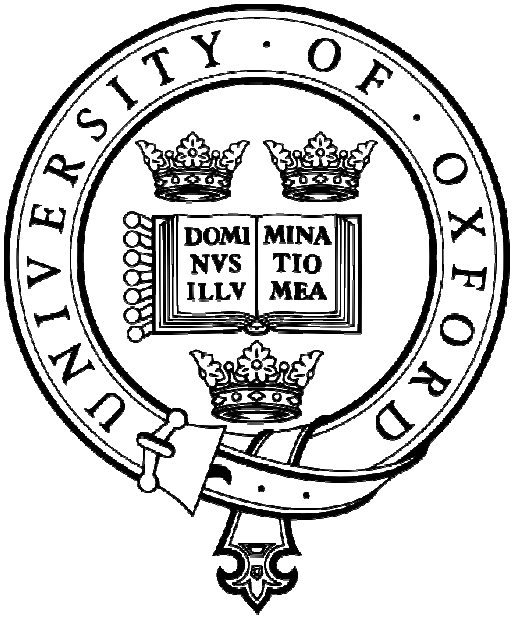


Mechanistic Studies of JMJD6, Fe(II) and 2OG dependent lysyl hydroxylase

Monica Mantri

Chemistry Research Laboratory

Department of Organic Chemistry

Wolfson College

Michealemas Term 2011

A thesis submitted to the board of the Faculty of Physical Sciences at University of Oxford in partial fulfilment of the requirements for the degree of Doctor of Philosophy

ABSTRACT

MECHANISTIC STUDIES OF JMJD6, FE(II) AND 2-OXOGLUTARATE DEPENDENT LYSYL HYDROXYLASE

JMJD6 or PSR (phosphatidyl serine receptor) was initially proposed to be a membrane receptor involved in apoptotic cell clearance by recognition of apoptotic cells. However, sequence analyses implied the presence of a jelly roll or double stranded beta helix (DSBH) structural domain in PSR/JMJD6 and similarity with JmjC family of enzymes which are involved in chromatin regulation. Subsequently, PSR was renamed as JMJD6 and was reported to be a histone arginine demethylase. Previous work from our group has shown that JMJD6 is a lysine hydroxylase that interacts with nuclear proteins including CROP and U2AF65 which are involved in mRNA splicing. Peptide screening and cell based assays led to the conclusion that JMJD6 catalyses lysine hydroxylation of splicing regulatory proteins containing arginine serine rich domains (SR proteins) including U2AF65 and Luc7like-2.

Studies were carried out to investigate the putative arginine demethylation activity of JMJD6 using MS analysis of histone peptides and luminescence-based assays. New substrates from SR proteins were identified by immunoprecipitation of JMJD6 expressed in human cell lines followed by LC-MS/MS analysis and MALDI-MS based assays of synthesised peptide substrates. Work then focussed on studying the mechanism of lysyl-hydroxylation from substrate and enzyme perspective. A crystal structure of seleno-methionine labelled JMJD6 was obtained and it provided insights into the JMJD6 active site and its substrate interactions. Based on this data, single point variants of JMJD6 were prepared and their substrate binding properties were studied by MALDI-MS and 2OG turnover assays.

Collagen lysyl-hydroxylases are also 2OG dependent oxygenases. Efforts to investigate the stereochemistry of JMJD6 catalysed hydroxylation, employing NMR and amino acid analyses were carried out. These studies led to the interesting finding that the C-5 stereochemistry of hydroxylysine in LUC7L2 peptide is opposite (2*S*,5*S-hydroxylysine)* to that present in collagen (2*S*,5*R-hydroxylysine*). It was found that JMJD6 undergoes autocatalytic self-hydroxylation. Lysine residues from both recombinant JMJD6 and that from HeLa cells at endogenous level were identified to be hydroxylated by amino acid and LC-MS/MS analyses.

JMJD6 has a strong tendency to form aggregates and gel electrophoresis always reveals multimeric bands of various JMJD6 constructs. Characterisation and identification of oligomeric states of JMJD6 was carried out using Electron Microscopy. Studies were initiated to identify possible inhibitors by screening a set of 2OG analogues. The results from this preliminary inhibition studies have identified the tricarboxylic acid (TCA) cycle intermediates, succinate and fumarate to be JMJD6 inhibitors and form a basis of further studies aimed at identifying selective inhibitors.

In the memory of beloved grandmother, Sunanda Mantri

Table of Contents

Acknowledgements vii

Abbreviations ix

Chapter 1. Introduction 1

1.1 Overview 1
1.2 JMJD6 as phosphatidyl serine receptor 1
1.3 Gene deletion studies of PSR/JMJD6 in model organisms 4
1.4 PSR/JMJD6 is not involved in apoptosis 6
1.5 PSR/JMJD6 is a nuclear protein 7
1.6 PSR/JMJD6 and the presence of JmjC domain 10
1.7 General overview of 2OG and Fe (II) dependent oxygenases 12
1.7.1 Biosynthesis of small molecules and antibiotics 15
1.7.2 Collagen post-translational modifications: lysine/proline hydroxylation 17
1.7.3 Oxygen sensing 19
1.7.4 Fatty acid biosynthesis 20
1.7.5 Histone modifications 22
1.7.6 Nucleic acid modifications 24
1.8 JMJD6 as arginine demethylase 26
1.9 JMJD6 is a lysine hydroxylase 27
1.10 Cellular localisation of JMJD6 28
1.11 Involvement of JMJD6 in splicing 30
1.12 Aims of the work described in this thesis 32
1.13 References 33

Chapter 2. JMJD6 as a lysyl hydroxylase 41

2.1 Introduction 41
2.2 Background: Arginine demethylation of histone tails 41
2.3 ALPHA Screen binding assay for investigating the association of JMJD6 with histone peptides 44
2.4 Histone peptides showing no arginine demethylation but lysine hydroxylation 49

2.5 Finding substrates for JMJD6.....56
2.5.1 IP of JMJD6.....56
2.5.2 LC-MS/MS results confirming the association of JMJD6 with arginine-serine rich (SR) proteins involved in pre-mRNA splicing.....60
2.6 Pre-mRNA Splicing and SR proteins.....63
2.7 Synthesis of potential peptide substrates for JMJD6.....67
2.8 MALDI-MS assays of peptides from SR proteins associated with JMJD6 in IP..70
2.8.1 U2AF65 and U2AF35 interaction and catalytic hydroxylation by JMJD6....70
2.8.2 Lysine hydroxylation in LUC7L2, SRSF11, ZRANB2 73
2.9 Lysine hydroxylation in ACINUS.....77
2.10 Conclusions and future work.....82
2.11 References 85

Chapter 3. Crystal structure of JMJD6 90

3.1 Introduction 90
3.2 Preparation of protein.....91
3.2.1 Choice of truncated construct.....91
3.2.2 Preparation of Seleno-methionine substituted JMJD6 92
3.3 Crystallisation of JMJD6.....93
3.3.1 High throughput crystallisation screen.....93
3.3.2 Optimisation of crystallisation conditions.....95
3.3.3 Crystallisation under hypoxic conditions. 96
3.4 Crystal structure analysis.....96
3.4.1 General structural features of JMJD6.....98
3.4.2 Details of active site 102
3.4.3 Comparison of active site residues with other 2OG oxygenases (FIH, PHF8 and JMJD2A).....104
3.5 Docking of LUC7L2 peptide substrate to JMJD6 crystal structure 106
3.6 Mutagenesis studies of JMJD6.....107
3.6.1 Comparing mutants by 2OG turnover assays.....108
3.6.2 Comparing mutants by MALDI-MS assays 109
3.7 Results and Discussion.....110
3.8 Conclusions 115
3.9 References 116

Chapter 4. Stereochemical determination of lysine hydroxylation by JMJD6.....119

4.1 Introduction 119
4.2 Synthesis of Fmoc protected hydroxylysine.....121
4.3 Peptide Synthesis.....122
4.4 Hydroxylation of LUC7L2 peptide by JMJD6.....123
4.5 NMR analysis of hydroxylated lysine 124
4.6 Amino acid analysis 126
4.6.1 HCl digestion of peptide after hydroxylation by JMJD6 and possible epimerisation of hydroxylysine 132
4.6.2 Development of enzyme based digestion of peptide after JMJD6 hydroxylation.....133
4.7 Discussion.....136
4.8 References 138

Chapter 5. Self-hydroxylation of JMJD6 144

5.1 Introduction 144
5.2 Uncoupled 2OG turnover by JMJD6.....145
5.2.1 Radioactive 2OG turnover assays 145
5.2.2 2OG turnover of JMJD6 as monitored by ^{1}H-NMR analysis.....146
5.2.3 MALDI-MS assays.....148
5.3 Amino acid analysis of JMJD6.....149
5.4 Proteomic studies to identify hydroxylated lysine residues in JMJD6.....152
5.4.1 Determination of lysine hydroxylation sites in recombinant JMJD6.....153
5.4.2 Determination of lysine hydroxylation sites in JMJD6 from HeLa cells.....155
5.5 Determination of lysine hydroxylation sites by MALDI-MS assays using JMJD6 peptide substrates.....157
5.6 Discussion.....160
5.6.1 Active site of JMJD6 and sites of self-hydroxylation 160
5.6.1 Biological relevance of self-hydroxylation in case of JMJD6 163
5.7 References 166

Chapter 6. Oligomerisation of JMJD6 169

6.1 Introduction 169
6.2 Gel electrophoresis showing higher bands 170
6.3 Analytical size exclusion chromatogarphy.....171

6.4 Analysis of JMJD6 oligomers by TEM 173
6.4.1 Transmission electron microscopy-A general overview 173
6.4.2 Treatment of samples for TEM 175
6.4.3 Image analysis 177
6.5 TEM analysis of various JMJD6 constructs by negative staining technique 178
6.6 Cryo EM analysis of native full length JMJD6 184
6.6.1 Preparation of wild-type JMJD6 for Cryo-freezing TEM 184
6.6.2 Image Analysis 186
6.7 Discussion 187
6.8 References 191

Chapter 7. Preliminary inhibition studies on JMJD6 195

7.1 Introduction 195
7.2 Kinetic studies of JMJD6 catalysed hydroxylation 196
7.2.1 Time-course analysis of JMJD6 catalysed hydroxylation 198
7.2.2 2OG dependence of JMJD6 for catalysing hydroxylation in $LUC7L2_{267-278}$ peptide 199
7.2.3 Fe(II) dependence of JMJD6 for catalysing hydroxylation in $LUC7L2_{267-278}$ peptide 200
7.2.4 JMJD6 does not require ascorbate for hydroxylation 201
7.3 Screening of inhibitor library for JMJD6. 202
7.3.1 Differential scanning fluorimetry (DSF) assays for JMJD6 202
7.3.2 Screening of library of inhibitors by MALDI-MS analysis 203
7.4 Inhibition of JMJD6 by TCA cycle intermediates 208
7.5 Development of MALDI-MS assay for determining IC_{50} of various inhibitors 210
7.6 Discussion 213
7.6.1 TCA cycle intermediates 213
7.6.2 Inhibition by other compounds by JMJD6 215
7.7 Future Work 217
7.8 References 219

Chapter 8. Materials and Methods 224

8.1 General methods used in chemical synthesis 224
8.2 General methods used for biological work 224
8.2.1 pH measurement 225

8.3 Genral methods used for molecular biology 225
8.3.1 Oligonucleotide primer synthesis and preparation 225
8.3.2 Plasmid DNA purification and determination of DNA concentration 225
8.3.3 DNA sequencing 225
8.4 Site directed mutagenesis 226
8.5 Microbial techniques 227
8.5.1 Bacterial growth media.....228
8.5.2 Competent cell transformation 229
8.5.3 Preparation of competent cells 229
8.5.4 Bacterial glycerol stocks.....230
8.5.5 Small scale growths.....230
8.5.6 Large scale growths.....230
8.6 Expression of Seleno-methionine exchanged JMJD6.....231
8.6.1 100ml starter culture:.....231
8.6.2 Large scale growth 232
8.7 Protein purification.....233
8.7.1 Nickel affinity chromatography 233
8.7.2 Gel filtration chromatography 234
8.7.3 Protein concentration determination.....235
8.8 SDS-PAGE analysis of protein 235
8.9 ALPHA Screen assays.....237
8.10 MALDI-MS assays.....237
8.11 Cell biology techniques 238
8.11.1Cell culture and transfection.....238
8.11.2 Cell lysis and immunoprecipitation.....239
8.11.3 In-gel trypsin digestion.....240
8.12 LC-MS/MS analysis 241
8.13 Peptide synthesis 242
8.14 HPLC purification of peptides.....245
8.15 Crystallisation of JMJD6.....246
8.15.1 Crystallisation.....246
8.15.2 Data Collection and Processing.....247
8.15.3 Structure Solution and Refinement 247
8.15.4 Structure analysis.....248
8.16 2OG turnover assays.....248
8.16.1 Radioactive 2OG turnover assays 248

8.16.2 2OG turnover by NMR....249
8.17 NMR analysis of peptides....249
8.18 Amino Acid Analysis....250
8.18.1 Sample preparation for amino acid analysis....250
8.18.2 Derivatisation with AQC and LC-MS analyses....251
8.19 Transmission Electron Microscopy....252
8.19.1 Negative staining EM....252
8.19.2 Cryo -EM....253
8.20 DSF analysis....253
8.21 MALDI-MS assays to determine IC_{50} values of inhibitors....254
8.22 References....255

Appendix A:....257

Appendix B:....262

Appendix C:....264

Appendix D:....269

Acknowledgements

I would like to thank Prof. Chris Schofield for his supervision and support during my D.Phil studies and for allowing me to pursue research in this exciting project. I have thoroughly enjoyed all aspects of my research within his group and shall be eternally grateful for all the help and advice that he has given me.

To quote Sir Isaac Newton “If I have seen further it is by standing on the shoulders of giants.”, I would sincerely like to thank Danica Butler, Celia Webby and Alexander Wolf who carried out the initial work of this research project. Special thanks to Celia for her supervision and sharing her tips on handling JMJD6 which lead the path of this challenging project. I am indebted to Alexander Wolf for his encouraging words and invaluable input on the project alongwith his patience while teaching cell biology.

Work in this thesis would not have been possible without contributions from many collaborators. Many thanks to our collaborator Prof. Angelika Böttger from Ludwig-Maximillians University, Germany for her valuable discussion and feedback on the project. I am also grateful to Dr. Catherine Vénien-Bryan for electron microscopy analysis. In particular, I am indebted to the many co-workers who shared their expertise to get this project working; Dr. Michael McDonough for crystallography, Dr. Refaat Hamed for NMR, Nikita Loik for amino acid analysis, Dr. Mukram Mckeen for LC-MS analysis. I would also like to thank Radhakrishnan Memorial Bequest for their financial help.

Thanks to everyone else in the CJS lab, in particular to Wendy and Manso for invaluable technical assistance. Thanks to ex-CJS group members Maggie, Aman, Esther, Steve and Nathan who helped me immensely in initial first year to get settled in lab. Additional thanks to Eleanor, Wei Shen, Louise, Inga, Anna, Andrea, Ruben, Wei, Zhihong, Akane, Adam, Emily, Luc, Richard, Katherine and all other CJS lab members that have helped make working in the lab so enjoyable. Thanks to Anna Boelingerer, Nikita Loik, Esther Woon and Anna Ridzick for being great friends.

Outside the lab, amongst the friends who shared all the ups and downs of Oxford days, I am especially thankful to Ashutosh, Titziana and Miralisa. I am thankful to Anil, Ritesh and Karthik for their constant support in last four years. Special thanks to Parag and Sangeeta who made sure that I did not miss India in last four years with their love and special treats.

Finally, my most heartfelt thanks must go to parents, my sister Manasi and my in-laws for their continued support towards my research career. Special thanks goes to my father-in-law, Virendra who believed in my dreams and supported me to achieve them. Last, but by no means least, a huge thanks to my husband Amol; someone who shared all ups and downs of research life, cooked for me, shared scientific ideas on project, tolerated my tantrums on bad protein prep days in last four years, even proofread the thesis and still managed to tell me that I am a great wife and researcher.

Abbreviations

2OG	2-Oxogutaric acid
2TY	2x tryptone /Yeast Extract medium
2HG	2-hydroxyglutarate
α-TM	α-tropomyosin
ABH	AlkB homologues
ACINUS	Apoptotic chromatin condensation inducer of the nucleus
ACCO	1-amino-1-cyclopropanecarboxylic acid oxidase
ALPHA Screen	Amplified luminescent proximity homogeneous
APS	Ammonium Persulfate
ASAP	Apoptosis and splicing-associated protein
Atsk	Alkylsulfatase
BBOX	γ-butyrobetaine hydroxylase
BOC	N-tert-butyloxycarbonyl group
bp	Base pairs (of DNA sequence)
BSA	Bovine Serum Albumin
CAS	Clavaminic acid synthase
CAD	*C*-terminal activation domain
CHCA	Alpha-cyano-4-hydroxycinnamic acid
CoA	Coenzyme A
CROP	Cisplatin resistance associated overexpressed protein
Cs	Spherical aberration constant
CTF	Contrast transfer function
CV	Column volumes
DFO	Desferrioxamine,
DHB	2,5-dihydroxybenzoic acid
DIC	*N*,*N*′-di-isopropylcarbodiimide
Dll4	Delta like-4
DMEM	Dulbecco´s Modified Eagles Medium
DMF	Dimethyl formamide

DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide 5'-triphosphate
ds	Double stranded
DSBH	Double-stranded β-helix
DSF	Differential scanning fluorometry
DTT	Dithiothreitol
EDTA	Ethelyne diamine tetraacetic acid
EMSA	Electrophoretic mobility shift assay
emPAI	Exponentially modified protein abundance index
FBXL	*F*-box and leucine-rich repeat containing protein
FEG	field emission gun
FIH	Factor-inhibiting hypoxia-inducible factor
Fmoc	Fluorenylmethyloxycarbonyl group
FPLC	Fast protein liquid chromatography
FTO	Fat mass and obesity associated) gene
GFP	Green fluorescent protein
HEK	Human Embyonic Kidney
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
HIF	Hypoxia-inducible factor
HOBt	*N*-hydroxybenzotriazole
HPLC	High Performance liquid chromatography
HPPE	*S*-hydroxypropylphosphonic acid epoxidase
HRE	Hypoxia response elements
HSQC	Heteronuclear Single Quantum Coherence
IC_{50}	Half-maximal inhibitory concentration
ICSGNM	International committee for standardized genetic nomenclature
IP	Immunoprecipitation
IPNS	Isopenicillin N synthase
IPTG	Isopropyl β-D-1-thiogalactopyranoside]
JmjC	Jumonji C
Jmjd	JmjC-domain-containing demethylase

LB	Luria-Bertani medium
LUC7L2	LUC7-like2
MALDI-TOF	Matrix assisted laser deorption/ionization time of flight
MOPS	3-(N-morpholino)propanesulfonic acid
NCBI	National Center for Biotechnology Information
NFκB	Nuclear factor-κB
NLS	Nuclear localisation signal
NMR	Nuclear magnetic resonance
NOG	*N*-oxalylglycine
NP40	Nonidet P40 (octylphenoxypolyethoxyethanol
OD	Optical Density
ODD	oxygen-dependent degradation domain
PAGE	Polyacrylamide Gel Electrophoresis
PAHX	Phytanoyl CoA Hydroxylase
PADI4	peptidyl arginine deiminase
PDB	RCSB Protein Data Bank
PEI	Poly(ethylenimine)
PHD	*trans*-4-prolyl hydroxylases
PHF8	PHD finger containing protein 8
PlGF	Placental growth factor
PRMT	Protein arginine methyl transferase
PSR	Phosphotidyl serine receptor
PUF60	Poly-U binding splicing factor 60
RS domain	Arginine-serine domains
RRMs	RNA recognition motifs
SA	3,5-dimethoxy-4-hydroxycinnamic acid
SAD	Single wavelength anomalous dispersion
SDS	Sodium Dodecyl Sulphate
SeMet	Seleno-methionine
SF2/ASF	Splicing regulatory protein SF2/ASF
snRNPs	Small nuclear ribonuclear protein complexes
SOC	Super Optimal Broth Catabolite Repression

SR protein	Proteins containing Serine Arginine rich domains
SRSF11	Serine/arginine-rich splicing factor 11
Sxl	Sex-lethal protein
TAP-tag	Tandem affinity tag purification
TAE	Tris-Acetate-EDTA
TauD	Taurine/α-ketoglutarate dioxygenase
TCA	Tricarboxylic acid cycle
TCEP	(*tris*(2-carboxyethyl)phosphine
TEM	Transmission electron microscope
TEMED	*N*,*N*,*N'*,*N'*-tetramethylethylenediamine
TET	Ten-Eleven-Translocation
TFA	Trifluroacetic acid
TfdA	2,4-dichlorophenoxyacetate oxygenase.
TGF-β	Transforming growth factor
THF	Tetrahydro furan
THP	3-(2,2,2-trimethylhydrazine)-propionate
Tris	Tris (hydroxyl methyl) amino methane
TLC	Thin layer chromatography
T_m	Melting temperature
TMLH	N^{ϵ}-trimethyllysine hydroxylase
TOCSY	1D total correlation nuclear magnetic resonance spectroscopy
U2AF65	U2 small nuclear ribonucleoprotein auxiliary factor 65-kDa subunit
U2AF35	U2 small nuclear ribonucleoprotein auxiliary factor 35-kDa subunit
UPLC	Ultra High Pressure Liquid Chromatography
VEGF	Vascular endothelial growth factor
v/v	Volume by volume

Chapter 1. Introduction

1.1 Overview

The gene encoding for Jumonji domain containing protein 6 (JMJD6, GI Number 125988395) was first identified by Fadock et al in 2000.[1] For sometime it was thought that JMJD6 is a receptor involved in recognition of phosphatidylserine present on apoptotic cells and thereby plays a major role in phagocytosis. Hence, JMJD6 was initially named as phosphatidylserine receptor (PSR). Since the discovery of PSR/JMJD6, contradictory studies on its cellular localisation and biological function have been reported. JMJD6 has been assigned as a transmembrane receptor involved in apoptosis but also as a nuclear protein involved in chromatin regulation.[2] Gene deletion studies in animals have shown that ablation of JMJD6 causes neonatal lethality and severe embryonic and developmental defects,[3-7] thereby emphasising its importance as a vital differentiation promoting gene. Bioinformatics studies have predicted the presence of a cupin fold in JMJD6 which is also present in members of the jumonji (JmjC) family of transcription factors; hence it was renamed from PSR to JMJD6.[8] In this chapter, I review the contradictory studies carried out to investigate the biological functions of JMJD6.

1.2 JMJD6 as phosphatidyl serine receptor

Unlike during the removal of lysed or necrotic cells which involves inflammation and the immune response, apoptotic cell clearance involves a complex mechanism of cell uptake by special cells like macrophages. Apoptotic cells undergo loss of phospholipid symmetry which exposes phosphatidylserine on the plasma membrane so that the

apoptotic cells can be recognised by phagocytes including macrophages.[9] During the course of their work, Fadock et al. generated a monoclonal antibody mAb217 against human macrophages that had been treated with transforming growth factor-β (TGF-β) and β-glucans to induce phosphatidylserine recognition.[1] This antibody bound more strongly to stimulated macrophages than non-stimulated macrophages and binding was inhibited by phosphatidylserine lipososmes. Inhibition of apoptotic cell binding by mAb217 corresponded with inhibition by phosphatidylserine liposomes. These observations led to the conclusion that mAb217 interacts with a receptor which binds with PS on the membrane. Fadock et al. further cloned an antigen for the mAb217 antibody from a phage display library and identified a protein of molecular weight 48 kDa. When this protein was ectopically expressed in cells that could not be stained with mAb217 and that could not phagocytose apoptotic cells (such as mouse M12.C3 cells and human Jurakat cells), an antibody signal on the plasma membrane of these cells was obtained by flow cytometric analysis. Furthermore, 25% of M12.C3 cells that were transiently transfected with the *PSR* gene gained the ability to bind apoptotic cells and some even engulfed them. Thus, this putative protein receptor was thought to be responsible for recognition of phosphatidylserine on outer membranes of apoptotic cells and was termed the phosphatidylserine receptor (PSR) or PtdSerR. PSR was also reported to be highly conserved in mice and *C. elegans*.

Subsequently, it was proposed by Hoffman et al. that PSR plays a role as a molecular switch in apoptosis in a “tether and tickle” mechanism.[10] In this process binding and ingestion of apoptotic cells by phagocytes are two separate events. Binding of PSR on phagocytes delivers “tickle” signals which stimulate the internalisation of apoptotic

cells that are tethered to it following the uptake of apoptotic cells by macropinocytosis as shown in Figure 1.1.[11, 12] They also showed direct ligation of PSR produced TGF-β suggesting that PSR modulation is critical for inducing an immunosuppressive response.

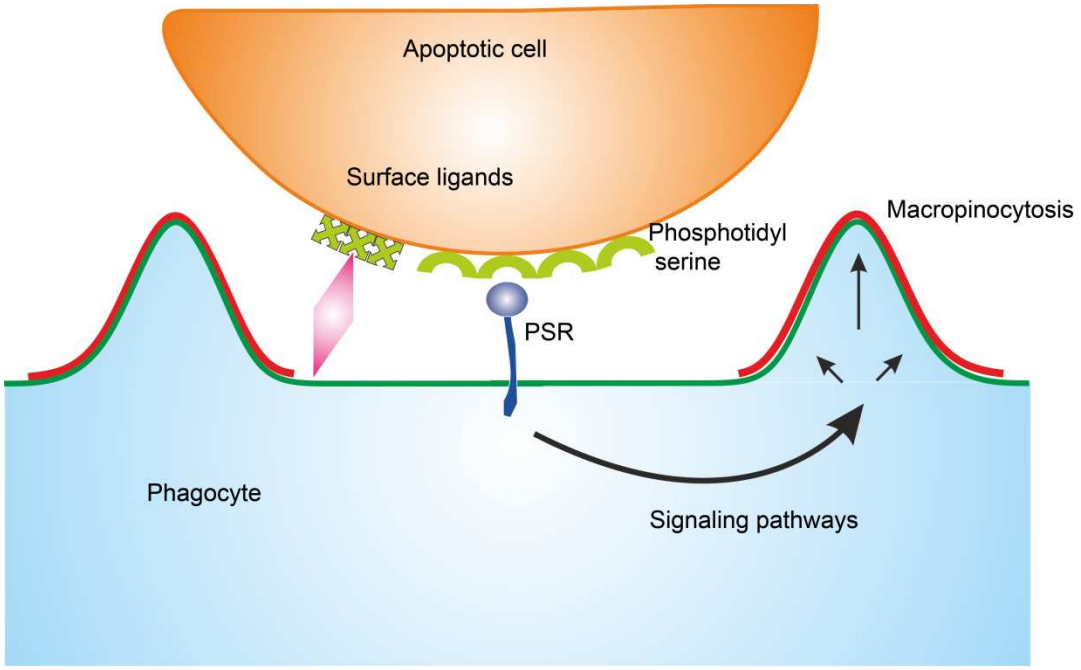


Figure 1.1 Tethering and tickling model of involvement of PSR in apoptotic cell engulfment by phagocytes. Image is modified from reference[11]

Arur et al. reported that PSR co-localises with annexin-1 on the outer plasma membrane of apoptotic cells and is necessary for engulfment of apoptotic cells.[13] They examined the possible clustering of PSR around apoptotic Jurkat T cells by overexpressing PSR in HEK293T cells co-cultured with apoptotic Jurkat T cells by immunoflouroscence microscopy. PSR clustering as complete rings around apoptotic Jurkat T cells was observed; this was diminished or was not detectable around annexin-1 silenced apoptotic cells or around viable Jurkat cells, even though they were attached to the PSR expressing HEK293T cells. When wild-type endothelial cells or PSR down- regulated endothelial cells were co-cultured with wild-type or annexin-1 silenced apoptotic Jurkat

cells, reduction of tethering was seen in single silencing experiments where PSR in engulfing cells or annexin-1 in apoptotic Jurkat cells were tested. These results suggested that PSR and annexin-1 both have crucial role in the same engulfment pathway.

1.3 Gene deletion studies of PSR/JMJD6 in model organisms

In order to investigate the role of PSR in apoptosis and its physiological relevance *in vivo*, gene deletion studies of PSR have been carried out in various model organisms such as *C. elegans*, zebrafish, *Drosophila* and mice.[3-6, 14]

Wang et al. investigated phagocytosis in *C. elegans* by deletion of the PSR1 gene locus thereby creating a phenotype lacking PSR1 protein.[14] These authors reported relatively mild engulfment defects and accumulation of cell corpses in mutant worm embryos. They further showed that mutant strains could be rescued from these defects by overexpression of PSR1. Hong et al. studied the effects of knock down of zebrafish PSR with morpholino oligonucleotides.[3] They reported severe developmental defects in somites, brain, heart and notochord in PSR deficient zebrafish embryos. Weakly defective phenotypes showed bending of notochord and slight delay in heart development prior to hatching while severely defective embryos displayed shrinkage of brain, loss of posterior somite development and failure in hatching. Defective zebrafish embryos at 3 days post fertilization stage had tube like heart and abnormal blood circulation rate with some having a heart cavity that was enlarged upto 60% as compared to normal embryos.

Three research groups have independently carried out PSR deletion studies in mice.[4-6] In all the studies ablation of PSR in knockout mice caused perinatal lethality and displayed severe differentiation defects causing malformation of vital organs during embryogenesis. PSR deficient mice displayed extensive brain malformations resulting in externally visible head abnormalities as shown in Figure 1.2. Both Böse et al. and Li et al. observed defects in ocular development with protrusion of the retinal neuroepithelium and underdeveloped lens with complete anophthalmia in severe cases which was associated with induction of ectopic eye structures in nasal cavities.[4, 6] Morphological analysis of PSR deficient embryos showed a significant delay in organ and tissue differentiation in the lungs, kidneys and intestine. Kunisaki et al. also reported impairment of thymocyte differentiation and poor development of the thymus.[5] Li et al. showed that during foetal development of PSR deficient mice, lumen formation of lungs was severely impaired and these mice were unable to breathe within few hours of birth. Kunisaki et al. and Böse et al. also observed blocked erythropoietic differentiation at erythroblast stage in liver of these mice explaining the anaemic appearance observed in all the studies. Böse et al. further reported that PSR deficient homozygotes died as a result of cardiopulmonary malformations.

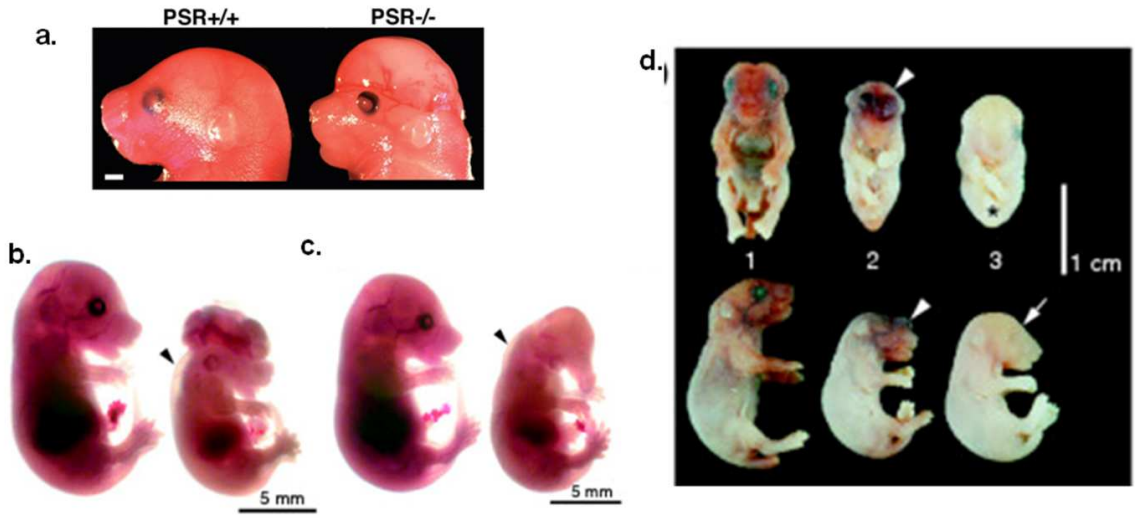


Figure 1.2 Developmental abnormalities in PSR deficient mice, wild-type littermates are located on the left and homozygous mutant mice on the right. (a) Embryos E17.5 with abnormal head shape. (b) Embryos E15.5 show exencephaly and an abnormal head shape with proboscis and absence of eyes. or (c) prosencephalic hernia in the forebrain region and absence of eyes. (d) Littermates at birth with uni- or bilateral absence of the eyes, edema and anaemia. Images are adapted from Li et al.[4] and Böse et al.[6]

Further, PSR gene knockout studies were carried out by Krieser et al. in *Drosophila*, in order to study the involvement of *Drosophila* PSR (dPSR) in apoptosis.[7] They reported increased apoptosis in developing pupal eye of the flies that lack dPSR. By extensive analysis of mutants of dPSR, they determined that PSR is not required for engulfment of apoptotic cells by macrophages in the developing embryos, on the contrary inhibition of apoptosis by increased dPSR was observed. These studies also indicated that dPSR is localised to the nucleus of drosophila cells and is unlikely to act as a cell surface receptor involved in apoptosis.

1.4 PSR/JMJD6 is not involved in apoptosis

In all of the PSR knockdown mice, phagocytosis analysis was carried out to investigate the involvement of PSR in apoptosis.[4-6] Li et al. and Kunisaki et al. reported apoptotic cell englufment defects in the case of PSR deficient mice. On the contrary, Böse et al.

did not observe any impairment in the uptake of apoptotic cells by macrophages from PSR deficient mice. Böse et al. used monoclonal antibody mAb217 for immunostaining of PSR in the wild-type and PSR deficient mice and they observed similar staining patterns in both genotypes. They also reported that the mAb217 antibody bound weakly to the PSR but also recognised an additional membrane associated protein. Many of the previously reported functional studies had been based on the use of monoclonal antibody mAb217; hence, this work raised important questions about the involvement of PSR in the recognition of apoptotic cells.

Later Mitchell et al. re-evaluated phagocytosis from the liver derived PSR deficient macrophages of mice embryos.[15] They used 3T3 fibroblast cell lines from $PSR^{+/+}$, $PSR^{+/-}$, $PSR^{-/-}$ mice and observed that these fibroblasts were fully competent for apoptotic phagocytosis, regardless of PSR expression. The absence of PSR expression had no effect on the target cell uptake indicating clearly that PSR serves no direct role in phagocytosis. Other signalling events involved in apoptosis such as Rac activation in response to αVβ5-integrin engagement, modulation of the nuclear factor-κB (NFκB)-dependent transcription after treatment with apoptotic target cells, activation of protein kinase B after incubation with apoptotic cells were also unaffected. This study clearly implies that PSR is not involved in the recognition or engulfment of apoptotic cells.

1.5 PSR/JMJD6 is a nuclear protein

Cui et al. and Cikala et al. independently carried out cellular localisation studies of PSR.[16, 17] It is known that proteins are targeted to various intracellular compartments by specific signals such as signal peptides or nuclear localisation signals (NLS).[18, 19] Bioinformatics analysis carried out by Cikala et al. for PSR sequence from *Hydra*

revealed several putative NLS, i.e. PKKRK (141-145), KRRPP (167-171) in the predicted cytoplasmic domain and KRRK (377-380) in earlier predicted extracellular domain.[17] These NLS were also found to be conserved in PSR sequences varying from *Hydra* to human. However, none of the transmembrane prediction programs used, identified a transmembrane domain in the previously predicted region or anywhere else in the PSR sequence of *Hydra*. Further, they expressed the PSR mutants by deleting these NLS and adding *N*-terminal GFP-tag (as GFP-PSRΔNLS141-145, GFP-PSRΔNLS167-171 and GFP-PSRΔNLS377-380). Subcellular localisation of these GFP fusion proteins was later observed in animals and after fixation by confocal microscopy. GFP-tagged full-length PSR was reported to be localised to the nucleus exclusively. Deletion of both NLS1 and NLS2 sequences did not change the nuclear localisation of the protein. However, deletion of the most *C*-terminal NLS compromised nuclear localisation of the fusion protein dramatically. In the case of GFP-PSRΔNLS377-380, GFP is found to be distributed throughout the cell, both in fixed and in non-fixed hydra cells. An additional GFP fusion protein only containing the JmjC domain of PSR showed nuclear localisation. Interestingly, none of the mutants were localised to the membrane. From these data it was concluded that *Hydra* PSR is normally localised in the nucleus and the nuclear localisation of *Hydra* PSR depends on a *C*-terminal NLS. Further analysis of GFP-tagged PSR in transfected *Hydra* cells, after induction of apoptosis showed that PSR remains in the nuclei of cells undergoing phagocytosis; thereby supporting the proposal that *Hydra* PSR is not a membrane receptor involved in phagocytosis.

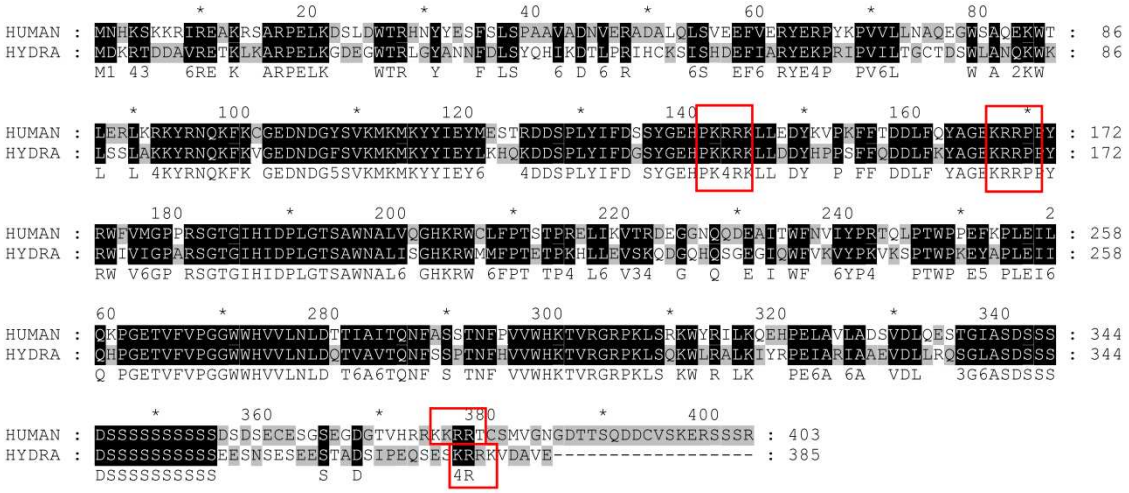


Figure 1.3 Sequence alignment of Human PSR and hydra PSR sequences as calculated by ClustalW. Three NLS from hydra which are conserved in human PSR sequence are shown in red block.

To investigate whether PSR is a cell surface protein, Cui et al. constructed a fusion between human PSR sequence and GFP and added a FLAG tag to the carboxyl end of GFP and transfected MDCK, PC3, HEK293T, HeLa and NIH 3T3 cells.[16] Interestingly, when these cells were analysed by fluorescence microscopy, this fusion protein was not localised on the cell surface as predicted, rather, it was concentrated exclusively in the nuclei of transfected cells as shown in Figure 1.4, suggesting that PSR may be is a nuclear protein. Cui et al. further predicted that human PSR can encode five potential NLS as PSR 6-10, PSR 91-95, PSR 141-145, PSR 167-170 and PSR 373-378, named NLS 1 to 5 and these signals can potentially deliver PSR to the nuclei either alone or in concert. Similar to the approach of Cikala et al., Cui et al. also performed deletion analysis and fused the deleted fragments to GFP at the *C*-termini as followed by analysis of their cellular localisations using fluorescent microscopy analysis of transfected HeLa cells. Except for the fragment which lacked any of the NLS, all of the fusions were localised in the nuclei, suggesting that any of the five NLS can target PSR

to the nuclei. To test this possibility, all five NLS were inserted individually into the *N*-terminus of GFP and resulting constructs were used to transfect the HeLa cells. Confocal microscopy of these transfected cells revealed that all five GFP fusions were localised in the nuclei and each of the NLS was sufficient to target overexpressed GFP-PSR to the nucleus of HeLa cells. This data supports the proposal that human PSR sequence is a nuclear protein.

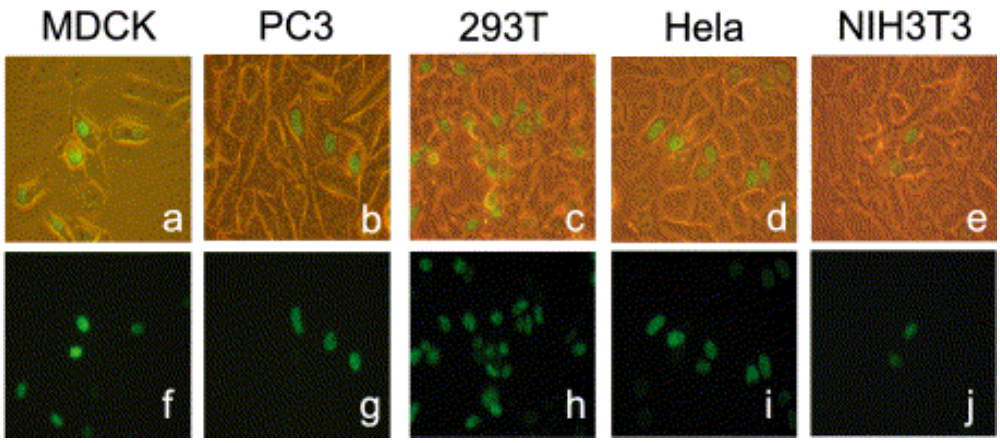


Figure 1.4 Nuclear localisation of PSR-GFP-Flag tagged construct as shown by Cui et al.[16] The PSR-GFP-Flag vector was introduced into (a) and (f) MDCK; (b) and (g) PC3; (c) and (h) HEK293T; (d) and (i) HeLa; (e) and (j) NIH3T3 cells. Images were taken in bright field and fluorescent field (f to j). Both bright field and fluorescent field were combined to show the localisation of PSR-GFP in the nuclei (panels a to e). This image has been adapted from reference [16].

1.6 PSR/JMJD6 and the presence of JmjC domain

Based on a bioinformatics analysis of the cloned *Hydra* homologue PSR, Cikala and colleagues were the first to suggest that the PSR might function as a 2-oxoglutarate (2OG) dependent dioxygenase that is capable of modifying unknown nuclear targets.[17] *Hydra* cDNA encoding PSR with high homology to the mammalian, *C. elegans* and *Drosophila* genes was identified and three NLS which were conserved between PSR from all species. The presence of an AT-hook DNA binding motif was also predicted.

Most importantly, the well conserved characteristic cupin fold which is present in members of the Jumonji C domain family (JmjC) of transcription factors was predicted in the PSR sequence. Comparing domain structures of proteins containing JmjC domain, revealed striking similarity of PSR with the factor-inhibiting hypoxia-inducible factor (FIH). FIH is an Fe(II) and 2OG dependent oxygenase catalysing asparginyl hydroxylation of hypoxia-inducible factor (HIF), thereby playing an important role in the cellular oxygen sensing mechanism.[20, 21]

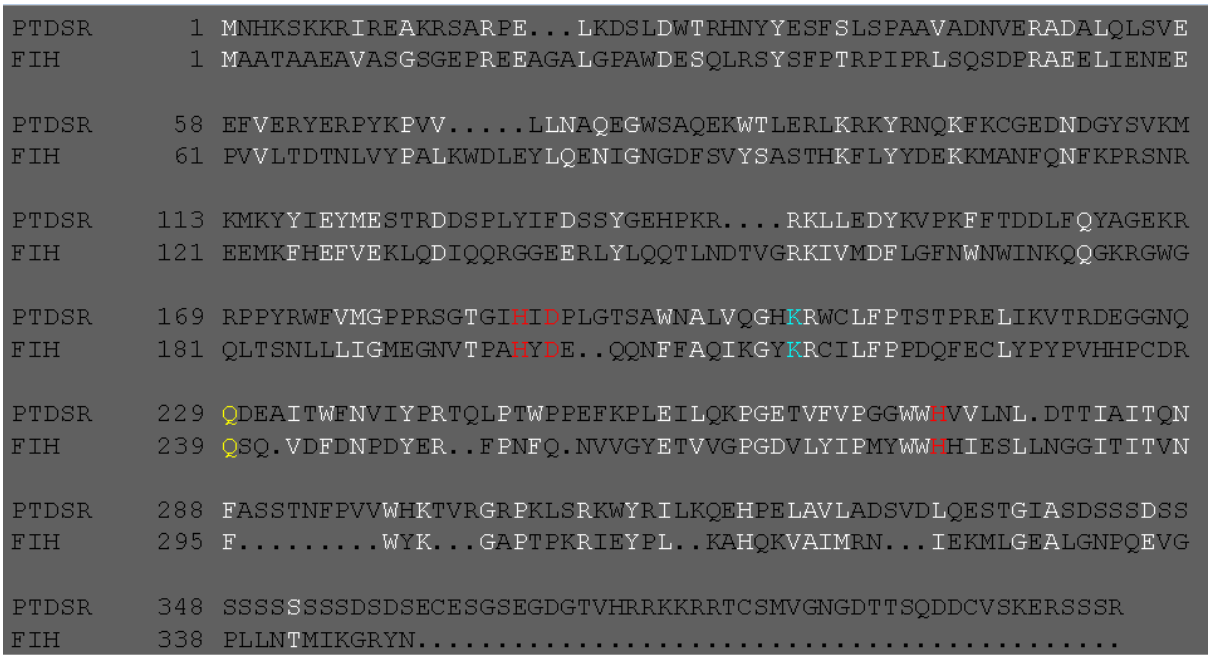


Figure 1.5 Sequence alignment of PSR and FIH showing conserved Fe(II) binding residues from HxD/E...H triad in red and 2OG binding lysine residues in blue.

There is a clear similarity in the predicted arrangement of β-strands between the central domain of PSR and that of FIH,[2] which was later confirmed by crystallography as described in Chapter 3. Moreover, closer analysis of the 3-D structure of the β-barrel that is formed by depicted β-strands of FIH revealed hydrophobic residues that stabilise the barrel and these residues are also found to be conserved between PSR and FIH,[17]

making it very likely that conserved core domains of both proteins adapt the same fold. Moreover, residues which form the Fe(II) binding HxD/E…H triad, a motif comprising two histdines and a carboxylate of 2OG oxygenases are also conserved. H199 and D201 and the *C*- terminal histidine H279 of FIH form this triad (Figure 1.6). Sequence alignment of FIH and PSR (Figure 1.5) predicts H187, D189 and H273 in PSR to coordinate Fe(II) at the active site. A lysine residue K214 that binds 2OG at the active site in FIH is also predicted to be conserved in PSR (K204).

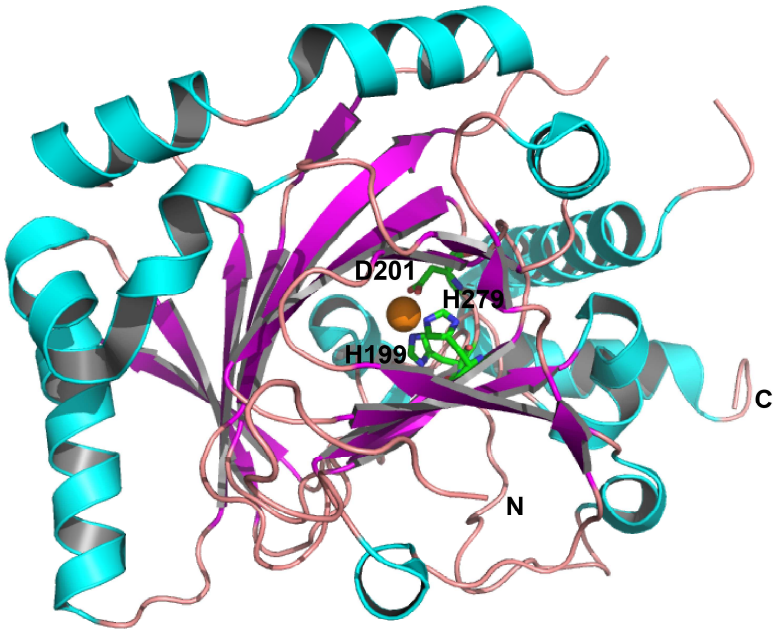


Figure 1.6 Ribbon representation of a crystal structure of FIH (PDB ID 1H2K).Fe(II) (shown in orange) is bound by conserved HxD...H motif formed by residues H199, D201 and H279 shown by green sticks. The active site is surrounded by double stranded β-helix structure represented by magenta sheets.

1.7 General overview of 2OG and Fe (II) dependent oxygenases

The family of oxygenases that employs 2OG as a cosubstrate and iron as a cofactor is ubiquitous in aerobic life forms.[22, 23] Members of the 2OG oxygenase family couple the

two electron oxidation of their substrate with the reaction of oxygen and 2OG, a Kreb's cycle intermediate to give carbon dioxide and succinate.[24] These enzymes primarily catalyse hydroxylation reactions of their primary substrates, during which one oxygen atom from molecular oxygen is incorporated into the alcohol product, with the other oxygen atom incorporated into the succinate co-product. The extensive crystallographic evidence implies that most, if not all 2OG oxygenases have conserved structures based on a double-stranded β-helix (DSBH) fold (or jelly roll or cupin fold) that supports the conserved iron and 2OG binding sites.[25, 26] The iron cofactor is normally coordinated by three protein ligands comprising the HxD/E…H triad formed by the well conserved motif of two histidines and one carboxylic acid, either aspartate or glutamate. During the catalytical cycle, ferrous iron binds first to the apo-enzyme followed by 2OG. The 2OG binds to the iron in a bidentate manner via its 1-carboxylate and ketone carbonyl groups.[24] Crystallographic studies reveal that the 2OG C-5 carboxylate is, at least typically, positioned to form an electrostatic interaction with a lysine or arginine residue and a hydrogen bonding interaction with one or more alcohol/phenol side chains.[25, 26] Subsequent to oxygen binding, oxidative decarboxylation of 2OG occurs to give carbon dioxide, succinate and a Fe(IV)-oxo intermediate for which there is spectroscopic evidence.[27, 28] The Fe(IV)-oxo species can then react with the substrate to effect hydroxylation or other oxidative chemistry. Release of the product followed by succinate then occurs. A general mechanism of 2OG oxygenases catalysed hydroxylation is outlined in Figure 1.7.

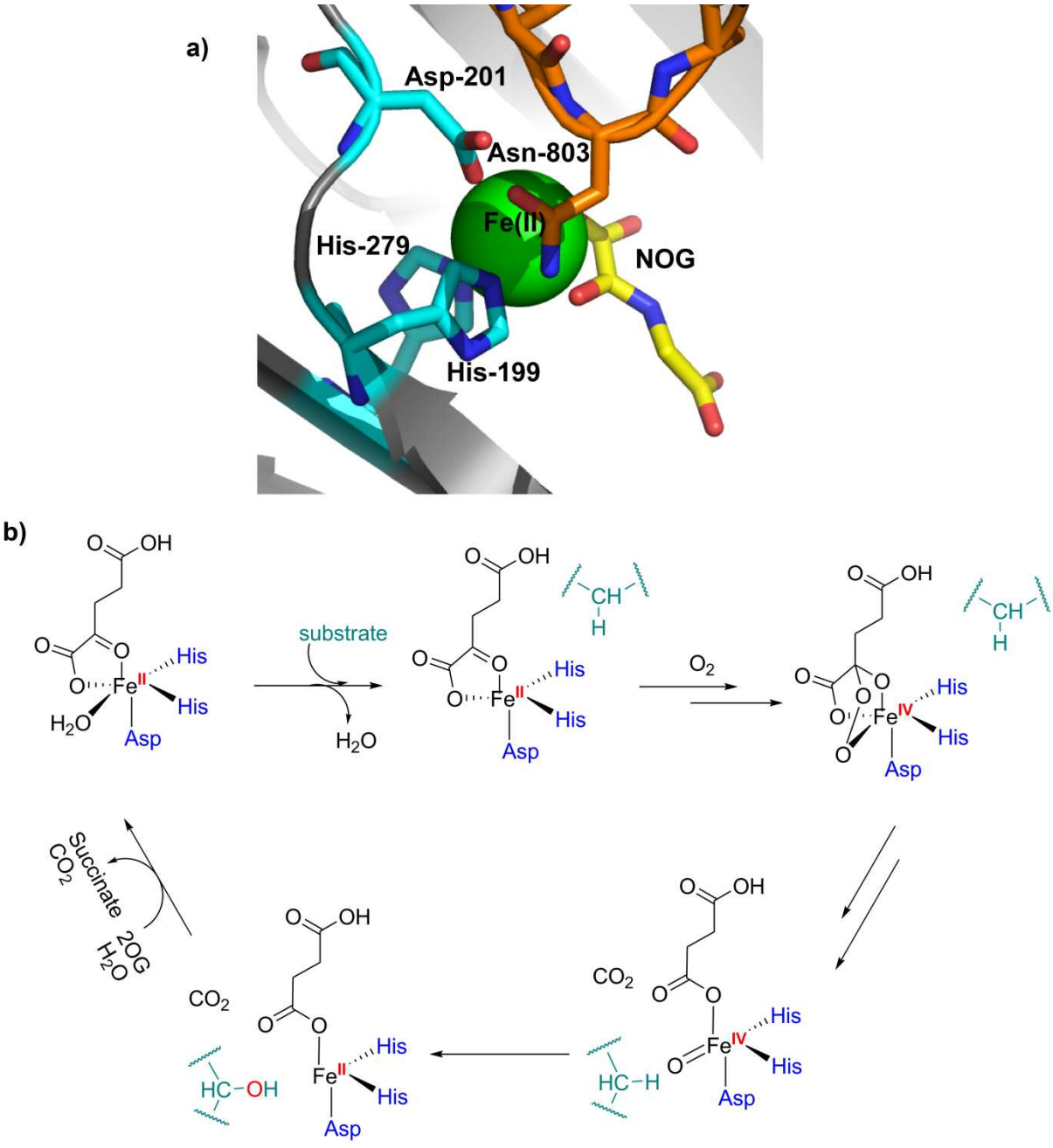


Figure 1.7 Catalysis by 2OG oxygenases. (a) View from a crystal structure showing the active site of FIH in complex with a fragment of HIF-1α (yellow), Fe(II) (orange) and *N*-oxalylglycine (NOG, green) which is a 2OG analogue. His199, Asp201 and His279 form the HxD…H motif that coordinates the Fe(II). The substrate residue Asn803 that is hydroxylated is adjacent to the Fe(II); (c) Outline of the proposed catalytic mechanism for 2OG oxygenases. The iron is bound by the HxD/E…H facial triad and the remaining coordination sites are occupied by 2OG and water. Water is displaced upon substrate binding, allowing oxygen to bind and subsequently react to give the Fe(IV)=O species.

2OG oxygenases catalyse a very wide range of oxidative reactions including *N*-methyl and *O*-methyl demethylation via initial hydroxylation, desaturation, oxygen insertion, epoxidation, oxidative rearrangement/epimerisation, oxidative halogenations.[23] In a recently reported case, one enzyme catalyses oxidation of a clinically used inhibitor via a Stevens type reaction.[29] The diversity of biological roles of the 2OG oxygenases apparently mirrors their catalytic flexibility. In humans, and likely most other animals including the simplest animal *Trichoplax adhaerens*,[30] they have a range of biological roles including in collagen biosynthesis, DNA repair, RNA modification, chromatin regulation/histone modification, epidermal growth factor domain like protein modification (aspartyl- and asparaginyl- hydroxylation), hypoxia sensing and in fatty acid metabolism.[31] 2OG oxygenases likely also have additional biological roles because not all of the 60-70 human 2OG oxygenases have been assigned biological functions.[23] In microorganisms and plants, 2OG oxygenases play many roles in the biosynthesis of secondary metabolites such as flavanoids and signalling molecules i.e. ethylene and gibberellins.

1.7.1 Biosynthesis of small molecules and antibiotics

The catalytic flexibility of 2OG oxygenases in terms of reaction type catalysed is perhaps best illustrated by considering their roles in antibiotic biosynthesis. During the biosynthesis of the cephalosporin nucleus, they catalyse formation of the bicyclic penicillin ring structure (isopenicillin N synthase, (IPNS)), the oxidative rearrangement of the penicillin nucleus to give a cephem ring, and modification of the latter by allylic hydroxylation (Figure 1.8a, deacetoxy-cephalosporin synthase (DAOCS) and deacetyl-cephalosporin synthase (DACS)).[32, 33] In the case of clavulanic acid biosynthesis, a

single 2OG oxygenase, clavaminic acid synthase (CAS) catalyses three steps comprising hydroxylation, oxidative ring closure and desaturation reactions (Figure 1.8b).[34] A similarly high degree of flexibility is demonstrated in plants in 2OG oxygenase catalysed reactions, viz. in the biosynthesis of the giberellin signalling molecules and the flavonoids.[35, 36] Several members of the 2OG oxygenase family that catalyse oxidative reactions do not use 2OG as a cosubstrate, but have clearly related structures. These enzymes include IPNS[37] and 1-amino-1-cyclopropanecarboxylic acid oxidase (ACC oxidase), which catalyses the final step in ethylene biosynthesis in plants.[38]

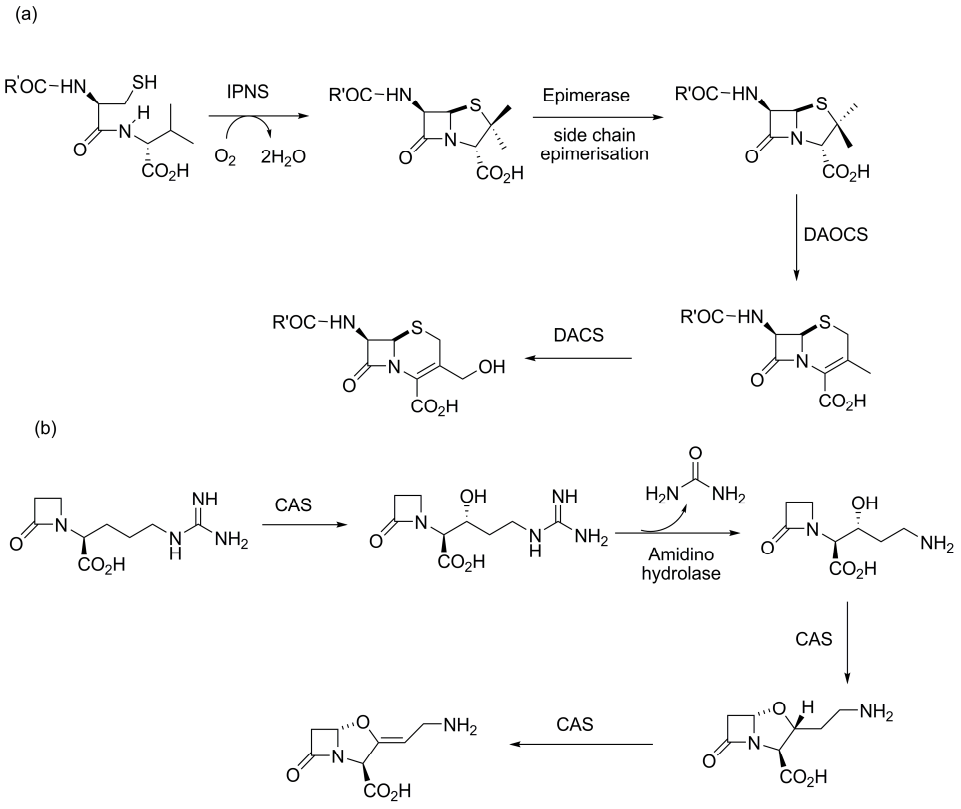


Figure 1.8 Examples of catalysis by 2OG oxygenases and related enzymes involved in antibiotic biosynthesis. Reactions catalysed by (a) IPNS, DAOCS, DACS and (b) CAS. Each of the CAS, DAOCS, DACS reactions is coupled to the conversion of 2OG to succinate and CO_2. H_2O is produced in the cases of reactions involving desaturation.

1.7.2 Collagen post-translational modifications: lysine/proline hydroxylation

The first 2OG oxygenases to be identified were involved in collagen biosynthesis. Collagens are a family of proteins found essentially in all animal tissues. They form a diverse range of highly organised supramolecular assemblies in the extra-cellular matrix and generate structural scaffolds to support cells within tissues, connect tissues within an organ and facilitate attachment and migration of cells.[39] The biosynthesis of collagens involves many post-translational modifications, which include hydroxylation of prolyl- and lysyl- residues and glycosylation of hydroxylysyl residues occurring in the endoplasmic reticulum.[40] Collagen prolyl and lysyl hydroxylases are 2OG oxygenases. Prolyl 4-hydroxylation is the most frequent collagen modification and occurs during elongation of the nascent chains in the endoplasmic reticulum. The 4-prolyl hydroxylases have important function because 4*R*-hydroxy proline in -X-Pro-Gly- triplet stabilises the triple helix of collagen under physiological conditions.[41] 3*S*-Prolyl hydroxylation also occurs, but is less common. Prolyl 3-hydroxylases have been reported to be an important component of a molecular chaperone complex that functions in moving the newly synthesised collagen through the endoplasmic reticulum.[42]

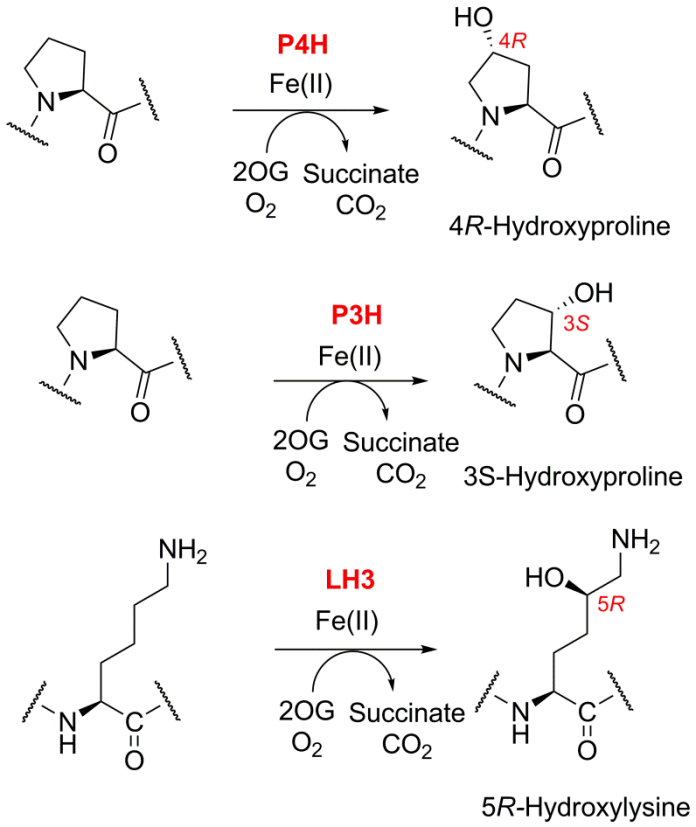


Figure 1.9 Reaction schemes highlighting the roles of the prolyl- and lysyl- hydroxylases, involved in the post-translational modification of collagen.

Lysyl hydroxylases catalyse the formation of hydroxylysine in collagens and other proteins with collagen-like sequences by the hydroxylation of lysine residues in -X-Lys-Gly- triplets.[43] In addition, the telopeptides of some fibril forming collagens contain hydroxylysine in an -X-Hyl-Ala- or -X-Hyl-Ser- sequence. The hydroxylysine residues have two important functions viz., they are essential for the formation of stable intra- and inter-molecular collagen crosslinks, and these hydroxy groups serve as attachment sites for galactose or glucosylgalactose groups.

1.7.3 Oxygen sensing

The maintenance of cellular oxygen homeostasis is a critical process as either excess (hyperoxia) or low oxygen (hypoxia) levels are harmful. Due to the vital importance of oxygen in both local and systemic oxygen homeostasis is tightly controlled in animals through a number of pathways capable of responding to either chronic or acute changes in oxygen availability. [44] The master regulator of one such cellular hypoxic response pathway involves a transcription factor called hypoxia-inducible factor (HIF). Transcriptional activity of HIF signalling pathway is regulated by the post-translational hydroxylation of specific residues which are catalysed by a set of Fe(II) and 2OG dependent oxygenases. HIF is a heterodimeric transcription factor that is composed of two basic helix-loop-helix proteins, HIF-α and HIF-β. HIF-α is selectively hydroxylated under normoxia at two proline residues (Pro402 and Pro564) within the oxygen-dependent degradation domain (ODD) by a family of three *trans*-4-prolyl hydroxylases (PHDs).[45, 46] Modification of these residues is necessary for the recognition of HIF-α subunit by the product of the von Hippel-Lindau gene, a component of the E3 ubiquitin ligase complex that targets HIF-α for proteasomal degradation (Figure 1.10).[47, 48] Under hypoxic, conditions PHD activity is inhibited by limiting oxygen concentrations.

FIH is another 2OG oxygenase which catalyses asparginyl hydroxylation in the *C*-terminal activation domain (CAD) of HIF-α.[49, 50] Hydroxylation of an asparagine residue Asn803 under normoxia creates steric clashes responsible for abrogating the interaction of HIF-α with the coactivator p300, thereby inhibiting hypoxia.[50] Thus, non hydroxylated HIF-α can form complex with HIF-β and p300, enabling binding to hypoxia response elements and thereby activating the hypoxic response cascade by

transcription of many genes such as erythropoietin, vascular endothelial growth factor (VEGF) etc.

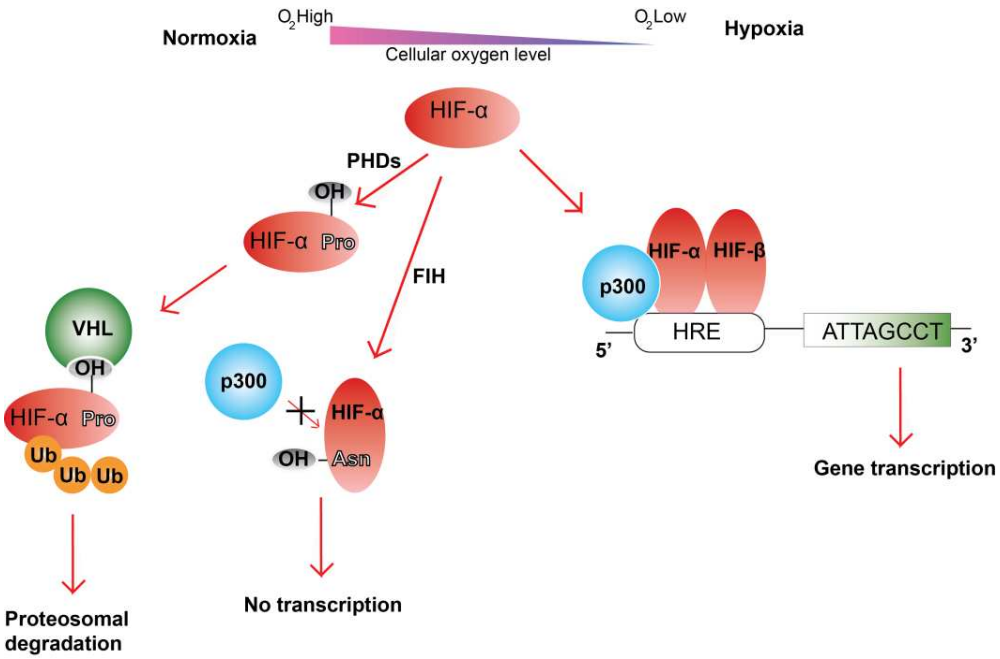


Figure 1.10 Schematic representation of the HIF regulated hypoxic response pathway. Under normoxic conditions, HIF is hydroxylated leading to either proteasomal degradation or blocking its interaction with p300. Under hypoxic conditions, the p300-HIFα-HIFβ complex that initiates gene transcription from hypoxia response elements (HREs) is formed.

1.7.4 Fatty acid biosynthesis

In animals, plants and microorganisms, carnitine (L-3-hydroxy-4-*N*,*N*,*N*-trimethylaminobutyrate) plays an essential metabolic role by enabling transport of long chain fatty acid into mitochondria.[51] It is also involved in transport of oxidation products, fatty acid storage, excretion of toxic products and maintaining acyl-coenzyme-A homoeostasis. In animals, carnitine is extracted from the diet and is produced endogenous and by biosynthesis from N^{ε}-trimethyl lysine which is produced from degradation of post-translationally modified proteins. Two Fe (II) and 2OG dependent oxygenases which catalyse the steps in this biosynthesis of carnitine (Figure 1.11a) are

N^{ε}-trimethyllysine hydroxylase (TMLH) and γ-butyrobetaine hydroxylase (BBOX).[52, 53] Recent crystal structure analyses and mechanistic studies have shown the mechanism of inhibition of BBOX by 3-(2,2,2-trimethylhydrazine)-propionate (THP) which is used to treat myocardial infarction.[29]

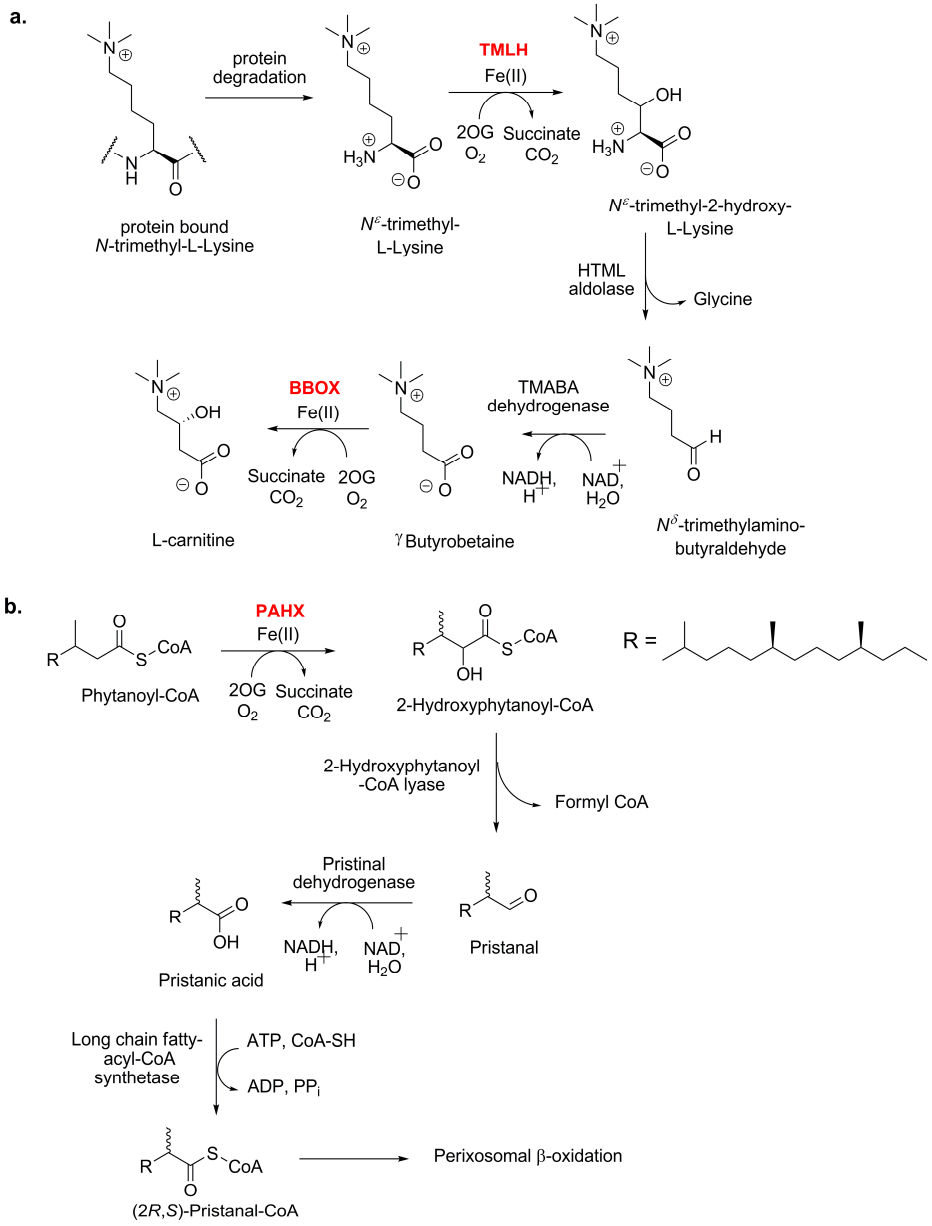


Figure 1.11 Reaction schemes highlighting the roles of 2OG oxygenases involved in fatty acid biosynthesis. (a) The role of BBOX in the carnitine biosynthesis pathway. (b) The role of PAHX in metabolic pathway of chlorophyll.

Phytanoyl CoA hydroxylase (PAHX or PHYH) catalyses an important step in the modification of fatty acid chlorophyll side chain in animals i.e. the hydroxylation of Phytanoyl CoA (Figure 1.11b).[54, 55] Mutations to the human DNA sequence encoding for PAHX can cause Refsum Disease, the symptoms of which are attributed to increased levels of phytanoic acid or its derivatives.[55] Phytanic acid is an intermediate in chlorophyll digestion and cannot be processed by normal β-oxidation mechanism as this is blocked by the β-methyl group. Hydroxylation of phytanoyl CoA allows removal of the α-carbon atom to form pristanic acid which is later metabolised by a β-oxidation mechanism.

1.7.5 Histone modifications

The basic unit of chromatin consists of a 146 bp of DNA wrapped around a histone octamer which is composed of two copies of each of four core histones H2A, H2B, H3 and H4. Covalent histone modifications play an important role in chromatin regulation, gene expression and epigenetic inheritance.[56, 57] Identified post-translational modifications on *N*-terminal histone tails include phosphorylation, ubiquitinylation, acetylation and methylation. N^{ε}-Methylations of lysines or arginines in histone tails can regulate transcription via steric factors or by facilitating or blocking interaction with other histone modifying/binding proteins thereby regulating a wide range of processes including gene activity, chromatin structure, dosage compensation and epigenetic memory.[58] Histone lysine methylations on H3K4, H3K9, H3K27, H3K36, H3K79, and H4K20 have different biological effects.[59] In general, methylations on H3K4, H3K36, and H3K79 have an activating effect on gene expression, whereas methylation of H3K9, H3K27, and H4K20 impose a repressing effect on transcription.[59]

Until recently, histone methylation was thought to be irreversible. Two families of histone N^{ε}-methyl lysine demethylases have been identified, the largest of which are members of JmjC subfamily of Fe(II) and 2OG dependent oxyegnases.[60] These enzymes catalyse hydroxylation of N^{ε}-methyl groups to give a hemiaminal intermediates which subsequently fragment to give formaldehyde and the demethylated product as shown in Figure 1.12.[61, 62] Different subfamilies of human JmjC histone demethylases have different selectivities for particular histone residues and methylation states. For example, members of the JmjC domain containing protein 2 (JMJD2) subfamily are selective for demethylation of the tri- and di-N^{ε}-methylation states of specific lysines on histone H3 (although they accept the mono-methylated state weakly *in vitro*),[61] whereas other subfamilies (e.g. the *F*-box and leucine-rich repeat containing protein (FBXL) and PHD finger containing protein 8 (PHF8) subfamilies) are selective for the di- and mono-N^{ε}-methylated states and do not accept the trimethylated state.[63-65] The functional characterisation of Fe(II) and 2OG dependent histone demethylation in chromatin regulation is currently ongoing.

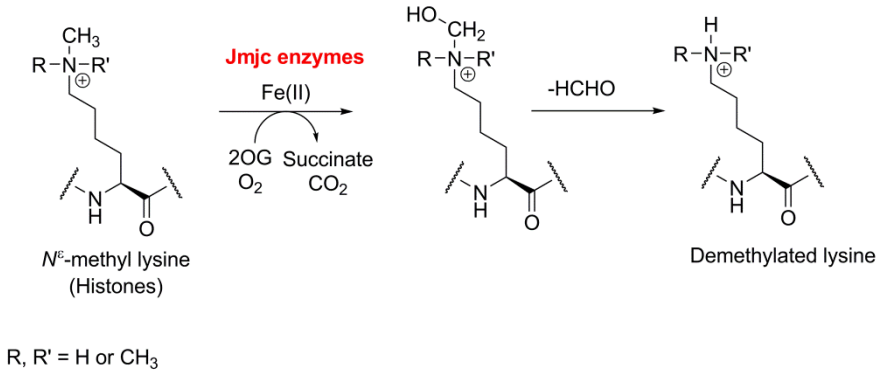


Figure 1.12 Outline of lysine demethylation catalysed by histone demethylases.

1.7.6 Nucleic acid modifications

One of the well characterised 'epigenetic' modification to DNA is the formation of 5-methyl cytosine from cytosine.[31] Many (but not all) other types of DNA methylations are toxic. In *Escherichia coli*, the 2OG oxygenase AlkB catalyses repair of *N*-alkylated single-stranded DNA caused by alkylating agents. AlkB catalysed reaction proceeds via methyl hydroxylation followed by fragmentation with the release of formaldehyde, similar to the proposed histone demethylase mechanism.[23] AlkB accepts 1-methyl adenine, 3-methyl cytosine and less efficiently 1-methyl guanine and 3-methyl thymine as substrates. Eight human AlkB homologues (ABH1-8) have been identified, of which ABH2 and ABH3 have similar functions to AlkB, catalysing the oxidative demethylation of 1-methyl adenine and 3-methyl cytosine lessions.[66] ABH2 and ABH3 have been reported to have *N*-demethylation activity for both DNA and RNA[66] where ABH2 prefers single stranded DNA while ABH3 shows distinct preference for single stranded DNA. Due to their involvement in repair mechanism of mutagenic and toxic DNA lesions, ABH2 and ABH3 have been linked to cancer and point mutations of ABH2 or ABH3 have been identified in paediatric brain tumours.[67] Elevation of ABH3 protein levels has been observed in case of human prostate carcinomas.[68] Another protein of human AlkB family, ABH8 has been linked to human bladder cancer.[69] ABH8 in addition to 2OG oxygenases core, comprises of *N*-terminal RNA recognition motif and *C*-terminal; S-adenosylmethionine (SAM) dependent methyl transferase domain which has been reported to catalyse the methylation of 5-carboxymethyluridine, specifically in tRNAs containing wobble position in the anticodon loop.[70]

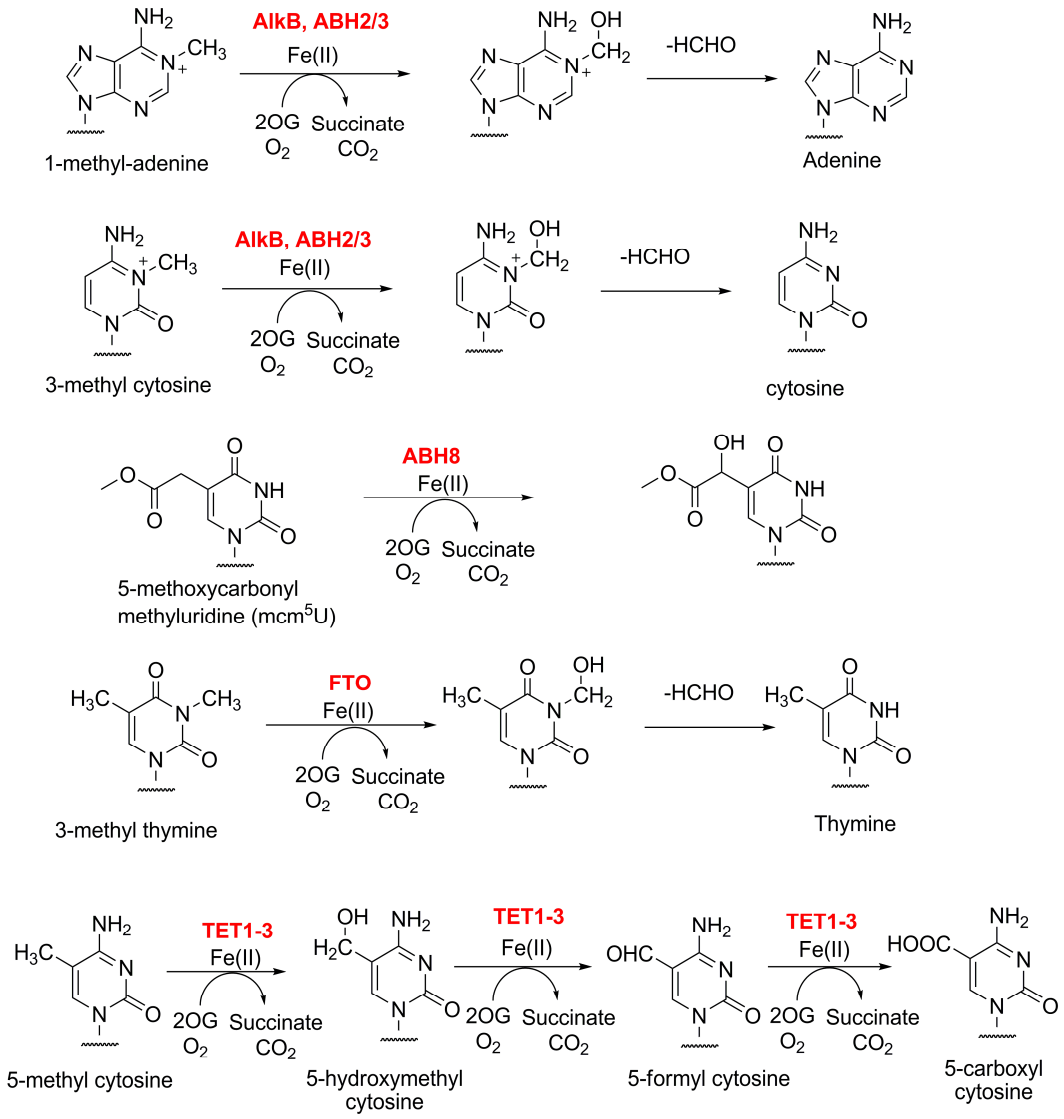


Figure 1.13 Outline of 2OG oxygenases catalysed oxidative modifications to DNA and/or RNA.

Recent work has identified the Ten-Eleven-Translocation 1-3 (TET1-3) protein family to be involved in nucleic acid modifications. The TET1-3 enzymes were found to catalyse hydroxylation of 5-methyl cytosine in DNA to 5-hydroxymethyl cytosine (Figure 1.13) which could be further oxidised to give 5-formyl cytosine and 5-carboxycytosine.[71] The mutations in TET2 are associated with reduced hydroxylation of

5-methyl cytosine in myeloid malignancies, thus suggesting a link between nucleic acid modification by TET2 and cancer.[72]

One of the other 2OG oxygenases enzyme involved in nucleic acid demethylation is the fat mass and obesity associated protein (FTO). FTO accepts *N*-methylated DNA, RNA and nucleosides as demethylation substrates.[73] Genome wide association studies on Type 2 diabetes had revealed an association between specific human variants in the *FTO* gene and obesity. It is known that adults who are homozygous for obesity related *FTO* variants weigh about 3 kg more than those homozygous for the low risk allele[74] and studies carried out in mice suggest that FTO inhibition can cause weight loss.[75] Bioinformatics analysis imply that the FTO sequence is closely related to that of ABHs.[73] However, the mechanism by which demethylation relates to body mass has not yet been identified.

1.8 JMJD6 as arginine demethylase

On the basis of the presence of JmjC domain in PSR, it was renamed as Jumonji domain containing gene 6 in 2007 by the International committee for standardised genetic nomenclature (ICSGNM);[8] due to its nuclear localisation in cells, it was proposed to be capable of catalysing histone demethylation similar to other JmjC domain containing enzymes. In 2007, Chang et al. reported histone arginine demethylation activity of JMJD6 for methylated histone peptides by *in vitro* assays and comparing them with products of inactive JMJD6 mutants.[76] To the extent of my knowledge, attempts to reproduce these results by our laboratory and other laboratories have failed, raising questions about *in vivo* substrates and binding partners of JMJD6. Details of this analysis are discussed in Chapter 2.

1.9 JMJD6 is a lysine hydroxylase

To investigate the endogenous function of JMJD6, when Webby et al. analysed binding partners of JMJD6 in HEK293T cells using tandem-affinity purification coupled with mass spectrometry (MS) analyses; many of the identified potential JMJD6 interacting proteins were found to be involved in RNA metabolism, processing, and splicing with ~25% being nuclear proteins possessing Arg-Ser (RS) domains.[77] The interaction of JMJD6 with the splicing factor U2 small nuclear ribonucleoprotein auxiliary factor 65 kDa subunit (U2AF65)[78] and cisplatin resistance associated overexpressed protein (CROP)[79] which are known to be involved in mRNA splicing was also confirmed using pull-down assays in HEK293T cells. Endogenous U2AF65 and JMJD6 were found to be co-immunoprecipitated from HeLa cells with an antibody against U2AF65 and this interaction was further found to be disrupted by treatment with RNase. LC-MS/MS analysis of endogenous U2AF65 from HeLa cells revealed hydroxylation at Lys15 and Lys276. Further, MALDI-MS based assays carried out by incubation of the RS domain sequence peptides from U2AF65, and LUC7-like2 (LUC7L2) in the presence of JMJD6 showed +16 Da mass shift corresponding to hydroxyl group. When these assays were carried out in the presence of $^{18}O_2$, the origin of introduced hydroxyl group in lysine was found to be from dioxygen. MS/MS analyses of $LUC7L2_{267\text{-}278}$ which was incubated with JMJD6 in the presence of Fe(II) and 2OG, indicated the oxidation of Lys269 in LUC7L2 peptide; nuclear magnetic resonance (NMR) analyses demonstrated oxidation to occur at C-5 on the side chain of lysine. This data indicated that JMJD6 is capable of catalysing lysine hydroxylation from RS domain of proteins involved in mRNA splicing in an Fe(II) and 2OG dependent manner.

Webby et al. also attempted to investigate effects of JMJD6 on splicing of various genes in endogenous cells by knockdown and siRNA studies.[77] JMJD6 knockdown had altered the alternative splicing pattern of the endogenous tumor antigen gene MGEA6, which is regulated by SF2/ASF, protein involved in mRNA splicing.[80] The authors also overexpressed α-tropomyosin (α-TM) minigene construct in mock-treated cells and cells treated with siRNA targeting JMJD6 and MAZ4. This α-TM minigene construct recapitulates splicing regulation of the α-TM gene and contains exons 1, 3, and 4 surrounded by their intronic regulatory sequences; the construct is capable of generating two mRNA isoforms: 1-3-4 and 1-4 on splicing.[81] It was observed that there was a clear increase in the amount of 1-4 spliced isoform in JMJD6 knockdown cells, but not in mock-treated cells or cells treated with siRNA targeting the MAZ4 protein. Overall, these results implied the influence of JMJD6 in the regulation of alternative splicing.

1.10 Cellular localisation of JMJD6

Recently Hahn et al. reported cellular localisation and distribution studies of JMJD6 using novel antibody mAB328 in different human cell lines and mouse embryonic fibroblasts (MEF).[82] In interphase MEFs, JMJD6 showed a discrete punctual immunofluorescence pattern as seen by mAb328. This pattern was not found in regions of heterochromatic DNA organization as revealed by Hoechst DNA staining. A similar staining pattern was found in human HEK293T cells, human lung carcinoma cell line A549 and MEF cells as shown in Figure 1.14.

In order to test the involvement of JMJD6 in the demethylation of histone lysine residues, mono- , di- , and tri- methylation states of H3K4, H3K9, H3K27, H3K36, and H4K20 histones in MEFs generated from wild-type and JMJD6 knockout mice were

compared by immunoblotting. However, no differences in methylation states between *JMJD6*$^{+/+}$ and *JMJD6*$^{-/-}$ MEFs were detected in western blots of cellular lysates and specific antibodies. Similarly, overexpressing JMJD6 with a *C*-terminal Halo-tag or JMJD6 active site mutants (histidines replaced by alanines as H187A and H189A) with Halo-tag in HEK293T cells had no effect on the amount of methylation of H3K4, H3K9, H3K27, H3K36, and H4K20. These results indicated that JMJD6 is not involved in histone lysine demethylation of H3K4, H3K9, H3K27, H3K36, and H4K20 residues under *in vivo* conditions.

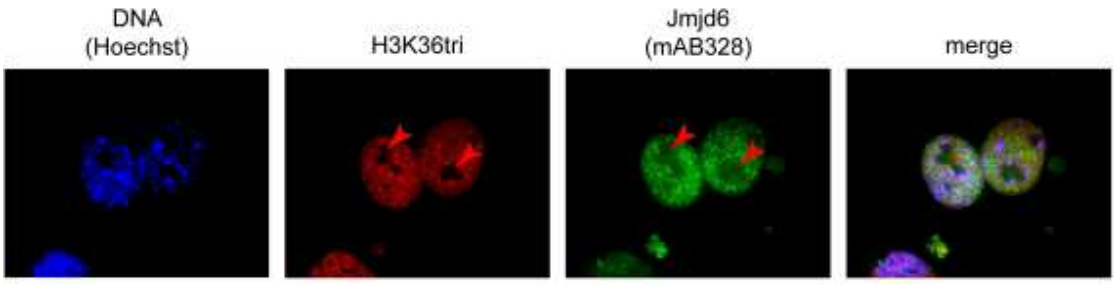


Figure 1.14 Confocal immunofluorescence image of wild-type MEF nuclei showing nuclear localisation of JMJD6. Images correspond to Hoechst DNA stain (blue), tri-methylated Histone H3 Lysine 36 (H3K36tri, red), JMJD6 (green), and a merge thereof (from left to right, respectively). Red arrowheads show the position of some nucleoli as indicated by H3K36 tri contra-staining. This image is adapted from Hahn et al.[82]

These authors further studied the distribution of JMJD6 during different stages of mitosis. The results showed that the endogenous JMJD6 protein is located in punctuated clusters in the nucleoplasm of human and mouse cells during the interphase of the cell cycle and this expression pattern is lost when cells start to divide and the nuclear membrane dissolves. During the interphase of the cell cycle, JMJD6 was occasionally observed in nucleoli suggesting that it is capable of shuttling in and out of nucleoli. Different intensities observed for JMJD6 nucleoli staining patterns, further imply that

JMJD6 can shuttle between the nucleoplasm and subnuclear compartments; the results of treatment with RNase and DNase-I indicated that the nuclear localisation of JMJD6 is dependent on an intact ribonuclear matrix in the nucleoplasm.

1.11 Involvement of JMJD6 in splicing

It has been reported recently by Boeckel et al. that JMJD6 is involved in the regulation of splicing of VEGF receptor-1 (Flt1) and thereby probably plays a role in angiogenesis in endothelial cells (ECs).[83] During these studies, when JMJD6 in ECs was silenced by siRNA, reduction in endothelial sprouting was observed and overexpression of JMJD6 gave rise to enhanced angiogenesis sprouting. Similarly, ECs derived from the lungs of $JMJD6^{+/-}$ mice showed impaired capillary network formation as compared to the cells isolated from wild-type mice.

These authors further carried out experiments to investigate effects of JMJD6 on gene expression using Affymatrix exon arrays; the results indicate that JMJD6 silencing reduces the expression level of genes which are known to play a major role in ECs while gene expression of angiogenesis inhibitors (Notch Ligand Delta like-4 (Dll4) and SEMA3A) was increased. Interestingly, further studies carried out on JMJD6 splice variant in ECs demonstrated that reduction of JMJD6 expression altered splicing of Flt1 and increased the levels of the soluble form of Flt1, which binds to VEGF and placental growth factor (PlGF) thereby inhibiting angiogenesis. Exposing ECs to hypoxic conditions or addition of 2OG inhibitor, desferrioxamine, (DFO) led to elevated levels of soluble Flt1. Immunoprecipitation studies showed that JMJD6 interacts with the splicing factors U2AF65 that binds to Flt1 mRNA, confirming the link between JMJD6 and splicing of Flt1 and angiogenesis. Thus JMJD6 is proposed to act as molecular

sensor for oxygen levels to coordinated endothelial response in case of hypoxic conditions.

1.12 Aims of the work described in this thesis

This chapter has reviewed studies on PSR/JMJD6 over the last decade. PSR/JMJD6 was initially assigned as a membrane receptor and subsequently identified as an arginine demthylase and nuclear protein involved in regulation of pre-mRNA splicing. Its exact role in RNA splicing mechanism is unknown and no work has been reported regarding the structure of JMJD6 or on its mechanism of lysine hydroxylation.

The broad aims of the work carried out in this thesis were to-

1. To test if any arginine histone demthylation activity exists for JMJD6 and to find novel *in vivo* substrates for JMJD6 involved in lysine hydroxylation by immunoprecipitation of JMJD6 from human cell lines and *in vitro* peptide analysis.
2. To attempt crystallisation of recombinant human JMJD6 in order to obtain the structural information with a view to understanding the selectivity of its interaction with SR proteins and the mechanism of lysyl hydroxylation.
3. To investigate the stereochemical assignment of the hydroxylysine product of JMJD6 catalysed hydroxylation reaction in case of LUC7L2 peptide by NMR and amino acid analyses.
4. To investigate the possible occurrence and extent of self-hydroxylation of JMJD6 as several members of the 2OG oxygenase family have been shown to catalyse self-hydroxylation in the absence of their substrates.[84-87]
5. To investigate the oligomerisation of JMJD6 by biophysical analyses.
6. To carry out preliminary inhibition studies on JMJD6.

1.13 References

1. Fadok, V. A.; Bratton, D. L.; Rose, D. M.; Pearson, A.; Ezekewitz, R. A. B.; Henson, P. M., A receptor for phosphatidylserine-specific clearance of apoptotic cells. *Nature* **2000,** 405, (6782), 85-90.

2. Wolf, A.; Schmitz, C.; Böettger, A., Changing story of the receptor for phosphatidylserine-dependent clearance of apoptotic cells. *EMBO Reports* **2007,** 8, (5), 465-469.

3. Hong, J.-R.; Lin, G.-H.; Lin, C. J.-F.; Wang, W.-p.; Lee, C.-C.; Lin, T.-L.; Wu, J.-L., Phosphatidylserine receptor is required for the engulfment of dead apoptotic cells and for normal embryonic development in zebrafish. *Development* **2004,** 131, (21), 5417-5427.

4. Li, M. O.; Sarkisian, M. R.; Mehal, W. Z.; Rakic, P.; Flavell, R. A., Phosphatidylserine Receptor Is Required for Clearance of Apoptotic Cells. *Science* **2003,** 302, (5650), 1560-1563.

5. Kunisaki, Y.; Masuko, S.; Noda, M.; Inayoshi, A.; Sanui, T.; Harada, M.; Sasazuki, T.; Fukui, Y., Defective fetal liver erythropoiesis and T lymphopoiesis in mice lacking the phosphatidylserine receptor. *Blood* **2004,** 103, (9), 3362-3364.

6. Böse, J.; Gruber, A.; Helming, L.; Schiebe, S.; Wegener, I.; Hafner, M.; Beales, M.; Kontgen, F.; Lengeling, A., The phosphatidylserine receptor has essential functions during embryogenesis but not in apoptotic cell removal. *Journal of Biology* **2004,** 3, (4), 15.

7. Krieser, R. J.; Moore, F. E.; Dresnek, D.; Pellock, B. J.; Patel, R.; Huang, A.; Brachmann, C.; White, K., The Drosophila homolog of the putative phosphatidylserine receptor functions to inhibit apoptosis. *Development* **2007,** 134, (13), 2407-2414.

8. Hahn, P.; Bose, J.; Edler, S.; Lengeling, A., Genomic structure and expression of Jmjd6 and evolutionary analysis in the context of related JmjC domain containing proteins. *BMC Genomics* **2008,** 9, (1), 293.

9. Fadok, V. A.; de Cathelineau, A.; Daleke, D. L.; Henson, P. M.; Bratton, D. L., Loss of Phospholipid Asymmetry and Surface Exposure of Phosphatidylserine Is Required for Phagocytosis of Apoptotic Cells by Macrophages and Fibroblasts. *Journal of Biological Chemistry* **2001,** 276, (2), 1071-1077.

10. Hoffmann, P. R.; deCathelineau, A. M.; Ogden, C. A.; Leverrier, Y.; Bratton, D. L.; Daleke, D. L.; Ridley, A. J.; Fadok, V. A.; Henson, P. M., Phosphatidylserine (PS) induces PS receptor mediated macropinocytosis and promotes clearance of apoptotic cells. *The Journal of Cell Biology* **2001,** 155, (4), 649-660.

11. Henson, P. M.; Bratton, D. L.; Fadok, V. A., The phosphatidylserine receptor: a crucial molecular switch? *Nature Reviews Molecular Cell Biology* **2001,** 2, (8), 627-633.

12. Somersan, S.; Bhardwaj, N., Tethering and tickling. *The Journal of Cell Biology* **2001,** 155, (4), 501-504.

13. Arur, S.; Uche, U. E.; Rezaul, K.; Fong, M.; Scranton, V.; Cowan, A. E.; Mohler, W.; Han, D. K., Annexin I Is an Endogenous Ligand that Mediates Apoptotic Cell Engulfment. *Developmental Cell* **2003,** 4, (4), 587-598.

14. Wang, X.; Wu, Y.-C.; Fadok, V. A.; Lee, M.-C.; Gengyo-Ando, K.; Cheng, L.-C.; Ledwich, D.; Hsu, P.-K.; Chen, J.-Y.; Chou, B.-K.; Henson, P.; Mitani, S.; Xue, D., Cell Corpse Engulfment Mediated by *C. elegans* Phosphatidylserine Receptor Through CED-5 and CED-12. *Science* **2003,** 302, (5650), 1563-1566.

15. Mitchell, J. E.; Cvetanovic, M.; Tibrewal, N.; Patel, V.; Colamonici, O. R.; Li, M. O.; Flavell, R. A.; Levine, J. S.; Birge, R. B.; Ucker, D. S., The Presumptive Phosphatidylserine Receptor Is Dispensable for Innate Anti-inflammatory Recognition and Clearance of Apoptotic Cells. *Journal of Biological Chemistry* **2006,** 281, (9), 5718-5725.

16. Cui, P.; Qin, B.; Liu, N.; Pan, G.; Pei, D., Nuclear localization of the phosphatidylserine receptor protein via multiple nuclear localization signals. *Experimental Cell Research* **2004,** 293, (1), 154-163.

17. Cikala, M.; Alexandrova, O.; David, C.; Proschel, M.; Stiening, B.; Cramer, P.; Böttger, A., The phosphatidylserine receptor from Hydra is a nuclear protein with potential Fe(II) dependent oxygenase activity. *BMC Cell Biology* **2004,** 5, (1), 26-36.

18. Dirk, G. r., Nuclear protein import. *Current Opinion in Cell Biology* **1997,** 9, (3), 412-419.

19. Blobel, G., Unidirectional and bidirectional protein traffic across membranes. In *Cold Spring Harbor Symposia on Quantitative Biology*, 1995; Vol. 60, pp 1-10.

20. Hewitson, K. S.; McNeill, L. A.; Riordan, M. V.; Tian, Y.-M.; Bullock, A. N.; Welford, R. W.; Elkins, J. M.; Oldham, N. J.; Bhattacharya, S.; Gleadle, J. M.; Ratcliffe, P. J.; Pugh, C. W.; Schofield, C. J., Hypoxia-inducible factor (HIF) asparagine hydroxylase is identical to factor inhibiting HIF (FIH) and is related to the cupin structural family. *Journal of Biological Chemistry* **2002,** 277, (29), 26351-26355.

21. Lando, D.; Peet, D. J.; Gorman, J. J.; Whelan, D. A.; Whitelaw, M. L.; Bruick, R. K., FIH-1 is an asparaginyl hydroxylase enzyme that regulates the transcriptional activity of hypoxia-inducible factor. *Genes & Development* **2002,** 16, (12), 1466-1471.

22. Flashman, E.; Schofield, C. J., The most versatile of all reactive intermediates? *Nature Chemical Biology* **2007,** 3, (2), 86-87.

23. Loenarz, C.; Schofield, C. J., Expanding chemical biology of 2-oxoglutarate oxygenases. *Nature Chemical Biology* **2008,** 4, (3), 152-156.

24. Hausinger, R. P., Fe(II)/α-Ketoglutarate-Dependent Hydroxylases and Related Enzymes. *Critical Reviews in Biochemistry and Molecular Biology* **2004,** 39, (1), 21-68.

25. Clifton, I. J.; McDonough, M. A.; Ehrismann, D.; Kershaw, N. J.; Granatino, N.; Schofield, C. J., Structural studies on 2-oxoglutarate oxygenases and related double-stranded β-helix fold proteins. *Journal of Inorganic Biochemistry* **2006,** 100, (4), 644-669.

26. McDonough, M. A.; Loenarz, C.; Chowdhury, R.; Clifton, I. J.; Schofield, C. J., Structural studies on human 2-oxoglutarate dependent oxygenases. *Current Opinion in Structural Biology* **2010,** 20, (6), 659-672.

27. Price, J. C.; Barr, E. W.; Tirupati, B.; Bollinger, J. M.; Krebs, C., The First Direct Characterization of a High-Valent Iron Intermediate in the Reaction of an α-Ketoglutarate-Dependent Dioxygenase: A High-Spin Fe(IV) Complex in Taurine/α-Ketoglutarate Dioxygenase (TauD) from *Escherichia coli Biochemistry* **2003,** 42, (24), 7497-7508.

28. Price, J. C.; Barr, E. W.; Glass, T. E.; Krebs, C.; Bollinger, J. M., Evidence for Hydrogen Abstraction from C1 of Taurine by the High-Spin Fe(IV) Intermediate Detected during Oxygen Activation by Taurine:α-Ketoglutarate Dioxygenase (TauD). *Journal of the American Chemical Society* **2003,** 125, (43), 13008-13009.

29. Leung, I. K. H.; Krojer, T. J.; Kochan, G. T.; Henry, L.; von Delft, F.; Claridge, T. D. W.; Oppermann, U.; McDonough, M. A.; Schofield, C. J., Structural and Mechanistic Studies on γ-Butyrobetaine Hydroxylase. *Chemistry & Biology* **2010,** 17, (12), 1316-1324.

30. Loenarz, C.; Coleman, M. L.; Boleininger, A.; Schierwater, B.; Holland, P. W. H.; Ratcliffe, P. J.; Schofield, C. J., The hypoxia-inducible transcription factor pathway regulates oxygen sensing in the simplest animal, Trichoplax adhaerens. *EMBO Reports* **2010,** 12, (1), 63-70.

31. Loenarz, C.; Schofield, C. J., Physiological and biochemical aspects of hydroxylations and demethylations catalyzed by human 2-oxoglutarate oxygenases. *Trends in Biochemical Sciences* **2011,** 36, (1), 7-18.

32. Baldwin, J. E.; Bradley, M., Isopenicillin N synthase: mechanistic studies. *Chemical Reviews (Washington, DC, United States)* **1990,** 90, (7), 1079-1088.

33. Schofield, C. J.; Baldwin, J. E., Chapter 1. The Biosynthesis of β-Lactams. In *The Chemistry of β-Lactams*

Page, M. I., Ed. Blackie Academic & Professional: London, 1992.

34. Baggaley, K. H.; Brown, A. G.; Schofield, C. J., Chemistry and biosynthesis of clavulanic acid and other clavams. *Natural Product Reports* **1997,** 14, (4), 309-333.

35. Lange, T.; Hedden, P.; Graebe, J. E., Expression cloning of a gibberellin 20-oxidase, a multifunctional enzyme involved in gibberellin biosynthesis. *Proceedings of the National Academy of Sciences* **1994,** 91, (18), 8552-8556.

36. Turnbull, J. J.; Nakajima, J.; Welford, R. W. D.; Yamazaki, M.; Saito, K.; Schofield, C. J., Mechanistic Studies on Three 2-Oxoglutarate-dependent Oxygenases of Flavonoid Biosynthesis. *Journal of Biological Chemistry* **2004,** 279, (2), 1206-1216.

37. Roach, P. L.; Clifton, I. J.; Hensgens, C. M. H.; Shibata, N.; Schofield, C. J.; Hajdu, J.; Baldwin, J. E., Structure of isopenicillinN synthase complexed with substrate and the mechanism of penicillin formation. *Nature* **1997,** 387, (6635), 827-830.

38. John, P., Ethylene biosynthesis: The role of 1-aminocyclopropane-1-carboxylate (ACC) oxidase, and its possible evolutionary origin. *Physiologia Plantarum* **1997,** 100, (3), 583-592.

39. David J.S, H., Building Collagen Molecules, Fibrils, and Suprafibrillar Structures. *Journal of Structural Biology* **2002,** 137, (1-2), 2-10.

40. Krane, S., The importance of proline residues in the structure, stability and susceptibility to proteolytic degradation of collagens. *Amino Acids* **2008,** 35, (4), 703-710.

41. Myllyharju, J., Prolyl 4-hydroxylases, the key enzymes of collagen biosynthesis. *Matrix Biology* **2003,** 22, (1), 15-24.

42. Schumacher, M. A.; Mizuno, K.; Bächinger, H. P., The Crystal Structure of a Collagen-like Polypeptide with 3(S)-Hydroxyproline Residues in the Xaa Position Forms a Standard 7/2 Collagen Triple Helix. *Journal of Biological Chemistry* **2006,** 281, (37), 27566-27574.

43. Heikkinen, J.; Risteli, M.; Wang, C.; Latvala, J.; Rossi, M.; Valtavaara, M.; Myllylä, R., Lysyl Hydroxylase 3 Is a Multifunctional Protein Possessing Collagen Glucosyltransferase Activity. *Journal of Biological Chemistry* **2000,** 275, (46), 36158-36163.

44. Schofield, C. J.; Ratcliffe, P. J., Oxygen sensing by HIF hydroxylases. *Nature Reviews Molecular Cell Biology* **2004,** 5, (5), 343-354.

45. Hon, W.-C.; Wilson, M. I.; Harlos, K.; Claridge, T. D. W.; Schofield, C. J.; Pugh, C. W.; Maxwell, P. H.; Ratcliffe, P. J.; Stuart, D. I.; Jones, E. Y., Structural basis for the recognition of hydroxyproline in HIF-1α by pVHL. *Nature* **2002,** 417, (6892), 975-978.

46. Masson, N.; Willam, C.; Maxwell, P. H.; Pugh, C. W.; Ratcliffe, P. J., Independent function of two destruction domains in hypoxia-inducible factor-α chains activated by prolyl hydroxylation. *EMBO Journal* **2001,** 20, (18), 5197-5206.

47. Huang, L. E.; Gu, J.; Schau, M.; Bunn, H. F., Regulation of hypoxia-inducible factor 1α is mediated by an O_2-dependent degradation domain via the ubiquitin-proteasome pathway. *Proceedings of the National Academy of Sciences* **1998,** 95, (14), 7987-7992.

48. Jaakkola, P.; Mole, D. R.; Tian, Y.-M.; Wilson, M. I.; Gielbert, J.; Gaskell, S. J.; Kriegsheim, A. v.; Hebestreit, H. F.; Mukherji, M.; Schofield, C. J.; Maxwell, P. H.; Pugh, â. C. W.; Ratcliffe, â. P. J., Targeting of HIF-1α to the von Hippel-Lindau Ubiquitylation Complex by O_2-Regulated Prolyl Hydroxylation. *Science* **2001,** 292, (5516), 468-472.

49. McNeill, L. A.; Hewitson, K. S.; Claridge, T. D.; Seibel, J. F.; Horsfall, L. E.; Schofield, C. J., Hypoxia-inducible factor asparaginyl hydroxylase (FIH-1) catalyses hydroxylation at the β-carbon of asparagine-803. *Biochemical Journal* **2002,** 367, (3), 571-575.

50. Mahon, P. C.; Hirota, K.; Semenza, G. L., FIH-1: a novel protein that interacts with HIF-1α and VHL to mediate repression of HIF-1 transcriptional activity. *Genes & Development* **2001,** 15, (20), 2675-2686.

51. Vaz, F. d. r. M.; Wanders, R. J. A., Carnitine biosynthesis in mammals. *Biochemical Journal* **2002,** 361, (3), 417-429.

52. Hulse, J. D.; Ellis, S. R.; Henderson, L. M., Carnitine biosynthesis. beta-Hydroxylation of trimethyllysine by an alpha-ketoglutarate-dependent mitochondrial dioxygenase. *Journal of Biological Chemistry* **1978,** 253, (5), 1654-1659.

53. Göran, L., Effect of metal ions on the hydroxylation of Î³-butyrobetaine to carnitine in rat liver homogenates. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1967,** 141, (3), 492-498.

54. Mukherji, M.; Chien, W.; Kershaw, N. J.; Clifton, I. J.; Schofield, C. J.; Wierzbicki, A. S.; Lloyd, M. D., Structure-function analysis of phytanoyl-CoA 2-hydroxylase mutations causing Refsum's disease. *Human Molecular Genetics* **2001,** 10, (18), 1971-1982.

55. McDonough, M. A.; Kavanagh, K. L.; Butler, D.; Searls, T.; Oppermann, U.; Schofield, C. J., Structure of Human Phytanoyl-CoA 2-Hydroxylase Identifies Molecular Mechanisms of Refsum Disease. *Journal of Biological Chemistry* **2005,** 280, (49), 41101-41110.

56. Klose, R. J.; Zhang, Y., Regulation of histone methylation by demethylimination and demethylation. *Nature Reviews Molecular Cell Biology* **2007,** 8, (4), 307-318.

57. Zhang, Y.; Reinberg, D., Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes & Development* **2001,** 15, (18), 2343-2360.

58. Lachner, M.; Jenuwein, T., The many faces of histone lysine methylation. *Current Opinion in Cell Biology* **2002,** 14, (3), 286-298.

59. Martin, C.; Zhang, Y., The diverse functions of histone lysine methylation. *Nature Reviews Molecular Cell Biology* **2005,** 6, (11), 838-849.

60. Klose, R. J.; Kallin, E. M.; Zhang, Y., JmjC-domain-containing proteins and histone demethylation. *Nature Reviews Genetics* **2006,** 7, (9), 715-727.

61. Ng, S. S.; Kavanagh, K. L.; McDonough, M. A.; Butler, D.; Pilka, E. S.; Lienard, B. M. R.; Bray, J. E.; Savitsky, P.; Gileadi, O.; von Delft, F.; Rose, N. R.; Offer, J.; Scheinost, J. C.; Borowski, T.; Sundstrom, M.; Schofield, C. J.; Oppermann, U., Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity. *Nature* **2007,** 448, (7149), 87-91.

62. Hopkinson, R. J.; Hamed, R. B.; Rose, N. R.; Claridge, T. D. W.; Schofield, C. J., Monitoring the Activity of 2-Oxoglutarate Dependent Histone Demethylases by NMR Spectroscopy: Direct Observation of Formaldehyde. *ChemBioChem* **2010,** 11, (4), 506-510.

63. Horton, J. R.; Upadhyay, A. K.; Qi, H. H.; Zhang, X.; Shi, Y.; Cheng, X., Enzymatic and structural insights for substrate specificity of a family of jumonji histone lysine demethylases. *Nature Structural & Molecular Biology* **2010,** 17, (1), 38-43.

64. Yue, W. W.; Hozjan, V.; Ge, W.; Loenarz, C.; Cooper, C. D.; Schofield, C. J.; Kavanagh, K. L.; Oppermann, U.; McDonough, M. A., Crystal structure of the PHF8 Jumonji domain, an N(ε)-methyl lysine demethylase. *FEBS Letters* **2010,** 584, (4), 825-830.

65. Yang, Y.; Hu, L.; Wang, P.; Hou, H.; Lin, Y.; Liu, Y.; Li, Z.; Gong, R.; Feng, X.; Zhou, L.; Zhang, W.; Dong, Y.; Yang, H.; Lin, H.; Wang, Y.; Chen, C. D.; Xu, Y., Structural insights into a dual-specificity histone demethylase ceKDM7A from *Caenorhabditis elegans*. *Cell Research* **2010,** 20, (8), 886-898.

66. Sedgwick, B.; Bates, P. A.; Paik, J.; Jacobs, S. C.; Lindahl, T., Repair of alkylated DNA: Recent advances. *DNA Repair* **2007,** 6, (4), 429-442.

67. Cetica, V.; Genitori, L.; Giunti, L.; Sanzo, M.; Bernini, G.; Massimino, M.; Sardi, I., Pediatric brain tumors: mutations of two dioxygenases (hABH2 and hABH3) that directly repair alkylation damage. *Journal of Neuro-Oncology* **2009,** 94, (2), 195-201.

68. Konishi, N.; Nakamura, M.; Ishida, E.; Shimada, K.; Mitsui, E.; Yoshikawa, R.; Yamamoto, H.; Tsujikawa, K., High Expression of a New Marker PCA-1 in Human Prostate Carcinoma. *Clinical Cancer Research* **2005,** 11, (14), 5090-5097.

69. Shimada, K.; Nakamura, M.; Anai, S.; De Velasco, M.; Tanaka, M.; Tsujikawa, K.; Ouji, Y.; Konishi, N., A Novel Human AlkB Homologue, ALKBH8, Contributes to Human Bladder Cancer Progression. *Cancer Research* **2009,** 69, (7), 3157-3164.

70. Songe-Møller, L.; van den Born, E.; Leihne, V.; Vågbø, C. B.; Kristoffersen, T.; Krokan, H. E.; Kirpekar, F.; Falnes, P. Ø.; Klungland, A., Mammalian ALKBH8 Possesses tRNA Methyltransferase Activity Required for the Biogenesis of Multiple Wobble Uridine Modifications Implicated in Translational Decoding. *Molecular and Cellular Biology* **2010,** 30, (7), 1814-1827.

71. Tahiliani, M.; Koh, K. P.; Shen, Y.; Pastor, W. A.; Bandukwala, H.; Brudno, Y.; Agarwal, S.; Iyer, L. M.; Liu, D. R.; Aravind, L.; Rao, A., Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science* **2009,** 324, (5929), 930-935.

72. Delhommeau, F.; Dupont, S.; Valle, V. D.; James, C.; Trannoy, S.; Massé, A.; Kosmider, O.; Le Couedic, J.-P.; Robert, F.; Alberdi, A.; Lécluse, Y.; Plo, I.; Dreyfus, F. J.; Marzac, C.; Casadevall, N.; Lacombe, C.; Romana, S. P.; Dessen, P.; Soulier, J.; Viguié, F.; Fontenay, M.; Vainchenker, W.; Bernard, O. A., Mutation in TET2 in Myeloid Cancers. *New England Journal of Medicine* **2009,** 360, (22), 2289-2301.

73. Gerken, T.; Girard, C. A.; Tung, Y.-C. L.; Webby, C. J.; Saudek, V.; Hewitson, K. S.; Yeo, G. S. H.; McDonough, M. A.; Cunliffe, S.; McNeill, L. A.; Galvanovskis, J.; Rorsman, P.; Robins, P.; Prieur, X.; Coll, A. P.; Ma, M.; Jovanovic, Z.; Farooqi, I. S.; Sedgwick, B.; Barroso, I. s.; Lindahl, T.; Ponting, C. P.; Ashcroft, F. M.; O'Rahilly, S.; Schofield, C. J., The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. *Science* **2007,** 318, (5855), 1469-1472.

74. Frayling, T. M.; Timpson, N. J.; Weedon, M. N.; Zeggini, E.; Freathy, R. M.; Lindgren, C. M.; Perry, J. R. B.; Elliott, K. S.; Lango, H.; Rayner, N. W.; Shields, B.; Harries, L. W.; Barrett, J. C.; Ellard, S.; Groves, C. J.; Knight, B.; Patch, A.-M.; Ness, A. R.; Ebrahim, S.; Lawlor, D. A.; Ring, S. M.; Ben-Shlomo, Y.; Jarvelin, M.-R.; Sovio, U.; Bennett, A. J.; Melzer, D.; Ferrucci, L.; Loos, R. J. F.; Barroso, I. s.; Wareham, N. J.; Karpe, F.; Owen, K. R.; Cardon, L. R.; Walker, M.; Hitman, G. A.; Palmer, C. N. A.; Doney, A. S. F.; Morris, A. D.; Smith, G. D.; Consortium, T. W. T. C. C.; Hattersley, A. T.; McCarthy, M. I., A Common Variant in the FTO Gene Is Associated with Body Mass Index and Predisposes to Childhood and Adult Obesity. *Science* **2007,** 316, (5826), 889-894.

75. Fischer, J.; Koch, L.; Emmerling, C.; Vierkotten, J.; Peters, T.; Brüning, J. C.; Rüther, U., Inactivation of the Fto gene protects from obesity. *Nature* **2009,** 458, (7240), 894-898.

76. Chang, B.; Chen, Y.; Zhao, Y.; Bruick, R. K., JMJD6 Is a Histone Arginine Demethylase. *Science* **2007,** 318, (5849), 444-447.

77. Webby, C. J.; Wolf, A.; Gromak, N.; Dreger, M.; Kramer, H.; Kessler, B.; Nielsen, M. L.; Schmitz, C.; Butler, D. S.; Yates, J. R., III; Delahunty, C. M.; Hahn, P.; Lengeling, A.; Mann, M.; Proudfoot, N. J.; Schofield, C. J.; Bottger, A., Jmjd6 Catalyses Lysyl-Hydroxylation of U2AF65, a Protein Associated with RNA Splicing. *Science* **2009,** 325, (5936), 90-93.

78. Zamore, P. D.; Patton, J. G.; Green, M. R., Cloning and domain structure of the mammalian splicing factor U2AF. *Nature* **1992,** 355, (6361), 609-614.

79. Nishii, Y.; Morishima, M.; Kakehi, Y.; Umehara, K.; Kioka, N.; Terano, Y.; Amachi, T.; Ueda, K., CROP/Luc7A, a novel serine/arginine-rich nuclear protein, isolated from cisplatin-resistant cell line. *FEBS Letters* **2000,** 465, (2-3), 153-156.

80. Karni, R.; de Stanchina, E.; Lowe, S. W.; Sinha, R.; Mu, D.; Krainer, A. R., The gene encoding the splicing factor SF2/ASF is a proto-oncogene. *Nature Structural & Molecular Biology* **2007,** 14, (3), 185-193.

81. Gromak, N.; Smith, C. W. J., A splicing silencer that regulates smooth muscle specific alternative splicing is active in multiple cell types. *Nucleic Acids Research* **2002,** 30, (16), 3548-3557.

82. Hahn, P.; Wegener, I.; Burrells, A.; Böse, J.; Wolf, A.; Erck, C.; Butler, D.; Schofield, C. J.; Böttger, A.; Lengeling, A., Analysis of Jmjd6 Cellular Localization and Testing for Its Involvement in Histone Demethylation. *PLoS One* **2010,** 5, (10), e13769.

83. Boeckel, J.-N.; Guarani, V.; Koyanagi, M.; Roexe, T.; Lengeling, A.; Schermuly, R. T.; Gellert, P.; Braun, T.; Zeiher, A.; Dimmeler, S., Jumonji domain-containing protein 6 (Jmjd6) is required for angiogenic sprouting and regulates splicing of VEGF-receptor 1. *Proceedings of the National Academy of Sciences of the United States of America* **2011,** 108, (8), 3276-3281.

84. Chen, Y. H.; Comeaux, L. M.; Eyles, S. J.; Knapp, M. J., Auto-hydroxylation of FIH-1: an Fe(II), alpha-ketoglutarate-dependent human hypoxia sensor. *Chemical Communications (Cambridge, United Kingdom)* **2008**, (39), 4768-4770.

85. Ryle, M. J.; Liu, A.; Muthukumaran, R. B.; Ho, R. Y.; Koehntop, K. D.; McCracken, J.; Que, L., Jr.; Hausinger, R. P., O_2- and alpha-ketoglutarate-dependent tyrosyl radical formation in TauD, an alpha-keto acid-dependent non-heme iron dioxygenase. *Biochemistry* **2003,** 42, (7), 1854-1862.

86. Liu, A.; Ho, R. Y. N.; Que, L.; Ryle, M. J.; Phinney, B. S.; Hausinger, R. P., Alternative Reactivity of an α-Ketoglutarate-Dependent Iron(II) Oxygenase: Enzyme Self-Hydroxylation. *Journal of the American Chemical Society* **2001,** 123, (21), 5126-5127.

87. Sundheim, O.; Vagbo, C. B.; Bjoras, M.; Sousa, M. M. L.; Talstad, V.; Aas, P. A.; Drablos, F.; Krokan, H. E.; Tainer, J. A.; Slupphaug, G., Human ABH3 structure and key residues for oxidative demethylation to reverse DNA/RNA damage. *EMBO Journal* **2006,** 25, (14), 3389-3397.

Chapter 2. JMJD6 as a lysyl hydroxylase

2.1 Introduction

JMJD6 has been reported by our group to be involved in the lysine hydroxylation of proteins involved in pre-mRNA splicing. Arginine demethylation activity of histone tails as catalysed by JMJD6 was also reported by Chang et al. in 2007.[1] The aim of the work described in this chapter is to investigate if JMJD6 has demethylation activity and to explore the possibly of novel substrates of JMJD6.

2.2 Background: Arginine demethylation of histone tails

Histone methylations play an important role in regulating transcription, genome integrity and epigenetic regulations.[2, 3] In earlier days, histone methylations were considered irreversible, but recent studies on various demethylase enzymes have revealed the dynamic nature of histone modifications and their involvement in chromatin regulation. In mammals, arginine methylation is found on residues 2, 8, 17, 26 of histone 3 (H3R2, H3R8, H3R17, H3R26) and residue 3 of histone H4 (H4R3).[3] Arginine methylations are catalysed by protein arginine methyl transferase (PRMT) enzymes.[4, 5] Both mono- and di-methylations of arginine residues are reported to be catalysed by these enzymes, though some methyl transferases catalyse specifically symmetrical or unsymmetrical di-methylation of arginine. Only two enzymes have been shown to alter the methylation state of histone arginines. One of them is peptidyl arginine deiminase, (PADI4). PADI4 catalyses the demethylamination of methyl arginine residues to produce citrulline and methylamine.[6] Because PADI4 converts methyl arginine to citrulline rather than unmodified arginine, it is not a true histone

arginine demethylase. Nevertheless, *in vivo* PADI4 appears to have a role in antagonising histone arginine methylation in certain instances.[7] Of the biologically identified three arginine methylation states, PADI4 specifically targets the deimination of the mono-methyl arginine modification state *in vitro*.

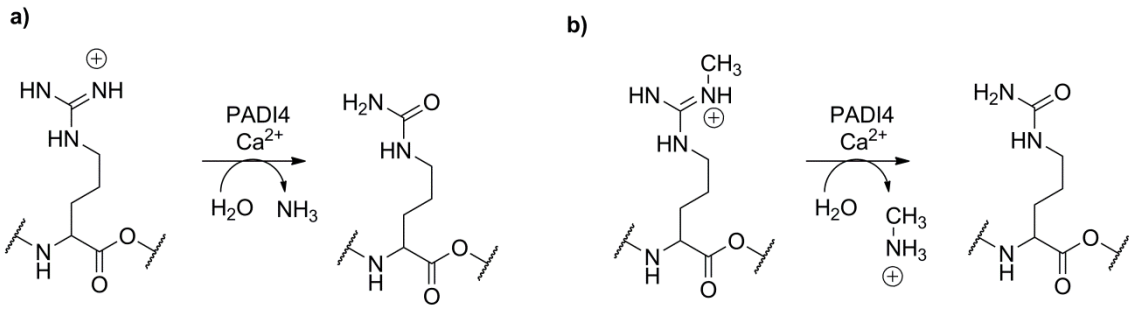


Figure 2.1 Aginine and methyl arginine demethylation is catalysed by PDI4 to give citrulline.

JMJD6 has extensive sequence and structural homology with FIH[8] as well as with the JmjC domain that is characteristic of 2OG dependent histone dmethylases.[9] Due to the predicted conservation of structural elements and key residues, JMJD6 might be expected to possess an analogous catalytic activity to the histone demethylases and thereby be involved in epigenetic regulation. In 2007, Chang et al. reported that JMJD6 possesses *in vitro* arginine demethylation activity.[1] A 30 residue H4 histone peptide with a symmetric dimethyl arginine residue at R3 (p2meR3H4) and histone H3 peptide at R2 position as p2meR2H3 were incubated with JMJD6. The crude product was immunoprecipitated with an antibody specific for the monomethylated product and then analysed by MALDI-MS. The results from the reported paper are shown in Figure 2.2. Chang et al. reported that in addition to expected demethylated peaks, additional peaks at +16 and +32 for likely oxidation were observed. These were shown to result from lysine hydroxylation by MS/MS analysis at the K5 and K8 positions of the R3H4

peptide. Webby et al. tried to replicate these results using various peptide fragments of H3 and H4 peptides. However, only hydroxylation was observed in the tested cases where reaction occurred.[10]

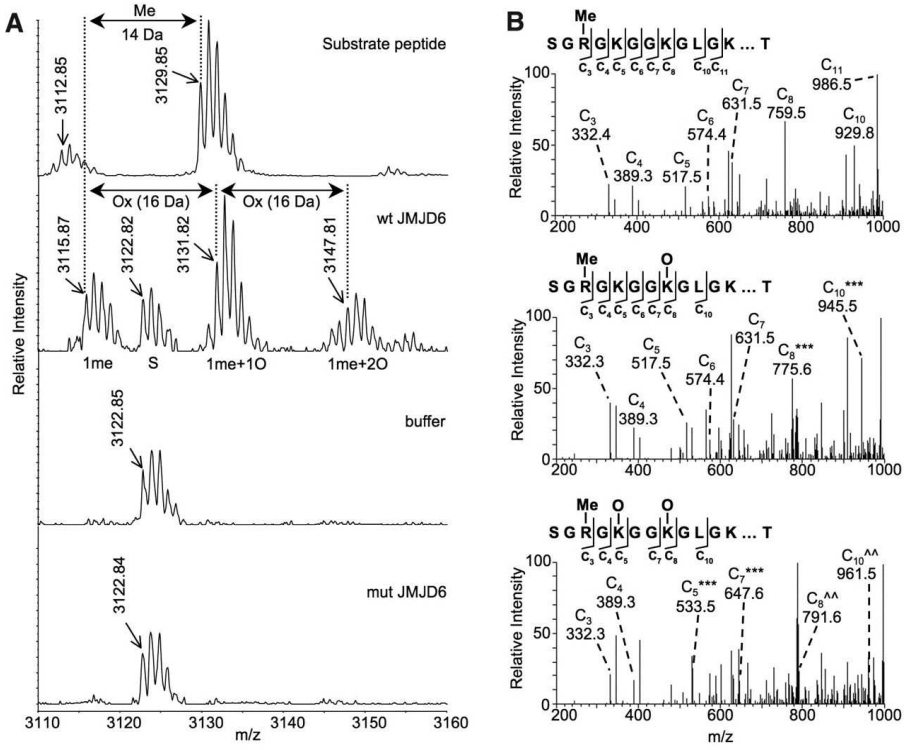


Figure 2.2 Analysis of demethylation products by MS as shown by Chang et al. Figure is taken from reference [1]. (A**) JMJD6 demethylation of H4R3 peptide with symmetric dimethylated arginine at R3. [(M + 1) = 3129.6] Peptide was incubated at various conditions. Products were immunoprecipitated with the α-H4R3me1 antibody before analysis. H4R3 peptide with monomethylated arginine at R3 synthesised with a stable isotope of leucine (containing six ^{13}C and one ^{15}N atoms) was added as internal standard (S) before immunoprecipitation. The peak with m/z 3112.85 in the top panel is a deaminated impurity that was removed during immunoprecipitation. (**B**) MS/MS spectrum of the methylated and oxidised products.**

During the course of this chapter, I have investigated the reported histone arginine demethylation catalytic activity of JMJD6 using amplified luminescent proximity homogeneous (ALPHA) screen assays and by MALDI-MS analyses on various histone peptides including the one which was used by Chang et al. in their work.

2.3 ALPHA Screen binding assay for investigating the association of JMJD6 with histone peptides

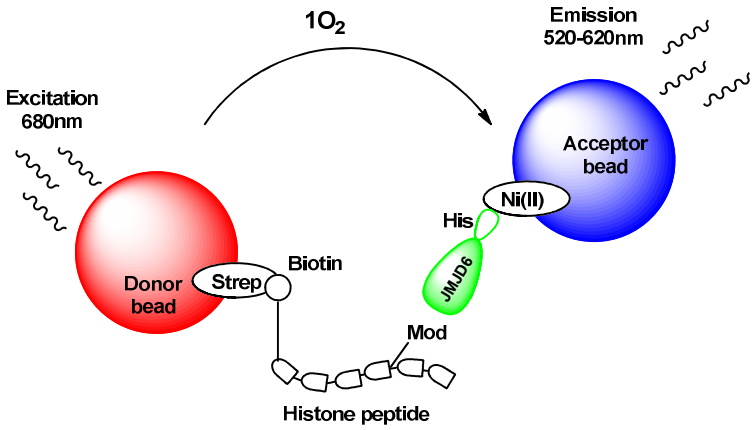


Figure 2.3 A typical ALPHA Screen setup of binding assays for His-tagged JMJD6.

ALPHA screen assays have been developed to investigate protein-protein or protein-peptide interactions of histone demethylases and many other systems.[11] In ALPHA Screen assays, laser excitation (680 nm) of a photosensitizer within the "donor bead" converts ambient oxygen to an excited O_2 singlet oxygen. If an "acceptor bead" containing a thioxene derivative is in proximity (~200 nm), energy transfer occurs, producing a luminescent signal at 520 to 620 nm.[12] In the absence of an acceptor bead, singlet oxygen returns to its ground state and no signal is detected. Both donor and acceptor bead surfaces can be conjugated to biomolecules, giving diversity in the binding partners to be investigated. To analyse the binding of protein and histone peptides, a streptavidin coated donor and Ni^{+2} chelate acceptor bead pairing was utilised. On binding of biotinylated histone peptide and His-tagged JMJD6, the donor and acceptor beads should be in proximity, allowing energy transfer and luminescence by excited beads. The binding assay setup is shown in Figure 2.3. The assays were carried out using a histone peptide library obtained from Alta Bioscience containing *N*-

terminally biotinylated peptides from H3, H4 and H2 histone sequences. The sequences of peptides used are given in Appendix A. Binding assays were carried out in 96 well plates in 20 μl of reaction volume by incubating His-tagged JMJD6 with biotinylated peptide in 50 mM Tris (pH 7.5), 0.01% Tween 20, and 0.1% BSA buffer. ALPHA screen streptavidin conjugated donor and nickel chelate conjugated acceptor beads were added to the wells at a final concentration of 10 μg ml^{-1} and incubated for 1 hour in the dark at 22°C. The resultant fluorescence was recorded using an Envision (PerkinElmer) plate reader.

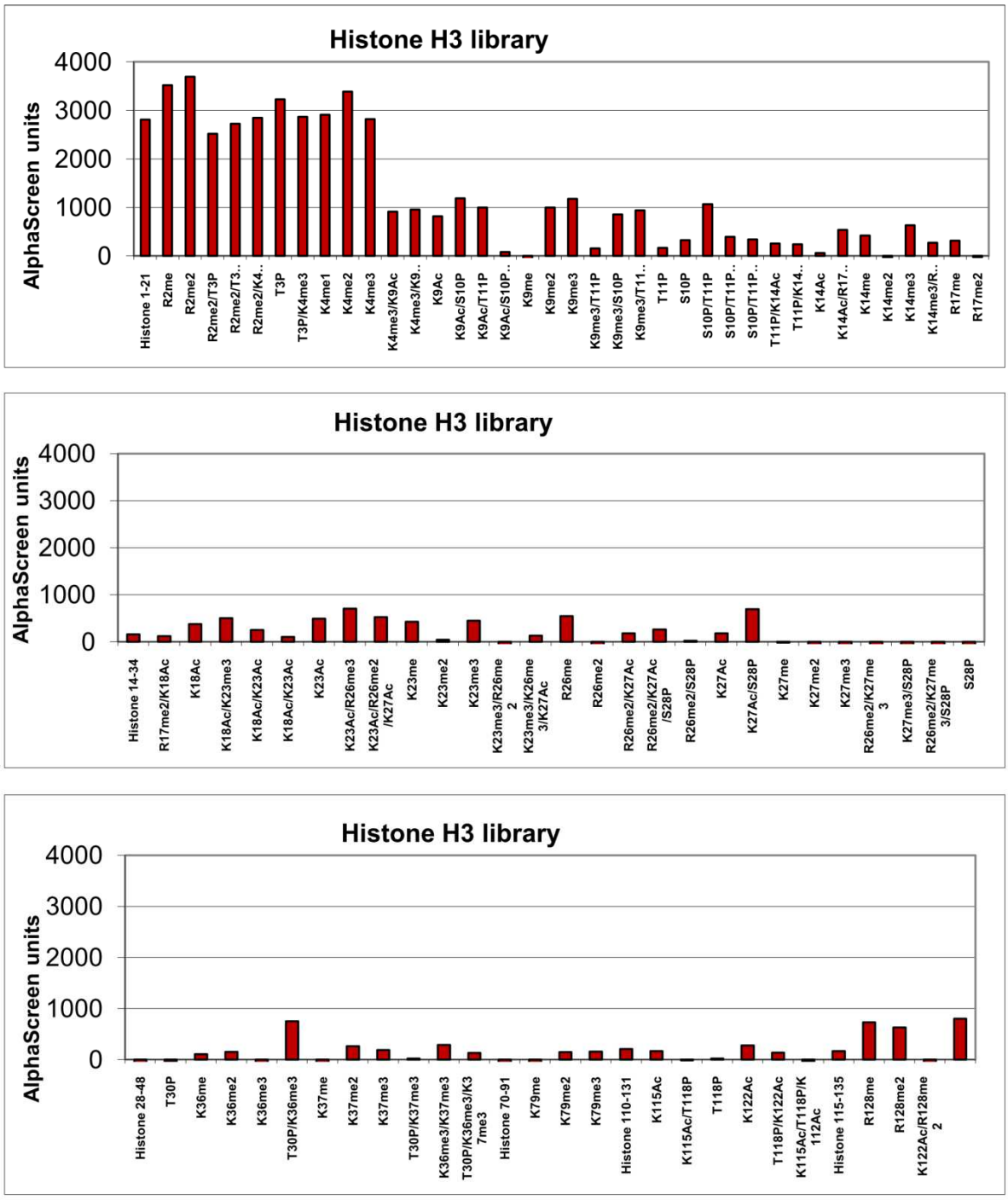


Figure 2.4 Results of ALPHA screen assays when JMJD6 was incubated with a library of biotinylated peptides from Histone H3 sequence.

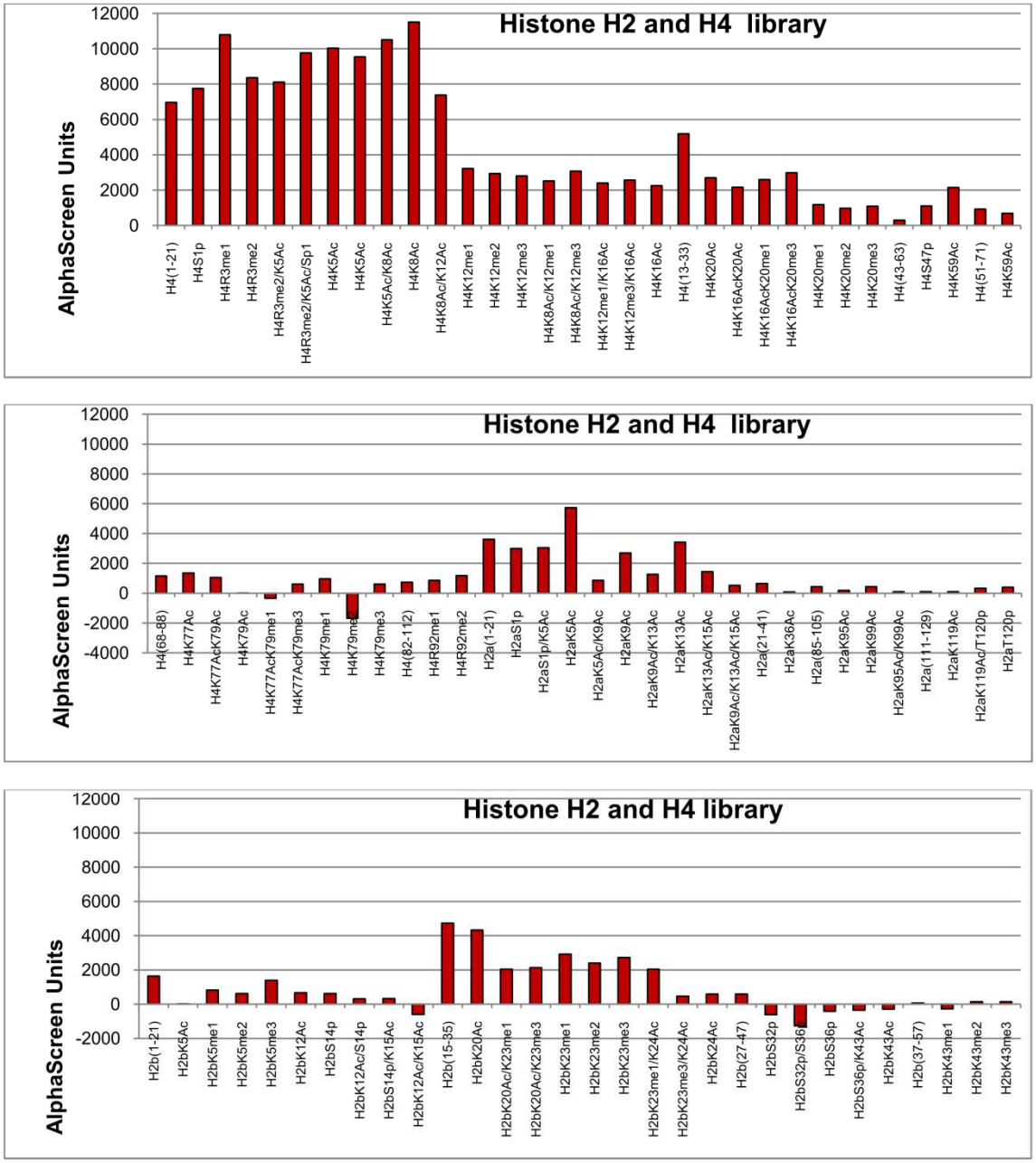
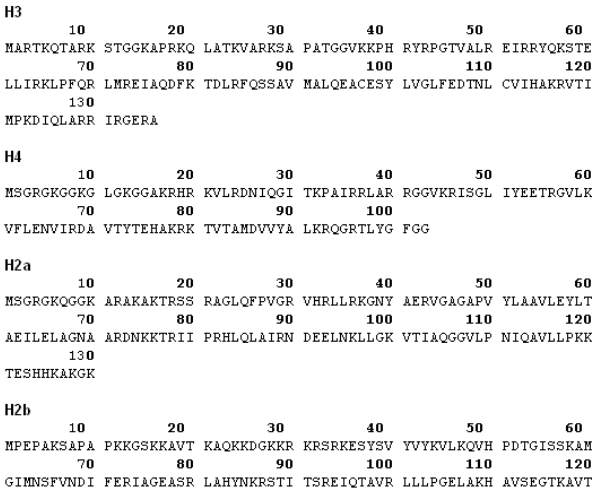


Figure 2.5 Results of ALPHA screen assays when JMJD6 was incubated with the library of biotinylated peptides from Histone H2 and H4 sequence.

The results at the ALPHA screen assays indicate that there may be stronger binding of JMJD6 with the $H3_{1-21}$ peptide and its R2 and K9 methylated derivatives (Figure 2.4). When the peptide library from H4 and H2 histone sequences was analysed for JMJD6

binding using the ALPHA screen assay, it also indicated stronger binding with the peptides at or closer to the *N*-terminus, i.e. $H4_{1-21}$, $H2a_{1-21}$ and $H2b_{15-35}$ peptides and their methylated/acetylated derivatives (Figure 2.5). These results correlate with the predicted positive charge on these residues due to the presence of arginine and lysine residues in the sequences as shown in Figure 2.6. The $H3_{1-21}$, $H4_{1-21}$, $H2a_{1-21}$ and $H3_{15-35}$ sequences are very basic. The presence of arginine and serines in the vicinity of lysines in the peptide may make them more prone to bind with JMJD6 because they may mimic the RS domain peptides which have been previously found to be substrates for JMJD6 by Webby et al.[10] Thus, this data indicates that not only basic charge but also presence of serine residues may promote binding of JMJD6 with peptides.



Peptide	No. of basic residues
H3 1-21	7
H3 14-34	6
H3 28-48	5
H3 70-91	4
H3 110-131	5
H3 115-135	7
H4 1-21	8
H4 13-33	7
H4 43-63	4
H4 51-71	3
H4 68-88	4

Peptide	No. of basic residues
H2a 1-21	8
H2a 21-41	5
H2a 85-105	2
H2a 111-129	3

Peptide	No. of basic residues
H2b 1-1	6
H2b 15-35	12
H2b 27-47	9
H2b 37-57	2

Figure 2.6 Sequences of the various histone peptides and the number of basic residues in peptides from the histone library that was tested for JMJD6 binding using the ALPHA screen assay.

2.4 Histone peptides showing no arginine demethylation but lysine hydroxylation

Various histone peptides including the one reported as a substrate by Chang et al. (H4R3 30 mer) were tested by Matrix assisted laser desorption ionisation (MALDI)-MS based assays to investigate potential arginine demethylation activity of JMJD6. These assays involved incubation of the peptide substrates with JMJD6 in the presence of other cofactors/cosubstrates (Fe(II) and 2OG) which are essential for 2OG oxygenase catalytic activity followed by MS analyses using MALDI-MS. For these assays peptide, 2OG and ascorbate were premixed separately in 50 mM Tris at pH 7.5. JMJD6 (20 μM) was incubated with Fe(II) (400 μM) in Tris buffer. To initiate the reaction, the peptide and enzyme/cosubstrate mixtures were mixed together. It was found to be important to follow this order of mixing so as to avoid loss of JMJD6 activity due to a precipitation or self-hydroxylation as described later in Chapter 5. The reaction mixture was then incubated at 37ºC for one hour and then analysed by MALDI-MS.

MALDI is often used for the analysis of biomolecules and large organic molecules which are more prone to fragmentation damage by conventional strong ionisation methods.[13, 14] MALDI is a soft ionisation technique in which an analyte is cocrystallised with a matrix on a MALDI plate prior to irradiation by a laser beam.[14] During ionisation, photo excited matrix molecules are stabilised though protein transfer to the analyte and due to the high laser energy, the matrix vaporises and carries analyte molecules with it. Molecules sensitive to the laser are thus protected from direct UV laser excitation. Once in the gas phase, the desorbed charged molecules are then directed electrostatically from the MALDI ionisation source into the mass analyser such as time-of-flight (TOF) mass analyser which separates ions according to their mass-to-

charge ratio (m/z). Thus, the matrix plays a very important role in MALDI analysis. Commonly used MALDI matrices are 3,5-dimethoxy-4-hydroxycinnamic acid (SA), 2,5-dihydroxybenzoic acid (DHB), and α-cyano-4-hydroxycinnamic acid (CHCA).[15] The CHCA matrix was extensively used during the course of this thesis as analytes of interest are peptides.

One of the advantages of MALDI-MS assays with peptides is that it requires a small amount of peptide (and thus, enzyme) and the analysis method is relatively simple and quick. A simple MALDI machine is also a relatively inexpensive. MALDI-MS technique also has the advantage of producing singly charged ions of peptides and proteins, minimising the spectral complexity so that it is easier to observe any modification of the peptide.

As previously observed by Webby et al., peptides from H4R3 and H3R2 with arginine substituted by symmetric and unsymmetric methyl arginine did not show any demethylation but instead showed hydroxylation.[10] It was shown by MS/MS analysis that hydroxylation in the the case of H3R2 peptide occurs at the K9 position. When this lysine was replaced by mono, di or tri-methyl lysine (Peptide 10, 9 and 8 from Table 2.1) and then incubated by JMJD6, it did not undergo any demethylation or hydroxylation. Peptides 1, 2, 6 to 11 from Table 2.1 were previously reported by Webby et al.[10] In addition the library of histone peptides with various modifications (symmetric and unsymmetric methylation at arginine and acetylated lysines) was tested for lysine hydroxylation by MALDI-MS assays. *This library of peptides was synthesised by Dr. Wei Ge*. The results are shown in Table 2.1, Figure 2.7 and Figure 2.8. In none of the histone peptides tested, demethylation activity was observed. $H4_{1\text{-}16}$ with symmetric and

unsymmetric dimethylated arginine at R3 (peptide 1, 2), $H4_{1-15}$ with (peptide 4) and without (peptide 3) symmetric dimethylated arginine at R3 and $H4_{5-19}$ peptide (Peptide 5) showed hydroxylation (Figure 2.7). $H3_{1-15}$ peptide with dimethylated arginine (symmetric) at R2 (peptide 6) had shown double hydroxylation. Hydroxylation was seen in the case of $H3_{1-15}$ peptide (Peptide 11). A $H3_{1-15}$ peptide with various lysine methylations and lysine acetylation (Peptides 12-16) was also assayed for JMJD6 catalysed lysine demethylase activity, in general no such activity was seen. Peptide 11 (unmodified) and peptide 12 (N^{ε}-trimethylated lysine at Lys 4) showed hydroxylation in this set of peptides. Peptides 18, 19 based on $H2a_{1-15}$ sequence and peptides 20, 21 based on $H2b_{7-17}$ and $H2b_{13-26}$ sequences showed hydroxylation (Figure 2.8) consistent with the binding of JMJD6 observed for the peptides in this region by ALPHA screen assays. These results are interesting as in no case was demethylation observed. However, hydroxylation as catalysed by JMJD6 was clearly observed (as indicated by a +16 Da mass shift) in the case of some peptides.

Table 2.1 Various histone peptides as analysed by MALDI-MS assays to check for arginine demethylation activity by JMJD6.

Peptide No.	Peptide Sequence	Mol. Wt.	Sequence from	MALDI Results
1	SGR(me2sym)GKGGKGLGKGGAK	1444.6	$H4_{1-16}$	Hydroxylation/ no demethylation
2	SGR(me2asym)GKGGKGLGKGGAK	1444.6	$H4_{1-16}$	Hydroxylation/ no demethylation
3	SGRGKGGKGLGKGGA	1286.4	$H4_{1-15}$	Hydroxylation
4	SGR(me2sym)GKGGKGLGKGGA	1316.4	$H4_{1-15}$	Hydroxylation
5	KGGKGLGKGGAKRHR	1506.7	$H4_{5-19}$	Hydroxylation

6	AR(me2sym)TKQTARKSTGGKAPPK	1913.6	$H3_{1-18}$	Double hydroxylation/ No demethylation
7	AR(me2asym)TKQTARKSTGGKAPPK	1913.6	$H3_{1-18}$	Double hydroxylation/ No demethylation
8	ARTKQTARK(me3)STGGKA	1605.7	$H3_{1-15}$	No
9	ARTKQTARK(me2)STGGKA	1590.7	$H3_{1-15}$	No
10	ARTKQTARK(me1)STGGKA	1575.7	$H3_{1-15}$	No
11	ARTKQTARKSTGGKA	1560.7	$H3_{1-15}$	Hydroxylation
12	ARTK(me3)QTARKSTGGKA	1605.7	$H3_{1-15}$	Hydroxylation
13	ARTKQTARK(Ac)STGGKA	1549.7	$H3_{1-15}$	No
14	ARTKQTARKSTGGK(Ac)A	1603.7	$H3_{1-15}$	No
15	ARTKQTARK(acetyl)STGGK(acetyl)A	1646.7	$H3_{1-15}$	No
16	ARTK(me3)QTARK(acetyl)STGGK(acetyl)A	1691.8	$H3_{1-15}$	No
17	QLATKAARKSAPATG	1470.6	$H3_{19-33}$	No
18	SGRGKQGGKARAKTR	1557.7	$H2a_{1-15}$	Hydroxylation
19	SGR(me2asym)GKQGGKARAKTR	1587.7	$H2a_{1-15}$	Hydroxylation
20	APAPKKGSKKAVTKA	1481.8	$H2b_{7-17}$	Hydroxylation
21	GSKKAVTKAQKKDSK	1603.8	$H2b_{13-26}$	Hydroxylation

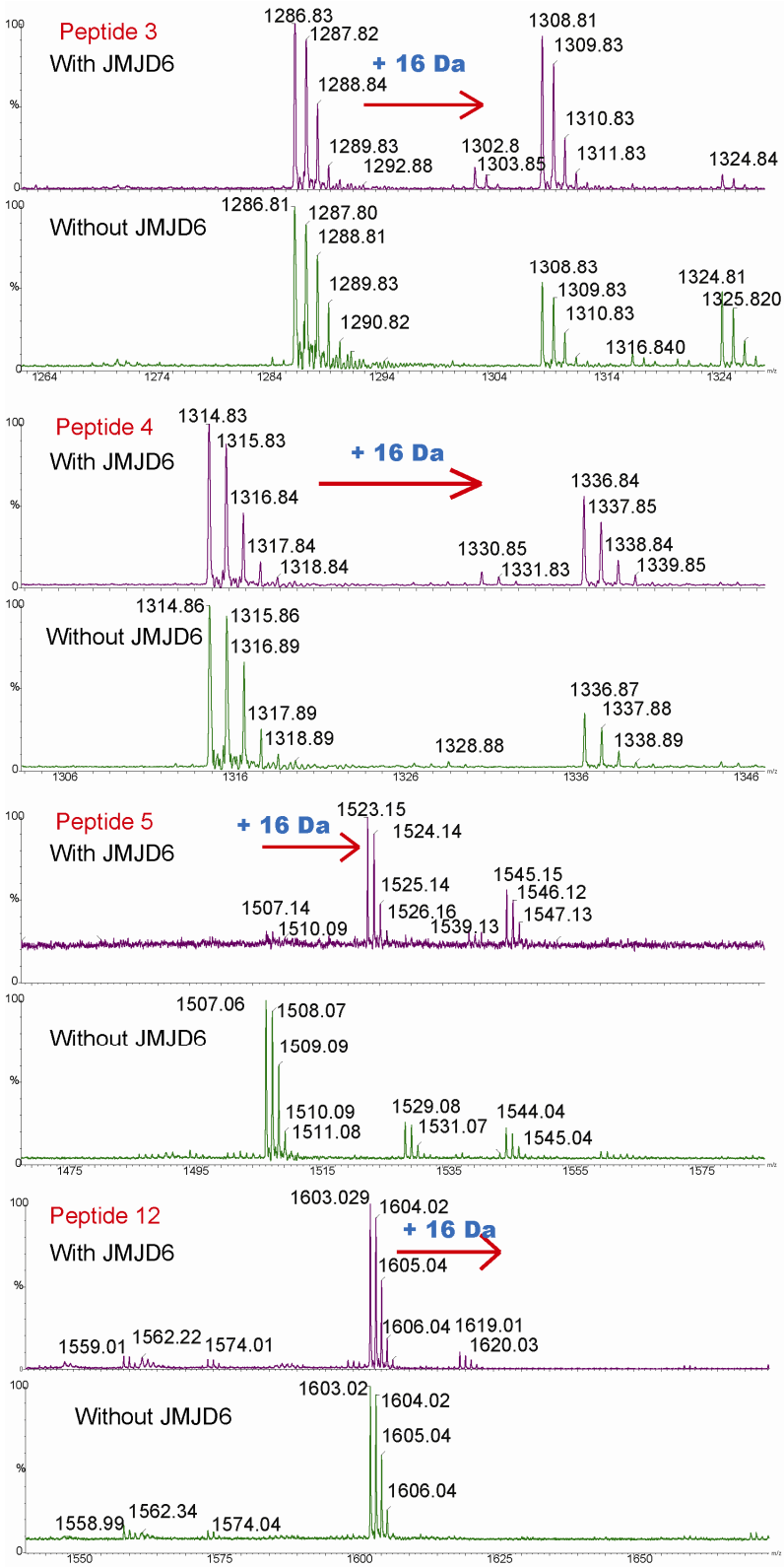


Figure 2.7 MALDI spectra of peptides from H4 and H3 histone sequences when incubated with JMJD6 and cofactors. The MALDI spectra from control samples in which peptides were incubated with cofactors and buffer without addition of JMJD6 are shown.

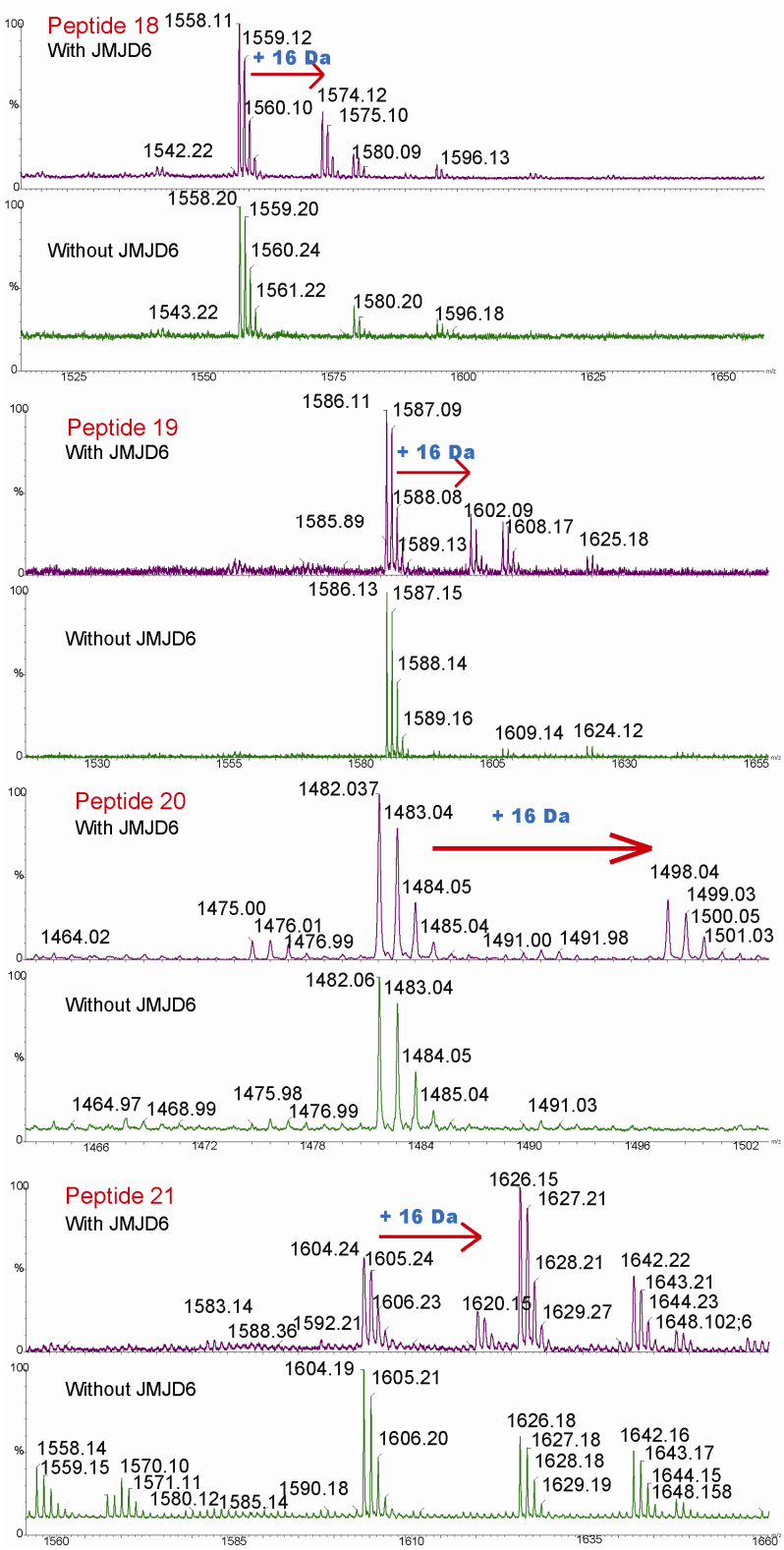


Figure 2.8 MALDI spectra of peptides from H2 histone sequences when incubated with JMJD6 and cofactors. The MALDI spectra from control samples in which peptides were incubated with cofactors and buffer without addition of JMJD6 are shown.

30 residue peptides corresponding to the $H4_{1-30}$ sequence, with and without a symmetric dimethylated arginine at R3 position were also tested by MALDI-MS assays with JMJD6. The H4R3 30 mer sequence with a symmetric dimethylated arginine was the same as that reported to have demethylation activity on incubation with JMJD6 by Chang et al.[1] Both the peptide sequences were purchased from Peptide Synthetic and incubated with JMJD6, Fe(II) and 2OG and then analysed by MALDI-MS. It can be seen from the Figure 2.9 that there is no observed arginine demethylation; instead peaks at 3142.53 and 3158.38 corresponding to lysine hydroxylations are observed.

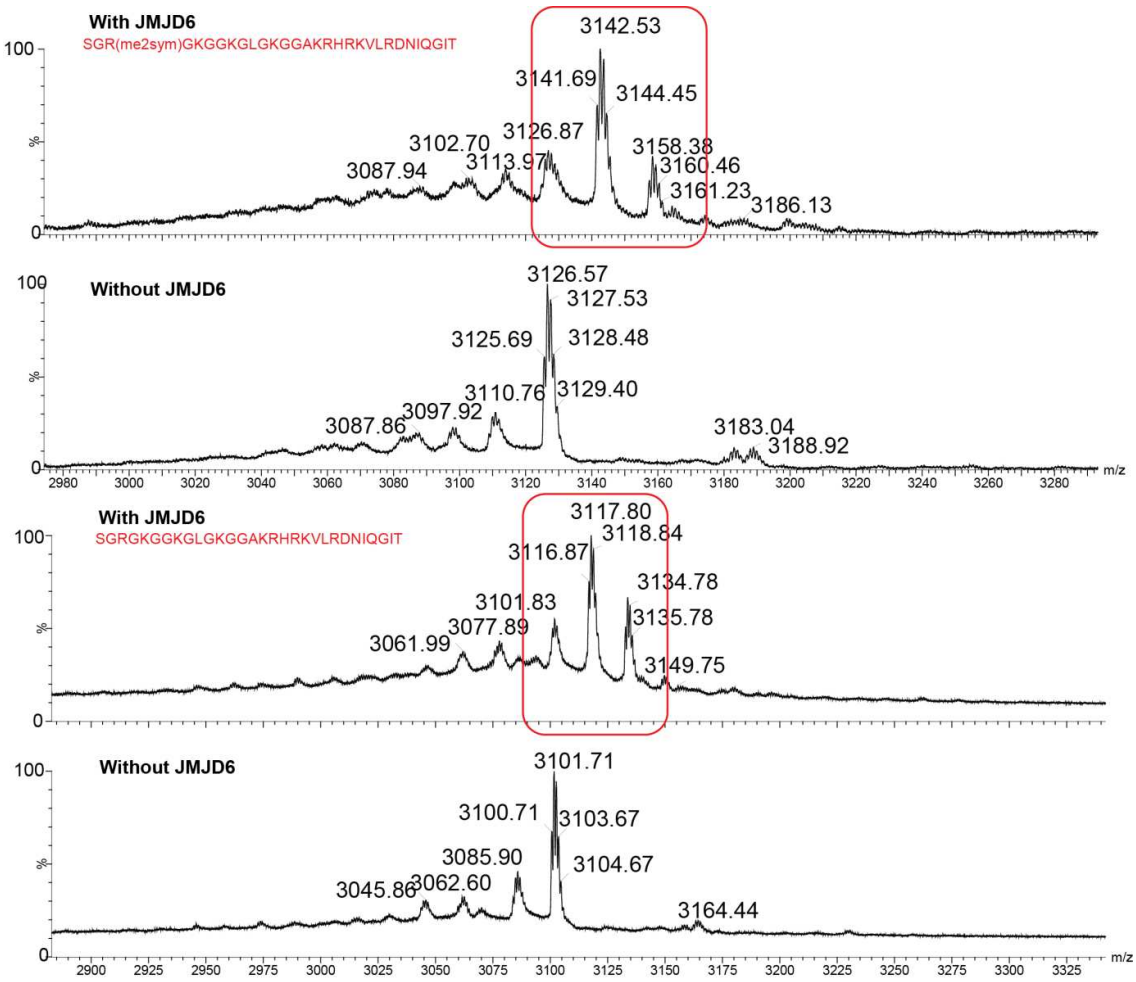


Figure 2.9 MALDI spectra of H4R3 peptides with and without methylated arginine at R3 position when incubated with JMJD6 and cofactors. The MALDI spectra from control samples in which peptides were incubated with cofactors and buffer without the addition of JMJD6 is also shown.

These results indicate that histone arginine demethylation activity for JMJD6 does not occur for any of the peptides that have been tested, at least under these assay conditions. However, there is a probable binding of JMJD6 and histone peptides as seen by ALPHA screen assays. Further, the lysine hydroxylation by JMJD6 is observed in a number of these histone peptides from MALDI-MS assays. Thus, in order to investigate *in vivo* substrates and to investigate lysine hydroxylation by JMJD6, further studies were carried out using human cell lines as described below in Section 2.4.

2.5 Finding substrates for JMJD6

Attempts were then made to find substrates for JMJD6 using *in vivo* immunoprecipitation (IP) as well as *in vitro* peptide substrate studies. IP was carried out in order to analyse binding partners by co-immunoprecipitation followed by LC-MS/MS analysis.

2.5.1 IP of JMJD6

Endogenous expression of JMJD6 and cell lysis

For *in vivo* studies of JMJD6, HeLa and HEK293T human epithelial cell lines were used. The HEK 293T cell line is derived from human embryonic kidney cells and additionally contains the SV40 Large T-antigen,[16] allowing episomal replication of plasmids containing the SV40 origin of replication. Therefore, transfected plasmids can be amplified allowing extended protein expression. The immortal HeLa cell line is derived from a rapidly proliferating cervical cancer cells.[17] Both cell lines were used for studies with endogenous JMJD6 as well as with the overexpression of the JMJD6 gene.

For transfection studies, DNA sequences encoding for full length JMJD6 and GFP-tagged JMJD6 were cloned into plasmids pcDNA3 and pEGFPN1 respectively by Dr. Alexander Wolf. $4x10^8$ cells (8x150 mm plates) were grown in plates using Dulbecco´s Modified Eagles Medium (DMEM) for 18 hours to give 70-80% cell density and then transfected by using Polyethylenimine (PEI).[18] Cells were harvested 24 hours post-transfection by scraping from a petri dishes and cell pellets were stored at -80°C until lysis. Cells were lysed using the previously optimised protocol by *Dr. Aleaxander Wolf* (reported in reference [10]). Cell lysis was carried out by resuspending cell pellet in lysis buffer containing 20 mM Tris.HCl pH 8.0, 500 mM NaCl, 0.5% NP40 supplemented with protease (Pefabloc (Boehringer), pepstatin A, Aprotinin, Leupeptin) and phosphatase inhibitors (100 nM okadaic acid and phosphatase inhibitor cocktails 1 + 2 (Sigma) followed by incubation on ice for 10 minutes.

IP of JMJD6 from cell lysates

IP is a protein isolation technique that involves precipitation of a protein from cell lysate using an antibody which can be immobilised on solid support (such as agarose beads or magnetic beads). IP can be carried out using either direct or indirect IP methods. In the direct method, the cell lysate is added to the beads which are pre-bound with antibody. In the case of the indirect IP method, the cell lysate is incubated with an antibody and then beads coated with protein G, for instance, are added, so that the antibody attached to the protein can form a conjugate with beads, thereby enabling the isolation procedure. Once attached to a desired protein, the beads can be easily isolated from solution by centrifugation (agarose beads) or by magnets (magnetic beads). Successful IP requires a strong, robust antibody-protein interaction which can survive the IP reaction conditions.

Commercially available antibodies to well characterised fusion tags (such as GFP and FLAG) have been developed to withstand such conditions. During the course of this work, both wild-type JMJD6 and GFP-tagged JMJD6 were used, ProteinG Sepharose as well as Agarose-GFP beads (GFP-trap-A, Chromotek, Germany) were used for the IP. A typical protocol in the case of the JMJD6 IP assays is depicted in Figure 2.10.

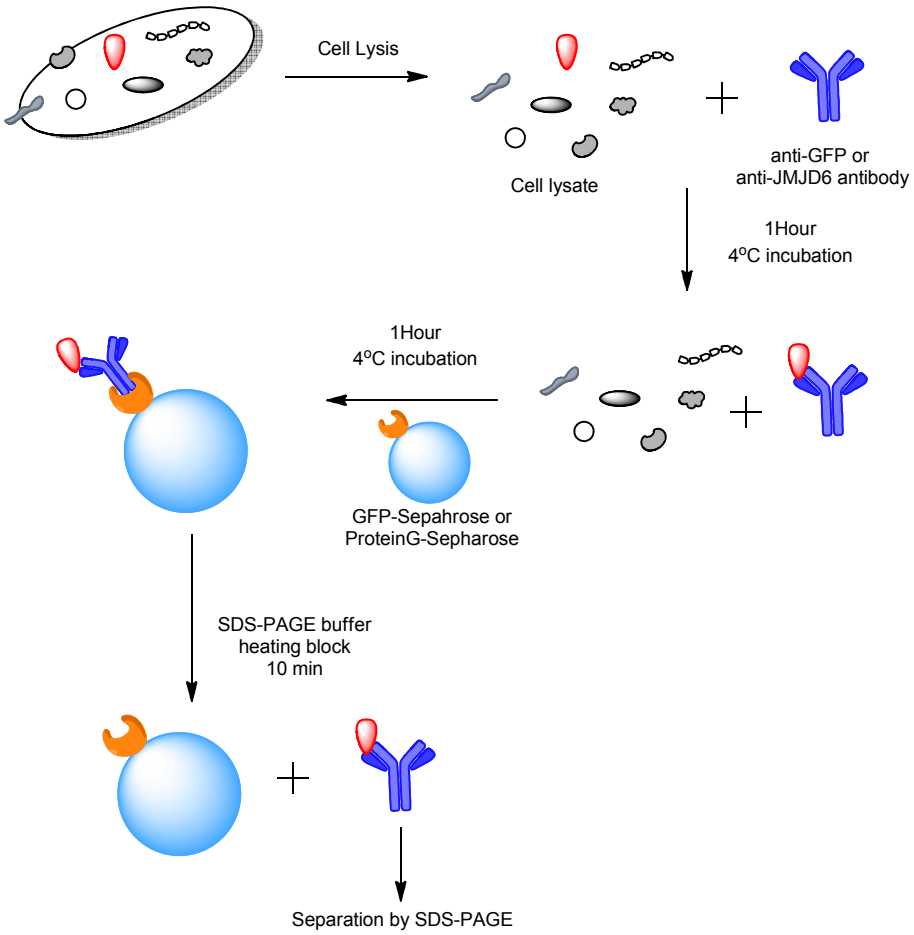


Figure 2.10 Outline of IP protocol for JMJD6 pulldown procedure from human cell line using Sepharose beads as a solid support. JMJD6 expressed in cells is shown in red.

Cell lysates prepared as described previously, were centrifuged and the supernatant was then incubated with either anti-JMJD6 or anti-GFP for 1 hour under agitation at 4°C. Protein-G Sepharose beads were added to the anti-JMJD6 antibody sample (agarose-

GFP beads were added to anti-GFP antibody sample) and incubation was carried for another hour. After washing with wash buffer (as explained in Chapter 8) thrice, the beads were re-suspended in SDS sample buffer followed by heating for 10 min on heating block which ensures the separation of antibody-protein complex from beads. The protein mixture was then subjected to SDS-PAGE electrophoresis using gradient 4-12% NuPAGE gel (Invitrogen).



Figure 2.11 4-12 % gradient NuPAGE gel run for immunoprecipitated GFP-JMJD6 when overexpressed in HeLa cells and stained with Colloidal Coomassie. The gel bands which were excised from gel and then subjected to trypsinolysis followed by LC-MS/MS analysis are labelled as 1-10.

Each lane was cut into 10 bands (as shown in Figure 2.11) which were further cut into 1mm pieces, washed thoroughly with water and acetonitrile. Cystine residues were reduced to their thiols by DTT and then selectively alkylated with iodoacetamide to reduce their nucleophilicity. Further treatment with specific peptidase was carried out before LC-MS/MS analysis as described in Chapter 8. Trypsin is a common serine protease used for MS based protein identification via peptide mass fingerprinting. In

some cases, other peptidases, endoproteinase Glu-C and elastase were also used for better coverage of sequence in proteomic analyses.

2.5.2 LC-MS/MS results confirming the association of JMJD6 with arginine-serine rich (SR) proteins involved in pre-mRNA splicing

LC-MS/MS analysis of tryptic digest of SDS-PAGE gel bands was carried out by *Dr. Mukram Mackeen at the Centre for Cellular and Molecular Physiology, Oxford*. The mass profiles obtained for each lane were merged, then searched using the MASCOT search engine version 2.3.01 (Matrix Science), which uses MS peptide data to identify protein matches present in the UniProtKB/Swiss-Prot database (Version 2010.07.16; 518,415 sequences). The LC-MS/MS results of immunoprecipitated JMJD6 from HeLa cells when overexpressed as GFP tagged JMJD6 and further subjected to trypsionolysis are shown in Table 2.2. The first 40 proteins when sorted by exponentially modified protein abundance index (emPAI) score are shown. It can be seen that most of the proteins (22 out of 40 are mRNA splicing proteins) associated with JMJD6 are proteins involved in pre-mRNA splicing mechanism including U2AF65 and U2AF35. Common likely contaminating proteins such as trypsin, keratin, heat shock proteins and immunoglobulins were also present in the data. This data correlates with the results obtained by Webby et al. using tandem affinity tag purification (TAP-tag) assays,[10] confirming the association of JMJD6 with the pre-mRNA splicing machinery. The results support the proposal that JMJD6 plays a role in the regulation of pre-mRNA splicing rather than its involvement in chromatin regulation.

Table 2.2 List of potential proteins interacting with JMJD6 as analysed LC-MS/MS carried out after in gel trypsinolysis of immunoprecipitated GFP tagged JMJD6 from HeLa cells. The first 40 proteins when sorted by exponentially modified protein abundance index (emPAI) score are shown. Proteins involved in RNA splicing are shown in red colour.

Acc. No.	emPAI	Score	Protein Name
Q6NYC1	9.17	2048	Bifunctional arginine demethylase and lysyl-hydroxylase JMJD6
Q9Y383	4.46	1572	Putative RNA-binding protein Luc7-like 2
Q15365	4.32	934	Poly(rC)-binding protein 1 or Heterogeneous nuclear ribonucleoprotein E1 (hnRNP)
O95218	3.92	830	Zinc finger Ran-binding domain-containing protein 2
P62316	3.77	296	Small nuclear ribonucleoprotein (snRNP core protein D2)
O43809	3.72	1056	Cleavage and polyadenylation specificity factor subunit 5
P05387	2.64	228	60S acidic ribosomal protein P2
Q14498	2.63	1206	RNA-binding protein 39 (Splicing factor HCC1)
Q01081	2.38	873	Splicing factor U2AF 35 kDa subunit
Q9BWJ5	2.22	99	Splicing factor 3B subunit 5 (Pre-mRNA-splicing factor SF3b 10 kDa subunit)
Q16630	2.1	1213	Cleavage and polyadenylation specificity factor subunit 6 (Pre-mRNA cleavage factor Im 68 kDa subunit)
Q05519	2.1	1051	Serine/arginine-rich splicing factor 11
P09661	2.02	570	U2 small nuclear ribonucleoprotein A' (U2 snRNP A')
Q9UHX1	1.75	1212	Poly(U)-binding-splicing factor PUF60; (60 kDa poly(U)-binding-splicing factor)
Q9NQ29	1.75	748	Putative RNA-binding protein Luc7-like 1
P08107	1.73	1036	Heat shock 70 kDa protein 1A/1B
P14174	1.61	101	Macrophage migration inhibitory factor;
P11142	1.47	987	Heat shock cognate 71 kDa protein
P07237	1.44	616	Protein disulfide-isomerase
Q5VYJ4	1.3	107	Putative small nuclear ribonucleoprotein polypeptide E-like protein 1
P07437	1.29	528	Tubulin beta chain
P50990	1.11	516	T-complex protein 1 subunit theta
P62857	1.09	107	40S ribosomal protein S28

Q15366	1.08	592	Poly(rC)-binding protein 2 (hnRNP E2)
A8MWD9	0.99	87	Small nuclear ribonucleoprotein G-like protein
P04264	0.97	471	Keratin, type II cytoskeletal 1
Q8WVK2	0.93	189	U4/U6.U5 small nuclear ribonucleoprotein 27 kDa protein (U4/U6.U5 snRNP 27 kDa protein)
P62318	0.92	136	Small nuclear ribonucleoprotein Sm D3 (snRNP core protein D3)
P62258	0.91	181	14-3-3 protein epsilon
P55769	0.89	78	High mobility group-like nuclear protein 2 homolog 1 (U4/U6.U5 tri-snRNP 15.5 kDa protein)
P37108	0.87	292	Signal recognition particle 14 kDa protein (SRP14) or 18 kDa Alu RNA-binding protein
Q9Y3B4	0.87	111	Pre-mRNA branch site protein p14 (SF3b 14 kDa subunit)
P78371	0.84	316	T-complex protein 1 subunit beta
P05455	0.84	199	Lupus La protein, La ribonucleoprotein
P26368	0.81	498	Splicing factor U2AF 65 kDa subunit
P39019	0.77	77	40S ribosomal protein S19
Q7L4I2	0.76	310	Arginine/serine-rich coiled-coil protein 2
P61604	0.74	72	10 kDa heat shock protein, mitochondrial
P34931	0.72	415	Heat shock 70 kDa protein 1-like
Q9BU76	0.7	293	Multiple myeloma tumor-associated protein 2

In addition to trypsin, endoproteinase Glu-C and elastase were also investigated in proteolytic cleavage protocols. When JMJD6 was expressed at endogenous levels and subjected to Glu-C (Roche) activity, the coverage of peptides from various protein sequences as determined by MASCOT was poor. It has been seen by *in vitro* assays that JMJD6 catalyses hydroxylation of lysine residues from arginine serine domain of SR proteins.[10] Thus, in order to maintain arginine serine domain of SR proteins intact, during peptidase cleavage, porcine pancreatic elastase (Sigma) was used. When elastase was used for immunoprecipitated GFP-tagged JMJD6 samples, it did show that number of proteins pulled down along with JMJD6 were SR proteins (Results shown in

Appendix B). But the region covering arginine serine domains in many of the proteins was still missing in these SR proteins.

2.6 Pre-mRNA Splicing and SR proteins

Most eukaryotic genes contain introns which are removed from pre-mRNA by the spliceosome to give mature mRNA which can later direct the protein synthesis. The biochemical mechanism by which splicing occurs has been studied in a number of systems and is now fairly well characterised. Introns are removed from primary transcripts by cleavage at conserved sequences called splice sites.[19] These sites are found at the 5′ and 3′ ends of introns. Most commonly, the intron that has to be removed starts with the dinucleotide GU at its 5′ end, and ends with AG at its 3′ end. These consensus sequences are known to be critical, because changing one of the conserved nucleotides results in inhibition of splicing. Upstream of the AG there is a region which is rich in pyrimidines and thus termed as polypyrimidine tract. A typical sequence is YNYYRAY.[19] Another important consensus sequence exists at the branch point, located anywhere from 18 to 40 nucleotides upstream from the 3′ end of an intron. The branch point always contains an adenine, but it is otherwise loosely conserved. A typical mRNA is depicted in Figure 2.11.

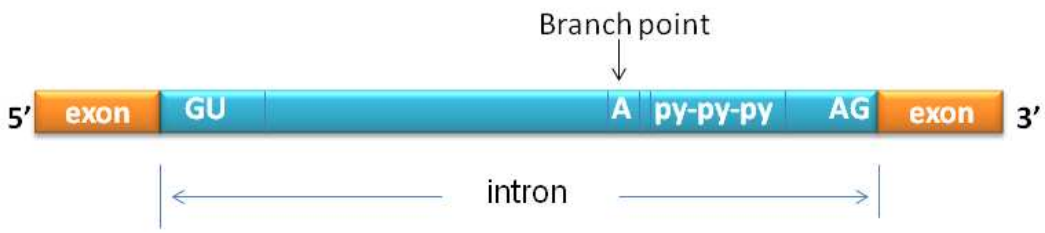


Figure 2.11 Pre-mRNA structure showing regions involved in splicing.

The removal of introns is catalysed by the spliceosome which consists of five small nuclear ribonuclear protein complexes (snRNPs) U1, U2, U4, U5 and U6 that are associated with a large number of additional proteins.[19-21] The spliceosome assembles onto the pre-mRNA in a coordinated series of intricate movements by binding to regulatory pre-mRNA sequences as shown in Figure 2.12. Spliceosome assembly is initiated by the stable associations of (i) U1 snRNP with the 5′ splice site, (ii) branch point binding protein SF1 with the branchpoint, and (iii) U2 snRNP auxiliary factor (U2AF) with the pyrimidine tract (E complex). ATP hydrolysis then leads to the formation of the A complex, which is characterised by the stable association of U2 snRNP at the branchpoint. After the incorporation of the U4/U6-U5 tri-snRNP (B complex), the catalytic spliceosome is formed by large structural rearrangements. Finally, a rearrangement occurs to form the catalytically active C complex, in which U2 and U6 interact making the 5' end of introns to ligate with the branchpoint adenine; U5 binds the exon at the 3' splice site. The C Complex of U2/U5/U6 remains bound to the intron and the 3' splice site is cleaved and exons are ligated using ATP hydrolysis, whereas the excised intron is released.[22] A common feature of all these rearrangements is the repeated breaking and formation of base-pair interactions either between snRNAs or between the pre-mRNA and snRNAs.

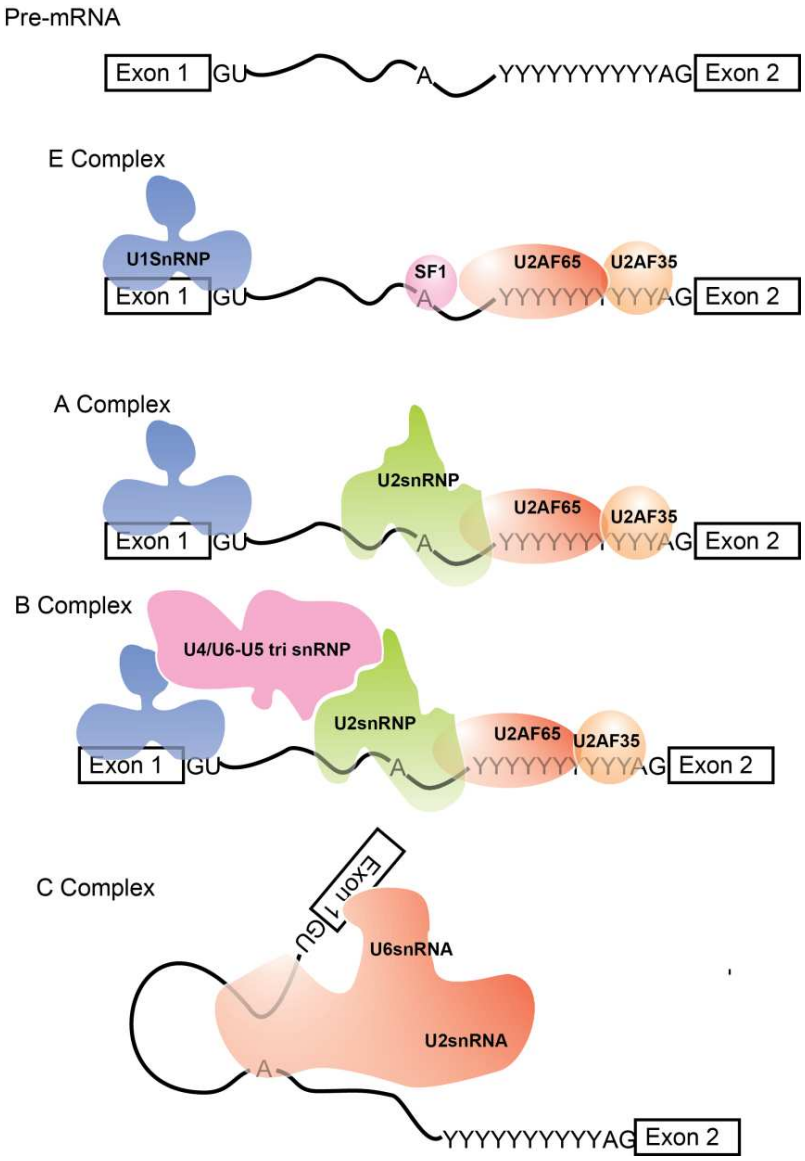


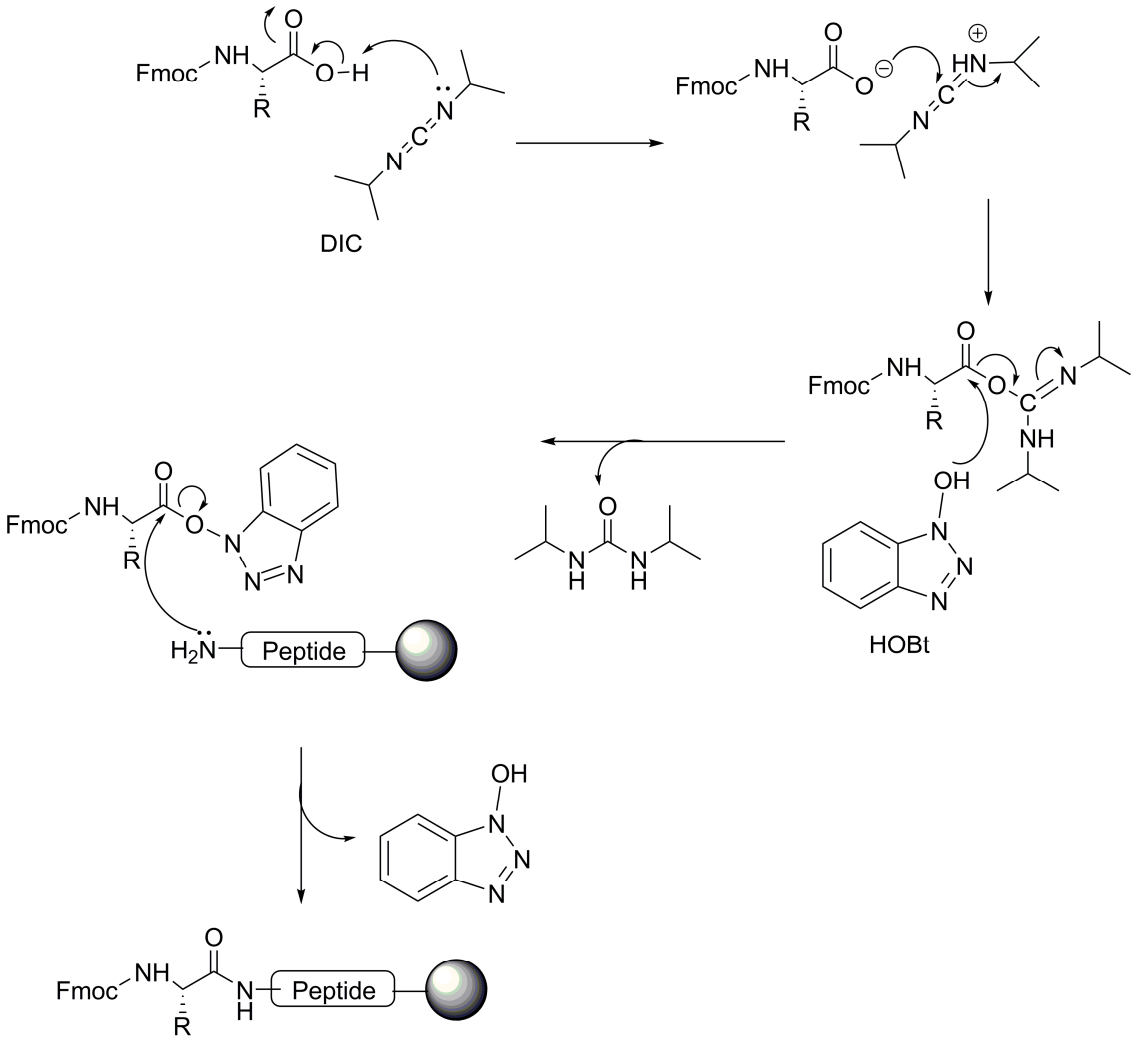
Figure 2.12 The spliceosome assembles onto pre-mRNA in a stepwise manner forming various structural complexes.

Splicing is a very complex phenomenon involving a large number of protein interactions and post-translational modifications occurring at various stages of the splicing process. One of the characteristic features of the proteins involved in RNA splicing is that they contain one or two RNA recognition motif (RRM) type RNA binding domain and variable *C*-terminal arginine-serine rich domains (RS domain); thus

these proteins are termed as SR proteins .[23, 24] Members of the SR protein family are known to be involved in alternative and constitutive splicing. SR proteins can be heavily phosphorylated on serine residues in their RS domain. These modifications play an important role in regulating subcellular localisation and activity of SR proteins.[24] Webby et al. have shown that JMJD6 catalyses hydroxylation of lysine residues from RS domains of Luc7-like protein 2 (LUC7L2), Luc7-like protein 3 (CROP) and U2AF65.[10] During the course of this thesis attempts have been made to investigate further lysine hydroxylation of SR proteins associated with JMJD6 by MALDI-MS based peptide assays.

LC-MS/MS analysis of immunoprecipitated JMJD6 showed probable lysine hydroxylation in various SR proteins. The lysine hydroxylation of these residues was further tested by carrying out MALDI-MS assays on synthesised peptides from these SR protein sequences.

2.7 Synthesis of potential peptide substrates for JMJD6



Scheme 2.1 Scheme showing activation of Fmoc protected amino acids by DIC and HOBt and coupling of activated carboxyl group on to free amino group of resin bound amino acid.

Various potential peptide substrates based on the results of the LC-MS/MS analysis were synthesised by solid phase peptide synthesis using Fmoc based protection chemistry. The peptides were synthesised as *C*-terminal amides coupled to a Rink amide resin (solid support) using a CS Bio peptide synthesiser as shown in Figure 2.13. The amino group at the *N*-terminal of the amino acid is protected by a base-labile Fmoc (9-fluorenylmethoxycarbonyl) protecting group. Side-chains are protected by acid-labile

groups, such as tertiary-butyl ester (tBu) or *N*-tert-butyloxycarbonyl group (BOC). The reaction was initiated with pre-swelling the resin in DMF, followed by the addition of the Fmoc protected amino acid. Activation of amino acids was carried out using *N*-hydroxybenzotriazole (HOBt) and *N,N′*-di-isopropylcarbodiimide (DIC) as shown in Scheme 2.1. The coupling of activated amino acid to the resin or previously coupled amino acid was carried out by mixing for 2 hours. The Fmoc group of amino acid bound to the solid support was removed by reaction with 20% piperidine in DMF/DMSO so that the free amino group could couple with the activated carboxylic acid group of newly added Fmoc protected amino acid. The entire cycle was repeated and the final peptide was cleaved from resin using trifluroacetic acid (TFA) in the presence of silane scavenger. The acid labile side chain protecting groups (Boc, t-butyl, trityl and Pbf groups) were also removed during the TFA cleavage step. Whenever required, peptides were purified by HPLC using reverse phase C18 column as described in Chapter 8.

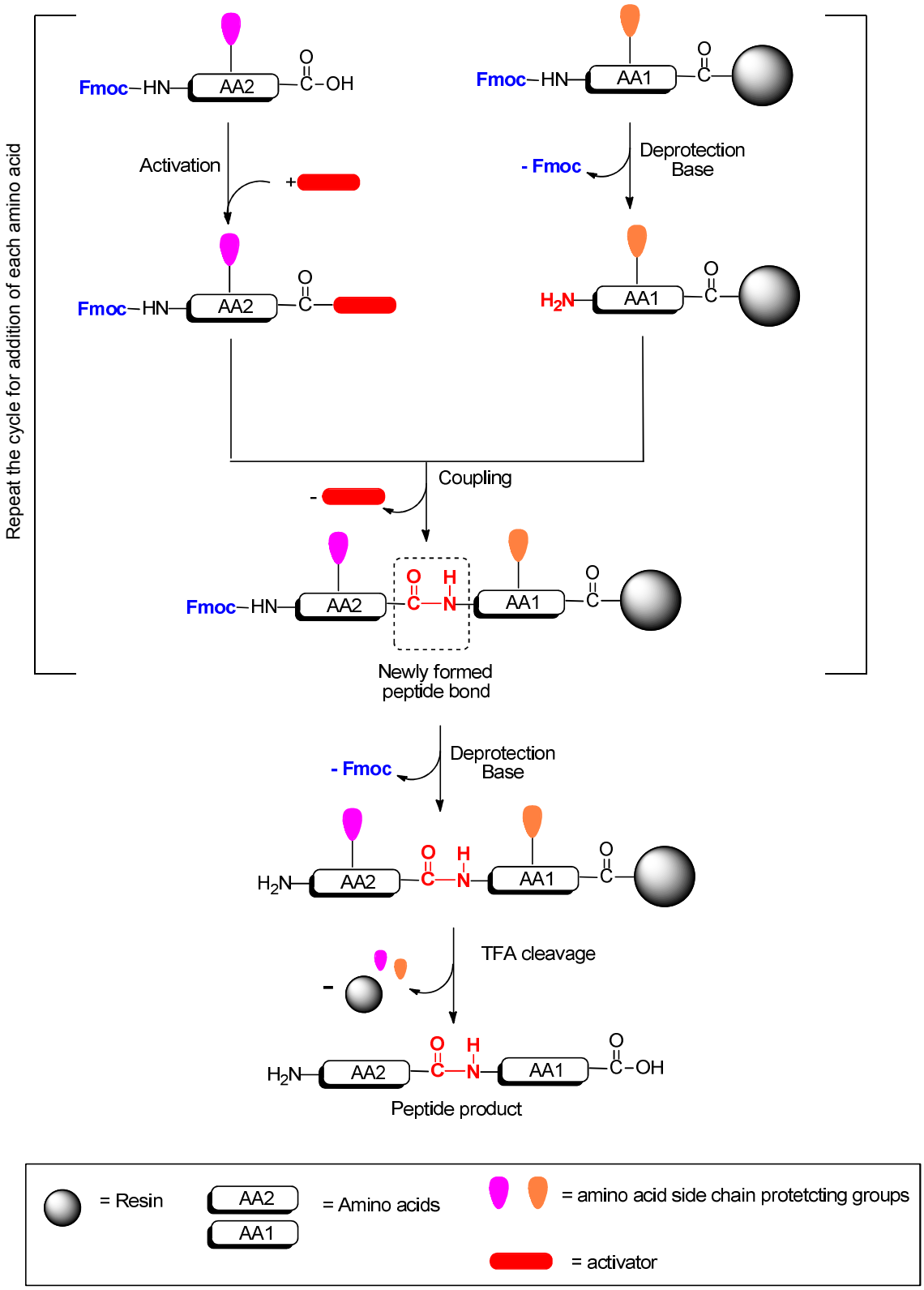


Figure 2.13 A scheme depicting the solid phase peptide synthesis cycle including deprotection, coupling, and cleavage steps.

2.8 MALDI-MS assays of peptides from SR proteins associated with JMJD6 in IP

2.8.1 U2AF65 and U2AF35 interaction and catalytic hydroxylation by JMJD6

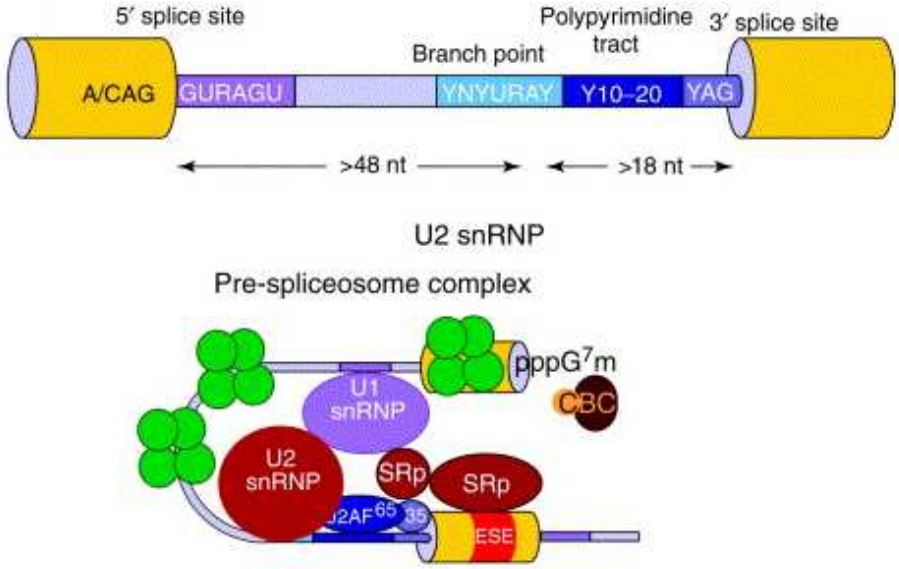


Figure 2.14. Splicesite elements and pre-spliceosome complex assembly. Consensus splice-site elements for a typical metazoan intron are shown. During the formation of pre-spliceosome U2AF65 binds polypyrimidine tract, and U2AF35 binds the 3′ splice site AG. These interactions allow assembly of an early or commitment complex. The figure is modified from ref [20].

In pre-mRNA splicing, spliceosome assembly is directed by conserved pre-mRNA sequences, including a short consensus sequence at each exon-intron junction and a branch point sequence followed by a polypyrimidine tract, located within the intron adjacent to the 3′ splice site.[25] U2 snRNP auxiliary factor is an essential pre-mRNA splicing factor complex, comprising a smaller 35 kDa (U2AF35) subunit and a larger 65 kDa (U2AF65) subunit. U2AF65 binds specifically to the pyrimidine tract via its *C*-terminal RNA-binding domain, which consists of three RNA recognition motifs (RRMs).[26] U2AF65 stabilises the interaction of U2 snRNP with the mRNA branchpoint by contacting the branch region through its *N*-terminal RS domain, promoting base pair

interactions between U2 snRNA and the branchpoint.[27-30] In addition to its well-characterised role in constitutive splicing, U2AF65 also plays an important role in alternative splicing.[20, 31-33] The U2AF35 binding domain of U2AF65 is located in the proline rich fragment between the RS domain and RNA-binding region.[34, 35]

The smaller subunit, U2AF35 interacts with the conserved AG dinucleotide at the 3′ splice site of introns.[36-38] Its primary structure comprises a central RRM that is flanked by two zinc fingers in the *N*-terminus,[39] an RS domain and a glycine tract at its *C*-terminus.[40, 41] In addition to constitutive splicing, U2AF35 plays an important role in the splicing of both regulated introns and introns that have weak polypyrimidine tracts.[42] U2AF35 binds both U2AF65 and the pre-mRNA through its RRM domain. Several studies support a model in which splicing enhancer-bound SR proteins function by recruiting U2AF65 to weak polypyrimidine tracts.[43, 44] In this model, recruitment requires interactions between the SR proteins and U2AF35.

In various JMJD6-IP assays, U2AF65 and U2AF35 were always found to co-purify with JMJD6. It was reported by Webby et al. that endogenous U2AF65 can be co-immunoprecipitated with JMJD6 and that the interaction is disrupted by application of RNase.[10] They also reported a hydroxylation at K15 and K276 of endogenous U2AF65. Recombinant U2AF65 did also show hydroxylation after incubation with JMJD6, 2OG and Fe(II), whereas the peptides $U2AF65_{12\text{-}26}$ and $U2AF65_{274\text{-}287}$ were not found to be substrates by *in vitro* assays.

Table 2.3 Peptides from U2AF65 and U2AF35 tested by MALDI-MS assays.

Peptide No.	Peptide Sequence	Mol. Wt.	Sequence from	MALDI Results
22	PPGFEHITPMQYKAMQAAG	2074.4	$U2AF65_{96\text{-}114}$	no
23	RELYGRRRKKHRSRSR	2141.4	$U2AF35_{183\text{-}198}$	Double hydroxylation

When LC-MS/MS data resulting from the IP assays was analysed by the central proteomics facilities pipeline (CPFP) search engine,[45] U2AF65 peptide was shown to have a potential lysine hydroxylation sites at K108. Thus, peptide $U2AF65_{96\text{-}114}$ was synthesised by solid phase peptide synthesis. The peptide was further incubated with JMJD6, 2OG and Fe(II) and analysed by MALDI-MS. But MALDI spectra did not show a peak for +16 (Figure 2.15).

The peptide sequences of U2AF35 that were identified in LC-MS/MS results did not show any lysine hydroxylation. Similarly to U2AF65, U2AF35 also has a RS domain including lysine residues. But this region was never covered in the LC-MS/MS analysis, possibly due to peptidase cleavage in this region due to high arginine content. Hence, a peptide covering U2AF35, RS-domain sequence corresponding to residues 183-198 was synthesised and tested by MALDI-MS assays. MALDI spectra for $U2AF35_{183\text{-}198}$ when incubated with/without JMJD6 and other cofactors necessary for catalytic hydroxylation are shown in Figure 2.15. New peaks at 2158.22 and 2174.22, corresponding to +16 and +32 Da relative to unmodified peptide indicate that $U2AF35_{183\text{-}198}$ is *in vitro* substrate for JMJD6 undergoing double hydroxylation (Figure 2.15).

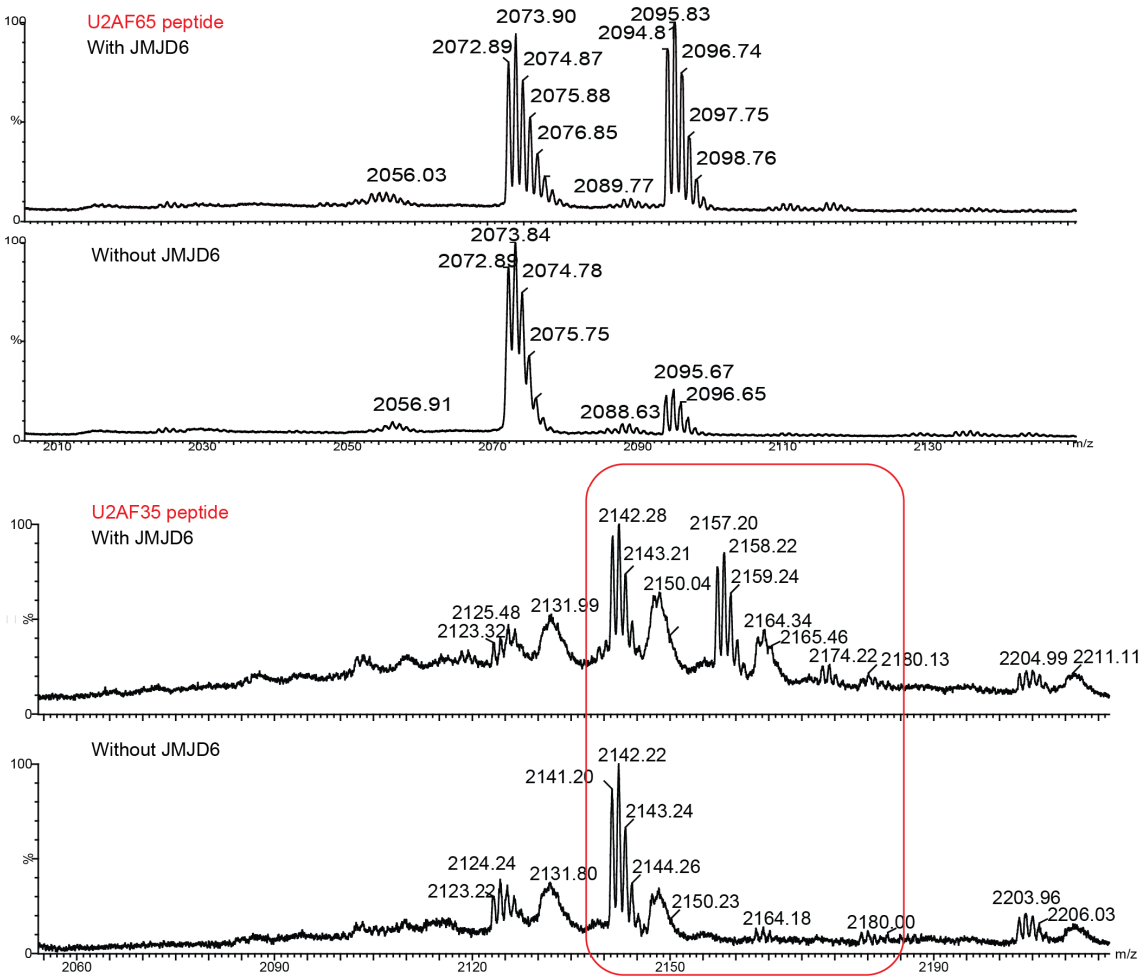


Figure 2.15 MALDI spectra of U2AF65$_{96\text{-}114}$ and U2AF35$_{183\text{-}198}$ peptides when incubated with JMJD6 and cofactors. The MALDI spectra from control samples in which peptides were incubated with cofactors and buffer without addition of JMJD6 is also shown. In case of U2AF35$_{183\text{-}198}$ peptides, MALDI-MS spectra shows peaks at +16 (2158.22) and +32 (2174.22) indicating occurrence of double hydroxylation.

2.8.2 Lysine hydroxylation in LUC7L2, SRSF11, ZRANB2

As described in Section 2.4, the IP assays indicate that JMD6 may be associated with a large number of SR proteins. Further analysis of these proteins by MASCOT MS/MS ion search showed potential lysine hydroxylation sites at K242 of serine/arginine-rich splicing factor 11 (SRSF11), K196 of zinc finger RAN-binding domain containing 2 (ZRANB2) and K122 of LUC7L2 (Figure 2.16, 2.17 and 2.18). To investigate *in vitro*

hydroxylation of lysine residues at these positions, peptides from these SR proteins were synthesised and then subjected to MALD-MS assays. The set of peptides shown in Table 2.4 were synthesised by solid phase peptide synthesis. Previously, it was shown that JMJD6 catalyses the hydroxylation of K269 from LUC7L2 using *in vitro* assays.[10] However, incubation of $LUC7L2_{111\text{-}130}$ peptide with JMJD6 did not show any +16 peak when analysed by MALDI-MS. SRSF11 is known to co-localise with the spliceosome[46] and it forms *in vitro* stable complex with poly-U binding splicing factor 60 (PUF60).[47] In all the IP-LC-MS/MS analyses carried out in this thesis and previously by Webby et al., PUF60 is also consistently found to be associated with JMJD6. However, the peptides $SRSF11_{234\text{-}255}$ with K242 did not show any hydroxylation. Similarly, peptide $ZRANB2_{191\text{-}210}$ did not show any peak corresponding to hydroxylation in MALDI-MS spectrum obtained after incubation with JMJD6 and other cofactors. ZRANB2 is a highly conserved protein that can regulate alternative splicing and has single-stranded RNA-binding domains that recognize 5' splice site-like sequences.[48]

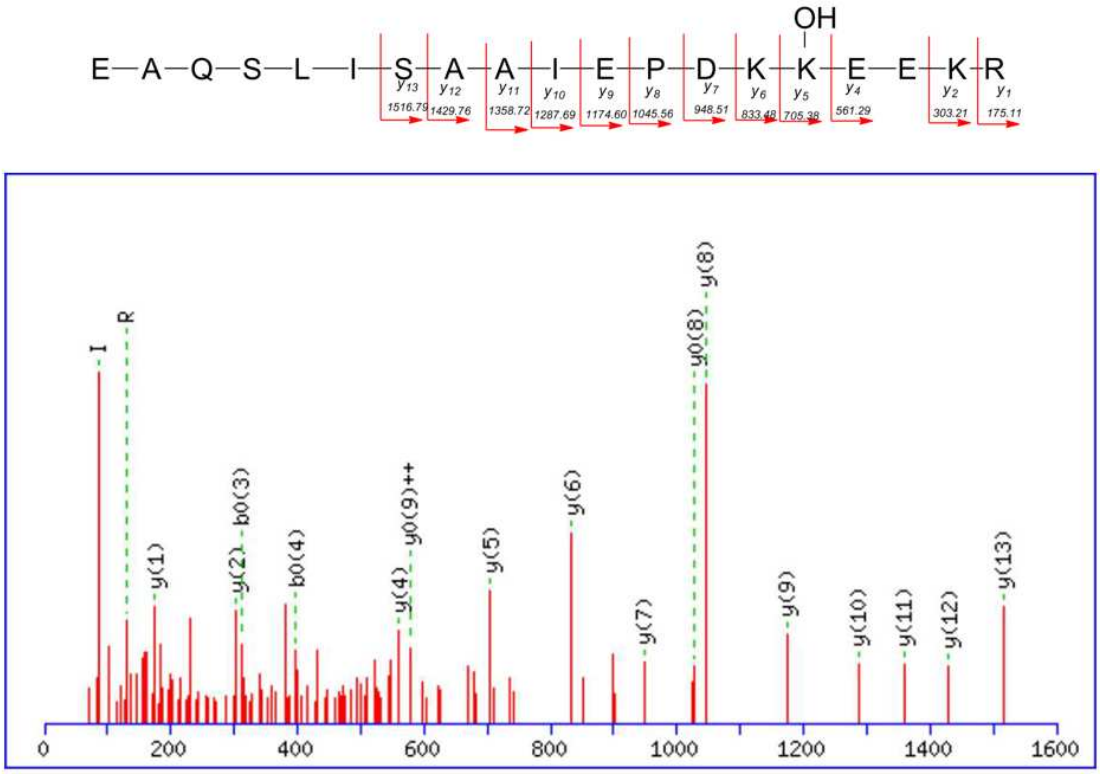


Figure 2.16 MASCOT MS/MS ion search result (with ion score 90) assigning the lysine oxidation at K242 on the SRSF11 sequence.

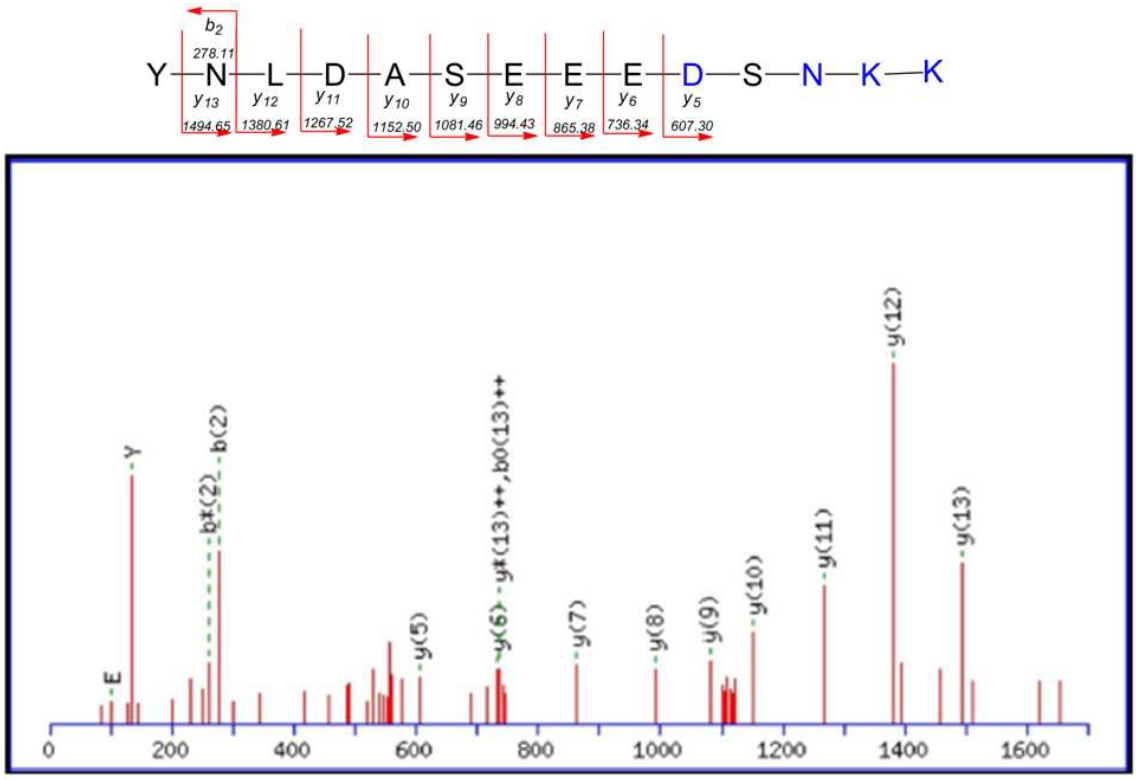


Figure 2.17 MASCOT MS/MS ion search result (with ion score 62) assigning probable (lysine) oxidation at ZRANB2$_{83-96}$ sequence. Possible sites of oxidation are shown in blue.

Table 2.4 Peptides designed from probable lysine hydroxylation seen by IP and LC-MS/MS assays. These peptides were tested by MALDI-MS assays.

Peptide No.	Peptide Sequence	Mol. Wt.	Sequence from	MALDI Results
24	TQEEISAEVAAKAERVHELN	2224.4	LUC7L2$_{111-130}$	No
25	EDSNKKKSNRRSRSKSRSSH	2374.5	ZRANB2$_{191-210}$	No
26	SAAIEPDKKEEKRRHSRSRS	2367.6	SRSF11$_{234-255}$	No

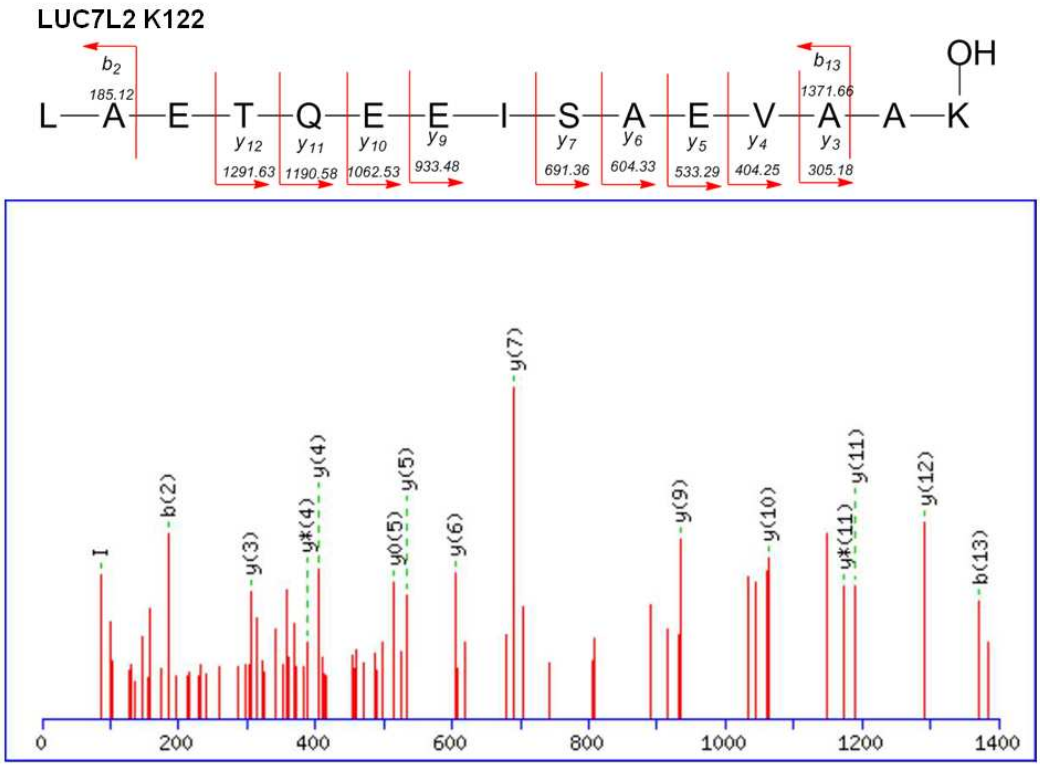


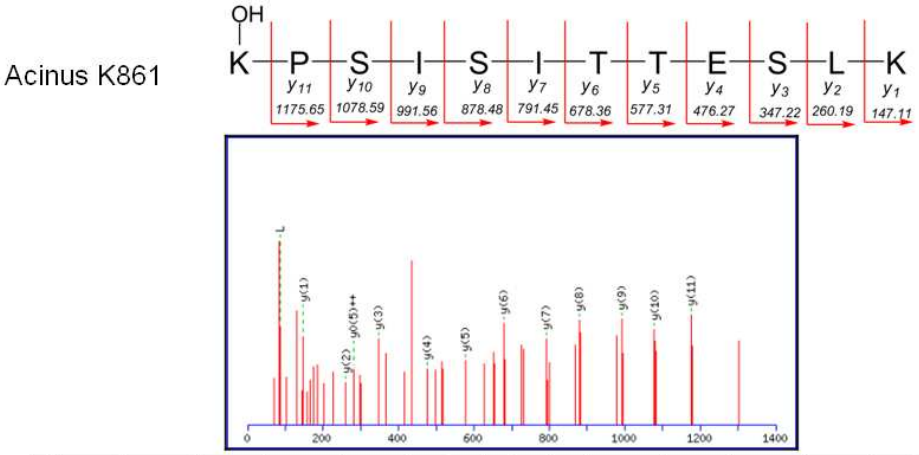
Figure 2.18 MASCOT MS/MS ion search result (with ion score 62) assigning the lysine oxidation at K122 on the LUCL2 sequence.

2.9 Lysine hydroxylation in ACINUS

Apoptotic chromatin condensation inducer of the nucleus (ACINUS) is a nuclear protein initially proposed to be involved in apoptotic chromatin condensation.[49] ACINUS is expressed as different isoforms, which are most likely generated by alternative splicing. During apoptosis ACINUS is cleaved by caspase-3 and a still unknown protease to produce a 23 kDa fragment (p23) that is contained in all described non apoptotic isoforms of ACINUS. Importantly, ACINUS-p23 has been reported to mediate apoptotic chromatin condensation prior to DNA fragmentation.[49] siRNA analysis of ACINUS indicates that loss of ACINUS inhibits oligonucleosomal DNA fragmentation but not chromatin condensation during apoptosis.[50] ACINUS has a P-

loop motif[51] and a region similar to the RNA-recognition motif of splicing regulator sex-lethal protein (Sxl)[52], indicating that ACINUS might have ATPase and DNA/RNA-binding activity. ACINUS is located in nuclear interchromatin granule clusters, which are thought to function as storage sites for splicing factors.[53] ACINUS was also identified as a component of functional spliceosomes and as a subunit of a novel apoptosis and splicing-associated protein (ASAP) complex.[54-56] Recently, ASAP has been identified as forming a part of the exon junction complex, a multiprotein complex that is deposited at exon-exon junctions during RNA splicing.[57] Thus ACINUS is proposed to be involved in both apoptosis and pre-mRNA processing, which are not necessarily exclusive.

ACINUS is one of the SR proteins which were present in all the JMJD6 IP experiments carried out by our collaborator, Prof. Angelika Böttger's laboratory as well as our previous attempts. Because of these interesting results, further experiments were carried out to identify lysine hydroxylation in ACINUS sequence. GFP-tagged ACINUS was cloned into the pEGFP-N1-plasmid and full length JMJD6 sequence was cloned into the pHA-plasmid. HEK293T cells were transfected with these plasmids and then ACINUS was immunoprecipitated using GFP-trap-A beads (Chromotek, Germany) and followed by separation using SDS-PAGE. These bands were further subjected to trypsinolysis and analysed by LC-MS/MS analysis. *Astrid Heim carried out cloning, overexpression and immunoprecipitation experiments and Dr. Mukram Mackeen carried out the LC-MS/MS analysis.*



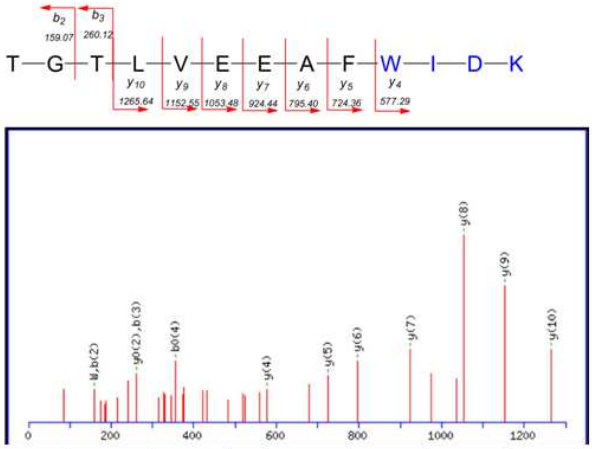
#	Immon.	b	b^{++}	b*	b^{*++}	b^0	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^0	y^{0++}	#
1	117.1022	145.0972	73.0522	128.0706	64.5389			K							12
2	70.0651	242.1499	121.5786	225.1234	113.0653			P	*1175.6518*	588.3295	1158.6252	579.8163	1157.6412	579.3243	11
3	60.0444	329.1819	165.0946	312.1554	156.5813	311.1714	156.0893	S	*1078.5990*	539.8032	1061.5725	531.2899	1060.5885	530.7979	10
4	86.0964	442.2660	221.6366	425.2395	213.1234	424.2554	212.6314	I	*991.5670*	496.2871	974.5405	487.7739	973.5564	487.2819	9
5	60.0444	529.2980	265.1527	512.2715	256.6394	511.2875	256.1474	S	*878.4829*	439.7451	861.4564	431.2318	860.4724	430.7398	8
6	86.0964	642.3821	321.6947	625.3556	313.1814	624.3715	312.6894	I	*791.4509*	396.2291	774.4244	387.7158	773.4403	387.2238	7
7	74.0600	743.4298	372.2185	726.4032	363.7053	725.4192	363.2132	T	*678.3668*	339.6871	661.3403	331.1738	660.3563	330.6818	6
8	74.0600	844.4775	422.7424	827.4509	414.2291	826.4669	413.7371	T	*577.3192*	289.1632	560.2926	280.6499	559.3086	280.1579	5
9	102.0550	973.5201	487.2637	956.4935	478.7504	955.5095	478.2584	E	*476.2715*	238.6394	459.2449	230.1261	458.2609	229.6341	4
10	60.0444	1060.5521	530.7797	1043.5255	522.2664	1042.5415	521.7744	S	*347.2289*	174.1181	330.2023	165.6048	329.2183	165.1128	3
11	86.0964	1173.6361	587.3217	1156.6096	578.8084	1155.6256	578.3164	L	*260.1969*	130.6021	243.1703	122.0888			2
12	101.1073							K	*147.1128*	74.0600	130.0863	65.5468			1

Figure 2.19 MASCOT MS/MS ion search result (with ion score 57) assigning the lysine oxidation at K861 in the ACINUS sequence

Subsequent LC-MS/MS analyses revealed probable lysine hydroxylation at K861 and K1047 in ACINUS. The MASCOT MS/MS ion search results for the K861 and K1047 containing peptides are shown in Figures 2.19 and 2.20 respectively. These hydroxylation sites were further tested for JMJD6 catalysed hydroxylation in MALDI-MS assays for peptides. ACINUS contains two RS domains but K861 and K1074 are not present in either of the RS rich regions. Thus, various peptides as shown in Table 2.5 including those corresponding to K861 and K1074 as well as three peptides from the RS domains were synthesised by solid phase peptide synthesis. These peptides were

incubated with full length JMJD6 in the presence of Fe(II) and 2OG at 37°C and then analysed by MALDI-MS.

Acinus K1047



#	Immon.	b	b^{++}	b^0	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^0	y^{0++}	#
1	74.0600	102.0550	51.5311	84.0444	42.5258	T							13
2	30.0338	**159.0764**	80.0418	141.0659	71.0366	G	1423.7104	712.3588	1406.6838	703.8456	1405.6998	703.3535	12
3	74.0600	**260.1241**	130.5657	242.1135	121.5604	T	1366.6889	683.8481	1349.6624	675.3348	1348.6783	674.8428	11
4	86.0964	373.2082	187.1077	355.1976	178.1024	L	***1265.6412***	633.3243	1248.6147	624.8110	1247.6307	624.3190	10
5	72.0808	472.2766	236.6419	454.2660	227.6366	V	***1152.5572***	576.7822	1135.5306	568.2689	1134.5466	567.7769	9
6	102.0550	601.3192	301.1632	583.3086	292.1579	E	***1053.4888***	527.2480	1036.4622	518.7347	1035.4782	518.2427	8
7	102.0550	730.3618	365.6845	712.3512	356.6792	E	***924.4462***	462.7267	907.4196	454.2134	906.4356	453.7214	7
8	44.0495	801.3989	401.2031	783.3883	392.1978	A	***795.4036***	398.2054	778.3770	389.6921	777.3930	389.2001	6
9	120.0808	948.4673	474.7373	930.4567	465.7320	F	***724.3665***	362.6869	707.3399	354.1736	706.3559	353.6816	5
10	159.0917	1134.5466	567.7769	1116.5360	558.7717	W	***577.2980***	289.1527	560.2715	280.6394	559.2875	280.1474	4
11	86.0964	1247.6307	624.3190	1229.6201	615.3137	I	391.2187	196.1130	374.1922	187.5997	373.2082	187.1077	3
12	88.0393	1362.6576	681.8324	1344.6470	672.8272	D	278.1347	139.5710	261.1081	131.0577	260.1241	130.5657	2
13	117.1022					K	163.1077	82.0575	146.0812	73.5442			1

Figure 2.20 MASCOT MS/MS ion search result (with ion score 61) assigning probable lysine oxidation on ACINUS 1035-1047 sequence. Possible sites of oxidation are shown in blue.

Table 2.5 Peptides designed from probable lysine hydroxylation of ACINUS seen by LC-MS/MS and RS domain of ACINUS. These peptides were tested by MALDI-MS assays.

Peptide No.	Peptide Sequence	Mol. Wt.	Sequence from	MALDI Results
27	RKRRWGASTATTQKKPSISIT	2373.7	$ACINUS_{847-867}$	No
28	TQKKPSISITTESLKSLIPD	2186.5	$ACINUS_{858-877}$	No
29	GASTATTQKKPSISITTESL	2021.2	$ACINUS_{852-871}$	No

30	LLGRTGTLVEEAFWIDKIKS	2276.6	ACINUS$_{1031-1050}$	No
31	WIDKIKSHCFVTYSTVEEAR	2412.75	ACINUS$_{1044-1063}$	No
32	LVEEAFWIDKIKSHCFVTYS	2412.7	ACINUS$_{1038-1057}$	No
33	RSRSRSRDRRRKERAKSKEK	2572.9	ACINUS$_{1175-1194}$	Double hydroxylation
34	DRDKVREGPRSRSRSRDRRR	2540.7	ACINUS$_{1166-1185}$	No
35	RDRRDTKRHSRSRSRS	2056.2	ACINUS$_{1316-1331}$	No

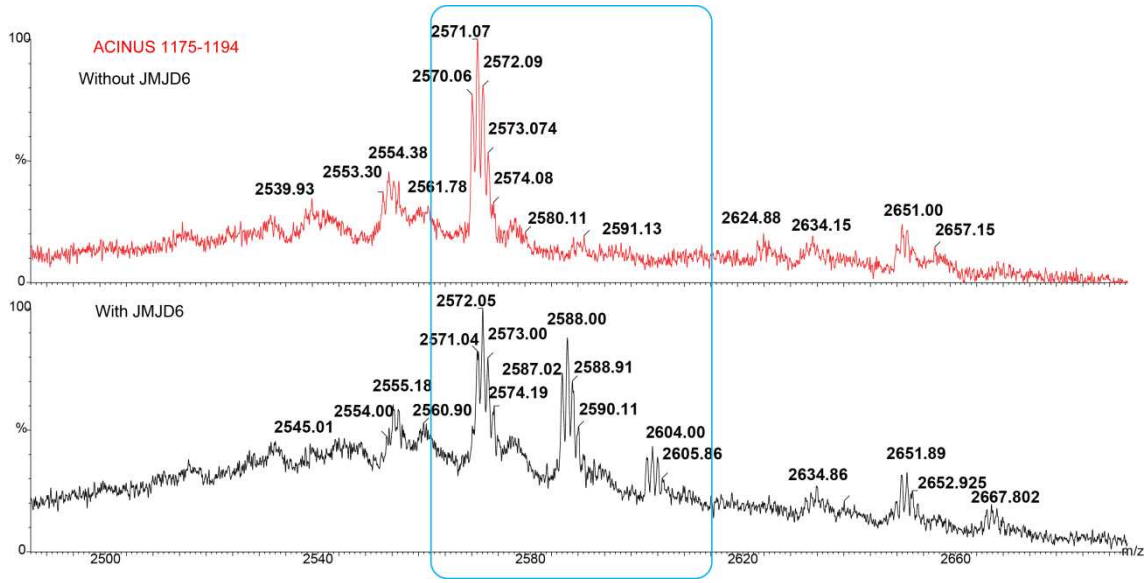


Figure 2.21 MALDI spectra of ACINUS$_{1175-1194}$ peptide when incubated with JMJD6 and cofactors. The MALDI spectra from control samples in which peptides were incubated with cofactors and buffer without addition of JMJD6 is also shown. Spectra show peaks at +16 and +32 indicating occurrence of double hydroxylation.

Interestingly, when peptides ACINUS$_{847-867}$, ACINUS$_{858-877}$, ACINUS$_{852-871}$ containing K861 were incubated with JMJD6, +16 Da peaks corresponding for lysine hydroxylation were not observed by MALDI-MS. Similarly, no hydroxylation was observed for peptides containing K1047, ACINUS$_{1031-1050}$, ACINUS$_{1044-1063}$,

ACINUS$_{1038-1057}$. However, peptide ACINUS$_{1175-1194}$ containing the RS domain showed +16 and +32 Da peak for lysine hydroxylation in MALDI-MS assays (Figure 2.21). This region corresponding to RS domain was not detected in the LC-MS/MS analysis, probably due to poor sequence coverage in this region.

2.10 Conclusions and future work

It can be concluded from the work carried out in this Chapter that JMJD6 likely does not possess histone demethylation activity as observed by MALDI-MS based *in vitro* assays. However, lysine hydroxylation was observed for the H4$_{1-15}$, H3$_{1-15}$ and H2a$_{1-15}$, H2b$_{7-17}$ and H2b$_{13-26}$ histone peptides. These results are consistent with the observations from the ALPHA screen assays which indicate relatively strong binding of JMJD6 with the biotinylated peptides from these regions. IP experiments coupled to LC-MS/MS studies carried out on JMJD6 expressed in human cells show the association of JMJD6 with SR proteins and probable lysine hydroxylations in the case of SRSF11, LUC7L2, U2AF65 and ZRANB2. Collectively these results support the proposal that JMJD6 plays a regulatory role in pre-mRNA splicing rather than being involved in chromatin regulation as previously described. The lysine hydroxylation activity seen for histone peptides maybe due to the fact that the abundant histone tails are very basic and the presence of arginines and serines in the histone peptides mimic the RS domain of SR proteins. However, there is no biological evidence that JMJD6 acts on histones in cells.

For the majority of the lysines which have been identified to be hydroxylated in SR proteins co-immunoprecipitated with JMJD6, the MALDI-MS based *in vitro* assays were carried out for synthesised peptides. But no hydroxylation was observed. This

observation indicates that rather than just amino acid sequence surrounding the lysine, additional structural factors are involved in lysine hydroxylation by JMJD6.

On the other hand, a number of peptides corresponding to RS domain with lysines, on incubation with JMJD6 were found to undergo hydroxylation. There is no particular consensus sequence that is readily indentified in case of RS rich regions of SR proteins, in part because these regions vary in length and posses repeats of arginine and serine residues.[23, 24] Thus, there is no overall consensus sequence linked to whether or not a particular protein or peptide is hydroxylated by JMJD6.

Under the current protocol conditions using various peptidases so far, attempts to identify further lysines from RS domains being hydroxylated in human cells have failed. Initially, in-gel peptidase activity was carried out using trypsin, a serine protease that selectively cleaves peptide bonds at the *C*-terminal side of lysine or arginine.[58] Probably due to high reactivity of trypsin, RS domain sequences were not observed by LC-MS/MS. To overcome this problem, the endoproteinase Glu-C which predominantly cleaves peptide bonds at the *C*-terminal side of glutamate, as well as at aspartate residues was used. However, Glu-C reacts at a much slower rate as compared to trypsin and it is not as effective as trypsin. It has also been reported that Glu-C generates mostly large peptides and a substantial decrease in number of detected peptides is observed for protein complexes.[59] Indeed the sequence coverage was poor when LC-MS/MS analyses were carried out on samples treated with Glu-C. An alternative to specific enzymes, low specificity proteases such as elastase has been applied to increase the sequence coverage of selected proteins. Elastase has been reported to generate mainly medium sized peptides in the mass range between 500 and 1500 Da.[60] The use of

elastase did support the association of JMJD6 with SR proteins but the overall sequence coverage was low. Previously *Dr. Celia Webby* investigated Lys-C proteinase for cleavage, but even those attempts failed to identify any lysine hydroxylation in SR domains (unpublished work from our group). On the other hand, because RS domains of SR proteins are highly phosphorylated on serine residues,[24] it is necessary to identify if phosphorylation influences the JMJD6 catalysed lysine hydroxylation. Thus, it is necessary to develop the cleavage protocol by use of combination of low specific peptidases to get better sequence coverage of intact RS domains.

On a final note, Hong et al. have reported that JMJD6 binds single stranded RNA *in vitro* using electrophoretic mobility shift assays.[61] It is now necessary to evaluate if RNA is truly a specific binding partner of JMJD6 *in vivo*. It will be interesting to investigate the influence of the presence of RNA to JMJD6 catalysed hydroxylation of SR proteins and hence its involvement in splicing.

2.11 References

1. Chang, B.; Chen, Y.; Zhao, Y.; Bruick, R. K., JMJD6 Is a Histone Arginine Demethylase. *Science* **2007,** 318, (5849), 444-447.

2. Klose, R. J.; Zhang, Y., Regulation of histone methylation by demethylimination and demethylation. *Nature Reviews Molecular Cell Biology* **2007,** 8, (4), 307-318.

3. Zhang, Y.; Reinberg, D., Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes & Development* **2001,** 15, (18), 2343-2360.

4. Chen, D.; Ma, H.; Hong, H.; Koh, S. S.; Huang, S.-M.; Schurter, B. T.; Aswad, D. W.; Stallcup, M. R., Regulation of Transcription by a Protein Methyltransferase. *Science* **1999,** 284, (5423), 2174-2177.

5. Yu, M. C.; Lamming, D. W.; Eskin, J. A.; Sinclair, D. A.; Silver, P. A., The role of protein arginine methylation in the formation of silent chromatin. *Genes & Development* **2006,** 20, (23), 3249-3254.

6. Wang, Y.; Wysocka, J.; Sayegh, J.; Lee, Y.-H.; Perlin, J. R.; Leonelli, L.; Sonbuchner, L. S.; McDonald, C. H.; Cook, R. G.; Dou, Y.; Roeder, R. G.; Clarke, S.; Stallcup, M. R.; Allis, C. D.; Coonrod, S. A., Human PAD4 Regulates Histone Arginine Methylation Levels via Demethylimination. *Science* **2004,** 306, (5694), 279-283.

7. Cuthbert, G. L.; Daujat, S.; Snowden, A. W.; Erdjument-Bromage, H.; Hagiwara, T.; Yamada, M.; Schneider, R.; Gregory, P. D.; Tempst, P.; Bannister, A. J.; Kouzarides, T., Histone Deimination Antagonizes Arginine Methylation. *Cell* **2004,** 118, (5), 545-553.

8. Cikala, M.; Alexandrova, O.; David, C.; Proschel, M.; Stiening, B.; Cramer, P.; Böttger, A., The phosphatidylserine receptor from Hydra is a nuclear protein with potential Fe(II) dependent oxygenase activity. *BMC Cell Biology* **2004,** 5, (1), 26-36.

9. Klose, R. J.; Kallin, E. M.; Zhang, Y., JmjC-domain-containing proteins and histone demethylation. *Nature Reviews Genetics* **2006,** 7, (9), 715-727.

10. Webby, C. J.; Wolf, A.; Gromak, N.; Dreger, M.; Kramer, H.; Kessler, B.; Nielsen, M. L.; Schmitz, C.; Butler, D. S.; Yates, J. R., III; Delahunty, C. M.; Hahn, P.; Lengeling, A.; Mann, M.; Proudfoot, N. J.; Schofield, C. J.; Bottger, A., Jmjd6 Catalyses Lysyl-Hydroxylation of U2AF65, a Protein Associated with RNA Splicing. *Science* **2009,** 325, (5936), 90-93.

11. Kawamura, A.; Tumber, A.; Rose, N. R.; King, O. N. F.; Daniel, M.; Oppermann, U.; Heightman, T. D.; Schofield, C., Development of homogeneous luminescence assays for histone demethylase catalysis and binding. *Analytical Biochemistry* **2010,** 404, (1), 86-93.

12. Ullman, E. F.; Kirakossian, H.; Singh, S.; Wu, Z. P.; Irvin, B. R.; Pease, J. S.; Switchenko, A. C.; Irvine, J. D.; Dafforn, A.; Skold, C. N., Luminescent oxygen

channeling immunoassay: measurement of particle binding kinetics by chemiluminescence. *Proceedings of the National Academy of Sciences* **1994,** 91, (12), 5426-5430.

13. El-Aneed, A.; Cohen, A.; Banoub, J., Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers. *Applied Spectroscopy Reviews* **2009,** 44, (3), 210-230.

14. Frecklington, D., General Method for MALDI-MS Analysis of Proteins and Peptides. *Cold Spring Harbor Protocols* **2007,** 2007, (3), pdb.prot4679.

15. Jaskolla, T. W.; Lehmann, W.-D.; Karas, M., 4-Chloro-α-cyanocinnamic acid is an advanced, rationally designed MALDI matrix. *Proceedings of the National Academy of Sciences* **2008,** 105, (34), 12200-12205.

16. Bissett, M. L.; Payne, F. E., Development of Antigens in Human Cells Infected with Simian Virus 40. *Journal of Bacteriology* **1966,** 91, (2), 743-749.

17. Gey, G. O.; Coffman, W. D.; Kubick, M. T., Tissue culture studies of the proliferative capacity of cervical carcinoma and normal epithelium. . *Cancer Research* **1952,** 12, 264-265.

18. Ehrhardt, C.; Schmolke, M.; Matzke, A.; Knoblauch, A.; Will, C.; Wixler, V.; Ludwig, S., Polyethylenimine, a cost-effective transfection reagent. *Signal Transduction* **2006,** 6, (3), 179-184.

19. Black, D. L., Mechanisms of alernative pre-messenger RNA splicing *Annual Review of Biochemistry* **2003,** 72, (1), 291-336.

20. Smith, C. W. J.; Valcárce, J., Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends in Biochemical Sciences* **2000,** 25, (8), 381-388.

21. Matlin, A. J.; Clark, F.; Smith, C. W. J., Understanding alternative splicing: towards a cellular code. *Nature Reviews Molecular Cell Biology* **2005,** 6, (5), 386-398.

22. Graveley, B. R.; Maniatis, T., Arginine/Serine-Rich Domains of SR Proteins Can Function as Activators of Pre-mRNA Splicing. *Molecular Cell* **1998,** 1, (5), 765-771.

23. Manley, J. L.; Tacke, R., SR proteins and splicing control. *Genes & Development* **1996,** 10, (13), 1569-1579.

24. Long, J. C.; Caceres, J. F., The SR protein family of splicing factors: master regulators of gene expression. *Biochemical Journal* **2009,** 417, (1), 15-27.

25. Robin, R., Mechanisms of fidelity in pre-mRNA splicing. *Current Opinion in Cell Biology* **2000,** 12, (3), 340-345.

26. Kramer, A., The Structure and Function of Proteins Involved in Mammalian Pre-mRNA Splicing. *Annual Review of Biochemistry* **1996,** 65, (1), 367-409.

27. Lee, C. G.; Zamore, P. D.; Green, M. R.; Hurwitz, J., RNA annealing activity is intrinsically associated with U2AF. *Journal of Biological Chemistry* **1993,** 268, (18), 13472-13478.

28. Gaur, R. K.; Valcárcel, J.; Green, M. R., Sequential recognition of the pre-mRNA branch point by U2AF65 and a novel spliceosome-associated 28-kDa protein. *RNA* **1995,** 1, (4), 407-417.

29. Valcárcel, J.; Gaur, R. K.; Singh, R.; Green, M. R., Interaction of U2AF65 RS Region with Pre-mRNA of Branch Point and Promotion Base Pairing with U2 snRNA. *Science* **1996,** 273, (5282), 1706-1709.

30. Kent, O. A.; Reayi, A.; Foong, L.; Chilibeck, K. A.; MacMillan, A. M., Structuring of the 3' Splice Site by U2AF65. *Journal of Biological Chemistry* **2003,** 278, (50), 50572-50577.

31. Graveley, B. R.; Hertel, K. J.; Maniatis, T., The role of U2AF35 and U2AF65 in enhancer-dependent splicing. *RNA* **2001,** 7, (6), 806-818.

32. Douglas L, B., Protein Diversity from Alternative Splicing: A Challenge for Bioinformatics and Post-Genome Biology. *Cell* **2000,** 103, (3), 367-370.

33. Singh, R., RNA-Protein Interactions That Regulate Pre-mRNA Splicing. *Gene Expression* **2002,** 10, (1&2), 79-92.

34. Rudner, D. Z.; Kanaar, R.; Breger, K. S.; Rio, D. C., Interaction between Subunits of Heterodimeric Splicing Factor U2AF Is Essential In Vivo. *Molecular and Cellular Biology* **1998,** 18, (4), 1765-1773.

35. Kielkopf, C. L.; Rodionova, N. A.; Green, M. R.; Burley, S. K., A Novel Peptide Recognition Mode Revealed by the X-Ray Structure of a Core U2AF35/U2AF65 Heterodimer. *Cell* **2001,** 106, (5), 595-605.

36. Merendino, L.; Guth, S.; Bilbao, D.; Martinez, C.; Valcarcel, J., Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3[prime] splice site AG. *Nature* **1999,** 402, (6763), 838-841.

37. Wu, S.; Romfo, C. M.; Nilsen, T. W.; Green, M. R., Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature* **1999,** 402, (6763), 832-835.

38. Zorio, D. A. R.; Blumenthal, T., Both subunits of U2AF recognize the 3' splice site in *Caenorhabditis elegans*. *Nature* **1999,** 402, (6763), 835-838.

39. Birney, E.; Kumar, S.; Krainer, A. R., Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Research* **1993,** 21, (25), 5803-5816.

40. Zhang, M.; Zamore, P. D.; Carmo-Fonseca, M.; Lamond, A. I.; Green, M. R., Cloning and intracellular localization of the U2 small nuclear ribonucleoprotein

auxiliary factor small subunit. *Proceedings of the National Academy of Sciences* **1992,** 89, (18), 8769-8773.

41. Kellenberger, E.; Stier, G.; Sattler, M., Induced folding of the U2AF35 RRM upon binding to U2AF65. *FEBS Letters* **2002,** 528, (1-3), 171-176.

42. Guth, S.; Martínez, C.; Valcárcel, J., Evidence for Substrate-Specific Requirement of the Splicing Factor U2AF35 and for Its Function after Polypyrimidine Tract Recognition by U2AF65. *Molecular and Cellular Biology* **1999,** 19, (12), 8263-8271.

43. Benjamin J, B., Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases. *Trends in Biochemical Sciences* **2000,** 25, (3), 106-110.

44. Graveley, B. R., Sorting out the complexity of SR protein functions. *RNA* **2000,** 6, (9), 1197-1211.

45. Trudgian, D. C.; Ridlova, G.; Fischer, R.; Mackeen, M. M.; Ternette, N.; Acuto, O.; Kessler, B. M.; Thomas, B., Comparative evaluation of label-free SINQ normalized spectral index quantitation in the central proteomics facilities pipeline. *Proteomics* **2011,** 11, (14), 2790-2797.

46. Chaudhary, N.; McMahon, C.; Blobel, G., Primary structure of a human arginine-rich nuclear protein that colocalizes with spliceosome components. *Proceedings of the National Academy of Sciences* **1991,** 88, (18), 8189-8193.

47. Page-McCaw, P. S.; Amonlirdviman, K.; Sharp, P. A., PUF60: a novel U2AF65-related splicing activity. *RNA* **1999,** 5, (12), 1548-1560.

48. Loughlin, F. E.; Mansfield, R. E.; Vaz, P. M.; McGrath, A. P.; Setiyaputra, S.; Gamsjaeger, R.; Chen, E. S.; Morris, B. J.; Guss, J. M.; Mackay, J. P., The zinc fingers of the SR-like protein ZRANB2 are single-stranded RNA-binding domains that recognize 5' splice site-like sequences. *Proceedings of the National Academy of Sciences* **2009,** 106, (14), 5581-5586.

49. Sahara, S.; Aoto, M.; Eguchi, Y.; Imamoto, N.; Yoneda, Y.; Tsujimoto, Y., Acinus is a caspase-3-activated protein required for apoptotic chromatin condensation. *Nature* **1999,** 401, (6749), 168-173.

50. Joselin, A. P.; Schulze-Osthoff, K.; Schwerk, C., Loss of Acinus Inhibits Oligonucleosomal DNA Fragmentation but Not Chromatin Condensation during Apoptosis. *Journal of Biological Chemistry* **2006,** 281, (18), 12475-12484.

51. Saraste, M.; Sibbald, P. R.; Wittinghofer, A., The P-loop: a common motif in ATP- and GTP-binding proteins. *Trends in Biochemical Sciences* **1990,** 15, (11), 430-434.

52. Samuels, M. E.; Schedl, P.; Cline, T. W., The complex set of late transcripts from the Drosophila sex determination gene sex-lethal encodes multiple related polypeptides. *Molecular and Cellular Biology* **1991,** 11, (7), 3584-3602.

53. Saitoh, N.; Spahr, C. S.; Patterson, S. D.; Bubulya, P.; Neuwald, A. F.; Spector, D. L., Proteomic Analysis of Interchromatin Granule Clusters. *Molecular Biology of the Cell* **2004,** 15, (8), 3876-3890.

54. Schwerk, C.; Prasad, J.; Degenhardt, K.; Erdjument-Bromage, H.; White, E.; Tempst, P.; Kidd, V. J.; Manley, J. L.; Lahti, J. M.; Reinberg, D., ASAP, a Novel Protein Complex Involved in RNA Processing and Apoptosis. *Molecular and Cellular Biology* **2003,** 23, (8), 2981-2990.

55. Zhou, Z.; Licklider, L. J.; Gygi, S. P.; Reed, R., Comprehensive proteomic analysis of the human spliceosome. *Nature* **2002,** 419, (6903), 182-185.

56. Rappsilber, J.; Ryder, U.; Lamond, A. I.; Mann, M., Large-Scale Proteomic Analysis of the Human Spliceosome. *Genome Research* **2002,** 12, (8), 1231-1245.

57. Tange, T. Ø.; Shibuya, T.; Jurica, M. S.; Moore, M. J., Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core. *RNA* **2005,** 11, (12), 1869-1883.

58. Olsen, J. V.; Ong, S.-E.; Mann, M., Trypsin Cleaves Exclusively C-terminal to Arginine and Lysine Residues. *Molecular and Cellular Proteomics* **2004,** 3, (6), 608-614.

59. Molina, H.; Horn, D. M.; Tang, N.; Mathivanan, S.; Pandey, A., Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry. *Proceedings of the National Academy of Sciences* **2007,** 104, (7), 2199-2204.

60. Schlosser, A.; Pipkorn, R.; Bossemeyer, D.; Lehmann, W. D., Analysis of Protein Phosphorylation by a Combination of Elastase Digestion and Neutral Loss Tandem Mass Spectrometry. *Analytical Chemistry* **2000,** 73, (2), 170-176.

61. Hong, X.; Zang, J.; White, J.; Wang, C.; Pan, C.-H.; Zhao, R.; Murphy, R. C.; Dai, S.; Henson, P.; Kappler, J. W.; Hagman, J.; Zhang, G., Interaction of JMJD6 with single-stranded RNA. *Proceedings of the National Academy of Sciences of the United States of America* **2010,** 107, (33), 14568-14572.

Chapter 3. Crystal structure of JMJD6

3.1 Introduction

As described in the preceding chapter, although JMJD6 accepts isolated histone peptide fragments as lysyl-hydroxylation substrates, the available evidence suggests that in cells JMJD6 catalyses the hydroxylation of arginine-serine rich regions of SR proteins. The finding that JMJD6 is an Fe(II) and 2OG dependent oxygenase catalysing lysyl hydroxylation of SR proteins involved in mRNA splicing, raises the question of how its sequence and three dimensional structure are related to its function and how its structure is related to other JmjC family members including the asparginyl hydroxylase, FIH and the histone N^{ϵ}-lysine demethylases.[1-6] The collagen lysyl hydroxylases were the first reported lysyl hydroxylase of 2OG oxygenase family to be identified. Variations in the genes encoding for some lysyl hydroxylases are associated with genetic diseases including Ehlers-Danlos syndrome, type VI.[7] However, to date there has been no structural information reported for the collagen lysyl hydroxylase or indeed any other oxygenase that catalyses the lysyl hydroxylation.

The aim of the work described in this chapter was to attempt crystallisation of recombinant human JMJD6 in order to understand its interaction with SR proteins and the mechanism of lysyl hydroxylation. On the basis of the X-ray diffraction results, site directed mutagenesis studies were carried out to study substrate binding by JMJD6.

3.2 Preparation of protein

3.2.1 Choice of truncated construct

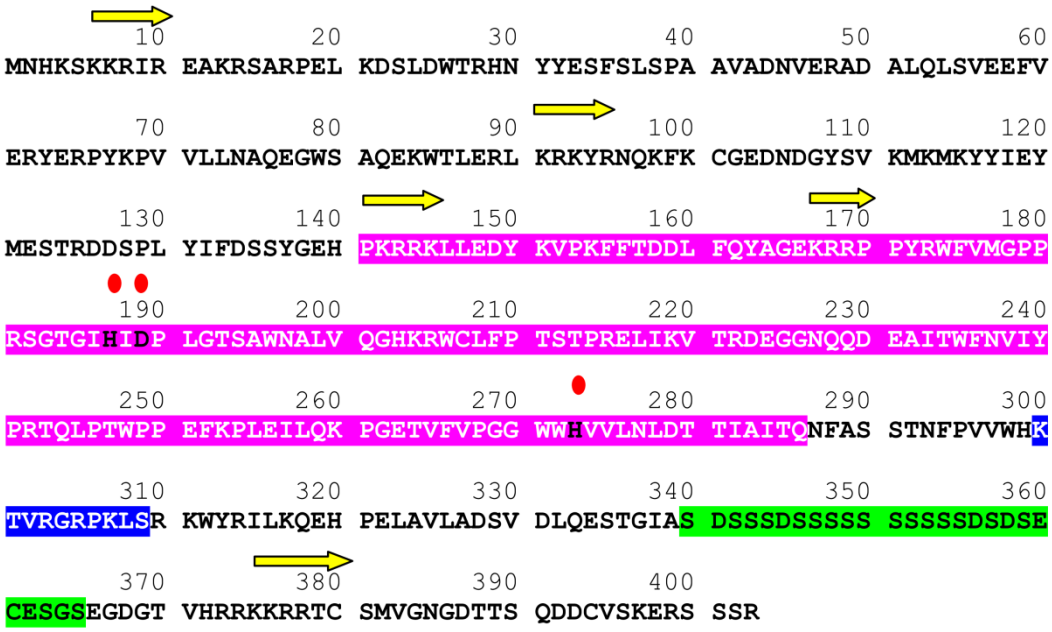


Figure 3.1 Domain architecture of JMJD6 showing JmjC domain (pink), poly-serine stretch (green), AT hook region (blue). The HXD...H facial triad residues are in red and the predicted NLS regions are shown by yellow arrows.

Crystallisation of JMJD6 was initially attempted with His-tagged full length sequence ($JMJD6_{1\text{-}403}$). However, these earlier experiments failed to produce any crystals. It has been found that many proteins contain intrinsically disordered regions which do not allow formation of well-defined three-dimensional structures in their native states.[8] The deletion of such regions is often necessary to develop a construct that can be crystallised in order to enable structural analysis. The *C*-terminal residues 335-420 of full length JMJD6 contains a polyserine rich region (residues 340-362) that is predicted to be disordered by DISPROT[9] and it was thought that this disorder might explain the inability to obtain crystals with the His-tagged full length $JMJD6_{(1\text{-}403)}$ construct (MW

47.5 kDa). A construct comprising of first 343 JMJD6 residues, but without polyserine region and including the catalytic JmjC domain (Figure 3.1) was then used for attempted crystallisation.

3.2.2 Preparation of Seleno-methionine substituted JMJD6

Sequence analysis of JMJD6 showed the presence of four non terminal methionine residues. The replacement of methionine by incorporating seleno-methhionine, a heavy atom substitute is a standard method used for determining the phase of X-ray diffraction data. The gene encoding for full length human JMJD6 was used to produce expression constructs of $JMJD6_{1-343}$ using the pET-28a(+) vector (*This construct was prepared by Dr. Danica Butler.*), which was expressed in *Escherichia coli* Rosetta cells. Seleno-methionine (SeMet) derivatised $JMJD6_{1-343}$ was produced in 4 L of LeMaster media[10, 11] supplemented with SeMet (50 mg ml^{-1}), kanamycin (30 µg ml^{-1}) and chloramphenicol (34 µg ml^{-1}), using 10 ml of starter culture grown overnight. Cells were grown at 37°C (while shaking at 200 rpm) until an OD_{600} of 0.6 was reached; the temperature was then reduced to 25°C and isopropyl β-*D*-1-thiogalactopyranoside (IPTG) (0.5 mM) was added to induce $JMJD6_{1-343}$ production. Cells were harvested after 16-18 hours by centrifugation and stored at -80°C. Cells were resuspended in lysis buffer and JMJD6 was purified as described in Chapter 8. Purity of protein was assessed by SDS-PAGE gel (Figure 3.2) which does show band at MW about ~43 kDa.

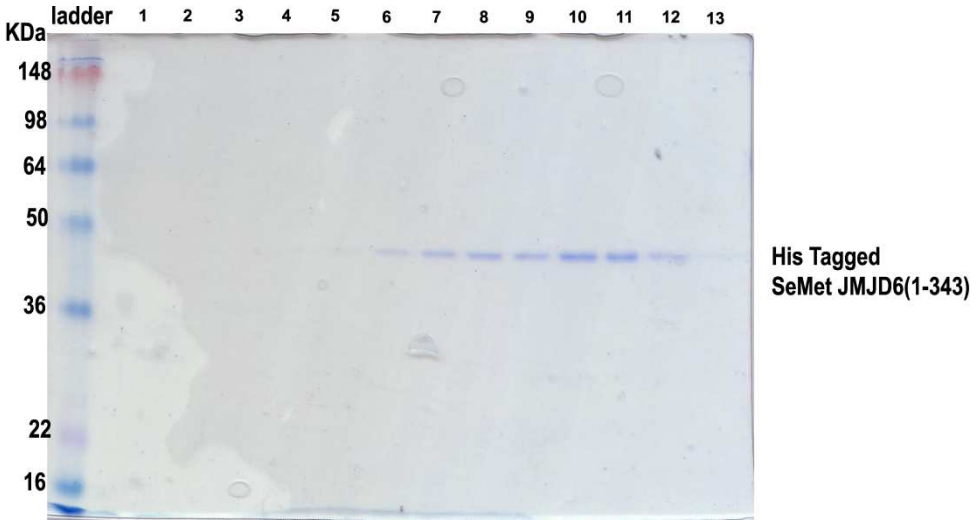


Figure 3.2 SDS-PAGE gel showing purity of SeMet His tagged JMJD6$_{1-343}$ after purification.

3.3 Crystallisation of JMJD6

3.3.1 High throughput crystallisation screen

A high throughput crystallisation screen was then carried out using Phoenix™ RE crystallisation robot and Minstrel-HT™ platform with CrystalTrek™ (Rigaku Inc). 20 mg ml^{-1} of SeMet JMJD6$_{1-343}$ was used for crystallisation using the sitting drop vapour diffusion method. In this method, either 0.1 µl or 0.2 µl of protein or protein-ligand solution was mixed with a reservoir solution containing the choice of precipitants in 1:1, 1:2 and 2:1 ratio such that the final drop volume was 0.3 µl in Greiner 96 well CrystalQuick™ SW plates. Because the protein/precipitant mixture in the drop is less concentrated than the reservoir solution, the drop is brought very gradually to super-saturation by loss of water from the droplet into the reservoir, allowing protein crystallisation to occur. The attempted crystallisations were carried out at 25°C and 4°C. Crystallisation screening was also carried out in this fashion of JMJD6 with the

commonly used 2OG analogue inhibitors NOG, 2,4 PDCA and 2,6 PDCA, with and without $LUC7L2_{267-278}$ peptide.

Rhomboidal JMJD6 crystals were obtained in 0.3 µl drops (0.2 µl protein (20 mg ml^{-1}) and 0.1 µl reservoir solution) with 80 µl 0.8 M sodium dihydrogen phosphate, 0.8 M potassium dihydrogen phosphate, 0.1 M sodium HEPES pH 7.5 (Molecular Dimension Structure Screen I & II HT-96, condition C4) and in the reservoir in the wells with 2:1 and 1:1 protein to screen solution ratios. Crystals of approximately 70 x 50 µm size formed within 18 hours at 20°C and then slowly grew over a period of four days at 4°C. Crystals were also obtained at these same conditions when 2,6-PDCA was added as a ligand. After nine days, crystals were transferred directly to well solution supplemented with 20% ($^{v}/_{v}$) glycerol (cryo-protectant) and cryo-cooled in liquid nitrogen for storage and data collection.

Data was collected on these crystals as described in Section 3.4. Some of the crystals were used for soaking experiments in which peptide in 50 mM Tris.HCl buffer, pH 7.5 was added in a drop containing crystals such that final concentration of peptide was 10 mM. Even if the crystallisation solution contained ligands and/or peptide, JMJD6 was found to be crystallised in the apo form (See below).

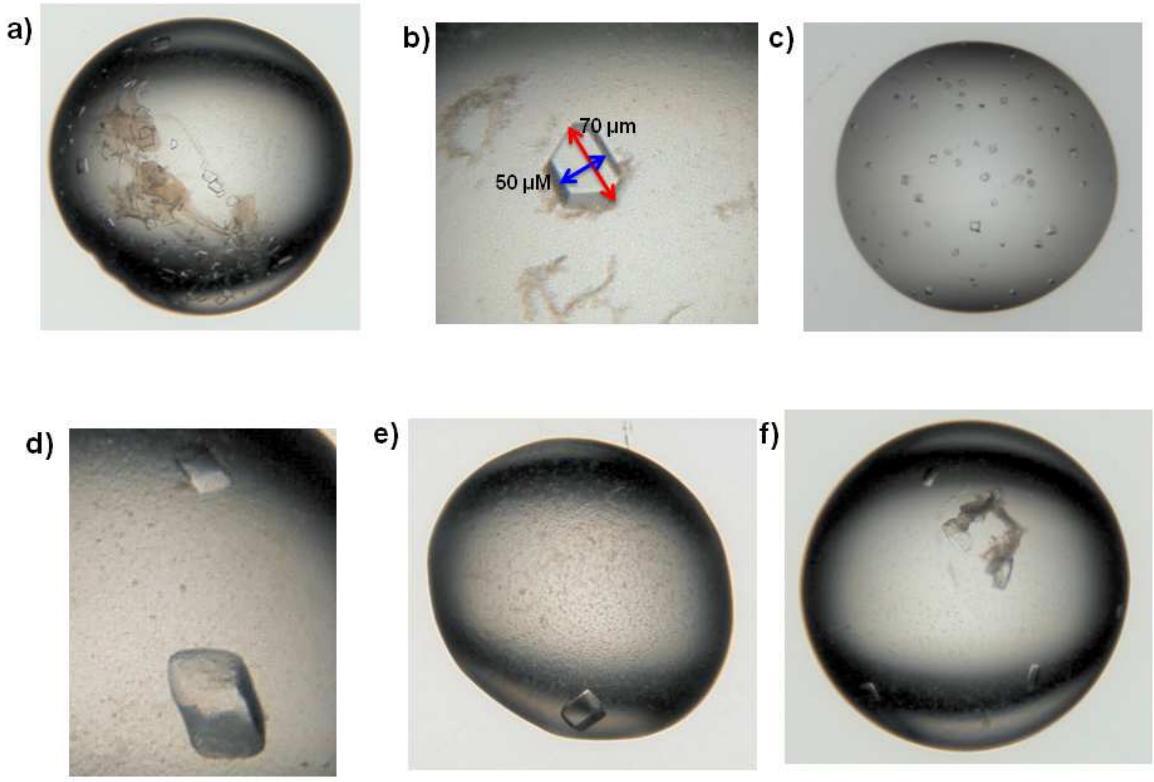


Figure 3.3 Various crystal morphologies of JMJD6. a) and b) shows the morphology of crystals which gave good quality diffraction data. d), e) and f) Types of crystals were obtained during optimisation screen in the presence of an inhibitor and peptide but did not diffract well.

3.3.2 Optimisation of crystallisation conditions

In order to attempt the preparation of crystals of JMJD6 complexed with a SR protein substrate fragment, optimisation of the crystallisation conditions was carried out using the sitting drop vapour diffusion technique in 24-well plates. Screening was carried out by altering concentration of sodium dihydrogen phosphate and potassium dihydrogen phosphate in HEPES buffer pH 7.5. Ligands including NOG, 2,4-PDCA and the methyl ester of 2,4-PDCA (Figure 3.4) were mixed with JMJD6 and LUC7L2$_{267\text{-}278}$ peptide in crystallisation plates such that the final concentration of ligand and peptide was 4 mM and 10 mM, respectively. Crystals with different morphologies were obtained in these

plates. Preliminary diffraction analyses were carried out under in-house conditions by *Dr. Michael McDonough*. However, the obtained crystals did not diffract well.

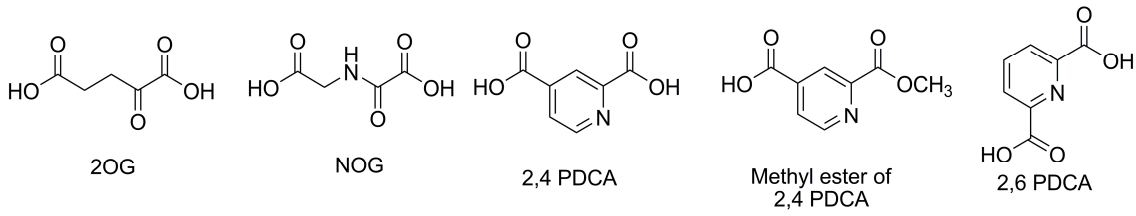


Figure 3.4 Structures of 2OG and 2OG analogue inhibitors which were used for crystallisation.

3.3.3 Crystallisation under hypoxic conditions.

In further attempts to obtain crystals of JMD6 along with its peptide substrate, crystallisation was also carried out using other methods. Because JMJD6 is an Fe(II) and 2OG dependent oxygenase, crystallisation was carried out under hypoxic conditions in an anaerobic glove box. Crystals in the presence of 2OG were obtained by adding 1 mM $FeSO_4$ and 1 mM 2OG to the crystallisation solution. However, the crystals obtained were of low quality and gave relatively poor diffraction.

3.4 Crystal structure analysis

Data collection and processing were carried out by Dr. Michael McDonough. Single wavelength anomalous dispersion (SAD) data using 0.9794Å wavelength X-rays were collected at DIAMOND beamline IO4 on a single $JMJD6_{1\text{-}343}$ SeMet crystal grown at 20°C in the presence of 0.8 M sodium dihydrogen phosphate, 0.8 M potassium dihydrogen phosphate and 0.1 M sodium Hepes, pH 7.5 Data for $JMJD6_{1\text{-}343}$ (SeMet) were integrated and scaled with the HKL2000 package.[12] High resolution data were also collected on a single crystal grown at 4°C and having slightly different unit cell

parameters. The high-resolution data were processed with XDS[13] and subsequently scaled using SCALA.[14]

Structure solution and refinement was carried out by Dr. Kathryne Kavanagh and Dr. Tobias Krojer at Structural Genomic Consortium, Oxford. Unit cell parameters were used to calculate the Matthews coefficient 2.55 $Å^3$/Dalton, which indicates two molecules per asymmetric unit. The direct methods program SHELXD[15] was used to locate nine out of ten possible selenium sites using the $JMJD6_{1-343}$ SeMet dataset. The nine selenium sites were refined and initial phases calculated using SHARP (Global Phasing Ltd). The initial electron density map after solvent flattening with SOLOMON[16] was of excellent quality and automated model building with ARP/wARP[17] resulted in a model where the central parts (60% of residues in the crystallised protein's sequence) of both molecules in the asymmetric unit were built. One subunit from the initial SeMet structure solution was used to locate both molecules in the asymmetric unit of the high resolution data by molecular replacement using PHASER.[18] Further refinement was performed against the $JMJD6_{1-343}$ data set using PHENIX.[19] After several rounds of manual rebuilding in COOT[20] and subsequent cycles of refinement, the model converged to a final R_{factor} and R_{free} of 19.2% and 22.5%, respectively. In this way the crystal structure of $JMJD6_{1-343}$ was obtained at 1.75 Å resolution.

3.4.1 General structural features of JMJD6

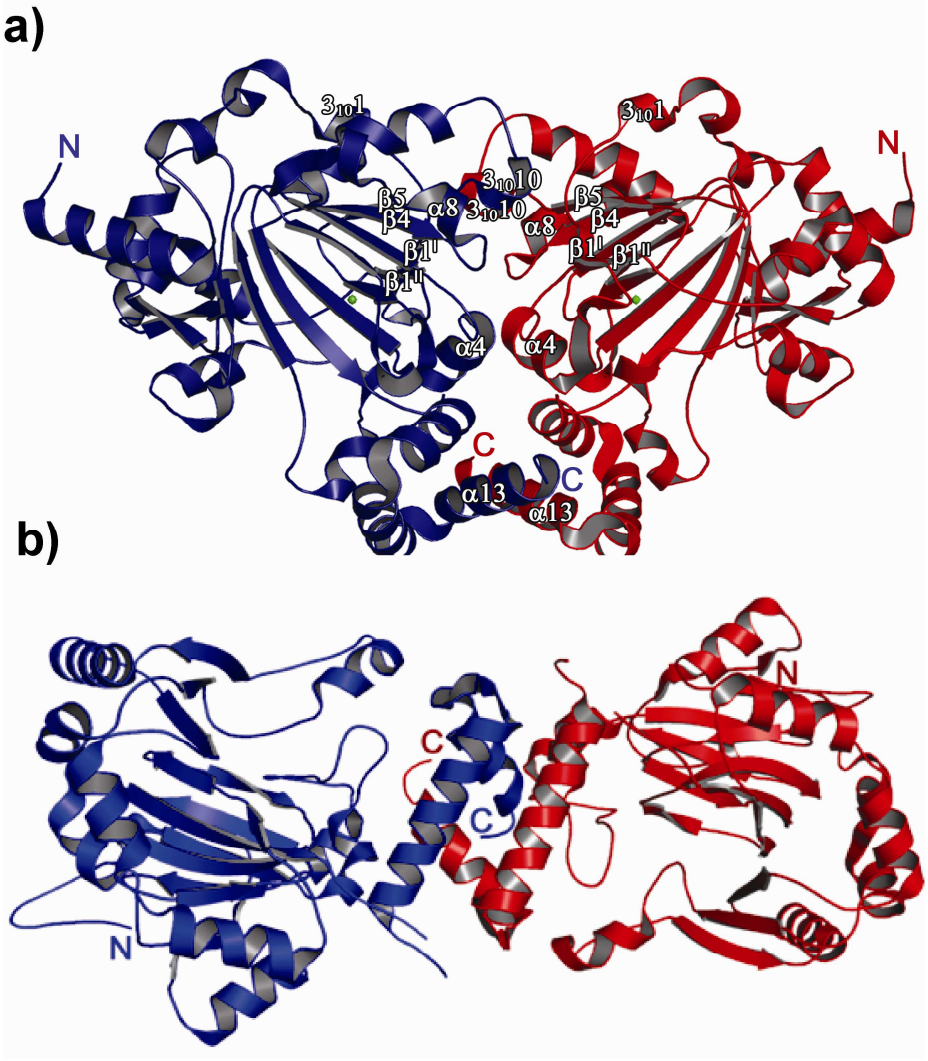


Figure 3.5 The overall fold of JMJD6 (PBD ID: 3K2O) and comparison with FIH dimer (PDB ID: 1H2K). (a) Ribbons representation of the JMJD6 dimer. Secondary structure elements involved at the dimer interface are labelled; (b) Ribbons representation of the FIH dimer.

JMJD6 crystallises as a homodimer, similar to FIH. The dimer interface in the crystalline state, was calculated to have a total buried surface area of 2050 $Å^2$ and a Δ^iG equal to -16.6 kcal mol^{-1} indicating a stable complex.[21] The interface is comprised of interactions between α13, α4, the loop between $3_{10}1/\beta1^I$, and the loop between $\alpha4/\beta1^{II}$ of one molecule and α13, the loop between β4/α8, $\beta5/3_{10}10$, and the loop between

$\alpha 4/\beta 1^{II}$ of the other molecule, respectively (Figure 3.5a). Comparison of the apparent JMJD6 dimer with that of the FIH dimer is shown in Figure 3.5. Note that though both enzymes dimerise using interactions between their *C*-terminal helices, FIH only interlocks via its *C*-terminal helices from each subunit while JMJD6 forms a pseudo-4-helix bundle (the helices are not contiguous) with its *C*-terminal helices that pack side by side between subunits without interlocking. The pseudo-4-helix bundle of JMJD6 is reminiscent of the fold at the *C*-terminal helices observed in the structure of histone demethylases, JMJD2A and PHF8 as shown in Figure 3.6. However, for all of these enzymes the *C*-terminal regions representing the full length protein is not present in the crystal structures are in fact extended by multiple domains. In solution, full length JMJD6[22, 23] and $JMJD6_{1\text{-}343}$ both undergo oligomerisation (discussed later in Chapter 6). However, for $JMJD6_{1\text{-}343}$ crystallographic construct, *C*-terminus is missing and the dimer formed is most probably a crystallographic feature rather than one with any biological relevance.

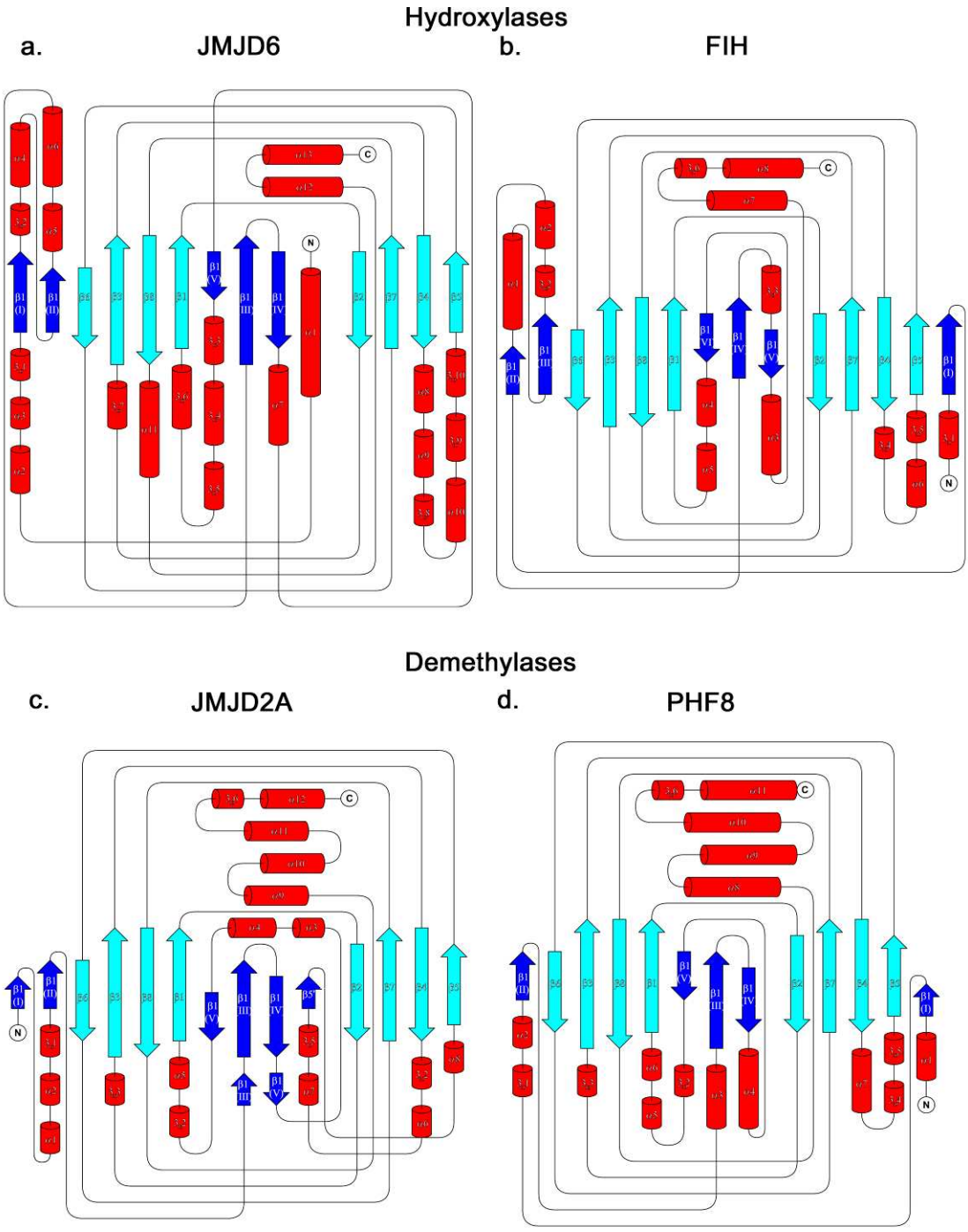


Figure 3.6. Topology diagrams of human 2OG oxygenases (hydroxylases and demethylases). (a) JMJD6 (b) FIH (c) JMJD2A (d) PHF8. Topologies were calculated using the HERA[24], subroutine of PROMOTIF[25], manually modified to show similarity between family members and drawn using TOPDRAW[26] within the CCP4 package.[14] ***This image was prepared by Dr. Michael McDonough.***

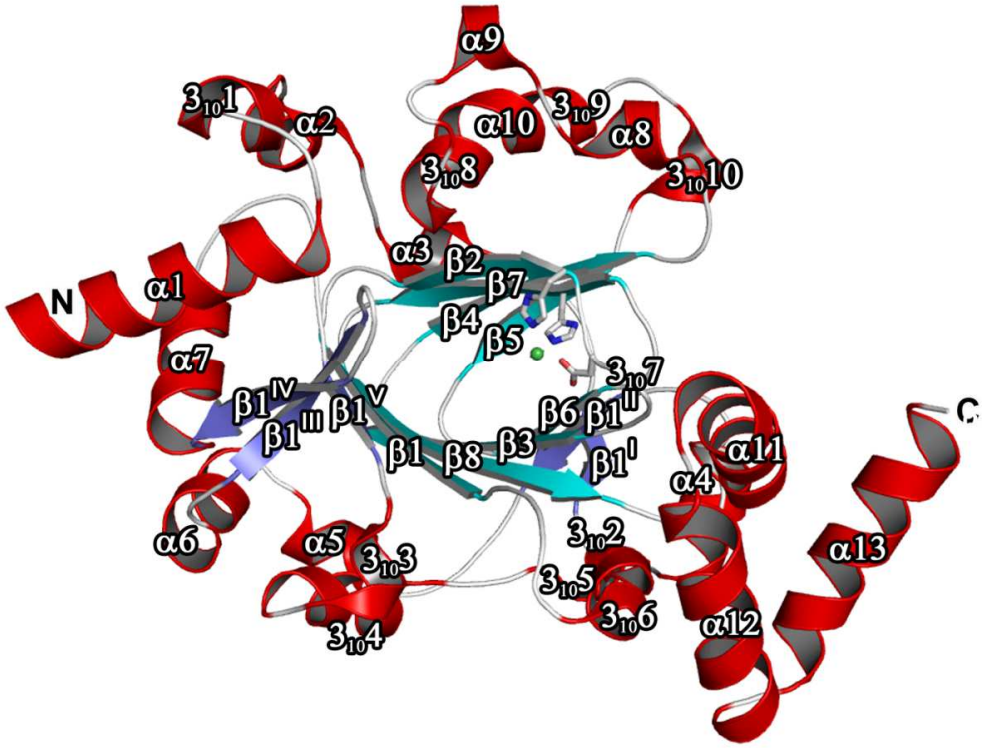


Figure 3.7 Stereoview of the core DSBH fold (cyan) surrounded by stabilising helices (red) of JMJD6 is typical of the 2OG dependent non-heme Fe(II) dependent oxygenases. β-strands extending the major β-sheet are in purple. Metal ligating residues are shown as white sticks and the metal as a green sphere. The metal at the active site is Ni(II).

The JMJD6 structure possesses the DSBH fold (sometimes referred as 'jelly roll' fold or cupin (barrel)) which is characteristic of all known 2OG oxygenases structures including the JmjC subfamily. The highly conserved DSBH fold is comprised of eight β-strands, four of which (β1, β3, β6 and β8) form a part of the major β-sheet with the other four DSBH β-strands (β2, β4, β5, β7) making up the minor β-sheet. The major β-sheet is extended at both ends of the DSBH by five additional β-strands ($\beta1^{I}$, $\beta1^{II}$, $\beta1^{III}$, $\beta1^{IV}$, $\beta1^{V}$) derived from the *N*-terminus. In the case of JMJD6, a total of thirteen β-strands form the central β-helix structure which is surrounded by thirteen α-helices and ten 3_{10}-helices is shown in Figure 3.7.

3.4.2 Details of active site

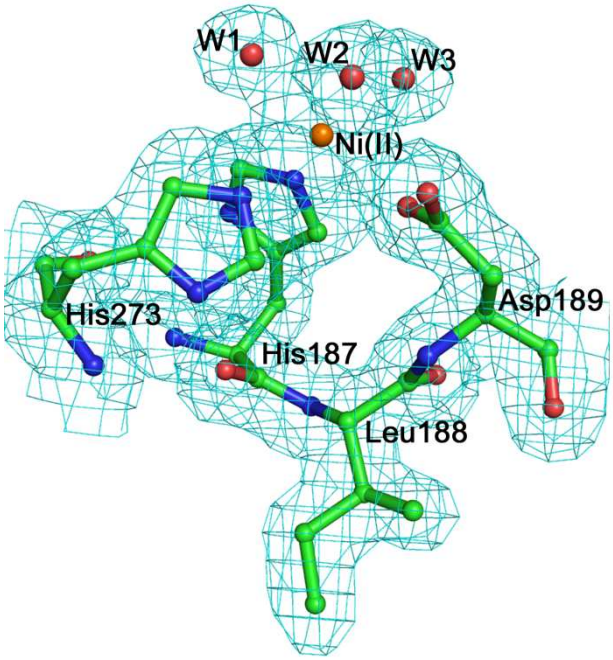


Figure 3.8 Representative electron density ($2F_o$-F_c electron density map (cyan mesh) contoured at 2σ) of the 1.75Å resolution structure. Stereoview showing the octahedral coordination of the metal (orange sphere) by three JMJD6 derived ligands (green sticks) and three water molecules (red spheres).

Similar to other 2OG oxygenases, the iron binding site of JMJD6 is located at the "open" end of the DSBH, between strands β1 and β2. In the case of this crystal structure, the active site metal was modelled as Ni(II) because of the high concentration of Ni(II) present during purification. Ni(II) is bound by a facial triad of residues His187, Asp189, and His273 that form a conserved HXD/E…H motif (Figure 3.8) located within minor β-sheet. His187, Asp189, and His273 together with three water molecules octahedrally coordinate the Ni(II) ion (substituting for Fe(II)).[27] The position of these water molecules likely reflects binding sites for 2OG and oxygen. The conformations of the two iron-ligating histidine side-chains are nearly identical to those observed for FIH[28] and PHF8 (Figure 3.10).[5, 6] The conformation of the Asp189 side-

chain carboxylate is slightly different; it is possible that this is due to the absence of primary substrate and/or the presence of Ni(II) ions.

Although *N*-oxalylglycine (NOG), a 2OG analogue and inhibitor of JMJD6, was present in the crystallisation buffer, it was not observed in the predicted 2OG binding site. It is possible that the acetate present in the crystallisation buffer (used to titrate buffer pH in Molecular Dimension crystallisation screens) competes with NOG for binding. Attempts to crystallise JMJD6 in the absence of acetate or to soak out and replace the acetate with NOG in pre-formed crystals were unsuccessful. Hence, the electron density near Lys204, Tyr131, Thr184 and Asn197 is modelled by an acetate ion as shown in Figure 3.9.

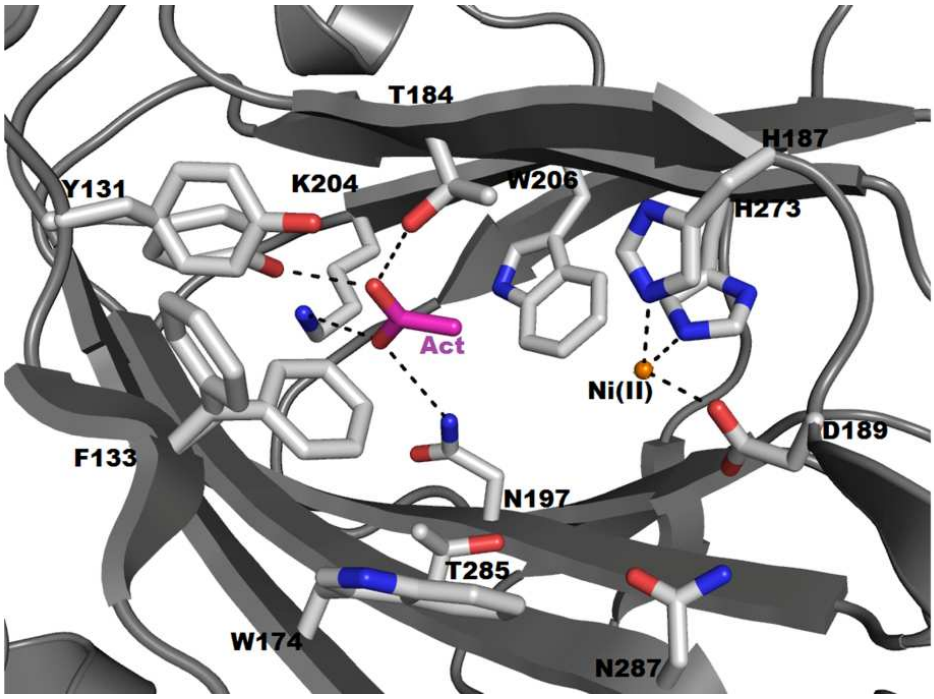


Figure 3.9 The 2-oxoglutarate binding site of JMJD6. (a) Ribbons representation of the JMJD6 active site with 2OG and metal binding residues shown as white sticks. The acetate ion (magenta sticks) is predicted to bind similarly to the 2OG C-5 carboxylate.

3.4.3 Comparison of active site residues with other 2OG oxygenases (FIH, PHF8 and JMJD2A)

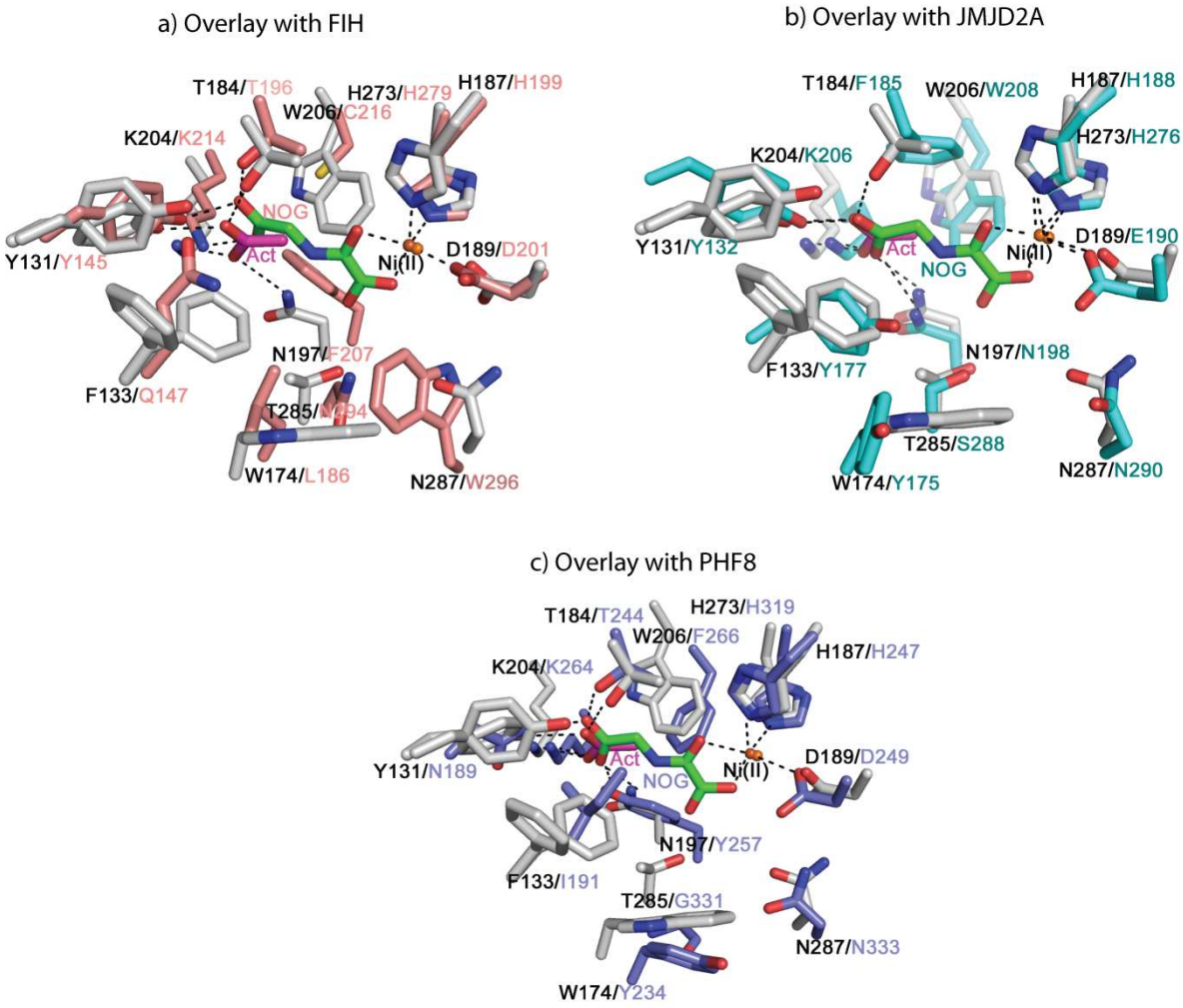


Figure 3.10 The predicted JMJD6 2OG and metal binding site superimposed on that of (a) FIH (salmon sticks)[28] (b) JMJD2A (cyan sticks)[1] and (c) PHF8 (purple sticks)[1]. It is apparent that the 2OG C-5 carboxlyate (green sticks) in the structures of other 2OG oxygenases is bound in a similar manner to the acetate ion (magenta sticks) observed in the JMJD6 structure (grey ribbons and white sticks).

The proposed 2OG oxygenase consensus mechanism, along with crystallographic and spectroscopic analysis of 2OG oxygenases indicates that 2OG binds to the active site Fe(II), with displacement of two water molecules to give an octahedral coordination complex in which the 2OG 2-oxo carbonyl and its C-1 carboxylate complex with Fe(II)

in bidentate manner.[29, 30] With previously reported structures, the side chain of 2OG is bound in a pocket in which C-5 carboxylate typically forms interactions with a basic lysine or arginine residue and an alcohol side chain. In case of JMJD6, Lys204, located on β-strand β4, extends into the deepest part of the 2OG binding pocket and by analogy with structural studies on other 2OG oxygenases is predicted to interact with the 2OG C-5 carboxylate via electrostatic and hydrogen bonding interactions as observed in structures of FIH, JMJD2A and PHF8.[2, 5, 6, 28, 31] Based on the observed acetate binding mode, the 2OG C-5 carboxylate oxygen atoms are predicted to form hydrogen bonds with Lys204 N^{ζ}(2.9 Å), Tyr131 OH (2.8 Å), Thr184 $O^{\gamma}1$ (3.2 Å), and Asn197 $N^{\delta}2$ (3.4 Å), as observed for similarly positioned residues in the structures of other 2OG oxygenases; e.g. $Tyr132_{JMJD2A}$, $Thr244_{PHF8}$, $Thr196_{FIH}$ and $Thr172_{CAS}$ (CAS, clavaminic acid synthase) (Figure 3.10).[2, 5, 28, 31] However, JMJD6 is unusual in that it has four polar side-chains positioned to potentially interact with the 2OG C-5 carboxylate where typically two or three polar side-chains are involved in binding of 2OG in reported 2OG oxygenase structures, thus its binding with 2OG may also be unusual.[29]

3.5 Docking of LUC7L2 peptide substrate to JMJD6 crystal structure

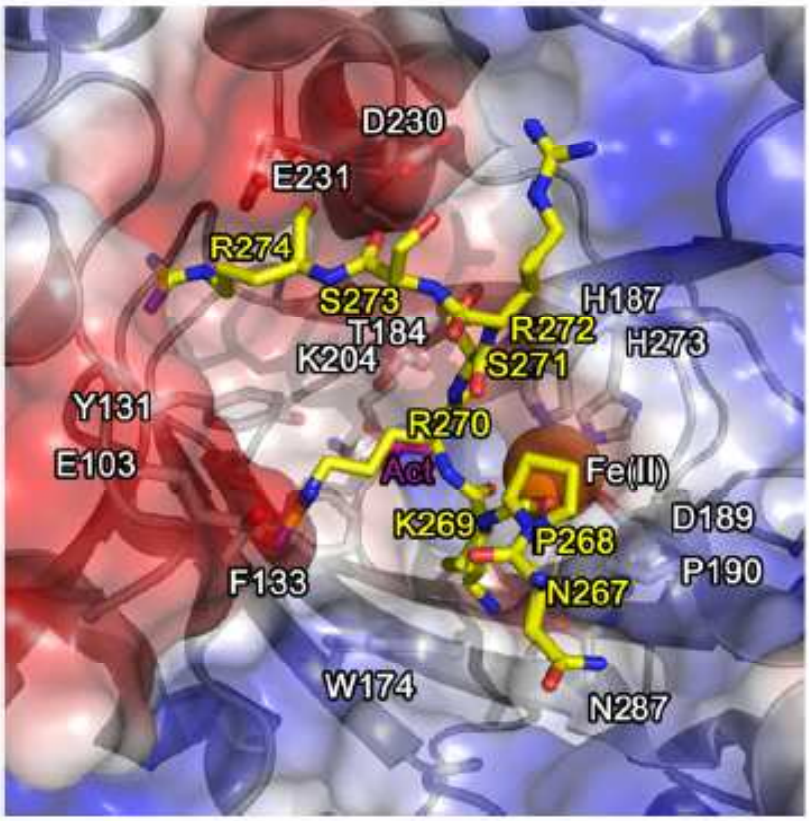


Figure 3.11 A peptide fragment of LUC7L2 peptide (yellow sticks and labels) docked on to the JMJD6 structure (white sticks and labels with transparent surface displayed). JMJD6 residues predicted to be involved in substrate recognition are shown as sticks. For the docking calculations, the Ni(II) observed in the crystal structure was substituted with Fe(II) and 2OG was modelled based on superposition with FIH, JMJD2A, and PHF8. The acetate ion shown is that observed in the JMJD6 crystal structure.

Since attempts to obtain a crystal structure of JMJD6 in complex with a substrate $LUC7L2_{267\text{-}278}$ were unsuccessful, modelling studies by docking LUC7L2 peptide were carried out to investigate the potential interaction of JMJD6 with the substrate. *Docking studies were carried out using AutoDock Vina*[32] *by Dr. Michael McDonough*. An 8 residue peptide $LUC7L2_{267\text{-}274}$ was used as the ligand with 32 rotatable bonds and a grid covering the entire surface of the JMJD6 dimer structure (with Fe(II) and 2OG

modelled, based on FIH, JMJD2A and PHF8 substrate structures in place of Ni(II) and acetate) was used as the receptor.

Results from 'blind' docking[33] of $LUC7L2_{267-274}$ peptide support a binding mode in which the target lysine is within ~4.5 Å of the catalytic metal in the active site. Four out of nine of the lowest energy binding modes were in the predicted binding site with all having the substrate lysine C-5 atom positioned within 5 Å of the metal suggesting these positions are suitable for hydroxylation (Figure 3.11). It appears that JMJD6 likely binds the substrate lysine residue in a manner more closely related to substrate binding by the histone demethylases JMJD2A and PHF8 rather than the asparaginyl hydroxylase FIH. Furthermore, as shown in Figure 3.11 it can be seen that there is a cluster of acidic residues at the "mouth" of the active-site which may be involved in interaction with RS rich domains of proteins involved in mRNA splicing i.e. substrates of JMJD6. These residues include Glu19, Glu103, along with Asp230 and Glu231 from the β4-β5 insert. Interaction of Glu231 with substrate was further studied by site directed mutagenesis.

3.6 Mutagenesis studies of JMJD6

As mentioned in Section 3.4, residues Lys204, Tyr131, Thr184, and Asn197 of JMJD6 are likely to be involved in 2OG binding and residue Glu231 is proposed to be involved in substrate binding. The effects of mutating these residues were analysed by comparing the activity of wild-type JMJD6 with the single point mutants by MALDI-MS and 2OG turnover assays. Single point substitutions for Lys204Ala, Tyr131Phe, Thr184Ala, Asn197Ala, Glu231Ala were generated from wild-type JMJD6 construct using the Quickchange Stratagene Site directed mutagenesis protocol (Stratagene) as described in Chapter 8. The constructs were transformed into Rosetta2DE3 (Invitrogen) cells,

plasmids were sequenced to confirm substitutions and all JMJD6 mutant constructs were expressed and purified using the nickel affinity chromatography similar to JMJD6 wild-type. All the resultant JMJD6 variants were purified to about 95% purity as judged by SDS-PAGE analyses.

3.6.1 Comparing mutants by 2OG turnover assays

Catalytic activity of 2OG oxygenases can be monitored by using radioactively labelled 1-[^{14}C]-2OG as 2OG is the cosubstrate used by these enzymes. During the catalytic cycle, 2OG is converted into succinate, thereby releasing carbon dioxide. While using labelled 1-[^{14}C]-2OG for carrying out enzyme catalysis reactions, radiolabelled $^{14}CO_2$ is released which can be trapped by hyamine hydroxide. Hyamine hydroxide forms the hydrogen carbonate adduct with carbon dioxide as shown in Figure 3.12, this adduct gives rise to beta emission which can be measured by liquid scintillation counter.

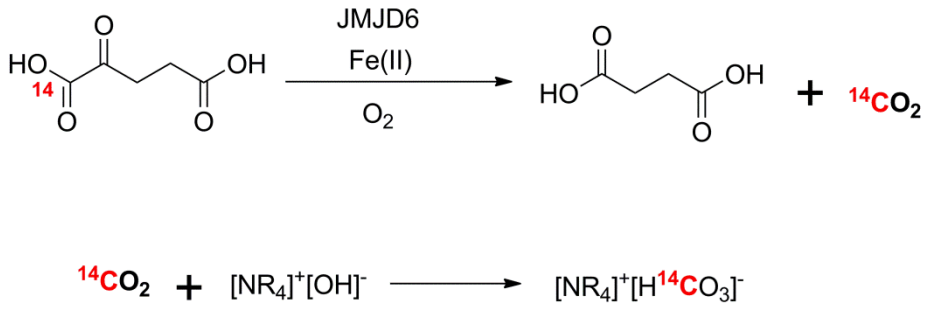


Figure 3.12 Conversion of radio-labelled 1-[^{14}C]-2OG into succinate in the presence of JMJD6 and Fe(II), $^{14}CO_2$ is released in the process which forms a radiolabelled adduct with hyamine hydroxide.

For comparing mutants, enzyme assays measuring the release of $^{14}CO_2$ from 1-[^{14}C] 2OG were employed.[34] Because JMJD6 displays 2OG turnover in the absence of substrate (results discussed in Chapter 5), JMJD6 (10 µM) was incubated with and without 50 µM LUC7L2$_{267\text{-}278}$ substrate for comparison in the presence of 400 µM

Fe(II), 160 µM 2OG, 1 mM ascorbate, 1mM dithiothreitol (DTT) and 0.3 mg ml^{-1} catalase, in 50 mM Tris.HCl pH 7.5 at 37ºC for 30 minutes. The final reaction volume was 100 µl. Reactions were quenched by adding 200 µl of methanol followed by 30 minute incubation on ice. Hyamine hydroxide was added to each reaction mixture and radiations emitted were measured.

3.6.2 Comparing mutants by MALDI-MS assays

For MALDI-MS assays, the JMJD6 catalysed hydroxylation reactions were initiated by mixing the 100 µM $LUC7L2_{267-278}$, 500 µM 2OG, 400 µM Fe(II), 100 µM ascorbic acid in 50mM Tris.HCl buffer pH 7.5 and 20 µM of either wild-type JMJD6 or mutant JMJD6 enzyme together in a final volume of 20 µl and incubated at 37°C for 30 minutes. The reaction was quenched by adding 10 µl of 0.1% HCOOH. MS analyses were carried out using a MALDI-TOF MS. The quenched reaction mixture (1 µl) was spotted directly onto the MALDI target plate, followed by CHCA matrix (1 µl), and allowed to air dry. Mass spectra were acquired in the positive reflectron mode with a 17 kV acceleration voltage.

3.7 Results and Discussion

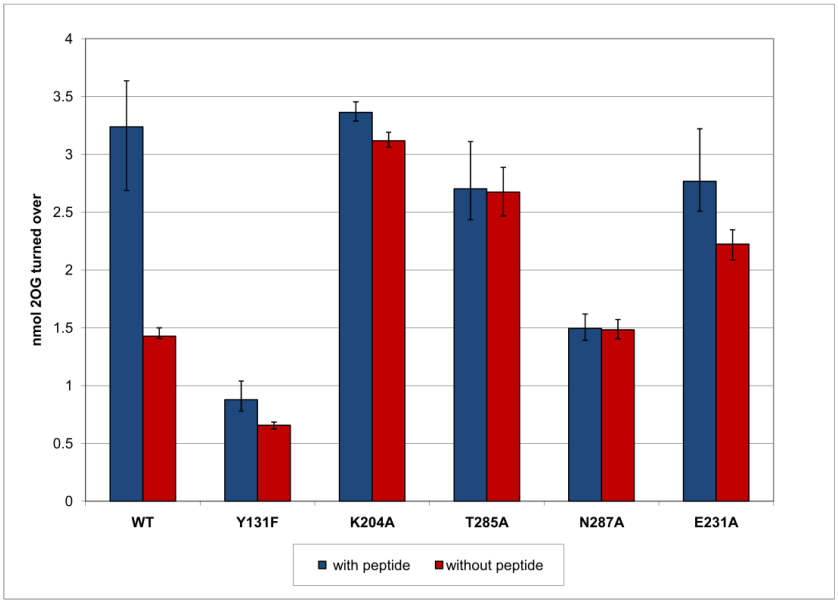


Figure 3.13 Activity of wild-type and variant JMJD6 enzymes analysed by 1-[^{14}C] 2OG turnover assay. Data for JMJD6 enzyme variants in the presence and absence of peptide is shown.

Consistent with the proposed mode of 2OG binding as discussed in Section 3.4, the Lys204Ala variant is inactive and the Tyr131Phe variant displays significantly reduced activity (about 10% hydroxylation as calculated from intensity of hydroxylated and unhydroxylated peptide peaks corresponding to 1595.64 and 1579.68 respectively) in MALDI-MS based hydroxylation detection assays as shown in Figure 3.14D. Interestingly, the Lys204Ala variant retained wild-type levels of 2OG turnover activity in the absence of peptide substrate as measured by the radiolabelled 2OG turnover assay (Figure 3.13G) suggesting that 2OG was still capable of binding to and reacting with the Lys204Ala variant, however, in a non-productive manner with respect to peptide hydroxylation. For most other 2OG oxygenases, replacing the basic (arginine or lysine) 2OG binding residue equivalent to Lys204 significantly reduces or ablates primary substrate-uncoupled 2OG turnover. These results suggest a role for Lys204 in the

coupling of 2OG turnover with substrate hydroxylation by JMJD6. The Tyr131Phe variant has substantially lower 2OG turnover than wild-type JMJD6, only a small amount of which is coupled to primary substrate hydroxylation (Figure 3.13) suggesting that the Tyr131Phe, unlike the Lys204Ala variant, can bind 2OG in a partially hydroxylation-productive manner.

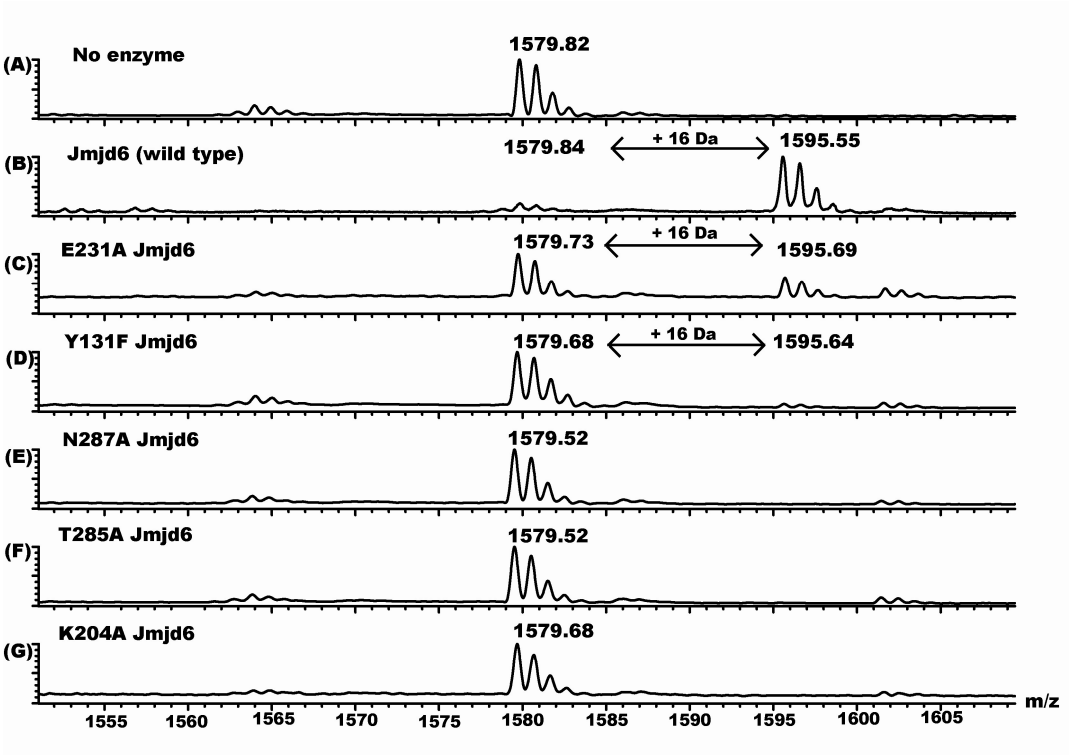


Figure 3.14 Activity of wild-type and variant JMJD6 enzymes as tested by MALDI-MS assays in the presence of LUC7L2 peptide.

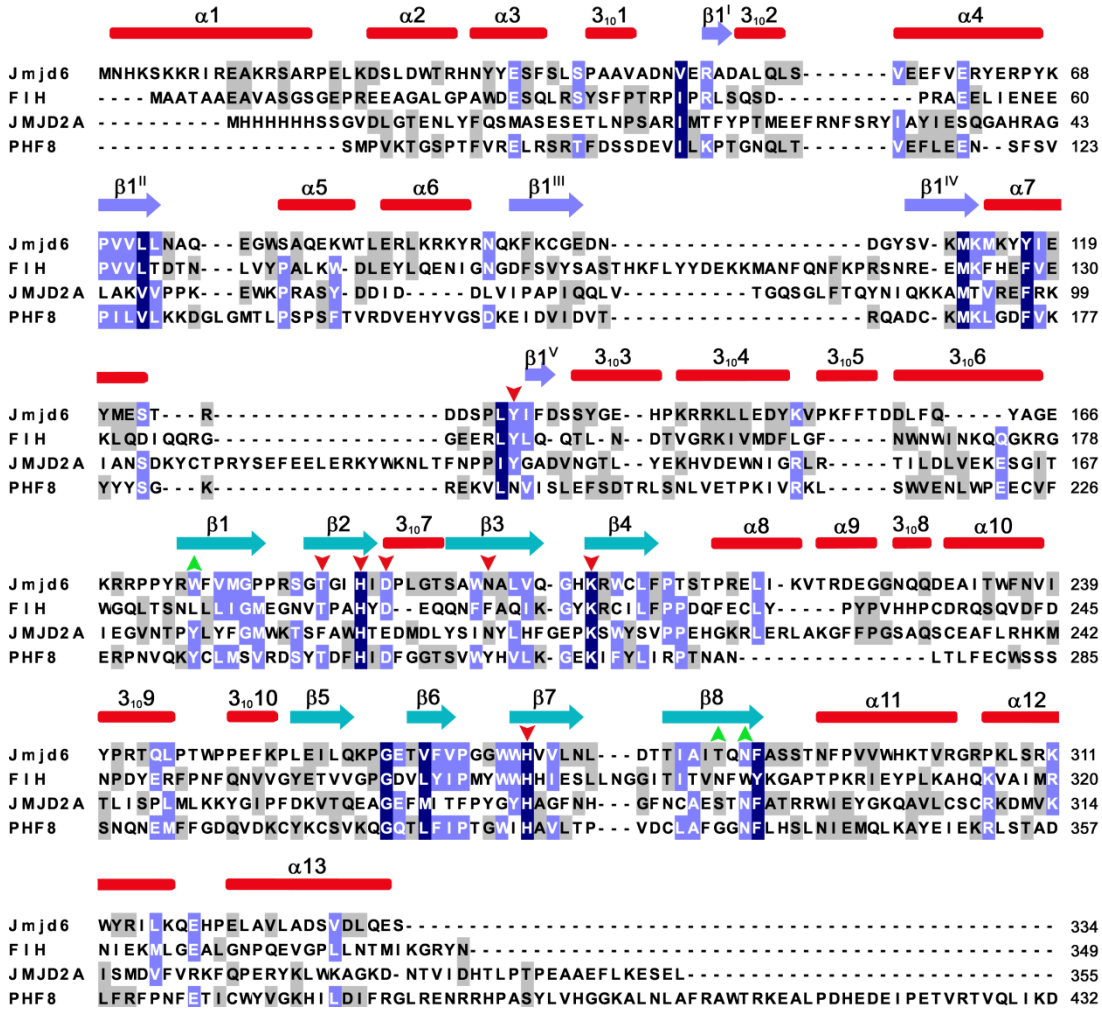


Figure 3.15 Structure based sequence alignment of JmjC hydroxylases and demethylases (JMJD6 PDB ID: 3K2O, FIH PDB ID: 1H2K, JMJD2A PDB ID: 2OQ6, and PHF8 PDB ID: 2WWU). The alignment was created using EXPRESSO (3D-Coffee)[35]. Conservation is indicated by background color; blue – strictly conserved, light blue – semi-conserved, grey – some conservation, and white – non-conserved. The JMJD6 secondary structure is indicated above the sequence as α-helix (red cylinders) and β-strand (blue/purple arrows). The double-stranded β-helix core is highlighted with cyan arrows. Red triangles indicate metal and predicted 2OG binding residues. Green triangles indicate the residue predicted to be involved in binding the hydroxylated lysine residue of the substrate. RefSeq IDs NP_001074930.1 (JMJD6), NP_060372.2 (FIH), NP_055478.2 (JMJD2A), NP_055922.1 (PHF8). R.m.s.d. between JMJD6 and: FIH (PDB ID;1H2K) 1.77 Å using 172 best matched residues; PHF8 (PDB ID;2WWU) 1.74 Å using 179 best matched residues; JMJD2A (PDB ID;2OQ6) 1.84 Å using 169 best matched residues. Sequence identity between JMJD6 and : FIH 7%; PHF8 8%; JMJD2A 12% based on a clustalW alignment of the crystallised domain sequence analysed above. *This image was prepared by Dr. Michael McDonough.*

Although a crystal structure of a substrate bound at the JMJD6 active site is not available, analyses of the variant activity levels and available 2OG oxygenase substrate

structures together with results from docking $LUC7L2_{267-274}$ substrate suggest how JMJD6 may bind the lysine residue of its substrate.[1, 3, 4, 6, 28, 29]

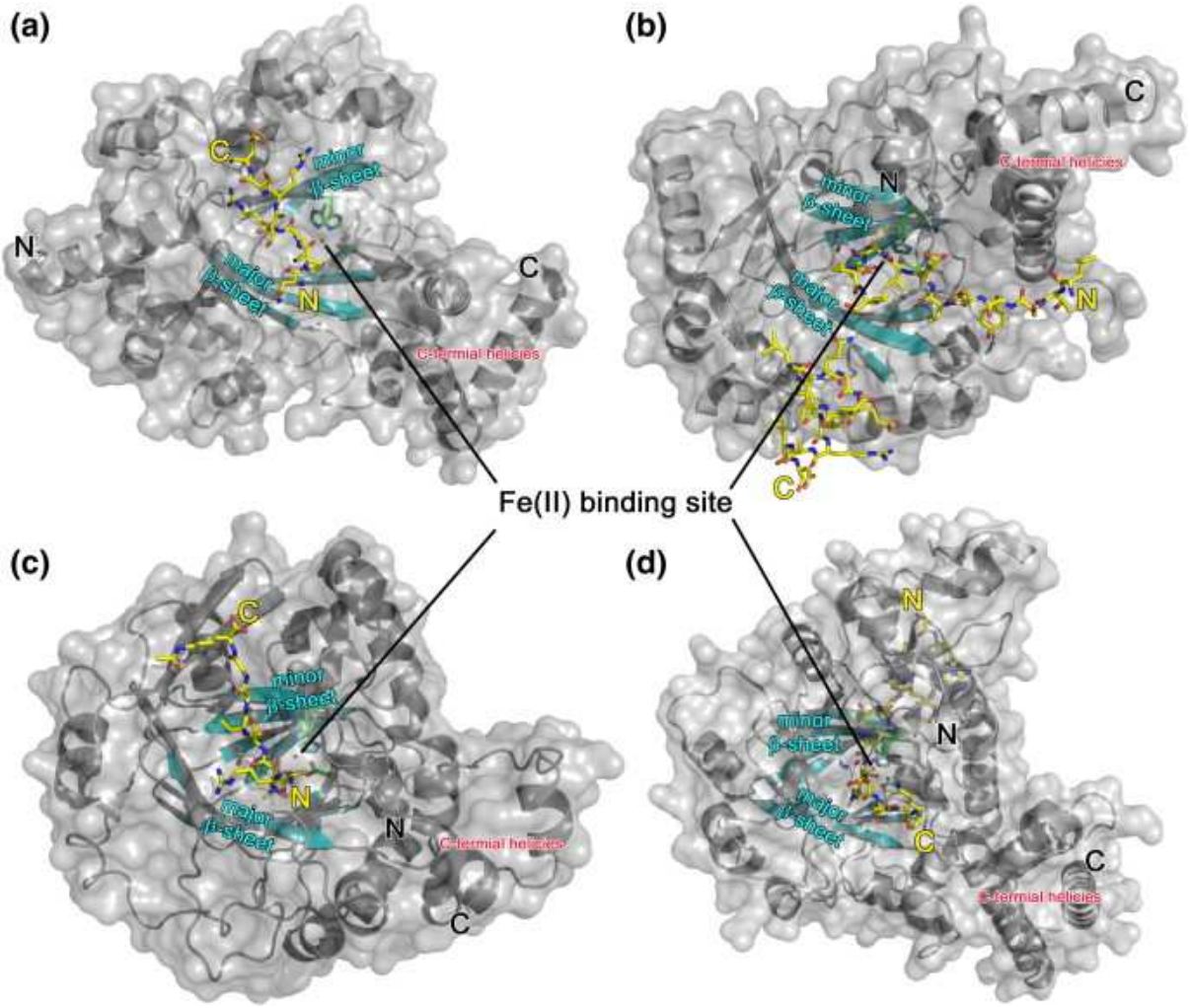


Figure 3.16 The JMJD6 substrate binding site in relation to the metal and 2OG binding sites: comparison with substrate complex structures for other JmjC oxygenases. All enzymes are represented by grey ribbons with the DSBH core highlighted in cyan, metal binding triad shown as green sticks, NOG as white sticks (where present), metal ion as orange sphere, substrate as yellow stick and a transparent light-grey protein surface (a) JMJD6 with predicted LUC7L2 peptide binding mode. (b) The asparaginyl hydroxylase FIH binding site with substrate (HIF-1α *C*-terminal transactivation domain). [28] (c) The substrate binding site of the histone demethylase JMJD2A. [1] (d) The substrate binding of histone N^{ϵ}-methyl lysine demethylase PHF8.

The proposed substrate lysine binding residues (Phe133, Trp174, and Asn287) are located on three beta-strands ($\beta1^{V}$, β1, and β8 in JMJD6 as shown in Figure 3.15) which form a part of the major beta-sheet. Phe133 was refined in two conformations,

suggesting that in JMJD6 some active-site residues may be involved in induced-fit binding, as proposed for FIH.[28] FIH crystallographic analyses have indicated that $Trp296_{FIH}$, which is analogous to Asn287 in JMJD6, undergoes movement upon substrate binding. Asn287, Tyr131, Phe133 and Trp174 of JMJD6 are conserved in JMJD2A with corresponding residues $Asp290_{JMJD2A}$, $Tyr132_{JMJD2A}$, $Tyr177_{JMJD2A}$, $Tyr175_{JMJD2A}$. One of the conformations of Phe133 appears to overlap with the side-chain of $Tyr177_{JMJD2A}$, a 2OG binding pocket residue in JMJD2A also predicted to be involved in substrate recognition.[1] It can be proposed that the overall effect of the differences between JMJD2A and JMJD6 is to allow the substrate lysine residue to project deeper into the JMJD6 active site than does the N^{ϵ}-methyl lysine residue substrate of JMJD2A[1] thus making the predicted substrate binding mode of JMJD6 more like that of the di-N^{ϵ}-methyl lysine residue substrate of PHF8.[6] This binding mode is predicted to enable the substrate lysine amino group to hydrogen bond with the side-chains of Asn287 and/or Thr285. Substitution of either of these residues with alanine (Asn287Ala, Thr285Ala) ablates hydroxylation activity under standard assay conditions (Figure 3.14).

With FIH, the asparaginyl substrate side-chain adopts a different conformation in the active site to that observed for JMJD2A and PHF8 substrate complex structures and that predicted for JMJD6. In part this appears to be enabled by the presence of the $Trp296_{FIH}$ side-chain which would block the N^{ϵ}-methyl lysine side-chain binding as observed for JMJD2A and PHF8. Additional hydrogen bonding interactions of the HIF asparagine substrate side chain with $Gln239_{FIH}$, derived from the "insert" loop that links the DSBH β-strands β4 and β5, contribute to its positioning in the active site. Elements of the analogous insertion loop of JMJD6 are also predicted to be involved in substrate

recognition. These include a pair of acidic residues, Asp230 and Glu231, that may help to recognise the basic region of substrate (JMJD6 substrates tend to be basic in the region around the target lysine[27]) (Figures 3.11. 3.13 and 3.14). Furthermore, acidic residues at the "mouth" of the active-site including Asp230 and Glu231 from the β4-β5 insert, may contribute to substrate binding in regions upstream or downstream of the hydroxylated lysine. The Glu231Ala variant displays reduced activity in both hydroxylation and 2OG turnover assays (Figure 3.13 and 3.14) consistent with the proposal that Glu231 is involved in substrate binding.

3.8 Conclusions

This chapter describes work leading to the first reported crystal structure of JMJD6. Overall, the combined structural and mutagenesis work supports the proposal that JMJD6 is closely related in overall fold to the other JmjC subfamily members. However, there are clear differences at the active site of JMJD6 compared both to the JmjC N^{ε}-methyl lysyl-demethylases and the HIF asparaginyl hydroxylase, FIH. It should be possible to exploit these differences to develop selective inhibitors of different members of the JmjC 2OG dependent oxygenases for use in therapeutic applications, ranging from ischemic disease (FIH) to cancer (lysyl-demethylases).

Modelling suggests that acidic residues are probably involved in substrate binding. However, to understand the actual interaction of SR substrates with JMJD6 and to obtain insight into how JMJD6 catalyses hydroxylation of lysine residue in stereospecific manner (described later in Chapter 4), it is necessary to obtain crystal structures of JMJD6 with substrate.

3.9 References

1. Ng Stanley, S.; Kavanagh Kathryn, L.; McDonough Michael, A.; Butler, D.; Pilka Ewa, S.; Lienard Benoit, M. R.; Bray James, E.; Savitsky, P.; Gileadi, O.; von Delft, F.; Rose Nathan, R.; Offer, J.; Scheinost Johanna, C.; Borowski, T.; Sundstrom, M.; Schofield Christopher, J.; Oppermann, U., Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity. *Nature* **2007,** *448* (7149), 87-91.

2. Chen, Z.; Zang, J.; Whetstine, J.; Hong, X.; Davrazou, F.; Kutateladze, T. G.; Simpson, M.; Mao, Q.; Pan, C. H.; Dai, S.; Hagman, J.; Hansen, K.; Shi, Y.; Zhang, G., Structural insights into histone demethylation by JMJD2 family members. *Cell* **2006,** *125* (4), 691-702.

3. Chen, Z.; Zang, J.; Kappler, J.; Hong, X.; Crawford, F.; Wang, Q.; Lan, F.; Jiang, C.; Whetstine, J.; Dai, S.; Hansen, K.; Shi, Y.; Zhang, G., Structural basis of the recognition of a methylated histone tail by JMJD2A. *Proceedings of the National Academy of Sciences* **2007,** *104* (26), 10818-10823.

4. Couture, J. F.; Collazo, E.; Ortiz-Tello, P. A.; Brunzelle, J. S.; Trievel, R. C., Specificity and mechanism of JMJD2A, a trimethyllysine-specific histone demethylase. *Nature Structural & Molecular Biology* **2007,** *14* (8), 689-695.

5. Yue, W. W.; Hozjan, V.; Ge, W.; Loenarz, C.; Cooper, C. D.; Schofield, C. J.; Kavanagh, K. L.; Oppermann, U.; McDonough, M. A., Crystal structure of the PHF8 Jumonji domain, an N(ε)-methyl lysine demethylase. *FEBS Letters* **2010,** *584* (4), 825-830.

6. Horton, J. R.; Upadhyay, A. K.; Qi, H. H.; Zhang, X.; Shi, Y.; Cheng, X., Enzymatic and structural insights for substrate specificity of a family of jumonji histone lysine demethylases. *Nature Structural & Molecular Biology* **2010,** *17* (1), 38-43.

7. Prockop, J. D., Collagens: Molecular Biology, Diseases, and Potentials for Therapy. *Annual Review of Biochemistry* **1995,** *64* (1), 403-434.

8. Zygmunt S, D., The use of recombinant methods and molecular engineering in protein crystallization. *Methods* **2004,** *34* (3), 354-363.

9. Sickmeier, M.; Hamilton, J. A.; LeGall, T.; Vacic, V.; Cortese, M. S.; Tantos, A.; Szabo, B.; Tompa, P.; Chen, J.; Uversky, V. N.; Obradovic, Z.; Dunker, A. K., DisProt: the Database of Disordered Proteins. *Nucleic Acids Research* **2007,** *35*, 786-793.

10. Guerrero, S. A., Hecht, H. J., Hofmann, B., Biebl, H., Singh, M., Production of selenomethionine-labelled proteins using simplified culture conditions and generally applicable host/vector systems. *Appl. Microbiol. Biotech.* **2001,** *56*, 718-723.

11. LeMaster, D. M., Richards, F. M., Proton-nitrogen-15 heteronuclear NMR studies of Escherichia coli thioredoxin in samples isotopically labeled by residue type. *Biochemistry* **1985,** *24* (25), 7263-7268.

12. Otwinowski Z., M. W., Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology* **1997,** *276* (part A), 307-326.

13. Kabsch, W., Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *Journal of Applied Crystallography* **1993,** *26* (6), 795-800.

14. Collaborative Computational Project, Number 4. The CCP4 Suite: Programs for protein crystallography. *Acta Crystallographica Section D, Biological ctystallography* **1994,** *50* (Pt. 5), 760-763.

15. Schneider, T. R.; Sheldrick, G. M., Substructure solution with SHELXD. *Acta Crystallographica Section D, Biological ctystallography* **2002,** *58* (10), 1772-1779.

16. Abrahams, J. P.; Leslie, A. G., Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallographica Section D, Biological ctystallography* **1996,** *52* (Pt.1), 30-42.

17. Perrakis, A.; Harkiolaki, M.; Wilson, K. S.; Lamzin, V. S., ARP/wARP and molecular replacement. *Acta Crystallographica Section D, Biological ctystallography* **2001,** *57* (10), 1445-1450.

18. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., Phaser crystallographic software. *Journal of Applied Crystallography* **2007,** *40* (Pt 4), 658-674.

19. Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L.-W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C., PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallographica Section D, Biological ctystallography* **2002,** *58* (Pt. 11), 1948-1954.

20. Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D, Biological ctystallography* **2004,** *60* (12), 2126-2132.

21. Krissinel, E.; Henrick, K., Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology* **2007,** *372* (3), 774-797.

22. Tibrewal, N.; Liu, T.; Li, H.; Birge, R., Characterization of the biochemical and biophysical properties of the phosphatidylserine receptor (PS-R) gene product. *Molecular and Cellular Biochemistry* **2007,** *304* (1), 119-125.

23. Han, G.; Li, J.; Wang, Y.; Li, X.; Mao, H.; Liu, Y.; Chen, C. D., The hydroxylation activity of Jmjd6 is required for its homo-oligomerization. *Journal of Cellular Biochemistry* **2012,** *113* (5), 1663-1670.

24. Hutchinson, E. G.; Thornton, J. M., HERA - A program to draw schematic diagrams of protein secondary structures. *Proteins: Structure, Function and Genetics* **1990,** *8* (3), 203-212.

25. Hutchinson, E. G.; Thornton, J. M., PROMOTIF - A program to identify and analyze structural motifs in proteins. *Protein Science* **1996,** *5* (2), 212-220.

26. Bond, C. S., TopDraw: a sketchpad for protein structure topology cartoons. *Bioinformatics* **2003,** *19* (2), 311-312.

27. Webby, C. J.; Wolf, A.; Gromak, N.; Dreger, M.; Kramer, H.; Kessler, B.; Nielsen, M. L.; Schmitz, C.; Butler, D. S.; Yates, J. R., III; Delahunty, C. M.; Hahn, P.; Lengeling, A.; Mann, M.; Proudfoot, N. J.; Schofield, C. J.; Bottger, A., Jmjd6 Catalyses Lysyl-Hydroxylation of U2AF65, a Protein Associated with RNA Splicing. *Science* **2009,** *325* (5936), 90-93.

28. Elkins, J. M.; Hewitson, K. S.; McNeill, L. A.; Seibel, J. F.; Schlemminger, I.; Pugh, C. W.; Ratcliffe, P. J.; Schofield, C. J., Structure of factor-inhibiting hypoxia-inducible factor (HIF) reveals mechanism of oxidative modification of HIF-1α. *Journal of Biological Chemistry* **2003,** *278* (3), 1802-1806.

29. Clifton, I. J.; McDonough, M. A.; Ehrismann, D.; Kershaw, N. J.; Granatino, N.; Schofield, C. J., Structural studies on 2-oxoglutarate oxygenases and related double-stranded β-helix fold proteins. *Journal of Inorganic Biochemistry* **2006,** *100* (4), 644-669.

30. McDonough, M. A.; Loenarz, C.; Chowdhury, R.; Clifton, I. J.; Schofield, C. J., Structural studies on human 2-oxoglutarate dependent oxygenases. *Current Opinion in Structural Biology* **2010,** *20* (6), 659-672.

31. Zhang, Z.; Ren, J.; Stammers, D. K.; Baldwin, J. E.; Harlos, K.; Schofield, C. J., Structural origins of the selectivity of the trifunctional oxygenase clavaminic acid synthase. *Nature Structural & Molecular Biology* **2000,** *7* (2), 127-133.

32. Trott, O.; Olson, A. J., AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry* **2010,** *31* (2), 455-461.

33. Csaba, H.; David van der, S., Efficient docking of peptides to proteins without prior knowledge of the binding site. *Protein Science* **2002,** *11* (7), 1729-1737.

34. Hewitson, K. S.; McNeill, L. A.; Riordan, M. V.; Tian, Y.-M.; Bullock, A. N.; Welford, R. W.; Elkins, J. M.; Oldham, N. J.; Bhattacharya, S.; Gleadle, J. M.; Ratcliffe, P. J.; Pugh, C. W.; Schofield, C. J., Hypoxia-inducible factor (HIF) asparagine hydroxylase is identical to factor inhibiting HIF (FIH) and is related to the cupin structural family. *Journal of Biological Chemistry* **2002,** *277* (29), 26351-26355.

35. Armougom, F.; Moretti, S.; Poirot, O.; Audic, S.; Dumas, P.; Schaeli, B.; Keduas, V.; Notredame, C., Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Research* **2006,** *34*, 604-608.

Chapter 4. Stereochemical determination of lysine hydroxylation by JMJD6

4.1 Introduction

Collagen prolyl- and lysyl- hydroxylases were the first reported enzymes to catalyse post-translational hydroxylations on protein carbons. Both collagen prolyl- and lysyl-hydroxylases belong to the family of Fe(II) and 2OG dependent oxygenases. Subsequently, 2OG oxygenases have been shown to catalyse other protein hydroxylation reactions.[1, 2] Epidermal growth factor domain type containing proteins undergo carbon hydroxylation of aspartyl and asparaginyl residues.[1] Intracellularly localised proteins also undergo hydroxylation on carbon. Following the identification of the HIF prolyl and asparaginyl hydroxylases,[3, 4] one of these enzymes, asparaginyl hydroxylase, FIH has been shown to catalyse multiple hydroxylations in ankyrin repeat domains.[5-7] There is also evidence for alternative substrates for the HIF prolyl-hydroxylases.[5-7] Further N^{ϵ}-methyl lysine demethylation is catalysed by a widespread family of histone demethylases[8] (JmjC subfamily) in a mechanism likely involving initial hydroxylation of the N^{ϵ}-methyl group.[1, 9, 10]

2OG oxyygenases can catalyse the hydroxylation of varied substrate residues due to the structural differences in their active sites. These differences also give rise to variations in the regio- and stereo-chemical specificity of identified 2OG dependent hydroxylases. In the case of collagen prolyl- hydroxylation, different regio/stereo chemical outcomes of catalysis are associated with different functions. *trans*-4-Prolyl hydroxylation stabilises the collagen triple helix, whereas *trans*-3-prolyl hydroxylation is proposed to

destabilise collagen, enabling further modifications.[11] Further, in the case of epidermal growth factor domain protein, hydroxylation of aspartyl and asparginyl residues results in the 3*R* products,[12] whereas HIF and ankyrin repeat domain, hydroxylation of aspartyl residues occurs to give the 3*S*-products (Figure 4.1).[6, 13] Thus, in addition to identifying the residue that is oxidised, it is important to define the regio- and stereo-chemical outcome of the hydroxylation reaction.

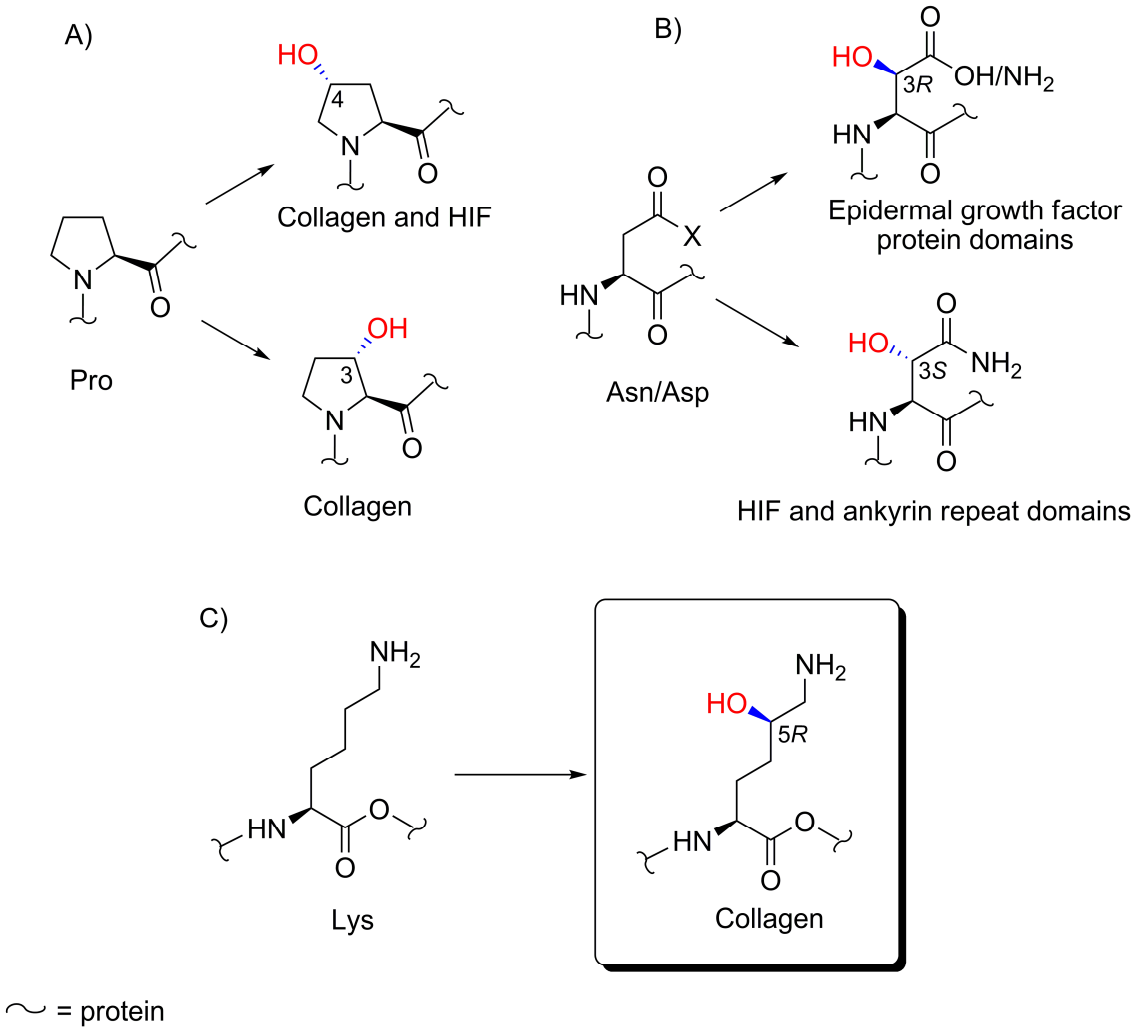


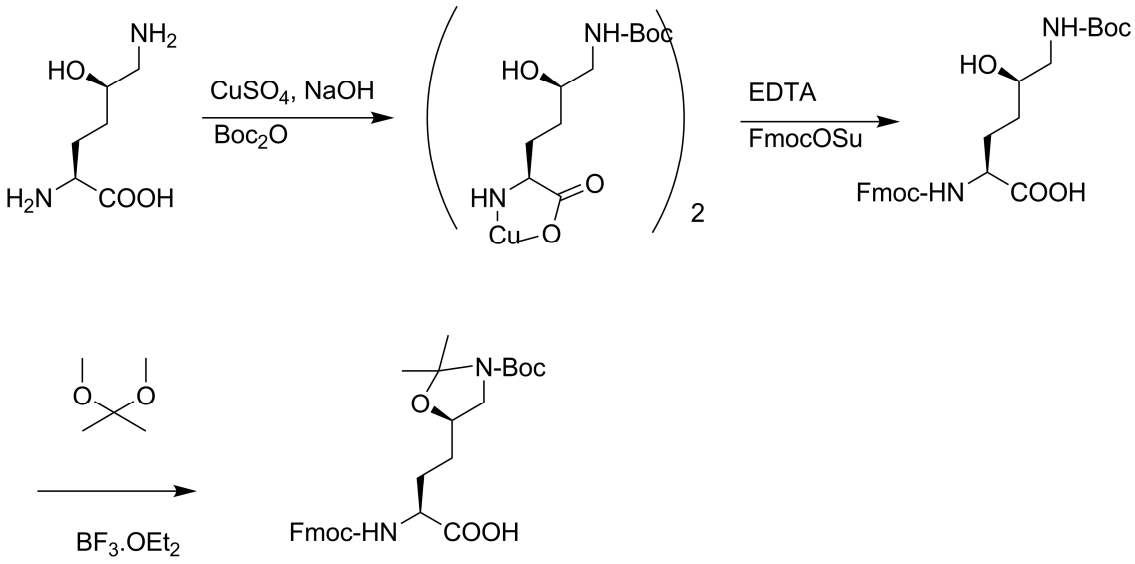
Figure 4.1 Different observed regio and stereo-chemical outcomes of hydroxylation products by 2OG dependent oxygenases in animals. (a) Products of collagen and HIF prolyl- hydroxylation; (b) Products of epidermal growth factor type domain and aspartyl- and asparaginyl- hydroxylations; (c) Products of collagen hydroxylation by lysyl hydroxylase.

This chapter describes investigations on the stereochemical determination of the hydroxylysine product of JMJD6 catalysis. It is known that the collagen lysyl hydroxylase catalyses hydroxylation of lysines from collagen to give the 2*S*,5*R* hydroxylysine residue.[14] Thus, in the case of JMJD6, the sterochemical determination of the product was compared to the hydroxylysine product in collagen. NMR and amino acid analyses were employed in these studies.

4.2 Synthesis of Fmoc protected hydroxylysine

Webby et al. have reported that the splicing regulatory protein LUC7L2 undergoes JMJD6 catalysed C-5 hydroxylation at Lys 269 and that a peptide fragment of LUC7L2 is a JMJD6 substrate ($LUC7L2_{267\text{-}278}$).[15] In an attempt to define the nature of the JMJD6 catalysed product, work began by synthesising the $LUC7L2_{267\text{-}278}$ peptide with a 2*S*,5*R* hydroxylysine at residue 269 and then comparing it with a hydroxylated peptide produced by JMJD6 catalysis using NMR and MS analyses. Fmoc-2*S*,5*R*-hydroxylysine (Boc-oxazolidine)) was synthesised essentially as described by Spetzler et al.[16] as represented in Scheme 4.1, although some optimisation of the synthesis was required to give sufficient amounts of product for peptide synthesis. Initially, the α-amino acid functional group of 5*R*-hydroxylysine was protected by treatment with $CuSO_4 \cdot 5H_2O$ under alkaline conditions. The generated copper complex was then treated with Boc anhydride in dioxane which gave the protected product as a precipitate after 3 days. The copper complex was dissociated by treatment with an aqueous solution of EDTA and the liberated α-amine was then protected by addition of N-(9-Fluorenylmethoxycarbonyloxy) succinimide (Fmoc-OSu). In order to avoid lactone formation in Fmoc-Hyl(Boc), Boc protected *N*-Fmoc-2*S*,5*R*-hydroxylysine was immediately treated with dimethoxypropane–$BF_3 \cdot OEt_2$ in acetone. The final desired

product Fmoc-2*S*,5*R*-hydroxylysine (Boc-oxazolidine) was isolated after work-up using ethyl acetate and water. The crude product was purified by flash column chromatography.



Scheme 4.1 Reaction scheme for Fmoc-2*S*,5*R*-hydroxylysine (Boc-oxazolidine)

4.3 Peptide Synthesis

A peptide fragment of LUC7L2$_{267\text{-}278}$ (sequence NPKRSRSREHRR where the lysine at position 269 is replaced by hydroxylysine) was synthesized by solid phase peptide synthesis. All the α-Fmoc protected amino acids and Fmoc-2*S*,5*R*-hydroxylysine (Boc-oxazolidine) were coupled on Rink Amide MBHA resin using the HOBt and DIC coupling reaction described in Chapter 2. Both the oxazolidine and Boc groups used for protecting the hydroxyl and amino groups in hydroxylysine are deprotected by cleavage with CF_3COOH in the presence of tri-isopropyl silane in order to trap the reactive cation intermediates. LUC7L2$_{267\text{-}278}$ was also synthesised in order to compare the synthetic hydroxylated peptide with *in vitro* catalysed product of JMJD6. Both these peptides

were then purified by reverse phase HPLC purification using a reverse phase C18 column.

4.4 Hydroxylation of LUC7L2 peptide by JMJD6

A large scale hydroxylation of purified $LUC7L2_{267-278}$ by recombinant JMJD6 was carried out in order to obtain the hydroxylation product in sufficient (milligram) amounts desired for NMR analyses. The concentrations of cofactors/cosubstrates were scaled by a factor of 10 and the reaction (final volume 1 ml) was set up in 30 separate Eppendorf tubes such that in total 10 mg of peptide substrate was reacted. The reaction was initiated by mixing the substrate, JMJD6, Fe(II) and 2OG mixtures together, then followed by incubation at 37°C. After two hours, hydroxylation of the peptide was analysed by MALDI-MS using the CHCA matrix. Hydroxylation was further carried out by leaving the assay mixture overnight at room temperature until the peptide showed substantial hydroxylation (90% as analysed by MALDI-MS).

JMJD6 was precipitated from reaction mixture by adding methanol and the resultant mixture was centrifuged (13,000 rpm for 10 minutes). The supernatant was then isolated, freeze dried and resuspended in 0.1% formic acid/water. Purified hydroxylated $LUC7L2_{267-278}$ was obtained by HPLC purification using a reverse phase C18 column with MS-based detection.

4.5 NMR analysis of hydroxylated lysine

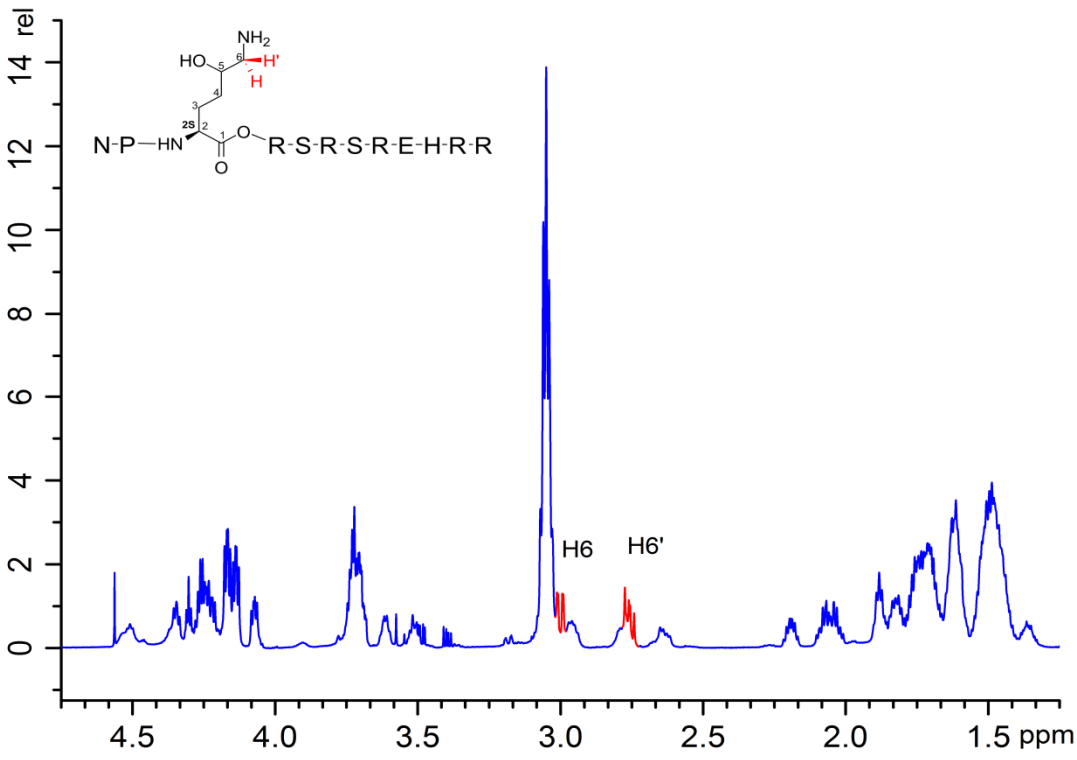


Figure 4.2 ^{1}H NMR spectra of LUC7L2$_{267-278}$ peptide with Lysine269, hydroxylated by JMJD6 enzyme.

All the NMR experiments for this Chapter were carried out by Dr. Refaat Hamed. All the NMR spectra in this chapter were recorded in D_2O. ^{1}H NMR spectra of hydroxylated peptide which was catalysed by JMJD6 *in vitro*, showed peaks at 3.0 ppm and 2.8 ppm which were assigned to protons at C6 of hydroxylysine (Figure 4.2). The synthetic and enzymatically hydroxylated peptides were then compared by heteronuclear single quantum coherence (HSQC) and 1D total correlation nuclear magnetic resonance spectroscopy (TOCSY). The HSQC spectra of 2*S*,5*R* hydroxylysine (Sigma) (brown and green) and JMJD6 catalysed LUC7L2$_{267-278}$ (red and blue) are overlayed in Figure 4.3. The protons from C5 (multiplet at 3.9 ppm) and C6 (CHN [3.0 ppm], CH'N [2.8 ppm]) of 2*S*,5*R* hydroxylysine are appearing in the similar positions to

that hydroxylysine from peptide, while the carbon resonances appear in the same position indicating and reconfirmating that JMJD6 catalyses hydroxylation at C5 position of Lys269 in LUC7L2$_{267-278}$.

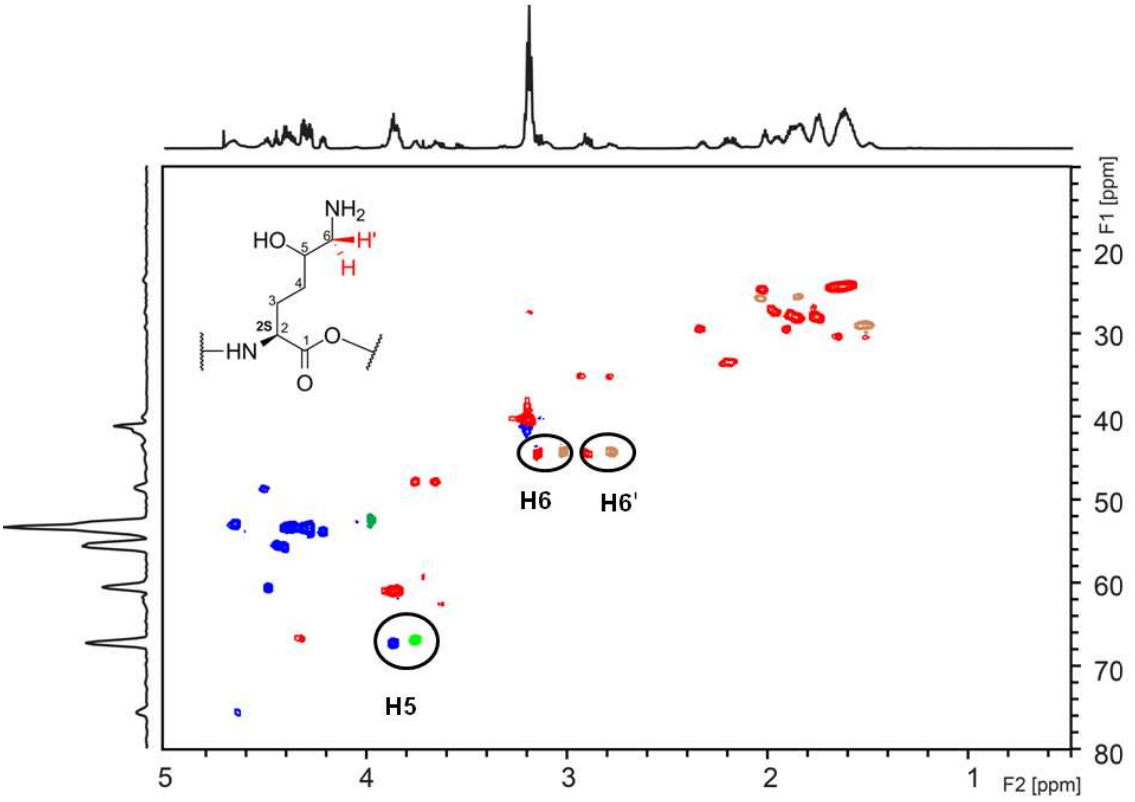


Figure 4.3 Superimposition of HSQC spectra of hydroxylysine containing LUC7L2$_{267-278}$ product of JMJD6 enzyme (shown in red and blue) and 2*S*,5*R* hydroxylysine (shown in brown and green).

A selective 1D-TOCSY experiment was then carried out on both the peptides using frequencies of protons at 3.0 ppm and 2.8 ppm. Comparison of these 1D-TOCSY spectra with ^{1}H NMR of 2*S*,5*R* hydroxylysine is shown in Figure 4.4. However, for very small quantity of enzymatically produced material available, the NMR analyses were unable to unambiguously distinguish between the 5*R*- and 5*S*-products due to the small differences in the observed coupling constants for the 5*R*- and 5*S* protons*;* for 2*S*,5*R*

hydroxylysine the *J* values for H-6 and H-6' are 13.3, 3.1 Hz (H-6, 2.97 ppm) and 13.3, 9.6 Hz (H-6', 2.73 ppm), respectively; for 2*S*,5*S*-hydroxylysine, the values are 13.1, 1.7 Hz (H6, 2.96 ppm) and 13.1, 10.0 Hz (H-6', 2.73 ppm).[17]

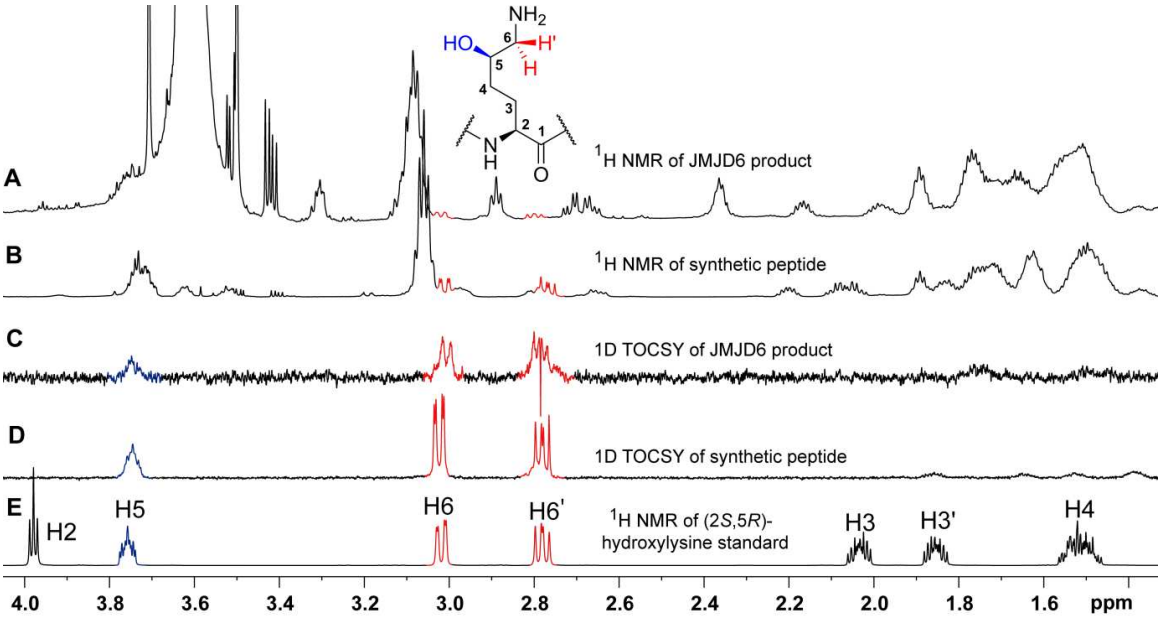


Figure 4.4 NMR analyses of JMJD6 catalysed production of a hydroxylysine residue in a LUC7L2$_{267\text{-}278}$ peptide. (A) ^{1}H NMR of JMJD6 catalysed hydroxylation of LUC7L2$_{267\text{-}278}$ peptide; (B) ^{1}H NMR of synthetic peptide LUC7L2$_{267\text{-}278}$ containing a 2*S*,5*R*-hydroxylysine residue at 269; (C) 1D TOCSY spectrum of the JMJD6 catalysed LUC7L2$_{267\text{-}278}$ product; (D) 1D TOCSY of synthetic peptide LUC7L2$_{267\text{-}278}$ containing 2*S*,5*R* hydroxylysine; (E) ^{1}H NMR of a 2*S*,5*R* hydroxylysine standard.

4.6 Amino acid analysis

Due to the difficulty in determining the stereochemistry of the C-5 hydroxylysine product by NMR studies, amino acid analyses were carried out on JMJD6 catalysed peptide LUC7L2 substrate. *The UPLC and MS studies for amino acid analyses were carried out by Nikita Loik.*

Amino acid analysis is a technique commonly used to determine the amino acid contents of proteins and peptides. Amino acid analysis can be used to quantify protein and peptides, to determine the identity of proteins or peptides based on their amino acid

composition, to support protein and peptide structure analysis, to evaluate fragmentation strategies for peptide mapping, and to detect atypical amino acids that might be present in a protein or peptide.[18] A typical wokflow for amino acid analyses involves initial hydrolysis of peptide bonds in a protein or peptide which produces a mixture of amino acids. This mixture is then separated by chromatography following the detection and quantification of amino acids.

Amino acid analysis was first developed by Moore, Stein and co-workers in the 1950s using HCl acid hydrolysis to determine the amino acid content of pure proteins. Their method involved use of 6 M HCl for hydrolysis in an oxygen-free environment at 110°C for 22 hours to liberate amino acids from proteins.[19] They then separated the free amino acids using ion-exchange chromatography and detected and quantified the eluted amino acids using post-column ninhydrin detection. Numerous studies have been conducted with the aim of reducing the analysis time, simplifying the methods, and reducing data variability. While considerable progress has been made with respect to the chromatography, the hydrolysis step has largely remained unchanged.

Many different chromatographic techniques have been used to separate and quantify amino acids. Ion-exchange chromatography of amino acids was originally developed by Moore and Stein in the early 1950s using a stepwise pH gradient to separate amino acids.[19, 20] With the development of HPLC, the separations are efficient and faster with run times of 1 to 2 hours for ion-exchange and 20 to 40 minutes for reversed-phase chromatography.[18] With recently developed Ultra High Pressure Liquid Chromatography (UPLC) systems, run time for amino acid hydrolysates is about 10 minutes.

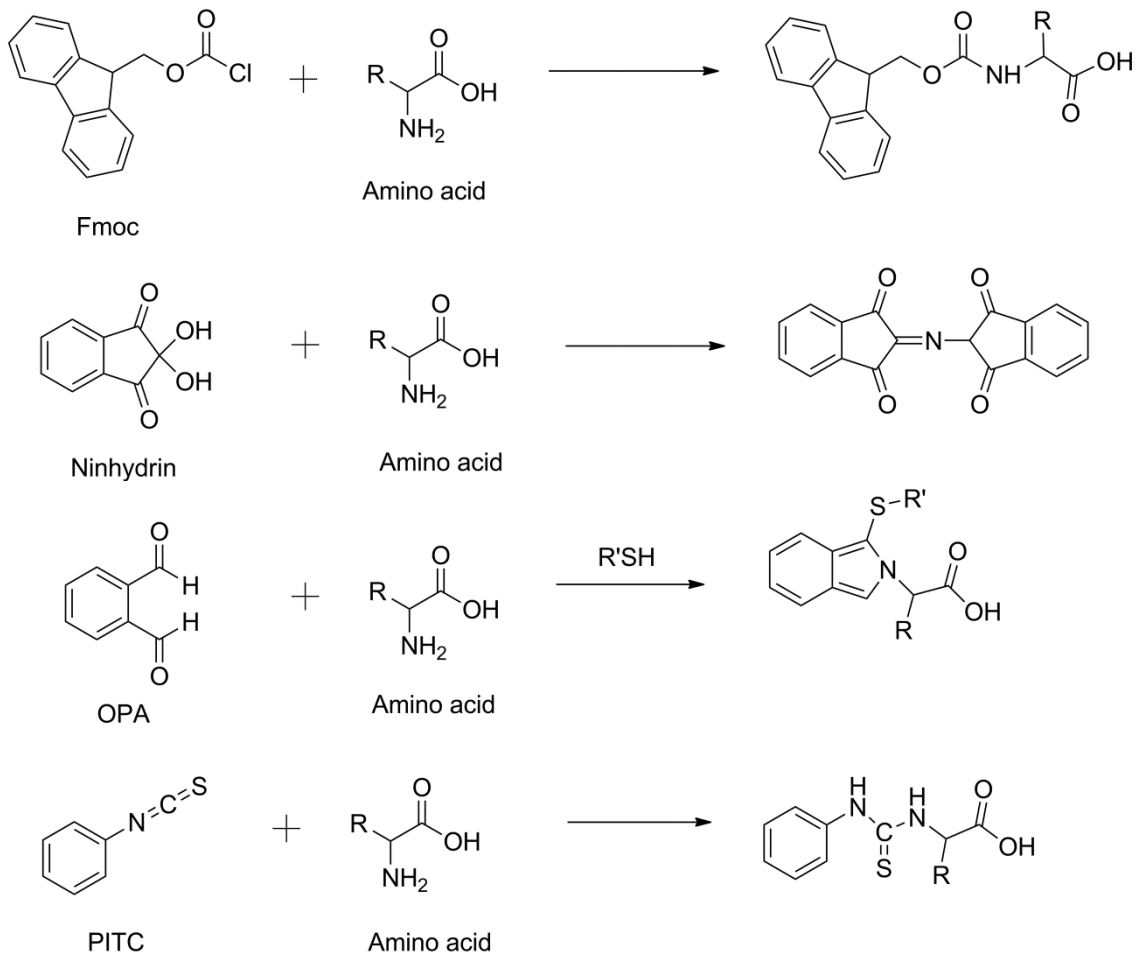


Figure 4.5 Various derivatising agents commonly used for amino acid analysis.

Due to the inherent lack of aromatic absorption for amino acids except for tyrosine, tryptophan and phenylalanine, it is difficult to detect amino acids by UV absorbance.[18] Thus amino acids need to be derivatised to enable their detection. The choice of derivatising agent depends on the sensitivity and specifications required for the assay. There are two ways in which derivatisation is carried out; i.e. pre-column or post-column dervatisation.[21] In the case of the post column derivatisation technique, separation of the free amino acids by ion-exchange chromatography is followed by derivatisation of amino acid with ninhydrin or *o*-phthalaldehyde (OPA).[18] Post-column detection techniques can be used with samples that contain small amounts of buffer

components, such as salts and urea, and generally require between 5 and 10 µg of protein sample per analysis. On the other hand the pre-column derivatisation technique involves derivatisation of the free amino acids by one of the derivatising agents such as phenyl isothiocyanate (PITC), 6-amino-quinolyl-*N*-hydroxysuccinimidyl carbamate (AQC), *o*-phthalaldehyde (OPA), 9-fluorenylmethylchloroformate (Fmoc) followed by reversed-phase HPLC.[18, 22] Pre-column derivatisation techniques are very sensitive and usually require between 0.5 and 1.0 µg of protein sample per analysis but may be influenced by buffer salts in the samples.[18]

In order to achieve the separation of 2*S*,5*R* and 2*S*,5*S* hydroxylysine diasteromers, AccQTag solution protocol (Waters), which is a pre-column derivatisation of hydrolysate by AQC was employed. AQC reacts with amino acids to form stable, fluorescent unsymmetric urea derivatives (λ excitation 250 nm, λ emission 350 nm) which are readily amenable to analysis by reversed-phase HPLC. The reaction is rapid, taking only 1 min at room temperature to complete. [23] After derivatisation, the mixture is heated to 55°C to convert the minor side product of tyrosine to the major mono-derivatised compound and to hydrolyse the AQC reagent to 6-aminoquinoline, *N*-hydroxysuccinimide, and carbon dioxide (Figure 4.6). The derivatives are stable at room temperature for up to a week, and the detection limit of the method is 40 to 800 fmol, depending on the amino acid.[23]

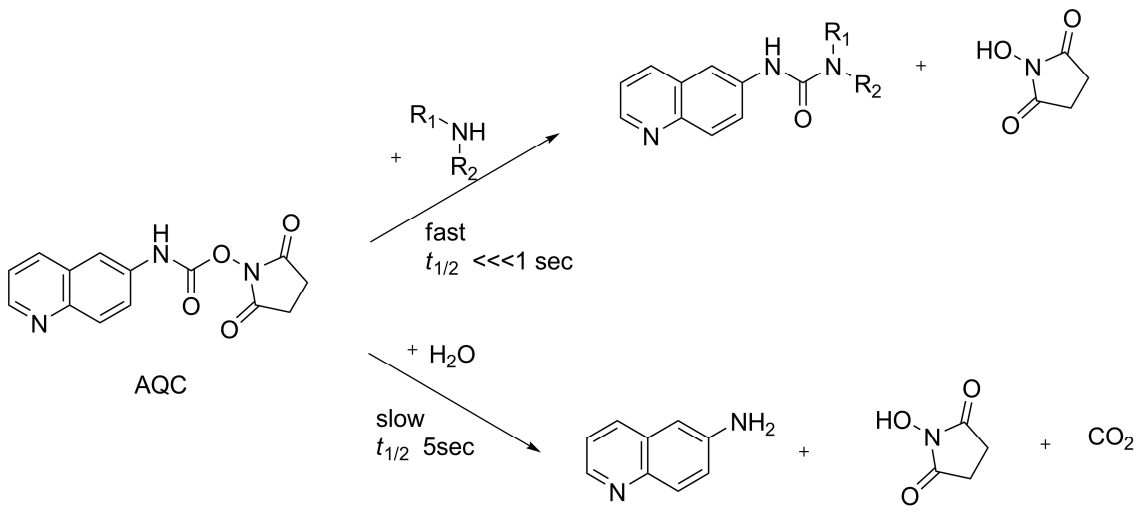


Figure 4.6 Reaction kinetics of amino acid with derivatising agent AQC.

Initially a method for amino acid analysis employing derivatisation using a mixture of all four stereoisomers of racemic C-5 hydroxylysine with AQC[23] was optimised and positive ion electrospray ionisation (ESI) MS analysis was carried out. Separation of the two sets of diastereomers (2*S*,5*S*/2*R*,5*R* and 2*S*,5*R*/2*R*,5*S*) was achieved by optimisation of HPLC conditions and a peak corresponding the MH^+ 503.1 Da (hydroxylysine derivatised on both the amino groups) was monitored by analysis of extracted ion chromatograms.

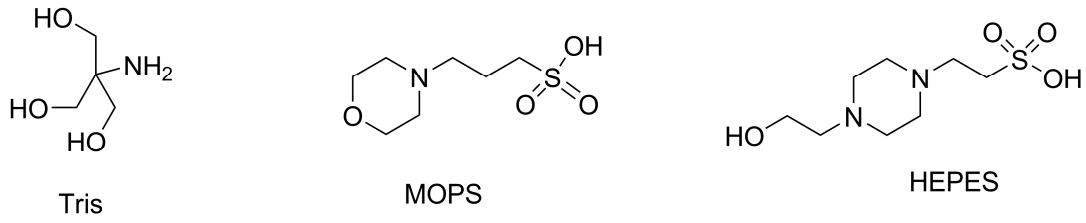


Figure 4.7 Chemical structures of various buffers which were used for lysine hydroxylation assays of JMJD6.

The peptide substrate $LUC7L2_{267\text{-}278}$ (100 µM) was hydroxylated by mixing 2OG (500 µM), Fe(II) (400 µM), ascorbic acid (100 µM) in 50mM Tris buffer pH 7.5 and 20 µM

JMJD6, followed by incubation at 37°C for 1 hour. In order to remove JMJD6 from the assay mixture, it was precipitated by addition of methanol and then removed by centrifugation at 13,000 rpm on benchtop centrifuge. The resultant supernatant containing peptide was isolated and freeze dried in a pyrex tube which was then placed in air tight vial with 6 M HCl. The vial was then incubated at 110°C for 24 hours. The hydrolysate was then resuspended in borate buffer (pH 9.0, 100 µl), derivatised by AQC and analysed by LC-MS. These initial experiments in Tris buffer were hampered by the reaction with AQC to give monoderivatised Tris. As Tris was present in excess (450 fold) relative to the peptide, a large excess of derivatising reagent was required which distorted the chromatogram and made quantitative analyses difficult. To overcome this problem, other buffers such as 3-(N-morpholino)propanesulfonic acid (MOPS), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) and phosphate buffers were tested. The higher extent of hydroxylation was observed for MOPS by MALDI-MS (Figure 4.8) and hence this buffer was used in subsequent analyses.

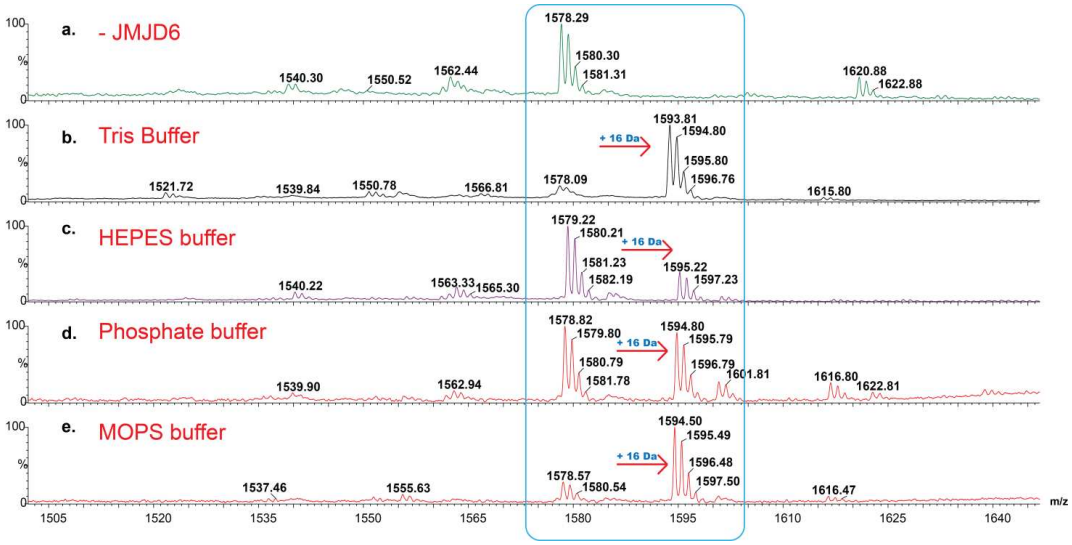


Figure 4.8 JMJD6 catalysed lysine hydroxylation of LUC7L2$_{267-278}$ peptide as carried out in various buffers and analysed by MALDI-MS.

4.6.1 HCl digestion of peptide after hydroxylation by JMJD6 and possible epimerisation of hydroxylysine

The hydroxylated LUC7L2$_{267\text{-}278}$ peptide sample was prepared by JMJD6 catalysis in 50 mM MOPS at pH 7.5 as described earlier in this section. In order to compare the stereochemistry from collagen peptides, cow bone collagen extracted from a femur, using the demineralisation and gelatinisation method of Longin,[24] was used. (*The sample was provided by Dr. James McCullagh*). HCl digestion of LUC7L2$_{267\text{-}278}$ peptide and collagen sample was carried out using 6 M HCl. Both the samples were analysed by LC-MS after dervatisation by AQC.

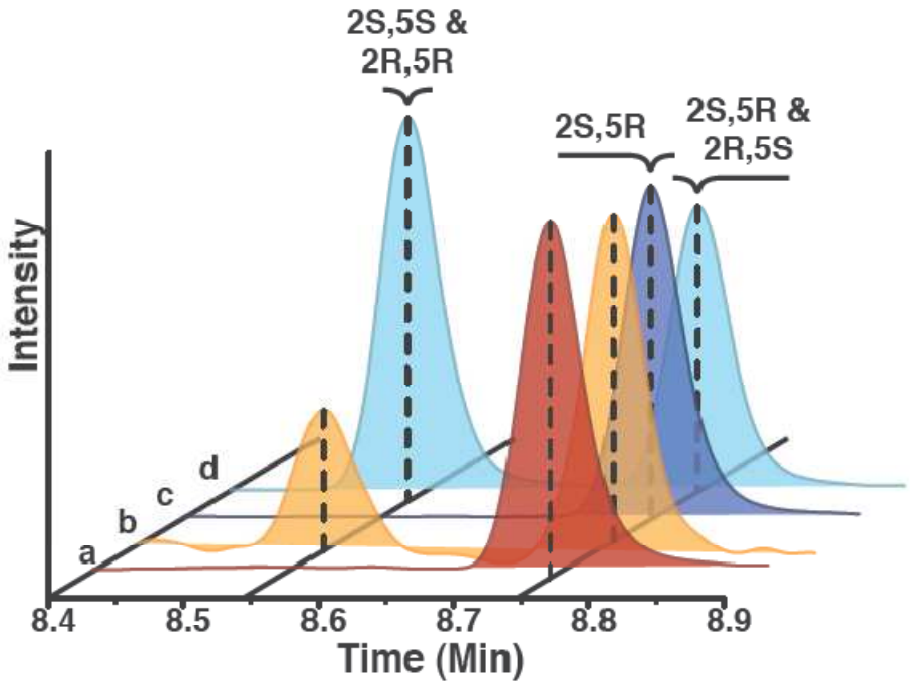


Figure 4.9. Amino acid analyses comparing the hydroxylysine products of JMJD6 catalysis with that in collagen. (a) Hydroxylysine product after treatment of collagen with trypsin and carboxypeptidase Y; (b) Collagen standard after HCl catalysed hydrolysis; (c) 2*S*,5*R*-Hydroxylysine; (d) Racemic mixture of 2*R*,5*R*/2*S*,5*S* and 2*R*,5*S*/2*S*,5*R*-Hydroxylysine

LC-MS analyses after HCl catalysed hydrolysis and derivatisation of both the JMJD6 product and collagen revealed peaks at both 8.54 min (corresponding to 2*S*,5*R* or 2*R*,5*S*

hydroxylysine) and 8.74 min (corresponding to 2*S*,5*S* or 2*R*,5*R* hydroxylysine) as shown in Figure 4.9. Although for the collagen sample the 2*S*,5*R*/2*R*,5*S* peak was larger, these results were not useful in stereochemical assignments. This is likely because under the hydrolysis acidic conditions, C-2 epimerisation of C-5 hydroxylysine occurs via formation of a 6-membered lactone involving stereochemically favoured 6-exo-trig cyclisation of the C-5 hydroxy group onto the protected C-2 carboxylic acid (Figure 4.10); lactone formation lowers the p*K*a of the C-2 hydrogen, enabling epimerisation via enolisation, as proposed by Anderson and Hamilton.[25]

4.6.2 Development of enzyme based digestion of peptide after JMJD6 hydroxylation

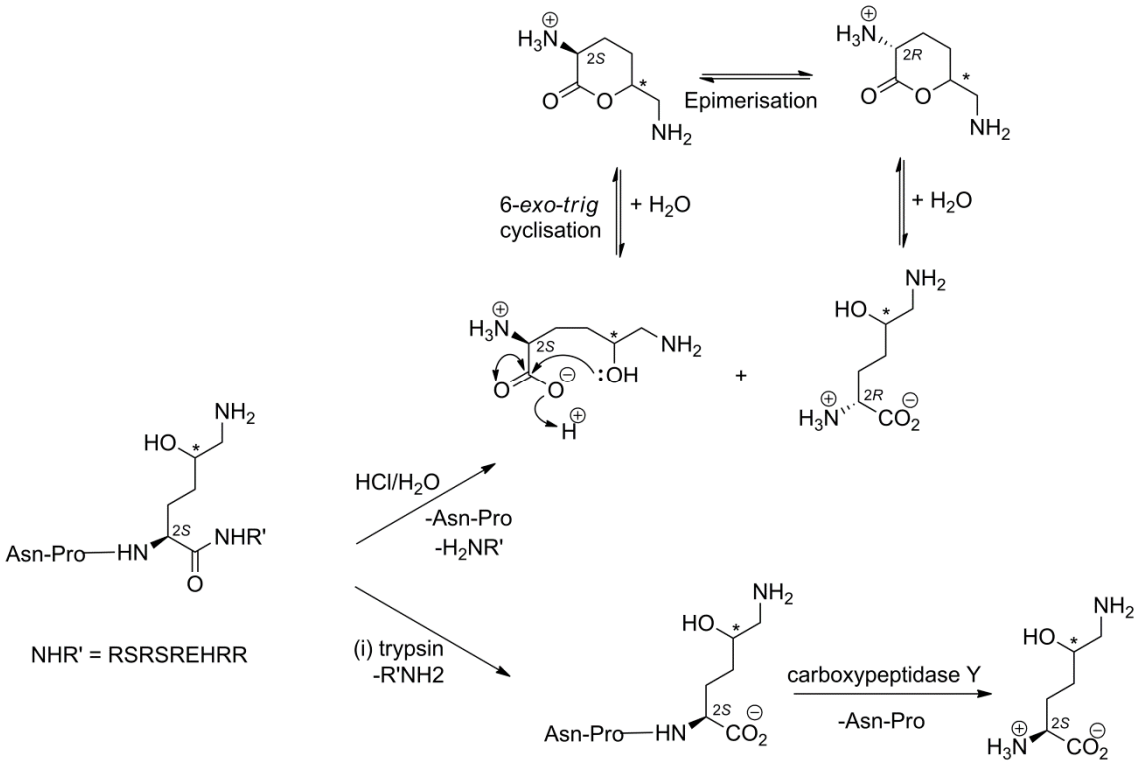


Figure 4.10 Outcomes of acid- and enzyme- catalysed hydrolysis of hydroxylysine containing peptides. Formation of 2*S*- and 2*R*-hydroxylysine products via epimerisation at C-2 in HCl catalysed hydrolysis[25] and formation of the product with retained C-2*S* stereochemistry when using an enzyme catalysed hydrolysis procedure.

In order to avoid/minimise the *C*-2 epimerisation, there was a need to develop a new hydrolysis protocol. An enzyme catalysed hydrolysis procedure involving treatment first with trypsin to cleave the amide bond *C*- terminal to the hydroxylysine, followed by treatment with a carboxypeptidase to release the hydroxylated lysine was developed as shown in Figure 4.10.

The $LUC7L2_{267-278}$ peptide was hydroxylated by JMJD6 in MOPS buffer. After testing that the reaction was complete by MALDI-MS, JMJD6 was precipitated by adding methanol. The sample was then centrifuged (10 min, 13000 rpm) and freeze dried. The assay mixtures were resuspended in 0.1% formic acid in water and purified by HPLC using Phenomenex Synergi Hydro RP C18 column (Phenomenex). The purified peptide was then suspended in 50 mM MOPS buffer and incubated with 2µl of trypsin (Promega, 100 µg ml^{-1},) for 2 hours at 37 ºC. The trypsin was removed by centrifugation (10min, 14000 rpm). Carboxypeptidase Y (Roche Applied science, 1 µl, 5 mg ml^{-1}) was then added to the supernatant and the mixture was incubated at 37ºC (overnight). The derivatisation of hydrolysate by AQC followed by LC-MS analysis was carried out as described previously.

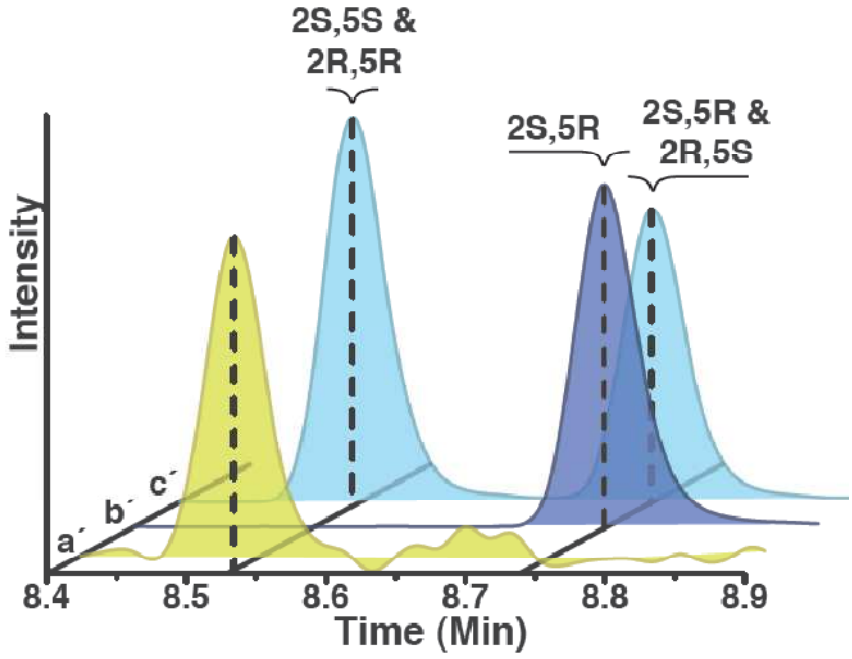


Figure 4.11. Amino acid analyses comparing the hydroxylysine products of JMJD6 catalysis with that in collagen. (a') Hydroxylysine product of JMJD6 catalysed LUC7L2$_{267-278}$ after treatment with trypsin and carboxypeptidase Y; (b') 2*S*,5*R*-Hydroxylysine; (c') Racemic mixture of 2*R*,5*R*/2*S*,5*S* and 2*R*,5*S*/2*S*,5*R*-hydroxylysine.

When treated in this manner, the collagen sample gave a single peak by amino acid analysis corresponding to the retention time (t = 8.74) for derivatised 2*S*,5*R* hydroxylysine as shown in Figure 4.9a; it is subsequently assumed that the 2*S* (*L*) stereochemistry of the substrate is maintained in the product. In contrast, and unexpectedly, when the hydroxylated LUC7L2$_{267-278}$ peptide was then subjected to the enzyme catalysed hydrolysis protocol, a single peak corresponding to 2*S*,5*S*-hydroxylysine at retention time 8.54 was observed (Figure 4.11a'). A set of controls ((i) without LUC7L2$_{267-278}$, (ii) without enzyme, (iii) without LUC7L2$_{267-278}$ and 2OG, (iv) without LUC7L2$_{267-278}$ and Fe(II), and (v) without LUC7L2$_{267-278}$ and without both 2OG and Fe(II)) was carried out (results of standards discussed in Chapter 5). Overall the analyses support the conclusion that JMJD6 catalyses hydroxylation to produce 2*S*,5*S* hydroxylysine.

4.7 Discussion

The work in this Chapter emphasises the assignment of stereochemistry of catalytic lysine hydroxylation carried out by JMJD6 and its comparison with collagen lysyl hydroxylases involved in post-translational modification of collagen and other extracellular matrix proteins. In conclusion, the results show that JMJD6 catalyses hydroxylation of lysyl residues to give the product with the 2*S*,5*S*- stereochemistry as shown in Figure 4.12, in contrast with the collagen lysyl hydroxylases which produce 2*S*,5*R*- products.

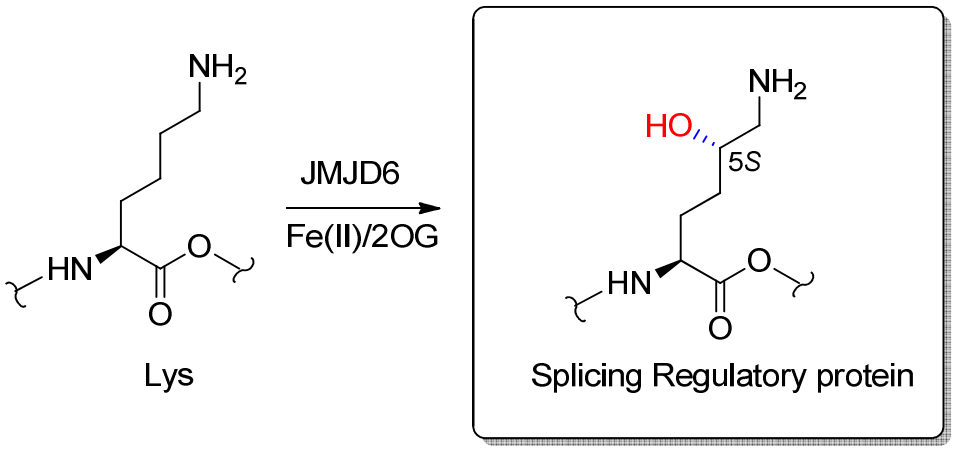


Figure 4.12 JMJD6 catalyses hydroxylation of lysine from its substrate peptide LUC7L2$_{267-278}$ to give product with 5*S* stereochemistry.

Additionally, the work carried out illustrates the application of amino acid analysis to assign the stereochemistry of amino acid modifications, thus emphasising the importance of proteomic analyses in stereochemical assignment. As a part of the work, a new enzymatic digestion protocol suitable for 5-hydroxylysine residues was developed in order to avoid the formation of a lactone, thereby conserving the stereochemistry at the C-2 position. This protocol can be further exploited to assign the stereochemistry of various 2OG oxygenases catalysed products.

Structurally informed bioinformatics and structural studies imply that JMJD6 is more closely related to the JmjC family of N^{ϵ}-methyl lysine demethylases and hydroxylases (FIH) than it is to the collagen hydroxylases.[26] Thus, whilst it seems likely that most animal 2OG oxygenases, if not all, evolved from a common ancestor, at least two different subfamilies of lysyl hydroxylases may have evolved with different stereochemical selectivities. It has been proposed that N^{ϵ}-methyl demethylases catalyse demethylation of histone tails via hydroxylation of the N^{ϵ}-methylated lysine followed by fragmentation of the N^{ϵ}-hydroxymethyllysine, to give formaldehyde and the demethylated lysine as shown in Figure 4.13.[10, 27] Given the proximity of the C-5 and N^{ϵ}-amino group of lysine, it is possible to envisage that N^{ϵ}-methyl demethylases are evolved from a C-5 hydroxylase or vice versa.

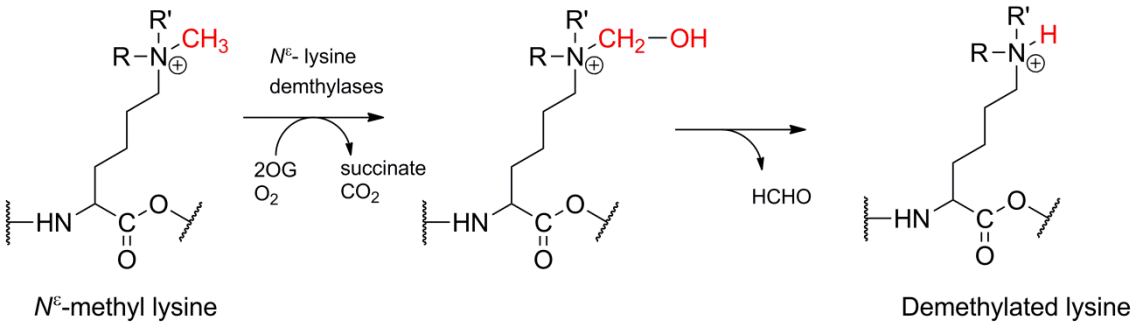


Figure 4.13 Stoichiometry of histone N^{ϵ}-methyl demethylases catalysed demethylation via hydroxylation of N^{ϵ}-methyl group.

2OG dependent oxygenases are currently of immense interest for development of small molecular inhibitors because of their roles in transcriptional regulation and oxygen sensing. This stereochemical assignment is also important for development of selective small molecule inhibitors for lysyl hydroxylases and related 2OG oxygenases for pharmaceutical and functional assignment purposes.

4.8 References

1. Loenarz, C.; Schofield, C. J., Expanding chemical biology of 2-oxoglutarate oxygenases. *Nature Chemical Biology* **2008,** *4* (3), 152-156.

2. Flashman, E.; Schofield, C. J., The most versatile of all reactive intermediates? *Nature Chemical Biology* **2007,** *3* (2), 86-87.

3. Webb, J.; Coleman, M.; Pugh, C., Hypoxia, hypoxia-inducible factors (HIF), HIF hydroxylases and oxygen sensing. *Cellular and Molecular Life Sciences* **2009,** *66* (22), 3539-3554.

4. Kaelin Jr, W. G.; Ratcliffe, P. J., Oxygen Sensing by Metazoans: The Central Role of the HIF Hydroxylase Pathway. *Molecular Cell* **2008,** *30* (4), 393-402.

5. Cockman, M. E.; Lancaster, D. E.; Stolze, I. P.; Hewitson, K. S.; McDonough, M. A.; Coleman, M. L.; Coles, C. H.; Yu, X.; Hay, R. T.; Ley, S. C.; Pugh, C. W.; Oldham, N. J.; Masson, N.; Schofield, C. J.; Ratcliffe, P. J., Posttranslational hydroxylation of ankyrin repeats in I*κ*B proteins by the hypoxia-inducible factor (HIF) asparaginyl hydroxylase, factor inhibiting HIF (FIH). *Proceedings of the National Academy of Sciences* **2006,** *103* (40), 14767-14772.

6. Coleman, M. L.; McDonough, M. A.; Hewitson, K. S.; Coles, C.; Mecinović, J.; Edelmann, M.; Cook, K. M.; Cockman, M. E.; Lancaster, D. E.; Kessler, B. M.; Oldham, N. J.; Ratcliffe, P. J.; Schofield, C. J., Asparaginyl Hydroxylation of the Notch Ankyrin Repeat Domain by Factor Inhibiting Hypoxia-inducible Factor. *Journal of Biological Chemistry* **2007,** *282* (33), 24027-24038.

7. Cockman, M. E.; Webb, J. D.; Ratcliffe, P. J., FIH-Dependent Asparaginyl Hydroxylation of Ankyrin Repeat Domain-Containing Proteins. *Annals of the New York Academy of Sciences* **2009,** *1177* (1), 9-18.

8. Xie, L.; Xiao, K.; Whalen, E. J.; Forrester, M. T.; Freeman, R. S.; Fong, G.; Gygi, S. P.; Lefkowitz, R. J.; Stamler, J. S., Oxygen-regulated β2-adrenergic receptor hydroxylation by EGLN3 and ubiquitylation by pVHL. *Science Signaling* **2009,** *2* (78), r33-r33.

9. Klose, R. J.; Kallin, E. M.; Zhang, Y., JmjC-domain-containing proteins and histone demethylation. *Nature Reviews Genetics* **2006,** *7* (9), 715-727.

10. Hopkinson, R. J.; Hamed, R. B.; Rose, N. R.; Claridge, T. D. W.; Schofield, C. J., Monitoring the Activity of 2-Oxoglutarate Dependent Histone Demethylases by NMR Spectroscopy: Direct Observation of Formaldehyde. *ChemBioChem* **2010,** *11* (4), 506-510.

11. Hodges, J. A.; Raines, R. T., Stereoelectronic and Steric Effects in the Collagen Triple Helix: Toward a Code for Strand Association. *Journal of the American Chemical Society* **2005,** *127* (45), 15923-15932.

12. Przysiecki, C. T.; Staggers, J. E.; Ramjit, H. G.; Musson, D. G.; Stern, A. M.; Bennett, C. D.; Friedman, P. A., Occurrence of beta-hydroxylated asparagine residues in non-vitamin K-dependent proteins containing epidermal growth factor-like domains. *Proceedings of the National Academy of Sciences of the United States of America* **1987,** *84* (22), 7856-7860.

13. McNeill, L. A.; Hewitson, K. S.; Claridge, T. D.; Seibel, J. F.; Horsfall, L. E.; Schofield, C. J., Hypoxia-inducible factor asparaginyl hydroxylase (FIH-1) catalyses hydroxylation at the β-carbon of asparagine-803. *Biochemical Journal* **2002,** *367* (3), 571-575.

14. Witkop, B., Interconversions and epimerizations of open and cyclic hydroxyamino acids. Building stones of collagen. *Chemical Society (London)* **1955,** *3*, 60-82.

15. Webby, C. J.; Wolf, A.; Gromak, N.; Dreger, M.; Kramer, H.; Kessler, B.; Nielsen, M. L.; Schmitz, C.; Butler, D. S.; Yates, J. R., III; Delahunty, C. M.; Hahn, P.; Lengeling, A.; Mann, M.; Proudfoot, N. J.; Schofield, C. J.; Bottger, A., Jmjd6 Catalyses Lysyl-Hydroxylation of U2AF65, a Protein Associated with RNA Splicing. *Science* **2009,** *325* (5936), 90-93.

16. Spetzler, J. C.; Hoeg-Jensen, T., Masked side-chain aldehyde amino acids for solid-phase synthesis and ligation. *Tetrahedron Letters* **2002,** *43* (12), 2303-2306.

17. Allevi, P.; Anastasia, M., A practical and simple synthesis of (2*S*,5*R*)- and (2*S*,5*S*)-5-hydroxylysine and of a related α-amino acid required for the synthesis of the collagen cross-link pyridinoline. *Tetrahedron: Asymmetry* **2004,** *15* (13), 2091-2096.

18. Rutherfurd, S. M.; Gilani, G. S., Amino Acid Analysis. In *Current Protocols in Protein Science*, John Wiley & Sons, Inc.: 2001.

19. Moore, S.; Stein, W. H., Photometric ninhydrin method for use in the chromatography of amino acids. *Journal of Biological Chemistry* **1948,** *176* (1), 367-388.

20. Moore, S.; Stein, W. H., Chromatography of amino acids on sulfonated polystyrene resins. *Journal of Biological Chemistry* **1951,** *192* (2), 663-681.

21. Schegg, K. M.; Denslow, N. D.; Andersen, T. T.; Bao, Y.; Cohen, S. A.; Mahrenholz, A. M.; Mann, K., Quantitation and Identification of Proteins by Amino Acid Analysis: ABRF-96AAA Collaborative Trial. In *Techniques in Protein Chemistry VIII*, Acdemic Press: San Diego, CA, 1997; pp 207-216.

22. Bidlingmeyer, B. A.; Cohen, S. A.; Tarvin, T. L., Rapid analysis of amino acids using pre-column derivatization. *Journal of Chromatography B: Biomedical Sciences and Applications* **1984,** *336* (1), 93-104.

23. Cohen, S. A.; Michaud, D. P.; D. Antonis, K., Compositional protein analysis using 6-aminoquinolyl-N-hydroxysuccinimidyl Carbamate, a Novel derivatizaion

reagent In *Techniques in Protein chemistry IV*, Academic Press, San Diego, CA: 1993; pp 289-298.

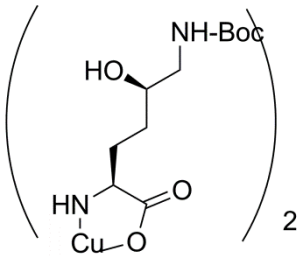
24. Longin, R., New Method of Collagen Extraction for Radiocarbon Dating. *Nature* **1971,** *230* (5291), 241-242.

25. Hamilton, P. B.; Anderson, R. A., Hydroxylysine: Isolation from gelatin and resolution of its diastereoisomers by Ion exchange chromatography. *Journal of Biological Chemistry* **1955,** *213* (1), 249-258.

26. Hahn, P.; Bose, J.; Edler, S.; Lengeling, A., Genomic structure and expression of Jmjd6 and evolutionary analysis in the context of related JmjC domain containing proteins. *BMC Genomics* **2008,** *9* (1), 293.

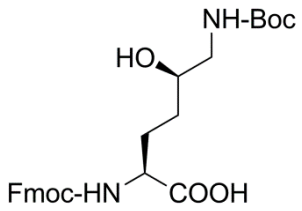
27. Ng, S. S.; Kavanagh, K. L.; McDonough, M. A.; Butler, D.; Pilka, E. S.; Lienard, B. M. R.; Bray, J. E.; Savitsky, P.; Gileadi, O.; von Delft, F.; Rose, N. R.; Offer, J.; Scheinost, J. C.; Borowski, T.; Sundstrom, M.; Schofield, C. J.; Oppermann, U., Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity. *Nature* **2007,** *448* (7149), 87-91.

Cu-Complex of Boc protected 2*S*,5*R*-hydroxy lysine:



2*S*,5*R*-hydroxyl lysine dihydrochloride (0.5 g, 1.97 mmol) was dissolved in 12 ml of water. NaOH (3 eq., 5.92 mmol) and copper sulphate (0.5 eq., 0.987 mmol) were added to resultant solution. The solution was then heated to 80°C and allowed to cool to room temperature. Boc anhydride (2 eq., 3.95 mmol) was added together with dioxane (15 ml). The resultant solution was then stirred at room temperature overnight. Additional Boc anhydride (0.5 eq, 0.98 mmol) was added and stirring was continued for two nights. The dark blue crystals formed were washed with water and ether. Crystals were dried over rotavapour and used in the crude form for the subsequent Fmoc protection step. Yield of crude product: 1.36 g, 75.5%.

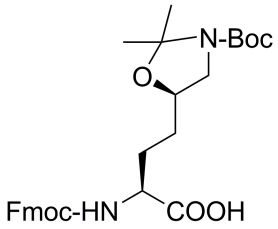
***N*-Fmoc-2*S*,5*R*-Hydroxy lysine:**



The crude copper complex (1 g), prepared as described above was dissolved in water (15 ml). EDTA-sodium salt (1.5 eq., 1.47 mmol.) and sodium hydroxide (3 eq., 2.94 mmol) were then added. After 2 hours, the reaction mixture was cooled at 0ºC, Fmoc-Osu (1.5 eq., 2.55 mmol) and dioxane (10 ml) were added. The pH was changed to 9 by

addition of Na_2CO_3. The reaction mixture was then stirred overnight with regular monitoring by TLC. Reaction mixture was extracted in ether (3X 50 ml), acidified to pH 2 with HCl. Then aqueous layer was extracted with (100 ml) ethyl acetate. The organic layer was then washed with brine, dried over $MgSO_4$ and concentrated using a rotavapour. The crude product was purified by flash column chromatography using 5% AcOH in chloroform: 5% AcOH in MeOH (19:1) (R_f = 0.73).Yield 0.400mg, 85.9%. White solid, Melting point 46°C. ESMS (*m/z*) : 506.8. δ_H: (DMSO, 400MHz) 1.25-1.45(4H, m, C*H*$_2$C*H*$_2$), 1.36 (9H, s, (C*H*$_3$)$_3$CO), 1.51-1.68 (2H, m, CH_2), 2.89 (2H, m, C*H*$_2$NH), 3.89 (1H, m, NH_2-C*H*-COOH), 4.21 (1H, t, J=6.5 Hz, C*H*CH_2O), 4.27 (1H, d, J=6.5 Hz, CHC*H*$_2$O), 6.77 (1H, t, J=6.0 Hz, N*H*CH_2), 7.31 (2H, t, J=7.5 Hz, Ar*H*), 7.41 (2H, t, J=7.5 Hz, Ar*H*), 7.72 (2H, d, J=8.0 Hz, Ar*H*), 7.88 (2H, d, J=7.5 Hz, ArH).

Hydroxy lysine (Boc-oxazolidine):



N-Fmoc-2*S*,5*R*-Hydroxy lysine was found to be unstable and undergoes rapid cyclisation to give a lactone. Therefore, the free hydroxyl group was protected by oxazolidine ring formation to the prevent cyclisation. Fmoc-2S,5R-Hydroxy lysine (0.5 g) was dissolved in pre-dried acetone (5 ml). The solution was then cooled using ice bath and treated with $BF_3.OEt_2$ (25 µl). The reaction was complete in half an hour as monitored by TLC. Reaction mixture work up was carried out by adding ethyl acetate and washing with water. The organic layer was washed with brine, dried over $MgSO_4$

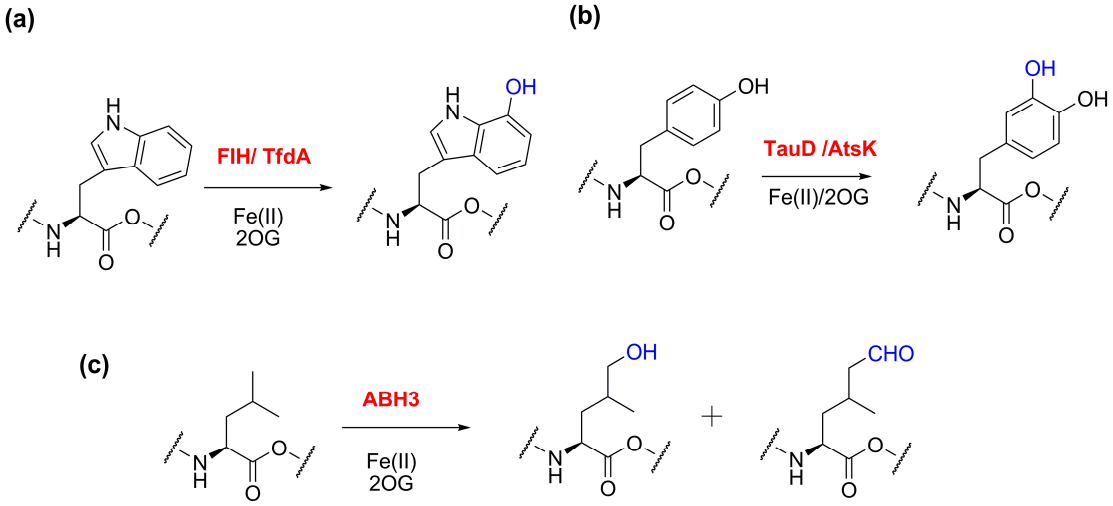
and concentrated in vacuo. The crude product was purified by flash column chromatography using 4% AcOH in hexane: 4% AcOH in THF (60:40). R_f = 0.45. Yield 0.452g, 90%. Melting point 102°C. ESMS (*m/z*) : 547.2 (M+Na), 425.2 (M+H-Boc). δ_H: ($CDCl_3$, 400MHz) : 1.21 (12H, s, C*H*$_3$ + (C*H*$_3$)$_3$CO), 1.23 (3H, s, C*H*$_3$), 1.47 (2H, m, CH_2C*H*$_2$CO), 1.66-1.92 (4H, m, C*H*$_2$C*H*$_2$CO), 2.07 (1H, m, C*H*N), 3.55 (1H, t, C*H*N), 3.98 (1H, m, C*H*O), 4.27 (2H, d, Ar-C*H*$_2$), 4.39 (2H, m, ArC*H* + NH-C*H*-COOH), 6.63 (1H, d, N*H*), 7.33 (2H, t, J=7.0 Hz, Ar*H*), 7.42 (2H, t, J=7.5 Hz, Ar*H*), 7.71 (2H, q, J= 7.5, 6.6 Hz, Ar*H*), 7.86 (2H, d, J= 7.5 Hz, Ar*H*). δ_c: ($CDCl_3$, 400MHz) 173.0, 156.3, 152.2, 144.4, 141.4, 127.8, 127.2, 125.5, 120.1, 93.7, 80.4, 73.7, 66.4, 53.9, 80.4, 73.7, 66.4, 53.9, 51.0, 47.3, 29.8, 28.8, 28.2, 27.9

Chapter 5. Self-hydroxylation of JMJD6

5.1 Introduction

Several members of the 2OG oxygenase family have been shown to catalyse self-hydroxylation in the absence of their substrates.[1-4] The closest human homologue of JMJD6 is FIH, which reduces the transcriptional activity of HIF by the post-translational asparaginyl hydroxylation in the *C*-terminal transcriptional activation domain (CAD).[5] FIH has been reported to undergo self-hydroxylation of an active site tryptophan residue in the absence of HIF.[1] Other members of the 2OG oxygenase family, taurine/α-ketoglutarate dioxygenase (TauD) and 2,4-dichlorophenoxyacetate oxygenase (TfdA), are apparently deactivated by hydroxylation of an active site tyrosine (TauD)[2] and tryptophan (TfdA)[3] residues in the absence of their prime substrates, taurine or 2,4-dichlorophenoxyacetate, respectively. In these cases, the self-hydroxylation has been shown to be Fe(II) and 2OG dependent. In the case of human AlkB homolog hABH3 (as well as AlkB, hABH2 and hABH6), there is an evidence that a leucine residue ($Leu177_{hABH3}$) undergoes oxidation to give an alcohol which is further oxidised to an aldehyde at δ- position (Scheme 5.1).[4] A MS-based study on oxidative modifications of a viral collagen prolyl- hydroxylase has revealed low-level hydroxylation of the residues in its protein sequence.[6] In the crystal structure analysis of the alkylsulfatase, (Atsk) as AtsK-Fe-succinate complex, a side chain of Tyr-168 was observed to be coordinate directly to the Fe(II) in the active site.[7] Oxidative cleavage has also been reported in case of the 1-aminocyclopropane-1-carboxylate (ACC) oxidase, the ethylene forming enzyme in plants which is closely related to 2OG

oxygenases.[8] The aim of the work discussed in this chapter was to investigate the occurrence and extent of such self-hydroxylation in the case of JMJD6.



Scheme 5.1 Self-hydroxylation products of various 2OG oxygenases.

5.2 Uncoupled 2OG turnover by JMJD6

5.2.1 Radioactive 2OG turnover assays

To investigate uncoupled turnover of 2OG by JMJD6, assays measuring the conversion of [1-^{14}C]-2OG into $^{14}CO_2$,[9] as discussed in Chapter 3 were employed. 2OG turnover was monitored in the presence and absence of peptide substrate LUC7L2$_{267\text{-}278,}$ which is a fragment of splicing regulatory protein, LUC7like2.[10] Under standard assay conditions it was found that the amount of 2OG consumed was 1.7 fold higher in the presence of LUC7L2$_{267\text{-}278}$ than in its absence (Figure 5.1), suggesting that uncoupled turnover by JMJD6 might be relatively high.

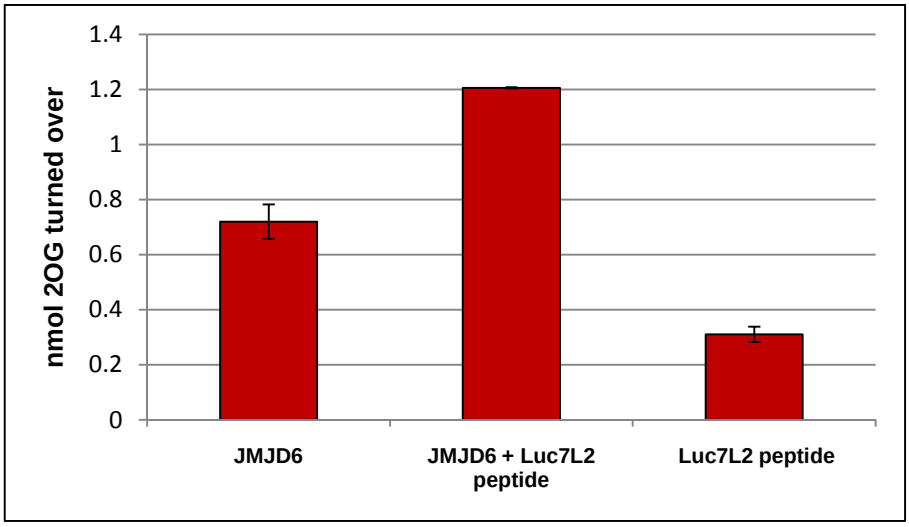


Figure 5.1 Results of [1-^{14}C]-turnover assay by JMJD6 under standard assay conditions with Fe(II), 2OG and LUC7L2$_{267-278}$ peptide.

5.2.2 2OG turnover of JMJD6 as monitored by ^{1}H-NMR analysis

2OG turnover by JMJD6 in the absence of primary substrate was also investigated by ^{1}H NMR analyses. Recombinant JMJD6 alone was incubated with Fe(II) and 2OG at 37°C and the reaction was monitored by Bruker AVIII 700 spectrometer. The assay mixture consisted of 2OG (500 µM) and ferrous ammonium sulphate (100 µM) in 50 mM deuteriated Tris buffer (pH 7.5). A solution of deuteriated Tris (50 mM) was made by dissolving deuteriated Tris in deuterium oxide and adjusting the pH to 7.5 with deuteriated HCl. 2OG (5 mM) and ascorbic acid (1 mM) solutions were made in 50 mM deuteriated Tris (pH 7.5). $FeNH_4SO_4$ was dissolved in 20 mM deuteriated HCl and then diluted to requisite concentration in deuterium oxide. The reaction was initiated by addition of JMJD6 (20 µM) to a 1 mm microprobe NMR tube. ^{1}H NMR spectra were recorded at 37ºC for 48 measurements over 70 minutes at 89 seconds intervals. Figure 5.2 shows the decrease in signal of 2-oxoglutarate (triplet at 2.42 ppm and 2.97 ppm)

and continuous increase in the succinate peak (singlet) at 2.39 ppm, indicating continuous conversion of 2OG to succinate by JMJD6.

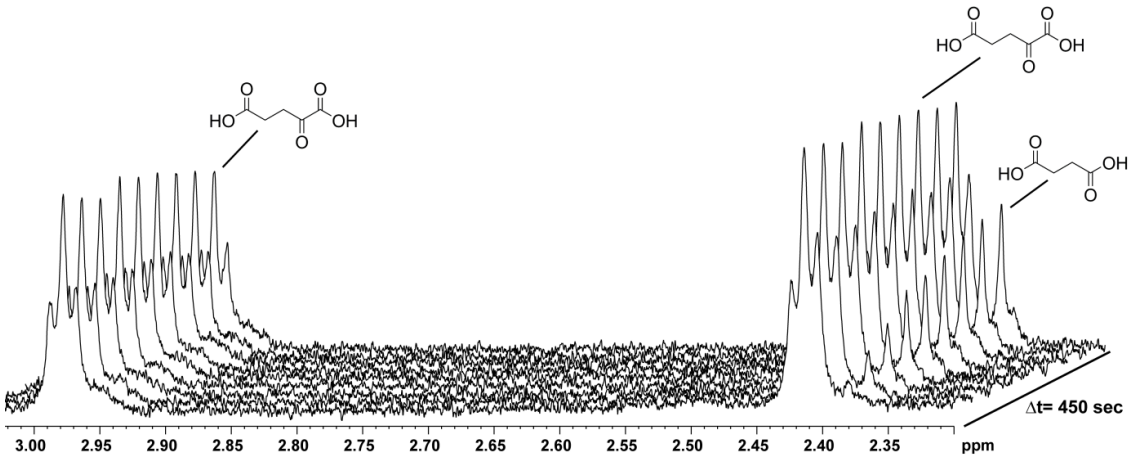


Figure 5.2 JMJD6 catalysed 2OG turnover in the absence of substrate as observed by ^{1}H-NMR.

The 2OG oxygenase substrate "uncoupled" turnover of 2OG to succinate has long been known.[11] The extent of uncoupled 2OG turnover varies between enzymes. In some but not all the cases, coupled turnover of 2OG is proportional to the oxidation of ascorbate e.g. the collagen prolyl-4-hydroxylases.[11-13] One role for ascorbate in promoting 2OG catalysis *in vitro* (though there are likely others)[12, 14] is proposed to be to reduce high (e.g. ferryl) oxidation state intermediates arising from uncoupled turnover, thus returning the enzyme to an active form.[15] However, evidence for specific requirement of ascorbate *in vivo* is lacking. Thus, 2OG turnover by ^{1}H NMR in the presence of ascorbate was monitored. However, a significant increase in succinate peak was not observed by addition of ascorbate to the assay mixture indicating that the 2OG turnover by JMJD6 in the absence of its primary substrate is independent of ascorbate oxidation.

5.2.3 MALDI-MS assays

The effect of 2OG turnover was seen when JMJD6 was pre-incubated with Fe(II) and 2OG for 30 minutes prior to the addition of substrate ($LUC7L2_{267-278}$) by MALDI-MS assays. JMJD6 20 µM was incubated with 500 µM 2OG and 400 µM Fe(II). After 30 minutes, 100 µM of $LUC7L2_{267-278}$ was added and assay mixture was incubated at 37°C. Standard assay condition in which JMJD6 and $LUC7L2_{267-278}$ were mixed at the same time with Fe(II) and 2OG without any prior incubation was also carried out. Both the samples were spotted on MALDI plate with CHCA matrix and MS were recorded in positive electron mode. The results shown in Figure 5.3 suggest that when JMJD6 is pre-incubated with 2OG, Fe(II) for 30 minutes prior to the addition of $LUC7L2_{267-278}$, a 40-50% decrease in substrate hydroxylation was observed in comparison to when all the components were mixed simultaneously.

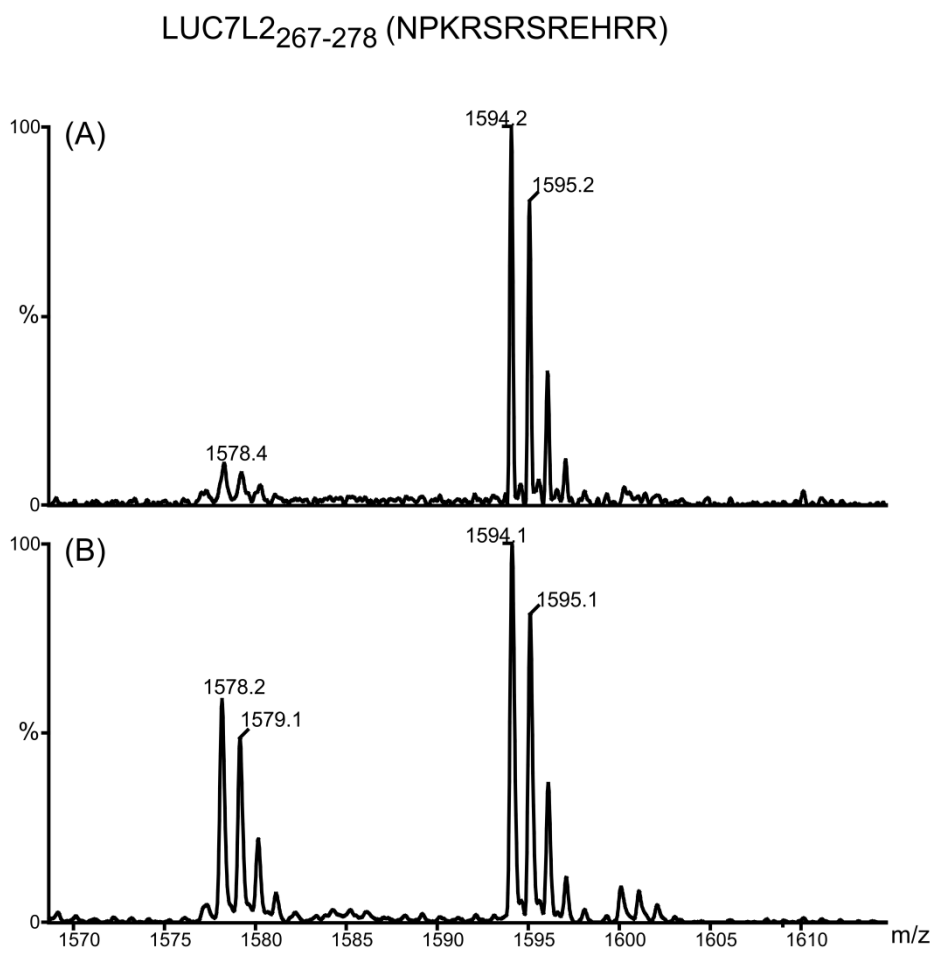


Figure 5.3 JMJD6-catalysed hydroxylation of LUC7L2$_{267-278}$ peptide is significantly decreased when JMJD6 is pre-incubated with 2OG and Fe(II). (A) MALDI-MS analysis of hydroxylation of LUC7L2$_{267-278}$ in the presence of Fe(II), 2OG and JMJD6 showing almost 90% hydroxylation. (B) Pre-incubating JMJD6 with 2OG and Fe(II) for 30 minutes prior to the addition of LUC7L2$_{267-278}$ results in a decrease in the level of peptide hydroxylation.

5.3 Amino acid analysis of JMJD6

To investigate the extent of JMJD6 self-hydroxylation, amino acid analysis was carried out using recombinant JMJD6 that had been incubated with or without Fe(II) and/or 2OG. For amino acid analyses, the reactions were initiated by mixing JMJD6 (20 μM), 2OG (500 μM), Fe(II) (400 μM) and ascorbic acid (100 μM) in MOPS buffer (50 mM, pH 7.5) followed by incubation at 37°C for 1 hour. The assay mixture was treated first

with trypsin and then with carboxypeptidase Y in an enzyme based hydrolysis procedure to give rise to mixture of peptides. The enzymatic hydrolysis procedure followed was the one which had been previously developed for analysis of peptide LUC7L2$_{267\text{-}278}$ as explained in Chapter 4, i.e. trypsin and carboxypeptidase Y were used rather than acid hydrolysis in order to avoid isomerisation of the hydroxylysine product. The mixture of peptides formed after hydrolysis were then derivatised by AQC in acetonitrile according to the AccQTag Solution Protocol (Waters) and analysed by LC-MS as reported.[16]

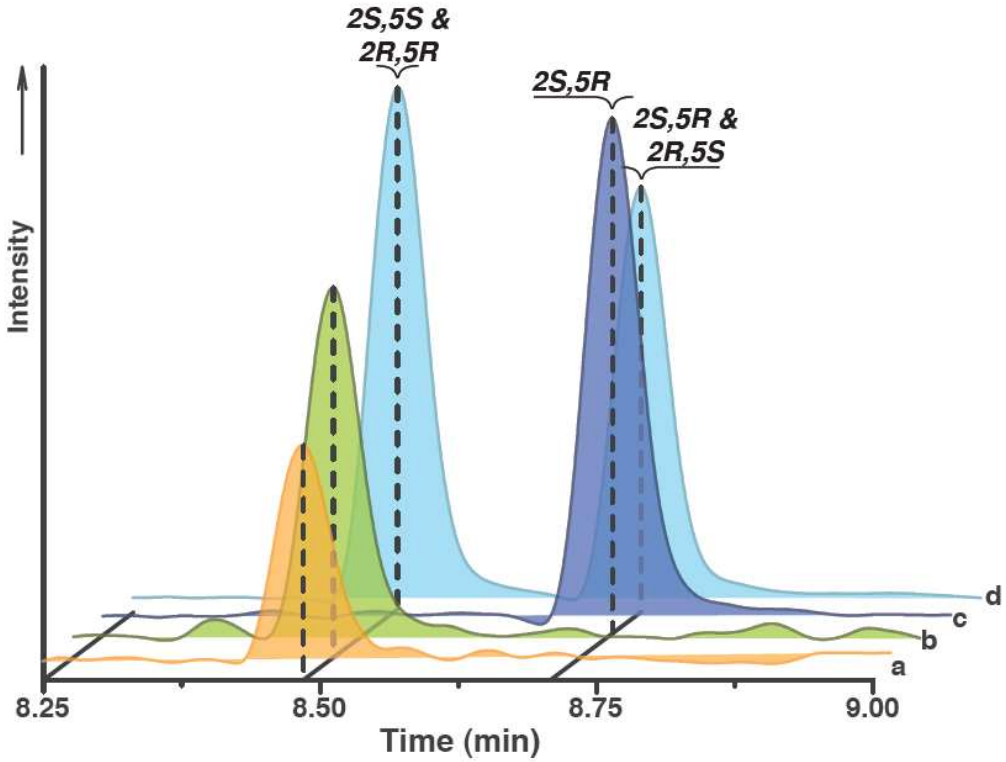
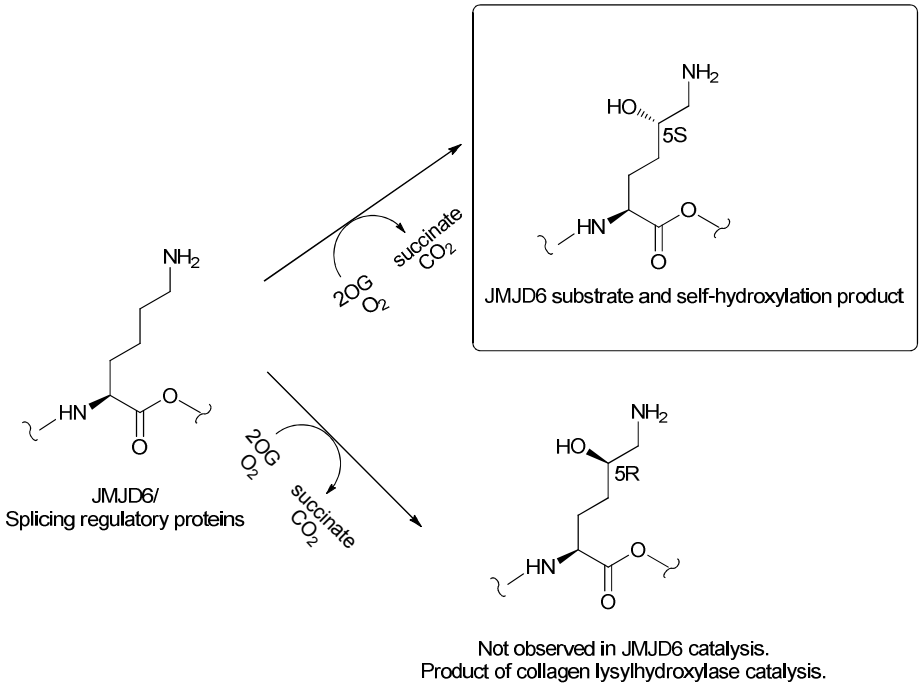


Figure 5.4 Amino acid analyses of JMJD6 catalysis. After incubation, the peptide product was hydrolysed using trypsin and carboxypeptidase Y. (a) Hydroxylsyine product obtained after hydrolysis of LUC7L2$_{267\text{-}278}$ peptide that had been incubated with JMJD6, Fe(II) and 2OG; (b) Hydroxylsyine product obtained after hydrolysis of JMJD6 that had been incubated with Fe(II) and 2OG; (c) 2*S*,5*R*-Hydroxylysine standard; (d) Mixture of 2*R*,5*R*/2*S*,5*S*- and 2*R*,5*S*/2*S*,5*R*-hydroxylysine standards.

The LC-MS analyses showed a peak at 8.50 corresponding to 2*S*,5*S* hydroxylysine (Figure 5.4(b) and Figure 5.5(d)) which is also observed for JMJD6 catalysed $LUC7L2_{267-278}$ hydroxylation as reported in Chapter 4. The lack of evidence for 5*R*-lysyl hydroxylation as shown in Scheme 5.2 supports the proposal that the self-hydroxylation of lysines in JMJD6 occurs at the active site in a manner similar to splicing protein hydroxylation. This result reveals that JMJD6 undergoes self-hydroxylation on one or more lysine residues giving rise to hydroxylysine product with stereochemistry 2*S*,5*S*, i.e. the same as that observed for JMJD6 catalysed splicing regulatory proteins.



Scheme 5.2 JMJD6 catalysed self-hydroxylation gives the 2*S*,5*S*-hydroxylysine product as observed for hydroxylation of splicing regulatory protein substrates.

Notably, it was found that as purified, the recombinant JMJD6 had already undergone a degree of self-hydroxylation, presumably in *E. coli* cells (Figure 5.4(a) and Figure 5.5(a)). The extent of hydroxylation was not substantially increased by incubation with

Fe(II) (Figure 5.5(b)), likely reflecting co-purification of JMJD6 with Fe(II). However, the extent of hydroxylation was increased substantially by the addition of 2OG (Figure 5.5(c)), indicating that self-hydroxylation of lysine residues is 2OG dependent.

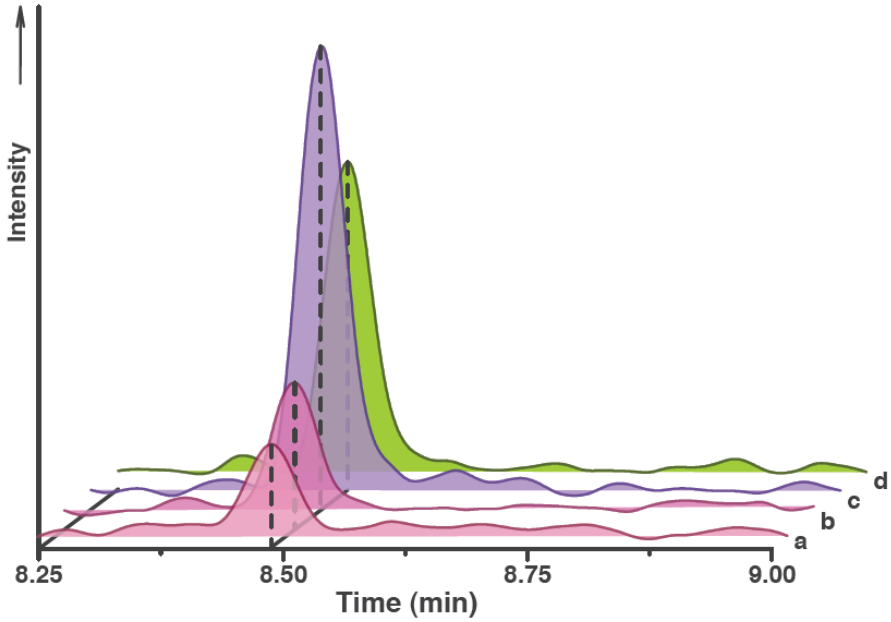


Figure 5.5 Set of standards analysed by amino acid analysis. (a) Hydroxylysine product from JMJD6 as purified after hydrolysis, i.e. without Fe(II) and 2OG incubation; (b) Hydroxylysine product obtained after incubation of JMJD6 incubated with Fe(II); (c) Hydrolysis product obtained after incubation of JMJD6 with 2OG; (d) Hydroxylysine product obtained after incubation of JMJD6 with Fe(II) and 2OG.

5.4 Proteomic studies to identify hydroxylated lysine residues in JMJD6

The sites of lysine hydroxylation in JMJD6 were identified by carrying out proteomic studies on recombinant JMJD6 expressed from *E. coli*., endogenous JMJD6 extracted from HeLa cells and a set of synthesised peptides from JMJD6 sequence.

The LysC protease protocol was optimised by Dr. Celia Webby. Immunoprecitation of JMJD6 from HeLa cells was carried out by Dr. Alexander Wolf at Prof. Angelika

Bottger's laboratory in Ludwig-Maximilians-University (Germany). LC-MS/MS analyses were carried out by Dr. Michael Nielsen at Mathias Mann's laboratory in Max-Planck Institute of Biochemistry (Germany).

5.4.1 Determination of lysine hydroxylation sites in recombinant JMJD6

In order to investigate the sites of lysine hydroxylation in recombinant JMJD6, it was first incubated with Fe(II) and 2OG at 37°C. The incubation mixture was then subjected to LysC protease (Roche) digestion using a standard protocol described in reference[17]and later analysed by LC-MS/MS.[10] These MS analyses support hydroxylation of $Lys111_{JMJD6}$ and $Lys167_{JMJD6}$ (Figure 5.6). No hydroxylation at $Lys111_{JMJD6}$ and $Lys167_{JMJD6}$ was observed in the absence of 2OG and Fe(II). Though the possibility of hydroxylations at other sites cannot be ruled out because the proteomic analyses were limited (at least for confident assignment) to residues Lys91, Lys100, Lys111, Lys113, Lys115, Lys142, Lys151, Lys154, Lys167, Lys219, Lys317 and Lys397. Indeed the observation that hydroxylysine is present in recombinant JMJD6 as determined by amino acid analysis reveal that this is the case (Figure 5.5(d)).

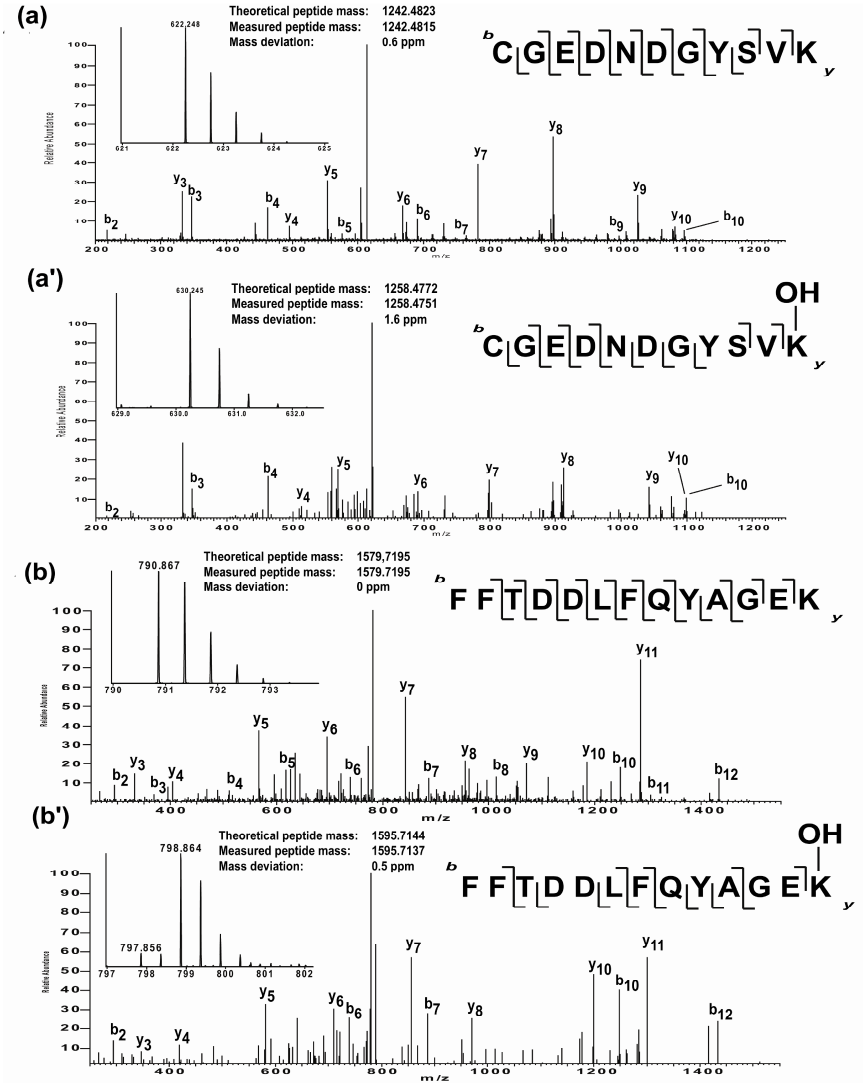


Figure 5.6 MS/MS spectra showing hydroxylation of recombinant JMJD6 in the presence of 2OG and Fe(II). LC-MS/MS supports hydroxylation at $Lys111_{JMJD6}$ (a') and $Lys167_{JMJD6}$ (b') of recombinant JMJD6 in the presence of Fe(II) and 2OG. Hydroxylation was not observed to be complete and the spectra corresponding to unhydroxylated $Lys111_{JMJD6}$ (a) and $Lys167_{JMJD6}$ (b) are also shown. The MH^{2+} peptide that was fragmented to give the resulting MS spectra is shown on the right hand side of each spectrum.

Recombinant JMJD6 (20 µM) was also incubated with Fe(II) (400 µM) and 2OG (500 µM) along with recombinant U2AF65 (40 µM), its primary substrate. After two hours incubation at 37°C, the JMJD6 was separated by SDS-PAGE. The protein band corresponding to the molecular weight of JMJD6 was excised, then digested with LysC

protease and analysed by LC-MS/MS. Interestingly even in this case, hydroxylation at $Lys167_{JMJD6}$ was also observed to occur (Figure 5.7).

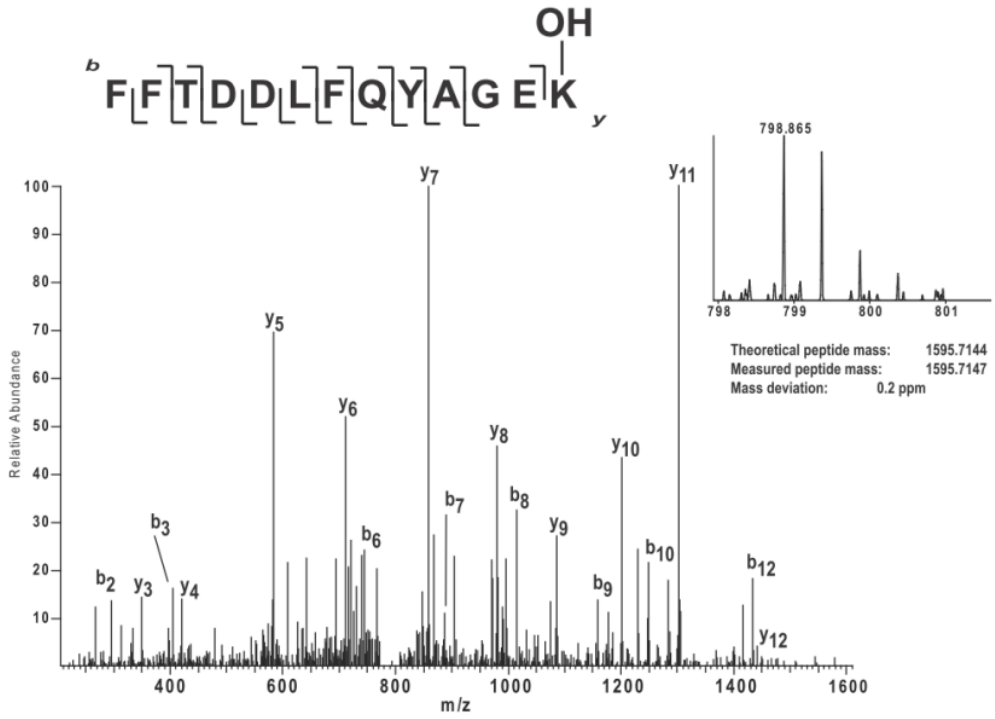


Figure 5.7 $Lys167_{JMJD6}$ of JMJD6 is hydroxylated in the presence of recombinant U2AF65 protein. LC-MS/MS analyses support hydroxylation of $Lys167_{JMJD6}$ of JMJD6 in the presence of 2OG, Fe(II) and U2AF65.

5.4.2 Determination of lysine hydroxylation sites in JMJD6 from HeLa cells

Finally, self-hydroxylation of endogenous JMJD6 in human cells was investigated. JMJD6 was extracted from HeLa (human) cells by immunoprecipitation with polyclonal anti-JMJD6 antibody (Abcam 10526). Extracts from 4x 10^8 cells were prepared in lysis buffer (20 mM Tris/HCl pH 8.0, 500 mM NaCl, 0.5% NP40 supplemented with protease (Pefabloc (Boehringer), pepstatin A, Aprotinin, Leupeptin) and phosphatase inhibitors (100 nM okadaic acid and phosphatase inhibitor cocktails 1 + 2 (Sigma)).

After centrifugation supernatants were incubated with 30 µl of the anti-JMJD6 antibody for 1 hour under agitation at 4°C. 60 µl of Protein-G Sepharose 4 Fast Flow (GE Healthcare, NJ, USA) was added for another hour. After three wash steps with 20 mM Tris.HCl pH 8.0, 500 mM NaCl, the beads were re-suspended in SDS sample buffer and subjected to SDS-PAGE (12% acrylamide). JMJD6 was then digested with LysC protease and analysed by LC-MS/MS. The fragment ion series also showed hydroxylation of $Lys167_{JMJD6}$ (Figure 5.8). As mentioned earlier, the possibility of hydroxylation at other sites cannot be ruled out.

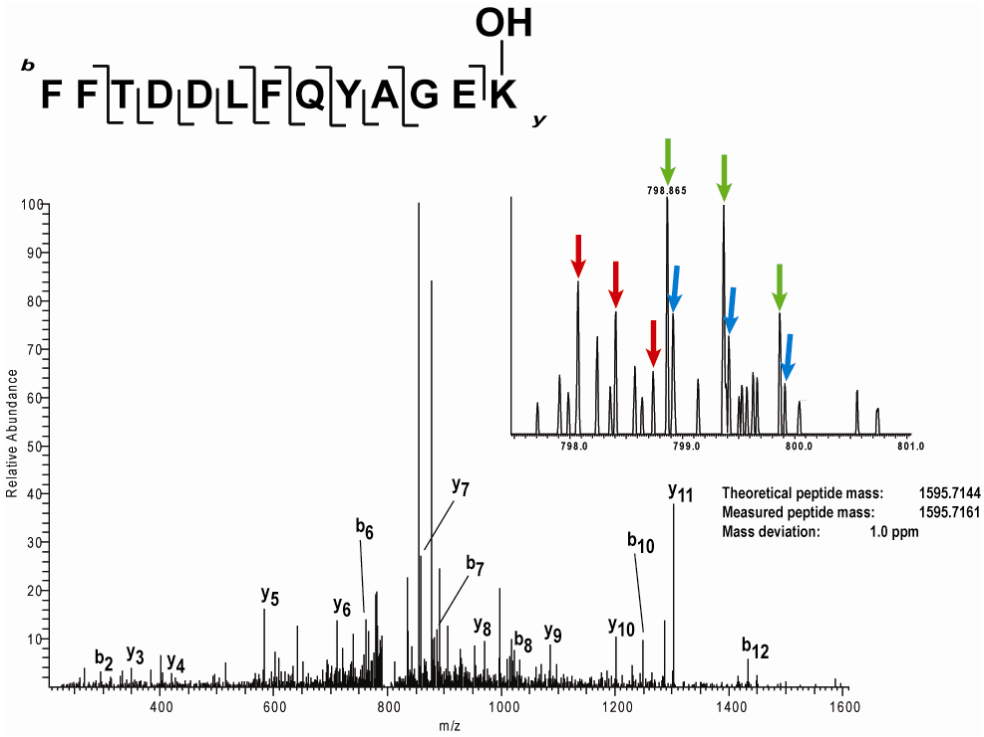


Figure 5.8 MS/MS analysis of endogenous JMJD6 supports hydroxylation of $Lys167_{JMJD6}$. JMJD6 was immunoprecipitated from HeLa cells, digested with LysC protease and analysed by LC-MS/MS. The MH^{2+} peptide that was fragmented to give the resulting MS spectra is shown on the right hand side of the spectrum (in green). The co-elution of two different peptides is shown by use of blue and red colours.

5.5 Determination of lysine hydroxylation sites by MALDI-MS assays using JMJD6 peptide substrates

A set of peptides as shown in Table 5.1, containing the identified hydroxylation sites as well as other lysine residues from the NLS were synthesised by solid phase peptide synthesis. All peptides were incubated with JMJD6 in the presence of Fe(II) and 2OG under standard assay conditions and then analysed by MALDI-MS. Out of the 10 peptides analysed only two peptides, corresponding to $JMJD6_{1\text{-}14}$ and $JMJD6_{301\text{-}314}$ were substrates as observed by MALDI-MS analysis (Figure 5.9). Notably the two peptides, $JMJD6_{105\text{-}120}$, $JMJD6_{167\text{-}176}$ containing the sites of hydroxylation identified in JMJD6 protein ($Lys111_{JMJD6}$ and $Lys167_{JMJD6}$) were not observed to be substrates (Figure 5.8). Although we cannot be certain, it is possible that hydroxylation of lysines corresponding to $JMJD6_{1\text{-}14}$ and $JMJD6_{301\text{-}314}$ observed in peptide work, may reflect the hydroxylation product of recombinant JMJD6 as obtained by amino acid analysis.

Table 5.1 Results of MALDI-MS analyses of peptides synthesised containing possible lysine hydroxylation sites within JMJD6

No.	JMJD6 Sequence	Peptide sequence	Molecular weight	+16 Da Shift in MS analysis
36	$JMJD6_{1\text{-}14}$	MNHKSKKRIREAKR	1782.1	YES
37	$JMJD6_{14\text{-}30}$	RSARPELKDSLDWTRHN	2081.2	NO
38	$JMJD6_{78\text{-}95}$	GWSAQEKWTLERLKRKYR	2335.7	NO
39	$JMJD6_{86\text{-}104}$	TLERLKRKYRNQKFKCGED	2412.8	NO
40	$JMJD6_{105\text{-}120}$	NDGYSVKMKMKYYIEY	2032.3	NO

41	$JMJD6_{135-150}$	SSYGEHPKRRKLLEDY	1978.1	NO
42	$JMJD6_{160-176}$	LFQYAGEKRRPPYRWFV	2213.5	NO
43	$JMJD6_{301-314}$	TVRGRPKLSRKWYR	1803.1	YES
44	$JMJD6_{363-379}$	SGSEGDGTVHRRKKRRT	1921.1	NO
45	$JMJD6_{363-381}$	SGSEGDGTVHRRKKRRTCS	2117.3	NO

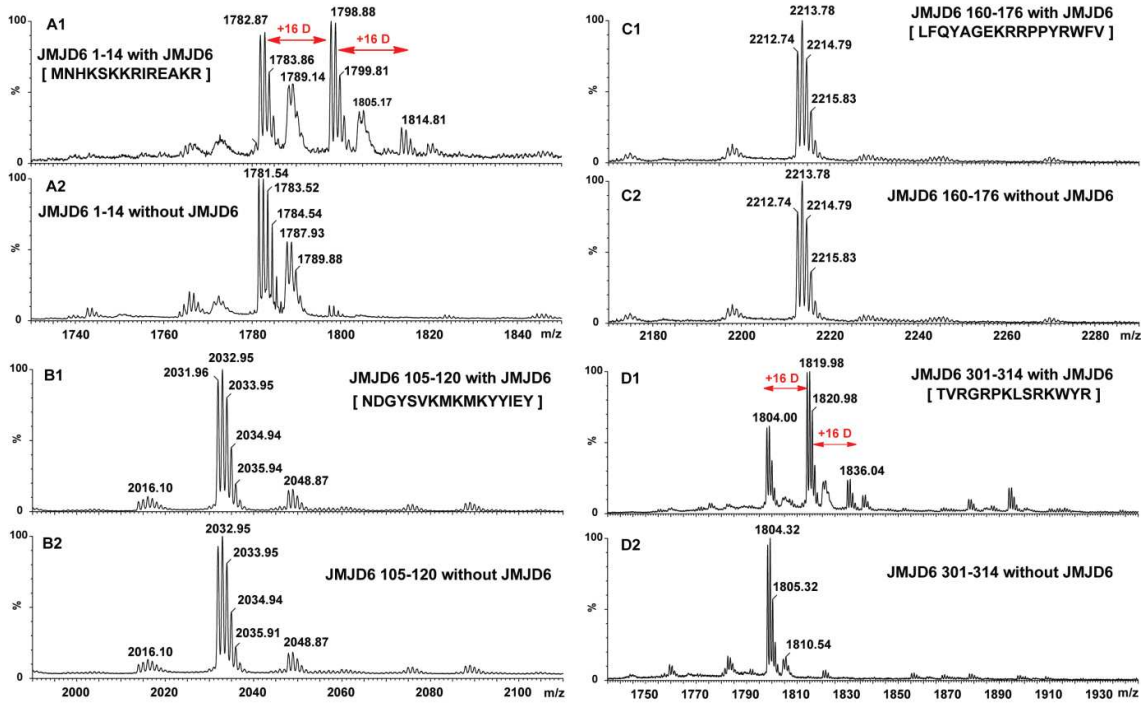


Figure 5.8 MALDI-TOF analysis of peptides synthesised for identifying possible lysine hydroxylation sites. Peptides were incubated with and without JMJD6, Fe(II) and 2OG. (A1) $JMJD6_{1-14}$ with JMJD6 (A2) $JMJD6_{1-14}$ without JMJD6 (B1) $JMJD6_{105-120}$ with JMJD6 (B2) $JMJD6_{105-120}$ without JMJD6 (C1) $JMJD6_{160-176}$ with JMJD6 (C2) $JMJD6_{160-176}$ without JMJD6 (D1) $JMJD6_{301-314}$ with JMJD6 (D2) $JMJD6_{301-314}$ without JMJD6.

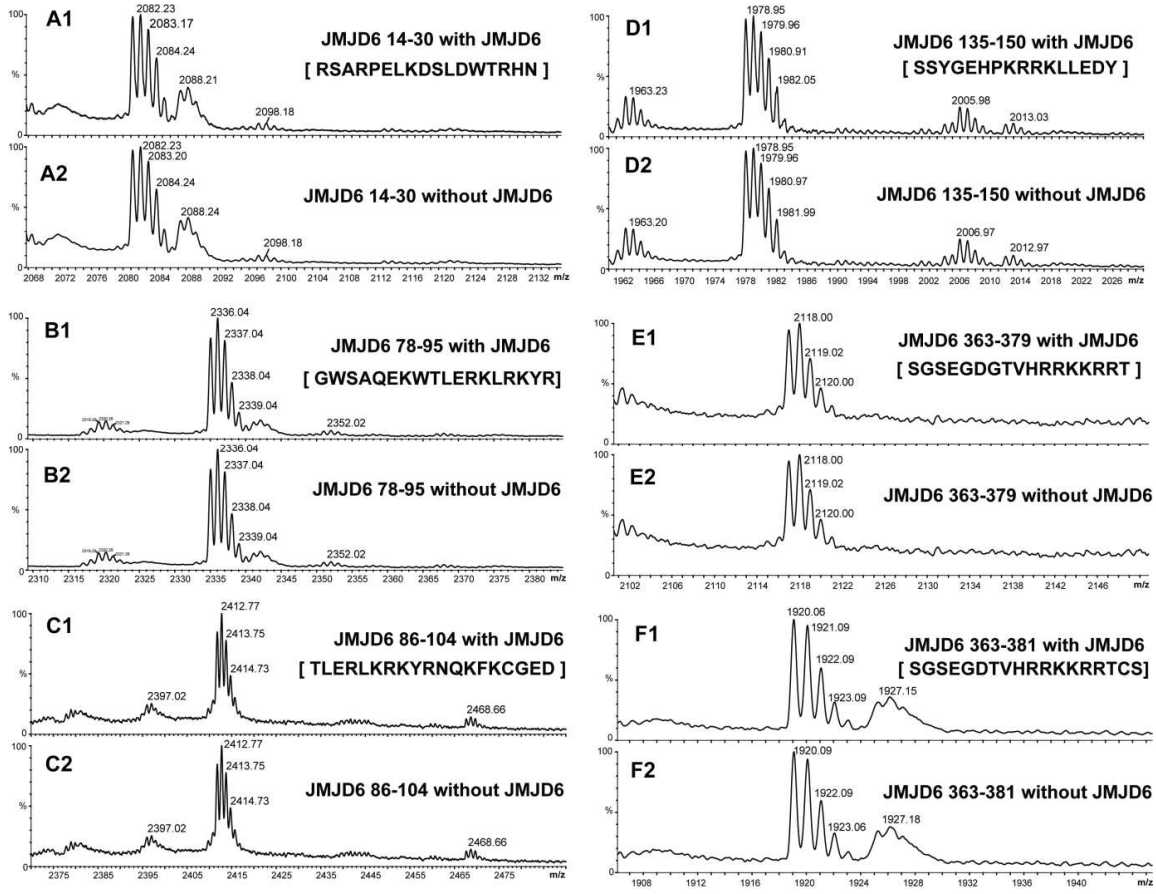


Figure 5.9 MALDI-TOF analysis of peptides synthesised for identifying possible lysine hydroxylation sites. All the peptides were incubated with and without JMJD6, Fe(II) and 2OG. (A1) and (A2) $JMJD6_{14\text{-}30}$ on incubation with and JMJD6 without JMJD6 (B1) and (B2) $JMJD6_{78\text{-}95}$ on incubation with and without JMJD6 (C1) and (C2) $JMJD6_{86\text{-}104}$ on incubation with and without JMJD6 (D1) and (D2) $JMJD6_{135\text{-}150}$ on incubation with and without JMJD6 (E1) and (E2) $JMJD6_{363\text{-}379}$ on incubation with and without JMJD6 (F1) and (F2) $JMJD6_{363\text{-}381}$ on incubation with and without JMJD6.

5.6 Discussion

5.6.1 Active site of JMJD6 and sites of self-hydroxylation

Peptides containing the hydroxylation sites $Lys111_{JMJD6}$ and $Lys167_{JMJD6}$ identified in JMJD6 recombinant protein were not observed to be substrates under the current assay conditions. Instead, two other peptides (corresponding to $JMJD6_{1-14}$ and $JMJD6_{301-314}$) that were not observed to be hydroxylated in the JMJD6 protein were observed to be substrates. Together these results reveal that factors other than the immediate sequence around the lysyl hydroxylation site are important in determining selectivity as has been found for FIH,[18] where factors relevant include secondary structure and overall protein stability.[19] It should also be noted that for the self-hydroxylation of JMJD6, it is unclear if the reaction is intra- or inter-molecular (for either of the $Lys111_{JMJD6}$, $Lys167_{JMJD6}$ or other hydroxylation sites), and the available JMJD6 crystal structures[20, 21] do not rule out either possibility.

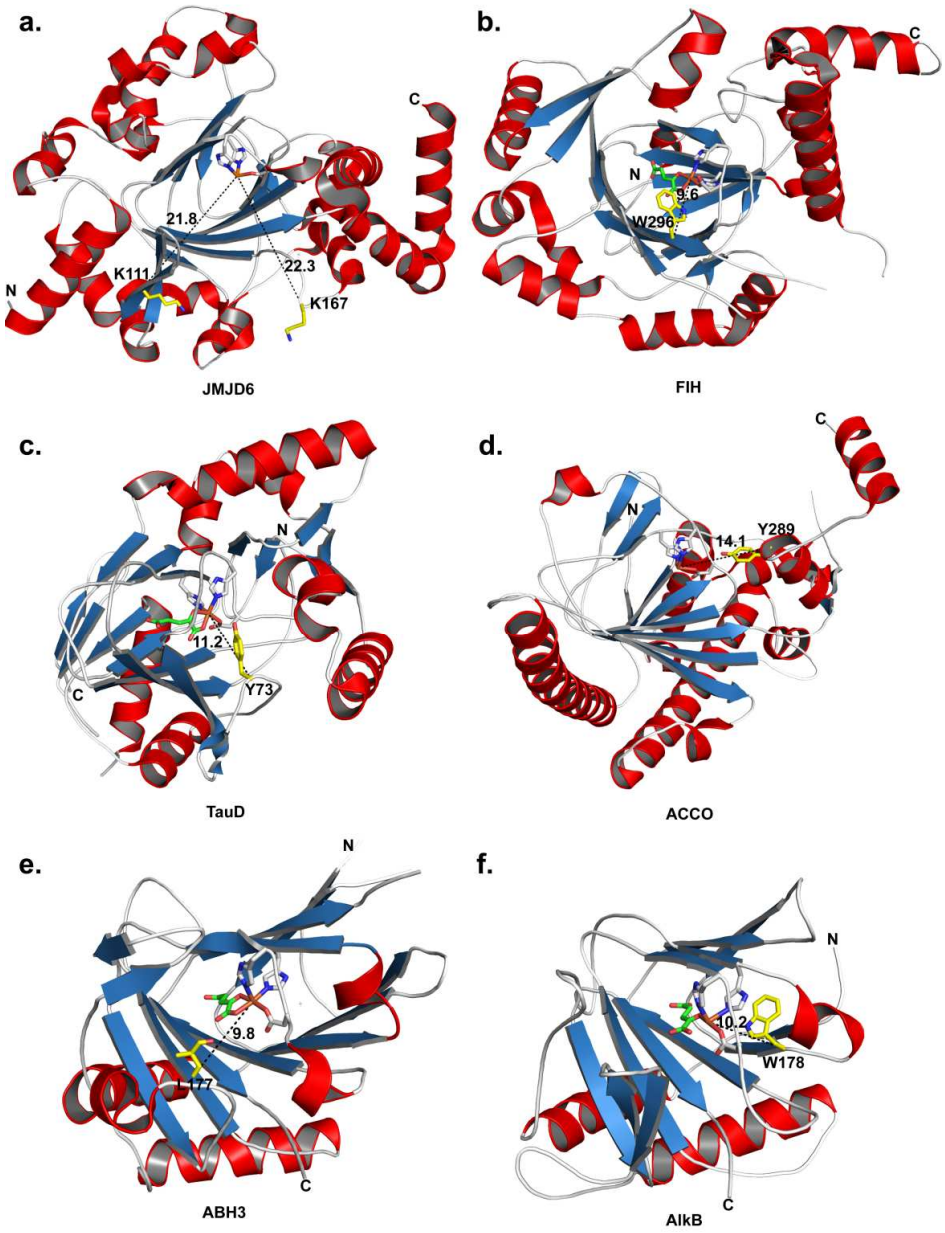


Figure 5.10 Positions of self-hydroxylation sites within various 2OG oxygenases and ACCO. Ribbons representation of 2OG oxygenases and ACCO showing the distances in Angstorms (black dashes) between the iron centre (white sticks for protein derived ligands, green for 2OG cofactors and orange for iron or surrogate metal) and *C*-alpha atom of the self-hydroxylated residues (yellow sticks). a. JMJD6 (PDB ID 3K2O), b. FIH (PDB ID 1H2K), c. TauD (PDB ID 1GQW), d. ACCO (PDB ID 1W9Y), e. ABH3 (PDB ID 2IUW), f. AlkB (PDB ID 2FD8). In the case of ACCO hydroxylation was not observed directly but is proposed to lead to backbone fragmentation. (*Figure is prepared by Dr. Michael A. McDonough*).

For other 2OG oxygenases for which oxidative modifications have been observed, they occur within ~10Å of active site. Whilst the mechanism of how these modifications occur is unknown, the proximity of modified residues to the active site implies that they are modified either directly by reactive oxidising species or by leakage of reactive oxidising species (e.g. H_2O_2, $O_2^{\cdot-}$) from the active site. However in the core of JMJD6, the identified site of hydroxylation at $Lys111_{JMJD6}$, $Lys167_{JMJD6}$ are both approximately 21.8 and 22.3 Å away from the active site as shown in Figure 5.10. Lys111 and Lys167 are conserved across higher eukaryotes (Figure 5.11). Both $Lys111_{JMJD6}$ and $Lys167_{JMJD6}$ are located in the loop regions that on the basis of the crystallographic analysis are likely to be relatively flexible. This may explain how these residues are able to fit into active site to be hydroxylated.

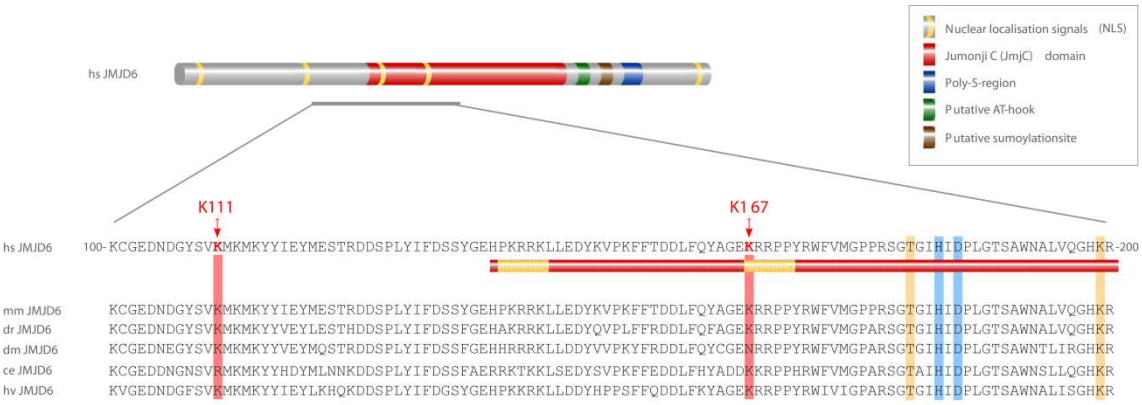


Figure 5.11 Multiple sequence alignment of JMJD6 proteins (amino acids 100-200) from different species. $Lys111_{JMJD6}$ and $Lys167_{JMJD6}$ are highlighted in orange, residues proposed to bind Fe(II) and 2OG are shown in green and pink, respectively. The domain structure of the full length human JMJD6 is shown at the top. Sequences from *Homo sapiens* (hs), *Mus musculus* (mm), *Danio rerio* (dr), *Drosophila melanogaster* (dm),) *Hydra vulgaris* (hv) and *Caenorhabditis elegans* (ce). *This image is prepared by Dr. Alexander Wolf*.

5.6.1 Biological relevance of self-hydroxylation in case of JMJD6

Overall the results demonstrate that JMJD6 can undergo self-hydroxylation both in recombinant form, where the hydroxylation is shown to be Fe(II) and 2OG dependent, as well as when isolated from human cells. Notably amino acid analysis studies imply that the self-hydroxylation proceeds to give the product with the 5*S*-stereochemistry, i.e. the same as for JMJD6 catalysed hydroxylation of arginine serine rich domain proteins (e.g. U2AF65). This differs from the 5*R*- hydroxylation stereochemistry observed for collagen lysyl hydroxylation which is catalysed by 2OG dependent oxygenases.[16] This observation is important because it implies that the JMJD6 catalysed hydroxylation is active site determined and is not a result of reactive oxidising species 'leaking' away from the iron centre. It is possible that Fe=O species is generated at the active site by uncoupled 2OG turnover and can cause self-hydroxylation. In some cases (TauD, TfdA), ^{18}O labelling experiments along with mass spectroscopic analysis have provided an evidence that oxidation of aromatic residues at the active site of these enzymes occurs without incorporation of ^{18}O from $^{18}O_2$.[2, 3] This contrasts with the case of substrate hydroxylation where the alcohol oxygen comes predominantly from O_2. Thus, for TfdA and TauD it is possible that exchange of an O_2 derived Fe-O species occurs with water prior to self-hydroxylation. Support for this proposal comes from the observation that for some 2OG oxygenases labelling studies have shown that substrate hydroxylation occurs with incomplete incorporation of oxygen from O_2.[22] In the case of JMJD6, as the stereochemistry of hydroxylated lysyl residues is same as that of substrate hydroxylation, it seems most probable that self-hydroxylation occurs via a

closely related mechanism to that occurring for lysyl hydroxylation of substrates i.e. via Fe=O mediated hydroxylation.

In this regard the JMJD6 catalysed self-hydroxylation may differ from that observed for other 2OG oxygenases to date, where the oxidised residues differ from the normal substrates. In the case of FIH, a tryptophan residue is oxidised via self-hydroxylation but a tryptophan residue has not been reported to be normal substrate for FIH.[18] Thus it is arguably more likely that the JMJD6 catalysed self-hydroxylation is biologically relevant, especially as it is observed in human cells.

It has been speculated that the self-hydroxylation of some 2OG oxygenases may protect cells from damaging oxygen species or may serve as a negative feedback mechanism.[1] Self-regulation is employed by many enzymes in nature, e.g. kinases autophosphorylate residues on their own protein scaffold to regulate their activity[23] and ubiquitinases are capable of targeting themselves for ubiquitin-dependent degradation.[24, 25] It was observed that self-hydroxylation and uncoupled turnover occur in the presence of JMJD6 substrate, U2AF65. It will be interesting to see if the extent of self-hydroxylation/ uncoupled turnover varies with different JMJD6 substrates and if it is possible that they have any biological role. As suggested for other 2OG oxygenases, it is possible that self-hydroxylation of JMJD6 is a part of negative feedback mechanism, though at this stage there is no evidence for any *in vivo* relevance of this. JMJD6 has been shown to play a role in the regulation of splicing of genes including VEGF receptor 1,[26] but the role of JMJD6 catalysed hydroxylation is so far unclear. At present the available evidence suggests that as proposed for hydroxylation catalysed by FIH,

JMJD6 may have multiple hydroxylation substrates and that the sum of many hydroxylation events control regulation in a context dependent manner.

The occurrence of self-hydroxylation by JMJD6 introduces another layer of complexity in oxygenase activity. It is notable that the hydroxylation of lysyl residues in collagen, both enables glycosylation and is involved in cross-linking.[27] Given that JMJD6 can form oligomers[28-30] and undergoes self-hydroxylation, it is possible that it may form intermolecular cross-linking. It is also important that modulation of self-hydroxylation should be taken into consideration when considering biological effects of small molecule inhibitors targeting JMJD6 and other 2OG oxygenases.

5.7 References

1. Chen, Y. H.; Comeaux, L. M.; Eyles, S. J.; Knapp, M. J., Auto-hydroxylation of FIH-1: an Fe(II), alpha-ketoglutarate-dependent human hypoxia sensor. *Chemical Communications (Cambridge, United Kingdom)* **2008,** (39), 4768-4770.

2. Ryle, M. J.; Liu, A.; Muthukumaran, R. B.; Ho, R. Y.; Koehntop, K. D.; McCracken, J.; Que, L., Jr.; Hausinger, R. P., O_2- and alpha-ketoglutarate-dependent tyrosyl radical formation in TauD, an alpha-keto acid-dependent non-heme iron dioxygenase. *Biochemistry* **2003,** *42* (7), 1854-1862.

3. Liu, A.; Ho, R. Y. N.; Que, L.; Ryle, M. J.; Phinney, B. S.; Hausinger, R. P., Alternative Reactivity of an α-Ketoglutarate-Dependent Iron(II) Oxygenase: Enzyme Self-Hydroxylation. *Journal of the American Chemical Society* **2001,** *123* (21), 5126-5127.

4. Sundheim, O.; Vagbo, C. B.; Bjoras, M.; Sousa, M. M. L.; Talstad, V.; Aas, P. A.; Drablos, F.; Krokan, H. E.; Tainer, J. A.; Slupphaug, G., Human ABH3 structure and key residues for oxidative demethylation to reverse DNA/RNA damage. *EMBO Journal* **2006,** *25* (14), 3389-3397.

5. Cikala, M.; Alexandrova, O.; David, C.; Proschel, M.; Stiening, B.; Cramer, P.; Böttger, A., The phosphatidylserine receptor from Hydra is a nuclear protein with potential Fe(II) dependent oxygenase activity. *BMC Cell Biology* **2004,** *5* (1), 26-36.

6. Ge, Y.; Lawhorn, B. G.; ElNaggar, M.; Sze, S. K.; Begley, T. P.; McLafferty, F. W., Detection of four oxidation sites in viral prolyl-4-hydroxylase by top-down mass spectrometry. *Protein Science* **2003,** *12* (10), 2320-2326.

7. Müller, I.; Stückl, C.; Wakeley, J.; Kertesz, M.; Usón, I., Succinate complex crystal structures of the α-Ketoglutarate-dependent dioxygenase AtsK. *Journal of Biological Chemistry* **2005,** *280* (7), 5716-5723.

8. Zhang, Z.; Ren, J.-S.; Clifton, I. J.; Schofield, C. J., Crystal Structure and Mechanistic Implications of 1-Aminocyclopropane-1-Carboxylic Acid Oxidase-The Ethylene-Forming Enzyme. *Chemistry & Biology* **2004,** *11* (10), 1383-1394.

9. Sabourin, P. J.; Bieber, L. L., Purification and characterization of an alpha-ketoisocaproate oxygenase of rat liver. *Journal of Biological Chemistry* **1982,** *257* (13), 7460-7467.

10. Webby, C. J.; Wolf, A.; Gromak, N.; Dreger, M.; Kramer, H.; Kessler, B.; Nielsen, M. L.; Schmitz, C.; Butler, D. S.; Yates, J. R., III; Delahunty, C. M.; Hahn, P.; Lengeling, A.; Mann, M.; Proudfoot, N. J.; Schofield, C. J.; Bottger, A., Jmjd6 Catalyses Lysyl-Hydroxylation of U2AF65, a Protein Associated with RNA Splicing. *Science* **2009,** *325* (5936), 90-93.

11. Myllylä, R.; Majamaa, K.; Günzler, V.; Hanauske-Abel, H. M.; Kivirikko, K. I., Ascorbate is consumed stoichiometrically in the uncoupled reactions catalyzed by prolyl 4-hydroxylase and lysyl hydroxylase. *Journal of Biological Chemistry* **1984,** *259* (9), 5403-5405.

12. Flashman, E.; Davies, S. L.; Yeoh, K. K.; Schofield, C. J., Investigating the dependence of the hypoxia-inducible factor hydroxylases (factor inhibiting HIF and prolyl hydroxylase domain 2) on ascorbate and other reducing agents. *Biochemical Journal* **2010,** *427* (1), 135-142.

13. Nytko, K. J.; Maeda, N.; Schlāfli, P.; Spielmann, P.; Wenger, R. H.; Stiehl, D. P., Vitamin C is dispensable for oxygen sensing in vivo. *Blood* **2011,** *17* (20), 5485-5493.

14. De Jong, L.; Albracht, S. P. J.; Kemp, A., Prolyl 4-hydroxylase activity in relation to the oxidation state of enzyme-bound iron: The role of ascorbate in peptidyl proline hydroxylation. *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology* **1982,** *704* (2), 326-332.

15. Buettner, G. R.; Jurkiewicz, B. A., Catalytic metals, ascorbate and free radicals: Combinations to avoid. *Radiation Research* **1996,** *145* (5), 532-541.

16. Mantri, M.; Loik, N. D.; Hamed, R. B.; Claridge, T. D. W.; McCullagh, J. S. O.; Schofield, C. J., The 2-Oxoglutarate-Dependent Oxygenase JMJD6 Catalyses Oxidation of Lysine Residues to give 5S-Hydroxylysine Residues. *ChemBioChem* **2011,** *12* (4), 531-534.

17. Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M., Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Analytical Chemistry* **1996,** *68* (5), 850-858.

18. Yang, M.; Ge, W.; Chowdhury, R.; Claridge, T. D. W.; Kramer, H. B.; Schmierer, B.; McDonough, M. A.; Gong, L.; Kessler, B. M.; Ratcliffe, P. J.; Coleman, M. L.; Schofield, C. J., Asparagine and Aspartate Hydroxylation of the Cytoskeletal Ankyrin Family Is Catalyzed by Factor-inhibiting Hypoxia-inducible Factor. *Journal of Biological Chemistry* **2011,** *286* (9), 7648-7660.

19. Kelly, L.; McDonough, M. A.; Coleman, M. L.; Ratcliffe, P. J.; Schofield, C. J., Asparagine β-hydroxylation stabilizes the ankyrin repeat domain fold. *Molecular BioSystems* **2009,** *5* (1), 52-58.

20. Mantri, M.; Krojer, T.; Bagg, E. A.; Webby, C. J.; Butler, D. S.; Kochan, G.; Kavanagh, K. L.; Oppermann, U.; McDonough, M. A.; Schofield, C. J., Crystal Structure of the 2-Oxoglutarate- and Fe(II)-Dependent Lysyl Hydroxylase JMJD6. *Journal of Molecular Biology* **2010,** *401* (2), 211-222.

21. Hong, X.; Zang, J.; White, J.; Wang, C.; Pan, C.-H.; Zhao, R.; Murphy, R. C.; Dai, S.; Henson, P.; Kappler, J. W.; Hagman, J.; Zhang, G., Interaction of JMJD6 with

single-stranded RNA. *Proceedings of the National Academy of Sciences of the United States of America* **2010,** *107* (33), 14568-14572.

22. Welford, R. W. D.; Kirkpatrick, J. M.; McNeill, L. A.; Puri, M.; Oldham, N. J.; Schofield, C. J., Incorporation of oxygen into the succinate co-product of iron(II) and 2-oxoglutarate dependent oxygenases from bacteria, plants and humans. *FEBS Letters* **2005,** *579* (23), 5170-5174.

23. Yamaguchi, H.; Hendrickson, W. A., Structural basis for activation of human lymphocyte kinase Lck upon tyrosine phosphorylation. *Nature* **1996,** *384* (6608), 484-489.

24. Fang, S.; Jensen, J. P.; Ludwig, R. L.; Vousden, K. H.; Weissman, A. M., Mdm2 is a RING finger-dependent ubiquitin protein ligase for itself and p53. *Journal of Biological Chemistry* **2000,** *275* (12), 8945-8951.

25. Tang, J.; Qu, L. K.; Zhang, J.; Wang, W.; Michaelson, J. S.; Degenhardt, Y. Y.; El-Deiry, W. S.; Yang, X., Critical role for Daxx in regulating Mdm2. *Nature Cell Biology* **2006,** *8* (8), 855-862.

26. Boeckel, J.-N.; Guarani, V.; Koyanagi, M.; Roexe, T.; Lengeling, A.; Schermuly, R. T.; Gellert, P.; Braun, T.; Zeiher, A.; Dimmeler, S., Jumonji domain-containing protein 6 (Jmjd6) is required for angiogenic sprouting and regulates splicing of VEGF-receptor 1. *Proceedings of the National Academy of Sciences of the United States of America* **2011,** *108* (8), 3276-3281.

27. Myllyharju, J.; Kivirikko, K. I., Collagens and collagen-related diseases. *Annals of Medicine* **2001,** *33* (1), 7-21.

28. Tibrewal, N.; Liu, T.; Li, H.; Birge, R., Characterization of the biochemical and biophysical properties of the phosphatidylserine receptor (PS-R) gene product. *Molecular and Cellular Biochemistry* **2007,** *304* (1), 119-125.

29. Hahn, P.; Wegener, I.; Burrells, A.; Böse, J.; Wolf, A.; Erck, C.; Butler, D.; Schofield, C. J.; Böttger, A.; Lengeling, A., Analysis of Jmjd6 Cellular Localization and Testing for Its Involvement in Histone Demethylation. *PLoS One* **2010,** *5* (10), e13769.

30. Han, G.; Li, J.; Wang, Y.; Li, X.; Mao, H.; Liu, Y.; Chen, C. D., The hydroxylation activity of Jmjd6 is required for its homo-oligomerization. *Journal of Cellular Biochemistry* **2012,** *113* (5), 1663-1670.

Chapter 6. Oligomerisation of JMJD6

6.1 Introduction

JMJD6 is a typical 2OG oxygenases enzyme which possesses a DSBH or jelly roll core structural motif as described in Chapter 3. Proteins that contain a DSBH fold, including 2OG oxygenases and related enzymes, are known to exist in various oligomeric forms.[1] For example, *S*-hydroxypropylphosphonic acid epoxidase (HPPE) which catalyses the formation of the epoxide antibiotic fosfomycin, exists as a tetramer,[2, 3] while germin forms hexameric structure.[4] On the other hand, FIH undergoes dimerisation via its *C*-terminal helical region which is essential for its hydroxylation activity.[5] Previously, it has been shown that JMJD6 has structural and sequence homology with FIH, hence, it was of interest to investigate its oligomerisation.

It has been proposed by Tibrewal et al. that JMJD6 undergoes probable post-translational protein cross-linking to form covalently linked multimers.[6] This group reported the observation of multimeric bands when various JMJD6 constructs with either *N*- terminal HA tag or *C*- terminally located V5 tag was produced in HEK293T cells. This apparent oligomerisation was also seen at endogenous expression level of JMJD6. Similar to these results, Hahn et al. also reported evidence for the production of JMJD6 as multimers in both nuclear and cytoplasmic extracts of MEF cells; they also reported probable dimerisation of YFP tagged JMJD6 when a JMJD6-YFP fusion construct was produced in HEK293T cells.[7] More recently, Han et al. reported that JMJD6 with a *N*- terminal Flag-tag produced in HEK293T cells undergoes homo-

oligomerisation.[8] The oligomerisation was shown to be dependent on intact Fe(II) binding histidine residue (H187) which is a part of HxD/E…H motif of JMJD6. In this chapter, the focus of my studies was to investigate the oligomerisation state of wild-type and various deletion constructs of JMJD6 by employing transmission electron microscopy for structural analysis.

6.2 Gel electrophoresis showing higher bands

Oligomerisation of JMJD6 was first analysed by gel electrophoresis. When recombinant JMJD6 full length construct with *N*-terminal His-tag was analysed by SDS-PAGE, gel bands corresponding to 120, 148, 200, 250 kDa were observed (Figure 6.1). When the same JMJD6 construct was analysed by native gel electrophoresis it accumulated at the top of the gel and no separation was observed. Further, carrying out western blot analyses, using the mAb10526 antibody did confirm that these higher bands correspond to JMJD6.

To investigate the effect of cofactors essential for 2OG oxygenase activity on JMJD6 oligomerisation, wild-type JMJD6 and *C*-terminal and *N*-terminal truncated JMJD6 constructs ($JMJD6_{1\text{-}343}$ and $JMJD6_{49\text{-}403}$) with *N*-terminal His-tag were incubated with Fe(II) and 2OG prior to separation by SDS-PAGE; the formation of oligomer states was not affected by addition of Fe(II) or 2OG. Active site mutants of JMJD6 ($JMJD6_{H187A}$ and $JMJD6_{H187AandH273A}$) also showed higher bands in SDS-PAGE gels indicating that probably oligomerisation of JMJD6 is independent of its oxygenase activity. These results contrast with recently reported results by Han et al., where western blot analysis of a JMJD6 H187A mutant overexpressed from HEK293T did not reveal formation of

higher bands suggesting that an intact HxD/E…H motif might be required for oligomerisation of JMJD6.[8]

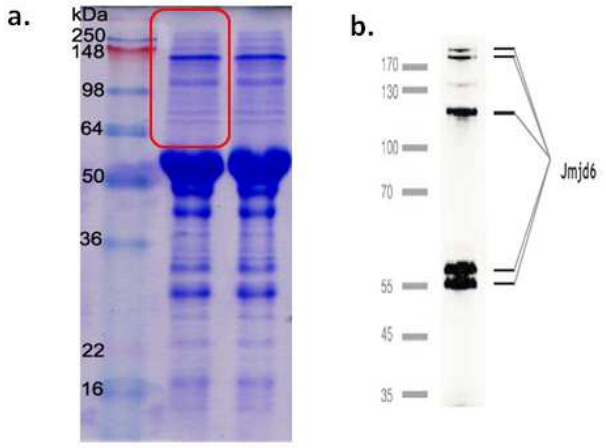


Figure 6.1 SDS-PAGE gels showing formation of oligomeric JMJD6 bands. (a) As observed by coomassie stain and (b) western blot analysis using JMJD6 antibody mAb10526.

6.3 Analytical size exclusion chromatogarphy

Analytical size exclusion chromatography was then used to study JMJD6 oligomerisation. The calculated molecular weight of monomeric full length JMJD6 with *N*-terminal His-tag is 48 kDa. Analytical size exclusion chromatography for JMJD6 was carried out using a Superdex 200 column (30 ml, GE Healthcare), that had been calibrated using blue dextran, thyroglobulin, apoferritin, ThnG and cytochrome *C* as standards. The calibration curve was plotted as the log of molecular weights of calibration standards against retention volume as shown in Figure 6.2, according to standard practise.[9, 10]

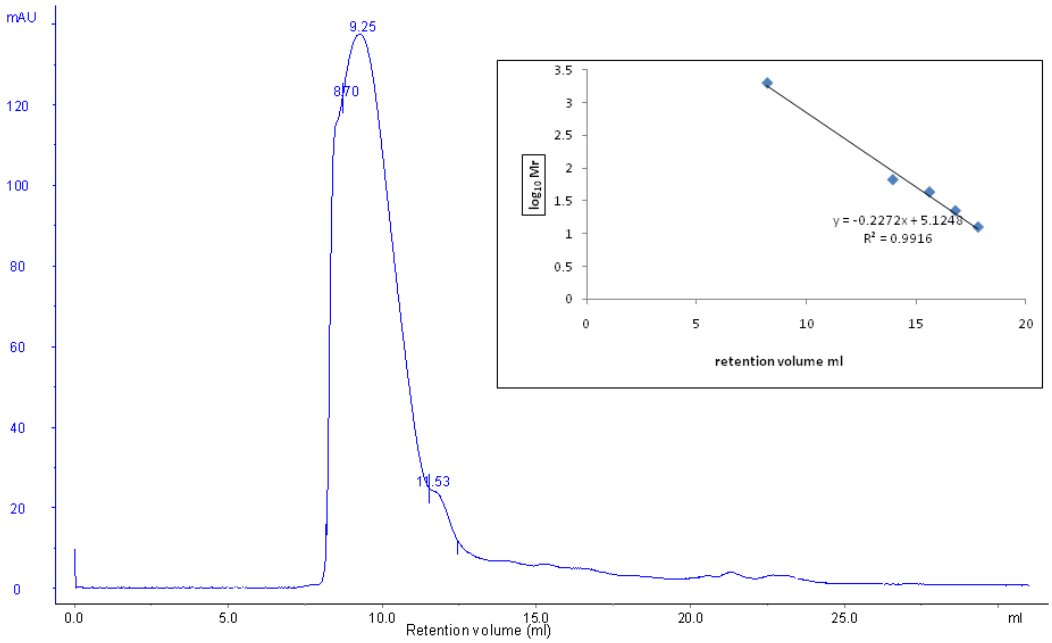


Figure 6.2 A_{280} chromatogram showing that JMJD6 exists as a mixture of higher molecular weight species corresponding to 8.70 and 9.25 ml and as hexamer (predicted MW=291 kDa; observed MW=310 kDa) forms. Insert shows the calibration plot obtained by blue dextran, thyroglobulin, apoferritin, ThnG and cytochrome *C* as standards using S200 analytical size exclusion column.

The purified His-tagged JMJD6 protein was then applied to the Superdex S200 column that had been pre-equilibrated with buffer (50 mM Tris-HCl, pH 7.5). The retention volume for JMJD6 was then compared with the curve obtained from the standards. The results of the analytical size exclusion chromatography predict that JMJD6 exists as a mixture of higher molecular weight species corresponding to retention volume 8.70 and 9.25 ml as shown in Figure 6.2. The retention volume 11.53 ml probably corresponds to a hexamer of JMJD6 (predicted MW=291 kDa; observed MW=310 kDa).

6.4 Analysis of JMJD6 oligomers by TEM

6.4.1 Transmission electron microscopy-A general overview

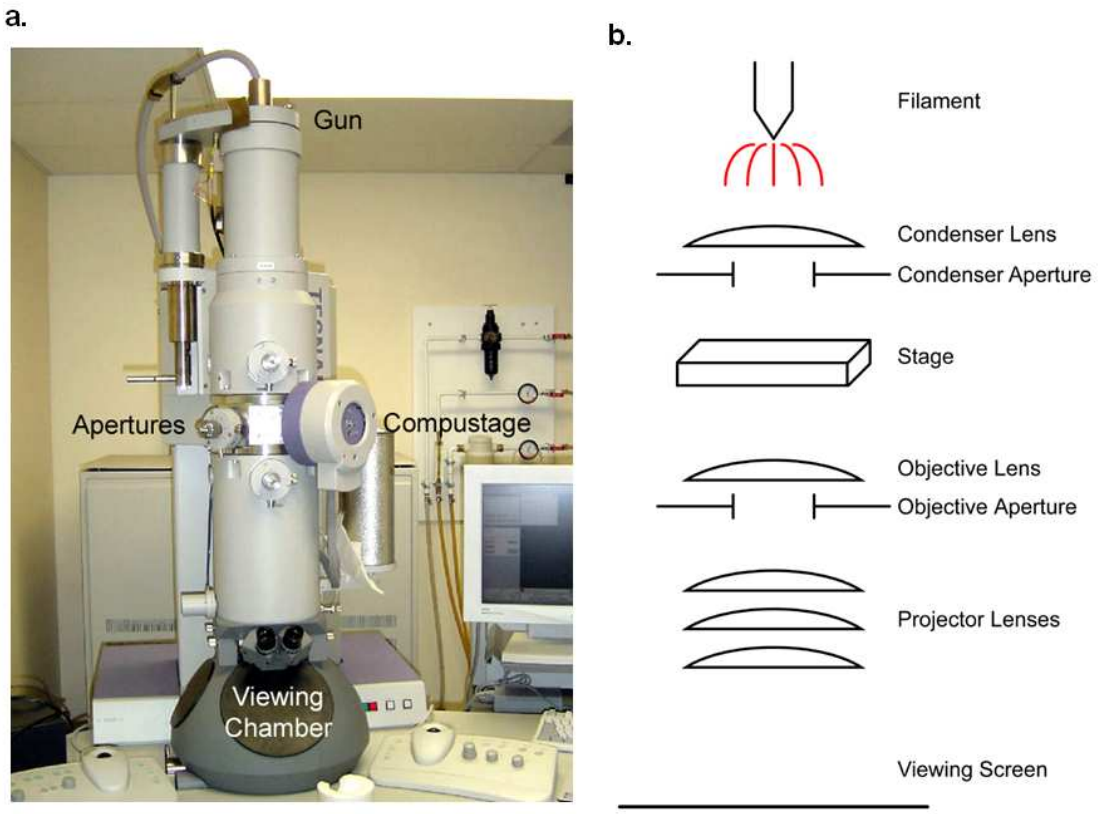


Figure 6.3 General features of TEM. a) A typical picture of TEM, Philips CM120 microscope (http://www.snaggledworks.com/em_for_dummies/background.html). b) Schematic of TEM showing the actual position of various electromagnetic lenses responsible for generating an image.

Transmission electron microscope (TEM) is a type of electron microscope which uses a high energy electron beam to obtain projection images of objects such as biological macromolecules at high resolution. The first TEM was built in 1931 by Max Knoll and Ernst Ruska for which Nobel prize was awarded in 1986. The basic principle underpinning the operation of TEM is the de Broglie's postulate that electrons have

wavelike character and these waves can be focused by electromagnetic lenses. In TEM, high energy electrons generated by an electron gun are directed onto the sample of interest.[11]

Three different types of scattering events occur when an electron beam reaches the specimen: no interaction, elastic scattering and inelastic scattering.[11, 12] Most electrons (about 90%) pass through the sample without any interaction. In elastic scattering, the incident electrons are deflected from their original path by the coulombic interaction with the relatively large positively charged atomic nucleus without loss of energy. Thus, in elastic scattering, the electrons maintain their initial wavelength but the phase is modified. Inelastic scattering occurs when high energy electrons interact with slow moving electrons from the atoms of the specimen. This event involves a transfer of energy resulting in a change of wavelength and as a consequence of phase. Inelastically scattered electrons contribute to the noise in the image and also deteriorate the specimen. All of these scattering events are used to generate contrast in the image of the specimen resulting in its visualisation.

Among the primary methods used to generate contrast in TEM is amplitude and phase contrast. Amplitude contrast is achieved by placing an aperture in front of the objective lens allowing only inelastically scattered electron beams to pass through, for image formation, while blocking the most deflected elastically scattered electrons.[11, 13] This generates image contrast because the sample regions wherein much scattering occurs remains dark and the region where there is less scattering, appears light. Amplitude contrast plays an important role in negative staining electron microscopy. Phase contrast

is achieved by utilising the wavelike property of electrons such that the objective aperture of the microscope is utilised to allow elastically scattered and unscattered electrons to interfere and generate contrast.[13] When electrons, in the form of waves pass through matter, the wavefront undergoes a phase shift. This causes interference when mixed with the wavefront of unscattered electrons.

A typical TEM set up is shown in Figure 6.3. At the top of the column, there is a high voltage electron gun that generates a high energy beam of electrons that travels down a high vacuum column which is lined with a series of electromagnetic lenses.[13] Electromagnetic condenser lenses, that can change the direction of the electrons, are placed to control the irradiation on the sample. These electrons further pass through the sample placed on the specimen stage. The objective lens, placed in the middle of the column is the most important lens in the microscope as it recombines the scattered and unscattered electrons to give rise to an image. A series of magnifying magnetic lenses are placed after the objective lens where electrons are ultimately focussed at the viewing screen at the bottom of the column. Different lenses can be used to change the magnification and focal point of the image. Apertures along the column are also used to change the contrast and resolution of the image.

6.4.2 Treatment of samples for TEM

TEM imaging of biological samples like proteins and macromolecules has specific challenges : (a) biological samples are easily damaged by high energy radiation (b) their structures can deteriorate in the high vacuum TEM column and (c) the primary constituents of biological samples are carbon, oxygen and nitrogen which have similar

scattering profiles causing low intrinsic contrast. These issues can be overcome by either negative staining or by cryo-freezing of samples.

The negative staining technique is used to help to overcome the above mentioned issues by applying the biological sample with a low concentration salt solution of heavy atoms such as uranyl acetate, phosphotungstic acid, uranyl formate, ammonium molybdate etc.[14] High contrast in the image is achieved by using heavy atoms and the actual image is cast generated by stain. A resolution of about 15 Å can be achieved by negative staining. In cryo-freezing technique, samples are allowed to freeze very rapidly in liquid ethane (at liquid nitrogen temperatures) such that water molecules (present in the sample) are trapped so as to avoid formation of crystalline ice which damages the sample and which readily absorbs electrons, obscuring the sample.[15] The cryo-freezing technique has a drawback as images have a low contrast, but due to high preservation of sample integrity, it can generate very high resolution. However, the samples need to be maintained at liquid nitrogen temperature throughout the imaging.

In TEM, samples are applied onto copper grids with a fine mesh (about 15 squares per millimetre) and then placed on a sample holder. Prior to sample application, a thin layer of carbon is deposited by evaporating carbon graphite onto the copper grid.[11] In order to have an evenly spread and stained specimen on the grid, the carbon film has to be made hydrophilic. This is done using the glow discharge method. For glow discharging, the carbon coated grids are placed inside a partly evacuated chamber connected to a power supply and high voltage is applied between the cathode and anode at each end of the chamber. The electron potential ionizes the gas within the chamber and negatively

charged ions are deposited on the carbon, giving the carbon film an overall hydrophilic surface.

6.4.3 Image analysis

The TEM image generated is a 2D projection containing information of the whole sample, which is a 3D object.[14, 16] When the sample under observation is a biological material, both due to the poor contrast and the fact that these samples are illuminated for a short time with a limited electron dose, the images are usually very noisy. In order to recover 3D biological object from noisy 2D images, extensive computational analysis is required. This involves two steps: firstly multiple images taken from the same view of the object are identified and averaged together to increase the signal/noise ratio. Secondly, different averaged views are combined together in the right way to recover the 3D information. This process is called reconstruction.

In the work described in this Chapter, 2D averaging and classification of negatively stained particles and a 3D reconstruction from frozen hydrated particles were carried out. For the negatively stained specimen and the analysis in 2D, electron micrographs were taken at a magnification of 45000. Selected images were digitised with a step size of 25 µM on a Nikon Super Coolscan 9000 (Nikon, London, UK). Individual particles were selected manually from TEM micrographs then 'windowed' into individual images followed by normalisation. The WEB and SPIDER software packages[17] were used for all image processing. The particles were then subjected to reference-free alignment, and sorted into classes using the partitional method (K-means method) of clustering.[18] For

the analysis of the specimen in cryo-condition and the reconstruction in 3D see the Section 6.6.

6.5 TEM analysis of various JMJD6 constructs by negative staining technique

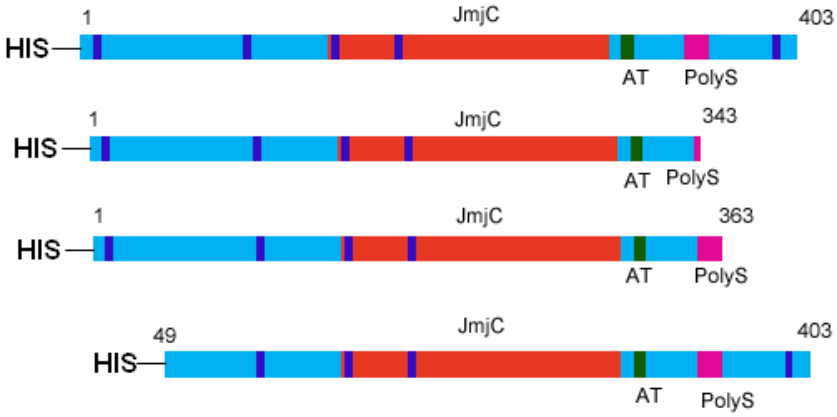


Figure 6.4 Various constructs of JMJD6 expressed for TEM analysis. Nuclear localisation signal regions are shown in blue, the JmjC domain is coloured in red, AT hook region is in green and the polyserine region is shown in pink.

As observed by electrophoresis and analytical size exclusion chromatography, JMJD6 undergoes oligomerisation to form a probable hexamer with a molecular weight of approximately 291 kDa. Due to the formation of high molecular weight oligomers, it was possible to visualise JMJD6 multimers by TEM. Various constructs of JMJD6 as shown in Figure 6.4 were proposed and then analysed by TEM. JMJD6 has a *C*-terminal conserved polyserine region ($JMJD6_{340\text{-}363}$), occurring after the catalytic JmjC domain in its sequence. In order to investigate the involvement of *N*-terminal and *C*-terminal sequences and the polyserine region of JMJD6 in the formation of oligomers, an *N*-terminal truncated construct from 49-403 amino acids ($JMJD6_{49\text{-}403}$) and two *C*-terminal

truncated sequences 1-362 ($JMJD6_{1-362}$) and 1-343 ($JMJD6_{1-343}$, without the polyserine region) were analysed by TEM. All the constructs were cloned into the pET28b(+) vector and then expressed with an *N*-terminal His-tag in *E. coli* cells, followed by purification using a nickel affinity column as described in Chapter 8. *The JMJD6 construct* $JMJD6_{1-343}$ *was prepared by Dr. Danica Butler while constructs* $JMJD6_{1-362}$ *and* $JMJD6_{49-403}$ *were prepared by Dr. Alexander Wolf.* In order to avoid aggregation of protein over the time and by repeated freeze thaw cycles, once purified, the protein was concentrated and then buffer exchanged with a low glycerol concentration buffer (50 mM Tris.HCl, 300 mM NaCl, 5% glycerol pH 7.5) using Micro BioSpin columns (Bio-Rad). Sample grids for TEM analysis were prepared immediately after protein preparation by mixing with uranyl acetate for negative staining on charged carbon coated copper grids. *All the TEM analyses in this chapter were carried out by Dr. Catherine-Véniene Bryan, Department of Biochemistry, Oxford. TEM images were acquired using a Philips CM120 microscope.*

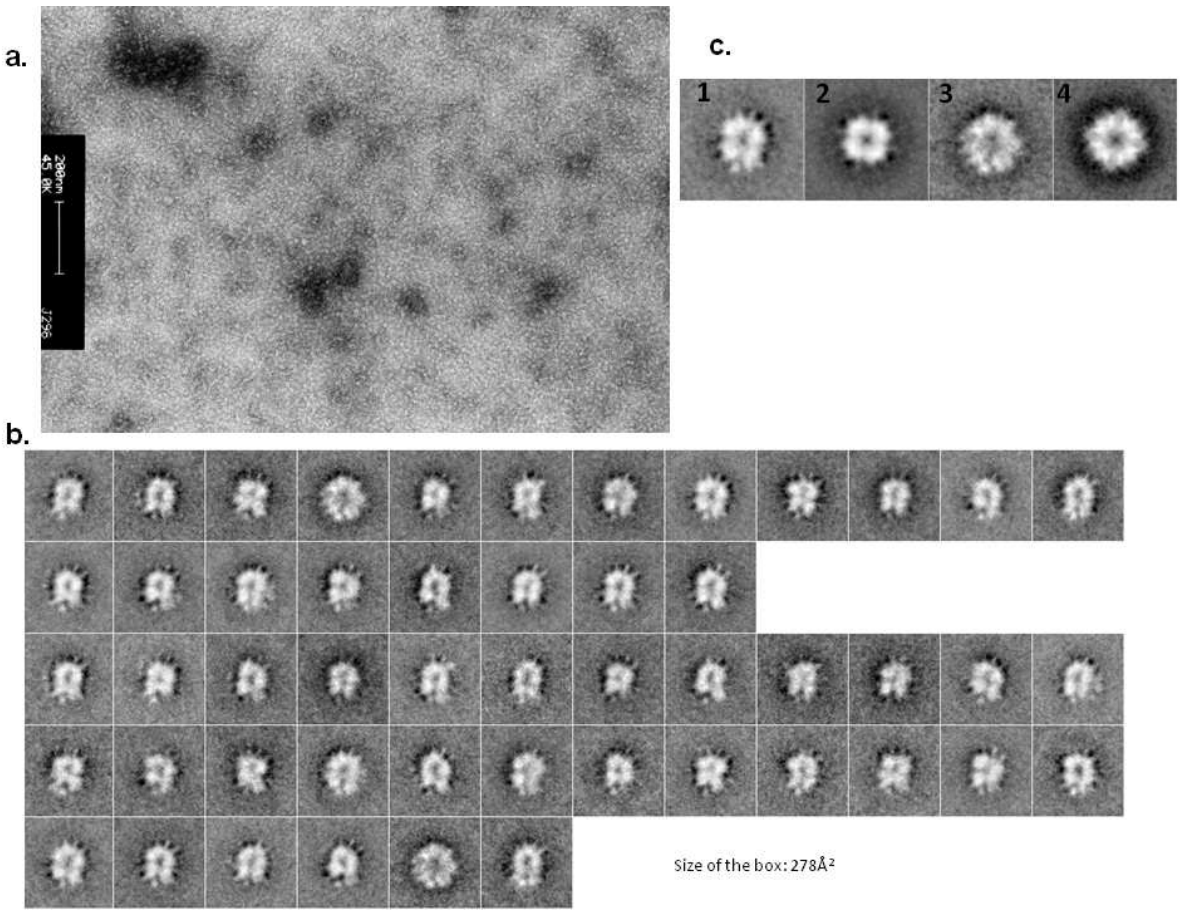


Figure 6.5 Image analysis of the full length His-JMJD6 construct. (a) TEM micrograph of the sample. (b) 3122 particles distributed into 20 and 30 classes. (c) Selected classes showing (1) Tetramer of dimer and (3) higher oligomer. (2) and (4) represent the reconstructions with the imposed symmetry.

Initially, the full length JMJD6 and truncated JMJD6$_{1\text{-}343}$ constructs with *N*-terminal His-tags were expressed and analysed by the negative staining technique using uranyl acetate as a metal salt to preserve the specimen and generate the contrast. Full length JMJD6 with a His-tag is observed to form ring like structures as observed in the TEM images (Figure 6.5). For image analysis of JMJD6 micrographs, 3122 particles were selected which were distributed into classes of 20 and 30 groups. The two interesting classes were selected; tetramer of dimers to which fourfold symmetry was imposed

(Figure 6.5c1) and the other class formed from a higher oligomeric ring (Figure 6.5c3) with a diameter of 161±15Å.

The filament-like structures (as shown in Figure 6.6) were observed in the case of the *C*-terminally truncated His-JMJD6$_{1-343}$ construct which does not have an intact polyserine region. Further image analysis was carried out on these filaments by aligning the images in order to identify common features among these filaments. However, the results indicated that there is no clear organisation of the protein inside these filaments. Comparing these results with the ring like oligomers formed by full length JMJD6 construct, it is probable that *C*-terminal sequence of JMJD6 from 343-403 with polyserine region maybe is involved in oligomeric ring formation.

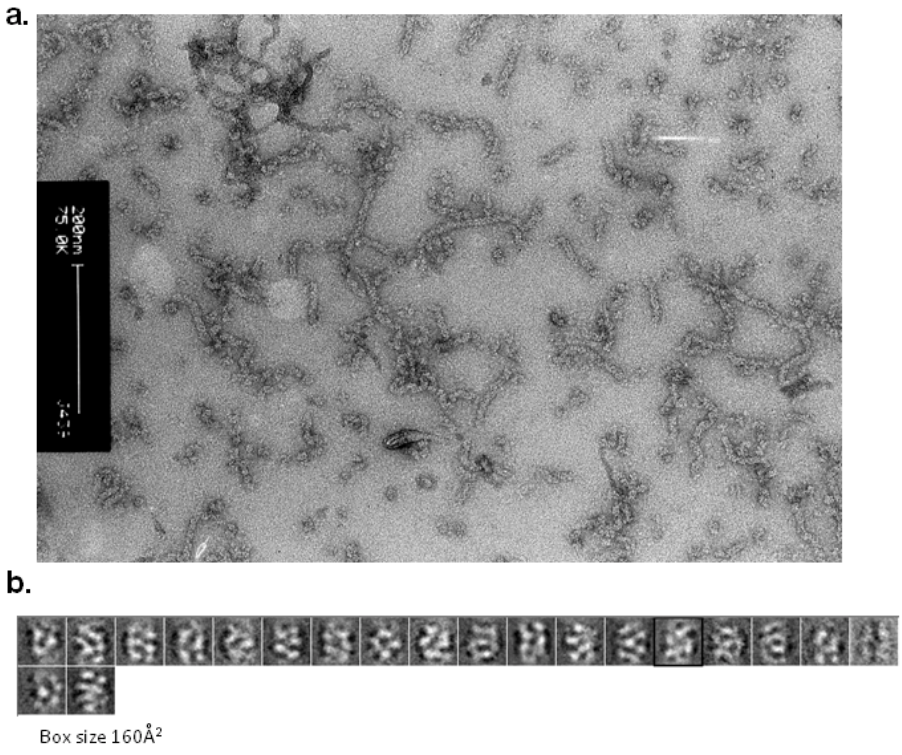


Figure 6.6 Image analysis of His-JMJD6$_{1-343}$ construct. (a) TEM micrograph of sample showing formation of filaments. (b) Sections of filament like structures distributed into classes.

Further, TEM analysis of a new *C*-terminally truncated construct JMJD6$_{1-362}$ which includes the polyserine region (JMJD6$_{340-362}$) present in the JMJD6 sequence was carried out. The formation of oligomeric rings was observed for this construct as shown in Figure 6.7. Image analysis was carried out by aligning 4194 particles and classifying them into 50 classes. The analysis showed that oligomers of JMJD6$_{1-362}$ exist in three classes probably as dimer of tetramers with smaller diameter 101Å and with bigger diameter 116Å and heptamer of dimer, diameter 153Å.

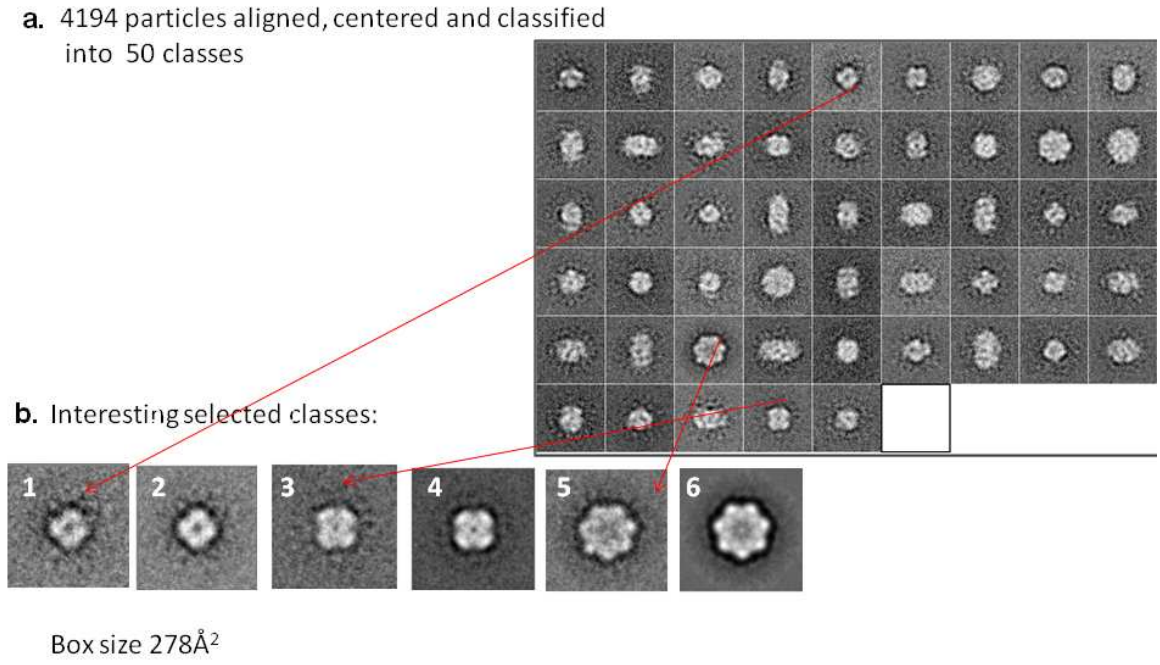


Figure 6.7 Image analysis of His-JMJD6$_{1-362}$ construct. (a) 4194 particles distributed into 50 classes. (b) Selected classes showing apperent (1) and (3) Tetramer of dimer and (5) Heptamer of dimer. (2), (4) and (6) correspond to the reconstruction with symmetry imposed on particles from (1), (3) and (5) respectively.

To investigate the potential involvement of the *N*-terminal region of JMJD6 in oligomerisation, TEM analysis of the *N*-terminally truncated His-tagged JMJD6$_{49-403}$

construct was carried out. Image analysis of $JMJD6_{49-403}$ indicated the formation of oligomers in the form of probable heptamer of dimers (Figure 6.8) similar to the one observed for $JMJD6_{1-362}$ (Figure 6.7b5). Furthermore, comparing images from constructs $JMJD6_{1-362}$ and $JMJD6_{49-403}$, it is seen that in the case of the heptameric structures of $JMJD6_{49-403}$, there is a lack of density inside the rings. These results indicate that the *N*-terminal region of JMJD6 (residues 1-48) is probably situated inside these heptameric rings.

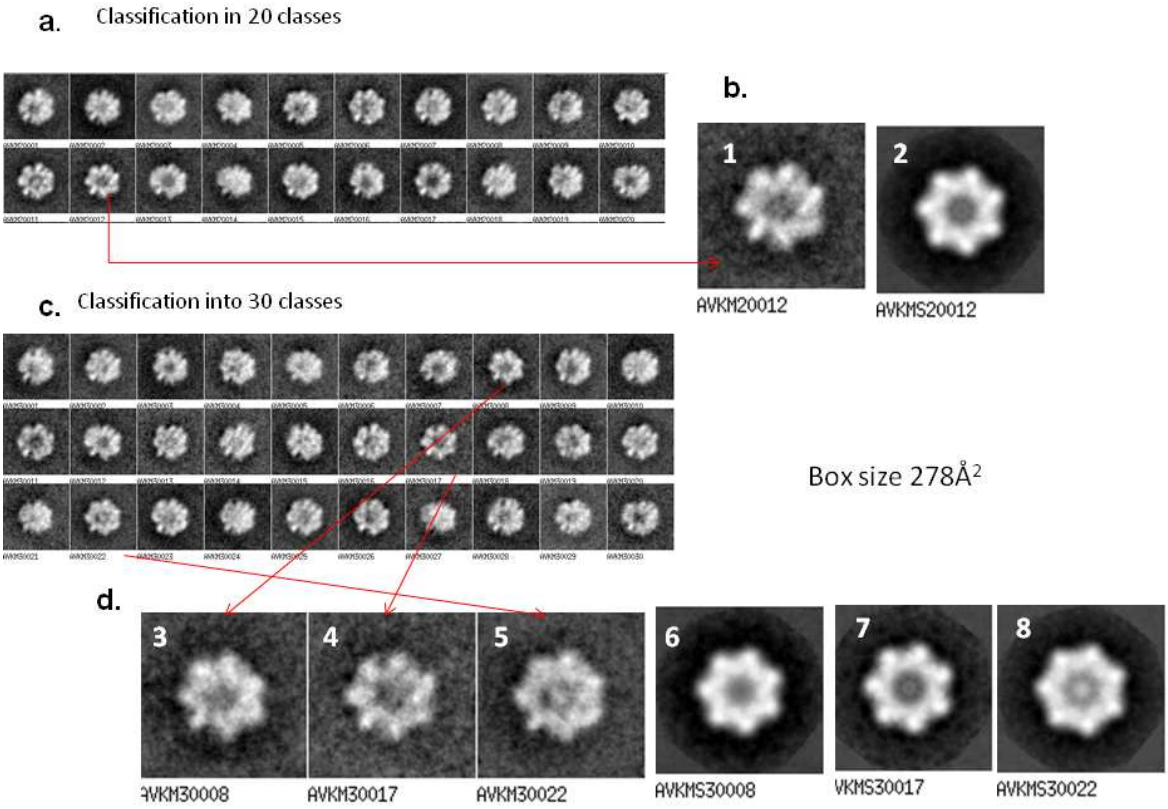


Figure 6.8 Image analysis of the His-$JMJD6_{49-403}$ construct. (a) and (c) Particles distributed into 20 and 30 classes. (b) and (d) Selected classes showing the heptamer of dimer chosen as (1), (3), (4) and (5). Reconstruction with symmetry imposed on these particles are shown as (2), (6), (7) and (8) respectively.

To investigate if active site residues of JMJD6 are necessary for oligomarisation, TEM analysis of the His-tagged full length JMJD6 construct with His187 and His198 substituted for alanine residues was carried out. The resultant analysis of preliminary images showed formation of oligomeric rings (results not shown). However, the sample was observed to be not very homogeneous and further image analysis was not carried out for the mutant.

Oligomerisation was also studied by TEM for full length His-tagged JMJD6 by adding recombinant U2AF65 substrate in the presence of NOG and Ni(II) in order to mimic the binding by 2OG and Fe(II). Even in the presence of U2AF65, it was seen that JMJD6 formed ring like oligomers (results not shown). However, the addition of U2AF65 gave rise to aggregation in the sample and thus further analysis couldn't be carried out.

6.6 Cryo EM analysis of native full length JMJD6

6.6.1 Preparation of wild-type JMJD6 for Cryo-freezing TEM

JMJD6 with an *N*-terminal His-tag was produced in *E. coli* cells and purified using a nickel affinity column as described in Chapter 8. The His-tag was cleaved using thrombin to generate wild-type JMJD6 (without any tag) for cryo-TEM analysis. The cleavage is carried out by addition of thrombin (1 unit of thrombin per 4 mg of protein), followed by incubation at 4ºC for 36 hours in the presence of 10x thrombin cleavage buffer (20 mM Tris-HCl, 150 mM NaCl, 2.5 mM $CaCl_2$, pH 8.4). The cleavage was monitored by SDS-PAGE (Figure 6.9). The His-tag cleaved protein is separated from His-tag attached protein by running another 5 ml nickel affinity column (HisTrap™ HP 5 ml column) which had been pre-equilibrated using 5CV of gel filtration buffer.

Protein present in the gel filtration buffer was loaded onto the column and unbound protein (His-tag cleaved) was eluted using 5CV of gel filtration buffer. Then the bound, His-tag attached protein was eluted using an elution buffer containing 500 mM imidazole. The His-tag cleaved protein was concentrated and buffer exchanged to the gel filtration buffer to remove imidazole, using a 10,000 MWCO Amicon concentrator (Millipore). A small amount of protein was further buffer exchanged to glycerol free buffer (50 mM Tris, 300 mM NaCl) using Micro BioSpin columns (Bio-Rad). The buffer exchange procedure was carried out in two steps because JMJD6 tends to undergo aggregation if directly buffer exchanged from elution buffer to the glycerol free buffer.

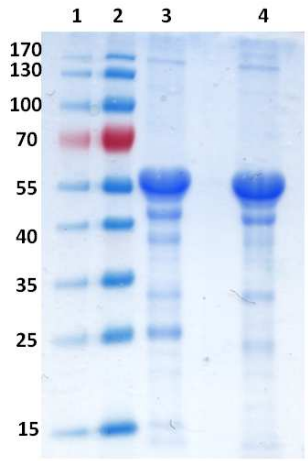


Figure 6.9 SDS-PAGE gel showing cleavage of His-tag after incubation with thrombin. Lane 1 and 2- MW markers Lane 3- His-tagged JMJD6 (wild-type) as purified by nickel affinity column. Lane 4- His-tag cleaved JMJD6 (wild-type) after purification from nickel affinity column and incubation with thrombin for 36 hours, showing displacement of JMJD6 band on gel.

Three microliters of JMJD6 (about 5 mg ml^{-1}) were placed on a 400-mesh copper grid coated with a thin holey-carbon film. The grid was blotted, vitrified by freeze plunging in liquid ethane, and transferred at liquid nitrogen temperature into the microscope

using a side-entry nitrogen cooled Gatan 626 cryoholder. CryoEM micrographs were obtained as images from a 4Kx4K CCD camera using a JEM-2100F (JEOL) microscope operating at 200 kV with a field emission gun (FEG) at a magnification of 67,000. The microscope has a spherical aberration constant (Cs) of 0.5 mm. The CCD images are 4096 by 4096 pixels and the CCD elements are 15 microns across.

6.6.2 Image Analysis

Image analysis was carried out by Dr. William Nicholson from Dr. Catherine-Véniene Bryan's group. 13553 particles were picked from 71 micrograph images. Particles were manually picked from micrograph images using the interactive particle picking program boxer (from the EMAN1 software package).[19] Contrast transfer function (CTF) determination i.e. determination of the defocus and other parameters was carried out in the EMAN2 workflow.[20] Particle images were CTF corrected by phase flipping (i.e. amplitude related effects of the CTF were not corrected) in SPIDER. 2D reference free alignment was carried out on the phase flipped particle images, followed by classification into 100 classes by K-means classification in SPIDER.[21] A SPIDER image stack with the class averages was imported into EMAN2 and then used to generate an initial model with c1 symmetry. Refinement of the 3D reconstruction was then carried out by multiple iterations of alignment of phase flipped particles versus the current 3D reconstruction. Image reconstruction results obtained by these preliminary analyses are shown in Figure 6.10. The resolution for the calculated 3D model is only 30Å due to the very high heterogeneity of the specimen. The calculated width of the reconstructed JMJD6 is 176Å and the dimension of the depth is roughly half the

dimension for the width. Further model refinement is currently ongoing by modifying the analysis method to improve the resolution of the model, to take the heterogeneity of sample into consideration.

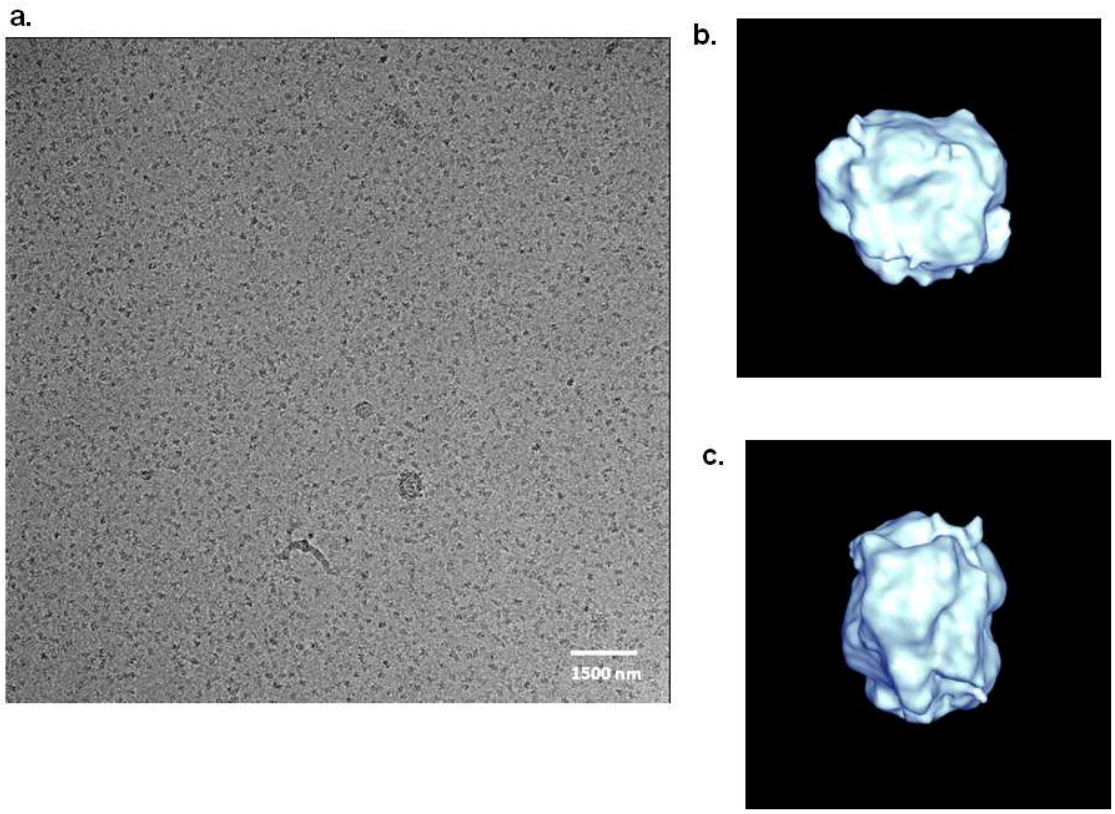


Figure 6.10 Cryo TEM analysis of wild-type JMJD6. (a) TEM micrograph of wild-type JMJD6 obtained after cryo freezing. (b) and (c) View of a particle 3D-reconstruction and visualisation using software package SPIDER. (*This image is provided by Dr. William Nicholson.*)

6.7 Discussion

This chapter describes the first structural analysis carried out on JMJD6 oligomers. The TEM data in this chapter shows that JMJD6 oligomerises giving rise to the ring like structures probably as tetramer of dimers (8 mer) and heptamer of dimers (14 mer). The

possibility of higher order structures cannot be ruled out as multiple oligomeric bands are observed by SDS-PAGE and western blot analysis for recombinant and endogenous JMJD6. On the other hand, analytical size exclusion chromatography indicates the probable formation of a hexamer. It is difficult to assign a particular oligomerisation state to JMJD6 because in the case of size exclusion chromatography, peaks corresponding to higher oligomeric states were also present and they could not be separated from peak assigned as hexamer by S200 Sepharose column. Attempts were made to separate the higher peaks using a S500 Sephacryl size exclusion column but the separation of higher oligomers of JMJD6 was not improved by the use of this column.

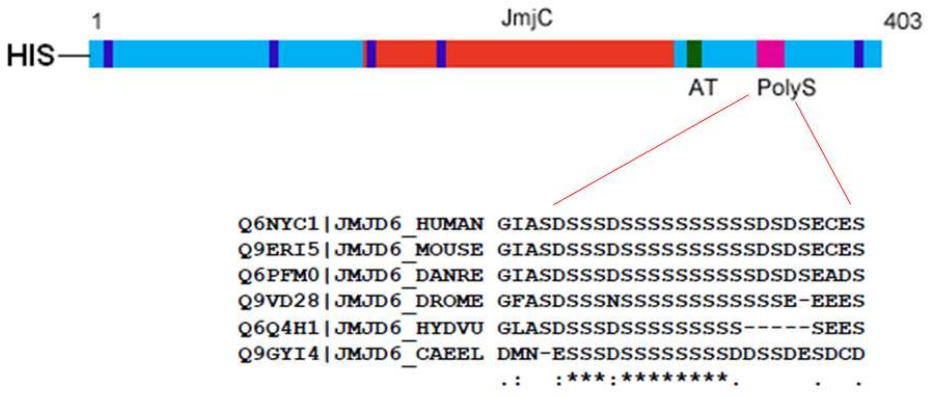


Figure 6.11 Polyserine region of JMJD6 is well conserved in various species. Sequence analysis of polyserine region obtained by ClustalW alignment of human JMJD6 sequence with JMJD6 from *Mus musculus, Danio rerio, Drosophila melanogaster, Hydra vulgaris* and *Caenorhabditis elegans*.

The comparison of the TEM data for His-tagged full length JMJD6, JMJD6$_{1\text{-}343}$ and JMJD6$_{1\text{-}362}$ reveals that the construct His-JMJD6$_{1\text{-}343}$ which lacks the polyserine region forms filamentous structures while the other two form oligomers with ring forms. This leads to the proposal that the polyserine region is either necessary for the formation of JMJD6 oligomer rings or it blocks filament formation. The region containing the

polyserine tract comprising of 16 consecutive serines interspersed with four aspartates is well conserved in JMJD6 as shown in Figure 6.11. It is known that many human proteins have homopolymeric amino acid tracts, which are involved in protein-protein interactions and also have intrinsic polymerisation properties.[22] The presence of a polyglutamine tract has been associated with neurological disorders, including Huntington's disease, spinobulbar muscular atrophy, dentatorubral pallidoluysian atrophy, and six forms of spinocerebellar ataxia.[22, 23] Not much work has been reported and little is known about the function of polyserine tracts even though it is commonly present in human proteins with the number of serines varying from 8 to 42. A recent study identified 59 proteins with polyserine regions in the human genome. In most cases it was proposed that replicative slippage at the DNA level was probably the driving force of their rapid evolution.[24] Many splicing related proteins such as pinin, RNPS1 and SRrm160/300, SRrp37 contain a successive array of serine residues.[25-28] So far, none of these polyserine regions has any assigned biological function. It is also plausible that phosphorylation of this region could interfere with its role in ring formation. Phosphorylation of the *C*-terminal polyserine region of nucleolar transcription factor UBF via the TOR pathway through p70 S6 kinase has been demonstrated.[29] In addition, phosphorylation of serine residues is supposed to be a major control mechanism in pre-mRNA splicing and it is necessary for formation of spliceosomal complex.[30] Thus, it is necessary to carry out further *in vivo* studies to understand the function of the polyserine region of JMJD6. However, the results presented in this chapter suggest that the polyserine region may exist, at least in part to prevent filament formation.

The observed ring-like structure of JMJD6 shows some similarity in shape with a number of protein complexes involved in RNA metabolism such as the exosome complex and the Sm and Sm-like (LSm) protein families.[31] The seven human Sm proteins form a hetero-heptameric ring-like complex, which binds to a conserved site found in single-stranded regions of the U1, U2, U4 and U5 snRNAs.[32, 33] The Sm ring is formed around RNA in the cytoplasm via a dedicated pathway and then transfered to the nucleus to act at the centre of snRNPs in pre-mRNA splicing[34, 35] and direct interaction with RNA occurs through RNA-binding sites within flexible loops facing the central pore of the Sm ring.[36] LSm proteins also form hetero-heptameric rings, but can assemble in the absence of RNA molecules.[37] The LSm rings bind to U6 snRNA also via the central pore[37] and process pre-tRNAs, pre-snoRNAs and pre-rRNAs.[38-41] JMJD6 forms homo-oligomeric rings *in vitro* without any obvious presence of RNA. However, Hong et al. have earlier reported the binding of a random 27 nt single-stranded RNA molecule to recombinant full-length JMJD6 in electrophoretic mobility shift assays (EMSA) and weaker binding of RNA fragment with a JMJD6 deletion mutation lacking the *C*-terminal amino acids 338-403 and thus lacking the polyserine domain.[42]

Thus, although it is clear from this work that JMJD6 is involved in the formation of higher order oligomeric ring like structures, it is necessary to understand if JMJD6 exists in similar form *in vivo* and if this oligomerisation is influenced by the presence of RNA. It is necessary to carry out further *in vivo* work to understand the roles of JMJD6 in splicing and its oligomeric states and the role of the conserved polyserine region.

6.8 References

1. Clifton, I. J.; McDonough, M. A.; Ehrismann, D.; Kershaw, N. J.; Granatino, N.; Schofield, C. J., Structural studies on 2-oxoglutarate oxygenases and related double-stranded β-helix fold proteins. *Journal of Inorganic Biochemistry* **2006,** *100* (4), 644-669.

2. Higgins, L. J.; Yan, F.; Liu, P.; Liu, H.-w.; Drennan, C. L., Structural insight into antibiotic fosfomycin biosynthesis by a mononuclear iron enzyme. *Nature* **2005,** *437* (7060), 838-844.

3. McLuskey, K.; Cameron, S.; Hammerschmidt, F.; Hunter, W. N., Structure and reactivity of hydroxypropylphosphonic acid epoxidase in fosfomycin biosynthesis by a cation- and flavin-dependent mechanism. *Proceedings of the National Academy of Sciences of the United States of America* **2005,** *102* (40), 14221-14226.

4. Woo, E.-J.; Dunwell, J. M.; Goodenough, P. W.; Marvier, A. C.; Pickersgill, R. W., Germin is a manganese containing homohexamer with oxalate oxidase and superoxide dismutase activities. *Nature Structural & Molecular Biology* **2000,** *7* (11), 1036-1040.

5. Lancaster, D. E.; McNeill, L. A.; McDonough, M. A.; Aplin, R. T.; Hewitson, K. S.; Pugh, C. W.; Ratcliffe, P. J.; Schofield, C. J., Disruption of dimerization and substrate phosphorylation inhibit factor inhibiting hypoxia-inducible factor (FIH) activity. *Biochemical Journal* **2004,** *383* (3), 429-437.

6. Tibrewal, N.; Liu, T.; Li, H.; Birge, R., Characterization of the biochemical and biophysical properties of the phosphatidylserine receptor (PS-R) gene product. *Molecular and Cellular Biochemistry* **2007,** *304* (1), 119-125.

7. Hahn, P.; Wegener, I.; Burrells, A.; Böse, J.; Wolf, A.; Erck, C.; Butler, D.; Schofield, C. J.; Böttger, A.; Lengeling, A., Analysis of Jmjd6 Cellular Localization and Testing for Its Involvement in Histone Demethylation. *PLoS One* **2010,** *5* (10), e13769.

8. Han, G.; Li, J.; Wang, Y.; Li, X.; Mao, H.; Liu, Y.; Chen, C. D., The hydroxylation activity of Jmjd6 is required for its homo-oligomerization. *Journal of Cellular Biochemistry* **2012,** *113* (5), 1663-1670.

9. Andrews, P., Estimation of the molecular weights of proteins by Sephadex gel-filtration. *Biochemical Journal* **1964,** *91* (2), 222-233.

10. Dian, E. A.; Rassow, J.; Motz, C., Separation of Proteins by Blue Native Electrophoresis (BN-PAGE). In *The Protein Protocols Handbook* 2009; pp 201-210.

11. Frank, J., *Three-dimensional electron microscopy of macromolecular assemblies: visualization of biological molecules in their native state*. Oxford University Press: 2006.

12. Hayat, M. A., *Principles and Techniques of Electron Microscopy: Biological Applications,*. 4th ed.; Cambridge University Press: 2000; p 564.

13. Williams, D. B., Carter, C. Barry, *Transmission Electron Microscopy: A Textbook for Materials Science*. 2nd ed.; Springer: 2009; Vol. 1, p 760.

14. Ohi, M.; Li, Y.; Cheng, Y.; Walz, T., Negative staining and image classification - Powerful tools in modern electron microscopy. *Biological Procedures Online* **2004,** *6* (1), 23-34.

15. Al-Amoudi, A.; Chang, J.-J.; Leforestier, A.; McDowall, A.; Salamin, L. M.; Norlén, L. P. O.; Richter, K.; Blanc, N. S.; Studer, D.; Dubochet, J., Cryo-electron microscopy of vitreous sections. *EMBO Journal* **2004,** *23* (18), 3583-3588.

16. Jonic, S.; Vénien-Bryan, C., Protein structure determination by electron cryo-microscopy. *Current Opinion in Pharmacology* **2009,** *9* (5), 636-642.

17. Frank, J.; Radermacher, M.; Penczek, P.; Zhu, J.; Li, Y.; Ladjadj, M.; Leith, A., SPIDER and WEB: Processing and Visualization of Images in 3D Electron Microscopy and Related Fields. *Journal of Structural Biology* **1996,** *116* (1), 190-199.

18. Frank, J., Classification of macromolecular assemblies studied as "single particles". *Quarterly Reviews of Biophysics* **1990,** *23* (03), 281-329.

19. Ludtke, S. J.; Baldwin, P. R.; Chiu, W., EMAN: Semiautomated Software for High-Resolution Single-Particle Reconstructions. *Journal of Structural Biology* **1999,** *128* (1), 82-97.

20. Tang, G.; Peng, L.; Baldwin, P. R.; Mann, D. S.; Jiang, W.; Rees, I.; Ludtke, S. J., EMAN2: An extensible image processing suite for electron microscopy. *Journal of Structural Biology* **2007,** *157* (1), 38-46.

21. Shaikh, T. R.; Gao, H.; Baxter, W. T.; Asturias, F. J.; Boisset, N.; Leith, A.; Frank, J., SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. *Nature Protocols* **2008,** *3* (12), 1941-1974.

22. Oma, Y.; Kino, Y.; Sasagawa, N.; Ishiura, S., Intracellular Localization of Homopolymeric Amino Acid-containing Proteins Expressed in Mammalian Cells. *Journal of Biological Chemistry* **2004,** *279* (20), 21217-21222.

23. Davies, J. E.; Rubinsztein, D. C., Polyalanine and polyserine frameshift products in Huntington's disease. *Journal of Medical Genetics* **2006,** *43* (11), 893-896.

24. Huntley, M. A.; Golding, G. B., Selection and Slippage Creating Serine Homopolymers. *Molecular Biology and Evolution* **2006,** *23* (11), 2017-2025.

25. Blencowe, B. J.; Baurén, G.; Eldridge, A. G.; Issner, R.; Nickerson, J. A.; Rosonina, E.; Sharp, P. A., The SRm160/300 splicing coactivator subunits. *RNA* **2000,** *6* (1), 111-120.

26. Ouyang, P., SRrp37, a novel splicing regulator located in the nuclear speckles and nucleoli, interacts with SC35 and modulates alternative pre-mRNA splicing in vivo. *Journal of Cellular Biochemistry* **2009,** *108* (1), 304-314.

27. Sakashita, E.; Tatsumi, S.; Werner, D.; Endo, H.; Mayeda, A., Human RNPS1 and Its Associated Factors: a Versatile Alternative Pre-mRNA Splicing Regulator In Vivo. *Molecular and Cellular Biology* **2004,** *24* (3), 1174-1187.

28. Wang, P.; Lou, P. J.; Leu, S.; Ouyang, P., Modulation of alternative pre-mRNA splicing in vivo by pinin. *Biochemical and Biophysical Research Communications* **2002,** *294* (2), 448-455.

29. Hannan, K. M.; Brandenburger, Y.; Jenkins, A.; Sharkey, K.; Cavanaugh, A.; Rothblum, L.; Moss, T.; Poortinga, G.; McArthur, G. A.; Pearson, R. B.; Hannan, R. D., mTOR-Dependent Regulation of Ribosomal Gene Transcription Requires S6K1 and Is Mediated by Phosphorylation of the Carboxy-Terminal Activation Domain of the Nucleolar Transcription Factor UBF. *Molecular and Cellular Biology* **2003,** *23* (23), 8862-8877.

30. Misteli, T., RNA splicing: What has phosphorylation got to do with it? *Current Biology* **1999,** *9* (6), R198-R200.

31. Pruijn, G. J. M., Doughnuts dealing with RNA. *Nature Structural & Molecular Biology* **2005,** *12* (7), 562-564.

32. Branlant, C.; Krol, A.; Ebel, J. P.; Lazar, E.; Haendler, B.; Jacob, M., U2 RNA shares a structural domain with U1, U4, and U5 RNAs. *EMBO Journal* **1982,** *1* (10), 1259-1265.

33. Liautard, J.-P.; Sri-Widada, J.; Brunel, C.; Jeanteur, P., Structural organization of ribonucleoproteins containing small nuclear RNAs from HeLa cells: Proteins interact closely with a similar structural domain of U1, U2, U4 and U5 small nuclear RNAs. *Journal of Molecular Biology* **1982,** *162* (3), 623-643.

34. Meister, G.; Eggert, C.; Fischer, U., SMN-mediated assembly of RNPs: a complex story. *Trends in Cell Biology* **2002,** *12* (10), 472-478.

35. Yong, J.; Wan, L.; Dreyfuss, G., Why do cells need an assembly machine for RNA protein complexes? *Trends in Cell Biology* **2004,** *14* (5), 226-232.

36. Kambach, C.; Walke, S.; Young, R.; Avis, J. M.; de la Fortelle, E.; Raker, V. A.; Lührmann, R.; Li, J.; Nagai, K., Crystal Structures of Two Sm Protein Complexes and Their Implications for the Assembly of the Spliceosomal snRNPs. *Cell* **1999,** *96* (3), 375-387.

37. Achsel, T.; Brahms, H.; Kastner, B.; Bachi, A.; Wilm, M.; Lührmann, R., A doughnut-shaped heteromer of human Sm-like proteins binds to the 3[prime]-end of U6

snRNA, thereby facilitating U4/U6 duplex formation in vitro. *EMBO Journal* **1999,** *18* (20), 5789-5802.

38. Kufel, J.; Allmang, C.; Petfalski, E.; Beggs, J.; Tollervey, D., Lsm Proteins Are Required for Normal Processing and Stability of Ribosomal RNAs. *Journal of Biological Chemistry* **2003,** *278* (4), 2147-2156.

39. Kufel, J.; Allmang, C.; Verdone, L.; Beggs, J.; Tollervey, D., A complex pathway for 3' processing of the yeast U3 snoRNA. *Nucleic Acids Research* **2003,** *31* (23), 6788-6797.

40. Kufel, J.; Bousquet-Antonelli, C.; Beggs, J. D.; Tollervey, D., Nuclear Pre-mRNA Decapping and 5' Degradation in Yeast Require the Lsm2-8p Complex. *Molecular and Cellular Biology* **2004,** *24* (21), 9646-9657.

41. Kufel, J.; Allmang, C.; Verdone, L.; Beggs, J. D.; Tollervey, D., Lsm Proteins Are Required for Normal Processing of Pre-tRNAs and Their Efficient Association with La-Homologous Protein Lhp1p. *Molecular and Cellular Biology* **2002,** *22* (14), 5248-5256.

42. Hong, X.; Zang, J.; White, J.; Wang, C.; Pan, C.-H.; Zhao, R.; Murphy, R. C.; Dai, S.; Henson, P.; Kappler, J. W.; Hagman, J.; Zhang, G., Interaction of JMJD6 with single-stranded RNA. *Proceedings of the National Academy of Sciences of the United States of America* **2010,** *107* (33), 14568-14572.

Chapter 7. Preliminary inhibition studies on JMJD6

7.1 Introduction

Whole genome sequencing efforts have revealed 2OG oxygenases as being present in most organisms; in humans there are estimated to be more than 60 2OG oxygenase genes.[1] In recent years, new functional roles for human 2OG oxygenases have been discovered, including in biomedicinally important processes such as the hypoxic response, nucleic acid repair, fatty acid metabolism, and chromatin modifications.[2] Several human 2OG oxyegnases are interesting therapeutic targets due to their involvement in cellular processes[3] and their links to cancer,[4] obesity[5] and ischemic diseases. Structural studies on 2OG oxygenases have revealed a conserved DSBH motif and conserved cofactor and cosubstrate binding sites,[6, 7] suggesting that they might be amenable to template based inhibitor design approaches.[8] There is currently an ongoing effort to develop selective chemical probes for *in vivo* studies of the histone demethylases to understand their roles in chromatin regulation.[9, 10] Inhibition studies on HIF hydroxylases[11, 12] and histone demethylases[10, 13, 14] have also been reported.

During the course of this thesis, preliminary inhibition studies on JMJD6 were carried out by screening an in-house set of potential inhibitors against JMJD6 using Differential Scanning Fluorimetry (DSF) and MALDI-MS based assays. MALDI-MS assays were developed to determine IC_{50} values of selected inhibitors including tricarboxylic acid (TCA) cycle intermediates.

7.2 Kinetic studies of JMJD6 catalysed hydroxylation

Various 2OG oxygenases have different requirements with respect to cosubstrate/cofactor concentration; for example, PHD2 requires relatively high ascorbate concentrations for optimal reaction[15] while the K_m value of FIH for 2OG is higher than for most other human 2OG oxygenases.[16, 17] Thus, it is necessary to determine the kinetic parameters to optimise the assay conditions for monitoring inhibitory effects of small molecules on the hydroxylation activity of JMJD6. In order to carry out kinetic studies of JMJD6, a peptide fragment of LUC7L2 ($LUC7L2_{267-278}$) was incubated with JMJD6 along with appropriate cofactor, Fe(II) and cosubstrate, 2OG at 37°C. The extent of hydroxylation was analysed by MALDI-MS using the CHCA matrix.

Assuming that a steady-state reaction occurs (as shown in Equation 7.1), the Michaelis-Menten equation can be applied to determine K_m values for each of the assay components for JMJD6 catalysed hydroxylation reaction. A K_m value is often used to estimate affinity of enzyme for a particular substrate (assuming $k_{-1} >> k_2$ in Equation 7.1), such that the lower the K_m value, the higher the catalytic efficiency of enzyme at lower substrate concentrations. In order to carry out efficient inhibition assays, it is necessary to ensure that assay components are in excess of their K_m values, so that they are not limiting.



$$\frac{d[P]}{dt} = \frac{V_{max}[S]}{K_m + [S]}$$

$$K_m = \frac{k_2 + k_{-1}}{k_1}$$

Equation 7.1 The Michaelis-Menten model for a single-substrate enzyme-catalysed reaction. The model employs steady-state approximation where $k_2 >> k_{-2}$ where E-enzyme, S-substrate, ES enzyme-substrate complex, P-Product, k_x rate constants for reaction stage indicated, [*P*]- Product concentration, V-Reaction rate, V_{max}- maximum rate at saturating substrate concentration, [S]- Substrate concentration, K_m-substrate concentration at which the reaction rate is half of V_{max}.

The rate constant k_2 also referred to as k_{cat} represents the maximum number of substrate molecules turned over by one enzyme molecule throughout a given time span, usually a second.[18] In this model, assuming the reversible formation of an enzyme-substrate intermediate, followed by the irreversible substrate turnover, k_{cat} can be defined as shown in Equation 7.2.

$$V_{max} = k_2[E]_0$$

Equation 7.2 V_{max} under saturating substrate conditions where k_2 is a rate constant as defined in equation 7.1 and $[E]_0$ is the initial enzyme concentration.

7.2.1 Time-course analysis of JMJD6 catalysed hydroxylation

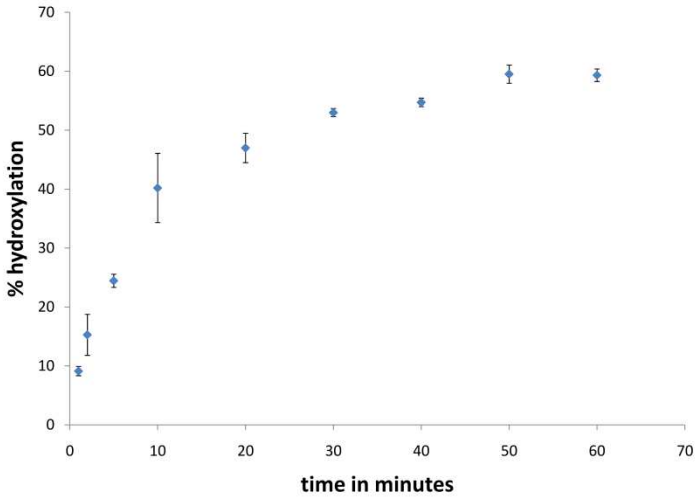


Figure 7.1 Time course analysis of JMJD6 catalysed hydroxylation of LUC7L2$_{267-278}$ peptide determined from intensities of hydroxylated and unhydroxylated peptide peaks obtained by MALDI-MS. Data points represent the mean of triplicate experiments, with error bars showing the standard deviation.

Initially, a time course analysis of JMJD6 catalysed hydroxylation of LUC7L2$_{267-278}$ was carried out at the following time points; 1min, 2min, 5min, 10 min, 20 min, 30 min, 40 min, 50 min and 60 min. Reactions were initiated by mixing 2OG (50 µM), Fe(II) (40 µM), LUC7L2$_{267-278}$ (10 µM) and ascorbate (10 µM) to 2 µM of JMJD6. Three reactions were carried out for each time-point. Reactions were quenched by freezing the solution in liquid nitrogen. The CHCA matrix was added to each Eppendorf vial and 1 µl of mixture was placed on MALDI plate. Figure 7.1 shows that *Vmax* was achieved around 30 minutes under these conditions. Following an initial lag phase, an approximately linear phase between 5-10 minutes was observed.

7.2.2 2OG dependence of JMJD6 for catalysing hydroxylation in LUC7L2$_{267-278}$ peptide

2OG, a TCA cycle intermediate is a cosubstrate for 2OG oxygenases. The K_m value of 2OG for JMJD6 catalysed hydroxylation of LUC7L2$_{267-278}$ was determined by varying the concentration of 2OG from 5 to 100 µM, maintaining other assay components at an excess to ensure that these are not rate limiting. The obtained K_m value of 2OG is about 10 µM (Figure 7.2). However, for monitoring hydroxylation of various peptides in previous chapters, 50 µM of 2OG was found to be required to obtain substantial hydroxylation. The need for higher 2OG concentration in the assays might in part be explained by the possible inhibition of JMJD6 by succinate, the product of 2OG turnover which has been further discussed later in this chapter.

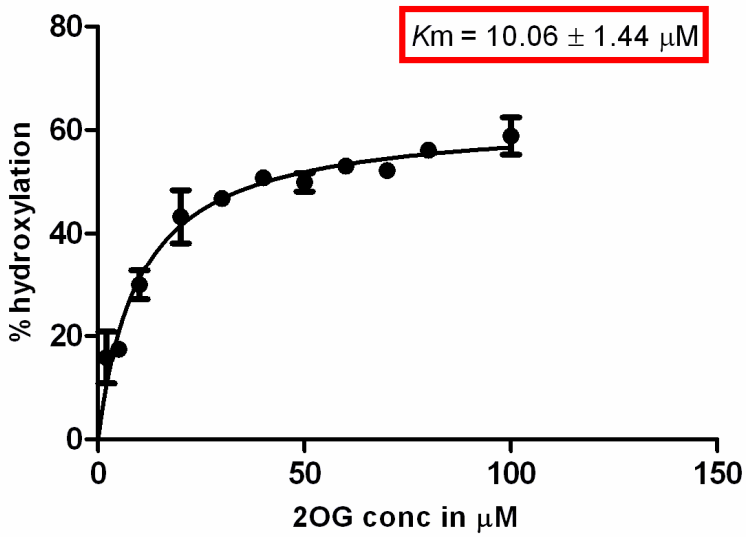


Figure 7.2 Michaelis-Menten plot of 2OG concentrations for JMJD6 catalysed hydroxylation of LUC7L2$_{267-278}$, determined using the MALDI-MS assays. K_m value is indicated on the graphs, and is derived from fitting the Michaelis-Menten equation to the data obtained, using GraphPad Prism. Data points represent the mean of triplicate experiments, with error bars showing the standard deviation.

7.2.3 Fe(II) dependence of JMJD6 for catalysing hydroxylation in LUC7L2$_{267-278}$ peptide

The K_m value of Fe(II) for JMJD6 catalysed hydroxylation of LUC7L2$_{267-278}$ was determined by varying concentrations of Fe(II) from 10 to 80 μM. Assays were carried out in triplicate. Addition of higher concentration of Fe(II) (more than 80 μM) to assay mixture caused precipitation of JMJD6. The K_m value for Fe(II) obtained by fitting the Michaelis-Menten equation to the data points is about 26 μM which is higher than those observed for other human 2OG oxygenases.[19, 20] This property might be related to the self-hydroxylation of JMJD6 as described in Chapter 5.

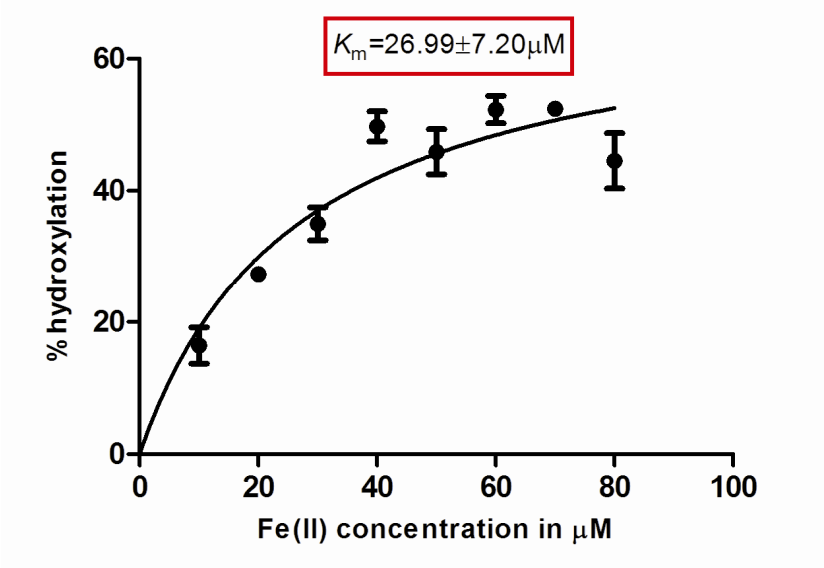


Figure 7.3 Michaelis-Menten plot of Fe(II) concentrations for JMJD6 catalysed hydroxylation of LUC7L2$_{267-278}$, using the MALDI-MS assays. The K_m is derived from fitting the Michaelis-Menten equation to the data obtained, using GraphPad Prism. Data points represent the mean of triplicate experiments, with error bars showing the standard deviation.

7.2.4 JMJD6 does not require ascorbate for hydroxylation

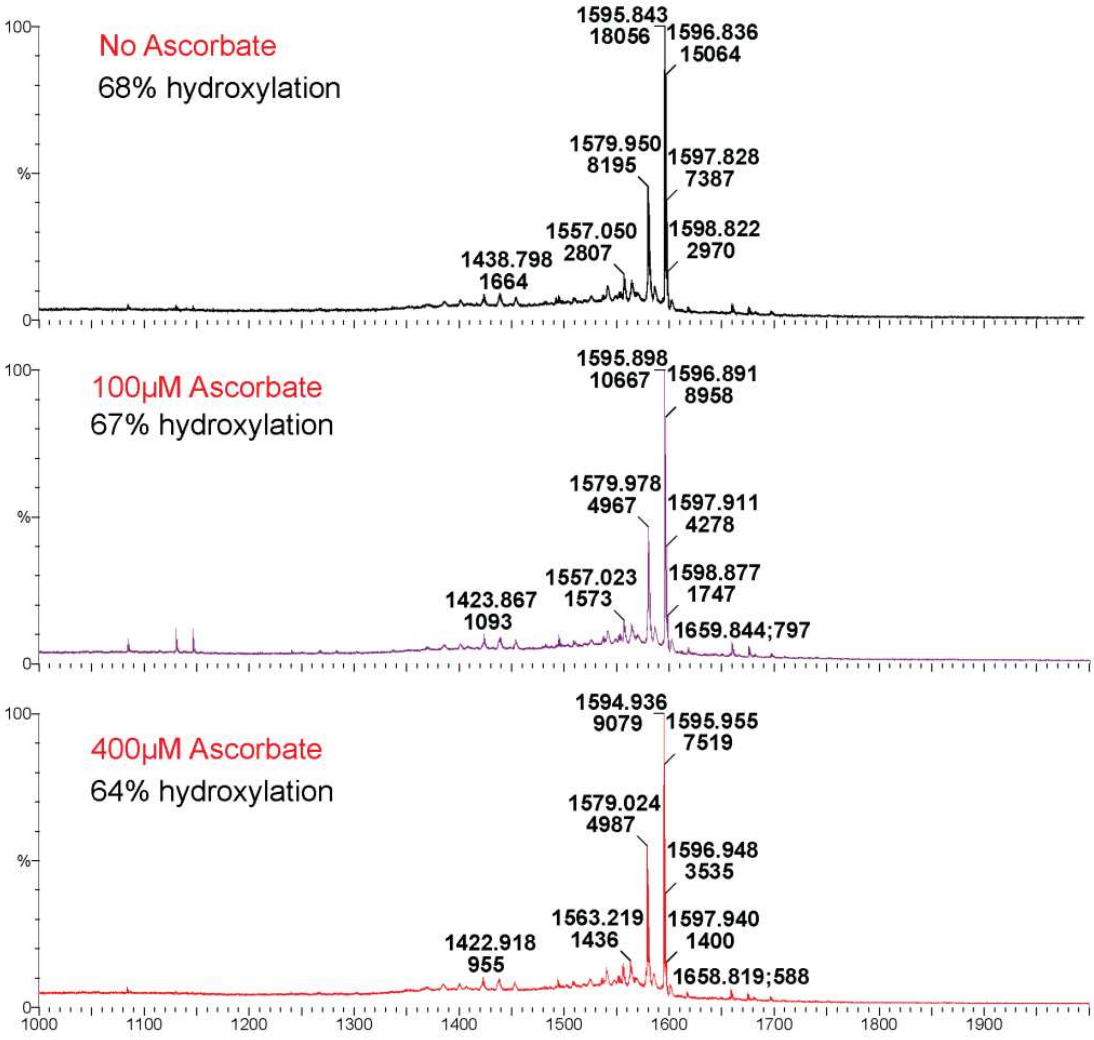


Figure 7.4 MALDI-MS spectra showing that JMJD6 catalysed hydroxylation of LUC7L2$_{267-278}$ peptide is independent of ascorbate concentration. Peaks at 1579 and 1595 correspond to unhydroxylated and hydroxylated peptide, respectively. Percent hydroxylation is calculated from the intensities of these peaks.

Ascorbate is an important factor in many reactions catalysed by 2OG oxygenases. However, the biological role of ascorbate in these reactions is poorly understood. The effect of ascorbate on 2OG oxygenases differs between different enzymes as well as between different substrates for the same enzyme. Some 2OG oxygenase enzymes such

as PHD2 require a higher concentration of ascorbate for their catalytic activity both *in vitro* and possibly, *in vivo*.[15, 21] It has been proposed that one role of ascorbate might be to promote enzyme function by reducing the Fe(III) or Fe(IV) state ions produced at the active site back to the active Fe(II) state.[22] In the case of JMJD6, hydroxylation of the $LUC7L2_{267-278}$ peptide was compared at two different ascorbate concentrations (100 µM and 400 µM) and then compared with the reaction which was carried out without adding any ascorbate. There was no observed difference in hydroxylation activity by adding ascorbate to the assay mixture as observed by MALDI-MS analyses (Figure 7.4).

7.3 Screening of inhibitor library for JMJD6.

7.3.1 Differential scanning fluorimetry (DSF) assays for JMJD6

DSF is a screening method developed by Niesen et al. to identify ligands that can bind to a protein.[23] DSF monitors thermal unfolding of proteins in the presence of a fluorescent dye (typically SYPRO dye (Invitrogen)) and is typically performed using a real time PCR instrument. DSF analysis is based on the principle that the stability of proteins decreases with the increase in temperature as Gibb's free energy of unfolding (ΔG_u) decreases.[24, 25] ΔG_u becomes zero at equilibrium where concentrations of folded and unfolded proteins are equal and the temperature at this point is described as the melting temperature, T_m. Thus, when a ligand binds and stabilises the protein, there is normally an increase in the T_m value. This can be monitored by addition of a hydrophobic fluorescent dye which undergoes fluorescence quenching in the presence of an aqueous environment.[23] Thus, when such a dye is mixed with a protein of interest and protein starts to unfold with the increase in temperature, exposing the hydrophobic

core regions, the dye binds to these regions thereby unquenching the fluorescence. This fluorescence can be measured by addition of a ligand; if a ligand stabilises the protein there will be an increase in T_m which can be monitored by plotting a sigmoidal curve of fluorescence intensity as a function of temperature. One of the advantages of DSF is that one can carry out analysis with smaller amount and lower concentrations of proteins and the method can easily thus be applied to samples which would undergo aggregation at higher concentrations,[23] one of the problems encountered with JMJD6.

DSF analysis of JMJD6 was carried out to screen an in-house library of 132 small molecule inhibitors of 2OG oxygenases in 96 well plates. The structures of the inhibitors screened are shown in Appendix C. The assays were carried out in the presence of 50 mM Tris (pH 7.5) buffer, using Ni(II) instead of Fe(II) such that final concentration of Ni(II) was 1 mM. The final concentration of JMJD6 was 2 µM and that of inhibitor was 250 µM. After adding the SYPRO dye, the melting curves were recorded for an increase in temperature from 5ºC to 95ºC.

7.3.2 Screening of library of inhibitors by MALDI-MS analysis

In order to screen the compounds from in-house library (for which T_m shift values were determined by DSF assays), hydroxylation of the $LUC7L2_{267\text{-}278}$ peptide by JMJD6 was carried out in the presence of inhibitors such that the enzyme to peptide concentration ratio was 1:5 and the peptide to inhibitor ratio was 1:10. The reactions were carried out at higher (50 µM) 2OG concentration (5x K_{m2OG}). Table 7.1 shows the final concentrations in the assay mixture.

Table 7.1 Standard composition of assay mixture used for screening of inhibitors by MALDI-MS assays.

Enzyme Mix	Substrate Mix
2µl JMJD6 (2 µM)	2 µl peptide LUC7L2$_{267-278}$ (10 µM)
2 µl Fe(II) (40 µM)	2 µl 2-OG (50 µM)
2 µl inhibitor (100 µM)	2 µl ascorbate (10 µM)
4µl Tris .HCl (50 mM, pH 7.5)	4 µl Tris.HCl (50 mM, pH 7.5)

The reaction plate was incubated at 37°C for one hour and quenched with methanol. 1 µl of assay mixture from each well was spotted onto a MALDI plate with 1 µl CHCA matrix and MALDI-MS analysis was carried out. The results obtained from T_m shift assays and MALDI-MS analysis are shown in Table 7.2. Assays were carried as duplicate for each inhibitor.

Table 7.2 Results from screening of an in-house library of inhibitors by T_m shift and MALDI-MS analysis. Data values for the T_m shift and MALDI percent hydroxylation are the mean value of duplicate experiments. Values for control reaction were determined by carrying out assays without any inhibitor. MALDI percent hydroxylation is calculated from intensity of hydroxylated and unhydroxylated LUC7L2$_{267-278}$ peptide peaks when MALDI assays were carried out at 100 µM inhibitor and 50 µM 2OG.

Compound	T_m Shift ºC	MALDI % hydroxylation
Control	**-0.02**	**48.0**
EW-T4	-0.46	46.1
OC0353	1.24	11.2
OC0358	-0.55	0
Succinate	2.42	9.3
Fumarate	3.13	15.4
Oxaloacetate	0.61	50.0
Malate	0.31	31.9
Citrate	-0.10	53.9
Isocitrate	-0.05	56.8
Pyruvate	0.14	47.7
D-2-hydroxyglutarate	0.32	46.0
L-2-hydroxuglutarate	0.41	38.5
EW-TNB	0.01	45.0
BIQ	0.44	16.1
BB-286	-0.47	48.0
JM104	-1.39	30.5
JM106	0.06	40.9
JM103	-1.61	28.7
2,3-PDCA	-2.98	59.7
3,4-PDCA	-0.34	48.5
3,5-PDCA	-0.31	47.3
2,6-PDCA	-4.28	55.5
2,5-PDCA	-2.65	58.3
2,4-PDCA	2.88	0
BMRL08012	0.69	33.4
Prohexadione	0.24	53.0
BB212	2.88	5.8
Malathion	-0.33	45.5
Malaoxon	-0.26	47.9
Minoxidil	-0.49	42.6
IS70/NOAL	0.01	51.6
IS71/NOAD	-0.04	50.0
NOG	1.41	43.1
CO-A2-5a	-0.19	48.9
CO-A1-6d	-0.06	43.6
CO-A7-6h	0.21	51.7
CO-B10-7a	-0.18	44.0

CO-C3-7f	0.21	33.3
CO-C10-10a	-1.13	10.2
CO-D8-10k	0.14	46.2
CO-D9-12	-0.37	47.7
CO-C9-8d	-0.35	42.7
CO-D4-8m	-0.09	53.6
NOVD	-0.11	46.3
NOLD	0.26	52.3
NOFD	0.08	51.3
NOVL	-0.1	49.3
NOLL	0	48.1
NOFL	-0.11	48.6
NOYL	-0.01	52.2
NOYD	-0.05	0
EW 29	-0.51	37.3
EW 30	-0.47	39.3
EW 31 / NOCD	-0.07	50.7
EW32 / NOCL	-0.1	47.0
EW 33	-0.11	34.0
EW 35	-0.17	37.0
EW 36	-0.22	47.7
EW 37	-0.04	52.1
ACG-263	0.38	9.5
BB-287	1.23	0
BB-294	-0.13	6.0
BB-315	-0.02	47.1
BB-316	1.59	31.1
JM87	-0.74	24.7
JM110	-1.01	22.7
JM124	-1.78	24.7
JM102	-1.75	28.1
JM109	-0.8	26.6
Ciclopirox Olamine	5.42	0
ISS44	-0.05	40.1
IS45	0.07	0
ACG-114c	0.11	12.0
LBE-332-6-3	-0.03	21.7
LBE-153	0.04	27.6
LBE-314	0.21	28.2
LBE-329	0.33	52.1
LBE-526	2.03	0
LBE-096	0.28	23.3
LBE-060	0.32	23.5
SAHA	-0.2	24.5
TSA	0.24	37.9
N-hydroxyphthalmide	0.57	38.5
IS52	0.63	44.4
Gallate	0.07	8.3
2,4-DHB	-0.25	0

2,5-DHB	-0.31	31.4
3,4-DHB	-0.35	0
R2/quercetin	3.16	45.7
EW-B17	-0.21	50.8
OC0248	2.82	15.6
OC0249	3.24	6.0
OC0206	4.01	0
OC0336	-0.14	50.1
OC0340	-2.41	8.7
BB-103	0.95	33.8
BMRL08014	2.91	8.5
ACG-194b	0.24	58.5
ACG-181a	-0.12	49.3
ACG-255d	0.09	30.2
ACG-253a	-0.02	42.0
ACG-227a	-0.43	25.2
ACG-223a	-0.55	0
ACG-226a	-0.01	41.6
ACG-228a	-0.18	48.1
ACG-237a	0.02	26.7
BB-307	-0.04	39.9
BB-308	-0.38	48.8
BB-318	2.2	32.5
BMRL08050	3.1	7.5
MD228	-0.11	0
MD236	0.14	41.0
MD104	0.92	49.7
OC0344	0	50.3
OC0345	0.1	0
HQNO	0.27	15.8
Daminozide	2.06	21.2
Hydralazine	-6.31	52.7
IS21	-0.47	0
IS37	-0.38	46.2
IS38	-0.62	33.1
IS54	-0.11	20.9
IS84	-0.24	38.2
ACG-148a	-0.08	25.5
ACG-109a	0.04	53.4
ACG-137a	-0.31	45.3
ACG144a	-0.31	45.0
ACG-138a	0.28	48.0
BB-110H	-1.14	41.7
BB-169	-1.03	0
BB-204	-0.4	46.0
IS3	0	23.3
3-hydroxymandelate	0.01	43.7

7.4 Inhibition of JMJD6 by TCA cycle intermediates

2OG oxygenases utilise a Kreb's (or TCA) cycle intermediate, 2OG as a cosubstrate for the catalysis of various reactions. In the TCA cycle, 2OG is produced by the oxidative decarboxylation of isocitrate as catalysed by isocitrate dehydrogenases 1 and 2 (IDH1 and IDH2).[26] In proliferating/cancer cells, excess 2OG enters into the TCA cycle by 'anaplerotic' processes involving the glutaminolysis pathway; in this pathway glutamine is metabolized into 2OG by catalytic activities of glutamate dehydrogenase (GDH)[27, 28] or aminotransferases.[29] Another major anaplerotic pathway that cancer cells employ for energy generation is aerobic glycolysis (a phenomenon known as the Warburg effect),[28] ensuring the high flux of 2OG in cancer cells. The increased amount of 2OG in the TCA cycle leads to the production of oxaloacetate which in turn reacts with acetyl-CoA to generate citrate for lipid synthesis.[30, 31] Further, the lipid synthesis products (i.e. fatty acids, cholesterol, and isoprenoids) are used as building blocks of proliferating/tumour cell membranes.[30, 31]

In some tumours, mutations in succinate dehydrogenases and fumarate hydratase, (enzymes involved in TCA cycle) give rise to an accumulation of succinate and fumarate, which are proposed to stimulate tumor growth by activation of the hypoxic response via inhibition of HIF hydroxylases. On the other hand, mutations in other TCA cycle enzymes, IDH1 and IDH2 are linked to brain tumours and acute leukaemia.[26, 32] Mutations to IDH block the conversion of isocitrate to 2OG,[33] and cause a gain of function comprising catalysis of 2OG to the 2(*R*)-hydroxyglutarate.[32, 34, 35] The accumulation of 2(*R*)-hydroxyglutarate is associated with malignant glioblastoma and acute myeloid leukaemia.[16] Germline mutations in *L*- or *D*-2-hydroxyglutarate

dehydrogenases cause elevated levels of *L*-2(*S*)-hydroxyglutarate or *D*-2(*R*)-hydroxyglutarate; an excess of 2(*S*)-hydroxyglutarate is being correlated with increased risk of brain tumour.[36] Due to their involvement in cellular regulation in cancer cells, TCA cycle intermediates and their related metabolites have been studied as potential inhibitors of 2OG oxygenases.

Thus, during the course of inhibition studies of JMJD6, TCA cycle intermediates including 2(*R*)- and 2(*S*)-hydroxyglutarate were selected to determine T_m shift and IC_{50} values for inhibition of JMJD6 catalysed hydroxylation. The data for T_m shift of TCA cycle intermediates is shown in Figure 7.5.

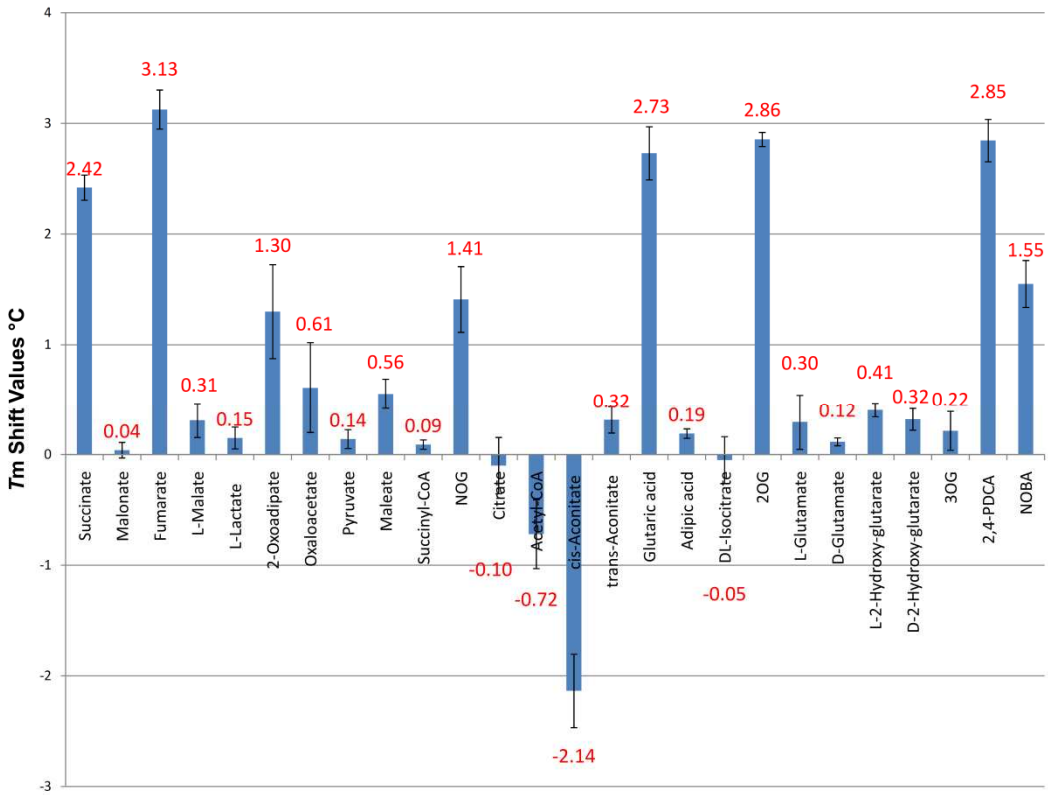


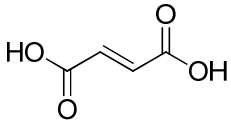
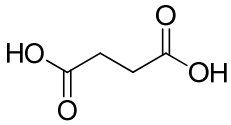
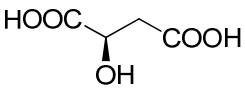
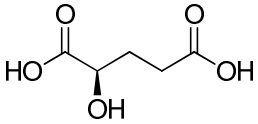
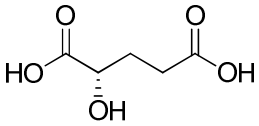
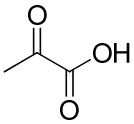
Figure 7.5 Results of T_m shift assays carried out for JMJD6 using TCA cycle intermediates and related compounds. Data points represent the mean of duplicate experiments, with error bars showing the standard deviation.

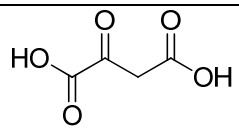
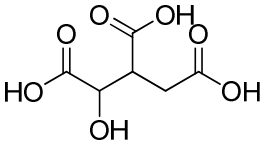
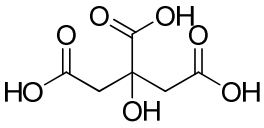
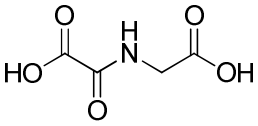
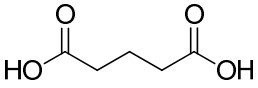
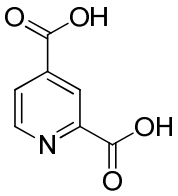
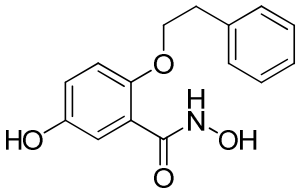
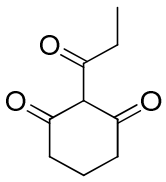
7.5 Development of MALDI-MS assay for determining IC_{50} of various inhibitors

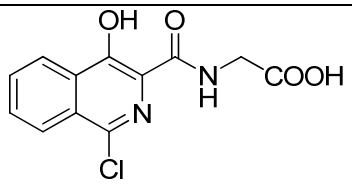
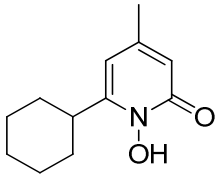
MALDI-MS based assays were further developed to determine the IC_{50} values of selected JMJD6 inhibitors. The $LUC7L2_{267-278}$ peptide substrate was used to measure the activity of JMJD6 and percentage hydroxylation is calculated from MALDI spectra by taking the ratio of hydoxylated state peak intensity to the sum of intensity of peaks corresponding to hydroxylated and unhydroxylated $LUC7L2_{267-278}$ peptide peaks. Although the peptide concentration would ideally be in excess in the assay mixture, the fact that measurement hydroxylation is by relative changes in the proportion of peptide hydroxylation states means that a large excess of peptide would lead to very small relative proportions of hydroxylated peptide products, with associated difficulties in accurate quantification. Thus, the peptide concentration was reduced to a level where a significant proportion of total peptide could be hydroxylated by JMJD6 (the ratio of peptide:enzyme is 5:1). As reported in Section 7.2, the K_m value for 2OG concentration for JMJD6 reaction is about 10 µM and at lower concentrations of 2OG , hydroxylation is very low, thus to get substantial hydroxylation activity (about 55-60%); a 2OG concentration equal to double its K_m value (i.e. 20 µM) was used for final assays. Hydroxylation of $LUC7L2_{267-278}$ by JMJD6 in the presence of Fe(II) and 2OG was determined for various inhibitor concentrations as described in Chapter 8. Three sets of reactions were set up for each inhibitor concentration by incubation at 37°C for 20 minutes. After 20 minutes, each reaction was quenched by adding methanol (1:1 by volume) and then spotted on the MALDI plate with CHCA matrix. IC_{50} values were calculated by plotting the percentage hydroxylation of $LUC7L2_{267-278}$ peptide by JMJD6

against the log of inhibitor concentration; the program GraphPad prism was then used to fit dose-response curve and calculate IC_{50} values.

Table 7.3 Inhibition data of selected inhibitors including TCA cycle intermediates for JMJD6 when 2OG concentration was 2x K_m(2OG).

Name of compound	Structure	IC_{50} µM
Fumarate		13.3
Succinate		7.6
L-Malate		95.0
2(*S*)-hydroxuglutarate		394
2(*R*)-hydroxyglutarate		338.7
Pyruvate		>10000

Oxaloacetate		>10000
Isocitrate		>10000
Citrate		>10000
NOG		4068
Glutaric acid		295
2,4-PDCA		2.4
LBE 526		53.1
BB212		12.2

BIQ		330.2
Ciclopirox olamine		64.2

7.6 Discussion

7.6.1 TCA cycle intermediates

Various studies have shown that succinate and fumarate inhibit 2OG oxygenases with varying degrees of potency with IC_{50} values ranging from the micromolar to milimolar range.[8] The T_m shift data suggest that both fumarate and succinate can bind relatively strongly to JMJD6 (Figure 7.5). The IC_{50} for succinate is 7.6 μM and for fumarate, it is 13.3 μM. Succinate is likely to be a competitive inhibitor with respect to 2OG as it binds to the 2OG and Fe(II) binding site of JMJD2A[37] and FIH.[38] Succinate does inhibit PLOD1, PHD1-3, FIH and JMJD2E as shown in Table 7.4; so far its greatest inhibitory affect had been reported for PHDs.[8] From the IC_{50} values it can be seen that JMJD6 is strongly inhibited by succinate. Interestingly, FIH and JMJD2E which are more closely related to JMJD6 as compared to PHDs,[39] are more weakly inhibited by succinate and fumarate. Inhibition of JMJD6 by succinate does explain the requirement of higher 2OG concentration for substrate hydroxylation in the assays as 2OG is converted to succinate during catalysis by 2OG oxygenases.

The inhibition of HIF hydroxylases by TCA cycle intermediates results in elevated HIF activity upregulating VEGF resulting in angiogenesis which enables tumour growth.[40] An interesting link between VEGF and JMJD6 has recently been reported by Boeckel et al.[41] JMJD6 is found to regulate the splicing of the VEGF receptor 1 (Flt1) which binds to VEGF or PIGF, thereby controlling angiogenic splicing. Further, cell based inhibition studies of JMJD6 can now be focussed on investigating a potential link between inhibition of JMJD6 by fumarate or succinate and angiogenesis.

Table 7.4 Available inhibition data for fumarate and succinate as inhibitors of 2OG oxygenases. IC_{50} or K_i values are given with K_i values in parenthesis. The table has been extended from reference [8].

	C-P4H	CP3H	PLOD1	PHD1	PHD2	PHD3	FIH	JMJD2E	JMJD6
Succinate	(1.5mM)	(0.8mM)	(5-15mM)	(350μM)	19μM	(430μM)	(>10mM)	320μM	7.6μM
Fumarate			2mM	(80μM)	3μM	(50μM)	(>10mM)	2.3mM	13.3μM
Malate			1mM	(>10mM)	(>10mM)	(>10mM)	(>10mM)		95μM
NOG	(0.5μM)			(50μM)	(8μM)	(10μM)	(2μM)	75μM	4mM
2,4PDCA	(2μM)	(3μM)		(40μM)	(7μM)	(8μM)	(30μM)	1.4μM	2.4μM

2(*R*)- and 2(*S*)-hydroxyglutarate were recently reported to be involved in inhibition of HIF hydroxylases and histone demethylases with higher inhibition being observed in the case of histone demethylases suggesting the involvement of 2-hydroxyglutarate inhibition in chromatin modifications.[16] The IC_{50} values for 2(*S*)- and 2(*R*)-hydroxyglutarate show moderate inhibition (394 μM and 338 μM) of JMJD6. The high concentrations of 2(*R*)-hydroxyglutarate (more than 10 mM) that have been observed in malignant brain tumour and leukaemia,[32, 34, 35] might be sufficient to inhibit even the less susceptible 2OG oxygenases.

Other TCA cycle intermediates, oxaloacetate, citrate and pyruvate have been reported to display varying degrees of inhibition for 2OG oxygenases,[8] but from IC_{50} assays for JMJD6, these inhibitors do not seem to inhibit JMJD6. The K_i value for malate has been previously reported to be greater than 10 mM for PHDs and FIH while the IC_{50} value for PLOD1 is reported to be 1mM.[8] Interestingly, on the other hand IC_{50} value for JMJD6 is found to be 94 µM.

7.6.2 Inhibition by other compounds by JMJD6

Inhibition studies carried out on 2OG oxygenases including collagen prolyl hydroxyylases, HIF hydroxylases and histone demthylases have identified *N*-oxalyl glycine (NOG) and dimethyl oxalyl glycine (DMOG) as generic 2OG oxygenase inhibitors.[42, 43] DMOG has been extensively used as a prodrug for NOG to mimic *in vivo* hypoxic conditions.[43] NOG was initially identified as inhibitor for collagen prolyl hydroxylase but later shown to inhibit other 2OG oxygenases as well.[20, 44] Structural studies have shown that NOG occupies the 2OG binding site coordinating Fe(II) via its amido and C-1 carboxylate oxygen, analogous to 2OG.[6, 8] In the case of JMJD6, as seen from Figure 7.5, the T_m shift value for NOG is 1.6ºC which is moderate as compared to that of 2OG (T_m shift= 2.8ºC). The IC_{50} of NOG for JMJD6 is also higher (4032 µM) indicating that NOG is not a very potent inhibitor of JMJD6. This explains the failure of earlier co-crystallisation experiments of JMJD6 with NOG as seen in Chapter 3.

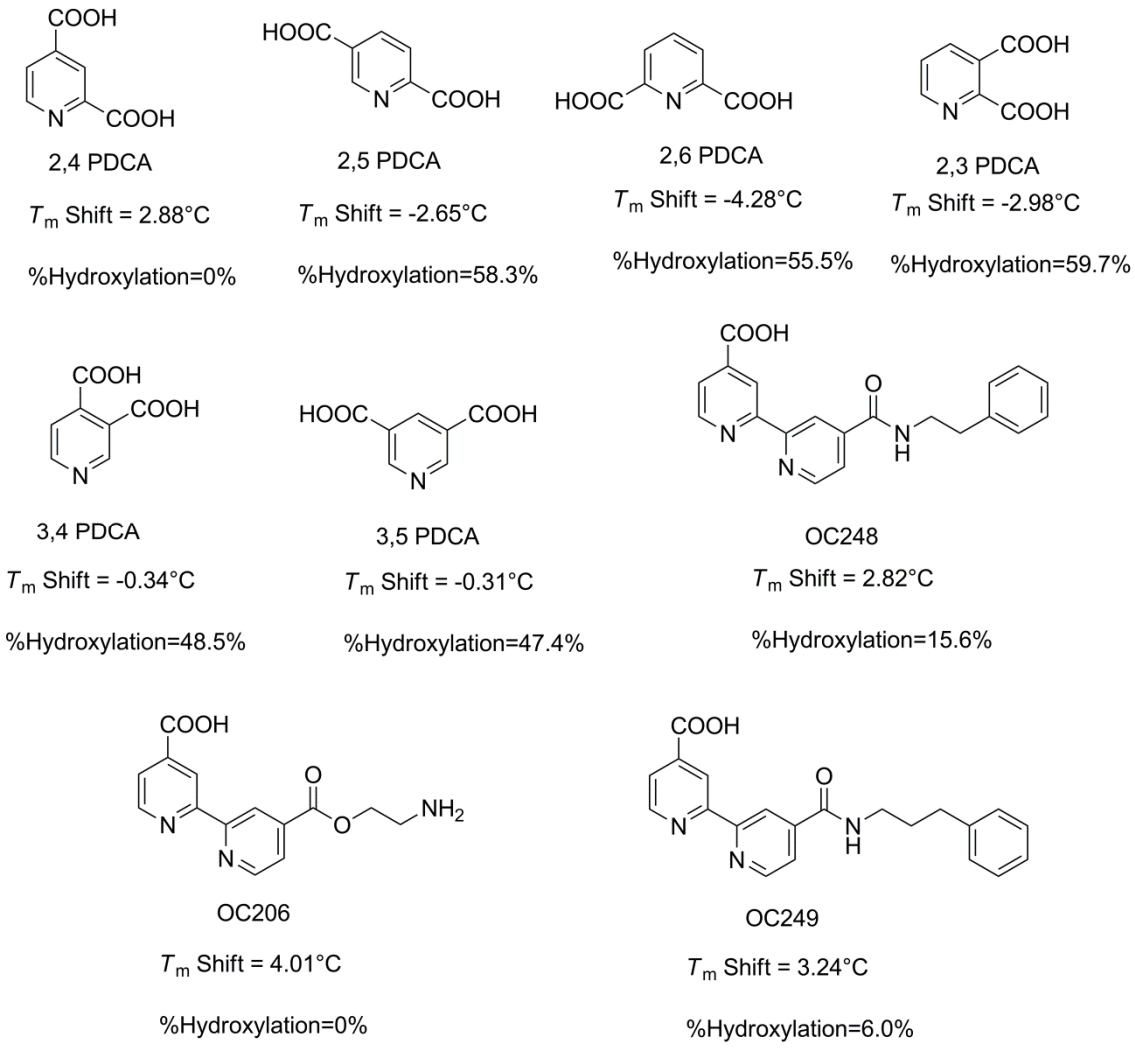


Figure 7.6 Pyridinecarboxylates and bipyridyl compounds from in-house library. T_m shift values and percentage hydroxylation analysed by MALDI-MS spectra of LUC7L2$_{267-278}$ when 100 µM inhibitor and 500 µM 2OG were used.

Pyridine carboxylates are inhibitors of collagen prolyl hydroxylase, PHDs, FIH and histone demthylases.[8, 20] These compounds chelate iron by pyridinyl nitrogen and 2-carboxylate groups. From the structural studies, it has been demonstrated that C-4 carboxylate of 2,4-pyridinedicarboxylate (2,4-PDCA) binds to the conserved basic residue analogous to the binding of C-5 carboxylic acid of 2OG in the active site.[8] Due to the similarity between 2OG and 2,4-PDCA, it is likely that 2,4-PDCA is a generic

inhibitor of 2OG oxygenases. Amongst other pyridine carboxylates, 2,5-PDCA displays some selectivity and is a potent inhibitor of collagen prolyl hydroxylases. 2,6-PDCA displays some inhibitory effects which are dependent on iron concentration in solution as it is very effective at removing Fe(II) from JMJD2 demethylases.[45] T_m shift assays carried out for JMJD6 using pyridinedicarboxylates and bipyridyl derivatives (Figure 7.6 and Table 7.2), shows higher shift in T_m value for 2,4-PDCA and three bipyridyl derivatives (OC248, OC249, OC206 from an in-house library) and these compounds also displayed significant inhibition in hydroxylation of $LUC7L2_{267\text{-}278}$ peptide at a 2OG concentration 5 times its K_m. The IC_{50} for 2,4-PDCA is determined to be 2.4 µM by MALDI-MS assays, indicating potent inhibition of JMJD6 by 2,4-PDCA.

Ciclopiroxolamine is one of the compounds from the screening library which displayed a significant shift in T_m value (5.4°C) and inhibition of JMJD6 catalysed hydroxylation of $LUC7L2_{267\text{-}278}$ peptide by MALDI-MS assays when the 2OG concentration is 5 times its K_m value. It has been reported that ciclopiroxolamine, a simple hydroxamic acid inhibits collagen prolyl hydroxylases[46] and PHD2 using *in vitro* and in cellular assays.[46, 47] In some cases it was shown that such compounds led to stabilisation of HIF-α, induction of VEGF and stimulation of angiogenesis.[47] For JMJD6, the IC_{50} determined for ciclopiroxolamine is 64 µM, displaying moderate inhibition.

7.7 Future Work

This chapter has described preliminary studies of inhibition of JMJD6 by small molecules. In the future, selective inhibitors for JMJD6 could be designed based on these studies. In order to determine binding constants (K_i) for inhibitors, it will be

necessary to optimise and design a new assay that can measure lysine hydroxylation more efficiently. These inhibition results might also be used to optimise crystallisation attempts to more easily obtain structures of JMJD6 with peptide substrate and inhibitors in the active site.

7.8 References

1. Loenarz, C.; Schofield, C. J., Expanding chemical biology of 2-oxoglutarate oxygenases. *Nature Chemical Biology* **2008,** 4, (3), 152-156.

2. Loenarz, C.; Schofield, C. J., Physiological and biochemical aspects of hydroxylations and demethylations catalyzed by human 2-oxoglutarate oxygenases. *Trends in Biochemical Sciences* **2011,** 36, (1), 7-18.

3. Hewitson, K. S.; Schofield, C. J., The HIF pathway as a therapeutic target. *Drug Discovery Today* **2004,** 9, (16), 704-711.

4. Kampranis, S. C.; Tsichlis, P. N., Chapter 4 histone demethylases and cancer. *Advances in Cancer Research* **2009,** 102, 103-169.

5. Tung, Y. C.; Yeo, G. S., From GWAS to biology: lessons from FTO. *Annals of the New York Academy of Sciences* **2010,** 1220, (1), 162-171.

6. McDonough, M. A.; Loenarz, C.; Chowdhury, R.; Clifton, I. J.; Schofield, C. J., Structural studies on human 2-oxoglutarate dependent oxygenases. *Current Opinion in Structural Biology* **2010,** 20, (6), 659-672.

7. Clifton, I. J.; McDonough, M. A.; Ehrismann, D.; Kershaw, N. J.; Granatino, N.; Schofield, C. J., Structural studies on 2-oxoglutarate oxygenases and related double-stranded β-helix fold proteins. *Journal of Inorganic Biochemistry* **2006,** 100, (4), 644-669.

8. Rose, N. R.; McDonough, M. A.; King, O. N. F.; Kawamura, A.; Schofield, C. J., Inhibition of 2-oxoglutarate dependent oxygenases. *Chemical Society Reviews* **2011,** 40, (8), 4364-4397.

9. Frye, S. V., The art of the chemical probe. *Nature Chemical Biology* **2010,** 6, (3), 159-161.

10. Rose, N. R.; Woon, E. C. Y.; Kingham, G. L.; King, O. N. F.; Mecinović, J.; Clifton, I. J.; Ng, S. S.; Talib-Hardy, J.; Oppermann, U.; McDonough, M. A.; Schofield, C. J., Selective Inhibitors of the JMJD2 Histone Demethylases: Combined Nondenaturing Mass Spectrometric Screening and Crystallographic Approaches. *Journal of Medicinal Chemistry* **2010,** 53, (4), 1810-1818.

11. McDonough, M. A.; McNeill, L. A.; Tilliet, M.; Papamicaël, C. A.; Chen, Q.-Y.; Banerji, B.; Hewitson, K. S.; Schofield, C. J., Selective Inhibition of Factor Inhibiting Hypoxia-Inducible Factor. *Journal of the American Chemical Society* **2005,** 127, (21), 7680-7681.

12. Mecinović, J.; Loenarz, C.; Chowdhury, R.; Schofield, C. J., 2-Oxoglutarate analogue inhibitors of prolyl hydroxylase domain 2. *Bioorganic & Medicinal Chemistry Letters* **2009,** 19, (21), 6192-6195.

13. King, O. N. F.; Li, X. S.; Sakurai, M.; Kawamura, A.; Rose, N. R.; Ng, S. S.; Quinn, A. M.; Rai, G.; Mott, B. T.; Beswick, P.; Klose, R. J.; Oppermann, U.; Jadhav, A.; Heightman, T. D.; Maloney, D. J.; Schofield, C. J.; Simeonov, A., Quantitative High-Throughput Screening Identifies 8-Hydroxyquinolines as Cell-Active Histone Demethylase Inhibitors. *PLoS One* **2010,** 5, (11), e15535.

14. Thalhammer, A.; Mecinovic, J.; Loenarz, C.; Tumber, A.; Rose, N. R.; Heightman, T. D.; Schofield, C. J., Inhibition of the histone demethylase JMJD2E by 3-substituted pyridine 2,4-dicarboxylates. *Organic & Biomolecular Chemistry* **2011,** 9, (1), 127-135.

15. Flashman, E.; Davies, S. L.; Yeoh, K. K.; Schofield, C. J., Investigating the dependence of the hypoxia-inducible factor hydroxylases (factor inhibiting HIF and prolyl hydroxylase domain 2) on ascorbate and other reducing agents. *Biochemical Journal* **2010,** 427, (1), 135-142.

16. Chowdhury, R.; Yeoh, K. K.; Tian, Y.-M.; Hillringhaus, L.; Bagg, E. A.; Rose, N. R.; Leung, I. K. H.; Li, X. S.; Woon, E. C. Y.; Yang, M.; McDonough, M. A.; King, O. N.; Clifton, I. J.; Klose, R. J.; Claridge, T. D. W.; Ratcliffe, P. J.; Schofield, C. J.; Kawamura, A., The oncometabolite 2-hydroxyglutarate inhibits histone lysine demethylases. *EMBO Reports* **2010,** 12, (5), 463-469.

17. Ehrismann, D.; Flashman, E.; Genn, D. N.; Mathioudakis, N.; Hewitson, K. S.; Ratcliffe, P. J.; Schofield, C. J., Studies on the activity of the hypoxia-inducible-factor hydroxylases using an oxygen consumption assay. *Biochemical Journal* **2007,** 401, (1), 227-234.

18. Lehninger, A. L.; Nelson, D. L.; Cox, M. M., *Lehninger principles of biochemistry*. 5 ed.; W.H. Freeman: New York, 2005.

19. McNeill, L. A.; Flashman, E.; Buck, M. R. G.; Hewitson, K. S.; Clifton, I. J.; Jeschke, G.; Claridge, T. D. W.; Ehrismann, D.; Oldham, N. J.; Schofield, C. J., Hypoxia-inducible factor prolyl hydroxylase 2 has a high affinity for ferrous iron and 2-oxoglutarate. *Molecular BioSystems* **2005,** 1, (4), 321-324.

20. Rose, N. R.; Ng, S. S.; Mecinović, J.; Liénard, B. t. M. R.; Bello, S. H.; Sun, Z.; McDonough, M. A.; Oppermann, U.; Schofield, C. J., Inhibitor Scaffolds for 2-Oxoglutarate-Dependent Histone Lysine Demethylases. *Journal of Medicinal Chemistry* **2008,** 51, (22), 7053-7056.

21. Knowles, H. J.; Raval, R. R.; Harris, A. L.; Ratcliffe, P. J., Effect of Ascorbate on the Activity of Hypoxia-inducible Factor in Cancer Cells. *Cancer Research* **2003,** 63, (8), 1764-1768.

22. Myllylä, R.; Majamaa, K.; Günzler, V.; Hanauske-Abel, H. M.; Kivirikko, K. I., Ascorbate is consumed stoichiometrically in the uncoupled reactions catalyzed by prolyl 4-hydroxylase and lysyl hydroxylase. *Journal of Biological Chemistry* **1984,** 259, (9), 5403-5405.

23. Niesen, F. H.; Berglund, H.; Vedadi, M., The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols* **2007,** 2, (9), 2212-2221.

24. Schellman, J. A., Temperature, stability, and the hydrophobic interaction. *Biophysical Journal* **1997,** 73, (6), 2960-2964.

25. Privalov, P. L.; Anfinsen, C. B.; Edsall, J. T.; Richards, F. M., Stability of Proteins Small Globular Proteins. In *Advances in Protein Chemistry*, Academic Press: 1979; Vol. Volume 33, pp 167-241.

26. Reitman, Z. J.; Yan, H., Isocitrate Dehydrogenase 1 and 2 Mutations in Cancer: Alterations at a Crossroads of Cellular Metabolism. *Journal of the National Cancer Institute* **2010,** 102, (13), 932-941.

27. DeBerardinis, R. J.; Lum, J. J.; Hatzivassiliou, G.; Thompson, C. B., The Biology of Cancer: Metabolic Reprogramming Fuels Cell Growth and Proliferation. *Cell Metabolism* **2008,** 7, (1), 11-20.

28. Vander Heiden, M. G.; Cantley, L. C.; Thompson, C. B., Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. *Science* **2009,** 324, (5930), 1029-1033.

29. Piva, T. J.; McEvoy-Bowe, E., Oxidation of glutamine in HeLa cells: Role and control of truncated TCA cycles in tumour mitochondria. *Journal of Cellular Biochemistry* **1998,** 68, (2), 213-225.

30. Kannan, R.; Lyon, I.; Baker, N., Dietary Control of Lipogenesis in Vivo in Host Tissues and Tumors of Mice Bearing Ehrlich Ascites Carcinoma. *Cancer Research* **1980,** 40, (12), 4606-4611.

31. Ookhtens, M.; Kannan, R.; Lyon, I.; Baker, N., Liver and adipose tissue contributions to newly formed fatty acids in an ascites tumor. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology* **1984,** 247, (1), R146-R153.

32. Dang, L.; White, D. W.; Gross, S.; Bennett, B. D.; Bittinger, M. A.; Driggers, E. M.; Fantin, V. R.; Jang, H. G.; Jin, S.; Keenan, M. C.; Marks, K. M.; Prins, R. M.; Ward, P. S.; Yen, K. E.; Liau, L. M.; Rabinowitz, J. D.; Cantley, L. C.; Thompson, C. B.; Vander Heiden, M. G.; Su, S. M., Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* **2009,** 462, (7274), 739-744.

33. Zhao, S.; Lin, Y.; Xu, W.; Jiang, W.; Zha, Z.; Wang, P.; Yu, W.; Li, Z.; Gong, L.; Peng, Y.; Ding, J.; Lei, Q.; Guan, K.-L.; Xiong, Y., Glioma-Derived Mutations in IDH1 Dominantly Inhibit IDH1 Catalytic Activity and Induce HIF-1α. *Science* **2009,** 324, (5924), 261-265.

34. Gross, S.; Cairns, R. A.; Minden, M. D.; Driggers, E. M.; Bittinger, M. A.; Jang, H. G.; Sasaki, M.; Jin, S.; Schenkein, D. P.; Su, S. M.; Dang, L.; Fantin, V. R.; Mak, T. W., Cancer-associated metabolite 2-hydroxyglutarate accumulates in acute myelogenous leukemia with isocitrate dehydrogenase 1 and 2 mutations. *The Journal of Experimental Medicine* **2010,** 207, (2), 339-344.

35. Ward, P. S.; Patel, J.; Wise, D. R.; Abdel-Wahab, O.; Bennett, B. D.; Coller, H. A.; Cross, J. R.; Fantin, V. R.; Hedvat, C. V.; Perl, A. E.; Rabinowitz, J. D.; Carroll, M.; Su, S. M.; Sharp, K. A.; Levine, R. L.; Thompson, C. B., The Common Feature of Leukemia-Associated IDH1 and IDH2 Mutations Is a Neomorphic Enzyme Activity Converting α-Ketoglutarate to 2-Hydroxyglutarate. *Cancer Cell* **2010,** 17, (3), 225-234.

36. Dang, L.; Jin, S.; Su, S. M., IDH mutations in glioma and acute myeloid leukemia. *Trends in Molecular Medicine* **2010,** 16, (9), 387-397.

37. Couture, J.-F.; Collazo, E.; Ortiz-Tello, P. A.; Brunzelle, J. S.; Trievel, R. C., Specificity and mechanism of JMJD2A, a trimethyllysine-specific histone demethylase. *Nature Structural & Molecular Biology* **2007,** 14, (8), 689-695.

38. Hewitson, K. S.; Liénard, B. M. R.; McDonough, M. A.; Clifton, I. J.; Butler, D.; Soares, A. S.; Oldham, N. J.; McNeill, L. A.; Schofield, C. J., Structural and Mechanistic Studies on the Inhibition of the Hypoxia-inducible Transcription Factor Hydroxylases by Tricarboxylic Acid Cycle Intermediates. *Journal of Biological Chemistry* **2007,** 282, (5), 3293-3301.

39. Hahn, P.; Bose, J.; Edler, S.; Lengeling, A., Genomic structure and expression of Jmjd6 and evolutionary analysis in the context of related JmjC domain containing proteins. *BMC Genomics* **2008,** 9, (1), 293.

40. Selak, M. A.; Armour, S. M.; MacKenzie, E. D.; Boulahbel, H.; Watson, D. G.; Mansfield, K. D.; Pan, Y.; Simon, M. C.; Thompson, C. B.; Gottlieb, E., Succinate links TCA cycle dysfunction to oncogenesis by inhibiting HIF-Î± prolyl hydroxylase. *Cancer Cell* **2005,** 7, (1), 77-85.

41. Boeckel, J.-N.; Guarani, V.; Koyanagi, M.; Roexe, T.; Lengeling, A.; Schermuly, R. T.; Gellert, P.; Braun, T.; Zeiher, A.; Dimmeler, S., Jumonji domain-containing protein 6 (Jmjd6) is required for angiogenic sprouting and regulates splicing of VEGF-receptor 1. *Proceedings of the National Academy of Sciences of the United States of America* **2011,** 108, (8), 3276-3281.

42. Mole, D. R.; Schlemminger, I.; McNeill, L. A.; Hewitson, K. S.; Pugh, C. W.; Ratcliffe, P. J.; Schofield, C. J., 2-Oxoglutarate analogue inhibitors of hif prolyl hydroxylase. *Bioorganic & Medicinal Chemistry Letters* **2003,** 13, (16), 2677-2680.

43. Elvidge, G. P.; Glenny, L.; Appelhoff, R. J.; Ratcliffe, P. J.; Ragoussis, J.; Gleadle, J. M., Concordant Regulation of Gene Expression by Hypoxia and 2-Oxoglutarate-dependent Dioxygenase Inhibition. *Journal of Biological Chemistry* **2006,** 281, (22), 15215-15226.

44. Stubbs, C. J.; Loenarz, C.; Mecinović□, J.; Yeoh, K. K.; Hindley, N.; Liénard, B. M.; Sobott, F.; Schofield, C. J.; Flashman, E., Application of a Proteolysis/Mass Spectrometry Method for Investigating the Effects of Inhibitors on Hydroxylase Structure. *Journal of Medicinal Chemistry* **2009,** 52, (9), 2799-2805.

45. Majamaa, K.; Hanauske-Abel, H. M.; GÜnzler, V.; Kivirikko, K. I., The 2-oxoglutarate binding site of prolyl 4-hydroxylase. *European Journal of Biochemistry* **1984,** 138, (2), 239-245.

46. Smirnova, N. A.; Rakhman, I.; Moroz, N.; Basso, M.; Payappilly, J.; Kazakov, S.; Hernandez-Guzman, F.; Gaisina, I. N.; Kozikowski, A. P.; Ratan, R. R.; Gazaryan, I. G., Utilization of an In Vivo Reporter for High Throughput Identification of Branched Small Molecule Regulators of Hypoxic Adaptation. *Chemistry & Biology* **2010,** 17, (4), 380-391.

47. Linden, T.; Katschinski, D. M.; Eckhardt, K.; Scheid, A.; Pagel, H.; Wenger, R. H., The antimycotic ciclopirox olamine induces HIF-1α stability, VEGF expression, and angiogenesis. *FASEB Journal* **2003,** 17, (6), 761-763.

Chapter 8. Materials and Methods

8.1 General methods used in chemical synthesis

Chemicals and solvents were from Sigma. Melting points were recorded on a Griffin apparatus. Proton and carbon Nuclear Magnetic Resonance (δ_H, δ_C) spectra were recorded using Bruker AV400 or AVCII500 spectrometers. Chemical shifts are quoted on the δ-scale in ppm using residual solvent as an internal standard. Coupling constant values (J) are reported to the nearest 0.5 Hz. Low resolution mass spectra were recorded on a Micromass Platform 1 spectrometer using electrospray ionisation in either positive or negative ion modes (ES^+ or ES^-) and high resolution mass spectra were recorded using a Waters 2790-Micromass LCT electrospray ionisation mass spectrometer. *M/z* values are reported in Daltons. Thin Layer Chromatography (TLC) was carried out using Merck Kieselgel $60F_{254}$ pre-coated glass-backed plates. Visualisation of the plates was achieved using a U.V. lamp (l_{max} = 254 or 365 nm), and/or ammonium molybdate (5% in 2 M sulphuric acid), or sulphuric acid (5% in ethanol). Flash column chromatography was carried out using Sorbsil C60 40/60 silica.

8.2 General methods used for biological work

Chemicals were supplied by Sigma-Aldrich unless otherwise stated. The water used was from a Millipore Elix® 10 system, and was purified further through a 0.22 µ filter at the outlet by a Millipore MilliQ system.

8.2.1 pH measurement

The pH of solutions was determined using a Hanna Instruments HI 9321 microprocessor pH meter fitted with either a 5 mm or 1.5 mm diameter electrode, calibrated to pH 7.0 and either pH 4.0 or 10.0 using standard calibration solutions (Fisher Scientific) before use. 4 M potassium chloride solution was used as a storage buffer for electrodes.

8.3 Genral methods used for molecular biology

8.3.1 Oligonucleotide primer synthesis and preparation

Oligonucleotide primers were from Sigma-Genosys Ltd. Primers were supplied as freeze-dried pellets which were later resuspended in TE buffer (10 mM Tris.HCl pH 7.5, 1 mM EDTA) to make a 100 µM stock and then diluted as necessary to 10 µM before use.

8.3.2 Plasmid DNA purification and determination of DNA concentration

Plasmid DNA was purified using the QIAprep Spin Miniprep Kit from QIAGEN. DNA concentration was estimated by measuring the absorbance at 260 nm using a Nanodrop® ND-1000 spectrophotometer (NanoDrop® Technologies, Inc.).

8.3.3 DNA sequencing

DNA sequencing was performed by the Geneservice DNA Sequencing facility in the Department of Biochemistry at Oxford University.

8.4 Site directed mutagenesis

Site-directed mutagenesis of plasmids was performed following Stratagene's QuikChange® Site-Directed Mutagenesis protocol. Sequences of all primers used in this work are presented in Appendix C. The forward and reverse primers, each incorporating the mutation and complementary to opposite strands of the vector were used to amplify the sequence in opposite directions by PCR. A typical PCR reaction was set up according to the following protocol from Table 8.1.

Table 8.1 Protocol used for PCR reaction to carry out site directed mutagenesis.

Reagents	Amount
Template plasmid (~50 ng μl^{-1})	1 µl
Forward and reverse primers	125 ng each
dNTP nucleotide mix (2.5 mM each dNTP)	5 µl
PfuTurbo® DNA polymerase (2.5 U/ µl)	1 µl
10 × reaction buffer	5 µl
MilliQ H_2O to final volume of 50 µl	

PCR reaction was carried out using Techne TC-312 PCR thermocycler. The thermal cycler was programmed for cycling parameters as shown in Table 8.2.

Table 8.2 PCR cycling conditions used for site directed mutagenesis.

Temperature	Function	Time	Cycle
95 °C	Denaturation of	2min	1
95 °C	Denaturation	1min	
*54 °C	Annealing	1min	16
72 °C	Extension	1 min/ kb	
72 °C	Final extension	10min	1

*The annealing temperature was varied from 45°C to 60°C depending on the design of the primers.

After a final extension for 10 min at 72°C, the samples were then held at 4°C, before removal and storage at -20°C. Template DNA, which is extracted from a bacterial clone and therefore methylated, was digested with methylation specific *DpnI*. The mutated DNA was then transformed into XL10-GOLD Ultracompetent cells (Stratagene) for growth and subsequent plasmid DNA purification was carried out.

8.5 Microbial techniques

Standard sterile techniques were followed throughout and all the procedures were carried out in a Biological Safety Hood (Heraeus KJ12). All equipments, pipette tips and media used were sterilised by autoclaving (Thermo Life Sciences MAT 490 LEI) at 121ºC for 20 min. Solutions of heat-sensitive substances were sterilized using 20 µ filters. Bacto Agar, Bacto Tryptone and yeast extract were from Difco and Oxoid. IPTG, kanamycin, ampicillin and chloremphenicol were supplied by Melford Laboratories Ltd.

8.5.1 Bacterial growth media

Various growth media were prepared as shown in Table 8.3 and then autoclaved at 121ºC for 20 min before use.

Table 8.3 Recipe for preparing 1 Ltr of LB and 2TY media.

	Bacto Tryptone	Yeats Extract	NaCl
Luria Broth (LB) Media (1Ltr)	10 g	5 g	10 g
2TY Media (1Ltr)	15 g	10 g	5 g

15 g of agar was added prior to autoclaving for solid media and plates were poured in a laminar flow hood. Selective antibiotics (Melford Laboratories Ltd.) were always added post-autoclaving as required (kanamycin to 30 µg ml^{-1}, ampicillin to 100 µg ml^{-1} and chloramphenicol to 30 µg ml^{-1}).

Super optimal broth with catabolite repression **(**SOC) medium was prepared by adding reagents from Table 8.4. Glucose was added to the mixture after autoclaving. Once prepared, the SOC medium was aliquoted and stored at -20ºC.

Table 8.4 Ingradients for preparing 1Ltr of SOC medium

Reagents	Concentration	Volume
2TY	-	985 ml
Glucose	2 M	10 ml
$MgCl_2$	1 M	5 ml

8.5.2 Competent cell transformation

XL10 GOLD Ultracompetent cells were from Stratagene and Rosetta2(DE3) Competent cells were from Invitrogen. The transformations were performed as described in the manufacturer's manual. An aliquot of competent cells was thawed on ice and gently resuspended. 1 µl of plasmid DNA (0.1-50 ng) was mixed with competent cells with careful pipette mixing. The mixture was then incubated on ice for 30 min and then the Eppendorf was slowly swirled in a 42ºC water bath for 30 seconds. The cells were then reincubated on ice for a further 2 min before 1 ml SOC medium was added. The sample was then incubated in a shaker at 37ºC for 1 hour (250 rpm) to let the bacteria recover from the heat shock prior to plating on agar plates. 100 µl of this culture was spread on LB agar plates containing the appropriate antibiotic and/or IPTG. The plate was then incubated at 37ºC overnight.

8.5.3 Preparation of competent cells

E. coli cells were streaked onto LB agar plate and colonies were grown at 37ºC overnight. One colony was used to inoculate 100 ml sterile LB medium at 37ºC overnight. 800 µl of this starter culture was used to inoculate 400 ml LB, which was grown with shaking at 37ºC to an OD_{600} = 0.6-0.8. Cells were centrifuged at 750 x g, at 4ºC for 5 min, and resuspended in 200 ml sterile 100 mM $CaCl_2$, 20% glycerol. After incubation on ice for 20 min, the cells were centrifuged as before and resuspended in 8 ml of 100 mM $CaCl_2$, 20% glycerol. The resulting competent cells were aliquoted (200 µl) in sterile Eppendorfs, then immediately stored at -80ºC.

8.5.4 Bacterial glycerol stocks

Cells were preserved by gently mixing 400 µl of cell culture with 800 µl of sterile 80% glycerol solution. They were frozen rapidly in liquid nitrogen and stored at -80ºC.

8.5.5 Small scale growths

5 ml of 2TY media containing the appropriate antibiotic were inoculated from a single colony on an agar plate or from a glycerol freeze and grown at 37ºC overnight. 1 ml of this starter culture was used to inoculate a 500 ml flask containing 100 ml of 2TY media and the appropriate antibiotic.

8.5.6 Large scale growths

For large scale growth of all the JMJD6 constructs (cloned in pET28b(+) vector (Invitrogen) and expressed from Rosetta2(DE3) cells), four baffled flasks containing 1 ltr 2TY medium each, were inoculated with 1% starter culture in the presence of appropriate antibiotics and incubated at 37ºC in New Brunswick Scientific G24 Environmental incubator shakers at 800 rpm. Cells were grown until an OD_{600} of 0.6-0.8 was achieved; the temperature was then reduced to 25ºC and 0.5 mM IPTG was added to induce the protein expression. Cells were then harvested after 16-18 hours by centrifugation using a JLA 9.100 rotor in a Beckman Avanti J25 centrifuge at 8000 rpm and 4ºC for 10 min. The resulting cell pellet was weighed and stored at -80ºC.

8.6 Expression of Seleno-methionine exchanged JMJD6

Seleno-methionine labelled *N*-terminally His-tagged $JMJD6_{1-343}$ construct cloned in pET-28a(+) vector was expressed from Rosetta2DE(3) cells in the same way as described above.

The Le Master media, used for this expression was prepared as follows:

Part A of this media was prepared by mixing alanine 1.0 g, arginine.HCl 1.16 g, aspartic acid 0.80 g, cysteine 0.06 g, glutamic acid 1.34 g, glutamine 0.66 g, glycine 1.08 g, histidine 0.12 g, leucine 0.46 g, isoleucine 0.46 g, lysine.HCl 0.84 g, phenylalanine 0.26 g, proline 0.2 g, serine 4.16 g, threonine 0.46 g, tyrosine 0.34 g, valine 0.46 g, adenine 1.0 g, guanosine 1.34 g, thymine 0.34 g, uracil 1.0 g, sodium acetate 3.0 g, succinic acid 3.0 g, ammonium chloride 1.50 g, sodium hydroxide 2.16 g and dipotassium mono phosphate 21.0 g in 2 ltr of MilliQ water. The mixture was heated to aid solvation and divided in 500 ml aliquots before autoclaving.

Part B (10% of Le Master Media) of media consists of glucose 20.0 g, magnesium sulphate 0.5 g, iron sulphate ($FeSO_4.7H_2O$) 8.35 g and 16.6 µl of concentrated sulphuric acid in 200 ml of MilliQ water. Part B was filter-sterilised using 2µ sterile syringe filters before being added to the autoclaved media i.e. part A.

8.6.1 100ml starter culture:

Overnight starter cultures were grown in a shaker-incubator at 37ºC using glycerol stock and 100 ml media prepared by following the recipe given in Table 8.5.

Table 8.5 Recipe for preparation of 100 ml starter culture for Seleno-methionine labelled protein.

Reagent	Amount
Le master media PartA	80 ml
2TY	10 ml
Antibiotic (1000x)	100 µl
Kao-Mitchylek vitamin solution	1 ml
Thiamine (1mg/ml)	100 µl
L-Methionine (50mg/ml)	100 µl
Le master media PartB	10 ml

8.6.2 Large scale growth

Cells were grown in the media prepared from the recipe described in Table 8.6, by adding 10 ml of starter culture at 37ºC until an OD_{600} of 0.6 was achieved; the temperature was then reduced to 25ºC and 0.5 mM IPTG was added to induce protein expression. Cells were then harvested after 16-18 hours in the same way as done for the expression of other JMJD6 constructs.

Table 8.6 Recipe for preparation of 1 ltr of media for large growth of *E. coli* expressing Seleno-methionine labelled protein.

Reagent	Amount
Le master media partA	1000 ml
Antibiotic (1000x)	1100 µl
Kao-Mitchylek vitamin solution	11 ml
Thiamine (1mg/ml)	1100 µl
L-Seleno Methionine (10mg/ml)	5.5 ml

Le master media part B	100 ml

8.7 Protein purification

All the protein constructs were purified at 4ºC using a FPLC system. FPLC was carried out on Amersham Biosciences Äkta FPLC systems controlled by Amersham Pharmacia Unicorn Software. All buffers were freshly prepared with MilliQ water and filtered before use. Buffers for purification were prepared as shown in Table. 8.7.

Table 8.7 Recipe to prepare the buffers required for protein purification.

(1 ltr)	NaCl	Tris.HCl (pH 7.5)	Imidazole	Glycerol	TECEP
Binding/Lysis	300 mM	50 mM	10 mM	10%	0.2mM
Elution Buffer	300 mM	50 mM	500 mM	10%	0.2 mM
Gel filtration Buffer	300 mM	50 mM	-	10%	0.2 mM

For protein purification, the cells were thawed and resuspended in lysis buffer (Table 8.7, 1 in 5 dilution of cells in lysis buffer). DNAse I (0.2 mg) was added to the lysis buffer. Cells were lysed by sonication (5-6 $\times$ 45 seconds bursts with 1 min rest between bursts) using a Vibra Cell VCX 500 with a 13 mm probe, on ice. After lysis, the lysate was centrifuged (23000 rpm for 20 min) to remove cell debris. The supernatant was filtered through a 0.2 µ sterile filter (Millipore) prior to loading onto purification columns.

8.7.1 Nickel affinity chromatography

Purifications of all the His6 tagged JMJD6 constructs were carried out using a 5 ml HisTrap FF™ nickel affinity column (GE Healthcare), with buffers pumped at a

flowrate of 1ml/min on an FPLC system. Before each purification, the resin was charged with 3 column volumes (CV) of NiSO4 (400 mM), followed by washing with water (5 CV) and then with lysis buffer (3 CV).

Cell lysate was applied to a nickel-affinity column mounted on an AKTA purifier at 1.5 ml/min flowrate and the column was then washed with lysis buffer until the UV trace (monitored at 280 nm) had returned to background levels. The protein was then eluted in a gradient of 0-100 % elution buffer (Table 8.7) over 8 CV at the flowrate of 1ml/min in 40 min and collected as 5 ml fractions. Fractions were analysed by SDS-PAGE and those judged to contain pure JMJD6 were pooled and concentrated using a 10,000 MWCO Amicon concentrator (Millipore) and buffer exchanged in gel filtration buffer (Table 8.7).

8.7.2 Gel filtration chromatography

JMJD6 fractions purified from nickel affinity chromatography were further purified, if required, using a 300 ml Superdex™ S200 size exclusion chromatography column (GE Healthcare). The column was equilibrated at 3 ml/min with 1 CV of gel filtration buffer and the concentrated protein was loaded onto the column using a Superloop™ (GE Healthcare). The protein was eluted over 1 CV of gel filtration buffer at 2 ml/min, and 5 ml fractions were collected. SDS-PAGE was used to determine which of the fractions contained the purified protein. Fractions with highest purity were pooled and concentrated using a 10,000 MWCO Amicon concentrator (Millipore).

8.7.3 Protein concentration determination

Protein concentrations were measured by measuring the absorbance at 280 nm using an ND-1000 Nanodrop spectrophotometer (Thermo Scientific). All molecular weights and extinction coefficients (ε) of the proteins of interest were calculated using the ProtParam tool. Once the protein solution had reached the desired concentration, it was transferred into 500 ml Eppendorf tubes in the required aliquot sizes, and flash-frozen in liquid nitrogen for storage at -80ºC.

8.8 SDS-PAGE analysis of protein

Protein purity and cell lysate expression levels were analysed by SDS-PAGE. SDS-PAGE gels were prepared as per recipe from Table 8.8 and 8.9. 10 µl of protein was mixed with 10 µl of loading buffer (10% 0.5 M Tris-HCl; pH 6.8, 0.5% bromophenol blue, 2% SDS, 12% Glycerol, 0.5% β-mercaptoethanol, 53% MilliQ water.), boiled at 110°C for 10 min and then loaded on gels. Gels were run on a Bio-Rad Mini Protean II system at a constant potential of 200V using running buffer as prepared following recipe from Table 8.10. Molecular weight markers for SDS-PAGE (SeeBlue™ Plus2) were obtained from Invitrogen. After running the gel, it was placed in staining buffer (0.5% Coomassie Brilliant Blue, 50% methanol, 10% glacial acetic acid) for 10 min and then transferred to destaining buffer (10% methanol, 10% glacial acetic acid) where it was allowed to undergo destaining with constant shaking for 30 min.

Native-PAGE materials were prepared as for SDS-PAGE, with SDS and β-mercaptoethanol replaced by water. Native gels were run at 4ºC and 80V, using appropriate protein standards of known molecular weight and oligomerisation state as

markers for comparison. The gels were dried between Gel Drying film from Promega. Images of gels were recorded (scanned using a CanoScan LiDE 25 scanner).

Table 8.8 Separating gel recipe for preparing two 12.5% acrylamide SDS-PAGE gels.

Reagents	Amount
30% (w/v) acrylamide	4.16 ml
1.5 M Tris-HCl, pH 8.8	2.5 ml
MilliQ water	3.14 ml
10% SDS solution	100 µl
APS solution (100 mg/ml)	100 µl
TEMED	8 µl

Table 8.9 Stacking gel recipe for SDS-PAGE

Reagents	Amount
30% (w/v) acrylamide	0.5 ml
1 M Tris-HCl, pH 6.8	1.25 ml
MilliQ water	3.14 ml
10% SDS solution	50 µl
APS solution (100 mg/ml)	50 µl
TEMED	5 µl

Table 8.10 Recipe for running buffer (10x) at pH 8.8 used in SDS-PAGE analysis.

Reagents	Amount
Tris	250 mM
Glycine	1.92 M
10% SDS solution	10%

Alternatively, the Invitrogen NuPAGE® system was used for analysis of immunoprecipitated protein samples to be prepared for MS analysis. NuPAGE® Bis-Tris 4-12% gels were run in NuPAGE® MOPS SDS-PAGE running buffer, with NuPAGE® antioxidant. Gels were stained using the Colloidal Blue Staining Kit (Invitrogen), following the manufacturers protocol.

8.9 ALPHA Screen assays

All assays were conducted with 400 nM peptide, 400 nM *N*-terminally His-tagged JMJD6, 500:1 dilution of ALPHA screen beads, with all components diluted in ALPHA screen buffer (50 mM Tris, pH 7.5; 0.1% BSA; 0.01% Tween10X). Enzyme and biotin linked peptide fragments were incubated in a 384 well STARLAB384 plate for 1 hour before adding the His-tag affinity and streptavidin beads. The reactions where incubated (1 hour, 25ºC) before resulting emission between 520-620 nm was recorded using an Envision (PerkinElmer) plate reader.

8.10 MALDI-MS assays

Mass analysis of peptides with a molecular mass of less than 4,000 Da was carried out using a MALDI micro MX™ mass spectrometer using a 96 well MALDI sample plate (Waters Corporation). 1 μl of the peptide sample (at a concentration of 1-10 μM, either

as part of a reaction mixture or dissolved in buffer) along with 1 µl of the appropriate matrix (CHCA) was deposited onto the target plate within a circular well and left to air dry before analysis. Spectra were recorded in the reflectron mode following the protocols as described in the MALDI Operator's Manual. Instrument parameters: Laser energy 150, Pulse 2400, Detector 2500, Suppression 1500. PepMix4 (Trypsinogen, Enolase, BSA) in CHCA matrix was used as a calibrant. Spectra were processed using MassLynx™ v4.0.

The JMJD6 catalysed hydroxylation reaction for MALDI assays was initiated by mixing the 100 µM of peptide, 500 µM 2OG, 400 µM Fe(II), 100 µM ascorbic acid in 50 mM Tris.HCl buffer pH 7.5 and 20 µM of JMJD6 together in a final volume of 20 µl and incubated at 37ºC for 30 min. The reaction was quenched by adding 10 µl of 0.1% formic acid.

8.11 Cell biology techniques

8.11.1Cell culture and transfection

Procedures were carried out in a Class II Biological Safety Cabinet (Nuaire Nu 437-400E). HeLa or HEK293T cells were grown in 150 mm Petri dish using Dulbecco's Modified Eagle Medium (DMEM, Lonza) supplemented with glutamine, 100 U/ml penicillin, 100 µg ml^{-1} streptamycin and 10% fetal bovine serum in a Binder BD 53 incubator with a humidified 5% CO_2 atmosphere at 37ºC. Confluent cells were washed with Dulbecco's phosphate-buffered saline (PBS) and allowed to undergo trypsinolysis by adding 0.2 ml/ml trypsin (Lonza) and incubating at 37ºC until cells were released from the plate surface. The reaction was stopped by addition of 10x fresh medium,

which containing the trypsin inhibitors Mg^{2+} and Ca^{2+}. Cells were resuspended fully by pipetting to break up cell clumps, then the appropriate volume of cells was added to new plates containing fresh prewarmed medium.

Two hours prior to transfection, media was removed from cells and exchanged for fresh media to initiate cell division. 2.25 ml of OptiMEM® was added to 50 µg of DNA to be transfected, followed by addition of 90 µl PEI. The mixture was incubated for 15 min at room temperature before adding dropwise to the plated cells. PEI can be toxic to cells so media was exchanged 3 hours post-transfection. Cells were harvested 24 hours post-transfection in PBS either using a rubber cell scraper, or by pipetting to release the cells from the plate surface. The cell suspension was centrifuged at 1000 g, the supernatant removed carefully and the pellet was either kept in ice for immediate use, or frozen at -80ºC.

8.11.2 Cell lysis and immunoprecipitation

Cell lysis was carried out by resuspending cell pellet in 1 ml lysis buffer containing 20 mM Tris.HCl pH 8.0, 500 mM NaCl, 0.5% NP40 supplemented with protease (Pefabloc (Boehringer), pepstatin A, Aprotinin, Leupeptin) and phosphatase inhibitors (100 nM okadaic acid and phosphatase inhibitor cocktails 1 + 2 (Sigma) followed by incubation on ice for 20 min.

Insoluble debris was pelleted by centrifugation at 14000 g, 10 min at 4ºC. Beads used for IP were GFP-TRAP-A (Chromotek), Protein G Sepharose 4 Fast Flow (GE Healthcare). 100 µl beads were equilibrated in lysis buffer lacking NP400 by washing them 3 times with 500 µl of buffer, followed each time by centrifugation at 2,000 rpm,

4ºC for 2 min. 30 µl of anti-JMJD6 antibody (Abcam 10526, Abcam) and anti-GFP antibody (Mouse monoclonal mixture anti-GFP, Roche) 1 in 1000 was added to the cell lysate and further incubated for 1 hour at 4ºC by rotating end-over-end. Cell lysate was applied to pre-equilibrated beads and further rotated end-over-end at 4 ºC for 1 hour. Protein bound beads were washed 3 times with 5 x bead volume of wash buffer (150 mM NaCl, 20 mM Tris.HCl pH 7.5), followed each time by centrifugation at 2000 g and 4ºC for 2 min. For a final wash beads were transferred to a fresh tube to remove any proteins adsorbed to the plastic tube surface. Protein mixture was then mixed with 2x SDS-PAGE loading buffer, boiled at 110ºC for 10 min and then separated on gradient 4-12% NuPAGE (Invitrogen) gel.

8.11.3 In-gel trypsin digestion

Trypsin digestion of samples run on 4-12% gardient SDS-PAGE gels were prepared for MS analysis as described previously.[1, 2] In brief, gel bands stained with Coomassie Blue (Imperial™ Protein Stain, Thermo/Pierce) were cut into small pieces (1 mm^3). Destaining solution (50% methanol, 5% acetic acid) was added, and the solution was shaken (~350 rpm) for 2-3 hours, the de-staining solution was then removed, 200 µl of the same solution added and left overnight under the same conditions. The de-staining solution was then removed, 200 µl of acetonitrile (MeCN) was added and the solution left for 5-15 min (until the gel pieces had shrunk and whitened). The MeCN was then pipetted off and the gel pieces dried by centrifugation *in vacuo* (Heto/Eppendorf vacuum concentrator) for 5 min. Subsequently, reduction was carried out with DTT (10 mM), followed by alkylation with iodoacetamide (100 mM). Trypsin digestion for MS

analysis was then carried out whereby the excised gel pieces were subjected twice to hydration and dehydration using de-staining solution and MeCN, respectively, followed by overnight incubation at 37^{o}C with trypsin (20 ng μl^{-1}) in 100 mM ammonium bicarbonate (pH 7.8) digestion buffer. The final samples were redissolved in 0.1% formic acid (HCOOH), 2% MeCN and stored at -20ºC until analysis.

Digestion of gel bands by Glu-C (Roche), elastase (Sigma) and LysC(Roche) was also carried out in the similar way using appropriate concentration of protease enzyme as mentioned in manufacturer's manual.

8.12 LC-MS/MS analysis

The digested material was subjected to nano-ultra performance liquid chromatography tandem MS analysis (nano-UPLC-MS/MS) using a 75 µm inner diameter x25 cm C_{18} nanoAcquity^{TM} UPLC^{TM} column (1.7 µm particle size; Waters) and a 90 min gradient of 2-45% solvent B (solvent A: 99.9% H_2O, 0.1% formic acid; solvent B: 99.9% MeCN, 0.1% formic acid) on a Waters nanoAcquity UPLC system (final flow rate, 250 nl/min; 7000 p.s.i.) coupled to a Q-TOF Premier tandem mass spectrometer (Waters) run in the positive ion mode. MS analyses were performed in data-directed analysis (DDA) mode (MS to MS/MS switching at precursor ion counts greater than 1 and MS/MS collision energy dependent on precursor ion mass and charge state). Raw MS data were processed using the PLGS software (version 2.3) including deisotoping and deconvolution (converting masses with multiple charge states to m/z = 1). The mass accuracy of the raw data was corrected using Glu-fibrinopeptide (200 fmol μl^{-1}; 700 nl/min flow rate; 785.8426 Da $[M + 2H]^{2+}$) that was infused into the mass spectrometer as a lock mass

during analysis. MS and MS/MS data were calibrated at intervals of 30 s. MS/MS spectra (peak lists) were searched against the UniProtKB/Swiss-Prot database (Version 2010.07.16; 518,415 sequences) database using Mascot version 2.3.01 (Matrix Science) and the following parameters: peptide tolerance, 0.2 Da; $^{13}C = 1$; fragment tolerance, 0.1 Da; missed cleavages, 2; instrument type, ESI-Q-TOF; fixed modification, carbamidomethylation (C); and variable modifications, deamidation (N,Q) and oxidation (M,D,K,N,P,R). Analytical runs were repeated with an inclusion list for identified peptides with the highest ion score if lysyl hydroxylation was detected. All database searches were restricted to human species because of the complexity of the searches when combined with multiple modifications. The interpretation and presentation of MS/MS data were performed according to published guidelines.[3] Assignments of hydroxylation on lysine sites identified by Mascot were verified by manual inspection. Ion chromatograms were extracted using the mass windows of ± 0.1 Da. The local "in-house" Mascot server used for this study is supported and maintained by the Computational Biology Research Group, University of Oxford.

8.13 Peptide synthesis

All the peptides were synthesised on CS-Bio automated solid phase peptide synthesisers (CS-Bio 336S and CS-Bio 336X). Synthesis conditions employed were initially optimised by Dr. J. Offer at Oxford/Scripps Research Laboratory. All amino acids were supplied as α-amine Fmoc protected amino acids. Standard amino acids (Table 8.11) were supplied by CS-Bio.

Table 8.11 Fmoc-protected amino acids used in solid-phase peptide synthesis.

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

Peptide synthesis was carried out in dimethylformamide (DMF), unless otherwise indicated. Fmoc Rink Amide MDHA resin (CS-Bio) was used for synthesis. Typically, synthesis was carried out using 100 µmol resin. Amino acids were used at 10x excess (1 mmol), and activating reagents (DIC, HOBt) were also used at 1 mmol. The reaction started with swelling of 100 µmol resin in 2 ml DMF by shaking for 1 hour. Activation of amino acids was carried out using HOBt (0.5 M solution in DMF) and DIC (0.5 M solution in DMSO). Amino acids were dissolved in DMF. DIC (2 ml) and HOBt (2 ml) solutions were added immediately to the resin. The activation reaction was carried out by shaking for 30 min. Each amino acid was allowed to react with the resin for 2 hours, with constant shaking, after which the resin was washed with DMF. Fmoc group on amino acid was removed by treatment with 20% piperidine for 8 min so that free amino group of amino acid attached to peptide could couple with free carboxylic acid on newly added Fmoc protected amino acid. The next Fmoc-protected activated amino acid was added, and the entire cycle was repeated. After the coupling of last amino acid, the *N*-terminal Fmoc group was removed by 20% piperidine.

After synthesis, the resin was transferred to a filter using DMF and then washed 3x by DCM to remove any remaining DMF from resin. The resin was then dried overnight in a vacuum dessicator. Peptides were cleaved from the resin and all side-chain protecting groups were removed by soaking the resin in 5 ml of cleavage cocktail (2.5% triisopropyl silane in TFA) for 3 hours. For peptides containing methionine or cysteine residues, the cleavage cocktail was amended slightly, to TFA (88%), water (5%), phenol (5%) and triisopropylsilane (2%). After 3 hours, the TFA solution containing free peptide was then filtered and the TFA volume was reduced to 0.5 ml by bubbling N_2 gas

through the solution. The peptide was then precipitated from solution by adding 10 ml ice-cold diethyl ether which was further centrifuged at 2400 rpm for 2 min and supernatant ether was decanted. The solid peptide was then lyophilised in water with 0.1% TFA and freeze dried overnight.

8.14 HPLC purification of peptides

HPLC purifications of peptides were carried out on a Dionex Ultimate 3000 HPLC, with variable wavelength detectors set at 220 nm (peptide bond) and 254 nm. Analytical HPLC of crude and purified peptides was carried out on a Phenomenex Synergi HydroRP 80A 00D-4375-E0 column (C18, 4 µ, 100 mm length, 4.6 mm internal diameter). Preparative HPLC for peptide purification was carried out on a Phenomenex Synergi HydroRP 80A 00D-4375P0-AX column (C18, 4 µ, 100 mm length, 21.20 mm internal diameter). Peptides were purified using water (0.1 % TFA) and acetonitrile (0.1 % TFA) as solvents with the typical purification gradient applied for purification of peptides from RS domain (rich in arginine and serine content) is shown in Table 8.12; as shallow gradient was applied for the purification of these highly basic peptides. Peaks were detected at 220 nm wavelength and collected using a Gilson automated fraction collector, and fractions were later analysed using MALDI-MS. Fractions containing purified peptide (> 90% purity) were pooled and lyophilised in 50 ml Falcon tubes. Peptides were stored either as lyophilised solids or as solutions in reaction buffer (typically 50 mM Tris pH 7.5), at -20ºC.

Table 8.12. Gradient of 0.1% TFA in water and 0.1% TFA in acetonitrile used for separation of peptides.

Time (min)	0.1%TFA in water (%)	0.1%TFA in acetonitrile(%)
0	100	0
10	94	6
15	92	8
25	90	10
30	80	20
35	0	100
40	100	0
45	100	0

8.15 Crystallisation of JMJD6

8.15.1 Crystallisation

A Phoenix^{TM} RE crystallisation robot and Minstrel-HT^{TM} platform with CrystalTrek^{TM} from Rigaku Inc. was used to set up and analyse crystallisation plates (Greiner 96 well CrystalQuick^{TM} SW) for coarse screening and optimisation of crystallisation conditions. SeMet-derivatised JMJD6$_{1-343}$ was crystallized at both 20ºC and 4ºC by the sitting drop vapour diffusion method. Crystals were obtained in 0.3 µl drops (0.2 µl protein (20 mg ml^{-1}) and 0.1 µl reservoir solution) with 80 µl 0.8 M sodium dihydrogen phosphate, 0.8 M potassium dihydrogen phosphate, 0.1 M sodium HEPES pH 7.5 (Molecular Dimension Structure Screen I & II HT-96, condition C4) in the reservoir. Crystals (approximately 70 x 50 µm) formed within 18 hours. On day 9, crystals were transferred directly to 20% ($^{v}/_{v}$) glycerol in well solution and cryo-cooled in liquid nitrogen for storage and data collection.

8.15.2 Data Collection and Processing

Single wavelength anomalous dispersion (SAD) data using 0.9794Å wavelength X-rays were collected at DIAMOND beamline IO4 on a single $JMJD6_{1-343}$ SeMet crystal grown at 20ºC. Data for $JMJD6_{1-343}$ (SeMet) were integrated and scaled with the HKL2000 package.[4] High resolution data were also collected on a single crystal grown at 4ºC and having slightly different unit cell parameters. The high-resolution data were processed with XDS [5] and subsequently scaled using SCALA.[6]

8.15.3 Structure Solution and Refinement

Unit cell parameters were used to calculate Mathews coefficient 2.55 $Å^3$/Dalton indicating two molecules per asymmetric unit. The direct methods program SHELXD [7] was used to locate nine out of ten possible selenium positions using the $JMJD6_{1-343}$ SeMet dataset. The nine sites were refined and initial phases calculated using SHARP (Global Phasing Ltd.). The initial electron density map after solvent flattening with SOLOMON [8] was of excellent quality and automated model building with ARP/wARP[9] resulted in a model where the central parts (60% of residues in the crystallised sequence) of both molecules in the asymmetric unit were built. One subunit from the initial SeMet structure solution was used to locate both molecules in the asymmetric unit of the high resolution data by molecular replacement using PHASER.[10] Further refinement was performed against the $JMJD6_{1-343}$ (HR) data set using PHENIX.[11] After several rounds of manual rebuilding in COOT [12] and subsequent cycles of refinement, the model converged to a final R_{factor} and R_{free} of 19.2% and 22.5%, respectively.

8.15.4 Structure analysis

Structure validation checks were performed using MolProbity.[13] Oligomerisation analysis was performed using PISA.[14] Secondary structure assignments and topology diagrams were generated using PROMOTIF[15] and HERA[16] and manually modified using TopDraw [17] to generate the topology.

8.16 2OG turnover assays

8.16.1 Radioactive 2OG turnover assays

Radioactive 2OG turnover assays measured the release of [^{14}C]-CO_2 from 1-[^{14}C] 2OG.[18] JMJD6 (10 µM) was incubated with 50 µM LUC7L2$_{267-278}$ and without substrate for comparison, in the presence of 400 µM Fe(II), 160 µM 2OG, 1 mM ascorbate and 1 mM dithiothreitol, 0.3 mg ml^{-1} catalase, in 50 mM Tris.HCl pH 7.5 at 37ºC for 30 min. The final reaction volume was 100 µl. Reactions were quenched by adding 200 µl of methanol followed by 30 min incubation on ice.

Reagents were pipetted as three separate spots in the base of a 5 ml tube: enzyme, substrate, and all other reagents prepared as a master mix. A 0.5 ml Eppendorf tube, with its lid removed, containing 200 µl hyamine hydroxide (Perkin Elmer) was placed into the tube, which was sealed with a bung and transferred to a 37°C shaker incubator for typically 30 min. Reactions were quenched by injection of 200 µl methanol and incubated on ice for 20 min to allow absorption of $^{14}CO_2$ by hyamine hydroxide. Following this the exterior of the Eppendorfs were wiped, transferred to scintillation vials, 5 ml scintillant (Opti-phase "Safe", Fisher) added and ^{14}C activity measured in a Beckman LS6500 counter. Assays were carried out as triplicate. Typically wild-type

JMJD6 was used as a positive control, and uncoupled turnover assessed without substrate. In each assay, counts were recorded for the master mix alone. This showed total counts for the known amount of 2OG added, allowing calculation of the number of counts per nmol 2OG used, and then the number of nmol 2OG turned over in the course of the reaction.

8.16.2 2OG turnover by NMR

For ^{1}H-NMR analyses, the assay mixture consisted of 2OG (500 µM), L-ascorbic acid (100 µM) and ferrous ammonium sulphate (100 µM) in 50 mM deuteriated Tris buffer (pH 7.5). A solution of D-Tris (50 mM) solution was made by dissolving deuteriated Tris.HCl (Sigma) in deuterium oxide and adjusting the pH to 7.5 with DCl. 2OG (5 mM) and ascorbic acid (1 mM) solutions were made in 50 mM D-Tris (pH 7.5). $FeNH_4SO_4$ was dissolved in 20 mM deuteriated HCl and then diluted to requisite concentration in deuterium oxide. The reaction was initiated by addition of JMJD6 (20 µM) in a 1mm microprobe NMR tube. The reaction was monitored using a Bruker AVIII 700 spectrometer. 1D ^{1}H-NMR experiments were performed at 37ºC for 48 measurements over 70 min. Readings were recorded at 89 second intervals. The reaction was also carried out without adding ascorbate to the assay mixture.

8.17 NMR analysis of peptides

NMR spectra were recorded using a Bucker AVIII 700 machine (equipment with an inverse cryoprobe optimised for ^{1}H observation) and reported in ppm relative to D_2O ($^{\delta}$H 4.72). The deuterium signal was used as an internal lock signal and the HDO signal was suppressed by pre-saturating its resonance. Samples were prepared in D_2O and

transferred to a 2 mm NMR tube (Brucker). Prior to analysis, the NMR tube was centrifuged for few seconds using a hand centrifuge. Data were analysed using TOPSPIN 2 (Brucker).

8.18 Amino Acid Analysis

8.18.1 Sample preparation for amino acid analysis

For amino acid analysis, the hydroxylation reaction was initiated by mixing 100 µM $LUC7L2_{267-278}$, 500 µM 2OG, 400 µM Fe(II), 100 µM ascorbic acid in 50 mM MOPS buffer pH 7.5 and 20 µM JMJD6, followed by incubation at 37ºC for 1 hour. After testing that the reaction was complete by MALDI-MS, protein was precipitated by adding methanol. The sample was centrifuged (10 min, 14000 rpm) and freeze dried. The assay mixtures were resuspended in 0.1% aqueous formic acid and purified by HPLC using Phenomenex Synergi Hydro RP C18 column. The purified peptide was then suspended in 50 mM MOPS and incubated with 2 µl of trypsin (Promega, 100 µg ml^{-1},) for 2 hours at 37ºC. The trypsin was removed by centrifugation (10 min, 14000 rpm). Carboxypeptidase Y (Roche Applied science, 1 µl, 5 mg ml^{-1}) was then added to the supernatant, and the mixture was incubated at 37ºC (overnight). The Carboxypeptidase Y was removed by centrifugation at (14000 rpm, 10 min). For carrying out acid hydrolysis, an excess of 6 M HCl was added to each sample followed by incubation at 110ºC in an inert atmosphere for 24 hours. All the samples were concentrated by using a vacuum centrifuge (Speed-Vac, Heto).

Modern cow bone collagen digests were carried out using either HCl or the trypsin-Carboxypetidase Y. For the acid catalysed hydrolysis, 400 µl aqueous collagen solution

(3 g ltr^{-1}) was dried in a Pyrex tube, which was placed in an air tight vial with 200 µl of 6 N HCl at 100 mbar pressure and 110ºC for 24 hours. For the enzyme protocol, 50 µl of aqueous collagen solution (3 g ltr^{-1}) was used and the enzyme catalysed hydrolysis procedure was carried out as above.

8.18.2 Derivatisation with AQC and LC-MS analyses

Hydrolysed peptides were reconstituted in borate buffer pH 9.0 (100 µl) and a solution of AQC in acetonitrile (50 µl, twice concentrated in comparison with AccQTag Solution Protocol, Waters). The mixture was reacted for 10 min at 55ºC. LC-MS analyses were performed using Waters Acquity (Waters, USA) ultra performance liquid chromatography system coupled to LCT Premier XE (Waters, USA) orthogonal acceleration time-of-flight mass spectrometer equipped with an electrospray ionization source. Separations were performed using a AccQ-Tag Ultra column (1.7 µm, 100×2.1 mm) (column temperature set at 55ºC) with flow rate of 0.7 ml/min. The injection volume was 1 µl. Samples were kept at 15ºC in an autosampler. The gradient conditions used for separation are given in the Table 8.12.

Table 8.12. The gradient conditions used for amino acid separation.

Time (min)	Percentage of Solvent A(95% (v/v) Water and 5% AccQ Tag Phase A concentrate	Percentage of solvent B (98.7% (v/v) acetonitrile and 1.3% (v/v) formic acid).
0.0	99.9	0.1
0.54	99.9	0.1
9.74	90.9	9.1
11.74	78.8	21.2

12.04	40.4	59.6
12.05	10.0	90.0
12.64	10.0	90.0
12.73	99.9	0.1
13.50	99.9	0.1

The conditions for ESI-MS detection were as follows: Positive-ion mode, V-mode analyser, desolvation temperature at 150ºC, source temperature at 120ºC, capillary voltage at 3000V, sample cone voltage at 25 V, cone and desolvation gas flow at 30 and 800 ltr/min respectively. The spectra for time span from 8 to 9 min were generated and centred at the median. The extracted ion chromatogram was created at the mass corresponding to bis-derivatised hydroxylysine (503.1 Da). Four chromatograms were acquired for every sample.

8.19 Transmission Electron Microscopy

8.19.1 Negative staining EM

For structural analysis of all JMJD6 constructs by negative staining, 3 µl of about 1mg ml^{-1} protein solution in 300 mM NaCl, 50 mM Tris, 5% glycerol, pH 7.5 was applied to a hydrophilic glow discharged copper grid coated with a thin carbon film. Grids were washed with 2 drops of water to remove excess detergents, blotted with filter paper, negatively stained with 0.75% uranyl formate solution, and air dried. Micrographs were recorded on Kodak SO 163 film using a Philips CM120 electron microscope equipped with a LaB6 filament (FEI) operating at 120 kV under low dose conditions with a

magnification of 45,000. The defocus values were 800 nm. Micrographs were digitised with a Super Coolscan 9000 ED (Nikon) using a step size of 25 µm.

8.19.2 Cryo -EM

3 µl of JMJD6 (about 5 mg ml^{-1}) were placed on a 400-mesh copper grid coated with a thin holey-carbon film. The grid was blotted, vitrified by freeze plunging in liquid ethane, and transferred at liquid nitrogen temperature into the microscope using a side-entry nitrogen cooled Gatan 626 cryoholder. All images in cryo conditions were taken on a JEM-2100F (JEOL) using an acceleration voltage of 200 kV at a magnification of 67,000 and a 4Kx4K CCD camera. The microscope has a spherical aberration constant (Cs) of 0.5 mm. The CCD images are 4096 by 4096 pixels and the CCD elements are 15µ across.

8.20 DSF analysis

DSF was used to determine T_m values for different inhibitors. DSF was performed using a MiniOpticon™ Real-Time PCR Detection System (BioRad). SYPRO orange (Invitrogen) dye was used to monitor unfolding of proteins. The protocol was developed from that presented in reference[19]. The corresponding changes in fluorescence were monitored as a function of time. $NiSO_4$ was used as a source of Ni(II) ions to final concentration of 10 µM. Inhibitor concentration used was 250 µM and JMJD6 was 2 µM. The Orange dye stock solution in DMSO (5000x concentrate) was first diluted by adding 1 µl to 2.5 ml of 50 mM Tris.HCl, pH 7.5 to produce 2x concentrated solution. This SYPRO Orange buffer solution was used to make up an enzyme mixture solution to get a final assay concentration of 2 µM. The fluorescence was recorded at every 1ºC

in the range 20-95ºC, with the temperature increased linearly by 1ºC/min. The software provided was used to perform global minimum subtraction. For determination of T_m values, melting curves for each triplicate data set were exported into GraphPad Prism. A Boltzmann curve was fitted between the values of minimum and maximum intensities. The midpoint (i. e. the T_m value) was calculated by the software.

8.21 MALDI-MS assays to determine IC_{50} values of inhibitors

Stock solutions of selected inhibitors were prepared by serial dilutions as follows: 100 mM, 30 mM, 10 mM, 3 mM, 1 mM, 300 µM, 100 µM, 30 µM, 10 µM, 3 µM and 1 µM. 100 mM stock solution of inhibitor was prepared in DMSO, while other dilutions were made using 50 mM Tris.HCl, pH 7.5. Final in-assay concentration of DMSO was 5% of assay mix. Assays were carried out in 96 well plates as triplicates for each inhibitor concentration. Assay mixture was prepared by mixing Fe(II) (40 µM), ascorbate (40 µM), 2OG (20 µM), and $LUC7L2_{267-278}$ peptide (10 µM) in 50 mM Tris.HCl, pH 7.5. 2 µl of inhibitor stock solution were added in each well. Reaction was initiated by mixing 2 µM JMJD6 and further incubated for 30 min at 37ºC, before 1:1 quenching with methanol followed by addition of four volumes of 20 mM triammonium citrate. 1 µl of the diluted assay mixture was then mixed with 1 µl of CHCA and spotted onto the MALDI plate before analysis. Relative intensities of hydroxylated and unhydroxylated peptide was determined by MALDI spectra and further used to calculate percentage hydroxylation. IC_{50} values were calculated by plotting percentage hydroxylation of $LUC7L2_{267-278}$ peptide by JMJD6 against log of inhibitor concentration, GraphPad prism was then used to fit dose-response curve and to calculate IC_{50} values.

8.22 References

1. Kinter, M.; Sherman, N. E., *Protein sequencing and identification using tandem mass spectrometry*. Wiley Interscience, New York.: 2000.

2. Mackeen, M. M.; Kramer, H. B.; Chang, K.-H.; Coleman, M. L.; Hopkinson, R. J.; Schofield, C. J.; Kessler, B. M., Small-Molecule-Based Inhibition of Histone Demethylation in Cells Assessed by Quantitative Mass Spectrometry. *Journal of Proteome Research* **2010,** *9* (8), 4082-4092.

3. Taylor, G. K.; Goodlett, D. R., Rules governing protein identification by mass spectrometry. *Rapid Communications in Mass Spectrometry* **2005,** *19* (23), 3420-3420.

4. Otwinowski Z., M. W., Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology* **1997,** *276* (part A), 307-326.

5. Kabsch, W., Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *Journal of Applied Crystallography* **1993,** *26* (6), 795-800.

6. Collaborative Computational Project Number 4, The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallographica Section D, Biological ctystallography* **1994,** *50* (Pt. 5), 760-763.

7. Schneider, T. R.; Sheldrick, G. M., Substructure solution with SHELXD. *Acta Crystallographica Section D, Biological ctystallography* **2002,** *58* (10), 1772-1779.

8. Abrahams, J. P.; Leslie, A. G., Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallographica Section D, Biological ctystallography* **1996,** *52* (Pt.1), 30-42.

9. Perrakis, A.; Harkiolaki, M.; Wilson, K. S.; Lamzin, V. S., ARP/wARP and molecular replacement. *Acta Crystallographica Section D, Biological ctystallography* **2001,** *57* (10), 1445-1450.

10. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., Phaser crystallographic software. *Journal of Applied Crystallography* **2007,** *40* (Pt 4), 658-674.

11. Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L.-W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C., PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallographica Section D, Biological ctystallography* **2002,** *58* (Pt. 11), 1948-1954.

12. Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D, Biological ctystallography* **2004,** *60* (12), 2126-2132.

13. Davis, I. W.; Leaver-Fay, A.; Chen, V. B.; Block, J. N.; Kapral, G. J.; Wang, X.; Murray, L. W.; Arendall, W. B.; Snoeyink, J.; Richardson, J. S.; Richardson, D. C., MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Research* **2007,** *35* (suppl 2), W375-W383.

14. Krissinel, E.; Henrick, K., Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology* **2007,** *372* (3), 774-797.

15. Hutchinson, E. G.; Thornton, J. M., PROMOTIF - A program to identify and analyze structural motifs in proteins. *Protein Science* **1996,** *5* (2), 212-220.

16. Hutchinson, E. G.; Thornton, J. M., HERA - A program to draw schematic diagrams of protein secondary structures. *Proteins: Structure, Function and Genetics* **1990,** *8* (3), 203-212.

17. Bond, C. S., TopDraw: a sketchpad for protein structure topology cartoons. *Bioinformatics* **2003,** *19* (2), 311-312.

18. Hewitson, K. S.; McNeill, L. A.; Riordan, M. V.; Tian, Y.-M.; Bullock, A. N.; Welford, R. W.; Elkins, J. M.; Oldham, N. J.; Bhattacharya, S.; Gleadle, J. M.; Ratcliffe, P. J.; Pugh, C. W.; Schofield, C. J., Hypoxia-inducible Factor (HIF) Asparagine Hydroxylase Is Identical to Factor Inhibiting HIF (FIH) and Is Related to the Cupin Structural Family. *Journal of Biological Chemistry* **2002,** *277* (29), 26351-26355.

19. Niesen, F. H.; Berglund, H.; Vedadi, M., The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols* **2007,** *2* (9), 2212-2221.

Appendix A:

Table A1. Library of peptides from Histone H3 sequence used in Chapter 2 for ALPHA screen assays, where Δ= acetyl lysine, Φ= monomethyl lysine, Π=dimethyl lysine, Θ= trimethyl lysine, Σ= phospho serine, Ω= phosphor threonine, Ξ= monomethyl arginine, Ψ= asymmetric dimethyl arginine.

Histone 1-21	ARTKQTARKSTGGKAPRKQLA
R2me	AΞTKQTARKSTGGKAPRKQLA
R2me2	AΨTKQTARKSTGGKAPRKQLA
R2me2/T3P	AΨΩKQTARKSTGGKAPRKQLA
R2me2/T3P/K4me3	AΨΩΘQTARKSTGGKAPRKQLA
R2me2/K4me3	AΨTΘQTARKSTGGKAPRKQLA
T3P	ARΩKQTARKSTGGKAPRKQLA
T3P/K4me3	ARΩΘQTARKSTGGKAPRKQLA
K4me1	ARTΦQTARKSTGGKAPRKQLA
K4me2	ARTΠQTARKSTGGKAPRKQLA
K4me3	ARTΘQTARKSTGGKAPRKQLA
K4me3/K9Ac	ARTΘQTARΔSTGGKAPRKQLA
K4me3/K9me3	ARTΘQTARΘSTGGKAPRKQLA
K9Ac	ARTKQTARΔSTGGKAPRKQLA
K9Ac/S10P	ARTKQTARΔΣTGGKAPRKQLA
K9Ac/T11P	ARTKQTARΔSΩGGKAPRKQLA
K9Ac/S10P/T11P	ARTKQTARΔΣΩGGKAPRKQLA
K9me	ARTKQTARΦSTGGKAPRKQLA
K9me2	ARTKQTARΠSTGGKAPRKQLA
K9me3	ARTKQTARΘSTGGKAPRKQLA
K9me3/T11P	ARTKQTARΘSΩGGKAPRKQLA
K9me3/S10P	ARTKQTARΘΣTGGKAPRKQLA
K9me3/T11P/S10P	ARTKQTARΘΣΩGGKAPRKQLA
T11P	ARTKQTARKΣTGGKAPRKQLA
S10P	ARTKQTARKSΩGGKAPRKQLA
S10P/T11P	ARTKQTARKΣΩGGKAPRKQLA
S10P/T11P/K14Ac	ARTKQTARKΣΩGGΔAPRKQLA
S10P/T11P/K14me3	ARTKQTARKΣΩGGΘAPRKQLA
T11P/K14Ac	ARTKQTARKSΩGGΔAPRKQLA
T11P/K14me3	ARTKQTARKSΩGGΘAPRKQLA
K14Ac	ARTKQTARKSTGGΔAPRKQLA
K14Ac/R17me2	ARTKQTARKSTGGΔAPΨKQLA
K14me	ARTKQTARKSTGGΦAPRKQLA
K14me2	ARTKQTARKSTGGΠAPRKQLA
K14me3	ARTKQTARKSTGGΘAPRKQLA
K14me3/R17me2	ARTKQTARKSTGGΘAPΨKQLA
R17me	ARTKQTARKSTGGKAPΞKQLA
R17me2	ARTKQTARKSTGGKAPΨKQLA
Histone 14-34	KAPRKQLATKAARKSAPATGG

R17me2/K18Ac	KAPΨΔQLATKAARKSAPATGG
K18Ac	KAPRΔQLATKAARKSAPATGG
K18Ac/K23me3	KAPRΔQLATΘAARKSAPATGG
K18Ac/K23Ac	KAPRΔQLATΔAARKSAPATGG
K18Ac/K23Ac	KAPRΔQLATΔAARKSAPATGG
K23Ac	KAPRKQLATΔAARKSAPATGG
K23Ac/R26me3	KAPRKQLATΔAAΨKSAPATGG
K23Ac/R26me2/K27Ac	KAPRKQLATΔAAΨΔSAPATGG
K23me	KAPRKQLATΦAARKSAPATGG
K23me2	KAPRKQLATΠAARKSAPATGG
K23me3	KAPRKQLATΘAARKSAPATGG
K23me3/R26me2	KAPRKQLATΘAAΨKSAPATGG
K23me3/K26me3/K27Ac	KAPRKQLATΘAAΨΔSAPATGG
R26me	KAPRKQLATKAAΞKSAPATGG
R26me2	KAPRKQLATKAAΨKSAPATGG
R26me2/K27Ac	KAPRKQLATKAAΨΔSAPATGG
R26me2/K27Ac/S28P	KAPRKQLATKAAΨΔΣAPATGG
R26me2/S28P	KAPRKQLATKAAΨKΣAPATGG
K27Ac	KAPRKQLATKAARΔSAPATGG
K27Ac/S28P	KAPRKQLATKAARΔΣAPATGG
K27me	KAPRKQLATKAARΦSAPATGG
K27me2	KAPRKQLATKAARΠSAPATGG
K27me3	KAPRKQLATKAARΘSAPATGG
R26me2/K27me3	KAPRKQLATKAAΨΘSAPATGG
K27me3/S28P	KAPRKQLATKAARΘΣAPATGG
R26me2/K27me3/S28P	KAPRKQLATKAAΨΘΣAPATGG
S28P	KAPRKQLATKAARKΣAPATGG
S28P/T32P (18-38)	KQLATKAARKΣAPAΩGGVKKP
Histone 28-48	SAPATGGVKKPHRYRPGTVAL
T30P	SAPAΩGGVKKPHRYRPGTVAL
K36me	SAPATGGVΦKPHRYRPGTVAL
K36me2	SAPATGGVΠKPHRYRPGTVAL
K36me3	SAPATGGVΘKPHRYRPGTVAL
T30P/K36me3	SAPAΩGGVΘKPHRYRPGTVAL
K37me	SAPATGGVKΦPHRYRPGTVAL
K37me2	SAPATGGVKΠPHRYRPGTVAL
K37me3	SAPATGGVKΘPHRYRPGTVAL
T30P/K37me3	SAPAΩGGVKΘPHRYRPGTVAL
K36me3/K37me3	SAPATGGVΘΘPHRYRPGTVAL
T30P/K36me3/K37me3	SAPAΩGGVΘΘPHRYRPGTVAL
Histone 70-91	VREIAQDFKTDLRFQSSAVMA
K79me	VREIAQDFΦTDLRFQSSAVMA
K79me2	VREIAQDFΠTDLRFQSSAVMA
K79me3	VREIAQDFΘTDLRFQSSAVMA
Histone 110-131	CAIHAKRVTIMPKDIQLARRI
K115Ac	CAIHAΔRVTIMPKDIQLARRI
K115Ac/T118P	CAIHAΔRVΩIMPKDIQLARRI

T118P	CAIHAKRVΩIMPKDIQLARRI
K122Ac	CAIHAKRVTIMPΔDIQLARRI
T118P/K122Ac	CAIHAKRVΩIMPΔDIQLARRI
K115Ac/T118P/K112Ac	CAIHAΔRVΩIMPΔDIQLARRI
Histone 115-135	KRVTIMPKDIQLARRIRGERA
R128me	KRVTIMPKDIQLAΞRIRGERA
R128me2	KRVTIMPKDIQLAΨRIRGERA
K122Ac/R128me2	KRVTIMPΔDIQLAΨRIRGERA

Table A2: Library of peptides from Histone H4 and H2 sequence used in Chapter 2 for ALPHA screen assays where Δ= acetyl lysine, Φ= monomethyl lysine, Π=dimethyl lysine, Θ= trimethyl lysine, Σ= phospho serine, Ω= phosphor threonine, Ξ= monomethyl arginine, Ψ= asymmetric dimethyl arginine.

H4 1-21	SGRGKGGKGLGKGGAKRHRKV
H4S1P	ΣGRGKGGKGLGKGGAKRHRKV
H4R3me1	SGΞGKGGKGLGKGGAKRHRKV
H4R3me2	SGΨGKGGKGLGKGGAKRHRKV
H4R3me2/K5Ac	SGΨGΔGGKGLGKGGAKRHRKV
H4R3me1/K5Ac/Sp1	ΣGΨGΔGGKGLGKGGAKRHRKV
H4K5Ac	SGRGΔGGKGLGKGGAKRHRKV
H4K5Ac	SGRGΔGGKGLGKGGAKRHRKV
H4K5Ac/K8Ac	SGRGΔGGΔGLGKGGAKRHRKV
H4K8Ac	SGRGKGGΔGLGKGGAKRHRKV
H4K8Ac/K12Ac	SGRGKGGΔGLGΔGGAKRHRKV
H4K12me1	SGRGKGGKGLGΦGGAKRHRKV
H4K12me2	SGRGKGGKGLGΠGGAKRHRKV
H4K12me3	SGRGKGGKGLGΘGGAKRHRKV
H4K8Ac/K12me1	SGRGKGGΔGLGΦGGAKRHRKV
H4K8Ac/K12me3	SGRGKGGΔGLGΘGGAKRHRKV
H4K12me1/K16Ac	SGRGKGGKGLGΦGGAΔRHRKV
H4K12me3/K16Ac	SGRGKGGKGLGΘGGAΔRHRKV
H4K16Ac	SGRGKGGKGLGKGGAΔRHRKV
H4(13-33)	GGAKRHRKVLRDNIQGITKPA
H4K20Ac	GGAKRHRΔVLRDNIQGITKPA
H4K16AcK20Ac	GGAΔRHRΔVLRDNIQGITKPA
H4K16AcK20me1	GGAΔRHRΦVLRDNIQGITKPA
H4K16AcK20me3	GGAΔRHRΘVLRDNIQGITKPA
H4K20me1	GGAKRHRΦVLRDNIQGITKPA
H4K20me2	GGAKRHRΠVLRDNIQGITKPA
H4K20me3	GGAKRHRΘVLRDNIQGITKPA
H4(43-63)	VKRISGLIYEETRGVLKVFLE
H4S47p	VKRIΣGLIYEETRGVLKVFLE
H4K59Ac	VKRISGLIYEETRGVLΔVFLE
H4(51-71)	YEETRGVLKVFLENVIRDAVT
H4K59Ac	YEETRGVLΔVFLENVIRDAVT

H4 (68-88)	DAVTYTEHAKRKTVTAMDVVY
H4K77Ac	DAVTYTEHAΔRKTVTAMDVVY
H4K77AcK79Ac	DAVTYTEHAΔRΔTVTAMDVVY
H4K79Ac	DAVTYTEHAKRΔTVTAMDVVY
H4K77AcK79Ame1	DAVTYTEHAΔRΦTVTAMDVVY
H4K77AcK79Ame3	DAVTYTEHAΔRΘTVTAMDVVY
H4K79me1	DAVTYTEHAKRΦTVTAMDVVY
H4K79me2	DAVTYTEHAKRΠTVTAMDVVY
H4K79me3	DAVTYTEHAKRΘTVTAMDVVY
H4(82-112)	TAMDVVYALKRQGRTLYGFGG-acid
H4R92me1	TAMDVVYALKΞQGRTLYGFGG-acid
H4R92me2	TAMDVVYALKΨQGRTLYGFGG-acid
H2a(1-21)	SGRGKQGGKARAKAKTRSSRA
H2aS1p	ΣGRGKQGGKARAKAKTRSSRA
H2aS1p/K5Ac	ΣGRGΔQGGKARAKAKTRSSRA
H2K5Ac	SGRGΔQGGKARAKAKTRSSRA
H2K5Ac/K9Ac	SGRGΔQGGΔARAKAKTRSSRA
H2K9Ac/K13Ac	SGRGKQGGΔARAΔAKTRSSRA
H2K13Ac	SGRGKQGGKARAΔAKTRSSRA
H2K13Ac/K15Ac	SGRGKQGGKARAΔAΔTRSSRA
H2K9Ac/K13Ac/K15Ac	SGRGΔQGGKARAΔAΔTRSSRA
H2a(21-41)	AGLQFPVGRVHRLLRKGNYAE
H2aK36Ac	AGLQFPVGRVHRLLRΔGNYAE
H2a(85-105)	LAIRNDEELNKLLGKVTIAQG
H2aK95Ac	LAIRNDEELNΔLLGKVTIAQG
H2aK99Ac	LAIRNDEELNKLLGΔVTIAQG
H2aK95Ac/K99Ac	LAIRNDEELNΔLLGΔVTIAQG
H2a(111-129)	IQAVLLPKKTESHHKAKGK-acid
H2aK119Ac	IQAVLLPKΔTESHHKAKGK-acid
H2aK119Ac/T20p	IQAVLLPKΔΩESHHKAKGK-acid
H2aT20p	IQAVLLPKKΩESHHKAKGK-acid
H2b(1-21)	PEPAKSAPAPKKGSKKAVTKA
H2bK5Ac	PEPAΔSAPAPKKGSKKAVTKA
H2bK5me1	PEPAΦSAPAPKKGSKKAVTKA
H2bK5me2	PEPAΠSAPAPKKGSKKAVTKA
H2bK5me3	PEPAΘSAPAPKKGSKKAVTKA
H2bK12Ac	PEPAKSAPAPKΔGSKKAVTKA
H2bS14p	PEPAKSAPAPKKGSKKAVTKA
H2bK12Ac/S14p	PEPAKSAPAPKΔGΣKKAVTKA
H2bS14p/K15Ac	PEPAKSAPAPKKGΣΔKAVTKA
H2bK12Ac/K15Ac	PEPAKSAPAPKΔGSΔKAVTKA
H2b(15-35)	KKAVTKAQKKDGKKRKRSRKE
H2bK20Ac	KKAVTΔAQKKDGKKRKRSRKE
H2bK20Ac/K23me1	KKAVTΔAQΦKDGKKRKRSRKE
H2bK20Ac/K23me3	KKAVTΔAQΘKDGKKRKRSRKE
H2bK23me1	KKAVTKAQΦKDGKKRKRSRKE
H2bK23me2	KKAVTKAQΠKDGKKRKRSRKE

H2bK23me3	KKAVTKAQΘKDGKKRKRSRKE
H2bK23me1/K24Ac	KKAVTKAQΦΔDGKKRKRSRKE
H2bK23me3/K24Ac	KKAVTKAQΘΔDGKKRKRSRKE
H2bK24Ac	KKAVTKAQKΔDGKKRKRSRKE
H2b(27-47)	KKRKRSRKESYSVYVYKVLKQ
H2bS32p	KKRKRΣRKESYSVYVYKVLKQ
H2bS32p/S36p	KKRKRΣRKEΣYSVYVYKVLKQ
H2bS36p	KKRKRSRKEΣYSVYVYKVLKQ
H2bS36p/K43Ac	KKRKRSRKEΣYSVYVYΔVLKQ
H2bK43Ac	KKRKRSRKESYSVYVYΔVLKQ
H2b(37-57)	YSVYVYKVLKQVHPDTGISSK
H2bK43me1	YSVYVYΦVLKQVHPDTGISSK
H2bK43me2	YSVYVYΠVLKQVHPDTGISSK
H2bK43me3	YSVYVYΘVLKQVHPDTGISSK

Appendix B:

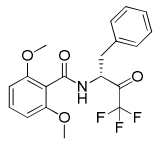
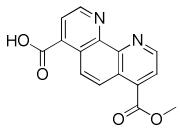
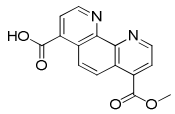
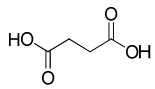
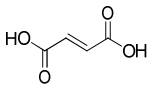
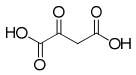
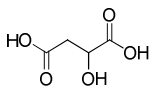
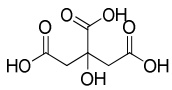
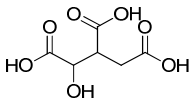
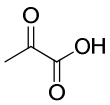
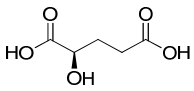
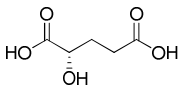
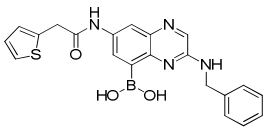
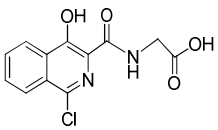
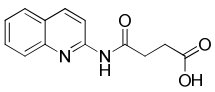
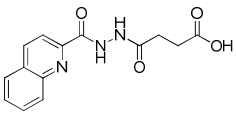
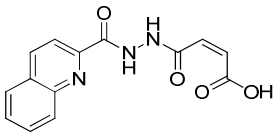
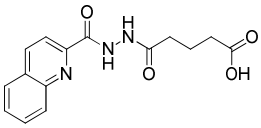
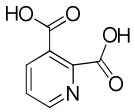
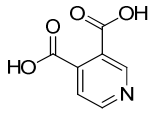
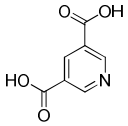
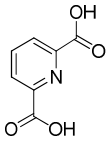
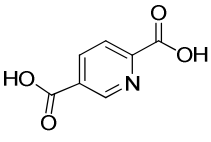
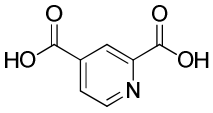
Table B: List of potential proteins interacting with JMJD6 as analysed LC-MS/MS carried out after in-gel peptidase activity by elastase of immunoprecipitated GFP tagged JMJD6 from HeLa cells. The first 40 proteins when sorted by exponentially modified protein abundance index (emPAI) score are shown.

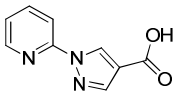
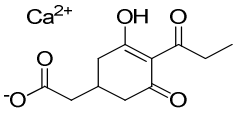
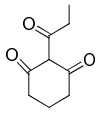
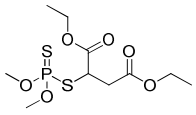
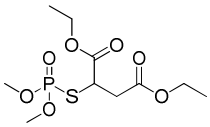
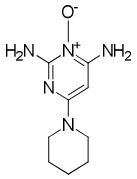
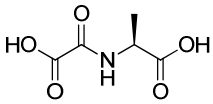
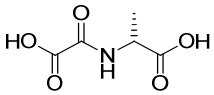
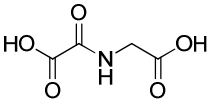
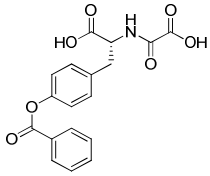
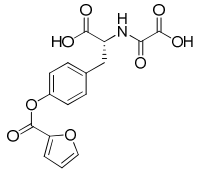
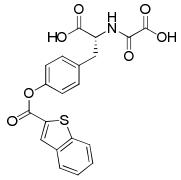
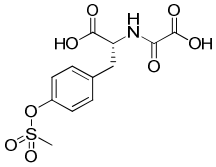
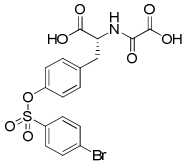
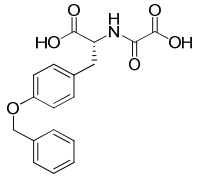
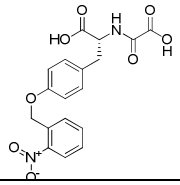
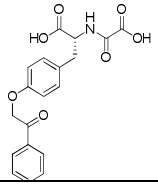
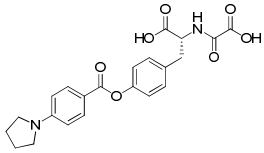
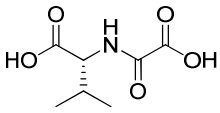
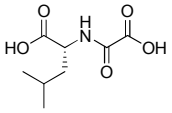
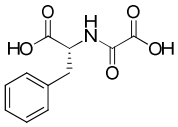
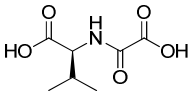
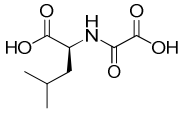
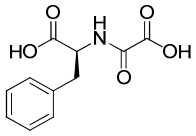
Acc. No.	Score	emPAI	Protein name
Q6NYC1	9112	43.85	Bifunctional arginine demethylase and lysyl-hydroxylase JMJD6
O43809	351	2.28	Cleavage and polyadenylation specificity factor subunit 5
Q01081	103	1.16	Splicing factor U2AF 35 kDa
Q16630	1127	1.01	Cleavage and polyadenylation specificity factor subunit 6
Q05519	478	0.71	Serine/arginine-rich splicing factor 11
Q14498	583	0.62	RNA-binding protein 39
P50151	29	0.48	Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-10
Q9Y383	80	0.4	Putative RNA-binding protein Luc7-like 2;
Q9UHX1	137	0.38	Poly(U)-binding-splicing factor PUF60
P62306	59	0.35	Small nuclear ribonucleoprotein F
Q9BRT7	39	0.31	Putative uncharacterized protein C1orf97;
P14174	52	0.27	Macrophage migration inhibitory factor
Q5SRE7	747	0.21	Phytanoyl-CoA dioxygenase domain-containing protein 1
O95218	47	0.18	Zinc finger Ran-binding domain-containing protein 2
Q15198	29	0.16	Platelet-derived growth factor receptor-like protein
Q06830	44	0.15	Peroxiredoxin-1
Q14197	43	0.14	Peptidyl-tRNA hydrolase ICT1
Q5PT55	30	0.14	Sodium/bile acid cotransporter 5
P26368	50	0.13	Splicing factor U2AF 65 kDa subunit
P53805	70	0.12	Calcipressin-1
Q9UNI1	48	0.12	Chymotrypsin-like elastase family member 1
P50135	42	0.1	Histamine N-methyltransferase
Q15365	59	0.09	Poly(rC)-binding protein 1
P11142	49	0.09	Heat shock cognate 71 kDa protein
P09525	40	0.09	Annexin A4
Q96IZ0	35	0.09	PRKC apoptosis WT1 regulator protein
Q9Y5K5	27	0.09	Ubiquitin carboxyl-terminal hydrolase isozyme L5
Q96J01	41	0.08	THO complex subunit 3

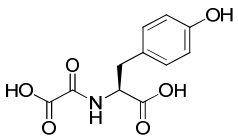
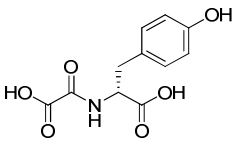
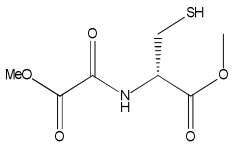
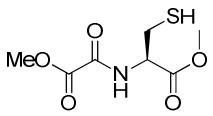
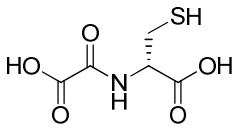
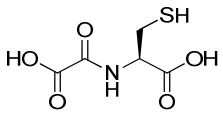
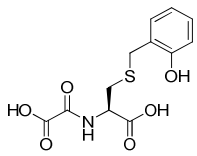
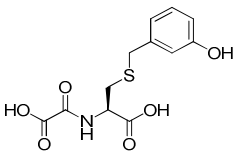
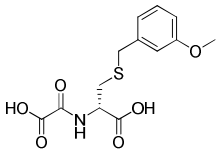
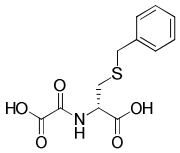
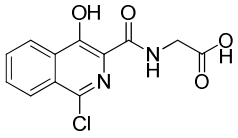
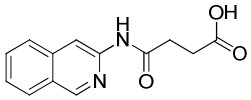
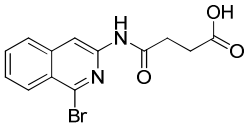
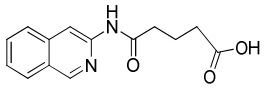
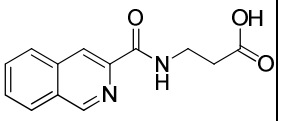
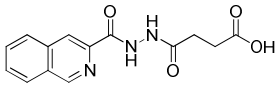
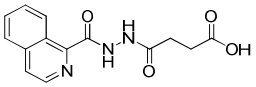
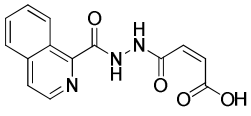
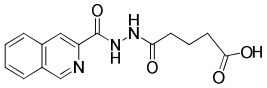
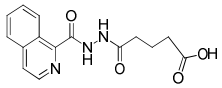
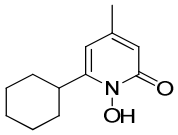
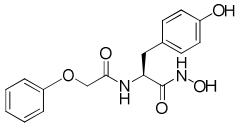
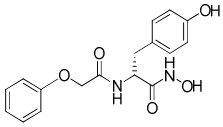
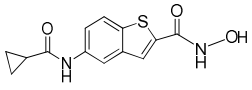
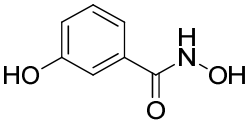
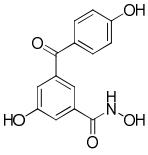
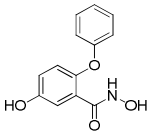
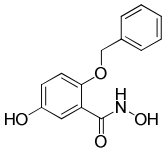
P05455	95	0.07	Lupus La protein
Q8N5F7	50	0.07	NF-kappa-B-activating protein
Q13569	49	0.07	G/T mismatch-specific thymine DNA glycosylaseFull=Thymine-DNA glycosylase;
Q6UX73	38	0.07	UPF0764 protein C16orf89; Flags: Precursor
Q9UKI9	31	0.07	POU domain, class 2, transcription factor 3
Q9Y2L6	72	0.06	FERM domain-containing protein 4B
P16233	67	0.06	Pancreatic triacylglycerol lipase
Q8N1N4	52	0.06	Keratin, type II cytoskeletal 78
Q9H3G5	44	0.06	Probable serine carboxypeptidase CPVL
P35557	39	0.06	Glucokinase
Q7L3T8	31	0.06	Probable prolyl-tRNA synthetase, mitochondrial
Q13228	28	0.06	Selenium-binding protein 1

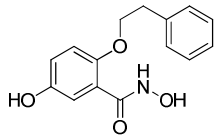
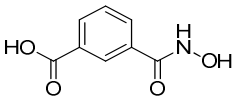
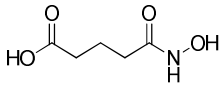
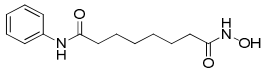
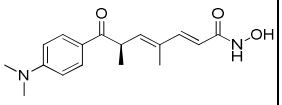
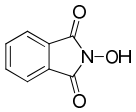
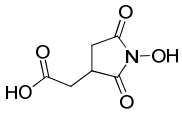
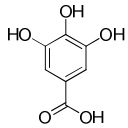
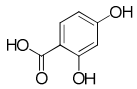
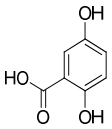
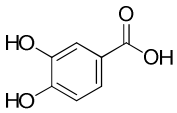
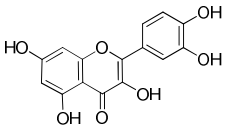
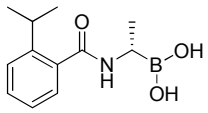
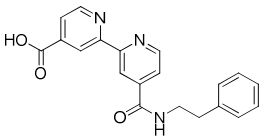
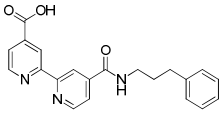
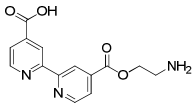
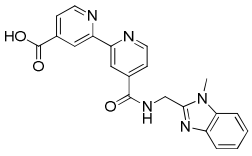
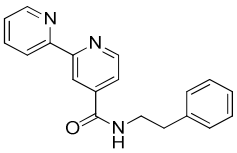
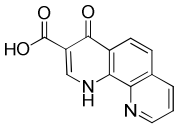
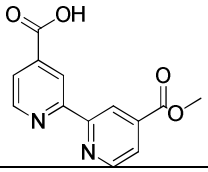
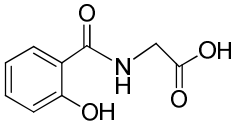
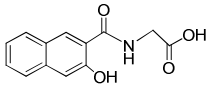
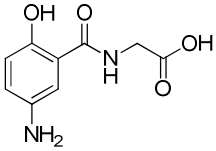
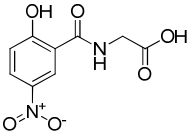
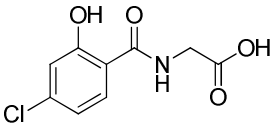
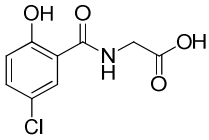
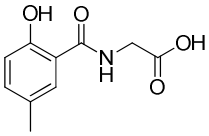
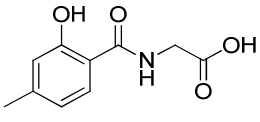
Appendix C:

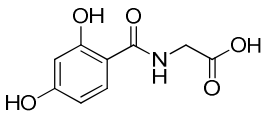
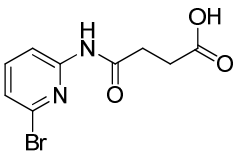
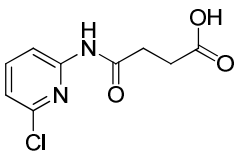
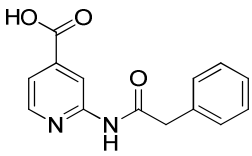
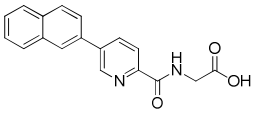
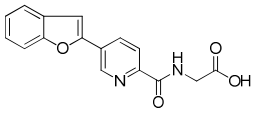
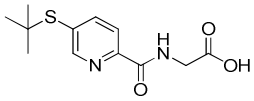
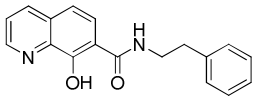
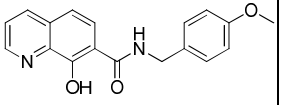
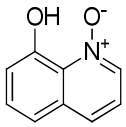
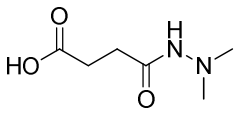
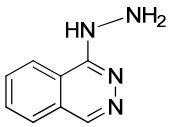
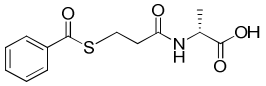
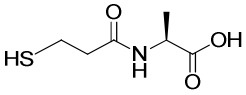
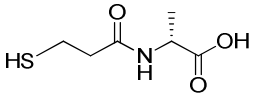
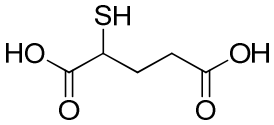
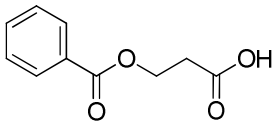
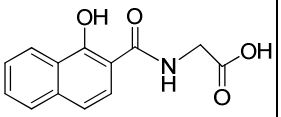
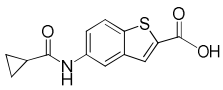
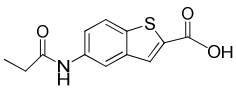
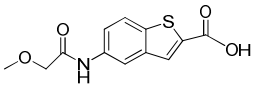
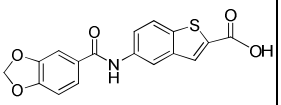
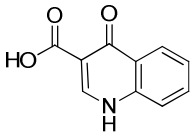
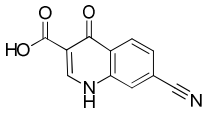
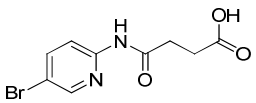
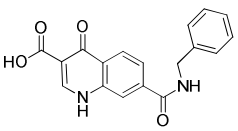
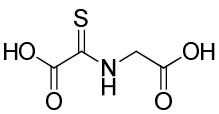
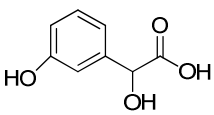
Table C: 2OG analogue compounds from in-house library of inhibitors which is tested in Chapter 7.

EW-T4 	OCO353 	OCO358 	Succinate 
Fumarate 	Oxaloacetate 	Malate 	Citrate 
Isocitrate 	Pyruvate 	D-2-hydroxyglutarate 	L-2-hydroxyglutarate 
EW-TNB 	FG2216 	BB286 	JM104 
JM106 	JM103 	2,3PDCA 	3,4PDCA 
3,5PDCA 	2,6PDCA 	2,5PDCA 	2,4PDCA 

BMRL08012 	Prohexadione 	BB212 	Malathione 
Malaoxon 	Minoxidil 	IS70/NOAL 	IS71/NOAD 
NOG 	CO-A2-5a 	CO-A1-6d 	CO-A7-6h 
CO-B10-7a 	CO-C3-7f 	CO-C10-10a 	CO-D8-10K 
CO-D9-8d 	CO-D4-8m 	NOVD 	NOLD 
NOFD 	NOVL 	NOLL 	NOFL 

NOYL 	NOYD 	EW29 	EW30 
EW31/NOCD 	EW32/NOCL 	EW33 	EW35 
EW36 	EW37 	ACG-263 	BB287 
BB294 	BB315 	BB316 	JM87 
JM110 	JM124 	JM102 	JM109 
Ciclopiroxolamine 	ISS44 	IS45 	ACG-114c 
LBE-332-6-3 	LBE-153 	LBE-314 	LBE-329 

LBE-526 	LBE-096 	LBE060 	SAHA 
TSA 	N-hydroxypthalamide 	IS52 	Gallate 
2,4-DHB 	2,5-DHB 	3,4-DHB 	R2/quercitine 
EW-B17 	OC0248 	OC0249 	OC0206 
OC0336 	OC0340 	BB-103 	BMRL08014 
ACG-194b 	ACG-181a 	ACG-255d 	ACG-253a 
ACG-227a 	ACG-223a 	ACG-226a 	ACG-228a 

ACG-237a 	BB-307 	BB-308 	BMRL08050 
MD228 	MD236 	MD104 	OC0344 
OC0345 	HQNO 	Daminizide 	Hydralazine 
IS21 	IS37 	IS38 	IS54 
IS84 	ACG-148a 	ACG-109a 	ACG-137a 
ACG-144a 	ACG-138a 	BB-110H 	BB-169 
BB-318 	BB-204 	IS3 	3-Hydroxymandelate 

Appendix D:

Oligonucleotide primers used in Chapter 3 for preparing point mutations.

K204A
Forward primer 5’ GTT CAG GGC CAC GCG CGC TGG TGC CTG 3’ Reverse primer 5’ CAG GCA CCA GCG CGC GTG GCC CTG AAC 3’
N287A
Forward primer 5’GCC ATC ACC CAA GCT TTT GCC AGC AGC 3’ Reverse primer 5’ GCT GCT GGC AAA AGC TTG GGT GAT GGC 3’
T285A
Forward primer 5’ ACT ATC GCC ATC GCC CAA AAT TTT GCC AG 3’ Reverse primer 5’ GGC AAA ATT TTG GGC GAT GGC GAT CGT CG 3’
Y131F
Forward primer 5’ GATAGTCCCCTTTTCATCTTTGACAGCAGC 3’ Reverse primer 5’ GCTGTCAAAGATGAAAAGGGGACTATCATC 3’
E231A
Forward primer 5’ CCAGCAAGACGCAGCTATTACCTGTTTTAATG 3’ Reverse primer 5’ AAACAGGTAATAGCTGCGTCTTGCAGGTTCCC 3’