Characterisation of a novel hydroxylase complex and its role in replication fidelity

By Tristan Kennedy

A thesis submitted to the University of Birmingham for the
degree of DOCTOR OF PHILOSOPHY

Institute of Cancer and Genomic Sciences

College of Medical and Dental Sciences

University of Birmingham

June 2023

UNIVERSITY OF BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

Abstract

Protein hydroxylation is an emerging and poorly characterised post-translational modification that is generally catalysed by enzymes called '2-oxoglutarate-dependent oxygenases'. JMJD5 is an 'orphan' hydroxylase, which has likely contributed to a relatively poor understanding of its cellular role. In this thesis we have attempted to further our understanding of JMJD5 by focusing on an activity-independent interaction between JMJD5 and the Regulator of Chromosome Condensation Domain containing protein 1 (RCCD1). We begin by identifying the specific amino acids in RCCD1 and the JMJD5 N-terminus required for this interaction. We subsequently show that RCCD1 knockdown causes replication stress and that this phenotype is epistatic with respect to JMJD5. Importantly, we demonstrate the importance of the JMJD5:RCCD1 interaction for DNA replication fidelity through reconstitution of the corresponding binding mutants into knockdown cells. We additionally demonstrate that JMJD5:RCCD1 form a stoichiometric 1:1 heterodimeric complex and develop a co-overexpression system for structural characterisation and interactome analysis. Through subsequent proteomic screens we identify a possible activity dependent interaction of the JMJD5:RCCD1 with the CX3 complex of RAD51 paralogs that prescribe the JMJD5:RCCD1 complex with a role in replication fork restart. Overall, our work identifies a novel heterodimeric complex with an important role in faithful DNA replication and genome stability. With genome instability being regarded as a key hallmark of cancer and both JMJD5/RCCD1 being implicated in cancer, our work is helping to understand a potential novel pathway that may contribute to tumorigenesis.

Acknowledgements

Firstly, I would like to thank Mat for taking me on as an MRes student in 2018 and for having enough faith to support me onto my PhD programme and for guiding me through the last 4 years. For all the beverages (both coffee and stronger), and scientific advice, I couldn't thank you enough.

To Sally, thank you so much for having the patience to train me throughout my masters and the early years of my PhD! Also, I need to thank Uncaar for the early morning football chats about all things Arsenal and for helping with cloning whenever needed! To the rest of TOG, Eline, Reg, Christian, Sonia and Arashpreet thank you for being great supportive lab mates. I also need to thank Professor Smerdon, for your help with all things protein structures and modelling!

I would also like to thank my Mum, Brother, and Sister, although you have no idea what I do, your support has been greatly appreciated.

A huge thanks must go to my friends especially Lisa, Poppy, Abbey, Gillian and Tom for all the support and the weekday socialising that helped to destress. To David, thank you being my lab partner on our first week of undergrad and my best friend for the last 9 years, who would've thought that we would both end up doing PhDs in Birmingham?

Lastly and most importantly, to my long-suffering partner Abbie, thank you for bearing with me through all the highs and lows across the last 4 years. Your support has meant the world.

Declaration

I declare that this thesis is entirely my own work, with the exception of the following:

Structural prediction used in Figure 2.4A was performed by Dr Chan Li.

Experimental data presented in Figure 4.6 was generated by Dr Chan Li, Dr Eilis Braggington and Professor Steve Smerdon.

Data presented in Figure 4.13B was generated by Dr Sally Fletcher.

Original pEF6-HA-JMJD5, pTIPZ 3XFLAG-JMJD5 (WT and H321A) and pc-DNA3-RCCD1 were generated by others including Professor Mathew Coleman and Mr Uncaar Boora. Generation of pIHZ RCCD1 SDM constructs was carried out in close collaboration with Mr Uncaar Boora.

Contents

List of Figures

List of Tables

List of Abbreviations

2HG	2-hydroxyglutarate
2OG	2-oxoglutarate
OGFOD1	2-oxoglutarate and Fe(II)-dependent oxygenase domain-containing protein 1
DAPI	4',6-diamidino-2-phenylindole
ATM	Ataxia Telangiectasia mutated kinase
ATR	ATM-Rad3-related kinase
ATRIP	ATR-interacting protein
BIR	Break-induced replication
CENPA	Centromere protein A
CENPF	Centromere protein F
CHK1	Checkpoint kinase 1
CFS	Common fragile sites
DRG	Developmentally regulated GTP-binding protein
DMOG	Dimethyloxalglycine
RS	DNA replication stress
DSB	Double strand break
DSBH	Double-stranded β-helix
EV	Empty vector
eRF1	Eukaryotic release factor 1
FIH	Factor inhibiting HIF
FTO	Fat mass and obesity-associated protein
FL	Full length
GST	Glutathione S-transferase
GFP	Green fluorescent protein
HSPBAP1	Heat shock protein beta-1 associated protein 1
HA	Hemagglutinin
HEK293T	Human embryonic kidney 293 T-antigen
HU	Hydroxyurea
HIF	Hypoxia inducible factor
IP	Immunoprecipitation
IDH	Isocitrate dehydrogenase
IPTG	Isopropyl ß-D-1-thiogalactopyranoside
JMJD	Jumonji domain containing protein
JMJD5	Jumonji domain containing protein 5
JmjC	Jumonji-C domain
KDM	Lysine demethylase
MS	Mass Spectrometry
MCM	Minichromosome maintenance
MSA	Multiple sequence alignment
MINA53	Myc induced nuclear antigen 53kDa
NOG	N-oxaylglycine
NLS	Nuclear localisation signal
NO66	Nucleolar protein 66 kDa
PIKK	Phosphoinositide 3-kinase related protein kinase

PARP	Poly ADP-ribose polymerase
PCR	Polymerase chain reaction
PTM	Post-translational modification
PHD	Prolyl-hydroxylase domain
RAD51C	RAD51 paralog C
RLD	RCC1-like domain
RCC1	Regulator of Chromosome Condensation 1
RCCD1	Regulator of Chromosome Condensation Domain containing protein 1
RPA	Replication protein A
RPS6	Ribosome protein S6
SNP	single nucleotide polymorphism
ssDNA	Single stranded DNA
siRNA	Small interfering RNA
TET	Ten-eleven translocation
TSG	Tumour suppressor gene
VHL	von Hippel-Lindau protein
WT	Wild-type
XRCC3	X-ray repair cross complementing 3

Chapter 1: Introduction

1.1 Hydroxylation is a Post-translational Modification catalysed by 2-OG oxygenases

1.1.1 Post-translational modifications

Proteins can be chemically modified through the covalent attachment of a chemical group, atom, or whole protein, termed post-translational modifications (PTM). PTMs play key roles in cellular physiology, helping to regulate protein stability, protein-protein interactions, and signalling pathways (Knorre et al., 2009). Cells require a complex network of PTMs, with crosstalk between PTMs coordinating the regulation of various cellular processes (Duan and Walther 2015). Examples of PTMs include phosphorylation, methylation, acetylation, and ubiquitination (Ribet and Cossart, 2010). Most PTMs are catalysed by a family of related enzymes.

1.1.2 Hydroxylation and the 2-OG Family

Hydroxylation is defined as the enzymatic addition of a single oxygen atom to create a hydroxyl or alcohol group (-OH) on small molecules, DNA, RNA, proteins, or lipids (Islam et al., 2018). Hydroxylation is a widespread modification that occurs across the evolutionary tree from prokaryotes to plants to humans and is generally catalysed by 2-Oxoglutarate (2OG) dependent oxygenases (Islam et al., 2018). In humans there are approximately 60 2OG-dependent oxygenases (Fig 1.1A). Although these enzymes are named for their reliance on the Krebs' cycle intermediate 2OG, they also require Fe(II) and O_2, with some enzymes additionally needing ascorbate as a reducing agent (Loenarz and Schofield, 2011). Eukaryotic 2OG oxygenases use these cofactors to catalyse stable hydroxylation or demethylation via hydroxylation (Fig 1.1B) (Islam et al., 2018).

2-oxoglutarate Oxygenases

A

Lysine histone demethylases

KDM4, KDM5, KDM6, KDM3, KDM7, KDM2

JMJD2E, JMJD2C, JMJD2F, JMJD2B, JMJD2D, JMJD2A, JARID1D, JARID1C, JARID1A, JARID1B, UTX, UTY, JMJD3, JARID2, JMJD1A, JMJD1B, JMJD1C, HR, PHF8, KIAA1718, PHF2, FBXL10, FBXL11

Small-molecule Hydroxylases

PHYD1, PHYH, BBOX1, TMLHE, ASPH

Prolyl Hydroxylases

PHD3, PHD1, PHD2, LEPREL2, FLJ2222, LEPRE1, LEPREL1, OGFOD1

Collagen prolyl Hydroxylases

PH4, P4HA3, P4HA1, P4HA2, PLOD2, PLOD, PLOD3

JmjC-only Hydroxylases

NO66, MINA53, JMJD4, JMJD6, JMJD8, JMJD7, FIH, TYW5, JMJD5, HSPBAP1

Nucleotide Hydroxylases

FTO, ALKBH1, ALKBH3, ALKBH2, TET1, TET3, TET2, ALKBH5, ALKBH7, ALKBH4, ALKBH6, ALKBH8

B

Hydroxylation

R-C-H → R-C-OH

2OG Oxygenase

R-NH-CH$_3$ → R-NH-CH$_2$OH → R-NH$_2$ + CHOH

Demethylation *via* hydroxylation

Figure 1.1. Hydroxylation is catalysed by the 2OG oxygenase family.
(A) Phylogenetic tree of 2OG dependent oxygenases. Each sub family is coloured according to substrate specificity. Lysine histone demethylases (KDMs) in yellow, JmjC-only hydroxylases in red, Nucleotide hydroxylases in blue, small-molecule hydroxylases in green, prolyl hydroxylases in pink and collagen prolyl hydroxylases in purple. Proteins in listed in black do not belong to other sub-groups. Figure adapted from Johansson et al 2014. **(B)** 2OG oxygenases catalyse either hydroxylation or demethylation via hydroxylation.

1.1.3   2OG oxygenases share conserved structure and catalytic mechanism

The catalytic domain structure of 2OG oxygenases consists of eight antiparallel β-strands folded into a double-stranded β-helix (DSBH) (Fig 1.2A). This compact structure allows the co-factor and substrate-binding residues to be orientated in the correct configuration for catalysis (McDonough et al., 2010). The catalytic domain contains two histidines and one carboxylate residue which form a 'facial triad' in the form of a HxD/E…H motif that is responsible for Fe(II) coordination (Fig 1.2B) (Loenarz and Schofield, 2008). Residues involved in the coordination of 2OG vary depending on the subfamily but are generally comprised of one basic (Arginine or Lysine) and one polar neutral (Serine or Tyrosine) residue (Islam et al., 2018). All subfamilies catalyse hydroxylation via the same general catalytic mechanism, as follows.
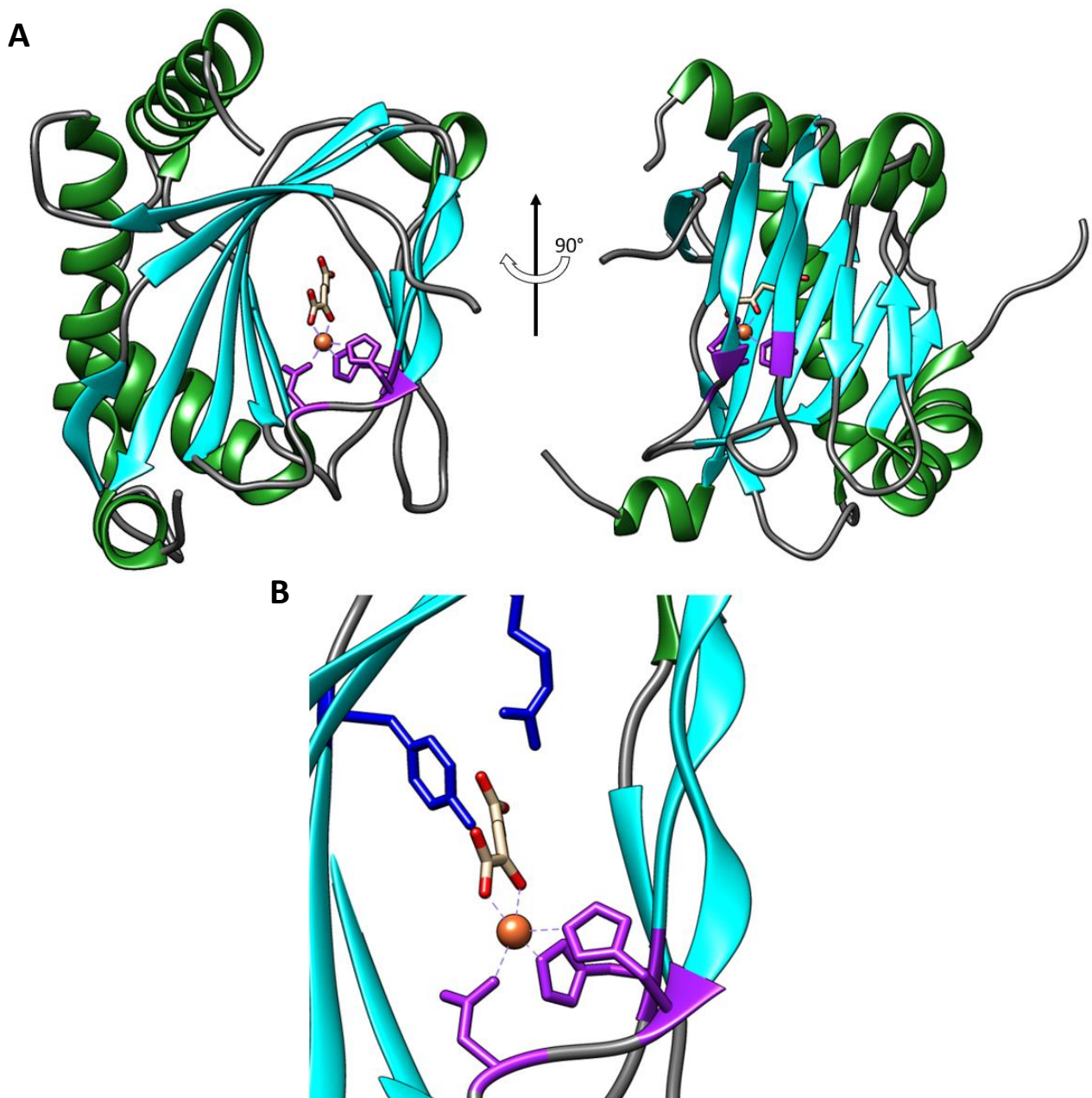
Figure 1.2. The catalytic domain of 2OG oxygenases is highly conserved across the family.
(A) 2OG oxygenases share a conserved catalytic structure comprised of eight antiparallel β-strands. **(B)** The catalytic site co-ordinates the Fe(II) co-factor through a catalytic 'facial' triad HxD/E...H motif (purple). The 2OG co-factor is co-ordinated by one basic (Arginine or Lysine) and one polar neutral (Serine or Tyrosine) residue (Blue). All images were made using pdb 3OUJ and the protein modelling software Chimera. Figure adapted from Ploumakis and Coleman 2014.

The binding of 2-oxoglutarate initiates the catalytic cycle by replacing two H_2O molecules occupying the remaining Fe(II) coordination sites (Fig 1.3). The substrate and O_2 then bind to generate a reactive oxygen species by displacing the final H_2O molecule. This reactive oxygen species attacks the C2 position of 2OG, initiating an oxidative decarboxylation reaction (Martinez and Hausinger, 2015),

converting 2OG to succinate and CO_2, which are released as by-products. This reaction also generates

an Fe(IV)=O (ferryl oxo) intermediate, which is highly reactive and oxidises the substrate (Loenarz and

Schofield, 2011). In the final step, the hydroxylated substrate and succinate are released. Whilst the

structure and catalytic mechanism are conserved across the 2OG family, the functional consequences

of the arising modifications are diverse.
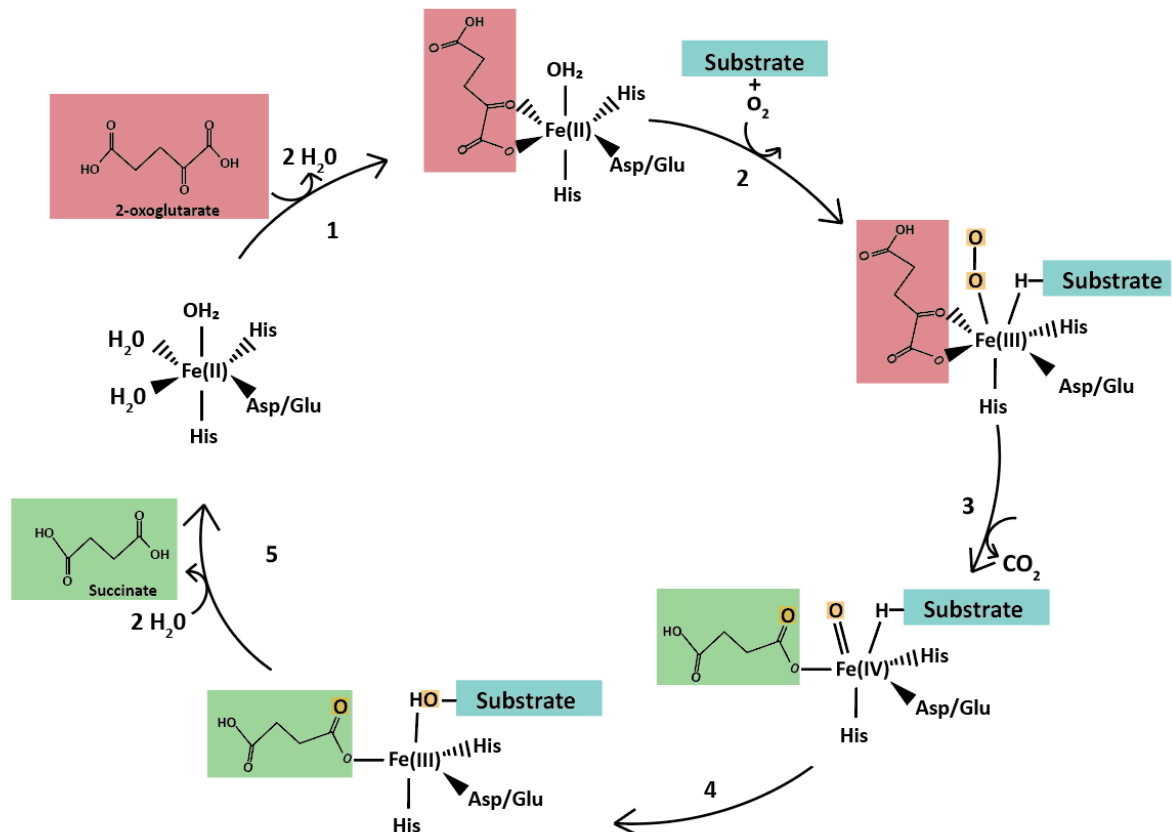
Figure 1.3. The catalytic cycle of 2OG oxygenases.
This figure depicts the facial catalytic triad HxD/E…H motif, the water molecule and the Fe(II) and 2OG co-factors. 2OG initiates the catalytic cycle by binding (1) and displacing two molecules of H_2O. Binding of the substrate and oxygen displaces the final H_2O, generating a reactive oxygen species (2). This triggers oxidative decarboxylation of 2OG generating succinate, releasing CO_2 and catalysing oxidation of the substrate (3,4). Succinate and hydroxylated substrate are released (5). The two oxygen atoms binding as O_2 are highlighted in orange.

1.1.4 2OG oxygenase exhibit diverse activities and functions

Hydroxylation was once considered rare; however, more than 60 2OG oxygenases have now been identified in humans. Outside the catalytic domain, 2OG oxygenases display structural diversity through additional functional domains, which together support a variety of interactions and the targeting of a wide range of substrates, including proteins, nucleotides, and lipids (Islam et al., 2018). Protein hydroxylation is also diverse, with 2OG oxygenases identified that hydroxylate prolyl, lysyl, histidyl, arginyl, asparaginyl and aspartyl residues (Ploumakis and Coleman, 2015). Hydroxylation was originally identified as a PTM in fibrillar collagen (Kivirikko and Prockop, 1967; Gjaltema and Bank, 2017); however, the functional diversity of these enzymes is now more fully understood. In addition to collagen hydroxylation, 2OG oxygenases regulate fatty acid metabolism (Loenarz and Schofield, 2011), hydroxylation of transmembrane proteins (Stenflo et al., 1989), and every step of the gene expression control pathway.

1.1.5 Hydroxylases are important regulators of gene expression

Hydroxylation plays a central role in the regulation of gene expression, involving 2OG oxygenases from several different subfamilies. This regulation includes the epigenetic control of DNA/RNA and modification of proteins involved in transcription and translation.

Critical levels of epigenetic regulation in cells involves covalent modification of DNA and histones within chromatin (Gibney and Nolan, 2010). Methylation is one of the most common epigenetic modifications and is dynamic, being reversed by demethylase enzymes of the 2OG oxygenase family using a hydroxylation reaction (Ploumakis and Coleman, 2015).

1.1.5.1 Nucleotide demethylation

Transcription can be repressed through DNA methylation catalysed by DNA methyltransferases (DNMTs), which modify the fifth carbon of cytosines to create 5-methylcytosine (5mC) (Greenberg and Bourc, 2019). This modification is reversed by the Ten-Eleven Translocation (TET) subfamily of 2OG-oxygenases. TET1 catalyses sequential hydroxylation reactions, converting 5mC to 5-

hydroxymethylcytosine (5hmC), 5hmC to 5-formylcytosine (5fC), and 5fC to 5-carboxylcytosine (5caC) (Fig 1.4) (Tahiliani et al., 2009). TET2 and 3 share similar kinetic activities with TET1, catalysing the same methylcytosine hydroxylation modifications, and appear to have similar physiological roles (Ko et al., 2010; He et al., 2011).
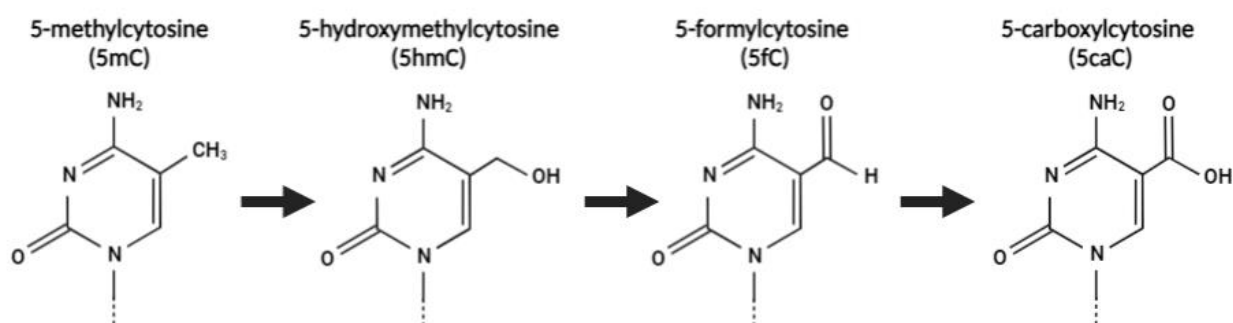
Figure 1.4. TET enzyme catalysed demethylation of 5-methylcytosine on DNA.
Ten-Eleven Translocation (TET) 1-3 enzymes catalyse the sequential conversion of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). Figure generated using BioRender.

DNA is not the only modified nucleic acid that regulates gene expression. mRNA is also methylated to create N^6-methyladenosine which controls mRNA stability, translation, and splicing (Meyer and Jaffrey, 2014). Reversal of this methylation mark is catalysed by two 2OG oxygenases, 'fat mass and obesity-associated' protein (FTO), and alkylated DNA repair protein alkB homologue 5 (ALKBH5) (Jia et al., 2011; Zheng et al., 2013).

1.1.5.2   Histone demethylation and the KDM subfamily

Another target of extensive epigenetic modification are histone tails, which are modified by acetylation, sumolyation, ubiquitination, phosphorylation, and methylation (Bannister and Kouzarides, 2011). These PTMs regulate gene transcription and chromatin structure and are key steps in epigenetic gene regulation (Bannister and Kouzarides, 2011). Methylation plays a key role in epigenetic inheritance, chromatin structure, cell differentiation, and transcriptional regulation (Martin and Zhang, 2005). Histone methylation mostly occurs on either lysine (K) or arginine (R) residues, both

of which can be mono- or dimethylated (arginine methylation may be symmetric or asymmetric), and lysine can be trimethylated.

The 2OG oxygenase family responsible for catalysing lysine histone demethylation is characterised by a specific variant of the archetypal DSBH fold, called the Jumonji-C (JmjC) domain. This domain was initially identified following characterisation of the jumonji gene, so called because jumonji mutant mice developed a cruciform-shaped neural groove (Jumonji means cruciform in Japanese) (Takeuchi et al., 1995). Two distinct subgroups of the 2OG oxygenase family contain a JmjC domain, including the lysine histone demethylase (KDM) subfamily and 'JmjC-only' protein hydroxylases, which will be discussed in more detail below (Klose et al., 2006). KDMs require additional functional domains which bind to DNA (e.g. zinc fingers and ARID domains) and are the largest subfamily, containing approximately 20 enzymes (Johansson et al., 2014). These KDMs catalyse the demethylation of mono-, di-, and tri-methylated lysine at positions 4, 9, 27, and 36 on histone 3 (H3K4, H3K9, H3K27, and H3K36) and lysine 20 on histone 4 (H4K20). Substrate specificity, sequence homology, and domain organisation further subdivides the histone lysine demethylase family into KDM2-7 (Johansson et al., 2014) (Fig 1.1A).

1.1.5.3   Oxygen sensing and regulation of transcription factors

Beyond epigenetics, a key determinant of gene expression control is transcription factor activity; one such factor which is heavily regulated by 2OG oxygenases is Hypoxia-Inducible Factor (HIF). The HIF heterodimer comprises two subunits (HIFα and HIFβ). The α subunit has three known isoforms (HIF1/2/3α), both HIF1α and HIF2α show strong homology, however, HIF3α is more distinctly related and contains no transactivation domain (Gu et al., 1998; Zhao et al., 2015a). Although HIF1α and HIF2α are closely related, HIF2α is only expressed in certain tissues at different developmental stages (Patel et al., 2010), therefore the following section will discuss HIF1α. The activity of the HIF heterodimer is regulated by three prolyl hydroxylase-domain containing proteins (PHD1-3 or EGLN1-3) (summarised in Fig 1.5). Under normoxia, Prolyl Hydroxylase Domain (PHD) enzymes catalyse the hydroxylation of

the conserved prolyl residues P402 and P564 on HIFα (Masson et al., 2001; Bruick and McKnight, 2001; Epstein et al., 2001; Ivan et al., 2001). This generates a recognition motif for the von Hippel Lindau (VHL) ubiquitin ligase, marking HIFα for proteasomal degradation (Jaakkola et al., 2001, Shen and Kaelin, 2013). During reduced oxygen availability (hypoxia), the activity of PHD enzymes decreases due to their requirement for O_2 in the catalytic cycle, resulting in stabilisation of HIFα and increased gene expression of its targets (Ratcliffe, 2013). Conversely, the HIFβ subunit is constitutively expressed, which allows a HIFα:HIFβ heterodimer to form in hypoxia (Ratcliffe, 2013). A further level of HIF regulation comes from a 2OG oxygenase called Factor Inhibiting HIF (FIH), a JmJC-only family 2OG oxygenase (described in more detail in Section 1.1.6), which hydroxylates asparagine 803 in the transactivation domain of HIFα (Lando et al., 2002; Hewitson et al., 2002). This hydroxylation prevents HIF from binding to the transcriptional co-activator p300/CBP (Fig 1.5) (Lisy and Peet, 2008). As with the PHD enzymes above, FIH activity is decreased under low-oxygen conditions, allowing p300/CBP binding and HIF transcription activation. The active HIF heterodimer binds to hypoxia response elements (HREs) to regulate an extensive network of transcription targets. Indeed, the network of transcriptional regulation of HIF is thought to comprise approximately 500 genes which regulate a vast number of cellular processes, including cell survival, metabolism, and differentiation (Schödel et al., 2013; Ratcliffe, 2013; Shen and Kaelin, 2013).
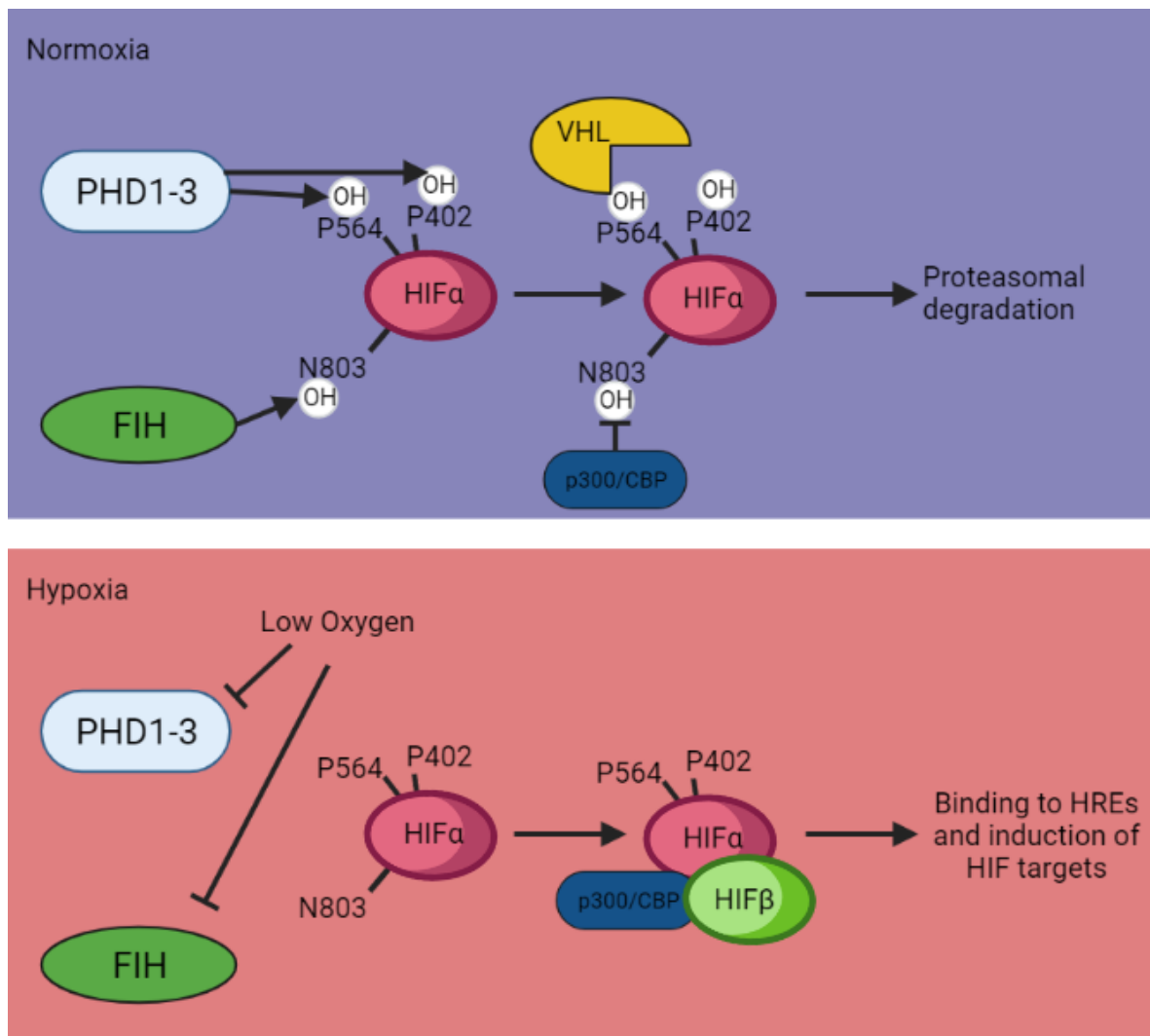
Figure 1.5. Regulation of the HIF transcription factor by 2OG oxygenases.
In normoxia, three prolyl hydroxylases (PHD1-3) and Factor Inhibiting HIF (FIH) hydroxylate P402, P564, and N803 of HIFα, respectively. Prolyl hydroxylation of HIFα creates a recognition motif for the von-Hippel Lindau (VHL) tumour suppressor which triggers proteasomal degradation of HIF. FIH hydroxylates N803 in the HIFα transactivation domain, which blocks its association with the p300/CBP transcriptional co-activator. Under hypoxia, low oxygen levels reduce PHD1-3 and FIH activity allowing HIF to form its active heterodimer and associate with p300/CBP to trigger transcription of HIF targets. Figure generated using BioRender.

1.1.5.4 Splicing control

Following transcription, newly generated pre-messenger RNA (mRNA) must be converted to mature mRNA through 5' capping, 3' polyadenylation, cleavage, and intron removal via splicing (Moore and Proudfoot, 2009). Interestingly, a 2OG oxygenase has been reported to play a key role in splicing. JMJD6, which was originally identified as a phosphatidylserine receptor (Böse et al., 2004), has subsequently been identified as a lysyl hydroxylase of U2AF/U2AF65 (targeting K15, K38 and K276) and LUC172. Both proteins are components of the spliceosome and regulate a subset of pre-mRNAs (Webby et al., 2009a). Interestingly, the binding of U2AF65 to its different pre-mRNA targets can be altered by mutation of different hydroxylation sites (Yi et al., 2017).

1.1.5.5 Regulation of translation

Regulation of translation is emerging as a fundamentally important level of gene expression control. Indeed, aberrant protein synthesis is regarded as being a hallmark of several diseases, including cancer (Silvera et al., 2010). Given the importance of various 2OG oxygenases across gene expression, it is perhaps unsurprising that multiple 2OG oxygenases have recently being implicated in translation (summarised in Fig 1.6).

One area of translation regulated by 2OG oxygenases involves direct modification of transfer RNAs (tRNA). To this end two oxygenases, TYW5 and ALKBH8, have been shown to hydroxylate tRNA. tRNA-yW synthesising enzyme 5 (TYW5) catalyses the hydroxylation of wybutosine (hypermodified guanosine) to hydroxybutosine (Noma et al., 2010). Another hypermodified nucleotide in arginine and glycine tRNAs is 5-methoxycarbonyl-hydroxymethyluridine (mchm^{5}U). mchm^{5}U is located within the anticodon loop and is synthesised by ALKBH8 (Fu et al., 2010). Such tRNA modifications are essential for translational fidelity, and their abnormal regulation is now thought to be key in human diseases, such as cancer, and neurological and mitochondrial conditions (Suzuki, 2021).

Translation also requires large protein factories called ribosomes that contain ribosomal RNA and ribosomal proteins (RPs). RPs are present in both the large and small ribosomal subunit, which are

also targeted by 2OG oxygenases. A hydroxylase called '2-oxoglutarate and Fe(II)-dependent oxygenase domain-containing protein 1' (OGFOD1) is distantly related to the HIF PHD enzymes. OGFOD1 hydroxylates the small subunit protein RPS23 on Proline 62 (Singleton et al., 2014).Furthermore,'Myc-induced nuclear antigen 53 kDa' (MINA53) and 'nucleolar protein 66 kDa' (NO66) catalyse histidyl hydroxylation of large subunit proteins. MINA53 hydroxylates RPL27A on histidine 39 and NO66 modifies histidine 216 on RPL8 (Ge et al., 2012). The exact role of ribosomal hydroxylation with respect to all three ribosomal hydroxylases is currently unknown.

Translation termination is also regulated by protein hydroxylation. Successful termination requires recognition of the stop codon by release factors. Eukaryotic release factor 1 (eRF1) is a class 1 release factor that recognises all three UGA, UAG, and UAA stop codons (Alkalaeva et al., 2006). JMJD4 modifies lysine 63 in a highly conserved NIKS motif located in the stop codon recognition domain of eRF1 (Feng et al., 2014). K63 hydroxylation was shown to be highly abundant and ubiquitous, and its inhibition resulted in stop codon readthrough and aberrant translation termination *in vivo* (Feng et al., 2014).
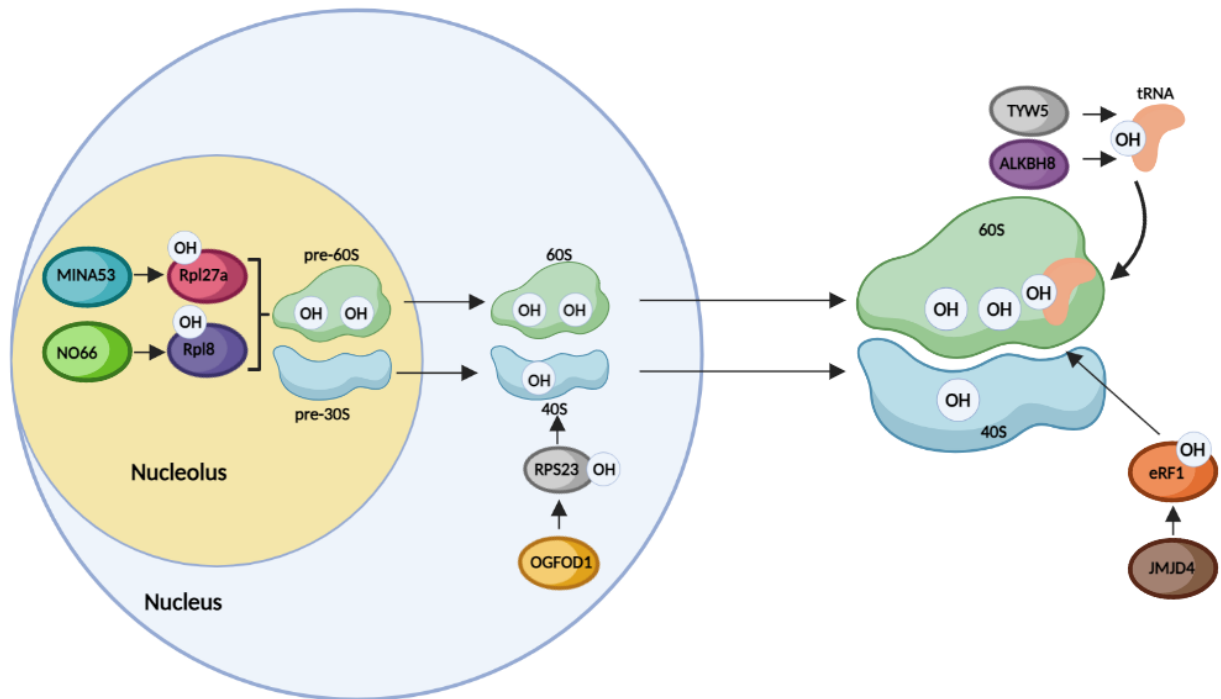
Figure 1.6. Overview of 2OG oxygenases regulating translation.
TYW5 and ALKBH8 catalyse hydroxylation of tRNA catalysing formation of hydroxybutosine and 5-methoxycarbonyl-hydroxymethyluridine (mchm^{5}U) respectively. The ribosomal oxygenases OGFOD1, MINA53 and NO66 hydroxylate RPS23, Rpl27a and Rpl8 respectively which form part of the final active ribosome. JMJD4 hydroxylates the translation termination factor eRF1. Figure generated using BioRender.

Like FIH, TYW5, MINA53, NO66 and JMJD4 all belong to the JmjC-only hydroxylase subfamily. The next section will review the literature surrounding this sub-family in more detail.

1.1.6 JmjC-only protein hydroxylases

Similar to the KDM subfamily, JmjC-only hydroxylases also contain the JmjC catalytic domain but lack the chromatin biology-related domains found in KDMs. This subfamily is comprised of 10 proteins – JMJD4-8, FIH, MINA53, NO66, TYW5, and HSPBAP1–and all except TYW5 (mentioned above) are thought to modify proteins.

The two ribosomal oxygenases in this group, MINA53 and NO66, have been discussed in relation to their roles in translation, and both are thought to be localised to the nucleolus (Eilbracht et al., 2004,

2005). Interestingly, Hendrix et al. recently suggested an alternative role for MINA53 beyond ribosomal hydroxylation, specifically in epithelial tight junction biology through interaction with the membrane-associated guanylate kinase MPP6 (Hendrix et al., 2023). This work suggested dual functions for MINA53 in actively proliferating versus confluent cells. It remains to be seen whether the highly related oxygenase NO66 may also possess similar dual functions.

Another previously mentioned JmjC-only hydroxylase with additional functions beyond that already mentioned is FIH. In addition to the regulation of HIF, FIH has been shown to hydroxylate asparagine residues within distinct protein domains called ankyrin repeat domains (ARD) (Cockman et al., 2006). Found in a diverse range of proteins, these ARDs contain 30-34 residue ankyrin repeats (AR) formed of two α-helices connected by a β-hairpin, which together are thought to facilitate protein-protein interactions (Li et al., 2006). ARD hydroxylation catalysed by FIH is found in a range of proteins, including NF-κB (Cockman et al., 2006), myosin phosphatase target subunit 1 (MYPT1) (Webb et al., 2009), Notch receptors (Coleman et al., 2007), AnkyrinR (Yang et al., 2011b), HECT domain and ankyrin repeat-containing E3 ubiquitin ligase (HACE1) (Kim et al., 2019), and suppressor of cytokine signalling box protein 4 (Ferguson et al., 2007). Indeed, this plethora of identified substrates and large-scale proteomic screens suggest that asparagine hydroxylation of ARDs may be ubiquitous (Cockman et al., 2009). Interestingly, the biochemical specificity of FIH may extend beyond asparagine, with hydroxylase activity against aspartate (Yang et al., 2011b) and histidine (Yang et al., 2011a) residues being identified within the ARDs. Despite the extensive catalogue of FIH targets, the functional consequences of ARD hydroxylation remain unknown.

The JmjC-only subfamily contains a large group (JMJD4-8), which includes several enzymes that have been intensively studied and for which substrates have been identified. In addition, new evidence suggests that JMJD4 and JMJD6 may have other physiological targets. A recent study suggests JMJD4 hydroxylates pyruvate kinase 2 on K66, which is reported to promote its degradation through heat shock protein 70 (HSP70) in cardiomyocytes (Tang et al., 2023). JMJD6 has also been recently shown

to catalyse lysine hydroxylation events in nearly 50 proteins in lysine-rich regions (Cockman et al., 2022).

JMJD7 has been shown to hydroxylate lysine 22 and 21 in two closely related proteins, Developmentally Regulated GTPases 1 and 2 (DRG1/2), respectively (Markolovic et al., 2018). Although the functional roles of these hydroxylation events are not fully understood, DRG2 has been identified as an RNA interactor (Ishikawa et al., 2003). Markolovic et al also showed that knockdown of JMJD7 or a hydroxylation site mutant reduced binding to RNA *in vitro*, providing a possible mechanism for the function of hydroxylated DRG2.

Beyond the biochemically assigned members of the JmjC-only subfamily, three exist that remain poorly characterised and with some controversy. For example, the cellular role of JMJD8 is not understood, but it has been implicated in transcriptional control (Khoueiry et al., 2017), cellular metabolism (Boeckel et al., 2016), and NF-κB signalling (Yeo et al., 2016). JMJD8 is unusual because it contains a variation within the aforementioned conserved catalytic triad: In place of the iron-coordinating aspartic acid, JMJD8 contains a histidine residue, which raises questions regarding its catalytic potential (Yeo et al., 2017). However, evidence from experimental mutations in FIH suggests that two histidines can be sufficient for oxygenase activity, indicating that JMJD8 could function as a hydroxylase (Hewitson et al., 2008). However, the substrates of its potential hydroxylase activity remain unclear.

The second of the three poorly characterised JmjC-only hydroxylases is 'heat shock protein beta-1 associated protein 1' (HSPBAP1), which was originally identified as interacting with heat shock protein 27 (Liu et al., 2000). A physiological substrate of HSPBAP1 has yet to be identified. The third protein which also has no validated substrate is JMJD5, which forms the subject of this thesis and will be extensively discussed in Section 1.2.

The evidence for physiologically critical roles of 2OG oxygenases is expansive. As such, it is perhaps unsurprising that these enzymes are widely implicated in the progression of human disease.

1.1.7   2OG oxygenases and disease

The role of 2OG oxygenases in fundamentally important cellular processes such as gene expression control may explain growing evidence supporting their contribution to disease pathogenesis. While 2OG oxygenases have been implicated in a variety of disease, including cardiac and pulmonary disorders, the next section will focus on developmental disorders and cancer because of their relevance to the hydroxylase studied in this thesis.

1.1.7.1   Developmental disorders

Some of the strongest evidence for a role of 2OG oxygenases in development comes from the KDM6 group. Two enzymes, KDM6A (UTX) and KDM6B (JMJD3), catalyse the demethylation of trimethylated K27 on histone 3 and are thought to be required for normal development (Agger et al., 2007). UTX inactivation through mutations or deletions is associated with Kabuki syndrome, a developmental disorder characterised by intellectual and developmental defects (Banka et al., 2015). Additionally, JMJD3 variants have been shown to be associated with neurodevelopmental delays, including motor function and speech, as well as dysmorphic facial features (Stolerman et al., 2019). The relationship between histone methylation and development disorders has been extensively reviewed by Kim et al. (2017), who discussed the role of further 2OG oxygenases, KDM5C and PHF8 in these disorders (Kim et al., 2017).

Beyond histone demethylation, JmjC-only hydroxylases are also beginning to be appreciated for their role in human developmental disorders. JMJD7 which catalyses the hydroxylation of DRG1/2, as discussed above (Markolovic et al., 2018), has been shown to have novel single nucleotide polymorphisms that increase the risk of autism spectrum disorders (Matsunami et al., 2014). Indeed, the role of the JMJD7 substrate DRG1/2 in disease was reviewed by Westrip et al., who describe evidence for roles of both DRG1 and 2 in developmental disorders (Westrip et al., 2021). Additionally, recent studies have demonstrated that inactivating DRG1 mutations that reduce GTPase activity led to developmental disorders (Westrip et al., 2023). This evidence suggests the JMJD7-DRG1/2 pathway

may be important in developmental disorders; however, the molecular mechanisms involved have yet to be elucidated.

1.1.7.2 Hydroxylation, demethylation, and cancer

Potential roles of the 2OG oxygenase family in cancer appear to be extensive. Indeed, all the gene expression functions outlined above (1.1.5) are heavily implicated in tumourigenesis. For example, the TET2 DNA demethylase is frequently mutated in acute myeloid leukaemia (Ko et al., 2015), as well as in solid tumours (Huang and Rao, 2014). The potential importance of HIF prolyl hydroxylases in cancer is also widely reported. For example, several lines of evidence suggest a tumour-suppressive role for both PHD2 and PHD3. Interestingly, the downstream effects of these enzymes on HIF transcription targets can be 'hijacked' by tumour cells owing to the hypoxic nature of many solid tumours and the oxygen dependency of the enzymes (Ploumakis and Coleman, 2015).

The large KDM family has also been widely implicated in tumourigenesis of multiple cancer subtypes (Johansson et al., 2014; D'Oto et al., 2016; Højfeldt et al., 2013). For example, the various enzymes of the KDM4 group (KDM4A-C/JMJD2A-C) have been shown to be overexpressed in aggressive cancer subtypes including lung, prostate, breast, and colorectal tumours (Berry and Janknecht, 2013). While the KDM4 subgroup showed predominantly pro-tumourigenic roles, the KDM5 family appeared to be more complex. As discussed by D'Oto et al. (2016), KDM5A is overexpressed in multiple cancers and its loss reduces the proliferation of retinoblastoma (Rb) heterozygous knockout mice (Lin et al., 2011). KDM5B is also thought to have an oncogenic role, owing to its overexpression in prostate, breast, and testicular cancer (Xiang et al., 2007; Barrett et al., 2002). In contrast, KDM5C is thought to have a tumour suppressive role; inactivating mutations in the protein were identified in clear cell renal carcinoma (Dalgliesh et al., 2010). This variation in the oncogenic behaviour of this subfamily raises the concept of context-dependent roles of these enzymes in cancer.

Potential context-dependent roles are also seen in JmjC-only oxygenases, including the ribosomal oxygenase MINA53. In support of its pro-oncogenic role, high expression of MINA53 has been

observed in multiple cancer subtypes, including lung (Lu et al., 2009), pancreatic (Kumar et al., 2017; Tan et al., 2014), and gastric (Xing et al., 2014; Aziz et al., 2020) tumours. Conversely, MINA53 has been suggested to have a tumour-suppressive role in lung adenocarcinoma cells, where overexpression of MINA53 reduced proliferation (Komiya et al., 2010). Questions remain regarding the underlying mechanisms involved, but these could relate to the dual functions of MINA53 outlined above (Bundred et al., 2018).

Other JmjC-only oxygenases appear to have a more definitive role in cancer. The well characterised hydroxylase FIH has been studied in multiple cancer subtypes, where it is implicated as a tumour suppressor, including in renal cell carcinoma (Kroeze et al., 2010), glioblastoma (Wang et al., 2014a), colorectal cancer (Chen et al., 2015a), ovarian cancer (Kang et al., 2018), and breast cancer (Kim et al., 2019). The assignment of FIH substrates has supported mechanistic studies of some of these roles; decreased FIH hydroxylation of HACE1 under hypoxia was reported to cause Rac1 hyperactivation and increased breast cancer invasion (Kim et al., 2019).

The less well-characterised JmjC-only oxygenases have also been linked to tumourigenesis. For example, JMJD8 was shown to have a pro-oncogenic role in colorectal cancer through the NF-κB pathway (L.Wang et al., 2019). Furthermore, a role for HSPBAP1 as an oncogene in prostate cancer (Saeed et al., 2015) and hepatocellular carcinoma (Yang et al., 2015) has been suggested. However, the mechanistic understanding of the role of either protein in cancer remains elusive owing to the absence of a physiological substrate. The final 'orphan' hydroxylase in the JmjC-only sub-family is JMJD5, which also has extensive links to cancer, in addition to developmental disorders.

1.2 JMJD5

1.2.1 JMJD5: histone demethylase, endopeptidase, or arginyl hydroxylase?

JMJD5 has been assigned contradictory catalytic activities. It was originally identified as a KDM that catalyses the demethylation of lysine 36 on histone 3 (H3K36me2), and as such, was named KDM8 (Hsia et al., 2010). Additionally, JMJD5 has been reported to act as an endopeptidase that clips the

ends of methylated histone tails: JMJD5 exhibited a preference to cleave after monomethylated H3K9 (Shen et al., 2017). Following this, a second study reported that the endoproteolytic activity of JMJD5 preferentially occurs after mono- or dimethylated arginine residues on H2, H3, and H4 (Liu et al., 2017a). Indeed, it has been suggested that the activity of JMJD5 is coupled to that of histone arginine methyltransferases (Liu *et al.*, 2022). Finally, JMJD5 was assigned as the first reported human arginine hydroxylase, catalysing hydroxylation of arginine at the C3 carbon *in vitro* (Fig 1.7A) (Wilkins et al., 2018)*.* This study resolved the crystal structure of the catalytic domain of JMJD5 in complex with 2OG, showing that JMJD5 contains a JmjC domain comprising eight β-strands, as well as seven α-helices, five further β-strands, and three 3_{10} helices, consistent with other JmjC hydroxylases (Fig 1.7B) (Wilkins et al., 2018). In addition, typical of 2OG oxygenases, JMJD5 contains an HXD/E…H motif comprised of H321, D323, and H400 that coordinates the Fe(II) co-factor (Fig 1.7C). JMJD5 was also demonstrated to have four residues involved in direct binding to the 2OG cofactor (Y272, K336, S318, and N327) as well as three residues which create a hydrophobic pocket through side chain properties (W310, L329, and V402) (Fig 1.7C) (Wilkins et al., 2018). Interestingly, and in contrast to other JmjC-only hydroxylases, the catalytic pocket for 2OG binding is much more compact in JMJD5 (Islam *et al.*, 2022).
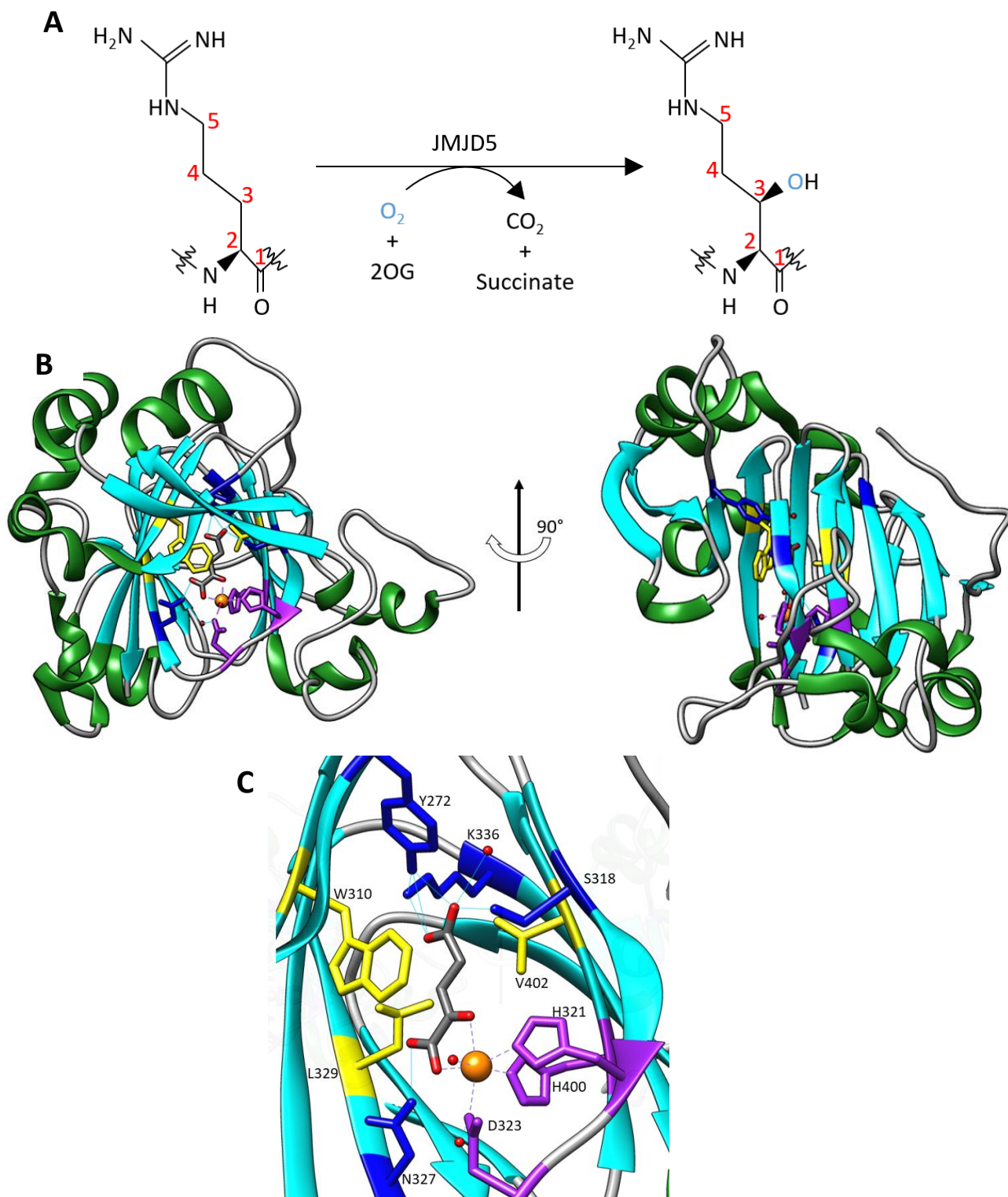
Figure 1.7. JMJD5 crystal structures demonstrates features typical of 2OG oxygenases.
(A) JMJD5 catalysed C3 arginyl hydroxylation. Carbon numbers of arginyl residue are indicated in red. **(B)** The JMJD5 catalytic domain contains eight β-strands, as well as seven α-helices, five further β-strands and three 3_{10} helices. Beta strands coloured in cyan, and all helices coloured in green. **(C)** View of JMJD5 catalytic pocket. Fe(II) binding residues coloured in purple (H321, H400 and D323), 2OG binding residues in blue (Y272, S318, K336 and N327), other residues involved in creating a hydrophobic environment for 2OG are coloured in yellow (W310, L329 and V402M). Manganese coloured in orange and 2OG in grey. H_2O molecules are depicted as red spheres. All images generated using PDB 6F4N and the protein modelling software chimera.

Current biochemical and structural analyses support the assignment of JMJD5 as an arginyl hydroxylase, for the following reasons. JMJD5 catalyses a 16 Da mass shift (indicative of a hydroxylation event) on a synthetic peptide derived from the Regulator of Chromosome condensation domain-containing protein 1 (RCCD1) and 40S ribosomal protein (RPS6) (Wilkins et al., 2018). Further analysis identified a hydroxylation event at a specific arginine residue in each case. Importantly, the same authors were unable to replicate the *in vitro* KDM activity of JMJD5 using synthetic peptide fragments of H3. Similarly, an independent study showed that JMJD5 overexpression had no effect on H3K36me2 levels *in vivo* (Youn *et al.*2012; Wilkins et al. 2018). Additionally, the initial study reporting endopeptidase activity for JMJD5 demonstrated mutation of the key Fe(II) binding residue (H321A) was not required for the proposed association with H3 (Shen et al., 2017). Furthermore, the second study provided no evidence that any known co-factors critical for the catalytic cycle of JMJD5 were required for the endoproteolytic activity (Liu et al., 2017a).

JMJD5 is not the only JmjC hydroxylase to be assigned with KDM activity that has not been subsequently verified. Two closely related enzymes MINA53 and NO66 have been assigned with H3K9me3 and H3K4/K36 demethylase respectively (Lu et al., 2009; Sinha et al., 2010). Why this apparent mis-assignment has occurred for multiple JmjC-only hydroxylases is unclear. One possible explanation is the conserved JmjC catalytic domain between the two sub-families may have led to targeted approaches to identify histone demethylation substrates leading to false positive results. Indeed, it is thought KDMs evolved from JmjC-only hydroxylases (Chowdhury et al., 2014), which suggests they might share substrate targets. Prior to its identification of a lysyl hydroxylase of the U2AF65 splicing factor, JMJD6 was described as a histone arginyl methyltransferase (Webby et al., 2009b; Chang et al., 2007). Subsequently, owing to some primary sequence similarities, JMJD6 lysyl hydroxylation events were reported on histone substrates (Unoki et al., 2013), which could also contribute to the confusion in catalytic assignments across the related JmjC-only hydroxylases.

JMJD5 also shows clear homology to other validated JmjC protein hydroxylases, including JMJD6 and FIH (Del Rizzo, Krishnan, and Trievel, 2012)(Fig 1.8A). Indeed, the residues involved in catalysis showed remarkable homology between FIH and JMJD5 (Fig 1.8B). JMJD5 structural analysis also suggested that the active site was not in the correct conformation or size to accommodate a large, methylated lysine residue (Del Rizzo et al., 2012). Additionally, the JMJD5 catalytic pocket contains three residues (Q275, W310, and W414) that would likely prevent methylated lysine from entering (Wang et al., 2013b), two of which are also conserved in FIH (Q147 and W296) (Fig 1.8C). Although the peptides of RCCD1 and RPS6 were shown to be hydroxylated *in vitro*, no evidence was provided to indicate that these proteins are physiologically relevant substrates of JMJD5 (Wilkins et al., 2018). Therefore, an important avenue for further studies is to establish the physiologically relevant substrates of JMJD5.
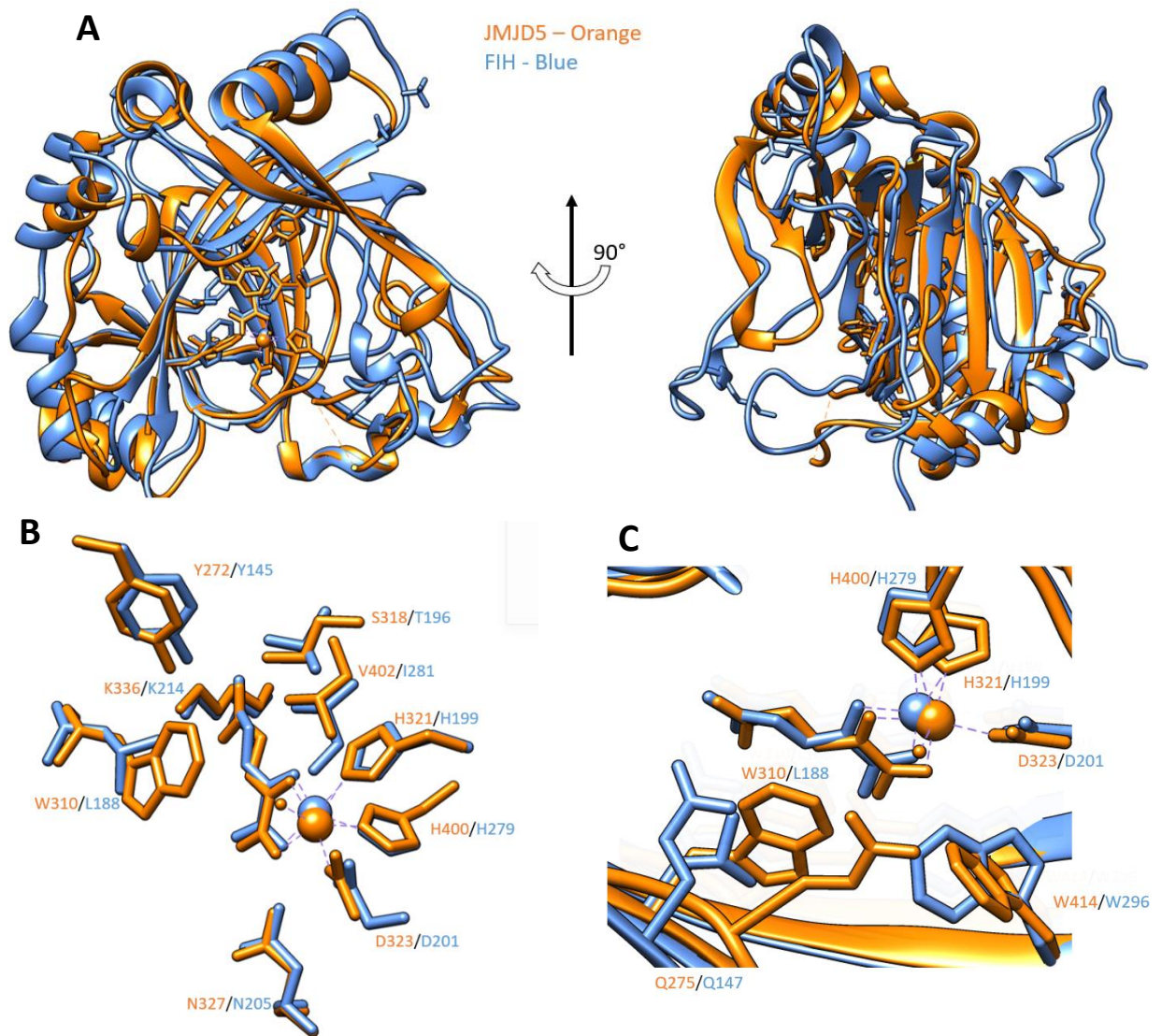
Figure 1.8. JMJD5 structure demonstrates homology with FIH.
(A) JMJD5 (orange) shows structural homology with FIH (Blue). Structures were overlayed using Chimera protein modelling software. **(B)** JMJD5 (orange) and FIH (blue) show homology in residues involved in catalysis. **(C)** JMJD5 (orange) contains residues (W414 and Q275) conserved in FIH (W296 and Q147) (blue) that may block methylated lysine entering the catalytic site. All images generated using PDB 6F4N (JMJD5) and 2ILM (FIH). Figures adapted from Del Rizzo et al 2012.

1.2.2   JMJD5 is required for normal development

The multiple contradictory assignments of JMJD5 biochemical activity have complicated an understanding of its role in cells. Indeed, studies have suggested that JMJD5 is important for a variety of seemingly disparate different pathways. That being said, it is clear that JMJD5 is biologically important, as gene knockout causes embryonic lethality at the mid-gestation stage in mouse embryos (Ishimura et al., 2012; Oh and Janknecht, 2012). Interestingly, knockout of *p53,* or its target *p21,*

partially reduced the severity of the JMJD5 knockout phenotype (Ishimura et al., 2012, 2016). Additional evidence has suggested that JMJD5 depletion leads to p53-dependent upregulation of 'TP53-induced glycolysis and apoptosis regulator' (TIGAR), leading to metabolic reprogramming (Liu et al., 2023). This relationship with p53 may indicate a role for JMJD5 in normal cell cycle regulation and cellular proliferation. Indeed, JMJD5 is thought to be important for mitosis; JMJD5 depletion leads to spindle assembly defects, decreased microtubule stability, and ultimately mitotic arrest (He et al., 2016). Additionally, JMJD5 depletion reduced pluripotency and caused G1 phase arrest in human embryonic stem cells, which was also linked to p21 (Zhu et al., 2014).

JMJD5 has also been implicated in DNA repair, with loss of *jmjd5* in *C. elegans* causing increased sensitivity to ionising radiation and defects in double-strand repair due to aberrant retention of RAD-51 at double-strand breaks (Amendola et al., 2017). Interestingly, depletion of JMJD5 also leads to defective mismatch repair in NIH3T3 cells (Suzuki et al., 2006).

Additional roles of JMJD5 have also been suggested in lower organisms. In *C. elegans*, JMJD5 has been proposed to regulate germline immortality by regulating gene expression through its controversial histone demethylase activity (Zaghet et al., 2021). The maintenance of germline immortality is essential for maintaining non-DNA genetic and epigenetic information across cell lineages to preserve species fertility (Furuhashi and Kelly, 2010).

JMJD5 is also implicated in regulation of the circadian system. In Human U20S cells, the depletion of JMJD5 resulted in a significantly shorter circadian length. Furthermore, in conjunction with TOC1, JMJD5 regulates the circadian system in Arabidopsis (Jones et al., 2010; Jones and Harmer, 2011). JMJD5 knockout in Drosophila also shortens the circadian period length and leads to a reduction in daytime sleep (Shalaby et al., 2018). While these initial studies were published in the context of JMJD5 as a KDM, more recent evidence suggests that the phenotype may not be associated with changes in H3K36 methylation (Jones et al., 2019).

Both cell cycle regulation and DNA repair play essential biological roles, and inefficiencies in either of these processes are widely associated with disease progression. Therefore, the assigned functions of JMJD5 in these pathways highlights the importance of better understanding the role of JMJD5 in the progression of human pathology.

1.2.3 JMJD5 and disease

1.2.3.1 JMJD5 has a context dependent role in cancer

As discussed above, mounting evidence implicates the majority of the JmjC-only protein family in tumourigenesis, including JMJD5 (Oh et al., 2019). However, JMJD5 appears to share characteristics with MINA53 in that both have been assigned with pro- and anti-tumourigenic functions. Evidence for JMJD5 as a context-dependent enzyme involved in cancer progression is discussed below.

First, *JMJD5* was identified through a tumour suppressor gene screen, in which JMJD5 depletion increased the number of mutational events (Suzuki et al., 2006). In a variety of different cancer types, including hepatocellular carcinoma, liver, and pancreatic cancers, lower expression levels of JMJD5 were associated with decreased patient survival (Wu et al., 2016; Chang, Forde, and Lai, 2019). Furthermore, in both lung cancer cells and intrahepatic cholangiocarcinoma cells, the expression of JMJD5 was found to be downregulated compared to non-cancerous lung and liver cells (Wang et al., 2012, 2016). Additionally, a new line of evidence has emerged for other cancer subtypes, suggesting that JMJD5 may play a role in regulating tumour metabolism to suppress proliferation. JMJD5 is reported to inhibit c-Myc, thereby preventing the expression of Myc-dependent target genes involved in glycolysis, which can drive pancreatic cancer proliferation (Wang et al., 2022). Furthermore, as discussed above, JMJD5 knockdown increased TIGAR expression, which was linked with increased glucose flux through glycolysis to drive lung cancer progression (Liu et al., 2023).

In contrast, other evidence supports a pro-tumourigenic function in some contexts. First, higher expression levels of JMJD5 correlate with decreased overall survival in colon cancer patients, and depletion of JMJD5 reduces invasion and proliferation of Caco-2 colon cancer cells (Zhang et al., 2015).

Furthermore, in patients with highly invasive oral squamous cell carcinoma (OSCC), JMJD5 expression is elevated compared to that in non-cancerous cell types. JMJD5 siRNA in OSCC cell lines led to a reduction in migration and invasion (Yao et al., 2019). JMJD5 overexpression has also been associated with increased metastasis and a poor prognosis in breast cancer (Zhao et al., 2015b). Lastly, JMJD5 was shown to directly interact with the androgen receptor and pyruvate kinase M2 which are key proteins involved in therapeutic resistance that drives the progression of prostate cancer (H.J. Wang et al., 2019). Additionally, the role of JMJD5 as a metabolic regulator of glycolysis appears to be context-dependent, as downregulation of JMJD5 in glioblastoma inhibits glycolysis and proliferation (Song et al., 2022).

The available evidence indicates important context dependence to the roles of JMJD5 in cancer progression. Interestingly, there is also emerging data suggesting that JMJD5 inhibition underlies novel developmental disorders.

1.2.3.2 Impaired JMJD5 activity contributes to a novel developmental disorder

A recent study by Fletcher et al. (2023) characterised the role of JMJD5 in a novel developmental disorder characterised by intellectual disability, severe failure to thrive, and relative macrocephaly (Fletcher et al., 2023). Interestingly, this developmental disorder was shown to be related to increased 'replication stress' (discussed further in Chapter 3), which was dependent on the catalytic activity of JMJD5. The disorder described arises from inherited biallelic mutations in the JMJD5 gene. In one of the families described, one mutation causes alternative splicing, resulting in the loss of a large portion of the catalytic domain (Fletcher et al., 2023), whilst the second mutation was a missense mutation in the poorly characterised N-terminus (Fig 1.9). This missense mutation was a cysteine-to-tyrosine substitution at amino acid position 123, which the authors demonstrated reduced JMJD5 protein stability *in vivo* and activity *in vitro*. This study suggests that the N-terminus is critical for the physiological function of JMJD5.

The absence of a proven and validated biological substrate means that understanding the role of JMJD5 in disease is complicated. The identification of a hydroxylation substrate for JMJD5 would help understand not only its biochemical activity, but also identify the JMJD5-dependent pathway(s) involved in disease progression. Likewise, functional studies of JMJD5 could help substrate discovery efforts and support a better understand of its role in disease. The Coleman group was particularly interested in attempting to understand the role of the N-terminal region of JMJD5 in these contexts.

1.2.4 The JMJD5 N-terminus is poorly characterised

Although a large proportion of published JMJD5 work has attempted to understand its catalytic activity, the protein also contains a large N-terminal region that remains poorly characterised. Unfortunately, the tertiary structure of the N-terminal region was not resolved in the recent crystal structure of JMJD5 (Wilkins et al., 2018). However, primary sequence analyses did identify a nuclear export sequence (AA 31-37) and a nuclear localisation sequence (AA 152-170) (Fig 1.9) (Huang et al., 2013). Additionally, the N-terminus is reported to contain an RNA polymerase binding domain, which is proposed to regulate RNA pol II pausing (Liu et al., 2020). It should be noted however that this study was conducted in the context of the controversial histone tail endopeptidase activity of JMJD5. One verified function of the N-terminus is facilitation of the interaction of JMJD5 with 'Regulator of chromosome condensation domain-containing protein 1' (RCCD1) (discussed further in Chapter 2). Overall, the potential regulatory roles and importance of the JMJD5 N-terminus are largely unknown.
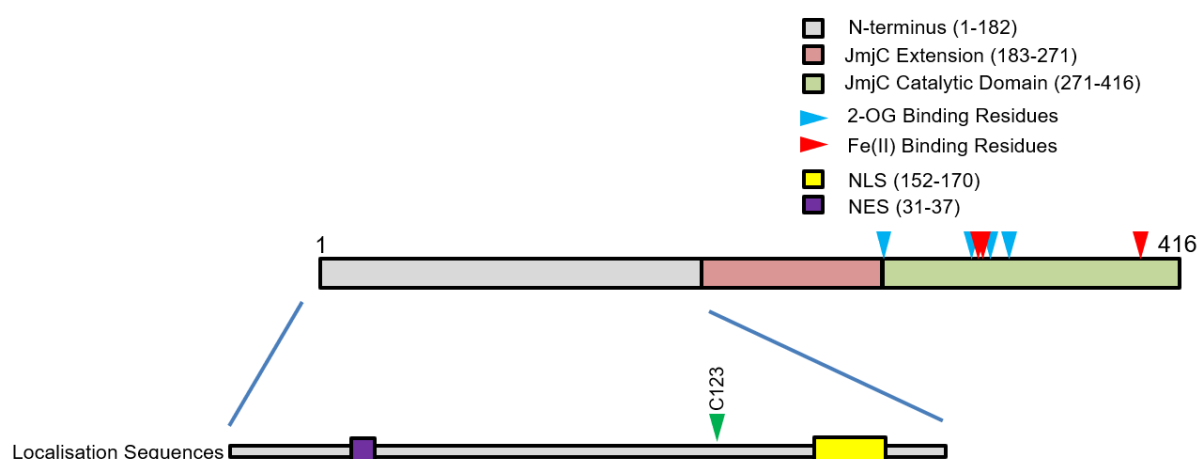
Figure 1.9 Domain architecture of JMJD5.
JMJD5 contains the indicated domain architecture. The JMJD5 N-terminus contains a nuclear localisation sequence (yellow) and a nuclear export sequence (purple). The green triangle indicates the location of the missense patient mutation C123Y.

1.3   Thesis aims

The overarching aim of this thesis is to build on our understanding of the cellular roles of JMJD5 and its importance in disease by focusing on the poorly characterised N-terminus. We hypothesize that the JMJD5 N-terminus has key functional roles in JMJD5 biology and aim to characterise this through the objectives below.

Our first objective was to define functional and structural domains within the N-terminus, with a particular focus on identifying specific residues required for the interaction of JMJD5 with the cancer-associated protein RCCD1 (Chapter 2). We then aimed to establish whether RCCD1 also regulates JMJD5-related cellular processes, specifically replication fidelity, and whether this requires its interaction with JMJD5 (and vice versa) (Chapter 3). Finally, we explored the physiological role of the JMJD5:RCCD1 complex through proteomic approaches and functional studies (Chapter 4).

Chapter 2:   Biochemical Characterisation of the JMJD5:RCCD1 Interaction

2.1   Introduction

As discussed in Chapter 1, the N-terminus of JMJD5 is poorly characterised in comparison to the catalytic domain. However, published work highlights potentially important functional roles for this region including a missense mutation in a novel developmental disorder (Fletcher et al., 2023) and binding to a protein called the Regulator of Chromosome Condensation Domain containing protein 1 (RCCD1) (Marcon et al., 2014; Wu et al., 2017). In this chapter we aimed to establish whether the N-terminus contained important structural regions and the potential functions of these. Because these analyses focused on the known interactor RCCD1, an overview of the RCCD1 literature is provided below before discussing the results obtained.

2.1.1   RCCD1

2.1.1.1   RCCD1 is an RCC1 super-family member

RCCD1 was originally named after Regulator of chromosome condensation 1 (RCC1) because its structural domains are homologous to RCC1. RCC1 contains seven consecutive RCC1 repeats each comprised of 50-60 amino acids which together form an RCC1-like domain (RLD), totalling 350-500 amino acids (Fig 2.1A). Since the identification of RCC1 in 1987, a superfamily comprised of almost 20 proteins containing at least one RLD has been reported (Fig 2.1B) (Ohtsubo et al., 1987; Hadjebi et al., 2008). The RCC1 superfamily contains functionally distinct subgroups (Fig 2.1B), with only the RCC1 subgroup being relatively well characterised. Members of the RCC1 subgroup contain a single RLD spanning the majority of the protein, and the absence of other recognisable functional domains and activities (Hadjebi et al., 2008).

A RCC1 Domain Structure

RCC1 repeat

RCC1 1 34-84
RCC1 2 85-136
RCC1 3 138-189
RCC1 4 191-257
RCC1 5 258-311
RCC1 6 312-362
RCC1 7 363-416

1 — RCC1 1 | RCC1 2 | RCC1 3 | RCC1 4 | RCC1 5 | RCC1 6 | RCC1 7 — 421

RCC1 Like Domain (RLD)

B RCC1 Superfamily

RCC1 Subgroup

RCC1   TD-60
DelGEF   WBSCR16
RCCD1?

HERC Subgroup

HERC1   HERC3
HERC2   HERC4
         HERC5
         HERC6

Large HERC proteins

Small HERC proteins

RCBTB Subgroup

RCBTB1   RCBTB2
IBtk

Kinase Subgroup

Nek9   Nek8

Miscellaneous Subgroup

Alsin   PAM
RPGR

Figure 2.1. RCC1 and the RCC1 superfamily of proteins.
(A) Domain structure of RCC1 highlighting RCC1 repeats and the formation of an RLD from seven consecutive RCC1 repeats. Figure annotated using the Uniprot bioinformatic databases. **(B)** Figure adapted from Hadjebi et al 2008. Proteins are grouped into 5 separate groups, the RCC1 subgroup, HERC subgroup, RCBTB subgroup, Kinase subgroup and the miscellaneous subgroup (proteins that do not fit into previous categories).

2.1.1.2   RCC1 and the RCC1 subgroup

RCC1 is a guanine nucleotide exchange factor (GEF) that catalyses GDP for GTP exchange on the nuclear Ras-related Ran GTPase, resulting in its activation (Bischoff and Ponstingl, 1991). The Ran GTPase cycle has key roles during mitosis, aiding both mitotic spindle formation and nuclear envelope reassembly (Bamba et al., 2002; Askjaer et al., 2002). Indeed, RCC1 localisation to chromosomes is required for proper mitotic spindle assembly (Moore et al., 2002). Structurally, the seven RCC1 repeats of the RLD form a seven bladed-propeller, each consisting of four antiparallel β-sheets, where the inner strand of each blade is aligned with the central shaft (Fig 2.2A+B) (Renault et al., 1998).
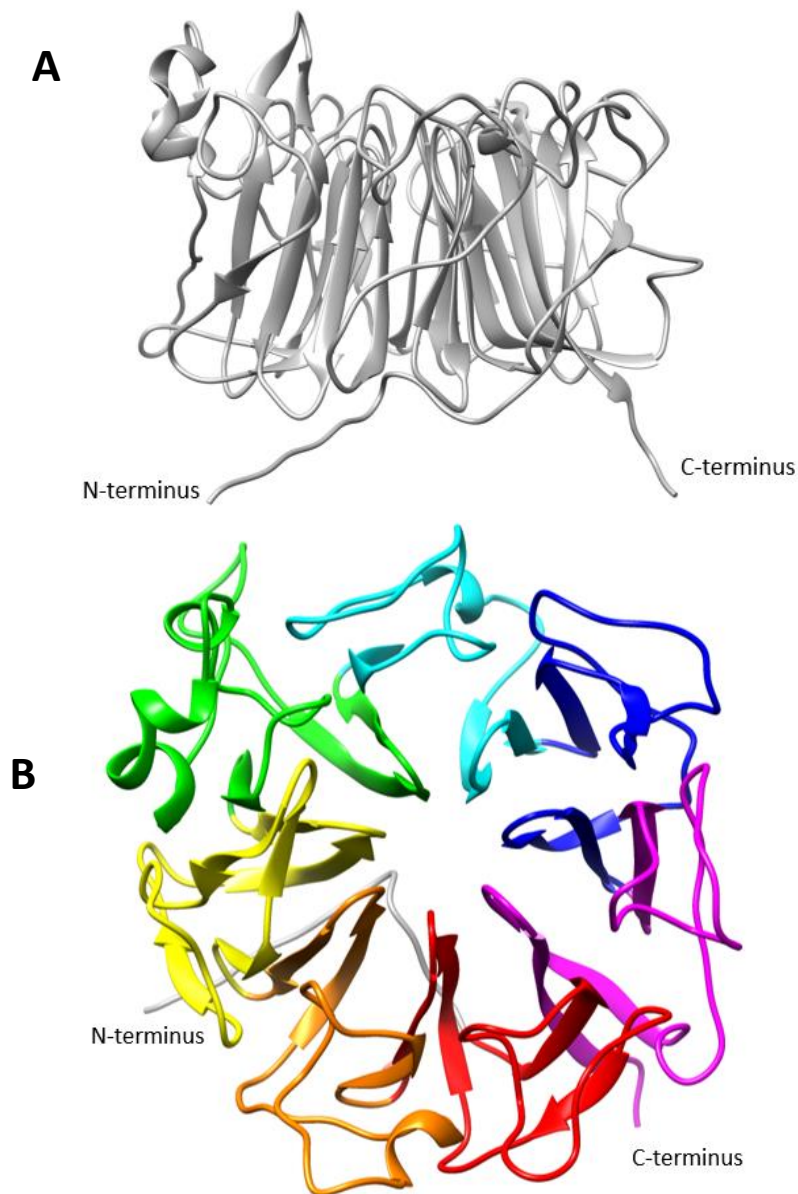
Figure 2.2. RCC1 has a seven bladed propeller structure comprised of its seven RCC1 repeats.
(A) Human RCC1. **(B)** Top view of human RCC1 with each RCC1 repeat highlighted in a different colour. (RCC1 repeat 1 = RED, repeat 2 = ORANGE, repeat 3 = YELLOW, repeat 4 = GREEN, repeat 5 = CYAN, repeat 6 = BLUE and repeat 7 = MAGENTA). Both images were generated from PDB file 1A12 using the protein modelling software chimera.

Other members of the RCC1 subgroup include telophase disk 60 (TD-60), deafness locus-associated putative guanine nucleotide exchange factor (delGEF) and William-Beuren syndrome critical-region 16 (WBSCR16). TD-60 (RCC2) is part of the chromosome passenger complex (CPC) important for prometaphase to metaphase transition (Mollinari et al., 2003). The crystal structure of TD-60 suggests homology to RCC1 with the RLD folding into a seven bladed propeller structure (Song et al., 2018).

DelGEF regulates proteoglycan secretion through binding to DelGEF interacting protein 1 (DelGIP1) and Sec5 (Sjölinder et al., 2004, 2002). WBSCR116 localises to the mitochondria and, like TD-60, also shows structural homology to RCC1 (Koyama et al., 2017). Finally, the structure of the 'retinitis pigmentosa GTPase regulator' (RPGR), which belongs to the miscellaneous subgroup, has been resolved also forms a seven bladed propeller (Wätzlich et al., 2013). Therefore, it seems highly likely that all reported RLDs form this structural conformation.

2.1.1.3   RCCD1

The domain architecture of RCCD1 in the UniProt database suggests it contains only four RCC1 repeats (Fig 2.3). However, as outlined above, all conventional RLDs contain seven consecutive RCC1 repeats, suggesting that RCCD1 would be unique in the family if this were the true (Hadjebi et al., 2008). Rather, structural prediction analysis (Professor Smerdon, personal communication) suggested that RCCD1 contains seven consecutive RCC1 repeats (Fig 2.3). This would be consistent with RCCD1 containing a single RLD that spans the whole protein, placing it in the RCC1 sub-family. Therefore, it is likely that RCCD1 forms a similar seven bladed propeller structural conformation. Currently, the biochemical role of the RCCD1 seven bladed propeller remains unclear and the cellular function of RCCD1 is poorly characterised. Studies in these areas are likely to be beneficial because it is increasingly clear that RCCD1 has important roles in cancer.
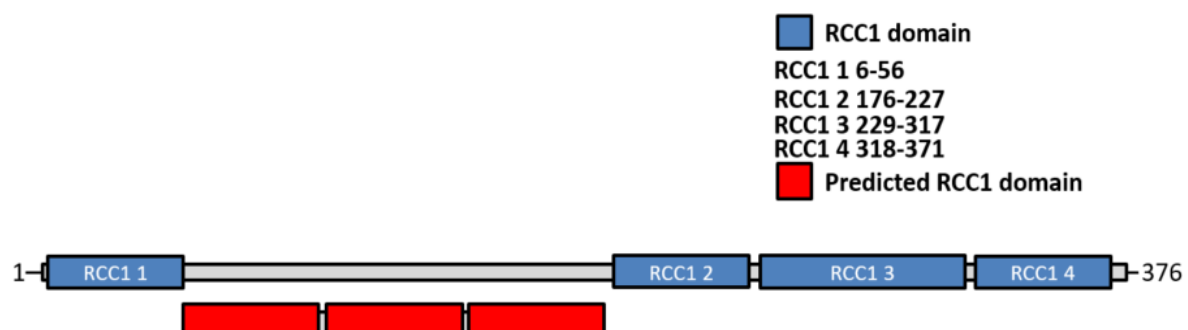
Figure 2.3. RCCD1 domain structure.
Graphical representation of the location of RCC1 repeats in RCCD1. Figure based on domain annotation data from the Uniprot server (blue rectangles). The additional RCC1 domains that we predict are indicated by the red rectangles.

2.1.2 RCCD1 and Cancer

RCCD1 was reported by Cai et al. (2014) as a possible tumour suppressor, through a genome wide association study identifying a novel single nucleotide polymorphism (SNP) that reduced RCCD1 expression correlating with increased breast cancer risk in East Asian women (Cai et al., 2014). This particular risk allele was also shown to correlate with reduced expression of RCCD1 in ovarian cancer (Kar et al., 2016; Gusev et al., 2019). Additional SNPs have also been reported in breast cancer that show the same relationship as above (Hoffman et al., 2017), indeed Ferreira et al. (2019) reported risk variants that increased RCCD1 expression correlated with decreased breast cancer risk (Ferreira et al., 2019). Further supporting the TSG function of RCCD1, decreased expression was also predicted to increase the risk of both prostate and pancreatic cancer respectively (Gusev et al., 2019; Zhong et al., 2020).

Conversely, there is some evidence suggesting, like JMJD5, RCCD1 may also act as an oncogene. Cheng et al (2019) reported that increased RCCD1 expression induced by the activity of the long non-coding RNA (LncRNA) LINC01419 was correlated with enhanced cell proliferation in lung adenocarcinoma (Cheng et al., 2019). Although the mechanism behind this remains unclear.

Interestingly, JMJD5 is implicated in many of the same tumour types as RCCD1 (Chapter 1). Furthermore, like JMJD5, the role of RCCD1 as a tumour suppressor versus oncogene appears to be complex and context dependent. Therefore, it is interesting to consider the possibility that the physical and/or function interaction of JMJD5 and RCCD1 may play a role in tumorigenesis.

2.1.3 JMJD5 and RCCD1

The JMJD5:RCCD1 interaction was originally identified as essential for proper chromosomal segregation. Depletion of JMJD5 and RCCD1 increased the number of multipolar spindles, a phenotype characteristic of spindle organisation defects and improper chromosome segregation (Marcon et al., 2014). Although the exact molecular mechanism is unknown, Marcon et al proposed that it relates to the controversial KDM activity of JMJD5. Whether the mitotic effects are direct, or indirect, also

remains unclear. The interaction of JMJD5 and RCCD1 is also purported to be important for the migration of non-small cell lung cancer (NSCLC) cells by stabilising cytoskeletal microtubules (Wu et al., 2017). Depletion of RCCD1 and JMJD5 led to an increase in the level of acetylated tubulin which is indicative of increased cytoskeletal microtubule stability. Again, the mechanism, which was also reported to involve the unconfirmed KDM activity, remains unclear.

Overall, there are a paucity of studies that consider the normal physiological role of RCCD1 in the context of the arginyl hydroxylase activity of JMJD5. Furthermore, the determinants and functional importance of the JMJD5:RCCD1 interaction, and the consequence of its deregulation in cancer, remain unclear. Therefore, in this chapter we aimed to characterise the biochemical nature of the JMJD5:RCCD1 interaction. We hypothesised that the methods utilised could help to identify the key amino acids on each protein responsible for facilitating the interaction.

2.2   Results

2.2.1   Bioinformatic prediction of the JMJD5 N-terminus

To begin to explore the function of the RCCD1:JMJD5 complex in more detail we first wanted to isolate the minimal region that RCCD1 binds on JMJD5. To support this, we aimed to generate a panel of JMJD5 truncation mutations that could be used in binding assays. Accidentally truncating a protein within an essential structural region could cause indirect effects and reduce protein stability and expression, thus complicating interpretation. Therefore, we first sought to undertake structural predictions of the JMJD5 N-terminus to better define structural regions of potential importance.

In collaboration with Chan Li (Smerdon Group) we submitted the full-length amino acid sequence of JMJD5 to the Xtalpred server. Although, XtalPred is predominantly used for predicting if a protein is crystallizable, it also provides predictions of secondary structure (Slabinski et al., 2007). XtalPred predicted regions of structure in the N-terminus (highlighted by grey rectangle) including one β-strand

and seven α-helices (Fig 2.4A+B) Interestingly, a region between amino acids 150-182 was also

identified as containing low complexity structure and disorder, potentially forming a flexible linker

between the N-terminus and catalytic domain (Fig 2.4A+B).
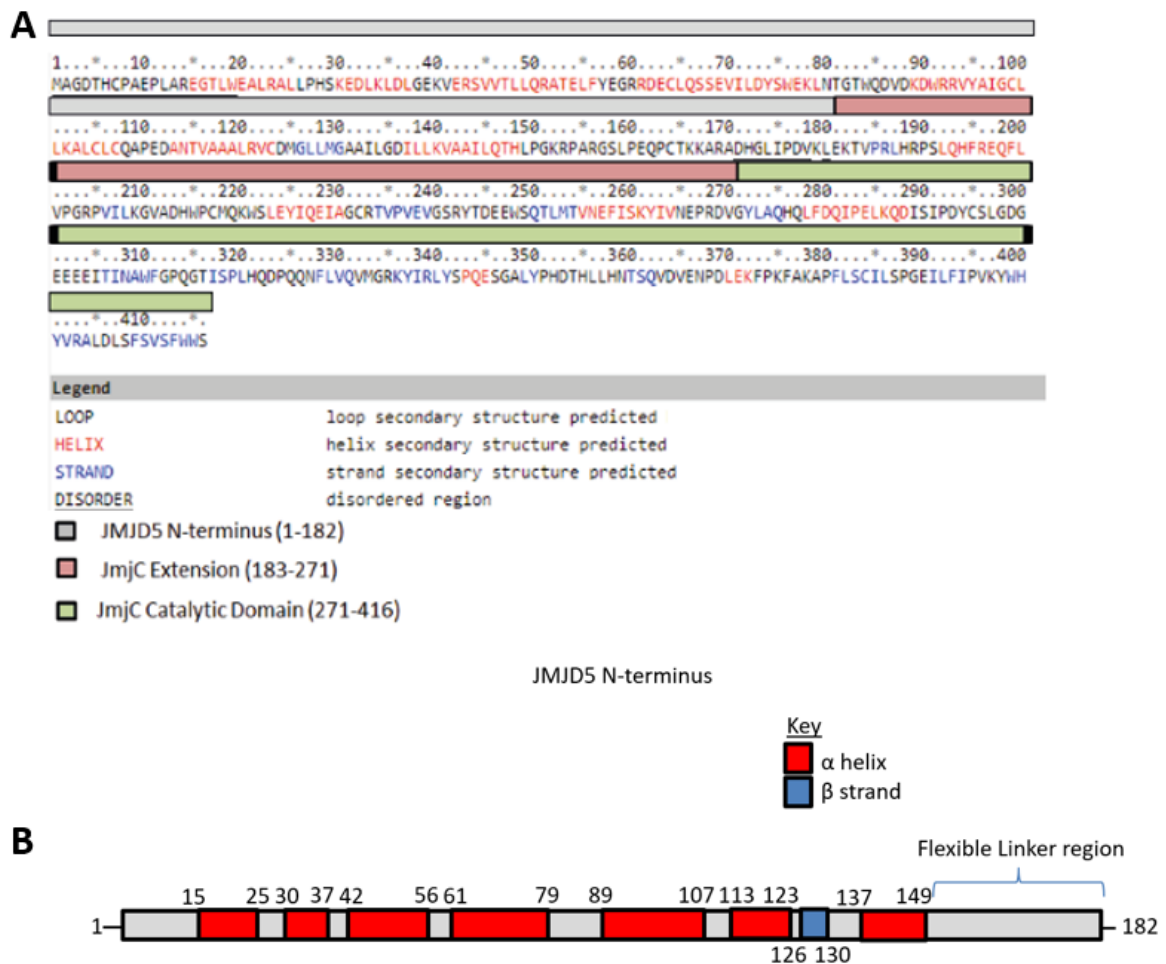
Figure 2.4. JMJD5 N-terminus contains regions of secondary structure according to predictions.
(A) Results of xtalpred server structural predictions for full length JMJD5. Red text indicates α-helices, blue text indicates β-strands and underlined text indicates regions of predicted disorder.
(B) Graphical representation of JMJD5 N-terminal tertiary structure predictions from Xtalpred.

2.2.2   Design and generation of panel of JMJD5 truncation mutations

Using our structure prediction for the JMJD5 N-terminus, we designed a panel of JMJD5 truncations

with N- and C-terminal deletions (Fig 2.5). The C-terminal deletions lacked the JmjC extension and

catalytic domains, consisting of only the N-terminus sequentially truncated after each predicted α-

helix. The N-terminal deletions contained the full JmjC extension and catalytic domains with varying lengths of the N-terminus included. The panel was then PCR cloned from full length JMJD5 into a HA-tagged mammalian expression vector (pEF6) (Fig 2.5).
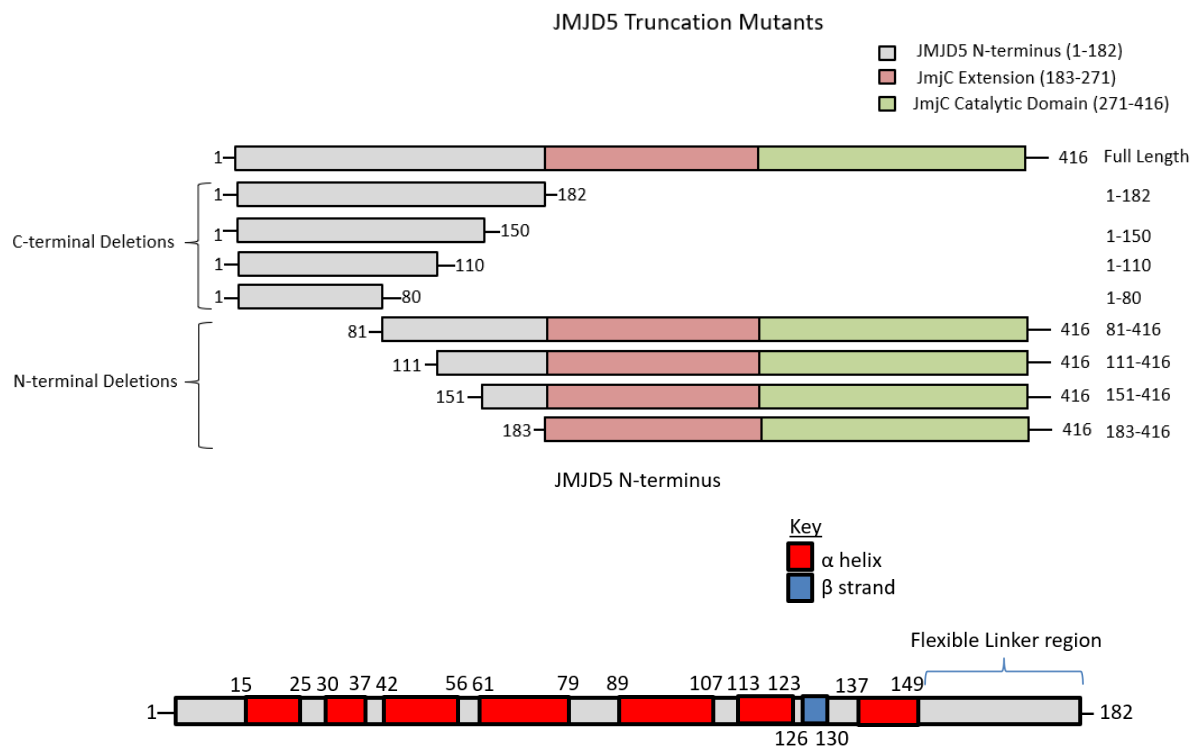
Figure 2.5. Panel of JMJD5 truncation mutants.
Graphical representation of JMJD5 truncation mutation panel. Representation of predicted N-terminus from Figure 2.4B is included for reference.

2.2.3 RCCD1 binds to the JMJD5 N-terminus

To test which JMJD5 truncations were able to bind RCCD1, the corresponding HA pEF6 constructs were transfected into HEK293T cells followed by anti-HA immunoprecipitation and Western blotting for endogenous RCCD1 (Fig 2.6). Full length JMJD5 was included as a positive control for RCCD1 binding.
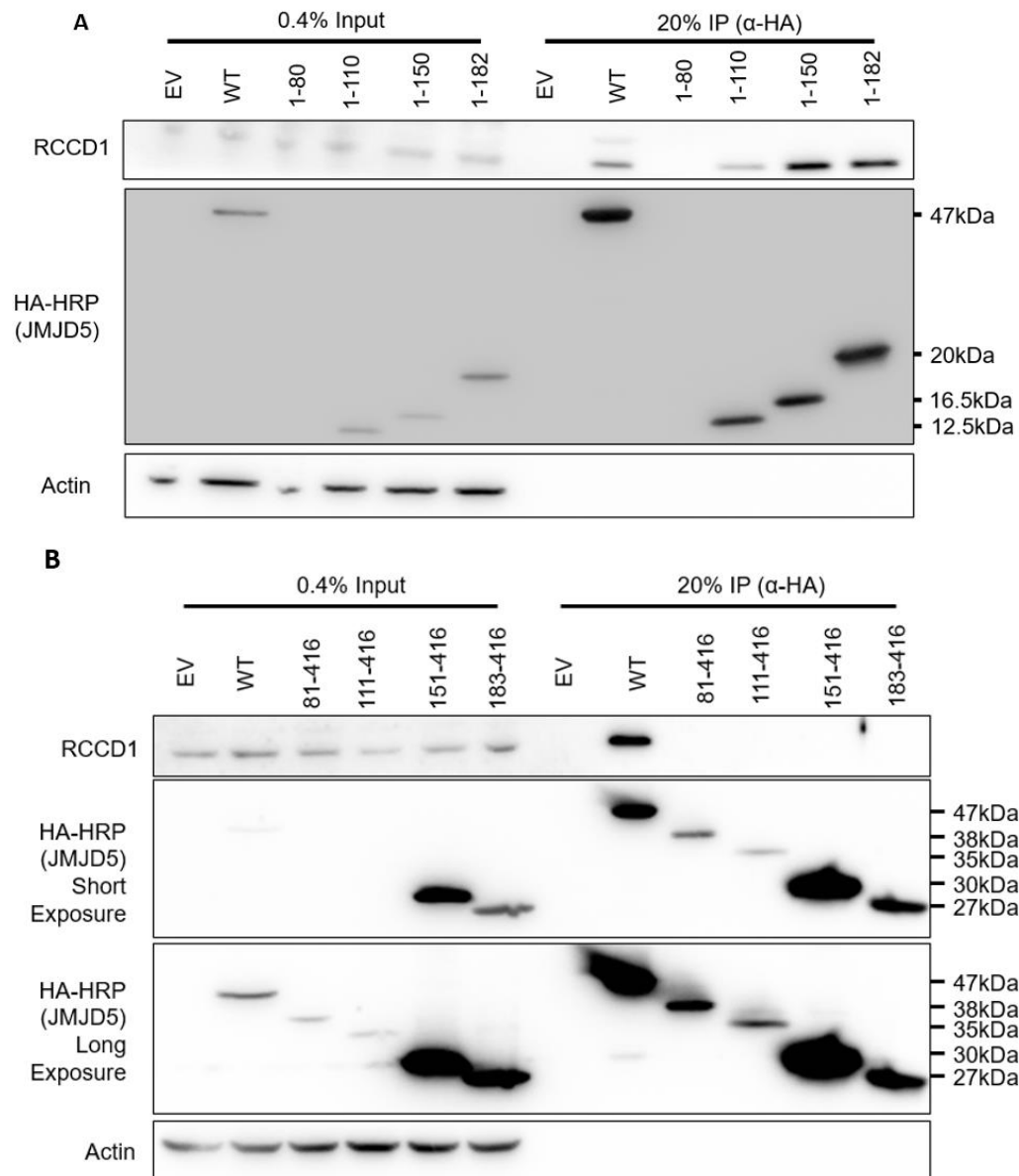
Figure 2.6. RCCD1 binds to the JMJD5 N-terminus.
(A) Anti-HA immunoprecipitates (IP) from HEK293T cell extracts expressing indicated C-terminal JMJD5 deletion constructs (see Figure 3.2) were Western blotted for HA, RCCD1 and Actin. Input is representative of the cell extract before anti-HA immunoprecipitation. **(B)** Anti-HA immunoprecipitates from HEK293T cell extracts expressing the indicated JMJD5 N-terminal deletion constructs were Western blotted as above. Western blots are representative of three biological repeats.

With the exception of JMJD5 1-80, which did not express, all of the C-terminal deletions were capable

of interacting with RCCD1 (Fig 2.6A). This is in line with previous findings which suggested that the

RCCD1 binding site is in the JMJD5 N-terminus (Marcon et al., 2014). Importantly, JMJD5 1-110 was

sufficient to bind RCCD1 (Fig 2.6A). Therefore, this suggests that the JMJD5:RCCD1 interaction is facilitated by amino acids within the 1-110 region of JMJD5.

In contrast, although each expressed, none of the N-terminal deletions supported binding to RCCD1 (Fig 2.6B). Of note, JMJD5 81-416 was not able to bind RCCD1. This suggests that amino acids within the first 80 are essential for the interaction. Taken together, the data suggests that JMJD5 residues 1-80 is necessary and sufficient for RCCD1 binding.

2.2.4   Testing recombinant protein expression of GST-tagged JMJD5

We sought to validate the results above using an orthogonal approach, specifically a 'pull-down' assay using recombinant JMJD5. The full library of truncations was cloned into a GST-tagged bacterial expression vector (pGEX-4T1). Before committing to a full purification process, we first sought to determine which constructs expressed through smaller scale expression trials. These smaller scale purifications generated insoluble, soluble, and GST-pulldown fractions with and without the addition of the inducing agent IPTG. These conditions aimed to provide a comprehensive analysis of the behaviour of each construct through the purification process, aiding any subsequent larger scale purifications.

Of the constructs tested, only 1-80 and 81-416 failed to express (Fig 2.7B + D). Of note, the 1-80 construct also failed to express in mammalian cells (Fig 2.6A). The remaining truncations all expressed and were the correct size (Fig 2.7). Therefore, six of our constructs were suitable for use batch purification and *in vitro* pull-down assays.
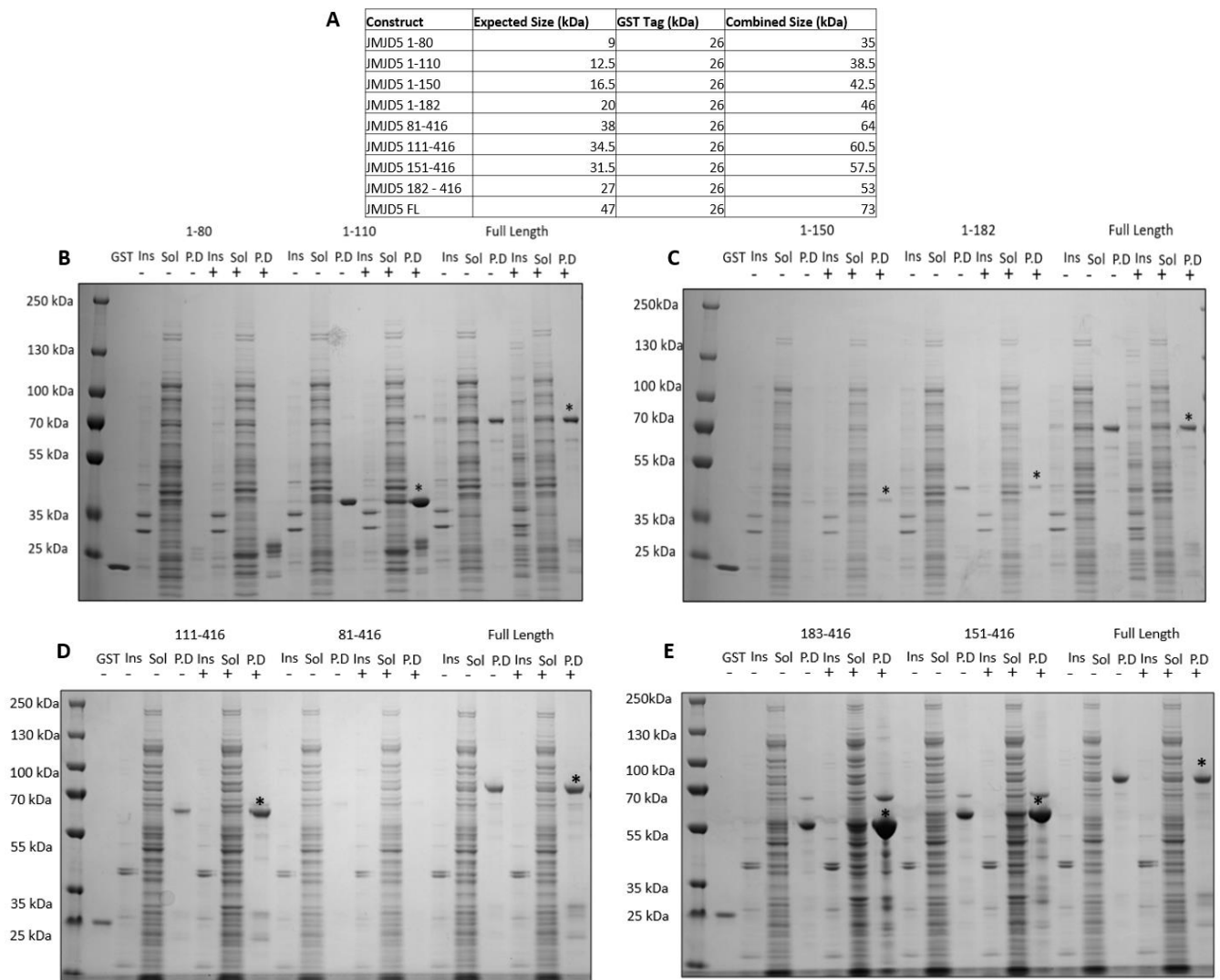
Figure 2.7. GST-tagged JMJD5 constructs express in E. coli.
(A) Table showing predicted size of JMJD5 N-terminal constructs **(B)** Results of protein expression test for JMJD5 truncations 1-80 and 1-110 in comparison to full length JMJD5 **(C)** Results of protein expression test for JMJD5 truncations 1-150 and 1-182 in comparison to full length JMJD5. **(D)** Results of protein expression test for JMJD5 truncations 111-416 and 81-416 in comparison to full length JMJD5. **(E)** Results of protein expression test for JMJD5 truncations 151-416 and 181-416 in comparison to full length JMJD5. Ins= insoluble fraction, Sol = soluble fraction, P.D = pulldown. Samples were induced with 0.5mM IPTG and left at 18°C ON. * Indicate the expressed and enriched JMJD5 truncation protein.

2.2.5 Recombinant JMJD5 1-110 is sufficient to bind RCCD1

To carry out *in vitro* pull-down assays, 12 µg of recombinant JMJD5 protein was incubated with cell lysate from HEK293T cells, followed by GST pulldown and Western blot for RCCD1. Coomassie staining was used to ensure that the amount of recombinant protein used in the assay was consistent across samples. Similar to the over-expression experiments in HEK293Ts (Fig 2.6), all recombinant C-terminal deletions (1-110, 1-150 and 1-182) were capable of binding RCCD1 (Fig 2.8). In contrast, and in line with my previous experiments, none of the N-terminal deletions (111-416, 151-426 and 183-416) showed binding to RCCD1 (Fig 2.8). Taken together, these observations support that JMJD5 1-110 contains the binding domain for RCCD1. Therefore, we next sought to further narrow down the minimal binding region within this domain.
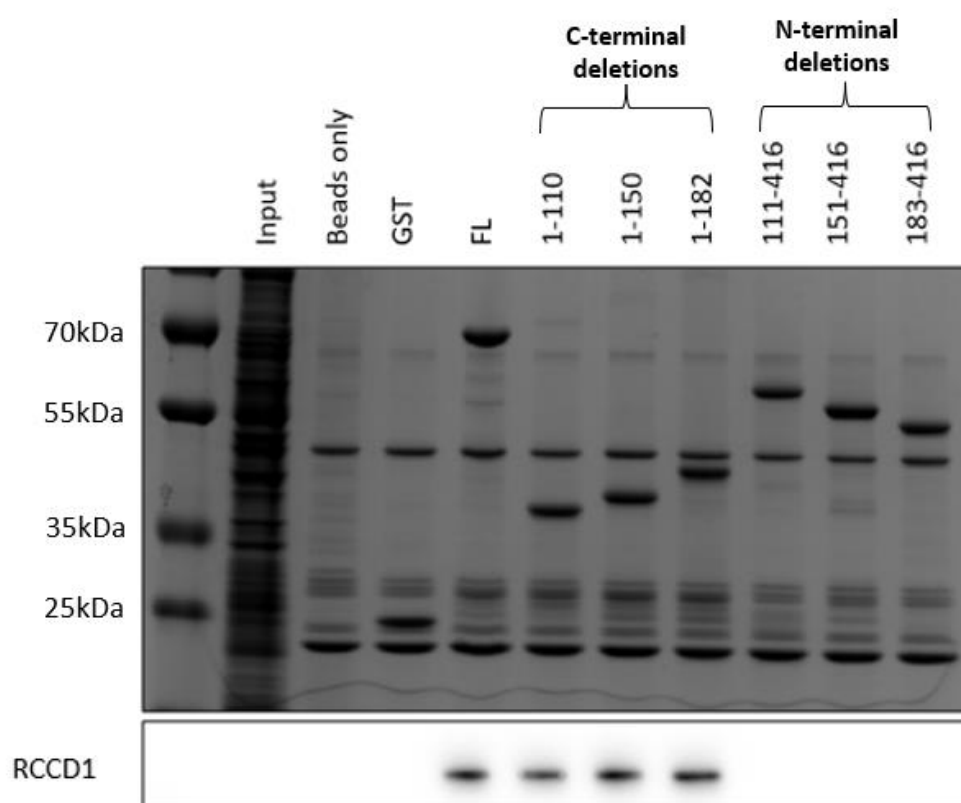
Figure 2.8. Recombinant JMJD5 1-110 is sufficient to bind RCCD1.
GST pulldown experiments using the indicated GST-tagged recombinant protein incubated in HEK293T cell lysates and Coomassie stained and Western blotted for RCCD1. FL = Full length recombinant JMJD5. Images representative of two independent biological repeats.

2.2.6 Design and generation of a panel of cancer mutations

Having determined that JMJD5 1-110 contains the region important for binding RCCD1, we wanted to identify the key amino acids required for the interaction. First, we considered how we might prioritise suitable candidates for analyses. Since both proteins have been implicated in overlapping cancer subtypes, as mentioned above, we hypothesised that the interaction might be important for tumour suppression and that the JMJD5:RCCD1 complex could be disrupted by specific cancer-associated missense mutations. Therefore, we created a list of JMJD5 cancer variants compiled from online databases COSMIC and cBIOportal (Appendix 1). To investigate the JMJD5-RCCD1 interaction in this way, we first focused on variants from this list located in the JMJD5 N-terminus, particularly within residues 1-110 (Fig 2.9A). When selecting our panel, we aimed to include mutations in predicted structural regions (Fig 2.9A), those in highly conserved regions (Fig 2.9B), and those predicted to be damaging by online bioinformatics tools. To support this, we generated a multiple sequence alignment (MSA) of full length JMJD5 from across several species (Appendix 2) and used PolyPhen2 to predict damaging cancer mutations (Fig 2.9C). We intended to test these mutations as with our truncations above, using immunoprecipitation of tagged constructs followed by analysing RCCD1 binding through Western blotting. The rationalisation for selecting each mutation is provided at the start of the next sections.

Figure 2.9. Design of JMJD5 N-terminal cancer mutation panel.
(A) The panel of cancer mutations selected for further study plotted against the predicted JMJD5 N-terminal structural regions. **(B)** Multiple sequence alignment (MSA) generated using MEGA7 and shaded using the Texshade LaTeX package. The red arrows indicate the panel of cancer mutations selected. **(C)** Summary table showing PolyPhen2 predictions for the panel of JMJD5 N-terminal cancer mutations. Mutations on PolyPhen2 scale closer to 1 are probably damaging, closer to 0 equals benign.

Figure 2.9. Design of JMJD5 N-terminal cancer mutation panel.
(A) The panel of cancer mutations selected for further study plotted against the predicted JMJD5 N-terminal structural regions. **(B)** Multiple sequence alignment (MSA) generated using MEGA7 and shaded using the Texshade LaTeX package. The red arrows indicate the panel of cancer mutations selected. **(C)** Summary table showing PolyPhen2 predictions for the panel of JMJD5 N-terminal cancer mutations. Mutations on PolyPhen2 scale closer to 1 are probably damaging, closer to 0 equals benign.

Mutation	Cancer subtype	PolyPhen2 Prediction	Score
D4N	Melanoma	Benign	0.244
P27L	Melanoma	Probably Damaging	1
L35F	Lung Squamous Cell Carcinoma	Probably Damaging	1
E39K	Melanoma	Possibly Damaging	0.483
V45M	Not Available	Benign	0.085
V70L	Not Available	Benign	0.006
S75F	Not Available	Probably Damaging	0.966
E77K	Cutaneous Melanoma	Probably Damaging	1
L79P	Large Intestine Carcinoma	Probably Damaging	1
D90N	Leukemia	Possibly Damaging	0.651
R92C	Uterine Endometrioid Carcinoma	Probably Damaging	1
R121Q	Not Available	Possibly Damaging	0.521
I144V	Breast Invasive Ductal Carcinoma	Benign	0.001
R153S	Non Small Cell Lung Cancer	Possibly Damaging	0.651

2.2.7 JMJD5 E77K reduces RCCD1 binding

The variants were grouped depending on their location within the JMJD5 primary sequence and tested in anti-HA immunoprecipitations, as follows. Each group was transfected into HEK293T cells with wildtype (WT) JMJD5 (JMJD5^{WT}), followed by anti-HA immunoprecipitation, SDS-PAGE and Western blot analysis for RCCD1, HA and Actin.

2.2.7.1 Group 1 variants (D4N to V45M)

Group 1 contained five variants (D4N, P27L, L35F, E39K and V45M) all located in the poorly conserved first 50 amino acids of JMJD5 (Fig 2.10A). Although D4N is a recurrent variant in the extreme N-terminus of JMJD5, the aspartate at this position is not well conserved (Fig 2.10A) and the mutation is not predicted to be damaging (Fig 2.9C). Conversely, the proline at position 27 is 100% conserved (Fig 2.10A), and mutation to leucine was assigned a 'probably damaging' score (Fig 2.9C). Both L35 and E39 are highly conserved and located in predicted regions of secondary structure (Fig 2.9A and Fig 2.10A), and their variants (L35F, E39K) are recurrent with high predicted damage score (Fig 2.9C). Although V45 was selected because it is highly conserved residue and in a predicted region of secondary structure (Fig 2.9A and Fig 2.10A), its mutation to methionine was not predicted to be damaging (Fig 2.9C).
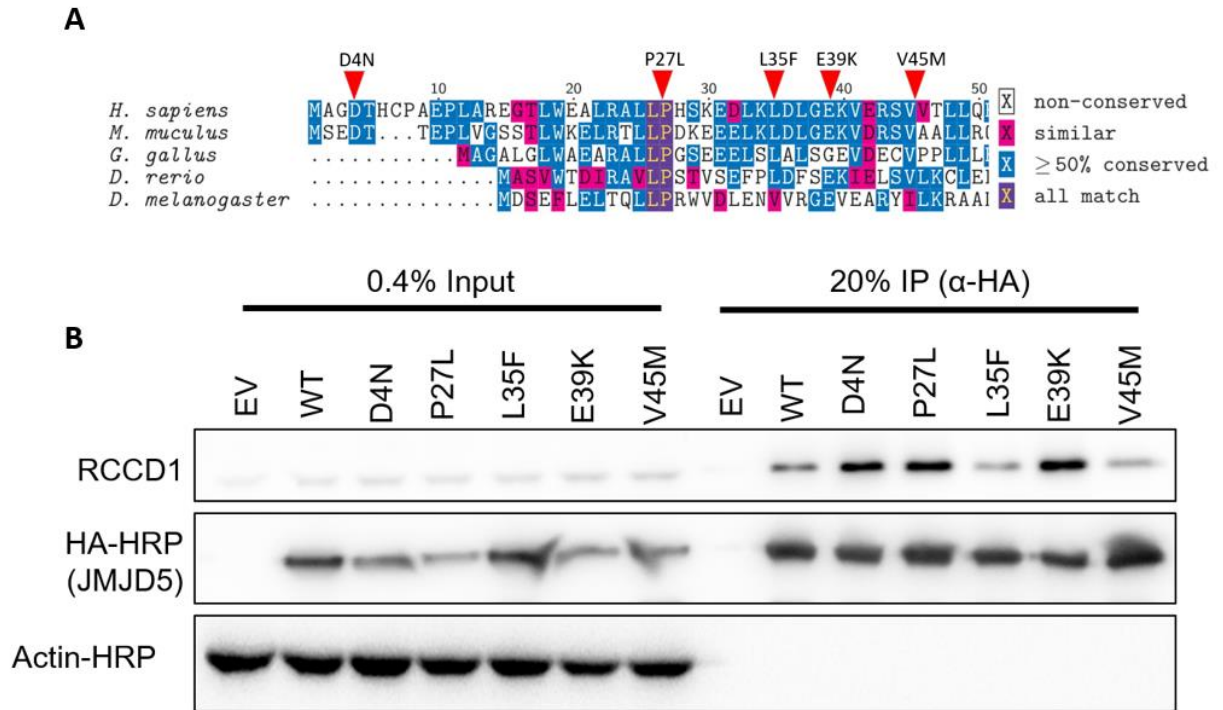
Figure 2.10. Group 1 D4N-V45M JMJD5 cancer variants IPs.
(A) MSA generated using MEGA7 and shaded using the Texshade LaTeX package. Red arrows indicated JMJD5 cancer variants contained in this group. **(B)** Anti-HA immunoprecipitates from HEK293T cell extracts expressing indicated JMJD5 variants were Western blotted for HA, RCCD1 and Actin. Input is representative of the cell extract before anti-HA immunoprecipitation. Western blots representative of three independent biological repeats.

Group 1 variants were transfected into HEK293T cells for anti-HA immunoprecipitation to monitor potential differences in RCCD1 binding. By first examining the input blots we observed reduced expression of the JMJD5 variants in comparison to wildtype. All with the exception of JMJD5[L35F] showed reduced expression with the most severe phenotype in JMJD5[P27L] (Fig 2.10B). This suggests an intrinsic effect of these variants on reducing protein stability. In addition, none of these variants caused a significant reduction in RCCD1 binding compared to the wildtype (Fig 2.10B). Although JMJD5[D4N], JMJD5[P27L], and JMJD5[E39K] may have shown modestly more RCCD1 binding compared to wildtype, the relevance of this is currently unclear (Fig 2.10B). Likewise, for JMJD5[L35F] and JMJD5[V45M], which only showed a slight reduction in RCCD1 binding (Fig 2.10B).

2.2.7.2 Group 2 variants (V70L to L79P)

The second group consists of four variants V70L, S75F, E77K and L79P, with all located in a highly conserved region between amino acids 70-80 in the N-terminus (Fig 2.11A). All variants were in a region of predicted secondary structure (Fig 2.9A) and all except JMJD5^{V70L} were predicted to be highly damaging by PolyPhen2 (Fig 2.9C). JMJD5^{E77} was fully conserved across the five species analysed (Fig 2.11A).
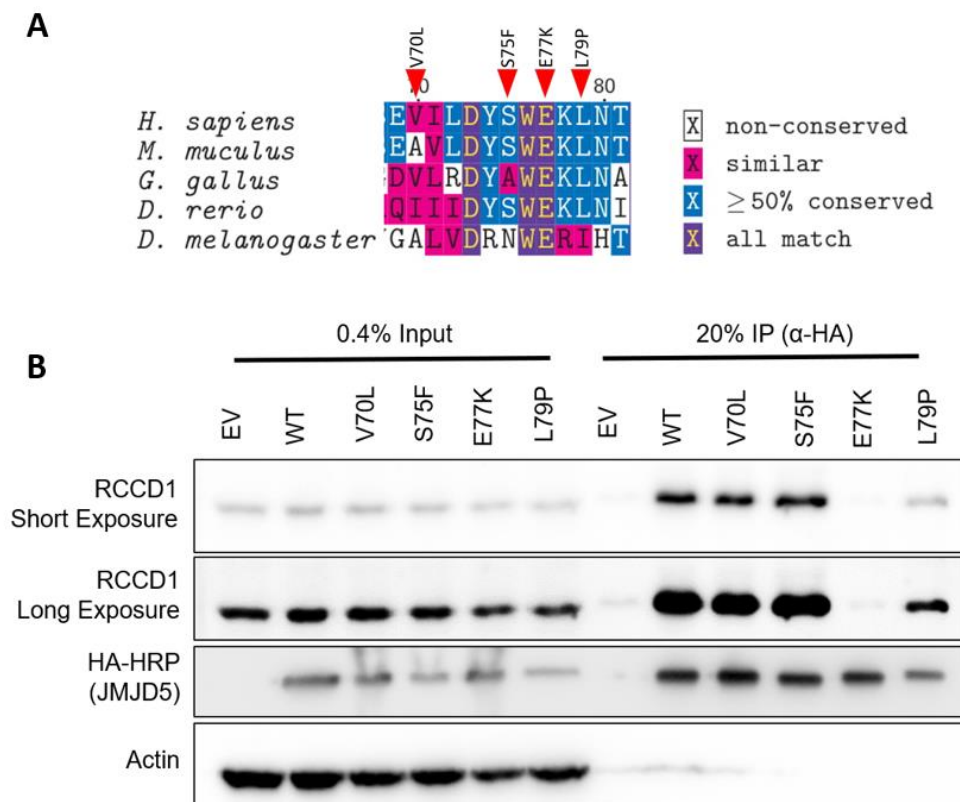
Figure 2.11. Group 2 V70L-L79P JMJD5 cancer variants IPs.
(A) MSA generated using MEGA7 and shaded using the Texshade LaTeX package. Red arrows indicated JMJD5 cancer variants contained in this group. **(B)** Anti-HA immunoprecipitates from HEK293T cell extracts expressing indicated JMJD5 mutations were Western blotted for HA, RCCD1 and Actin. Input is representative of the cell extract before anti-HA immunoprecipitation. Western blots representative of three independent biological repeats.

As above, some mutations showed an intrinsic effect on JMJD5 stability, in the input samples, less JMJD5^{S75F} and JMJD5^{L79P} was observed compared to wildtype JMJD5, suggesting reduced protein expression (Fig 2.11B). JMJD5^{V70L} and JMJD5^{S75F} had no effect on the RCCD1 interaction (Fig 2.11B). Although JMJD5^{L79P} showed reduced expression, it also appeared to show reduced RCCD1 binding (Fig

2.11B). Interestingly, RCCD1 binding to JMJD5^E77K was undetectable above background levels (Fig

2.11B). Importantly, this occurred in the absence of effects on expression, suggesting that E77 is likely

to be an important residue for the JMJD5:RCCD1 interaction.

2.2.7.3   Group 3 variants (D90N to R153S)

Group three consists of five variants (D90N, R92C, R121Q, I144V and R153S), all located at the C-

terminal end of the JMJD5 N-terminus. Although JMJD5^D90N was predicted to be possibly damaging

(Fig 2.9C) the D90 residue is poorly conserved (Fig 2.12A). However, it was selected for its negatively

charged side chain capable of facilitating electrostatic interactions and location in a predicted

structural region (Fig 2.9A). R92 is highly conserved (Fig 2.12A) and its mutation in cancer is recurrent

(Fig 2.9A). Furthermore, substitution of R92 to a cysteine was predicted to be damaging (Fig 2.9C).

R121 is well conserved (Fig 2.12A) in a region of predicted structure (Fig 2.9A) and mutation to

glutamine was predicted to be possibly damaging (Fig 2.9C). Although I144 was only partially

conserved (Fig 2.12A) and mutation to valine was not predicted to be damaging (Fig 2.9C), it is located

within a region of predicted secondary structure (Fig 2.9A). Although R153 is not well conserved (Fig

2.12A) and only 'possibly damaging' when mutated to serine (Fig 2.9C), it was selected for analysis

because it is located within a region not covered by other variants (Fig 2.9A).
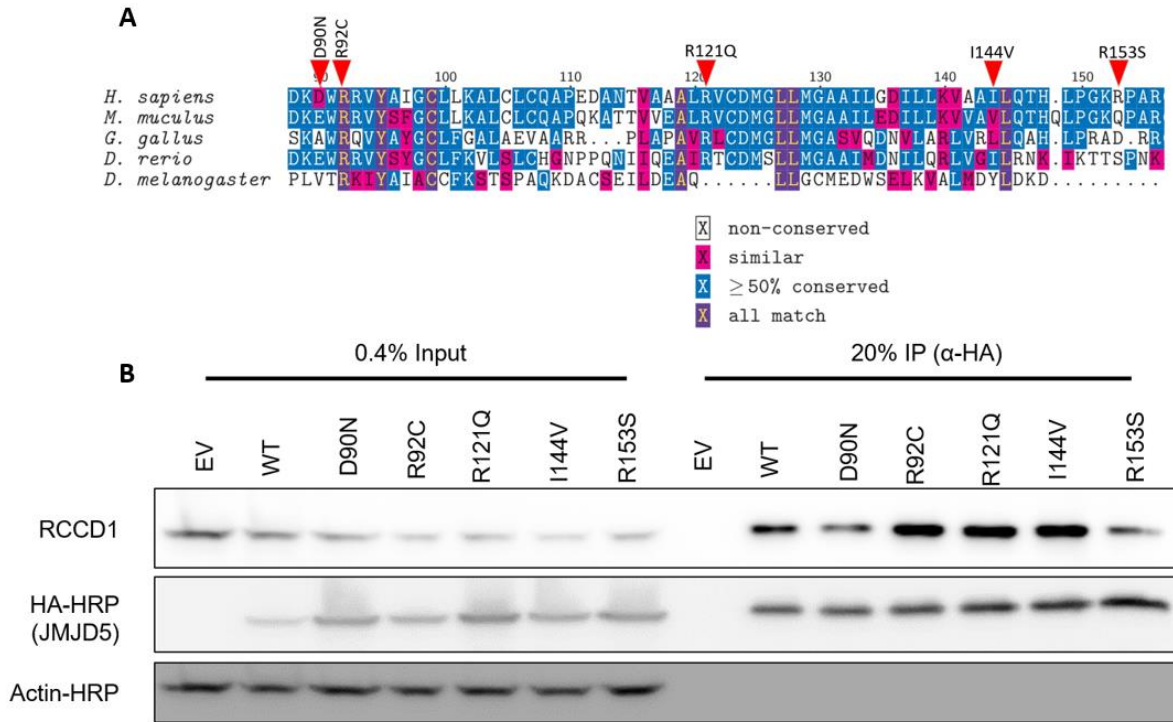
Figure 2.12. Group 3 D90N-R153S JMJD5 cancer variants IPs.
(A) MSA generated using MEGA7 and shaded using the Texshade LaTeX package. Red arrows indicated JMJD5 cancer variants contained in this group. **(B)** Anti-HA immunoprecipitates from HEK293T cell extracts expressing indicated JMJD5 mutations were Western blotted for HA, RCCD1 and Actin. Input is representative of the cell extract before anti-HA immunoprecipitation. Western blots representative of three independent biological repeats.

First by analysing the input samples, none of this group of variants showed a reduction in protein expression, however a caveat of this interpretation is that the anti-HA-blot of the input samples indicates that wildtype JMJD5 may have been less well expressed (Fig 2.12B). Additionally, none of the variants showed a substantial reduction in RCCD1 binding when expressed in HEK293T cells for immunoprecipitation assays (Fig 2.12B). JMJD5^{R92C}, JMJD5^{R121Q} and JMJD5^{I144V} may have shown a modest increase in RCCD1 binding compared to JMJD5^{WT} (Fig 2.12B). However, the lower expression of wildtype JMJD5 raises questions about the validity of this interpretation. Interestingly, both JMJD5^{D90N} and JMJD5^{R153S} appear to show a partial reduction of RCCD1 binding (Fig 2.12B), although the effect is substantially less than that observed for JMJD5^{E77K} (Fig 2.11B).

2.2.7.4   Recombinant E77K JMJD5 shows reduced RCCD1 Binding

Overall, the findings above suggest that JMJD5^{E77K} and JMJD5^{L79P} showed the most striking effects on reduced interaction with RCCD1. To investigate whether these residues are critical for the JMJD5-RCCD1 interaction using an orthogonal approach we aimed to use the GST-pulldown approach presented above. We cloned JMJD5^{E77K} and JMJD5^{L79P} into a bacterial expression vector to generate recombinant protein. Whereas JMJD5^{E77K} was successfully purified, we were unable to purify enough JMJD5^{L79P} for use in our experiments because it was expressed very poorly (data not shown). Therefore, we focussed the GST-pulldown experiment on the JMJD5^{E77K} variant.

As above, 12 µg of recombinant JMJD5$^{WT \ and \ E77K}$ was incubated with HEK293T lysate followed by GST pulldown and Western blot analysis for RCCD1. Consistent with the HEK293T immunoprecipitation experiments (Fig 2.11B), recombinant JMJD5^{E77K} showed a severe reduction in the interaction with RCCD1 (Fig 2.13). Coomassie staining of SDS-PAGE gels showed equal amounts of wildtype and JMJD5^{E77K} protein (Fig 2.13). Because we cannot conclude whether it is the loss of a glutamate residue at position 77 that causes reduced RCCD1 binding, or substitution to the oppositely charged lysine residue (E77K), we sought to interrogate the area around E77 in more detail.
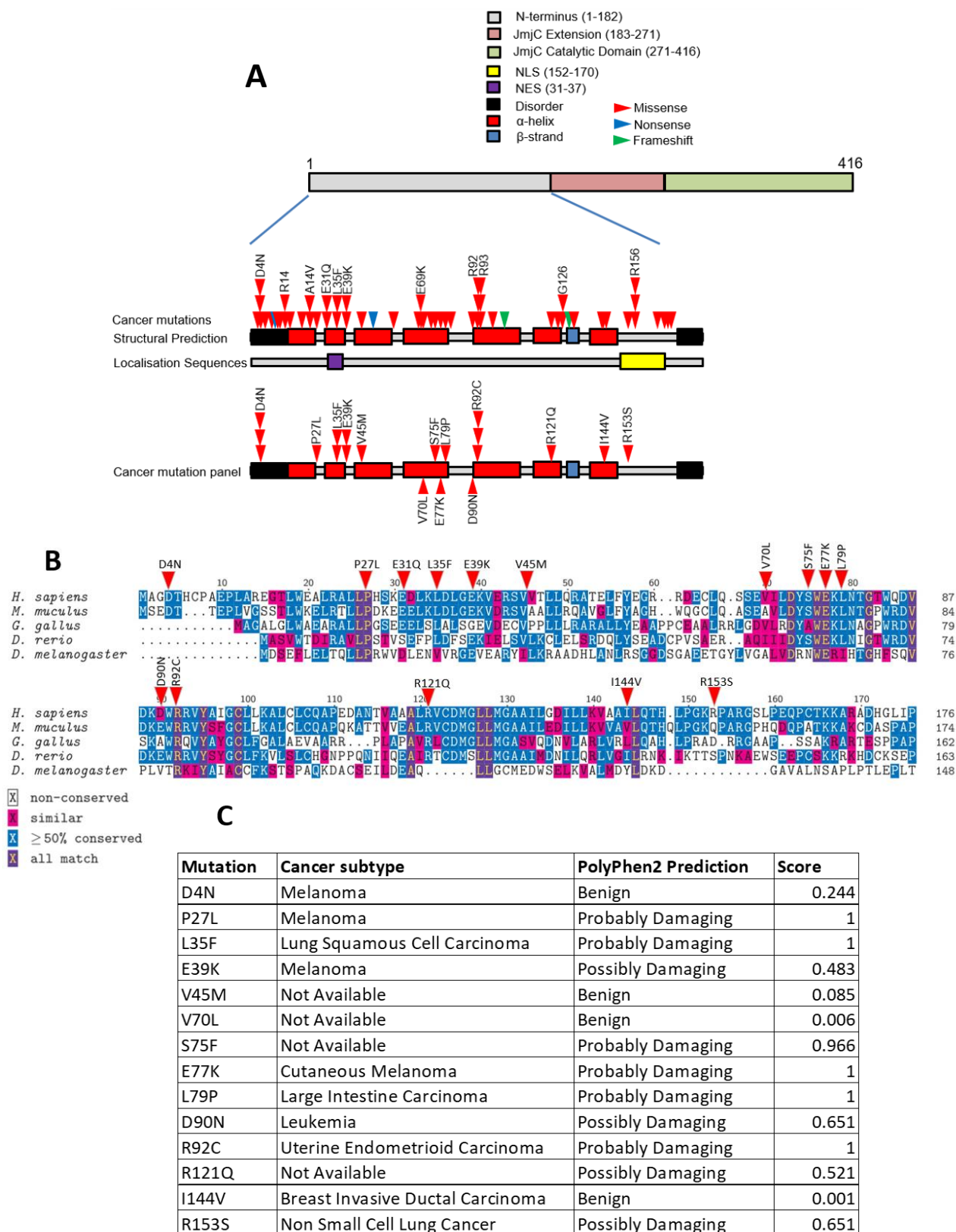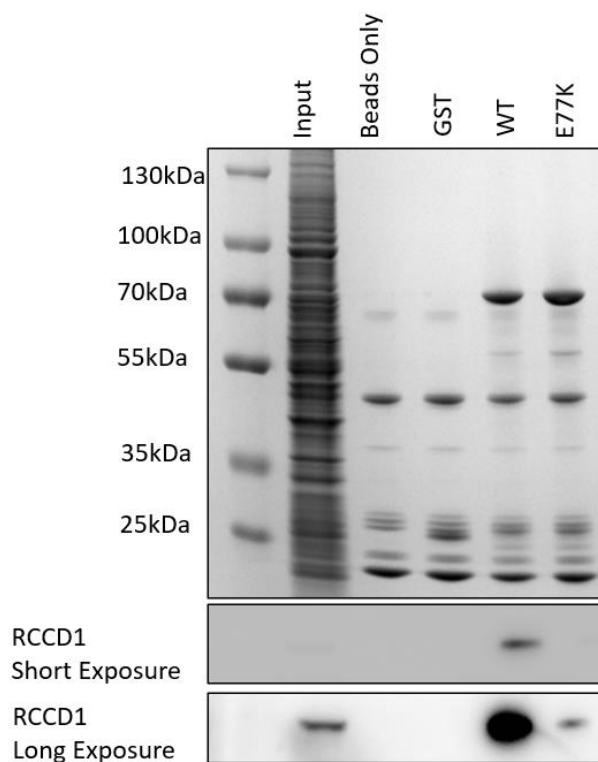
Figure 2.13. Recombinant JMJD5 E77K reduces binding to RCCD1.
GST pulldown experiments using the indicated GST-tagged recombinant protein incubated in HEK293T cell lysates and coomassie stained and Western blotted for RCCD1. Images representative of two independent biological repeats.

2.2.8 D73 and E77 form an RCCD1 binding surface in the N-terminus of JMJD5

When examining the region around JMJD5 E77 we observed a high degree of sequence conservation between residues 70 and 80 (Appendix 2). We therefore wanted to generate experimental mutants of key residues within this region. We chose to use an alanine scanning approach in order to try and help dissect effects of loss of wildtype residue versus missense substitution.

2.2.8.1 D73A and E77A reduce RCCD1 binding

A panel of experimental mutants was created consisting of D73A, W76A, E77A, K78A and L79A. JMJD5^{E77A} and JMJD5^{L79A} were chosen because the corresponding cancer variants showed reduced RCCD1 binding. Conservation was the primary selection criteria for the remaining three, with both JMJD5^{D73} and JMJD5^{W76} being fully conserved and JMJD5^{K78} highly conserved Fig 2.14A). All mutations were cloned into pEF6 as above and transfected into HEK293T cells followed by anti-HA immunoprecipitation and Western blot for RCCD1.

49
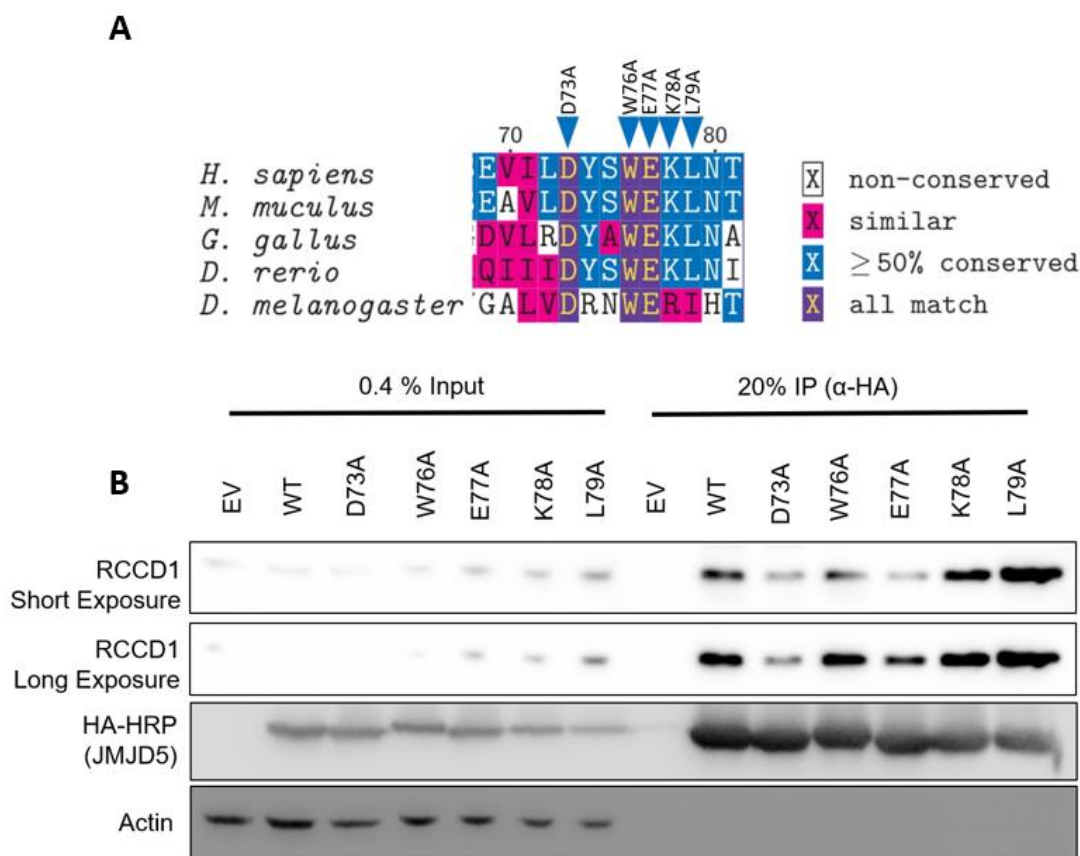
Figure 2.14. JMJD5^D73A and JMJD5^E77A have reduced interaction with RCCD1.
(A) MSA generated using MEGA7 and shaded using the Texshade LaTeX package. Blue arrows indicated JMJD5 experimental mutations assayed. **(B)** Anti-HA immunoprecipitates from HEK293T cell extracts expressing indicated JMJD5 mutations were Western blotted for HA, RCCD1 and Actin. Input is representative of the cell extract before anti-HA immunoprecipitation. Western blots representative of two independent biological repeats.

Neither JMJD5^W76A, JMJD5^K78A, or JMJD5^L79A showed any reduction in RCCD1 binding (Fig 2.14B). Interestingly, JMJD5^E77A did show a reduction in binding (Fig 2.14B), but not to the same extent as that of the JMJD5^E77K cancer variant (Fig 2.12B). Strikingly, JMJD5^D73A also reduced RCCD1 binding (Fig 2.14B).

2.2.8.2   D73A and E77A mutations are cooperative in reducing RCCD1 binding

As D73 and E77 are both negatively charged, we hypothesised that these two amino acids could form a charged binding surface for docking RCCD1. Indeed, D73 and E77 are four residues apart on a predicted α-helix (Fig 2.9A): Because there are approximately 3.6 amino acids per turn of an α-helix,

it is possible that D73 and E77 could be on the same surface of a solvent exposed α-helix available for

RCCD1 binding (structural modelling is presented in section 2.2.11). Therefore, we next sought to

investigate D73 and E77 in more detail using a double alanine mutant. We cloned the double mutation

using SDM of our HA-tagged mammalian vector (pEF6) and tested as above, with comparison to the

existing single alanine variants (D73A and E77A).
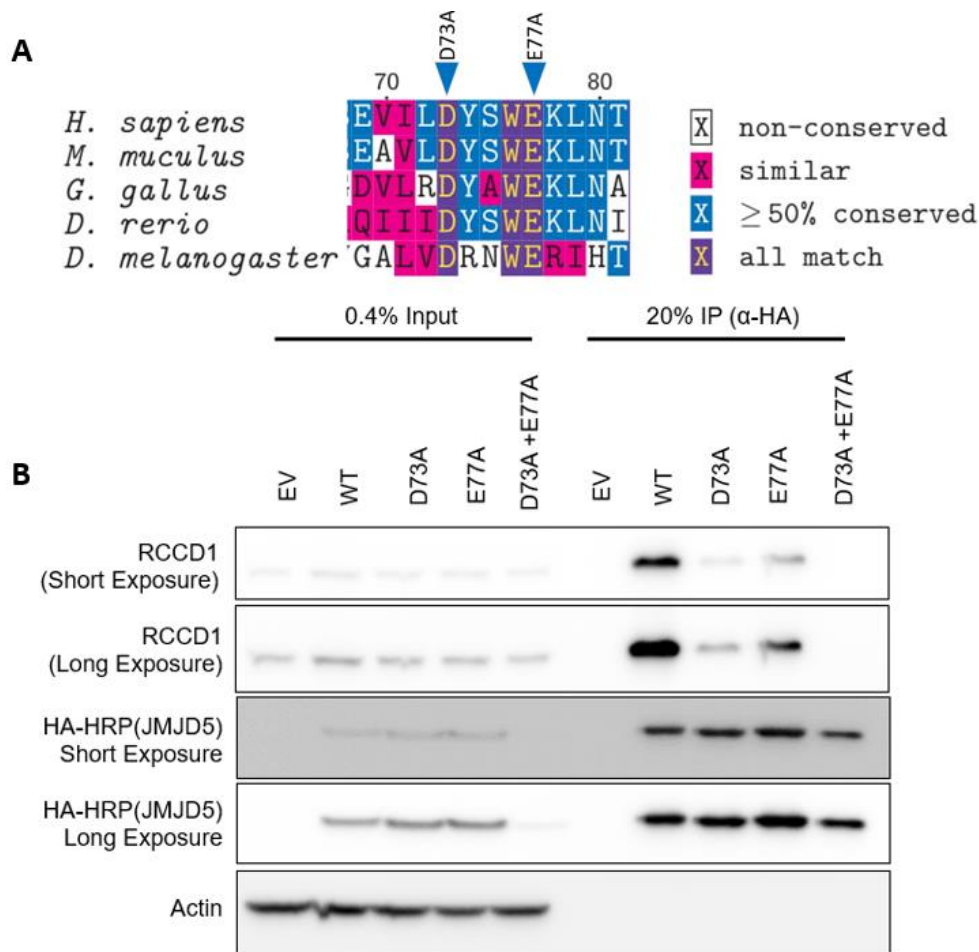
Figure 2.15. JMJD5$^{D73A+E77A}$ drastically reduces RCCD1 binding.
(A) MSA generated using MEGA7 and shaded using the Texshade LaTeX package. Blue arrows
indicated JMJD5 experimental mutations assayed. **(B)** Anti-HA immunoprecipitates from HEK293T
cell extracts expressing indicated JMJD5 mutations were Western blotted for HA, RCCD1 and Actin.
Input is representative of the cell extract before anti-HA immunoprecipitation. Western blots
representative of three independent biological repeats.

JMJD5^{D73A} and JMJD5^{E77A} showed the same binding profile as the previous experiment, with the single alanine mutations causing a reduction in RCCD1 binding (Fig 2.15B). Importantly, the double alanine mutation showed an additive effect, with RCCD1 binding being completely undetectable. However, it is important to highlight that this was in the context of reduced expression of the JMJD5 double mutant, which might have complicated the interpretation (Fig 2.15B). Therefore, we sought to use the GST-pulldown assay to further study the interaction.

2.2.8.3   Recombinant JMJD5$^{D73A/E77A}$ is unable to bind RCCD1

We cloned the single (D73A and E77A) and double alanine mutations into our bacterial expression vector and successfully generated recombinant protein (Fig 2.16, top panel). As above, 12 µg of recombinant protein was incubated with HEK293T cell lysate followed by GST pulldown and Western blotting   for   RCCD1.
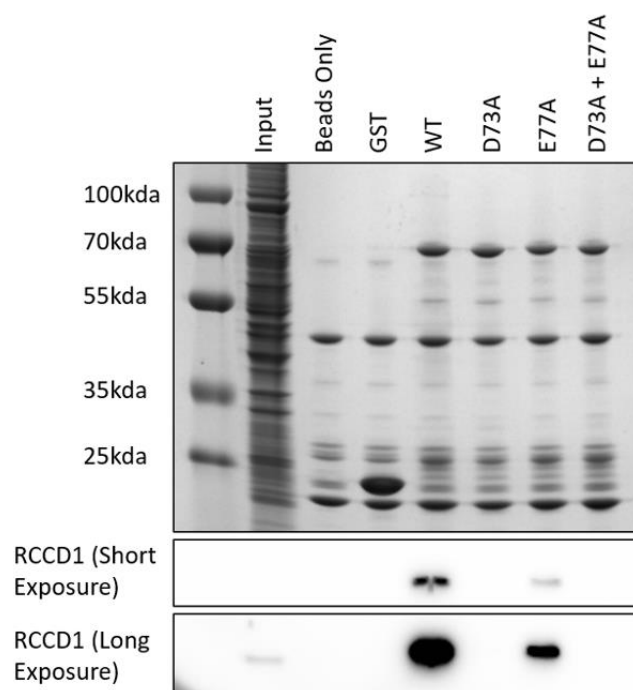
Figure 2.16. Recombinant JMJD5 experimental mutations reduce binding to RCCD1.
GST pulldown experiments using the indicated GST-tagged recombinant protein incubated in HEK293T cell lysates and coomassie stained and Western blotted for RCCD1. Images representative of two independent biological repeats.

Interestingly, in this assay the single JMJD5^{D73A} mutation appeared to be sufficient to completely ablate RCCD1 binding, whilst JMJD5^{E77A} also significantly reduced the interaction (Fig 2.16, bottom panels). The JMJD5 double mutant also showed a complete loss of RCCD1 binding. Importantly, and in contrast to the cell-based experiment presented in Fig 2.15, this was in the context of equal amounts

of JMJD5. Overall, the data suggests that D73 and E77 are both key to mediating the interaction of JMJD5 with RCCD1. However, it remains to be seen what the functional consequences of reduced JMJD5:RCCD1 interaction are.

2.2.9 JMJD5 binding regulates RCCD1 localisation

2.2.9.1 RCCD1 localises to the nucleus with wildtype JMJD5

Our findings suggest that overexpressed JMJD5^{E77K} and JMJD5$^{D73A/E77A}$ are unable to bind RCCD1. To begin to investigate the potential consequences of reduced JMJD5:RCCD1 interaction we first investigated their subcellular localisation. Whilst RCCD1 localisation has not been reported in the literature, JMJD5 has been shown to be predominantly nuclear but can shuttle between the nucleus and the cytoplasm using localisation sequences (Huang et al., 2013). We aimed to use immunofluorescence of fixed cells to investigate the localisation of exogenous RCCD1 in the presence or absence of wildtype or mutant JMJD5. Because HEK293T cells do not efficiently adhere to glass coverslips, we intended to use the lung cancer cell line A549, which are more amenable to immunofluorescence analyses. Initial attempts using transient transfection of RCCD1 expression constructs were disappointing because of variable transfection efficiency. Therefore, we decided to generate stable cell lines expressing RCCD1. To this end, full length RCCD1 was PCR cloned with a C-terminal 3XFLAG-tag into the lentiviral expression vector 'pIHZ', a derivative of 'pGIPZ' that lacks the GFP cDNA and substitutes puromycin (P) for hygromycin (H) resistance. Stable A549 lines were generated by lentiviral infection and hygromycin selection followed by validation of FLAG-RCCD1 expression by Western blot and immunofluorescence (data not shown). We then transiently transfected these cells with pEF6 HA-tagged wildtype (WT) or RCCD1 binding mutant (E77K and D73A/E77A) JMJD5 expression constructs. After 48 h of transfection the cells were fixed using methanol and stained for the two epitope tags, HA and FLAG.

Consistent with the literature (Huang et al., 2013), JMJD5^{WT} staining was predominantly nuclear in this model (Fig 2.17, HA panel). The two RCCD1 binding mutants (JMJD5^{E77K} and JMJD5$^{D73A/E77A}$) showed

similar localisation to JMJD5^{WT}, consistent with these mutations, and RCCD1 binding, not having a major impact of JMJD5 subcellular localisation. Interestingly however, anti-FLAG staining in the same cells suggested that JMJD5, and its interaction with RCCD1, regulate RCCD1 localisation. FLAG-RCCD1 was predominantly cytoplasmic in cells transfected with a control vector (pEF6-EV), whereas JMJD5^{WT} overexpression was associated with a shift to nuclear FLAG-RCCD1 localisation (Fig 2.17, FLAG panel – HA positive cells indicated by red boxes). Consistent with this being a consequence of their interaction, RCCD1 remained cytoplasmic in cells expressing JMJD5^{E77K} or and JMJD5$^{D73A/E77A}$.
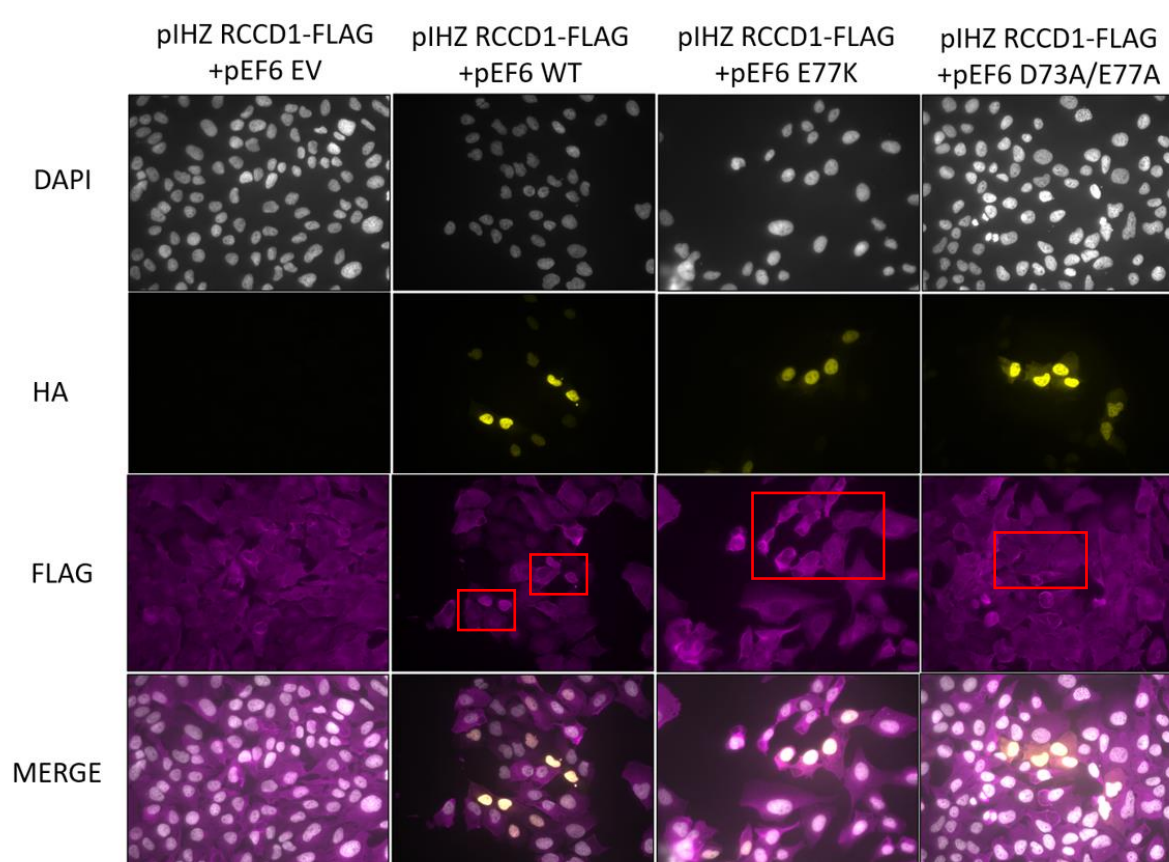
Figure 2.17. RCCD1 does not co-localise with N-terminal JMJD5 RCCD1 binding mutants.
Localisation of 3xFLAG-RCCD1 in A549. Stable A549 cells were transfected with indicated pEF6 HA-JMJD5 constructs and fixed with MeOH and stained for FLAG (magenta) and HA (yellow). Nuclei were visualised with DAPI. Images representative of three independent biological repeats.

2.2.9.2   JMJD5 N-terminal RCCD1 binding mutants do not give a proximity ligation assay signal

Although co-localisation by immunofluorescence can be consistent with proteins interacting in cells, other imaging methods, such as Proximity Ligation Assay (PLA), measures this more directly. PLA uses

secondary antibodies labelled with oligonucleotides (PLA probes) that can bind to primary antibodies (HA and FLAG in this experiment) (Fig 2.18A). If the two probes are within the same physical space (<40 nm), the two oligonucleotides can be ligated (Fig 2.18A). DNA polymerase then enhances the signal with rolling-circle amplification incorporating fluorescently labelled oligonucleotides that form discreet foci visualised by fluorescent microscopy (Fig 2.18A). We applied PLA to the same A549 cell model as above, combining stable FLAG-RCCD1 expression with transient pEF6 HA-JMJD5 transfection. After 48 h post-transfection, the cells were fixed using methanol followed by the PLA protocol.

As expected, FLAG-RCCD1 cells not transfected with HA-JMJD5 did not show any PLA foci. In contrast, we observed a large number of PLA foci in the nuclei of FLAG-RCCD1 cells (Fig 2.18B), consistent with the co-staining results above. No PLA foci were observed for the RCCD1 binding mutants, consistent with our interaction data and the co-localisation data. However, the same low transfection efficiency as above was observed, gauged by the low number of cells positive for PLA foci.
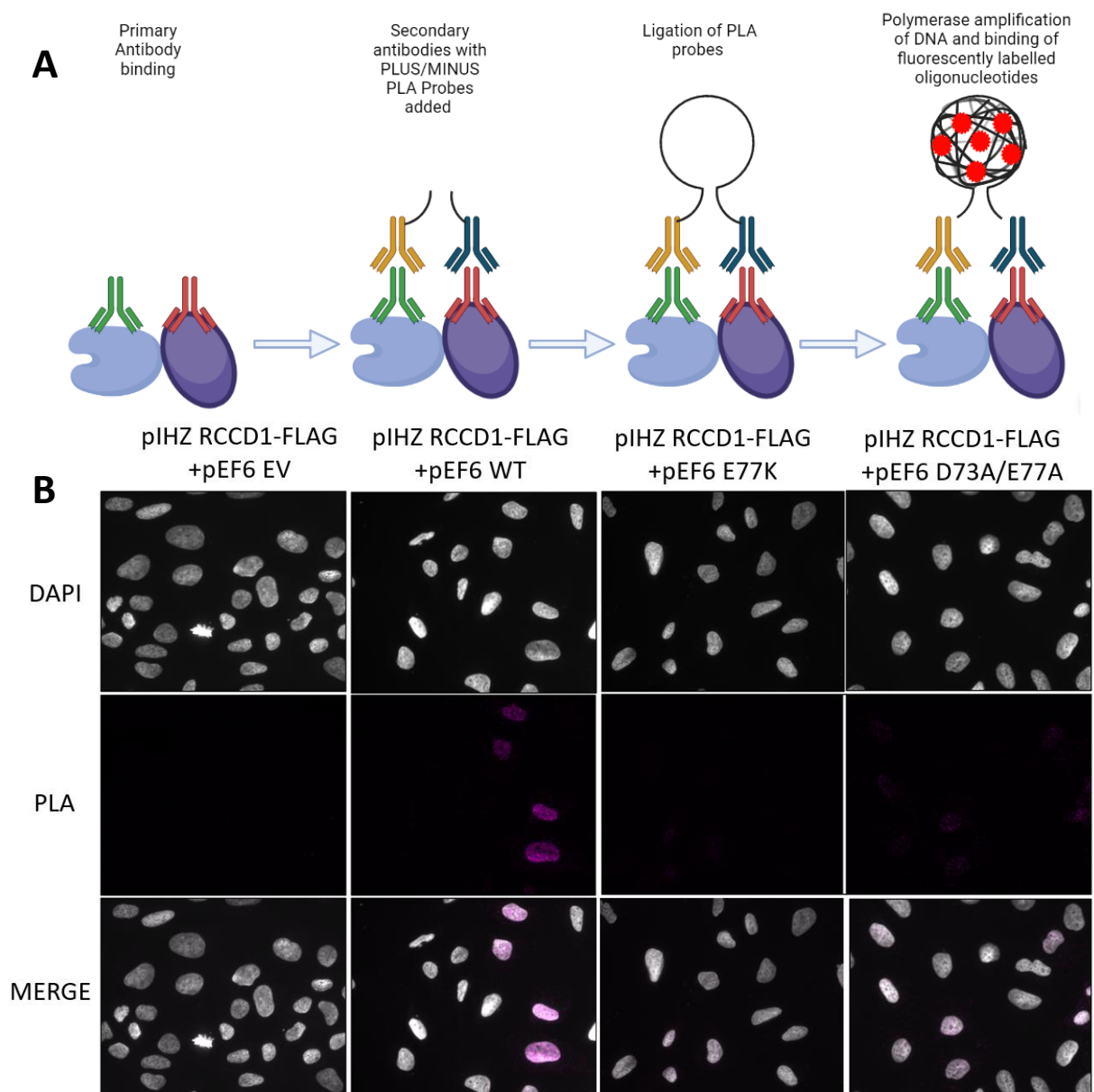
Figure 2.18. RCCD1 and JMJD5 co-localise to the nucleus.
Localisation of 3xFLAG-RCCD1 and HA-JMJD5 in A549. Stable A549 cells were transfected with indicated pEF6 HA-JMJD5 constructs and fixed with MeOH and stained for PLA Foci (Magenta). PLA foci were visualised using anti-HA and anti-FLAG primary antibodies. Images representative of three independent biological repeats.

2.2.10 RCCD1 binding mutations do not affect JMJD5 catalytic activity *in vitro*

Having confirmed that JMJD5 D73 and E77 mutants ablate RCCD1 binding, we were intrigued by the possibility that these could be used as 'separation of function' mutants to investigate the importance of JMJD5:RCCD1 complex formation in cells. However, because of the known importance of JMJD5 hydroxylase activity, including, for example, in replication fidelity (Chapter 1), we first wanted to

determine whether enzyme activity was maintained in the JMJD5 D73 and E77 mutants. Therefore, we used the recombinant enzymes generated above in an *in vitro* hydroxylation assay. This assay uses a synthetic peptide containing a portion of ribosomal protein RPS6 which was previously identified as an experimental JMJD5 substrate (Wilkins et al., 2018). 20 µg of recombinant JMJD5 was incubated with the required co-factors plus RPS6 peptide at 37°C for 1 hour. The Promega Succinate-Glo™ assay was then used to detect succinate production, which is a by-product of 2OG oxygenase activity (Section 1.1.3). This activity can be 'coupled', i.e. resulting from a productive hydroxylation event, or 'uncoupled', due to 2OG conversion to succinate in the absence of substrate hydroxylation. To focus on coupled activity, we converted raw assay values to 'fold change' in hydroxylase activity +/- RPS6 peptide.

In these assays wildtype JMJD5 was included as a positive control, whereas the catalytically dead H321A Fe(II)-binding mutant was used as a negative control. As expected, JMJD5 H321A had a fold change in activity in the presence of RPS6 peptide of one, indicative of an absence of hydroxylation above uncoupled turnover (Fig 2.19). Importantly, none of the RCCD1 binding mutant proteins had a statistically significant effect on JMJD5 catalytic activity *in vitro* (Fig 2.19). These data indicate that the RCCD1 binding mutations are unlikely to have a strong intrinsic effect on JMJD5 hydroxylase activity and may therefore be viable separation of function mutants.
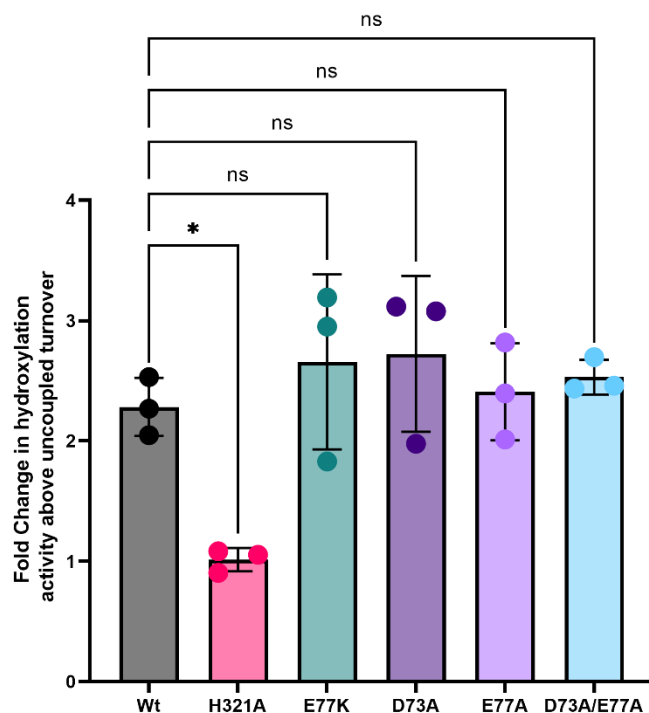
Figure 2.19. JMJD5 mutations that reduce RCCD1 binding in cells do not affect JMJD5 activity *in vitro*.
The indicated recombinant JMJD5 proteins were incubated with Fe(II), 2OG, ascorbate and RPS6 peptide for 1 hr at 37°C. Activity was then monitored using the succinate-Glo assay, which measures succinate production as a by-product. Statistical analysis used One way ANOVA with Bonferroni's post hoc test (with p-values of ≤0.05 (*) and ≥0.05 (ns)). N=3

Overall, our work identifies important JMJD5 residues that facilitate binding to RCCD1. When mutated, these residues showed a clear reduction in RCCD1 binding without effecting JMJD5 activity or localisation. However, the structural basis for the JMJD5:RCCD1 interaction remains unclear.

2.2.11 Structural modelling of the JMJD5:RCCD1 complex using AlphaFold2

During the course of the biochemical studies above, the AlphaFold2 server was published (Jumper et al., 2021), which provided us with an opportunity to model the structure of the JMJD5 N-terminus and RCCD1, for the first time. Here, we describe the Alphafold 2 model of JMJD5 (the RCCD1 model is described in section 2.2.13).

2.2.11.1 D73 and E77 are surface accessible on an N-terminal α-helix

The AlphaFold2 model supported our previous predictions (Fig 2.4) that the JMJD5 N-terminus is attached to the catalytic domain through a long flexible linker (Fig 2.20A). The model suggests that the N-terminus consists of seven α-helices that lie alongside the catalytic domain. By comparison, our initial prediction also included one β-sheet (Fig 2.4), which was not included in the AlphaFold2 model

(Fig 2.20B). Of particular interest, we have hypothesised that D73 and E77 were likely to be on the

same side of an α-helix, which was corroborated by AlphaFold (Fig 2.20C).
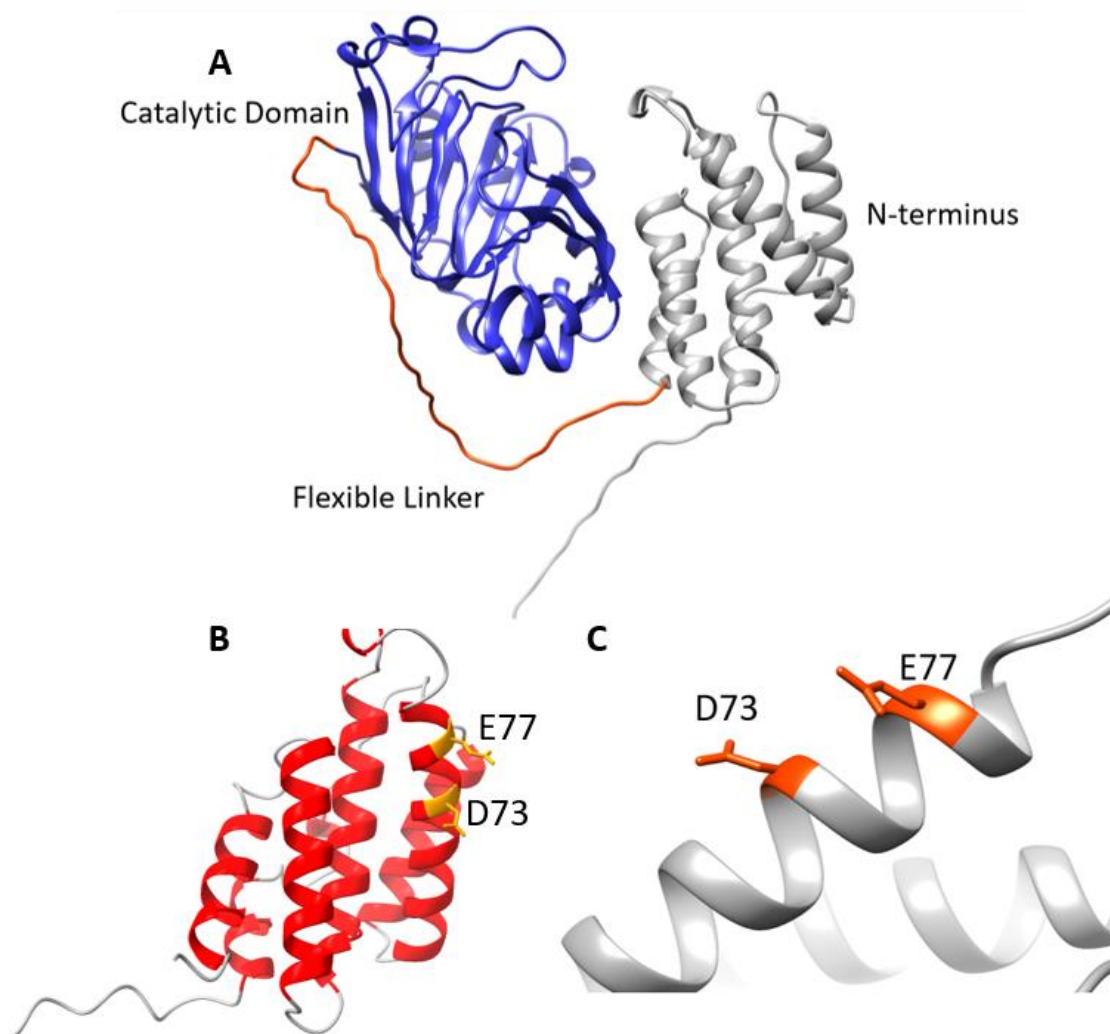
Figure 2.20. D73 and E77 are on a surface accessible α-helix according to AlphaFold2 prediction.
(A) The JMJD5 N-terminus (grey) is connected to the catalytic domain (blue) via a flexible linker (orange). **(B)** The N-terminus is comprised of seven α-helices (coloured in red)**. (C)** D73 and E77 are surface accessible on the same side of an α-helix. All images using predicted AlphaFold structure and chimera software. D73 and E77 are highlighted in orange.

2.2.12 RCCD1 interacts with JMJD5 via RCCD1 residues R310 and K328

Following the discovery of the key RCCD1 binding residues within the N-terminus of JMJD5 we wanted to attempt to identify the reciprocal binding region on RCCD1. As with JMJD5, we wanted to begin by generating a panel of truncation mutants for domain mapping that could enable us to identify functionally discrete domains. To structurally rationalise these truncations, we investigated our AlphaFold2 model. Consistent with the predicted homology to RCC1 and RCC2 (Renault et al., 1998; Mollinari et al., 2003), AlphaFold2 predicted RCCD1 to form a seven-bladed β-propellor structure (Fig 2.21A+B). In contrast to the RCC1/2 structures, AlphaFold2 predicted RCCD1 to include a unique large α-helical loop that projects from the base of the main ring and many longer loops between the RCC1 domains (Fig 2.21A).

However, the predicted β-propellor structure makes the use of domain mapping approaches inadvisable, because of the ring shape each RCC1-like domain could be essential in maintaining it. Therefore, we hypothesised that removal of any of the domains could cause a complete destabilisation of the protein (Professor Steve Smerdon – Personal Communications). Therefore, we started with a targeted site-directed mutagenesis approach.
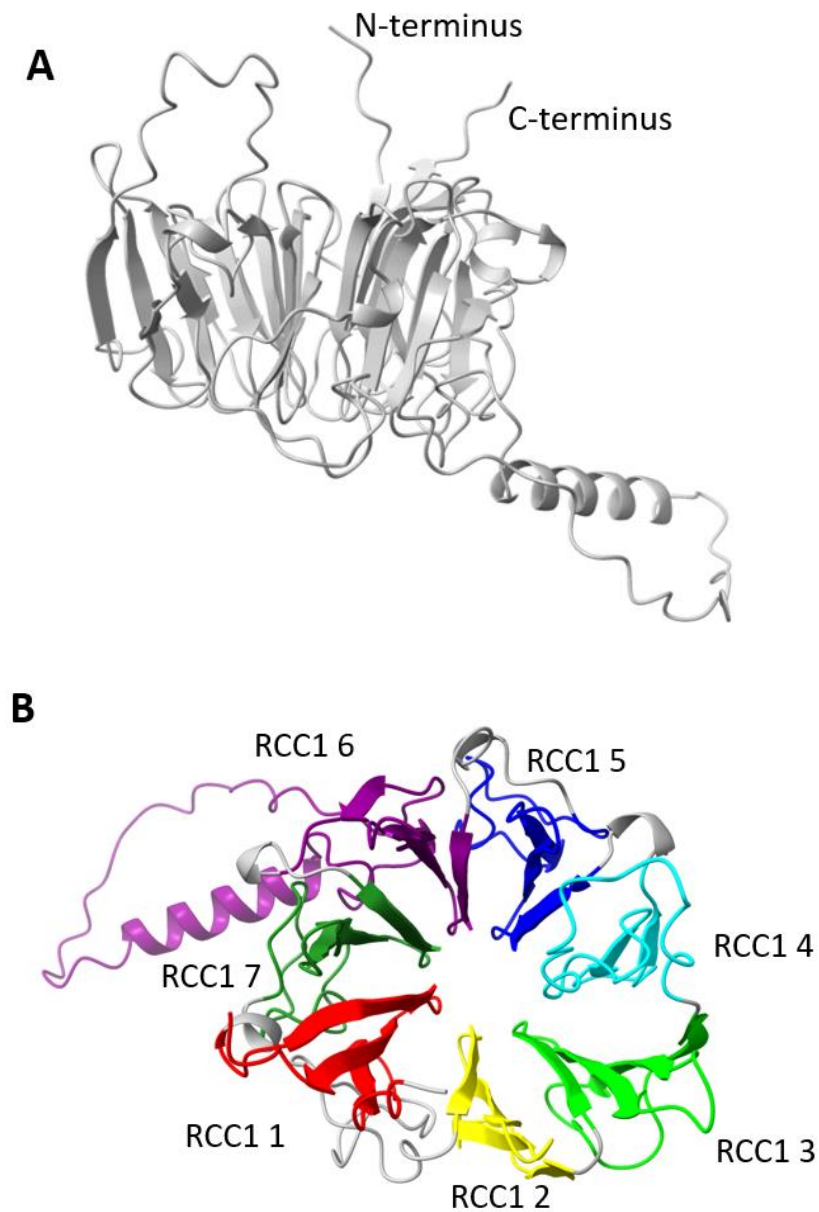
Figure 2.21. Predicted RCCD1 structure by AlphaFold2 shows strong homology to RCC1.
(A) RCCD1 is predicted to form a seven bladed β- propellor structure by AlphaFold. RCCD1 prediction coloured in grey). **(B)** RCCD1 RCC1 like domains (RLDs) coloured from one to seven.

2.2.12.1 Generation of an RCCD1 panel of cancer mutations

In section 2.2.6, we hypothesised that the JMJD5:RCCD1 interaction might be important in cancer, and that if this were the case, that JMJD5 cancer variants would disrupt RCCD1 binding. By extension, we predict here that RCCD1 cancer variants will disrupt JMJD5 binding, and that identifying such variants could help identify the critical interface. To explore this, we used COSMIC and cBIOportal to generate a list of reported variants in RCCD1 and plotted them against the RCCD1 domain structure (Fig 2.22A

and Appendix 3). Interestingly, we observed a significant number of recurrent mutations and clustering in key hotspots within specific RCC1 domains (Fig 2.22A). Considering the large number of mutations to consider, we needed a strategy to shortlist a small number of candidates to test in JMJD5 binding assays. Because D73 and E77 are negatively charged residues we hypothesised that positively charged RCCD1 amino acids (arginine and lysine) would be interacting residues. Investigating our full list of cancer mutations revealed 10 initial variants in positively charged residues, we began cloning these mutations but were unsuccessful in generating three (R202Q, R272I and R345C). Therefore, we had generated a panel of seven variants that spanned the length of the protein (Fig 2.22A). Subsequent updates to the list of cancer variants revealed an additional five variants in positively charged amino acids, that due to timing could not be considered during this investigation (Appendix 3).

Once successfully generated we used similar bioinformatic analysis to JMJD5 to ascertain if any were either; (i) Predicted damaging by PolyPhen2 score, and (ii) sequence conserved (RCCD1 MSA - Appendix 4). It should be noted that if any variants selected had low predicted damage or sequence conservation, it did not invalidate their selection as they were all in positively charged residues. The rationale for each variant selected is outlined in more detail below.
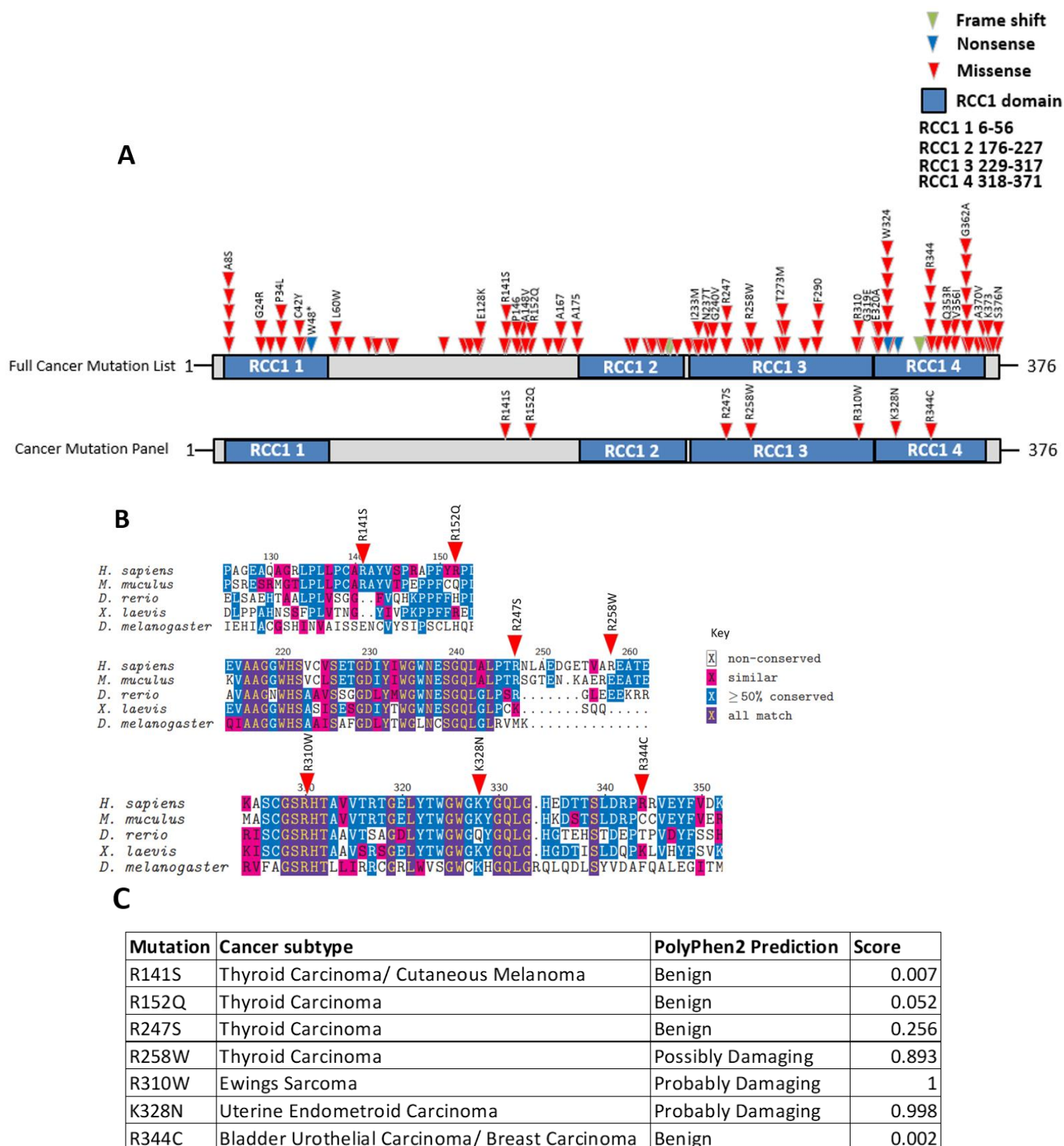
Figure 2.22. Design of RCCD1 cancer variant panel.
(A) The panel of cancer variants selected for further study plotted against the RCCD1 domain structure. **(B)** MSA generated using MEGA7 and shaded using the Texshade LaTeX package. The red arrows indicate the panel of cancer mutations selected. **(C)** Summary table showing PolyPhen2 predictions for the panel of RCCD1 cancer variants.

Mutation	Cancer subtype	PolyPhen2 Prediction	Score
R141S	Thyroid Carcinoma/ Cutaneous Melanoma	Benign	0.007
R152Q	Thyroid Carcinoma	Benign	0.052
R247S	Thyroid Carcinoma	Benign	0.256
R258W	Thyroid Carcinoma	Possibly Damaging	0.893
R310W	Ewings Sarcoma	Probably Damaging	1
K328N	Uterine Endometroid Carcinoma	Probably Damaging	0.998
R344C	Bladder Urothelial Carcinoma/ Breast Carcinoma	Benign	0.002

2.2.12.2 Group 1 variants (R141S to R258W)

The first group consists of four variants (R141S, R152Q, R247S and R258W), none of which are highly conserved (Fig 2.23A). R247 is conserved in three out of the five species and functionally conserved in the fourth. Interestingly, all the variants in this group were reported as recurrent. RCCD1[R141S] and RCCD1[R247K/S] mutations were detected in three tumour samples each, whereas RCCD1[R152Q] and RCCCD1[R258W] were both reported in two tumours (Fig 2.22A and Appendix 3). Three of the variants in this group were also predicted to be benign by PolyPhen2 with R141S, R152Q and R247S scoring 0.007, 0.052 and 0.256 respectively. RCCD1[R258W] was predicted to be possibly damaging with a score of 0.893 (Fig 2.22C). The variants were overexpressed in HEK293T cells, followed by anti-FLAG immunoprecipitation and Western blotting for JMJD5, FLAG and Actin.
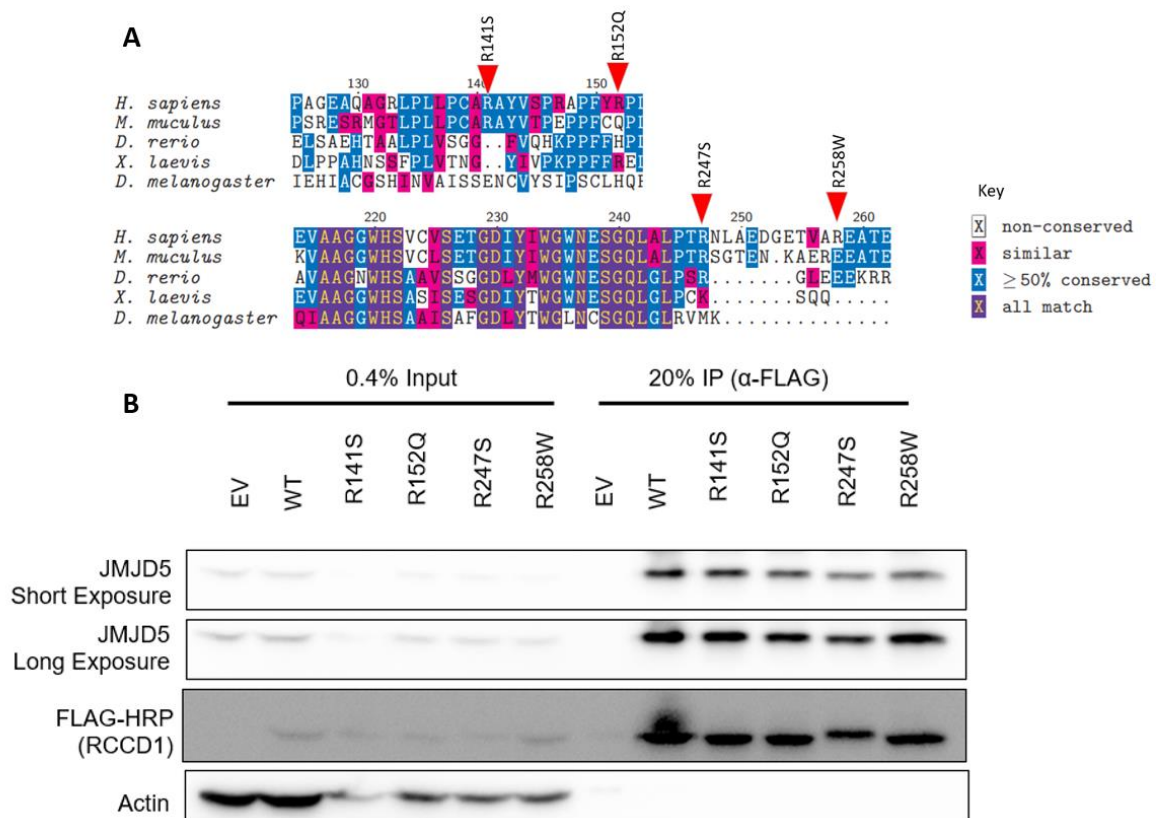
Figure 2.23. Group 1 Variants R141S – R258W.
(A) Multiple sequence alignment (MSA) generated using MEGA7 and shaded using the Texshade LaTeX package. Red arrows indicated RCCD1 cancer variants assayed. **(B)** Anti-FLAG immunoprecipitates from HEK293T cell extracts expressing indicated RCCD1 variants were Western blotted for FLAG, JMJD5 and Actin. Input is representative of the cell extract before anti-HA immunoprecipitation. Western blots representative of three independent biological repeats.

Firstly, by analysing the input samples three of the variants (R141S, R152Q and R247S) appear to reduce RCCD1 expression compared to wildtype (Fig 2.23B). However, the resolution of the input blot is very low caveating this observation. Of these mutants, none-appeared to cause a significant reduction in JMJD5 binding, with levels of JMJD5 in the IP samples similar to wildtype (Fig 2.23B).

2.2.12.3 Group 2 variants (R310W to R344C)

The second group contained three mutations (R310W, K328N and R344C). R310 is fully conserved (Fig 2.24A), is a recurrently mutated residue (one mutation to W and one to Q) and is predicted to be highly damaging (Fig 2.22C). K328 is also highly conserved and predicted to be highly damaging (Fig 2.24A and Fig 2.22C). R344 is poorly conserved and not predicted to be damaging but is mutated in six independent tumours (four R344C and two R344H) (Fig 2.22A+C and Appendix 3). This group were overexpressed in HEK293T cells and tested as above.
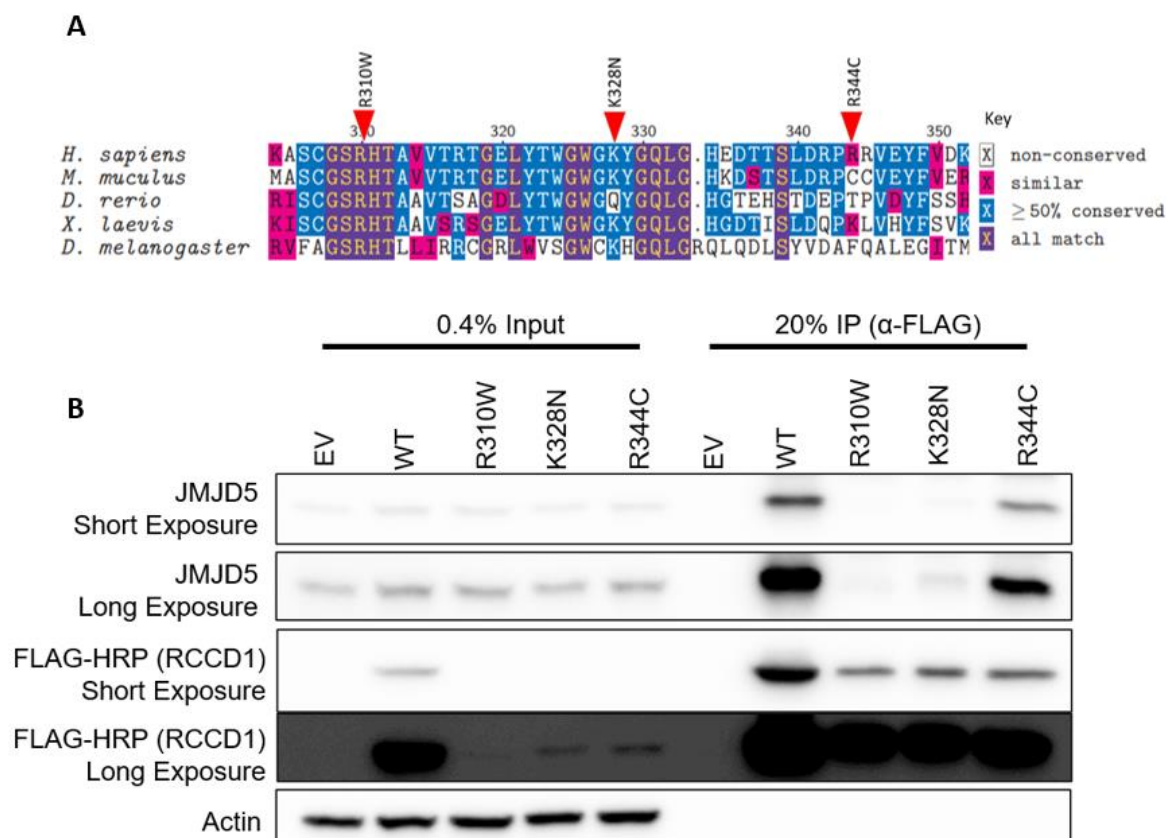
Figure 2.24. Group 2 variants R310W to R344C.
(A) Multiple sequence alignment (MSA) generated using MEGA7 and shaded using the Texshade LaTeX package. Red arrows indicated RCCD1 cancer variants assayed. **(B)** Anti-FLAG immunoprecipitates from HEK293T cell extracts expressing indicated RCCD1 variants were Western blotted for FLAG, JMJD5 and Actin. Input is representative of the cell extract before anti-HA immunoprecipitation. Western blots representative of three independent biological repeats.

As above, all three variants in this group drastically reduced protein expression in the input samples compared to wildtype RCCD1 (Fig 2.24B). Interestingly, RCCD1^{R310W} and RCCD1^{K328N} appeared to drastically reduce JMJD5 binding (Fig 2.24B). It should be noted that this was in the context of all three mutations reducing RCCD1 protein expression however, which slightly confounds accurate interpretation. However, given that all three mutants show similarly reduced expression, but only RCCD1^{R310W} and RCCD1^{K328N} show almost complete loss of JMJD5 binding, it is possible to conclude that they likely reduce JMJD5 binding. Given the confounding effects of the cancer missense mutations on expression we sought to verify our observation through an alternate approach.

2.2.12.4 RCCD1 R310M/K328M variants reduce JMJD5 binding

Following the above observations, we first investigated the RCCD1 AlphaFold2 model to determine whether R310 and K328 are surface accessible and available for JMJD5 binding. Indeed, R310 and K328 are surface accessible on one plane of the 'doughnut' created by the RCCD1 β-propellors (Fig 2.25A). Interestingly, R310 and K328 are located within a ring of tryptophan residues (Fig 2.25A). Because of the hydrophobic nature of tryptophan residues, they prefer not to be solvent exposed, suggesting that they might be 'protected' by an interface with JMJD5.
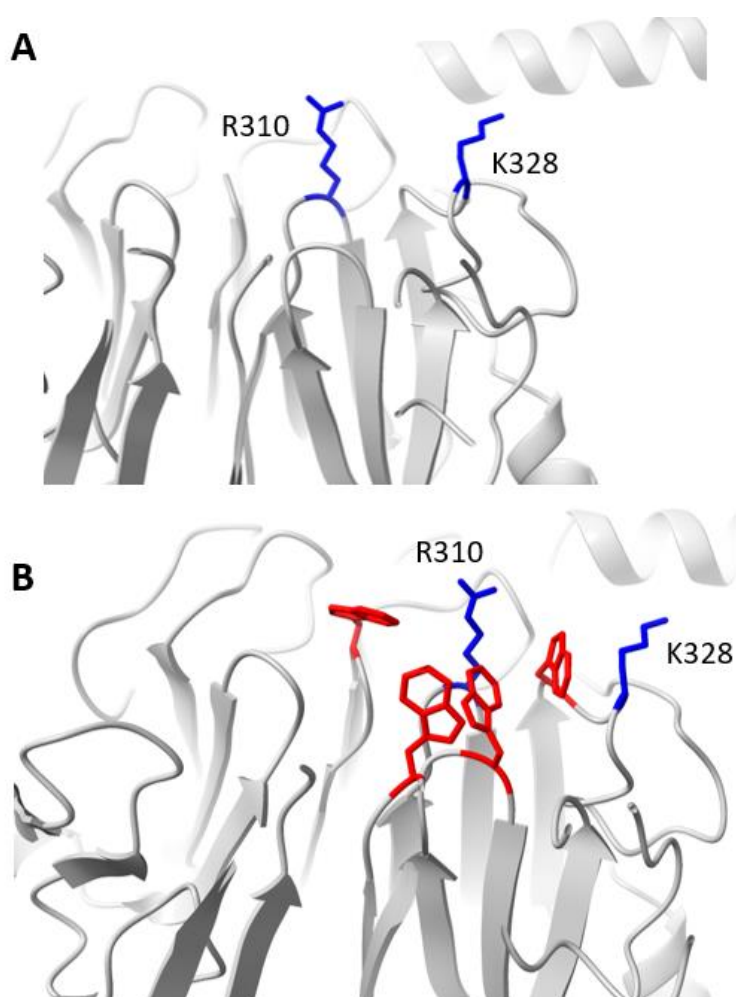
Figure 2.25. R310 and K328 are surface accessible in a hydrophobic region of RCCD1.
(A) R310 and K328 (Blue) are surface accessible on the same side of the ring-shaped structure of RCCD1. **(B)** R310 and K328 (blue) are located around a ring of hydrophobic tryptophan residues (red). All images using predicted AlphaFold structure and chimera software.

Since both RCCD1^{R310W} and RCCD1^{K328N} had drastically reduced protein expression we sought to generate experimental mutations that would minimise this impact to support more accurate interpretation. In addition to the potential electrostatic interaction with JMJD5, the aliphatic portion of the Arginine sidechain of R310 makes hydrophobic interactions with the outer core of RCCD1 in our AlphaFold2 model (Fig 2.25A). Therefore, rather than substitution with alanine, which would not preserve this interaction, we chose methionine, which is predicted to prevent the electrostatic interaction with JMJD5 whilst supporting the hydrophobic interaction with the outer core. To this end, we created an RCCD1$^{R310M/K328M}$ double mutant. This experimental mutant was generated by cloning into the pIHZ vector containing a 3XFLAG-tag, as above. Once successfully generated, the binding of JMJD5 was assayed using the same immunoprecipation approach.

Consistent with our hypothesis, the double methionine substitutions 'rescued' the severe impact of the R310 and K328 cancer variants on expression (Fig 2.26). Furthermore, the double methionine mutant was completely incapable of binding JMJD5 (Fig 2.26). These results suggest that RCCD1^{R310} and RCCD1^{K328} are the positively charged residues that interact with JMJD5^{D73} and JMJD5^{E77} on JMJD5, forming an electrostatic interaction between the two proteins.
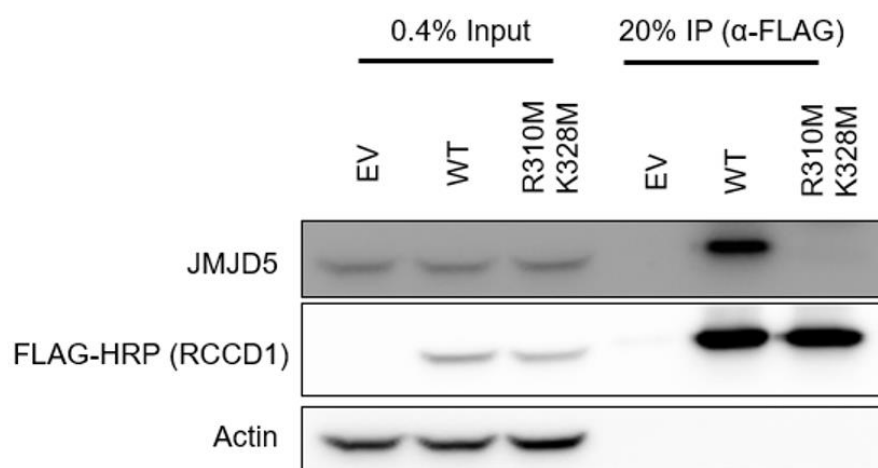
Figure 2.26. R310M/K328M double mutation reduces JMJD5 binding.
Anti-FLAG immunoprecipitates from HEK293T cell extracts expressing indicated JMJD5 mutations were Western blotted for FLAG, JMJD5 and Actin. Input is representative of the cell extract before anti-HA immunoprecipitation. Western blots representative of three independent biological repeats.

2.2.13 AlphaFold multimer predicts interaction surface between JMJD5 D73 and E77 with RCCD1 R310 and K328

After carrying out the initial individual analyses of JMJD5 and RCCD1, the AlphaFold Multimer server was published, which allows for prediction of protein complexes (Evans et al., 2022). We used the multimer server to predict the structure of the RCCD1:JMJD5 complex in order to visualise the potential binding interface (Fig 2.27A). Interestingly, and consistent with all of our binding assays, the model places the N-terminus of JMJD5 in an interaction with the ring shape of RCCD1 (Fig 2.27A). Critically, the model is also consistent with our biochemical studies, by placing R310 and K328 of RCCD1 in close proximity to D73 and E77 (Fig 2.27B). This prediction supports the results of our biochemical investigation, confirming the importance of these residues for the JMJD5:RCCD1 interaction and alluding to a role for the complex in cancer.
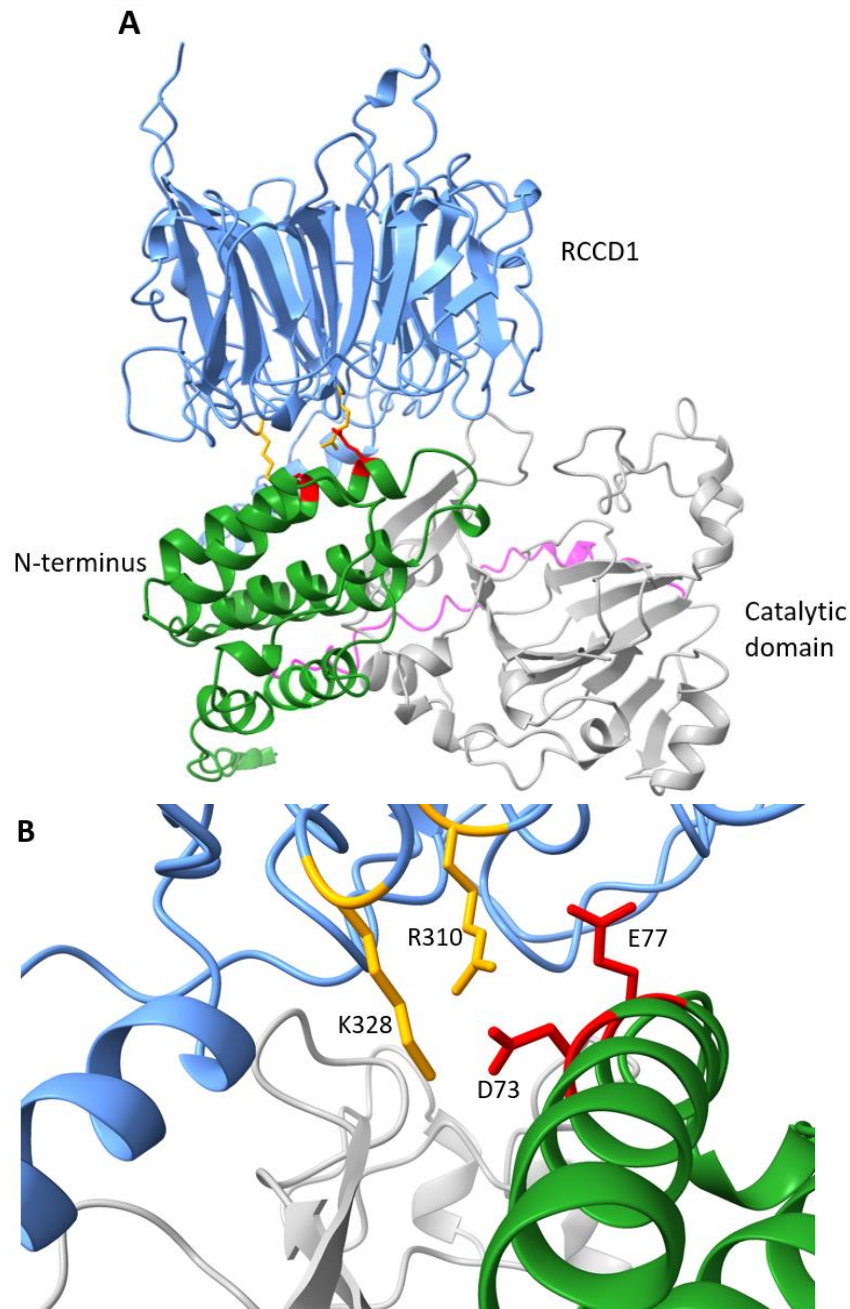
Figure 2.27. AlphaFold Multimer aligns JMJD5 D73 and E77 with RCCD1 R310 and K328.
(A) RCCD1 (blue) is predicted to interact with the N-terminus of JMJD5 (green). The JMJD5 Catalytic domain is coloured in grey and the flexible linker region in purple. **(B)** RCCD1 (blue) R310 and K328 (orange) are predicted to be close to D73 and E77 (red) on the JMJD5 N-terminus (green). All images using predicted AlphaFold structure and chimera software. The JMJD5 catalytic domain is coloured in grey.

2.3   Discussion

In this chapter we have characterised the interaction between RCCD1 and JMJD5 using multiple

approaches including *in vivo* and *in vitro* biochemistry, imaging, and structural modelling. We have

demonstrated that this interaction is facilitated through two negatively charged residues in the JMJD5 N-terminus (D73 and E77) and two positively charged residues (R310 and K328) at the base of the RCCD1 ring shape.

2.3.1 Investigating the binding regions of RCCD1 and JMJD5

To begin our investigation into how JMJD5 and RCCD1 interact, we used structure predictions to generate a panel of JMJD5 truncation mutations. We showed the minimal JMJD5 binding domain for RCCD1 interaction was within the first 110 amino acids. This region has been previously identified as being important for binding to RCCD1 by two groups who used similar truncation methods to show loss of RCCD1 binding in 111-416 and 151-416 (Marcon et al., 2014; Wu et al., 2017). In addition, we used a truncation mutation of 81-416 which also did not interact with RCCD1 (Fig 2.6B). From this we hypothesised that the RCCD1 binding site was likely within the first 80 amino acids of JMJD5. However, as the reciprocal truncation (1-80) did not express in either HEK293T cells (Fig 2.6A) or *E. coli* cells (Fig 2.7B) for recombinant protein we were unable to narrow down the region using our truncation-based approach.

Whilst our JMJD5 truncations support assignments made in the literature and further our understanding of the N-terminal region responsible for binding RCCD1, the data arising from our studies involving RCCD1 mutations are at odds with the literature. Marcon et al. compared an RCCD1 N-terminal deletion mutant that lacked the first 169 residues to full-length RCCD1 (Marcon et al., 2014). In their experiments JMJD5 was only able to interact with full-length RCCD1, leading the authors to conclude that N-terminus of RCCD1 facilitates the interaction with JMJD5. Wu et al also tested a panel of RCCD1 truncation mutants (1-164, 1-228, 165-376, 318-376 and 165-318) (Wu et al., 2017). Again, JMJD5 was only able to interact with full-length RCCD1. We propose that these observations are due to the truncation approach not being compatible with a protein such as RCCD1 that is predicted to have a largely 'singular' tertiary structure that it may not be possible to dissect into subdomains. Indeed, our own modelling confirms that RCCD1 is predicted to form a seven bladed

β-propellor structure, similar to all known structures for proteins containing RCC1-like domains (Koyama et al., 2017; Renault et al., 1998). Therefore, it seems likely that in both RCCD1 studies described above that the reported RCCD1 truncations have disrupted the overall ring structure, leading to indirect effects on the JMJD5 interaction. Indeed, both studies used the Uniprot assigned domain architecture to design their truncations (Marcon et al., 2014; J. Wu et al., 2017),which, based on our AlphaFold2 predictions (Fig 2.21) are likely not to be accurate. Partly because of these concerns we proposed an alternative approach, using a combination of preliminary biochemical data on critical JMJD5 residues, and recurrent RCCD1 cancer mutations, to test missense substitutions within the context of the full-length RCCD1 protein. Using this approach, we successfully identified R310 and K328 as critical JMJD5 binding residues (Fig 2.24 + Fig 2.26). The AlphaFold Multimer prediction suggests that the N-terminal JMJD5 helix that contains D73 and E77 lies across one plane of the RCCD1 ring to make contact with R310 and K328 (Fig 2.27A). As AlphaFold2 is only a prediction, further work to generate actual structures and biophysical data would be supremely beneficial. The current structural understanding of JMJD5 is limited to the catalytic domain, with the N-terminus not being resolved in the crystal structure (Wilkins et al., 2018). We have demonstrated here a clear functional role for the N-terminus, our understanding of which could be elucidated through structure:function analysis. Therefore, through collaboration with the Smerdon group we aimed to answer some of these questions, not just for JMJD5 but also RCCD1, of which there is no published structure. We have aimed to use high resolution structural biology techniques i.e X-ray crystallography and or cryo-Electron Microscopy to resolve structures, indeed the investigation into this will be discussed further in Chapter 4.

2.3.2 What is the physiological role of the JMJD5:RCCD1 interaction?

Whilst our biochemical investigation has identified the critical residues that facilitate the JMJD5:RCCD1 interaction, we still do not know the physiological role of the complex. Although previous studies have suggested that JMJD5 and RCCD1 are important for maintaining accurate mitotic division and microtubule stability (Marcon et al., 2014; Wu et al., 2017), the importance of

their physical interaction was not formally proven. Furthermore, it is not clear whether the mitotic effects observed were direct, or potentially indirect, through effects on S-phase DNA replication. Whether it could relate to our reported role of JMJD5 hydroxylase activity in replication fidelity (Fletcher et al., 2023) is not clear. Indeed, the potential importance of RCCD1, and the JMJD5:RCCD1 interaction in faithful DNA replication has not yet been investigated (see Chapter 3).

One potential role of the JMJD5:RCCD1 interaction may be to help localise RCCD1 to the nucleus. For example, we demonstrated that the localisation of overexpressed RCCD1 is largely restricted to the cytoplasm in the absence of JMJD5 overexpression (Fig 2.17). However, when wildtype JMJD5 was co-overexpressed, RCCD1 appeared to become enriched in the nucleus (Fig 2.17). Using PLA and binding deficient mutants we showed that this co-localisation was dependent on the JMJD5:RCCD1 interaction (Fig 2.17 and Fig 2.18). Of the two proteins, only JMJD5 is thought to contain any localisation sequences, with both a predicted nuclear localisation sequence (NLS) and a nuclear export sequence (NES) reported (Huang et al., 2013) (Fig 2.9A). This suggests that when bound to JMJD5, RCCD1 is capable of translocating to the nucleus by 'piggy backing' on the JMJD5 NLS. Further work is required to explore these dependencies in the context of the endogenous proteins. The JMJD5 and RCCD1 endogenous antibodies were not compatible with immunofluorescence (data not shown) but did show sufficient signal in Western blotting. Therefore, two orthogonal approaches to investigate endogenous protein localisation could be used. Firstly CRISPR-CAS9 tagging of the endogenous proteins with HA and FLAG-tags used above for IF and PLA, or secondly sub-cellular fractionation followed by Western blotting.

2.3.3   Are JMJD5 and RCCD1 cancer mutations pathological?

We began this investigation into the relationship between JMJD5 and RCCD1 to potentially shed light on a novel pathway contributing to cancer progression. As part of this, we generated panels of cancer mutations in both the JMJD5 N-terminus (Fig 2.9A) and full-length RCCD1 (Fig 2.22A). We used these panels to show that cancer variants JMJD5^{E77K}, RCCD1^{R310W}, and RCCD1^{K328N} significantly impair the

JMJD5:RCCD1 interaction (Fig 2.11B and Fig 2.24B). One limitation of our investigation of the RCCD1 residues required for JMJD5 binding was the candidacy approach based largely on recurrency and charge. Testing a larger panel of RCCD1 cancer mutations would likely identify other residues of importance. In the meantime, it would be interesting to test the importance of the tryptophan residues which appear to form hydrophobic region in the RCCD1 structure (Fig 2.25B). Investigating the RCCD1:JMJD5 multimer model also shows these tryptophan residues would be covered by the JMJD5 N-terminus maintaining an optimal hydrophobic environment. If the hydrophobic environment is essential for creating the conditions for the RCCD1:JMJD5 interaction, we hypothesised they might have been identified as variants in our list of reported RCCD1 cancer mutations. Investigating our full list of cancer variants revealed no missense variants in any of the identified tryptophan residues (Appendix 3). However, generating experimental mutations to alanine could be used to profile the effects of mutating the tryptophan residues including assaying protein stability and JMJD5 binding.

In addition to effects on interactions, we have shown that some mutations significantly disrupt protein expression, including RCCD1^{R344C} (Fig 2.24), JMJD5^{P27L} (Fig 2.10B), JMJD5^{S75F} (Fig 2.11B), and JMJD5^{L79P} (Fig 2.11B). Moreover, in the IP experiments both RCCD1^{R310W} and RCCD1^{K328N} also showed reduced protein expression (Fig 2.24B), which might suggest a role for the JMJD5:RCCD1 interaction in maintaining RCCD1 stability (Chapter 4). Although the post-transcriptional pathways controlling RCCD1 and JMJD5 protein expression require further investigation, our studies using cancer mutants suggest that one mechanism by which the JMJD5:RCCD1 pathway might be suppressed in tumours is through altered expression. Future studies could involve measuring protein half-life and proteosomal degradation of cancer mutant RCCD1 and JMJD5 for example.

For JMJD5, another potential functional consequence of cancer mutations could be loss of hydroxylase activity. A large panel of JMJD5 cancer mutations have been cloned and tested for activity in the Coleman group. Although, this analysis was predominantly based on residues in the catalytic domain, a large proportion of the cancer mutations tested were shown to reduce JMJD5 catalytic activity (Sally

Fletcher – data not shown). Of the cancer mutations in the N-terminus, only the activity of JMJD5^{E77K} was tested as part of this investigation (Fig 2.19), showing normal levels of hydroxylation. However, the limitations of our *in vitro* JMJD5 assay may not detect more complex effects of distal mutations on hydroxylation activity. For example, the assay uses a short synthetic substrate peptide, which is unlikely to be representative of more complex enzyme:substrate interactions in cells. Additionally, the RPS6 peptide used is not thought to be a physiologically relevant substrate, and so the effects observed may not accurately reflect the endogenous system (Fletcher et al., 2023). Considering our AlphaFold Multimer modelling (Fig 2.27) suggests that the JMJD5 N-terminus, and by extension, RCCD1, pack against the catalytic domain, it remains possible that N-terminal JMJD5 cancer mutations could regulate JMJD5 hydroxylase activity. It remains to be seen whether mutations in the N-terminus of JMJD5 could also impact its hydroxylase activity.

Overall, our work on JMJD5 and RCCD1 cancer mutations may be consistent with the complex being important for tumour suppression. However, further work is required to explore this, including for example in tumour models such as xenografts or transgenic mice. Because the biological functions of RCCD1 and JMJD5 remain enigmatic, a deeper understanding of the mechanistic roles of both could also significantly improve our understanding of their role in tumourigenesis (discussed more in Chapter 4).

2.3.4   Does the JMJD5 N-terminus contain other functionally important domains?

Whilst we have generated evidence suggesting that the N-terminus contains an important α-helix that facilitates RCCD1 binding (Fig 2.20C), both the initial prediction (Fig 2.4A) and the AlphaFold model (Fig 2.20B) suggest that the N-terminus contains additional regions of secondary structure. The function of these other structures remains unclear. Interestingly, the Coleman group have already demonstrated that a JMJD5^{C123Y} variant present in a novel developmental disorder causes reduced protein stability and a modest decrease in JMJD5 activity (Fletcher et al., 2023). Interestingly, the

C123Y variant was shown not to reduce RCCD1 binding, suggesting that the N-terminus contains at least two domains with functionally distinct roles.

Another group have proposed that the N-terminus of JMJD5 (residues 1-165), acts as an RNA Pol II binding domain that helps to recruit JMJD5 to nucleosomes to facilitate histone tail cleavage (Liu et al., 2020). However, the physiological relevance of this observation, and its reliability, remain unclear. Indeed, structural and biochemical studies question the protease activity assignment (Wilkins et al., 2018), and our own proteomic screens have not yet identified a specific RNA Pol II interaction (Chapter 4).

2.3.5   Chapter Conclusions

In this chapter we have demonstrated that the interaction between RCCD1 and JMJD5 is facilitated through oppositely charged amino acids. The negatively charged amino acids D73 and E77 in the JMJD5 N-terminus form an electrostatic interaction with the positively charge R310 and K328 located in RCCD1. Additionally, our work suggests that the JMJD5 N-terminus does indeed contain structurally and functionally important domains that may play many different aspects of JMJD5 biology that are relevant to tumourigenesis. Our structural modelling helps to visualise the JMJD5 N-terminus, RCCD1, and the JMJD5:RCCD1 complex, for the first time. The physiological importance of the JMJD5:RCCD1 in interaction is investigated in the chapters that follow.

Chapter 3:   The RCCD1 and JMJD5 interaction is important for replication stress

3.1   Introduction

In Chapter 2, we showed that the RCCD1:JMJD5 interaction is mediated by specific electrostatic interactions. However, the role of this interaction and its physiological importance remains to be investigated. The Coleman group previously reported a role for JMJD5 loss of function in a novel developmental disorder characterised by increased 'replication stress' (Fletcher et al., 2023), which we describe in more detail below. We hypothesised that if JMJD5 is important for suppressing replication stress, RCCD1 might also play a role. Therefore, we investigated whether RCCD1 loss of function causes replication stress here in this chapter. Before presenting the results, we provide an overview of the causes and cellular responses to replication stress, beginning with a brief description of DNA replication.

3.1.1   DNA replication

DNA replication is the synthesis of a complementary DNA strand to the original template, with accurate DNA replication being essential to ensure that daughter cells contain a complete genome copy of high fidelity. The DNA double helix is 'unzipped' by a DNA helicase to create a fork-shaped structure known as a replication fork, creating two single-stranded templates (Fig 3.1) (Leman and Noguchi, 2013). The 'leading' strand is orientated 3' to 5' and DNA polymerase ε progresses along this strand continuously adding complementary nucleotides (Fig 3.1) (Leman and Noguchi, 2013). Conversely, the 'lagging' strand is synthesised 'discontinuously' using short RNA primers which are extended by DNA polymerase δ forming 'Okazaki' fragments which are subsequently joined by a DNA ligase (Fig 3.1)(Leman and Noguchi, 2013).
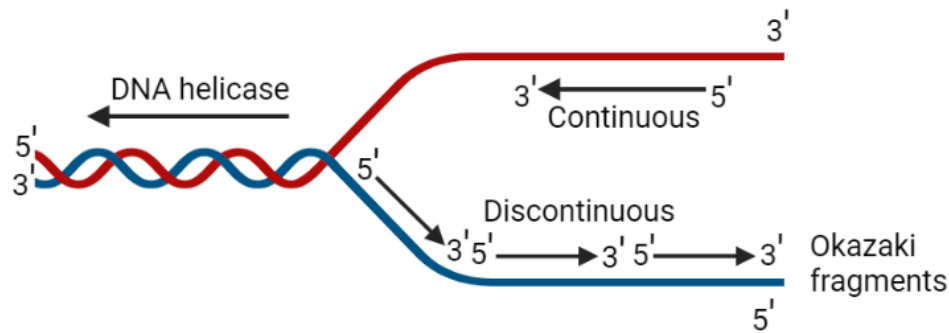
Figure 3.1. Progression of replication fork during DNA replication.
DNA replication on the leading strand occurs continuously in a 5' to 3' direction with DNA polymerase ε adding complementary nucleotide bases continuously. Synthesis of the 'lagging' stand occurs discontinuously, with short RNA primers used to synthesised short DNA segments known as 'Okazaki' fragments ligated together.

DNA replication occurs bidirectionally from sites known as origins of replication arranged along chromosomes, which are licensed prior to initiation (Bell and Dutta, 2002). Licencing of a replication origin occurs via the loading of the MCM helicase by the 'pre-replicative complex', which comprises the origin recognition complex (ORC), Cdc10 dependent transcript 1 (CDT1), and cell division cycle 6 (CDC6) (Mizushima et al., 2000; Speck and Stillman, 2007; Frigola et al., 2017). The MCM complex is a hexamer unit comprised of six subunits: MCM2-7, which couples with Cdc45 and GINS to create an active CMG helicase that unwinds the DNA to create a single-stranded template (Moyer et al., 2006; Ilves et al., 2010; Van Deursen et al., 2012). The replisome containing DNA polymerase can then proceed as described above (Fig 3.1).

3.1.2   DNA Replication stress

Replication stress is defined as the slowing or stalling of DNA replication forks under stress conditions (Zeman and Cimprich, 2014). If DNA progression is not restored, it can result in under-replicated DNA or even DNA damage, which can, in turn, contribute to genome instability. Therefore, it is essential for cells to have pathways to resolve replication stress; these pathways are discussed below.

3.1.2.1   Causes of DNA replication stress

DNA replication stress can arise from many sources, broadly categorised into endogenous and exogenous, both of which can have severe cellular consequences. For endogenous sources, the DNA replication fork can encounter many obstacles during its progression along DNA. These may be physical barriers, such as sites of DNA damage, DNA lesions, or secondary DNA structures (e.g. hairpins) which can block fork progression (Fig 3.2) (Thys and Wang, 2015; Zeman and Cimprich, 2014). Rather than being physical barriers, some DNA sequences may have intrinsic properties that make DNA replication more difficult. These regions which include telomeres and repetitive sequences, are termed common fragile sites (CFS) (Glover et al., 1984; Krasilnikova and Mirkin, 2004). The balance between transcription and replication can also contribute to replication stress (RS), as highly transcribed regions can also impede fork progression (Fig 3.2). During transcription, RNA can be hybridised with the DNA template, causing collisions between transcription machinery and the replisome, and DNA may become positively supercoiled between them (Fig 3.2) (García-Muse and Aguilera, 2016; Bermejo et al., 2012). Other sources of endogenous RS that are not due to physical barriers can be improper origin firing and depletion of essential factors such as nucleotides and new histones (Mejlvang et al., 2014; Toledo et al., 2013; Anglana et al., 2003).

RS due to endogenous sources can by exacerbated by exogenous stressors. DNA-damaging agents, such as ionising radiation, chemotherapies, and ultraviolet light, all cause DNA lesions, creating more physical barriers for the DNA replication fork (Fig 3.2) (Vesela et al., 2017). Other chemical agents such as aphidicolin inhibit DNA polymerases, and hydroxyurea limits the available nucleotide pool by diminishing nucleotide production through ribonucleotide reductase inhibition (Fig 3.2) (Vesela et al., 2017).

One result of RS is the functional uncoupling of the DNA helicase from the replisome. When the replisome stalls at an obstacle, the helicase continues to move and unwind the DNA (Byun et al., 2005). This continuous DNA unwinding can lead to long stretches of single stranded DNA (ssDNA). As such,

the presence of ssDNA localised to replication forks is a hallmark of RS (Gelot et al., 2015). Overall, several sources of RS exist, and cells are equipped with a number of responses that can help restore DNA replication fork progression.
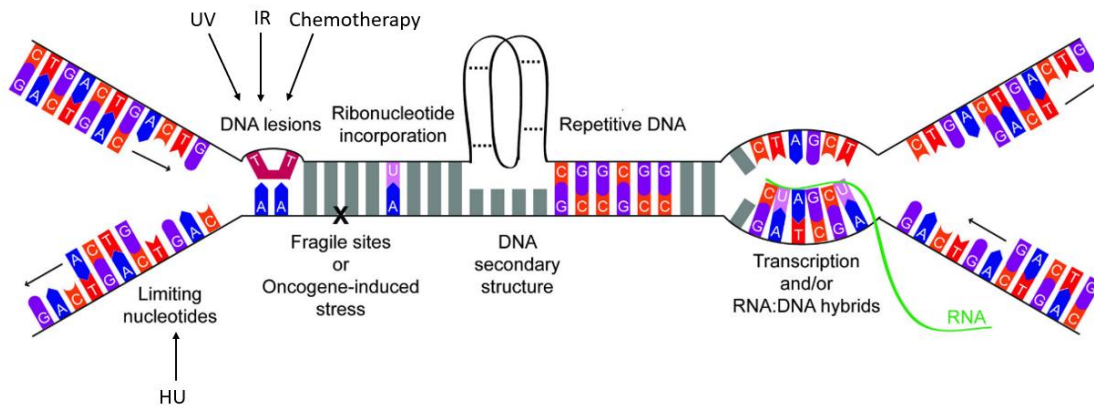
Figure 3.2. Causes of Replication Stress.
The sources of endogenous replication stress include fragile sites, limiting of nucleotides, DNA lesions, DNA structures, transcription:replication conflicts and DNA:RNA hybrids. These endogenous sources maybe exacerbated by hydroxyurea (HU), ionizing radiation (IR), ultraviolet (UV) and chemotherapies (e.g cisplatin). Figure taken and modified from Zeman and Cimprich, 2014.

3.1.2.2   Cellular response to RS

Tracts of ssDNA associated with stalled replication forks are targeted in one of the first steps of the response to RS (Fig 3.3). Replication Protein A (RPA) rapidly binds to ssDNA to prevent nuclease degradation and activate the response pathway (Fig 3.3) (Zou et al., 2006; Fanning et al., 2006). This signal is amplified by recruitment of the phosphoinositide 3-kinase related protein kinase (PIKK) 'ATM-Rad3-related' (ATR), which is facilitated by ATR-interacting protein (ATRIP) binding to RPA on ssDNA (Fig 3.3) (Cimprich and Cortez, 2008; Zou et al., 2006; Ball et al., 2007). The ATR:ATRIP complex is joined at ssDNA sites by RAD9-RAD1-HUS1 (the '9-1-1' complex), which is loaded onto the primer-template junction on 5' ssDNA (Fig 3.3) (Majka et al., 2006).

After ATR is recruited, it has two main functions: phosphorylation of the histone variant H2AX (γH2AX) to amplify the signal, and activation of the RS response by phosphorylating checkpoint kinase 1 (CHK1) (Ward and Chen, 2001; Lopez-Girona et al., 2001). Phosphorylated CHK1 is activated and released

from chromatin, regulating downstream cellular pathways through phosphorylation of its substrates, which causes rapid cell cycle arrest (Smits et al., 2006). This rapid cycle arrest is key to providing time for DNA repair and efficient replication fork restart, allowing the source of RS to be resolved and DNA replication to proceed. This activity of ATR-CHK1 also prevents stalled replication forks from collapsing, resulting in single-ended DSBs (Cortez, 2015).

Another response to RS that ensures complete DNA replication, especially when two converging forks are stalled, is dormant origin activation. During replication there is an excess of MCM helicase complexes loaded onto the DNA termed the 'MCM Paradox'; If DNA replication occurs unperturbed then most MCM complexes are not activated (Das et al., 2014). Under RS conditions, these dormant origins can be licenced and activated to complete DNA replication (Woodward et al., 2006; Ge et al., 2007). ATR-CHK1 has been proposed to activate these dormant origins through phosphorylation of MCM2 (Cortez et al., 2004; Ge and Blow, 2010).
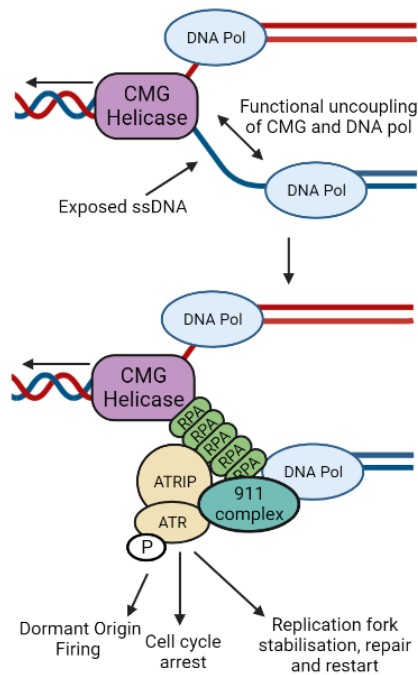
Figure 3.3. Cellular responses to replication stress.
The functional uncoupling of the CMG helicase unwinding the parental DNA duplex from the synthesizing DNA polymerase can lead to exposed tracts of ssDNA. This ssDNA can be bound by RPA, which in turn stimulates ATR recruitment through the ATRIP protein, the 9-1-1 complex (RAD9-RAD1-HUS1) is additionally recruited. ATR then phosphorylates downstream targets to trigger the RS response.

If cells are unable to repair any damage and complete DNA replication during the S Phase, an additional method exists to rectify this during the G2 phase and mitosis. Mitotic DNA synthesis, or MiDAS, is a major pathway to complete DNA replication before the completion of mitosis to prevent chromosome separation defects (Minocherhomji et al., 2015). CFS mentioned above can form ultrafine DNA anaphase bridges (UFBs) between sister chromatids which are commonly removed during mitosis for MiDAS repair (Garribba et al., 2018). Overall, these pathways and responses are important for preventing severe consequences of RS.

3.1.2.3   RS and disease

Unresolved RS can lead to increased genomic instability which is a key contributing factor to human diseases. There is a plethora of evidence demonstrating the link between RS and a group of developmental disorders broadly characterised as Primordial Dwarfism (PD), the key phenotypes of

which include intrauterine growth retardation and abnormal postnatal growth (Khetarpal et al., 2016).

Many studies have shown that inactivation of key proteins in the response to RS can lead to a severe

PD phenotype (Vetro et al., 2017; Harley et al., 2016; Reynolds et al., 2017). This is epitomised by ATR,

where mutations cause a faulty RS response and lead to a specific PD disorder known as Seckel

Syndrome (O'Driscoll et al., 2003). Interestingly, the presence of elevated replication stress in patients

with a novel developmental disorder supports a role for JMJD5 in replication fidelity (Fletcher et al.,

2023). Genome instability is also a hallmark of cancer, and RS is now understood as a driving force

behind this instability. The role of RS in cancer progression has been widely studied (Macheret and

Halazonetis, 2015; Gaillard et al., 2015; Hanahan and Weinberg, 2011). The Coleman group has also

begun investigating whether JMJD5 plays a role in cancer progression through increased RS, and we

have shown that depletion of JMJD5 increases markers of RS in cancer cell lines (Sally Fletcher,

personal communication).

In this chapter, we investigate the importance of RCCD1 and the RCCD1:JMJD5 interaction in

suppressing replication stress using human tumour cell lines as a model.

3.2   Results

3.2.1   Validation of RCCD1 siRNA as a knockdown model

To begin our investigation, we developed a loss-of-function method for RCCD1. To this end, we

ordered two independent siRNA sequences and validated their efficacies. We tested our siRNAs in the

bone osteosarcoma cell line U2OS and the lung adenocarcinoma cell line A549 because studies on

JMJD5 and replication stress were undertaken in these models (Sally Fletcher, personal

communication). The cells were seeded into 10 cm dishes, transfected after 24 h, and harvested 72 h

post-transfection. Once harvested, the cell pellets were lysed, and Western blotting was performed

for RCCD1 and Actin as a loading control.

Both siRNA sequences caused efficient RCCD1 knockdown (KD) in both cell models (Fig 3.4), suggesting that they are a suitable method of RCCD1 depletion, allowing further investigation of RCCD1 loss-of-function.
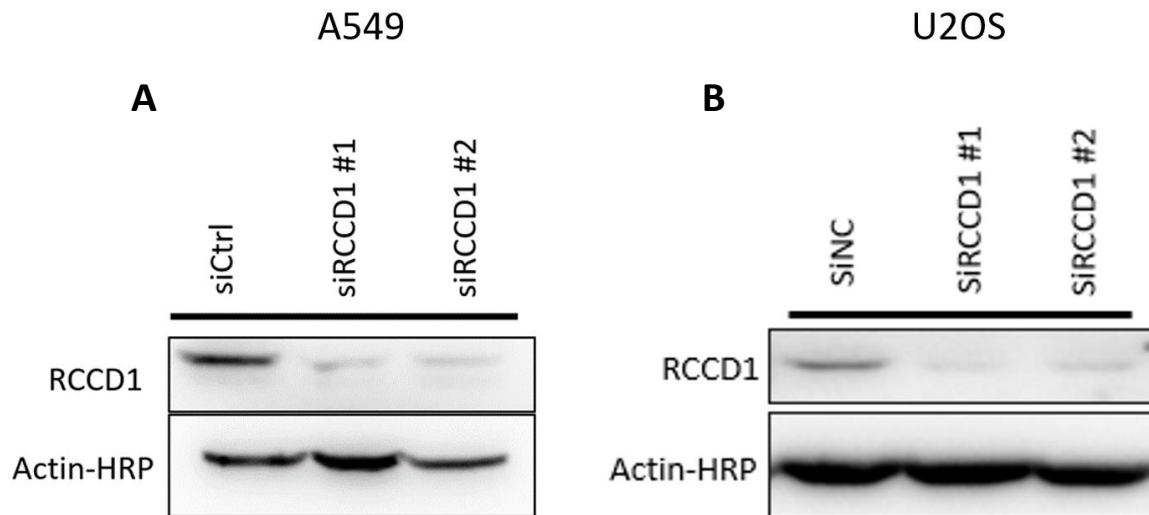
Figure 3.4. RCCD1 siRNA successfully knockdown RCCD1 in A549 and U2OS cancer cell lines.
(A) A549 and **(B)** U2OS cells were transfected with the indicated siRNA for 72 h before the cells were harvested, lysed, and Western blotted with the indicated antibodies.

3.2.2   RCCD1 depletion increases indirect markers of replication stress

Methods to measure the replication stress phenotype are broadly characterised into indirect and direct (DNA fibre assay – discussed in section 3.2.3) methods. The indirect methods provide a cheaper and faster alternative for initial RS phenotyping. We began by quantifying two indirect markers: micronuclei and p53 binding protein 1 (53BP1) bodies. Micronuclei are small extra-nuclear fragments detected by DAPI staining, which are formed from either whole or fragmented chromosomes that cannot be incorporated into daughter nuclei following mitosis (Fig 3.5A) (Fenech et al., 2011; Luzhna et al., 2013). These can arise from many sources, including errors in mitosis and DNA damage, but are also hallmarks of prolonged periods of RS (Fenech et al., 2011). 53BP1 bodies are large nuclear bodies that form in response to RS and are often identified in the G1 phase of daughter cells resulting from RS in the previous S-phase (Fig 3.6A) (Fernandez-Vidal et al., 2017). 53BP1 can be visualised

throughout the cell cycle in small nuclear foci, however, 53BP1 bodies as a marker of RS are exclusive

to G1 phase (Anderson et al., 2001; Ward et al., 2003; Fernandez-Vidal et al., 2017).

3.2.2.1 siRNA knockdown of RCCD1 increases levels of micronuclei

Consistent with increased replication stress, transfection of the two independent RCCD1 siRNA

sequences in A549 (Fig 3.5B) or U20S (Fig 3.5C) caused a statistically significant increase in the number

of micronuclei.
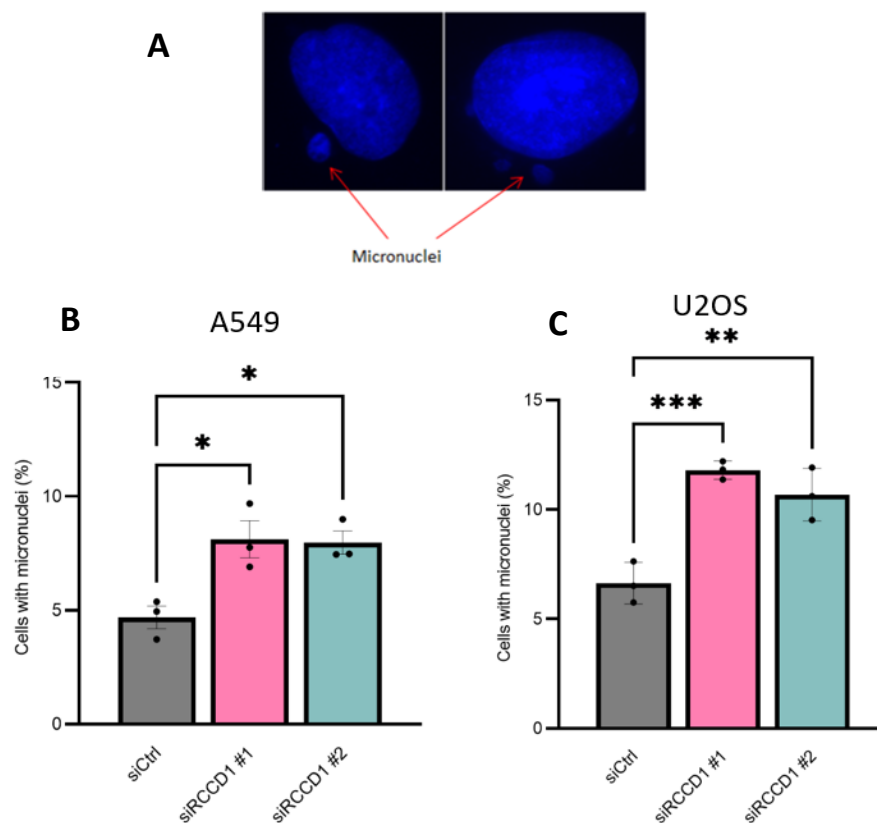
Figure 3.5. RCCD1 siRNA knockdown increases spontaneous micronuclei formation.
(A) Example images highlighted observed micronuclei. Cells were stained with DAPI (blue). **(B)** A549 cells transfected with RCCD1 siRNA have increased micronuclei. **(C)** U2OS cells transfected with RCCD1 siRNA have increased micronuclei. Both cell lines were transfected with the indicated siRNAs; siCtrl (grey), siRCCD1 #1 (pink) and siRCCD1 #2 (blue) for 72 h before the cells were harvested and stained. Statistical analysis used One way ANOVA with Bonferroni's post hoc test (with p-values of ≤0.05 (*), ≤0.01 (**), ≤0.001 (***) and ≥0.05 (ns)). Data shown is the mean of three independent biological repeats -/+ SEM with a minimum of 500 cells were counted per sample in each biological repeat.

3.2.2.2 Levels of 53BP1 bodies increase after RCCD1 depletion

Because the presence of 53BP1 bodies in the G1 phase of the cell cycle are a marker of RS, we co-stained cells with an antibody to a cell cycle marker called 'Centromere protein-F' (CENPF). CENPF is absent in G1 cells, but gradually accumulates across the cell cycle and peaks in the G2/M phase. Therefore, we counted 53BP1 bodies in cells that were negative for CENPF staining (Fig 3.6A) (Liao et al., 1995). Consistent with the micronuclei data, there was a statistically significant increase in 53BP1 bodies following transfection of both siRCCD1 sequences in A549 (Fig 3.6B) and U2OS (Fig 3.6C) cells.
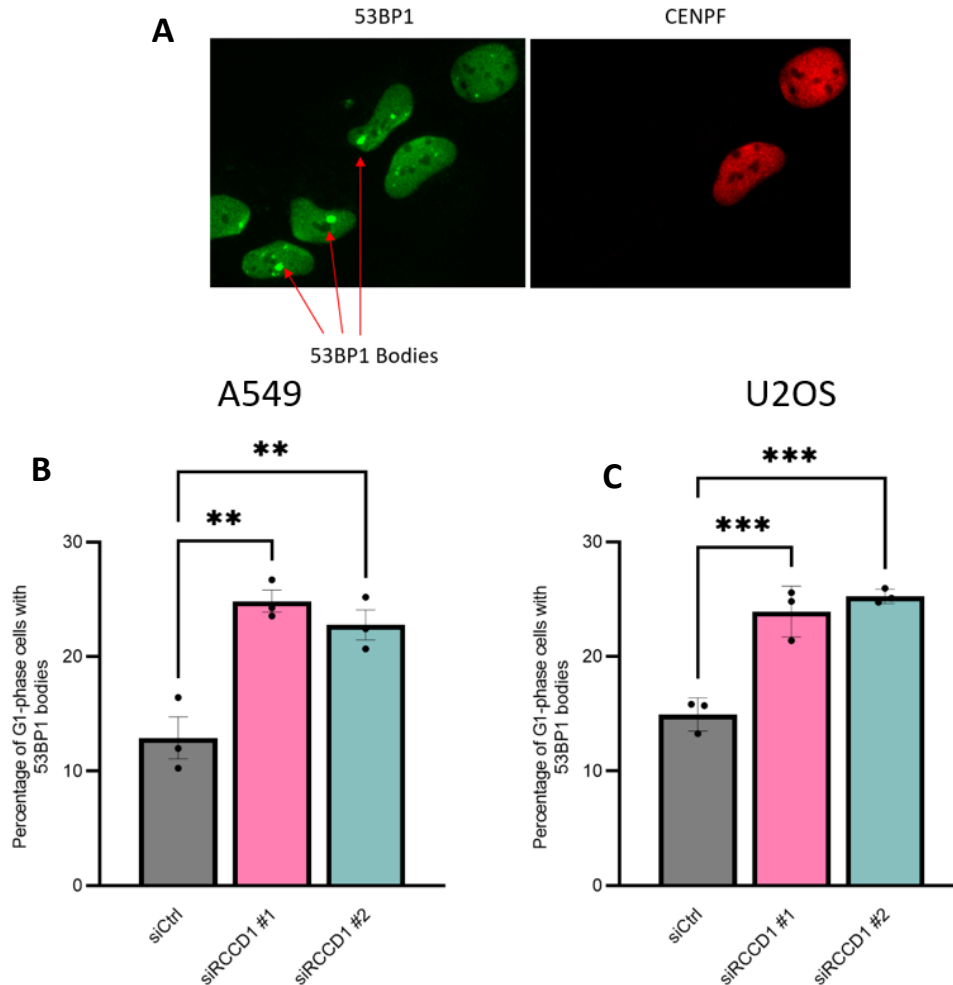
Figure 3.6. RCCD1 knockdown increases spontaneous 53BP1 bodies.
(A) Example images highlighting observed 53BP1 bodies. Cells were stained with 53BP1 (green).
(B) A549 cells transfected with RCCD1 siRNA have increased G1-phase 53BP1. **(C)** U2OS cells transfected with RCCD1 siRNA have increased G1-phase 53BP1 bodies. Both cell lines were transfected with indicated siRNAs; siCtrl (grey), siRCCD1 #1 (pink) and siRCCD1 #2 (blue) for 72 h before the cells were harvested and stained. Statistical analysis used One way ANOVA with Bonferroni's post hoc test (with p-values of ≤0.01 (**), ≤0.001 (***) and ≥0.05 (ns)). Data shown represents the mean of 3 biological repeats -/+ SEM. A minimum of 300 cells were counted per sample in each biological repeat.

Since both indirect markers of RS were elevated by RCCD1 knockdown, we next investigated more direct markers.

3.2.3 Direct markers of replication stress are increased by RCCD1 loss of function

DNA replication progression can be measured directly using the DNA fibre assay, which utilises incorporation of halogenated thymidine nucleotide analogues into replicating DNA in cells. These analogues are visualised using immunofluorescence which can provide resolution at the single-

molecule level (Nieminuszczy et al., 2016). Full details of the protocol can be found in the Materials and Methods section (Chapter 6). Briefly, cells were first incubated with chlorine-labelled thymidine (CldU) for 20 min, followed by incubation with iodine-labelled thymidine analogue (IdU). After incubation, the cells were harvested, lysed and spread onto microscope slides before staining with primary antibodies against the two analogues. CldU was then detected using Alexa Fluor® 555 (red) secondary antibody, and IdU was detected using Alexa Fluor® 488 (green) secondary antibody.

The relationship between these respective stains allows us to identify different fibre structures (Fig 3.7), including ongoing Forks, second label origins, stalled forks, first label origins, and second label terminations (Fig 3.7A). From these structures, we can quantify the number of stalled forks (Section 3.2.3.1) and fork asymmetry (Section 3.2.3.2), which are both key markers of replication stress. A condition with addition of hydroxyurea was included for each siRNA (Fig 3.7B): As discussed above, HU induces replication stress, providing a positive control for abnormal replication fork progression (Section 3.1.2.1).
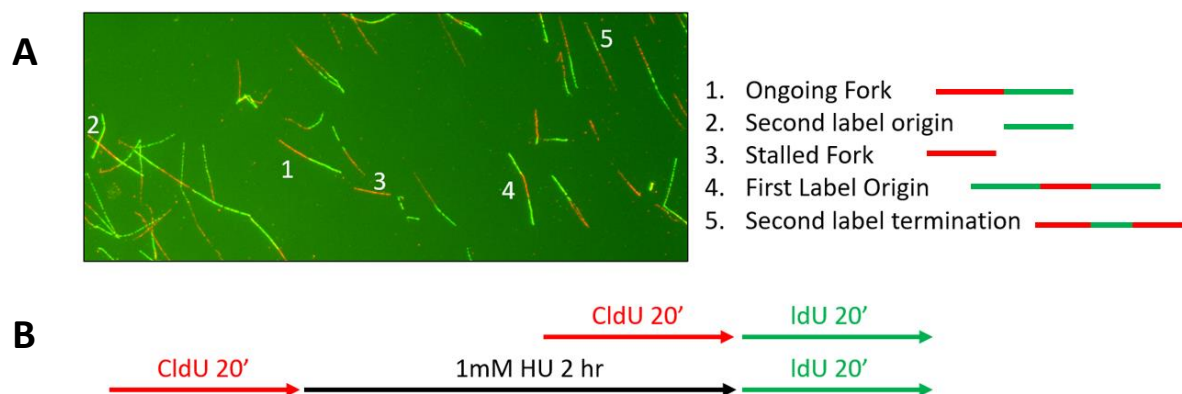
Figure 3.7. DNA fibre assay for direct monitoring of replication stress.
(A) Example fibre structures quantified using DNA fibre assay (1. Ongoing fork, 2. Second label origin, 3. Stalled Fork, 4. First Label origin and 5. Second label termination). **(B)** Assay schematic detailing length of incubations for CldU, HU and IdU treatments. Red indicates Alexa Fluor® 555 secondary used to visualise CldU and Green is the Alexa Fluor® 488 secondary used to visualise the IdU.

3.2.3.1 siRNA knockdown of RCCD1 increases levels of stalled forks

To investigate the direct impact of RCCD1 depletion on replication fidelity, we first quantified the levels of stalled forks. Stalled forks are indicated by the presence of the first label only (Fig 3.7) and are indicative of a replication fork meeting a barrier and either stalling or terminating the incorporation of nucleotides into the nascent DNA strand. The number of stalled forks is expressed as a percentage of the DNA fibres analysed, with the remaining structures quantified and classified as ongoing forks.

In both A549 and U2OS cells, siRNA knockdown of RCCD1 significantly increased the level of stalled forks (Fig 3.8A + B). As expected, the addition of HU increased the level of stalled forks observed under all conditions; however, the phenotype observed for siRCCD1 #1 and #2 was even further elevated (Fig 3.8A + B).
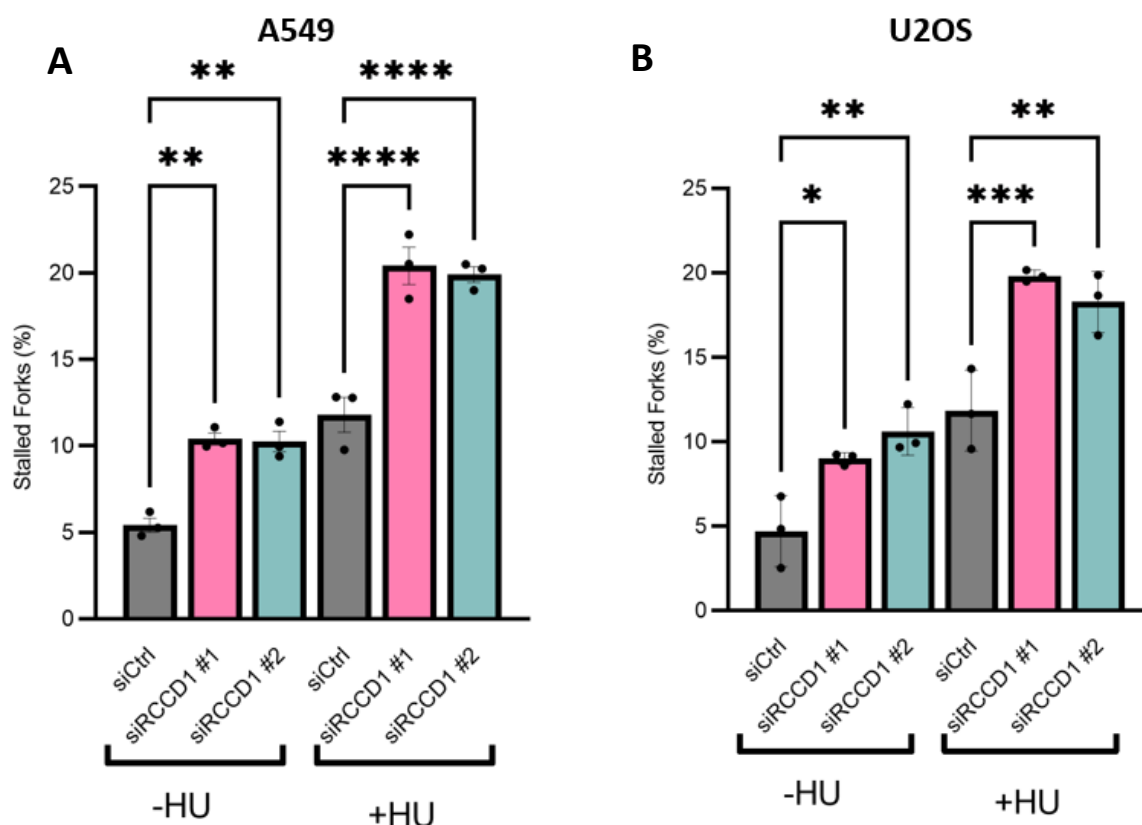
Figure 3.8. siRCCD1 knockdown causes increased prevalence of stalled replication forks.
The indicated cell lines were transfected with either siCtrl (grey), siRCCD1 #1 (pink), or siRCCD1 #2 (blue) and treated with HU where indicated. Levels of stalled forks increased in both **(A)** A549 and **(B)** U2OS following siRCCD1 knockdown**.** Statistical analysis used One way ANOVA with Bonferroni's post hoc test (with p-values of ≤0.01 (**), ≤0.001 (***), ≤0.0001 (****) and ≥0.05 (ns)). The data shown represents the mean of 3 biological repeats -/+ SEM. A minimum of 250 forks were counted per condition in each biological repeat.

3.2.3.2 Fork asymmetry is increased following RCCD1 depletion

Normal DNA replication occurs bidirectionally from the same origin and approximately equal speed, resulting in 'fork symmetry'. However, replication stress perturbs this, which can be quantified as fork asymmetry using the DNA fibre assay. To measure this, we used first label origins (Fig 3.7A), which are structures that contain green-red-green labelling, indicative of bidirectional replication that started during the first label. By measuring the length of the green tracts on either side of the red tract we can calculate their ratio to one another. A ratio close to one suggests equal fork speed consistent with fork symmetry, whereas a ratio greater than one suggests differential speeds and tract length, indicative of fork asymmetry.

Consistent with the other markers presented above, RCCD1 knockdown increased the levels of fork asymmetry in both A549 and U2OS cells (Fig 3.9A + B). This phenotype was exacerbated by the addition of HU (Fig 3.9A + B). Overall, we have shown that RCCD1 depletion results in an RS phenotype similar to JMJD5.
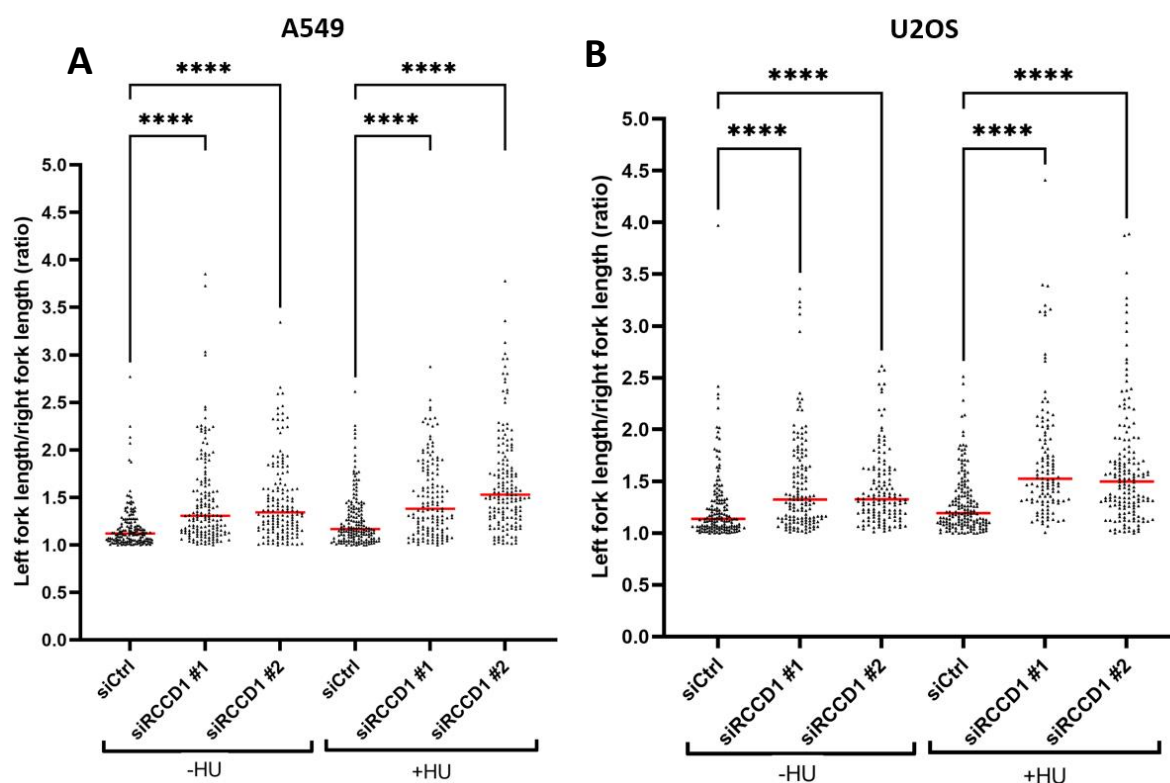
Figure 3.9. Fork asymmetry is increased following siRCCD1 knockdown.
The indicated cell lines were transfected with either siCtrl, siRCCD1 #1, or siRCCD1 #2 and treated with HU where indicated. Fork asymmetry was increased in both **(A)** A549 and **(B)** U2OS following siRCCD1 knockdown**.** Statistical analysis used Kruskal-Wallis with Dunn's correction with p-values of ≤0.0001 (****). The data shown is the average of 3 biological repeats. A minimum of 50 asymmetric forks were counted per condition per repeat.

3.2.4 Micronuclei arising from JMJD5 or RCCD1 knockdown result from a DNA replication defect

Although micronuclei can result from RS in S-phase, they can also potentially arise from mitotic defects (Luzhna et al., 2013). Both JMJD5 and RCCD1 have been shown to potentially play a role in accurate mitotic division (Marcon et al., 2014). Therefore, it was possible that our JMJD5 (Fletcher et al., 2023)

and RCCD1 (Fig 3.5) micronuclei phenotypes were an indirect result of a deficiency in another cellular pathway that operates in a distinct cell cycle phase. In Fletcher et al. (2023), we used a method to identify the source of the micronuclei we quantified, showing that in cells derived from the affected JMJD5 patients, the micronuclei came from a DNA replication/repair defect (Fletcher et al., 2023). This involved co-staining micronuclei for two specific markers. 53BP1 foci observed in micronuclei are indicative of a deficiency in DNA replication/repair, as the micronuclei are likely to contain fragments of damaged chromosomes. Additionally, if micronuclei are positive for centromeric protein A (CENPA), they are likely to contain whole chromosomes, indicating chromosome mis-segregation during mitosis.

We applied this methodology to our cancer cell line models, focusing on A549 cells. We transfected the cells with RCCD1 siRNA, as described above, and stained the cells for CENPA and 53BP1. The MN were then categorised as either CENPA or 53BP1 positive and quantified as a percentage of the MN counted. Similar to data arising from our JMJD5 patient-derived cell models, we observed a statistically significant increase in the number of 53BP1 positive micronuclei following RCCD1 knockdown (Fig 3.10). Importantly, this was in the context of no increase in the number of CENPA-positive micronuclei. Therefore, the RS phenotype for JMJD5 and RCCD1 depletion likely arises from a DNA repair/replication pathway defect. However, it remains to be seen whether these effects of JMJD5 and RCCD1 knockdown reflect a role in a common pathway that might be consistent with them forming a complex.
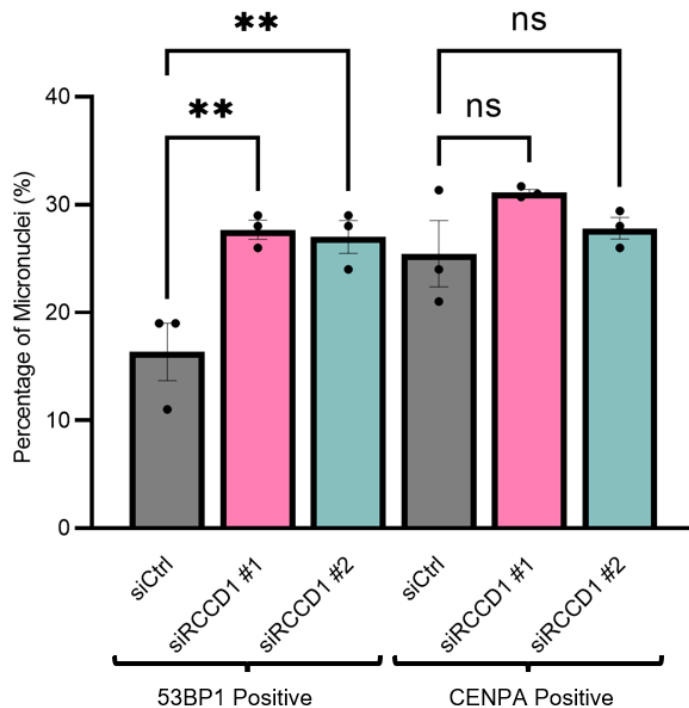
Figure 3.10. The RCCD1 knockdown micronuclei phenotype arises from a DNA replication/repair defect. A549 cells were transfected with either siCtrl (grey), siRCCD1 #1 (pink), or siRCCD1 #2 (blue) and 53BP1/CENPA positive MN were quantified. Statistical analysis used One way ANOVA with Bonferroni's post hoc test (with p-values of ≤0.01 (**) and ≥0.05 (ns)). The data shown is the average of 3 biological repeats.

3.2.5 JMJD5 and RCCD1 RS phenotypes are epistatic

Following the discovery that RCCD1 KD results in a similar RS phenotype to that of JMJD5, we next tested whether this relationship was epistatic. We hypothesised that their physical interaction underlies their role in a common biological pathway which, when lost, causes replication stress. To explore this, we performed an epistasis analysis by using siRNA depletion of the two proteins together, or separately, followed by quantification of both indirect and direct markers of RS. Any additional increase in the RS phenotype following the double knockdown of the two proteins would indicate that they are not in the same pathway and that the relationship is not epistatic. Conversely, no additive effect of double knockdown would be consistent with an epistatic relationship indicative of them acting in the same biological pathway together.

Before we explored RS epistasis, we first wanted to test whether our siRNAs could be used in combination with high efficacy. A549 and U2OS cells were transfected with siCtrl, siRCCD1/Ctrl, siJMJD5/Ctrl, or siRCCD1/JMJD5, followed by Western blot analysis for RCCD1, JMJD5, and Actin. Reassuringly, we observed efficient knockdown of JMJD5 and RCCD1, both in isolation and in

combination (Fig 3.11). Interestingly, siRCCD1 transfection caused a slight reduction in JMJD5 expression in A549 cells (Fig 3.11A). There was a significant reciprocal effect on RCCD1 levels following JMJD5 knockdown (Fig 3.11A). Although we observed a similar reduction in JMJD5 expression following RCCD1 knockdown in U2OS cells (Fig 3.11B), JMJD5 knockdown did not reduce RCCD1 expression in these cells.
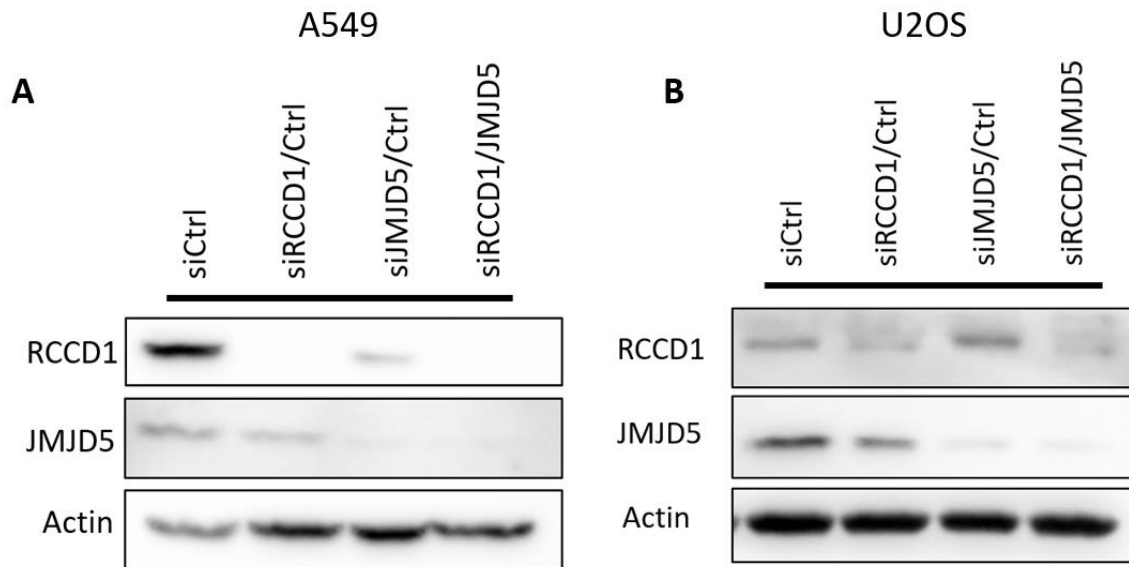
Figure 3.11. siRNA combinations successfully knockdown RCCD1 and JMJD5 in A549 and U2OS cancer cell lines.
(A) A549 and **(B)** U2OS cells were transfected with the indicated siRNA for 72 h before the cells were harvested, lysed, and Western blotted with the indicated antibodies.

3.2.5.1 Increase in 53BP1 bodies and MN after RCCD1 and JMJD5 knockdown is Epistatic

After successful siRNA validation, we began our epistatic analysis by quantifying 53BP1 bodies and micronuclei in siRNA transfected A549 and U2OS cells. As expected, and consistent with prior results, A549 cells showed a statistically significant increase in both 53BP1 bodies and micronuclei after RCCD1 or JMJD5 knockdown (Fig 3.12A + B). Importantly, we observed no additional increase in either marker in RCCD1/JMJD5 double knockdown cells (Fig 3.12A + B). Similar observations were made in U2OS cells (Fig 3.12C + D). Overall, these analyses suggest that the relationship between JMJD5 and RCCD1 with respect to 53BP1 bodies and micronuclei is epistatic in both cell lines, consistent with them acting together in the same pathway.
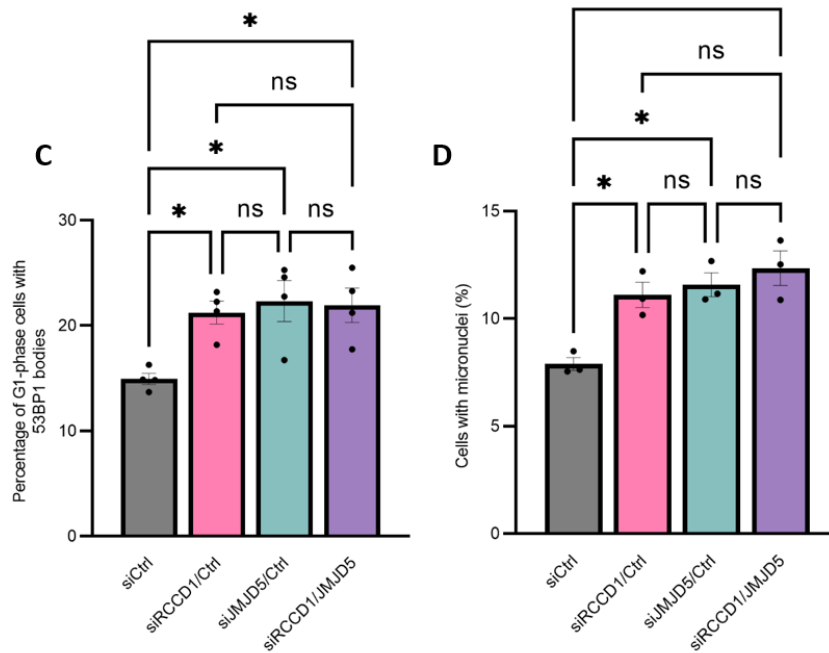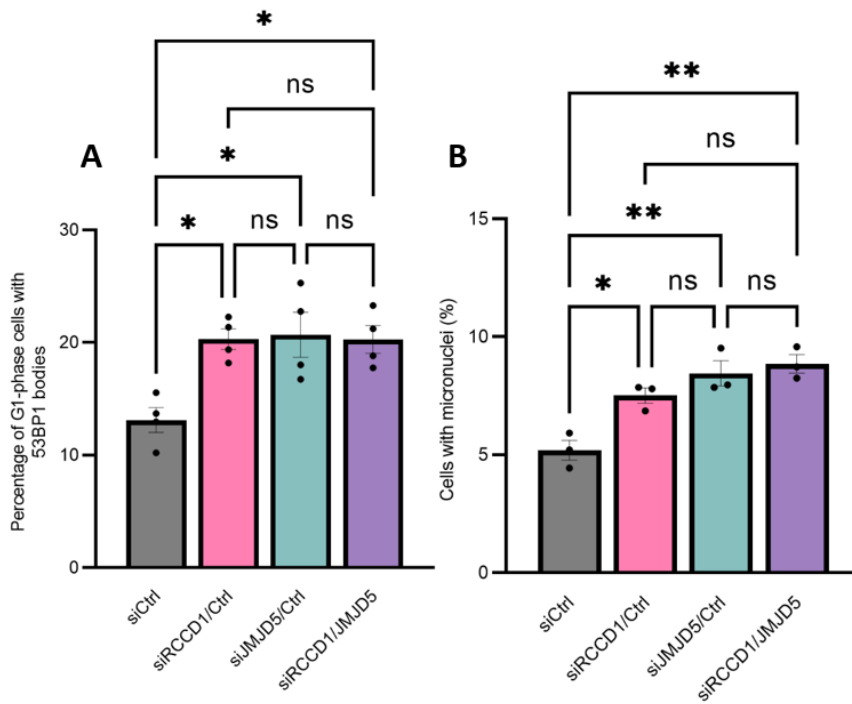
Figure 3.12. JMJD5 and RCCD1 are epistatic with respect to 53BP1 and micronuclei induction following knockdown.
Cell lines were transfected with the indicated siRNA; siCtrl (grey), siRCCD1/Ctrl (pink), siJMJD5/Ctrl (blue), or siRCCD1/JMJD5 (purple) for 72 h before the cells were harvested and stained. RCCD1 and JMJD5 RS phenotypes are epistatic with A549 **(A)** and **(B)** and U2OS **(C)** and **(D)** cells showing **(A)** and **(C)** increased spontaneous 53BP1 bodies and **(B)** and **(D)** increased spontaneous micronuclei. Statistical analysis used One way ANOVA with Bonferroni's post hoc test (with p-values of ≤0.05 (*) ≤0.01 (**) and ≥0.05 (ns)). The data shown is the mean of 3 biological repeats **(B)** and **(D)** or 4 biological repeats **(A)** and **(C)** -/= SEM. For micronuclei, a minimum of 500 cells were counted per sample in each biological repeat. For 53BP1 bodies, a minimum of 300 cells were counted per sample in each biological repeat.

3.2.5.2   The role of JMJD5 and RCCD1 in replication fork dynamics is epistatic

Following our epistasis work using indirect markers, we extended our investigation using the DNA fibre assay, focussing on A549 cells, as above.

Consistent with our previous observations, the number of stalled forks increased after the knockdown of either RCCD1 or JMJD5, which was further exacerbated by the addition of HU (Fig 3.13). In line with our epistasis analysis using indirect markers, there was no additional phenotypic increase over single knockdown when applying RCCD1/JMJD5 double knockdown, even in the presence of HU (Fig 3.13). This suggests that the increase in stalled forks caused by RCCD1 and JMJD5 loss of function is also epistatic.
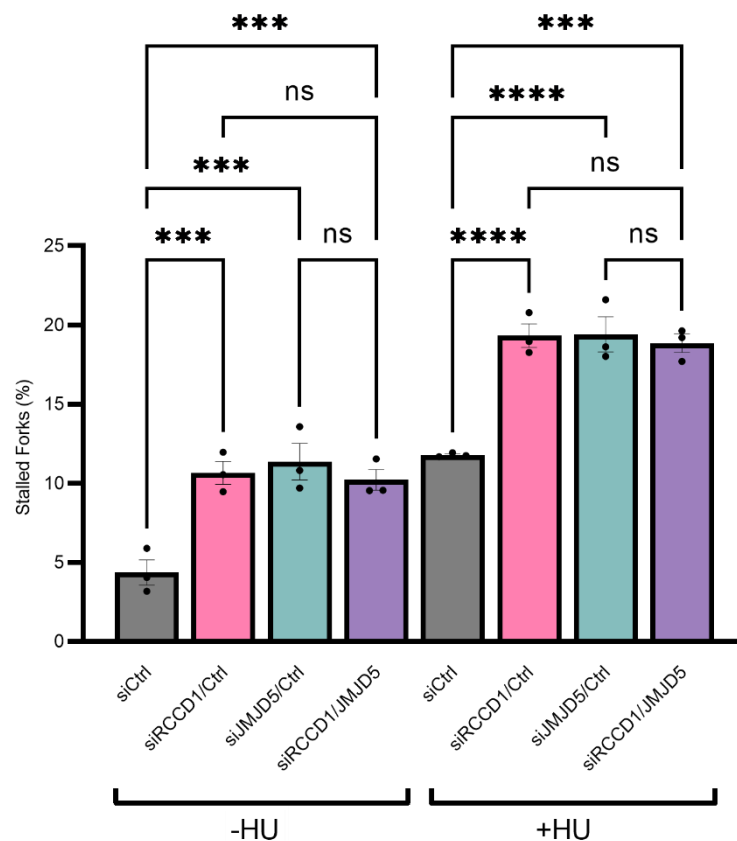
Figure 3.13. JMJD5 and RCCD1 are epistatic with respect to stalled replication forks following knockdown.
A549 cells were transfected with either siCtrl (grey), siRCCD1/Ctrl (pink), siJMJD5/Ctrl (blue), and siRCCD1/JMJD5 (purple) and treated with HU where indicated. Levels of stalled forks increased A549. Statistical analysis used One way ANOVA with Bonferroni's post hoc test (with p-values of ≤0.001 (***), ≤0.0001 (****) and ≥0.05 (ns)). Data shown is the mean of 3 biological repeats -/+ SEM. For each condition a minimum of 250 forks were counted per biological repeat.

We next quantified fork asymmetry, which showed an increase following single RCCD1 or JMJD5

knockdown (+/- HU), as expected (Fig 3.14). Consistent with the other markers of RS, fork asymmetry

was not further induced by RCCD1/JMJD5 double knockdown in A549 cells (Fig 3.14). Overall, our

combined analyses of multiple markers of RS suggest that the roles of RCCD1 and JMJD5 in replication

fidelity are epistatic to one another, suggesting that they likely act in a common pathway.
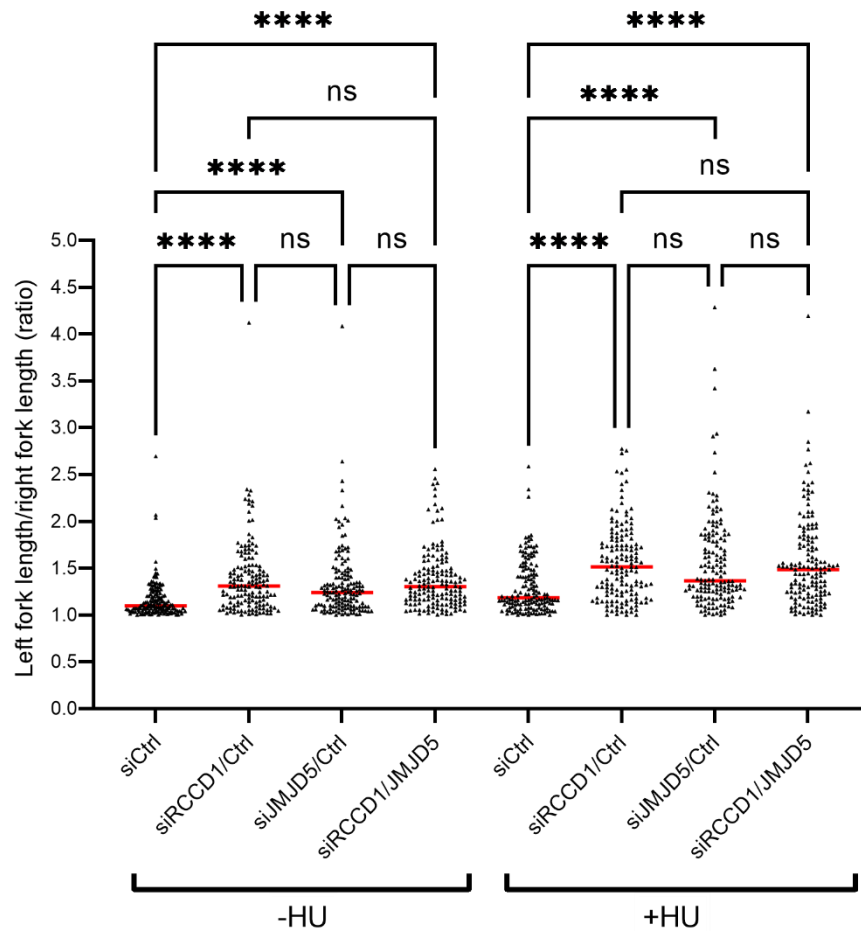
Figure 3.14. JMJD5 and RCCD1 are epistatic with respect to fork asymmetry following knockdown.
A549 cell lines were transfected with either siCtrl, siRCCD1/Ctrl, siJMJD5/Ctrl, or siRCCD1/JMJD5 and
treated with HU where indicated. Fork asymmetry was increased in A549 cells**.** Statistical analysis
used Kruskal-Wallis with Dunn's correction (with p-values of ≤ 0.0001 (****) and ≥ 0.05 (ns)). The data
shown is the average of 3 biological repeats. A minimum of 50 asymmetric forks were counted per
condition per repeat.

3.2.6 RCCD1 binding is required for the role of JMJD5 in replication fidelity

Considering our epistasis results in the context of the interaction studies presented in Chapter 2 led us to hypothesis that the JMJD5 and RCCD1 interaction could be required to maintain replication fidelity. To explore this, we wanted to apply a 'reconstitution' system where siRNA-resistant wildtype (WT) or RCCD1 binding defective JMJD5 could be expressed in JMJD5 knockdown cells.

Fletcher et al. (2023) used a pTIPZ doxycycline -inducible expression vector to restore JMJD5^{WT} expression in patient-derived fibroblasts (Fletcher et al., 2023). Whereas reconstitution of JMJD5^{WT} was sufficient to reduce RS markers, a catalytically dead iron-binding mutant (H321A) was not. We therefore considered that this system could potentially be applied to assess the importance of other JMJD5 structure-function relationships in RS. For example, the importance of RCCD1 binding could be determined by expressing JMJD5$^{D73A/E77A}$ mutant created in Chapter 2.

Importantly, the pTIPZ vector system was also predicted to be compatible with RNA interference because our JMJD5 siRNA sequence targets the 3'UTR of the endogenous JMJD5 transcript (which is not present in the vector expressed cDNA).

3.2.6.1 Generation and characterisation of the JMJD5 rescue system

We generated the pTIPZ rescue vector by cloning JMJD5$^{D73A/E77A}$ with the addition of an N-terminal 3XFLAG tag. Following sequence verification, lentiviral infection was used to generate cells expressing pTIPZ EV, or JMJD5 WT, H321A, or D73A/E77A. We decided to focus our study on A549 cells because we had already characterised the role of JMJD5 and RCCD1 in replication fidelity using these cells (sections 3.2.2 and 3.2.3), and because this reconstitution system had also been successfully applied in that context (Sally Fletcher, personal communication).

We first characterised 3XFLAG JMJD5 expression in A549 cells following doxycycline treatment (data not shown). Consistent with our observations in Chapter 2 (Fig 2.15), JMJD5$^{D73A/E77A}$ was not well expressed, requiring a higher Dox concentration to achieve comparable expression to that of JMJD5^{WT} and JMJD5^{H321A} (data not shown). Following optimisation, we found that 10 ng/mL doxycycline was

sufficient to achieve physiological expression of JMJD5^{WT} and JMJD5^{H321A}, whereas 1000 ng/mL was required for comparable expression levels of JMJD5$^{D73A/E77A}$ mutant (data not shown).

To validate 3X-FLAG JMJD5 re-expression in the context of a knockdown experiment, we grew the pTIPZ A549 cell lines in the relevant dose of doxycycline for 1 week (consistent with previous assays; Sally Fletcher, personal communication), before transfecting with control or JMJD5 siRNA for 72 h. Figure 3.13 presents a representative Western blot that demonstrates expression of 3XFLAG-JMJD5 constructs in the context of endogenous JMJD5 knockdown. Although we were able to confirm similar JMJD5 reconstitution levels by anti-FLAG Western blot (Fig 3.15), this was not possible using the anti-JMJD5 antibody, which was unable to detect the D73A/E77A mutant, presumably because these mutations destroy the antigen. Consistent with our previous results (Fig 3.11), JMJD5 knockdown caused reduced RCCD1 expression (Fig 3.15). Overall, this expression trials confirms we have a functioning reconstitution system that can be used for RS analysis to determine the importance of the JMJD5:RCCD1 interaction.
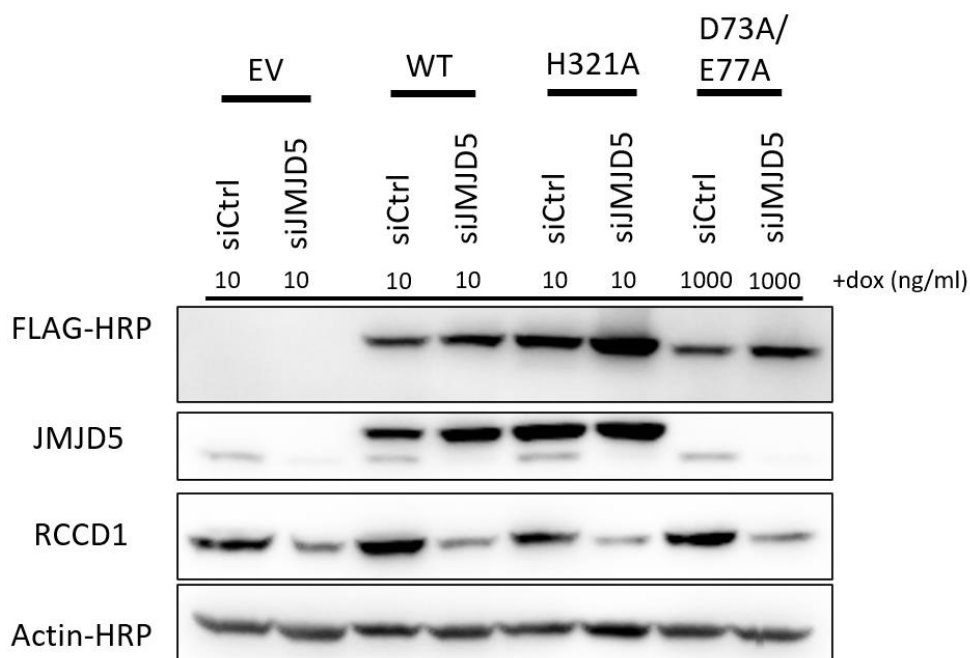
Figure 3.15. Representative Western blot of A549 pTIPZ 3XFLAG-JMJD5 reconstitution system. pTIPZ A549 cells were treated with the indicated amount of dox (ng/mL) for one week. Cells were then transfected with the indicated siRNA for 72 h before being harvested. Samples were then Western blotted for FLAG, JMJD5, RCCD1 and Actin as a loading control.

3.2.6.2 The RCCD1:JMJD5 interaction is required for the role of JMJD5 in maintaining replication fidelity

Using the same knockdown/reconstitution conditions presented in Figure 3.13, we measured MN and 53BP1 bodies, as above (Section 3.2.2). As expected, both RS markers were increased in pTIPZ EV cell lines following siJMJD5 transfection (Fig 3.16A +B). In the presence of reconstituted WT JMJD5, there was no increase in the levels of 53BP1 bodies or MN after siJMJD5 transfection (Fig 3.16A + B). This confirms the results that others have demonstrated with this approach (Sally Fletcher, personal communication) and validate the JMJD5 siRNA RS phenotype as 'on target'. Importantly, JMJD5^{H321A} reconstitution failed to prevent RS induced by JMJD5 knockdown, as observed by others (Fletcher et al., 2023). Critically, the JMJD5$^{D73A/E73A}$ RCCD1 binding mutant was also completely unable to restore normal replication fidelity in cells with endogenous JMJD5 knockdown, as evidenced by elevated levels of 53BP1 and MN (Fig 3.16A + B). This suggests that RCCD1 binding is required for the role of JMJD5 in maintaining replication fidelity.
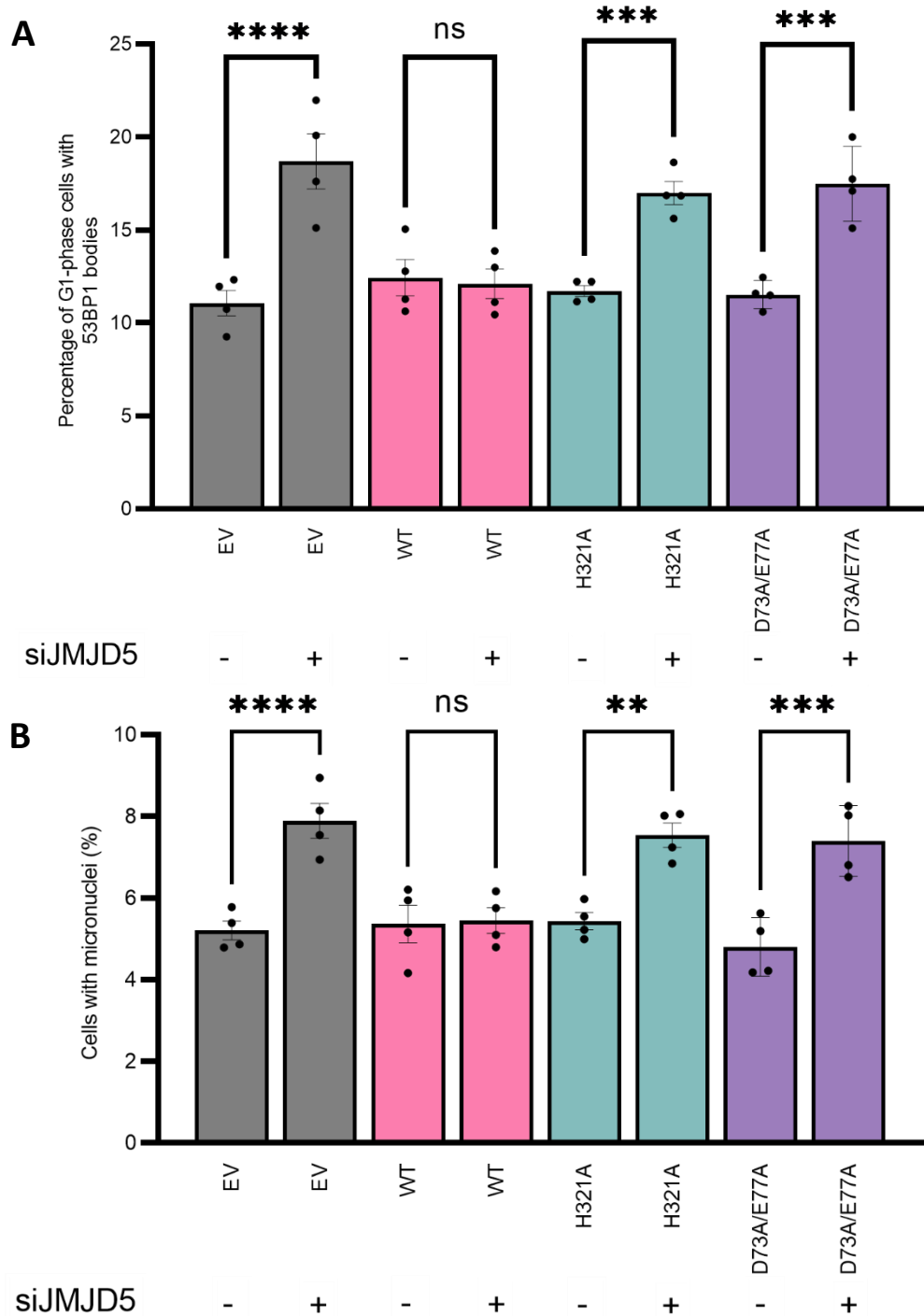
Figure 3.16. RCCD1 binding is required to rescue siJMJD5 induced replication stress.
pTIPZ FLAG-JMJD5 cell lines pTIPZ EV (grey), WT JMJD5 (pink), H321A JMJD5 (blue), or D73A/E77A JMJD5 (purple) were transfected with indicated siRNA for 72 h before the cells were harvested and stained. Cell lines expressing WT JMJD5 but not catalytically dead or RCCD1 binding mutants were able to rescue **(A)** 53BP1 bodies and **(B)** MN. Statistical analysis used One way ANOVA with Bonferroni's post hoc test (with p-values o≤0.01 (**), ≤0.001 (***), ≤0.0001 (****) and ≥0.05 (ns)). Data shown represents the meanof 4 biological repeats -/+ SEM. For micronuclei, a minimum of 500 cells were counted per condition per biological repeat. For 53BP1 bodies, a minimum of 300 cells were counted per condition per biological repeat.

Next, we quantified stalled forks and fork asymmetry using the DNA fibre assay. Because of the complex nature of this assay however, we first considered whether it would be possible to remove some of the conditions to simplify the analysis. For example, each pTIPZ cell line in Figure 3.16 included an siCtrl to determine whether there were any striking differences in basal replication stress between the cell lines following the lentiviral transfection process. Because we did not observe any significant difference between them (Fig 3.16A + B), we concluded that one siCtrl condition (pTIPZ EV siCtrl) would be sufficient to demonstrate the baseline RS.

As expected, pTIPZ EV cells showed an increase in stalled forks after JMJD5 knockdown (Fig 3.17). Whereas JMJD5^{WT} reconstitution (pTIPZ WT) prevented this increase, expression of the H321A or D73A/E77A JMJD5 mutants did not (Fig 3.17), indicating that both catalytic activity and RCCD1 binding are required for the role of JMJD5 in maintaining normal replication fork dynamics. As expected, treatment with HU had an additive effect on all the increases observed (Fig 3.17).
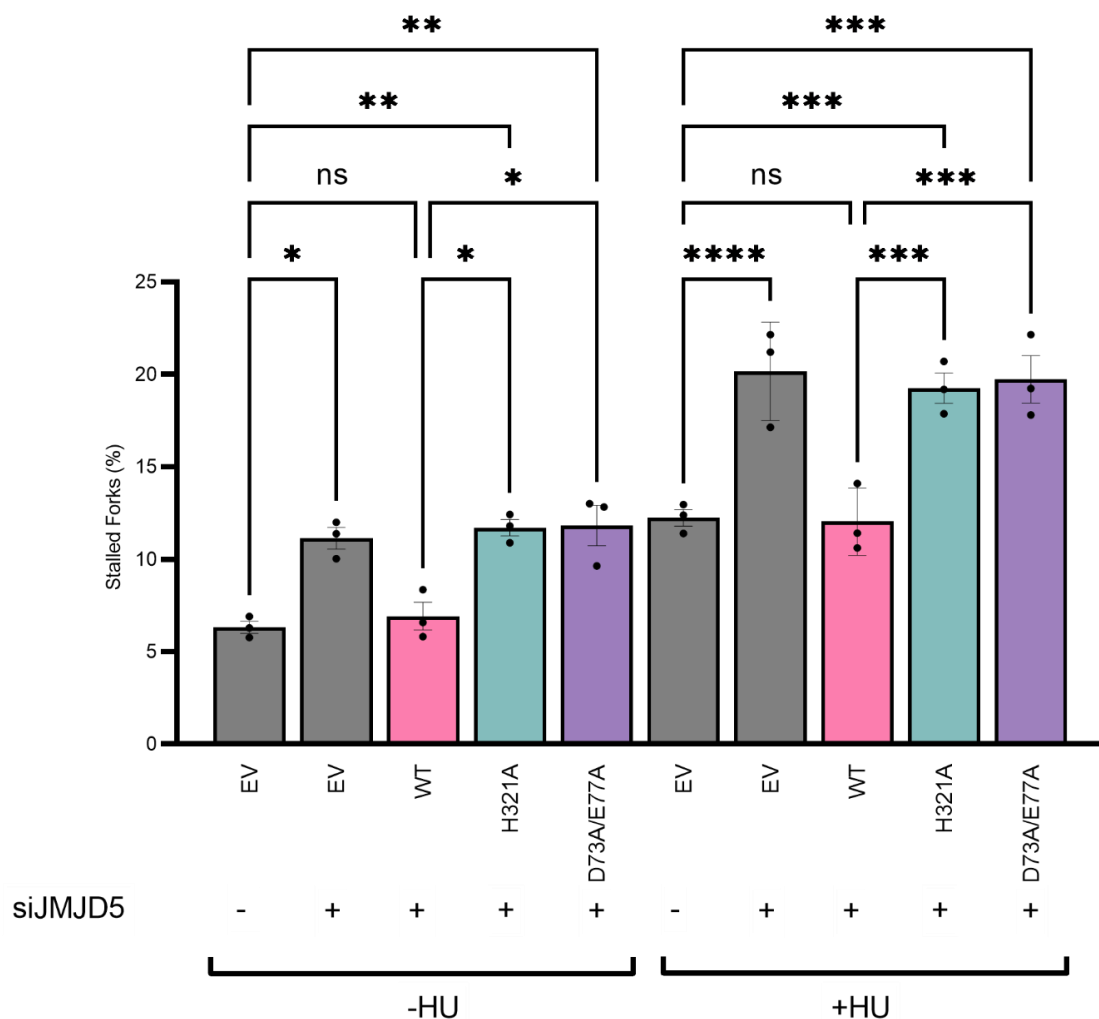
Figure 3.17. RCCD1 binding is required to rescue siJMJD5 induced replication stress.
pTIPZ FLAG-JMJD5 cell lines pTIPZ EV (grey), WT JMJD5 (pink), H321A JMJD5 (blue), or D73A/E77A (purple) were transfected with the indicated siRNA for 72 h before the cells were harvested and DNA fibres were spread. Cell lines expressing WT JMJD5 but not catalytically dead or RCCD1 binding mutants were able to rescue observed increases in stalled forks. Statistical analysis used One way ANOVA with Bonferroni's post hoc test (with p-values of ≤0.05 (*), ≤0.01 (**), ≤0.001 (***), ≤0.0001 (****) and ≥0.05 (ns)). Data shown is the mean of 3 biological repeats -/+ SEM. A minimum of 250 forks were counted per condition per biological repeat.

Following the fork stalling analysis, we measured fork asymmetry, as above. As expected, the analysis showed an increase in fork asymmetry following JMJD5 knockdown in the pTIPZ EV cells (Fig 3.18), that was again exacerbated by HU treatment. Similar to the fork stalling data, expression of JMJD5^{WT} was capable of preventing increased fork asymmetry by endogenous JMJD5 knockdown Fig 3.18), whereas the H321A and D73A/E77A JMJD5 mutants were not.
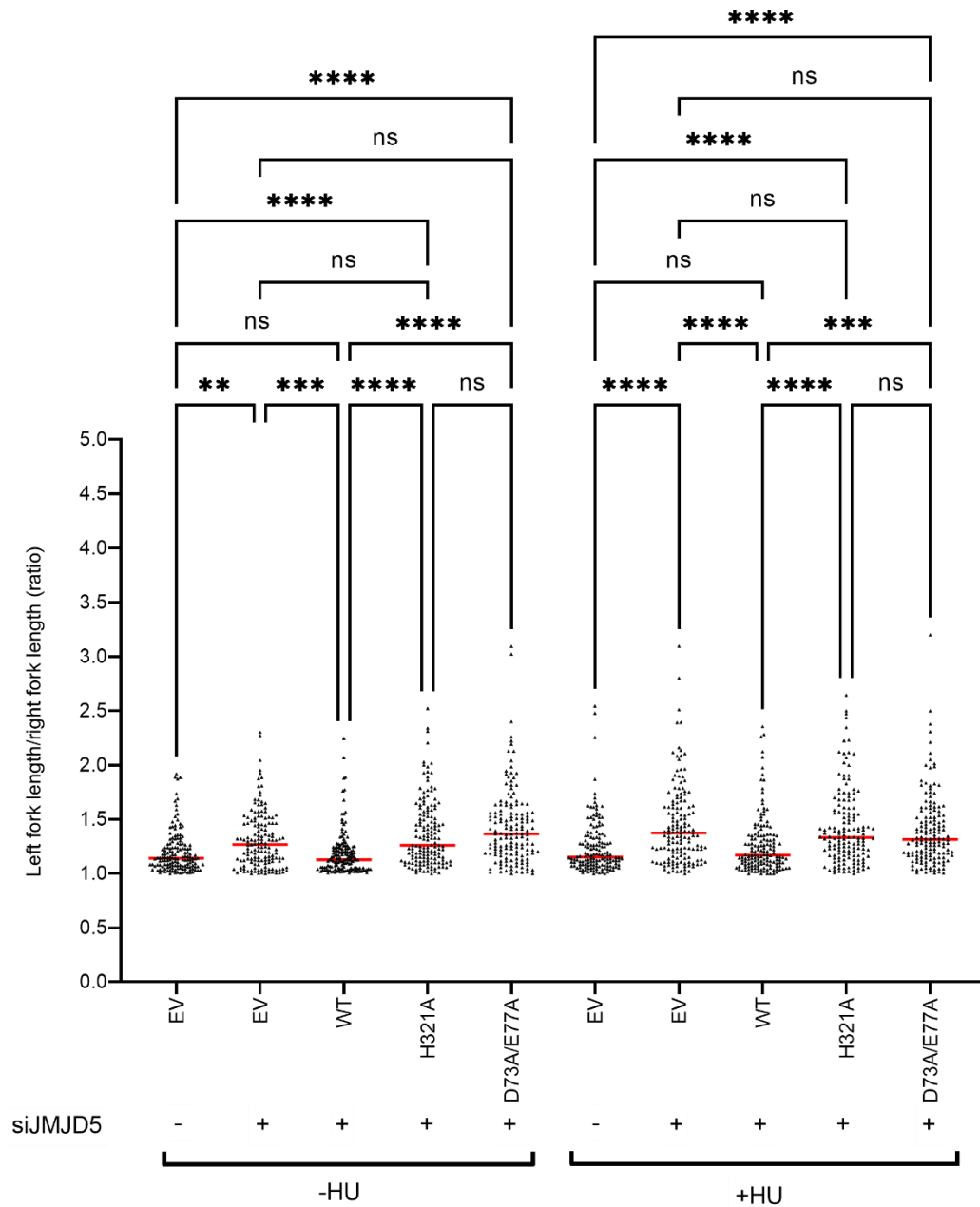
Figure 3.18. Fork asymmetry increases following JMJD5 knockdown are not rescued by JMJD5 D73A/E77A.
A549 pTIPZ cell lines (EV, WT, H321A or D73A/E77A) were transfected with the indicated siRNA and treated with HU where indicated. Fork asymmetry was increased in A549 cells. Statistical analysis used Kruskal-Wallis with Dunn's correction (with p-values of ≤0.01 (**), ≤0.001 (***), ≤0.0001 (****) and ≥0.05 (ns)). The data shown is the average of 3 biological repeats. A minimum of 50 asymmetric forks were counted per condition per repeat.

Taken together, our findings demonstrate that JMJD5$^{D73A/E77A}$ is incapable of rescuing the RS phenotype induced by JMJD5 loss of function. As the D73A/E77A mutation cannot bind RCCD1, we concluded that RCCD1 binding is required for the role of JMJD5 in maintaining replication fidelity. Therefore, we hypothesise that JMJD5 binding is required for the role of RCCD1 in maintaining replication fidelity, which we explore in the next section.

3.2.7   JMJD5 binding is required for the role of RCCD1 in maintaining replication fidelity

Following the discovery that RCCD1 binding is required for the role of JMJD5 in maintaining replication fidelity, we investigated whether the reciprocal was true. Because the cancer variants in RCCD1 which reduced JMJD5 binding showed poor expression (Fig 2.24), we focussed our study on the experimental double methionine RCCD1 mutant (R310M/K328M), which was expressed normally (Fig 2.26).

3.2.7.1   Generation of an RCCD1 reconstitution system

Initially, we attempted to generate an RCCD1 reconstitution system using the same doxycycline-inducible pTIPZ vector used above for JMJD5. However, during the characterisation of those cell lines, we discovered that we could not achieve physiological re-expression levels of RCCD1 using the pTIPZ system (data not shown). Because we had previously used a constitutive lentiviral expression vector (pIHZ) to successfully overexpress RCCD1 in A549 cells (Section 2.2.9), we decided to return to that model here. Unlike JMJD5, where the siRNA targets the 3'UTR, our RCCD1 siRNAs target the coding regions of RCCD1. Therefore, we used site-directed mutagenesis to introduce site-specific silent mutations that would render the RCCD1 cDNAs siRNA-resistant. These cDNAs were then PCR cloned into the pIHZ lentiviral vector with the addition of a C-terminal 3XFLAG tag. After sequence verification, we generated stable pIHZ A549 cells expressing Empty Vector control, or 3XFLAG RCCD1^{WT} or RCCD1$^{R310M/K328M}$. Initial characterisation demonstrated physiological levels of 3XFLAG-RCCD1 expression that were comparable in the RCCD1^{WT} and RCCD1$^{R310M/K328M}$ cell lines (data not shown).

Next, we checked 3XFLAG RCCD1 re-expression levels in the context of siRNA transfection. Figure 3.19 shows a representative Western blot demonstrating successful expression of both RCCD1^{WT} and

RCCD1$^{R310M/K328M}$ in cells transfected with RCCD1 siRNA (Fig 3.19). Despite modest underexpression of

RCCD1$^{R310M/K328M}$ mutant compared to RCCD1^{WT} in this system (compare lanes 3 and 5), the level of

reconstituted 3XFLAG RCCD1$^{R310M/K328M}$ in RCCD1 knockdown cells (lane 6) was similar to the

endogenous RCCD1 levels (lanes 1, 3 and 5) (Fig 3.19). We therefore considered this a viable system

for exploring the role of the RCCD1:JMJD5 complex in replication fidelity.
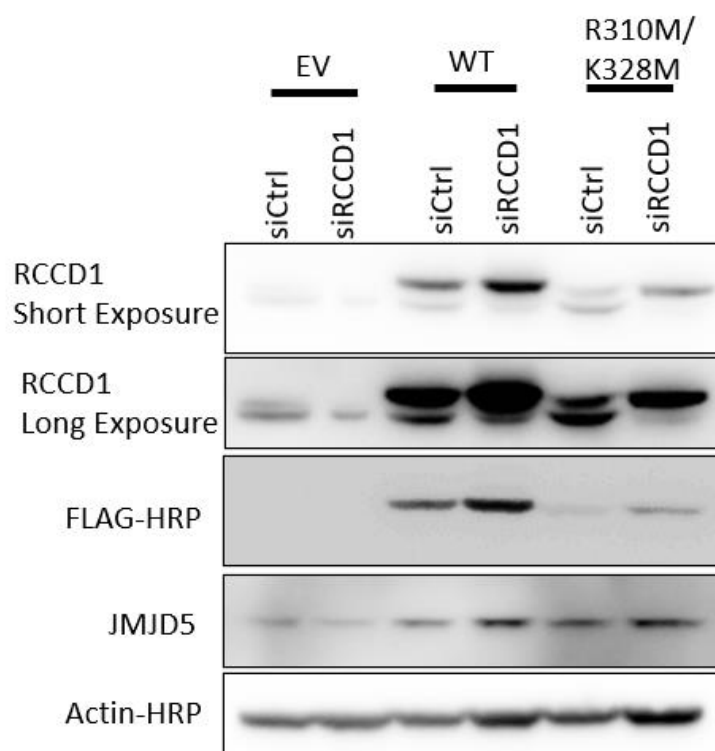
Figure 3.19. Representative Western blot of A549 pIHZ RCCD1-FLAG rescue system.
The indicated pIHZ A549 cells were then transfected with the indicated siRNA for 72 h before
being harvested. Samples were then Western blotted for FLAG, JMJD5, RCCD1 and Actin as a
loading control.

3.2.7.2 The RCCD1 R310M/K328M mutant is unable to suppress MN and 53BP1 bodies levels induced by endogenous RCCD1 knockdown

Having validated the RCCD1 reconstitution system, we used it to explore the response of indirect RS

markers to RCCD1 knockdown. As expected, RCCD1 siRNA caused a statistically significant increase in

both 53BP1 bodies (Fig 3.20A) and MN (Fig 3.20B) in pIHZ EV cells (grey bars). Reconstitution with

3XFLAG RCCD1^{WT} (pIHZ WT) prevented the increase in both markers by RCCD1 knockdown (Fig 3.20A

+ B), which demonstrates that the system is functional, and proves that the RS phenotype is 'on

target'. Interestingly, reconstitution with the 3XFLAG RCCD1$^{R310M/K328M}$ mutant failed to suppress 53BP1 bodies (Fig 3.20A) and MN (Fig 3.20B) under the same conditions. This suggests that JMJD5 binding is important for the role of RCCD1 in maintaining replication fidelity.
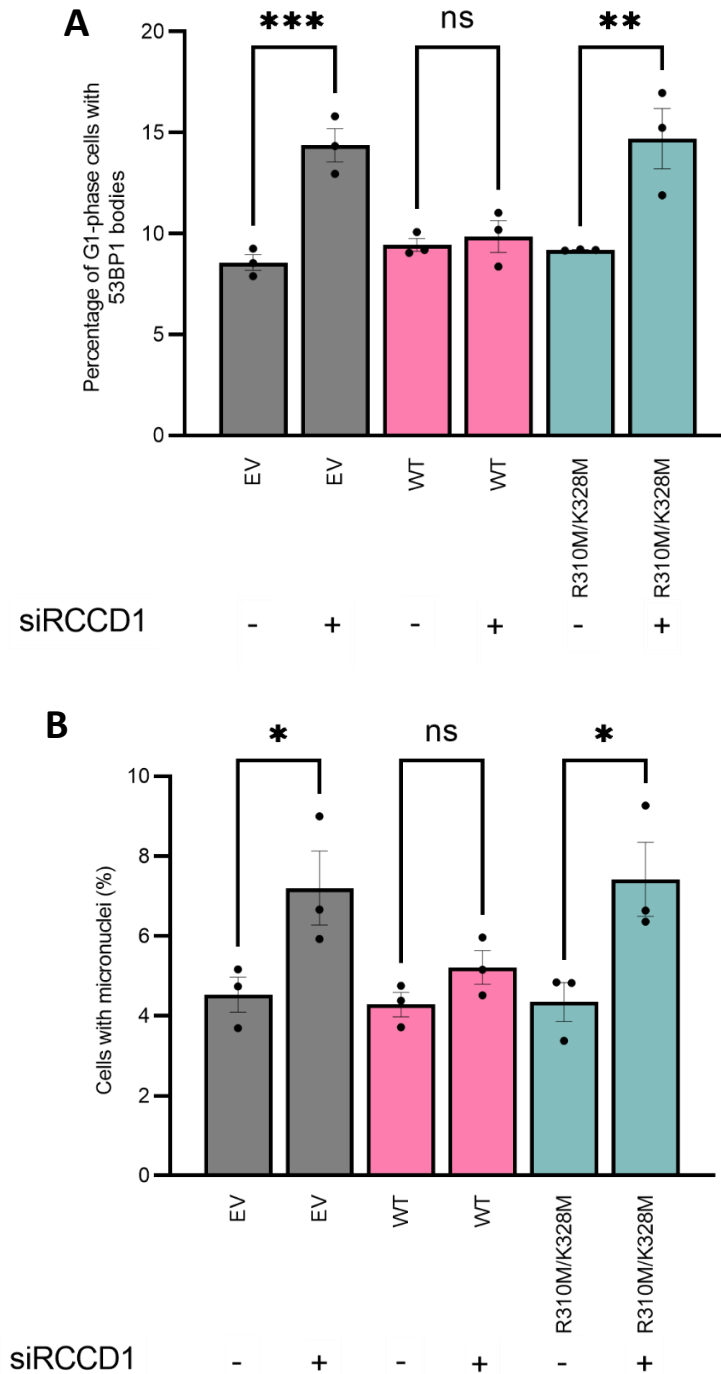
Figure 3.20. JMJD5 binding is required for 3XFLAG RCCD1 to rescue siRCCD1-induced replication stress.

pIHZ EV (grey), WT RCCD1 (pink), or R310M/K328M RCCD1 (blue) cell lines were transfected with the indicated siRNA for 72h before the cells were harvested and stained. Cell lines expressing WT 3XFLAG RCCD1, but not the R310M/K328M JMJD5 binding mutant, were able to rescue **(A)** 53BP1 bodies and **(B)** MN. Statistical analysis used One way ANOVA with Bonferroni's post hoc test (with p-values of ≤0.05 (*), ≤0.01 (**), ≤0.001 (***) and ≥0.05 (ns)). Data shown is the mean of 3 biological repeats -/+ SEM. For micronuclei, a minimum of 500 cells were counted per condition per biological repeat. For 53BP1 bodies, a minimum of 300 cells were counted per condition per biological repeat.

3.2.7.3 The RCCD1 R310M/K328M mutant is unable to suppress replication fork abnormalities induced by endogenous RCCD1 knockdown

Following the observations above, we used the DNA fibre assay to quantify stalled forks and fork asymmetry in the RCCD1 reconstitution model. As expected, the pIHZ EV cell line showed a statistically significant increase in stalled forks following siRCCD1 transfection (with and without HU treatment) (Fig 3.21). Expression of 3XFLAG RCCD1^{WT} suppressed any increase in stalled forks caused by RCCD1 siRNA, both in HU-treated and -untreated cells (Fig 3.21). RCCD1$^{R310M/K328M}$ mutant failed to suppress the stalled forks caused by siRCCD1 transfection however, indicating that this JMJD5 binding mutant was unable to rescue the stalled fork phenotype (Fig 3.21).
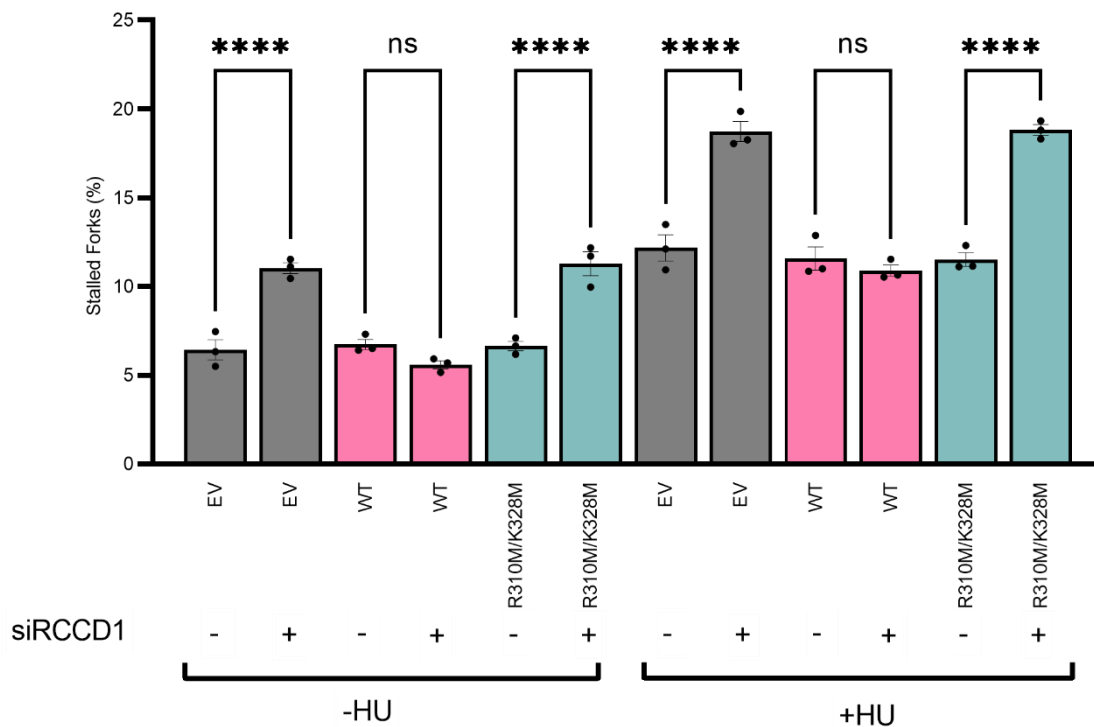
Figure 3.21. JMJD5 binding is required for 3XFLAG RCCD1 to rescue siRCCD1-induced increases in stalled forks.
pIHZ EV (grey), WT RCCD1 (pink), or R310M/K328M RCCD1 (blue) were transfected with the indicated siRNAs for 72 h before the cells were harvested and DNA fibres were spread. Cell lines expressing WT 3XFLAG RCCD1, but not the R310M/K328M JMJD5 binding mutant, were able to rescue the observed increase in stalled forks caused by RCCD1 siRNA. Statistical analysis used One way ANOVA with Bonferroni's post hoc test (with p-values of ≤0.0001 (****) and ≥0.05 (ns)). Data shown represents the mean of 3 biological repeats -/+ SEM. A minimum of 250 forks were counted per sample per biological repeat.

Following the analysis above, we next quantified fork asymmetry as above. As expected, the knockdown of RCCD1 increased fork asymmetry (Fig 3.22), which was further exacerbated by addition of HU. Additionally, as with the markers above expression of RCCD1^{WT} could rescue this observed increase, but the RCCD1 binding mutant could not (Fig 3.22).
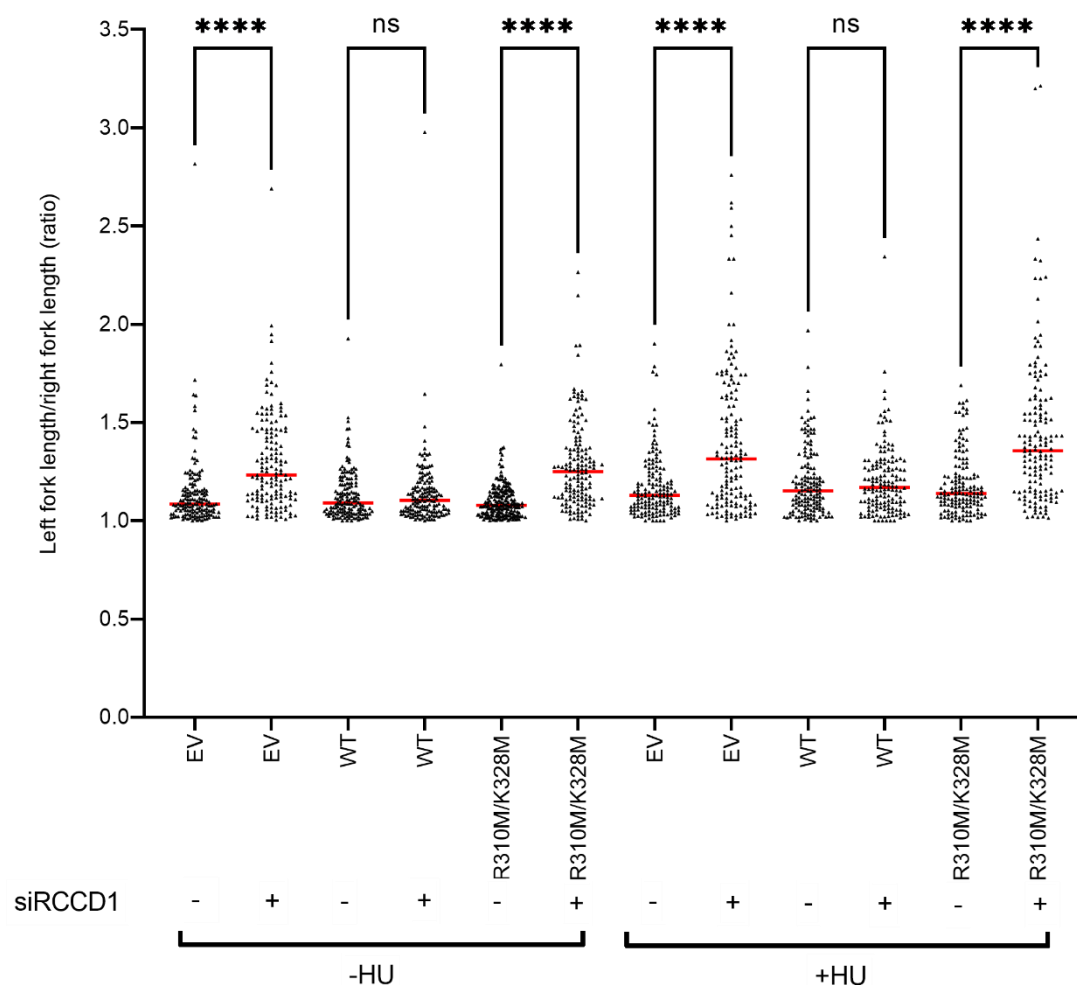
Figure 3.22. Fork asymmetry increases following RCCD1 knockdown are not rescued by RCCD1 R320M/K328M.
A549 pIHZ cell lines (EV, WT or R310M/K328M) were transfected with the indicated siRNA and treated with HU where indicated. Fork asymmetry was increased in A549 cells. Statistical analysis used Kruskal-Wallis with Dunn's correction (with p-values of ≤0.0001 (****) and ≥0.05 (ns)). The data shown is the average of 3 biological repeats. A minimum of 50 asymmetric forks were counted per condition per repeat.

3.3 Discussion

Here in this chapter, we have demonstrated a role for RCCD1 in maintaining replication fidelity that is epistatic with that of JMJD5 and, through reconstitution experiments using binding deficient mutants, shown the importance of the JMJD5:RCCD1 interaction.

3.3.1 Are the roles of JMJD5 and RCCD1 in replication fidelity direct, or does it relate to regulation of an associated cellular process?

Although it is clear from our published work on JMJD5, and the results presented in this chapter, that JMJD5, RCCD1, and their interaction, are required for maintenance of replication fidelity, the mechanism involved remains unclear. The purported roles of JMJD5 and RCCD1 in chromosome segregation and mitotic spindle stability (Marcon et al., 2014; He et al., 2016; Wu et al., 2017) raised the possibility that some of the markers analysed, specifically MN, may have arisen from a mitotic defect. However, we showed that MN induced by RCCD1 knockdown are unlikely to arise from a mitotic defect and are more consistent with replication stress during S-phase (Fig 3.10). This raises the possibility that the reported mitotic phenotype could actually be a consequence of a defect in replication fidelity and unresolved DNA damage following JMJD5 and RCCD1 inhibition. Indeed, prolonged replication stress has been correlated with centrosome amplification, which in turn can result in multipolar spindles and chromosome mis-segregation (Wilhelm et al., 2020). Marcon et al. reported a role for JMJD5 and RCCD1 in increasing the number of multipolar spindles, leading to chromosome segregation defects (Marcon et al., 2014). Analyses of centrosomes and multipolar spindles could be considered in future work using the epistasis and reconstitution approaches described here.

Depletion of JMJD5 and RCCD1 has also been reported to increase the stability of microtubules (He et al., 2016; Wu et al., 2017). Whether this is a direct effect or an indirect consequence of replication stress is unclear however, because replication stress has been shown to stabilise microtubules (Wilhelm et al., 2019). JMJD5 has also been implicated in DNA repair processes. For example, JMJD5

has been suggested to contribute to mismatch repair (Suzuki et al., 2006). Furthermore, Amendola et al reported a role for JMJD5 in the late stages of homologous recombination (HR) (Amendola et al., 2017). HR is now understood to be a vital process in the protection, repair and restart of stalled replication forks (Ait Saada et al., 2018). Therefore, it is possible that JMJD5 may contribute to the replication stress response through a role in HR, and future work could attempt to explore this hypothesis (Discussed further in Chapter 4). Whether RCCD1 might also play a role in DNA repair processes remains unclear but would also be an interesting line of future investigation.

Whether through a direct effect on replication, or through an associated process such as HR, the molecular mechanism by which JMJD5:RCCD1 regulates replication fidelity requires further investigation. Key knowledge gaps include identification of the relevant JMJD5 substrate and other JMJD5 and RCCD1 interacting proteins. These issues are investigated in Chapter 4 and discussed further in Chapter 5.

3.3.2   Do JMJD5 and RCCD1 regulate each other?

One of the interesting observations made in this chapter was the potential for mutual regulation of JMJD5 and RCCD1 expression by one another. In A549 cells, we reproducibly observed a decrease in RCCD1 following JMJD5 knockdown and a more subtle decrease in JMJD5 following RCCD1 depletion (Fig 3.11A). The decrease in JMJD5 and RCCD1 was more variable in U2OS but was still apparent (Fig 3.11B). This relationship was also observed with our reconstitution system: JMJD5 knockdown caused a clear decrease RCCD1, and in the RCCD1 rescue system, RCCD1 depletion caused a subtle reduction in JMJD5 expression (Fig 3.15 and Fig 3.19).

These observations come with some caveats however. Firstly, both JMJD5 and RCCD1 are expressed at very low levels and are difficult to detect, likely due in part to the antibodies for these proteins being of low sensitivity, which can make the results and their interpretation variable. Secondly, it is not yet clear whether these regulator effects are 'on target'. For example, reconstitution with 3XFLAG-JMJD5 may not rescue the reduction in endogenous RCCD1 expression by JMJD5 siRNA (Fig 3.15).

Whether this is indicative of an 'off target' effect of the siRNAs used or is related to the timepoint analysed or some other variable of the experimental conditions unrelated to siRNA specificity, is not yet clear. If our results are explained by co-regulation by JMJD5 and RCCD1, the mechanism(s) involved also require further investigation. Whether they might be related to direct effects on protein stability or some level of gene expression control are unknown. The potential for JMJD5 and RCCD1 post-transcriptional co-regulation will be discussed in more detail in Chapter 4.

3.3.3   The physiological and pathological importance of the RCCD1:JMJD5 interaction

In this chapter, we have demonstrated, using mutants identified and characterised in Chapter 2, that the JMJD5:RCCD1 interaction is essential for the role of each protein in maintaining replication fidelity (Results 3.2.6 and 3.2.7). Because the binding residues were originally discovered using cancer variants the work raises the possibility that the JMJD5:RCCD1 interaction and its role in replication fidelity is important in tumourigenesis. Our data would suggest that cancer variants of JMJD5 and RCCD1 could inactivate their interaction and function, causing enhanced replication stress and genomic instability, both of which are now recognised as hallmarks of cancer (Hanahan and Weinberg, 2011);(Gaillard et al., 2015). Future work could explore whether other cancer variants impact the JMJD5:RCCD1 interaction, and the impact of this on tumourigenesis in cancer models such as xenografts.

Altered replication fidelity and genome stability play important roles in diseases beyond cancer. Indeed, we have demonstrated a role for JMJD5 inactivation and replication stress in a novel human neurodevelopmental disorder (Fletcher et al., 2023). The work presented in this chapter would predict that related neurodevelopmental disorders could be explained by inactivating variants in RCCD1. Interestingly, The Coleman group recently had a RCCD1 match on 'GeneMatcher', an online portal for connecting clinicians and researchers working on rare disease, where the clinical phenotypes of the patients showed some overlap with those of the JMJD5 patients (Mat Coleman – personal communication). Investigating cells derived from these patients, or the associated missense variants

using the RCCD1 reconstitution model described here, particularly in relation to replication stress, would be an interesting line of future investigation.

3.3.4   Chapter Conclusions

In this chapter, we demonstrated the role of RCCD1 in replication fidelity and shown that this role is dependent on its interaction with JMJD5, and vice versa. Through detailed analysis, we have shown that this is associated with roles for the complex in replication fidelity. However, the specific mechanisms involved remained elusive. In the next chapter, we aimed to begin addressing this by investigating the JMJD5:RCCD1 interactome in more detail.

Chapter 4:   JMJD5 and RCCD1 form a stoichiometric 1:1 complex involved in replication fork restart

4.1   Introduction

In the previous chapters of this thesis, we identified sequence determinants of the RCCD1:JMJD5 interaction and demonstrated a role for the complex in replication stress fidelity. The work raised questions about the molecular mechanisms involved, which we aimed to explore further here in this chapter through combined structural, proteomic, and phenotypic analyses. This chapter describes our structural biology efforts and the development of a JMJD5:RCCD1 co-overexpression system, which we apply in proteomic screens that suggest a potential role of the complex in replication fork restart.

4.2   Results

4.2.1   JMJD5 and RCCD1 form a stoichiometric 1:1 complex

The Coleman group use unbiased proteomic screens as a powerful approach for identifying hydroxylase interactomes that can support the discovery of biological functions and elucidate the molecular mechanisms involved. Although the Coleman group have performed several such screens for JMJD5, these have not yet delivered high confidence substrate candidates that could explain the role of JMJD5 in replication fidelity. Similar screens using RCCD1 as the bait have not yet been undertaken but could prove useful as an alternative strategy.

Prior to large scale RCCD1 proteomic screens and mass spectrometry (MS) analyses, we tested whether exogenous RCCD1 complexes could be immunopurified on a smaller scale. We transfected two 15 cm plates of HEK293T cells with pcDNA3-EV or pcDNA3-RCCD1-HA, followed by anti-HA immunoprecipitation, SDS-PAGE, and Coomassie staining (Fig 4.1A). Unfortunately, this staining showed low levels of immunopurified exogenous RCCD1 protein in our immunoprecipitated samples (compared to similar projects; Mat Coleman, personal communication). In addition to RCCD1-HA, we

identified several unknown protein bands of different molecular weights (Fig 4.1A). Because these bands were not detected in the control immunoprecipitation, they could be indicative of specific RCCD1-interacting proteins. We excised the indicated segments from the polyacrylamide gel and sent them for digestion and liquid chromatography MS/MS (LC-MS/MS) identification (Fig 4.1B). This analysis confirmed the interaction between exogenous RCCD1 and endogenous JMJD5 (identified in unknown 3 (UNK3) band) and identified potentially novel interactors with RCCD1, including 'Sperm associated antigen 9' (SPAG9), BTB/POZ domain-containing protein KCTD9 (KCTD9), and SLC25A3-6 proteins. Although it is not clear how the candidate interactors from this pilot experiment could explain JMJD5-related functions, particularly in replication fidelity, they will be discussed in more detail later.
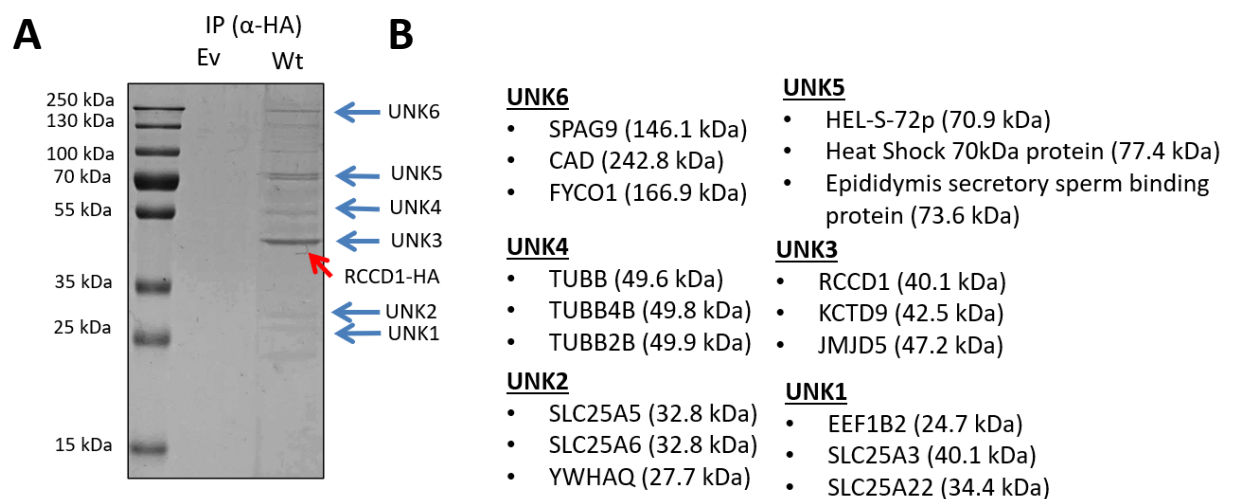
Figure 4.1. Interactors of exogenous RCCD1 in HEK293T cells.
(A) Coomassie gel following pcDNA3-EV or pcDNA3-RCCD1-HA transfection and anti-HA immunoprecipitation. Bands indicated were excised and sent for liquid chromatography and MS/MS analysis (UNK = Unknown). **(B)** Results of MS analysis showing the three most abundant proteins in each band.

One of the most abundant interactors was identified as a molecular chaperone called Heat shock 70 kDa protein (HSP70; identified in band UNK5) (Fig 4.1B), which binds to misfolded proteins (Mayer and Bukau, 2005). The HSP70 interaction might suggest that exogenous RCCD1 is not appropriately folded, which could perhaps explain by its lower-than-expected expression in this system. We considered the possibility that RCCD1 folding, and stability might require its interaction with a binding partner, which

would not have been exogenously co-expressed in this transient transfection experiment. Because

siRNA depletion of JMJD5 can reduce RCCD1 expression (Section 3.2.5), we hypothesised that JMJD5

could represent such an interactor. To test this, we transfected pIHZ-RCCD1-FLAG, with or without

pEF6 HA-JMJD5, into HEK293T cells, followed by anti-FLAG immunoprecipitation, SDS-PAGE, and

Coomassie staining. Strikingly, when the two proteins were co-overexpressed, we observed an

increase in the expression of RCCD1-FLAG compared with the expression of RCCD1-FLAG alone (Fig

4.2). Indeed, Western blot analysis of the input samples confirmed increased RCCD1-FLAG

overexpression following HA-JMJD5 transfection (Fig 4.2). Additionally, the UNK5/HSP70 double band

appeared to be reduced after HA-JMJD5 co-overexpression with RCCD1-FLAG (Fig 4.2).
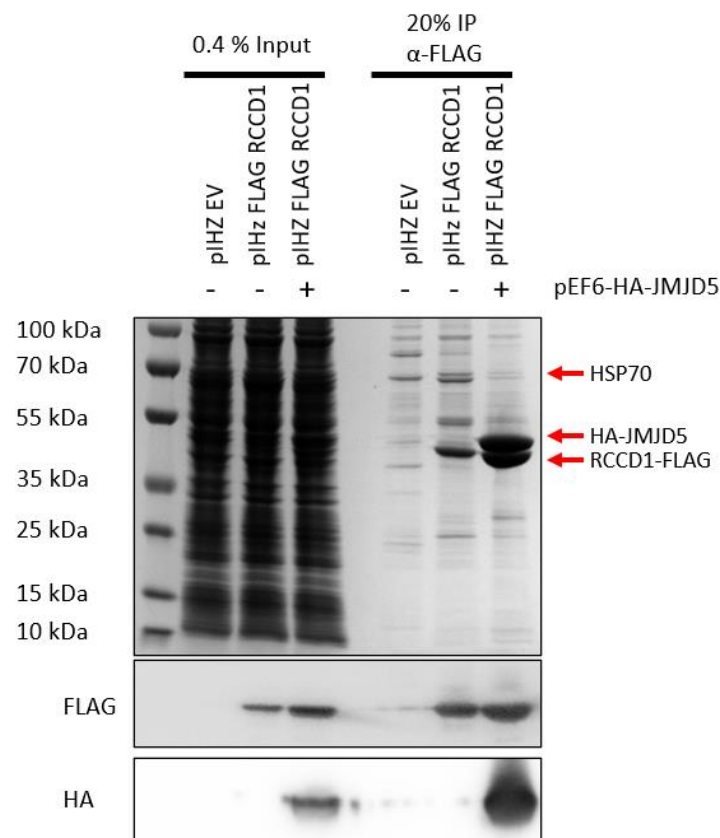
Figure 4.2. HA-JMJD5 expression increases RCCD1-FLAG expression.
HEK293T were transfected with indicated constructs followed by anti-FLAG IP, SDS-PAGE,
Coomassie staining and Western blot analysis for HA and FLAG. Inputs are indicative of cell lysate
prior to IP reaction.

Overall, these data suggest that JMJD5 may be able to promote the expression of RCCD1. We were

intrigued by the possibility that the reciprocal might also be true, i.e. that RCCD1 might be required

117

for optimal JMJD5 protein expression. Whilst addressing this hypothesis we aimed to further refine our experimental approach towards large scale purification of soluble complexes. To this end we tested purification of tagged-RCCD1 or -JMJD5, in the presence or absence of the other, followed by peptide elution of complexes, and Coomassie staining of SDS-PAGE gels (Fig 4.3). Consistent with the previous results (Fig 4.2), expression of HA-JMJD5 significantly increased the expression of RCCD1-FLAG (Fig 4.3A) (compare right 3 lanes to middle 3 lanes). Interestingly, and consistent with the hypothesis, the reverse was also true: expression of RCCD1-FLAG significantly increased the expression of HA-JMJD5 (Fig 4.3B) (compare right 3 lanes to middle 3 lanes).

Interestingly, in both FLAG (Fig 4.3A) and HA (Fig 4.3B) immunoprecipitations, the amounts of eluted JMJD5 and RCCD1 appeared equimolar across each elution, demonstrating a stoichiometric 1:1 relationship (Fig 4.3). This relationship was confirmed by analytical gel filtration, where a major peak was observed at approximately 90 kDa (Chan Li - data not shown), consistent with a 1:1 heterodimer of RCCD1 (40 kDa) and JMJD5 (47 kDa).

Overall, these data suggest that JMJD5 and RCCD1 may be 'obligate' binding partners that co-exist in a heterodimeric 1:1 complex, one function of which may be to promote optimal protein stability, folding, and/or expression.
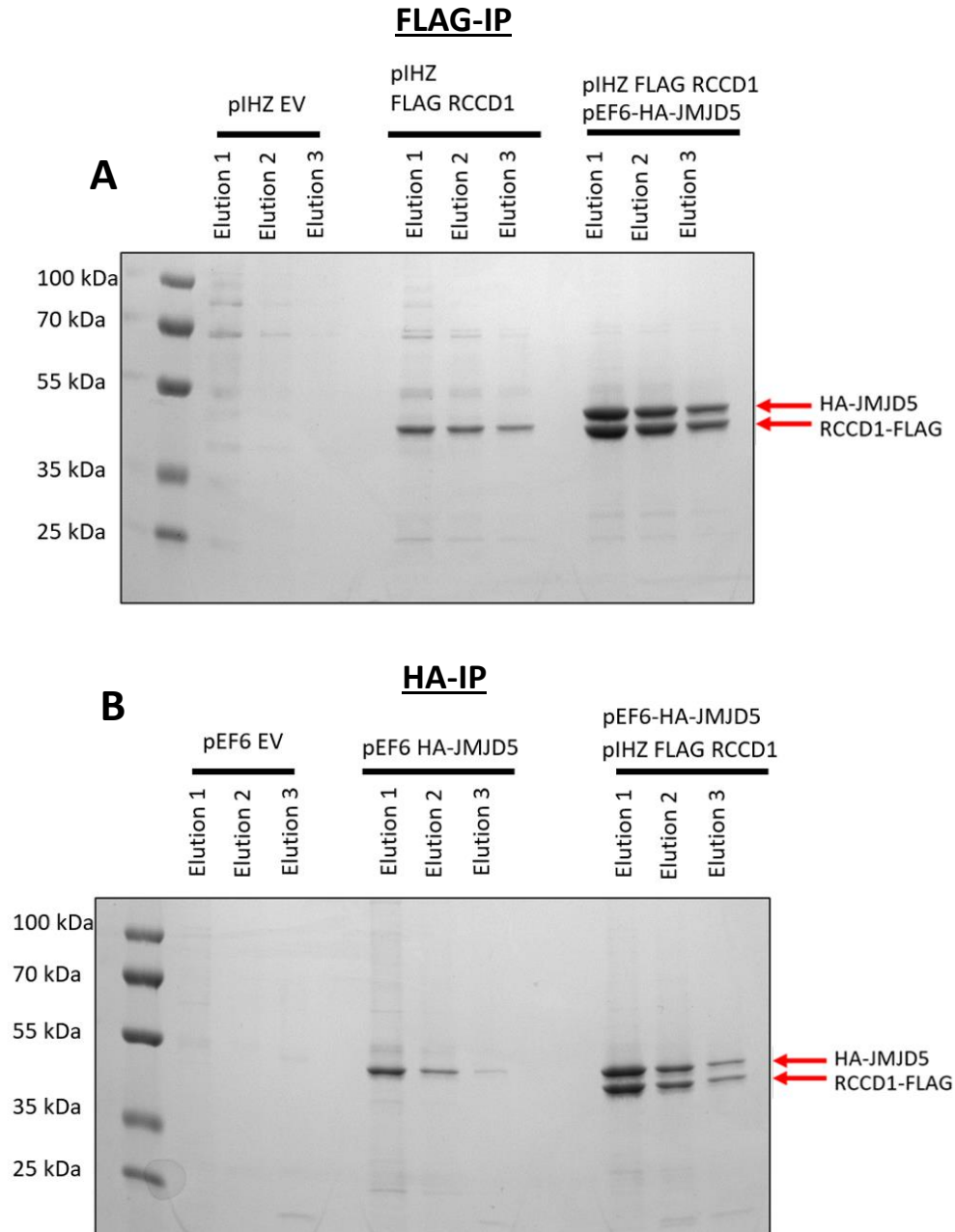
Figure 4.3. Overexpressed HA-JMJD5 and RCCD1-FLAG form a stoichiometric 1:1 complex.
HEK293T cells transfected with indicated pIHZ (EV or RCCD1-FLAG) or pEF6 (EV or HA-JMJD5) followed by anti-FLAG IP **(A)** or anti-HA IP **(B)**, SDS-PAGE and Coomassie staining.

4.2.2 Design and validation of a P2A co-overexpression system

The results presented in the previous section suggested that it might be possible to purify soluble 1:1 JMJD5:RCCD1 complexes from mammalian cells for structural biology, biophysical, and proteomic studies. However, because transient transfection methods can be variable and very expensive, we sought a more cost-effective and scalable approach. In doing so, we considered developing a system

that could simultaneously support equal co-expression of two proteins. 2A sequences are 'self-cleaving' viral peptides that can be used to generate two separate proteins from a single nucleotide sequence (Liu et al.2017) (Fig 4.4). For our system, we used the porcine teschovirus-1 2A (P2A) sequence. A multicistronic vector containing a P2A sequence is transcribed as a single mRNA before the P2A sequence is translated into two separate proteins via ribosome skipping between the final glycine and proline residues of the P2A sequence (Fig 4.4). To utilise this system, we had a nucleotide sequence encoding HA-JMJD5 and RCCD1-FLAG, separated by a P2A sequence (Fig 4.4), commercially synthesised. The JMJD5 inactivating mutation (H321A) was subsequently introduced into the vector via site-directed mutagenesis. To support the generation of stable cell lines we cloned the relevant sequences into the doxycycline inducible pTIPZ vector, to allow lentiviral infection and selection of stable HEK293T cells.
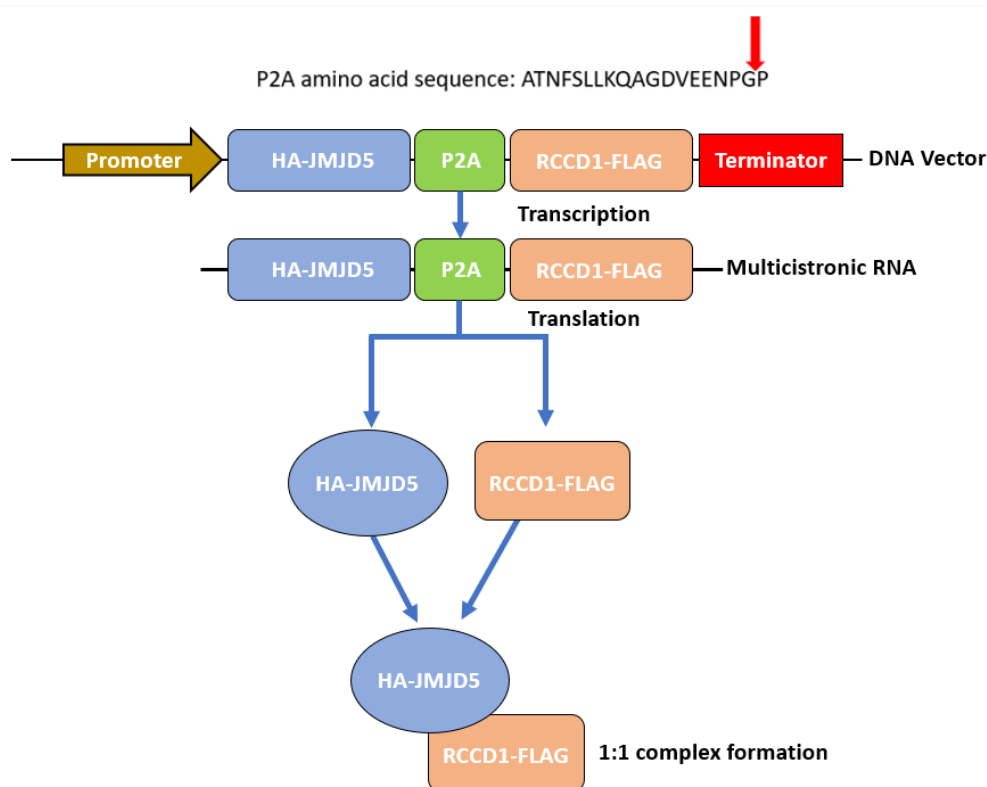
Figure 4.4. P2A multicistronic vector generates HA-JMJD5 and RCCD1-FLAG proteins from a single mRNA. HA-JMJD5 and RCCD1-FLAG sequences separated by a porcine adenovirus (P2A) were cloned into a pTIPZ doxycycline inducible lentiviral expression vector. This DNA sequence would be transcribed into a single multi-cistronic RNA, which would then be translated into two separate proteins. Following translation, we hypothesised the proteins would form a 1:1 heterodimer as seen above.

After generation of control (pTIPZ EV), HA-JMJD5(WT)/RCCD1-FLAG, and HA-JMJD5(H321A)/RCCD1-FLAG inducible cell lines, we validated the efficacy of the P2A sequence and the ability to form the 1:1 complex. Two 15 cm dishes of each HEK293T cell line (EV, WT, and H321A) were seeded, followed by incubation with 1 µg/ml of doxycycline for 48 h before anti-FLAG immunoprecipitation, SDS-PAGE, and Coomassie staining. This validation showed a high P2A 'cleavage' efficiency and clear 1:1 complex formation for both WT and H321A JMJD5 (Fig 4.5).
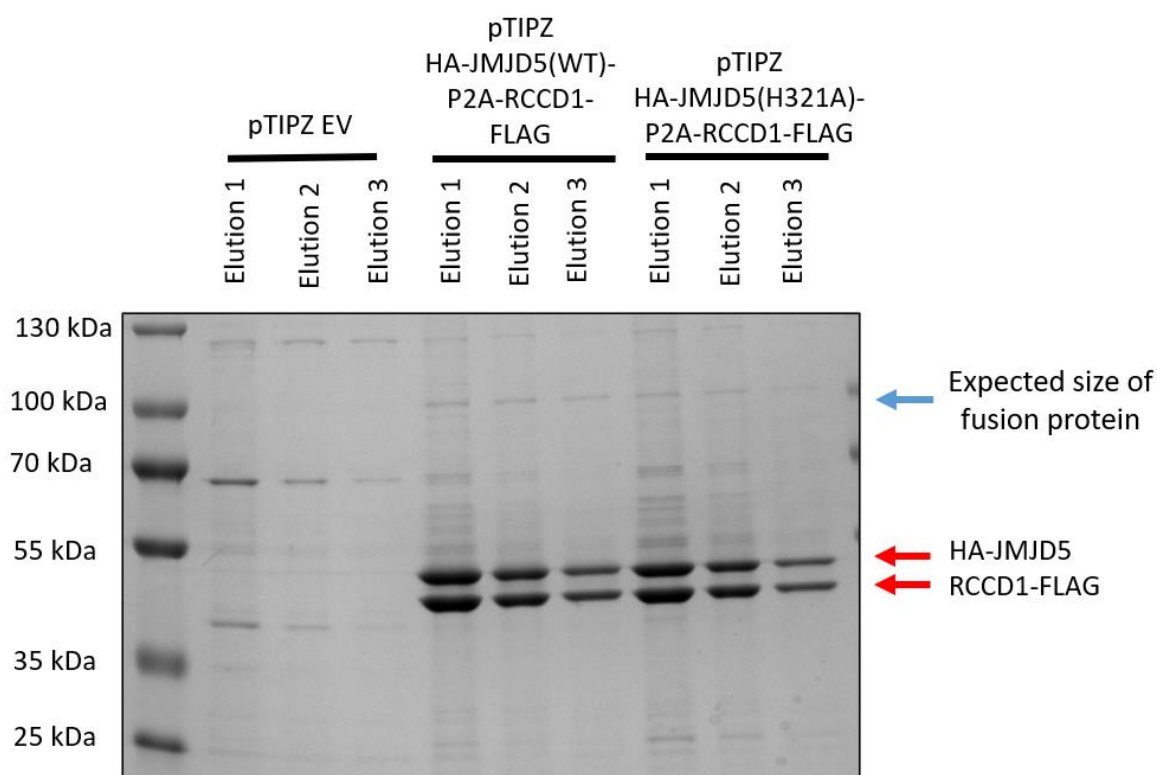
Figure 4.5. pTIPZ HA-JMJD5-P2A-RCCD1-FLAG is highly efficient for 1:1 heterodimer formation. HEK293T pTIPZ EV, HA-JMJD5(WT)/RCCD1-FLAG, and HA-JMJD5(H321A)/RCCD1-FLAG cells were treated with 1 µg doxycycline for 48 h before harvesting and anti-FLAG immunoprecipitation, SDS-PAGE and Coomassie staining.

Because the P2A system had proved highly efficient, and scalable, we aimed to use it to generate recombinant protein for preliminary structural characterisation (Section 4.2.3) and investigation of the JMJD5:RCCD1 interactome (Section 4.2.4).

4.2.3 Preliminary structural characterisation of the JMJD5:RCCD1 complex

Following the validation of the P2A system as a method for generating recombinant JMJD5:RCCD1, we attempted a larger-scale purification for structural characterisation. The complex was expressed, immunoprecipitated, and released from anti-FLAG beads via peptide elution (as above, 4.2.2). Gel filtration chromatography was then used to further isolate the complex and to confirm the stoichiometric relationship (performed by Dr Chan Li, Fig 4.6A). As with transiently transfected JMJD5:RCCD1 analysed by Coomassie staining (Fig 4.3), the gel filtration identified a major peak consistent in size with a 1:1 heterodimer formation of JMJD5:RCCD1 (Fig 4.6A). This isolated complex was then used for cryo-electron microscopy structural characterisation, by freezing onto cryo-EM grids and imaging using a Titan Krios equipped with a Gatan K3 detector (performed by Dr Eilis Braggington). Images were combined into 2D class averages (Fig 4.6B, left) and subsequently generated into a 3D reconstruction using Relion 4 software (Kimanius et al., 2021) (performed by Professor Stephen Smerdon) (Fig 4.6B, right). This reconstruction was then overlayed with the AlphaFold2 multimer complex, which showed reasonable fit to the density map (Fig 4.6B, right). This suggest that the AlphaFold2 model of the RCCD1:JMJD5 complex is largely correct and validates purification of the JMJD5:RCCD1 from human cells as a viable strategy for higher resolution structural investigations in the future.
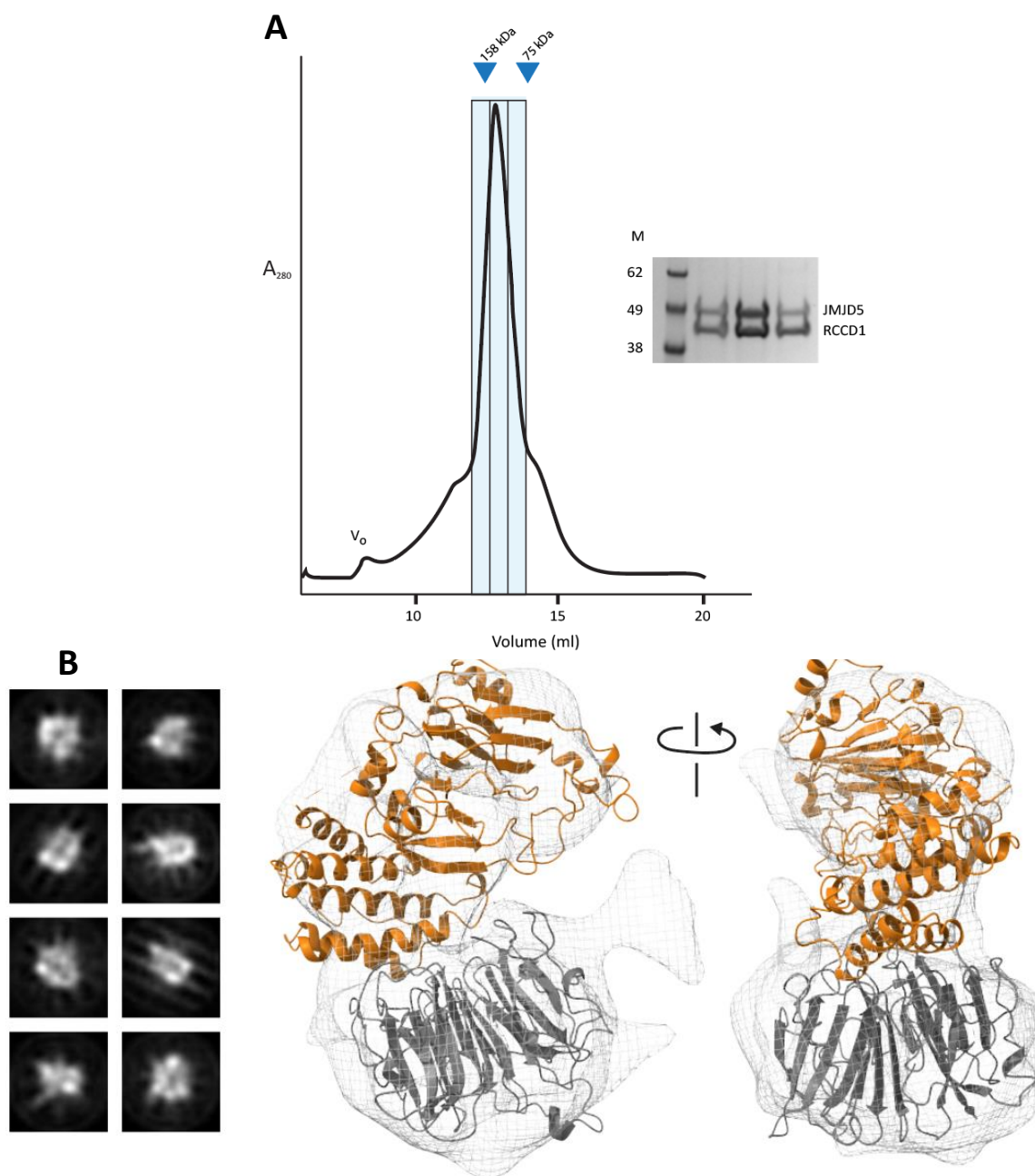
Figure 4.6. Recombinant JMJD5:RCCD1 forms a heterodimer with structural homology to AlphaFold2 prediction.

(A) Gel filtration chromatography of recombinant JMJD5:RCCD1 isolated from HA-JMJD5(WT)/RCCD1-FLAG HEK293T cells. Gel filtration was carried out on a Superdex S-200 Increase 10/300 column. Coomassie gel is of the fractions indicated by the blue markers. The blue arrows indicated the calibration molecular weight markers. **(B)** Cryo-Electron Microscopy 'representative' 2D class averages collected from the apo JMJD5:RCCD1 complex (left). The 2D class averages were combined into a 3D reconstruction (right). Orthogonal views of a 3D reconstruction generated from the 2D classifications were docked with AlphaFold Multimer complex model overlayed. All work here was carried out by the Smerdon group, in particular Dr Chan Li, Dr Eilis Braggington and Professor Stephen Smerdon.

4.2.4 Investigation of the RCCD1:JMJD5 Interactome

The second intended application of the P2A vector was the investigation of the JMJD5:RCCD1 interactome. We reasoned that if JMJD5:RCCD1 exist as a 1:1 heterodimer, then proteomics to identify functionally relevant interactors should apply to the complex rather than the individual subunits. Although we were interested in identifying a variety of different interactors, we had a particular interest in candidate substrates because of the importance of JMJD5 hydroxylase activity in replication fidelity (Fletcher et al., 2023) (Section 3.2.6).

The Coleman group have developed a proteomic workflow to identify candidate substrates of JMJD5-related protein hydroxylases (Fig 4.7A). This strategy is based on two parallel approaches. The first involves substrate blocking, using an iron-binding mutant of the hydroxylase: The inability of the mutant to bind substrates identifies proteins whose interaction is activity-dependent (Fig 4.7A). The second involves treating cells with a pan-hydroxylase inhibitor to trap substrates: Dimethyl-N-oxalylglycine (DMOG) is a non-hydrolysable cell-permeable analogue of 2OG called N-oxalylglycine (NOG) (Jaakkola et al., 2001) (Fig 4.7B). Because the Coleman group have successfully applied this approach to JMJD5-related hydroxylases (Feng et al., 2014; Markolovic et al., 2018), we applied it here to the JMJD5:RCCD1 complex in the hope of identifying candidate JMJD5 substrates and activity-independent interactors.
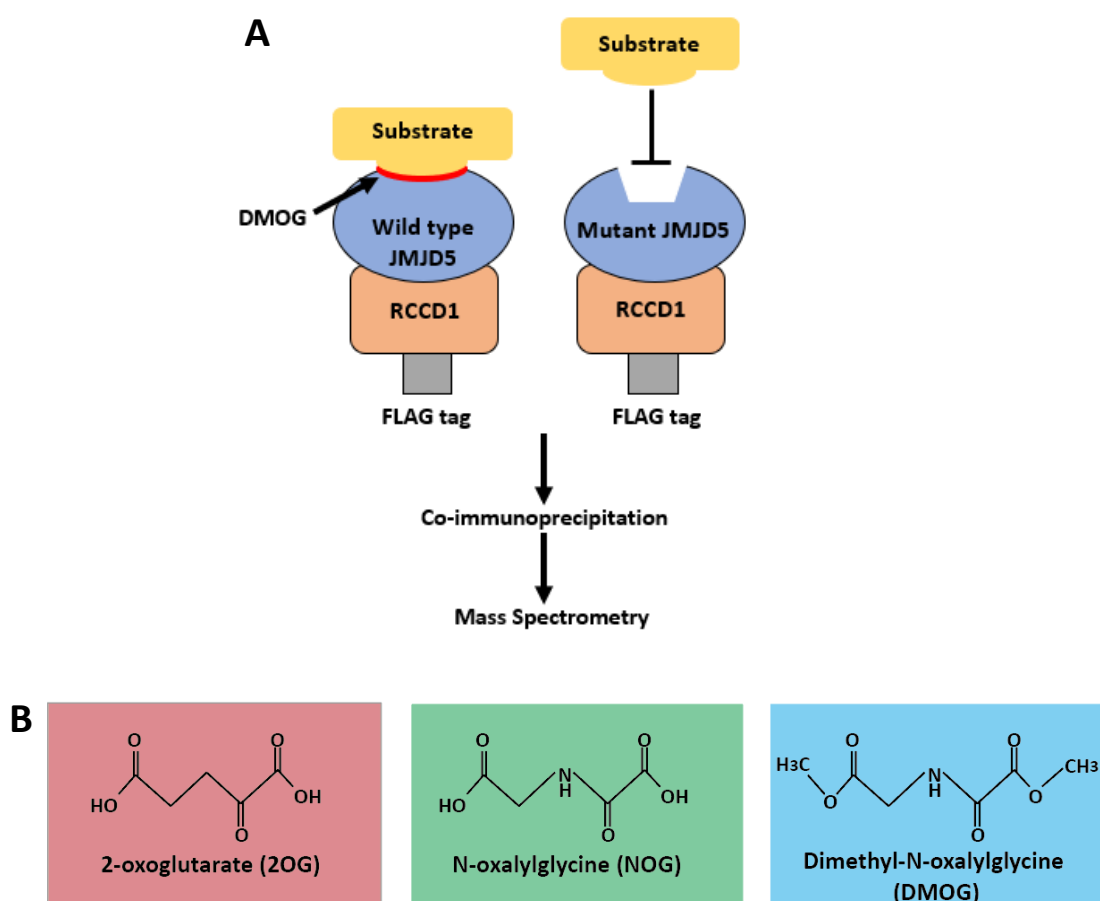
Figure 4.7. Experimental proteomic methodology to identify JMJD5:RCCD1 interactors.
(A) HA-JMJD5(WT)/RCCD1-FLAG and HA-JMJD5(H321A)/RCCD1-FLAG HEK293T cells were treated with doxycycline for 48 h followed by 16 h treatment with dimethyl-N-oxalylglycine (DMOG). DMOG is converted to NOG in cells and can trap candidate substrates within the active site of 2OG oxygenases. The cells are harvested followed by anti-FLAG immunoprecipitation and mass spectrometry identification of interacting proteins. **(B)** Chemical structure of 2OG co-factor and closely related inhibitor analogues N-oxalylglycine (NOG) and Dimethyl-N-oxalylglycine (DMOG).

4.2.4.1 Initial proteomic screen

Control, WT, and H321A JMJD5/RCCD1-FLAG HEK293T cells were treated with doxycycline for 48 h followed by anti-FLAG immunoprecipitation, peptide elution, methanol/chloroform extraction of eluted proteins, and liquid chromatography tandem mass spectrometry (LC-MS/MS) identification. We filtered the full list of identified proteins using a stringent set of criteria, as follows. First, we removed any proteins identified in the control (EV) samples lacking HA-JMJD5:RCCD1-FLAG. We also identified and removed proteins listed in the 'Contaminant Repository for Affinity Purification' (CRAPome) which is a collation of contaminants from proteomics screens (Mellacheruvu et al., 2013).

We used this database to remove contaminants identified in over 50% of the screens and remove them from our sample. We subsequently sorted the remaining proteins present in the WT JMJD5-HA/RCCD1-FLAG sample based on the number of unique peptides: The top 50 proteins are present in Table 4.1. Consistent with the controversy surrounding the histone demethylase assignment of JMJD5 activity (Hsia et al., 2010), no histone proteins were identified as specific JMJD5/RCCD1 interactors in our screen (Table 4.1). Likewise, the RPS6 ribosomal protein, which has also been proposed as a potential JMJD5 substrate, was not identified by our MS/MS analysis (Wilkins et al., 2018). Liu et al. also described an interaction between JMJD5 and RNA pol II, which was also not confirmed in our system (Liu et al., 2020). Earlier we identified SPAG9 and KCTD9 as potential RCCD1 interactors (discussed in Chapter 5), which we did confirm here (Table 4.1).

		WT JMJD5/RCCD1			H321A JMJD5/RCCD1		
Accession	Gene Name	Coverage [%]	# Unique Peptides	Sequest Score	Coverage [%]	# Unique Peptides	Sequest Score
Q8N371	JMJD5	82	44	689.72	82	42	575.58
A6NED2	RCCD1	97	34	1439.72	95	34	1083.72
O60271	SPAG9	20	17	64.82	40	36	167.49
Q7L273	KCTD9	40	10	41.88	57	15	70.73
P22413	ENPP1	18	10	39.08	36	19	78.05
P54819	AK2	55	9	54.43	15	2	0
Q9UJS0	SLC25A13	17	8	25.76	37	12	54.68
P08574	CYC1	35	7	28.68	35	7	34.83
Q9P035	HACD3	23	7	34.3	32	11	46.43
Q7Z6Z7	HUWE1	4	7	24.7	13	33	102.65
P53985	SLC16A1	12	6	29.17	12	5	39.46
Q07065	CKAP4	17	6	31.23	31	12	63.28
P13994	CCDC130	12	4	16.2	4	1	3.66
P22695	UQCRC2	18	4	11.32	21	4	17.08
Q9HCU5	PREB	17	4	13.88	17	4	18.48
Q92616	GCN1	3	4	4.66	4	6	11.55
Q9Y5M8	SRPRB	24	4	6.43	43	8	22.96
Q96CS3	FAF2	17	4	12.65	36	10	33.02
O15067	PFAS	5	3	2.74	0	0	0
P23526	AHCY	7	3	6.18	0	0	0
O75475	PSIP1	12	3	4.28	3	1	4.59
P57088	TMEM33	12	3	6.11	3	1	0
Q14318	FKBP8	13	3	6.26	10	2	10.85
Q9GZT3	SLIRP	33	3	7.28	22	2	6.87
O94766	B3GAT3	11	3	9.04	14	2	1.61
P62834	RAP1A	23	3	13.55	15	3	13.24
Q93008	USP9X	2	3	7.97	2	3	7.23
P36542	ATP5F1C	18	3	7.25	18	3	7.89
Q9Y277	VDAC3	11	3	5.15	11	3	5.15
Q8N428	GALNT16	13	3	11.23	15	3	4.69
Q53GQ0	HSD17B12	15	3	14.25	19	4	17.57
O00264	PGRMC1	39	3	10.23	44	5	16.66
P60891	PRPS1	9	3	7.42	25	5	24.54
P63000	RAC1	17	3	7.14	26	5	7.78
Q9Y6C9	MTCH2	22	3	9.55	35	7	25.69
Q96TA2	YME1L1	6	3	9.86	13	7	22.34
P61619	SEC61A1	10	3	11.55	25	8	19.74
Q96P70	IPO9	6	2	6.93	3	1	6.47
P51114	FXR1	4	2	5.92	2	1	3.25
Q99714	HSD17B10	18	2	7.23	18	2	5.79
P63173	RPL38	36	2	10.62	36	2	11.03
P52815	MRPL12	13	2	3.05	13	2	2.04
Q96T76	MMS19	3	2	4.57	3	2	5.62
O00410	IPO5	4	2	4.98	4	2	4.98
P11310	ACADM	10	2	0	14	3	4.61
P62879	GNB2	14	2	5.38	14	3	6.41
Q86VU5	COMTD1	19	2	7.27	31	3	11.97
Q14739	LBR	6	2	10.81	7	3	12.87
P42704	LRPPRC	2	2	5.11	6	7	16.97
Q9UBX3	SLC25A10	14	2	10.66	36	7	30.21

Table 4.1. List of fifty top interactors of the JMJD5:RCCD1 complex.
Displayed is the accession number and gene name, peptide coverage (%), unique peptides, and Sequest Score. The coverage is the percentage of the protein covered by detected versus predicted tryptic fragments. The unique peptides are the number of peptides identified in each protein that belong solely to that protein in the reference proteome used (Homo Sapiens). The Sequest Score is the cumulative value obtained from scoring and adding each peptide. MS data is from one biological repeat.

To test the validity of our results we wanted to investigate one of the interactions in more detail. SPAG9 encodes 'C-Jun-amino-terminal kinase-interacting protein 4' (Bouwmeester et al., 2004). We attempted to validate this interaction by immunoprecipitation followed by Western blotting. Reassuringly, we were able to confirm a specific interaction of the JMJD5:RCCD1 complex with SPAG9 (Fig 4.8). However, the physiological relevance of this interaction is currently unclear because of the absence of a reported role for Jun kinase signalling in RCCD1 biology (and vice versa). The RCCD1:SPAG9 interaction is discussed further in Section 5.3.
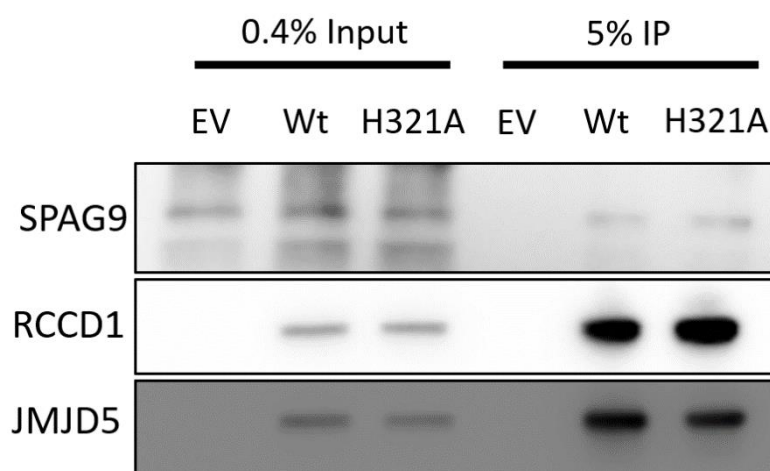
Figure 4.8. The JMJD5:RCCD1 complex interacts with SPAG9.
pTIPZ EV, HA-JMJD5(WT)/RCCD1-FLAG and HA-JMJD5(H321A)/RCCD1-FLAG cells were treated with doxycycline for 48 h and DMOG for 16 h before harvesting and anti-FLAG immunoprecipitation. Input samples are indicative of cell lysate prior to IP. Samples were run on an SDS-PAGE gel and Western blotted with the indicated antibodies.

Next, we further analysed this top 50 group to identify potential activity-dependent JMJD5 interactors: To do so we used the number of unique peptides identified for each protein (a crude indicator of abundance) to calculate the ratio between the WT and H321A HA-JMJD5/RCCD1-FLAG samples. If a protein was not identified in the H321A HA-JMJD5/RCCD1-FLAG sample (and the interaction appeared to be entirely activity-dependent), the 'activity-dependence ratio' was assigned as infinite (∞). We ranked candidates based on activity-dependence ratio and present the top 10 proteins in Table 4.2.

Accession	Gene Name	WT JMJD5/ RCCD1 # Unique Peptides	H321A JMJD5/ RCCD1 # Unique Peptides	Activity Dependence
O15067	PFAS	3	0	∞
P23526	AHCY	3	0	∞
P54819	AK2	9	2	4.5
P13994	CCDC130	4	1	4
O75475	PSIP1	3	1	3
P57088	TMEM33	3	1	3
Q96P70	IPO9	2	1	2
P51114	FXR1	2	1	2
Q14318	FKBP8	3	2	1.5
Q9GZT3	SLIRP	3	2	1.5

Table 4.2. List of top ten activity dependent interactors.
Displayed is the accession number and gene name of the identified protein, the unique peptides, and the activity-dependence, which was expressed as the ratio of unique peptides in the WT to H321A sample. For samples where no protein was detected in the H321A the activity-dependence ratio is indicated by the infinity symbol. The unique peptides are the number of peptides identified in each protein that belong solely to that protein in the reference proteome used (Homo Sapiens).

Next, STRING analysis (Szklarczyk et al., 2019) was used to identify potential protein networks and functional enrichment of any shared pathways in the top ten activity-dependent interactors (plus RCCD1 and JMJD5). The network analysis showed only one potential cluster, which involved RCCD1, JMJD5, and a protein called coiled coil domain containing protein 130 (CCDC130) (Fig 4.9). The BioGRID database indicated that CCDC130 was identified as an interactor of RCCD1 and JMJD5 in a large-scale screen called BioPlex 2.0 (Huttlin et al., 2017). Because the CCDC130 protein is almost completely uncharacterised, it was not immediately clear how its JMJD5:RCCD1 interaction might relate to a role in replication fidelity (Chapter 3). This interaction will be discussed further in Section 5.2.1.
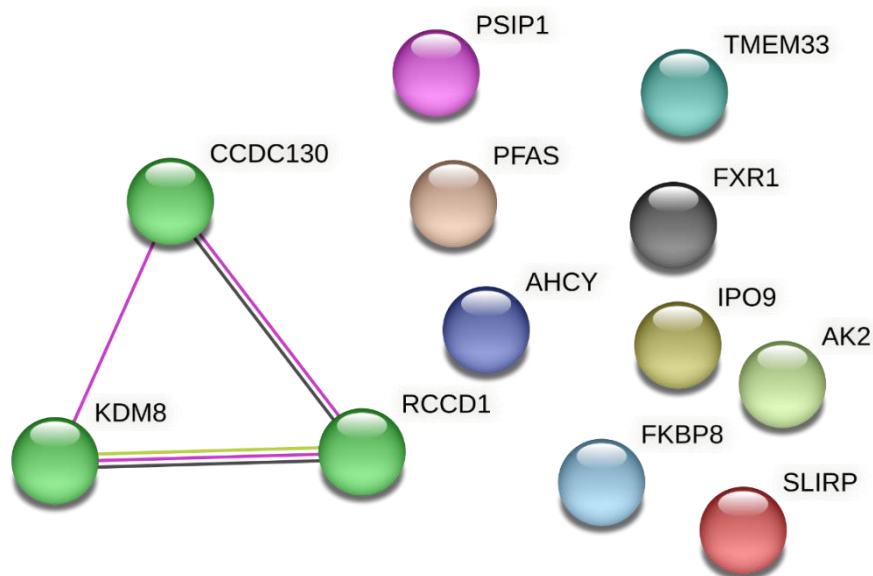
Figure 4.9. Network Analysis of JMJD5 activity-dependent interactors.
STRING network analysis (https://string-db.org). The identified cluster between JMJD5 (KDM8), CCDC130 and RCCD1 is coloured in green. The other nodes are not associated and randomly coloured. Network associations indicated by pink lines indicate experimental evidence, black is indicative of co-expression and lime green association results from text mining.

4.2.4.2   Proteomic screen with addition of Hydroxyurea

Our initial proteomic screen using the JMJD5:RCCD1 complex as bait did not appear to identify interactors or networks that were obviously linked with replication fidelity, DNA damage repair, or genome stability. Although further detailed investigation could elucidate such a role, we also considered it possible that the conditions used were not conducive to identify a replication-relevant interaction, particularly those that might be weak, rare, and/or insoluble. We hypothesised that it might be possible to enrich for replication relevant JMJD5:RCCD1 interactors by perturbing normal replication. Therefore, we used the same methodology as described above (Section 4.2.4.1), with the addition of 2 mM HU for 24 h prior to cell harvesting. Once the MS/MS analysis was completed, the full list of proteins was filtered as described above (Section 4.2.4.1). The top 50 interactors are presented below (Table 4.3).

Accession	Gene Name	WT JMJD5/RCCD1			H321A JMJD5/RCCD1		
		Coverage [%]	# Unique Peptides	Sequest Score	Coverage [%]	# Unique Peptides	Sequest Score
Q8N371	JMJD5	82	48	1613.03	82	47	1037.47
A6NED2	RCCD1	97	37	2082.62	97	38	1736.47
O60271	SPAG9	42	30	153.09	47	35	170.17
Q7Z6Z7	HUWE1	6	16	53.45	11	26	65.85
Q7L273	KCTD9	56	15	64.31	67	16	67.63
P22413	ENPP1	25	13	42.22	34	19	82.23
P13994	CCDC130	30	12	34.99	0	0	0
Q96CS3	FAF2	38	11	39.91	46	13	55.9
Q9P035	HACD3	28	10	34.94	32	11	43.87
Q10469	MGAT2	28	9	24	19	5	16.77
Q8N428	GALNT16	29	9	38.53	33	9	33.18
O94766	B3GAT3	55	8	36.42	22	5	21.33
Q53GQ0	HSD17B12	41	8	24.92	36	7	22.07
O94923	GLCE	18	7	20.48	8	3	4.82
P57088	TMEM33	30	5	21.71	26	4	12.12
P08574	CYC1	32	5	35.41	32	5	23.18
O95573	ACSL3	13	5	18.81	14	5	14.6
Q9HCU5	PREB	22	5	23.37	22	5	19.23
Q6UB35	MTHFD1L	6	4	10.62	2	1	2.76
Q13724	MOGS	9	4	8.01	5	2	7.14
P63173	RPL38	36	4	17.97	8	3	4.82
Q8WVX9	FAR1	9	4	11.14	12	3	14.03
Q9UNE7	STUB1	16	4	9.68	16	4	9.01
P62834	RAP1A	23	4	13.14	23	4	17.23
Q9NVI1	FANCI	5	4	11.91	6	5	9.1
Q9Y5M8	SRPRB	23	4	10.82	18	6	14.99
Q8N2K0	ABHD12	12	4	8.88	25	7	16.01
O43542	XRCC3	10	3	7.59	0	0	0
Q8NBS9	TXNDC5	12	3	7.9	5	1	2.01
Q9Y6K0	CEPT1	11	3	8.25	3	1	2.27
Q9HBH5	RDH14	18	3	8.11	5	1	4.59
O00767	SCD	11	3	8.33	9	2	2.55
P48047	ATP5PO	20	3	7.72	15	2	4.54
Q4VCS5	AMOT	5	3	9.82	3	2	6.13
Q96CX2	KCTD12	9	3	7.9	9	3	5.16
P46379	BAG6	5	3	9.46	5	3	8.15
O96008	TOMM40	17	3	11.98	17	3	10.49
Q3SXM5	HSDL1	15	3	11.2	12	3	14.03
Q8IY26	PLPP6	14	3	7.81	25	4	9.04
P08237	PFKM	7	3	12.32	11	5	11.59
Q99615	DNAJC7	8	3	7.71	13	5	13.44
O75746	SLC25A12	13	3	15.6	24	5	35.67
O43502	RAD51C	10	2	8.77	5	1	2.02
P43487	RANBP1	17	2	12.14	11	1	4.75
P58546	MTPN	32	2	8.08	18	1	3.79
P22061	PCMT1	14	2	9.3	7	1	4.39
Q7Z7H8	MRPL10	23	2	7.92	23	2	5.65
P08754	GNAI3	16	2	12.52	19	2	17.81
P63000	RAC1	22	2	12.62	30	4	12.65
O95816	BAG2	14	2	8.17	33	4	21.82

Table 4.3. List of top fifty JMJD5:RCCD1 interactors following HU treatment.
Displayed is the accession number and gene name, peptide coverage (%), unique peptides, and Sequest Score. The coverage is the percentage of the protein covered by detected versus predicted tryptic fragments. The unique peptides are the number of peptides identified in each protein that belong solely to that protein in the reference proteome used (Homo Sapiens). The Sequest Score is the cumulative value obtained from scoring and adding each peptide. MS data is from one biological repeat.

To compare the effects of HU treatment on the JMJD5:RCCD1 interactome directly, we generated a table comparing the WT JMJD5/RCCD1 before and after HU treatment (Table 4.4). Although HU treatment did not appear to dramatically alter the previously described interactions with SPAG9 and KCTD9, it may have led to others becoming more prominent. For example, HUWE1, an E3 ubiquitin ligase which targets many proteins involved in the DNA damage response (Kao et al., 2018), was identified in our first screen (Table 4.1) but may have been a more abundant interactor in the presence of HU (Table 4.4). HUWE1 appeared to be enriched, rather than depleted, in the H321A JMJD5/RCCD1 sample, suggesting that it may be more likely to be a regulator of the JMJD5:RCCD1 complex than a substrate (Table 4.3) (discussed further in section 5.3).

Accession	Gene Name	Coverage [%]	# Unique Peptides	Accession	Gene name	Coverage [%]	# Unique Peptides
WT JMJD5/RCCD1 -HU				**WT JMJD5/RCCD1+HU**			
Q8N371	JMJD5	82	44	Q8N371	JMJD5	82	48
A6NED2	RCCD1	97	34	A6NED2	RCCD1	97	37
O60271	SPAG9	20	17	O60271	SPAG9	42	30
Q7L273	KCTD9	40	10	Q7Z6Z7	HUWE1	6	16
P22413	ENPP1	18	10	Q7L273	KCTD9	56	15
P54819	AK2	55	9	P22413	ENPP1	25	13
Q9UJS0	SLC25A13	17	8	P13994	CCDC130	30	12
P08574	CYC1	35	7	Q96CS3	FAF2	38	11
Q9P035	HACD3	23	7	Q9P035	HACD3	28	10
Q7Z6Z7	HUWE1	4	7	Q10469	MGAT2	28	9
P53985	SLC16A1	12	6	Q8N428	GALNT16	29	9
Q07065	CKAP4	17	6	O94766	B3GAT3.	55	8
P13994	CCDC130	12	4	Q53GQ0	HSD17B12	41	8
P22695	UQCRC2	18	4	O94923	GLCE	18	7
Q9HCU5	PREB	17	4	P57088	TMEM33	30	5
Q92616	GCN1	3	4	P08574	CYC1	32	5
Q9Y5M8	SRPRB	24	4	O95573	ACSL3	13	5
Q96CS3	FAF2	17	4	Q9HCU5	PREB	22	5
O15067	PFAS	5	3	Q6UB35	MTHFD1L	6	4
P23526	AHCY	7	3	Q13724	MOGS	9	4
O75475	PSIP1	12	3	P63173	RPL38	36	4
P57088	TMEM33	12	3	Q8WVX9	FAR1	9	4
Q14318	FKBP8	13	3	Q9UNE7	STUB1	16	4
Q9GZT3	SLIRP	33	3	P62834	RAP1A	23	4
O94766	B3GAT3	11	3	Q9NVI1	FANCI	5	4
P62834	RAP1A	23	3	Q9Y5M8	SRPRB	23	4
Q93008	USP9X	2	3	Q8N2K0	ABHD12	12	4
P36542	ATP5F1C	18	3	O43542	XRCC3	10	3
Q9Y277	VDAC3	11	3	Q8NBS9	TXNDC5	12	3
Q8N428	GALNT16	13	3	Q9Y6K0	CEPT1	11	3
Q53GQ0	HSD17B12	15	3	Q9HBH5	RDH14	18	3
O00264	PGRMC1	39	3	O00767	SCD	11	3
P60891	PRPS1	9	3	P48047	ATP5PO	20	3
P63000	RAC1	17	3	Q4VCS5	AMOT	5	3
Q9Y6C9	MTCH2	22	3	Q96CX2	KCTD12	9	3
Q96TA2	YME1L1	6	3	P46379	BAG6	5	3
P61619	SEC61A1	10	3	O96008	TOMM40	17	3
Q96P70	IPO9	6	2	Q3SXM5	HSDL1	15	3
P51114	FXR1	4	2	Q8IY26	PLPP6	14	3
Q99714	HSD17B10	18	2	P08237	PFKM	7	3
P63173	RPL38	36	2	Q99615	DNAJC7	8	3
P52815	MRPL12	13	2	O75746	SLC25A12	13	3
Q96T76	MMS19	3	2	O43502	RAD51C	10	2
O00410	IPO5	4	2	P43487	RANBP1	17	2
P11310	ACADM	10	2	P58546	MTPN	32	2
P62879	GNB2	14	2	P22061	PCMT1	14	2
Q86VU5	COMTD1	19	2	Q7Z7H8	MRPL10	23	2

Table 4.4 HU treatment alters interactome of WT JMJD5/RCCD1.
Displayed is the accession number and gene name, peptide coverage (%), unique peptides, and Sequest Score. The coverage is the percentage of the protein covered by detected versus predicted tryptic fragments. The unique peptides are the number of peptides identified in each protein that belong solely to that protein in the reference proteome used (Homo Sapiens). The Sequest Score is the cumulative value obtained from scoring and adding each peptide. Proteins enriched following HU treatment are highlighted in blue. MS data is from one biological repeat.

Interestingly, three proteins associated with DNA damage repair appeared as interactors following HU treatment; 'Fanconi anaemia group I' (FANCI) and the RAD51 paralogs 'X-ray repair cross-complementing protein 3' (XRCC3) and 'RAD51 homolog C' (RAD51C) (Table 4.3). Subsequently, the application of our activity dependence filtering criteria further isolated both RAD51C and XRCC3, with XRCC3 completely absent from the H321A list (Table 4.5). CCDC130 again emerged as a prominent substrate candidate under these conditions (Table 4.5) (discussed further in Section 5.2.1).

Accession	Gene Name	WT JMJD5/ RCCD1 # Unique Peptides	H321A JMJD5/ RCCD1 # Unique Peptides	Activity Dependence
P13994	CCDC130	12	0	∞
O43542	XRCC3	3	0	∞
Q6UB35	MTHFD1L	4	1	4
Q8NBS9	TXNDC5	3	1	3
Q9Y6K0	CEPT1	3	1	3
Q9HBH5	RDH14	3	1	3
O94923	GLCE	7	3	2.333333333
Q13724	MOGS	4	2	2
O43502	RAD51C	2	1	2
P43487	RANBP1	2	1	2

Table 4.5. List of top ten activity dependent interactors after HU treatment.
Displayed is the accession number and gene name of the identified protein, the unique peptides, and the activity-dependence, which was expressed as the ratio of unique peptides in the WT to H321A sample. For samples where no protein was detected in the H321A the activity-dependence ratio is indicated by the infinity symbol. The unique peptides are the number of peptides identified in each protein that belong solely to that protein in the reference proteome used (Homo Sapiens).

We applied STRING analysis (Szklarczyk et al., 2019) to identify potential protein networks and functional enrichments, which now showed two clusters: the previously mentioned JMJD5(KDM8)-RCCD1-CCDC130 cluster and a new cluster formed by RAD51C-XRCC3 (Fig 4.10).
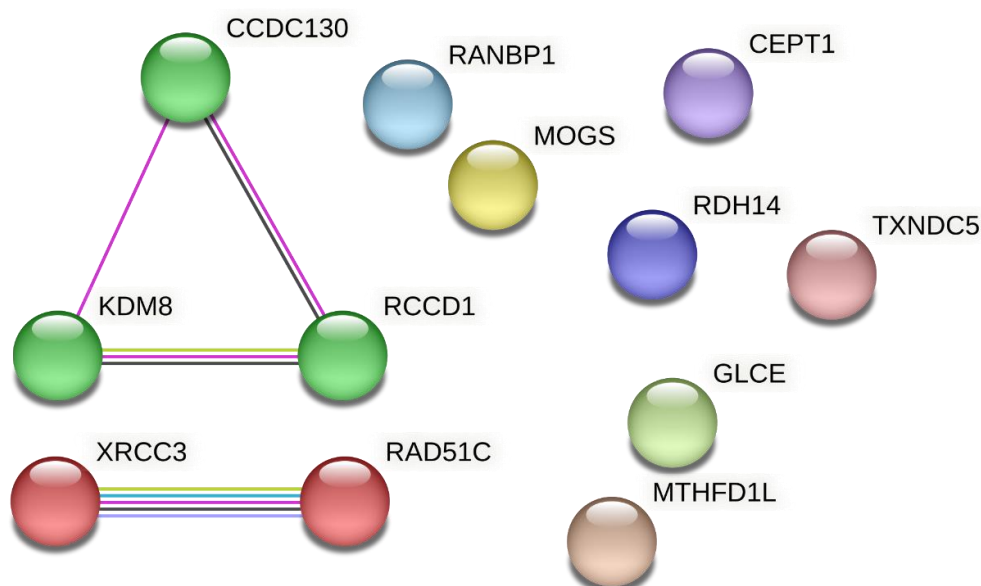
Figure 4.10. Network Analysis of JMJD5:RCCD1 activity-dependent interactors following HU treatment.
STRING network analysis (https://string-db.org). The identified cluster between JMJD5 (KDM8), CCDC130 and RCCD1 is coloured in green. The XRCC3-RAD51C cluster is coloured in red. The other nodes are not associated and are randomly coloured. Network associations indicated by pink lines indicate experimental evidence, black is indicative of co-expression and lime green association results from text mining, turquoise is interactions found in curated databases and lilac indicates protein homology.

RAD51C and XRCC3 are RAD51 paralogs (reviewed in detail in Section 4.3.4) that are known to form a distinct sub-complex (termed 'CX3') (Liu et al., 2002), that function in the DNA damage response. The two roles of the CX3 complex of particular interest are 1) functions in the efficient restart of stalled replication forks (Petermann et al., 2010; Berti et al., 2020) and 2) importance in the late stages of HR (Chun et al., 2013). These directly overlap with the evidence provided here regarding the role of the JMJD5:RCCD1 complex in maintaining replication fidelity (Chapter 3) and the published role of JMJD5 in HR (Amendola et al., 2017).

Considering the known importance of the CX3 complex in the above processes, their interaction with JMJD5:RCCD1 warranted further investigation. We obtained p3XFLAG-myc-CMV™-26 vectors that express RAD51C or XRCC3 with an N-terminal 3XFLAG tag and a C-terminal Myc tag. We transfected these vectors into pTIPZ EV, HA-JMJD5(WT)/RCCD1-FLAG, and HA-JMJD5(H321A)/RCCD1-FLAG HEK293T cells for 48 h, followed by HA-immunoprecipitation. Western blotting for Myc (RAD51C and

XRCC3), FLAG (RAD51C, XRCC3, and RCCD1), and HA (JMJD5) successfully confirmed the specific interaction between the JMJD5:RCCD1 and CX3 complexes (Fig 4.11). The relative signal in the input and IP samples suggests this interaction is quite weak under the conditions used, however. Interestingly, the anti-Myc blot in the IP samples suggested some potential activity-dependence of the interaction (Fig 4.11), consistent with the proteomic data (Table 4.5). However, the interpretation is complicated by the H321A JMJD5 mutant being modestly underexpressed compared to WT JMJD5 (Fig 4.11), and the anti-Myc blot for the input samples being overexposed. Although further work is needed to resolve the activity-dependence of the interaction, we were sufficiently encouraged by the specificity of the interaction, using two different approaches, to explore its potential functional relevance further.
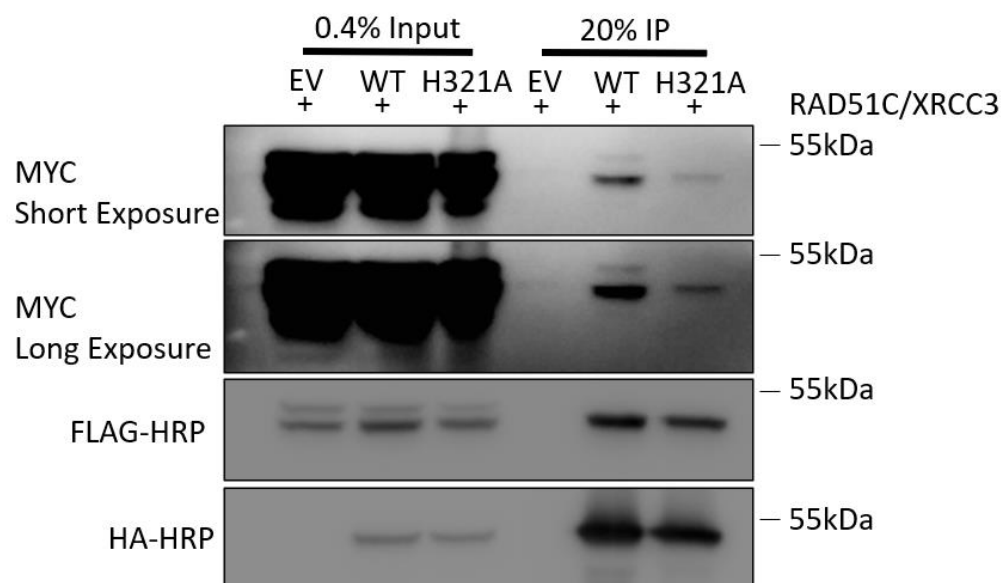
Figure 4.11. JMJD5:RCCD1 interacts with the CX3 complex of RAD51 paralogs.
pTIPZ EV HA-JMJD5(WT)/RCCD1-FLAG and HA-JMJD5(H321A)/RCCD1-FLAG cells were treated with doxycycline for 48 h and transfected with p3XFLAG-myc-CMV™-26 RAD51C and XRCC3 before harvesting and anti-HA immunoprecipitation. Input samples are indicative of cell lysate prior to IP. Samples were run on an SDS-PAGE gel and Western blotted with the indicated antibodies.

4.2.5 RCCD1 and JMJD5 are required for efficient replication fork restart and this is epistatic with respect to that of the CX3 complex

As mentioned above, the CX3 complex shares overlapping functional roles with JMJD5:RCCD1, indicating that further investigation is required to elucidate this correlation; a more detailed review of the RAD51 paralogs is provided in section 4.3.4. Although both the HR and fork restart roles of CX3 were relevant, due to time constraints, we focused on elucidating the connection between JMJD5:RCCD1 and replication fork restart, whereas a potential role in HR is discussed later (Section 4.3.4).

When a replication fork stalls at an obstacle, mechanisms exist to restart the fork and complete DNA replication if the blockage is removed (Petermann and Helleday, 2010). Inefficient replication fork restart could contribute to replication stress induced by JMJD5:RCCD1 loss of function (Chapter 3), raising the possibility that this might be related to altered function of the CX3 complex. If true, we hypothesised that RCCD1/JMJD5 knockdown should cause a replication fork restart phenotype.

Investigating replication fork restart uses a variation of the DNA fibre assay methodology (section 3.2.3). Rather than incubating the two thymidine analogues for equal amounts of time, the incubation time for the IdU label (green) is doubled (Fig 4.12A). This allows time after the removal of the HU treatment for stalled forks to restart efficiently. Therefore, any fibre stained with only the first label is classified as a stalled fork (red label only). Those forks with red-green orientations have successfully incorporated the IdU analogue (green) and are indicative of ongoing forks that have restarted (Fig 4.12B).
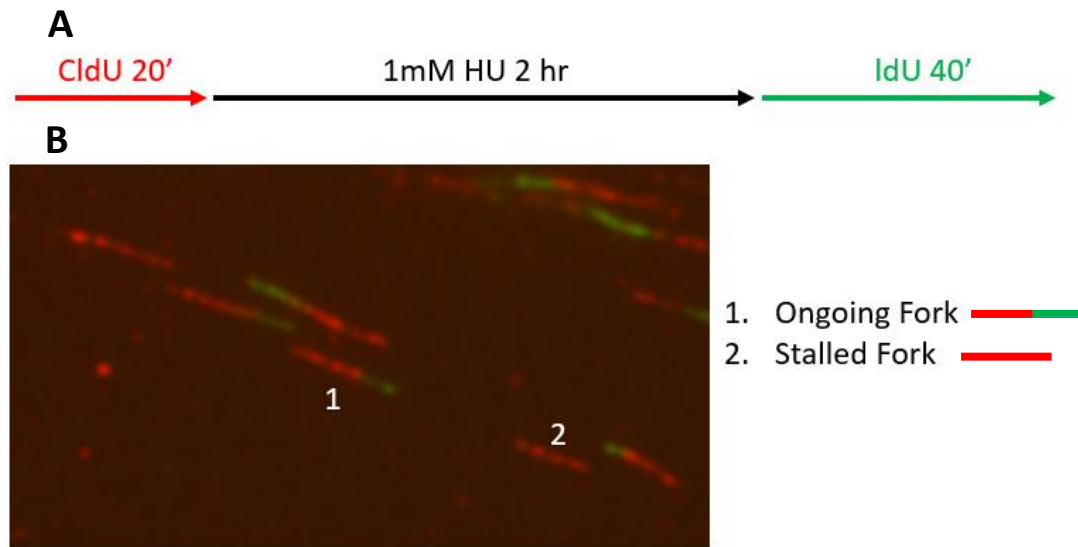
Figure 4.12. DNA fibre methodology for examining efficient replication fork restart.
(A) Assay schematic detailing length of incubations for CldU, HU, and IdU treatments. Red indicates Alexa Fluor® 555 secondary used to visualise CldU, and Green is the Alexa Fluor® 488 secondary used to visualise the IdU. **(B)** Example fibre structures quantified using DNA fibre assay for Replication fork restart. (1. Ongoing fork, 2. Stalled Fork)

4.2.5.1   RCCD1 and JMJD5 depletion decreases replication fork restart

To test for a potential role of RCCD1 and JMJD5 in replication fork restart, we used the siRNA knockdown models described in the previous section (Section 3.2.1) and the specific DNA fibre methodology outlined in Figure 4.12. Consistent with the hypothesis, siRNA knockdown of RCCD1 or JMJD5 resulted in a statistically significant increase in the number of stalled forks, indicative of deficient replication fork restart (Fig 4.13A and B). While providing preliminary insight into the potential role of JMJD5:RCCD1 at a replication fork in response to replication stress, its functional relationship with the CX3 complex requires further analysis.
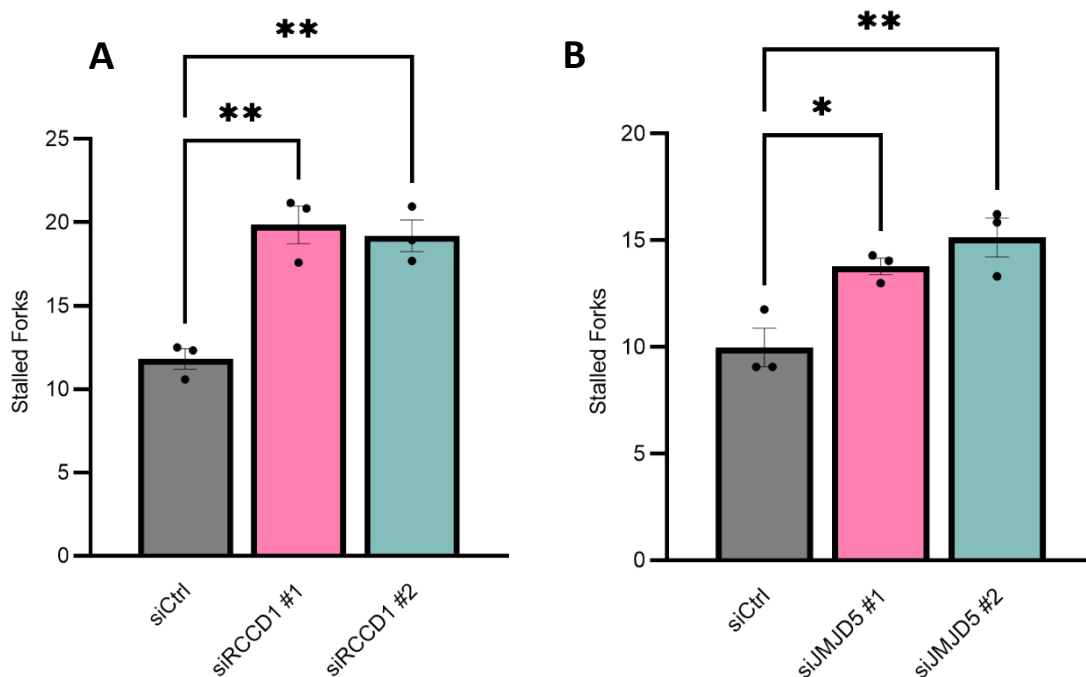
Figure 4.13. RCCD1 and JMJD5 depletion increase deficiencies in replication fork restart.
With this specific fibre labelling protocol (Fig 5.12) the number of stalled forks indicative of replication fork restart was increased after **(A)** RCCD1 (siCtrl (grey), siRCCD1 #1 (pink), and siRCCD1 #2 (blue)) knockdown and **(B)** JMJD5 (siCtrl (grey), siJMJD5 #1 (pink) and siJMJD5 #2 (blue)) knockdown in A549 cells. Statistical analysis used One way ANOVA with Bonferroni's post hoc test (with p-values of ≤0.05 (*) and ≤0.01 (**)). The data shown represents the mean of 3 biological repeats -/+ SEM. A minimum of 250 forks were quantified per condition per biological repeat.

4.2.5.2   The JMJD5:RCCD1 replication fork restart phenotype is epistatic with that of XRCC3

In an attempt to further explore the relationship between the JMJD5:RCCD1 and RAD51C:XRCC3 complexes in replication fork restart, we sought to use a similar epistasis analysis to that applied in Chapter 3 (Section 3.2.5). We hypothesised that if the CX3 complex is important for the role of JMJD5:RCCD1 in replication fork restart (or vice versa), the phenotypes would be epistatic, that is to say that JMJD5:RCCD1 loss of function would not exacerbate the phenotype in response to CX3 knockdown. To specifically investigate the role of the CX3 complex, it is necessary to target only XRCC3 because RAD51C also functions in a 'BCDX2' complex involved in the earlier stages of HR and replication fork reversal (Chun et al., 2013; Berti et al., 2020). Two independent XRCC3 siRNA

sequences were tested by transfection and Western blot using an endogenous antibody against XRCC3 (data not shown), with one of the sequences being taken forward to the epistasis analysis.

We transfected A549 cells with the following siRNA combinations: Ctrl, RCCD1/Ctrl, JMJD5/Ctrl, XRCC3/Ctrl, RCCD1/XRCC3, JMJD5/XRCC3, and RCCD1/JMJD5. Successful siRNA depletion was confirmed by Western blotting (Fig 4.14A). Interestingly, XRCC3 depletion seemed to increase the levels of both RCCD1 and JMJD5, and co-depletion of JMJD5/RCCD1 increased XRCC3 levels (Fig 4.14A) (discussed further in Section 4.3.4.3).

Consistent with the results presented in Figure 4.13, RCCD1 or JMJD5 siRNA increased the number of stalled forks associated with fork restart deficiency (Fig 4.14B). In line with the literature, XRCC3 depletion also increased the number of stalled forks associated with fork restart deficiency (Fig 4.14B) (Berti et al., 2020; Petermann et al., 2010). Importantly, there was no additional increase in these fork restart phenotypes under conditions where both RCCD1:JMJD5 and CX3 were targeted simultaneously (Fig 4.14B), suggesting that the functional relationship between these complexes is indeed epistatic. The results may be consistent with the physical association of the RCCD1:JMJD5 and CX3 complexes that underlies a role in a common pathway involved in replication fidelity.
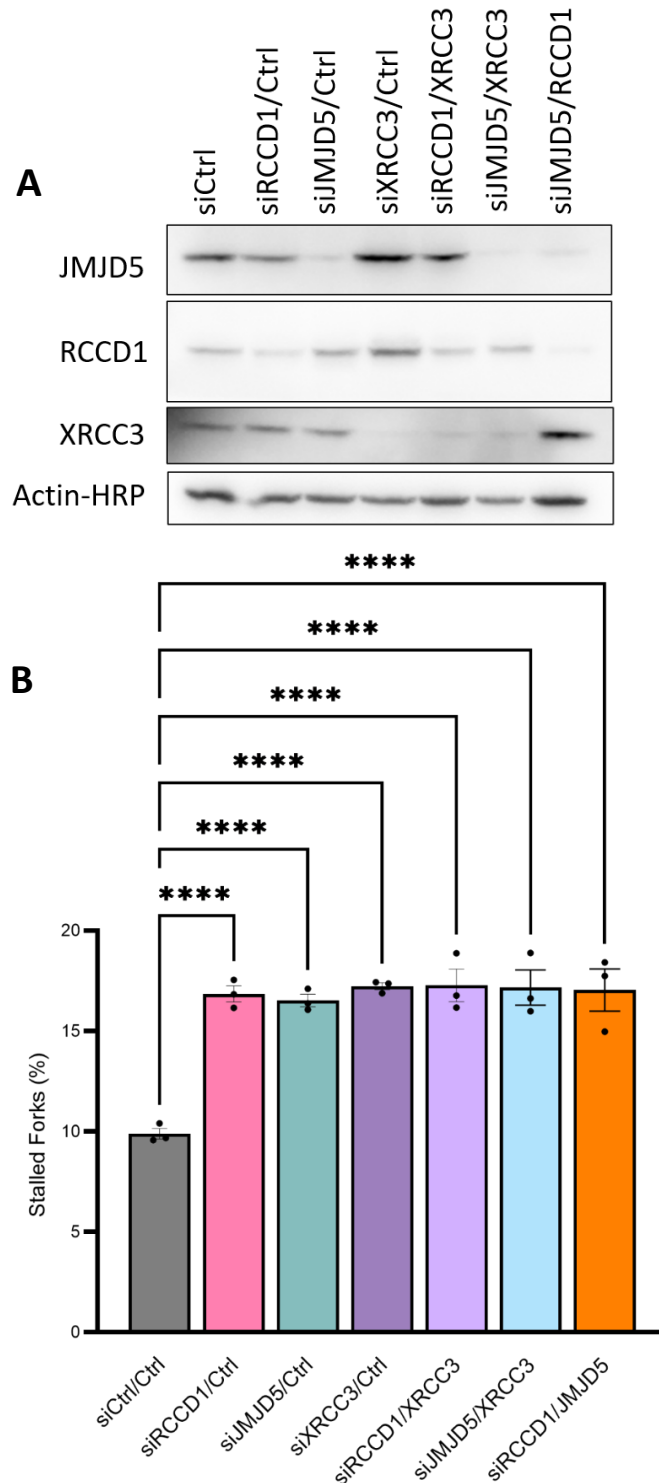
Figure 4.14. Replication fork restart deficiency due to JMJD5:RCCD1 loss of function is epistatic with respect to that of CX3.

(A) Epistasis siRNA combinations successfully knockdown RCCD1, JMJD5, and XRCC3 in A549 cancer cells. Cells were transfected with the indicated siRNA sequences for 72 h, harvested, and then analysed by SDS-PAGE and Western blotting with the indicated antibodies. **(B)** JMJD5, RCCD1, and XRCC3 depletion increases the number of stalled forks indicative of replication fork restart deficiency. A549 cells were transfected with either siCtrl (grey), RCCD1/Ctrl (pink), siJMJD5/Ctrl (teal), siXRCC3/Ctrl (purple), siRCCD1/XRCC3 (lilac), siJMJD5/XRCC3 (light blue) and siRCCD1/JMJD5 (orange). Statistical analysis used One way ANOVA with Bonferroni's post hoc test (with p-values of ≤0.0001 (****). The data shown represents the mean of 3 biological repeats -/+ SEM. A minimum of 250 forks were quantified per condition per biological repeat.

4.3 Discussion

In this chapter, we uncovered the stoichiometric nature of the RCCD1:JMJD5 interaction. This in turn led us to develop a unique co-expression system that we applied in proteomic screens to identify an interaction with the CX3 complex and a role for JMJD5 and RCCD1 in replication fork restart.

4.3.1 JMJD5:RCCD1 form a heterodimeric 1:1 complex

During the work in Chapter 3 that characterised the JMJD5:RCCD1 complex it was not clear whether the interaction was transient in nature, or more stable, consistent with a heterodimer. In this chapter, we discovered that JMJD5 and RCCD1 form a stoichiometric 1:1 heterodimer (Section 4.2.1). This was validated via analytical gel filtration, which demonstrated a single peak for the isolated complex at 90 kDa (Fig 4.6A), which was consistent with the combined molecular weight of one RCCD1 molecule and a JMJD5 monomer. The latter is interesting because several closely-related JmjC-only protein hydroxylases discussed earlier (section 1.1.5) demonstrate dimerisation, including; JMJD7 (Markolovic et al., 2018), ribosomal hydroxylases MINA53 and NO66 (Chowdhury et al., 2014), and FIH (Lancaster et al., 2004). Furthermore, Shen et al proposed that the JmjC-domain of JMJD5 facilitates dimerisation, the functional relevance of which was unclear (Shen et al., 2017). As such, we might have expected to observe JMJD5 dimerisation via our analytical gel filtration, which could have presented as a heteromultimer comprised of two monomers of JMJD5 and one molecule of RCCD1. However, the gel filtration profile observed suggests that JMJD5 may not dimerise, at least under the conditions tested. Whether it may dimerise in response to a stimulus, such as DNA damage or replication stress, would be interesting to explore.

A priority for future investigation is to establish the function of the JMJD5:RCCD1 interaction (discussed in more detail in section 5.1.). The stoichiometric nature of the complex might suggest that the interaction is essential for the function of JMJD5 and RCCD1 in replication fidelity. Such complexes can also demonstrate interdependencies whereby each subunit is an 'obligate' binding partner that is essential for the optimal stability and function of the other subunits.

4.3.2   Is the JMJD5:RCCD1 interaction required for optimal expression of the complex?

Several lines of evidence suggest that JMJD5 and RCCD1 might be required for each other's optimal expression. Firstly, we discovered that mutations in RCCD1 that reduce JMJD5 binding also cause reduced RCCD1 expression (Section 2.2.12.3). Furthermore, knockdown of endogenous JMJD5 reduced RCCD1 expression (Section 3.2.5), although it is currently unclear if these observed effects were 'on-target' (as discussed in 3.3.2). Additionally, attempts to generate bacterially expressed RCCD1 in isolation also failed, due to instability and aggregation (Dr Chan Li – personal communication). Efforts to overexpress RCCD1 in HEK293T cells were also not very successful because the protein was poorly expressed and bound to several other proteins, including molecular chaperones (Fig 4.1B). Based on this evidence, we hypothesised that RCCD1 protein expression might be reliant on JMJD5; indeed, co-overexpression with JMJD5 significantly increased RCCD1 protein expression (Fig 4.2). While these correlative observations suggest a positive role for JMJD5 in regulating RCCD1 expression, possibly at the level of protein stability, we have not formally proven that this regulation occurs at the post-transcriptional level. Future studies aiming to address this could measure the half-life of each protein individually versus in the complex using radiolabelled methionine pulse chase or cycloheximide assays. Furthermore, qPCR should be used to measure potential changes in JMJD5 and RCCD1 mRNA abundance in response to knockdown.

4.3.3   Proteomic characterisation of the JMJD5:RCCD1 interactome

In this chapter we have used our multicistronic P2A system to overexpress the JMJD5:RCCD1 heterodimeric complex and identify a complex associated interactome (Table 4.1 and Table 4.3), that could provide new functional insight. Proteins identified include the potential RCCD1 interactors SPAG9 and KCTD9, both discussed in section 5.3. Additionally, we have identified potential regulators of the complex in HUWE1, providing a potential insight into upstream signalling regulating the JMJD5:RCCD1 complex (discussed further in section 5.3). We also identified a HU induced interaction with FANCI, a key member of the DNA damage response that will be discussed further in Chapter 5.

Lastly, we have also identified potential activity dependent interactors of the complex in CX3 (discussed in Section 4.3.4) and CCDC130 (section 5.2.1).

Although our MS approach to identify JMJD5:RCCD1 interactors and JMJD5 substrates yielded several interesting candidates that will be developed into new projects, including the CX3 complex (which was functionally rationalised with respect to replication fidelity here), several technical limitations may have hampered a more sensitive analysis. For example, our approach was only semi-quantitative, relying on indirect measures of abundance such as the number of unique peptides detected. Although this approach showed an activity-dependent interaction with the CX3 complex (Table 4.4) that was confirmed by Western blotting (Fig 4.11), it does not factor in the potential expression differences between WT and H321A JMJD5. Alternative more quantitative methods of proteomics that could be considered in the future include 'stable isotope labelling by amino acids in cell culture' (SILAC) and 'tandem mass tags' (TMT). Such approaches could enable any differences in JMJD5 expression to be accounted for and potential activity-dependent and -independent interactors to be detected with greater sensitivity.

Another potential limitation of our proteomic approach is that it may have been limited to relatively soluble interactors due to the mild lysis conditions used (method described in Chapter 6). If the JMJD5:RCCD1 complex plays a role in either DNA repair or replication fidelity, as reported here and in the literature (Amendola et al., 2017; Fletcher et al., 2023), it may be localised to chromatin. Proteins localised to chromatin may be harder to isolate in soluble complexes such as those analysed here. Specific proteomic methods have been developed that can be applied in such circumstances. For example, Bioitin Identification (BioID) uses a biotin ligase fused to a targeting protein (in this case JMJD5 or RCCD1), which can ligate biotin to neighbouring proteins when biotin is added to the cell media (Roux et al., 2018). These proteins can be isolated and identified using streptavidin-based pull-down and mass spectrometry. Because the method employs covalent proximity tagging it enables the use of much harsher cell lysis methods to be used, which could support more efficient solubilisation

and detection of JMJD5:RCCD1 interactors present in chromatin. Given the transient nature of enzyme-substrate interactions, it could also help identify potential JMJD5 substrates with greater sensitivity.

4.3.4 The CX3 complex of RAD51 paralogs link the JMJD5:RCCD1 complex to replication fork restart

Although they have some limitations, the MS methods used in this chapter led to the identification of the CX3 complex as an activity-dependent JMJD5:RCCD1 interactor (Table 4.4), that was validated by Western blotting (Fig 4.11). Limitations of this interaction analyses include the use of overexpression systems, future work should address whether the interactions of the endogenous proteins behave as described above. However, current endogenous antibodies for JMJD5, RCCD1 and the CX3 are unsuitable for endogenous immunoprecipitation (data not shown), which makes this challenging at present. CRISPR-CAS9 mediated epitope tagging of the endogenous proteins could provide a viable way forward for analyses of the endogenous interactions. Alternative approaches to immunoprecipitation such as Western blotting of fractions collected from gel filtration of endogenous proteins in cell extracts, or proximity ligations assays, could also be useful.

Future work should explore whether this activity-dependent interaction profile is explained by CX3 being a JMJD5 substrate. First, the CX3 amino acid sequence should be assessed for homology to the experimental JMJD5 substrate (RPS6 peptide) that has been used to characterise JMJD5 activity, both in the literature (Wilkins et al., 2018) and this thesis (Section 2.2.10). If one or more regions of homology are identified, short synthetic peptides could be generated and used for *in vitro* hydroxylation assays. Because this interaction was identified in the presence of RCCD1, such *in vitro* assays should consider using recombinant JMJD5:RCCD1 complex rather than JMJD5 alone (potential regulation of JMJD5 activity by RCCD1 is discussed in Section 5.1). Additionally, the CX3 overexpression vectors used in section 4.2.4.2 could be transfected into the JMJD5-P2A-RCCD1 HEK293T cells, followed by IP and MS/MS screening for arginyl hydroxylation.

Regardless of whether the CX3 complex is a substrate or binding partner of the JMJD5:RCCD1 complex, the work has enabled us to refine the role of JMJD5 and RCCD1 in replication fidelity. By considering the known function of CX3 we identified that JMJD5 and RCCD1 are required for efficient restart of stalled replication forks (Fig 4.13), which could also be consistent with a role in HR. Because fork restart and HR pathways are intrinsically linked through the RAD51 protein, we provide an overview of RAD51 and related biological pathways below.

4.3.4.1   RAD51 and its paralogs

RAD51 is a key protein involved in DNA repair, where it forms nucleoprotein filaments, and uses homology searches and strand invasion to identify the homologous template required for HR (Bonilla et al., 2020). RAD51 contains an ATPase domain with Walker A and Walker B motifs that bind and hydrolyse ATP, which is dependent on DNA binding (Short et al., 2016; Tombline and Fishel, 2002). RAD51 has five paralogs in humans, which show approximately 20-30% homology to RAD51 and retention of the Walker motifs (Suwaki et al., 2011). According to Suwaki et al. (2011), these five paralogs (RAD51B, RAD51C, RAD51D, XRCC2, and XRCC3) form two biologically distinct complexes, the 'BCDX2' complex (RAD51**B**, RAD51**C**, RAD51**D** and **X**RCC**2**) and the CX3 complex (RAD51**C** and **X**RCC**3**), with RAD51C found in both complexes (Fig 4.15) (Suwaki et al., 2011). In this chapter we describe a specific interaction between JMJD5:RCCD1 and the CX3 complex, with other components of the BCDX2 complex not being detected (Table 4.3). Since RAD51C is common to both complexes, this might suggest that JMJD5:RCCD1 is interacting with CX3 via XRCC3, not RAD51C, although this requires further investigation.
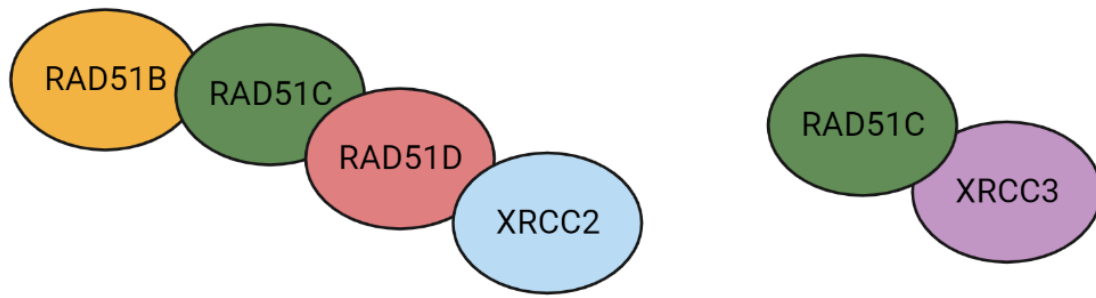
Figure 4.15. Schematic of RAD51 paralog complexes.
The BCDX2 consists of RAD51B, RAD51C, RAD51D, and XRCC2 and CX3 consists of RAD51C and XRCC3.

4.3.4.2 Homologous recombination

Double-strand breaks (DSBs) can arise from multiple sources, including but not limited to replication stress, radiation, and chemotherapy. DSBs are very cytotoxic, and their mis repair is an important contributory factor to genome instability. HR is a DSB repair pathway predominantly active in S and G2 phases of the cell cycle, which uses a homologous DNA template to ensure accurate repair (Wright et al., 2018). Non-homologous end joining (NHEJ) is another DSB repair method, providing faster repair throughout across the cell cycle (Pannunzio et al., 2018). However, HR offers a more accurate form of DNA repair than NHEJ, which is potentially error-prone.

Mechanistically, HR begins with the recognition and binding of DSBs by the MRE11-RAD50-NBS1 (MRN) complex in conjunction with CtBP-interacting protein (CtIP), which facilitates resection to generate short 3' ssDNA overhangs (Symington and Gautier, 2011). Additional resection events by either the exonuclease EXO1 or helicases DNA2 and BLM lengthen the exposed ssDNA (Symington and Gautier, 2011). The exposed ssDNA is protected from degradation and formation of ssDNA structures by the rapid binding of RPA (Chen et al., 2013), which is subsequently displaced by formation of RAD51 filaments (Fig 4.16). RAD51 nucleoprotein filaments form by each RAD51 protein binding three nucleotides of DNA, which is then extended by the addition of RAD51 dimers (Bonilla et al., 2020). RAD51 binding to ssDNA and filament formation are facilitated by a group of proteins known as 'RAD51 mediators', which include the RAD51 paralogs (discussed further below) and 'Breast Cancer

type 2 susceptibility protein' (BRCA2). BRCA2 accelerates the formation of RAD51 filaments following its recruitment to DSBs by the 'Partner and localiser of BRCA2' (PALB2) (Prakash et al., 2015), where it facilitates RAD51 recruitment and suppresses its ATPase activity, enhancing its affinity for ssDNA (Carreira et al., 2009). RAD51 filaments subsequently search and invade a homologous region of DNA to generate a 'displacement loop' (D-loop), allowing DNA synthesis to occur (Wright et al., 2018) (Fig 4.16).
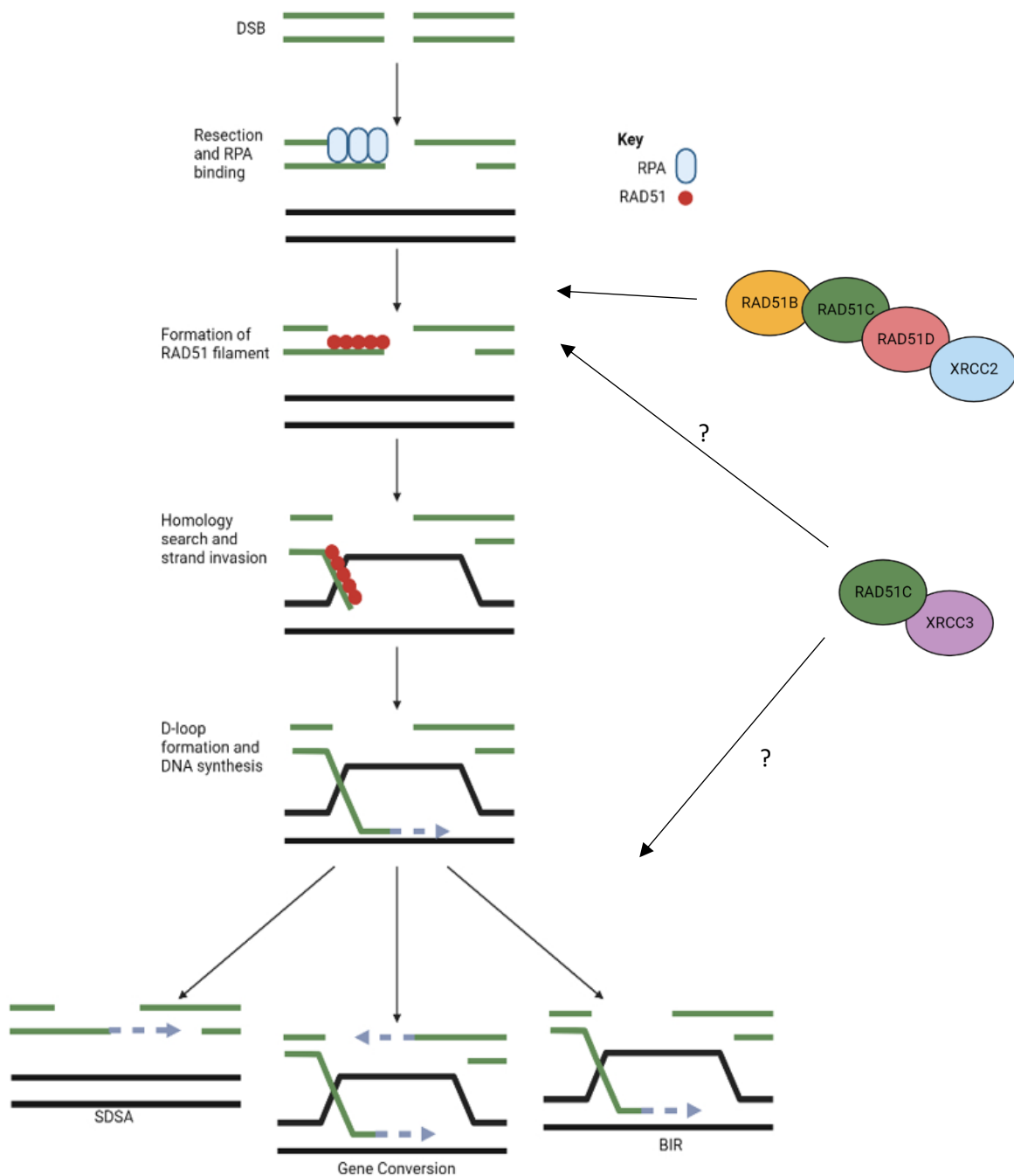
Figure 4.16. Model for DSB repair via homologous recombination (HR).
A DSB is recognised and resected into a 3' ssDNA overhang by the actions of the MRN complex, with additional resection catalysed by EXO1 and DNA2. The exposed ssDNA is bound rapidly by RPA, which is displaced to form RAD51 nucleoprotein filaments. The RAD51 filament then carries out homology search and strand invasion to form a D-loop for DNA synthesis. D-loops are then resolved via either SDSA, Gene conversion or BIR. The two RAD51 paralog complexes play sequential roles in HR. The BCDX3 complex functions to support RAD51 filament formation and CX3 function in HR remains controversial with both early and late-stage roles reported although the precise mechanism remains unclear. Figure adapted from Bonilla et al 2020 and generated using BioRender.

Following DNA synthesis, the resolution of the D-loop can occur via three main pathways: (1) synthesis dependent strand annealing (SDSA), (2) RAD51-dependent break-induced replication (BIR) or (3) gene conversion (Fig 4.16) (Bonilla et al., 2020; Wright et al., 2018). SDSA occurs when the D-loop structure is disrupted, which allows annealing of nascent DNA to the other end of the DSB (Li and Heyer, 2008). Gene conversion is characterised by the formation of a double Holliday junction (dHJ), resulting in the extension of the D-loop and invasion and capture of the second strand by the nascent DNA, this dHJ can then be processed by either dissolution or resolution (Wright et al., 2018). Break-induced replication occurs when the second end fails to be captured, resulting in the progression of DNA synthesis (Sakofsky and Malkova, 2017).

Of particular interest here is the role of RAD51 paralogs as RAD51 mediators during HR, where functionally distinct roles for the two complexes (BCDX2 and CX3) have been reported. The BCDX2 complex was shown to be important for the formation of RAD51 filaments (Fig 4.16), whereas the function of the CX3 complex is controversial. One report has suggested CX3 is not essential for RAD51 filament formation, but depletion results in severely reduced HR (Chun et al., 2013). However, there is additional evidence suggesting that the CX3 complex was required for RAD51 filament formation (Garcin et al., 2019). Taken together, there could be a mechanistic role of the CX3 complex in the early and late stages of HR (Fig 4.16).

In support of a function for CX3 in the late stages of HR, RAD51C foci appear concurrently with RAD51, but persist after RAD51 foci disappear (Badie et al., 2009), possibly indicative of the presence of RAD51C in the two distinct paralogue complexes (Fig 4.16). Furthermore, both RAD51C and XRCC3 depletion have been shown to result in reduced resolution of Holliday junctions (Liu et al., 2004). The potential late-stage HR role of CX3 is of particular interest, because JMJD5 loss-of-function did not reduce RAD51 foci formation in *C. elegans* but did cause persistence of foci that were not removed (Amendola et al., 2017), which may be consistent with a function of JMJD5 after RAD51 filament

formation. However, it remains unclear whether JMJD5 and RCCD1 directly play a role in eukaryotic HR, or at what stage, and future studies should attempt to elucidate this.

4.3.4.3   Replication fork restart

When a stalled fork encounters a blockage, it can be restarted efficiently if the blockage is removed (Petermann and Helleday, 2010). Of particular interest here is the non-DSB repair-based role of the central recombinase RAD51 and its paralogs in fork reversal and restart.

An active replication fork progresses as a three-way junction. However, following fork stalling the replication fork is remodelled into a four-way junction via a process called fork reversal, which rewinds the parental DNA duplex with the nascent strand protruding to generate a 'chicken foot' (Fig 4.17) (Bhat and Cortez, 2018). Fork reversal is of central importance in promoting genomic stability (Neelsen and Lopes, 2015; Bhat and Cortez, 2018). Additionally, numerous factors have been shown to contribute to fork reversal, including the chromatin remodeller 'SWI/SNF-related, Matrix-associated, Actin-dependent, Regulator of Chromatin, and subfamily A-Like 1' protein (SMARCAL1), 'Zinc finger, RAN-Binding domain containing 3' (ZRANB3), and RAD51.
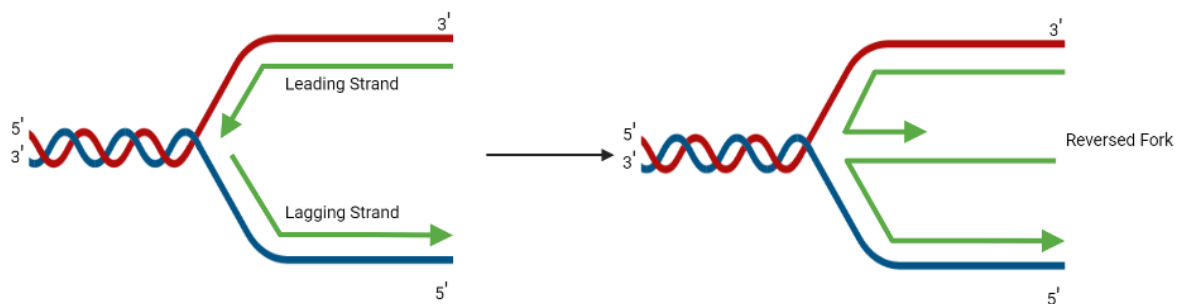
Figure 4.17. Stalled Replication forks can be converted to a reversed fork.
A stalled three-way replication fork can be converted to a four-way 'chicken-foot'
structure following fork reversal.

Although the role of RAD51 in fork reversal has been demonstrated to be universally involved in response to a wide range of genotoxic agents (Zellweger et al., 2015), the exact mechanism remains unclear. Current proposed models include RAD51 binding to ssDNA on the arm of the 'chicken foot' to promote fork reversal, or RAD51 invading the newly synthesised DNA to displace the template and reanneal the parental DNA (Bhat and Cortez, 2018). Regardless of the precise mechanism, extensive

evidence has shown RAD51 mediated fork reversal does not require either the strand exchange activity of RAD51 or interaction with BRCA2 (Kolinjivadi et al., 2017; Mijic et al., 2017; Mason et al., 2019). However, BRCA2-dependent stabilisation of RAD51 filaments on the regressed arm is essential for protection against fork degradation (Mijic et al., 2017; Kolinjivadi et al., 2017).

Following reversal, the replication fork can be restarted to complete DNA replication, which is directly facilitated by the RecQ Helicase RECQL1 (Berti et al., 2013) or SMARCAL1 (Kolinjivadi et al., 2017). RAD51 has also been implicated in the restart of stalled forks (Petermann et al., 2010), which is proposed to be facilitated by the RAD51 filaments on the regressed arm driving an HR-mediated restart by invading the DNA ahead of the replication fork (Petermann et al., 2010; Ait Saada et al., 2018). Of particular interest to the observations presented in this thesis is a study by Berti et al. (2020) on the role of RAD51 paralogs in the above processes (Berti et al., 2020). A sequential role of the two RAD51 paralogue complexes in RAD51-mediated reversal and restart was proposed, with the BCDX2 complex facilitating fork reversal, and CX3 functioning in fork restart (Berti et al., 2020).

We have demonstrated that the increase in fork restart deficiency observed after JMJD5/RCCD1 depletion is epistatic with that of the CX3 complex, indicative of these two complexes functioning in the same biological pathway (Fig 4.14B). Indeed, Western blot analysis showed increased RCCD1:JMJD5 expression after XRCC3 knockdown and vice versa (Fig 4.14A) potentially indicating a functional relationship between the proteins. However, further investigation is required to elucidate this connection and the potential mechanism behind it. Beyond CX3, whether the JMJD5:RCCD1 complex might also play a role in fork reversal is not clear, however. The absence of other BCDX2 components from our JMJD5:RCCD1 interactomics suggests that this may be unlikely, but this needs to be investigated. Perhaps consistent with no major role in fork reversal, investigation of the JMJD5 mutant patient cells showed no decrease in replication fork speed (Fletcher et al., 2023), which would not be expected if there was a significant deficit (Ray Chaudhuri et al., 2012).

Petermann et al. (2010) reported that RAD51 and XRCC3 depletion reduced the number of replication forks restarting after a short 2 h HU block, but longer HU treatment (24 h) led to collapsed replication forks that needed RAD51 mediated repair (Petermann et al., 2010). This is of particular interest here because the long HU treatment conditions used to uncover the discrepancy are similar to those used in our proteomic screens. However, our JMJD5 and RCCD1 fork restart assays involved a release from a much shorter HU block (Section 4.2.5), as is standard for this method. Therefore, future proteomic screens could include short HU treatments to investigate the JMJD5:RCCD1 interactome under conditions when CX3 is known to play a role in restart, as opposed to longer HU treatment where fork collapse may be more prevalent.

4.3.5   Chapter Conclusions

In this chapter, we have shown that RCCD1 and JMJD5 form a 1:1 heterodimeric complex, indicating that they may be 'obligate' binding partners. We used this new understanding to develop a multicistronic overexpression system for recombinant protein generation, preliminary structural characterisation, and interactome analysis. Our identification of CX3 as a novel JMJD5:RCCD1 interactor led us to demonstrate a potential role in replication fork restart, which further supports our observations presented in Chapter 3. The work raises important questions regarding the function of a RCCD1:JMJD5-CX3 pathway in replication fidelity and DNA repair and its role in neurodevelopmental disorders and cancer, which will be explored more in the Final Discussion (Chapter 5).

Chapter 5:   Final Discussion

The work presented in this thesis has attempted to elucidate how JMJD5 contributes to replication fidelity and, by extension, genome stability and tumour suppression. We have demonstrated that JMJD5 forms a heterodimeric complex with RCCD1 (Section 4.2.1) and that formation of this complex is required for the observed role in faithful DNA replication (Section 3.2.6 and 3.2.7). Although our work suggests that JMJD5 and RCCD1 act in a complex together, the targets and regulators of this complex, and their functional roles, are not fully understood. We hypothesise that the JMJD5:RCCD1 complex acts in a pathway (Fig 5.3), which we predict involves several upstream and downstream factors that would also control replication fidelity and be altered in cancer and neurodevelopmental disorders.

5.1   What is the function of the JMJD5:RCCD1 interaction?

Although we have demonstrated functional consequences of suppressing the expression and/or interaction of JMJD5 and RCCD1, including altered subcellular localisation (Section 2.2.9), protein expression (Section 4.2.1), and replication fidelity (Chapter 3), there are other areas that warrant further investigation. One key area not yet discussed is the potential relationship between heterodimeric complex formation and JMJD5 enzymatic activity. Indeed, our structural modelling and biochemical data suggests that the N-terminus of JMJD5 interacts with the catalytic domain (Fig 2.20). The JMJD5 N-terminus has a possible role in JMJD5 catalytic activity as the C123Y patient variant displayed reduced activity *in vitro* (Fletcher et al., 2023). Although the mechanistic role of the JMJD5 N-terminus in catalytic activity is unclear, it does contain the interaction site for RCCD1 (Chapter 2). This suggests the potential for RCCD1 binding to be communicated to the catalytic domain of JMJD5 through the N-terminus, but it is unclear what effect this may have on JMJD5 catalytic activity.

Others have reported a potential role for the N-terminus in regulating JMJD5 activity. Hsia et al suggested it may act as an autoinhibitory domain (Hsia et al., 2010), and others reported that RCCD1

binding released this inhibition to stimulate JMJD5 activity (Marcon et al., 2014). However, both studies were focused on the controversial histone demethylase activity of JMJD5, which remains in dispute. In the Coleman group, we have used full-length GST-tagged JMJD5 to study *in vitro* hydroxylase activity (Fig 2.19), however, we have not yet tested if this is increased upon removal of the N-terminus. We could explore this by comparing full-length GST-JMJD5 to the GST-JMJD5 183-416 fusion (Fig 2.5), which includes only the catalytic domain. Ideally, we would do this in the presence or absence of recombinant RCCD1 to explore the importance of RCCD1 binding to regulate any role of the JMJD5 N-terminus. However, this would require a source of purified recombinant RCCD1, which thus far has remained elusive. An alternative approach might be to compare the activity of JMJD5 versus JMJD5:RCCD1 *in vitro*, using protein purified from transient transfection or from the P2A system (Section 4.2.2). Although such approaches would be useful to uncover any direct effects of RCCD1 on JMJD5 hydroxylase activity, results would need to be carefully considered in the context of the potential limitations. For example, because the *in vitro* assay is based on a short linear peptide, which is unlikely to be representative of the physiological substrate, effects of RCCD1 on the binding and/or conformational change of a full-length substrate might be missed. In relation to this, it would be interesting to explore whether RCCD1 promotes the binding of candidate substrates such as CX3 (and CCDC130). In doing so, RCCD1 could help to target JMJD5 to its target(s) within specific subcellular domains, such as DSBs and/or replication forks. Indeed, there is precedent for RCCD1-like domains (RLDs) acting as subcellular targeting factors of enzymes: One side of the RCC1 RLD binds to histones H2A and H2AB, whilst the other side binds to and regulates the enzymatic activity of the Ran GTPase (Bamba et al., 2002). Future work should aim to characterise the precise localisation(s) of JMJD5 and its substrates, including foci formation in response to DNA damage and replication stress, and to determine the importance of RCCD1 in any observations made.

5.2 What are the targets of the JMJD5:RCCD1 complex?

With regards to factors acting downstream of JMJD5:RCCD1, we might envisage the existence of several activity-dependent (substrates) and -independent interactors involved in replication, HR, and/or other as yet unidentified cellular processes.

5.2.1 Novel candidate JMJD5 substrates

We have already discussed the candidacy of the CX3 complex as a novel, potentially HU-inducible, JMJD5 substrate in Chapter 4. However, our screens may have identified an additional candidate substrate in both untreated and HU-treated conditions. 'Coiled-coil domain containing 130' (CCDC130) was identified as a highly activity-dependent interactor, with no interaction detected with the catalytically dead JMJD5 mutant after HU treatment (Table 4.4). As presented in Section 4.2.4, this binding profile can be indicative of an enzyme:substrate interaction, which could be explored in future biochemical experiments using the same cellular and *in vitro* assays proposed (Section 4.3.4) for exploring the substrate candidacy of the CX3 complex.

Because CCDC130 is an orphan protein with no clear links to replication fidelity or DNA damage repair we did not investigate it further in this thesis, choosing to focus on the CX3 complex due to time constraints. CCDC130 is a poorly characterised protein with the only available literature suggesting it may be a splicing factor, having been identified as a spliceosome component via mass spectrometry (Schmidt et al., 2014). Because there is currently no evidence linking either JMJD5 or RCCD1 to splicing, the physiological relevance of this interaction with CCDC130 remains unclear. However, a closely related JmjC hydroxylases (JMJD6, Chapter 1) has been heavily studied in the context of splicing, so future studies should perhaps explore a role for the JMJD5:RCCD1 complex in splicing via CCDC130. It is equally plausible that CCDC130 plays an as yet undiscovered role in replication fork restart and/or HR. A role in replication fidelity and fork restart could be explored using the assays presented here, including with epistasis analyses to pinpoint the functional relationship with respect to JMJD5 and RCCD1 (Chapters 3 and 4).

5.2.2 Novel activity-independent JMJD5 targets?

Another interacting protein that appeared only after HU treatment, but was not activity dependent, was Fanconi anaemia group I (FANCI), which is part of the Fanconi anaemia (FA) pathway that functions to repair interstrand crosslinks (ICLs) (Ishiai, 2021). The FA pathway is encoded by 22 genes that are mutated in an inherited disorder characterised by congenital abnormalities and cancer predisposition (Nalepa and Clapp, 2018). These genes encode FA proteins that fall into four functional Groups 1) lesion recognition 2) the large E3 ligase core complex 3) the substrates of the ligase complex (FANCI:FANCD2) and 4) DNA damage associated proteins (Ishiai, 2021). The recognition of a lesion by Group 1 activates the core FANC complex, resulting in mono-ubiquitylation of FANCI:FANCD2, which is a central step in FA pathway activation (Ceccaldi et al., 2016). The activated FANCI:FANCD2 complex localises to damaged chromatin and interacts with many DNA damage proteins including BRCA1 (FANCS) and RAD51 (FANCR) (Ishiai, 2021). Of particular interest is the link between FANCI and the proposed roles of JMJD5 and RCCD1 in HR and replication fork restart. Many proteins involved in HR, including those discussed above (RAD51, BRCA1, PALB2, BRCA2 and RAD51C), are involved in the FA pathway due to the need for HR in ICL repair (Ceccaldi et al., 2016). Additionally, phosphorylation of FANCI by ATR was shown to be essential for replication fork restart, although the precise mechanism remains unclear (Chen et al., 2015b). Overall, the links between FANCI, replication, and HR, make it an intriguing potential JMJD5:RCCD1 interactor that warrants further investigation.

5.3 Regulation of the JMJD5:RCCD1 complex?

With regards to factors acting upstream of JMJD5:RCCD1, we might envisage the existence of enzymes that regulate the functional activity of the complex. One level of regulation applied to biological pathways can be through PTM.

To look for this potential regulation we searched our MS/MS screens for phosphorylation and detected a modification on RCCD1^{S268}, but only after HU treatment (data not shown), suggesting the phosphorylation event may be induced by replication stress and/or DNA damage. As several

phosphoproteomic screens have been undertaken to analyse the DDR and deposited to the online

database 'PhosphoSite Plus' (Hornbeck et al., 2015), we decided to search this database for

modifications on the JMJD5:RCCD1 complex. Although there were no highly reported modifications

on JMJD5, the phosphorylation on RCCD1^{S268} had been demonstrated in ten independent MS studies

(Fig 5.1). PhosphoSite Plus also listed one methylation site (R56) and one ubiquitination site (K271) in

RCCD1 (Fig 5.1), both reported five times.
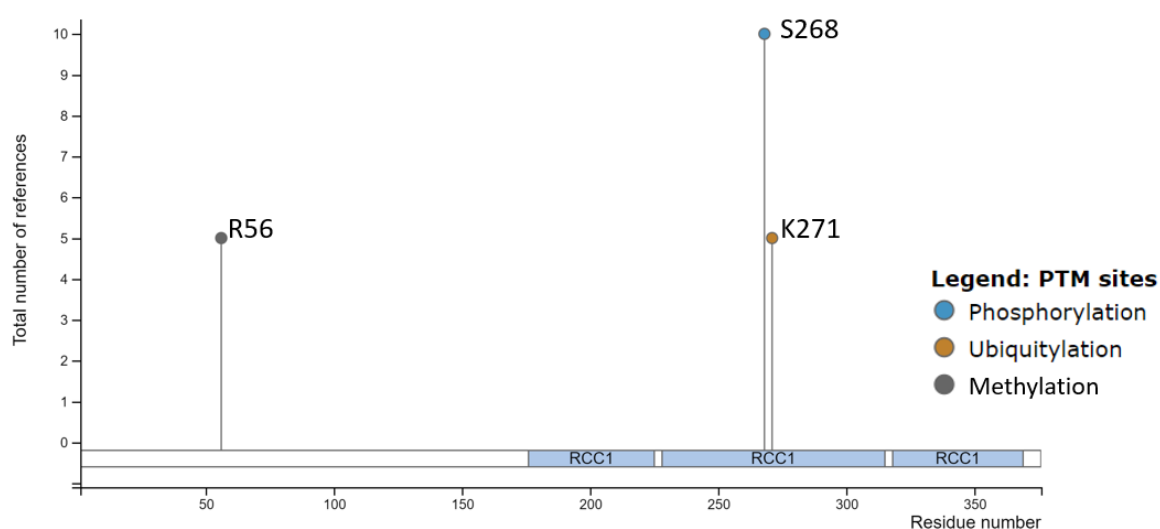
Figure 5.1. RCCD1 modifications reported on PhosphoSite Plus.
Prominent modifications of RCCD1 listed in the PhosphoSite Plus database, reference cut off was set at a minimum of 5. Image taken from PhosphoSite Plus (Hornbeck et al 2015).

As PhosphoSite had confirmed the RCCD1^{S268} phosphorylation, we further investigated the references

and found many of the curated screens had focused on identifying 'Ataxia telangiectasia mutated'

(ATM) and 'ATM-Rad3-related' (ATR) substrates. RCCD1^{S268} is located in an SQ motif, which ATM/ATR

are known to target (Traven and Heierhorst, 2005). The potential inducible nature of this

phosphorylation site is of interest given the roles of ATM and ATR in homologous recombination and

the replication stress response (Blackford and Jackson, 2017). ATM is central in the response to double

strand breaks where it phosphorylates hundreds of proteins including CtIP (discussed in section

4.3.4.2) to promote homologous recombination (Wang et al., 2013a). As mentioned in section 3.1.2.2,

RPA binding to ssDNA triggers the recruitment of ATR, which in turn phosphorylates downstream targets in the RS response. These kinases having key roles in HR and accurate DNA replication, suggests that they could possibly be regulators of the JMJD5:RCCD1 complex that warrant further investigation. However, it is unclear which of the above kinases is likely to catalyse this modification. Indeed, many ATM substrates are thought to be phosphorylated by ATR in response to RS (Blackford and Jackson, 2017; Matsuoka et al., 2007).

Also worthy of further investigation in relation to potential phosphoregulation is SPAG9, which was reproducibly identified in our screens and validated via Western blotting (Fig 4.8). SPAG9 is a C-Jun-amino-terminal kinase (JNK)-interacting protein, shown to be involved in both JNK and p38 signalling by acting as a scaffold that aggregates components of the MAPK cascades (Engström et al., 2010). It would be interest to explore whether SPAG9 scaffolds components of the JMJD5:RCCD1 pathway or regulates phosphorylation of JMJD5:RCCD1 by MAPKs, and the potential role of such biology in replication fidelity and HR.

Future work on RCCD1^{S268} phosphorylation should initially focus on attempting to validate this modification. One potential method could be to carry out repeat MS/MS screens under the presence of other DNA damaging agents, with orthogonal approaches using phospho-specific antibodies and Western blot analysis. The subsequent analysis should focus on attempting to identify the kinase responsible for this modification, with a particular focus on ATM/ATR.

Future work should also consider the potential site of RCCD1^{K271} ubiquitination, this could be verified through orthogonal approaches such as ubiquitin pull down assays and *in vitro* ubiquitylation assays. The identification of multiple proteins involved in the ubiquitin-proteasome pathway in our proteomic screens support the potential for ubiquitylation of the JMJD5:RCCD1 complex. Indeed, we identified KCTD9 which is relatively poorly characterised protein thought to belong to the cullin-3 ubiquitin ligase substrate specific adaptors (Ji et al., 2016). Also of particular interest is the identification of the large E3 ubiquitin ligase HUWE1/MULE/ARF-BP1 which targets a huge range of proteins, including some

involved in the DNA damage response (Kao et al., 2018). DDR-associated HUWE1 targets include BRCA1 (Wang et al., 2014b) and γH2AX (Atsumi et al., 2015). Whilst the functional relevance of the JMJD5:RCCD1 interaction with HUWE1 remains unclear, the possible link between JMJD5:RCCD1 and the repair of DSBs (Section 4.3.4.2) means HUWE1 warrants further investigation. Interestingly, investigation of the PhosphoSite references for RCCD1^{K271} ubiquitylation demonstrated one screen was done in the presence of UV treatment (Boeing et al., 2016) further suggesting inducible modification of RCCD1 in response to DNA damage. In this thesis, we have observed variable expression of RCCD1 and JMJD5 when overexpressed alone, potentially indicative of reduced protein stability and post-transcriptional regulation (Section 4.2.1). Is it possible these ubiquitin-proteasome components have a functional role in the above observations?

The proximity of the ubiquitination site on RCCD1^{K271} to the RCCD1^{S268} phosphorylation site raises the possibility for cross talk between these modifications. As it appears both modifications are inducible, is there a functional difference in the JMJD5:RCCD1 complex in the presence of each modification. For example, we have demonstrated JMJD5 activity is required to rescue increases in RS markers (Section 3.2.6 and (Fletcher et al., 2023)) and RCCD1^{S268} phosphorylation occurs in response to increased RS caused by HU treatment. Does this suggest RCCD1^{S268} phosphorylation is important for active JMJD5:RCCD1 complex? Conversely, is RCCD1^{K271} ubiquitination required to 'switch off' the active complex once its mechanistic role has been fulfilled? Interestingly, our AlphaFold2 multimer model (Section 2.2.13) shows K271 and S268 to be localised in the large loop region extending off the base of the RCCD1 ring shape (Fig 5.2A). Of further interest is the fact that this loop, and both these modified residues, may be in close proximity to the interaction interface between RCCD1 and JMJD5 (Fig 5.2B). This raises the possibility that these modifications could have key functional roles for the JMJD5:RCCD1 complex, including regulation of the interaction, activity of JMJD5 or being the signal for a conformational change to facilitate substrate binding.
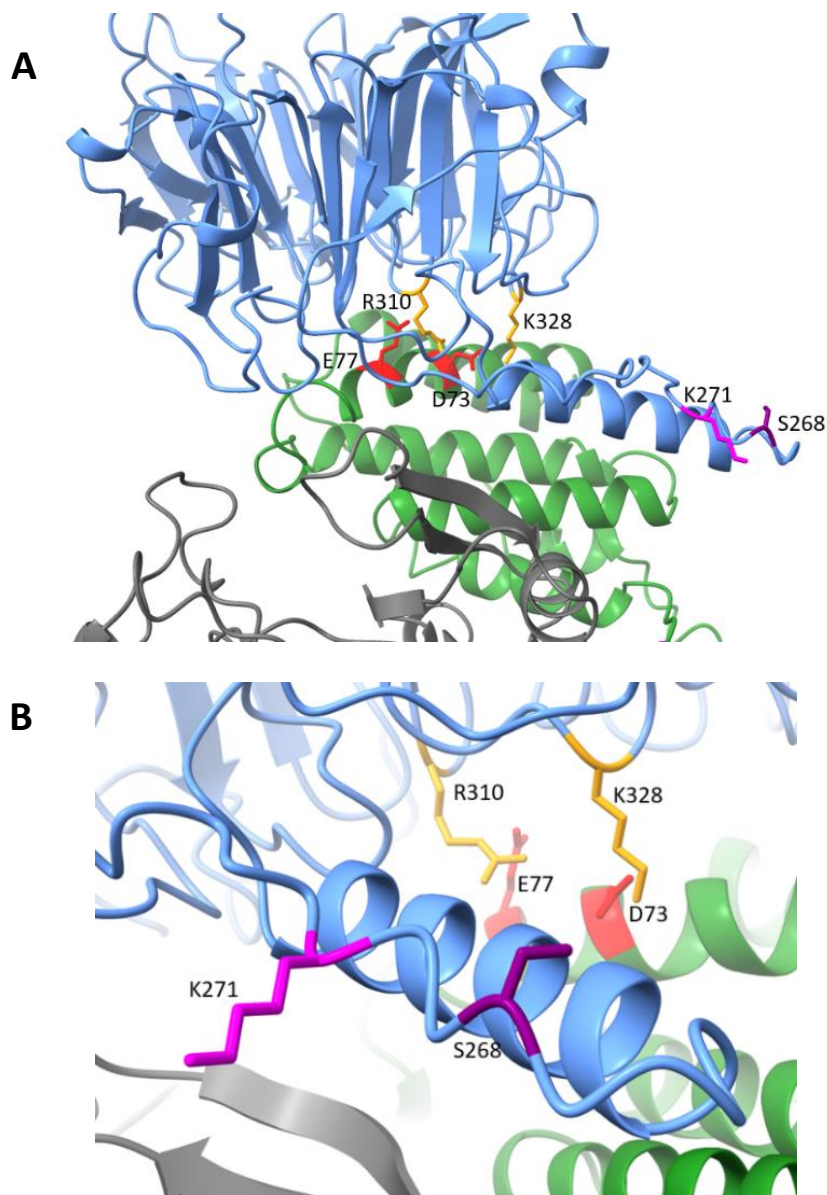
Figure 5.2. Potential regulatory RCCD1 modifications.
(A) Location of RCCD1 modification sites in the context of the predicted AlphaFold Multimer complex. **(B)** S268 and K271 are located in proximity with the RCCD1:JMJD5 interaction interface. Images coloured as follows RCCD1 (blue), N-terminus of JMJD5 (green), S268 in purple and K271 in magenta. Interacting residues R310 and K328 (RCCD1) are coloured in orange and D73 and E77 in red (JMJD5).

The final RCCD1 PTM listed on Phosphosite Plus is arginine methylation at position 56, which was reported in five MS studies. During our MS/MS analysis we did not search for arginine methylation, and so it was not detected. Therefore, to explore the potential for this modification in our samples we could start by reanalysing our existing data, and adding methylation as a variable modification to the data analyses of any new screens.

Arginine methylation is catalysed by 'Protein arginine methyltransferases' (PRMTs), which catalyse mono or dimethylation, with dimethylation being symmetric or asymmetric (Blanc and Richard, 2017). These different methylation marks are catalysed by specific PRMTs (Blanc and Richard, 2017). Type I enzymes (PRMT1, PRMT2, PRMT3, PRMT4, PRMT6, and PRMT8) catalyse asymmetric dimethylation, type II enzymes (PRMT5 and PRMT9) catalyse symmetric dimethylation, and both types form monomethylated arginine as an intermediate (Blanc and Richard, 2017). PRMT7 is the only type III enzyme and is unique amongst PRMTs as it only catalyses formation of monomethylated arginine (Blanc and Richard, 2017). Interestingly, the RCCD1[56] arginine methylation was reported to be a monomethylation (Guo et al., 2014), which could suggest it was catalysed by PRMT7. Although, the exact PRMT responsible for this modification requires further investigation, it is interesting to note that R56 is surface accessible and located on the opposite side of RCCD1 to the JMJD5 interaction surface (data not shown). Although the molecules interacting with this surface are currently unknown, it is interesting to speculate that R56 methylation could regulate their binding affinity. Interestingly, many reports link arginine methylation with DSB repair and accurate DNA replication (Mersaoui et al., 2019; Boisvert et al., 2005; Sanchez-Bailon et al., 2021; Clarke et al., 2017). This correlation with JMJD5/RCCD1 biology suggests that further investigation of the RCCD1 methylation site and its potential regulatory impact on the function of the RCCD1:JMJD5 complex, and its interactions, would be of great interest.

5.4   Inactivation of the JMJD5:RCCD1 pathway in human disease

A key aim of our investigation into JMJD5 was to further explore the connection between this poorly characterised protein hydroxylase and human disease, particularly cancer and developmental disorders. Through our application of cancer variants to help map determinants of the JMJD5:RCCD1 interaction, we have provided evidence of the functional impact of disease mutations (Chapter 3), which in turn may support the role of JMJD5 and RCCD1 as novel tumour suppressors. TSGs are known to be inactivated by a range of mechanisms in addition to genetic mutation, including epigenetic

silencing, altered splicing and deletion, which could be fruitful areas of future investigation for JMJD5 and RCCD1. Of particular interest, due to the unique nutrient and cofactor dependencies of 2OG-oxygenases, is the possibility that the tumour microenvironment could also impair the TSG function of the JMJD5:RCCD1 complex.

'Oncometabolites', such as fumarate, succinate and D-2-hydroxyglutarate (D-2HG), are now thought to have important roles in tumourigenesis. These metabolic intermediates occur due to mutations in specific metabolic enzymes including the Krebs cycle enzymes Fumarate Hydratase (FH), Succinate Dehydrogenase (SDH), and Isocitrate Dehydrogenase 1 or 2 (IDH1/2) (Collins et al., 2017). These oncometabolites share structural homology to 2OG and can thereby act as competitive inhibitors of 2OG oxygenases. Indeed, there is increasing evidence implicating the inhibition of certain 2OG oxygenases by oncometabolites as an important driver in some cancers (Selak et al., 2005; Figueroa et al., 2010; Xu et al., 2011). Of particular importance to the work in this thesis is the evidence that accumulation of 2-HG (Sulkowski et al., 2017), succinate, and fumarate (Sulkowski et al., 2018) can lead to deficiencies in HR. Although the latter has been assigned to inhibition of KDM4A/B (Sulkowski et al., 2018)(Sulkowski et al., 2018)(Sulkowski et al., 2018)(Sulkowski et al., 2018), it is unclear whether inhibition of other 2OG-oxygenases, such as JMJD5, might also contribute. Therefore, further investigation into possible JMJD5 inactivation by oncometabolites is of great interest.

Another potential mechanism of JMJD5 inactivation in tumours is reduced oxygen availability, particularly because hypoxia has been shown to suppress HR, alter DNA replication and promote genome instability (Bristow and Hill, 2008). Although some 2OG oygenases, such as the HIF PHD hydroxylases, are very sensitive to hypoxia and have been widely studied (Section 1.1.5.3), there has been limited investigation into the regulation of the JmjC-only subfamily by oxygen tension. We propose that this could be a valuable avenue of further investigation for JMJD5, particularly because of increasing evidence demonstrating that hypoxia contributes to replication stress: Several studies have demonstrated that acute severe hypoxia ($\leq 0.1\%$ O_2) increases stalled forks and reduces fork

speed, indicative of RS, which in turn induces activation of ATM and ATR in the absence of DNA damage (Hammond et al., 2002, 2004; Bencokova et al., 2009; Pires et al., 2010). The precise molecular mechanism involves inactivation of the ribonucleotide reductase (RNR) that generates dNTPs for use in DNA replication (Ng et al., 2018). Interestingly, chronic periods (>12 hours) of severe hypoxia also prevent fork restart following reoxygenation (Pires et al., 2010). Also as previously discussed, deficient fork restart was also observed by Petermann et al (2010) following a long DNA replication block (24 hours) through inhibition of RNR by HU treatment (Petermann et al., 2010).

Although the mechanisms involved in the response of replication to chronic severe hypoxia remain unclear, the authors observed disassembly of the replisome (Pires et al., 2010). It would be of interest to explore whether impaired JMJD5 function under these conditions might also be a contributing factor, considering the results presented in this thesis. Exploring the oxygen sensitivity of JMJD5 in cells will be important for such studies, which will be greatly facilitated by identification of the physiological substrate(s).

As mentioned above, we began investigating the relationship between JMJD5 and RCCD1 to elucidate a possible pathway that contributes to human disease, and in the process identified new interactors. Interestingly, many of these interactors have clear links to both developmental disorders and tumourigenesis much like JMJD5. For example, aberrant HUWE1 function through inherited missense mutations contributes to neurodevelopmental disorders including but not limited to Kabuki-like and Say-Meyer syndromes (Giles and Grill, 2020). Additionally, HUWE1 has been shown to have TSG function, suppressing proliferation of prostate and thyroid cancer (Gong et al., 2020). RAD51C and FANCI both have intrinsic roles in Fanconi Anaemia which leads to congenital defects and susceptibility to cancer (Ceccaldi et al., 2016). Additionally, RAD51C germline pathogenic variants have been extensively catalogued over the past decade increasing the evidence for RAD51C as a cancer predisposition gene (Boni et al., 2022). A polymorphism in XRCC3 that causes mutation of threonine at position 241 to leucine is associated with increased risk of numerous cancer types including bladder

(Li et al., 2011), glioma (Tan et al., 2022)(Tan et al., 2022)(Tan et al., 2022)(Tan et al., 2022) and ovarian cancer (Hu et al., 2022). Further investigation into the mechanistic role of these relationships, could provide further insight to mechanisms underlying developmental disorders and cancer progression. Therefore, increasing the possibilities for therapeutic intervention.

5.5 Final Conclusions

Here, in this thesis, we have expanded our understanding of both the function of JMJD5 and the biochemical basis and importance of its interaction with RCCD1, demonstrating that the two proteins function as obligate binding partners in a heterodimeric 1:1 complex. We have shown that this complex may have a central role in a novel pathway facilitating successful replication fork restart and efficient HR (Fig 5.3), through interaction with the RAD51 paralog complex CX3, and other newly identified binding partners. The work provides a solid foundation for further characterisation of this new pathway, providing many avenues to investigate its role and regulation in health and disease in the future (Fig 5.3).
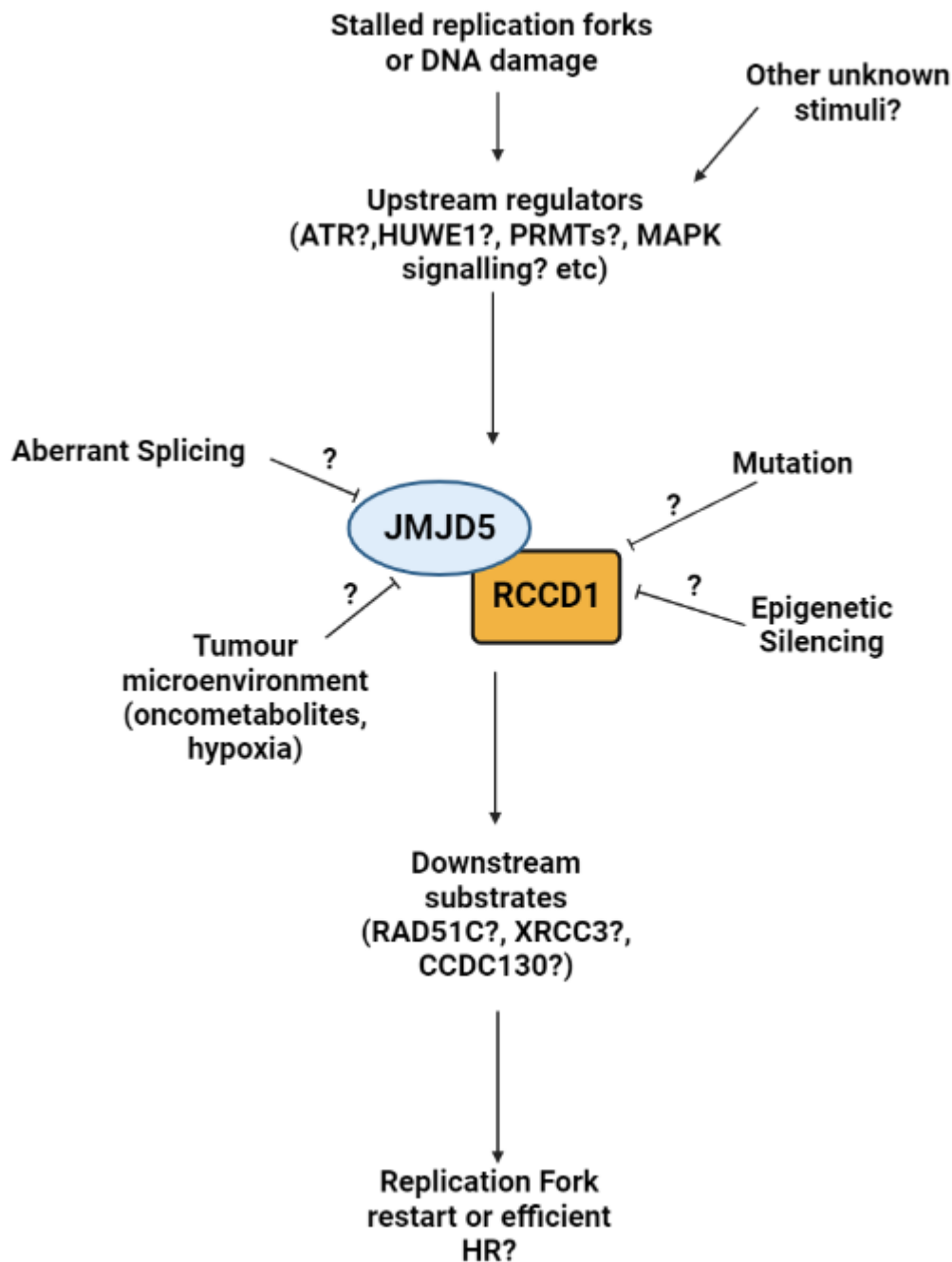
Figure 5.3. Working model for JMJD5:RCCD1 complex pathway in replication fork restart.
The activity of the JMJD5:RCCD1 complex may be regulated by upstream proteins in response to replication stress or DNA damage. JMJD5:RCCD1 in turn may hydroxylate downstream substrates to promote replication fork restart and efficient HR. Additionally, this complex may be inhibited through mutation, aberrant splicing, epigenetic silencing or the tumour microenvironment.

Chapter 6: Materials and Methods

6.1 Cloning procedures

6.1.1 Vectors and Plasmids

All constructs used in this study were cloned from existing pEF6 HA-JMJD5 (full-length/wildtype) or pcDNA3 RCCD1-HA (full-length/wildtype) vectors with each construct cloned into pEF6, pGEX-4T1, pIHZ or pTIPZ. pIHZ is a derivative of 'pGIPZ' (Dharmacon) that lacks the green fluorescent protein (GFP) cDNA and substitutes puromycin (P) for hygromycin (H) resistance. Additionally, pTIPZ is a modification of the pTRIPZ commercial vector with the red fluorescent protein (RFP) cDNA removed (Dhamarcon). The p3XFLAG-myc-CMV™-26 (RAD51C and XRCC3) vectors were a kind gift from Professor Grant Stewart.

6.1.2 Polymerase chain reactions

Polymerase chain reactions (PCRs) were carried out using Phusion® High-fidelity DNA polymerase (NEB). After PCR, samples were run on a 1% (w/v) agarose gel, the amplified DNA was excised and solubilised using the Gen-Elute™ Gel Extraction kit (Sigma-Aldrich). The resulting PCR products and desired vectors were digested at 37°C for one hour with relevant restriction enzymes. The digested vectors were additionally treated with calf intestinal alkaline phosphatase (CIP, NEB, 0290S). The resulting digests were run on a 1% (w/v) agarose gel and recovered through the Gen-Elute™ Gel Extraction kit. The inserts were ligated into plasmids using T4 DNA ligase (NEB) at 25°C for 10 minutes followed by heat inactivation at 65°C for 10 minutes. Ligation mixtures were subsequently transformed into *E. coli* DH5α competent bacteria and the plasmid DNA was purified using the Gen-Elute Plasmid Miniprep kit (Sigma-Aldrich), and sequence verified using SourceBioscience. Primers used for PCR cloning are listed in Table 6.1.

Table 6.1. List of primers used for PCR cloning

Construct	Primers 5' to 3'
pEF6 and pGEX_4T1 JMJD5 1-80	Forward: TTTATCGAATTCGCTGGAGACACCCACTGC
	Reverse: GATAAAGTCGACCTAGTTGAGCTTCTCCCAGGAGTAGTCCA
pEF6 and pGEX_4T1 JMJD5 1-110	Forward: TTTATCGAATTCGCTGGAGACACCCACTGC
	Reverse: GATAAAGTCGACCTAAGGTGCCTGGCACAGACACAG
pEF6 and pGEX_4T1 JMJD5 1-150	Forward: TTTATCGAATTCGCTGGAGACACCCACTGC
	Reverse: GATAAAGTCGACCTAAGGGAGGTGTGTCTGGAGGATGG
pEF6 and pGEX_4T1 JMJD5 1-182	Forward: TTTATCGAATTCGCTGGAGACACCCACTGC
	Reverse: GATAAAGTCGACCTATTTTTCTAACTTCACATCTGGAATCAAACCATGGTC
pEF6 and pGEX_4T1 JMJD5 81-416	Forward: TTTATCGAATTCACGGGCACATGGCAGGACG
	Reverse: GATAAAGTCGACCTACGACCACCAGAAGCT
pEF6 and pGEX_4T1 JMJD5 111-416	Forward: TTTATCGAATTCGAGGATGCCAACACTGTGGCCG
	Reverse: GATAAAGTCGACCTACGACCACCAGAAGCT
pEF6 and pGEX_4T1 JMJD5 151-416	Forward: TTTATCGAATTCGGAAAGAGGCCTGCCCGTGG
	Reverse: GATAAAGTCGACCTACGACCACCAGAAGCT
pEF6 and pGEX_4T1 JMJD5 183-416	Forward: TTTATCGAATTCACAGTCCCCCGGCTGCACC
	Reverse: GATAAAGTCGACCTACGACCACCAGAAGCT
pIHZ RCCD1-3xFLAG	Forward: TTTATCATCGATATGGCGGAGGAGCGG
	Reverse: GATAAAGCGGCCGCCTACTTATCATCGTCATCCTTGTAATCAATGTCATGATCTTTATAATCACCATCATGGTCTTTGTAGTCGCTCTTCCCTTTCTCCACAGC
pTIPZ 3xFLAG-JMJD5	Forward: TATGCAGGCGCGCCATGGGAGACTACAAAGACCATGATGGTGATTATAAAGATCATGACATCGACTACAAGGATGACGATGATAAGGCTGGAGACACCCACT
	Reverse: AATATACCTGCAGGCTACGACCACCAGAAGCT

6.1.3 Site-directed mutagenesis

Primers for site-directed mutagenesis (SDM), were designed with the point mutation in the centre of the forward primer with the reverse primer annealing back-to-back. Prior to use, the primers were phosphorylated using a T4 polynucleotide kinase (ThermoFisher, EK0031). A Phusion SDM kit (ThermoFisher, F541) was used to generate the point mutations according to manufacturer's instructions. JMJD5^{L79A} was generated using JMJD5^{L79P} as a template and JMJD5$^{D73A/E77A}$ used JMJD5^{D73A} as a template. The RCCD1$^{R310M/K328M}$ vector was generated through sequential SDM reactions of RCCD1^{WT}. Primer sequences are listed in Table 6.2. All constructs were confirmed by sequencing as above (6.1.2).

Table 6.2. Primers used for site-directed mutagenesis

Mutation	Primers 5' to 3'
JMJD5 D4N	Forward: GAATTCGCTGGAAACACCCACTGC
	Reverse: GGATCCAGCGTAATCTGGAACATCGTATG
JMJD5 P27L	Forward: GCGCTCCTGCTGCACAGTAAA
	Reverse: CCTGAGGGCCTCCCATAAAGTG
JMJD5 L35F	Forward: GAAGACCTGAAGTTTGACCTCGGGGA
	Reverse: TTTGTCTACGTCCTGCCATGTGCCC
JMJD5 E39K	Forward: GACCTCGGGAAGAAAGTGGAG
	Reverse: CAACTTCAGGTCTTCTTTACTGTGC
JMJD5 V45M	Forward: GTGGAGAGGAGCATGGTGACATTGTTG
	Reverse: TTTCTCCCCGAGGTCCAACTTCAG
JMJD5 V70L	Forward: AGCAGCGAGTTGATCCTGGAC
	Reverse: CTGCAGACACTCGTCCCTCC
JMJD5 S75F	Forward: ATCCTGGACTACTTCTGGGAGAAG
	Reverse: CACCTCGCTGCTCTGCAG
JMJD5 E77K	Forward: GACTACTCCTGGAAGAAGCTCAACAC
	Reverse: CAGGATCACCTCGCTGCTCTG
JMJD5 L79P	Forward: TGGGAGAAGCCCAACACGGGC
	Reverse: GGAGTAGTCCAGGATCACCTCGCT
JMJD5 D90N	Forward: GTAGACAAAAACTGGCGCCGG
	Reverse: GTCCTGCCATGTGCCCGT
JMJD5 R92C	Forward: AAAGACTGGTGCCGGGTCTAC
	Reverse: GTCTACGTCCTGCCATGTGC
JMJD5 R121Q	Forward: GCAGCCCTGCAGGTCTGTGAC
	Reverse: GGCCACAGTGTTGGCATCCTCA
JMJD5 I144V	Forward: TGTGACATGTGCCTGCTGATG
	Reverse: TTTAAGAAGGATGTCCCCCAGGATGG
JMJD5 R152S	Forward: CCTGGAAAGAGTCCTGCCCGT
	Reverse: GAGGTGTGTCTGGAGGATGGCAG
JMJD5 D73A	Forward: GTGATCCTGGCCTACTCCTGGGA
	Reverse: CTCGCTGCTCTGCAGACACTC
JMJD5 W76A	Forward: GACTACTCCGCGGAGAAGCTC
	Reverse: CAGGATCACCTCGCTGC
JMJD5 E77A	Forward: GACTACTCCTGGGCGAAGCTCAAC
	Reverse: CAGGATCACCTCGCTGCTCTG
JMJD5 K78A	Forward: TCCTGGGAGGCGCTCAACACG
	Reverse: GTAGTCCAGGATCACCTCGCT
JMJD5 L79A (uses JMJD5 L79P as template)	Forward: TGGGAGAAGGCCAACACGGGC
	Reverse: GGAGTAGTCCAGGATCACCTCGCT
JMJD5 D73A/E77A (uses JMJD5 D73A as template)	Forward: TACTCCTGGGCGAAGCTCAAC
	Reverse: GGCCAGGATCACCTCGC

Table 6.2 Primers used for site-directed mutagenesis continued

Mutation	Primers 5' to 3'
RCCD1 R141S	Forward: CTGCGCCAGTGCCTACG
	Reverse: GGCAGCAGGGGTAGCCT
RCCD1 R152Q	Forward: CCCTTCTACCAGCCTCTGGCT
	Reverse: CGCCCGCGGGCTCAC
RCCD1 R247S	Forward: CTGCCCACCAGCAACCTGGC
	Reverse: GGCCAGCTGCCCTGATTCATTCCA
RCCD1 R258W	Forward: AGACTGTCGCATGGGAAGCCACAGA
	Reverse: TCTCTCCATCCTCTGCCAGG
RCCD1 R310W	Forward: TGTGGATCCTGGCACACAGCT
	Reverse: GCTGGCCTTGACTGCATCTGAG
RCCD1 K328N	Forward: TGGGGCTGGGGTAACTATGGACAGCTG
	Reverse: GGTGTAGAGCTCCCCTGTTCGTGT
RCCD1 R310M	Forward: TGTGGATCCATGCACACAGCT
	Reverse: GCTGGCCTTGACTGCATCTGAG
RCCD1 K328M	Forward: TGGGGCTGGGGTATGTATGGACAGCTG
	Reverse: GGTGTAGAGCTCCCCTGTTCGTGT
RCCD1 WT siRCCD1 #1 resistant	Forward: GCTAGAGAAGCCACAGAACTGAATGAAGAT
	Reverse: GACTGTTTCTCCATCCTCTGCCAGG

6.2 Cell biology techniques

6.2.1 Cell culture

Human embryonic kidney 293 T-antigen (HEK293T), A549 and U20S cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) with the addition of 1% penicillin/streptomycin and 10% (v/v) Fetal Bovine Serum (FBS). Stable cell lines were cultured as above with the addition of 200 µg/mL Hygromycin B or 1 µg/mL puromycin. All cell lines were cultured in a humidified incubator at 5% CO_2 and 37°C.

6.2.2 DNA plasmid Transfection

HEK293T or A549 cells were seeded into 10 cm dishes 24 hours in advance, for between 30-50% confluency at the time of transfection. Transfections were carried out using FuGENE® HD (Promega) according to manufacturer's instructions. FuGENE® HD was used at a ratio of 3 µL per 1 µg of plasmid DNA. The DNA was added to Opti-MEM® reduced serum medium and briefly vortexed, FuGENE was then added, vortexed and incubated at room temperature (RT) for 30 minutes. The transfection

mixture was added dropwise to the cell culture and gently mixed. Cells were harvested 48 hours after transfection.

6.2.3   Small Interfering RNA (siRNA) Transfection

siRNA transfections were carried out using the transfection reagent Lipofectamine™ RNAiMAX (ThermoFisher) as per manufacturer's instructions. U2OS or A549 cells were seeded into 10 cm dishes 24 hours prior to transfection for 30-50% confluency at time of transfection. 8 μL of siRNA (Final concentration 30 nM, Stock 50 μM) was added to 500 μL of Opti-MEM and a pre-mix of RNAiMAX was prepared containing 10 μL of RNAiMAX and 500 μL of Opti-MEM per transfection. 500 μL of this pre-mix was added to each tube containing siRNA and incubated at RT for 15 minutes. The transfection mixture was then added dropwise to the cells, after the media on the cells was replaced with fresh media. Cells were harvested 72 hours after transfection. This protocol was scaled according to cell culture plate size.

Table 6.3. List of siRNAs

siRNA target	Sigma siRNA ID	Referred to in text as:
SIC001 MISSION® siRNA Universal Negative Control #1	SIC001	siCtrl
RCCD1	SASI_Hs01_00084850	siRCCD1 #1 / siRCCD1
RCCD1	SASI_Hs01_00084852	siRCCD1 #2
JMJD5	SASI_Hs01_00082889	siJMJD5 #1
JMJD5	SASI_Hs01_00082891	siJMJD5 #2 /siJMJD5
XRCC3	SASI_Hs02_00321675	siXRCC3

6.2.4   Generation of stable cell lines

To produce lentivirus, HEK293T cells were transfected, using FuGENE (see Section 6.2.2), with 0.5 μg of the lentiviral vector and the packaging vectors pMD2.G (0.15 μg) and psPAX2 (0.35 μg). The cells were incubated for 48 hours, and the virus-containing medium was collected and filtered. The recipient A549 cells were incubated with the viral supernatant for 24 hours. The successfully

transduced cells were then selected using either 200 μg/mL Hygromycin B (Invitrogen) or 1 μg/mL Puromycin (Gibco).

6.3 Biochemical techniques

6.3.1 Whole cell lysate preparation

Cells were washed twice in cold Phosphate Buffered Saline (PBS) and then lysed in an appropriate volume of JIES-buffer (100 mM $NaCl_2$, 20 mM Tris-HCl pH 7.4, 5 mM $MgCl_2$, 0.5% (v/v) IGEPAL CA-630 and complete protease inhibitors (Roche) or RIPA (150 mM NaCl, 25 mM Tris pH 8, 1% (v/v) NP40, 0.5% (w/v) Sodium deoxycholate, 0.1% (w/v) SDS). The samples were then pelleted at 4°C, 14,000 rpm for 10 minutes. For Western blotting, protein concentration was then calculated using the Pierce 660 nm reagent (ThermoFisher) and normalised using the appropriate buffer.

6.3.2 Immunoprecipitation

Samples were harvested in JIES buffer and lysed by rotation at 4°C for 1 hour. Lysates were then pelleted as above (Section 6.3.1). An input fraction was subsequently collected and 6x Laemmli buffer (125 mM Tris-HCl [pH6.8], 6% (w/v) SDS, 50% (v/v) glycerol, 225 mM dithiotheitol (DTT), 0.1% (w/v) bromophenol blue) was added to a 1x final concentration, after which the samples were boiled at 95°C for five minutes. Then anti-HA agarose beads (Sigma-Aldrich Cat.No: A2095) or anti-FLAG® M2 Magnetic Beads (Sigma-Aldrich, M8823) were added and incubated at 4°C, rotating overnight (ON). The beads were washed 6 times in JIES buffer, followed by elution using appropriate volume of elution buffer (100 μg HA or 100 μg FLAG peptide, 20 mM Tris pH8 and 100 mM NaCl). The elution buffer was added to the lysate and the samples were left shaking at RT in a Thermomixer Compact (Eppendorf) at 2000 rpm for 15 minutes. The supernatant was removed from the beads and an appropriate volume of 6x Laemmli buffer was added, and the samples boiled at 95°C for 5 minutes.

6.3.3 Western blotting

Protein samples, boiled in Laemmli buffer, were resolved using sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). Samples were prepared as in 6.3.1 or 6.3.2 and loaded onto a homemade 12% polyacrylamide gel or a Novex™ WedgeWell™ 8-16% gels (Invitrogen). Electrophoresis was performed in Tris-Glycine SDS running buffer (25 mM Tris-HCl pH 8.3, 192 mM glycine, 0.1% (w/v) SDS) using the Mini PROTEAN Tetra apparatus (Bio-Rad), at 160 volts. Proteins were then transferred onto an Immobilon-FL polyvinylidene difluoride (PVDF) membrane activated in 100% methanol. To assemble the transfer cassette, the gel and membrane were layered in between 4 pieces of Whatman filter paper whilst submerged in transfer buffer (25 mM Tris pH 8.3, 192 mM glycine, 20% (v/v) methanol). Gels were transferred at 320 mA at 25 minutes per gel. Membranes were then blocked in 5% (w/v) milk at RT for one hour. After blocking, membranes were incubated in primary antibody made up in 5% (w/v) milk in PBS-T (Phosphate-Buffered Saline with 0.1% (v/v) Tween) for either 1 hour at RT or ON at 4°C. Following this incubation, membranes were washed three times in PBS-T and incubated with HRP-tagged secondary antibodies for 1 hour and washed as before. Membranes were then developed using the Clarity Max Western ECL (Bio-Rad) or Femto (ThermoFisher) blotting substrates and imaged using a Vilber Lourmat FusionFX imager. The full list of antibodies used is shown in Table 6.4.

Table 6.4. List of all antibodies used for Western blotting in this project

Antigen	Source	Catalogue Number	Dilution used
HA HRP	Roche	12CA5	1 in 2000
FLAG HRP	Sigma	A8592	1 in 10000
β-actin HRP	Abcam	Ab49900	1 in 25000
Myc tag	Abcam	Ab9132	1 in 1000
RCCD1	Abcam	Ab122570	1 in 250
JMJD5	Collaborators Matsuura Yoshiharu Lab, Japan		1 in 250
SPAG9	Cell Signalling Technology	5519S	1 in 1000
XRCC3	Novus Biologicals	NB100-165SS	1 in 250
Anti-Rabbit HRP	NEB	7074	1 in 2000
Anti-Mouse HRP	NEB	7076	1 in 2000
Anti-Goat HRP	Santa Cruz Biotechnology	sc-2028	1 in 2000

6.3.4 Recombinant protein expression

BL21 competent *E. coli* were transformed with each pGEX-4T1 GST-JMJD5 vector and inoculated into 2x Luria Broth (LB) and grown in a shaking incubator at 37°C and 200 rpm ON as a starter culture. The following day, the starter cultures were back diluted 1:100 into the appropriate volume of 2xLB in a conical flask (100 mL for expression trials, 1 L for larger scale purification). This culture was grown at 37°C until the culture reached an optical density (OD) of 0.6 at 600 nm.

For small scale expression trials, once the culture reached 0.6 OD, 1 mL of culture was collected in an Eppendorf and attached to the conical flask (non-induced fraction). Isopropyl β- d-1-thiogalactopyranoside (IPTG) was added to the culture at a concentration of 0.5 mM and incubated at 18°C shaking at 200 rpm ON. The induced fraction of 1 mL of culture was collected and both the –IPTG and +IPTG fractions were then centrifuged at 6000 rpm for five minutes. The –IPTG and +IPTG fractions were then lysed in 0.5 mL Lysis buffer (Table 6.5). These samples were frozen at -80°C for one hour before thawing (freeze-thaw to promote lysis) and incubated rotating at 4°C for one hour. The samples were then pelleted at 14000 rpm, 4°C for 10 minutes. The supernatant was collected and 50 µL was taken from this to represent the soluble fraction, and 10 µL of 6x Laemmli buffer was then added. 20

µL of glutathione-sepharose beads (GE, 17-07560-01) were added to the remaining supernatant and incubated rotating at 4°C for one hour. The insoluble fraction was prepared by resuspension of the insoluble pellet in 500 µL of 1xSB and sonicated. The glutathione beads were washed 6 times in lysis buffer (Table 6.5) and all the supernatant was removed. 50 µL of 1xSB was then added to the dried beads forming the pulldown fraction. All fractions were then boiled at 95°C for five minutes.

Table 6.5. Lysis buffer recipe for recombinant protein purification

Reagent	Stock	Final
Tris pH 8	1M	50mM
NaCl	5M	300mM
Tris(2-carboxyethyl)phosphine		0.5mM
Protease Inhibitor Cocktail		1 Tablet
Turbonuclease		1:10000
Lysozyme		0.5mg/ml
Triton X-100	100%	0.10%

The larger scale purifications proceeded as above, with no 1 mL fractions collected. Instead, after ON induction at 18°C the culture was pelleted at 4000 rpm for 20 minutes and lysed in 100 mL of lysis buffer (Table 6.4). Following lysis, the samples were treated as above with freeze-thaw, pelleting at 14000 rpm for 15 minutes and incubation of the supernatant with 1 mL glutathione beads (ON, rotating at 4°C). The beads were washed six times in wash buffer (0.1% (v/ v) Triton X-100, 50 mM Tris pH 8.0, 300 mM NaCl, and 0.5 mM TCEP) before elution using 1 mL of 10 mM glutathione (Sigma-Aldrich), 300 mM NaCl, 50 mM Tris (pH 8.0), and 0.5 mM TCEP. The samples were incubated for three sequential elutions, rotating at RT for 10 minutes. These elutions were combined and added to a 7,000-molecular-weight-cutoff Slide-A-Lyzer dialysis cassette (Pierce) for ON dialysis against 50 mM Tris (pH 8.0), 300 mM NaCl, 0.5 mM TCEP. Finally, the dialyzed samples were concentrated using a 10,000-molecular-weight-cutoff protein concentrator (Pierce).

6.3.5 Hydroxylation Assays

RPS6 peptide ([129]VPRRLGPKRASRIRKL[144]) was prepared at a stock concentration of 400 µM in substrate buffer (150 mM NaCl, 1 mM DTT and 50 mM Tris-HCl pH 7.5). A stock concentration of $(NH_4)_2Fe(SO_4)_2$ was prepared at 500 mM by dilution in 20 mM HCl, which was further diluted in MilliQ H_2O prior to use at 10 mM. A 2X master mix of co-factors and buffers was prepared containing 400 µM 2-oxoglutarate, 1 mM ascorbate, 200 µM Fe(II), and 50 mM HEPES pH 7.5. A final reaction volume of 20 µL consisted of 10 µL 2X co-factor mix, 5 µL of control buffer (300 mM NaCl, 0.5 mM TCEP and 50 mM Tris pH 8.0) or 20 µg GST-tagged recombinant JMJD5 (WT, H321A, E77K, D73A, E77A or D73A/E77A), and 5 µL of substrate buffer or RPS6 peptide. The reaction was briefly mixed before incubation at 37°C for one hour, following which the reactions were quenched in 1% (v/v) formate.

The quenched samples were used in the Succinate-Glo™ JmjC demethylase/hydroxylase assay (Promega). 5 µL of quenched sample was loaded into triplicate wells on a 384 white walled plate, 5 µL of Succinate-Glo detection reagent was added to each well and the plate was sealed with parafilm. The plate was briefly mixed on a thermomixer for 15 seconds at 600 rpm and incubated at RT for one hour. 10 µL of Succinate-Glo detection reagent II was added, the plate was resealed followed by a subsequent incubation at RT for 10 minutes. Luminescence signal was then read using a PerkinElmer Enspire plate reader.

6.4 Microscopy techniques

All images were taken using a Leica DM6000 microscope with a 40x objective lens using the LAS X software. All immunofluorescence images of similarly stained experiments were acquired using identical illumination settings. Images were processed with ImageJ and the final images were assembled using Microsoft PowerPoint.

6.4.1   Immunofluorescence staining

Cells were grown on coverslips and fixed using either methanol for 10 minutes at -20°C or 4% (w/v) Paraformaldehyde (PFA) in PBS for 15 minutes at RT. For 53BP1 staining, coverslips were incubated in pre-extraction buffer (20 mM NaCl, 3 mM MgCl$_2$, 300 mM sucrose, 10 mM PIPES, 0.5% (v/v) Triton X-100) on ice for five minutes prior to fixing in PFA. The cells were subsequently permeabilised in 0.1% (v/v) Triton X-100 in PBS at RT for 10 minutes. The coverslips were then blocked in 1% (w/v) Bovine serum albumin (BSA) in PBS for one hour at RT. The coverslips were then incubated in primary antibodies in 1% (w/v) BSA/PBS for one hour at RT, in a humidified dark plate. After this, coverslips were washed three times in 1% (w/v) BSA/PBS before incubation with secondary antibodies at 1:1000 dilution in 1% (w/v) BSA/PBS for one hour at RT. To stain nuclei, 4',6 diamidino-2-phenylindole (DAPI)(Invitrogen™) was used, and the coverslips were mounted onto microscope slides with Prolong® Gold Antifade Reagent (Cell Signaling Technology). All antibodies used are listed in Table 6.6.

Table 6.6. List of Antibodies used in immunofluorescence experiments

Antigen	Source	Catalogue Number	Dilution used
53BP1	Bio-Techne	NB100-904V	1 in 500
CENPF	BD Biosciences	610768	1in 500
HA	Cell Signaling Technology	C29F4	1in 500
FLAG	Sigma-Aldrich	F1804-1MG	1 in 1000
CENPA	Life Technologies	MA1-20832	1 in 500
BrdU (Mouse)	BD Biosciences	347580	1 in 500
BrdU (Rat)	Abcam	ab6326	1 in 500
Rat 555nm	Invitrogen	A21434	1 in 500
Mouse 488nm	Thermo Fisher Scientific	A11029	1 in 500
Rabbit 488nm	Life Technologies	A11070	1 in 1000
Mouse 555nm	Life Technologies	A32570	1 in 1000

6.4.2   Proximity Ligation Assay

Proximity ligation assay was carried out using the Duolink® assay kit from Sigma-Aldrich under manufacturer's instructions. The samples were initially fixed and permeabilised as in Section 6.4.1, before being blocked in Duolink® blocking solution and subsequently incubated in primary anti-HA and anti-FLAG antibodies ON at 4°C. The samples were then incubated for one hour at 37°C with Duolink® PLUS and MINUS secondary antibody probes, the conjugated probes were subsequently ligated at 37°C for 30 minutes. Lastly, the samples were incubated at 37°C for 100 minutes with polymerase. The coverslips were then mounted using Duolink® Mounting Media with DAPI (cat. no. DUO82040). All antibodies used are listed in Table 6.6.

6.4.3   DNA Fibre Assay

6.4.3.1   Spreading DNA fibres

DNA synthesis was measured using halogenated thymidine analogues CldU and IdU (Sigma-Aldrich). IdU was placed in an incubator at 37°C and 5% CO_2 for one hour prior to use to equilibrate. A549 or U2OS cells were seeded in 6-well plates at a density to ensure active proliferation at the time of the assay. The following day cells were incubated for 20 minutes with 25 µM CldU at 37°C, before washing with media. Cells were then incubated in 1 mM Hydroxyurea (Sigma-Aldrich) for 2 hours at 37°C. The cells were washed with media before incubating at 37°C for 20 minutes with 250 µM pre-equilibrated IdU. Control cells were not incubated with HU, instead each analogue was added for 20 minutes sequentially at 37°C. After incubation, cells were washed using ice-cold PBS and harvested by trypsinisation before collecting in 10% (v/v) FBS in PBS. Cell suspension was then diluted to $5x10^5$ cells per mL in PBS. 2 µL of cell suspension was dropped at the top of a microscope slide and left to dry for 6 minutes. 7 µL of spreading buffer (0.5% (w/v) SDS, 50 mM EDTA, 200 mM Tris pH 7.4) was pipetted onto the cell suspension and mixed. To promote lysis this mixture was allowed to incubate for 2 minutes. Slides were then tilted allowing the mixture to run down the slide. After spreading, the slides

were left to dry for 2 minutes before being fixed using a 3:1 methanol to acetic acid solution for 10 minutes. Slides were left to dry for 10 minutes before storage at 4°C.

6.4.3.2 Staining of DNA fibres

To stain the DNA fibres, slides were washed twice with H_2O before being incubated in 2.5 M HCl for one hour and 15 minutes for DNA denaturation. Microscope slides were then washed twice with PBS before incubation with a blocking solution (0.1% (v/v) Tween-20, 1% (w/v) BSA Fraction V (Roche) in PBS) for 30 minutes. Primary antibodies Mouse α-BrdU (Becton Dickinson) and Rat α-BrdU (BioRad) were diluted 1:500 in blocking solution and added to the slides and incubated for one hour. Following three washes with PBS, antibodies were then fixed to the DNA by adding 4% Paraformaldehyde for 10 minutes. Slides were then washed three times in PBS and three times in blocking solution sequentially. Secondary antibodies α-Rat 555 nm and α-Mouse 488 nm from Alexa Fluor were diluted 1:500 in blocking solution and incubated with the slides for 1.5 hours. Slides were then washed in PBS, before being mounted (Sigma-Aldrich). All antibodies used are listed in Table 6.6.

6.4.3.3 DNA fibre analysis

DNA fibres were imaged as above (Section 6.4) with DNA fibre structures analysed using ImageJ (NIH) and fork asymmetry analysed in LasX software (Leica). Replication fork asymmetry was calculated by measuring the ratio between IdU tract lengths resulting from a first-label origin structure.

6.5 Interactome analysis

6.5.1 Co-IP for proteomics

The immunoprecipitation reactions for proteomics were carried out predominantly as described above (Section 6.3.2) with the following variations. pTIPZ-HA-JMJD5(WT)-P2A-RCCD1-FLAG and pTIPZ-HA-JMJD5(H321A)-P2A-RCCD1-FLAG HEK293T cells were incubated with 1 μg doxycycline for 48 hours, before treatment with 1 mM Dimethyl-N-oxalylglycine (DMOG – Sigma-Aldrich D3695) ON before harvesting as above (Section 6.3.1). The harvested cell pellets were frozen at -80°C for a freeze-thaw cycle to improve lysis. After thawing, cell pellets were lysed in JIES buffer as above with the addition

of turbonuclease (Sigma-Aldrich) at a final dilution of 1:1000 and incubated with end over end rotation at 4°C for 2 hours. The lysate was subsequently pelleted, and anti-FLAG immunoprecipitation was carried out as above (Section 6.3.2).

6.5.2   Methanol Chloroform Extraction

Following co-immunoprecipitation and elution, the samples were subjected to methanol chloroform extraction. 600 μL of methanol and 150 μL of chloroform was added to 200 μL of peptide elution and briefly vortexed. Subsequently, 450 μL of MilliQ-H_2O was added to the mixture and vortexed before centrifugation at 14000 rpm for one minute. The upper aqueous phase was then removed without disturbing the protein precipitate at the interface with the organic phase. Next, 450 μL of methanol was added to the organic phase (including the precipitate) and vortexed before centrifuging at 14000 rpm for two minutes. All supernatant was then removed, and the samples incubated at RT in an open Eppendorf to allow the protein pellet to fully dry. Once fully dry, the protein pellets were sent for Mass Spectrometry analysis.

6.5.3   Mass Spectrometry

Following methanol-chloroform extraction, the pellets were sent to the Advanced Mass Spectrometry facility at the University of Birmingham for resuspension, in-solution trypsin digestion, liquid chromatography tandem mass spectrometry (LC-MS/MS), data analysis and protein identification.

6.6   Bioinformatics analysis

6.6.1   Multiple sequence alignments

JMJD5 and RCCD1 protein sequences used for evolutionary conservation analysis were identified through BLAST searches of NCBI databases. Sequences were aligned in MEGAX using the MUSCLE algorithm (Kumar et al., 2018)(Kumar et al., 2018)(Kumar et al., 2018)(Kumar et al., 2018) and shaded using the Texshade LaTeX package.

6.6.2   Structural analysis

Structural analysis and imaging was done using the protein modelling software ChimeraX (Pettersen et al., 2021)(Pettersen et al., 2021)(Pettersen et al., 2021)(Pettersen et al., 2021). PDB files used are indicated in figure legends. The AlphaFold2 database was used to retrieve predicted structures for JMJD5 and RCCD1 (Jumper et al., 2021)(Jumper et al., 2021)(Jumper et al., 2021)(Jumper et al., 2021) and the JMJD5:RCCD1 complex was predicted using the CollabFold server (Evans et al., 2022; Mirdita et al., 2022)(Evans et al., 2022; Mirdita et al., 2022)(Evans et al., 2022; Mirdita et al., 2022)(Evans et al., 2022; Mirdita et al., 2022).

6.6.3   Cancer variant analysis

The list of cancer variants was generated through searching COSMIC and cBIOportal databases for JMJD5 and RCCD1, the separate results were then compiled into one full list of cancer variants per protein. The predicted impact of cancer variants was analysed using PolyPhen2 (Adzhubei et al., 2010)(Adzhubei et al., 2010)(Adzhubei et al., 2010)(Adzhubei et al., 2010).

6.7   Statistical analysis

Statistical analysis was done using GraphPad Prism for one-way ANOVA (with Bonferroni Post-hoc test) and Kruskal-Wallis (with Dunns' correction) tests of experiments with at least three independent biological repeats. The statistical differences were considered significant at p<0.05. All histograms were displayed with the standard error of the mean as error bars, for fork asymmetry the red lines indicate the median.

Appendices

Appendix 1 – Full list of JMJD5 cancer variants

AA Change	Codon Change	Cancer Type	Mutation ID	Mutation Type
p.G3E	c.8G>A		3108616	Substitution - missense
p.D4N	c.10G>A	Malignant melanoma	1708999	Substitution - missense
D4N	c.10G>A	Malignant melanoma	2121577	Substitution - missense
D4N	c.10G>A	Malignant melanoma	2013639	Substitution - missense
p.H6N	c.16C>A	Adenocarcinoma	5162392	Substitution - missense
p.A9T	c.25G>A	Adenocarcinoma	2549404	Substitution - missense
E10*	G>T	Glioblastoma Multiforme	TCGA-06-5416-01	Substitution - coding silent
P11S	C>A	Cutaneous Melanoma	TCGA-3N-A9WC-06	
L12M	C>A	Carcinoma	TCGA-85-6561-01	
p.R14I	c.41G>T	Acinar Adenocarcinoma	2640418	Substitution - missense
p.R14K	c.41G>A	Carcinoma	2747081	Substitution - missense
p.G16S	c.46G>A	melanoma	137044	Substitution - missense
p.A21S	c.61G>T	Carcinoma	77923	Substitution - missense
P27L	C>T	Melanoma (TCGA)	TCGA-OD-A75X-06	Substitution - missense
p.A24V	c.71C>T		3108618	Substitution - missense
A24V		Colorectal Adenocarcinoma	HCC_2998	Substitution - missense
p.E31Q	c.91G>C	Carcinoma	344421	Substitution - missense
E31Q	c.91G>C	Lung adeno (Broad)	LUAD-S01315	Substitution - missense
E31Q	c.91G>C	NSCLC (TCGA 2016)	LUAD-S01315-Tumor	Substitution - missense
p.L35F	c.105G>T	Carcinoma	702842	Substitution - missense
L35F	c.105G>T	Lung Squamous Cell Carcinoma	TCGA-60-2698-01	Substitution - missense
L35F	c.105G>T	Lung Squamous Cell Carcinoma	TCGA-60-2698-01	Substitution - missense
L35F	c.105G>T	Lung Squamous Cell Carcinoma	TCGA-60-2698-01	Substitution - missense
p.E39K	c.115G>A	Malignant melanoma	3508165	Substitution - missense
E39K	c.115G>A	Melanoma (TCGA)	TCGA-EE-A2MI-06	Substitution - missense
p.V45M	c.133G>A		3108622	Substitution - missense
p.Q50*	c.148C>T		6580351	Substitution - coding silent
Q50*	c.148C>T	Carcinoma	MBC_151	Substitution - coding silent
p.E58D		Neoplasm	2634310	Substitution - missense
p.S68N	c.203G>A	Carcinoma	4695062	Substitution - missense
p.E69K	c.205G>A	Carcinoma	969201	Substitution - missense
E69K	c.205G>A	Carcinoma	TCGA-D1-A17Q-01	Substitution - missense
E69K	c.205G>A	Carcinoma	TCGA-D1-A17Q-01	Substitution - missense
p.V70L	c.208G>T	Carcinoma	6845441	Substitution - missense
D73E	C>A	Carcinoma	TCGA-63-A5MB-01	Substitution - missense
p.S75F	c.224C>T	Carcinoma	5017199	Substitution - missense
E77K	G>A	Melanoma	TCGA-EE-A2MI-06	Substitution - missense
p.L79P	c.236T>C	Carcinoma	6692778	Substitution - missense
p.T81K	c.242C>A	Carcinoma	3672206	Substitution - missense
D90N	G>A	B-Lymphoblastic Leukemia/Lymphoma	PIP13-95124-T1	Substitution - missense
D90N	G>A	Acute Myeloid Leukemia	PIP13-95124-T2	Substitution - missense
p.R92C	c.274C>T	Uterine Endometrioid Carcinoma	TCGA-EO-A22R-01	Substitution - missense
p.R92C		Adenocarcinoma	TCGA-NH-A5IV-01	Substitution - missense
p.R92H	c.275G>A	Carcinoma	6692769	Substitution - missense
p.R93L	c.278G>T	lung	5333831	Substitution - missense
R93L	c.278G>T	NSCLC (TCGA 2016)	TCGA-L9-A8F4-01	Substitution - missense
p.R93W	c.277C>T	Carcinoma	6692776	Substitution - missense
p.G98C	c.292G>T	Carcinoma	4881452	Substitution - missense
p.A103fs*26	c.304delA		3108626	Frameshift
p.R121Q	c.362G>A		3108628	Substitution - missense
M125R	T>G	AML (TCGA pub)	TCGA-AB-2983-03	Substitution - missense
M125R	T>G	AML (TCGA)	TCGA-AB-2983-03	Substitution - missense
p.G126C	c.376G>T	Carcinoma	6692756	Substitution - missense
G126D	G>A	Uterine Endometrioid Carcinoma	TCGA-EY-A2OP-01	Substitution - missense
p.M129fs*45	c.384_387delGATG	lung	312120	Frameshift
M129Gfs*45	c.384_387delGATG	Small Cell Lung (CLCGP)	S00946	Frameshift
A131P	G>C	Prostate Neuroendocrine Carcinoma	WCMC212_1_N	Substitution - missense
p.A142T	c.424G>A		3108630	Substitution - missense

AA Change	Codon Change	Cancer Type	Mutation ID	Mutation Type
p.A143T	c.427G>A		3108632	Substitution - missense
p.I144V	c.430A>G	Carcinoma	3817643	Substitution - missense
I144V	c.430A>G	Breast Invasive Ductal Carcinoma	TCGA-BH-A18G-01	Substitution - missense
R153S	G>T	NSCLC (TCGA 2016)	TCGA-21-1083-01	Substitution - missense
p.R156H	c.467G>A	lung	557329	Substitution - missense
R156H	c.467G>A	NSCLC (TCGA 2016)	TCGA-64-5815-01	Substitution - missense
R156C	C>T	Uterine Endometrioid Carcinoma	TCGA-AJ-A3EK-01	Substitution - missense
R156S	C>A	Papillary Renal Cell Carcinoma	TCGA-BQ-7048-01	Substitution - missense
p.T165I	c.494C>T	large intestine	3108636	Substitution - missense
A168T	G>A	Bladder Urothelial Carcinoma	TCGA-XF-AAN2-01	Substitution - missense
A169P				Substitution - missense
A170E	C>A	Lung adeno (Broad)	LUAD-YINHD	Substitution - missense
A170E		NSCLC (TCGA 2016) Carcinoma	LUAD-YINHD-Tumor	Substitution - missense
p.T183fs*31	c.547_548insA		3108640	Frameshift
p.P185L	c.554C>T		4257938	Substitution - missense
p.P185S	c.553C>T	Malignant melanoma	5850564	Substitution - missense
R186W	C>T	Melanoma (TCGA)	TCGA-GN-A4U7-06	Substitution - missense
R186W		Uterine Serous Carcinoma/Uterine Papillary Serous Carcinoma	TCGA-A5-A0G2-01	Substitution - missense
p.R189H	c.566G>A	Haematopoietic neoplasm	5802727	Substitution - missense
p.P190Q/L	c.569C>A	melanoma	302448	Substitution - missense
p.P190L	c.569C>A	Carcinoma	302448	Substitution - missense
P190Q	C>A	Medulloblastoma	MD-051	Substitution - missense
P190L	C>T	Cutaneous Melanoma (Yale)	YUSWI	Substitution - missense
p.P202L	c.605C>T		3108642	Substitution - missense
p.R204M	c.611G>T	Carcinoma	5952332	Substitution - missense
p.L208P	c.623T>C	Carcinoma	4695068	Substitution - missense
p.V211M	c.631G>A		4257940	Substitution - missense
p.L223V	c.667T>G	Malignant melanoma	4893201	Substitution - missense
L223V		Cutaneous melanoma	TCGA-EE-A2MF-06	Substitution - missense
p.E224K	c.670G>A	Carcinoma	4558899	Substitution - missense
p.Y225*	c.675T>G		3108644	Substitution - coding silent
p.Q227R	c.680A>G	Lymphoid neoplasm	1290497	
Q227R	c.680A>G	Chronic Lymphocytic Leukemia	CLL047	Substitution - missense
p.A230T	c.688G>A	Carcinoma	4059615	Substitution - missense
A230T	c.688G>A	Stomach (TCGA pub)	TCGA-D7-6527-01	Substitution - missense
A230T	c.688G>A	Stomach (TCGA)	TCGA-D7-6527-01	Substitution - missense
A230T	c.688G>A	Stomach/Esophageal	TCGA-D7-6527-01	Substitution - missense
p.R233Q	c.698G>A	Carcinoma	969207	Substitution - missense
R233Q	c.698G>A	Uterine Endometrioid Carcinoma	TCGA-BS-A0UV-01	Substitution - missense
R233Q	c.698G>A	Uterine Endometrioid Carcinoma	TCGA-BS-A0UV-01	Substitution - missense
T234I	C>T	Uterine Endometrioid Carcinoma	TCGA-AJ-A3BH-01	Substitution - missense
S241L	C>T	Uterine (TCGA pub)	TCGA-AP-A056-01	Substitution - missense
S241L		Uterine (TCGA pub)	TCGA-AX-A0J0-01	Substitution - missense
S241L		Uterine (TCGA)	TCGA-AP-A056-01	Substitution - missense
S241L		Uterine (TCGA)	TCGA-AX-A0J0-01	Substitution - missense
S241L		Cutaneous Melanoma	LSD4744_T	Substitution - missense
E246K	G>A	Stomach (Esophagogastric)	PGM40	Substitution - missense
p.V255I	c.763G>A	Carcinoma	4695070	Substitution - missense
p.E257K	c.769G>A		3108650	Substitution - missense
S260R	C>G	Carcinoma	TCGA-XF-AAMT-01	Substitution - missense
p.Y262C	c.785A>G	Carcinoma	4059617	Substitution - missense
Y262C	c.785A>G	TCGA data for Esophagus-Stomach Cancers (TCGA, Nature 2017) (s	TCGA-BR-4361-01	Substitution - missense
Y262C	c.785A>G	Stomach Adenocarcinoma (TCGA, Nature 2014)	TCGA-BR-4361-01	Substitution - missense
Y262C	c.785A>G	Stomach Adenocarcinoma (TCGA, Provisional)	TCGA-BR-4361-01	Substitution - missense
V264M	G>A	Breast Invasive Ductal Carcinoma	MBC-MBCProject_RKf1	Substitution - missense
V270I	G>A	Esophagus (Broad)	ESO-0029	Substitution - missense
V270I	G>A	NCI-60	SK_OV_3	Substitution - missense
p.G271W	c.811G>T	Carcinoma	5015363	Substitution - missense
G271W	c.811G>T	Renal Non-Clear Cell Carcinoma	9266940	Substitution - missense
p.L273R	c.818T>G	Carcinoma	6692760	Substitution - missense

Appendix 1 – Full list of JMJD5 cancer variants continued

AA Change	Codon Change	Cancer Type	Mutation ID	Mutation Type
p.P283L	c.848C>T		969211	Substitution - missense
P283L	c.848C>T	Uterine (TCGA pub)	TCGA-BS-A0TI-01	Substitution - missense
P283L	c.848C>T	Uterine (TCGA)	TCGA-BS-A0TI-01	Substitution - missense
p.D288N	c.862G>A	Malignant melanoma	5850566	Substitution - missense
p.I289T	c.866T>C	Carcinoma	3732260	Substitution - missense
p.S290N	c.869G>A	Carcinoma	245108	Substitution - missense
p.D293N	c.877G>A	Neoplasm	2634355	Substitution - missense
p.Y294C	c.881A>G	Carcinoma	4695072	Substitution - missense
D299				Substitution - missense
D299Y	G>T	Cutaneous Melanoma	TCGA-BF-A1PZ-01	Substitution - missense
p.G300W	c.898G>T		3108656	Substitution - missense
G300V	G>T	Cutaneous Melanoma	TCGA-D3-A2JN-06	Substitution - missense
L311R	T>G	Colorectal (DFCI 2016)	coadread_dfci_2016_1	Substitution - missense
P319L	C>T	Hepatocellular Adenoma	H112501	Substitution - missense
p.D323N	c.967G>A	Carcinoma	4520876	Substitution - missense
Q325K	C>A	Cutaneous Melanoma	TCGA-EE-A2MD-06	Substitution - missense
p.F328L	c.984C>G		3108660	Substitution - missense
p.Q331H	c.993G>T	Carcinoma	6692764	Substitution - missense
p.M333T	c.998T>C	Malignant melanoma	4897625	Substitution - missense
M333T	T>C	Acral melanoma	TCGA-ER-A19T-01	Substitution - missense
p.Y337fs*60	c.1009delT		5958897	Frameshift
p.R339Q	c.1016G>A		3108662	Substitution - missense
P343T	C>A	Lung squ (TCGA pub)	TCGA-43-3920-01	Substitution - missense
P343Q	c.1028C>A	Carcinoma	1898101	Substitution - missense
p.A348T	c.1042G>A	Carcinoma	4791132	Substitution - missense
A348T	c.1042G>A	Hepatocellular Carcinoma	CHC1079T	Substitution - missense
L349M	C>A	Uterine Serous Carcinoma/Uterine Papillary Serous Carcinoma	TCGA-A5-A0G2-01	Substitution - missense
Y350H	T>C	Uterine Endometrioid Carcinoma	TCGA-AP-A1DV-01	Substitution - missense
p.T354M	c.1061C>T	Carcinoma	4630187	Substitution - missense
p.H355Y	c.1063C>T	Carcinoma	969215	Substitution - missense
H355Y	c.1063C>T	Uterine (TCGA pub)	TCGA-B5-A11E-01	Substitution - missense
H355Y	c.1063C>T	Uterine (TCGA)	TCGA-B5-A11E-01	Substitution - missense
p.L357delL	c.1069_1071delCTC		3108666	Deletion
p.T360M	c.1079C>T		3108668	Substitution - missense
p.Q362H	c.1086G>T	Carcinoma	4059619	Substitution - missense
Q362H	c.1086G>T	Stomach (TCGA pub)	TCGA-CG-5733-01	Substitution - missense
Q362H	c.1086G>T	Stomach (TCGA)	TCGA-CG-5733-01	Substitution - missense
p.V365L	c.1093G>T	Carcinoma	2662305	Substitution - missense
V365M	G>A	Uterine Endometrioid Carcinoma	TCGA-FI-A2D6-01	Substitution - missense
P368L	C>T	Cutaneous Melanoma	CR06670_T	Substitution - missense
p.D369Y	c.1105G>T	Malignant melanoma	3108670	Substitution - missense
D369N	G>A	Uterine Endometrioid Carcinoma	TCGA-AP-A1E1-01	Substitution - missense
E371K	G>A	Cutaneous Melanoma	TCGA-GN-A4U5-01	Substitution - missense
p.K375M	c.1124A>T		98317	Substitution - missense
K375M	c.1124A>T	Carcinoma	HN27PT	Substitution - missense
A379V	C>T	Uterine Endometrioid Carcinoma	TCGA-AJ-A3EL-01	Substitution - missense
P380S	C>T	Uterine Endometrioid Carcinoma	TCGA-AP-A1DK-01	Substitution - missense
p.E390D	c.1170G>C		3108672	Substitution - missense
p.E390K	c.1168G>A		6567997	Substitution - missense
E390K	G>A	Carcinoma	MBC_45	Substitution - missense
p.I391F	c.1171A>T		3108674	Substitution - missense
p.L392M	c.1174C>A	Carcinoma	6143975	Substitution - missense
p.F393L	c.1179C>A	Carcinoma	3420903	Substitution - missense
P395L	C>T	Head & neck (TCGA)	TCGA-F7-A624-01	Substitution - missense
P395L		Uterine Serous Carcinoma/Uterine Papillary Serous Carcinoma	TCGA-A5-A0G1-01	Substitution - missense
P395Q	C>A	Lung Squamous Cell Carcinoma	TCGA-37-3789-01	Substitution - missense
P395Q	C>A	Lung Squamous Cell Carcinoma	TCGA-66-2744-01	Substitution - missense
p.Y398H	c.1192T>C		3108678	Substitution - missense

Appendix 1 – Full list of JMJD5 cancer variants continued

AA Change	Codon Change	Cancer Type	Mutation ID	Mutation Type
p.Y401C	c.1202A>G	Carcinoma	471546	Substitution - missense
p.Y401*	c.1203C>A		3108680	Substitution - coding silent
p.V402M	c.1204G>A	Carcinoma	434990	Substitution - missense
V402M	G>A	Breast Invasive Ductal Carcinoma	TCGA-A8-A07R-01	Substitution - missense
V402M	G>A	Breast Invasive Ductal Carcinoma	TCGA-A8-A07R-01	Substitution - missense
V402M		Mucinous Adenocarcinoma of the Colon and Rectum	TCGA-A6-6780-01	Substitution - missense
p.R403Q	c.1208G>A	Carcinoma	6016100	Substitution - missense
p.L407F	c.1221G>C		3108682	Substitution - missense
S410L	C>T	Uterine Endometrioid Carcinoma	TCGA-AP-A1DV-01	Substitution - missense
F413V	T>G	Uterine Endometrioid Carcinoma	TCGA-EO-A3B0-01	Substitution - missense
p.W414R	c.1240T>C	Carcinoma	4913236	Substitution - missense
W414R	T>C	Liver (TCGA) Hepatocellular carcinoma	TCGA-ED-A4XI-01	Substitution - missense
p.S416L	c.1247C>T	Carcinoma	969217	Substitution - missense
S416L	C>T	Uterine (TCGA pub)	TCGA-B5-A11E-01	Substitution - missense
S416L	C>T	Uterine (TCGA)	TCGA-B5-A11E-01	Substitution - missense

```
                        10        20        30        40        50        60        70        80
H. sapiens      MAGDTHCPAEPLAREGTLWEALRALLPHSKEPLKLDLGEKVERSVVTLLQRATELFYEGR..RDECLQ.SSEVILDYSWEKLNTGIWQDV  87
M. muculus      MSEDT...TEPLVGSSTLWKELRTLLPDKEEELKLDLGEKVDRSVAALLRQAVGLFYAGH..WQGCLQ.ASEAVLDYSWEKLNTGPWRDV  84
G. gallus       .........MAGALGLWAEARALLPGSEEELSLALSGEVDECVPPLLLRARALLYEAAPPCEAALRRLGDVLRDYAWEKLNAGPWRDV  79
D. rerio        ..............MASVWTDIRAVLPSTVSEFPLDFSEKTELSVLKCLELSRDQLYSEADCPVSAER..AQIIDYSWEKLNIGTWRDV  74
D. melanogaster ..............MDSELLELTQLLPRWVDLENVVRGEVEARYILKRAADHLANLRSGGDSGAEETGYLVGALVDRNWERIHTGHISQV  76

                        90       100       110       120       130       140       150       160       170
H. sapiens      DKDWRRVYAIGCLLKALCLCQAPEDANTVAAALRVCDMGLLMGAAILGDILLKVAAILQTH.LPGKRPARGSLPEQPCTKKARADHGLIP  176
M. muculus      DKEWRRVYSFGCLLKALCLCQAPQKATTVVEALRVCDMGLLMGAAILEDILLKVVAVLQTHQLPGKQPARGPHQDQPATKKAKCDASPAP  174
G. gallus       SKAWRQVYAYGCLFGALAEVAARR...PLAPAVRLCDMGLLMGASVQDNVLARLVRLLQAH.LPRAD.RRGAAP..SSAKRARTESPPAP  162
D. rerio        DKEWRRVYSYGCLFKVLSLCHGNPPQNIIQEAIRTCDMSLLMGAAIMDNILQRLVGILRNK.IKTTSPNKAEWSEEPCSKKRKHDCKSEP  163
D. melanogaster PLVTRKIYAIACCFKSTSPAQKDACSEILDEAQ......LLGCMEDWSELKYALMDYLDKD............GAVALNSAPLPTLEPLT  148

                       180       190       200       210       220       230       240       250       260
H. sapiens      DVKLEKTVPRLHRPSLQHFREQFLVPGRPVILKGVADHWPCMQK..WSLEYIQEIAGCRTVPVEVGSRYTDEEWSQTLMTVNEFISKYIV  264
M. muculus      DVMLERMVPRLRCPPLQYFKQHFLVPGRPVILEGVADHWPCMKK..WSLQYIQEIAGCRTVPVEVGSRYTDEDWSQTLMTVDEFIQKFIL  262
G. gallus       VVRPEDTVPIERCPSLEHFRDRYLIPQKPVVLEGITDHWPCMKK..WSVDYYRQVAGCRTVPVEIGSRYTDEEWSQKLMTVNDFINQYIV  250
D. rerio        VLNPTKEVPRTHCPSLERFRSDFLDSKKPVILEGITDHWPAFTQHPWSIDYLRTVAGCRTVPIEVGSKYTDEEWSQKLITVNDFIDRYIT  253
D. melanogaster RVTSNCDIPQLDAPSLEEFQTKCFEAGQPTLLLNTIQHWPALHK.WLDLNYLLQVAGNRTVPIEIGSNYASDEWSQQLVKTRDFLSHQFG  237

                       270       280       290       300       310       320       330       340
H. sapiens      NEP....RDVGYLAQHQLFDQIPELKQDISIPDYCSLGDGEEE.EITINAWFGPQGTISPLHQDPQQNFLVQVMGRKYIRLYSPQESGAL  349
M. muculus      SEA....KDVGYLAQHQLFDQIPELKRDISIPDYCCLGNGEEE.EITINAWFGPQGTISPLHQDPQQNFLVQVLGRKYIRLYSPQESEAV  347
G. gallus       NE.....NSVGYLAQHQLFDQIPELKEDISIPDYCCLGPGEED.DITINAWFGPAGTISPLHQDPQQNFLAQVFGRKYIRLCSPQDSENL  334
D. rerio        GTEE...DGVGYLAQHQLFDQVPELKEDIRIPDYCCLGPGDED.DITINAWFGPGGTVSPLHQDPQQNFLAQVVGRKYIRLYSPEETKSL  339
D. melanogaster KEPSKAGQNIEYLAQHELFAQIPALKEDISIPDYCTISNEDTPGAVDIKAWLGPAGTVSPMHYDPKHNLLCQVFGSKRITLAAPADTDNL  327

                       350       360       370       380       390       400       410
H. sapiens      YPHDTHLLHNTSQVDVENPDLEKFPKFAKAPFLSCILSPGEILFIPVKYWHYVRALDLSFSVSFWWS  416
M. muculus      YPHETHILHNTSQVDVENPDLEKFPKFTEAPFLSCILSPGDTLFIPAKYWHYVRSLDLSFSVSFWWS  414
G. gallus       YPHESQLLHNTSQVDVEDPDLTKFPNFRKVAFQSCILMPGQVLFIPVKYWHYIRSLDISFSVSFWWS  401
D. rerio        YPHESCLLHNTSQVEVENPDLVKFPDFSRASMEECVLCPGDVLFIPLQHWHYVRSLELSFSVSFWWS  406
D. melanogaster YPHDSEFLANTARIDAALLDPETYPLVAKVKFYQLLLQPGDCLYMEPKWWHYVRSEAPSFSVSFWWE  394
```

X	non-conserved
X	similar
X	≥ 50% conserved
X	all match

Appendix 3 – Full list of RCCD1 cancer variants

AA change	Codon change	Cancer Type	Mutation ID		Mutation type
p.A2=	c.6G>T		COSM6303068		Substitution - coding silent
p.G7=	c.21G>A		COSM5523661		Substitution - coding silent
p.A8S	c.22G>T	Thyroid Neoplasm		2185970	Substitution - Missense
p.A8S	c.22G>T	Large intestine Adenocarcinoma	TCGA-AU-6004-01		Substitution - Missense
p.A8S	c.22G>T	Large intestine Adenocarcinoma	TCGA-AG-4022-01		Substitution - Missense
p.A8S	c.22G>T	Large intestine Adenocarcinoma	TCGA-A6-5660-01		Substitution - Missense
p.A8S	c.22G>T	Thyroid Neoplasm		2186083	Substitution - Missense
p.L21=	c.63G>C	Liver Hepatocellular Carcinoma	TCGA-DD-AADV-01		Substitution - coding silent
p.G24R	c.70G>A	Thyroid Carcinoma	COSM6409082		Substitution - Missense
p.G24R	c.70G>A	Thyroid Carcinoma		2635078	Substitution - Missense
p.V29L	c.85G>T	Liver Neoplasm	COSM6286334		Substitution - Missense
p.P34L	c.101C>T	Thyroid Carcinoma	COSM6409216		Substitution - Missense
p.P34L	c.101C>T	Thyroid Carcinoma		2635084	Substitution - Missense
P34L	C>T	Colorectal Adenocarcinoma	SP18121		Substitution - Missense
p.L35=	c.105G>A	Large intestine Adenocarcinoma	COSM6764503		Substitution - coding silent
p.C42Y	c.125G>A	Thyroid Carcinoma	COSM6388470		Substitution - Missense
p.C42Y	c.125G>A	Thyroid Carcinoma		2635111	Substitution - Missense
R43C	C>T	Bladder Urothelial Carcinoma	WCM117_7		Substitution - Missense
W48*	G>A	Hepatocellular Carcinoma	TCGA-DD-AADB-01		Nonsense
X56_splice	T>G	Pediatric Low Grade Gliomas	7316-3570		Splice Region
R59H	G>A	Bladder Urothelial Carcinoma	WCM325_1		Substitution - Missense
p.L60W	c.179T>G	Prostate Adenocarcinoma	COSM5559816		Substitution - Missense
L60W	T>G	Prostate Adenocarcinoma	MO_1221		Substitution - Missense
A66V	C>T	Diffuse Large B-Cell Lymphoma, NOS	PCNSL_3		Substitution - Missense
p.D75G	c.224A>G	Biliary Tract Carcinoma	COSM5498182		Substitution - Missense
p.A78T	c.232G>A	Squamous cell carcinoma		2912471	Substitution - Missense
p.S79L	c.236C>T	Large cell neuroendocrine carcinoma		2864338	Substitution - Missense
p.L83=	c.249C>T	Cervix Squamous Cell Carcinoma		2386224	Substitution - coding silent
p.L86M	c.256C>A	Large intestine Adenocarcinoma	COSM4634037		Substitution - Missense
R87H	G>A	Hepatocellular Carcinoma	SCB-MERiC-0037-D433		Substitution - Missense
p.P92=	c.276G>A	Large intestine Adenocarcinoma		1651230	Substitution - coding silent
p.W111C	c.333G>T	Large intestine Adenocarcinoma	COSM6037797		Substitution - Missense
p.G121R	c.361G>A	Basal cell carcinoma		2869871	Substitution - Missense
D124H	G>C	Hepatocellular Carcinoma	TCGA-DD-AAEA-01		Substitution - Missense
p.A126=	c.378C>T	Large intestine Adenocarcinoma		1651580	Substitution - coding silent
G127A	G>C	Esophageal Squamous Cell Carcinoma	TCGA-LN-A7HZ-01		Substitution - Missense
p.E128K	c.382G>A	Thyroid Carcinoma	COSM6431646		Substitution - Missense
p.E128K	c.382G>A	Thyroid Carcinoma		2635080	Substitution - Missense
p.A129=	c.387C>A	Large intestine Adenocarcinoma	COSM4632676		Substitution - coding silent
p.A131=	c.393T>A	Papillary renal cell carcinoma		2386151	Substitution - coding silent
p.A140T	c.418G>A	Ovary Mixed adenosquamous carcinoma	COSM3944335		Substitution - Missense
p.R141S	c.421C>A	Thyroid Carcinoma	COSM6370896		Substitution - Missense
R141S	C>A	Cutaneous Melanoma	TCGA-D3-A2JP-06		Substitution - Missense
R141S	C>A	Thyroid Carcinoma		2635107	Substitution - Missense
P146Q	C>A	Cutaneous Melanoma	TCGA-ER-A19Q-06		Substitution - Missense
P146T	C>A	Upper Tract Urothelial Carcinoma	UTUC_6T		Substitution - Missense
p.A148V	c.443C>T	Thyroid Carcinoma	COSM6391150		Substitution - Missense
p.A148V	c.443C>T	Thyroid Carcinoma		2635065	Substitution - Missense
p.Y151H	c.451T>C	Breast ER-PR-positive carcinoma	COSM6574381		Substitution - Missense
p.R152Q	c.455G>A	Thyroid Carcinoma	COSM6398861		Substitution - Missense
p.R152Q	c.455G>A	Thyroid Carcinoma		2635098	Substitution - Missense
p.R161C	c.481C>T	Endometrioid carcinoma		2898368	Substitution - Missense
p.G166A	c.497G>C	Large intestine Adenocarcinoma	COSM4654748		Substitution - Missense
p.A167V	c.500C>T	Large intestine Adenocarcinoma	COSM4647869		Substitution - Missense
p.A167G	c.500C>G	Ovarian Mucinous carcinoma		2897900	Substitution - Missense
p.A175V	c.524C>T	Oesophageal Squamous Cell Carcinoma	COSM5439100		Substitution - Missense
A175T	G>A	B-Lymphoblastic Leukemia/Lymphoma	ALL-B-14-R		Substitution - Missense
p.G183=	c.549C>T	Large intestine Adenocarcinoma	COSM7266095		Substitution - coding silent

AA change	Codon change	Cancer Type	Mutation ID	Mutation type
p.L190=	c.570G>T	Lung adenocarcinoma	1913982	Substitution - coding silent
p.L190=	c.570G>A	Biliary tract Carcinoma	2866083	Substitution - coding silent
E200K	G>A	Acral Melanoma	MEL-IPI_Pat58-Tumor-SM-4DK2J	Substitution - Missense
p.R202Q	c.605G>A	Large intestine Adenocarcinoma	COSM4720882	Substitution - Missense
G209C	G>T	Serous Ovarian Cancer	TCGA-13-0807-01	Substitution - Missense
L210P	T>C	Prostate	SCB-PCBM-0036-4-1-19-M-S12	Substitution - Missense
p.V211=	c.633C>A	Central nervous system glioma	COSM2151529	Substitution - coding silent
V211A	T>C	Melanoma	ME032	Substitution - Missense
p.V215A	c.644T>C	Stomach Adenocarcinoma	COSM4057974	Substitution - Missense
A216V	C>T	Esophageal Squamous Cell Carcinoma	Pt4	Substitution - Missense
G219Afs*7		Cutaneous Melanoma	TCGA-D3-A8GM-06	Frame_Shift_Del
W220Lfs*8		Cutaneous Melanoma	TCGA-D3-A8GK-06	Frame_Shift_Ins
V223L	G>C	Tubular Stomach Adenocarcinoma	TCGA-VQ-AA6D-01	Substitution - Missense
G229E	G>A	Uterine Endometrioid Carcinoma	TCGA-FI-A2D0-01	Substitution - Missense
Y232C	A>G	Uterine Endometrioid Carcinoma	TCGA-FI-A2D5-01	Substitution - Missense
I233M	C>G	Bladder Urothelial Carcinoma	TCGA-4Z-AA7W-01	Substitution - Missense
p.I233M	c.699C>G	Bladder Urothelial Carcinoma	COSM7080648	Substitution - Missense
p.N237T	c.710A>C	Thyroid Carcinoma	COSM6425178	Substitution - Missense
p.N237T	c.710A>C	Thyroid Carcinoma	2635106	Substitution - Missense
p.G240V	c.719G>T	Thyroid Carcinoma	COSM6393093	Substitution - Missense
p.G240V	c.719G>T	Thyroid Carcinoma	2635095	Substitution - Missense
p.Q241=	c.723G>A	Prostate Adenocarcinoma	COSM6001601	Substitution - coding silent
p.R247K	c.740G>A	Central nervous system glioma	COSM5731457	Substitution - Missense
p.R247S	c.741G>T	Thyroid Carcinoma	COSM6410040	Substitution - Missense
p.R247S	c.741G>T	Thyroid Carcinoma	2635089	Substitution - Missense
A257T	G>A	Uterine Endometrioid Carcinoma	TCGA-AX-A1CE-01	Substitution - Missense
p.R258W	c.772A>T	Thyroid Carcinoma	COSM6422576	Substitution - Missense
p.R258W	c.772A>T	Thyroid Carcinoma	2635094	Substitution - Missense
X260_splice	C>G	Non-Small Cell Lung Cancer	CRUK0043-R2	Splice Region
E262K	G>A	Lung Adenocarcinoma	C3N-01893	Substitution - Missense
R272I	G>T	Rectal Adenocarcinoma	TCGA-AG-A002-01	Substitution - Missense
p.T273M	c.818C>T	Large intestine adenoma	2809971	Substitution - Missense
p.T273M	c.818C>T	Large intestine Adenocarcinoma	2809979	Substitution - Missense
T273M	C>T	Cutaneous Melanoma	MEL-JWCI-14	Substitution - Missense
G274C	G>T	Lung Squamous Cell Carcinoma	TCGA-34-8455-01	Substitution - Missense
p.G274S	c.820G>A	Malignant melanoma	2894833	Substitution - Missense
p.D278=	c.834T>C	Oesophageal Squamous Cell Carcinoma	COSM1240027	Substitution - coding silent
p.F284L	c.850T>C	Lung non small cell carcinoma	COSM2140570	Substitution - Missense
p.F290V	c.868T>G	Lung adenocarcinoma	COSM355654	Substitution - Missense
p.F290L	c.870C>A	Thyroid Carcinoma	COSM6406801	Substitution - Missense
p.F290L	c.870C>A	Haemotopoietic and Lymphoid Plasma Cell Myeloma	2635105	Substitution - Missense
S309F	C>T	Acral Melanoma	MEL-IPI_Pat58-Tumor-SM-4DK2J	Substitution - Missense
p.R310W	c.928C>T	Ewings sarcoma-peripheral primitive neuroectodermal tumour	COSM4578628	Substitution - Missense
p.R310Q	c.929_930delinsAA	Skin Squamous Cell Carcinoma	COSM4567932	Substitution - Missense
p.G319E	c.956G>A	Large intestine Adenocarcinoma	COSM1375499	Substitution - Missense
G319E	G>A	Colon Adenocarcinoma	TCGA-AA-3663-01	Substitution - Missense
E320A	A>C	Renal Clear Cell Carcinoma	TCGA-B0-5075-01	Substitution - Missense
p.E320A	c.959A>C	Renal Clear Cell Carcinoma	COSM471256	Substitution - Missense
p.W324*	c.972G>A	Basal Cell Carcinoma	COSM5927742	Nonsense
W324L	G>T	Lung Squamous Cell Carcinoma	TH3_N1_11R	Substitution - Missense
W324L	G>T	Lung Squamous Cell Carcinoma	TH3_Primary_LUL	Substitution - Missense
W324L	G>T	Lung Squamous Cell Carcinoma	TH3_N2_4R	Substitution - Missense
W324C	G>T	Lung Squamous Cell Carcinoma	TH3_N1_11R	Substitution - Missense
W324C	G>T	Lung Squamous Cell Carcinoma	TH3_Primary_LUL	Substitution - Missense
W324C	G>T	Lung Squamous Cell Carcinoma	TH3_N2_4R	Substitution - Missense

Appendix 3 – Full list of RCCD1 cancer variants continued

AA change	Codon change	Cancer Type	Mutation ID		Mutation type
G325C	G>T	Neuroblastoma	NB-S-559		Substitution - Missense
X327_splice	A>G	Cutaneous Melanoma	MEL-UKRV-Mel-24		Splice Region
K328N	A>C	Uterine Endometrioid Carcinoma	TCGA-EO-A22U-01		Substitution - Missense
Y329*	T>G	Breast Invasive Ductal Carcinoma	MO_1065		Nonsense
G240Afs*22		Esophageal Adenocarcinoma	SP111031		Frame shift
R344C	C>T	Bladder Urothelial Carcinoma	TCGA-G2-A2EL-01		Substitution - Missense
p.R344C	c.1030C>T	Breast Carcinoma	COSM163975		Substitution - Missense
R344C	C>T	Bladder Urothelial Carcinoma	BLCA-015-Tumor-SM-CUCGP		Substitution - Missense
R344H	c.1031G>A	Central nervous system glioma		2813453	Substitution - Missense
R344H	c.1031G>A	Central nervous system glioma		2813454	Substitution - Missense
p.R344=	c.1032C>T	Haemotopoietic and Lymphoid Plasma Cell Myeloma			Substitution - coding silent
R345C		Adrenocortical Carcinoma	TCGA-PK-A5HB-01		Substitution - Missense
p.Y348C	c.1043A>G	Biliary tract Carcinoma		2866080	Substitution - Missense
p.Q353R	c.1058A>G	Malignant Melanoma	COSM5850083		Substitution - Missense
p.Q353R	c.1058A>G	Malignant Melanoma		2492838	Substitution - Missense
p.V356I	c.1066G>A	Nasopharyngeal carcinoma		2857345	Substitution - Missense
p.V356I	c.1066G>A	Nasopharyngeal carcinoma		2857355	Substitution - Missense
p.V359=	c.1077C>A	Lung adenocarcinoma		2194826	Substitution - coding silent
p.T360=	c.1080C>T	Breast Metaplastic Carcinoma	COSM6195049		Substitution - coding silent
p.G362A	c.1085G>C	Haemotopoietic and Lymphoid Plasma Cell Myeloma		2809769	Substitution - Missense
p.G362A	c.1085G>C	Lung adenocarcinoma		2776232	Substitution - Missense
G362A	G>C	Prostate	SCB-PCBM-0044-7-2-43-T-S27		Substitution - Missense
G362A	G>C	Prostate	SCB-PCBM-0044-8-2-43-T-S28		Substitution - Missense
G362A	G>C	Prostate	SCB-PCBM-0044-10-2-43-M-S29		Substitution - Missense
G362A	G>C	Prostate	SCB-PCBM-0044-12-2-43-M-S30		Substitution - Missense
G362A	G>C	Prostate	SCB-PCBM-0044-13-2-43-M-S31		Substitution - Missense
p.P363=	c.1089G>A	Endometrioid carcinoma		2198465	Substitution - coding silent
P363T	C>A	Colon Adenocarcinoma	11CO059		Substitution - Missense
p.T366=	c.1098C>A	Stomach Adenocarcinoma	COSM4057975		Substitution - coding silent
p.V368M	c.1102G>A	Skin Squamous Cell Carcinoma	COSM2140579		Substitution - Missense
p.A370V	c.1109C>T	Breast Invasive Ductal Carcinoma	COSM434555		Substitution - Missense
A370V	C>T	Breast Invasive Ductal Carcinoma	TCGA-AO-A0JD-01		Substitution - Missense
K373T	A>C	Germinal Center B-Cell Type	DLBCL-LS3866		Substitution - Missense
K373R	A>G	Angiosarcoma	RP-1447_ASCProject_NdUxUwCM_T1_v2		Substitution - Missense
G374E	G>A	Glioblastoma	GLSS-SF-0005-R1		Substitution - Missense
K375R	A>G	Angiosarcoma	Angio-ASCProject_NdUxUwCM-Tumor-SM		Substitution - Missense
p.S376N	c.1127G>A	Large intestine Adenocarcinoma	COSM298149		Substitution - Missense
S376N	G>A	Mucinous Adenocarcinoma of the Colon and Rectum	TCGA-AA-3994-01		Substitution - Missense

Appendix 4 – RCCD1 multiple sequence alignment

```
                                            10        20        30        40
H. sapiens      ..........................MAEERPGAWFGFGFCGFGQELG......SGRGHQ.VHSFSPLRAGVDICR........VSASW    48
M. muculus      ..........................MAEKRHGAWFGFGFCGFGQALG......SGNSHHSVYSPEPLHASDDICQ........VSAGW    49
D. rerio        MTKYTDNRHRFGFLTTFRAQKQLTVLCVDQSKSMSWFGFGFNGFGQTSRAD...VSGQVMCKVNSPVLISDD.ECKSGVNRSDRRIRACW    86
X. laevis       ..........................MSWYGFGYRGFEQLG......LGGKISLELPEPILLEEDADAPEGIKVS..KAVPSW    49
D. melanogaster ..........................MPLLFTGFNAFGQHECVSGDVDCAAGFSELNAPVSTQNQCTISIGWRYAALAFGRKL    57

                 50        60        70        80        90       100       110       120
H. sapiens      SYTAFVTRGGR....LELSGSASGAAGR........CKDAWASEGLLAVLRAGPGPEALFQVWAAESALRGEPLWAQNVVPEAEG..EDD    124
M. muculus      SYTALVTRGGR....VELSGSVSGAADG........CRDVWASEELLVLLRNKGGSSTEVQAWVPGSALQGEPLWVQNLVSGAKGQGEDE    127
D. rerio        SSRADLQQTQSGSG.VYLSGCVCGPSAQVCASQC..CTDALISETHLTLSFTDRVELWELKPQ......QNQLVWKR...........EH    156
X. laevis       SYSAFLTEDGS....LLLSGSIAVSPNKYLHFKGLHCVDVLPAEKYLVLQLKHGLQCWETAAFTAE.GPQAEPMWK............M    121
D. melanogaster CLRGLLDDGPNECVTLEATGNIRALAAADS.....HCLVLLQSGQLYRVQPKLQAELVAVRLEAAP...RSNSGTKRSIFGAAKAPSSPI    139

                     130       140       150       160       170       180       190       200       210
H. sapiens      PAGEAQAGRLPLLPCARAYVSPRAPFYRPLAP.ELRARQLELGAEHALLLDAAGQVFSWGCGRHGQLGHGTLEAELEPRLLEALQGLVMA    213
M. muculus      PSRESRMGTLPLLPCARAYVTPEPPFCQPLAP.ELRVRQLELGAEHVLLLCAAGQVFSWGAGRHGQLGHGTLEAELEPRLLEALQGLRMA    216
D. rerio        ELSAEHTAALPLVSGG..FVQHKPPFFHPL...KLCAVSLVLGSEHALLLTADGTLYSWGSGSHGQLGHGVLTSLEDPQAVEALWGVPIK    241
X. laevis       DLPPAHNSSFPLVTNG..YIVPKPPFFRELPS.KIQAHKLALGNEHAVLLTSHWTLLTWGAGRHGQLGHGDLEDVEEPRIVDALHGVPMR    208
D. melanogaster IEHIACGSHINVAISSENCVYSIPSCLHQFSERQFRVKQLQCGHEHAVLLNANGDVFTWCNGLRGQLGLAELRVEETPQLLEALAGIKIT    229

                     220       230       240       250       260       270       280       290       300
H. sapiens      EVAAGGWHSVCVSETGDIYIWGWNESGQLALPTRNLAEDGETVAREATELNEDGSQVKRTGGAEDGAPAPFIAVQPFPALLDLPMGSDAV    303
M. muculus      KVAAGGWHSVCLSETGDIYIWGWNESGQLALPTRSGTEN.KAEREEATELNEDG.LKEELAVADAGAPAHFIAIQPFPALLDLPLGSDAV    304
D. rerio        AVAAGNWHSAAVSSGGDLYMWGWNESGQLGLPSR.......GLEEEKRRGNGSCNDDQPINTDGKSRTDVFISIQAFPALVDIANMSEIS    324
X. laevis       EVAAGGWHSASISESGDIYTWGWNESGQLGLPCK.......SQQ........CTSTEQSHLEDELGNIDEFITIQAFPALIDLPQESDAS    283
D. melanogaster QIAAGGWHSAAISAFGDLYTWGLNCSGQLGLRVMK..............PGGVLKEPTVFPLPQLQD..LPECACSQSGESNDDCAPL    301

                     310       320       330       340       350       360       370
H. sapiens      KASCGSRHTAVVTRTGELYTWGWGKYGQLG.HEDTTSLDRPRRVEYFVDKQLQVKAVTCGPWNTYVYAVEKGKS.............    376
M. muculus      MASCGSRHTAVVTRTGELYTWGWGKYGQLG.HKDSTSLDRPCCVEYFVERQLEVRAVTCGPWNTYVYAMERDKS.............    377
D. rerio        RISCGSRHTAAVTSAGDLYTWGWGQYGQLG.HGTEHSTDEPTPVDYFSSHSLSVKDVVCGSWNTFVSVVPKETSSVLTETHPGLLN    409
X. laevis       KISCGSRHTAAVSRSGELYTWGWGKYGQLG.HGDTISLDQPRLVHYFSVKHLCVNDVICRNWSTYVFCAER..............    353
D. melanogaster RVFAGSRHTLLIRRCGRLWVSGWCKHGQLGRQLQDLSYVDAFQALEGITMNPTVDDVLCGPWSTLLHLKCTA..............    373
```


non-conserved
similar
≥ 50% conserved
all match

Bibliography

Adzhubei, I.A., Schmidt, S., Peshkin, L., et al. (2010) A method and server for predicting damaging missense mutations. *Nature Methods*, 7 (4): 248–249. doi:10.1038/nmeth0410-248.

Agger, K., Cloos, P.A.C., Christensen, J., et al. (2007) UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature*, 449 (7163): 731–734. doi:10.1038/nature06145.

Ait Saada, A., Lambert, S.A.E. and Carr, A.M. (2018) Preserving replication fork integrity and competence via the homologous recombination pathway. *DNA Repair*, 71 (August): 135–147. doi:10.1016/j.dnarep.2018.08.017.

Alkalaeva, E.Z., Pisarev, A. V, Frolova, L.Y., et al. (2006) In Vitro Reconstitution of Eukaryotic Translation Reveals Cooperativity between Release Factors eRF1 and eRF3. *Cell*, 125 (6): 1125–1136. doi:10.1016/j.cell.2006.04.035.

Amendola, P.G., Zaghet, N., Ramalho, J.J., et al. (2017) JMJD-5/KDM8 regulates H3K36me2 and is required for late steps of homologous recombination and genome integrity. *PLoS Genetics*, 13 (2): 1–24. doi:10.1371/journal.pgen.1006632.

Anderson, L., Henderson, C. and Adachi, Y. (2001) Phosphorylation and Rapid Relocalization of 53BP1 to Nuclear Foci upon DNA Damage. *Molecular and Cellular Biology*, 21 (5): 1719–1729. doi:10.1128/mcb.21.5.1719-1729.2001.

Anglana, M., Apiou, F., Bensimon, A., et al. (2003) Dynamics of DNA replication in mammalian somatic cells: Nucleotide pool modulates origin choice and interorigin spacing. *Cell*, 114 (3): 385–394. doi:10.1016/S0092-8674(03)00569-5.

Askjaer, P., Galy, V., Hannak, E., et al. (2002) Ran GTPase Cycle and Importins alpha and beta Are Essential for Spindle Formation and Nuclear Envelope Assembly in Living Caenorhabditis elegans Embryos. *Molecular Biology of the Cell*, 13 (November): 4100–4109. doi:10.1091/mbc.E02.

Atsumi, Y., Minakawa, Y., Ono, M., et al. (2015) ATM and SIRT6/SNF2H Mediate Transient H2AX Stabilization When DSBs Form by Blocking HUWE1 to Allow Efficient γH2AX Foci Formation. *Cell Reports*, 13 (12): 2728–2740. doi:10.1016/j.celrep.2015.11.054.

Aziz, N., Hong, Y.H., Jo, M.K., et al. (2020) Molecular Signatures of JMJD10/MINA53 in Gastric Cancer. *Cancers*, 12 (5): 1141. doi:10.3390/cancers12051141.

Badie, S., Liao, C., Thanasoula, M., et al. (2009) RAD51C facilitates checkpoint signaling by promoting CHK2 phosphorylation. *Journal of Cell Biology*, 185 (4): 587–600. doi:10.1083/jcb.200811079.

Ball, H.L., Ehrhardt, M.R., Mordes, D.A., et al. (2007) Function of a Conserved Checkpoint Recruitment Domain in ATRIP Proteins. *Molecular and Cellular Biology*, 27 (9): 3367–3377. doi:10.1128/mcb.02238-06.

Bamba, C., Bobinnec, Y., Fukuda, M., et al. (2002) The GTpase Ran regulates chromosome positioning and nuclear envelope assembly in vivo. *Current Biology*, 12 (6): 503–507. doi:10.1016/S0960-9822(02)00741-8.

Banka, S., Lederer, D., Benoit, V., et al. (2015) Novel KDM6A (UTX) mutations and a clinical and molecular review of the X-linked Kabuki syndrome (KS2). *Clinical Genetics*, 87 (3): 252–258. doi:10.1111/cge.12363.

Bannister, A.J. and Kouzarides, T. (2011) Regulation of chromatin by histone modifications. *Nature Publishing Group*, pp. 381–395. doi:10.1038/cr.2011.22.

Barrett, A., Madsen, B., Copier, J., et al. (2002) PLU-1 nuclear protein, which is upregulated in breast cancer, shows restricted expression in normal human adult tissues: A new cancer/testis antigen? *International Journal of Cancer*, 101 (6): 581–588. doi:10.1002/ijc.10644.

Bell, S.P. and Dutta, A. (2002) DNA Replication in Eukaryotic Cells. *Annual Review of Biochemistry*, 71 (1): 333–374. doi:10.1146/annurev.biochem.71.110601.135425.

Bencokova, Z., Kaufmann, M.R., Pires, I.M., et al. (2009) ATM Activation and Signaling under Hypoxic Conditions. *Molecular and Cellular Biology*, 29 (2): 526–537. doi:10.1128/mcb.01301-08.

Bermejo, R., Lai, M.S. and Foiani, M. (2012) Preventing Replication Stress to Maintain Genome Stability: Resolving Conflicts between Replication and Transcription. *Molecular Cell*, 45 (6): 710–718. doi:10.1016/j.molcel.2012.03.001.

Berry, W.L. and Janknecht, R. (2013) KDM4/JMJD2 Histone demethylases: Epigenetic regulators in cancer cells. *Cancer Research*, 73 (10): 2936–2942. doi:10.1158/0008-5472.CAN-12-4300.

Berti, M., Chaudhuri, A.R., Thangavel, S., et al. (2013) Human RECQ1 promotes restart of replication forks reversed by DNA topoisomerase I inhibition. *Nature Structural and Molecular Biology*, 20 (3): 347–354. doi:10.1038/nsmb.2501.

Berti, M., Teloni, F., Mijic, S., et al. (2020) Sequential role of RAD51 paralog complexes in replication fork remodeling and restart. *Nature Communications*, 11 (1). doi:10.1038/s41467-020-17324-z.

Bhat, K.P. and Cortez, D. (2018) RPA and RAD51: Fork reversal, fork protection, and genome stability. *Nature Structural and Molecular Biology*, 25 (6): 446–453. doi:10.1038/s41594-018-0075-z.

Bischoff, F.R. and Ponstingl, H. (1991) Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. *Nature*, 354 (6348): 80–82. doi:10.1038/354080a0.

Blackford, A.N. and Jackson, S.P. (2017) ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Molecular Cell*, 66 (6): 801–817. doi:10.1016/j.molcel.2017.05.015.

Blanc, R.S. and Richard, S. (2017) Arginine Methylation: The Coming of Age. *Molecular Cell*, 65 (1): 8–24. doi:10.1016/j.molcel.2016.11.003.

Boeckel, J.N., Derlet, A., Glaser, S.F., et al. (2016) JMJD8 Regulates Angiogenic Sprouting and Cellular Metabolism by Interacting with Pyruvate Kinase M2 in Endothelial Cells. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 36 (7): 1425–1433. doi:10.1161/ATVBAHA.116.307695.

Boeing, S., Williamson, L., Encheva, V., et al. (2016) Multiomic Analysis of the UV-Induced DNA Damage Response. *Cell Reports*, 15 (7): 1597–1610. doi:10.1016/j.celrep.2016.04.047.

Boisvert, F.M., Déry, U., Masson, J.Y., et al. (2005) Arginine methylation of MRE11 by PRMT1 is required for DNA damage checkpoint control. *Genes and Development*, 19 (6): 671–676. doi:10.1101/gad.1279805.

Boni, J., Idani, A., Roca, C., et al. (2022) A decade of RAD51C and RAD51D germline variants in cancer. *Human Mutation*, 43 (3): 285–298. doi:10.1002/humu.24319.

Bonilla, B., Hengel, S.R., Grundy, M.K., et al. (2020) RAD51 Gene Family Structure and Function. *Annual Review of Genetics*, 54: 25–46. doi:10.1146/annurev-genet-021920-092410.

Böse, J., Gruber, A.D., Helming, L., et al. (2004) The phosphatidylserine receptor has essential functions during embryogenesis but not in apoptotic cell removal. *Journal of Biology*, 3 (4).

Bouwmeester, T., Bauch, A., Ruffner, H., et al. (2004) A physical and functional map of the human TNF-α/NF-κB signal transduction pathway. *Nature Cell Biology*, 6 (2): 97–105. doi:10.1038/ncb1086.

Bristow, R.G. and Hill, R.P. (2008) Hypoxia and metabolism: Hypoxia, DNA repair and genetic instability. *Nature Reviews Cancer*, 8 (3): 180–192. doi:10.1038/nrc2344.

Bruick, R.K. and McKnight, S.L. (2001) A conserved family of prolyl-4-hydroxylases that modify HIF. *Science*, 294 (5545): 1337–1340. doi:10.1126/science.1066373.

Bundred, J.R., Hendrix, E. and Coleman, M.L. (2018) The emerging roles of ribosomal histidyl hydroxylases in cell biology, physiology and disease. *Cellular and Molecular Life Sciences*, 75 (22): 4093–4105. doi:10.1007/s00018-018-2903-z.

Byun, T.S., Pacek, M., Yee, M.C., et al. (2005) Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. *Genes and Development*, 19 (9): 1040–1052. doi:10.1101/gad.1301205.

Cai, Q., Zhang, B., Sung, H., et al. (2014) Genome-wide association analysis in East Asians identifies breast cancer susceptibility loci at 1q32.1, 5q14.3 and 15q26.1. *Nature Genetics*, 46 (8): 886–890. doi:10.1038/ng.3041.

Carreira, A., Hilario, J., Amitani, I., et al. (2009) The BRC Repeats of BRCA2 Modulate the DNA-Binding Selectivity of RAD51. *Cell*, 136 (6): 1032–1043. doi:10.1016/j.cell.2009.02.019.

Ceccaldi, R., Sarangi, P. and D'Andrea, A.D. (2016) The Fanconi anaemia pathway: New players and new functions. *Nature Reviews Molecular Cell Biology*, 17 (6): 337–349. doi:10.1038/nrm.2016.48.

Chang, B., Chen, Y., Zhao, Y., et al. (2007) JMJD6 Is a Histone Arginine Demethylase. *Science*, 318 (5849): 444–447. doi:10.1126/science.1145801.

Chang, W.H., Forde, D. and Lai, A.G. (2019) Dual prognostic role of 2-oxoglutarate-dependent oxygenases in ten cancer types: Implications for cell cycle regulation and cell adhesion maintenance. *Cancer Communications*, 39 (1): 1–14. doi:10.1186/s40880-019-0369-5.

Chen, H., Lisby, M. and Symington, L.S. (2013) RPA Coordinates DNA End Resection and Prevents Formation of DNA Hairpins. *Molecular Cell*, 50 (4): 589–600. doi:10.1016/j.molcel.2013.04.032.

Chen, T., Ren, Z., Ye, L.C., et al. (2015a) Factor inhibiting HIF1α (FIH-1) functions as a tumor suppressor in human colorectal cancer by repressing HIF1α pathway. *Cancer Biology and Therapy*, 16 (2): 244–252. doi:10.1080/15384047.2014.1002346.

Chen, Y.H., Jones, M.J.K., Yin, Y., et al. (2015b) ATR-Mediated Phosphorylation of FANCI Regulates Dormant Origin Firing in Response to Replication Stress. *Molecular Cell*, 58 (2): 323–338. doi:10.1016/j.molcel.2015.02.031.

Cheng, Z., Hou, S., Wu, Y., et al. (2019) LINC01419 promotes cell proliferation and metastasis in lung adenocarcinoma via sponging miR-519b-3p to up-regulate RCCD1. *Biochemical and Biophysical Research Communications*, 520 (1): 107–114. doi:10.1016/j.bbrc.2019.09.090.

Chowdhury, R., Sekirnik, R., Brissett, N.C., et al. (2014) Ribosomal oxygenases are structurally conserved from prokaryotes to humans. *Nature*, 510 (7505): 422–426. doi:10.1038/nature13263.

Chun, J., Buechelmaier, E.S. and Powell, S.N. (2013) Rad51 Paralog Complexes BCDX2 and CX3 Act at Different Stages in the BRCA1-BRCA2-Dependent Homologous Recombination Pathway. *Molecular and Cellular Biology*, 33 (2): 387–395. doi:10.1128/mcb.00465-12.

Cimprich, K.A. and Cortez, D. (2008) ATR: An essential regulator of genome integrity. *Nature Reviews Molecular Cell Biology*, 9 (8): 616–627. doi:10.1038/nrm2450.

Clarke, T.L., Sanchez-Bailon, M.P., Chiang, K., et al. (2017) PRMT5-Dependent Methylation of the TIP60 Coactivator RUVBL1 Is a Key Regulator of Homologous Recombination. *Molecular Cell*, 65 (5): 900-916.e7. doi:10.1016/j.molcel.2017.01.019.

Cockman, M.E., Lancaster, D.E., Stolze, I.P., et al. (2006) Posttranslational hydroxylation of ankyrin repeats in IκB proteins by the hypoxia-inducible factor (HIF) asparaginyl hydroxylase, factor inhibiting HIF (FIH). *Proceedings of the National Academy of Sciences of the United States of America*, 103 (40): 14767–14772. doi:10.1073/pnas.0606877103.

Cockman, M.E., Sugimoto, Y., Pegg, H.B., et al. (2022) Widespread hydroxylation of unstructured lysine-rich protein domains by JMJD6. *Proceedings of the National Academy of Sciences of the United States of America*, 119 (32): 1–10. doi:10.1073/pnas.2201483119.

Cockman, M.E., Webb, J.D., Kramer, H.B., et al. (2009) Proteomics-based identification of novel factor inhibiting hypoxia-inducible factor (FIH) substrates indicates widespread asparaginyl hydroxylation of ankyrin repeat domain-containing proteins. *Molecular and Cellular Proteomics*, 8 (3): 535–546. doi:10.1074/mcp.M800340-MCP200.

Coleman, M.L., McDonough, M.A., Hewitson, K.S., et al. (2007) Asparaginyl hydroxylation of the notch ankyrin repeat domain by factor inhibiting hypoxia-inducible factor. *Journal of Biological Chemistry*, 282 (33): 24027–24038. doi:10.1074/jbc.M704102200.

Collins, R.R.J., Patel, K., Putnam, W.C., et al. (2017) Oncometabolites: A new paradigm for oncology, metabolism, and the clinical laboratory. *Clinical Chemistry*, 63 (12): 1812–1820. doi:10.1373/clinchem.2016.267666.

Cortez, D. (2015) Preventing replication fork collapse to maintain genome integrity. *DNA Repair*, 32: 149–157. doi:10.1016/J.DNAREP.2015.04.026.

Cortez, D., Glick, G. and Elledge, S.J. (2004) Minichromosome maintenance proteins are direct targets of the ATM and ATR checkpoint kinases. *Proceedings of the National Academy of Sciences of the United States of America*, 101 (27): 10078–10083. doi:10.1073/pnas.0403410101.

Dalgliesh, G.L., Furge, K., Greenman, C., et al. (2010) Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature*, 463 (7279): 360–363. doi:10.1038/nature08672.

Das, M., Singh, S., Pradhan, S., et al. (2014) MCM Paradox: Abundance of Eukaryotic Replicative Helicases and Genomic Integrity. *Molecular Biology International*, 2014: 1–11. doi:10.1155/2014/574850.

Van Deursen, F., Sengupta, S., De Piccoli, G., et al. (2012) Mcm 10 associates with the loaded DNA helicase at replication origins and defines a novel step in its activation. *EMBO Journal*, 31 (9): 2195–2206. doi:10.1038/emboj.2012.69.

D'Oto, A., Tian, Q.-W., Davidoff, A.M., et al. (2016) Histone demethylases and their roles in cancer epigenetics. *Journal of medical oncology and therapeutics*, 1 (2): 34–40. doi:10.1053/j.gastro.2016.08.014.CagY.

Duan, G. and Walther, D. (2015) The Roles of Post-translational Modifications in the Context of Protein Interaction Networks. *PLoS Computational Biology*, 11 (2): 1–23. doi:10.1371/journal.pcbi.1004049.

Eilbracht, J., Kneissel, S., Hofmann, A., et al. (2005) Protein NO52-a constitutive nucleolar component sharing high sequence homologies to protein NO66. *European Journal of Cell Biology*, 84 (2–3): 279–294. doi:10.1016/j.ejcb.2004.12.022.

Eilbracht, J., Reichenzeller, M., Hergt, M., et al. (2004) NO66, a Highly Conserved Dual Location Protein in the Nucleolus and in a Special Type of Synchronously Replicating Chromatin. *Molecular Biology of the Cell*, 15 (4): 1816–1832. doi:10.1091/mbc.e03-08-0623.

Engström, W., Ward, A. and Moorwood, K. (2010) The role of scaffold proteins in JNK signalling. *Cell Proliferation*, 43 (1): 56–66. doi:10.1111/j.1365-2184.2009.00654.x.

Epstein, A.C.R., Gleadle, J.M., McNeill, L.A., et al. (2001) C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell*, 107 (1): 43–54. doi:10.1016/S0092-8674(01)00507-4.

Evans, R., O'Neill, M., Pritzel, A., et al. (2022) Protein complex prediction with AlphaFold-Multimer. *bioRxiv*, p. 2021.10.04.463034.

Fanning, E., Klimovich, V. and Nager, A.R. (2006) A dynamic model for replication protein A (RPA) function in DNA processing pathways. *Nucleic Acids Research*, 34 (15): 4126–4137. doi:10.1093/nar/gkl550.

Fenech, M., Kirsch-Volders, M., Natarajan, A.T., et al. (2011) Molecular mechanisms of micronucleus, nucleoplasmic bridge and nuclear bud formation in mammalian and human cells. *Mutagenesis*, 26 (1): 125–132. doi:10.1093/mutage/geq052.

Feng, T., Yamamoto, A., Wilkins, S.E., et al. (2014) Optimal Translational Termination Requires C4 Lysyl Hydroxylation of eRF1. *Molecular Cell*, 53 (4): 645–654. doi:10.1016/j.molcel.2013.12.028.

Ferguson, J.E., Wu, Y., Smith, K., et al. (2007) ASB4 Is a Hydroxylation Substrate of FIH and Promotes Vascular Differentiation via an Oxygen-Dependent Mechanism. *Molecular and Cellular Biology*, 27 (18): 6407–6419. doi:10.1128/mcb.00511-07.

Fernandez-Vidal, A., Vignard, J. and Mirey, G. (2017) Around and beyond 53BP1 nuclear bodies. *International Journal of Molecular Sciences*, 18 (12): 1–20. doi:10.3390/ijms18122611.

Ferreira, M.A., Gamazon, E.R., Al-Ejeh, F., et al. (2019) Genome-wide association and transcriptome studies identify target genes and risk loci for breast cancer. *Nature Communications*, 10 (1): 1–18. doi:10.1038/s41467-018-08053-5.

Figueroa, M.E., Abdel-Wahab, O., Lu, C., et al. (2010) Leukemic IDH1 and IDH2 Mutations Result in a Hypermethylation Phenotype, Disrupt TET2 Function, and Impair Hematopoietic Differentiation. *Cancer Cell*, 18 (6): 553–567. doi:10.1016/j.ccr.2010.11.015.

Fletcher, S.C., Hall, C.L., Kennedy, T.J., et al. (2023) Impaired protein hydroxylase activity causes replication stress and developmental abnormalities in humans. *Journal of Clinical Investigation*, 133 (7). doi:10.1172/jci152784.

Frigola, J., He, J., Kinkelin, K., et al. (2017) Cdt1 stabilizes an open MCM ring for helicase loading. *Nature Communications*, 8: 1–10. doi:10.1038/ncomms15720.

Fu, Y., Dai, Q., Zhang, W., et al. (2010) The AlkB domain of mammalian ABH8 catalyzes hydroxylation of 5-methoxycarbonylmethyluridine at the wobble position of tRNA. *Angewandte Chemie - International Edition*, 49 (47): 8885–8888. doi:10.1002/anie.201001242.

Furuhashi, H. and Kelly, W.G. (2010) The Epigenetics of Germ-line Immortality: Lessons from an Elegant Model System. *Development, Growth & Differentiation*, 52 (6): 527–532. doi:10.1111/j.1440-169X.2010.01179.x.

Gaillard, H., García-Muse, T. and Aguilera, A. (2015) Replication stress and cancer. *Nature Reviews Cancer*, 15 (5): 276–280. doi:10.1038/nrc3916.

García-Muse, T. and Aguilera, A. (2016) Transcription-replication conflicts: How they occur and how they are resolved. *Nature Reviews Molecular Cell Biology*, 17 (9): 553–563. doi:10.1038/nrm.2016.88.

Garcin, E.B., Gon, S., Sullivan, M.R., et al. (2019) Differential requirements for the RAD51 paralogs in genome repair and maintenance in human cells. *PLoS Genetics*, 15 (10): 1–29. doi:10.1371/journal.pgen.1008355.

Garribba, L., Wu, W., Özer, Ö., et al. (2018) Inducing and Detecting Mitotic DNA Synthesis at Difficult-to-Replicate Loci. *Methods in Enzymology*, 601: 45–58. doi:10.1016/bs.mie.2017.11.025.

Ge, W., Wolf, A., Feng, T., et al. (2012) Oxygenase-catalyzed ribosome hydroxylation occurs in prokaryotes and humans. *Nature Chemical Biology*, 8 (12): 960–962. doi:10.1038/nchembio.1093.

Ge, X.Q. and Blow, J.J. (2010) Chk1 inhibits replication factory activation but allows dormant origin firing in existing factories. *Journal of Cell Biology*, 191 (7): 1285–1297. doi:10.1083/jcb.201007074.

Ge, X.Q., Jackson, D.A. and Blow, J.J. (2007) Dormant origins licensed by excess Mcm2-7 are required for human cells to survive replicative stress. *Genes and Development*, 21 (24): 3331–3341. doi:10.1101/gad.457807.

Gelot, C., Magdalou, I. and Lopez, B.S. (2015) Replication stress in mammalian cells and its consequences for mitosis. *Genes*, 6 (2): 267–298. doi:10.3390/genes6020267.

Gibney, E.R. and Nolan, C.M. (2010) Epigenetics and gene expression. *Heredity*, pp. 4–13. doi:10.1038/hdy.2010.54.

Giles, A.C. and Grill, B. (2020) Roles of the HUWE1 ubiquitin ligase in nervous system development, function and disease. *Neural Development*, 15 (1): 1–18. doi:10.1186/s13064-020-00143-9.

Gjaltema, R.A.F. and Bank, R.A. (2017) *Molecular insights into prolyl and lysyl hydroxylation of fibrillar collagens in health and disease in health and disease.*, 9238. doi:10.1080/10409238.2016.1269716.

Glover, T.W., Berger, C., Coyle, J., et al. (1984) DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. *Human genetics*, 67 (2): 136–142. doi:10.1007/BF00272988.

Gong, X., Du, D., Deng, Y., et al. (2020) The structure and regulation of the E3 ubiquitin ligase HUWE1 and its biological functions in cancer. *Investigational New Drugs*, 38 (2): 515–524. doi:10.1007/s10637-020-00894-6.

Greenberg, M.V.C. and Bourc, D. (2019) The diverse roles of DNA methylation in mammalian development and disease. *Nature Reviews Molecular Cell Biology*, 20 (October). doi:10.1038/s41580-019-0159-6.

Gu, Y.Z., Moran, S.M., Hogenesch, J.B., et al. (1998) Molecular characterization and chromosomal localization of a third α- class hypoxia inducible factor subunit, HIF3α. *Gene Expression*, 7 (3): 205–213.

Guo, A., Gu, H., Zhou, J., et al. (2014) Immunoaffinity enrichment and mass spectrometry analysis of protein methylation. *Molecular and Cellular Proteomics*, 13 (1): 372–387. doi:10.1074/mcp.O113.027870.

Gusev, A., Lawrenson, K., Lin, X., et al. (2019) A transcriptome-wide association study of high-grade serous epithelial ovarian cancer identifies new susceptibility genes and splice variants. *Nature Genetics*, 51 (5): 815–823. doi:10.1038/s41588-019-0395-x.

Hadjebi, O., Casas-Terradellas, E., Garcia-Gonzalo, F.R., et al. (2008) The RCC1 superfamily: From genes, to function, to disease. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1783 (8): 1467–1479. doi:10.1016/j.bbamcr.2008.03.015.

Hammond, E.M., Denko, N.C., Dorie, M.J., et al. (2002) Hypoxia Links ATR and p53 through Replication Arrest. *Molecular and Cellular Biology*, 22 (6): 1834–1843. doi:10.1128/mcb.22.6.1834-1843.2002.

Hammond, E.M., Dorie, M.J. and Giaccia, A.J. (2004) Inhibition of ATR leads to increased sensitivity to hypoxia/reoxygenation. *Cancer Research*, 64 (18): 6556–6562. doi:10.1158/0008-5472.CAN-04-1520.

Hanahan, D. and Weinberg, R.A. (2011) Hallmarks of cancer: The next generation. *Cell*, 144 (5): 646–674. doi:10.1016/j.cell.2011.02.013.

Harley, M.E., Murina, O., Leitch, A., et al. (2016) TRAIP promotes DNA damage response during genome replication and is mutated in primordial dwarfism. *Nature Genetics*, 48 (1): 36–43. doi:10.1038/ng.3451.TRAIP.

He, Y.F., Li, B.Z., Li, Z., et al. (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*, 333 (6047): 1303–1307. doi:10.1126/science.1210944.

He, Z., Wu, J., Su, X., et al. (2016) JMJD5 (jumonji domain-containing 5) associates with spindle microtubules and is required for proper mitosis. *Journal of Biological Chemistry*, 291 (9): 4684–4697. doi:10.1074/jbc.M115.672642.

Hendrix, E., Andrijes, R., Boora, U., et al. (2023) A protein hydroxylase couples epithelial membrane biology to nucleolar ribosome biogenesis. *bioRxiv*, pp. 1–32. doi:https://doi.org/10.1101/2023.03.15.532818.

Hewitson, K.S., Holmes, S.L., Ehrismann, D., et al. (2008) Evidence that two enzyme-derived histidine ligands are sufficient for iron binding and catalysis by factor inhibiting HIF (FIH). *Journal of Biological Chemistry*, 283 (38): 25971–25978. doi:10.1074/jbc.M804999200.

Hewitson, K.S., Mcneill, L.A., Riordan, M. V, et al. (2002) Hypoxia-inducible Factor (HIF) Asparagine Hydroxylase Is Identical to Factor Inhibiting HIF (FIH) and Is Related to the Cupin Structural Family *. *Journal of Biological Chemistry*, 277 (29): 26351–26355. doi:10.1074/jbc.C200273200.

Hoffman, J.D., Graff, R.E., Emami, N.C., et al. (2017) Cis-eQTL-based trans-ethnic meta-analysis reveals novel genes associated with breast cancer risk. *PLoS Genetics*, 13 (3): 1–19. doi:10.1371/journal.pgen.1006690.

Højfeldt, J.W., Agger, K. and Helin, K. (2013) Histone lysine demethylases as targets for anticancer therapy. *Nature Reviews Drug Discovery*, 12 (12): 917–930. doi:10.1038/nrd4154.

Hornbeck, P. V., Zhang, B., Murray, B., et al. (2015) PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations. *Nucleic Acids Research*, 43 (D1): D512–D520. doi:10.1093/nar/gku1267.

Hsia, D.A., Tepper, C.G., Pochampalli, M.R., et al. (2010) KDM8, a H3K36me2 histone demethylase that acts in the cyclin A1 coding region to regulate cancer cell proliferation. *Proceedings of the National Academy of Sciences*, 107 (21): 9671–9676. doi:10.1073/pnas.1000401107.

Hu, S., Jing, Y., Liu, F., et al. (2022) Association between XRCC3 rs861539 Polymorphism and the Risk of Ovarian Cancer: Meta-Analysis and Trial Sequential Analysis. *BioMed Research International*, 2022. doi:10.1155/2022/3915402.

Huang, X., Zhang, L., Qi, H., et al. (2013) Identification and functional implication of nuclear localization signals in the N-terminal domain of JMJD5. *Biochimie*, 95 (11): 2114–2122. doi:10.1016/j.biochi.2013.08.002.

Huang, Y. and Rao, A. (2014) Connections between TET proteins and aberrant DNA modification in cancer. *Trends in Genetics*, 30 (10): 464–474. doi:10.1016/j.tig.2014.07.005.

Huttlin, E.L., Bruckner, R.J., Paulo, J.A., et al. (2017) Architecture of the human interactome defines protein communities and disease networks. *Nature*, 545 (7655): 505–509. doi:10.1038/nature22366.

Ilves, I., Petojevic, T., Pesavento, J.J., et al. (2010) Activation of the MCM2-7 Helicase by Association with Cdc45 and GINS Proteins. *Molecular Cell*, 37 (2): 247–258. doi:10.1016/j.molcel.2009.12.030.

Ishiai, M. (2021) Regulation of the fanconi anemia DNA repair pathway by phosphorylation and monoubiquitination. *Genes*, 12 (11). doi:10.3390/genes12111763.

Ishikawa, K., Azuma, S., Ikawa, S., et al. (2003) Cloning and characterization of Xenopus laevis drg2, a member of the developmentally regulated GTP-binding protein subfamily. *Gene*, 322 (1–2): 105–112. doi:10.1016/j.gene.2003.08.016.

Ishimura, A., Terashima, M., Kondoh, G., et al. (2012) Jmjd5, an H3K36me2 histone demethylase, modulates embryonic cell proliferation through the regulation of Cdkn1a expression. *Development*, 139 (4): 749–759. doi:10.1242/dev.074138.

Ishimura, A., Terashima, M., Tange, S., et al. (2016) Jmjd5 functions as a regulator of p53 signaling during mouse embryogenesis. *Cell and Tissue Research*, 363 (3): 723–733. doi:10.1007/s00441-015-2276-7.

Islam, M.S., Markoulides, M., Chowdhury, R., et al. (2022) Structural analysis of the 2-oxoglutarate binding site of the circadian rhythm linked oxygenase JMJD5. *Scientific Reports*, 12 (1): 1–12. doi:10.1038/s41598-022-24154-0.

Islam, S., Leissing, T.M., Chowdhury, R., et al. (2018) 2-Oxoglutarate-Dependent Oxygenases. *Annual Review of Biochemistry*, 87: 585–620.

Ivan, M., Kondo, K., Yang, H., et al. (2001) HIF ⌣ Targeted for VHL-Mediated Destruction by Proline Hydroxylation : Implications for O 2 Sensing. *Science*, 292 (March): 464–469.

Jaakkola, P., Mole, D.R., Tian, Y.M., et al. (2001) Targeting of HIF-α to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. *Science*, 292 (5516): 468–472. doi:10.1126/science.1059796.

Ji, A.X., Chu, A., Nielsen, T.K., et al. (2016) Structural Insights into KCTD Protein Assembly and Cullin3 Recognition. *Journal of Molecular Biology*, 428 (1): 92–107. doi:10.1016/j.jmb.2015.08.019.

Jia, G., Fu, Y., Zhao, X., et al. (2011) N6-Methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nature Chemical Biology*, 7 (12): 885–887. doi:10.1038/nchembio.687.

Johansson, C., Tumber, A., Che, K., et al. (2014) The roles of Jumonji-type oxygenases in human disease. *Epigenomics*, 6 (1): 89–120. doi:10.2217/epi.13.79.

Jones, M.A., Covington, M.F., DiTacchio, L., et al. (2010) Jumonji domain protein JMJD5 functions in both the plant and human circadian systems. *Proceedings of the National Academy of Sciences*, 107 (50): 21623–21628. doi:10.1073/pnas.1014204108.

Jones, M.A. and Harmer, S.L. (2011) JMJD5 functions in concert with TOC1 in the arabidopsis circadian system. *Plant Signaling and Behavior*, 6 (3): 445–448. doi:10.4161/psb.6.3.14654.

Jones, M.A., Morohashi, K., Grotewold, E., et al. (2019) Arabidopsis JMJD5/JMJ30 Acts Independently of LUX ARRHYTHMO Within the Plant Circadian Clock to Enable Temperature Compensation. *Frontiers in Plant Science*, 10 (February): 1–12. doi:10.3389/fpls.2019.00057.

Jumper, J., Evans, R., Pritzel, A., et al. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, 596 (7873): 583–589. doi:10.1038/s41586-021-03819-2.

Kang, J., Shin, S.H., Yoon, H., et al. (2018) FIH is an oxygen sensor in ovarian cancer for G9a/GLP-driven epigenetic regulation of metastasis-related genes. *Cancer Research*, 78 (5): 1184–1199. doi:10.1158/0008-5472.CAN-17-2506.

Kao, S.H., Wu, H.T. and Wu, K.J. (2018) Ubiquitination by HUWE1 in tumorigenesis and beyond. *Journal of Biomedical Science*, 25 (1): 1–15. doi:10.1186/s12929-018-0470-0.

Kar, S.P., Beesley, J., Al Olama, A.A., et al. (2016) Genome-wide meta-analyses of breast, ovarian, and prostate cancer association studies identify multiple new susceptibility loci shared by at least two cancer types. *Cancer Discovery*, 6 (9): 1052–1067. doi:10.1158/2159-8290.CD-15-1227.

Khetarpal, P., Das, S., Panigrahi, I., et al. (2016) Primordial dwarfism: overview of clinical and genetic aspects. *Molecular Genetics and Genomics*, 291 (1): 1–15. doi:10.1007/s00438-015-1110-y.

Khoueiry, R., Sohni, A., Thienpont, B., et al. (2017) Lineage-specific functions of TET1 in the postimplantation mouse embryo. *Nature Genetics*, 49 (7): 1061–1072. doi:10.1038/ng.3868.

Kim, I., Shin, S.H., Lee, J.E., et al. (2019) Oxygen sensor FIH inhibits HACE1-dependent ubiquitination of Rac1 to enhance metastatic potential in breast cancer cells. *Oncogene*, 38 (19): 3651–3666. doi:10.1038/s41388-019-0676-y.

Kim, J.H., Lee, J.H., Lee, I.S., et al. (2017) Histone lysine methylation and neurodevelopmental disorders. *International Journal of Molecular Sciences*, 18 (7): 1–20. doi:10.3390/ijms18071404.

Kimanius, D., Dong, L., Sharov, G., et al. (2021) New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochemical Journal*, 478 (24): 4169–4185. doi:10.1042/BCJ20210708.

Kivirikko, K.I. and Prockop, D.J. (1967) Enzymatic Hydroxylation of Proline and Lysine in Protocollagen. *Proceedings of the National Academy of Sciences*, 57 (3): 782–789. doi:10.1073/pnas.57.3.782.

Klose, R.J., Kallin, E.M. and Zhang, Y. (2006) JmjC-domain-containing proteins and histone demethylation. *Nature Reviews Genetics*, 7 (9): 715–727. doi:10.1038/nrg1945.

Knorre, D.G., Kudryashova, N.V. and Godovikova, T.S. (2009) Chemical and Functional Aspects of Posttranslational Modification of Proteins. *Acta Naturae*, 1 (3): 29–51. doi:10.32607/20758251-2009-1-3-29-51.

Ko, M., An, J., Pastor, W.A., et al. (2015) TET proteins and 5-methylcytosine oxidation in hematological cancers. *Immunological Reviews*, 263 (1): 6–21. doi:10.1111/imr.12239.

Ko, M., Huang, Y., Jankowska, A.M., et al. (2010) Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature*, pp. 10–14. doi:10.1038/nature09586.

Kolinjivadi, A.M., Sannino, V., De Antoni, A., et al. (2017) Smarcal1-Mediated Fork Reversal Triggers Mre11-Dependent Degradation of Nascent DNA in the Absence of Brca2 and Stable Rad51 Nucleofilaments. *Molecular Cell*, 67 (5): 867-881.e7. doi:10.1016/j.molcel.2017.07.001.

Komiya, K., Sueoka-Aragane, N., Sato, A., et al. (2010) Expression of Mina53, a novel c-Myc target gene, is a favorable prognostic marker in early stage lung cancer. *Lung Cancer*, 69 (2): 232–238. doi:10.1016/j.lungcan.2009.10.010.

Koyama, M., Sasaki, T., Sasaki, N., et al. (2017) Crystal structure of human WBSCR16, an RCC1-like protein in mitochondria. *Protein Science*, 26 (9): 1870–1877. doi:10.1002/pro.3210.

Krasilnikova, M.M. and Mirkin, S.M. (2004) Replication Stalling at Friedreich's Ataxia (GAA) n Repeats In Vivo . *Molecular and Cellular Biology*, 24 (6): 2286–2295. doi:10.1128/mcb.24.6.2286-2295.2004.

Kroeze, S.G.C., Vermaat, J.S., Van Brussel, A., et al. (2010) Expression of nuclear FIH independently predicts overall survival of clear cell renal cell carcinoma patients. *European Journal of Cancer*, 46 (18): 3375–3382. doi:10.1016/j.ejca.2010.07.018.

Kumar, S., Stecher, G., Li, M., et al. (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, 35 (6): 1547–1549. doi:10.1093/molbev/msy096.

Kumar, S.A., Thakur, C., Li, L., et al. (2017) Pathological and prognostic role of mdig in pancreatic cancer. *Genes and Cancer*, 8 (7–8): 650–658. doi:10.18632/genesandcancer.149.

Lancaster, D.E., McNeill, L.A., McDonough, M.A., et al. (2004) Disruption of dimerization and substrate phosphorylation inhibit factor inhibiting hypoxia-inducible factor (FIH) activity. *Biochemical Journal*, 383 (3): 429–437. doi:10.1042/BJ20040735.

Lando, D., Peet, D.J., Gorman, J.J., et al. (2002) FIH-1 is an asparaginyl hydroxylase enzyme that regulates the transcriptional activity of hypoxia-inducible factor. *Genes and Development*, 16 (12): 1466–1471. doi:10.1101/gad.991402.

Leman, A.R. and Noguchi, E. (2013) *The replication fork: Understanding the eukaryotic replication machinery and the challenges to genome duplication*. doi:10.3390/genes4010001.

Li, F., Li, C., Jiang, Z., et al. (2011) XRCC3 T241M polymorphism and bladder cancer risk: A meta-analysis. *Urology*, 77 (2): 511.e1-511.e5. doi:10.1016/j.urology.2010.07.003.

Li, J., Mahajan, A. and Tsai, M.D. (2006) Ankyrin repeat: A unique motif mediating protein-protein interactions. *Biochemistry*, 45 (51): 15168–15178. doi:10.1021/bi062188q.

Li, X. and Heyer, W.D. (2008) Homologous recombination in DNA repair and DNA damage tolerance. *Cell Research*, 18 (1): 99–113. doi:10.1038/cr.2008.1.

Liao, H., Winkfein, R.J., Mack, G., et al. (1995) CENP-F is a protein of the nuclear matrix that assembles onto kinetochores at late G2 and is rapidly degraded after mitosis. *Journal of Cell Biology*, 130 (3): 507–518. doi:10.1083/jcb.130.3.507.

Lin, W., Cao, J., Liu, J., et al. (2011) Loss of the retinoblastoma binding protein 2 (RBP2) histone demethylase suppresses tumorigenesis in mice lacking Rb1 or Men1. *Proceedings of the National Academy of Sciences of the United States of America*, 108 (33): 13379–13386. doi:10.1073/pnas.1110104108.

Lisy, K. and Peet, D.J. (2008) Turn me on: Regulating HIF transcriptional activity. *Cell Death and Differentiation*, 15 (4): 642–649. doi:10.1038/sj.cdd.4402315.

Liu, C., Gilmont, R.R., Benndorf, R., et al. (2000) Identification and characterization of a novel protein from Sertoli cells, PASS1, that associates with mammalian small stress protein hsp27. *Journal of Biological Chemistry*, 275 (25): 18724–18731. doi:10.1074/jbc.M001981200.

Liu, G., Qi, H. and Shen, J. (2023) JMJD5 inhibits lung cancer progression by regulating glucose metabolism through the p53/TIGAR pathway. *Medical Oncology*, 40 (5): 145. doi:10.1007/s12032-023-02016-7.

Liu, H., Ramachandran, S., Fong, N., et al. (2020) JMJD5 couples with CDK9 to release the paused RNA polymerase II. *Proceedings of the National Academy of Sciences of the United States of America*, 117 (33): 19888–19895. doi:10.1073/PNAS.2005745117.

Liu, H., Wang, C., Lee, S., et al. (2017a) Clipping of arginine-methylated histone tails by JMJD5 and JMJD7. *Proceedings of the National Academy of Sciences*, p. 201706831. doi:10.1073/pnas.1706831114.

Liu, H., Wei, P., Zhang, Q., et al. (2022) The Novel Protease Activities of JMJD5–JMJD6–JMJD7 and Arginine Methylation Activities of Arginine Methyltransferases Are Likely Coupled. *Biomolecules*, 12 (3). doi:10.3390/biom12030347.

Liu, N., Schild, D., Thelen, M.P., et al. (2002) Involvement of Rad51C in two distinct protein complexes of Rad51 paralogs in human cells. *Nucleic Acids Research*, 30 (4): 1009–1015. doi:10.1093/nar/30.4.1009.

Liu, Y., Masson, J.Y., Shah, R., et al. (2004) RAD51C Is Required for Holliday Junction Processing in Mammalian Cells. *Science*, 303 (5655): 243–246. doi:10.1126/science.1093037.

Liu, Z., Chen, O., Wall, J.B.J., et al. (2017b) Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Scientific Reports*, 7 (1): 1–9. doi:10.1038/s41598-017-02460-2.

Loenarz, C. and Schofield, C.J. (2008) Expanding chemical biology of 2-oxoglutarate oxygenases. *Nature Chemical Biology*, 4 (3): 152–156. doi:10.1038/nchembio0308-152.

Loenarz, C. and Schofield, C.J. (2011) Physiological and biochemical aspects of hydroxylations and demethylations catalyzed by human 2-oxoglutarate oxygenases. *Trends in Biochemical Sciences*, 36 (1): 7–18. doi:10.1016/j.tibs.2010.07.002.

Lopez-Girona, A., Tanaka, K., Chen, X.B., et al. (2001) Serine-345 is required for Rad3-dependent phosphorylation and function of checkpoint kinase Chk1 in fission yeast. *Proceedings of the National Academy of Sciences of the United States of America*, 98 (20): 11289–11294. doi:10.1073/pnas.191557598.

Lu, Y., Chang, Q., Zhang, Y., et al. (2009) Lung cancer-associated JmjC domain protein mdig suppresses formation of tri-methyl lysine 9 of histone H3. *Cell Cycle*, 8 (13): 2101–2109. doi:10.4161/cc.8.13.8927.

Luzhna, L., Kathiria, P. and Kovalchuk, O. (2013) Micronuclei in genotoxicity assessment: From genetics to epigenetics and beyond. *Frontiers in Genetics*, 4 (JUL): 1–17. doi:10.3389/fgene.2013.00131.

Macheret, M. and Halazonetis, T.D. (2015) DNA replication stress as a hallmark of cancer. *Annual Review of Pathology: Mechanisms of Disease*, 10: 425–448. doi:10.1146/annurev-pathol-012414-040424.

Majka, J., Binz, S.K., Wold, M.S., et al. (2006) Replication protein a directs loading of the DNA damage checkpoint clamp to 5'-DNA junctions. *Journal of Biological Chemistry*, 281 (38): 27855–27861. doi:10.1074/jbc.M605176200.

Marcon, E., Ni, Z., Pu, S., et al. (2014) Human-chromatin-related protein interactions identify a demethylase complex required for chromosome segregation. *Cell Reports*, 8 (1): 297–310. doi:10.1016/j.celrep.2014.05.050.

Markolovic, S., Zhuang, Q., Wilkins, S.E., et al. (2018) The Jumonji-C oxygenase JMJD7 catalyzes (3S)-lysyl hydroxylation of TRAFAC GTPases. *Nature Chemical Biology*, 14 (7): 688–695. doi:10.1038/s41589-018-0071-y.

Martin, C. and Zhang, Y. (2005) The diverse functions of histone lysine methylation. *Nature Reviews Molecular Cell Biology*. 6 (11) pp. 838–849. doi:10.1038/nrm1761.

Martinez, S. and Hausinger, R.P. (2015) Catalytic mechanisms of Fe(II)- and 2-Oxoglutarate-dependent oxygenases. *Journal of Biological Chemistry*, 290 (34): 20702–20711. doi:10.1074/jbc.R115.648691.

Mason, J.M., Chan, Y.L., Weichselbaum, R.W., et al. (2019) Non-enzymatic roles of human RAD51 at stalled replication forks. *Nature Communications*, 10 (1). doi:10.1038/s41467-019-12297-0.

Masson, N., Willam, C., Maxwell, P.H., et al. (2001) Independent function of two destruction domains in hypoxia-inducible factor-α chains activated by prolyl hydroxylation. *EMBO Journal*, 20 (18): 5197–5206. doi:10.1093/emboj/20.18.5197.

Matsunami, N., Hensel, C.H., Baird, L., et al. (2014) Identification of rare DNA sequence variants in high-risk autism families and their prevalence in a large case/control population. *Molecular Autism*, 5 (1): 1–18. doi:10.1186/2040-2392-5-5.

Matsuoka, S., Ballif, B.A., Smogorzewska, A., et al. (2007) ATM and ATR Substrate Analysis Reveals Extensive Protein Networks Responsive to DNA Damage. *Science*, 316 (5828): 1160–1166. doi:10.1126/science.1140321.

Mayer, M.P. and Bukau, B. (2005) Hsp70 chaperones: Cellular functions and molecular mechanism. *Cellular and Molecular Life Sciences*, 62 (6): 670–684. doi:10.1007/s00018-004-4464-6.

McDonough, M.A., Loenarz, C., Chowdhury, R., et al. (2010) Structural studies on human 2-oxoglutarate dependent oxygenases. *Current Opinion in Structural Biology*, 20 (6): 659–672. doi:10.1016/j.sbi.2010.08.006.

Mejlvang, J., Feng, Y., Alabert, C., et al. (2014) New histone supply regulates replication fork speed and PCNA unloading. *Journal of Cell Biology*, 204 (1): 29–43. doi:10.1083/jcb.201305017.

Mellacheruvu, D., Wright, Z., Couzens, A.L., et al. (2013) The CRAPome: a contaminant repository for affinity purification–mass spectrometry data. *Nature Methods*, 10 (8): 730–736. doi:10.1038/nmeth.2557.

Mersaoui, S.Y., Yu, Z., Coulombe, Y., et al. (2019) Arginine methylation of the DDX 5 helicase RGG / RG motif by PRMT 5 regulates resolution of RNA:DNA hybrids . *The EMBO Journal*, 38 (15): 1–20. doi:10.15252/embj.2018100986.

Meyer, K.D. and Jaffrey, S.R. (2014) The dynamic epitranscriptome: N6-methyladenosine and gene expression control. *Nature Reviews Molecular Cell Biology*. 15 (5) pp. 313–326. doi:10.1038/nrm3785.

Mijic, S., Zellweger, R., Chappidi, N., et al. (2017) Replication fork reversal triggers fork degradation in BRCA2-defective cells. *Nature Communications*, 8 (1): 1–11. doi:10.1038/s41467-017-01164-5.

Minocherhomji, S., Ying, S., Bjerregaard, V.A., et al. (2015) Replication stress activates DNA repair synthesis in mitosis. *Nature*, 528 (7581): 286–290. doi:10.1038/nature16139.

Mirdita, M., Schütze, K., Moriwaki, Y., et al. (2022) ColabFold: making protein folding accessible to all. *Nature Methods*, 19 (6): 679–682. doi:10.1038/s41592-022-01488-1.

Mizushima, T., Takahashi, N. and Stillman, B. (2000) Cdc6p modulates the structure and DNA binding activity of the origin recognition complex in vitro. *Genes and Development*, 14 (13): 1631–1641. doi:10.1101/gad.14.13.1631.

Mollinari, C., Reynaud, C., Martineau-Thuillier, S., et al. (2003) The mammalian passenger protein TD-60 is an RCC1 family member with an essential role in prometaphase to metaphase progression. *Developmental Cell*, 5 (2): 295–307. doi:10.1016/S1534-5807(03)00205-3.

Moore, M.J. and Proudfoot, N.J. (2009) Pre-mRNA Processing Reaches Back toTranscription and Ahead to Translation. *Cell*, 136 (4): 688–700. doi:10.1016/j.cell.2009.02.001.

Moore, W.J., Zhang, C. and Clarke, P.R. (2002) Targeting of RCC1 to chromosomes is required for proper mitotic spindle assembly in human cells. *Current Biology*, 12 (16): 1442–1447. doi:10.1016/S0960-9822(02)01076-X.

Moyer, S.E., Lewis, P.W. and Botchan, M.R. (2006) Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proceedings of the National Academy of Sciences*, 103 (27): 10236–10241. doi:10.1073/pnas.0602400103.

Nalepa, G. and Clapp, D.W. (2018) Fanconi anaemia and cancer: An intricate relationship. *Nature Reviews Cancer*, 18 (3): 168–185. doi:10.1038/nrc.2017.116.

Neelsen, K.J. and Lopes, M. (2015) Replication fork reversal in eukaryotes: From dead end to dynamic response. *Nature Reviews Molecular Cell Biology*, 16 (4): 207–220. doi:10.1038/nrm3935.

Ng, N., Purshouse, K., Foskolou, I.P., et al. (2018) Challenges to DNA replication in hypoxic conditions. *FEBS Journal*, 285 (9): 1563–1571. doi:10.1111/febs.14377.

Nieminuszczy, J., Schwab, R.A. and Niedzwiedz, W. (2016) The DNA fibre technique – tracking helicases at work. *Methods*, 108 (2016): 92–98. doi:10.1016/j.ymeth.2016.04.019.

Noma, A., Ishitani, R., Kato, M., et al. (2010) Expanding role of the jumonji C domain as an RNA hydroxylase. *Journal of Biological Chemistry*, 285 (45): 34503–34507. doi:10.1074/jbc.M110.156398.

O'Driscoll, M., Ruiz-Perez, V.L., Woods, C.G., et al. (2003) A splicing mutation affecting expression of ataxia-telangiectasia and Rad3-related protein (ATR) results in Seckel syndrome. *Nature Genetics*, 33 (4): 497–501. doi:10.1038/ng1129.

Oh, S. and Janknecht, R. (2012) Histone demethylase JMJD5 is essential for embryonic development. *Biochemical and Biophysical Research Communications*, 420 (1): 61–65. doi:10.1016/j.bbrc.2012.02.115.

Oh, S., Shin, S. and Janknecht, R. (2019) The small members of the JMJD protein family: Enzymatic jewels or jinxes? *Biochimica et Biophysica Acta - Reviews on Cancer*, 1871 (2): 406–418. doi:10.1016/j.bbcan.2019.04.002.

Ohtsubo, M., Kai, R., Furuno, N., et al. (1987) Isolation and characterization of the active cDNA of the human cell cycle gene (RCC1) involved in the regulation of onset of chromosome condensation. *Genes & development*, 1 (6): 585–593. doi:10.1101/gad.1.6.585.

Pannunzio, N.R., Watanabe, G. and Lieber, M.R. (2018) Nonhomologous DNA end-joining for repair of DNA double-strand breaks. *Journal of Biological Chemistry*, 293 (27): 10512–10523. doi:10.1074/jbc.TM117.000374.

Patel, J., Landers, K., Mortimer, R.H., et al. (2010) Regulation of Hypoxia Inducible Factors (HIF) in Hypoxia and Normoxia during Placental Development. *Placenta*, 31 (11): 951–957. doi:10.1016/j.placenta.2010.08.008.

Petermann, E. and Helleday, T. (2010) Pathways of mammalian replication fork restart. *Nature Reviews Molecular Cell Biology*, 11 (10): 683–687. doi:10.1038/nrm2974.

Petermann, E., Orta, M.L., Issaeva, N., et al. (2010) Hydroxyurea-Stalled Replication Forks Become Progressively Inactivated and Require Two Different RAD51-Mediated Pathways for Restart and Repair. *Molecular Cell*, 37 (4): 492–502. doi:10.1016/j.molcel.2010.01.021.

Pettersen, E.F., Goddard, T.D., Huang, C.C., et al. (2021) UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science*, 30 (1): 70–82. doi:10.1002/pro.3943.

Pires, I.M., Bencokova, Z., Milani, M., et al. (2010) Effects of acute versus chronic hypoxia on DNA damage responses and genomic instability. *Cancer Research*, 70 (3): 925–935. doi:10.1158/0008-5472.CAN-09-2715.

Ploumakis, A. and Coleman, M.L. (2015) OH, the Places You'll Go! Hydroxylation, Gene Expression, and Cancer. *Molecular Cell*, 58 (5): 729–741. doi:10.1016/j.molcel.2015.05.026.

Prakash, R., Zhang, Y., Feng, W., et al. (2015) Homologous recombination and human health: The roles of BRCA1, BRCA2, and associated proteins. *Cold Spring Harbor Perspectives in Biology*, 7 (4): 1–27. doi:10.1101/cshperspect.a016600.

Ratcliffe, P.J. (2013) Oxygen sensing and hypoxia signalling pathways in animals: The implications of physiology for cancer. *Journal of Physiology*, 591 (8): 2027–2042. doi:10.1113/jphysiol.2013.251470.

Ray Chaudhuri, A., Hashimoto, Y., Herrador, R., et al. (2012) Topoisomerase i poisoning results in PARP-mediated replication fork reversal. *Nature Structural and Molecular Biology*, 19 (4): 417–423. doi:10.1038/nsmb.2258.

Renault, L., Nassar, N., Vetter, I., et al. (1998) The 1.7 Å crystal structure of the regulator of chromosome condensation (RCC1) reveals a seven-bladed propeller. *Nature*, 392 (6671): 97–101. doi:10.1038/32204.

Reynolds, J.J., Bicknell, L.S., Carroll, P., et al. (2017) Mutations in DONSON disrupt replication fork stability and cause microcephalic dwarfism. *Nature Genetics*, 49 (4): 537–549. doi:10.1038/ng.3790.

Ribet, D. and Cossart, P. (2010) Post-translational modifications in host cells during bacterial infection. *FEBS Letters*, 584 (13): 2748–2758. doi:10.1016/j.febslet.2010.05.012.

Del Rizzo, P.A., Krishnan, S. and Trievel, R.C. (2012) Crystal Structure and Functional Analysis of JMJD5 Indicate an Alternate Specificity and Function. *Molecular and Cellular Biology*, 32 (19): 4044–4052. doi:10.1128/MCB.00513-12.

Roux, K.J., Kim, D.I., Burke, B., et al. (2018) BioID: A Screen for Protein-Protein Interactions. *Current Protocols in Protein Science*, 91 (1): 19.23.1-19.23.15. doi:10.1002/cpps.51.

Saeed, K., Östling, P., Björkman, M., et al. (2015) Androgen receptor-interacting protein HSPBAP1 facilitates growth of prostate cancer cells in androgen-deficient conditions. *International Journal of Cancer*, 136 (11): 2535–2545. doi:10.1002/ijc.29303.

Sakofsky, C.J. and Malkova, A. (2017) Break induced replication in eukaryotes: mechanisms, functions, and consequences. *Critical Reviews in Biochemistry and Molecular Biology*, 52 (4): 395–413. doi:10.1080/10409238.2017.1314444.

Sanchez-Bailon, M.P., Choi, S.Y., Dufficy, E.R., et al. (2021) Arginine methylation and ubiquitylation crosstalk controls DNA end-resection and homologous recombination repair. *Nature Communications*, 12 (1): 1–18. doi:10.1038/s41467-021-26413-6.

Schmidt, C., Grønborg, M., Deckert, J., et al. (2014) Mass spectrometry-based relative quantification of proteins in precatalytic and catalytically active spliceosomes by metabolic labeling (SILAC), chemical labeling (iTRAQ), and label-free spectral count. *Rna*, 20 (3): 406–420. doi:10.1261/rna.041244.113.

Schödel, J., Mole, D.R. and Ratcliffe, P.J. (2013) Pan-genomic binding of hypoxia-inducible transcription factors. *bchm*, 394 (4): 507–517. doi:10.1515/hsz-2012-0351.

Selak, M.A., Armour, S.M., MacKenzie, E.D., et al. (2005) Succinate links TCA cycle dysfunction to oncogenesis by inhibiting HIF-α prolyl hydroxylase. *Cancer Cell*, 7 (1): 77–85. doi:10.1016/j.ccr.2004.11.022.

Shalaby, N.A., Pinzon, J.H., Narayanan, A.S., et al. (2018) JmjC domain proteins modulate circadian behaviors and sleep in Drosophila. *Scientific Reports*, 8 (1): 1–9. doi:10.1038/s41598-017-18989-1.

Shen, C. and Kaelin, W.G. (2013) The VHL/HIF axis in clear cell renal carcinoma. *Seminars in Cancer Biology*, 23 (1): 18–25. doi:10.1016/j.semcancer.2012.06.001.

Shen, J., Xiang, X., Chen, L., et al. (2017) JMJD5 cleaves monomethylated histone H3 N-tail under DNA damaging stress. *EMBO reports*, 18 (12): 2131–2143. doi:10.15252/embr.201743892.

Short, J.M., Liu, Y., Chen, S., et al. (2016) High-resolution structure of the presynaptic RAD51 filament on single-stranded DNA by electron cryo-microscopy. *Nucleic Acids Research*, 44 (19): 9017–9030. doi:10.1093/nar/gkw783.

Silvera, D., Formenti, S.C. and Schneider, R.J. (2010) Translational control in cancer. *Nature Reviews Cancer*. 10 (4) pp. 254–266. doi:10.1038/nrc2824.

Singleton, R.S., Liu-Yi, P., Formenti, F., et al. (2014) OGFOD1 catalyzes prolyl hydroxylation of RPS23 and is involved in translation control and stress granule formation. *Proceedings of the National Academy of Sciences of the United States of America*, 111 (11): 4031–4036. doi:10.1073/pnas.1314482111.

Sinha, K.M., Yasuda, H., Coombes, M.M., et al. (2010) Regulation of the osteoblast-specific transcription factor Osterix by NO66, a Jumonji family histone demethylase. *EMBO Journal*, 29 (1): 68–79. doi:10.1038/emboj.2009.332.

Sjölinder, M., Uhlmann, J. and Ponstingl, H. (2002) DelGEF, a homologue of the Ran guanine nucleotide exchange factor RanGEF, binds to the exocyst component Sec5 and modulates secretion. *FEBS Letters*, 532 (1–2): 211–215. doi:10.1016/S0014-5793(02)03677-3.

Sjölinder, M., Uhlmann, J. and Ponstingl, H. (2004) Characterisation of an evolutionary conserved protein interacting with the putative guanine nucleotide exchange factor DelGEF and modulating secretion. *Experimental Cell Research*, 294 (1): 68–76. doi:10.1016/j.yexcr.2003.09.033.

Slabinski, L., Jaroszewski, L., Rychlewski, L., et al. (2007) *Structural bioinformatics XtalPred : a web server for prediction of protein crystallizability*., 23 (24): 3403–3405. doi:10.1093/bioinformatics/btm477.

Smits, V.A.J., Reaper, P.M. and Jackson, S.P. (2006) Rapid PIKK-dependent release of Chk1 from chromatin promotes the DNA-damage checkpoint response. *Current Biology*, 16 (2): 150–159. doi:10.1016/j.cub.2005.11.066.

Song, C., Liang, L., Jin, Y., et al. (2018) RCC2 is a novel p53 target in suppressing metastasis. *Oncogene*, 37 (1): 8–17. doi:10.1038/onc.2017.306.

Song, J., Zheng, J., Liu, X., et al. (2022) A novel protein encoded by ZCRB1 - induced circHEATR5B suppresses aerobic glycolysis of GBM through phosphorylation of JMJD5. *Journal of Experimental & Clinical Cancer Research*, pp. 1–20. doi:10.1186/s13046-022-02374-6.

Speck, C. and Stillman, B. (2007) Cdc6 ATPase activity regulates ORC·Cdc6 stability and the selection of specific DNA sequences as origins of DNA replication. *Journal of Biological Chemistry*, 282 (16): 11705–11714. doi:10.1074/jbc.M700399200.

Stenflo, J., Holme, E., Lindstedt, S., et al. (1989) Hydroxylation of aspartic acid in domains homologous to the epidermal growth factor precursor is catalyzed by a 2-oxoglutarate-dependent dioxygenase. *Proceedings of the National Academy of Sciences of the United States of America*, 86 (2): 444–447. doi:10.1073/pnas.86.2.444.

Stolerman, E.S., Francisco, E., Stallworth, J.L., et al. (2019) Genetic variants in the KDM6B gene are associated with neurodevelopmental delays and dysmorphic features. *American Journal of Medical Genetics, Part A*, 179 (7): 1276–1286. doi:10.1002/ajmg.a.61173.

Sulkowski, P.L., Corso, C.D., Robinson, N.D., et al. (2017) 2-Hydroxyglutarate produced by neomorphic IDH mutations suppresses homologous recombination and induces PARP inhibitor sensitivity. *Science Translational Medicine*, 9 (375). doi:10.1126/scitranslmed.aal2463.

Sulkowski, P.L., Sundaram, R.K., Oeck, S., et al. (2018) Krebs-cycle-deficient hereditary cancer syndromes are defined by defects in homologous-recombination DNA repair. *Nature Genetics*, 50 (8): 1086–1092. doi:10.1038/s41588-018-0170-4.

Suwaki, N., Klare, K. and Tarsounas, M. (2011) RAD51 paralogs: Roles in DNA damage signalling, recombinational repair and tumorigenesis. *Seminars in Cell and Developmental Biology*, 22 (8): 898–905. doi:10.1016/j.semcdb.2011.07.019.

Suzuki, T. (2021) The expanding world of tRNA modifications and their disease relevance. *Nature Reviews Molecular Cell Biology*. 22 (6) pp. 375–392. doi:10.1038/s41580-021-00342-0.

Suzuki, T., Minehata, K.I., Akagi, K., et al. (2006) Tumor suppressor gene identification using retroviral insertional mutagenesis in Blm-deficient mice. *EMBO Journal*, 25 (14): 3422–3431. doi:10.1038/sj.emboj.7601215.

Symington, L.S. and Gautier, J. (2011) Double-strand break end resection and repair pathway choice. *Annual Review of Genetics*, 45: 247–271. doi:10.1146/annurev-genet-110410-132435.

Szklarczyk, D., Gable, A.L., Lyon, D., et al. (2019) STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, 47 (D1): D607–D613. doi:10.1093/nar/gky1131.

Tahiliani, M., Koh, K.P., Shen, Y., et al. (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, 324 (5929): 930–935. doi:10.1126/science.1170116.

Takeuchi, T., Yamazaki, Y., Katoh-Fukui, Y., et al. (1995) Gene trap capture of a novel mouse gene, jumonji, required for neural tube formation. *Genes and Development*, 9 (10): 1211–1222. doi:10.1101/gad.9.10.1211.

Tan, S.C., Low, T.Y., Hussain, H.M.J., et al. (2022) Association between XRCC3 p.Thr241Met polymorphism and risk of glioma: A systematic review and meta-analysis. *PLoS ONE*, 17 (10 October). doi:10.1371/journal.pone.0276313.

Tan, X. ping, Dong, W. guo, Zhang, Q., et al. (2014) Potential Effects of Mina53 on Tumor Growth in Human Pancreatic Cancer. *Cell Biochemistry and Biophysics*, 69 (3): 619–625. doi:10.1007/s12013-014-9841-7.

Tang, Y., Feng, M., Su, Y., et al. (2023) Jmjd4 Facilitates Pkm2 Degradation in Cardiomyocytes and Is Protective Against Dilated Cardiomyopathy. *Circulation*, pp. 1–21. doi:10.1161/CIRCULATIONAHA.123.064121.

Thys, R.G. and Wang, Y.H. (2015) DNA replication dynamics of the GGGGCC repeat of the C9orf72 gene. *Journal of Biological Chemistry*, 290 (48): 28953–28962. doi:10.1074/jbc.M115.660324.

Toledo, L.I., Altmeyer, M., Rask, M.B., et al. (2013) XATR prohibits replication catastrophe by preventing global exhaustion of RPA. *Cell*, 155 (5): 1088. doi:10.1016/j.cell.2013.10.043.

Tombline, G. and Fishel, R. (2002) Biochemical characterization of the human RAD51 protein. I. ATP hydrolysis. *Journal of Biological Chemistry*, 277 (17): 14417–14425. doi:10.1074/jbc.M109915200.

Traven, A. and Heierhorst, J. (2005) SQ/TQ cluster domains: Concentrated ATM/ATR kinase phosphorylation site regions in DNA-damage-response proteins. *BioEssays*, 27 (4): 397–407. doi:10.1002/bies.20204.

Unoki, M., Masuda, A., Dohmae, N., et al. (2013) Lysyl 5-hydroxylation, a novel histone modification, by jumonji domain containing 6 (JMJD6). *Journal of Biological Chemistry*, 288 (9): 6053–6062. doi:10.1074/jbc.M112.433284.

Vesela, E., Chroma, K., Turi, Z., et al. (2017) Common chemical inductors of replication stress: Focus on cell-based studies. *Biomolecules*, 7 (1). doi:10.3390/biom7010019.

Vetro, A., Savasta, S., Russo Raucci, A., et al. (2017) MCM5: A new actor in the link between DNA replication and Meier-Gorlin syndrome. *European Journal of Human Genetics*, 25 (5): 646–650. doi:10.1038/ejhg.2017.5.

Wang, E., Zhang, C., Polavaram, N., et al. (2014a) The role of Factor Inhibiting HIF (FIH-1) in inhibiting HIF-1 transcriptional activity in glioblastoma multiforme. *PLoS ONE*, 9 (1): 1–9. doi:10.1371/journal.pone.0086102.

Wang, H., Shi, L.Z., Wong, C.C.L., et al. (2013a) The Interaction of CtIP and Nbs1 Connects CDK and ATM to Regulate HR-Mediated Double-Strand Break Repair. *PLoS Genetics*, 9 (2): 25–27. doi:10.1371/journal.pgen.1003277.

Wang, H., Wang, J., Liu, J., et al. (2022) Jumonji-C domain-containing protein 5 suppresses proliferation and aerobic glycolysis in pancreatic cancer cells in a c-Myc-dependent manner. *Cellular Signalling*, 93 (February): 110282. doi:10.1016/j.cellsig.2022.110282.

Wang, H., Zhou, X., Wu, M., et al. (2013b) Structure of the JmjC-domain-containing protein JMJD5. *Acta Crystallographica Section D*, 69 (10): 1911–1920. doi:10.1107/S0907444913016600.

Wang, H.J., Pochampalli, M., Wang, L.Y., et al. (2019a) KDM8/JMJD5 as a dual coactivator of AR and PKM2 integrates AR/EZH2 network and tumor metabolism in CRPC. *Oncogene*, 38 (1): 17–32. doi:10.1038/s41388-018-0414-x.

Wang, J., Xie, H., Ling, Q., et al. (2016) Coding-noncoding gene expression in intrahepatic cholangiocarcinoma. *Translational Research*, 168: 107–121. doi:10.1016/j.trsl.2015.07.007.

Wang, L., Jiang, F., Ma, F., et al. (2019b) MiR-873-5p suppresses cell proliferation and epithelial–mesenchymal transition via directly targeting Jumonji domain-containing protein 8 through the NF-κB pathway in colorectal cancer. *Journal of Cell Communication and Signaling*, 13 (4): 549–560. doi:10.1007/s12079-019-00522-w.

Wang, X., Lu, G., Li, L., et al. (2014b) HUWE1 interacts with BRCA1 and promotes its degradation in the ubiquitin-proteasome pathway. *Biochemical and Biophysical Research Communications*, 444 (4): 549–554. doi:10.1016/j.bbrc.2014.01.075.

Wang, Z., Wang, C., Huang, X., et al. (2012) Differential proteome profiling of pleural effusions from lung cancer and benign inflammatory disease patients. *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1824 (4): 692–700. doi:10.1016/j.bbapap.2012.01.016.

Ward, I.M. and Chen, J. (2001) Histone H2AX Is Phosphorylated in an ATR-dependent Manner in Response to Replicational Stress. *Journal of Biological Chemistry*, 276 (51): 47759–47762. doi:10.1074/jbc.C100569200.

Ward, I.M., Minn, K., van Deursen, J., et al. (2003) p53 Binding Protein 53BP1 Is Required for DNA Damage Responses and Tumor Suppression in Mice. *Molecular and Cellular Biology*, 23 (7): 2556–2563. doi:10.1128/mcb.23.7.2556-2563.2003.

Wätzlich, D., Vetter, I., Gotthardt, K., et al. (2013) The interplay between RPGR, PDEδ and Arl2/3 regulate the ciliary targeting of farnesylated cargo. *EMBO Reports*, 14 (5): 465–472. doi:10.1038/embor.2013.37.

Webb, J.D., Murányi, A., Pugh, C.W., et al. (2009) MYPT1, the targeting subunit of smooth-muscle myosin phosphatase, is a substrate for the asparaginyl hydroxylase factor inhibiting hypoxia-inducible factor (FIH). *Biochemical Journal*, 420 (2): 327–333. doi:10.1042/BJ20081905.

Webby, C.J., Wolf, A., Gromak, N., et al. (2009a) *Jmjd6 Catalyses Lysyl-Hydroxylation of U2AF65, a Protein Associated with RNA Splicing.*, 325 (JULY): 90–94.

Webby, C.J., Wolf, A., Gromak, N., et al. (2009b) Jmjd6 catalyses lysyl-hydroxylation of U2AF65, a protein associated with RNA splicing. *Science*, 325 (5936): 90–93. doi:10.1126/science.1175865.

Westrip, C.A.E., Paul, F., Al-Murshedi, F., et al. (2023) Inactivation of DRG1, encoding a translation factor GTPase, causes a Recessive Neurodevelopmental Disorder. *Genetics in Medicine*, p. 100893. doi:10.1016/j.gim.2023.100893.

Westrip, C.A.E., Zhuang, Q., Hall, C., et al. (2021) Developmentally regulated GTPases: structure, function and roles in disease. *Cellular and Molecular Life Sciences*, 78 (23): 7219–7235. doi:10.1007/s00018-021-03961-0.

Wilhelm, T., Olziersky, A.M., Harry, D., et al. (2019) Mild replication stress causes chromosome mis-segregation via premature centriole disengagement. *Nature Communications*, 10 (1): 1–14. doi:10.1038/s41467-019-11584-0.

Wilhelm, T., Said, M. and Naim, V. (2020) Dna replication stress and chromosomal instability: Dangerous liaisons. *Genes*, 11 (6): 1–35. doi:10.3390/genes11060642.

Wilkins, S.E., Islam, S., Gannon, J.M., et al. (2018) JMJD5 is a human arginyl C-3 hydroxylase. *Nature Communications*, 9 (1). doi:10.1038/s41467-018-03410-w.

Woodward, A.M., Göhler, T., Luciani, M.G., et al. (2006) Excess Mcm2-7 license dormant origins of replication that can be used under conditions of replicative stress. *Journal of Cell Biology*, 173 (5): 673–683. doi:10.1083/jcb.200602108.

Wright, W.D., Shah, S.S. and Heyer, W.D. (2018) Homologous recombination and the repair of DNA double-strand breaks. *Journal of Biological Chemistry*, 293 (27): 10524–10535. doi:10.1074/jbc.TM118.000372.

Wu, B.-H., Chen, H., Cai, C.-M., et al. (2016) Epigenetic silencing of JMJD5 promotes the proliferation of hepatocellular carcinoma cells by down-regulating the transcription of CDKN1A 686. *Oncotarget*, 7 (6): 6847–63. doi:10.18632/oncotarget.6867.

Wu, J., He, Z., Yang, X.M., et al. (2017) RCCD1 depletion attenuates TGF-β-induced EMT and cell migration by stabilizing cytoskeletal microtubules in NSCLC cells. *Cancer Letters*, 400: 18–29. doi:10.1016/j.canlet.2017.04.021.

Xiang, Y., Zhu, Z., Han, G., et al. (2007) JARID1B is a histone H3 lysine 4 demethylase up-regulated in prostate cancer. *Proceedings of the National Academy of Sciences of the United States of America*, 104 (49): 19226–19231. doi:10.1073/pnas.0700735104.

Xing, J., Wang, K., Liu, P.W., et al. (2014) Mina53, a novel molecular marker for the diagnosis and prognosis of gastric adenocarcinoma. *Oncology Reports*, 31 (2): 634–640. doi:10.3892/or.2013.2918.

Xu, W., Yang, H., Liu, Y., et al. (2011) Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. *Cancer Cell*, 19 (1): 17–30. doi:10.1016/j.ccr.2010.12.014.

Yang, M., Chowdhury, R., Ge, W., et al. (2011a) Factor-inhibiting hypoxia-inducible factor (FIH) catalyses the post-translational hydroxylation of histidinyl residues within ankyrin repeat domains. *FEBS Journal*, 278 (7): 1086–1097. doi:10.1111/j.1742-4658.2011.08022.x.

Yang, M., Ge, W., Chowdhury, R., et al. (2011b) Asparagine and aspartate hydroxylation of the cytoskeletal ankyrin family is catalyzed by factor-inhibiting hypoxia-inducible factor. *Journal of Biological Chemistry*, 286 (9): 7648–7660. doi:10.1074/jbc.M110.193540.

Yang, Z., Zhuang, L., Szatmary, P., et al. (2015) Upregulation of heat shock proteins (HSPA12A, HSP90B1, HSPA4, HSPA5 and HSPA6) in tumour tissues is associated with poor outcomes from HBV-related early-stage hepatocellular carcinoma. *International Journal of Medical Sciences*, 12 (3): 256–263. doi:10.7150/ijms.10735.

Yao, Y., Zhou, W.Y. and He, R.X. (2019) Down-regulation of JMJD5 suppresses metastasis and induces apoptosis in oral squamous cell carcinoma by regulating p53/NF-κB pathway. *Biomedicine and Pharmacotherapy*, 109 (May 2018): 1994–2004. doi:10.1016/j.biopha.2018.07.144.

Yeo, K.S., Tan, M.C., Lim, Y.Y., et al. (2017) JMJD8 is a novel endoplasmic reticulum protein with a JmjC domain. *Scientific Reports*, 7 (1): 1–10. doi:10.1038/s41598-017-15676-z.

Yeo, K.S., Tan, M.C., Wong, W.Y., et al. (2016) JMJD8 is a positive regulator of TNF-induced NF-κB signaling. *Scientific Reports*, 6 (February): 1–10. doi:10.1038/srep34125.

Yi, J., Shen, H.F., Qiu, J.S., et al. (2017) JMJD6 and U2AF65 co-regulate alternative splicing in both JMJD6 enzymatic activity dependent and independent manner. *Nucleic Acids Research*, 45 (6): 3503–3518. doi:10.1093/nar/gkw1144.

Youn, M.Y., Yokoyama, A., Fujiyama-Nakamura, S., et al. (2012) JMJD5, a Jumonji C (JmjC) domain-containing protein, negatively regulates osteoclastogenesis by facilitating NFATc1 protein degradation. *Journal of Biological Chemistry*, 287 (16): 12994–13004. doi:10.1074/jbc.M111.323105.

Zaghet, N., Madsen, K., Rossi, F., et al. (2021) Coordinated maintenance of H3K36/K27 methylation by histone demethylases preserves germ cell identity and immortality. *Cell Reports*, 37 (8): 110050. doi:10.1016/j.celrep.2021.110050.

Zellweger, R., Dalcher, D., Mutreja, K., et al. (2015) Rad51-mediated replication fork reversal is a global response to genotoxic treatments in human cells. *Journal of Cell Biology*, 208 (5): 563–579. doi:10.1083/jcb.201406099.

Zeman, M.K. and Cimprich, K.A. (2014) Causes and consequences of replication stress. *Nature Cell Biology*, 16 (1): 2–9. doi:10.1038/ncb2897.

Zhang, R., Huang, Q., Li, Y., et al. (2015) JMJD5 is a potential oncogene for colon carcinogenesis. *International Journal of Clinical and Experimental Pathology*, 8 (6): 6482–6489.

Zhao, J., Du, F., Shen, G., et al. (2015a) The role of hypoxia-inducible factor-2 in digestive system cancers. *Cell Death and Disease*, 6 (1): 1–9. doi:10.1038/cddis.2014.565.

Zhao, Z., Sun, C., Li, F., et al. (2015b) Overexpression of histone demethylase JMJD5 promotes metastasis and indicates a poor prognosis in breast cancer. *International Journal of Clinical and Experimental Pathology*, 8 (9): 10325–10334.

Zheng, G., Dahl, J.A., Niu, Y., et al. (2013) ALKBH5 Is a Mammalian RNA Demethylase that Impacts RNA Metabolism and Mouse Fertility. *Molecular Cell*, 49 (1): 18–29. doi:10.1016/j.molcel.2012.10.015.

Zhong, J., Jermusyk, A., Wu, L., et al. (2020) A Transcriptome-Wide Association Study (TWAS) Identifies Novel Candidate Susceptibility Genes for Pancreatic Cancer. *JNCI: Journal of the National Cancer Institute*, 0: 1–10. doi:10.1093/jnci/djz246.

Zhu, H., Hu, S. and Baker, J. (2014) JMJD5 Regulates Cell Cycle and Pluripotency in Human Embryonic Stem Cells. *Stem Cells*, pp. 533–542. doi:10.1634/stemcells.2008-0596.

Zou, Y., Liu, Y., Wu, X., et al. (2006) Functions of human replication protein A (RPA): From DNA replication to DNA damage and stress responses. *Journal of Cellular Physiology*, 208 (2): 267–273. doi:10.1002/jcp.20622.