3)QTARKSTGGKAPRKQLATKAARK(me2)SAPATGGV
H3 $_{21-44}$ K27me2	ATKAARK(me2)SAPATGGVKKPHRYRPG

Peptide synthesis

Some peptides were synthesised in-house by solid phase peptide synthesis (SPPS) using standard procedures (i.e. from their C-termini) using a LibertyBlue™ mircrowave peptide synthesiser (CEM). Fluorenylmethyloxycarbonyl (FMOC) protected proteinogenic amino acids were purchased from Alfa Aesar, Novabiochem, Sigma-Aldrich, Merck, TCI Europe, or AGTC Bioproducts and were used in 10 equivalent excess relative to the enzyme (except for Kme3 and Kme2 residues). FMOC-Rink-Amid-methylbenzhydrylamine resin (100-200 mesh, > 0.5 mmol/g) and N^{α}-FMOC-amino acids were dissolved in dimethyformamide (DMF) (peptide synthesis grade, AGTC Bioproducts) prior to loading onto a peptide synthesiser. *N*, *N*'-diisopropylcarbodiimide (DIC, purchased from TCI Europe) and Oxyma Pure (Merck) were used as coupling reagents to activate N^{α}-FMOC-amino acids before addition to the resin.[1] N^{α}-FMOC deprotection was achieved by treating the resin with 20% peperidine in DMF. The process was then repeated for each amino acid required for the desired sequence. Peptide products were cleaved from the resin by incubating with the Peptide Cleaving Solution (Table 5.3) for 3 hours. The peptides were then precipitated with diethyl ether from the mixture. The resultant solid was dissolved in MilliQ water and lyophilised until dry.[2,3] Crude peptides were purified by HPLC (see the next section for method). The purity and mass of each peptide were confirmed by MALDI-TOF MS.

Table 5.3 Peptide Cleaving Solution.

Reagent	Percentage in mixture (%)	Volume (mL)
TFA (trifluoroacetic acid)	92.5	4.625
Water	2.5	0.125
TIPS (triisopropylsilane)	2.5	0.125
1,3 dimethoxybenzene	2.5	0.125

Peptide purification

Peptides were purified *via* Reverse Phase High Performance Liquid Chromatography (HPLC) using a JASCO LC Net II / ADC HPLC system (PU-2086 Plus Pumps, MX-2080-32 Dynamic Mixer and DG-2080-54 Degasser). The HPLC system was controlled with JASCO ChromNAC Control Centre software. Crude peptides were dissolved in 2 mL MilliQ water and up to 30% MeCN to aid solubility, when required. The crude peptide solution was filtered through a 0.22 µM mini filter to remove solids, so protecting the HPLC column from being blocked by insoluble particles. The resultant clear peptide solution (50 µL) was loaded on a Grace/Vydac 218GCC1520 column (C18, 10-15 µM, Prep Guard Cartridge 10 × 20 mm) for a trial purification. The mobile phase was a gradient of Solution A (0.1% TFA in H_2O) + Solution B (0.1% TFA in MeCN), with the percentage of Solution B increasing over time. Peaks were detected at 220 nm and collected automatically.[4] Sample injection and fraction collection were conducted using a JASCO AS-2055 Plus autosampler and a CHF1225C fraction collector. Fractions containing purified peptide were pooled and lyophilised in 50 mL Falcon tubes. The dried fractions were analysed by MALDI MS to confirm their predicted masses.

pH measurements

The pH of solutions was determined using a Jenway pH Meter 3305, with Aldrich glass/calomel combination electrode. Prior to measurements, calibration was carried out using standard buffer solutions of phthalate (pH 4.0), phosphate (pH 7.0) and borate (pH 10.0) from Fisher Scientific. Electrodes were stored in a 4 M potassium chloride solution at room temperature.

Centrifugation

Samples of volume less than 1.5 mL were centrifuged using a Bechman Coulter Ltd. benchtop centrifuge, at 4 ^{o}C or 20 ^{o}C (speed varied from 1000 rpm to 14,000 rpm). Samples of volumes from 2 mL to 15 mL, including protein solutions, were centrifuged and concentrated using a Beckmann Allegra X-30R at room temperature or 4 ^{o}C (22,000 rpm). Cell lysates for protein purification and large-scale cell growth cultures were centrifuged at 4 ^{o}C using a Beckmann Avanti J-25 at 4 ^{o}C (8000 rpm).

5.2 DNA constructs generation

Oligonucleotides

Oligonucleotides (or primers) were from Sigma Life Science and dissolved in sterile MilliQ water to a concentration of 100 μM. The sequences of the primers used in this thesis are given in Table 5.4. The PHF8 $_{PHD\text{-}JmjC}$ and PHF8 $_{JmjC}$ plasmids were kind gifts from the Structural Genomics Consortium (SGC).[5] The constructs of PHF8 $_{PHD}$ and PHF2 $_{PHD}$ were generated and gifted by Dr. Finn Wolfreys.

Table 5.4 Primer sequences. All DNA bases are shown in capitals, except for ligation independent cloning (LIC) overhangs. Base changes from the wild-type DNA sequences are shown in **bold**.

Construct	Name	Sequence
KIAA1718 $_{PHD-JmjC}$	Fwd	tacttccaatccatgGTGTACTGTGTGTGC
	Rev	tatccacctttactgtcaGCCGTTTTCCTCCTC
KIAA1718 $_{PHD}$	Fwd	tacttccaatccatgGTGTACTGTGTGTGC
	Rev	tatccacctttactgtcaCCTCCTTTTTTTCATCAA
PHF2 $_{PHD-JmjC}$	Fwd	tacttccaatccatgGCGACGGTGCC
	Rev	tatccacctttactgtcaGACGGCTTTGGAGGCATTCTCACTGAG
PHF2 $_{JmjC}$	Fwd	tacttccaatccatgTCCACCTTAAAGAAGAA
	Rev	tatccacctttactgtcaGACGGCTTTGGAGGCATTCTCACTGAG
PHF2Y321H	Fwd	CAGGCTGGATC**CAC**GCCACACTCACC
	Rev	TGAGTGTGGC**GTG**GATCCAGCCTGA
PHF2H249A	Fwd	TTACACCGACTTC**GCC**ATCGACTCTGG
	Rev	AGAGTCGAT**GGC**GAAGTCGGTGTAAC
PHF8H319Y	Fwd	CACAGGGTGGATC**TAT**GCTGTGCTGAC
	Rev	GTCAGCACAGC**ATA**GATCCACCCTTGTG
PHF8Y48C	Fwd	CATTGACCTC**TGC**CACTGCCCC
	Rev	GGGGCAGTG**GCA**GAGGTCAATG
PHF8Y48H	Fwd	CATTGACCTC**CAC**CACTGCCCC
	Rev	GGGGCAGTG**GTG**GAGGTCAATG
PHF8Y48A	Fwd	CATTGACCTC**GCC**CACTGCCCC
	Rev	GGGGCAGTG**GGC**GAGGTCAATG

DNA concentration determination

DNA concentrations were determined using a NanoDrop ND-1000 UV/Vis spectrophotometer (NanoDrop Technologies lnc.) by measuring their absorbance at 260 nm. The purity of DNA was assessed by reference to the A260/A280 ratio.[6]

Protein concentration determination

Protein concentrations were determined using 2 µL protein solution and a NanoDrop ND-1000 UV/Vis spectrophotometer (NanoDrop Technologies lnc.) measuring the absorbance at 280 nm. Molecular weight and extinction coefficients (ε) were calculated using the ProtParam tool on the ExPASY Proteomics Server (www.expasy.org).

Plasmid DNA purification

Plasmid DNA was purified from 5 mL (15 mL for low-copy constructs) cell culture using a GeneJET Plasmid Miniprep Kit (Thermo Scientific) following the manufacturer's protocol. In the final step, DNA was eluted with 30 µL sterile MilliQ water instead of 50 µL elution buffer, as described in the manufacturer's protocol, to increase the DNA concentration. A GeneJET Plasmid Maxiprep Kit (Thermo Scientific) was used for large-scale transfection-grade DNA purification.

DNA amplification by polymerase chain reaction (PCR)

DNA plasmids were amplified by the polymerase chain reaction (PCR). A typical reaction mixture contained the following reagents in 50 µL (total):

Table 5.5 PCR reaction mixture.

Reagent	Stock	Volume (µL)
***Q5*® reaction buffer (NEB)**	10 x	5
Forward primer (Sigma)	10 µM	1.25
Reverse primer (Sigma)	10 µM	1.25
DNA template	100 ng/µL	1
dNTP mix (NEB)	10 mM	3
***Q5*® DNA polymerase (NEB)**	2.5 U/ µL	0.5
MilliQ water	-	38
Total	-	50

A Techne Genius thermal cycler (Techne) was programmed according to the temperature requirement at each stage of the desired PCR reactions, the T_m of primers and the length of target DNA. An online T_m Calculator tool was used to assist temperature determination.[7]

A general programme for PCR involving plasmids is shown in Table 5.5. Optimisations were applied according to reaction efficiency and product purity analysed by agarose gel electrophoresis.

Table 5.6 Typical PCR programme, giving cycles and temperature.

Segment	Cycles	Temperature	Time
Initial denaturation	1	98 °C	1 min
Denaturation	12	98 °C	30 s
Annealing		5 °C below the T_m of primers	30 s
Extension		72 °C	1 min per kb of plasmid
Hold	1	4 °C	Hold

Products from PCR reactions were stored at -20 °C (if not used immediately) before being purified using a GeneJet PCR Purification Kit (Thermo Scientific) at room temperature. 30 µL of MilliQ water replaced the elution buffer in the purification kit to elute the final plasmid product.

Ligation independent cloning (LIC-cloning)

LIC cloning provides a framework for cloning the same PCR product into multiple different vectors when LIC compatible adaptor sequences are included at either ends of the PCR product.[8] LIC cloning makes use of the exonuclease activity of T4 DNA polymerase to generate specific 10 - 15 base single overhangs in the expression vector. The annealing of inserts and vectors with overhangs is independent of ligase and can be achieved by simple mixing of the DNA fragments, which is more efficient than the cloning using restriction enzymes.[8] The LIC compatible vectors used in this Thesis contain a SacB gene which allows sucrose negative selection for the insertion of one of the gene flanks.[9] The DNA sequences

of interest were amplified from template plasmids using PCR. The protocol of DNA amplification for LIC cloning was based on the general PCR protocol described above with minor optimisations.[10]

Insert - PCR recipe:

Table 5.7 PCR reaction mix for insert DNA.

Reagent	Volume/µl
DNA template (100ng/µl)	10
Primers mix (Sigma, 10µM each)	2.5
dNTPs (NEB, 10mM)	1
Q5® reaction buffer (5 X)	10
Q5® DNA polymerase (2000U/ml)	0.5
Sterile MiliQ water	26

Insert - PCR cycles and temperatures

Table 5.8 PCR programme for insert DNA, giving cycles and temperature.

Entry	Temperature/°C	Time
1	98	2 min
2*	98	10 s
3*	60	20 s
4*	72	2 min
5	72	2 min
6	4	Hold

*Entry 2,3,4 were repeated 40 times as a cycle.

Insert - gel electrophoresis

The presence of the desired PCR product of insert DNA was confirmed by agarose gel electrophoresis. The bands where a single PCR product was observed on the agarose gel and within the desired molecular weight range were cut from the gel and purified using a GeneJET Gel Extraction kit (Thermo Scientific), according to the manufacturer's protocol before further use.

Vector

The LIC compatible vector DNA (see Table 5.19) was digested with the BsaI restriction enzyme in NEB buffer 3 with 0.1 µg BSA at room temperature for 2 hours and then at 50 ^{o}C for 3 hours (Table 5.9).[10] The linearised vector DNA was purified using a GeneJET PCR Purification kit (Thermo Scientific). The purified linearised vector DNA was incubated with the T4 DNA polymerase (Novagen) and dGTPs at 22 °C for 30 min to generate sticky ends before heat inactivation of the polymerase at 75 °C for 20 min (Table 5.10, left column). At the same time, the inserted DNA purified using a GeneJET Gel Extraction kit (Thermo Scientific) was treated with dCTP instead of dGTP to produce complementary sticky ends (Table 5.10, right column). 2 µL of the insert DNA was incubated with 1 µL of the linear vector DNA for 20 min at room temperature and the mixture was transformed into competent *E. coli* cells and plated onto agar plates containing 5% sucrose in addition to the required antibiotic to select for clones that had been correctly linearized (and thus had lost the insert that makes them susceptible to sucrose).

Digestion of vector DNA

Table 5.9 Reaction mixture for vector digestion.

Reagent	Amount
DNA template	5 µg
BsaI	3 µL/3µg
Reaction buffer 3 (NEB)	10 µL
BSA	1 µL
Sterile MilliQ	Up to 100 µL

Generation of cohesive ends

Table 5.10 Reaction mixture for generating cohesive ends.

Vector		**Insert**	
Reagent	**Volume/µL**	**Reagent**	**Volume/µL**
Sterile MilliQ	29	Sterile MilliQ	1.15
Digested plasmid	50	Insert DNA	5
Reaction buffer2 (Novagen)	10	Reaction buffer2 (Novagen)	1
dGTP (100mM)	2.5	dCTP (100mM)	2.75
BSA (10mg/ml)	1	BSA (10mg/ml)	0.1
DTT (100mM)	5	DTT (100mM)	0.5
T4 DNA polymerase, LIC qualified (Novagen)	2.5	T4 DNA polymerase, LIC qualified (Novagen)	0.25

*mixture was left at room temperature (22 ℃) for 30 min, then 75 ℃ for 20 min.

Transformation procedure of LIC

Transformation procedure of LIC follows a standard transformation protocol.[10] A plate with 5% sucrose in addition to antibiotics was used to select the colonies containing the desired cloned product.

Site-directed mutagenesis (SDM)

The protocol for site-directed mutagenesis was optimised by Dr Md. Saiful Islam based on the standard protocol established in the research group.[1] The PCR recipe of SDM is shown in Table 5.11 and the PCR programme is shown in Table 5.12.

Table 5.11 PCR mixture for SDM reactions.

Reagent	Volume or Amount
dsDNA template	20 – 40 ng
Primer (Sigma, 10 µM stock)	2.5 µL (Fwd 1.25 µL, Rev 1.26 µL)
dNTPs (NEB, 10 µM stock)	2.5 µL
Q5® Reaction Buffer	5 µL
Q5® High GC Enhancer	5 µL
Q5® DNA Polymerase	0.5 µL
Sterile MilliQ	Up to 25 µL

PCR programme for SDM:

Table 5.12 PCR programme for SDM, giving time, cycles and temperature.

Entry	Temperature/°C	Time
1	94	2 min
2*	94	45 s
3*	55	1 min
4*	68	20 min
5	68	20 min
6	4	Hold

*Entries 2, 3 and 4 were repeated 19 times as a cycle.

Following the PCR reaction, 1 µL of DpnI (New England Biolabs) was added to the PCR product to digest any methylated DNA (i.e. original template DNA). The digest was conducted for 3 hours at 37 °C. 2 µL of the resultant mixture was transformed into 30 µL of XL10 Gold competent cells and grew on LB agar plates treated with appropriate antibiotics. Mutations were confirmed by sequencing the DNA extracted from selected colonies.

Agarose gel electrophoresis

Reagents:

50 x DNA running buffer (a.k.a. TAE buffer):

Table 5.13 Reagents for making 100 mL of 50 × TAE buffer.[1]

Reagent	Amount
Tris-Base	24.2 g
Glacial acetic acid	5.71 mL
EDTA	1.86 g

5 x DNA sample loading buffer (5 mL):

Table 5.14 Reagents for making 5 mL of 5 × DNA sample loading buffer.[1]

Reagent	Amount
TAE buffer (1 x)	1.5 mL
Glycerol	3.5 mL
Bromophenol blue (0.25% (w/v))	12.5 mg

Agarose gel for DNA visualisation

1% (w/v) agarose gels were made from 0.4 g agarose (Bioline) in 40 mL running buffer with heating for a few minutes in a SuperWave 750 microwave (Tecnolec) until the crystals dissolved. The resultant mixture was allowed to cool to around 60 °C, before addition of 4 µl SYBR SafeTM DNA gel stain (x10,000) (Invitrogen). Gels were cast in an H2-SET gel tray (Anachem) to a depth of 0.7 cm with a comb inserted at one end of the gel to enable sample loading and allowed to cool at room temperature for the gel to set. Gels were run on a Bio-Rad system at a constant potential of 80 V for approximately 45 min. Bands were visualised by UV light illumination using a Gel Logic 200 imaging system (Kodak) and a 535 nm emission filter. Samples were run against a 2-log DNA Ladder 0.1-10 kb (Invitrogen) allowing molecular weight determination. [11]

SDS-PAGE gel electrophoresis

Reagents:

Table 5.15 Reagents for making two pieces of SDS-PAGE gels.[1]

Reagent	Resolving gel (18%/15 %/12 %)	Stacking gel
Tris-HCl (1 M, pH 7.5)	1.25 mL	-
Tris-HCl (0.5 M, pH 6.8)	-	0.63 mL
Milli-Q water	0.65 mL / 1.15 mL / 1.65 mL	1.57 mL
30%(w/v) acrylamide/1.034%(w/v) bis-acrylamide stock	3 mL / 2.5 mL / 2 mL	0.25 mL
SDS (10% (w/v))	50 µL	25 µL
10% (w/v) APS	50 µL	25 µL
TEMED	4 µL	2.5 µL

Tetramethylethylenediamine (TEMED) and ammonium persulfate (APS) were added just prior to gel casting. The resolving gel % was determined based on the size of the proteins to be separated.

10 x SDS-PAGE gel running buffer

Table 5.16 Reagents for making 1 L of 10 x SDS-PAGE gel running buffer.

Reagent	Amount /g
Tris-Base	30
Glycine	144
SDS	10

SDS-PAGE gel destaining buffer

Table 5.17 Reagents for making 1 L of SDS-PAGE gel destaining buffer.

Reagent	Amount
Acetic acid	10% (w/v)
Methanol	30% (w/v)

SDS-PAGE gel staining buffer

The mixture used for SDS-PAGE gel staining buffer was the same as the gel destaining buffer with the addition of 0.25% (w/v) Coomassie® Brilliant Blue Dye.

SDS-PAGE sample loading buffer 100 mL

Table 5.18 Reagents for making 100 mL of SDS-PAGE sample loading buffer.

Reagent	Amount
Tris-HCl (0.5 M, pH 6.8)	10 mL
Bromophenol blue (0.2% (w/v))	0.2 g
SDS (10% (w/v))	20 mL
Glycerol	12 mL
β-mercaptoethanol	5 mL

SDS-PAGE gel electrophoresis

Samples were heated for 5 min at 100 °C before use. Polyacrylamide gels were run under denaturing conditions for 40-60 min at RT, 180 V using a Bio-Rad Mini-PROTEAN II system. Gels were stained with SDS-PAGE gel stain for 20 min and then destained with SDS-PAGE destain overnight.

5.3 Recombinant enzyme gene expression in *E. coli*

Large-scale incubations were carried out in a New Brunswick Scientific Innova 44 incubator at 37 °C with 200 rpm shaking. Small-scale incubations (100 mL or less) were carried out in a New Brunswick Series 25 incubator.

Vectors

Table 5.19 Vectors used in this Thesis and the proteins that they were used to express. Vectors marked with asterisks are LIC compatible vectors.[12–14]

Vector name	Antibiotics resistance	Tag	Tag cleavage site	Promotor	Proteins
pET28a	Kanamycin	N-terminal 6×His tag	Thrombin	T7	TEV
pNIC28-BSA4*	Kanamycin	N-terminal 6×His tag	TEV	T7	PHF8/KIAA1718
pNH-TrxT*	Kanamycin	N-terminal 6×His-TrxT tag	TEV	T7	PHF8/KIAA1718
pNIC-ZB*	Kanamycin	N-terminal 6×His-ZB tag	TEV	T7	PHF2

Antibiotics

Stock solutions (x1000) of antibiotics were made in MilliQ water and sterile filtered and stored at -20°C. Antibiotics were used at a final concentration of 100 µg/mL of Ampicillin or 30 µg/mL of Kanamycin.

2× Tryptone-Yeast (2-TY) media (1 L)

2-TY media was prepared by Mr H. M. Jubier.

Table 5.20 Reagents for making 1 L of 2-TY media.

Reagent	Amount / g
Bacto-tryptone	16
Yeast extract	10
NaCl	5

To produce solid medium for plates, 20 g agar was added to the media before autoclaving. The autoclaved media were mixed with desired antibiotics when cooled to 55 – 60 °C and pooled into clear sterile plates to set in 4 °C.

Terrific Broth (TB) media (1 L)

Table 5.21 Reagents for making 1 L of Terrific Broth.

Reagent	Amount / mL
2-TY media	900
0.17 M KH_2PO_4, 0.72 M K_2HPO_4	10

Optical density measurements (OD_{600})

OD_{600} values were measured in a 1.6 mL cuvette, against a reference sample of initial growth medium in a Novaspec® II spectrophotometer (Pharmacia).

Standard DNA plasmid transformations

XL10-Gold ultracompetent cells were used for production of DNA and BL21 (DE3) cells for bacterial expression of His-tagged proteins. Competent cells were thawed on ice. 2 µL of plasmid DNA (100 ng/ µL) was added to 25 µL cells in a pre-chilled sterile 1.5 mL Eppendorf tube. Tubes were left on ice for 20 minutes before the cells were transformed by heat shock in a water bath at 42 °C for 30 s. Cells were left on ice for a further 1 minute before 0.4 mL 2-TY media was added. The mixture was incubated for 1 hr at 37 °C with shaking. 250 µL of the mixture was plated onto agar plates containing kanamycin or ampicillin. For LIC cloning plates also contained 5 % sucrose. The plates were incubated at 37 °C overnight before use.[1]

Production of competent cells

Competent cells were prepared from commercial stocks (Stratagene). Bacteria were streaked onto antibiotic-free agar plates and grown overnight. A single colony from the freshly-prepared plates was used to inoculate 100 mL sterile antibiotic-free 2TY medium and grown overnight (37 °C, 220 rpm). Four 100 mL 2TY cultures were inoculated with 200 µL starter culture and the medium grown (37 °C, 220 rpm) to an OD600 of 0.6. Pre-chilled centrifuge tubes were used to pellet cells (750 × g, 5 min, 4 °C). Pellets were resuspended in 200 mL ice-cold 100 mL $CaCl_2$, 20 % glycerol and incubated on ice for 20 min. Cells were pelleted (750 × g, 5 min, 4 °C) and resuspended in 8 mL of the same solution. 50 µL aliquots of cells were flash frozen in liquid nitrogen and stored at -80 °C for future use.[11]

Starter cultures

100 mL 2-TY media containing the appropriate antibiotic (kanamycin or ampicillin) was inoculated with a single colony from an agar plate and then incubated overnight in a New Brunswick Scientific G25 environmental shaker at 37 °C and 250 rpm shaking.

Recombinant enzyme expression trials

500 mL flasks containing 100 mL 2-TY media were inoculated with the appropriate antibiotic (kanamycin or ampicillin) and 1 mL starter culture. At OD_{600} ~ 0.6, cultures were separated into incubators held at a variety of temperatures (typically 18 °C, 30 °C and 37 °C) and each induced with variable concentrations of IPTG (typically 0.5 mM or 0.1 mM). Cultures were incubated for four hours at 37 °C or at 18 °C and 30 °C overnight. 1.5 mL of each culture was transferred to a 1.5 mL Eppendorf tube and pelleted using a Bechman Coulter Ltd. benchtop centrifuge (4 °C, 14,000 rpm, 5 min). The media was removed and the cells resuspended in 300 µL binding buffer. 30 µL was removed for analysis as whole cell extract and the remainder was lysed by sonication (2 × 15 s with 20 s interval) on ice. Cell debris was removed by centrifugation (22,000 rpm, 4 °C, 30 min) and the supernatant removed for analysis as cell lysate. Soluble expression levels were assessed by SDS-PAGE gel electrophoresis using 15 µL samples of the whole cell extract and cell lysate.

Large-scale growths using *E. coli*

12 × 2 L flasks containing 600 mL 2TY media and the appropriate antibiotic (kanamycin or ampicillin) were each inoculated with 6 mL starter culture and grown at 37 °C with shaking (220 rpm). At OD_{600} ~ 0.8 – 1.2, recombinant protein production was induced using 0.1 mM of Isopropyl β-D-1-thiogalactopyranoside (IPTG) final (Melford Laboratories Ltd), using

optimal induction conditions as determined during expression trials. Cultures were incubated overnight at 18 °C. Cells were then pelleted by centrifugation (8000 rpm, 4 °C, 10 min) and stored in a resealable plastic bag at –80 °C.

5.4 Recombinant enzyme purification

Binding buffer (see below) was added according to the weight of pellet (5 mL/mg). The suspension was sonicated (3s with 4s interval, 4 min in total) followed by centrifugation (22,000 rpm, 4 °C, 30 min), as for procedure used during expression trials. The supernatant was filtered with a 0.45 µM filter prior to column purification. All purifications were carried out by two-step purification module – either a Histrap™ column or an ion exchange column followed by a size exclusion column in both cases.[5,15]

5.4.1 Immobilised Metal Affinity Chromatography (IMAC) - Histrap columns

Buffer

Table 5.22 Components for Histrap column buffers.

Reagents	Charge	Binding	Wash	Elution	Strip
HEPES	-	50 mM	50 mM	50 mM	-
NaCl	-	500 mM	500 mM	500 mM	-
Imidazole	-	20 mM	40 mM	250 mM	-
$NiSO_4$	500 mM	-	-	-	-
EDTA	-	-	-	-	100 mM
Glycerol	-	5%	5%	5%	-

The lysate was loaded onto a pre-charged Histrap™ FF Crude 5 mL column (GE Healthcare) using an ÄKTA FPLC system (GE Healthcare). The column was then washed with Histrap binding buffer and Histrap wash buffer until the UV absorption stabilised (approximately 15 column volumes of binding buffer and 5 column volumes of wash buffer). Elution buffer was applied for two column volumes. The eluted fractions were treated with 1 mM EDTA (3 mM for PHF8 wild-type) and TEV protease (m:m= 1:50) over 12-16 hours before loading on the size exclusion column.

5.4.2 Ion Exchange Chromatography (IEX) – Sulphopropyl Sepharose (SP Seph) Columns

Buffer

Table 5.23 Components for IEX column buffers.

Reagents	HEPES	NaCl	TCEP	EDTA	glycerol
A	100 mM	50 mM	1 mM	1 mM	5%
B	100 mM	1 M	1 mM	1 mM	5%

The lysates were loaded on an SP Seph column (50 mL) using ÄKTA FPLC system (GE Healthcare). A linear washing protocol from 0% to 100% IEX buffer B through 750 mL was applied. Fractions contained the enzyme of interest were treated by TEV protease (w/w = 1/50) over 12 hours at 4 °C before loading onto the size exclusion column.

5.4.3. Size Exclusion (SEC) Chromatography

The SEC column chromatography buffer contained HEPES 50 mM pH 7.0, NaCl 200 mM and TECP 1 mM. All buffers, including cleaning buffer (100 mM NaOH), MilliQ water and running buffer, were degassed for at least 30 minutes prior to loading onto SEC columns.

The fractions from Histrap or IEX column were concentrated to 2 mL or 10 mL depending on the amount of crude protein present before loading onto a S75 column (Packed by Dr Adam Hardy). One column volume of SEC buffer was applied to elute protein. The buffer was run at a flow-rate of 1 mL/min to protect the column from high pressure and to achieve best separation of protein and contaminants.

5.5 Activity Assays

Demethylase activity assays using MALDI-TOF mass spectrometry

Single concentration demethylation assays were carried out in 1.5 mL microcentrifuge tubes. The Reaction Buffer contained 50mM HEPES, pH 7.0 and 50 mM NaCl. Each reaction contained 5 µM enzyme, 10 µM H3 peptide, 200 µM 2-oxoglutarate (2OG, acid form), 1 mM (+)-sodium-L-ascorbate and 50 µM ammonium iron (II) sulfate hexahydrate ($(NH4)_2Fe(SO4)_2$). Reaction was incubated at 37 ℃ and quenched with methanol (v /v = 1/1) at the desired time points. 1 µL quenched solution was mixed with 1 µL of 10 mg/mL α-cyano-4-hydroxycinnamic acid (CHCA) solution in 50% acetonitrile (v/v) and 0.1% TFA (v/v) on a MALDI plate. Each 96-well MALDI plate was analysed using positive reflection mode using a Water MALDI Micro MX instrument. All reagents were freshly made prior to the experiments.[16]

For K_M value determinations, assays were carried out in 96-well clear PCR plates. A shaker that was thermally controlled (ThermoFisher Scientific) was used to keep the reactions at 37 °C. Each reaction contained 2.5 µM enzyme, various concentration of peptide (for K_M of substrate), various concentration of 2OG (for K_M of 2OG), 1 mM (+)-sodium-L-ascorbate, 50 µM $(NH4)_2Fe(SO4)_2$.[1] All reagents were pre-incubated to the same temperature before mixing to minimise changes in temperature.

KDM7 peptide turnover assay using AlphaScreen

Assays using the AlphaScreen method were carried out in 384-well white ProxiPlates (PerkinElmer) at 22 °C. 3 nM KDM7A (or 10 nM KDM7B) was incubated with test compounds dissolved in dimethyl sulfoxide (DMSO) for 20 minutes. A mixture of 100 µM (+)-sodium-L-ascorbate, 10 µM $(NH4)_2Fe(SO4)_2$, 20 µM 2OG (acid form), and 30 nM $H3_{1–21}$K9me2-biotin peptide (300 nM $H3_{1–21}$K4me3K9me2-biotin for KDM7B) in 50 mM HEPES buffer (pH 7.0) containing 0.01% (v/v) Tween 20 and 0.1% (w/v) bovine serum albumin (BSA) (final concentrations in a 10 µL reaction). Reactions were quenched by the addition of 30 mM ethylenediaminetetraacetic acid (EDTA) after 10 minutes. Streptavidin-conjugated donor and protein-A-conjugated acceptor beads were pre-incubated with monoclonal antibody (Ab8896, Abcam), and added into the quenched reaction. The final bead concentration was 20 µg/mL and the antibody concentration was 0.05 µg/mL, in a 20 µL final reaction volume. The plate was analysed after 1-hour incubation at dark, using an Envision (PerkinElmer) plate reader.[17]

Fluorescence based demethylation assay

A formaldehyde dehydrogenase (FDH) – coupled continuous fluorescence based assay - was also used to determine the K_M values for KDM7 enzymes. Reactions were carried out at room temperature and at 37 °C in 50 mM HEPES buffer (pH 7.0), containing 50 µM $(NH_4)2FeSO_4$, 1 mM 2OG (acid form) and 2 mM (+)-sodium-L-ascorbate, 1 mM β-nicotinamide adenine dinucleotide (NAD^+) and 0.05 units of FDH (gift from SGC). KDM7 enzyme was pre-incubated in the reaction mix (without peptide) for 30 minutes at room temperature. The desired amount of peptide was added into the reaction mixture to initiate the FDH assay reaction. Fluorescence was monitored (355 nm excitation, 460 nm emission) every 30 seconds using a PheraStar FX plate reader (BMG).[16,18]

Fluorescence polarisation (FP) binding assay

Assays were carried out in 96-well, black, low-binding microplate (Greiner Bio-One) at 22 °C. Reaction buffer contained 50 mM HEPES, pH 7.0, 50 mM NaCl, 0.1% Tween-20. Each reaction contained fluorescent labeled molecule at 50 nM ($H3_{1-15}$K4me3K9me2-FITC or $H3_{1-15}$K9me2-FITC) or 12.5 nM (methylstat-FI), $NiSO_4$ 100 µM, 2-fold serial dilution of enzyme from 50 µM across the plate. The reactions were incubated in the dark (r.t.) for 20 minutes and analysed by PHERAstar FX plate reader (BMG).[19] Data was fitted into an one-site specific binding fitting model using GraphPad Prism 7 software (Equation 5.1).

$$Y = \frac{B_{max} \times X}{K_d + X}$$

Equation 5.1 The one-site specific binding fitting model built in GraphPad Prism 7.

Fluorescence polarisation (FP) displacement assay

Assays were carried out in 384-well, black, non-binding microplate (Greiner Bio-One) at 22 ^{o}C. KDM7s (150 nM KDM7A, 600 nM KDM7B or 50 nM KDM7C) were incubated with 12.5 nM FI-methylstat or 50 nM $H3_{1\text{-}15}$K4me3K9me2-FITC, 100 µM $NiSO_4$, and compounds at various concentrations in 50 mM HEPES buffer (pH7.0), containing 50 mM NaCl and 0.1% Tween-20. Enzyme was not pre-incubated with compounds to avoid stabilising effect. The plate was incubated in the dark for 30 minutes and analysed using a PHERAstar FX plate reader (BMG).[19] Data was fitted into an dose-response inhibition log(inhibitor) vs. response fitting model using GraphPad Prism 7 software (Equation 5.2).

$$Y = Bottom + \frac{Top - Bottom}{1 + 10^{(X - LogIC50)}}$$

Equation 5.2 The log(inhibitor) vs. response fitting model built in GraphPad Prism 7.

Isothermal titration calorimetry (ITC) binding assays

ITC measurements were carried out using a MicroCal PEAQ-ITC machine (Malvern). Before loading the samples, the protein was buffer exchanged against reaction buffer which contained 20 mM HEPES, pH 7.5, NaCl 200 mM (500 mM for mutants), 1 mM TCEP. Ligands including metal ion and peptide were dissolved in the same buffer to ensure matched buffers. KDM7 enzymes were applied to the reference cell and a 40 µL syringe filled with titrant. Aliquots of titrant were injected into the cell on each injection programmed by the controlling software provided by the manufacturer.[20]

Differential scanning fluorimetry (DSF) assay

Differential scanning fluorimetry (DSF) assays were carried out using a MiniOpticon Real-Time PCR System (Bio-Rad) to determine the melting temperature (T_m) and T_m shift values of KDM7 enzymes. The protocol was adapted from the work of Niesen *et al.*[21] Apo enzymes were loaded in triplicate into a white 96-well PCR plate (Bio-Rad) in combination with metal ion solutions and SYPRO Orange (Invitrogen). Each reaction contained 2 µM of apo enzyme, 50 µM of metal ion and 1X SYPRO Orange in 50 mM HEPES buffer (pH 7.0). The samples were heated from 25 °C to 95 °C at the rate of 0.5 °C/min. A single fluorescence measurement was taken each minute. The system was monitored using an excitation at λ = 492 nm and emission at λ = 610 nm. The melting curve was fitted into the Boltzmann equation (Equation 5.1). The T_m shift value was calculated by comparing the Tm of KDM7s in the presence of metal ion against that of apo enzyme.

$$y = LL + \frac{(UL - LL)}{1 + \exp(\frac{T_m - x}{a})}$$

Equation 5.3 The Boltzmann equation. LL and UL are the values of minimum and maximum intensities, respectively, and a denotes the slope of the curve within T_m.[21]

NMR Spectroscopy

With the help from Dr Richard Hopkinson, NMR assays were performed using a Bruker Avance AV111 700 MHz spectrometer equipped with an inverse TCI CryoProbe, optimised for ^1H observation, and installed with Topspin 2 software. Samples were made up in deuterated ammonium formate buffer. Experiments were carried out in 2 mm MATCH NMR tubes (Wilgenberg) containing 5 µM enzyme, 100 µM ammonium iron (II) sulfate hexahydrate, 100 µM, 2-oxoglutarate (2OG) and 100 µM (+)-sodium L-ascorbate. The solvent deuterium signal was used as an internal lock signal.

5.6 References

1 Walport L. J., Structural and Functional Studies of Chromatin Modifying Enzymes, D.Phil Thesis, University of Oxford, 2013.

2 W. Ge, A. Wolf, T. Feng, C. H. Ho, R. Sekirnik, A. Zayer, N. Granatino, M. E. Cockman, C. Loenarz, N. D. Loik, A. P. Hardy, T. D. W. Claridge, R. B. Hamed, R. Chowdhury, L. Gong, C. V. Robinson, D. C. Trudgian, M. Jiang, M. M. MacKeen, J. S. McCullagh, Y. Gordiyenko, A. Thalhammer, A. Yamamoto, M. Yang, P. Liu-Yi, Z. Zhang, M. Schmidt-Zachmann, B. M. Kessler, P. J. Ratcliffe, G. M. Preston, M. L. Coleman and C. J. Schofield, *Nat. Chem. Biol.*, 2012, **8**, 960–962.

3 L. J. Walport, R. J. Hopkinson, R. Chowdhury, R. Schiller, W. Ge, A. Kawamura and C. J. Schofield, *Nat. Commun.*, 2016, **7**, 1–12.

4 Hancock R. L., Investigating the Effect of Hypoxia on the JmjC Histone Lysine Demethylase, D.Phil Thesis, University of Oxford, 2016.

5 N. R. Rose, S. S. Ng, J. Mecinović, B. M. R. Liénard, S. H. Bello, Z. Sun, M. A. McDonough, U. Oppermann and C. J. Schofield, *J. Med. Chem.*, 2008, **51**, 7053–7056.

6 260/280 and 260/230 Ratios, Manual for Thermo Scientific Nanodrop Spectrophotometer, http://www.nhm.ac.uk/content/dam/nhmwww/our-science/dpts-facilities-staff/Coreresearchlabs/nanodrop.pdf.

7 Tm Calculator, ThermoFisher, https://www.thermofisher.com/uk/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/tm-calculator.html.

8 Arie Geerlof, *EMBL - Hambg.*

9 R. Balhana, N. G. Stoker, M. H. Sikder, F. X. Chauviac and S. L. Kendall, *J. Microbiol. Methods*, 2010, **83**, 34–41.

10 P. et al. Gileadi, O., Burgess-Brown, N. A., Colebrook, S. M., Berridge, G., Savitsky, in *Structural Proteomics: High-Throughput Methods*, ed. T. Kobe, B., Guss, M., & Huber, Humana Press, 2008, pp. 221–246.

11 Langley G. W., Studies on Collagen Post-Translational Modifications, D.Phil Thesis, University of Oxford, 2017.

12 O. Gileadi, *pNIC28-Bsa4 Vector Sheet, Structural Genomics Consortium*, .

13 O. Gileadi, *pNH-TrxT Vector Sheet, Structural Genomics Consortium*, .

14 P. Savitsky, *pNIC-ZB Vector Sheet, Structural Genomics Consortium*, 2013.

15 C. Loenarz, W. Ge, M. L. Coleman, N. R. Rose, C. D. O. Cooper, R. J. Klose, P. J. Ratcliffe and C. J. Schofield, *Hum. Mol. Genet.*, 2009, **19**, 217–222.

16 J. R. Horton, A. K. Upadhyay, H. H. Qi, X. Zhang, Y. Shi and X. Cheng, *Nat. Struct. Mol. Biol.*, 2010, **17**, 38–44.

17 A. Kawamura, A. Tumber, N. R. Rose, O. N. F. King, M. Daniel, U. Oppermann, T. D. Heightman

and C. Schofield, *Anal. Biochem.*, 2010, **404**, 86–93.

18 R. J. Hopkinson, A. Tumber and C. Yapp, *Chem Sci*, 2015, **4**, 3110–3117.

19 W. Xu, J. D. Podoll, X. Dong, A. Tumber, U. Oppermann and X. Wang, *J Med Chem*, 2013, **56**, 5198–5202.

20 J. R. Horton, A. K. Upadhyay, H. Hashimoto, X. Zhang and X. Cheng, *J. Mol. Biol.*, 2011, **406**, 1–8.

21 F. H. Niesen, H. Berglund and M. Vedadi, *Nat. Protoc.*, 2007, **2**, 2212–2221.

Chapter 6 Summary and Conclusions

The histone lysine demethylases (KDMs) are involved in epigenetic regulation, at least in part, because they catalyse the removal of methyl groups from methylated lysine residues, in particular on the *N*-terminal histone tails.[1-4] The JmjC domain-containing KDMs (JmjC-KDMs) belong to the 2OG oxygenase superfamily and can catalyse the demethylation of tri-, di- and mono-methylated lysine residues.[4,5] The KDM7 subfamily of human JmjC-KDMs consists of three members: KIAA1718 (KDM7C), PHF8 (KDM7B) and PHF2 (KDM7C). Despite the sequences of each KDM7 enzyme sharing significant identity/similarity, the demethylation activities markedly differ across the subfamily.[5] PHF8 is a H3K9me2 demethylase while KIAA1718 is proposed to demethylate H3K9me2 and H3K27me2.[6] Although studies have been carried out for PHF2, whether it has catalytic activity remains a question.[7-10]

The work described in this Thesis aimed to investigate the substrate specificity of the KDM7 subfamily by carrying out kinetic and mechanistic studies. The work described in Chapter 2 concerned the development of a MALDI-TOF MS-based assay for the evaluation of the *in vitro* enzymatic activities of the human KDM7 enzymes. The results suggested that isolated PHF8 and KIAA1718 prefer a neutral pH buffer to achieve optimal demethylation activity. PHF8 preferentially demethylates H3K9me2. Importantly, the results support the proposal that this demethylation reaction is enhanced by the binding to H3K4me3 *via* its PHD finger.[6] On the other hand, KIAA1718 accepts H3K9me2 and H3K27me2. The demethylations at H3K9me2 as catalysed by KIAA1718 is inhibited by the presence of H3K4me3, consistent

with a previous report.[6] However, as the results in Chapter 2 (Section 2.3.4) suggest, the effect of the presence of H3K4me3 on the demethylation of H3K27me2 as catalysed by KIAA1718 is still uncertain. In addition to the previously reported substrates of PHF8 and KIAA1718,[6,11] H3K14me2 and H3K23me2 were identified as substrates for both enzymes. KIAA1718$_{PHD\text{-}JmjC}$, but not PHF8$_{PHD\text{-}JmjC}$, was also observed to catalyse demethylation of H3K18me2 and H3K36me2 (Table 6.1). Despite extensive efforts of potential substrate screening for PHF2, no demethylation activity for it was found.

Table 6.1 Summary of demethylation activities of the KDM7 enzymes with various H3 substrates. + indicates demethylation observed; - indicates no demethylation observed.

	H3K9me2	H314me2	H3K18me2	H3K23me2	H3K27me2	H3K36me2
KIAA1718	+	+	+	+	+	+
PHF8	+	+	-	+	+	-
PHF2	-	-	-	-	-	-

A fluorescence polarisation (FP) assay was developed by using fluorescently labelled H3 peptides (H3$_{1\text{-}15}$K4me3K9me2 and H3$_{1\text{-}15}$K9me2) and FI-methylstat. The assay was used to investigate the substrate binding by the PHD finger and the JmjC domains of the KDM7 enzymes individually.[12,13] The PHD finger of both KDM7 enzymes tightly binds to H3$_{1\text{-}15}$K4me3K9me2-FITC with K_D values of 150.4 $\pm$ 4.3 nM (KIAA1718), 0.83 $\pm$ 0.08 µM (PHF8) and < 1.0 µM (PHF2). The PHD finger and the PHD-JmjC protein of a KDM7 (e.g. PHF8$_{PHD}$ and PHF8$_{PHD\text{-}JmjC}$) have similar binding affinities to the substrates, as represented by the binding K_D values to H3$_{1\text{-}15}$K4me3K9me2-FITC and H3$_{1\text{-}15}$K9me2-FITC (Table 6.2). A PHF8 protein construct with the JmjC domain alone does not bind to the histone 3 (H3) peptides (Table 6.2), suggesting that interactions with the PHD finger (rather than the JmjC domain) 'dominates' substrate binding by the KDM7 enzymes. The substituted proteins Y321H (PHF2)

and H319Y (PHF8), possessing substituted metal interacting residues, manifest only small differences in binding to $H3_{1-15}$K4me3K9me2, further supporting the proposal that the JmjC domain only has a minor influence on, at least initial, KDM7 substrate binding. Finally, recombinant PHF2 enzymes were found to be catalytically inactive in isolated form (as evidenced by the activity assays); however, both $PHF2_{PHD-JmjC}$ and $PHF2_{JmjC}$ bind to FI-methylstat with K_D values in the micromolar range (Table 6.2), providing an opportunity to study these non-catalytic proteins.

In addition to the FP method, an AlphaScreen platform was established to facilitate high-throughput inhibitor screening of the KDM7 enzymes.[14] Using the AlphaScreen screening method, the IC_{50} values of a series of amiodarone (AMI) derivatives, which are reported to inhibit the PHD finger 3 of KDM5A ($KDM5A_{PHD3}$),[15] were measured. The potent AMI derivatives that inhibit the KDM activity of PHF8 and KIAA1718 have been found not directly bind to the PHD finger which they were designed to target.

Together with the FP screening, the two assay platforms helped to identify a novel cyclic peptide inhibitor of PHF8 and an 'enhancer' of KIAA1718 that binds selectively to their PHD fingers. A further example of the application of the FP and AlphaScreen screening platforms was in the investigation of the mode of inhibition of a selective KDM2/7 inhibitor (*S,S*)-6.[16] Although (*S,S*)-6 has been found to inhibit the enzymatic activity of KDM2A and KIAA1718 with IC_{50} values of 0.16 μM and 0.19 μM, respectively (using AlphaScreen), FP competition assays reveals that the compound does not target the PHD finger or the JmjC domain of KIAA1718 and is likely to bind to the linker region.

Table 6.2 Summary of the FP binding assays carried out during the work of this Thesis. NM indicates that the K_D values are not measurable under the tested conditions.

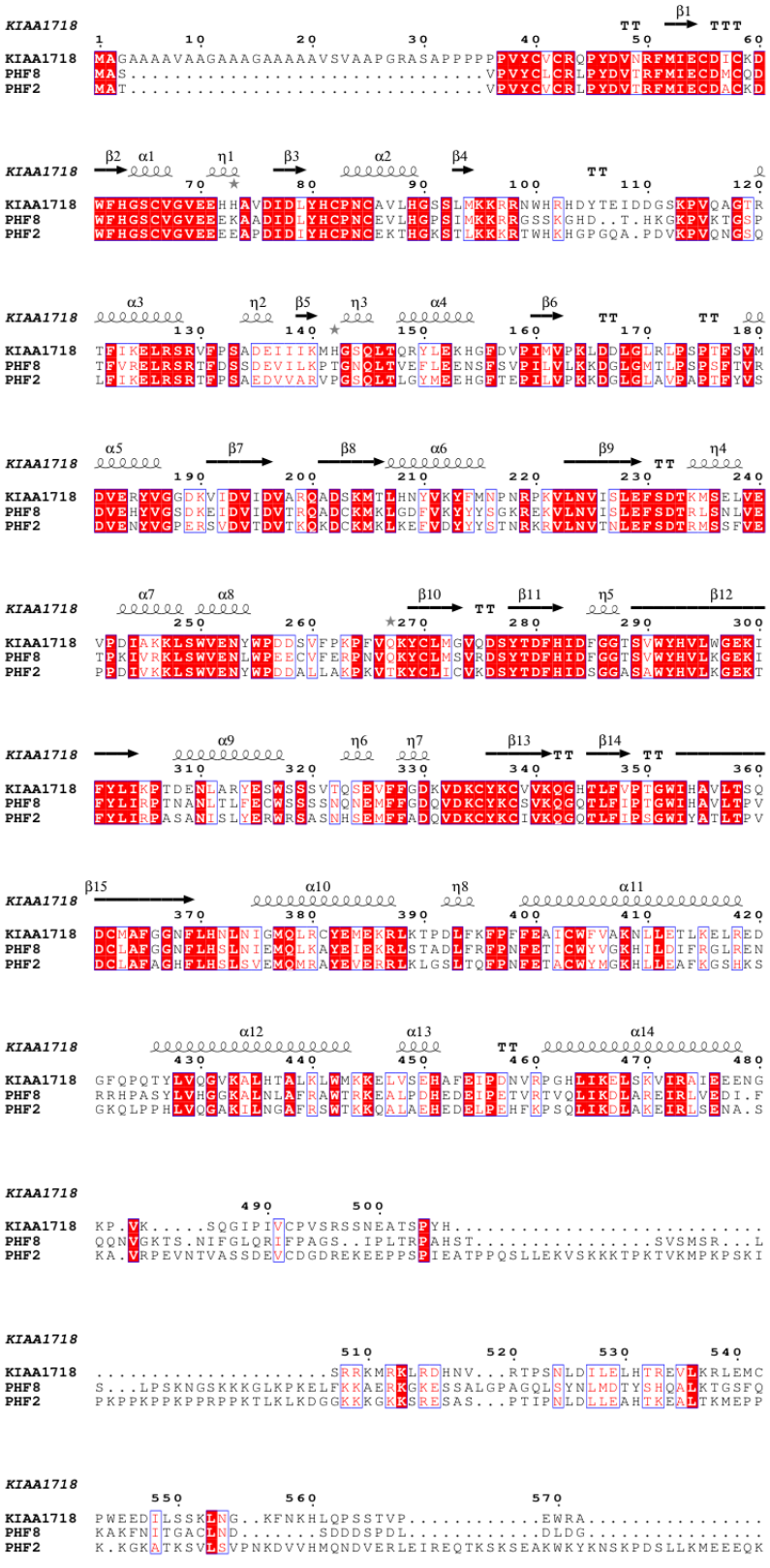
Protein	Construct	K_D H3K4me3K9me2/nM no 2OG	K_D H3K4me3K9me2/nM with 2OG	K_D H3K9me2/µM no 2OG	K_D H3K9me2/µM with 2OG	K_D Fl-methylstat/µM
KIAA1718	PHD	419.2 ± 54.0	-	-	-	-
	PHD-JmjC	150.4 ± 4.3	106.9 ± 3.7	> 50	> 50	125 ± 11.7 nM
PHF2	JmjC	NM	NM	NM	NM	6.37 ± 1.03
	PHD	< 1 µM	-	-	-	-
	PHD-JmjC	31.3 ± 1.4	32.1 ± 1.3	-	-	2.5 ± 0.21
PHF8	JmjC	NM	NM	NM	NM	1.33 ± 0.04
	PHD	159.5 ± 14.1	135.4 ± 18.2	> 50	> 50	-
	PHD-JmjC	826.2 ± 75.8	638.4 ± 68.1	NM	NM	0.98 ± 0.07
PHF8H319Y	PHD-JmjC	1.10 ± 0.1 µM	-	-	-	1.56 ± 0.03
PHF2Y321H	PHD-JmjC	< 1 µM	-	-	-	NM

Overall, the results described in this thesis support the proposal that the PHD finger of KDM7s is critical for substrate binding and can fine-tune the substrate selectivity of PHF8 and KIAA1718. Targeting the KDM's PHD finger may thus provide a novel strategy for inhibitor (PHF8) and, maybe, enhancer (KIA1718B) design. The establishment of the two complementary inhibitor screening platforms also enables opportunities to screen new inhibitors in a high-throughput manner and to efficiently investigate their modes of inhibition.

References:

1. Kooistra, S. M. & Helin, K. Molecular mechanisms and potential functions of histone demethylases. *Nat. Rev. Mol. Cell Biol.* **13,** 297 (2012).
2. Markolovic, S. *et al.* Structure–function relationships of human JmjC oxygenases—demethylases versus hydroxylases. *Curr. Opin. Struct. Biol.* **41,** 62–72 (2016).
3. Ponnaluri, V. K. C., Vavilala, D. T., Putty, S., Gutheil, W. G. & Mukherji, M. Identification of non-histone substrates for JMJD2A–C histone demethylases. *Biochem. Biophys. Res. Commun.* **390,** 280–284 (2009).
4. Hancock, R. L., Dunne, K., Walport, L. J., Flashman, E. & Kawamura, A. Epigenetic regulation by histone demethylases in hypoxia. *Epigenomics* **7,** 791–811 (2015).
5. Klose, R. J., Kallin, E. M. & Zhang, Y. JmjC-domain-containing proteins and histone demethylation. *Nat. Rev. Genet.* **7,** 715–727 (2006).
6. Horton, J. R. *et al.* Enzymatic and structural insights for substrate specificity of a family of jumonji histone lysine demethylases. *Nat. Struct. Mol. Biol.* **17,** 38–44 (2010).
7. Horton, J. R., Upadhyay, A. K., Hashimoto, H., Zhang, X. & Cheng, X. Structural Basis for Human PHF2 Jumonji Domain Interaction with Metal Ions. *J. Mol. Biol.* **406,** 1–8 (2011).
8. Baba, A. *et al.* PKA-dependent regulation of the histone lysine demethylase complex PHF2-ARID5B. *Nat. Cell Biol.* **13,** 668–675 (2011).
9. Hata, K. *et al.* Arid5b facilitates chondrogenesis by recruiting the histone demethylase Phf2 to Sox9-regulated genes. *Nat. Commun.* **4,** 1–11 (2013).
10. Shi, G. *et al.* PHD finger protein 2 (PHF2) represses ribosomal RNA gene transcription by antagonizing PHF finger protein 8 (PHF8) and recruiting methyltransferase SUV39H1. *J. Biol. Chem.* **289,** 29691–29700 (2014).
11. Loenarz, C. *et al.* PHF8, a gene associated with cleft lip/palate and mental retardation, encodes for an Nε-dimethyl lysine demethylase. *Hum. Mol. Genet.* **19,** 217–222 (2009).
12. Luo, X. *et al.* A selective inhibitor and probe of the cellular functions of jumonji C domain-containing histone demethylases. *J. Am. Chem. Soc.* **133,** 9451–9456 (2011).
13. Xu, W. *et al.* Quantitative analysis of histone demethylase probes using fluorescence polarization. *J Med Chem* **56,** 5198–5202 (2013).
14. Kawamura, A. *et al.* Development of homogeneous luminescence assays for histone demethylase catalysis and binding. *Anal. Biochem.* **404,** 86–93 (2010).
15. Bhushan, B. *et al.* Investigations on small molecule inhibitors targeting the histone H3K4 tri-methyllysine binding PHD-finger of JmjC histone demethylases. *Bioorganic Med. Chem.* **26,** 2984–2991 (2018).
16. Gerken, P. A. *et al.* Discovery of a Highly Selective Cell-Active Inhibitor of the Histone Lysine Demethylases KDM2/7. *Angew. Chemie - Int. Ed.* **56,** 15555–15559 (2017).

Appendix 1: Sequence alignment of KDM7 subfamily



PHF2(NP_005383.3),PHF8(NP_055922),KIAA1718(NP_085150).

Appendix 2: Recombinant KDM7 constructs

```
PHF2          MATVPVYCVCRLPYDVTRFMIECDACKDWFHGSCVGVEEEEAPDIDIYHCPNCEKTHGKS   60
KIAA1718      ----MVYCVCRQPYDVNRFMIECDICKDWFHGSCVGVEEHHAVDIDLYHCPNCAVLHGSS   92
PHF8          MASVPVYCLCRLPYDVTRFMIECDMCQDWFHGSCVGVEEEKAADIDLYHCPNCEVLHGPS   60
                   ***:** ****.******* *:************..* ***:******   ** *

PHF2          TLKKKRTWHKHGPGQ-APDVKPVQNGSQLFIKELRSRTFPSAEDVVARVPGSQLTLGYME   119
KIAA1718      LMKKRRNWHRHDYTEIDDGSKPVQAGTRTFIKELRSRVFPSADEIIIKMHGSQLTQRYLE   152
PHF8          IMKKRRGSSKGHD---THKGKPVKTGSPTFVRELRSRTFDSSDEVILKPTGNQLTVEFLE   117
               :**:*   :          ***: *:  *::*****.* *::::: :  *.***  ::*

PHF2          EHGFTEPILVPKKDGLGLAVPAPTFYVSDVENYVGPERSVDVTDVTKQKDCKMKLKEFVD   179
KIAA1718      KHGFDVPIMVPKLDDLGLRLPSPTFSVMDVERYVGGDKVIDVIDVARQADSKMTLHNYVK   212
PHF8          ENSFSVPILVLKKDGLGMTLPSPSFTVRDVEHYVGSDKEIDVIDVTRQADCKMKLGDFVK   177
              ::.*  **:* * *.**: :*:*:* * ***.*** :: :** **::* *.**.* ::*.

PHF2          YYYSTNRKRVLNVTNLEFSDTRMSSFVEPPDIVKKLSWVENYWPDDALLAKPKVTKYCLI   239
KIAA1718      YFMNPNRPKVLNVISLEFSDTKMSELVEVPDIAKKLSWVENYWPDDSVFPKPFVQKYCLM   272
PHF8          YYYSGKREKVLNVISLEFSDTRLSNLVETPKIVRKLSWVENLWPEECVFERPNVQKYCLM   237
              *: . :* :**** .******::*.:** *.*.:******* **::.:: :* * ****:

PHF2          CVKDSYTDFHIDSGGASAWYHVLKGEKTFYLIRPASANISLYERWRSASNHSEMFFADQV   299
KIAA1718      GVQDSYTDFHIDFGGTSVWYHVLWGEKIFYLIKPTDENLARYESWSSSVTQSEVFFGDKV   332
PHF8          SVRDSYTDFHIDFGGTSVWYHVLKGEKIFYLIRPTNANLTLFECWSSSSNQNEMFFGDQV   297
               *:********* **:*.***** *** ****:*:. *:: :* * *: .:.*:**.*:*

PHF2          DKCYKCIVKQGQTLFIPSGWIYATLTPVDCLAFAGHFLHSLSVEMQMRAYEVERRLKLGS   359
KIAA1718      DKCYKCVVKQGHTLFVPTGWIHAVLTSQDCMAFGGNFLHNLNIGMQLRCYEMEKRLKTPD   392
PHF8          DKCYKCSVKQGQTLFIPTGWIHAVLTPVDCLAFGGNFLHSLNIEMQLKAYEIEKRLSTAD   357
              ****** ****:***:*:***:*.**  **:**.*:***.*.: **::.**:*:**.  .

PHF2          LTQFPNFETACWYMGKHLLEAFKGSHKSGKQLPPHLVQGAKILNGAFRSWTKKQALAEHE   419
KIAA1718      LFKFPFFEAICWFVAKNLLETLKELREDGFQPQTYLVQGVKALHTALKLWMKKELVSEHA   452
PHF8          LFRFPNFETICWYVGKHILDIFRGLRENRRHPASYLVHGGKALNLAFRAWTRKEALPDHE   417
              * :** **: **::.*::*: ::  ::.  :   :**:* * *: *:: * :*: : :*

PHF2          DELPEHFKPSQLIKDLAKEIRLSENASKAV       449
KIAA1718      FEIPDNVRPGHLIKELSKVIRAIEEENG--       480
PHF8          DEIPETVRTVQLIKDLAREIRLVEDIFQQN       447
               *:*: .:  :***:*:: **  *:
```

PHD finger regions are highlighted in green, JmjC domains are highlighted in blue and linker regions are highlighted in purple.

Appendix 3: Publications arising from work described in this Thesis

1. Walport, L. J., Hopkinson, R. J., Chowdhury, R., Zhang Y., Bonnici, J., Schiller, R., Kawamura, A., and Schofield, C. J. (2018) Mechanistic and structural studies of KDM-catalysed demethylation of histone 1 isotype 4 at lysine 26. *FEBS Lett.* DOI:10.1002/1873-3468.13231

2. Gerken P. A., Wolstenhulme J. R., Tumber A., Hatch S. B., Zhang Y., Müller S., Chandler S. A., Mair B., Li F., Nijman S. M. B., Konietzny R., Szommer T., Yapp C., Fedorov O., Benesch J. L. P., Vedadi M., Kessler B. M., Kawamura A., Brennan P. E. and Smith M. D. (2017) Discovery of a highly selective cell-active inhibitor of the histone lysine demethylases KDM2/7. *Angew. Chemie - Int. Ed.*, 2017, **56**, 15555–15559

3. Bhushan B., Erdmann A., Zhang Y., Belle R., Johannson C., Oppermann U., Hopkinson R. J., Schofield C. J. and Kawamura A. (2018) Investigations on small molecule inhibitors targeting the histone H3K4 tri-methyllysine binding PHD-finger of JmjC histone demethylases. *Bioorganic Med. Chem.*, 2018, **26**, 2984–2991