# Identification of inhibitors for a family of human oxoglutarate-dependent oxygenases


Radosław Piotr Nowak

Lady Margaret Hall

University of Oxford

Degree of Doctor of Philosophy

Supervisors: Professor Udo Oppermann, Dr Catrine Johansson, Dr Anthony Tumber (University of Oxford)

Dr Dafydd Owen (Pfizer)

Identification of inhibitors for a family of human oxoglutarate-dependent oxygenases
Radoslaw Piotr Nowak, Lady Margaret Hall
Submitted for the degree of DPhil. Systems Approaches in Biomedical Science

## Abstract

The family of human iron- and 2-oxoglutarate - dependent oxygenases encompasses over 60 members and a subclass of the enzymes, containing Jumonji C (JmjC) domain, is directly involved in epigenetic control of gene transcription, through demethylation of methyl marks present on histone tails. Recently, JmjC-type enzymes, namely MINA53, NO66 and OGFOD, were identified as ribosomal protein hydroxylases (ROXs), broadening the known substrate set and indicating a function in the control of mRNA translation. Oxygenases are frequently implicated as key regulatory factors in various types of cancers and targeted inhibition of this class of enzymes can lead to suppressed cancer cell proliferation, opening an epigenetic avenue for therapeutic intervention by small molecules.

In order to identify novel chemical starting points for JmjC- type oxygenases we have profiled representative members of the family with two libraries of fragments by activity assays and used X-ray crystallography to explore the binding mode of identified inhibitors. The obtained X-ray crystal structure complexes with JARID1B demethylase represent 7 novel chemotypes exploring different regions of the catalytic binding site as well as a previously unidentified putative allosteric pocket. We have then focussed on the ribosomal hydroxylases NO66 and MINA53, and identified novel chemical lead compounds through a mass spectrometry based enzymatic activity assay and obtained X-ray crystal structure complexes with JARID1B demethylase of two active inhibitors of MINA53 and JARID1B (IC50 of 0.9-5 μM). The most active hydroxamic acid series has an $IC_{50}$ of 900 nM towards JARID1B with a favourable selectivity profile and shows anti-proliferative effects in human myeloma cell lines.

We anticipate that our comprehensive set of fragment library activity data together with X-ray crystal structure complexes identified in this thesis can be used as starting points for further structure-based inhibitor development programs for JmjC- type oxygenases.

## Acknowledgements

I am especially grateful to my supervisors for providing continued support and guidance throughout this DPhil: Prof Udo Oppermann, who trusted an engineer in a biochemistry laboratory and taught me the meaning of the word ‘detail’; Dr Catrine Johansson, who, despite my lack of laboratory experience prior to this DPhil, was confident in me and spent countless hours explaining the details of protein production and crystallisation; Dr Anthony Tumber, for his immense patience in teaching me all aspects of enzymology; Dr Dafydd Owen for the guidance in medicinal chemistry.

I would like to thank Prof Frank von Delft, Dr Tobias Krojer, Dr Melanie Vollmar, Dr Srikannathasan Velupillai and Dr Joseph Newman, for introducing me to the field of X-ray crystallography and for providing excellent company during the late night synchrotron shifts; Dr Romain Talon and Dr Patrick Collins, for introduction to the X-ray fragment soaking methods at Lab36, Diamond Synchrotron; Prof Christopher Schofield, for providing oxygenase inhibitors; Prof Antonello Mai and Dr Dante Rotili, for fruitful collaboration; Dr Stanley Ng, who worked previously on the MINA53/NO66 project and his practical advice; The DTC members and fellow DPhil students for stimulating discussions and all staff at the SGC who supported every part of my work.

Special thanks goes to the EPSRC Research Council and Pfizer for the generous funding for this project.

Finally, I would like to thank my family, and especially my parents, my brother and my wife, Klaudia for huge support in the pursuit of my dreams.

# Table of Contents

# Table of Figures

# Table of Tables

# List of abbreviations

***Institutions***

SGC – *Structural Genomics Consortium*

PDB – *Protein Data Bank*

FDA – *Food and Drug Administration*

Diamond/Diamond synchrotron/Dmnd – *Diamond Light Source, Harwell, UK.*

***Techniques and methods***

DSF – *Differential Scanning Fluorimetry*

BLI – *BioLayer Interferometry*

SPR – *Surface Plasmon Resonance*

ITC – *Isothermal Titration Calorimetry*

RF MS – *RapidFire mass spectrometry*

MMS – *Matrix microseeding*

IMAC – *immobilised affinity chromatography*

***Peptides, domains, proteins and protein families***

H3K4me2 – *Histone 3 di-methylated lysine 4; other modifications follow this style*

JmjC – *Jumonji C domain*

ROX – *Ribosomal hydroxylases*

KAT – *Histone lysine acetyltransferases*

HAT – *Histone acetyltransferase*

KDM – *K- demethylases (lysine demethylases)*

HDAC – *Histone deacethylases*

PRMT – *Protein arginine methyltransferases*

DNMT – *DNA methyltransferase*

OGT – *O-GlcNAc transferase*

OGA – *O-GlcNAcase hydrolase*

PARP – *Poly (ADP-ribose) polymerases*

PAD – *Protein arginine deiminase*

SUMO – *Small ubiquitin like modifier*

SET – *Methyltransferase SET domain*

MBD – *Methyl-CpG-binding domain*

PHD – *Plant homeodomain*

MBT – *Malignant brain tumor domain*

PWWP – *Proline-tryptophan-tryptophan-proline domain*

BRCT – *BRCA1 C terminus domain*

UIM – *Ubiquitin interaction motif*

IUIM – *Inverted ubiquitin interaction motif*

SIM – *Sumo interaction motif*

PBZ – *Poly ADP-ribose binding zinc finger*

ARD – *Ankyrin repeat domain*

MINA53 – *Myc-induced nuclear antigen with mass 53kDa*

NO66 – *Nucleolar protein 66*

Osx – *Osterix transcription factor*

DOT1L – *DOT1-Like Histone H3K79 Methyltransferase*

EZH2 – *Enhancer of Zeste Homolog 2*

PRC1 – *Polycomb repressive complex 1*

PRC2 – *Polycomb repressive complex 2*

LSD demethylase – *Lysine-specific demethylase*

LSD1 – *Lysine-specific demethylase 1*

TSGA – *Testis specific gene A*

SCAI – *Suppressor of cancer cell invasion*

E2F1 – *E2F Transcription Factor 1*

FGF4 – *Fibroblast growth factor 4*

pVHL – *von Hippel–Lindau tumour suppressor*

HR – *Hairless*

AR – *Androgen receptor*

TRPV3 – *transient receptor potential vanilloid 3 ion channel*

TEV – *Tobacco Etch Virus protease*

***Chemical compounds***

NAD – *Nicotinamide adenine dinucleotide*

FAD – *Flavin adenine dinucleotide*

2,4-PDCA – *2,4-pyridinedicarboxylic acid*

NOG – *N-oxalylglycine*

LAA – *L-ascorbic acid*

FAS – *Ferrous ammonium sulphate*

2OG – *2-oxoglutarate (α-ketoglutarate)*

FA – *Formic acid*

DMSO – *Dimethyl sulfoxide*

TCEP – *tris(2-carboxyethyl)phosphine*

***Other***

DSBH – *Double stranded beta helix*

TCA cycle – *Tricarboxylic acid cycle*

DDR pathway – *DNA-damage response pathway*

AML – *Acute myeloid leukaemia*

MSC – *Mesenchymal stem cells*

TSSG – *Transient site-specific copy gains*

Hox genes – *homeotic genes*

XLID – *X-linked intellectual disability*

MEF – *Mouse embryonic fibroblasts*

T-ALL – *T-cell acute lymphoblastic leukaemia*

JCSG – *Joint Center for Structural Genomics sparse matrix screen*

LFS – *PEG/ION/pH (PACT) equivalent sparse matrix screen*

MIDAS – *Modern Intelligent Dynamic Alternative Screen*

HIN – *Hampton Index sparse matrix screen*

HCS – *Hampton Research Crystal Screen*

TB – *terrific broth*

PVDV – *Polyvinylidene Fluoride filter*

BSA – *Bovine serum albumin*

FRER – *Rigaku 'home' rotating anode X-ray source*

## Work performed in collaboration

*Cloning, Insect Cell expression:*

All cloning and insect cell expression steps in this thesis were done by the Biotechnology Group at the SGC.

*Crystallisation mutants:*

MINA53 and NO66 mutants designed to generate crystal contacts were picked by R. Nowak and cloned by Dr C. Strain-Damerell, expressed, purified and crystallised by Dr M. Fairhead. Crystals were mounted by R. Nowak and Dr M. Fairhead, and datasets were collected by R. Nowak.

*JARID1B crystallisation:*

Design of construct, purification and initial crystallisation of JARID1B were performed by Dr Catrine Johansson.

Compound soaking and data collection JARID1B - GSK-J1 complex was performed by R. Nowak and the structure was solved and modelled by Dr S. Velupillai.

*In vitro activity assays and DSF of FBXL11:*

RapidFire based *in vitro* activity assay with FBXL11 and the corresponding DSF data described in 4.1.1 was performed by Dr A. Tumber.

*Cellular assays:*

The cellular proliferation assay with human multiple myeloma cell lines was performed by N. Wu. The demethylation assay was performed and analysed by Dr C. Yapp and R. Nowak.

# 1. Introduction

## 1.1. Chromatin and epigenetics

The definition of “epigenetics” is still debated in the scientific community. The term was first used by Conrad Waddington to describe heritable changes in cellular phenotypes independent of alterations in the underlying DNA sequence [1]. DNA is packaged into cells in an ordered fashion, forming chromatin, a macromolecular complex of DNA with structural proteins called histones. The basic building block of chromatin is the nucleosome. It is a complex of 147 base pairs of DNA wrapped around a histone octamer containing two of each histones H2A, H2B, H3, H4 and linked by a single histone H1 [1]. Chromatin can be present in two distinct forms: heterochromatin, a highly condensed and transcriptionally “inactive” form, and euchromatin, a relatively open and transcriptionally “active” type of chromatin. Each part of chromatin can be covalently modified leading to changes in its organisation and function. These modifications include, methylation, hydroxylation of cytosine in DNA, and acetylation, methylation, phosphorylation, ubiquitination, sumoylation of amino acids in histones as summarised in Table 1 [1]. Presence and absence of these specific marks influence gene expression and is believed to drive epigenetic processes in organisms [2]. For example, methylation of the 5-carbon on cytosine residues in DNA is especially prevalent in regions of DNA rich in cytosine and guanine residues called CpG islands. DNA methylation was initially linked to transcriptional silencing, but the discovery that some hyper-methylated genes are actively transcribed challenged this dogma, making DNA methylation a context-dependent transcriptional regulator [1]. Indeed, inhibition of DNA methylation led to approval by the Food and Drug Administration (FDA) of azacitidine as a hypomethylating agent in the treatment of the myelodysplastic syndrome [1]. The wealth of histone modifications points to a complexity of chromatin regulation, which is dynamically changing and dependent on the cellular context and environment [1].

Epigenetic regulators can be divided into three categories: “writers” – enzymes responsible for adding modifications (e.g. histone methyltransferases), “readers” – proteins that detect modifications (e.g. bromodomains detecting acetylated lysine residues) and “erasers” – enzymes responsible for removing modifications (e.g. histone demethylases, histone deacetylases) (Table 1) [3]. Interestingly, many enzymes that modify histones also contain histone “reader” domains or catalyse modifications of non-histone substrates. For example, in conjunction with the methyltransferase NSD1, the JmjC demethylase FBXL11 dynamically regulates the methylation status of NF-κB, a key transcription factor of the immune and inflammatory response [4]. There are multiple therapeutic attempts or developments (as summarized below) where successful targeting of proteins directly involved with histone acetylation, deacetylation, acetylation readers and methylation [1] is reported.

Histone acetylation neutralises the positive charge found on lysine side-chains, thereby weakening the electrostatic interaction with negatively charged DNA and histones. Acetylation is installed by histone lysine acetyltransferases (KATs) and deacetylation is catalysed by histone deacetylases (HDACs). Similarly, as mentioned above for methyltransferases and demethylases, both KATs and HDACs often have alternative substrates besides histone proteins [1]. Based on sequence homology HDACs are divided into four classes: class I (HDAC 1-3, HDAC8), class II (HDAC 4-7 and 9-10) and class IV (HDAC11) which act in cofactor independent manner requiring a $Zn^{2+}$ ion, and class III (sirtuins 1-7), where the catalytic mechanism is dependent on $NAD^{+}$ [1]. HDAC inhibition has been explored therapeutically leading to FDA approval of the largely unspecific HDAC inhibitors Vorinostat and Romidepsin in cutaneous T cell lymphoma [5], [6].

More recently, the proteins involved in “reading” of the acetyl marks have also been subject to inhibition studies and drug development [7]. The largest class of acetyl-lysine “readers” are proteins with a bromodomain motif and this protein-protein interaction has been very

successfully targeted in the BET bromodomain family as demonstrated by advancement into clinical trials in haematological cancers and in NUT-midline carcinoma [8], [9].

Methylation is found on side-chain moieties of arginine, lysine and histidine residues of histones and unlike acetylation and phosphorylation it does not alter the original charge. The most studied histone methylation is found on lysyl residues of H3K4, H3K9, H4K20, H3K27, H3K36 and H3K79. Lysine methylation is catalysed by enzymes containing a methyltransferase SET domain with the exception of H3K79 methylation which is catalysed by DOT1L [1]. Scientific studies with methyltransferases have revealed therapeutic potential for histone methyltransferases such as for example Enhancer of Zeste Homolog 2 (EZH2) or DOT1L. H3K27 methylation is mediated through Polycomb repressive complex 2 (PRC2) of which EZH2 methyltransferase is a member [10]. Multiple EZH2 inhibitors have been reported to date including chemical probes, with EPZ011989 showing promising results in xenograft models of B cell lymphoma [11], [12]. Encouraging developments are also based on the DOT1L methyltransferase as demonstrated by EPZ-5676, a compound that has been successful in Phase I clinical trials for the treatment of MLL-rearranged leukaemia [12].

These exciting clinical developments highlight the therapeutic potential of targeting epigenetic regulators. This excitement might also extend to targeting the family of histone demethylases. Enzymes demethylating methyl-lysine residues in histone tails comprise two distinct families: the FAD-dependent LSD demethylases (KDM1A and KDM1B) and a subset of Jumonji containing 2OG and $Fe^{2+}$ oxygenases. Based on sequence similarity histone demethylases are split into ‘KDM’ (K-demethylases) subfamilies, with KDM1 representing LSD demethylases and KDM2-8 JmjC- type demethylases [13]. Jumonji-containing demethylases are frequently found to be multi-domain proteins catalysing reactions with substrates outside of histones [14]. Whereas KDM1 inhibitors are currently advanced into clinical trials in oncology, there are currently no approved Jumonji-demethylase inhibitors, however the field develops rapidly attracting interest from both

biotechnology and pharmaceutical sectors. This thesis aims to contribute to our understanding and development of strategies for inhibitor identification for the Jumonji-type oxygenases. The following sub-chapters summarize the involvement of Jumonji-type oxygenases in chromatin modification (1.2), provide an attempt to describe their biological roles (1.3), structural features (1.4) and inhibitor development (1.6).

| Chromatin modification | Chromatin writer | Chromatin reader domain | Chromatin eraser | Associated function |
|---|---|---|---|---|
| **DNA modification** | | | | |
| 5-methylcytosine | DNMT [15] [16] | MBD domain [16] | unknown | transcription (X-inactivation, imprinting, long-term silencing, development, gene regulation) [15] |
| 5-hydroxymethylcytosine | TET [17] | unknown | unknown | transcription [16] |
| 5-formylcytosine | TET [17] | unknown | c5-MTases [18] | unknown |
| 5-carboxylcytosine | TET [17] | unknown | | unknown |
| **Histone modifications** | | | | |
| acetylation | KAT [16] | Bromodomain PHD fingers [16] | HDAC [16] | transcription, repair, replication and condensation [16] |
| Methylation (lysine) | KMT [16] | Chromodomain, Tudor domain, MBT domain, PWWP domain, PHD fingers, WD40/β propeller[16] | KDM | transcription and repair [16] |
| Methylation (arginine) | PRMT [16] | Tudor domain [16] | unknown* | transcription [16] |
| Phosphorylation (serine and threonine) | Kineases (eg. Aurora, JAK2) [16] | SH2 [16] | Protein phosphatases [16] | transcription and repair [16] |
| Ubiquitylation | Ubiqutinases [19] | UIM, IUIM [16] | Deubiqutinases [19] | |
| Sumoylation | SUMO [20] | SIM [16] | | |
| ADP ribosylation | PARPs [21] | Macro domain, PBZ domain [16] | unknown | |
| Deimination | PAD [22] | unknown | unknown | Transcription and decondensation [16] |
| Proline isomerisation | Proline isomerase [23] | unknown | unknown | transcription [16] |
| Crotonylation | unknown | unknown | unknown | Transcription, chromatin structure [16] |
| Propionylation, Butyrylation, | HAT [21] | unknown | HDAC [21] | unknown |
| Formylation, Malonylation, Succinylation, Hydroxylation | unknown | unknown | unknown | unknown |
| O-GlcNAcylation | OGT [15] | unknown | OGA [15] | unknown |

**Table 1. Posttranslational modifications found in chromatin, their readers, writers, erasers and associated function adapted from** [1], [15]**. *demethylation activity of JMJD6 is disputed [24]**

**Protein families**: DNA methyltransferase (DNMT), histone acetyltransferase (HAT), histone deacetylase (HDAC), protein arginine methyltransferases (PRMT), O-GlcNAc transferase (OGT), O-GlcNAcase hydrolase (OGA), poly(ADP-ribose) polymerases(PARP), protein arginine deiminase (PAD), small ubiquitin like modifier (SUMO), lysine demethylases (KDM)

**Reader domains**: methyl-CpG-binding domain (MBD), plant homeodomain (PHD), malignant brain tumor domain (MBT), proline-tryptophan-tryptophan-proline domain (PWWP), BRCA1 C terminus domain (BRCT), ubiquitin interaction motif (UIM), inverted ubiquitin interaction motif (IUIM), sumo interaction motif (SIM), poly ADP-ribose binding zinc finger (PBZ), chromo domains, macro domains, Tudor domains.

## 1.2. 2OG oxygenases are epigenetic modifiers

*LSD and JmjC demethylases modify chromatin*

The discovery of the first histone demethylase protein, Lys-specific demethylase 1 (LSD1) was described in 2004 by Shi *et al.* [25] and shortly followed by Schüle *et al.* [26]. This landmark discovery changed the way the scientific community thought about histone methylation: previously histone demethylation was considered a permanent modification of largely unknown significance; a modification which could only be removed by histone exchange or “dilution” during DNA replication [27]. The process of demethylation is now considered to be highly dynamic and regulated [27]. Many histone demethylases are associated with promoters and other transcriptional regulators contributing to transcriptional activation or suppression. It is thought that instead of tight regulation of transcription, histone demethylases create certain chromatin states, which provide a permissive environment for transcriptional complexes. The demethylase activity of the enzymes can in turn be regulated by their expression levels, modulated by cofactors, cosubstrates, and small molecules (including endogenous TCA cycle intermediates) and altered by association with other transcriptional activators or repressors [27].

The LSD-type demethylases, LSD1 and LSD2, belong to the larger amine oxidase family, which are dependent on the Flavin Adenine Dinucleotide (FAD) cofactor for catalysis. LSD1 and LSD2 are able to demethylate mono-, di- but not tri-methylated lysine 4 and/or lysine 9 on histone H3 [28]. This enzyme family contains also monoamine oxidases A

(MAO-A) and MAO-B, which are involved in degradation of neurotransmitters such as dopamine [28], [29]. The second family of demethylases, the $Fe^{2+}$- and 2OG- dependent JmjC family, with over 30 members in humans, contains a highly conserved catalytic *JmjC* domain composed of a core double-stranded beta-helix fold (DSBH), and shows activity towards various methylated states of H3K4, H3K9, H3K27 and H3K36 [14] (Figure 1).

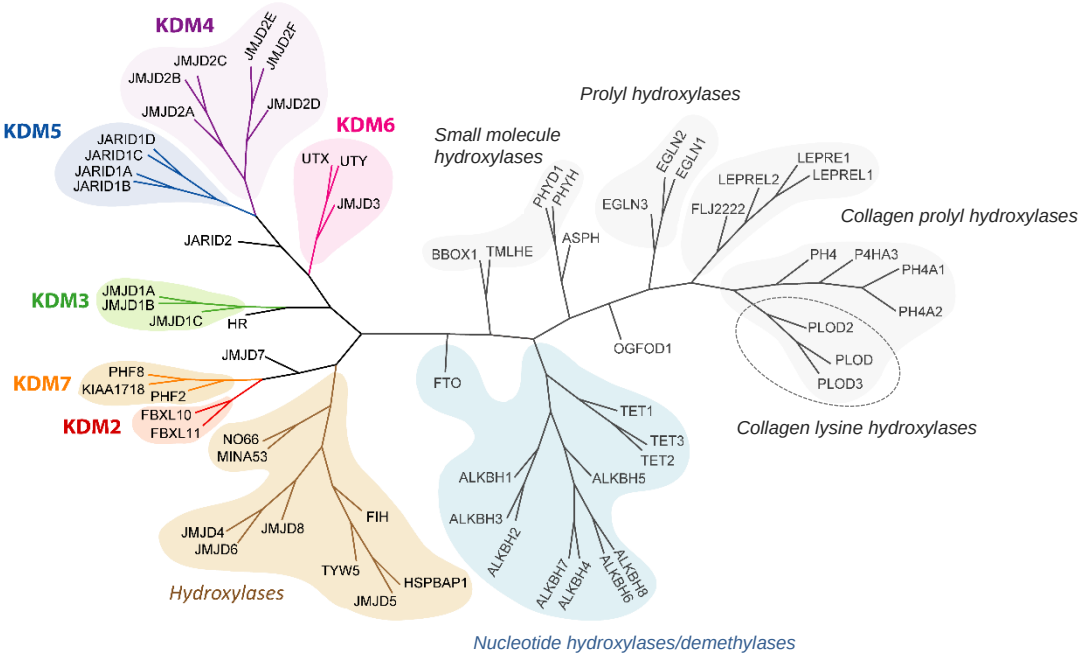



**Figure 1. Phylogenetic tree of human 2-oxoglutarate- and iron- dependent oxygenases based on the alignment of conserved *JmjC* domain double-stranded beta-helix (DSBH) fold, adapted from [14] .** The lysine demethylase families are highlighted in red – KDM2, green – KDM3, violet – KDM4, blue – KDM5, pink – KDM6 and orange – KDM7. JmjC- type hydroxylases are highlighted in brown, nucleotide hydroxylases/demethylases – cyan and other small molecule hydroxylases, prolyl hydroxylases and collagen prolyl hydroxylases in grey.

Other members of the JmjC family contain the oxygenases MINA53, NO66 and OGFOD1 which were found to catalyse ribosomal hydroxylation [30], broadening the scope of substrates and suggesting importance not only in transcriptional regulation of chromatin, but also in protein translation. This is only one example, demonstrating that the discovery of the substrate spectrum continues. The following paragraph gives a brief outline of the growing evidence for important biological roles of the Jumonji-type demethylases and hydroxylases in human physiology.

## 1.3. Roles of JmjC- type oxygenases in biology and disease

Although the post-translational modifications catalysed by the JmjC family of enzymes are explored with great effort [27], the link between these modifications and the biological function in human disease is still ill-defined in many cases.

**The KDM2 subfamily**

The KDM2 subfamily contains two members FBXL10 and FBXL11 which both catalyse demethylation of H3K36me2/1 methyl marks with FBXL10 also found to demethylate H3K4me3 [31]. The KDM2 family places itself on the same branch of the phylogenetic tree of JmjC- type oxygenases together with KDM7 implying significant similarity (Figure 1). In addition to histone substrates FBXL11 has been demonstrated to regulate the activity of the transcription factor NF-κB by demethylation of methylated Lys218 and Lys221 residues [14]. FBXL10 and FBXL11 both contain CxxC domains, a motif that is known to recognise unmethylated cytosine residues in CpG islands. Indeed FBXL10 and 11 bind to unmethylated CpG islands changing its chromatin environment, for example, FBXL10 mediates ubiquitination of histone 2A, through recruitment of Polycomb repressive complex 1 (PRC1) [32]. KDM2 enzymes have been also implicated in control of proliferation, biological aging, regulation of stem cells and are also involved in oncogenic processes [14].

**The KDM3 subfamily**

The KDM3 subfamily consists of the enzymes JMJD1A-B and HR (hairless) with H3K9me2/me1 demethylation activity and JMJD1C which has been shown to demethylate the MDC1 protein involved in the DNA- damage response (DDR) pathway [14], [33], [34]. The JMJD1A-C enzymes are widely expressed in human tissues and are localised primarily in the nucleus [14]. JMJD1A, also known as testis specific gene A (TSGA), is

expressed during spermatogenesis and co-activates androgen receptor mediated transcription [14]. Indeed, knockout of JMJD1A in male mice resulted in infertility, smaller testes and severe reduction of sperm count compared to control litter mates. Mice with a JMJD1A-knockout show an adult obesity phenotype implying an important role in control of transcription of metabolic genes [14]. Expression of JMJD1A is regulated by HIF-1α and is up-regulated in cells exposed to hypoxia [35]. Overexpression of JMJD1A was observed in a wide spectrum of human cancer cells derived from bladder, lung, hepatocellular, prostate, colorectal and renal cell carcinomas. A JMJD1A siRNA treatment of hepatocellular carcinoma cell lines resulted in suppression of proliferation [36] and inhibition of JMJD1A in a mouse xenograft model was shown to downregulate angiogenesis [37]. The JMJD1B gene locus on chromosome 5q31 is frequently deleted in breast cancer and myeloid leukaemias [38], [39]. JMJD1B is found to regulate expression of the leukemic oncogene *lmo2* and interacts with suppressor of cancer cell invasion (SCAI) [40]. JMJD1C is in terms of primary structure similar to JMJD1A and is widely expressed in human tissues [14]. The JMJD1C gene is important in development of mouse testes and regulation of the expression of marker of steroidogenesis p450c17 via SF-1-mediated transcription [41]. Expression of JMJD1C is found to be reduced in human breast cancer tumours in comparison to healthy tissue suggesting a role in tumour suppression [14]. JMJD1C is required for maintenance of acute myeloid leukaemia (AML) and identified as a potential therapeutic target for the disease [42]. The last member of the KDM3 family is a protein critical for maintenance of hair growth and named HR (*hairless*) after the murine knockout phenotype [43]. HR has been shown to interact with histone deacetylases and co-represses multiple nuclear receptor activities, including thyroid hormone receptor, vitamin D receptor and retinoic acid receptor-related orphan receptors [14].

**The KDM4 subfamily**

The JMJD2A-D enzymes belong to the KDM4 family and catalyse H3K9me3/me2 and less well linker histone H1.K26me demethylation [14], [44]. JMJD2A-C enzymes are also active

against H3K36me3, although at a lower rate [45]. The KDM4 family contains two further members, JMJD2E-F, which due to a lack of genomic intron structures are currently considered as pseudogenes [46]. The involvement of JMJD2 demethylases in cancer has been extensively studied and in many cases linked to co-regulation of hormone receptor activities. KDM4 enzymes have been shown to co-activate the androgen receptor (AR), and as in the case of JMJD2B, to regulate AR transcriptional activity and prevent proteosomal degradation of AR [14]. Overexpression of JMJD2A-D is observed in many cancers, such as JMJD2C in estrogen, progesterone and HER2 receptor negative breast cancers or JMJD2B in ER-positive breast cancer cells [14]. Gene silencing on the other hand leads to reduced cell proliferation, as in the case of JMJD2A in breast cancer and ER positive breast cancer cells, in line with the overexpression studies [14]. JMJD2B and JMJD2C have also been implicated in regulating the transcriptional response of cancer cell lines in hypoxia, through hypoxia inducible transcription factor HIF1α [47]. JMJD2D interacts with the p53 tumour suppressor [48]. Apart from their roles in cancer, an important role in epigenetic control of tissue remodelling and embryonic development begins to emerge, such as in the case of JMJD2A (promotion of cardiac hypertrophy in response to cardiac stresses) [49], or JMJD2B (osteogenic differentiation of mesenchymal stem cells (MSC)) which controls repressive K3K9me3 marks at osteogenic-lineage selective genes [50]. Copy number variation is a hallmark of cancer and JMJD2A has been recently shown to play an important role in genomic stability by inducing JMJD2A-dependent transient site- specific copy gains (TSSGs) [51], [52]. Interestingly the TSSGs are also induced by hypoxia in a JMJD2A-dependent manner and were shown to be inhibited by inhibition of JMJD2A with small molecules or the TCA cycle metabolite succinate [53]. Taken together, the well-documented involvement of KDM4 sub-family

members in the development of various type of cancers makes them interesting drug targets [54], [55].

**The KDM5 subfamily**

The KDM5 family consists of four members JARID1A-D, all of which are multi-domain proteins known to demethylate a H3K4me3/me2 substrate [56]–[59]. This particular histone mark is a hallmark of active gene transcription, making KDM5 enzymes transcriptional co-repressors [14]. Members of the KDM5 family participate in chromatin complexes important in establishing repressive marks in chromatin- for example JARID1A can transiently associate with the PRC2 (Polycomb repressive complex 2), or JARID1C is found in complex with HDAC1/2/REST and the H3K9 methyltransferase G9a [14]. Recent reports on the roles of the KDM5 family members in biology and disease paint an increasingly complex picture: for example, JARID1A can promote cellular differentiation through interaction with retinoblastoma protein (Rb), a master regulator of cell cycle and differentiation [60], in stem cell biology JARID1A is shown to interact with HOX genes during development [56], and is also found to influence the circadian clock [61]. In cancer, JARID1A is involved in paediatric acute megakaryoblastic leukaemia through a NUP98/JARID1A gene fusion [62]. Interestingly, the emergence of drug-tolerant subpopulations of cancer cells was linked to a specific chromatin state generated by the JARID1A demethylase activity [63]. JARID1B was shown to block differentiation in embryonic and hematopoietic stem cells, however JARID1B was also shown to be dispensable for embryonic stem cell maintenance in a separate study and essential for neural differentiation [14], [64]. JARID1B is specifically up-regulated in melanoma and breast cancer, thereby suppressing tumour development and angiogenesis for example in breast cancer cells [14]. The JARID1C gene is located on the X-chromosome and implicated in X-linked intellectual disability (XLID) through mutations, whereas JARID1D is found on the Y-chromosome in a genomic region previously linked to spermatogenesis [14]. JARID1C is expressed in brain tissue and polymorphic variants are linked to autism

spectrum disorders, implicating an important role in normal brain development and function [14]. Manipulation of JARID1C expression levels was shown to alter gene expression profiles and to suppress tumour growth by interaction with the von Hippel-Lindau protein [65]. JARID1D regulates the process of spermatogenesis through interaction with the MSH5 meiosis-regulatory protein and deletions of the JARID1D gene have been discovered in prostate cancer [66], [67].

**The KDM6 subfamily**

The KDM6 family of 2OG oxygenases contains three histone H3K27me3 demethylases: whereas JMJD3 is found on chromosome 17, UTX is located on chromosome X, and UTY is located on chromosome Y [14], [68]. UTY is less well characterised but implicated in embryonic development through male-specific H3K27me3 demethylation [69]. JMJD3 is important in developmental biology, due to the significance of the H3K27me3 mark for Polycomb mediated repression of developmental genes [70], for example JMJD3 enables expression of genes in the Wnt and Nodal pathways which are crucial for development [14]. Knockdown of JMJD3 is known to impact development of the blastocyst, endoderm and mesoderm structures through inhibition of H3K27me3 demethylation [71]. JMJD3 plays an important role in cellular differentiation and induced pluripotent stem cell (iPS) biology. JMJD3 is involved in body and limb morphogenesis at later stage of embryo development, in mouse embryonic fibroblasts (MEFs) overexpression of JMJD3 inhibited iPS formation during iPS reprogramming, whereas silencing of JMJD3 enhanced it [72]. In osteoblast development, for example, knockdown of JMJD3 reduced mesenchymal stem cell differentiation to osteoblasts [50]. JMJD3 is also involved in control of osteoclast development via regulation of the NFATc1 transcription factor whose expression is dependent on the H3K27me3 level in its promoter region [73]. In neural cells, JMJD3 was shown to regulate key genes involved in neural differentiation and commitment [74], [75]. A JMJD3 knockout was found to be lethal in mice at the perinatal stage and shown to be important for maintenance of an embryonic respiratory neuronal network [74]. Expression

of JMJD3 in LPS-stimulated macrophages was found to be regulated through the NF-κB pathway, leading to the discovery of important roles of KDM6 enzymes in inflammation [76]. In macrophages, H3K27 demethylation mediated by JMJD3 was inhibited by the KDM6 inhibitor GSK-J1 regulating the TNF promoter and TNF-alpha levels [77]. In T-cells JMJD3 influences expression of HPK1/MAP4K1, a kinase negatively regulating immune response, and is also involved in T-helper cell lineage commitment [78]. Overexpression of JMJD3 was demonstrated to suppress the T-cell response via HPK1 upregulation, making it a possible drug target for treatment of systemic lupus erythematosus, an autoimmune disorder [78]. Recently JMJD3 and UTX have been shown to be crucial in differentiation of T-cells through their H3K27 demethylase activity [79]. Both JMJD3 and UTX have been found to be up-regulated during wound healing [80], [81]. A wealth of biological studies also points to the importance of KDM6 proteins in cancer development. Depending on the cellular context JMJD3 can have suppressive or pro-proliferative effects. For example overexpression of JMJD3 is associated with increased proliferation and regulation of anti-apoptotic genes [82], [83], but in cells with oncogenic stress JMJD3 activated by RAS/RAF signalling pathway causes p53-dependent cell cycle arrest [84]–[86]. Interestingly, UTX is found mutated in patients with Kabuki syndrome, a rare intellectual disability syndrome [87]. Mutations and deletions of UTX have been identified through genome-wide analyses in human cancers including leukaemias as well as bladder, prostate, breast and renal cancer [88]–[91]. UTX is also identified as gender-specific tumour suppressor in T-cell acute lymphoblastic leukaemia (T-ALL) [92], [93] and as tumour suppressor of human breast cancer [94], [95].

**The KDM7 subfamily**

The KDM7 family comprises three members: KIAA1718, PHF8 and PHF2, enzymes which are related to the KDM2 family members. In addition to the catalytically active *JmjC* domain, KDM7 enzymes have an N-terminal plant homeodomain (PHD). Members of this subfamily were shown to demethylate H3K9me1/2, H3K27me1/2 or H4K20me1

substrates [96]. The C-terminal region is composed of coiled-coiled domains, containing phosphorylation and nuclear localisation sites, and has been shown to enable association with RNA polymerase, with transcription factors such as the retinoic acid receptor and likely with cell cycle regulators such as E2F1 [96]. KDM7 enzymes are involved in the regulation of a pro-inflammatory response via demethylation of H4K20me3. They also play a developmental role, for example by regulating neural differentiation through FGF4 [97], [98]. In cancer biology, PHF8 is regulating the retinoic acid response in acute promyelocytic leukaemia [99], and PHF2 has been implicated in development of breast cancer [100]. Mutations in PHF8 have been determined as causative in the development of X-linked mental retardation called Hamel-Siderius syndrome [101], [102].

**Other related JmjC- type hydroxylases and demethylases**

This functionally diverse group of JmjC- type oxygenases encompasses very well studied members such as factor inhibiting HIF (FIH), and poorly characterised proteins such as JMJD8 or JMJD4 [14]. FIH is an asparagine hydroxylase which catalyses $Asp^{803}$ hydroxylation in the C-terminal transactivation domain of the HIF1α transcription factor which undergoes proteolytic regulation by ubiquitin ligase via von Hippel-Lindau tumour suppressor (pVHL) recognition [103]. HIF1α proteins regulate the transcriptional response to hypoxia involving genes responsible for metabolic adaptation to low $O_2$, hypoxia tolerance, redox and pH homeostasis, and others [104]. The activity of HIF1α is regulated in normoxic conditions by prolyl hydroxylases PHD1 and PHD2 in tandem with FIH, which continuously inactivate and degrade HIF1α. Prolyl hydroxylation of HIF1α leads to its higher affinity for pVHL and asparaginyl hydroxylation of HIF1α blocks the p300 co-activator resulting in reduced activity of HIF1α. In hypoxic conditions PHD enzymes become less active leading to accumulation of HIF1α, which however continues to be hydroxylated by FIH. In severe hypoxia even FIH enzymes are inactive and HIFα remains unmodified [103] which results in high levels of HIF1α transcription factor. It is important to note that hydroxylation is an irreversible process in this context, and only rapid re-

synthesis of HIF1α allows for the inactive HIF1α subunits to be replaced [103]. Interestingly FIH-null mice display significant metabolic changes such as resistance to high fat diet, reduced weight and increased metabolic rate, but no increased hypoxia responses [105]. Since this initial discovery the list of substrates of FIH has increased significantly and now includes hydroxylation of asparagine and aspartate in ankyrin repeat domains (ARDs) in proteins such as NF-κB or Notch, and hydroxylation of $Asp^{242}$ within the cytoplasmic ARD domain of transient receptor potential vanilloid 3 (TRPV3) ion channel [106][107]. The activity of TRPV3 ion channel was found to be inhibited by FIH activity and to be dependent on hypoxia [107]. The enzymatic amino acid hydroxylation activity of FIH appears to be widespread and the list of substrates is likely to grow further. In non-small-cell lung cancer overexpression of FIH and PHD hydroxylases is linked to poor prognosis [108], whereas in clear cell renal cell carcinoma it is low nuclear expression that is linked to poor prognosis [109]. FIH is also expressed in invasive breast carcinomas, and inhibition of FIH in clear cell renal cell carcinoma decreases cell expansion and increases apoptosis via HIF1α mediated responses [110]. FIH is inhibited by an oncoprotein expressed in hepatocellular carcinomas called gankyrin, resulting in hemangioma caused by increased HIF1α and VEGF expression [111].

JMJD4 shows approximately 30% sequence similarity to JMJD6, but as in the case of JMJD7 its catalytic activity has not yet been reported in literature. JMJD5 was classified as a H3K36me2 demethylase [112], although the activity needs to be confirmed in further studies. It is postulated that JMJD5 regulates the cell cycle through demethylation mediated inhibition of HDAC recruitment and inhibition of the cyclin A1 locus [112]. In addition to its demethylase activity, JMJD5 was shown to hydroxylate transcriptional regulators, such as NFATc1, and control its proteolytic degradation through promotion of E3-ligase association [113]. The circadian clock of cells with JMJD5 deficiency have a shorter period as demonstrated both in mammalian cultures and *Arabidopsis* plants [114]. The function of JMJD5 in cancer appears to be related to the activity of the tumour

suppressor p53. JMJD5 was shown to form a complex with p53 negatively regulating its activity, and conversely regulating cell cycle progression and proliferation [115]. Overexpression of JMJD5 was reported in leukaemia and breast cancer, where knockout studies in MCF7 breast cancer cell lines showed reduced proliferation [112]. Ablation of the JMJD5 gene in mice results in severe growth retardation and embryonic lethality [116], [117]. The member JMJD6 was initially characterised as phospatidyl serine receptor and as a methyl arginine demethylase, functions which have been disputed in the literature [14]. Controversies on the functional assignments of JMJD6 are summarised in a recent review [24]. JMJD6 catalyses hydroxylation of lysyl residues on their carbon side chains of arginine-serine rich regions present in, for example, splicing factors [118], [119] and JMJD6 was also suggested to modify single-stranded RNA [120]. JMJD6 is essential during embryogenesis for normal development of tissues such as brain, eyes, heart, lung, kidney, liver and intestine as demonstrated in JMJD6 knockout mice which died before birth [24]. Overexpression in breast cancer and lung carcinoma has been linked to poor prognosis [121], [122]. JMJD8 is another poorly studied JmjC- type oxygenase with unknown enzymatic function. In cancer, cell invasion was inhibited by 80% in JMJD8 siRNA knockdown in squamous cell carcinoma SCC23/MET cells [123]. HSPBAP1 is a JmjC- oxygenase with unclear enzymatic function expressed in thymus and pancreas, and shown to inhibit the function of heat shock protein hsp27. Hsp27 was found to be neuroprotective in animal models of peripheral nerve injury and motor neuron disease [14]. Interestingly HSPBAP was also significantly expressed in the temporal neocortex of patients with intractable epilepsy [124]. TYW5 is a tRNA hydroxylase with unknown biological function [125].

MINA53 is localised to the nucleus and nucleolus [126], and closely linked to expression of the transcription factor c-Myc [127]. The expression of MINA53 follows expression of c-Myc in human leukaemia cells, and siRNA methods that specifically inhibit expression of MINA53 results in severe suppression of cell proliferation [127]. A series of papers

confirmed MINA53 effects on cell differentiation in various cancers, and its overexpression appears to be linked to oncogenesis [128]–[132]. MINA53 has been especially well characterised in lung cancer making the protein a possible biomarker and potential drug target for the disease [131], [132]. The role of MINA53 was also explored in non-neoplastic tissues where the expression was found to be induced by silica particles [132]. This suggested a key regulatory role for MINA53 in the immune system e.g. in the allergen induced inflammatory response [133], and importantly in differentiation of pro-inflammatory TH17 cells [134]. Taken together these studies conclude that MINA53 is an important regulator in inflammation and oncology. However, the underlying molecular mechanism both in oncology and the immune system remains enigmatic. Initially it was reported that MINA53 could demethylate H3K9me3 *in vivo* [135], but this activity has been controversial and no demethylase activity has been detected *in vitro* neither by us or others, e.g. Ge *et al*. [30] and Wang *et al.* [136]. Recently MINA53 has been shown to catalyse histidyl hydroxylation of ribosomal protein L27a *in vitro* and *in vivo*, suggesting a function in ribosomal biogenesis [30]. Thus the mechanistic link between protein hydroxylation and the cellular phenotypes reported upon deletion or overexpression of MINA53 awaits clarification.

The closely related oxygenase NO66 (also known as nucleolar protein 66 (NO66) or Myc-associated protein with *JmjC* domain (MAPJD)) has significant sequence homology to MINA53 (34% sequence identity) and the protein structure is highly conserved throughout evolution [137]. NO66 is localised in the granular part of nucleoli, nucleoplasm and nucleoplasmic foci [153], and was first identified and characterised in 2004 by Eilbracht *et al*. in purified nucleoli from *Xenopus laevis* oocytes [153]. NO66 has been shown to regulate osteoblast differentiation and bone formation by inhibiting Osterix (Osx) [154], a key transcription factor essential at later stages of bone development. Accordingly, Osx-null mice have a normal cartilage development, but lack bone formation. Conversely, knockdown of NO66 in differentiating osteoblasts derived from mesenchymal stem cells

causes accelerated differentiation and mineralisation of osteoblasts [154]. Further studies in mouse model confirmed the effect of NO66 modulation on bone development in-vivo [138], [139]. Prx1-Cre-dependent mesenchymal deletion of NO66 in mice promotes bone formation by increasing the number of pre-osteoblasts and osteoblasts [155]. Conversely, overexpression of NO66 in mice driven by the Prx1 promoter inhibited skeletal growth and bone formation and reduced the H3K36me3 in bones by ~50% [140]. In neoplastic tissues, NO66 is found to be over-expressed in non-small cell lung cancer [142]. Taken together these studies conclude that NO66 has an important role in bone formation and constitutes a potential drug target. During mouse embryonic stem cell differentiation NO66 is recruited by PHF19, a component of polycomb repressive complex PRC2, to stem cell genes resulting in loss of H3K36me3, transcriptional silencing and PRC-2- dependent methylation of H3K27 [156]. Although NO66 has been reported to demethylate lysine residues H3K4me3/me1 and H3K36me3/me2 *in vivo* [154], this activity has (similar to MINA53) been controversial since several groups have reported lack of demethylase activity *in vitro*. NO66 was recently shown to catalyse histidinyl hydroxylation of the ribosomal L8 protein [30], implicating a role in ribosome biogenesis.

## 1.4. Structural studies of 2-oxoglutarate- and iron- dependent oxygenases

The catalytic domain of JmjC-type oxygenases contains a highly conserved double-stranded β-helix (DSBH) fold with eight β-strands, also referred to as cupin, ‘jelly-roll’ or ‘beta barrel’ fold [140]. Additional structural elements are frequently found in the DSBH fold including additional C-terminal helices, β-strands, as in NO66, or ‘inserts’ of up to 400 amino acids between the fourth and fifth β-strands as depicted in Figure 2 [140], [141]. Crystallographic studies reveal the $Fe^{2+}$ binding site is found at the more open end of the DSBH ‘barrel’ and the β-strands I,II and surrounding loops frequently involved in peptide binding [141].

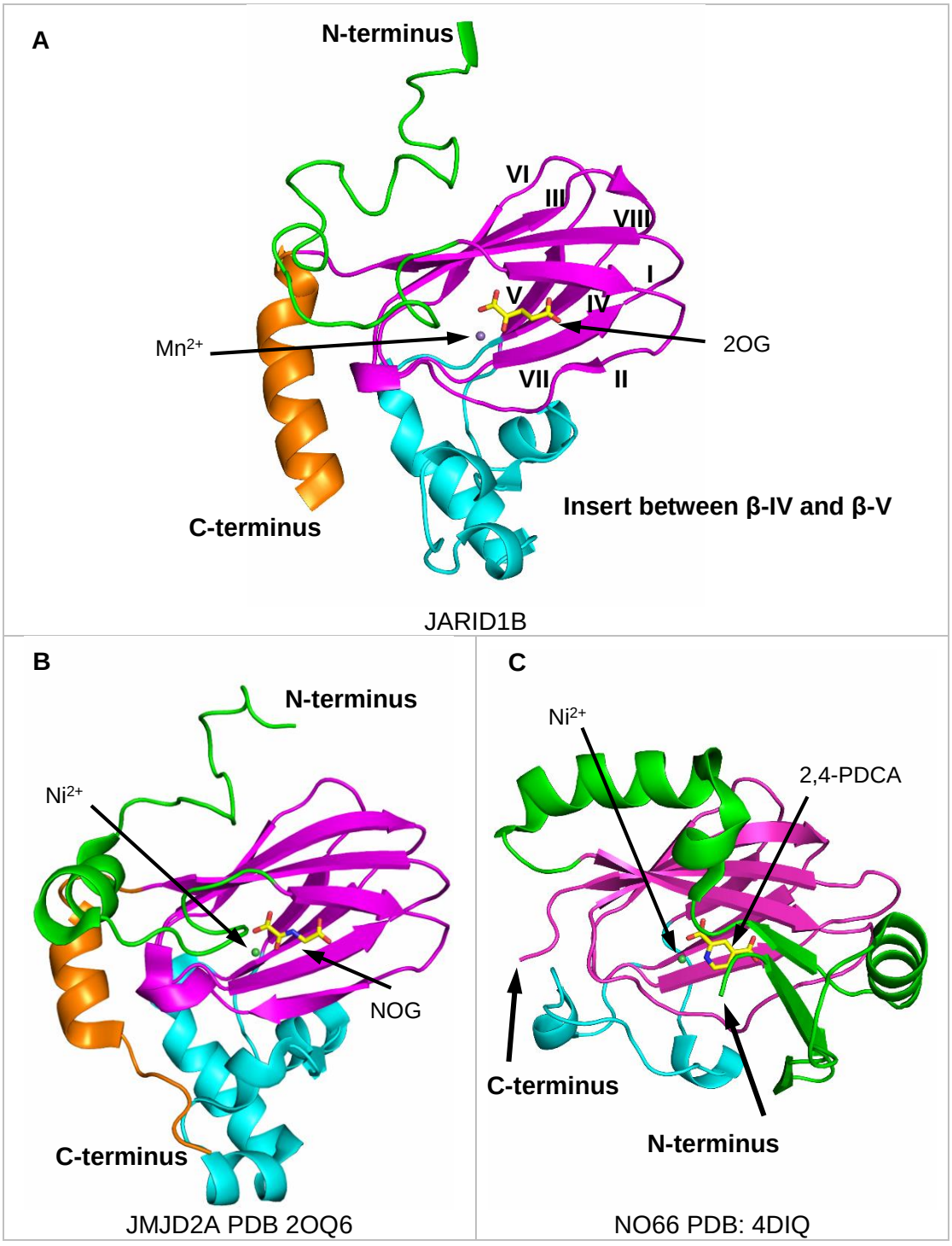
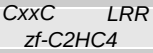
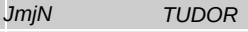
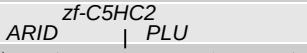



**Figure 2. A. Catalytic *JmjC* domain of JARID1B (Glu453-Arg619) in complex with cosubstrate 2OG and cofactor $Fe^{2+}$ substituted by $Mn^{2+}$ (unpublished) displays the conserved DSBH fold (magenta). Roman numerals denote conserved beta-strands present in the DSBH fold [141]. A typical insert between beta strands IV and V is displayed cyan. The C-terminal alpha helix is coloured in orange and the N-terminal loop in green. B. The *JmjC* domain in JMJD2A PDB ID: 2OQ6 (Glu142-Cys308) with NOG bound and C. NO66 PDB ID: 4DIQ (Asn260-Arg425) in complex with 2,4-PDCA. NO66 contains two additional β-strands depicted in green.**

2OG oxygenases are commonly present in monomeric form or as dimers, although higher oligomerisation states are observed, for example with NO66 which exists as tetramer in solution, and the oligomerisation is required for catalytic activity [140], [142]. Many of the Jumonji-type demethylases contain other conserved domains such as plant homeodomain (PHD), Tudor, AT-rich interaction domain (ARID) or a CxxC zinc finger motif. In many cases substrate selectivity is achieved mostly through the *JmjC* domain or is mediated via additional domains where an interesting example is PHF8, which contains both *JmjC* domain and PHD domains (Table 2). The PHD domain binds to tri-methylated lysine-4 in histone H3 substantially increasing the demethylation activity of PHF8 towards di-methylated lysine-9 on histone H3 [143]. Table 2 presents the overview of structural domains present in Jumonji-type demethylases and hydroxylases.

| Name | Domain Architecture |
|---|---|
| **KDM2** |   |
| FBXL10 | |
| FBXL11 | |
| **KDM3** |   |
| JMJD1A | |
| JMJD1B | |
| JMJD1C | |
| *HR* | |
| **KDM4** |   |
| JMJD2A | |
| JMJD2B | |
| JMJD2C | |
| JMJD2D | |
| JMJD2E | |
| JMJD2F | |
| **KDM5** |   |
| JARID1A | |
| JARID1B | |
| JARID1C | |
| JARID1D | |
| *JARID2* | |
| ***KDM6*** |   |
| UTX | |
| UTY | |
| JMJD3 | |
| **KDM7** | |
| PHF2 | |
| PHF8 | |
| KIAA1718 | |
| **HYDROXYLASES** |   |
| NO66 | |
| MINA53 | |
| JMJD4 | |
| JMJD5/KDM8 | |
| JMJD6 | |
| JMJD7 | |
| JMJD8 | |
| FIH | |
| HSPBAP1 | |
| TYW5 | |

- JmjC, - PHD, - JmjN, - TUDOR, - LRR, - TPR, - ARID, - FBOX, - wHTH, -GATAL, -zf-C2HC4, -CxxC, -zf-C5HC2, - PLU, -Zn-binding

**Table 2. Domain organisation of JmjC- type oxygenases and hydroxylases adapted from [14].**

In crystal structures determined in a cosubstrate-free state, the ferrous iron is coordinated by two to three water molecules, two histidine residues and depending on the specific enzyme either a Glu or Asp residue, which are present in the conserved HxE/DxH motif [140], [141] (Figure 3). The 2OG cosubstrate binds in a bidentate manner chelating the metal with the C1 carboxylic acid oxygen and the 2-oxo group, whereas the C5 carboxylic acid is stabilised by electrostatic interactions with basic residues such as Arg/Lys and/or through hydrogen bond interactions with Tyr/Thr/Ser residues [14]. A typical 2OG binding is shown in the structure of JARID1B in Figure 3.

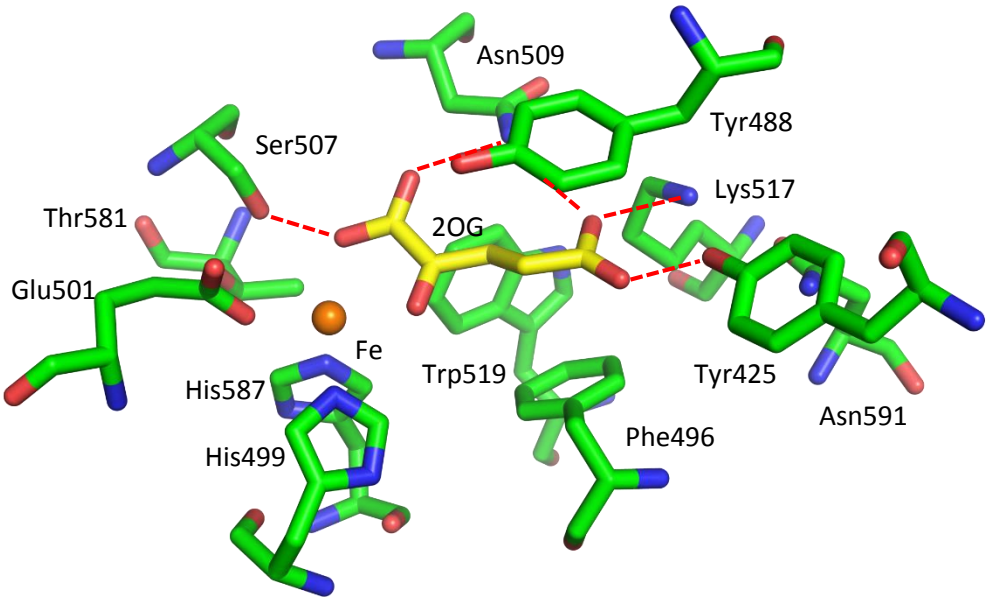



**Figure 3. Catalytic active site of JARID1B showing typical 2OG binding to $Fe^{2+}$ coordinated by two histidine residues and glutamic acid.**

The substrates of JmjC- type enzymes range from ribosomal proteins, tRNA, transcription factors to histone tails and their binding varies considerably between the enzymes as determined by multiple crystallographic studies with the peptide substrates [136], [137], [144]. The peptide binding has been especially well characterised in case of lysine demethylases such as JMJD2A where crystal structures of tri- and di- and mono-methylation states of lysine 9 and lysine 36 of histone H3 are available [144]. The tri-methylated lysine in the peptide substrate (PDB ID: 2OQ6) fits tightly in the binding site with the tri-methyl amino group positioned directly towards the active site metal as

presented in Figure 4 [144]. The structural studies reveal that many of the histone demethylases or ribosomal hydroxylases such as NO66 or MINA53 may have in fact other protein substrates that await identification [30], [136].

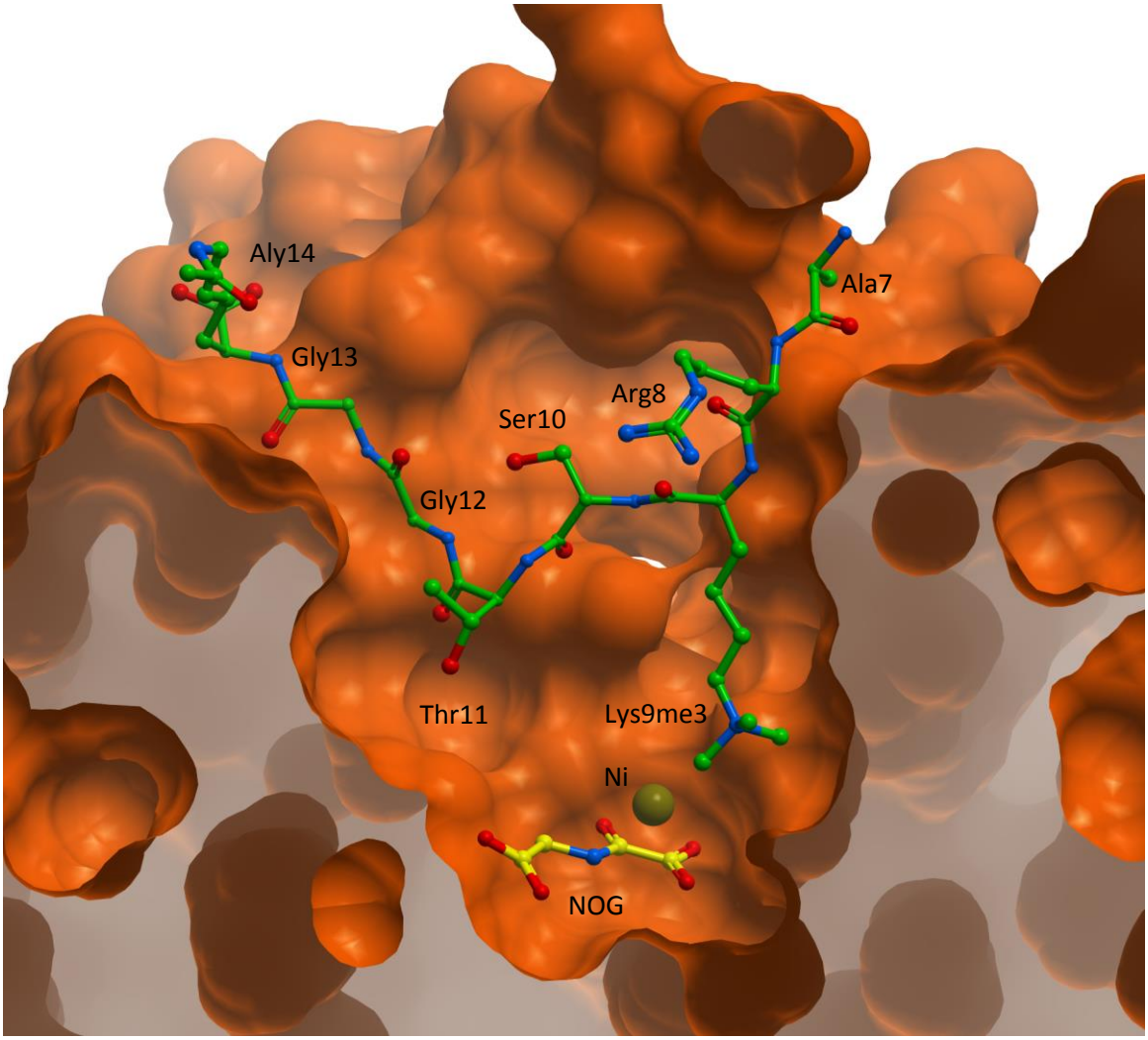



**Figure 4. The H3K9me3 peptide substrate binding site in JMJD2A PDB ID: 2OQ6.** The H3K9me3 histone peptide substrate is shown as green sticks, NOG as yellow sticks and $Ni^{2+}$ replacing the $Fe^{2+}$ cofactor as brown sphere.

## 1.5. Catalytic mechanism of 2-oxoglutarate- and iron-dependent oxygenases

The general mechanism of hydroxylation by 2-oxoglutarate- and iron- dependent oxygenases is shown in Figure 5. The 2-oxoglutarate cosubstrate binds to the metal in a bidentate manner via its C1 carboxylate and ketone oxygen displacing two metal-ligated

water molecules (2) (Figure 5). The primary substrate then binds to the active site, weakening the binding of the remaining water molecule coordinated to the metal and triggering the further reaction (3). Oxygen then binds to Fe (II) reacting with 2-oxoglutarate in an oxidative decarboxylation process generating $CO_2$, succinate and a highly reactive oxo-ferryl intermediate, which in turn hydroxylates the substrate (4-7) [109]. A methyl-lysine demethylation reaction happens via methyl group hydroxylation generating an unstable hemiaminal intermediate, which fragments into formaldehyde and a lysine side-chain reduced by a methyl group [110]. This mechanism allows demethylation of all three (me3, me2 and me1) methylation states found in methylated lysine residues.

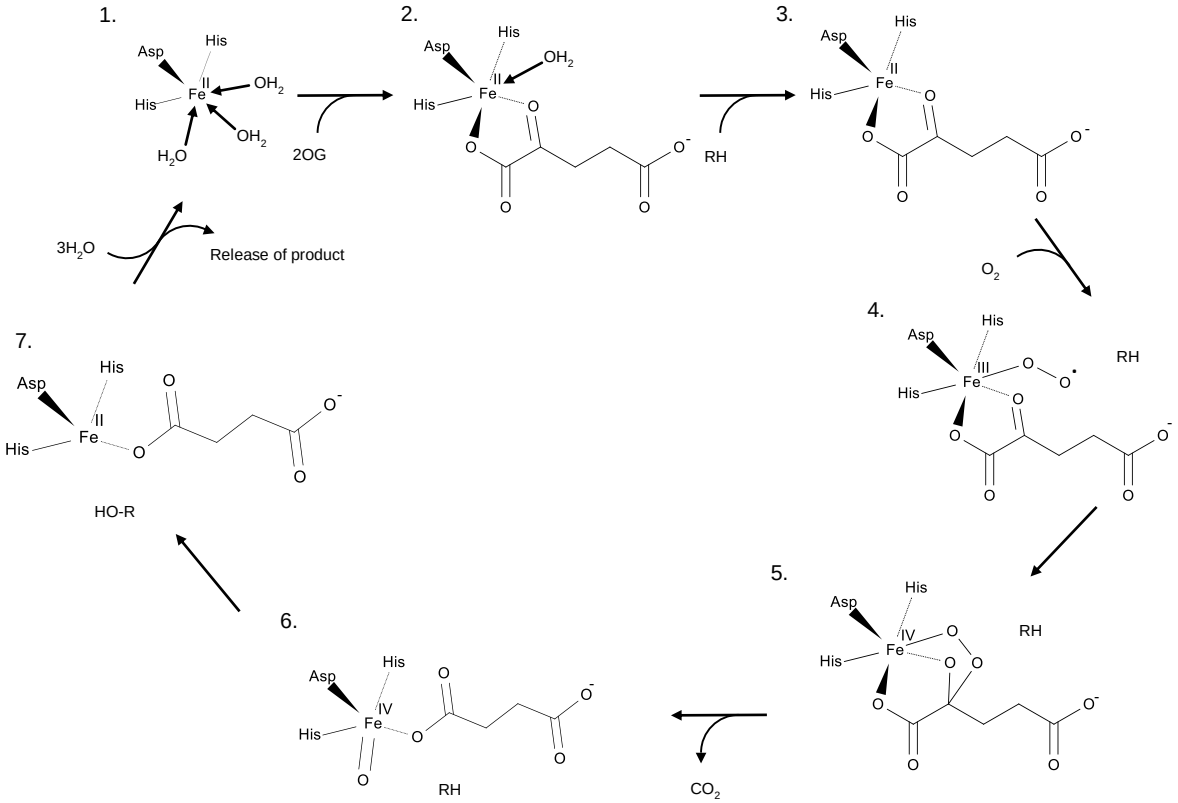



**Figure 5. Proposed catalytic mechanism for substrate hydroxylation by 2-oxoglutarate- and ferrous iron-dependent oxygenases** [145]**.** The $Fe^{2+}$ ion is coordinated by the conserved HxE/DxH motif. Steps of hydroxylation are described in the main text. Step 7 shows hydroxylated substrate prior to release and reaction by-product succinate bound to the $Fe^{2+}$ ion.

## 1.6. Inhibitors of JmjC-type lysine demethylases and hydroxylases

Analysis of crystal structures of human JmjC-type oxygenases highlights a highly conserved 2OG and Fe(II) binding site, however it also points to sub-family specific differences around the substrate binding site and to a lesser extent within the 2OG pocket itself [146]. The catalytic activity of 2OG oxygenases depend on the presence of the cosubstrates 2OG and oxygen, cofactor $Fe^{2+}$, accessibility of the substrate and is often stimulated by ascorbate and other reducing agents [147]. Multiple approaches of inhibition have been successfully demonstrated including mimicking the peptide substrate or the cosubstrate 2OG [9, 10][149]. Many of the 2OG oxygenases are multi-domain proteins, which in turn allows for indirect activity modulation, such as inhibition of JMJD2A catalytic activity by ejection of structural Zn in the accessory zinc binding site [150]. The status of inhibition studies is briefly summarised below, as in reviews [28], [146] and [148] which provide a comprehensive description on inhibition studies of 2OG oxygenases.

**Inhibition by natural products**

2OG oxygenases are inhibited by several natural products including flavonoids (e.g. quercetin) and various catechols (such as: 3,4-dihydrobenzoic acid, epinephrine, gallic acid and catechol) [148]. These compounds may inhibit 2OG oxygenases by sequestering iron in solution, removing iron from the active site or by binding the iron competitively with the cosubstrate 2OG or the substrate. Many of the flavonoids, such as quercetin are known to be “non-specific” inhibitors [151] for other enzyme classes including oxidoreductases. At present there are no reported crystal structures with either catechols or flavonoids with JmjC-type demethylases.

### Inhibition by metal ions

2OG oxygenases are inhibited by divalent transition metal ions, with $Mn^{2+}$, $Ni^{2+}$, $Co^{2+}$, $Cu^{2+}$ and $Zn^{2+}$ identified as the most potent inhibitors [148]. The mode of inhibition can be simply competition with $Fe^{2+}$ for the binding site, or degradation of the supplementary cofactor ascorbate as in the case of $Ni^{2+}$ and $Co^{2+}$ [152].

### Inhibition by TCA cycle intermediates and related metabolites

Tricarboxylic acid (TCA) cycle intermediates and metabolites thereof are known to inhibit 2OG oxygenases [146], [148] due to their structural similarity with 2OG [153]. The fact that oxygenases are regulated by citric acid cycle intermediates positions them on an interesting interface of nutrient availability and metabolism [153]. Nutrient balance can be altered by point mutations to enzymes involved in the TCA cycle. This is of particular biological relevance, since many cancers exhibit mutations impairing the activity of TCA cycle enzymes such as succinate dehydrogenase, fumarate hydratase and isocitrate dehydrogenase [148]. For example, gain of function point mutations in isocitrate dehydrogenase are frequently found in cancers [154] and divert the enzyme activity from native oxidation of the isocitrate substrate to the 2OG product to produce the 'oncometabolite' 2-(R)-2-hydroxyglutarate (D-2HG). 2HG is a chiral molecule with two enantiomers D and L. The L(S)-2-hydroxyglutarate (L-2HG) isomer is produced by reduction of 2OG by malate dehydrogenase and both enantiomers were found to inhibit oxygenases including enzymes in the TET family of 5-methylcytosine hydroxylases which also belong to the wider family of 2OG oxygenases [155], [156]. Interestingly L-2HG is produced under hypoxia were it was found to increase the levels of histone methylation, in line with its inhibitory activity on human oxygenases [157], [158]. The growing body of evidence suggests a role of TCA cycle intermediates and oncometabolites D- and L-2HG in regulation of metabolism in hypoxia partially through inhibition of oxygenases [159].

**Towards therapeutic inhibitors**

The generic inhibitor templates for 2OG oxygenases include 2OG mimetics such as N-oxalyl amino acids, where N-oxalylglycine (NOG) has been shown to be a pan-oxygenase inhibitor. Likewise pyridinedicarboxylates, such as 2,4-pyridinedicarboxylate (2,4-PDCA) (Figure 6) demonstrate a widespread inhibition pattern across the 2OG oxygenases [148] [160].

As mentioned, N-oxalylglycine, is a broad spectrum inhibitor of 2OG- and iron- dependent oxygenases with an $IC_{50}$ in the low micromolar range (Figure 6). It was first described as a prolyl hydroxylase (PHD) inhibitor and since then proven to inhibit the majority of demethylases [160], [161]. 2, 4-pyridine dicarboxylic acid (2,4-PDCA), a more rigid 2OG analogue, is another broad spectrum inhibitor of 2OG oxygenases. It was also identified as an inhibitor of prolyl-4-hydroxylases [160], is a 2OG competitive inhibitor, which binds to the metal through the pyridine nitrogen and carboxylic acid oxygen as demonstrated through multiple crystal structures [160]. Both NOG and 2,4-PDCA, although potent in *in vitro* assays contain more than one highly polar carboxylic acid moieties which reduce their cellular permeability [151]. Indeed cell permeable methyl-ester pro-drug derivatives of 2,4-PDCA and NOG – namely dimethyl-2,4-PDCA and dimethyloxalylglycine (DMOG) were used as tools to observe inhibition of demethylation *ex vivo* [151]. Screening of a set of 2OG analogues identified daminozide as a potent inhibitor of the KDM2 (FBXL11 $IC_{50}$ of 1.5 μM) and KDM7 (PHF8 $IC_{50}$ of 0.55 μM) families (Figure 6). Daminozide is a herbicide, the use of which is now discontinued in agriculture due to possible carcinogenic effects [148]. Daminozide chelates the active site metal through the dimethylamino nitrogen and the C-4 carbonyl group as identified in crystal structure complexes (PDB ID:4DO0 in complex with PHF8, PDB ID:4AI8 in complex with FIH and PDB ID:4AI9 in complex with JMJD2A) [146].

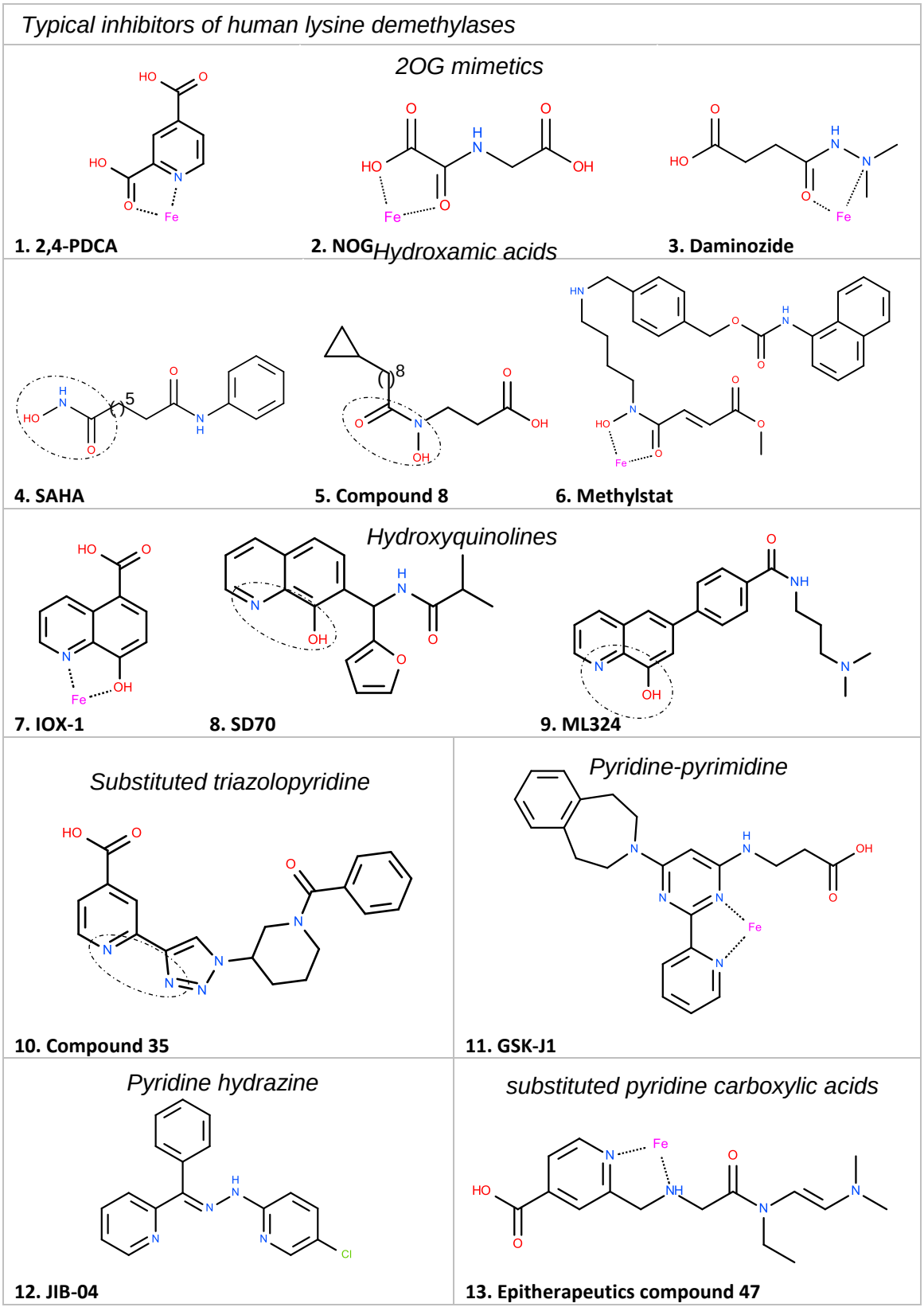



**Figure 6. Representative selection of histone lysine demethylase inhibitors, not including natural products and peptide mimetics.** The catalytic iron is depicted whenever a crystal structure of the complex was available. For the remaining compounds a dotted ellipse shows the expected metal binding motif.

**Hydroxamic acids**

Early studies observed inhibition of 2OG oxygenases by hydroxamic acids, with SAHA (a nM inhibitor for histone deacetylases) displaying 10-20 μM $IC_{50}$ for JMJD2E (Figure 6) [160]. Hydroxamic acids are known to be good metal chelating groups and indeed design of inhibitors containing the hydroxamic acid motif has been carried out as in the case of methylstat (Figure 6). Methylstat was developed with a strategy of mimicking both the cosubstrate 2OG and the peptide, by carbon-linking the carboxyhydroxamic metal binding scaffold with a protonated lysine mimetic [162]. It was demonstrated that methylstat inhibits JMJD2A/C/E with low micromolar $IC_{50}$, and shows some degree of selectivity with 30-50 μM $IC_{50}$ values for PHD1-3 and 22 μM inhibition of FIH, without inhibiting LSD1 and HDACs [162]. In order to achieve optimal cellular activity an ester prodrug strategy was adopted for methylstat. Another interesting example of the hydroxamic acid based scaffold is compound 8 (Figure 6) resulting in inhibitors of JMJD2A/C with low micromolar $IC_{50}$ values [163].

**Hydroxyquinolines**

The hydroxyquinoline scaffold was first identified as a hit from a high throughput screen of 230,000 compounds against JMJD2E [151]. Subsequently 5-carboxy-8-hydroxyquinoline (IOX-1) was identified as a broad-spectrum 2OG oxygenase inhibitor shown to inhibit activity of other demethylases with *in vitro* $IC_{50}$ ranging from sub-micromolar (JMJD2A/C/D/E, JMJD1A, JMJD3) to low micromolar concentrations (UTX, KDM2A, PHF8, JARID1B, FIH and PHD2) (Figure 6) [161]. Further studies have shown that IOX-1 has a cellular $EC_{50}$ of 80-100 μM, which can be further optimised with its n-octyl ester form to show inhibition of demethylation with low micromolar $EC_{50}$ in cells [164]. IOX-1 chelates the active site metal through its pyridine nitrogen and phenolic hydroxyl groups in a bidentate manner, as observed in crystal complexes (JMJD3 PDB ID:2XXZ, FIH PDB ID:3OD4, 4BIO) [161]. Interestingly, IOX-1 induced a metal movement observed by

comparison to crystal structures without IOX-1 [161], suggesting a certain flexibility in the metal binding location.

ML324, a cell-permeable 6-substituted quinolin-8-ol, inhibits JMJD2E with an $IC_{50}$ of 920 nM and also shows potent anti-viral activity against herpes simplex virus and human cytomegalovirus via inhibition of viral gene expression (Figure 6) [165]. This suggests the potential of epigenetic therapies for treatment of infectious diseases. SD70, another hydroxyquinoline derivative, was initially discovered as an inhibitor of dihydrotestosterone (DHT)-induced androgen receptor-mediated gene translocation in prostate cancer (Figure 6) [166]. JMJD2C was confirmed as its target through Chem-seq, a method to determine compound targets within chromatin. In Chem-seq cells are treated with biotinylated compound, which binds target chromatin interacting protein, the complex is then cross-linked to the DNA by treatment with formaldehyde, followed by the compound capture with streptavidin beads; elution and de-cross-linking isolates bound DNA fragments which are then sequenced [166], [167]. *In vitro* demethylase assays for JMJD2C show an $IC_{50}$ for SD70 of approximately 30 µM [166] and 293T cells show increased H3K9me2 levels upon treatment with SD70 [166]. SD70 inhibits prostate cancer growth *in vitro* (CWR22Rv1 cells) and tumour growth *in vivo* (CWR22Rv1 cell xenograft mouse model) [166].

**Substituted pyridine-pyrimidine compounds**

A variation of the pyridine-pyrimidine was used in the development of GSK-J1, which initially was described as a H3K27me3/me2 selective demethylase chemical tool by inhibiting the activities of JMJD3 and UTX (Figure 6) [77]. GSK-J1 inhibits JMJD3 with an *in vitro* $IC_{50}$ value of 60 nM in the alpha screen assay [77]. GSK-J1 is often used in conjunction with its inactive regioisomer form GSK-J2. Since the cellular permeability of GSK-J1 is restricted by its polar carboxylic acid group, the pro-drug ethyl esters GSK-J4 and GSK-J5 (for the GSK-2 compound) have been developed [77]. GSK-J4 inhibits H3K27 demethylation in cells at concentrations of 1-10 µM [77] and it was also proven to be active

in glioma xenografts at a dose of 100 mg/kg/day [168] [77]. Later studies on GSK-J1 selectivity *in vitro* showed that it inhibits JARID1B/A only 5-10-fold weaker than the targeted JMJD3 and UTX [169]. The cell-active form – GSK-J4 – was also shown to inhibit JMJD3, JARID1B and JMJD2C in cells with similar $IC_{50}$ values [169]. Nevertheless GSK-J1 remains one of the most selective demethylase inhibitors.

**Substituted pyridine carboxylic acids**

One of the most potent lysine demethylase inhibitors has been developed by Epitherapeutics - EPT-103182 [28], [170]. EPT-103182 inhibits JARID1B in the low nanomolar range with an $IC_{50}$ of 1.8 nM for H3K4me3 accumulation in U2OS cells [28]. It has also been shown to be 20-50-fold selective over the JMJD2 family and 3000-fold selective over JMJD3/UTX/UTY [28]. The *in vitro* activity of EPT-103182, the structure of which has not yet been disclosed, translated to good *in-vivo* efficacy in multiple myeloma xenografts [170]. The patent describing the Epitherapeutic series (Figure 6) [170] has been explored by multiple academic groups and pharmaceutical companies leading to the development of KDOAM25 which inhibits the JARID family of demethylases with an i*n vitro* $IC_{50}$ value in the nanomolar range [Dr A. Tumber, personal communication]. KDOAM25 is an analogue of Epitherapeutics compound 47 presented in Figure 6, with a carboxamide group replacing the carboxylic acid.

**Substituted triazolopyridine compounds**

Compound 35 is an FBXL11 inhibitor based on a substituted triazolopyridine scaffold with reported *in vitro* potency $IC_{50}$ of 63 nM with excellent selectivity over members of JMJD1, JMJD2 families and JMJD3 (Figure 6). Cellular activity, although not reported, is expected to require a pro-drug form due to the presence of a highly polar carboxylic acid moiety [171].

### Pyridine hydrazine compounds

JIB-04 is a pyridine-hydrazine based compound which was initially identified in a cellular screen of epigenetic modifiers and has shown a sub-micromolar inhibition for lysine demethylases in *in vitro* assays (Figure 6) [172]. The *in vitro* activity of demethylases was measured by an ELISA assay at 150 nM concentration of iron. In fact JIB-04 was determined to be not competitive with 2OG, but competitive with iron [172]. This was independently verified in a RapidFire mass spectrometry activity assay with lysine demethylases where compound inhibition was dependent on the concentration of active iron [Dr A. Tumber, personal communication]. Interestingly JIB-04 has shown to inhibit demethylase activity in cancer cells as well as in tumours *in vivo*, prolonging survival in mice bearing 4T1 mammary tumours [172].

### Benzisothiazole scaffold

The benzisothiazole based compound 2-4(4-Methylphenyl)-1,2-bezisothiazol-3(2H)-one (PBIT) was identified as a JARID inhibitor through a high throughput screen of 15,134 compounds (Figure 7) [173]. The *in vitro* $IC_{50}$ value was reported to be 3 µM for JMJD1B, 6 µM for JARID1A and 4.9 µM for JARID1C, with no inhibition of UTY or JMJD3 [173]. HeLa cells overexpressing JARID1B treated with PBIT show increased level of H3K4me3, suggesting inhibition of JARID1B enzymatic activity [173]. Interestingly, treatment with 10 µM PBIT of human breast cancer UACC-812, a cell line with high JARID1B expression, promoted apoptosis of the cells [173]. The same treatment of MCF7 human breast cancer cells with low JARID1B expression showed minimal toxicity [173]. Similar effects were observed by downregulating JARID1B in these cell lines with shRNA [173]. A close analogue of PBIT,

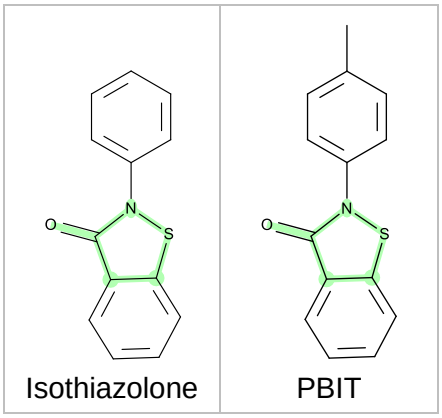



**Figure 7. Structures of compound containing the isothiazolone motif identified as PAIN (left) and PBIT (right) are near identical.**

Ebselen, exhibits similar *in vitro* activity and was shown to inhibit JMJD2A activity by ejection of the structural Zn [150], suggesting a possible mechanism of action for these compounds. However, isothiazolones, the scaffold represented by PBIT, have been reported to be Pan Assay Interference Compounds (PAINS), active largely through unspecific covalent modifications to proteins [174]. PAINS are structural motifs which frequently emerge as hits in biophysical assays, but their activity is largely attributed to an unspecific assay interference type effect [175]. In light of these reports it is likely that the reported activity of PBIT is not specific to demethylases, but much broader than previously thought.

**Diazepin-quinazoline-amine derivative**

BIX-01294 (Figure 8) was identified in a high throughput screen as H3K9 methyltransferase G9a inhibitor with *in-vivo* activity [176]. BIX-01294 was also characterised as an active inhibitor of KIAA1718 with an *in-vitro* $IC_{50}$ of 16.5 µM and its close analogue E67 with an $IC_{50}$ of 3.7 µM with no effect on activity of JARID1C [177]. The binding mode of compound E67 was confirmed by a crystal structure with KIAA1718 (PDB ID: 3KV4).

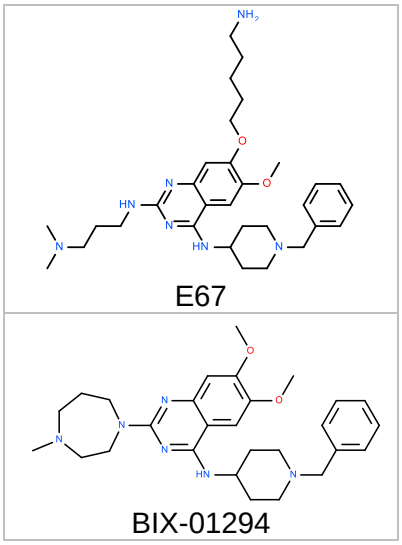



**Figure 8. Compound BIX-01294 identified as histone methyltransferase inhibitor and its analogue E67.**

**Crystal structures of inhibitor complexes**

The family of JmjC-containing oxygenases is well described by X-ray and NMR structural studies, thanks to the efforts of multiple scientific groups over the last 10 years, however the number of deposited ligand complexes remains relatively low. At the date of thesis writing (July 2015), the Protein Data Bank (PDB, http://www.rcsb.org/) [178] contains 68

high resolution X-ray crystal structures of the human *JmjC* domain proteins[1]. In terms of ligands, it can be deduced from these 68 structures that there are 19 small molecule binders present when excluding solvent molecules, cryoprotectants and crystallisation condition ligands. Within the set of 19 small molecule ligands, 9 represent 2OG, daminozide and 2,4-PDCA, TCA cycle intermediates (succinate, L and D 2HG), N-oxalyl amino acids (N-oxalylglycine (NOG), N-oxalylcysteine and N-oxyalyl-D-tyrosine), two 'fragment like' compounds (5,6-dimethylbenzimidazole and O-toluenesulfonamide) and the remaining 8 are published inhibitors of JmjC demethylases including hydroxyquinolines, GSK-J1, triazolopyridine, bipyridyls, the Epitherapeutic compound 47 and inhibitor of KIAA1718 compound BIX-01294.

The above review suggests that novel chemical scaffolds and/or structural information are needed to accelerate the development of more potent and selective chemical tools for the family of JmjC-type oxygenases, especially with the view to validate several members for therapeutic candidate development.

## 1.7. Aim of the thesis

We first aim to characterise the family of JmjC-type oxygenases with a library of fragments in activity assays and X-ray crystallography in order to identify novel hit matter for further chemical optimisation. Secondly we aim to establish a platform for drug discovery for two poorly characterised ribosomal oxygenases, NO66 and MINA53, consisting of activity assays and X-ray crystallography, and to identify small molecule inhibitors of the proteins. Thirdly, by using combined activity profiling and X-ray structural data collection, we attempt development of active inhibitor molecules for members of the human JmjC-type oxygenases.

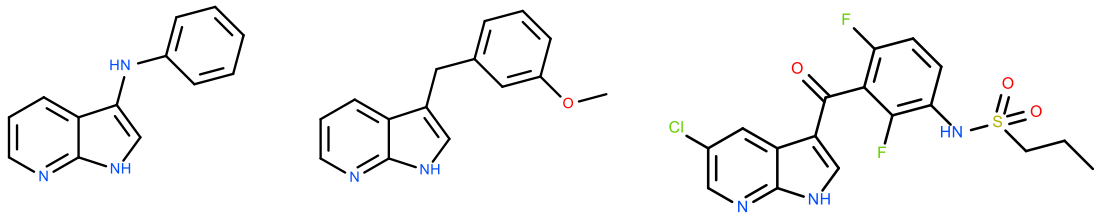
[1] Query was performed on the 28th of July 2015 in the PDB website (http://www.rcsb.org/pdb/home/home.do) with PFAM (http://pfam.xfam.org/) accession code PF02373, which represents family of proteins containing JmjC domain.

# 2. Identification and characterisation of inhibitor lead compounds using activity fragment screening and X-ray crystallography

## 2.1. Introduction

In order to identify novel starting points for inhibition of JmjC-type oxygenases we utilised a fragment-based drug discovery strategy, a method that is increasingly popular in the development of novel lead compounds [179]. The principle of the technique relies on the premise that fragments, with a molecular weight typically less than 300 Da, explore the chemical space more efficiently than libraries found in conventional high-throughput screens (HTS) [180]. Structural information derived from these diverse sets of small molecules can be combined to produce more potent inhibitors of the target. The chemical optimisation of identified hits is usually done through fragment linking or fragment building strategies, with the latter proving more successful [181].

There are more than a dozen successful fragment-based campaigns leading to clinical molecules to date [179], with the first molecule originating from a fragment based screen (PLX4032; vemurafenib) and being approved by the FDA in 2011 for the treatment of patients with unresectable or metastatic melanoma with the $BRAF^{V600E}$ mutation [182], [183]. The development of PLX4032 started with an initial screen of 20,000 compounds using an AlphaScreen based activity assay and resulted in a set of 238 hits of which 100 were co-crystallised leading to the identification of the 3-aminophenyl-7-azaindole scaffold bound to the ATP-binding site of the BRAF kinase domain. Creation of the final compound PLX4032 was achieved through iterative cycles of co-crystal structures with follow-up compounds (Figure 9). Other molecules derived from fragment-based projects in clinical trials are described in the recent review by Murray *et al*. [179].




3-aminophenyl-7-azaindole | 3-(3-methoxybenzyl)-7-azaindole | PLX4720

**Figure 9. Structures of compounds leading to development of PLX4720.**

From the multiple techniques developed to identify weakly binding fragment hits, nuclear magnetic resonance (NMR) methods developed at Abbott Laboratories [184] and X-ray crystallography were applied first [185]. More recently surface plasmon resonance (SPR), differential scanning fluorimetry (DSF), isothermal titration calorimetry (ITC) or functional screening have also been used [186]. X-ray fragment screening is often considered a low-throughput and resource-intensive method, however recent technological developments including robot assisted crystal mounting, fast detectors and automated data processing alleviate those hurdles [187].

In this work we first screened fragments for inhibition with several subfamilies of human JmjC-type oxygenases in order to identify novel inhibitor classes and prioritise fragments for X-ray crystallography. Once the structures of hits were identified we have determined the $IC_{50}$ values for the compounds in order to progress hits for further lead development. We have used two libraries: a subset of the 3D fragment library (182 fragments) from the 3D Fragment Consortium [188] and a subset of the Maybridge Rule of 3 (Ro3) Fragment Library (581 fragments) [189]. The 3D fragment library was developed to enhance the 3D characteristics of the compounds increasing the diversity of the set [188] whereas the focus of the Maybridge Ro3 Fragment Library was in compliance with the 'Rule of 3' for fragments [190]. 'Rule of 3' was defined by retrospective analysis of fragment hits on a range of targets at Astex Pharmaceuticals where the hits obeyed chemical constraints: molecular weight of <300 Da, number of hydrogen bond donors ≤ 3 and number of hydrogen bond acceptors ≤3 and ClogP ≤3.

## 2.2. JmjC family fragment inhibition screening

We screened the Maybridge library and the 3D fragment library in an AlphaScreen activity assay with representatives of KDM2-6 demethylases (FBXL11, JMJD1B, JARID1B, JMJD2A/D, JMJD3) at a concentration of 250 μM. AlphaScreen technology [191] relies on detection of the product methylation state by an antibody through chemiluminescence. For the ribosomal hydroxylases (NO66 and MINA53) we utilised the RapidFire mass spectrometry assay at a concentration of 250 μM for the 3D fragment library and 100 μM for the Maybridge library. Lower concentration of 100 μM was chosen for the Maybridge library in order to avoid detector saturation observed with some fragments from the 3D fragment set screened at 250 μM concentration. Compounds in both fragment libraries were solubilised in DMSO.

### 2.2.1. 3D fragment library

The heatmap displaying the results of the activity profiling of the 3D fragment screen is shown in Figure 10.

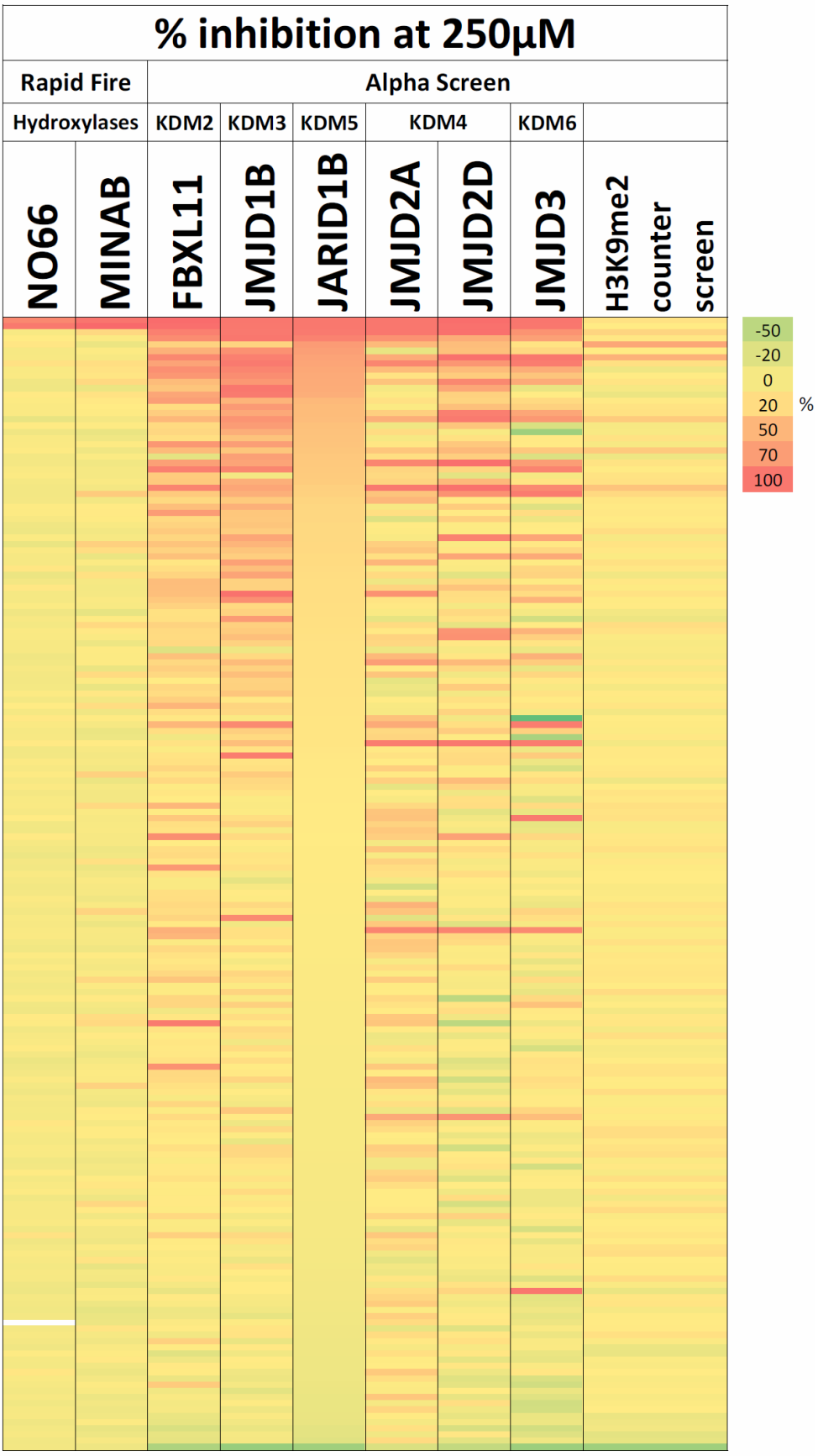



**Figure 10. Heatmap of activity screen with JmjC-type enzymes against compounds in the 3D fragment library screened at concentration of 250 µM, n=1. Inhibition is coloured from red: 100% to green: -50%.**

RapidFire screening with MINA53 and NO66 (see 3.4 for assay development) resulted in a visibly lower hit rate of 1.1% for both enzymes in comparison to hit rates in the AlphaScreen assay of 9.3-19.8% for KDM2-6 enzymes (Table 3). In order to eliminate assay interference compounds in the AlphaScreen-based assay we have used a counter screen with a H3K9me2 peptide which is the product of the activity of JMJD2A and JMJD2D. This strategy allowed the successful elimination of false-positive hits which non-specifically interfere in the counter screen assay. Based on the results of the counter screen false positives constituted 13-24% of the identified hits (Table 3).

| | **3D Fragment Library Activity Assay (% inhibition)** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **RapidFire (250 µM)** | | **AlphaScreen (250 µM)** | | | | | | |
| | **MINA53** | **NO66** | **FBXL11** | **JMJD1B** | **JARID1B** | **JMJD2A** | **JMJD2D** | **JMJD3** | **H3K9me2 counter screen** |
| *Absolute Hit Rate* | | | | | | | | | |
| **Hit number** | 2 | 2 | 23 | 36 | 13 | 17 | 21 | 24 | 8* |
| **% Hit rate** | 1.1% | 1.1% | 12.6% | 19.8% | 7.1% | 9.3% | 11.5% | 13.2% | 4.1%* |
| *Number of False Positives* | | | | | | | | | |
| False positives (% of false positive hits) | - | - | 3 (13.0%) | 5 (13.9%) | 3 (23.1%) | 4 (23.5%) | 5 (23.8%) | 5 (20.8%) | 67 |
| *Hit Rate corrected for false positives* | | | | | | | | | |
| **Hit number** | 2 | 2 | 20 | 31 | 10 | 13 | 16 | 19 | |
| **% Hit rate** | 1.1% | 1.1% | 11.5% | 17.8% | 5.7% | 7.5% | 9.2% | 10.9% | |
| | | | | | | | | | |
| **Total number of fragments screened** | 182 | | | | | | | | |

**Table 3. Success-rate and false-positive rate of the activity screen with the 3D fragment library. A t**otal of 182 compounds were screened and a "hit" is defined as a compound with >50% inhibition. All enzymes were screened at a single concentration of 250 µM. *False positive is assigned for >20% inhibition in the H3K9me2 counter screen.

Typical selectivity profiles identified by the activity profiling are shown in Table 4. Compounds N10075a and N10042a have shown complete or near complete inhibition of all tested enzymes with less than 8% and 16% inhibition in the counter screen respectively.

N10042a is the most potent inhibitor with broad inhibition of JmjC-type oxygenases and an $IC_{50}$ of 5.9 µM against JARID1B (Table 4). Compounds N09952a, N10072a and N10003a are of interest because of their apparent selectivity. N09952a selectively inhibits JMJD3, N10072a shows complete inhibition of JMJD1B and 40-70% inhibition of FBXL11, JARID1B and JMJD2D, and N10003a is selectively inhibiting FBXL11.

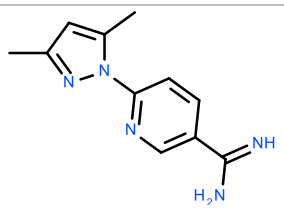
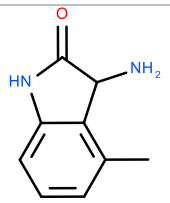
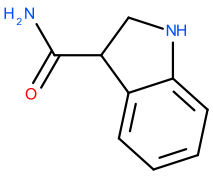
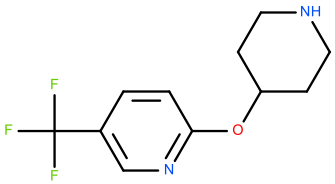
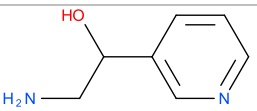
| | % inhibition at 250 µM | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | RapidFire | | Alpha Screen | | | | | | |
| | NO66 | MINA53 | FBXL11 | JMJD1B | JARID1B | JMJD2A | JMJD2D | JMJD3 | H3K9me2 counter screen |
|  N10075a | 99 | 111 | 106 | 100 | 99 | 100 | 105 | 101 | 8 |
|  N10042a | 86 | 102 | 109 | 100 | 100 | 100 | 107 | 101 | 16 |
|  N09952a | -7 | 4 | -10 | 9 | -2 | 18 | 25 | 100 | -9 |
|  N10072a | -6 | -8 | 38 | 100 | 58 | 0 | 62 | -11 | 0 |
|  N10003a | -3 | -8 | 74 | 18 | 7 | 20 | 4 | 4 | 7 |

**Table 4. Selected hits from the screening of the 3D fragment library show broad spectrum inhibition (N10075a, N10042a) and a high degree of apparent selectivity (N09952a, N10072a, N10003a).**

### 2.2.2. Maybridge library

The hit rates of the activity screen with compounds in the Maybridge library are shown in Table 5. MINA53 and NO66 show a hit rate of 0.9% and 0.7%, respectively, similar to the one obtained for the 3D fragment library, despite screening at a lower compound concentration of 100 µM. Enzymes in the KDM2-6 families present a hit rate of 8.8-12.6% (Table 5). 11 compounds inhibited the KDM2-6 enzymes in the AlphaScreen and only one compound inhibited all of the tested enzymes (see compound N05798b in Table 7).

| | Maybridge Library Activity Assay (% inhibition) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | RapidFire (100 µM) | | AlphaScreen (250 µM) | | | | | | |
| | MINA53 | NO66 | FBXL11 | JMJD1B | JARID1B | JMJD2A | JMJD2D | JMJD3 | H3K9me2 counter screen |
| *Absolute Hit Rate* | | | | | | | | | |
| **Hit number** | 5 | 4 | 94 | 105 | 85 | 100 | 91 | 79 | 67* |
| **% Hit rate** | 0.9% | 0.7% | 16.2% | 18.1% | 14.6% | 17.2% | 15.7% | 13.6% | 11.5%* |
| *Number of False Positives* | | | | | | | | | |
| False positives (% of false positive hits) | - | - | 43 (45.7%) | 40 (38.1%) | 36 (42.4%) | 40 (40.0%) | 41 (45.1%) | 34 (43.0%) | 67 |
| *Hit Rate corrected for false positives* | | | | | | | | | |
| **Hit number** | 5 | 4 | 51 | 65 | 49 | 60 | 50 | 45 | |
| **% Hit rate** | 0.9% | 0.7% | 9.9% | 12.6% | 9.5% | 11.7% | 9.7% | 8.8% | |
| | | | | | | | | | |
| **Total number of fragments screened** | 581 | | | | | | | | |

**Table 5. Success rate and false positive rate of the activity screen with the Maybridge fragment library.** Hits were defined as inhibition > 50% and *false positive is assigned for >20% inhibition in the H3K9me2 counter screen.

The number of screened compounds allowed an assessment of the selectivity profile of the inhibitors between the protein classes. Pairwise enzyme comparison presented in Table 6 indicates compound-cross reactivity within the library with other tested JmjC-type oxygenases. Interestingly, JMJD2A and JMJD2D have 40 overlapping hits which constitute a high percentage of JMJD2D (80%) and JMJD2A (67%) hits. JMJD1B shares 57% of JARID1B hits and JARID1B shares 76% of JMJD1B hits. Hits of JMJD3 appear to be most distinct from those identified for other JmjC-type oxygenases showing a maximal overlap of 47% (Table 6).

| | | RapidFire (100 µM) | | AlphaScreen (250 µM) | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MINA53 | NO66 | FBXL11 | JMJD1B | JARID1B | JMJD2A | JMJD2D | JMJD3A |
| RapidFire (100 µM) | MINA53 | 4 (100%) | 2 (67%) | 2 (4%) | 3 (5%) | 3 (6%) | 3 (5%) | 2 (4%) | 1 (2%) |
| | NO66 | 2 (50%) | 3 (100%) | 2 (4%) | 2 (3%) | 2 (4%) | 3 (5%) | 1 (2%) | 1 (2%) |
| AlphaScreen (250 µM) | FBXL11 | 2 (50%) | 2 (67%) | 51 (100%) | 25 (38%) | 20 (41%) | 22 (37%) | 17 (34%) | 17 (38%) |
| | JMJD1B | 3 (75%) | 2 (67%) | 25 (49%) | 65 (100%) | 37 (76%) | 35 (58%) | 24 (48%) | 21 (47%) |
| | JARID1B | 3 (75%) | 2 (67%) | 20 (39%) | 37 (57%) | 49 (100%) | 29 (48%) | 22 (44%) | 18 (40%) |
| | JMJD2A | 3 (75%) | 3 (100%) | 22 (43%) | 35 (54%) | 29 (59%) | 60 (100%) | 40 (80%) | 20 (44%) |
| | JMJD2D | 2 (50%) | 1 (33%) | 17 (33%) | 24 (37%) | 22 (45%) | 40 (67%) | 50 (100%) | 18 (40%) |
| | JMJD3A | 1 (25%) | 1 (33%) | 17 (33%) | 21 (32%) | 18 (37%) | 20 (33%) | 18 (36%) | 45 (100%) |

**Table 6. Number of compounds inhibiting a pair of JmjC-type oxygenases and the percentage hit overlap shown in brackets normalised for each column/enzyme.** For example 37 hits are shared between JARID1B and JMJD1B, which comprises 57% of all JMJD1B hits and 76% of all JARID1B hits. The total number of identified hits for each enzyme is shown on the diagonal. The table is coloured according to the number of hits.

Table 7 shows examples of compounds (such as N05798b) inhibiting all tested enzymes, compounds selective for KDM2-6 demethylases such as N05859b, or compounds displaying a distinct selectivity profile such as N11228a or N06169b.

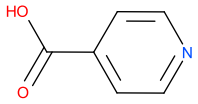
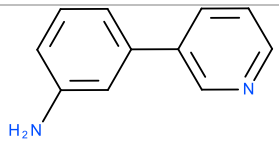
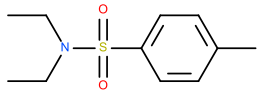
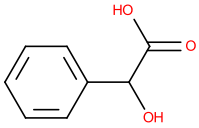
| | % inhibition at 250 µM | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | RapidFire | | Alpha Screen | | | | | | |
| | NO66 | MINA53 | FBXL11 | JMJD1B | JARID1BA | JMJD2A | JMJD2D | JMJD3 | H3K9me2 counter screen |
|  N05859b | 5 | -7 | 108 | 99 | 99 | 104 | 104 | 72 | -8 |
|  N05798b | 93 | 71 | 107 | 100 | 100 | 99 | 106 | 101 | 6 |
|  N11228a | 2 | -4 | 16 | 98 | 99 | -19 | -4 | 37 | -5 |
|  N06169b | -2 | 13 | 27 | -1 | 34 | 81 | 100 | 74 | -13 |

**Table 7. Examples of activity profiles for fragments derived from the Maybridge library.**

## 2.3. Structure determination of protein bound fragments using soaking

### 2.3.1. JARID1B as a fragment soaking system

In order to determine the binding mode of the prioritised fragments (such as those in Table 4 and Table 7) we employed a JARID1B crystallisation system that reproducibly grew well diffracting apo-crystals to 1.9-2.2 Å resolution in the *P*6$_5$22 space group. Typically 30-100 crystals per plate were grown in HEPES buffer supplemented with potassium phosphate after 3-5 days and were cryoprotected with 25% (v/v) ethylene glycol. The crystals were then soaked either manually or by using a semi-automated protocol developed at the I04-1 beamline

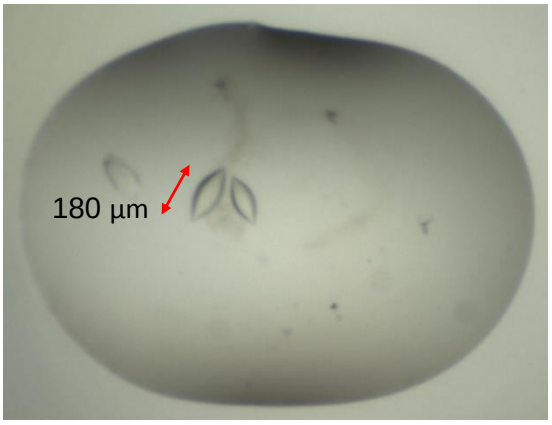



**Figure 11. Typical crystals of JARID1B grown at 4˚C in 0.1 M HEPES pH 7.5, 0.8 M potassium phosphate dibasic, 0.8 M sodium phosphate monobasic and 2 mM $MnCl_2$ in 300 nl drops with 1:2 (v/v) protein (9 mg/ml) to precipitant ratio.**

(Diamond synchrotron). Details on data collection and refinement can be found in supplementary section C.

Semi-automated soaking was performed using an Echo acoustic dispenser (Labcyte) which allowed for targeted dispensing of nanoliter volumes of compound and cryoprotectant solutions directly to crystals grown in Swiss-CI (Hampton Research) plates. Following a 1-2 h incubation period, crystals were harvested and flash-cooled using a robot-assisted method which allows for 150 individual crystals to be harvested within 1 hour. In order to optimise the soaking procedure we determined the maximum tolerated concentration of DMSO by incubating the crystals with final concentrations of 0%, 10%, 20%, 30% and 40% (v/v) DMSO. Crystals with concentrations of 30-40% DMSO dissolved, whereas those incubated with concentration of DMSO less than 20% diffracted with a resolution limit of 2.1-2.2 Å. Since both fragment libraries were available in concentrations of 250 mM and 200 mM (3D and Maybridge library, respectively), we decided to limit the final DMSO concentration to 5%, resulting in a final compound concentration of 10-12.5 mM.

Manual soaking was performed by transferring the crystal into a series of drops, containing 25% (v/v) ethylene glycol in crystallisation buffer, with increasing concentrations of compound based on a 2-fold dilution starting from 100-500 μM to 10-25 mM, resulting in a maximum DMSO concentration of ~10%. This method allowed for gradual increases of both compound and DMSO concentrations in consecutive transfer steps and for control over individual compound-crystal behaviour.

### 2.3.2. Validation of the soaking system with TCA cycle intermediates

In order to validate the system we have soaked a set of TCA cycle intermediates which included 2OG, succinate, fumarate, malate, oxaloacetate, pyruvate, citrate and D-isocitrate and the metabolites D-2HG and L-2HG with JARID1B crystals. We obtained ligand structure complexes for 2-oxoglutarate, succinate, fumarate, acetate, malate, pyruvate, D- and L-2HG with the catalytic domain of JARID1B (Figure 12). The manual

method was used in this soaking as described above with a maximal compound concentration of 25 mM. TCA cycle intermediates were dissolved in water, whereas D- and L-2HG were prepared in DMSO.

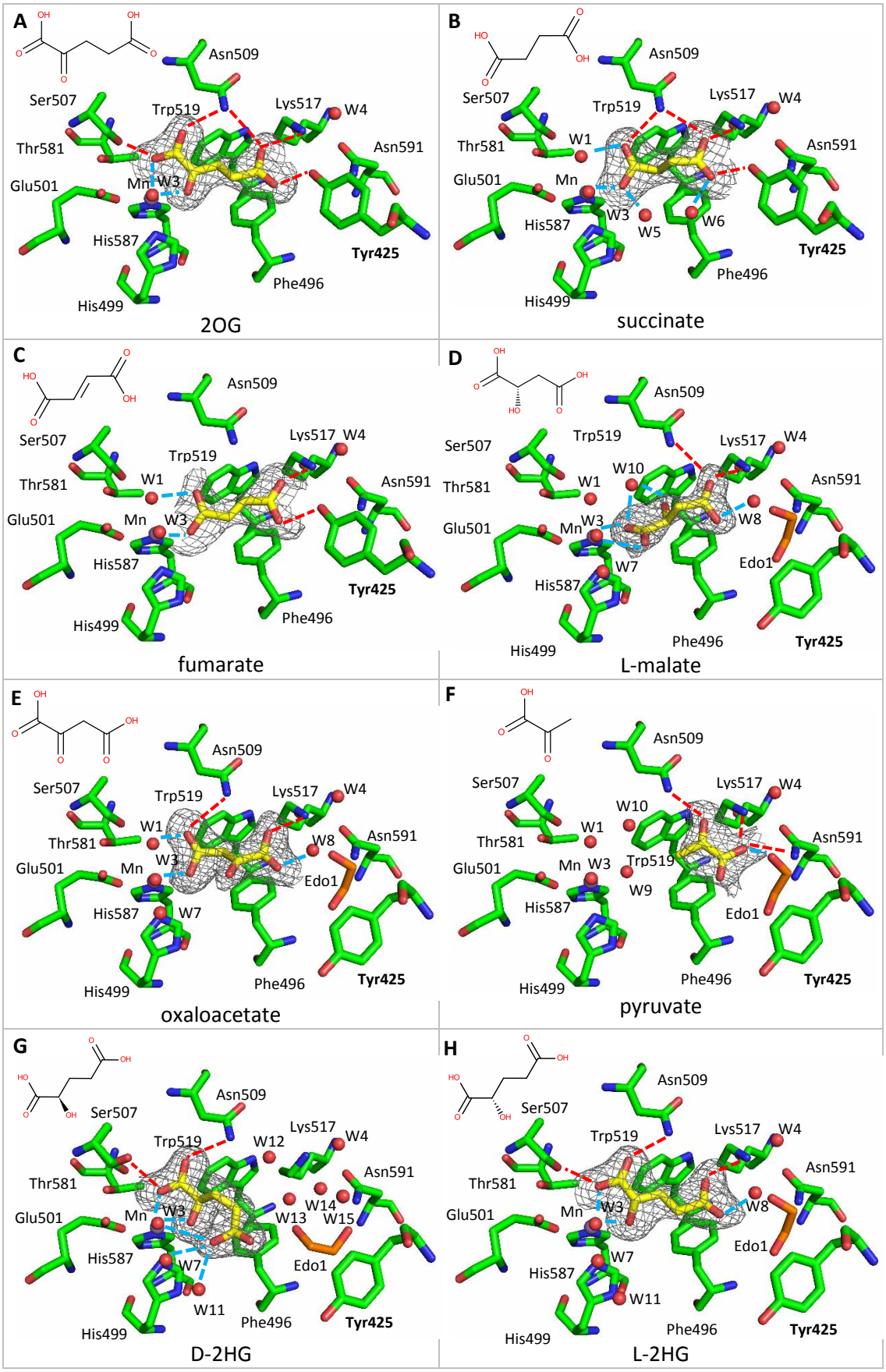



**Figure 12. Crystal structures of TCA cycle intermediates (A-F) and L-and D-2HG (G-H) (yellow) with JARID1B. Residue** Tyr425 alternates between two positions: facing the metal (2OG, succinate, fumarate)

and away from the metal (D-2HG, L-2HG, malate, pyruvate, acetate). The Fo-Fc OMIT density map for the ligand is contoured at 3σ. Water molecules are shown in red and ethylene glycol derived from the cryoprotectant in orange. Hydrogen bond interactions between ligand atoms and the residue side-chains are depicted as red dashed lines and solvent mediated interactions as cyan dashed lines. The cut-off distance for polar interactions is 3.2 Å.

The enzyme cosubstrate 2OG and D- and L-2HG chelate the metal in a bi-dentate manner with the C5 carboxylic acid and ketone oxygens (Figure 12 A). Whereas the other intermediates (Figure 12 B-E, G and H) chelate the metal in a mono-dentate manner. An exception is pyruvate (Figure 12 F), the smallest of the TCA cycle intermediates, where no direct interaction with the metal can be detected.

For 2OG, succinate and fumarate (Fig13. A-C) the carboxylic acid motif (C5 in 2OG) forms hydrogen bond interactions with $N^{\varepsilon}$ of Lys517 and the hydroxyl oxygen of Tyr425. The distance between the carboxylic acid oxygen atoms in fumarate and the carboxamide nitrogen of Asn509 is 3.8 Å, which indicates a weak polar interaction, instead fumarate is stabilised by hydrogen bond and electrostatic interactions detected between the carboxylic acid oxygen and with $N^{\varepsilon}$ of Lys517, the carboxamide nitrogen of Asn591 and ethylene glycol oxygen, and hydroxyl group with $N^{\varepsilon}$ of Lys517 and carboxamide nitrogen of Asn509.

In the crystal structures of L-malic acid, pyruvate, oxaloacetate, D-2HG and L-2HG (Figure 12 D-H) the Tyr425 side-chain/phenolic group points away from the catalytic site and its previous 2OG/succinate/fumarate position is occupied with ethylene glycol and set of ordered water molecules. The C5 carboxylic acid of L-2HG (Figure 12 H) places itself in the position of the C5 carboxylic acid of the 2OG, unlike the carboxylic acid of D-2HG (Fig13. G), which is stabilised by interactions with water molecules.

This study indicated a highly “plastic” active site in comparison to crystal structures of JmjC-type oxygenases in complex with TCA cycle intermediates such as succinate, fumarate or D/L-2HG [155], [192]. Importantly, the study allowed successful validation of the crystallography system and showed that small molecules could be successfully soaked into the 2OG-binding site of apo-crystals of human JARID1B.

### 2.3.3. Soaking of active fragments

We soaked JARID1B crystals with fragments of the Maybridge library showing >30% inhibition of any of the tested enzymes (262 out of a library of 581) and nearly all of the 3D fragments (169 out of a library of 182). We have screened most of the 3D fragments since the X-ray fragment screening allowed for a medium throughput soaking and data collection, whereas Maybridge fragments were stratified to enrich for potential inhibitors.

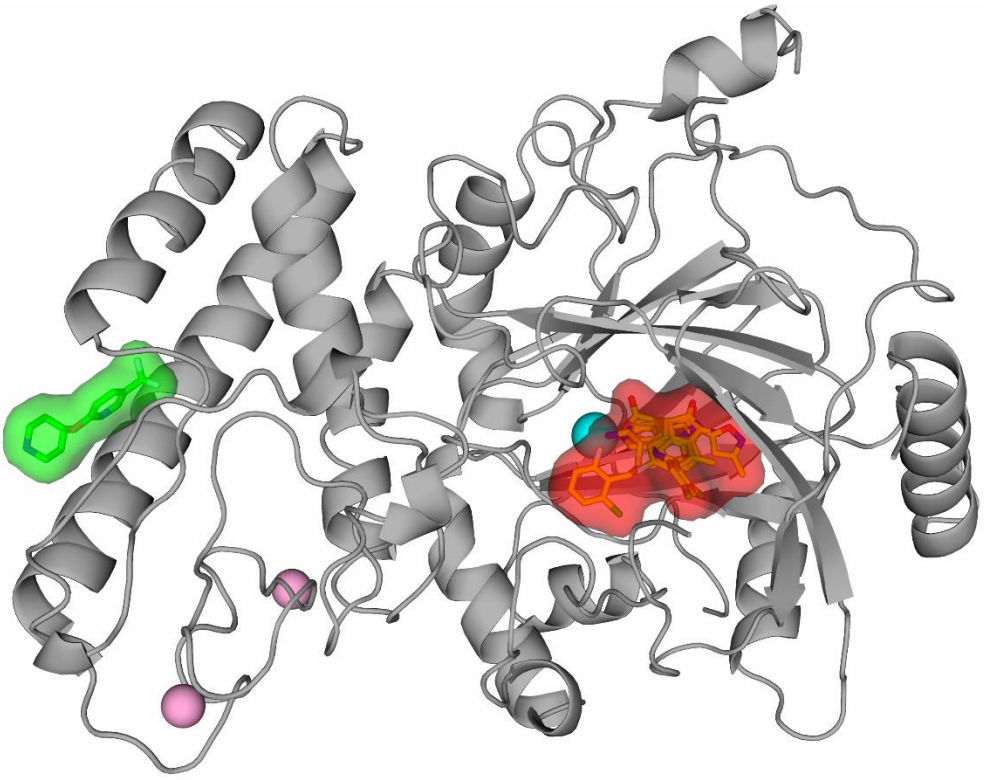
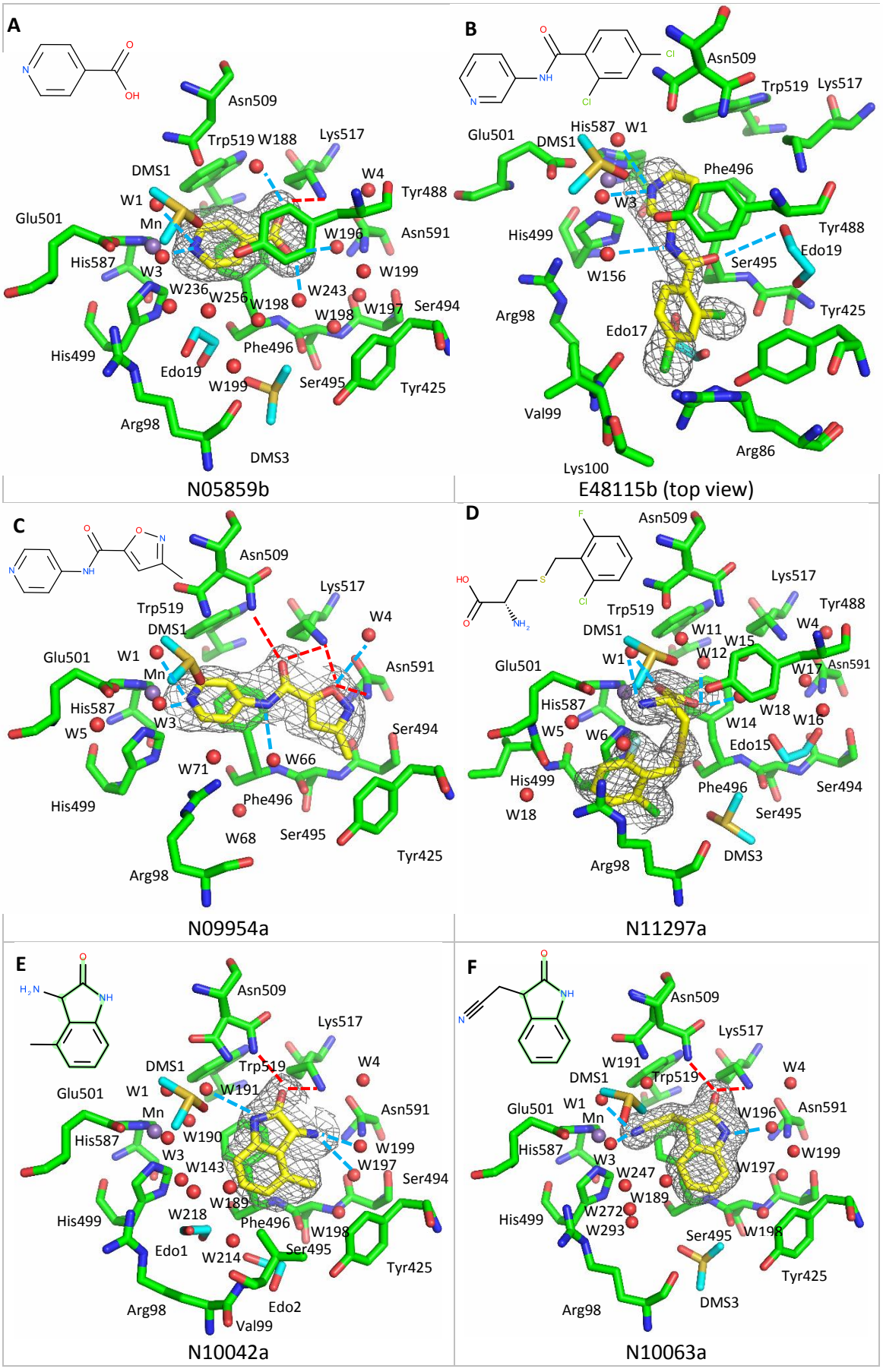


**Figure 13. Crystal structure of the catalytic core of human JARID1B in complex with 8 fragments showing binding hotspots. Red – 2OG binding site, green – putative allosteric binding pocket within the helical domain explored by N10072a.** $Mn^{2+}$ atoms are shown as cyan spheres and $Zn^{2+}$ atoms are shown as pink spheres.

This resulted in eight crystal structure complexes. The binding location of the fragments is shown in Figure 13, and the details of the binding in Figure 14. Out of 8 fragments, 7 were bound in the active site and one was found in a putative allosteric pocket of the alpha helical domain of JARID1B outside of crystal contacts. We further determined the $IC_{50}$ values for the crystallised fragments and the data is presented in Table 8.


A
Asn509
Trp519
W188
Lys517
DMS1
W4
Tyr488
W1
Glu501
Mn
W196
Asn591
His587
W3
W199
W236
W256
W198
W243
W198
W197
Ser494
Phe496
His499
Edo19
W199
Ser495
Tyr425
Arg98
DMS3
N05859b
B
Asn509
Trp519
Lys517
Glu501
DMS1
His587
W1
Phe496
W3
Tyr488
His499
Edo19
W156
Ser495
Arg98
Tyr425
Edo17
Val99
Arg86
Lys100
E48115b (top view)
C
Asn509
Lys517
Trp519
W4
DMS1
Glu501
W1
Mn
Asn591
His587
W3
W5
Ser494
W71
W66
His499
Phe496
W68
Ser495
Tyr425
Arg98
N09954a
D
Asn509
Lys517
Trp519
Tyr488
W4
DMS1
W11
W15
W1
W12
Glu501
W17
Asn591
His587
W18
W16
W5
W6
W14
Edo15
His499
Phe496
Ser494
W18
Ser495
Arg98
DMS3
N11297a
E
Asn509
Lys517
DMS1
Trp519
W4
Glu501
W1
W191
Mn
Asn591
His587
W190
W199
W3
W143
W197
Ser494
W189
W218
Phe496
His499
W198
Edo1
W214
Ser495
Tyr425
Arg98
Edo2
Val99
N10042a
F
Asn509
Lys517
W191
W4
DMS1
Trp519
Glu501
W1
Mn
W196
Asn591
His587
W199
W3
W247
W197
W189
His499
W272
W293
Ser495
W198
Tyr425
Arg98
DMS3
N10063a

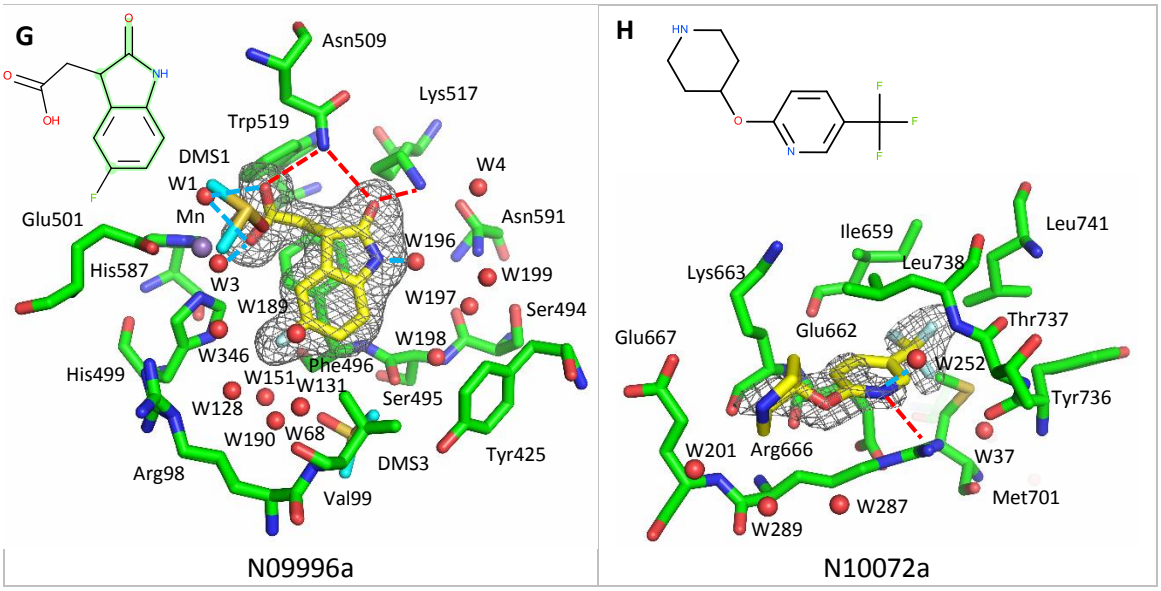



**Figure 14. Crystal structures of JARID1B in complex with soaked fragments. A-G) Fragments bound in the active site were compounds N05859b, E48115b, N09954a, N11297a, N10042a, N10063a, N09996a, H) A fragment bound directly to the helical domain (N10072a).** The Fo-Fc OMIT density map for each ligand is contoured at 3σ. Hydrogen bond interactions and polar interactions formed between the fragment and the residue side-chains are depicted as red dashed lines and solvent mediated interactions as cyan dashed lines. The cut-off distance for polar interactions is 3.2 Å. The oxindole based scaffold is highlighted in green.

A comparison of the 2OG binding site in the determined fragment complexes shows the conserved hydrogen bond interactions with the carbonyl group of N10042a, N10063a, N09954a and N09996a (or carboxylic in case of N05859b) with $N^{\varepsilon}$ of Lys517 and the carboxamide nitrogen of Asn509 (Figure 14 C and G). As previously observed with the TCA cycle intermediates, the active site accommodates dual conformations of Asn509 and alternative positions of Arg98 (see N09954a and N10042a). We also observed two "conserved" DMSO molecules (DMS1 and DMS3, Figure 14) where DMS1 is present in all the crystal structures and stabilised by interaction with Tyr488, and DMS3 appears to be more amenable to displacement by ethylene glycol (Edo2 and Edo17, Figure 14 B and E) or water molecules (Figure 14 C).

Among the pyridine-based metal chelators, isonicotinic acid (N05859b Figure 14 A) is the most active compound of all tested fragments with an $IC_{50}$ value of 0.5 µM against JMJD2A and 1.0 µM against JARID1B (Table 8). N05859b binds to the catalytic metal-ion with its pyridine nitrogen and the carboxylic acid oxygen forms a hydrogen bond with $N^{\varepsilon}$ of Lys517 (Figure 14 A). Compounds E48115b and N09954a (Figure 14 B and C) contain the same

pyridine motif but with meta- and para- substitutions of the pyridine nitrogen respectively, which results in dramatically different binding poses. In the structure of N09954a the carbonyl group forms a hydrogen bond with the carboxamide nitrogen of Asn508 and $N^{\varepsilon}$ of Lys517, whereas in the case of E48115b the carbonyl group places itself away from these residues and interacts weakly with the oxygen atom in an ethylene glycol molecule (Edo17, distance ~3 Å). The dichlorobenzene of E48115b places itself closely to an ethylene glycol molecule which replaces the conserved DMS3 molecule. The N09954a (Figure 14 C) is stabilised by hydrogen bond interaction between the carbonyl oxygen and the $N^{\varepsilon}$ of Lys517 and the carboxamide nitrogen of Asn591, and displacing three active site water molecules conserved in structures of N09996a, N10042a, N10063a and N05859b and a water molecule present in N10063a and N09996a.

Interestingly, compound N10042, the most potent of the oxindole analogues identified within the library with an $IC_{50}$ value of 5.9 µM towards JmjC-type oxygenases, does not bind to the metal directly but through the connecting water network forming hydrogen bond interactions between the oxindole nitrogen and a water molecule (Figure 14 E). N10042 is further stabilised by hydrogen bond interactions between the amino group nitrogen and water molecules, and also by carbonyl oxygen hydrogen bond with $N^{\varepsilon}$ of Lys517 and weak interaction (distance ~3 Å) with the carboxamide nitrogen of Asn509. Substitution of the amino group with the nitrile motif with a one-carbon linker (N10063a Figure 14 F), or the carboxylic acid group with a one-carbon linker (N09996a Figure 14 G) flips the conformation of the fragment in comparison with N10042a and enables a mono-dentate metal chelation, thereby decreasing the activity against all enzymes more than 100-fold (Table 8). Both N10063a and N09996a are stabilised by hydrogen bond interactions between the oxindole nitrogen and $N^{\varepsilon}$ of Lys517 and water molecules, and carboxyl oxygen forms hydrogen bond with Lys517 and albeit weaker interaction with the nitrogen of carboxamide in Asn509 (distance ~3 Å). In addition, the carboxylic acid oxygen of

N09996a forms a hydrogen bond with the carboxamide nitrogen of Asn509 and water molecules.

An alternative metal binding motif is observed for N11297a (Figure 14 D): the ligand binds to the catalytic metal ion through an “alanine-like” motif in a bi-dentate manner thereby displacing a water molecule and forming hydrogen bond interactions between its carboxylic acid oxygen and water molecules as well as through the amino nitrogen and a water molecule. The 1-chloro-3-fluoro-benzene motif found in N11297a is found in a binding pocket adjacent to the conserved DMS3 molecule. Compound N11297a is however inactive in the dose-response studies (maximum concentration of 500 μM), with the exception of FBXL11 where an $IC_{50}$ of ~500 μM was determined.

Finally, compound N10072a (Figure 14 H) binds in a putative allosteric hydrophobic pocked formed between α-helices 18 and 19 within the helical domain of JARID1B. The pyridine nitrogen forms a polar interaction (distance ~3 Å) with a water molecule (Figure 14 H). The binding of N10072a induces a slight conformational change, shifting the C-terminal helix (helix 19, residues Leu738-Ala752) by up to 1.3 Å and thus opening the binding pocket (Figure 15). The average pairwise calculated all-atom RMSD of helix 19 in the N09954a complex between other active site complexes reveals an average RMSD of 0.323 Å with a standard deviation of 0.024 Å, whereas the all-atom RMSD of helix 19 calculated between N09954a and N10072a complexes is equal to 0.539 Å indicating significant difference from the mean (Z-score of 8.9). Despite absent electron density for the compound in the 2OG binding site N10072a inhibits JARID1B with a high $IC_{50}$ of 81.2 μM, indicating a possible allosteric inhibitor site.

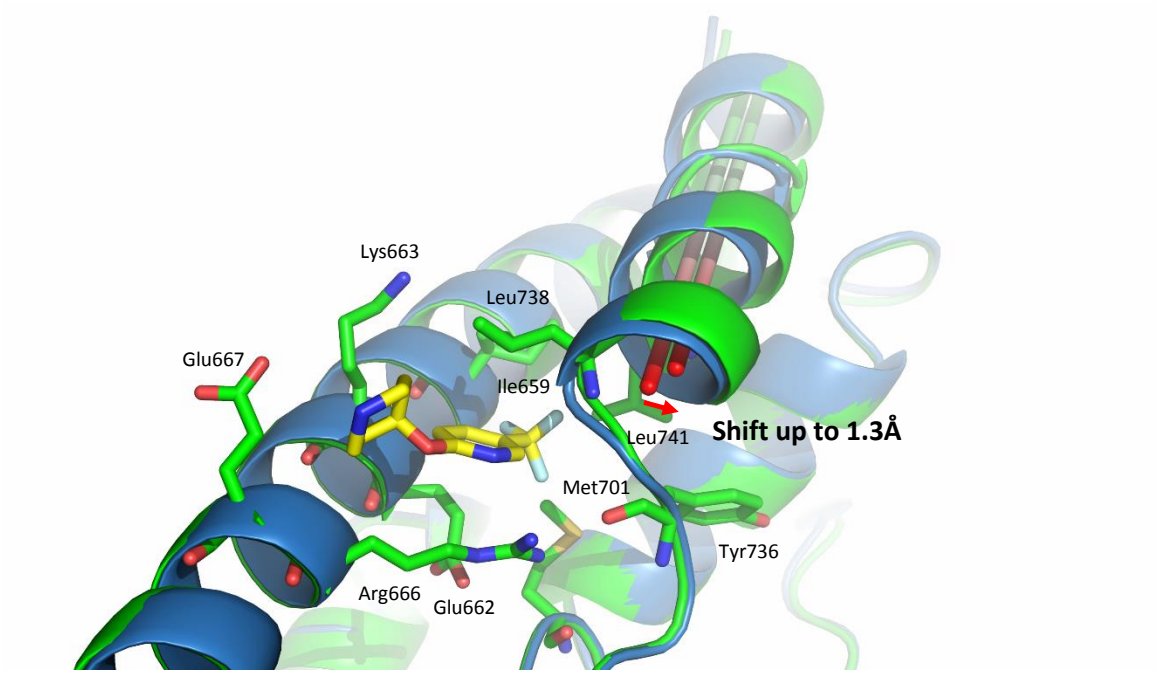



**Figure 15. Overlay of a section of the helical domain of JARID1B in complex with fragment N10072a (green) and JARID1B in complex with N09954a (blue).** N10072a (yellow) induces a shift of the C-terminal helix (Leu738-Ala752 green) in comparison to the ligand free state (within the putative allosteric pocket; blue). The two lines indicate the centre of mass between bound and unbound states of the C-terminal α-helix.

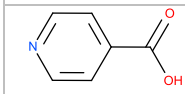
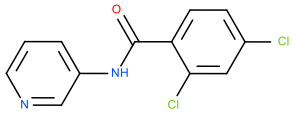
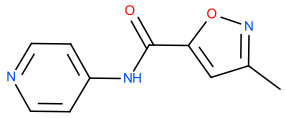
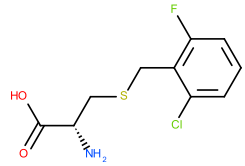
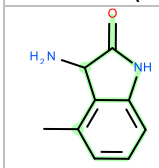
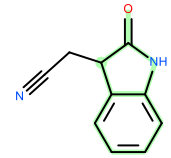
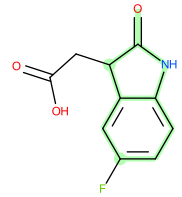
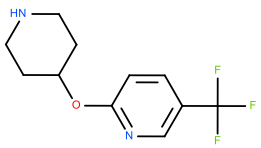
| | % inhibition at 250 µM | | IC$_{50}$ [µM] | | | | | |
|---|---|---|---|---|---|---|---|---|
| | RapidFire | | Alpha Screen | | | | | |
| | NO66 | MINA53 | FBXL11 | JMJD1B | JARID1B | JMJD2A | JMJD2D | JMJD3 |
|  N05859b (MB) | 5%* | -7%* | 28.6 | 38.2 | 1.4 | 0.5 | 1.9 | 50.1 |
|  E48115b (MB) | 0%* | 6%* | >500 | >500 | >500 | >500 | 350.7 | >500 |
|  N09954a (3D) | 2% | 6% | -19% | -5% | 16% | -6% | 2% | -7% |
|  N11297a (MB) | 9%* | 3%* | 570.5 | >500 | >500 | >500 | >500 | >500 |
|  N10042a (3D) | 86% | 102% | 8.7 | 34.1 | 5.9 | 119.7 | 28.0 | 48.6 |
|  N10063a (3D) | 4% | 6% | >1000 | >1000 | 584 | >1000 | >1000 | ~1321 |
|  N09996a (3D) | -5% | -23% | >1000 | >1000 | ~1200 | >1000 | >1000 | >1000 |
|  N10072a (allosteric pocket) (3D) | -6% | -8% | 41.4 | 246.9 | 81.2 | ~1065 | 107.5 | 318 |

**Table 8. Activity profile of bound fragments to human JARID1B.** The oxindole based scaffold is highlighted in green. % inhibition at 250 µM compound concentration is shown based on single data point. *screened at 100 µM, n=1. Abbreviations: MB – Maybridge library, 3D – 3D fragments.

## 2.4. Discussion

Both libraries used in this fragment-based screening approach constitute a diverse group of molecules, a diversity which translates to the identified chemotypes. Selectivity profiling identified fragments with a broad spectrum of inhibition against several JmjC members, as well as compounds with a higher degree of selectivity. A pairwise comparison of hit numbers of Maybridge fragments within the tested proteins reveals an elevated hit rate and overlap between the proteins within the KDM4 family, suggesting that similarity of identified hits correlates with the similarity of the active site architecture.

Approximately 40% of hits from the Maybridge library and ~20% of hits from the 3D fragment library identified in the AlphaScreen assay turned out to be false positives based on the H3K9me2 counter screen (see Table 3 and Table 5). It is indeed possible that some false positives have been missed since the remaining demethylation products have not been tested in a counter screen manner, however the activity assay was performed as an initial screen in order to identify potential novel lead compounds and also prioritise compounds for the initial X-ray soaking experiments with Maybridge library, where obtained complexes serve as the ultimate validation of compound binding. High throughput X-ray fragment soaking platform developed at Diamond synchrotron allows for screening of entire 581 fragment library, in the initial fragment soaking we however focused on the subset of fragments which display inhibitory activity. The remaining fragments could therefore be tested in subsequent experiments. The high number of false positive underlines the weaknesses of this chemoluminescence based AlphaScreen assay, which requires considerable validation and counter-screening activities. The hits identified in this screen would ideally be validated by an orthogonal assay such as RapidFire mass spectrometry assay. The RapidFire mass spectrometry based assay does not have the high false positive rate problem, since the measurement is label free, but we have found that fragments in the 3D library screened at 250 μM concentration interfere with the assay decreasing the effective peptide detection range by reducing the ionisation

of the peptide or dominating the ionisation spectra. These shortfalls were however compound-specific and were easily identifiable events in the data analysis. In the subsequent activity screen of Maybridge library we have decided to limit compound concentration to 100 μM.

The crystal structures obtained in this project show a high degree of "plasticity" of the 2OG binding site of JARID1B-this was indicated by two conformations of side-chain moieties of residues Tyr425 and Asn509 and these different conformations were successfully induced by various fragments. The active site also shows binding of two "conserved" DMSO molecules whose presence could influence the exploration of the active site by soaked fragments. This possible artefact could be avoided by solubilising compounds directly in ethylene glycol, which is used as a cryoprotectant, instead of DMSO. It is likely that the ethylene glycol molecule would then occupy the DMSO sites, but could possibly be easier to displace by fragments. The features of the soaking protocol are also important to note. In the semi-automated method, the compound and cryoprotectant are allowed to diffuse into a crystallisation drop, whereas in the manual soaking protocol, compound and cryoprotectant solution are pre-mixed which results in a uniform concentration. However, manual soaking allows for more control over the soaking procedure, in contrast to the high-throughput nature of the semi-automated method.

From the bound fragments isonicotinic acid (N05859b) is the most active compound with an $IC_{50}$ value below 2 μM for JARID1B, JMJD2A and JMJD2D. Derivatives of compound N05859b such as isonicotinamides may therefore present themselves as potent inhibitors of JmjC-type enzymes. The lack of inhibition of N05859b in NO66 and MINA53 RapidFire assay is surprising since its analogue 2,4-PDCA binds to both enzymes as verified by X-ray crystallography and inhibits the activity as described in 3.4.2. Since the inhibition data were performed in single experiments (n=1), the determination of a full $IC_{50}$ dose-response with NO66 and MINA53 should clarify this result. Due to external factors it was not possible to complete this experiment during the preparation of this thesis.

Since most inhibitors against the family of JmjC-type oxygenases act through chelation with the active site metal (see 1.6), fragment N10042a found in the 2OG binding site is especially interesting since it does not bind to the metal directly, but rather through a network of water molecules (another example of indirect metal binding is compound MC3098 described in 4.1.2). Interestingly oxindole scaffold with additional metal chelating motif of nitrile (N10063a) or carboxylate (N09996a) decreases the activity substantially making the N10042a the most active compound in the series. Compounds with metal chelating motifs are likely to inhibit other metalloenzymes such as HDACs which may impair their selectivity from the start, and development of inhibitors which are not metal chelating is therefore of substantial interest. Interestingly, in the identified set of metal chelators only one compound contains the frequently problematic carboxylic acid motif (N09996a). Carboxylate containing molecules such as 2,4-PDCA, NOG, IOX-1 or GSK-J1, although highly active *in-vitro,* frequently have low cellular permeability and require a pro-drug form in order to penetrate the cell membrane. Diversity of the metal chelators of this study may therefore enable the development of a new class of non-carboxylate inhibitors with an increased chance of efficacy in cells.

Fragment soaking also enabled identification of a putative allosteric binding site in JARID1B by fragment N10072a. Binding caused a shift towards the C-terminal α-helix in the alpha helical domain. It is interesting that the N10072a allosteric binding translates to an albeit weak inhibition of activity of JmjC-type enzymes with a range of $IC_{50}$ values around ~100 μM or below for JARID1B, JMJD2D and FBXL11. There is no electron density present in the active site to support the binding of the ligand, thereby indicating a possible allosteric site that could be further explored for inhibitor development. However, it is still possible that the compound is also inhibiting the reaction through binding to the active site, which could be assessed by repeating the soaking experiment. Furthermore, the mode of inhibition of N10072a should be evaluated through detailed kinetic experiments by

measuring the $V_{max}$ and $K_M$ of substrate (2OG or peptide) in a range of inhibitor concentrations [193].

Taken together, this fragment soaking study identified several novel indirect metal binding motifs for JmjC-type enzymes, identified a putative allosteric inhibitor site, and provided multiple starting points for further inhibitor development.

# 3. Inhibitor chemotype identification for NO66 and MINA53

## 3.1. Introduction

MINA53 and NO66 are two 2OG-oxygenases that are regulated by the oncogene c-Myc and both have been shown to hydroxylate ribosomal proteins [30]. Both proteins were initially reported to be histone demethylases (H3K9Me3 and H3K36Me3, respectively), but this activity is disputed in the literature [30], [136]. Recent publications show that overexpression of NO66 *in vivo* inhibited skeletal growth and bone formation [138], whereas inhibition of MINA53 results in suppression of cell proliferation in a variety of different cancers [128]–[132] (see also section 1.3). Thus, modulation of MINA53 and NO66 has therapeutic potential in inflammation and oncology, and our aim was therefore to identify and characterise specific inhibitors for both enzymes.

***Structural studies of NO66 and MINA53***

Several structures have been determined for MINA53 and NO66 both in an apo form (NO66 PDB: 4CCJ, 4E4H), in complex with 2OG (MINA53 PDB: 4BU2) or the generic inhibitors NOG and 2,4-PDCA (NO66 PDB: 4DIQ, 4CCK, 4Y33; MINA53 PDB: 2XDV) and with ribosomal peptides of Rpl8 and Rpl27a (NO66 PDB: 4CCN, 4CCM, 4CCO, 4Y33; MINA53 PDB: 4BXF). The crystal structures of MINA53 and NO66 reveal a similar folding pattern (Figure 16) with a N – terminus containing a flexible loop region that is followed by the catalytic *JmjC* domain that harbours the $Fe^{2+}$ and 2OG binding pockets, a dimerisation domain involved in oligomerisation, a linker region, and a C-terminal helix turn helix domain (wHTH) involved in substrate binding [137]. Interestingly, structural studies within the family of JmjC-containing proteins (such as FIH, TYW5, JMJD4, JMJD6) reveal that the wHTH domain might be unique to ribosomal oxygenases (NO66, MINA53, OGFOD in eukaryotes and ycfD in prokaryotes) across the family of JmjC-type oxygenases [137].

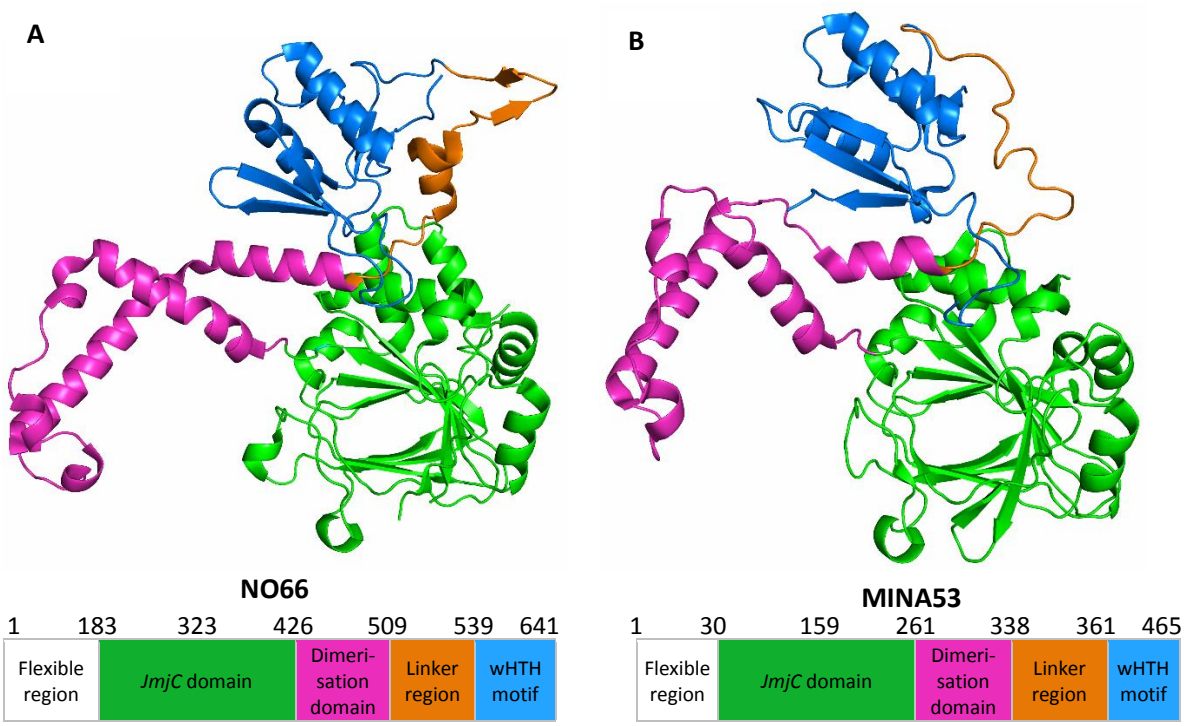



**Figure 16. A: Domain architecture of NO66 (PDB: 4E4H chain A). B: Domain architecture of MINA53 (PDB: 4BXF chain A). The catalytic *JmjC* domain is coloured in green.**

Superimposition of the MINA53 (PDB: 2XDV chain A) and NO66 (PDB: 4DIQ chain A) structures reveals a RMSD value[2] of 3.35 Å, and the major differences are found in the dimerisation domain, the linker as well as in the loop regions [194]. Both MINA53 and NO66 form homodimers in solution stabilised by polar interactions, hydrogen bond and hydrophobic interactions between helices in the dimerisation domain (Figure 17 interface 1) [137], whereas NO66 also exists as a tetramer through both polar and hydrophobic interactions between the linker regions (Figure 17 A, interface 2) [142]. Mutation studies of NO66 show that the catalytic activity is lost for NO66 in the monomeric form, and is reduced for dimers formed through interface 1 (dimerisation domain) and further reduced in dimers formed by interface 2 (linker domain) [136]. Interestingly, as demonstrated by ITC studies, binding of the peptide GST-Rpl8$^{193\text{-}C}$ or the 2OG cannot be detected for the monomer and dimer of NO66 formed through interface 2. The native tetramer resulted in a $K_D$ for the peptide of 2.59 ± 0.23 μM and 2OG of 1.0 ± 0.18 μM, whereas dimers formed

[2] TM-align on-line http://zhanglab.ccmb.med.umich.edu/TM-align/

through interface 1 showed a reduced $K_D$ towards both peptide (4.67 ± 0.29 μM) and 2OG (6.25 ± 0.13 μM) [136]. Reduced activity has also been observed with mutants in interface1 (dimerisation domain) of MINA53 [137]. Together these studies suggest that the activity of NO66 and MINA53 can be modulated by oligomerisation states.

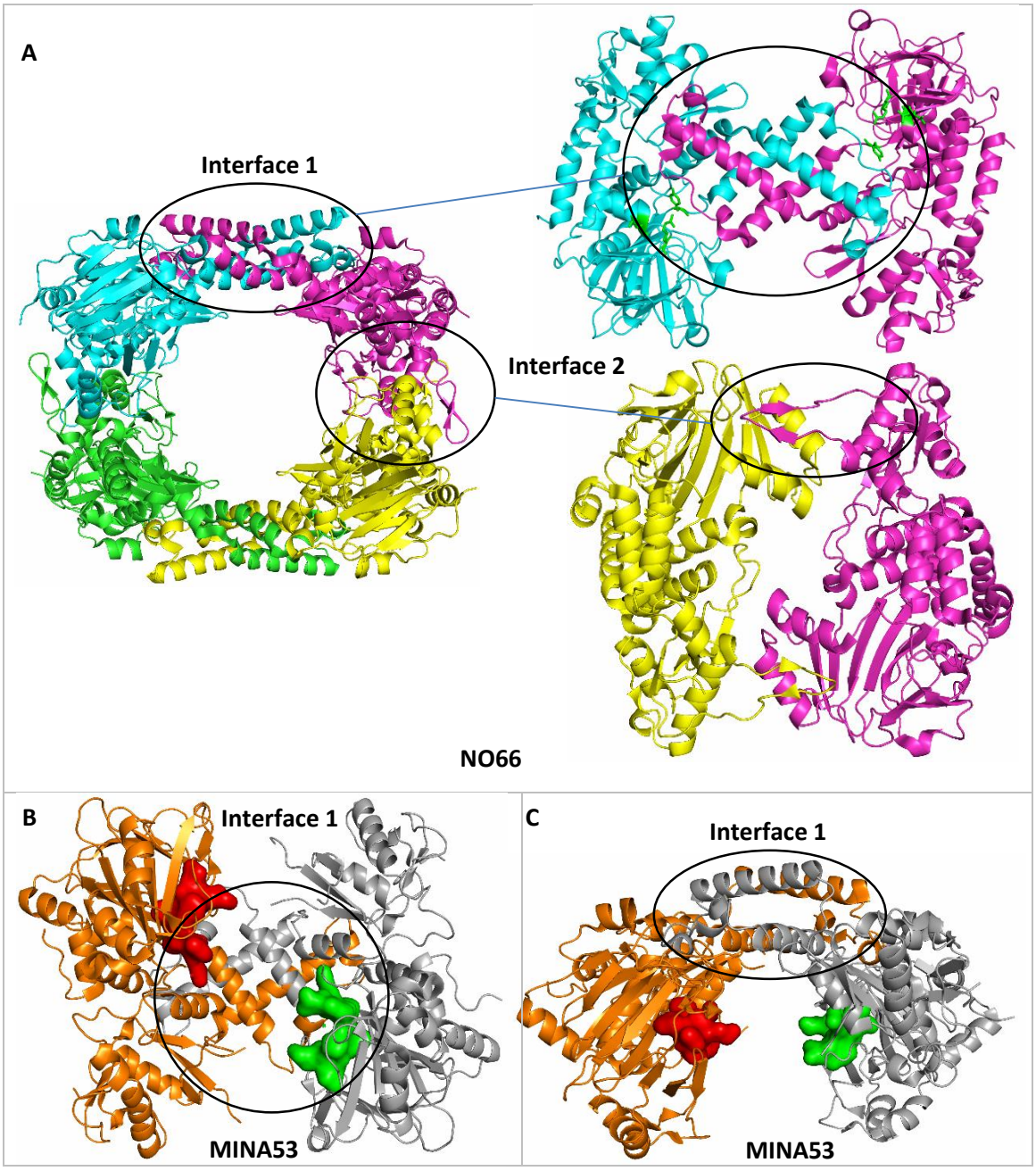



**Figure 17 A. Tetrameric assembly of NO66 (4E4H). For clarity the four chains have been coloured differently in green, cyan, magenta and yellow. Dimerisation interface 1 and 2 are indicated with a circle and residues interacting with the transcription factor Osterix highlighted in green. B-C. Dimeric assembly of MINAB (4BXF) cross –linked with Rpl8 peptide (red, green). The two chains have been coloured differently in grey and orange respectively. B. face view and C. top view of interface 1; dimerisation domains involved in dimerisation are circled.**

The electrostatic potential for the structures of NO66 and MINA53 was calculated in APBS [195] and presents a conserved electronegative inner core which is considerably larger in NO66 and more localised in MINA53 (Figure 18).

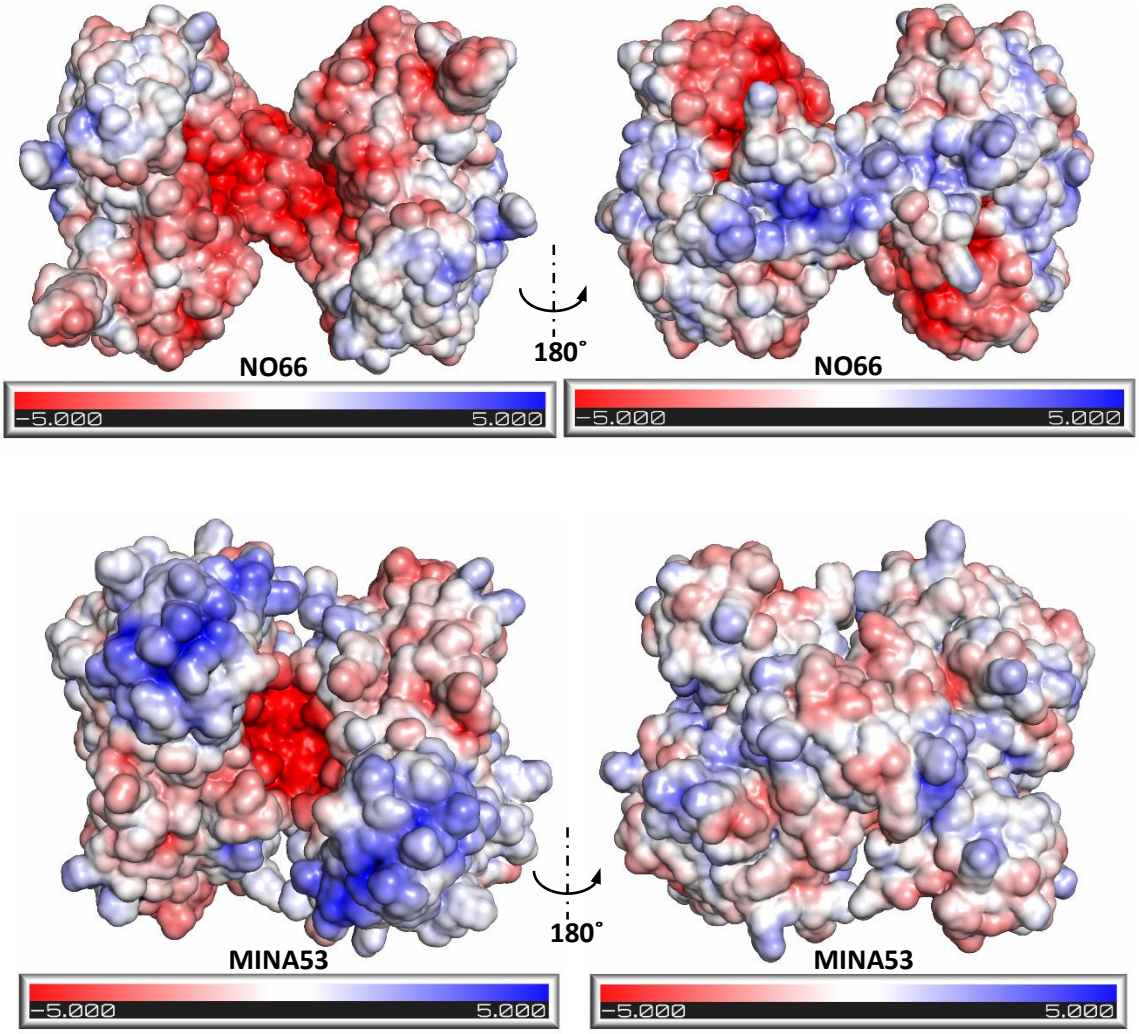



**Figure 18. NO66 (4DIQ) and MINA53 (4BXF) present distinct surface charges.** The electrostatic potential was calculated by using the boundary solution to Poisson equation in APBS [195] with standard settings and is shown by red to blue colour scale representing the ± 5 kcal/electron units (+5 blue and -5 red).

Oligomerisation of NO66 has also been shown to be necessary for interaction with the Osterix (Osx) transcription factor, which results in repression of Osx-mediated gene expression. The proposed binding location of Osx has been established through mutation studies where a triple point mutant R221A, R225A, and Y423A showed reduced binding of Osx to NO66 (Figure 17 A).

Despite similar active sites, both NO66 and MINA53 hydroxylate different ribosomal proteins (Rpl8 and Rpl27a respectively) with the specificity achieved through interactions with residues outside of the immediate 2OG binding pocket [137]. The 2OG binding site of MINA53 (PDB: 2XDV) and NO66 (PDB: 4CCK) is highly conserved as shown in Figure 19. The pairwise residue differences are modest with non-polar hydrophobic residue pairs Ile187/Val348, and Leu176/Phe337 and polar residues Thr134/Ser295 in MINA53/NO66 respectively. Tyr328 (NO66) and Tyr167 (MINA53) are involved in peptide substrate binding [136], [137].

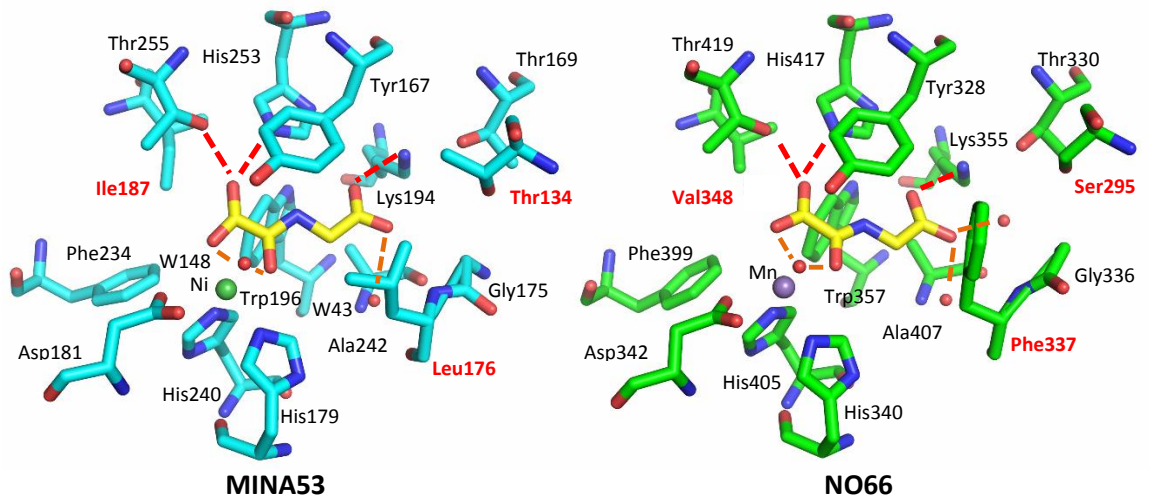



**Figure 19. Catalytic active site residues of MINA53 PDB ID 2XDV (cyan) and NO66 PDB ID 4CCK (green) in complex with NOG (yellow).** The residue differences are highlighted in red labels, and water molecules are shown as red spheres. Hydrogen bond interactions between ligand and protein residues are shown as red dashed lines and water-mediated interactions are depicted by orange dashed lines.

Crystal structures of MINA53 in complex with cross-linked peptide Rpl27a and NO66 with Rpl8 show that both peptides bind in shallow channels on the surface of proteins forming multiple hydrogen bond interactions and hydrophobic contacts with the wHTH domains of MINA53 (Met405, Met406) and NO66 (Tyr577, Val576) [137] (Figure 20). The backbone of both peptides shows significant overlap with an RMSD of 0.8 Å [137] and both Rpl8 and Rpl27a peptide substrates have two conserved protein binding residues: His39 (MINA53), His216 (NO66) modified on the α-carbon and His41 (MINA53), His218 (NO66). The hydroxylated histidine residues (His39 of Rpl8 peptide and His216 of Rpl27a) are stabilised by hydrogen bond interactions between the imidazole nitrogen atoms and

hydroxyl oxygens of Tyr167/Ser257 in MINA53, and Tyr328/Ser421 in NO66, and additional hydrophobic interactions are found in NO66 formed with Ile344.

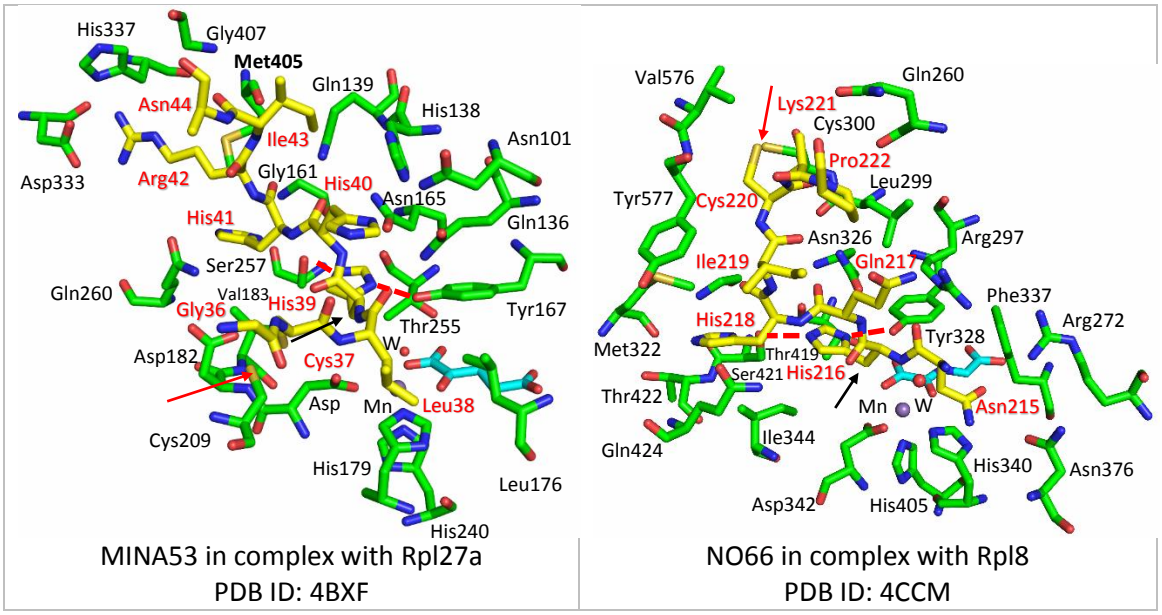



**Figure 20. Binding site of Rpl27a with MINA53 (left) and Rpl8 with NO66 (right).** Rpl8 (Gly214-Pro222) residues GNHQHI**C**KP and Rpl27a (Gly36-Asn44) residues G**C**LHHHRIN shown in yellow sticks, 2OG in cyan sticks and $Mn^{2+}$ bound water molecule is shown as a violet sphere. Black arrow points to the β-carbon of the histidine substrate of hydroxylation His39 in MINA53 and His216 in NO66. The red arrow points to the location of a disulphide Cys209 in MINA53 and Cys300 in NO66 used to obtain the complexes. Dashed red lines indicate the hydrogen bond interactions stabilising the hydroxylated histidine residues.

The recently obtained native crystal structure complex of NO66 and Rpl8 (Met204-Thr224) (PDB: 4Y3O) with residues Gly212-Pro222 visible in the electron density map indicates a different conformation of the peptide for the region Ile219-Pro222, which is likely due to the strain of the cross-linking (PDB: 4CCM).

The exact location of the ribosomal proteins Rpl8 and Rpl27a and the resulting peptides fragments in the ribosomal 60S subunit is shown in Figure 21.

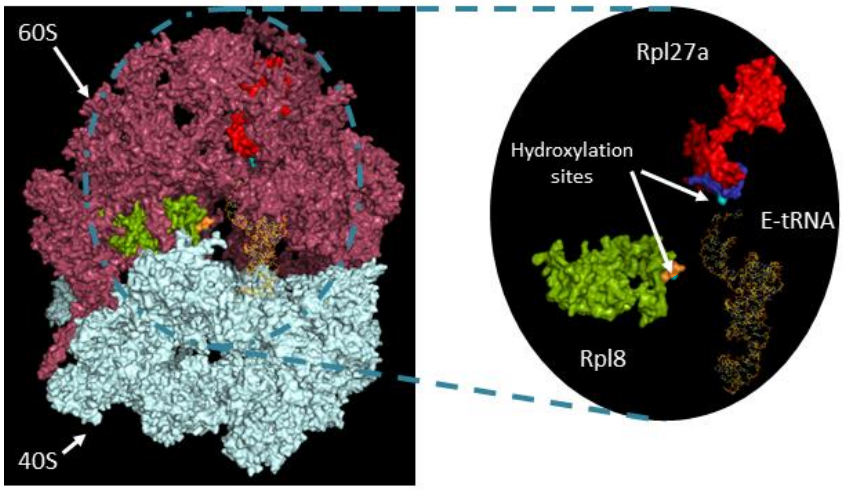



**Figure 21. Structure of human 60S and 40S subunits of ribosome together with E-tRNA showing location of NO66 and MINA53 substrates (PDB: 3J3A, 3J3B, 3J3D** [196]**). The ribosomal protein Rpl8, substrate for NO66, is highlighted in green and Rpl27a, the substrate for MINA53 is highlighted in red. Peptide fragments used in this work are shown in orange (Rpl8) and dark blue (Rpl27a), with hydroxylation site His 216 in Rpl8 and His 39 in Rpl27a coloured in cyan.**

The studies also identified a conserved peptide recognition motif, HxHR and NHxH, for NO66 and MINA53 [30], respectively, indicating that the substrate set is likely not limited to ribosomal proteins. A *BLAST* search performed by Tao *et al.* identified 20 possible substrates for NO66 [136]. Such substrate promiscuity is already observed with other JmjC-type enzymes such as FIH which can hydroxylate Asn, Asp, His, Leu and Ser residues [197].

### *Inhibitor development strategy*

To establish an inhibitor development system for NO66 and MINA53, we decided to utilise *in vitro* assays complimented by a crystallography system. We first screened available small molecule libraries in a readily available DSF assay format, identifying ligands increasing thermal stability upon binding as described in section 3.3. We then progressed to development of an activity assay based on the ribosomal hydroxylation. Assay development and screening for activity modulating compounds is described in paragraph 3.4. Although discovery and optimisation of hits is possible with just biophysical/biochemical assays, information about the binding mode through crystallography is highly desirable. The second challenge was therefore to establish a

crystallography system that reproducibly grew well diffracting crystals (< 2.5 Å) which could be used for either co-crystallisation or soaking experiments with the identified inhibitors. Although we were unable to obtain high resolution crystal structures with identified inhibitors with NO66/MINA53, we obtained high-resolution structures of several low micromolar MINA53 inhibitors (described in Chapter 4) in the established crystallography system of JARID1B.

## 3.2. Expression and purification

*E. coli Expression and Purification*

Prior to this project both MINA53 and NO66 had been cloned into bacterial expression vectors with N- and C- terminal truncations. Extensive work had identified soluble variants of both proteins suitable for activity and crystallography studies (Table 9), but both proteins were known to be difficult to purify with tendencies to aggregate and precipitate. Initially these truncated variants of NO66 and MINA53 were expressed in the *E. coli* Rosetta strain with either a C-terminal or N-terminal histidine tag as shown in Table 9. The typical yield of MINA53 was about 3.3 mg/L for both full-length and truncated constructs. In case of NO66, a truncated construct encompassing residue Ala167-Asn641, resulted in maximum yield of 2.8 mg/L whereas the construct covering Ser183-Asn641 had a much lower yield of ~1mg per 6L expression.

| Construct | Sequence | Tag | Yield (mg/L) |
|---|---|---|---|
| **NO66-c102** | Ala167-Asn641 | C-terminal His-6 tag | 0.8-2.8 mg |
| **NO66-c103** | Ser183-Asn641 | | 0.2 mg |
| **MINA53-c001** | Met1-Val464 | N-terminal His-6 tag | 0.8-3.7 mg |
| **MINA53-c002** | Ala26-Val464 | | 2.7-4.2 mg |

**Table 9. NO66 and MINA53 variants expressed in *E. coli* indicating the total yield per litre expression. Multiple other sequence variants were tested previously by the SGC, data not shown.**

The proteins were purified by nickel-affinity chromatography using a stepwise gradient of imidazole followed by size-exclusion chromatography. The histidine tag was cleaved by Tobacco Etch Virus protease (TEV) followed by nickel-affinity chromatography to remove uncleaved proteins, and the mass and purity of the proteins were confirmed by mass spectrometry. Details on the purification procedure are described in Appendix A. The mass of NO66 purified from *E. coli* cells was -131 Da lower than predicted, indicating removal of the initiator methionine with a mass of -131.040485 Da. For MINA53, we could not detect any posttranslational modifications. We noticed that both NO66 and MINA53 are temperature sensitive, and by keeping the proteins on ice during the purification

procedure, and by chilling tubes and pipette tips, both precipitation and aggregation was avoided. The typical purification yield is shown in (Table 9 - Table 11).

*Biotinylated NO66 and MINA53*

In order to use assays based on streptavidin-biotin interactions, both NO66 and MINA53 were biotinylated as described in [198]. The proteins were expressed in an *E. coli* Rosetta-R3-BirA strain which co-expresses BirA, a biotin ligase that enables *in vivo* biotinylation of the biotin acceptor site. Biotinylated NO66 and MINA53 were purified as described in the Appendix A and biotinylation was confirmed by mass spectrometry. Typical yield is shown in (Table 10).

| Construct | Sequence | Tag | Yield (mg/L) |
|---|---|---|---|
| **NO66-c044** | Ala167-Asn641 | N-terminal His-10 tag, C-terminal biotin | 1.5-2.2 mg |
| **MINA53-c063** | Met1-Val464 | N-terminal His-10 tag C-terminal biotin | 2.8 mg |

**Table 10. Constructs of NO66 and MINA53 used for *in vivo* biotinylation indicating the total yield per litre expression.**

*Expression in insect cells*

Since it proved difficult to obtain high quality crystals of both MINA53 and NO66 (see 3.5), and since we, as well as others [30], [136], were unable to detect any *in-vitro* histone demethylase activity (see 3.4.2), we decided to express both proteins in parallel in Sf9 cells using a baculovirus expression system.

The cloning of full-length NO66 failed, but two shorter variants were successfully cloned and purified. NO66 Ala116-Asn641 showed a +42.1 Da mass shift by mass spectrometry, which indicates that the protein might be acetylated at the N-terminal amino group. In contrast to NO66, both full-length and truncated variants of MINA53 were successfully cloned and expressed in SF9 cells and no posttranslational modifications could be detected in the purified protein.

| Construct | Sequence | Tag | Yield (mg/L) | Comments |
|---|---|---|---|---|
| **NO66-c014** | Met1-Asn641 | N-terminal His-6 tag | - | Cloned/no expression* |
| **NO66-c015** | Ala6-Asn641 | | - | Cloned/no expression* |
| **NO66-c016** | Gln116-Asn641 | | - | Cloned/no expression* |
| **NO66-c017** | Ala167-Asn641 | | - | Cloned/no expression* |
| **NO66-c022** | Met1-Asn641 | C-terminal His-10 tag | - | Cloning failed |
| **NO66-c023** | Ala6-Asn641 | | - | Cloning failed |
| **NO66-c024** | Gln116-Asn641 | | 5 | Purified |
| **NO66-c025** | Ala167-Asn641 | | 2.8-5.8 | Purified |
| **MINA53-c064** | Met1-Val464 | N-terminal His-6 tag | 4.5 | Purified |
| **MINA53-c065** | Ala5-Val464 | | - | Cloned/high expression* |
| **MINA53-c066** | Ala26-Val464 | | 13 | Purified |
| **MINA53-c067** | Met1-Val464 | C-terminal His-10 tag | 6.7 | Purified |
| **MINA53-c068** | Ala5-Val464 | | - | Cloned/no expression* |
| **MINA53-c069** | Ala26-Val464 | | 2.2 | Purified |

**Table 11. NO66 and MINA53 constructs expressed in insect cells (Sf9). Total yield per litre expression. Cloning and test expression was done by the Biotechnology Team at the SGC. *based on a 3 ml test expression elution yield.**

## 3.3. Ligand screening by thermal stability assay

### 3.3.1. Differential Scanning Fluorimetry assay

Differential Scanning Fluorimetry (DSF, often called $T_m$ shift assay) is a quick method of assessing stability of proteins by measuring the melting temperature change upon ligand binding [199]. A native protein unfolds as a result of temperature increase exposing the hydrophobic core and this process can be observed with addition of a dye which is highly-fluorescent in hydrophobic environment but quenched in aqueous solution. The measurement of the DSF assay is the melting temperature shift (hence the name '$T_m$ shift') in comparison to a reference curve which in case of small molecule screening is usually the DMSO solvent control. The assay allows to test for various protein stabilising agents such as small molecules or different buffer compositions [200]. Since an increase in melting temperature (stabilisation) observed after addition of the ligand often correlates with binding affinity [199] the DSF assay is used as an initial screen for small molecules.

All proteins in the DSF assays had the affinity tag removed. A typical melting temperature curve for MINA53 is shown in Figure 22. The DSF temperature profile of unfolding of NO66 and MINA53 indicates a monophasic transition. We have observed increased stabilisation for both NO66 and MINA53 with the presence of $MnCl_2$ for MINA53 or $NiCl_2$ for NO66 which likely replaces the active site iron as indicated by available crystal structures (see paragraph 3.5.2). The stabilisation of NO66/MINA53 depends on both concentration of the ligand, such as the natural cosubstrate 2OG and an added metal salt such as $MnCl_2$. The increased stability of MINA53 as a result of increasing concentration of both 2OG and $MnCl_2$ is shown in Table 12 and Figure 22. The $T_m$ shift of MINA53 with 1.56 µM $MnCl_2$ increases with increased concentration of 2OG exceeding 2.0˚C only at 25 µM of 2OG (Table 12). Lack of 2OG with increasing concentration of $MnCl_2$ induces a 1.4-1.8˚C temperature shift only at high concentrations of 50-100 µM of $MnCl_2$. Presence of both 2OG and $MnCl_2$ induces significant concentration-dependent stability with the maximum of 9.5-9.7˚C $T_m$ shift at 100 µM of $MnCl_2$ and 25-100 µM of 2OG.

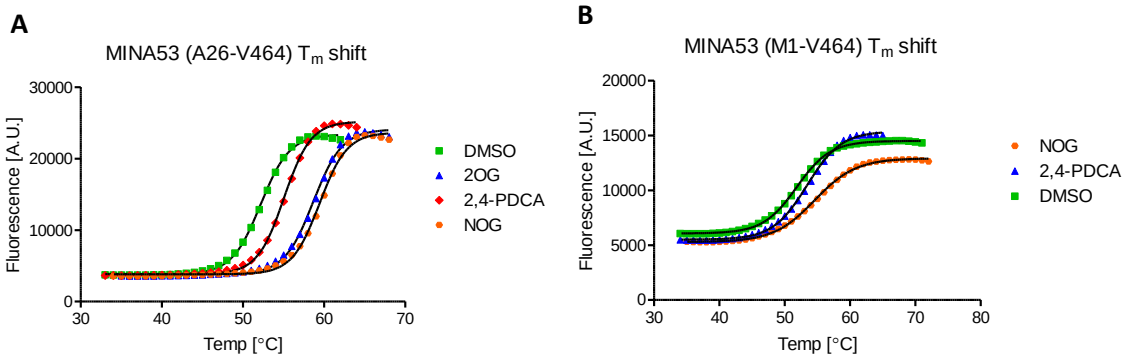



**Figure 22. A. Typical melting temperature for 2 µM MINA53 (Ala26-Val464) with 20 µM concentrations of freshly prepared 2OG, NOG and 2,4-PDCA in 50mM HEPES buffer pH7.5, 50mM NaCl supplemented with 10 µM $MnCl_2$.** The observed melting temperature for MINA53 with DMSO was 52.6˚C resulting in $T_m$ shift of 2.8˚C for 2,4-PDCA, 6.4˚C for 2OG and 7.2˚C for NOG. B. Typical melting temperature for method used in the screening assay with 1 µM MINA53 (M1-V464) with 10 µM concentrations of 2,4-PDCA and NOG in 50 mM HEPES buffer pH 7.5, 150 mM NaCl supplemented with 50 µM $MnCl_2$. The observed melting temperature for MINA53 with DMSO was 51.8˚C resulting in $T_m$ shift of 1.6˚C for 2,4-PDCA, 2.8˚C for NOG. The curve fitting is indicated with a solid black line.

| | | 2OG [µM] | | | | | | | | | No 2OG |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 100.00 | 50.00 | 25.00 | 12.50 | 6.25 | 3.13 | 1.56 | 0.78 | 0 | No $MnCl_2$ |
| $MnCl_2$ [µM] | 100.00 | 61.6 | 61.0 | 61.8 | 59.8 | 58.4 | 55.8 | 56.0 | 54.5 | 53.9 | **51.9** |
| | 50.00 | 58.9 | 57.8 | 59.7 | 58.8 | 57.1 | 56.4 | 54.3 | 54.2 | 53.5 | **51.8** |
| | 25.00 | 60.1 | 58.8 | 58.5 | 57.6 | 56.2 | 55.0 | 54.5 | 53.7 | 51.6 | **52.0** |
| | 12.50 | 59.4 | 57.8 | 56.8 | 56.5 | 55.0 | 54.6 | 54.0 | 53.2 | 52.8 | **52.0** |
| | 6.25 | 58.5 | 56.9 | 56.3 | 55.1 | 54.4 | 54.5 | 53.7 | 53.0 | 52.8 | **52.1** |
| | 3.13 | 58.0 | 55.4 | 55.4 | 54.5 | 53.8 | 53.5 | 53.2 | 52.8 | 52.7 | **52.3** |
| | 1.56 | 55.9 | 54.3 | 54.3 | 53.5 | 53.3 | 53.1 | 52.8 | 52.8 | 52.5 | **52.3** |

| | | 2OG [µM] | | | | | | | | | No 2OG |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 100.00 | 50.00 | 25.00 | 12.50 | 6.25 | 3.13 | 1.56 | 0.78 | 0 | No $MnCl_2$ |
| $MnCl_2$ [µM] | 100.00 | 9.5 | 8.9 | 9.7 | 7.7 | 6.3 | 3.7 | 3.9 | 2.4 | 1.8 | -0.2 |
| | 50.00 | 6.8 | 5.7 | 7.6 | 6.7 | 5.0 | 4.3 | 2.2 | 2.1 | 1.4 | -0.3 |
| | 25.00 | 8.0 | 6.7 | 6.4 | 5.5 | 4.1 | 2.9 | 2.5 | 1.6 | -0.5 | -0.1 |
| | 12.50 | 7.3 | 5.7 | 4.7 | 4.4 | 2.9 | 2.5 | 1.9 | 1.1 | 0.7 | -0.1 |
| | 6.25 | 6.4 | 4.8 | 4.2 | 3.0 | 2.3 | 2.4 | 1.6 | 1.0 | 0.7 | 0.1 |
| | 3.13 | 5.9 | 3.3 | 3.3 | 2.4 | 1.7 | 1.4 | 1.1 | 0.7 | 0.6 | 0.2 |
| | 1.56 | 3.8 | 2.2 | 2.2 | 1.4 | 1.2 | 1.0 | 0.7 | 0.7 | 0.4 | 0.2 |

**Table 12. DSF experiment with 2 µM MINA53 (Ala26-Val464) showing protein melting temperature (top) and $T_m$ shift (bottom) with reference to pure enzyme under various concentrations of 2OG and $MnCl_2$.** The assay was carried out in 50 mM HEPES buffer pH 7.5 with addition of 50 mM NaCl. Data is presented as a heatmap with values coloured with red – high and green – low. Values in the table are from a single experiment.

### 3.3.2. Compound libraries

A library of ~1200 compounds targeted for JmjC-type oxygenases was present at the SGC. The library consisted of a set of potential 2OG inhibitors available through collaboration with Professor Christopher Schofield (Chemistry Research Laboratories,

University of Oxford, UK), analogues of those from commercial sources, and compounds from Professor Antonello Mai A. Mai (Drug Chemistry and Technologies Department at the Sapienza University, Rome, Italy), which included hydroxamic acids and uracil based derivatives. The library included typical scaffolds of which many are described in Chapter 1 paragraph 1.6 such as: quinolines and isoquinolines including IOX1, pyridine dicarboxylates, pyrazolopyridine, N-oxalyl amino acids including hydrazine analogues such as Daminozide, TCA cycle intermediates, hydroxamic acids including SAHA and TSA, catechols, flavonoids, bipyridyls, benzoylglycines, amidopyridines and analogues thereof. The library was available in a 96-well plate format with compounds solubilised in DMSO at concentrations of 10 mM, 2 mM or 0.5 mM.

### 3.3.3. Library screening

The compound libraries were screened in single experiments at 10, 25 or 50 μM concentration with 1 μM concentration of protein in HEPES buffer supplemented with $NiCl_2$ for NO66, and $MnCl_2$ for MINA53, which resulted in optimal signal to noise ratios.

#### 3.3.3.1. MINA53

Full-length MINA53 from bacterial expression was used in the DSF experiments. The typical melting temperature of the DMSO reference was similar between the two purifications used for DSF, with a melting temperature of 50.1˚C ± 0.7˚C (n=18) and 51.3˚C ± 0.9˚C (n=99), respectively.

The highest scoring hits of the MINA53 DSF screen included 2OG mimetics such as IS100 or JM387 showing $T_m$ shift of ~3˚C with NOG inducing stabilisation of 1.5-2.8˚C at 10 μM concentration. We also identified a series of hydroxamic acids MC2402 and MC2421, supplied by professor Antonello Mai, which showed melting temperature increases in a region of 2-3°C for MINA53 (Table 13). The hydroxamic acids were primarily designed as histone deacetylases (HDAC) inhibitors [201]. The well-established chemistry relating to HDACs makes the hydroxamic acid class of compounds readily accessible and

could yield useful hits for both NO66 and MINA53. Flavonoids such as quercetin resulted in $T_m$ shift of 1.5-2˚C. The other tested compound classes including bipyridyls, hydroxyquinolines or triazolopyridines did not show any increase in melting temperature.

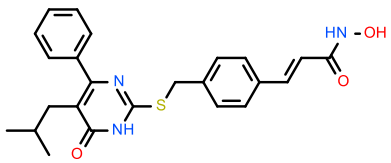
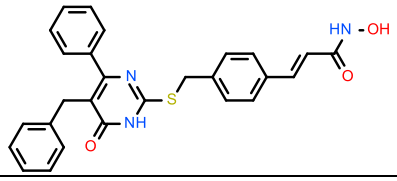
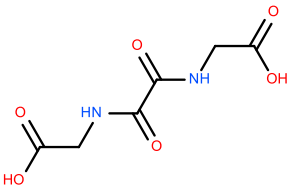
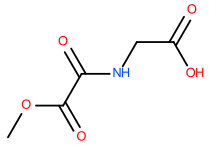
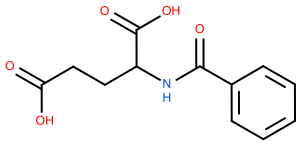
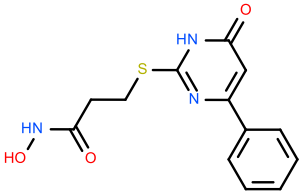
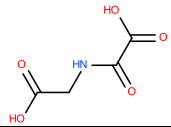
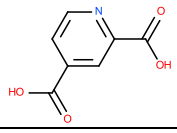
| Description | $T_m$ shift/ concentration | Structure and ID | MW [Da] | $K_d$ [µM] |
|---|---|---|---|---|
| MC2402<br><br>HDAC inhibitor | 3.3°C<br>50 µM | S10011a  | 435.16 | 9.1±1.9 |
| MC2421<br><br>HDAC inhibitor | 3.2°C<br><br>50 µM | S09899a  | 469.15 | 42±13 |
| IS100 | 3.1°C<br>25 µM | S10961a  | 204.04 | - |
| JM387 | 2.8°C<br>25 µM | S10179a  | 161.03 | - |
| N-oxalyl-L-alanine | 2.4°C<br>10 µM | S01369a  | 251.10 | - |
| MC1648 | 1.7°C<br>50 µM | S09926a  | 291.10 | - |
| **NOG** | 1.5-2.8°C<br>10 µM | S00123  | 147.02 | CND |
| **2,4-PDCA** | 1.5-2.8°C<br>10 µM | S00174d  | 167.02 | CND |

**Table 13. Best hits stabilising MINA53 identified by DSF at concentrations of 10-50 µM, ordered by highest $T_m$ shift. The corresponding $K_d$ determined with BLI is added in the last column. CND- cannot be determined.**

#### 3.3.3.2. NO66

DSF experiments were performed with two constructs of NO66, comprising residue ranges Ala167-Asn641 and Ser183-Asn641, with a typical result shown in Figure 23. For NO66 the DMSO reference melting temperature was variable both between the constructs and between purifications. For the construct covering residues Ser183-Asn641 the average melting temperature was 65˚C ± 2˚C (n=199). Thus, for the slightly larger protein covering residue Ala167-Asn641, an average melting temperature of 48˚C ± 1.9˚C (n=169) was determined from one purification whereas another purification of the same construct had an average melting temperature of 62.8˚C ± 1.0˚C (n=60). This effect may be a result of variation in purification quality or protein degradation/precipitation, which might be due to the N-terminal residues which are located within a flexible region of the protein. However, the observed variation in melting temperature should not have large effects on the results since the $T_m$ shift assay in a multi-well format is a relative measure, and the whole 96-well plate would be affected uniformly.

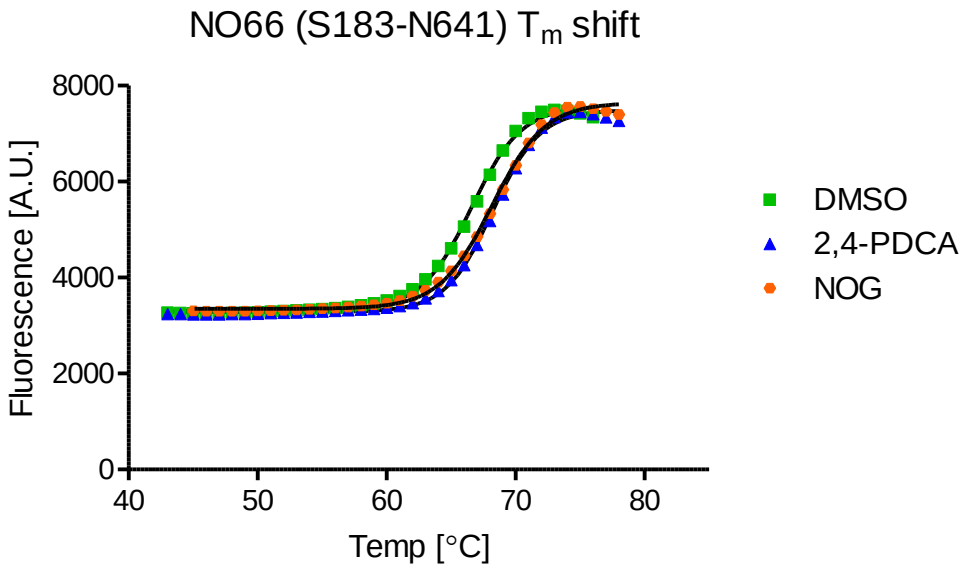



**Figure 23. Typical melting curve for 1 µM NO66 (Ser183-Asn641) with 10 µM concentrations of NOG and 2,4-PDCA in 50 mM HEPES buffer pH 7.5, 150 mM NaCl supplemented with 50 µM $NiCl_2$.** The observed melting temperature for NO66 with DMSO was 66.6˚C with $T_m$ shift of 1.8˚C for both NOG and 2,4-PDCA. The curve fitting is indicated with a solid black line.

Representative scaffolds displaying melting temperature increases for NO66 are shown in Table 14. The best hits included hydroxamic acids with long aliphatic chains such as M334 histone deacetylase inhibitor III and Oxamflatin with a $T_m$ shift of 5.2-5.7˚C or TSA and Scriptaid with ~3˚C stabilisation. Another large set of stabilising agents were 2OG mimetics such as JM134 with the top melting temperature stabilisation of 5.9˚C or IS100 with a $T_m$ shift of 3.1˚C, whereas NOG increased melting temperature by 1.8-3˚C and 2,4-PDCA by 1.5-1.8˚C. Flavonoids such as quercetin have shown an average of 2.7˚C stabilisation. The DSF screen has also identified possibly alkylating agents T5590498 and T0514-2968 with $T_m$ shift of ~4.5˚C, which however did not alkylate NO66 or MINA53 as determined by mass spectrometry. Other scaffolds, which have increased the melting temperature in a range of 1.5-2.5˚C were triazolopyridines, bipyridyls, aminopiridines and N-oxalyl amino acids.

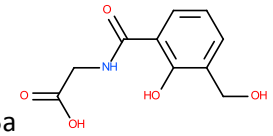
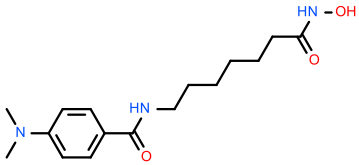
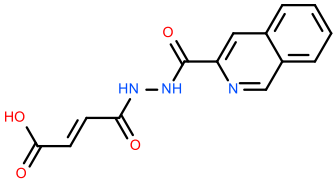
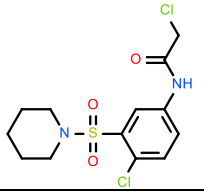
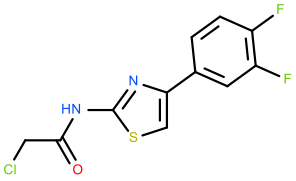
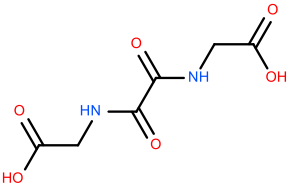
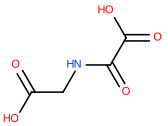
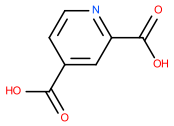
| Description | $T_m$ shift/ concentration | Structure and ID | MW |
|---|---|---|---|
| JM124 | 5.9°C<br>10 µM |  <br>S01385a | 225.1 |
| M344, Histone Deacetylase Inhibitor III | 5.7°C<br>10 µM |  <br>S01121a | 307.2 |
| JM134 | 5.2°C<br>10 µM | S00860a<br>  | 285.1 |
| T5590498 | 4.8°C<br>25 µM | E05912a<br>  | 350 |
| T0514-2968 | 4.6°C<br>25 µM | E05906a<br>  | 288 |
| IS100 | 3.1°C<br>25 µM | S10961a<br>  | 204 |
| **NOG** | 1.8-3.0°C<br>10 µM | S00123<br>  | 147.02 |
| **2,4-PDCA** | 1.5-1.8°C<br>10 µM | S00174d<br>  | 167.02 |

**Table 14. Selection of best scoring compounds for NO66 identified by DSF at a concentration range of 10-30 µM, ordered by highest $T_m$ shift.**

### 3.3.4. Verification of identified DSF hits by BioLayer Interferometry (BLI)

The affinity of selected compounds identified by DSF was tested further by BLI [202] for MINA53 (see data in Table 13). The assay uses streptavidin coated sensors which immobilise biotinylated protein. The continuous readout is generated as nanometer shift in wavelength of light reflected between the tip of the sensor and the immobilised protein. An increase in the nm shift corresponds to molecule binding. Association and dissociation curves can be measured for multiple concentrations of compound allowing direct determination of the binding constant – $K_D$. The assay is more sensitive for compounds with molecular masses greater than 300 Da. The best scoring hits from DSF, showing a $T_m$ shift greater than 2.5°C were selected for BLI analysis. In conclusion, the group of hydroxamic acids showed clear results with binding constants in a range of 9 to 42 µM such as S01011a shown in Figure 24 (DSF results in Table 13).

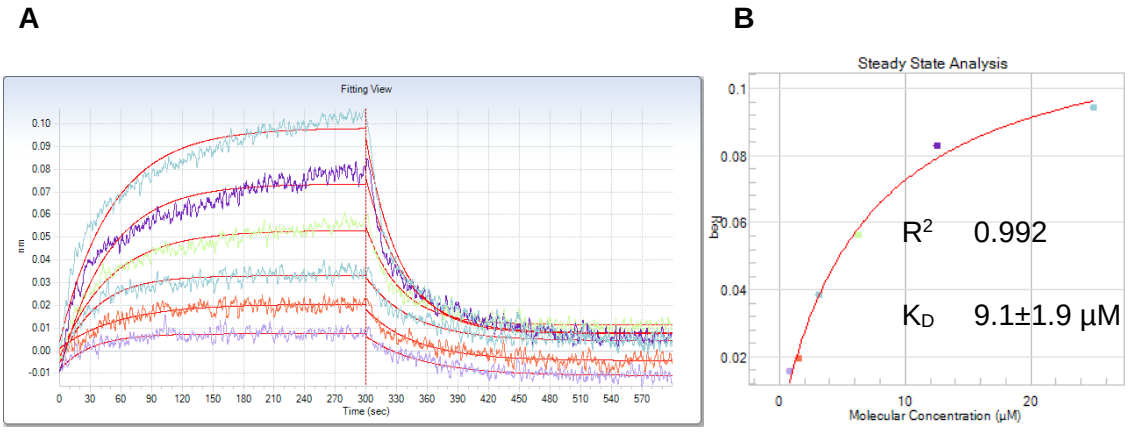



**Figure 24. A: Association and dissociation experiments show binding of S01011a (HDAC inhibitor, $T_m$ shift stabilisation 3.3°C) with MINA53. B: Determination of the binding constant of S01011a (HDAC inhibitor). Y-axis is the nm shift (measurement), X-axis is the compound concentration.**

## 3.4. Mass spectrometry RapidFire activity assay for NO66 and MINA53

### 3.4.1. Introduction

Multiple medium- and high-throughput assays for $Fe^{2+}$- and 2OG- dependent oxygenases have been reported to date [191], [203]. The fluorescence-based coupled FDH enzyme assay measures the oxidation of the by-product formaldehyde to formic acid by the coupling enzyme formaldehyde dehydrogenase (FDH) with a simultaneous reduction of nicotinamide adenine dinucleotide ($NAD^+$) to NADH [203]. Fluorescence and chemiluminescence (AlphaScreen) based assays allow for high-throughput screening, although auto-fluorescence compounds can result in many false-positives [204]. As discussed above, the tested constructs of MINA53 and NO66 lack *in vitro* demethylase activity (see 3.4.2, Figure 25), but efficiently hydroxylate ribosomal peptides [28]. Neither the FDH coupled enzyme assay nor the AlphaScreen based assays were applicable to detection of ribosomal hydroxylation since (1) no specific antibodies have been reported for either of the hydroxylated ribosomal peptide products and (2) formaldehyde is not a by-product of the hydroxylation reaction. Therefore we wanted to establish an *in vitro* assay that would allow characterising the selectivity and potency of the compounds identified by DSF and BLI. Consequently, we resorted to mass spectrometry assays, detecting the mass shift of +16 Da upon formation of the hydroxylated product. A label-independent mass spectrometry based assay such as matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) is popular in the demethylase field since both substrate and product can be monitored directly, however the achievable throughput is limiting. The recent development of RapidFire mass spectrometry (RF MS) has overcome the throughput limitation of the mass spectrometry approach [204].

RapidFire is a high throughput method very suitable for targets such as NO66 and MINA53, and it has also proven to be a method of choice for other targets in the epigenetic

field [204], [205]. RF MS has the advantage of an automated sample injection system that can be coupled to either a q-TOF or triple quad mass spectrometer [126].

Enzymatic turnover catalysed by 2OG oxygenases depends on the presence of 2OG, substrate, oxygen and $Fe^{2+}$ which in turn can be controlled by various reducing agents such as ascorbic acid [145]. Both ascorbic acid and iron are often kept in excess relative to the enzyme concentration, whereas the concentration of 2OG around its $K_M$ value is used in order to focus on identification of 2OG competitive inhibitors. Assays conducted at different iron concentrations are also used to identify inhibitors that depend on the concentration of active $Fe^{2+}$ and to filter out compounds that simply chelate $Fe^{2+}$ in solution [151]. In order to develop an activity-based assay it is necessary to appreciate the multi-substrate enzyme kinetics that govern 2OG oxygenases.

### 3.4.2. RapidFire mass spectrometry activity assay

No *in vitro* activity of NO66 and MINA53 was detected towards H3K4me3, H3K36me2, H3K9me3 and H3K27me3 (Figure 25) and we therefore progressed to implementation of a hydroxylation activity assay with ribosomal peptides [30] and established a RF MS based assay for both NO66 and MINA53 monitoring simultaneously the direct appearance of the hydroxylated product (Figure 27). For NO66 we utilised a synthetic peptide containing 20 amino acids of Rpl8 (Asn205 - Thr224, NPVEHPFGGGN**H**QHIGKPST) and for MINA53 a synthetic peptide containing 20 – residues of Rpl27a (Gly31 - Pro49, GRGNAGGL**H**HHRINFDKYHP), and both the disappearance of substrate and the simultaneous formation of Rpl8-His216OH and Rpl27a-His39OH, respectively, were detected. These activity assays have used TEV-cleaved enzymes.

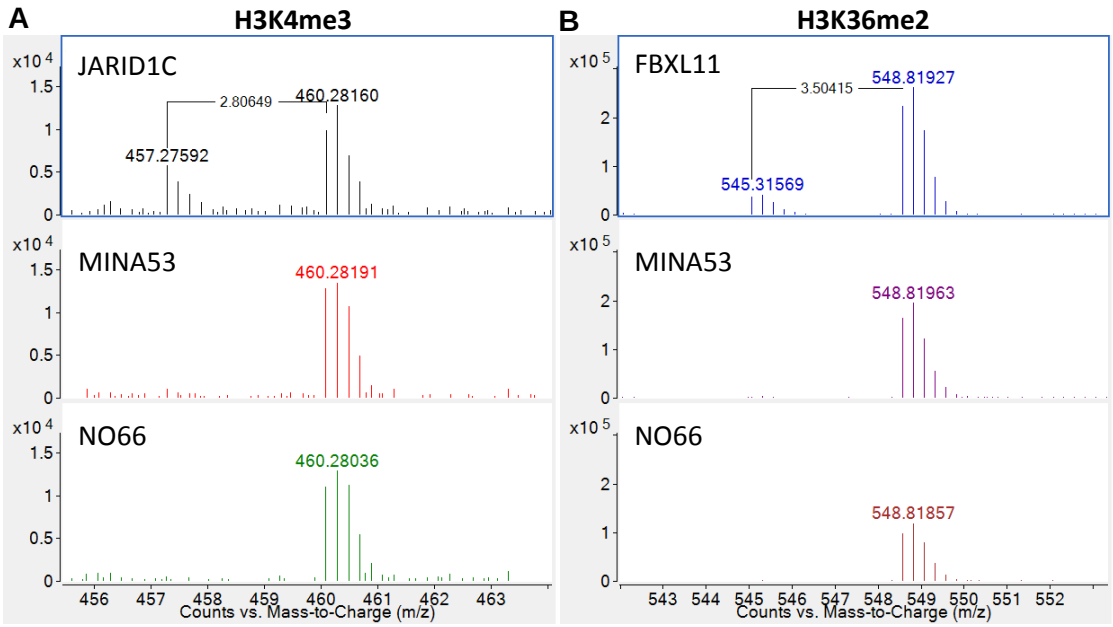



**Figure 25. Mass spectra showing no *in vitro* activity of NO66 (Ala167-Asn641) and MINA53 (Ala26-Val464) against: A. charge state 5 of H3 Ala1-Ala21 K4me3 with JARID1C as positive control. B. charge state 4 of H3 Ser28-Leu48 K36me2 with FBXL11 as positive control.**

RF MS employs a low volume clean up cartridge and allows for samples to be aspirated under vacuum directly onto a solid phase extraction (SPE) cartridge substantially speeding up the sample processing. The sample is first injected onto a SPE cartridge, followed by a wash with aqueous wash buffer ($H_2O$, 0.1% (v/v) Formic Acid) and then eluted with elution buffer (85% (v/v) acetonitrile, 0.1% (v/v) formic acid in $H_2O$) directly onto a QTOF mass spectrometer. The whole cycle time of aspiration to elution takes 6 – 8 seconds. A 384-well plate can therefore be analysed in a time of about 45 min.

The injection system allows for optimisation of wash and elution times of the sample. In order to maximise the readout signal we have tested a range of peptide concentrations using a C4 SPE cartridge with wash and elution times ranging from 3000 ms to maximum allowed 6000 ms. Optimum times for both MINA53 and NO66 peptides has shown little variability and was determined to be between 5500-6000 ms for elution and 5500 ms for wash.

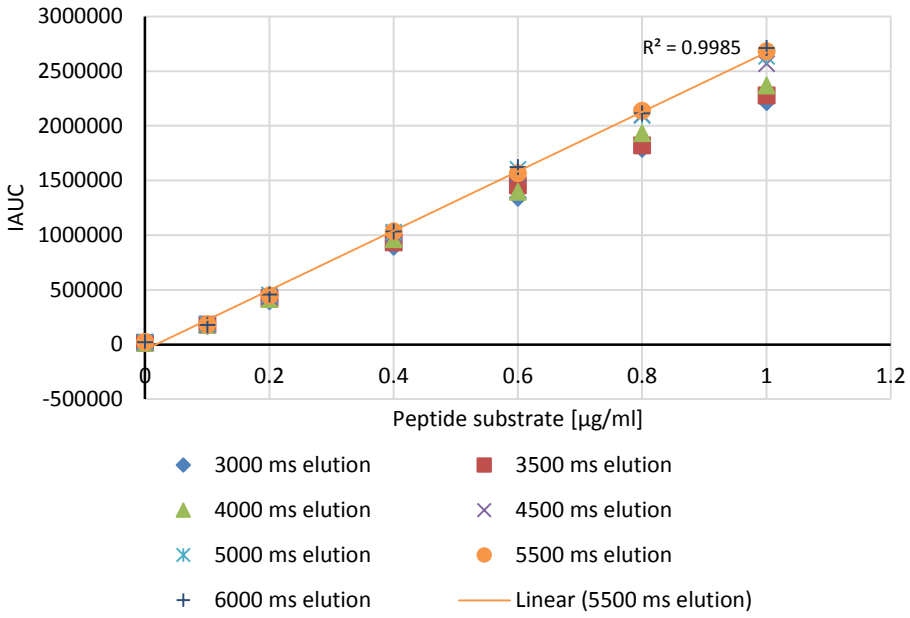



**Figure 26. Typical elution optimisation standard curve for Rpl27a peptide with cartridge C4.** The elution time of 5500 ms gives maximum signal and is linear with $R^2$=0.9985. Peptide concentration ranges from 0.1µg/ml to 1µg/ml with blank injection of water representing 0 concentration.

A typical enzyme progress curve is shown in Figure 27. The resulting hydroxylated product is quantified by extracting the charge state with the highest signal for the peptide (Extracted Ion Chromatogram – EIC) and integrating the area under the curve (IAUC) (see supplementary G for data analysis methods). Proteins used in assay development were initially purified from *E. coli*, but when insect cell expression became available, we have moved to longer constructs (Gly116 truncation for NO66 and full-length MINA53) from baculovirus.

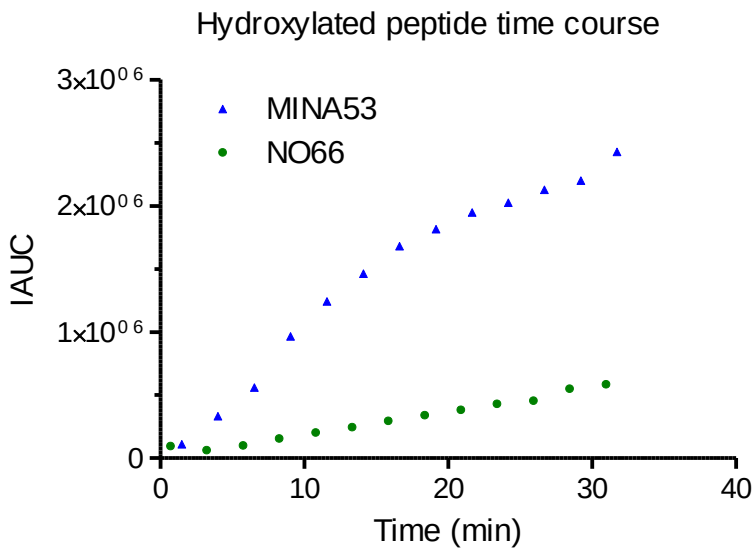



**Figure 27. Enzyme progress curve, monitoring appearance of hydroxylated Rpl8-His216OH and Rpl27a-His39OH peptides for NO66 (Gln116-Asn641) and MINA53 (Met1-Val464), respectively.** The reaction was stopped on the mass spectrometer by injection onto the C4 cartridge. Both reactions are conducted at the final conditions specified in Table 15, n=1. The MS signal represents an integrated EIC of the peptide based on the dominant charge state.

To validate that the inhibition of hydroxylation is purely an effect of the tested compound and not of the DMSO solvent, we evaluated the assay sensitivity to 1% DMSO, representing the final concentration in the reaction mixture. Both NO66 and MINA53 showed very little or no inhibition of hydroxylation as depicted in Figure 28. That gave us confidence to assume inhibition as sole effect of the ligand.

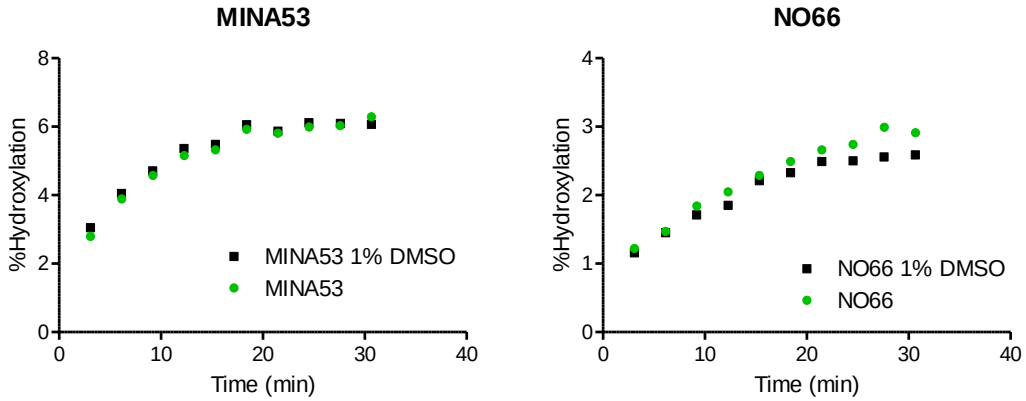
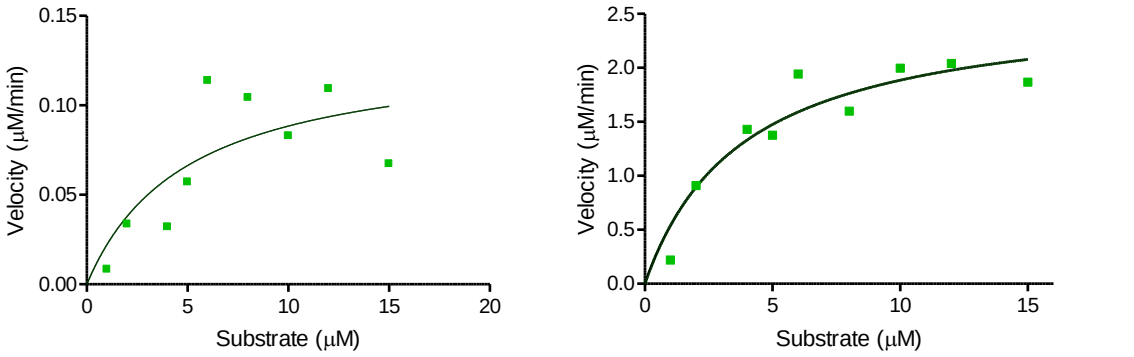



**Figure 28. Enzyme activity of MINA53 (M1-V464) is not inhibited by addition of 1% DMSO, n=1.** NO66 (Gln116-Asn641) appears to be more sensitive but tolerates 1% DMSO well. Data presented here came from initial purification of both enzymes, later purifications have proven to be more active.

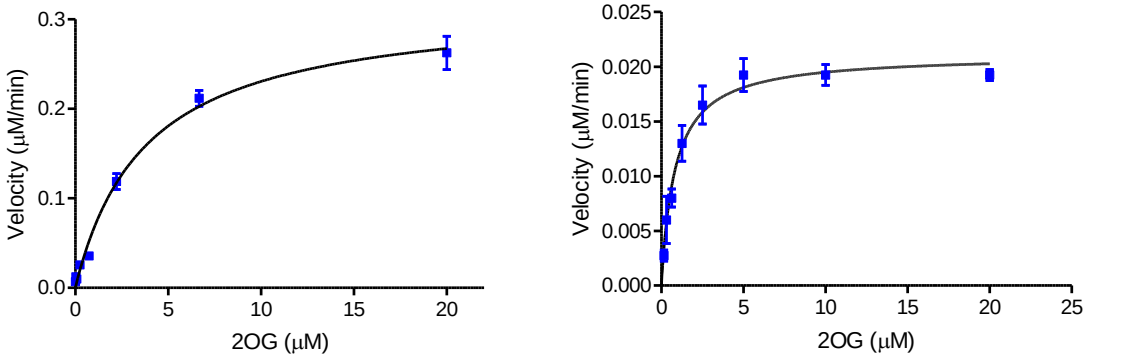
We also determined the affinity for the synthetic peptides by titrating the peptide concentration from 1 µM to 15 µM, yielding $K_M$ values of 5.0 ± 8.1 µM for the MINA53 substrate and 3.8 ± 3.4 µM for the NO66 substrate (Figure 29).




| MINA53 peptide $K_M$ | NO66 peptide $K_M$ |
|---|---|
| $K_M$ = 5.0 ± 8.1 µM | $K_M$ = 3.8 ± 3.4 µM |
| $V_{max}$ = 0.13 ± 0.12 µM/min | $V_{max}$ = 2.6 ± 0.8 µM/min |

**Figure 29. Determination of the enzymatic constant $K_M$ for substrate with NO66 (Gln116-Asn641) and MINA53 (Met1-Val464). Values are means ± 95% confidence, n=1.**

For assay optimisation we also determined the affinity for 2OG yielding a Michaelis-Menten constant $K_M$ of 3.2 ± 0.6 µM and 0.8 ± 0.2 µM for MINA53 and NO66 respectively (Figure 30).




| MINA53 2OG $K_M$ | NO66 2OG $K_M$ |
|---|---|
| $K_M$ = 3.2 ± 0.6 µM | $K_M$ = 0.8 ± 0.2 µM |
| $V_{max}$ = 0.32 ± 0.03 µM/min | $V_{max}$ = 0.021 ± 0.001 µM/min |

**Figure 30. Determination of the $K_M$ constant for 2OG with MINA53 (Met26-Val464) and NO66 (Ser183-Asn641). Values are means ± 95% confidence, n=4.**

We have also calculated the $K_M$ constant for $Fe^{2+}$ which was 0.5 ± 0.2 µM for MINA53 and 1.1 ± 0.3 µM for NO66 (Figure 31).

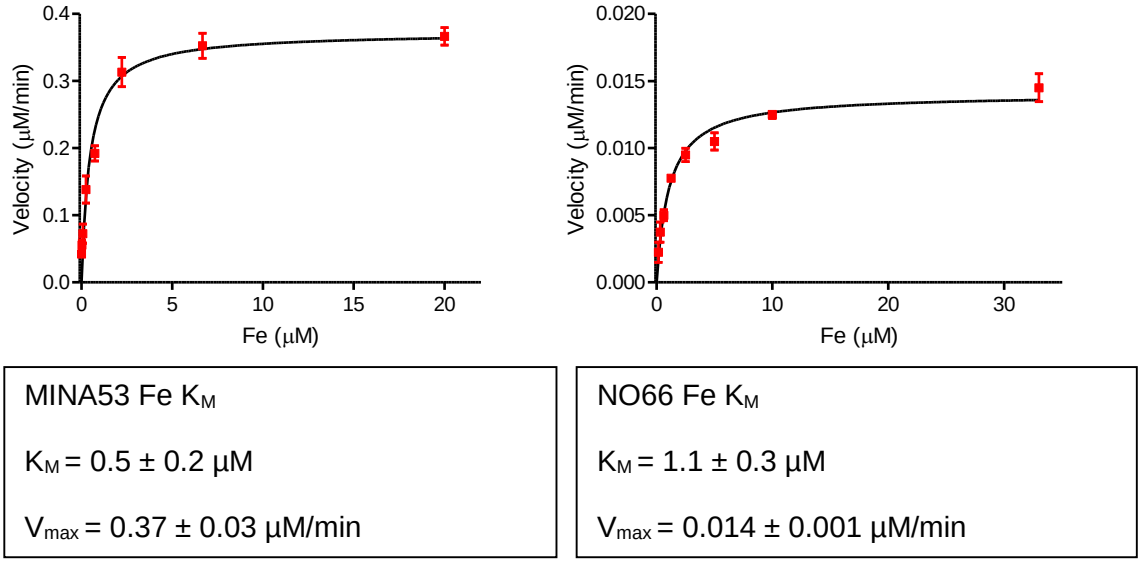



**Figure 31. Determination of the $K_M$ constant for $Fe^{2+}$ with MINA53 (Met26-Val464) and NO66 (Ser183-Asn641). Values are means ± 95% confidence, n=4.**

Initially reported hydroxylation assay parameters [30] were therefore optimised to the final assay conditions (Table 15). Based on the time courses with the final assay conditions we have decided to stop the reaction at the end of the linear region of catalytic turnover which was 30 min for MINA53 and ~60 min for NO66. We have also decided to use 5 µM peptide for both assays since this concentration resulted in an excellent mass spectrometry signal.

| | *Final Assay Concentration* |
|---|---|
| Enzyme | 0.150 µM MINA53<br>0.300 µM NO66 |
| Substrate | 5 µM |
| 2OG | 2 µM |
| Fe(II) (FAS) | 50 µM |
| Ascorbate (LAA) | 100 µM |

**Table 15. Final concentration of the Rapid Fire assay optimised for 2OG and peptide substrate. LAA – L-ascorbic acid, FAS – Ferrous ammonium sulphate.**

### *Buffer optimisation and temperature dependence*

Different buffers, pH, salt and other additives can have an effect on enzyme stability and thus enzyme activity [206]. For example, in a turnover progress curve the loss of enzyme

activity during the reaction leads to lower maximum plateau value of product formation as shown in Figure 32. The same effect can also be observed with altered concentration of one of the cofactors, cosubstrate 2OG or peptide leading to a possible source of error, however these errors can be limited by experimental design, for example by using the same pool of cofactors and cosubstrates for parallel experiments, which was used in this case. The effect of enzyme stability in various buffer compositions can therefore be measured by monitoring the final enzymatic turnover.

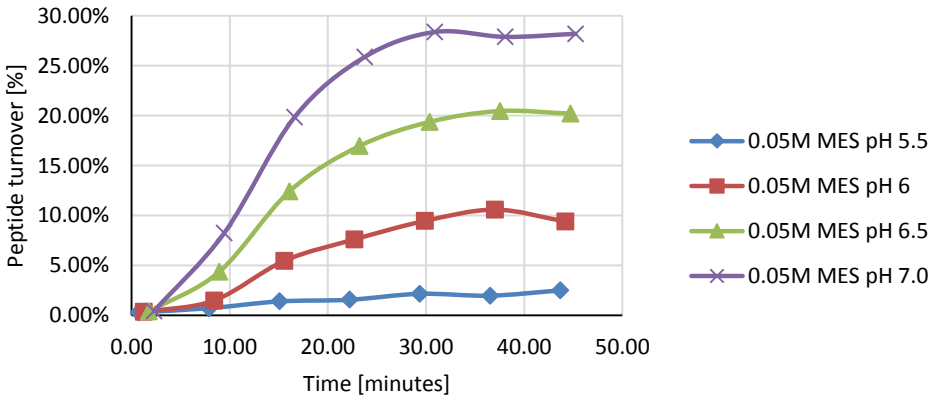



**Figure 32. Typical enzyme time-course with MINA53 (Met1-Val464) for a range of pH 5.5 – 7.0 in MES buffer.** Maximum peptide turnover decreases with decreasing pH indicating loss of enzyme activity during the reaction. Time-courses were measured in parallel in a single RapidFire experiment with enzyme concentration of 300 nM and remaining cofactors and cosubstrates under final conditions of Table 15.

The buffer composition of the activity assay was optimised in a range of pH, buffers and salt concentrations under final assay conditions from Table 15. We have refrained from using other enzyme stabilisation reagents such as detergents, also used to avoid aggregation [186], or bovine serum albumin additives as they were reported to interfere with the readout of RapidFire mass spectrometer [Dr A. Tumber, personal communication]. Results of optimisation of buffer compositions are shown in Figure 33 for NO66 and Figure 34 for MINA53.

NO66 appears to be sensitive to buffer types and maximum activity appears in a narrow pH range oscillating around pH 7.0 with both HEPES and MES buffers (Figure 33). Tris, BIS-TRIS and phosphate buffers display limited activity. The optimal assay conditions for NO66 were MES buffer pH 7.0 with a range of NaCl or KCl concentrations, or HEPES

buffer pH 7.0 with no NaCl (Figure 33). Both buffers were suitable for the assay with respect to optimal enzyme activity, but in RapidFire mass spectrometry MES buffer is preferable to HEPES as the ionisation signal is substantially lower than that of HEPES decreasing the assay interference. Although additional NaCl salt has shown to decrease the enzymatic turnover by ~20%, we decided to keep NaCl in the buffer as a general protein stabilisation agent at 150 mM. The final NO66 assay buffer was therefore 50 mM MES pH 7.0 supplemented with 150 mM NaCl.

| | | pH | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 5.5 | 6 | 6.5 | 7 | 7.5 | 8 | 8.5 | 9 | |
| NaCl | 0 mM | 3% | 11% | 15% | 30% | | | | | MES |
| | 50 mM | 2% | 4% | 12% | 25% | | | | | |
| | 100 mM | 1% | 4% | 8% | 24% | | | | | |
| | 150 mM | 1% | 3% | 6% | 24% | | | | | |
| KCl | 50 mM | 2% | 4% | 9% | 24% | | | | | |
| | 100 mM | 1% | 3% | 9% | 26% | | | | | |
| NaCl | 0 mM | | | | 27% | 15% | 14% | | | HEPES |
| | 50 mM | | | | 18% | 12% | 12% | | | |
| | 100 mM | | | | 14% | 12% | 11% | | | |
| | 150 mM | | | | 13% | 13% | 10% | | | |
| NaCl | 0 mM | | | | | 3% | 2% | 2% | 1% | Tris |
| | 50 mM | | | | | 4% | 2% | 1% | 1% | |
| | 100 mM | | | | | 5% | 2% | 5% | 1% | |
| | 150 mM | | | | | 6% | 1% | 1% | 1% | |
| KCl | 50 mM | | | | | 5% | 1% | 1% | 1% | |
| | 100 mM | | | | | 5% | 1% | 1% | 1% | |
| NaCl | 0 mM | | 6% | 6% | 3% | 1% | | | | BIS-TRIS |
| | 50 mM | | 6% | 6% | 3% | 2% | | | | |
| | 100 mM | | 5% | 5% | 4% | 2% | | | | |
| | 150 mM | | 4% | 5% | 4% | 2% | | | | |
| KCl | 50 mM | | 4% | 7% | 3% | 1% | | | | |
| | 100 mM | | 4% | 6% | 3% | 1% | | | | |

| pH | 5.4 | 5.6 | 5.8 | 6 | 6.2 | 6.4 | 6.6 | 6.8 | 7 | 7.2 | 7.4 | 7.5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Phosphate | 1% | 2% | 3% | 4% | 4% | 3% | 3% | 3% | 4% | 2% | 2% | 2% |

**Figure 33. Percentage peptide turnover for NO66 (Ala167-Asn641) after 60 minutes incubation with 50 mM MES, HEPES, Tris, BIS-TRIS or phosphate buffers in a range of NaCl or KCl salt concentrations.** NO66 was used at concentration of 300 nM, with 5 µM 2OG, 5 µM peptide substrate, 50 µM FAS and 100 µM LAA.

The buffer conditions for MINA53 were optimised similar to NO66 (Figure 34). MINA53 appears active over a range of pH 6.5-8.0 in MES, HEPES, Tris and BIS-TRIS buffers.

Phosphate buffer shows the lowest turnover of all tested buffers with a maximum between pH 5.8 and 6.4. A marginal increase in activity could be observed with salt concentrations higher than 150 mM, although the activity trend with salt concentration appears to be buffer specific. In BIS-TRIS buffer a higher salt concentration leads to increased activity, whereas in HEPES buffer a low salt concentration is preferred. The activity in HEPES buffer shows only a small variation in pH 7.0 - 8.0 and also a preference for low salt conditions. The final assay condition for MINA53 was chosen based on maximum turnover to be 50 mM HEPES pH 7.5 with addition of 50 mM NaCl.

| | | pH | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 5.5 | 6 | 6.5 | 7 | 7.5 | 8 | 8.5 | 9 | |
| NaCl | 0 mM | 3% | 15% | 24% | 22% | | | | | MES |
| | 50 mM | 3% | 11% | 32% | 27% | | | | | |
| | 100 mM | 3% | 13% | 26% | | | | | | |
| | 150 mM | 2% | 6% | 17% | 23% | | | | | |
| KCl | 50 mM | 3% | 11% | 26% | 36% | | | | | |
| | 100 mM | 2% | 10% | 26% | 30% | | | | | |
| NaCl | 0 mM | | | | 36% | 40% | 39% | | | HEPES |
| | 50 mM | | | | 39% | 41% | 38% | | | |
| | 100 mM | | | | 38% | 35% | 32% | | | |
| | 150 mM | | | | 31% | 29% | 29% | | | |
| NaCl | 0 mM | | | | | 26% | 24% | 11% | 7% | TRIS |
| | 50 mM | | | | | 25% | 22% | 9% | 5% | |
| | 100 mM | | | | | 26% | 21% | 9% | 6% | |
| | 150 mM | | | | | 19% | 21% | 4% | 2% | |
| KCl | 50 mM | | | | | 27% | 21% | 8% | 3% | |
| | 100 mM | | | | | 30% | 19% | 7% | 2% | |
| NaCl | 0 mM | | 16% | 25% | 21% | 15% | | | | BIS-TRIS |
| | 50 mM | | 22% | 36% | 28% | 19% | | | | |
| | 100 mM | | 16% | 36% | 33% | 22% | | | | |
| | 150 mM | | 25% | 33% | 27% | 16% | | | | |
| KCl | 50 mM | | 22% | 37% | 27% | 17% | | | | |
| | 100 mM | | 26% | 38% | 32% | 18% | | | | |

| pH | 5.4 | 5.6 | 5.8 | 6 | 6.2 | 6.4 | 6.6 | 6.8 | 7 | 7.2 | 7.4 | 7.5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Phosphate | 2% | 8% | 12% | 15% | 11% | 13% | 10% | 10% | 8% | 6% | 6% | 5% |

**Figure 34. Activity (measured as peptide turnover) for MINA53 (Met1-Val464) with 50 mM MES, HEPES, Tris, BIS-TRIS or phosphate buffers in a range of NaCl or KCl salt concentrations.** Full-length MINA53 construct was used at concentration of 300 nM, with 5 µM 2OG, 5 µM peptide substrate, 50 µM FAS and 100 µM LAA.

We have also determined the relationship between activity and incubation temperature with both NO66 and MINA53. The maximum turnover for NO66 was observed at 30˚C and 37˚C for MINA53, followed by a sharp drop in activity at 37˚C for NO66 and 42˚C for MINA53. Overall enzymatic turnover for NO66 was low and oscillating around 9-11%, in comparison to previous tests with buffer optimisation. This may be due to time-dependent enzyme degradation, which was observed with NO66. These results would suggest to conduct the enzymatic activity assays at 21-30˚C for NO66 and 37˚C for MINA53. However, high and medium throughput assays are usually carried out in 96- or 384-well plate formats which requires appropriate equipment to keep the sensitive reaction at constant temperature and at the same time avoid edge effects caused by not uniform heating, even by using a water bath. We have therefore decided to conduct both assays at room temperature.

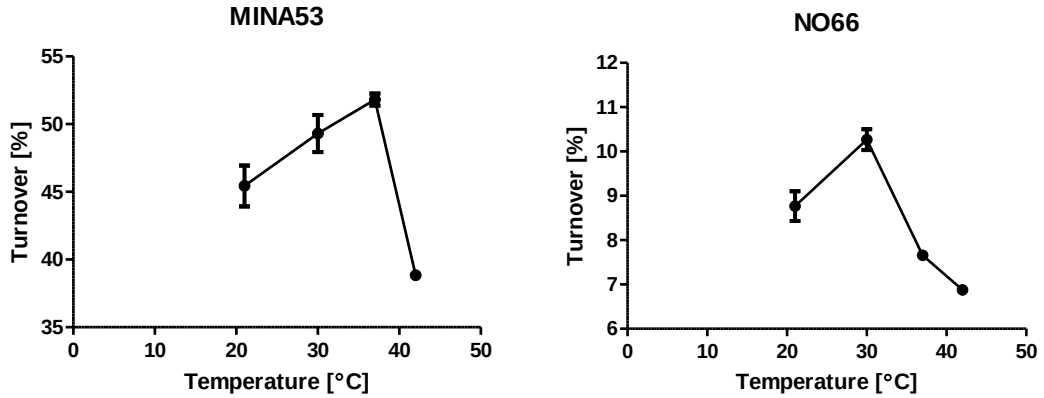



**Figure 35. Temperature-dependent activity of MINA53 (Met1-Val464) and NO66 (Ala167-Asn641).** Percentage turnover with 12 replicates per data point for NO66, and 8 replicates for MINA53 with exception of 21˚C where 5 replicates were used. Assays were performed in a water bath equilibrated to 21, 30, 37 and 42 ˚C and the reaction was stopped after 30 min for MINA53 and after 60 min for NO66.

### *Assay validation with known inhibitors*

The assay was validated with 2,4-PDCA and NOG, two known inhibitors of $Fe^{2+}$- and 2OG-dependent oxygenases showing $IC_{50}$ values of 0.11 ± 0.01 µM and 3.5 ± 0.5 µM for NO66 respectively, and 1.3 ± 0.5 µM and 1.8 ± 0.7 µM for MINA53 respectively (Figure 36). In addition we have also determined the $IC_{50}$ for IOX1, another pan-JmjC inhibitor, to be 38.5 ± 18.0 µM and 101.8 ± 35.1 µM against NO66 and MINA53 respectively. The chemical

probe GSK-J1 [77] showed ~30% inhibition at 100 µM concentration, and was found inactive against MINA53.

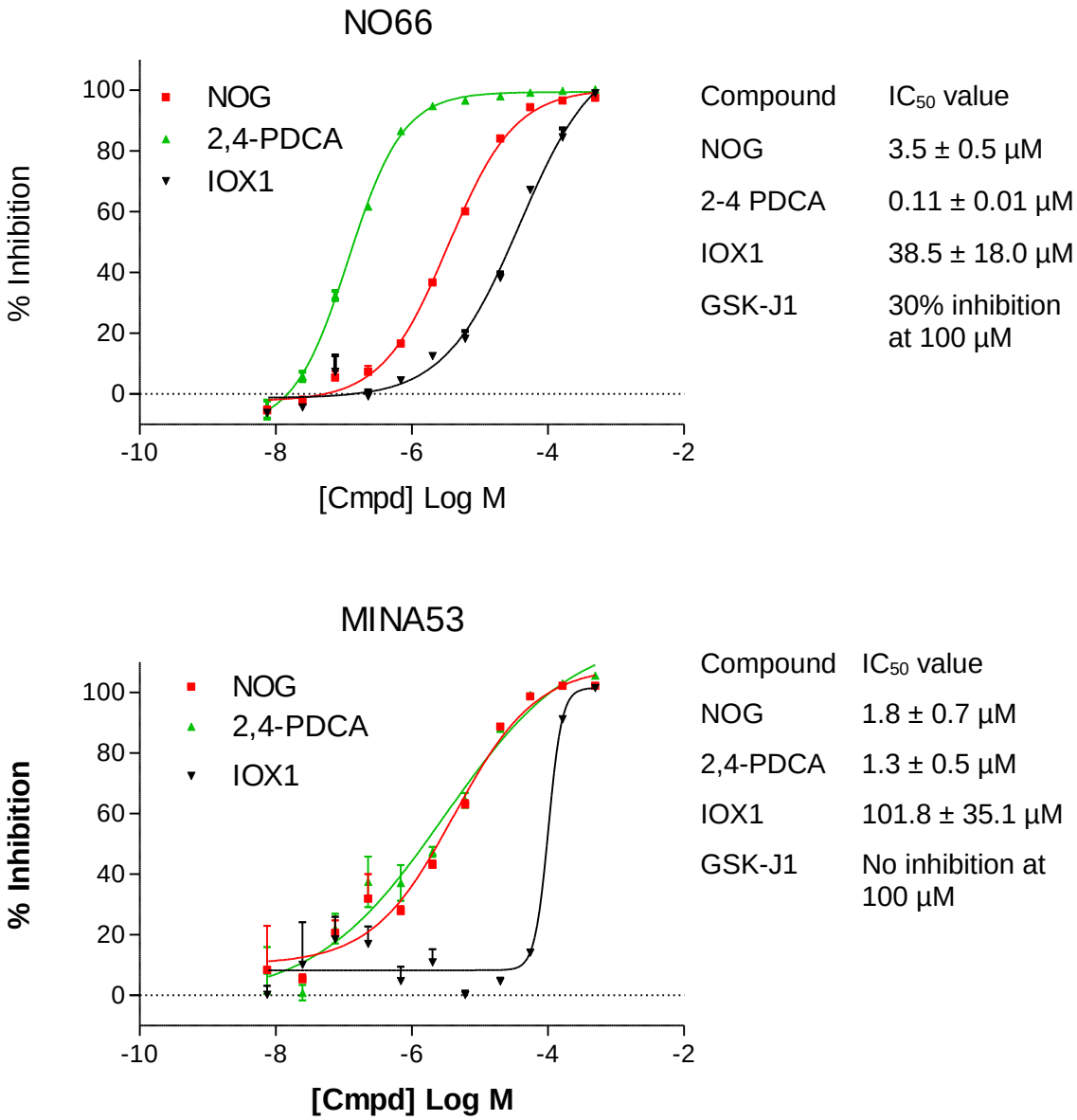



**Figure 36. $IC_{50}$ curves with MINA53 (Met1-Val464) and NO66 (Ser183-Asn641) for 2,4-PDCA and NOG, two known inhibitors of 2-OG oxygenases and IOX1.** GSK-J1 is not shown for clarity. Values are means ± 95% confidence.

### 3.4.3. Inhibition by TCA cycle intermediates

The TCA cycle intermediates pyruvate, succinate, acetate, citrate, isocitrate, fumarate and malate have been tested for inhibition of NO66 and MINA53. The resulting $IC_{50}$ constants are presented in Table 16. NO66 showed inhibition by succinate and fumarate with an $IC_{50}$

in a range of 200-300 µM and oxaloacetate with an $IC_{50}$ of 480 µM (Table 16). MINA53 appears to be more amenable to inhibition by TCA cycle intermediates and is inhibited by D-threo-isocitrate, oxaloacetate, succinate and fumarate in a 10-30 µM $IC_{50}$ range and 11 µM $IC_{50}$ with D-threo-isocitrate, which interestingly does not inhibit NO66 (Table 16). Malic acid and citrate show $IC_{50}$ values of 380 and 150 µM against MINA53, and 1200 µM and ~2700 µM for NO66 respectively. Both NO66 and MINA53 are however not inhibited at the tested concentration with pyruvate. $IC_{50}$ determination with the oncometabolite L-2-hydroxyglutarate (L-2HG) has shown no inhibition against NO66 and MINA53 at all tested concentrations starting with 500 µM. The isomer D-2-hydroxyglutarate (D-2HG) showed no inhibition for NO66 in the $IC_{50}$ experiments and 20-30% inhibition was observed for MINA53, although no clear dose-response could be established.

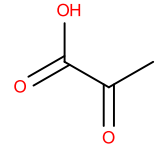
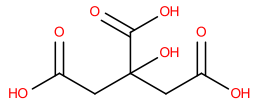
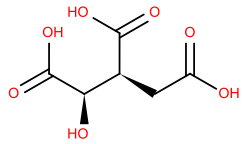
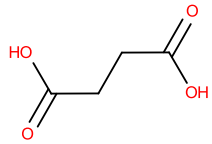
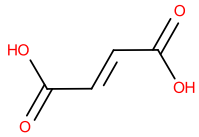
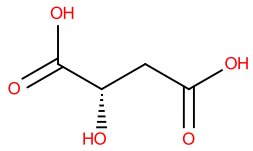
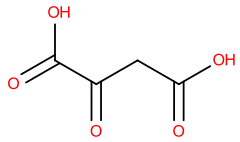
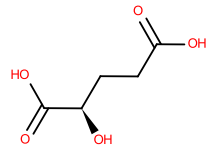
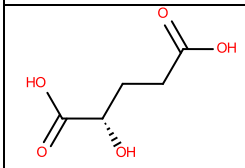
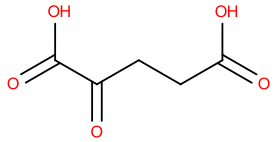
| TCA cycle intermediate | NO66 $IC_{50}$ | MINA53 $IC_{50}$ |
|---|---|---|
|  Pyruvate | inactive | inactive |
|  Citrate | ~2700 µM | 152 ± 68 µM |
|  D-threo-isocitrate | inactive | 11.4 ± 7.5 µM |
|  Succinic acid | 293 ± 57 µM | 19.4 ± 5.6 µM |
|  Fumarate | 211 ± 52 µM | 31.3 ± 16.8 µM |
|  L-(-)-Malic acid | 1212 ± 656 µM | 381 ± 188 µM |
|  Oxaloacetate | 479 ± 91 µM | 26.2 ± 11.2 µM |
|  D-2HG | inactive | Inactive* |
|  L-2HG | inactive | inactive |
|  2OG | Structure depicted for reference. 5 µM 2OG was used in the assay. | |

**Table 16. $IC_{50}$ values of the TCA cycle intermediates with NO66 (Ala167-Asn641) and MINA53 (Met1-Val464).** The $IC_{50}$ curve is composed of triplicate experiments of 11-data point 3-fold dilution series with a starting concentration of 1 mM. Assay conditions were 300 nM enzyme, 5 µM peptide substrate, 5 µM 2OG, 50 µM FAS and 100 µM LAA. $IC_{50}$ is mean ± 95% confidence interval. *marginal inhibition – see main text.

### 3.4.4. Inhibition of MINA53 and NO66 by other known inhibitors of 2OG oxygenases

We decided to screen a representative set of scaffolds known to inhibit 2OG oxygenases. The set included typical scaffolds including many described in chapter 1.6 such as: quinolines and isoquinolines including IOX1, pyridine dicarboxylates, pyrazolopyridine, PLOD inhibitors, N-oxalyl amino acids including Daminozide, hydroxamic acids, catechols, bipyridyls, benzoylglycines, amidopyridines and analogues thereof. Other tested compounds included tricyclic derivatives.

The set of compounds was first tested in NO66 and MINA53 assays in duplicate experiments at a single concentration of 20 μM taking 20 μM 2,4-PDCA as a positive control and a corresponding volume of DMSO as control. With this format 50% inhibition at a concentration of 20 μM would roughly correspond to an $IC_{50}$ value of 20 μM, making our single concentration study sensitive to low micromolar inhibitors. Compounds used for percentage inhibition and $IC_{50}$ determination experiments described in this paragraph were from the same batch, i.e. source plate. Percentage inhibition is reported as an average of two measurements. Standard deviation of percentage inhibition was found to be less than 10% in all of the reported values and is not shown for clarity. $IC_{50}$ experiments were performed in duplicate with 11-point 3-fold dilutions starting from 100 μM concentration unless otherwise stated. 2,4-PDCA was used as a positive control at a concentration of 100 μM and the corresponding volume of DMSO was used as a negative control.

The inhibitory activity of a series of pyridine dicarboxylates is shown in Table 17. 2,4-PDCA is the most active inhibitor of both NO66 and MINA53, as described above (Figure 36) and from the analogues only 2,5-PDCA shows weak inhibition of NO66 at 24% inhibition. 2,4-PDCA binds to the active site iron in oxygenases via the pyridine nitrogen and carboxylic acid oxygen, which may explain the loss of inhibition of 3,4- and 3,5-PDCA where the distance between the pyridine nitrogen and carboxylic acid is increased thus abolishing

the typical metal chelation. 2,4-PDCA binding is further stabilised by hydrogen bond interactions of the carboxylic acid in position 4 with side-chains of active site residues (typically lysine and tyrosine in JmjC-type oxygenases), for which interactions are likely missing in 2,3- 2,5- and 2,6-PDCA, thereby explaining this structure-activity relationship.

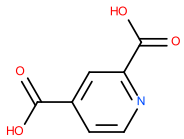
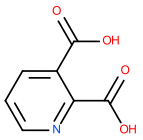
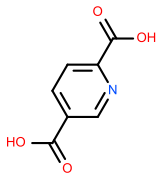
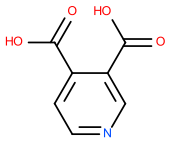
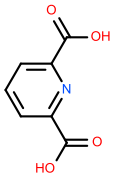
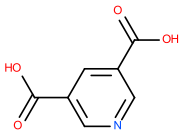
| | NO66 | MINA53 | | NO66 | MINA53 |
|---|---|---|---|---|---|
|  2,4-PDCA | 100% $IC_{50}$ 0.2 µM* | 100% $IC_{50}$ 3.0 µM* |  2,3-PDCA | 8% | -16% |
|  2,5-PDCA | 24% | -5% |  3,4-PDCA | -19% | -1% |
|  2,6-PDCA | -4% | -9% |  3,5-PDCA | -3% | -2% |

**Table 17. Inhibition of NO66 and MINA53 with a series of pyridine dicarboxylic acids.** 2,4-PDCA was used as a positive control compound and therefore shows maximal inhibition. *$IC_{50}$ values calculated from triplicates of an 11-point $IC_{50}$ curve starting at a maximal concentration of 500 µM.

The most potent compound in the series of N-oxalyl amino acids was NOG. Its analogues presented in Table 18 are also shown to weakly inhibit NO66 and MINA53. Interestingly the activity of N-oxalyl-alanine depends on the stereo-chemical configuration, with the L-enantiomer showing low micromolar $IC_{50}$ values and being more potent than the D-enantiomer (Table 18). Other tested N-oxalyl amino acids were N-oxalyl leucine (L/D), valine (L/D), phenylalanine (L), cysteine (L/D) which did not show any inhibition (data not shown).

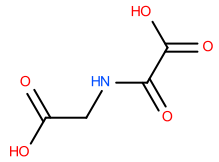
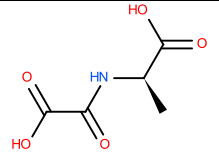
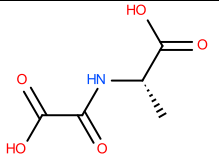
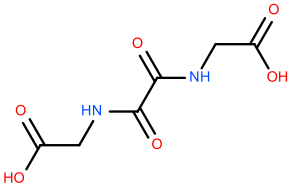
| | NO66 | MINA53 |
|---|---|---|
|  N-oxalyl-glycine (NOG) | 72% $IC_{50}$ 3.4 µM* | 84% $IC_{50}$ 4.2 µM* |
|  N-oxalyl-D-alanine | 4% | 36% |
|  N-oxalyl-L-alanine | 86% $IC_{50}$ 3.8 µM | 73% $IC_{50}$ 7.7 µM |
|  S10961a | 78% $IC_{50}$ 2.8 µM | 31% |

**Table 18. Inhibition of NO66 and MINA53 with N-oxalyl amino acids and analogues.** *$IC_{50}$ calculated from triplicates of an 11-point $IC_{50}$ curve starting at a maximal concentration of 500 µM.

A large group of tested compounds were catechols, known inhibitors of 2OG oxygenases [148] and also PAINS [175], with both gallate and quercetin inhibiting in the 20% range for NO66 and 50-96% for MINA53. The $IC_{50}$ for gallate is ~38 µM for NO66 and 13.0 µM for MINA53 (Table 19).

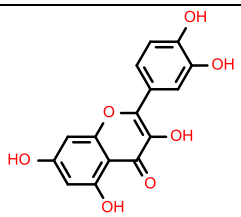
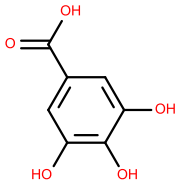
| | NO66 | MINA53 |
|---|---|---|
|  quercetin | 23% | 96% $IC_{50}$ 9.1 µM |
|  gallate | 23% $IC_{50}$ ~38 µM | 56% $IC_{50}$ 13.3 µM |

**Table 19. Inhibition of MINA53 and NO66 by catechol compounds.**

5-carboxy-8-hydroxyquinoline (IOX1) inhibited NO66 at 35% inhibition with a negligible 5% inhibition of MINA53, which translated to an $IC_{50}$ value of 25.2 µM for NO66 and ~80 µM for MINA53. Analysis of $IC_{50}$ measurements for MINA53 using IOX1 (see Table 20) resulted in a steep Hill slope coefficient of 3-5, in contrast with NO66 where the estimated Hill slope coefficient was found to be ~1.5. A steep Hill slope is frequently indicative of aggregation [186]. A set of other quinoline analogues did not show any inhibition for NO66 or MINA53. Representative compounds are shown in Table 20

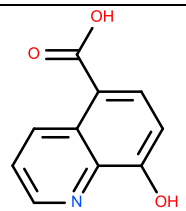
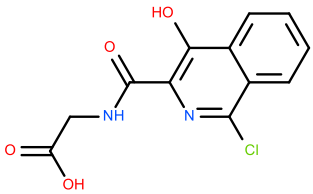
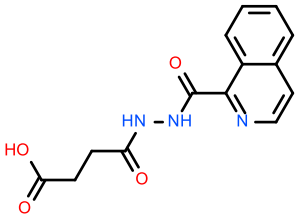
| | NO66 | MINA53 |
|---|---|---|
|  5-carboxy-8-hydroxyquinoline (IOX1) | 35%<br>$IC_{50}$ 38.2 µM* | 5%<br>$IC_{50}$ ~80 µM* |
|  S01391b | -4% | -2% |
|  S00858b | 12% | 17% |

**Table 20. Tested quinoline analogues proved inactive against NO66 and MINA53.** *The IOX1 $IC_{50}$ value for MINA53 was determined to be ~50 µM and 25.3 µM for NO66 using the compound source present in the library used to assess the % inhibition, however a fresh batch of IOX1 resulted in the $IC_{50}$ of ~80 µM for MINA53 and 38.2 µM for NO66, n=3.

Both NO66 and MINA53 were found to be inhibited by hydroxamic acids as described in Table 21. Two clinically approved drugs with the hydroxamic acid motif, namely SAHA and TSA have shown no inhibition of NO66 and MINA53, with the exception of very weak inhibition of SAHA towards MINA53.

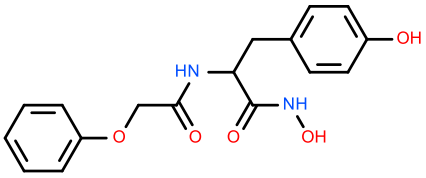
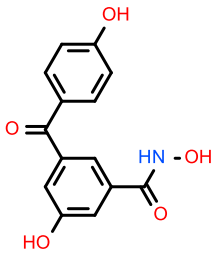
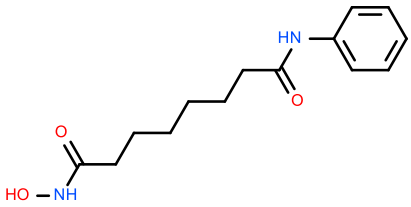
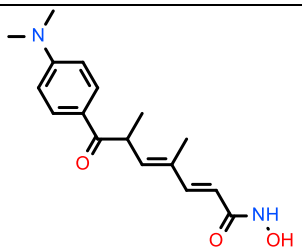
| | NO66 | MINA53 |
|---|---|---|
|  S00133b | 47%<br>$IC_{50}$ ~92 µM | 42%<br>$IC_{50}$ ~28 µM |
|  S00862c | 10%<br>$IC_{50}$~89 µM | 62%<br>$IC_{50}$ 16.0 µM |
|  SAHA | -13% | 24% |
|  TSA | -5% | 5% |

**Table 21. Inhibition of NO66 and MINA53 by hydroxamic acid analogues.**

Other compounds such as bipyridyl derivatives exemplified by S01434b, the pyrazolopyridine scaffolds, benzoylglycine analogues, amidopyridines, tricyclic derivatives, daminozide, and PLOD inhibitors such as Malathion, Malaoxon or Minoxidil were also found to show no inhibition at the tested concentration (Table 22).

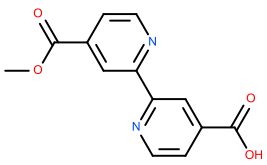
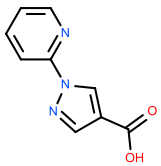
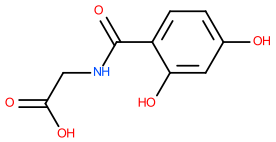
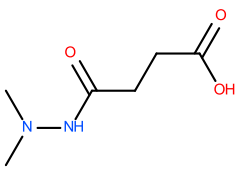
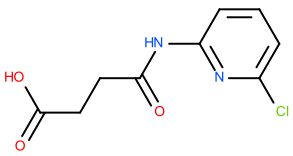
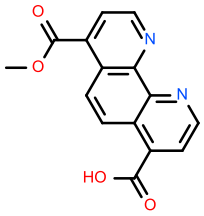
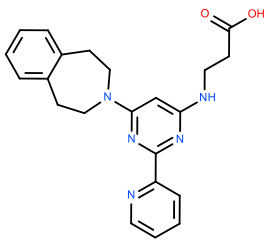
| | | | |
|---|---|---|---|
|  S01434b (bipyridyl) |  S01123b (pyrazolopyridine) |  S00243b (benzoylglycine) |  Daminozide |
|  S00841b (aminopyridine) |  S09477a (tricyclic derivative) |  GSK-J1* | |

**Table 22. Set of compounds inactive against NO66 and MINA53 at tested concentrations.** *$IC_{50}$ calculated in a fresh batch of compound in triplicates of 11-point curve with a 500 µM starting concentration. No inhibition was observed at highest concentration with NO66 or MINA53.

Two chemical probes and tool compounds were also tested, but showed no inhibition: IOX2, a selective inhibitor of the Hypoxia Inducible Factor (HIF) Prolyl-Hydroxylases [151] and GSK-J1, an inhibitor of the KDM6 and KDM5 histone demethylases [77].

Finally, the most promising series of compounds screened in this set included hydroxamic acids and uracil based derivatives shown in Table 23. The top-scoring hits of this series inhibited MINA53 with 15- to 44- fold selectivity against NO66.

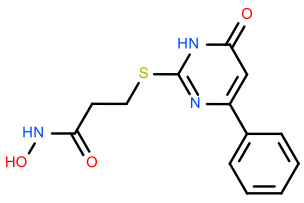
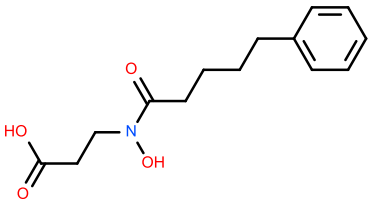
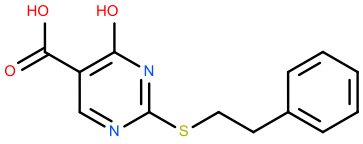
| | | |
|---|---|---|
|  MC1648 NO66 $IC_{50}$ 166 µM MINA53 $IC_{50}$ 6.0 µM |  MC3096 NO66 $IC_{50}$ 18.9 µM MINA53 $IC_{50}$ 160 µM |  MC3098 NO66 $IC_{50}$ 71.7 µM MINA53 $IC_{50}$ 1.6 µM |

**Table 23. Two series of compounds identified as active inhibitors of MINA53 (MC1648, MC3098) and NO66 (MC3096).**

## 3.5. Crystallisation of MINA53 and NO66

### 3.5.1. Aims

In order to elucidate the binding mode of the identified inhibitors we aimed to develop crystallisation systems for both MINA53 and NO66. Since crystal structures of both proteins were already determined, the main focus of this work was to reproducibly grow crystals diffracting to high-resolution, which could be used for co-crystallisation or soaking experiments with identified inhibitors. The aim was to obtain resolution limits below 2.5 Å and ideally below 2 Å, where ligand electron density could be unambiguously identified and interpreted.

### 3.5.2. Previous crystallisation experiments of MINA53 and NO66

At the beginning of this project only one crystal structure each of NO66 (PDB: 4DIQ) and MINA53 (PDB: 2XDV) was available in the PDB, but during this study additional structures of both proteins became accessible (Table 24 and Table 25). These structures were determined from crystals grown at 20˚C using the sitting drop vapour diffusion method.

All available crystal structures of MINA53 have been determined from a protein construct encompassing residues Ala26-Val465. The initial structure of MINA53 in complex with $Ni^{2+}$ and NOG was solved at 2.57 Å resolution, and followed by a complex with 2OG determined to 2.78 Å resolution in the same $P4_332$ space group but with an unusually high solvent content of 76-77% [137]. The resolution of both complexes does not exceed 2.5 Å and no structures other than with the natural cosubstrate 2OG, its analogue pan-demethylase inhibitor NOG, and substrate peptides have been reported to date (Table 24). The structure of MINA53 in complex with Rpl27 peptide was obtained by cross-linking a MINA53 Tyr209Cys variant with Rpl27 Gly37Cys variant peptide. The latter structure was determined to 2.05 Å resolution in $P2_12_12_1$ space group [137], and it is likely that both the higher resolution and the different space group are due to a stabilising effect of the Rpl27 peptide.

The reported structures of MINA53 are all obtained from crystal conditions based on PEG3350 in either HEPES pH 7.5 or Bis-Tris propane pH 6.5/8.5 buffer, with addition of either a divalent metal cocktail mix or either $Mn^{2+}$ or $Ni^{2+}$. The crystallisation condition for the two latter structures also contained addition of either Na/K-phosphate or ammonium sulphate (Table 24).

| **Protein/PDB ID/sequence/ Release date** | **Resolution/ Space group/ Unit cell** | **Active site metal/ ligand** | **Peptide** | **Crystallisation condition** | **Solvent content** |
|---|---|---|---|---|---|
| MINA53<br>(2XDV)<br>A26-V465<br><br>May 2010 | 2.57Å/ $P4_332$/<br>a=b=c=185.18 Å<br>α=β=γ=90˚ | $Ni^{2+}$/<br>NOG | - | 12% (w/v) PEG 3350,<br>0.005 M $CoCl_2$,<br>0.005 M $MgCl_2$,<br>0.005 M $CdCl_2$,<br>0.005 M $NiCl_2$,<br>0.1 M HEPES pH 7.5<br>20˚C | 77% |
| MINA53<br>(4BU2)<br>A26-V465<br><br>May 2014 | 2.78Å/ $P4_332$/<br>a=b=c=185.63 Å<br>α=β=γ=90˚ | $Ni^{2+}$/<br>2OG | - | 0.1 M BIS-TRIS propane pH 6.5,<br>19-22% (w/v) PEG 3350,<br>0.2 M $(NH_4)_2SO_4$,<br>0.005 M $NiCl_2$<br>20˚C | 76.1% |
| MINA53<br>(4BXF)<br>A26-V465<br>T209C<br><br>May 2014 | 2.05Å/ $P2_12_12_1$/<br>a=70.34 Å<br>b=88.39 Å<br>c=167.18 Å<br>α=β=γ=90˚ | $Mn^{2+}$/<br>2OG | Rpl27a<br>(R32-P50)<br>(cross linked G37C) | 0.1 M BIS-TRIS propane pH 8.5,<br>0.02 M Na-K-phosphate,<br>20-22% (w/v) PEG 3350,<br>0.002 M $MnCl_2$<br>20˚C | 51.66% |

**Table 24. Deposited crystal structures of MINA53.**

The initial structure of NO66 (PDB: 4DIQ) in complex with 2,4-PDCA and $Ni^{2+}$ was determined to 2.40 Å resolution in the $P2_12_12$ space group. No electron density was observed from Ala167-Ala180, which indicates flexibility of this N-terminal region. The latter structures were obtained from slightly shorter protein constructs (Ser183-Asn641 and recently Gly176-Asn641) crystallised in different space groups (Table 25).

Five crystallisation conditions have been reported to date for NO66. The initial condition (PDB: 4DIQ) was based on BIS-TRIS buffer pH 6.5, ammonium sulphate and pentaerythritol ethoxylate with $Ni^{2+}$ as active site metal. The shorter construct of NO66

(Ser183-Asn641) crystallised in its apo form with $Fe^{2+}$ in the active site in Tris-HCl buffer pH7.0, PEG8000 and ammonium sulphate (PDB: 4E4H). Five crystal structures were obtained in the apo form, with NOG and/or peptide, in a condition based on BIS-TRIS propane buffer at a pH range of 5.6 – 7.0, magnesium formate and $MnCl_2$ (Table 25). Finally, the structures that became available in the end of this project (PDB: 3Y33, 3Y3O) were obtained in imidazole and sodium acetate trihydrate and a structure of a NO66 mutant (PDB: 4Y4R) in ammonium phosphate dibasic and acetate buffer (Table 25).

| **Protein/PDB ID/sequence/ Release date** | **Resolution/ Space group/ Unit cell** | **Active site metal/ ligand** | **Peptide** | **Crystallisation condition** | **Solvent content** |
|---|---|---|---|---|---|
| NO66 (4DIQ)<br>A167-N641<br><br>Mar 2012 | 2.40Å/ $P2_12_12$/<br>a=78.67 Å<br>b=149.55 Å<br>c=105.31 Å<br>α=β=γ=90˚ | $Ni^{2+}$/<br>2,4-PDCA | - | 0.05 M BIS-TRIS pH 6.5, 0.05 M $(NH_4)_2SO_4$, 30% (v/v) pentaerythritol ethoxylate 15/4<br>20˚C | 55.79% |
| NO66 (4E4H)<br>S183-N641<br><br>Jan 2013 [142] | 2.28Å/ $P3_2$/<br>a=b=89.35 Å<br>c=304.86 Å<br>α=β=90˚<br>γ=120˚ | $Fe^{2+}$/- | - | 0.1 M TRIS-HCl pH 7.0, 0.2 M $(NH_4)_2SO_4$,<br>10% (w/v) PEG 8000,<br>20˚C | 63.13% |
| NO66 (4CCJ)<br>S183-N641<br><br>May 2014 | 2.15Å/ $P12_11$/<br>a=102.44 Å<br>b=81.56 Å<br>c=151.96 Å<br>α=γ=90˚<br>β=94.66˚ | -/- | - | 0.1 M BIS-TRIS propane pH 5.6-6.5, 0.5-0.7 M magnesium formate, 0.002 M $MnCl_2$<br>20˚C | 59.6% |
| NO66 (4CCK)<br>S183-N641<br><br>May 2014 | 2.15Å/ $P12_11$/<br>a=101.41 Å<br>b=80.71 Å<br>c=151.54 Å<br>α=γ=90˚<br>β=94.54˚ | $Mn^{2+}$/<br>NOG | - | 0.1 M BIS-TRIS propane pH 5.6-6.5, 0.5-0.7 M magnesium formate, 0.002 M $MnCl_2$<br>20˚C | 58.65% |
| NO66 (4CCN)<br>S183-N641<br>L229C/<br>C300S<br><br>May 2014 | 2.23Å/ $C121$/<br>a=155.16 Å<br>b=83.72 Å<br>c=97.03 Å<br>α=γ=90˚<br>β=100.35˚ | $Mn^{2+}$/<br>NOG | Rpl8 (N205-A239) (Cross linked G220C) | 0.1 M BIS-TRIS propane pH 7.4, 0.4 M magnesium formate, 0.002 M $MnCl_2$<br>20˚C | 57.87% |

| | | | | | |
|---|---|---|---|---|---|
| NO66 (4CCM)<br>S183-N641<br><br>May 2014 | 2.51Å/ *C*121/<br>a=154.77 Å<br>b=83.50 Å<br>c=96.87Å<br>α=γ=90˚<br>β=100.20˚ | $Mn^{2+}$/<br>NOG | Rpl8<br>(N205-A239)<br>(Cross linked G220C) | 0.1 M BIS-TRIS propane pH 6.8, 0.25 M magnesium formate, 0.002 M $MnCl_2$<br>20˚C | 57.66% |
| NO66 (4CCO)<br>S183-N641<br>S373C<br><br>May 2014 | 2.30Å/ *C*121/<br>a=155.39Å<br>b=84.94Å<br>c=97.30Å<br>α=γ=90˚<br>β=100.30˚ | $Mn^{2+}$/<br>NOG | Rpl8<br>(N205-A239)<br>(Cross linked G214C) | 0.1 M BIS-TRIS propane pH 7.4, 0.36 M magnesium formate, 0.002 M $MnCl_2$<br>20˚C | 58.53% |
| NO66 (4Y33)*<br>G176-N641<br><br>Sept 2015 | 2.70Å/ $P12_11$<br>a=82.59 Å<br>b=150.88 Å<br>c=106.90 Å<br>α=γ=90˚<br>β=91.32˚ | $NiCl^{2+}$/<br>NOG | - | 15% (v/v) ethanol, imidazole pH 8.0, 0.2 M $MgCl_2$ | 61.34% |
| NO66 (4Y3O)*<br>G176-N641<br><br>Sept 2015 | 2.20Å/ *C*2<br>a=88.50 Å<br>b=202.93 Å<br>c=85.68 Å<br>α=γ=90˚<br>β=118.94˚ | $NiCl^{2+}$/<br>NOG | Rpl8<br>(M204-T224) | 0.1 M imidazole pH 6.5, 0.5 M sodium acetate trihydrate | 60.64% |
| NO66 (4Y4R)*<br>G176-N641<br>*M2 mutant*<br><br>Sept 2015 | 3.3Å/ $P2_12_12_1$<br>a=86.06 Å<br>b=144.21 Å<br>c=144.21 Å<br>α=γ=β=90˚ | - | - | 0.1 M acetate pH 4.5, 1.0 M ammonium phosphate dibasic, | 71.65% |

**Table 25. Deposited crystal structures of NO66. *The structure was not available in the PDB at the time of writing of this thesis.**

### 3.5.3. Outline of crystallography methods

We have attempted crystallisation with both full-length and truncated constructs of MINA53 (Ala26-Val465), and since the initial purification of NO66 with Ser183-Asn641 resulted in low yield we focussed on the longer Ala167-Asn641 construct. Once the insect cell expression became available, we have also generated crystals with a longer NO66 (Gln116-Asn641) construct and full-length MINA53. All crystallisation experiments were prepared with proteins without the histidine tag, utilising the sitting drop vapour diffusion method. Three protein to precipitant ratios (2:1, 1:1 and 1:2) with standard drop size of 150 nl were used, at both 4°C and 20°C, using the commercial sparse matrix crystallisation screens JCSG, LFS, MIDAS, HIN and HCS.

*Optimisation of crystallisation conditions*

Crystal optimisation was performed both in the presence of identified inhibitors and with known binders such as NOG, 2,4-PDCA or the cosubstrate 2OG. The crystallisation conditions were optimised with grid screens ('fine screens') prepared with stepwise gradients of pH, PEG and precipitating agents around the initial condition. We also attempted addition of reagents from two commercially available libraries: the Silver Bullet and Additive Screen from Hampton Index, which were added to 10% (v/v) concentration of the reservoir solution of the crystallisation plates. These reagents are known to affect the solubility and crystallization of proteins by manipulating sample-sample and sample-solvent interactions (for example induce crystallisation contacts), as well as perturb water molecule mediated structures which can alter and improve both the solubility and crystallization of a sample [207].

*Crystal formation*

In addition to optimise crystallisation conditions we have also used microseeding, a technique where previously nucleated crystals are used as seeds and introduced into new drops equilibrated at lower levels of supersaturation. We also attempted matrix microseeding (MMS) [208], [209], a seeding method applied in sparse matrix screens, where crystal seeds are used to reduce the energy barrier of crystallisation resulting in crystal formation in conditions where crystals were previously not observed [208].

*Mounting, cryoprotecting and dehydration*

The generated crystals were tested for diffraction with and without cryoprotectant, as well as with different concentrations of ethylene glycol and glycerol, which have successfully been used previously as cryoprotectants. Finally, the diffraction of MINA53 crystals was also assessed by dehydration experiments on the Diamond Light Source I02 beamline using a high precision crystal humidifier/dehumidifier (HC1b). Dehydration can affect the crystal lattice, resolution and mosaicity of the diffraction data. The HC1b system uses a

humidifier to generate an air stream of chosen relative humidity. The designed humidity is then achieved at room temperature by replacing the cryo-stream nozzle with the stream of air at destined relative humidity and monitoring the size of the crystal drop until equilibration. The diffraction images can be then collected on RT sample or at 100 K achieved by moving the cryo-stream back onto the sample.

### 3.5.4. Crystallisation of NO66

Crystallisation trials, set up with sparse matrix screens, resulted in crystals grown in mainly PEG based conditions (Table 26) and microcrystals were observed in conditions of the previously deposited SGC structure (PDB: 4DIQ). The initial crystals reported in Table 26 were thin layers of plates diffracting to resolution limits of 2.9-3.2 Å (Diamond synchrotron), were indexed to the *P*222 space group with cell dimensions similar to PDB ID: 4DIQ, and obtained in either the HIN or HCS matrix screens. Other crystals obtained in the JCSG sparse matrix screen diffracted to a resolution of 10-14 Å. No crystals where obtained in the LFS screen.

| Screen | Well | Protein concentration (mg/ml) | Condition | Compound | Diffraction resolution | Source |
|---|---|---|---|---|---|---|
| **HIN** | C11c | 9.6 | 0.1 M HEPES pH 7.0, 1 M $(NH_4)_2SO_4$, 0.5% (w/v) PEG 8000 | - | 8.5/3 Å<br>Indexed to *P*222 | FRER/ Dmnd I02 |
| **HIN** | H04c | 9.6 | 20% (w/v) PEG 3350, 0.2M ammonium citrate | - | 4.5/2.9 Å Indexed to *P*222 | FRER/ Dmnd I02 |
| **HIN** | G05 | 9-10.6 | 0.1 M tris pH 8.5, 0.2 M $Li_2SO_4$, 25% (w/v) PEG 3350 | - | microcrystals | |
| **HIN** | G06 | 9-10.6 | 0.1 M BIS-TRIS pH 5.5, 0.2 M ammonium acetate, 25% (w/v) PEG 3350 | - | microcrystals | |
| **HCS** | D04a | 9 | 0.1 M citrate pH 5.6, 20% (w/v) PEG 4000, 20% (v/v) 2-propanol | S01094a | 7.8/3.2 Å | FRER/ Dmnd I02 |

**Table 26. Crystals of NO66 obtained in the coarse screens.** FRER – Rigaku home source, Dmnd – Diamond synchrotron. Experiment temperature in all presented conditions was 20°C.

We then progressed to crystallisation optimisation by setting up grid screens around the identified conditions and also the conditions reported in literature during the course of this study [210] (Table 27). Optimisation screens generated from the HIN sparse matrix screen included wells C11, G5, G6 and H4 (see Table 26 for conditions) where microcrystals or diffracting crystals were observed. Optimisation screens HIN-C11, HIN-H4 and HIN-G6 produced only microcrystals whereas crystals in HIN-G5 optimisation screen resulted in a 3 Å resolution dataset in the previously identified space group $P2_12_12$ (as in PDB: 4DIQ Table 25). Despite multiple repeats we were unable to reproduce the initial 2.9-3.2 Å resolution crystals from wells C11 and H04.

The optimisation screen based on HCS well D4 resulted in crystals consistently diffracting to 3 Å resolution in space group $P2_12_12$. To further improve the quality of the crystals in the identified conditions, we used additives, but no improvement of crystal diffraction was observed.

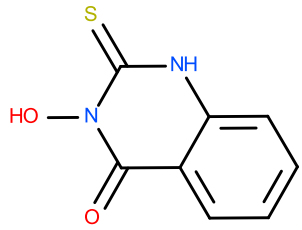



**Figure 37. S01094a used for co-crystallisation stabilises NO66 by 3.5°C and results in 17% inhibition of NO66 at 20 µM in the Rapid Fire assay, and a 60 µM $IC_{50}$ value in the MINA53 Rapid Fire assay.**

Crystals from the identified conditions were also optimised using microseeding, and seeds were added to the crystallisation drop in a serial dilution. Micro seeding of HCS-D4 resulted in thick and uniform diamond-like shaped crystal plates (Figure 38) that initially diffracted to 3 Å and through a second round of seeding resulted in a crystal which diffracted to 1.95 Å in space group $P2_12_12$ (Figure 39 and dataset statistics in Appendix B). However, no electron density was observed that supported the presence of the inhibitor fragment that was used in co-crystallisation (S01094a, Figure 37). Instead, the 2OG binding site was occupied by a glycerol molecule present in the purification buffer.

| Screen | Well | Protein Conc. (mg/ml) | Temp. | Condition | Compound | Diffraction resolution | Source |
|---|---|---|---|---|---|---|---|
| **HIN-G5** | B05a | 10 | 20°C | 0.1 M tris pH 7.9, 0.15 M $Li_2SO_4$, 23% (w/v) PEG 3350 | S01094a | 5/3 Å | FRER/ Dmnd I04 |
| **HCS-D4 (seeding)** | C09c | 12 | 20°C | 19% (w/v) PEG4000, 0.1 M citrate pH 5.7, 10% 2-propanol | S01094a | 3.5/**1.95 Å** | FRER/ Dmnd I03 |
| **4E4H based** | G05a | 8 | 4°C | 0.1 M BIS-TRIS pH 7.0, 0.3 M $(NH_4)_2SO_4$, 10% (w/v) PEG8000 | - | 9/4 Å *P*321 | FRER/ Dmnd I03 |
| **4DIQ based** | - | 13.4 | 20°C | 0.05 M $(NH_4)_2SO_4$, 0.02-0.1 M, BIS-TRIS pH 6.0-7.0, 30% (v/v) pentaerythritol ethoxylate 15/4 or 17/8 or 5/4 | NOG | microcrystals | |
| **4CCK based** | - | 11-12 | 20°C | 0.1 M BIS-TRIS propane pH 6.1-7.5, 0.2-0.8 M magnesium formate, 0.002 M $MnCl_2$ | -/NOG | 3.3 Å | I24 |

**Table 27. Crystals obtained by optimisation screens for NO66. NOG or S01094a were added to the protein in solution at a final concentration of 2 mM. Best diffracting crystal is highlighted in bold.**

Other grid screens included conditions based on crystal structures available in the PDB with ID 4DIQ, 4E4H and 4CCK as shown in Table 27. Grid screens based on condition of PDB ID 4CCK resulted in diffraction to a resolution of 3.3 Å indexed in the $P2_1$ space group in line with the published data, and condition of PDB ID 4E4H resulted in rod-like crystals with anisotropic diffraction to 4 Å in *P*321 space group (Figure 39 D).

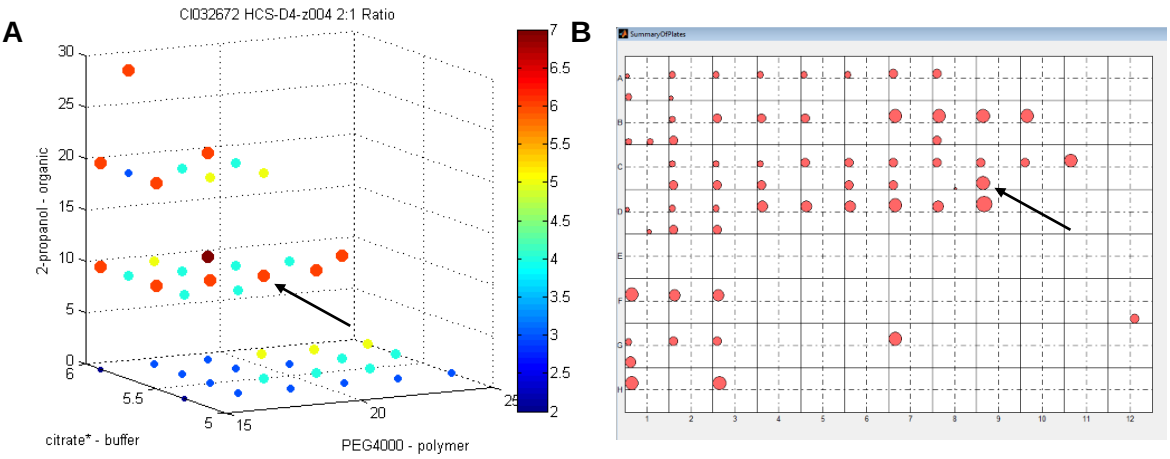



**Figure 38. Visualisation of NO66 optimisation screen based on HSC-D4 condition.** A. 3D view of the crystal plate B. with an axis based on a range of conditions present in the grid screen. The colour scale indicates arbitrary visual assessment of the crystal size/quality with a scale ranging from: 2-crystalline precipitate to 7-large crystal, B. HCS-D4 grid screen with circle size indicating arbitrary relative scale of crystal size/quality. Arrow points to the condition of the crystal resulting in a 1.95 Å dataset.

Although we have attempted crystal optimisation of NO66 in the presence of either 2OG, NOG or 2,4-PDCA, the best crystals of NO66 identified so far were the apo crystals grown in the presence of the S01094a inhibitor fragment. Thus, despite considerable efforts we were unable to reproduce crystals diffracting below 1.95 Å (see Appendix D for data collection and refinement statistics).

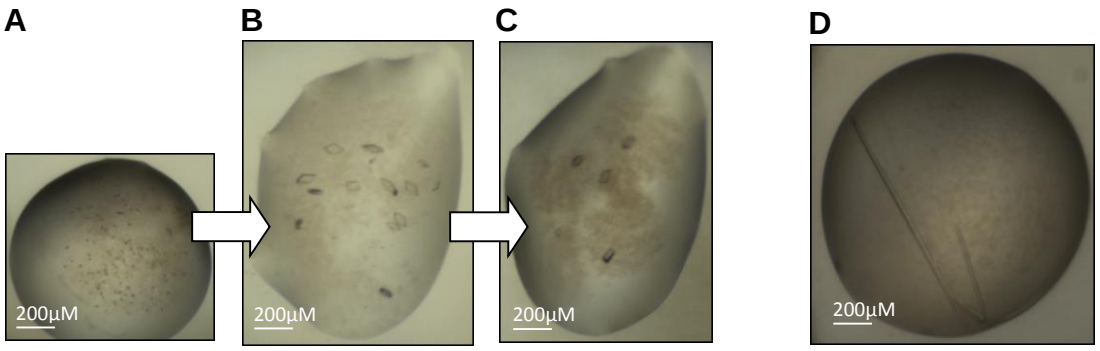



**Figure 39. Optimisation of NO66 crystals. A. Crystals obtained from sparse matrix screen HCS well D4 (Table 26 row 2), B. Crystals in grid follow-up screen around the HCS-D4 condition, and C. Crystals after a second round of microseeding using seeds from B. diffracted to 1.95 Å in $P2_12_12$. D: Crystals grown in a condition based on the published NO66 crystal structure PDB ID 4E4H from** [210]**, diffracted to 4Å at Diamond I03 in the *P*3 space group.**

### 3.5.5. Crystallisation of MINA53

Crystals of MINA53 were known to appear in many conditions but the diffraction was frequently not exceeding 5-7 Å. Previous efforts at the SGC which resulted in the crystal

structure complex with NOG (PDB: 2XDV) required mounting of 500 crystals in order to obtain the 2.57 Å structure. In this work we have mounted and screened over 812 crystals, most of which diffracted to resolution limits above 5 Å.

***Crystals of Ala26-Val464 MINA53***

Crystallisation trials, that were set up with sparse matrix screens with MINA53 resulted in crystals in multiple PEG and salt-based conditions in all tested screens (Table 28). Crystals were also observed in conditions without PEG in BIS-TRIS buffer pH 5.5, 0.1-0.3 M magnesium formate and also in pure BIS-TRIS buffer pH 5.5 with no additives (data not shown). The typical morphology of MINA53 crystals is shown in Figure 40. Despite their large size and apparent homogeneity, only 18 out of 173 crystals that were obtained in the coarse screen conditions diffracted to 5-7 Å using an in house X-ray source. Selected crystals were also screened at the Diamond synchrotron, resulting in a resolution between 3.2 and 6 Å (Table 28).

| Screen | Well | Protein Conc. (mg/ml) | Condition | Compound/ peptide | Diffraction resolution | Source |
|---|---|---|---|---|---|---|
| **HIN** | C11c | 11 (*E. coli*) | 0.1 M HEPES pH 7.0, 1 M $(NH_4)_2SO_4$, 0.5% (w/v) PEG 8000, 0.008 M $MnCl_2$ | NOG | 9.3/6 Å | FRER/ Dmnd I02 |
| **HIN** | F10a | 11 (*Insect cells*) | 0.1 M BIS-TRIS pH 5.5, 0.2 M NaCl, 25% (w/v) PEG 3350 | NOG/ 1mM Rpl27a | 9/6.5 Å | FRER/ Dmnd I02 |
| **HCS** | H06c | 11 (*E. coli*) | 0.1 M Tris pH 8.5, 1.5 M $(NH_4)_2SO_4$, 12% (v/v) glycerol, 0.008 M $MnCl_2$ | - | 9/4 Å | FRER/ Dmnd I02 |
| **HCS** | G02a | 11 (*Insect cells*) | 0.2 M $(NH_4)_2SO_4$, 30% (w/v) PEG5000MME, 0.1 M MES pH 6.5, 0.005 M $CdCl_2$ | NOG | 8.6/3.5 Å | FRER/ Dmnd I03 |
| **HCS** | D10a | 11 (*Insect cells*) | 18% (w/v) PEG8000, 0.2 M calcium acetate, 0.1 M cacodylate pH 6.5, 0.005 M $CdCl_2$ | NOG | 8.5/4 Å | FRER/ Dmnd I03 |
| **JCSG6** | H09a | 11 (*Insect cells*) | 25% (w/v) PEG3350, 0.2 M $Li_2SO_4$, 0.1 M BIS-TRIS pH 5.5, 0.005 M $MnCl_2$ | NOG/ 1mM Rpl27a | 7.5/4.2 Å | FRER/ Dmnd I04-1 |
| **LFS5** | F09c | 11 (*Insect cells*) | 20% (w/v) PEG3350, 10% (v/v) ethylene glycol, 0.1 M BIS-TRIS-propane pH 6.5, 0.2 M sodium/potassium tartrate, 0.005 M $MnCl_2$, 0.005M $CdCl_2$ | NOG | 4.3/3.2 Å | FRER/ Dmnd I03 |
| **MIDAS** | C04d | 11 (*Insect cells*) | 5 % (w/v) polyvinyl alcohol type II, 10 % (v/v) Jeffamine T403, 0.2 M potassium acetate, 0.1 M HEPES pH 7.0, 0.005 M $MnCl_2$, 0.005 M $CdCl_2$ | NOG | 7.2/3.5 Å | FRER/ Dmnd I03 |

**Table 28. Best MINA53 crystals from coarse screens. FRER – Rigaku in house X-ray generator.** Dmnd – Diamond synchrotron. NOG was added to the protein before crystallisation to the final concentration of 2 mM unless otherwise stated. Experiment temperature in all presented conditions was 20°C. All crystal conditions described in this table refer to the shorter MINA53 construct of Ala26-Val464.

Optimisation grid screens were based on identified conditions in the coarse screens and also those from the published MINA53 crystal structures (Table 24). Initial optimisation screen condition with 16-22% (w/v) PEG3350 and HEPES 7.0-7.5 pH (PDB: 2XDV) and metal cocktail resulted in precipitates and microcrystals, and optimisation attempts using microseeding, proved to be unsuccessful. We also decided to separate the metal additives

to test the effect of individual metals. Grid screens based on the initial condition were designed, using the individual metals, in combination pairs, as well as the original condition as a reference. Interestingly, addition of only $CdCl_2$ resulted in hexagonal crystals that diffracted to 3.2 Å in the $P4_332$ space group as in the original PDB ID 2XDV (Figure 40 A), where $CdCl_2$ was found at crystal contacts.

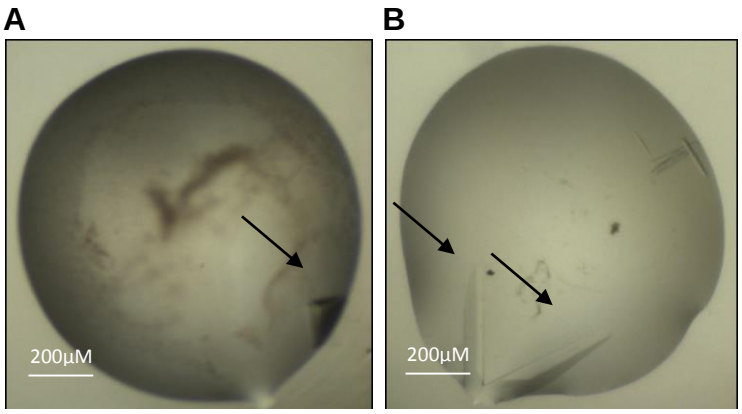



**Figure 40. Typical crystals of MINA53.**
**A: Crystals grew after 4-7 days and diffracted to 2.8-3.2 Å at Diamond I02 beamline. JCSG5 screen condition: 18% (w/v) PEG3.35K, 0.1 M HEPES pH 7.5, 0.005 M $CdCl_2$.**
**B: Crystals obtained after 4-14 days. The crystal diffracted to 8.7 Å. Well condition: 23% (w/v) PEG 3.3K, 0.1 M citrate pH 5.2, 15% (v/v) tacsimate.**

The identified condition (0.1 M HEPES pH 7.5, 18% (w/v) PEG3350, 0. 005 M $CdCl_2$) was repeated across a 96-well crystallisation plate, resulting in crystals appearing in 56 out of 96 2:1 ratio sub-wells indicating the reproducibility of crystallisation. A total of 15 datasets were collected from the best crystals in this condition with only 3 diffracting to 2.6, 2.5 and 2.8 Å resolution and the remaining around 3 Å. We attempted further optimisation by replacing the HEPES buffer with MES pH 5.6-6.6, citrate pH 4.0-5.8 or BIS-TRIS propane pH 6.5-9.0, but this turned out to be unsuccessful. The replicate condition was also tested with addition of a range of salts identified in previous diffracting crystals such as: 0.1-1.2 M ammonium sulphate, 0.05-0.35 M sodium acetate, 0.05-0.35 M sodium chloride, 0.05-0.35 M sodium malonate, 2-15% (v/v) tacsimate, 0.05-0.35 M ammonium phosphate. Crystals grown with 0.3 M sodium acetate diffracted to a resolution range of 3.2-3.4 Å similar to the original condition, with remaining salt additives leading to low diffraction. We also attempted addition of reagents from the Hampton Index Additive and Silver Bullet screens, but neither resulted in improved resolution. This series of experiments has lead us to the conclusion that the original condition of 0.1 M HEPES pH 7.5, 18% (w/v)

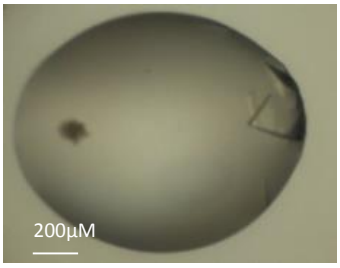
PEG3350, 0.005 M $CdCl_2$ (condition 1) is likely to be the optimal condition and that the addition of $CdCl_2$ is the crucial element for crystallisation.

| Screen | Well | Protein Conc. (mg/ml) | Temp. | Condition | Compound | Diffraction | Source |
|---|---|---|---|---|---|---|---|
| **JMJD2's screen** | E12c (*Insect cells*) | 10.7 | 4°C | 25% (w/v) PEG 3350, 0.1 M citrate pH 5.5, 5% (v/v) tacsimate | NOG | 4.1/2.7 Å | FRER/ Dmnd I02 |
| **HIN-F10** | C05c (*E. coli*) | 10.7 | 20°C | 0.1 M BIS-TRIS-propane pH 6.5, 0.2 M NaCl, 20% (w/v) PEG 3350 | - | 7.6/3.2 Å | FRER/ Dmnd I04-1 |
| **4BXF based +/- Additives/ Silver Bullets/ seeding** | - (*Insect cells*) | 10-20 | 20°C | 0.1 M BIS-TRIS-propane pH 6.5, 0.02 M $NaH_2/K_2HPO_4$, 17-20% (w/v) PEG3350, 0.002/0 M $MnCl_2$ | apo/2OG/ NOG/ MC3098/ MC2200/ MC3095/ MC2209/ | 3 Å (MC2209) | Dmnd I04 |
| **JCSG7-H1 +/- Additives/ seeding** | - (*Insect cells*) | 10 | 20°C | 0.1 M BIS-TRIS pH 5.5, 0.05 M Mg formate, 0.1/0 M $LiCl_2$ | 2OG (bound)/ MC3098 | 3.0 Å $P12_11$ (2OG) | Dmnd I04 |
| **JCSG5-G5 +/- Additives/ Silver Bullets** | - (*Insect cells/ E. coli)* | 10-20 | 20°C | 0.1 M HEPES pH 7.5, 18% (w/v) PEG3350, 0.005 M $CdCl_2$ | NOG/ 2,4-PDCA/ MC2200/ MC3098/ MC3444/ | 7/2.8 Å $P4_332$ **Reproducible 3Å** | BRUKER/ Dmnd I02 |
| **LFS-F9** | - (*Insect cells*) | 11 | 20°C | 22% (w/v) PEG3350, 10% (v/v) ethylene glycol, 0.1 M BIS-TRIS-propane pH 7.0, 0.1 M Na/K tartrate, 0.005 M $CdCl_2$ | NOG | 4.3/3.2 Å | FRER/ Dmnd I02 |
| **JCSG7 (MMS*)** | H01a (*Insect cells*) | 10 | 20°C | 0.3 M magnesium formate, 0.1 M BIS-TRIS pH 5.5, 0.005 M $MnCl_2$ | 5mM 2OG **(bound)** | 3.09 Å dataset *P*21 | Dmnd I04-1 |

**Table 29. Best crystals from follow-up screens for MINA53. FRER – Rigaku in house X-ray generator.** Dmnd – Diamond synchrotron. Compounds were added to the protein before crystallisation to final concentration of 0.1-2 mM. All crystal conditions described in this table refer to shorter MINA53 construct of Ala26-Val464. Notable comments are highlighted in bold – compound bound and diffraction reproducibility.

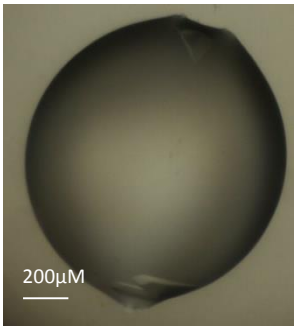
In addition to the crystallisation described above, we also attempted a JMJD2 screen (Table 29) containing PEG based conditions proven to be successful with JmjC-type enzymes JMJD2A, JMJD2D and JMJD3, and crystals obtained in this screen diffracted to 4.0-10 Å resolution in house with a single crystal resulting in 2.9 Å diffraction at the Diamond synchrotron. Despite significant efforts, we were unable to reproduce a crystal that diffracted to 2.9 Å resolution. We then moved back to the coarse screens and used matrix microseeding (MMS) to identify additional crystallisation conditions that grew higher quality crystals. Using this method, we identified a condition in the JCSG7 sparse matrix screen with 0.3 M magnesium formate and 0.1 M BIS-TRIS buffer pH 5.5 (condition 2), that grew crystals diffracting to 3.2 Å resolution (Diamond synchrotron). Further optimisation using a grid screen as well as seeding, resulted only in a slight improvement to 3.09 Å resolution dataset in the $P12_11$ space group with 2OG bound in the active site (Figure 42, see Appendix D for data and refinement statistics). Interestingly the resulting model included electron density allowing to model a loop which was not present in the electron density of previous structures as well as showing substantial conformational changes in comparison to the peptide bound structure PDB ID 4BXF (Figure 42).

A



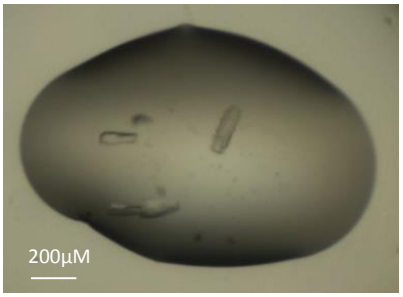

25% (w/v) PEG3350
0.2M $MgCl_2$
0.1M BIS-TRIS pH 5.5,
Resolution 3.6Å (Diamond)
8.6Å (Bruker)
Cpd: NOG

B



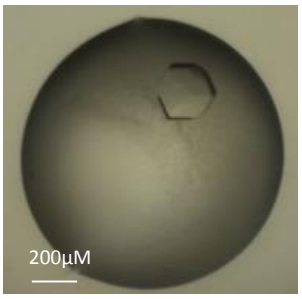

0.1M BIS-TRIS pH 5.5, 0.05M
magnesium formate
Resolution 3.09Å (Diamond)
Cpd: 2OG (bound)

C




0.1 M BIS-TRIS-propane pH 7.0
0.02 M sodium/potassium
phosphate
20% (w/v) PEG3350
0.002 M manganese chloride
Resolution 3.32Å (Diamond)
Cpd: S11208a

D




0.1 M HEPES pH 7.5
8% (w/v) PEG3350
0.005 M $CdCl_2$,
Resolution 2.8-3.5Å (Diamond)
3.7-4.5Å (Bruker)
Cpd: 2,4-PDCA

**Figure 41. Optimised MINA53 (Ala26-Val464) crystallisation conditions.**

The crystals grown in the identified conditions are shown in Figure 41. The model of the 2OG complex in Figure 42 requires further refinement and the initial statistics are presented in supplementary section D.

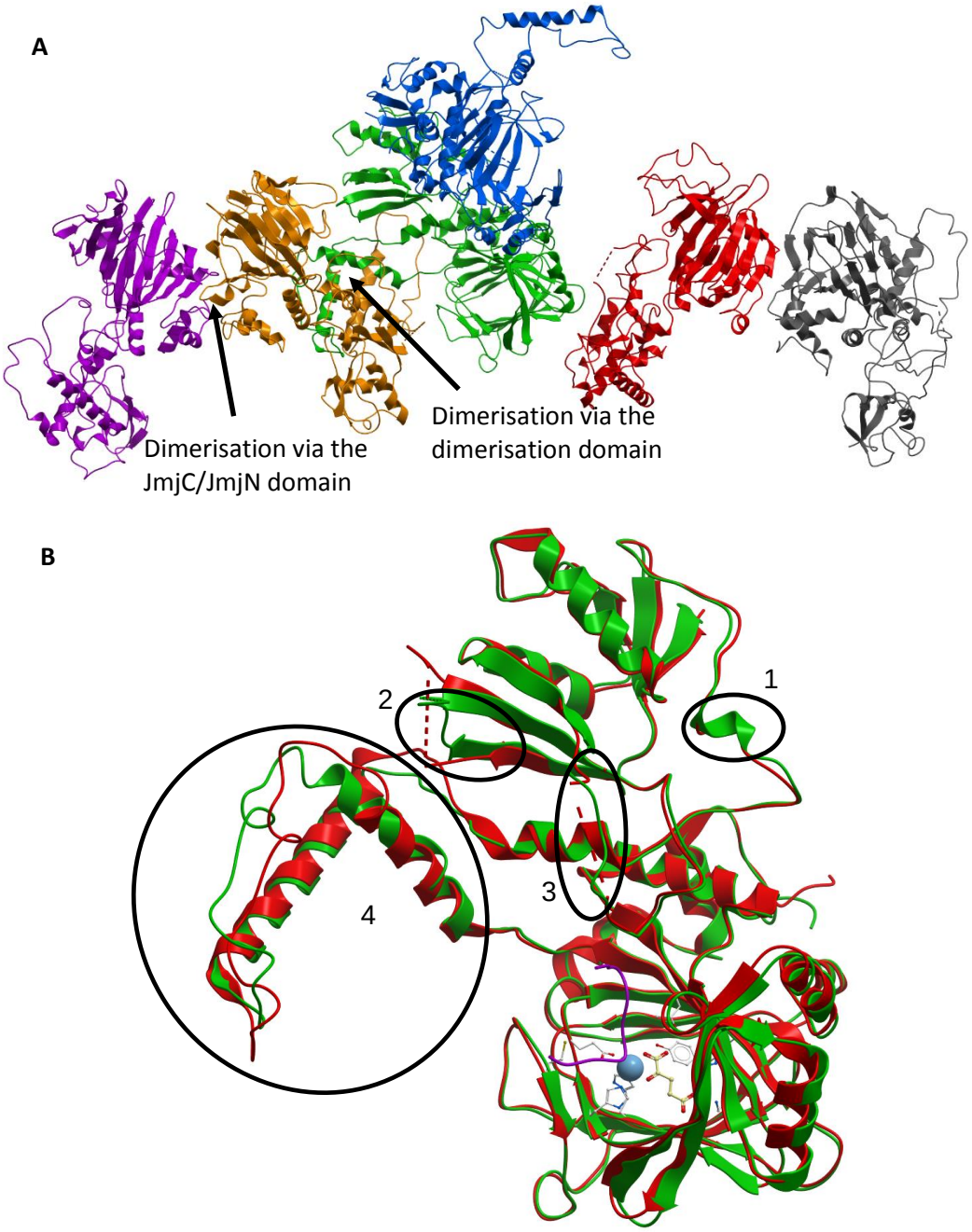



**Figure 42 A. The asymmetric unit of a MINA53 crystal having space group *P*12$_1$1 contained 6 molecules that forms two trimers through intermolecular interactions via loops in the *JmjC* domain and the dimerisation domain.** Alignment of the crystal structure with the peptide-bound structure revealed an RMSD of 1.73 Å. For clarity the individual monomers are coloured differently. B. Monomer of MINA53 from A. (green) aligned with chain A of PDB 4BXF (red). Regions with substantial differences are labelled: 1. Formation of a $3_{10}$ helix in a previously disordered loop, 2. Additional residues indicate a longer β-strand, 3. Loop which was not visible in the electron density maps in previous structures, 4. Conformational changes in the dimerization domain.

### *Crystals of full-length MINA53*

We also attempted crystallisation of the full-length MINA53 (c067) using the coarse screens in the presence of NOG at both 4˚C and 20˚C. The highest diffraction of 3.69Å

was obtained from a crystal in a $P2_1$ space group, grown at 20˚C in 25% (w/v) PEG3350, 0.2 M $MnCl_2$, 0.1 M Bis-Tris buffer pH 5.5 condition. The resulting electron density map could only be used to trace the Ala26-Val465 region which indicates that the N-terminal is flexible.

***Dehydration of MINA53 crystals***

In order to optimise the MINA53 crystals diffracting in $P4_332$ space group (Figure 41 D) we have attempted dehydration experiments. This was performed by Dr C. Lobely using the HC1b at I02 beamline at Diamond synchrotron. When the crystals were dehydrated and datasets collected at room temperature the unit cell did not change, whereas for cryo-cooled crystals a 3.6% contraction in the unit cell was observed and up to 6.1% when dehydrated at 91% relative humidity and then cryo-cooled. Thus, the dehydration of crystals from $P4_332$ condition (Table 29) did not result in improved diffraction.

### 3.5.6. Crystallisation of surface mutants

Since our attempts to grow crystals of MINA53 and NO66 that reproducibly diffracted below 2.5Å were unsuccessful, we resorted to optimisation of the construct by introducing surface mutations, or epitopes, which were based on analysis of crystal contacts of protein structures found in the PDB [211], [212].

Seven surface mutants based on MINA53 (Ala26-Val464) and NO66 (Ala167-Asn641) were chosen for crystallisation based on an output of total of ~120 mutations identified in the clusters of total of ~20 locations in the protein suggested by an in-house SGC surface mutation algorithm (Table 30 and Table 31) [211]. Mutants which were exposed to solvent and induced charge were preferred since those could potentially improve protein solubility and promote crystal contact interactions. The expression, purification and crystal plate preparation was performed by Dr M. Fairhead and the datasets were collected by this DPhil candidate (R. Nowak).

| ID | Mutation | Charge change | Mutation start | Mutation end | WT sequence | Mutated sequence |
|---|---|---|---|---|---|---|
| c076 | G345C:D346K:A348G:L350T | 0 | 345 | 350 | GDGAEL | CKGGET |
| c077 | L350T:S351V:T352A:G355E | 1 | 350 | 355 | LSTPGG | TVAPGE |
| c078 | S300T:T302D:T305R | 2 | 300 | 305 | STTVAT | TTDVAR |
| c079 | K56R:F58H:W59G* | 1 | 56 | 59 | KEFW | REHG |
| c080 | F368H:D370E | 1 | 368 | 370 | FKD | HKE |
| c081 | C105A:V106L:N107S | 0 | 105 | 107 | CVN | ALS |
| c082 | H222R | 0 | 222 | 222 | HEF | REF |

**Table 30. Mutants of MINA53 designed to induce crystallisation. *No expression.**

| ID | Mutation | Charge change | Mutation start | Mutation end | WT sequence | Mutated sequence |
|---|---|---|---|---|---|---|
| c036 | P637N:A639T:L640P:A642L | 0 | 637 | 642 | PLALNA | NLTPNL |
| c037 | W181G:S183D:R186Q | 0 | 181 | 186 | WDSPLR | GDDPLQ |
| c038 | D516E | 0 | 516 | 516 | LTDRER | LTERER |
| c039 | Y588H:Q590H | 2 | 588 | 590 | YPQ | HPH |
| c045 | T306H:Q310R | 2 | 306 | 310 | TTVWQ | HTVWR |
| c046 | R487H:F491Y | 0 | 487 | 491 | RLGHF | HLGHY |
| c047 | Q291R* | 1 | 291 | 291 | WSLYQ | WSLYR |

**Table 31. Mutants of NO66 designed to induce crystallisation. *No expression.**

Full details of expression, purification and crystallisation of the mutants can be found in Appendix A. Crystal plates were set up with all mutants in the JCSG sparse matrix screen without addition of divalent metals or inhibitors. Best crystals of apo NO66 surface mutants were obtained from construct c046 and diffracted to 6 Å resolution, whereas crystals of MINA53 resulted in two datasets in *C*121 space group to resolution of 2.42 Å and 2.51 Å for constructs c076 and c077, respectively. Further optimisation is required to assess the reproducibility of the diffraction resolution.

| ID | Mutation | Condition | Screening/dataset |
|---|---|---|---|
| c076 | G345C:D346K:A348G:L350T | 20% (w/v) PEG3350<br>0.15 M DL- malic acid | 2.51Å dataset<br>*C*121<br>a=118.21,b=180.35,c=91.82,<br>α=90.00,β=116.90,γ=90.00 |
| c077 | L350T:S351V:T352A:G355E | 20% (w/v) PEG3350<br>0.2 M potassium formate | 2.42 Å dataset<br>*C*121<br>a=118.31,b=180.99,c=91.94,<br>α=90.00, β=117.04, γ=90.00 |
| c078 | S300T:T302D:T305R | 25% (w/v) PEG3350<br>0.2 M $MnCl_2$<br>0.1 M BIS-TRIS pH 5.5 | 4.78 Å |
| c079 | K56R:F58H:W59G* | - | - |
| c080 | F368H:D370E | - | - |
| c081 | C105A:V106L:N107S | 25% (w/v) PEG3350<br>0.2 M $MnCl_2$<br>0.1 M BIS-TRIS pH 5.5 | 4.6 Å |
| | | 20% (w/v) PEG3000<br>0.1 M citrate pH 5.5 | 2.66 Å<br>Indexes to *P*3 |
| c082 | H222R | - | - |

**Table 32. MINA53 surface mutant crystallisation.*No expression**

| ID | Mutation | Condition | Screening/dataset |
|---|---|---|---|
| c036 | P637N:A639T:L640P:A642L | - | - |
| c037 | W181G:S183D:R186Q | 20% (w/v) PEG3350<br>0.2 M potassium citrate tribasic | 8 Å |
| c038 | D516E | - | - |
| c039 | Y588H:Q590H | 30% (w/v) PEG2000MME<br>0.15 M potassium bromide | 8 Å |
| c045 | T306H:Q310R | 10% (w/v) PEG8000<br>8% ethylene glycol<br>0.1 M HEPES pH 7.5 | 18 Å |
| c046 | R487H:F491Y | 10% (w/v) PEG8000<br>8% (v/v) ethylene glycol 0.1 M HEPES pH 7.5 | 6.3 Å |
| c047 | Q291R* | - | - |

**Table 33. NO66 surface mutant crystallisation. *No expression**

## 3.6. Discussion

We performed a DSF screen with a library of ~1200 compounds targeted for JmjC- type oxygenases. Both NO66 and MINA53 are stabilised by 2OG mimetics and hydroxamic acids with long aliphatic chains. In addition, NO66 is found to be stabilised by other reported 2OG inhibitor scaffolds including triazolopyridines, bipyridyls, aminopyridines and N-oxalyl amino acids. Although straightforward to perform, a DSF assay carries the limitations of the fluorescence readout which can be altered by auto-fluorescent compounds or quench effects. In addition, a high melting temperature shift is only indicative of compound binding and depends on the quality of tested compound and purified protein. An unusual degree of variability of $T_m$ shift and melting temperature between purifications was also observed which indicates that the quality of protein could have an impact on the result of the screen. The variation in the calculated melting temperature of replicates of the same compound derived from different source plates was also observed which could indicate that certain compounds have degraded or that the plates could contain differing compound concentrations. Nevertheless, the results of this library screen are valuable in prioritising the compound classes for crystallisation trials or for follow up screening with an activity assay. Despite the mentioned difficulties encountered, we were able to identify and validate a set of compounds, such as hydroxamic acids, which subsequently turned out to be inhibitors as validated by the RapidFire assay.

The success in developing a robust RF MS assay opened up the possibility to assess the effect of small molecules on the activity of the enzymes as well as to characterise the hydroxylation reaction by determining the enzymatic constants. The RapidFire assay, although resulting in high quality label-free quantification of the hydroxylation activity, has a relatively narrow assay window with regards to peptide concentrations. The peptide concentration should be in a range of 1-15 µM, otherwise the instrument detector is saturated in the high concentration range, or the signal to noise ratio detected in the ion

spectrum is insufficient at low concentrations- resulting in uncertain quantification of the hydroxylated product. Since a dilution of the assay solution can avoid the saturation problem, these assay constraints have an impact on the determination of the $K_M$ enzymatic constant for the peptides at low concentration, which can be below the detection limit of the RapidFire qTOF mass spectrometer. Although the peptide solution can be concentrated in order to be within the assay detection limit, this has not been attempted in this work. Both peptide $K_M$ determinations would need to be repeated since the data is based on single experiment and in order to fully characterise the NO66 and MINA53 enzymes additional characterisation of the $O_2$ dependence is necessary, which however does not impact the assay development.

The activity assay was used to screen a set of known inhibitor scaffolds for 2OG oxygenases first by validating 2OG analogues such as NOG and 2,4-PDCA as potent inhibitors of both NO66 and MINA53. Another class of compounds were two series of hydroxamic acids from collaborator A. Mai which inhibited MINA53 at low micromolar concentrations and NO66 at the $IC_{50}$ level of ~19 μM. We decided to perform a selectivity profile of these series across representative members of the JmjC-type oxygenases, and the development of these inhibitors is described in further detail in Chapter 4.

The second aim of the work in this chapter was to develop a crystallography system for both MINA53 and NO66 which would allow to reproducibly grow crystals with diffraction of < 2.5 Å, and ideally < 2 Å. As for many other human oxygenases, obtaining high resolution diffracting crystals (<2.5 Å) has proven very challenging and requires protein of high purity and quality, as well as the utilisation of several crystallisation and optimisation techniques. Crystallisation appeared to be sensitive to purification quality, and we therefore attempted crystallisation with protein expressed in either *E. coli* or Sf9 cells, tried various crystallisation optimisation techniques, starting from coarse screens and optimisation screens, seeding, additive screens, as well as collecting data at room temperature and

under cryogenic temperatures including variation of the cryoprotectant and its concentration.

The crystallisation of NO66 was focussed on a construct encompassing residues Ala167-Asn641, which was used previously at the SGC to obtain the first NO66 crystal structure (PDB: 4DIQ). The co-crystallisation experiments resulted in a single 1.95 Å structure of NO66, which we, despite considerable efforts, were unable to reproduce. During the course of this study additional structures of NO66 became available, which were determined from shorter constructs (Ser183-Asn641 and Gly176-Asn461) with the flexible residues in the N-terminus removed. These crystals were obtained in a different space group and resulted in resolution limits of 2.15 Å - 2.5 Å and therefore the focus of future work should be put on development of crystallisation system based on these constructs.

Crystallisation of MINA53 resulted in crystals in a $CdCl_2$ based condition that reproducibly diffracted to 3.0-3.2 Å, and less frequently to 2.8 Å resolution in the $P4_332$ space group. The best diffraction was obtained in the presence of NOG, whereas crystallisation in the presence of all the other tested compounds resulted in a resolution oscillating around 3.2 Å with no ligand bound. Utilising matrix microseeding, we also identified crystals in a new space group *P*21, which diffracted to 3 Å resolution. Thus, despite intensive optimisation attempts the diffraction resolution appeared to be limited to 3 Å.

Towards the end of this project we resorted to design surface mutants of both MINA53 and NO66, a method that has proved to be successful for crystallisation of other proteins [213]. To this end this attempt has resulted in a 2.4 Å resolution dataset obtained from a crystal grown from a triple surface mutant of MINA53 (c077, L350T:S351V:T352A:G355E) in a JCSG screen. The reproducibility of the diffraction remains to be assessed, but the initial results are indeed very promising. Thus far, crystals of the NO66 surface mutants did not result in diffraction limits better than 6 Å. The protein used in crystallisation was purified without the usual size-exclusion chromatography step, and it is possible that

optimisation of the purification procedure to obtain monodisperse protein and optimisation of crystallisation conditions will result in crystals of higher quality.

Since the current crystal systems for MINA53 and NO66 do not allow crystallographic validation of binding poses of identified inhibitors we resorted to the use of well-diffracting crystals with the JmjC- enzyme JARID1B. Other means of identification of the binding mode such as prediction of the binding mode by docking studies have not been attempted in this work. We have however developed initial structure activity relationships (SAR) of inhibitors identified in MINA53/NO66 activity assay which are described in detail in Chapter 4.

# 4. Development of selective inhibitors

## 4.1. Hydroxamic acid- and uracil- based inhibitors

Using a subset of our inhibitor library, three series of compounds from the laboratory of Prof. A. Mai were initially identified as low micromolar inhibitors of MINA53 (MC3098, MC1648) and low micromolar inhibitors of NO66 (MC3126) and are represented by compounds in Figure 43.

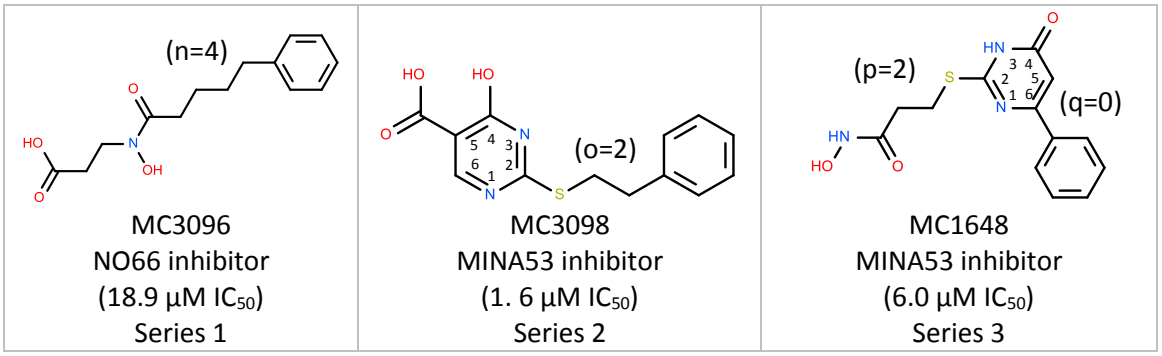


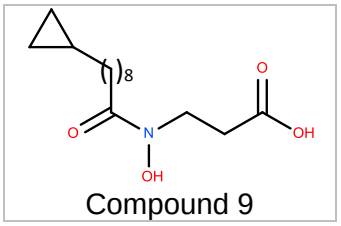

**Figure 43. Representative inhibitors identified in the RapidFire activity assay of MINA53 and NO66.** The carbon linker length n, o, p and q is main variable in those series. Optimal linker lengths are depicted. Series 2 represents 2,5 substitutions and series 3 2,6 substitutions of the core 2-thiouracil motif. The $IC_{50}$ against respective enzyme is provided in brackets. Data from the RapidFire MS activity assay.

In order to assess selectivity of identified hits we profiled representative members of the different families of demethylases and hydroxylases with available AlphaScreen or RapidFire activity assays. This approach included FBXL11 (KDM2), JMJD1B (KDM3), JMJD2A and JMJD2D (KDM4), JARID1B (KDM5), JMJD3 (KDM6), NO66 and MINA53. Activity assays for members of the KDM7 family (PHF8/PHF2/KIAA1718) were not available. The development of hydroxamic acids series 1 is presented in chapter 4.1.1 uracil based derivatives of series 2 in chapter 4.1.2 and optimisation of uracil based hydroxamic acid of series 3 in chapter 4.1.3.

### 4.1.1. Series 1 – FBXL11 inhibitors

**Introduction**

A compound series similar to MC3126 has been published by Suzuki, *et al.* as KDM2/7 inhibitors [214]. Suzuki, *et al.* have profiled eight-carbon linker cycloalkanes (cyclopropyl, cyclopentane, cyclohexane, cycloheptane) and eight-carbon linker phenyl-group analogues [214]. The study identified the eight-carbon linker cyclopropyl Compound 9 (Figure 44) as a KDM2/7 inhibitor with $IC_{50}$'s of 6.8 μM towards KDM2A (FBXL11), 0.2 μM towards KDM7A (KIAA1718) and 1.2 μM with KDM7B (PHF8) [214]. Western blot analysis of the H3K27me2 mark in mouse neuroblastoma N2a cells treated with compound 9 confirmed the dose-dependent effect consistent with inhibition of KDM7 family enzymes [214]. The antiproliferative activity of compound 9 in its carboxylic acid form was detected in the N2a cell line with a 50% growth inhibition parameter ($GI_{50}$) of 86 μM, in the human esophageal squamous cell carcinoma KYSE150 with a $GI_{50}$ of 16 μM and in HeLa cells with a $GI_{50}$ of 40 μM [214].



Compound 9

**Figure 44. Compound 9 with 8 carbon linker identified by Suzuki *et al*. as KDM2/7 inhibitor.**

**Activity profile**

Series 1 is based on compound MC3126 (Figure 43) within which analogues with variable length of the carbon linker between the hydroxamic acid and the phenyl-group were available (Table 34). Selectivity screening of MC3126 analogues were initially seen to inhibit NO66 with an $IC_{50}$ value in the range of 17-40 μM, but were later also found as potent inhibitors of FBXL11 with sub-micromolar $IC_{50}$ values (Table 34). Compound MC3126 appears to inhibit FBXL11 with an $IC_{50}$ of 0.5 μM and shows fair selectivity with 16-fold selectivity over JMJD2A, 44-fold selectivity over JMJD3 and over 80-fold selectivity over JMJD2D, JARID1BA, MINA53 and NO66.

| | | $IC_{50}$ [µM] | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | RapidFire | | Alpha Screen | | | | | |
| Compound | *n* | NO66 | MINA53 | FBXL11 | JMJD1B | JARID1BA | JMJD2A | JMJD2D | JMJD3 |
| MC3126 | 1 | 87.4 | >200 | 0.5 | | 73.2 | 8.1 | 42.3 | 22.1 |
| MC3125* |  | 294.7 | >200 | 2.3 | | 72.3 | 10.1 | 44.7 | 16.8 |
| MC3054 | 2 | 90.0 | 136.0 | 1.0 | | 26.4 | 9.2 | 30.8 | 15.1 |
| MC3096 | 4 | 18.9 | 160.0 | 1.0 | | 64.4 | 12.6 | 31.4 | 16.6 |
| MC3097 | 5 | 25.5 | 67.0 | 1.4 | | 54.0 | 9.1 | 31.5 | 14.6 |
| MC3110 | 6 | 53.5 | 198.0 | 1.9 | | 38.4 | 8.3 | 24.6 | 11.7 |
| MC3112 | 7 | 138.0 | >200 | 5.2 | | 47.6 | 16.3 | 36.9 | 30.7 |

**Table 34. Selectivity heatmap of series 1 with $IC_{50}$ values coloured from red (low) to green (high).** All measurements were done in duplicate of an 11-point 3-fold dilution series. The length of the carbon linker between the hydroxamic acid and phenyl ring moieties is represented by *n*. *linkers are carbon-based with exception of MC3125.

Unfortunately an eight-carbon linker analogue of MC3126 was not available for direct comparison with results of the screening performed by Suzuki, *et al.* We have however determined the $IC_{50}$ for compound 9 in a RapidFire assay for FBXL11 and also compared the DSF melting temperature shift between Compound 9, representatives of MC3126 series and daminozide which is a known inhibitor of KDM2/7 enzymes with reported 1.5 µM $IC_{50}$ for FBXL11 [215] (Table 35). Compound 9 used in the RapidFire assay was ordered from a commercial source (Tocris) and was diluted in DMSO on the day of the experiment to minimise potential compound instability effects.

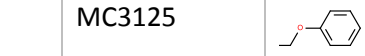
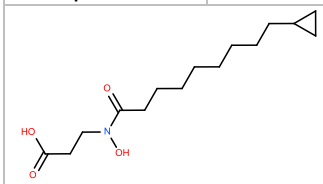
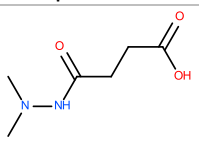
| | | FBXL11 Activity Assay $IC_{50}$ [µM] | | | FBXL11 DSF [20 µM] |
|---|---|---|---|---|---|
| | *n* | Published in literature | RapidFire* | AlphaScreen‡ | Melting temperature shift [˚C]* |
| MC3126 | 1 | | 0.35 | 0.55 | 6.7 ˚C |
| MC3125 |  | | | | 3.9 ˚C |
| MC3054 | 2 | | 0.43 | 0.97 | 3.6 ˚C |
| MC3096 | 4 | | 0.16 | 1.01 | 4.3 ˚C |
| MC3097 | 5 | | | | 4.5 ˚C |
| MC3110 | 6 | | | | 4.5 ˚C |
| MC3112 | 7 | | 0.10 | 5.18 | 4.4 ˚C |
| Compound 13 | 8 | 2.9 [214] | | | |
|  Compound 9 | | 6.8 [214] | 0.107 | | 3.67 ˚C |
|  Daminozide | | 1.5 [215] | | | 5.77 ˚C |

**Table 35. Comparison of $IC_{50}$ values of identified and known FBXL11 inhibitors.** $IC_{50}$ in RapidFire assay and AlphaScreen is determined as duplicate of an 11-point 3-fold dilution series. All compounds in the DSF assay were tested as n=1 in the same experiment. *Assay was performed by A. Tumber. ǂAlphaScreen data from Table **34** for comparison.

The RapidFire mass spectrometry based assay determined the $IC_{50}$ for MC3126 analogues to be in the 100-400 nM range, which is lower than the initial AlphaScreen results which identified MC3126 as the most potent compound with an $IC_{50}$ of 0.5 µM (Table 35). The melting temperature increase determined by DSF showed significant thermal stabilisation across all tested compounds with MC3126 showing maximum stabilisation of 6.7˚C, 1˚C higher than daminozide and 4˚C higher than Compound 9. AlphaScreen $IC_{50}$ data positively correlate with the determined melting temperature shift for the MC3126 analogues. Interestingly, the RapidFire activity assay with Compound 9 resulted in an $IC_{50}$ of 100 nM which is nearly 2 orders of magnitude lower than the published $IC_{50}$ of 6.8 µM [214].

Since no crystallography system was available for FBXL11 and crystallisation of PHF8 as a surrogate was shown to be problematic [Dr C. Johansson, personal communication], we

have attempted to soak compounds from series 1 at 12 mM concentration with JARID1B ($IC_{50}$ 26-73 µM) following the semi-automated method established in paragraph 2.3. Unfortunately no electron density supporting the presence of the compounds has been observed in any of the electron density maps of the collected JARID1B datasets.

**Discussion**

The $IC_{50}$ of Compound 9 obtained from a commercial source is in line with the RapidFire $IC_{50}$ values of series 1, where an increase of the methylene linker length correlates with a decrease of the $IC_{50}$ to 100 nM for MC3112. The published $IC_{50}$ with Compound 9 and Compound 13 for FBXL11 was determined with the AlphaScreen assay [214]. The difference in the activity could partially be a result of differences between the assays: the direct peptide readout in mass spectrometry and the necessity to use antibodies in AlphaScreen assay with detection of a chemiluminescence signal might produce divergent values. RapidFire and AlphaScreen assays use the same concentrations of 2OG, $Fe^{2+}$ and ascorbate, but RapidFire uses 10 µM peptide substrate with 150 nM of enzyme and AlphaScreen only 0.1 µM peptide substrate with 25 nM enzyme. The RapidFire $IC_{50}$ of MC3112 and Compound 9 was determined to be 100 nM, and is below the nominal free enzyme concentration of 150 nM used in the assay and is therefore on the border of the current analysis model which assumes that the compound concentration has to exceed the enzyme concentration. In practice $IC_{50}$ values can be measured to ~0.5 x enz. conc reliably, but not lower. Ideally the concentration of the active enzyme would be determined by active-site titration with a potent binder such as Compound 35 [216], it is however likely that the actual enzyme concentration is lower than the nominal value of 150 nM. The combined activity data and melting temperature shift data indicate that in fact a range of lengths of the linker results in potent and selective FBXL11 inhibitors although assessment of inhibition of the KDM7 family of enzymes is necessary to establish the full selectivity profile.

### 4.1.2. Series 2 – MINA53 inhibitors

**Introduction**

2-substituted 4-hydroxypyrimidine-5-carboxylic acids such as MC3098 (series 2) were initially identified as MINA53 inhibitors in the RapidFire activity assay. Compounds in this series have not been published before and represent a novel scaffold for inhibition of JmjC-type enzymes.

**Activity studies**

The selectivity profile for compounds in series 2 is presented in Table 36. The precursor of the series, compound MC2230, is inactive at the tested concentrations of 100 µM in AlphaScreen and 200 µM in RapidFire NO66 assays, showing only 200 µM $IC_{50}$ value against MINA53. Interestingly, the addition of a single methyl group (Table 36, MC2209) decreases the $IC_{50}$ towards MINA53 by two orders of magnitude to the low micromolar concentration range. The next demethylase inhibited by MC2209 is JMJD2A with an $IC_{50}$ of 17 µM. The selectivity profile of MC2314 containing a five-carbon long chain is similar to MC2209 with one-carbon chain. The tautomers of MC2209 are shown in Figure 45.

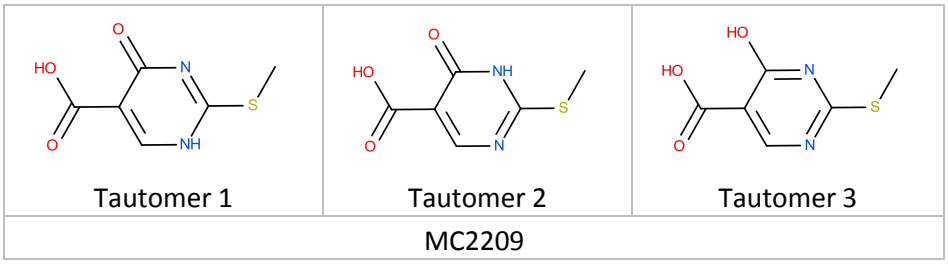



**Figure 45. Tautomers of series 2 represented by compound MC2209.**

Addition of extra carbon atoms results in loss of potency for MINA53, with the four-carbon linker compound MC2164 having an $IC_{50}$ of 13.1 µM, and the nine-carbon linker compound MC3359 an $IC_{50}$ of 65.4 µM. Increased length of the carbon chain results in loss of activity against NO66, but does not affect the other tested demethylases, where inhibition with nine-carbon linker is observed (Table 36).

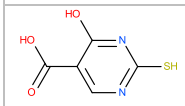
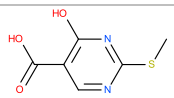
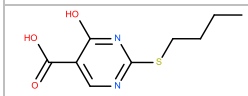
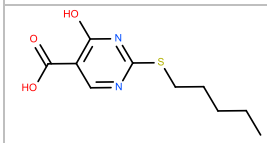
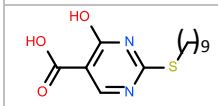
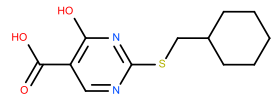
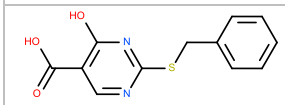
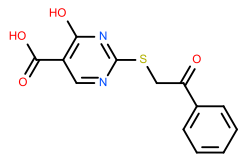
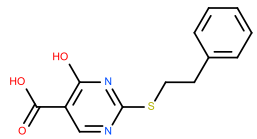
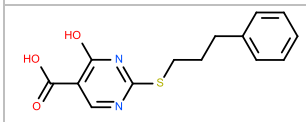
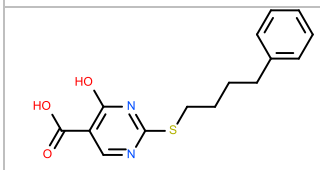
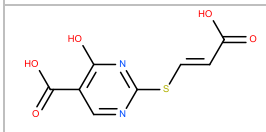
| | IC$_{50}$ [µM] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | RapidFire | | Alpha Screen | | | | | |
| | NO66 | MINA53 | FBXL11 | JMJD1B | JARID1B | JMJD2A | JMJD2D | JMJD3 |
|  MC2230 | >200 | 208.3 | >100 | NA | >100 | >100 | >100 | >100 |
|  MC2209 | 40.2 | 3.0 | 69.6 | NA | 81.4 | 17.6 | 31.9 | 43.3 |
|  MC2164 | 163.0 | 13.1 | 24.2 | NA | 21.5 | 9.1 | 15.5 | 21.5 |
|  MC2314 | 81.2 | 4.3 | 49.6 | NA | 31.7 | 16.5 | 27.5 | 35.1 |
|  MC3359* | >200 | 65.4 | 58.4 | NA | 20.3 | 13.4 | 19.8 | 34.1 |
|  MC2320 | 182.0 | 5.0 | 97.9 | NA | >100 | 20.8 | 30.3 | 38.1 |
|  MC2200 | 48.0 | 3.8 | 67.2 | NA | 57.6 | 14.6 | 34.7 | 49.2 |
|  MC3095 | 23.5 | 3.0 | 45.5 | NA | 46.7 | 12.8 | 26.7 | 36.3 |
|  MC3098 | 71.7 | 1.6 | 46.9 | NA | 41.4 | 13.4 | 30.8 | 35.1 |
|  MC2204 | 36.0 | 1.5 | >100 | NA | >100 | >100 | >100 | >100 |
|  MC3234 | 22.9 | 2.4 | 64.0 | NA | 36.4 | 14.3 | 26.0 | 39.3 |
|  MC2223 | 160.4 | 6.0 | 163.3 | NA | 117.3 | >100 | 123.1 | >100 |

**Table 36. Series 2 selectivity heatmap. All $IC_{50}$ are duplicates of 11-point 3-fold dilution. *Nine-carbon chain**

The second systematic modification in this series is addition of a phenyl ring system with various methylene linker lengths as in MC2200 (one-carbon linker), MC3098 (two-carbon linker), MC2204 (three-carbon linker), and MC3234 (four-carbon linker). Interestingly an increase in the linker length does not have a significant effect on the measured $IC_{50}$, with a modest 1.5-fold increase of potency for JARID1B, JMJD2D and JMJD3 with the four-carbon linker. Activity measurements for compound MC2204 needs to be repeated, since it appears to be inactive with FBXL11, JARID1B, JMJD2D/2A and JMJD3. This experiment will be reproduced in the future work. The lowest $IC_{50}$ value for MINA53 is seen with a linker length of 2-3 carbons. Interestingly, replacing the phenyl ring system (MC2200) with a non-aromatic cyclohexane system (MC2320) increases the $IC_{50}$ for MINA53 to 5 μM and significantly increases the selectivity over NO66 from 16-fold for MC2200 to 36-fold for MC2320, whereas it is inactive at 100 μM concentration in the JARID1B AlphaScreen assay. The third interesting 2-substitution of the 4-hydroxypyrimidine-5-carboxylic acid core is represented by compound MC2223, where the phenyl ring is replaced with a carboxylic acid with an alkene linker. MC2223 appears to be over 23-fold selective for MINA53 against all the other tested JmjC-type oxygenases with a 6 μM $IC_{50}$ for MINA53, suggesting that MINA53 selectivity can be achieved by replacing the lipophilic with a hydrophilic 2-substituted motif. Comparison of phenyl-based and alkane-based linkers suggests that shorter 2-3 carbon phenyl-based linkers are preferred, with an observed increase in MINA53 selectivity with a cyclohexane system replacing the aromatic phenyl ring. This limited series of compounds points to possibly interesting outcomes with the 2-substituted 4-hydroxypyrimidine-5-carboxylic acid motif.

*Crystal structure of JARID1B in complex with MC3095*

Since the attempts to generate crystal structure complexes of these compounds with NO66 and MINA53 were unsuccessful, we resorted to an established crystallography

system of JARID1B soaking all compounds listed in Table 36 following the method described in paragraph 2.3. Despite the high $IC_{50}$ of 46 μM against JARID1B we obtained a complex of JARID1B with MC3095 to 1.95 Å resolution. The active site $Mn^{2+}$ ion (used to replace the active site $Fe^{2+}$ in the crystallisation experiment) is coordinated by the catalytic triad His499, His587, Glu501 and 3 water molecules. The crystal structure reveals that MC3095 is bound to the 2OG binding site, but does not coordinate the metal directly. The carboxylic acid of MC3095 forms a hydrogen bond with Lys517 and Asn509 (Figure 46 A) and the compound is further stabilised by a complex network of water molecules involving the pyrimidine nitrogen (Figure 46), and the carboxylic acid moiety (Figure 46 B). Tyr488 forms a π-π stacking interaction with the pyrimidine ring of MC3095, which stabilises the planar pyrimidine ring. Finally the ketone oxygen present in the linker of MC3095 forms polar interactions with the backbone nitrogen of Val99 and a solvent DMSO molecule (DMS3 in Figure 46 B), and this functional group is missing in tested analogues of MC3095 with lipophilic linkers. The aromatic ring of the linker positions itself between residues Arg98 and His499. The DMSO molecule (DMS3 in Figure 46 B) is conserved across multiple crystal structures (see paragraph 2.3.3) and it is likely that the DMSO molecule may in fact be partially responsible for stabilisation of this conformation of the linker due to steric hindrance.

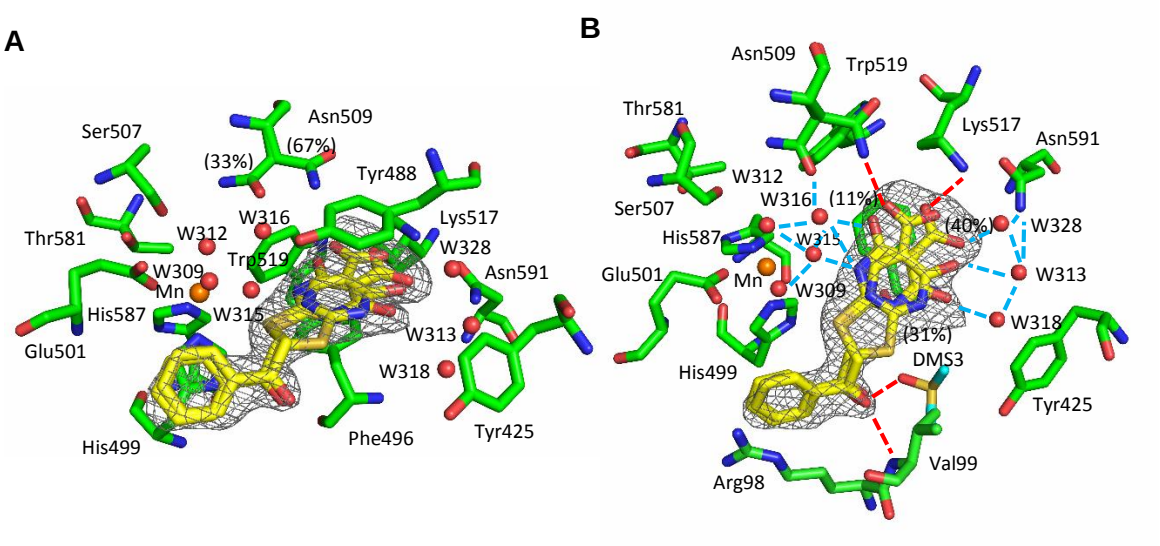



**Figure 46. Crystal structure of JARID1B in complex with MC3095.** The residues within the 2OG binding pocket in JARID1B are shown with green sticks and MC3095 with yellow sticks. $Mn^{2+}$ that was used in crystallisation is shown as an orange sphere whereas the water molecules are shown as red spheres. Hydrogen bond interactions with the side-chains are depicted as red dashed lines and solvent mediated interactions as cyan dashed lines. Three alternative conformations with occupancies of 0.1, 0.3 and 0.4 for MC3095 are built to fit the density, and the Fo-Fc OMIT density map is contoured at 3σ A) MC3095 is π-π stacking with Tyr488 and Phe496. Asn509 is found in two alternative conformations B) Top view showing complex water network surrounding the ligand and water mediated indirect metal binding through W315. Tyr488 not shown for clarity. Phe496 in B lies below the ligand.

### *Discussion*

Compound MC2230 could create a dimer due to the disulphide formation and addition of a methyl group to the sulphur atom would stop the dimerisation, which could explain the observed increase of the inhibition with MC2209. MC3095 inhibits NO66 with an $IC_{50}$ value of 25.3 µM. MINA53 is inhibited by MC3095 with an $IC_{50}$ value of 3.0 µM. We have used JARID1B as a surrogate of MINA53 in this study since the soaking system of JARID1B was readily available, validated and the active site has proven to accommodate various binding modes of ligands. However, other enzymes such as FIH could potentially be a more related surrogate of MINA53. The alignment of JARID1B and MINA53/NO66 points to considerable structural differences within the 2OG binding site that likely imply different binding poses for MC3095 in MINA53 and NO66 (Figure 47). In MINA53 (PDB: 2XDV) Leu176 is found in a similar position as Phe496 in JARID1B, and in NO66 (PDB: 4DIQ) Phe337 is found in a similar location, but is flipped and pointing towards the outside of the metal binding site (Figure 47 A and B). The position of Phe496 in JARID1B stabilises the

ligand whereas in the alignment with NO66 and MINA53 MC3095 would form steric clashes with the current poses of the corresponding residues possibly indicating conformational changes are necessary for this binding pose. Amino acid differences within the 2OG binding site may explain the selectivity of this compound towards MINA53 and NO66. In JARID1B Ala427 is found in the same position as Gln136 and Arg297 in MINA53 and NO66, respectively, and it is possible that these residues form polar interactions with the 4-hydroxypyrimidine-5-carboxylic acid motif. Conversely, in JARID1B, Asn509 is found in a similar position as Ile187 and Val187 in MINA53 and NO66. Finally, despite the fact that the hydrogen bond interaction mediated by Asn509 in JARID1B is abolished by the lipophilic residues Ile187/Val348 in MINA53/NO66 it is likely that Ala597 in JARID1B replaced with His253/417 in MINA53/NO66 could also mediate this interaction. Hence, the binding of MC3095 in its JARID1B pose to MINA53/NO66 requires multiple conformational changes and it is possible that the ligand conformation in addition its pose changes in order to accommodate the differences between the active sites. Nevertheless, the identified compounds in series 2 provide promising novel leads for optimisation of MINA53 inhibitors. The structure activity relationships of series 2 are mainly carbon substitutions with the exception of MC3095 and MC2223, and further SAR should include introducing hydrogen bond acceptors, such as ketone oxygen in MC3095 identified to interact with backbone nitrogen of Val99.

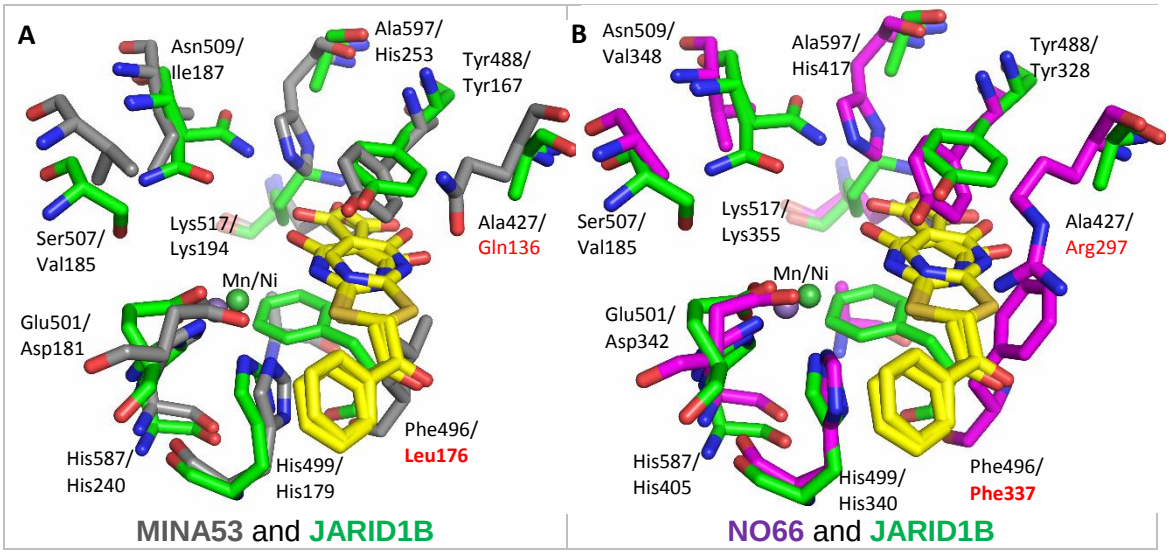


**Figure 47. Superposition of the MC3095 JARID1B crystal structure with A) MINA53 PDB ID 2XDV (grey) and B) NO66 PDB ID 4DIQ (purple). NO66 Leu176 and MINA53 Phe337 would form steric clashes with the ligand and are highlighted in red.**

### 4.1.3. Series 3 – hydroxamic acid - uracil based inhibitors

**Introduction**

Compounds in series 3 and their analogues have been characterised as inhibitors of histone deacetylases by Mai *et al.* [201]. MC1648 was found to inhibit maize HDAC HD2, HD1-B (class I HDAC) and HD1-A (class II HDAC) *in-vitro* with an $IC_{50}$ value of ~1 µM as shown in Table 37 [201]. Extension of the carbon linker between the hydroxamic acid and the uracil moiety by one extra carbon (linker *p*), and also between the uracil moiety and the phenyl ring (linker *q*) shifts the $IC_{50}$ values against maize HDACs to 20-100 nM [201] (see Table 37 for comparison). A further increase of the length of both linkers to up to 5 carbons keeps the nanomolar $IC_{50}$ for maize HDACs [201].

MC1648, a hydroxamic acid member of series 3, was identified in the initial screening of potential MINA53 and NO66 inhibitors (Table 37). The core 2-thiouracil motif is shared between series 2 and series 3 and in fact the same tautomeric states can be observed (Figure 45). Series 3 represents 2,6 substitutions and series 2 2,5 substitutions of common 2-thiouracil core (Table 37).

**Activity screening of MC1648 analogues**

Similar to series 1 and 2, we determined the $IC_{50}$ values for MC1648 and available analogues with representative members of JmjC-oxygenases as shown in Table 37. Compound MC1648 inhibits MINA53 with an $IC_{50}$ of 6.0 µM, and is 27-fold selective over NO66. Thus, the inhibition of JMJD2A and JMJD2D is very similar to MINA53 with an $IC_{50}$ of 8.7 µM and 10.5 µM respectively, while JARID1B is inhibited with an $IC_{50}$ of 900 nM.

The first structure-activity relationship (SAR) studied was the variation of the length of the carbon linker between the phenyl ring and the uracil moiety. Extension of the phenyl group linker to one (MC1630) or two carbons (MC1654) decreases the activity across the panel of JmjC-oxygenases with the exception of JMJD2A, which keeps $IC_{50}$ values of 5.6-10.5

μM. The extended linker of two-carbons found in MC1654 shifts the selectivity profile towards JMJD2A with an $IC_{50}$ value of 5.6 μM, yielding MC1654 inactive against MINA53 and NO66 at a concentration of 200 μM.

The second systematic modification of the MC1648 template was variation of the length of the carbon linker between the hydroxamic acid motif and the uracil moiety. Replacing the two-carbon linker of MC1648 with a three-carbon linker as in MC1640, shifts the JARID1B $IC_{50}$ from 900 nM to 39.9 μM, with no change observed in the JMJD2A $IC_{50}$ and a 3-fold decrease in the MINA53 $IC_{50}$ value. Replacement of the alkane-based hydroxamic acid linker with a more rigid alkene linker as in compound MC1711, dramatically reduces the activity for JARID1B by reducing the $IC_{50}$ to 41.8 μM.

In summary, any extension of the phenyl or hydroxamic linker has detrimental effects on the inhibition of oxygenases shifting the compound activity towards inhibition of maize HDAC's in the nanomolar range. Based on the activity data the initially identified MC1648 compound is the "lead" in the series with optimal lengths of the linkers.

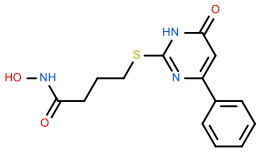
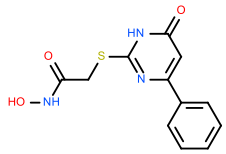
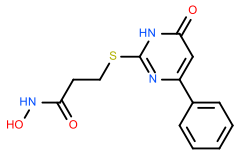
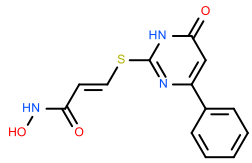
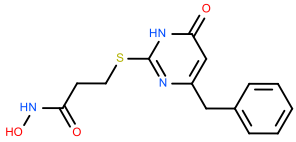
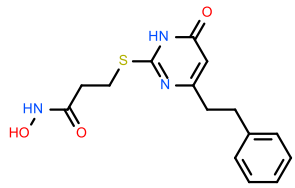
| | IC$_{50}$ [µM] | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | RapidFire | | Alpha Screen | | | | | | As in [201]* | | |
| | NO66 | MINA53 | FBXL11 | JMJD1B | JARID1BA | JMJD2A | JMJD2D | JMJD3 | HD2* | HD1-B* | HD1-A* |
|  MC1640 | 188.0 | 15.9 | 38.9 | | 39.9 | 8.6 | 16.6 | 22.7 | 0.027 | 0.031 | 0.056 |
|  MC3962 | | | 70.9 | >100 | 31.61 | | | 33.38 | ND | ND | ND |
|  MC1648 | 166.0 | 6.0 | 30.8 | 34.3 | 0.9 | 8.7 | 10.5 | 25.3 | 0.822 | 0.890 | 1.020 |
|  MC1711 | | | >100 | 78.3 | 41.8 | | | 41.3 | NI[a] | NI[a] | NI[a] |
|  MC1630 | >200 | 18.1 | 33.1 | | 13.3 | 10.5 | 14.3 | 30.9 | 0.038 | 0.041 | 0.126 |
|  MC1654 | >200 | >100 | 24.2 | | 12.7 | 5.6 | 10.8 | 39.4 | 0.037 | 0.060 | 0.023 |

**Table 37. Selectivity heatmap of series 3 against JmjC-type oxygenases in comparison with published maize HDAC activity** [201]**.** IC$_{50}$ values for JmjC-type oxygenases were performed in duplicate of an 11-point 3-fold dilution series and are coloured from red (low) to green (high). *Published data on inhibition of Maize HDACS HD2, HD1-B and HD1-A extracted from [201] for comparison [a] no inhibition (0%) at 20.9 µM [201].

### *Crystal structures of JARID1B in complex with either MC1648 or MC3962*

We have attempted soaking of all the compounds presented in Table 37 into JARID1B crystals. Soaking of MC1711 and MC3962 was performed in the manual method with

maximum soaking concentration of 10 mM and remaining compounds were soaked using the automated protocol at 12.5 mM concentration, both methods are described in section 2.3. Crystal structures of all soaked crystals were determined and resulted in successful determination of complex structures for MC1648 and MC3962, whereas no ligand density was observed for any of the other soaked compounds. Both MC1648 and MC3962 bind within the active site and also on the surface of the protein, away from crystal contacts, as shown in Figure 48.

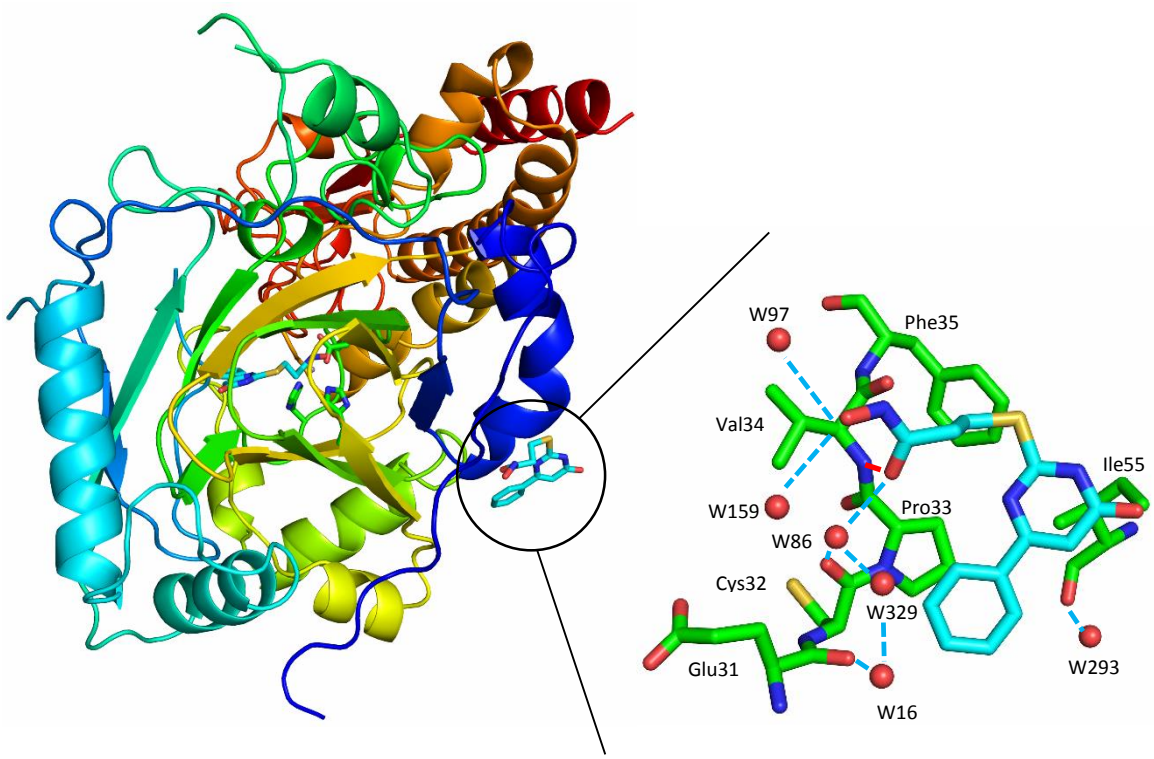



**Figure 48. The binding of MC1648 to the surface of JARID1B.** Hydrogen bond interactions between MC1648 and residue side-chains are depicted as red dashed lines and solvent mediated interactions as cyan dashed lines. The hydroxamic acid is stabilised by interactions with water molecules 97, 159, 86 and a hydrogen bond between the backbone nitrogen of Val34 and the carbonyl moiety of the hydroxamic acid.

MC1648 chelates the active site $Mn^{2+}$ ion in a bidentate manner with the hydroxamic acid motif displacing two metal bound water molecules as shown in Figure 49. The hydroxamic acid is further stabilised by hydrogen bond interactions to the side-chain of Asn509, which frequently is found in two conformations (see Figure 46 A) and Ser507. The 2-thiouracil motif is stabilised by three hydrogen bond interactions: first between $N^{\varepsilon}$ in Lys517 and pyrimidine nitrogen, second between the carboxamide nitrogen of Asn591 and the 2-

thiouracil oxygen, and third between a water molecule and the 2-thiouracil oxygen. In addition to stabilisation by hydrogen bond interactions the uracil- and the phenyl- moiety fit perfectly into a hydrophobic pocket formed by Tyr488, Phe496 and Tyr425.

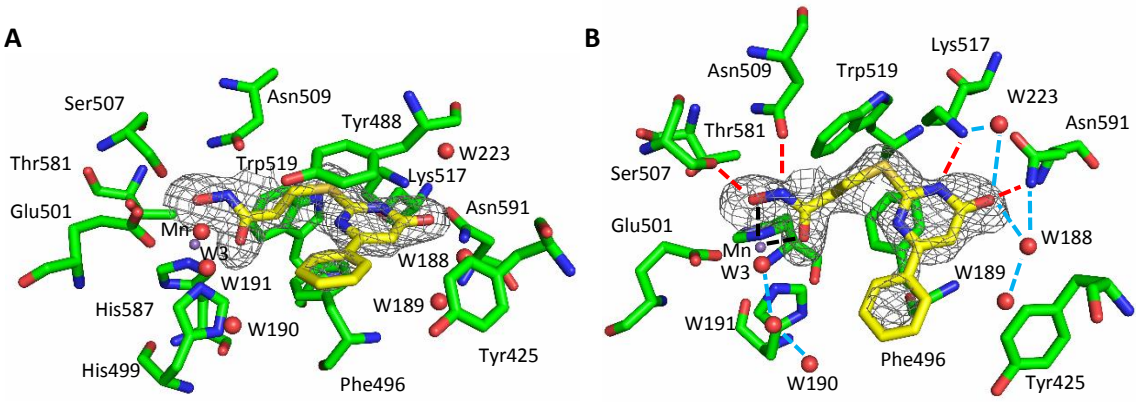



**Figure 49. Crystal structure of JARID1B (green) in complex with MC1648 (yellow).** Fo-Fc OMIT electron density map for the ligand is contoured to 3σ. Two DMSO molecules derived from the crystallisation are not shown for clarity. **A)** Side view of the active site of MC1648 showing the π-π stacking with Tyr488 and Phe496. **B)** Top view showing a water network surrounding the ligand and its bidentate metal binding. Tyr488 is not shown for clarity. Hydrogen bond interactions with the residue side-chains are depicted as red dashed lines, solvent mediated interactions as cyan dashed lines and metal coordination in black dashed lines.

The binding interactions of MC3962 and amino acid residues are the same as in MC1648 with the exception of Arg98 which in MC3962 points towards the metal centre and forms a hydrogen bond with water molecules as shown in Figure 50. Arg98 in MC3962 positions itself in place of water molecules present in MC1648 (labelled in orange, Figure 50). The water network of MC3962 includes an additional water molecule coordinated by pyrimidine nitrogen of the MC3962 and a water molecule (labelled in green, Figure 50) which is no longer clashing with the sulphur atom of MC1648.

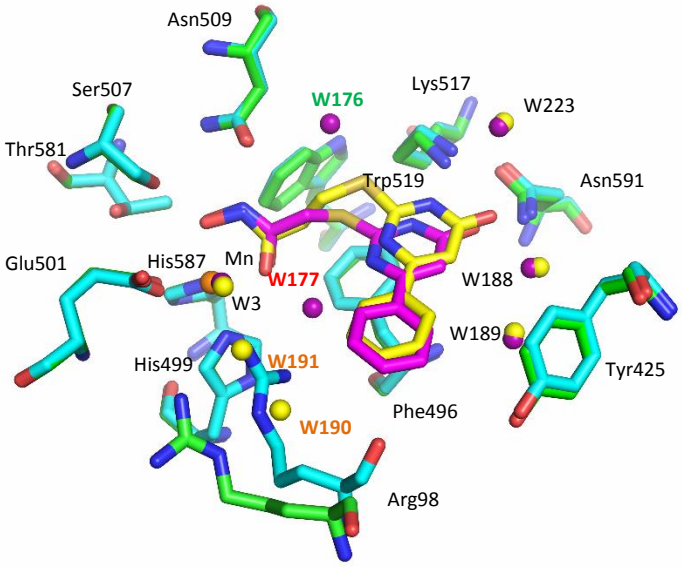



**Figure 50. Overlap of JARID1B complexes of MC1648 (compound and water molecules are shown in yellow sticks and spheres) and MC3962 (compound and water molecules are shown in magenta coloured sticks and spheres).** JARID1B residues of the MC1648 complex are shown in green and residues of MC3962 are shown in cyan. Additional water molecules are present in MC3962; W176 and W177 are highlighted in green and red respectively, whereas water molecules of MC1648 displaced by Arg98 in MC3962 are highlighted in orange.

### *Exploration of alternative metal binding groups of MC1648 based on fragment hits*

Hydroxamic acid based compounds are extensively used as a metal chelating motif in medicinal chemistry projects based on targeted inhibition of human HDACs [217]. Since compounds of series 3 were known to inhibit maize HDACs we decided to replace the hydroxamic acid motif in an attempt to generate JmjC-selective molecules. Based on the crystal structures of MC1648 and MC3962 we optimised compound MC1648 further by replacing the hydroxamic acid motif with alternative metal binding groups identified in the fragment complexes as described in chapter 2.3.

The available fragment X-ray crystal structures identify distinct metal binding motifs as presented in Figure 51.

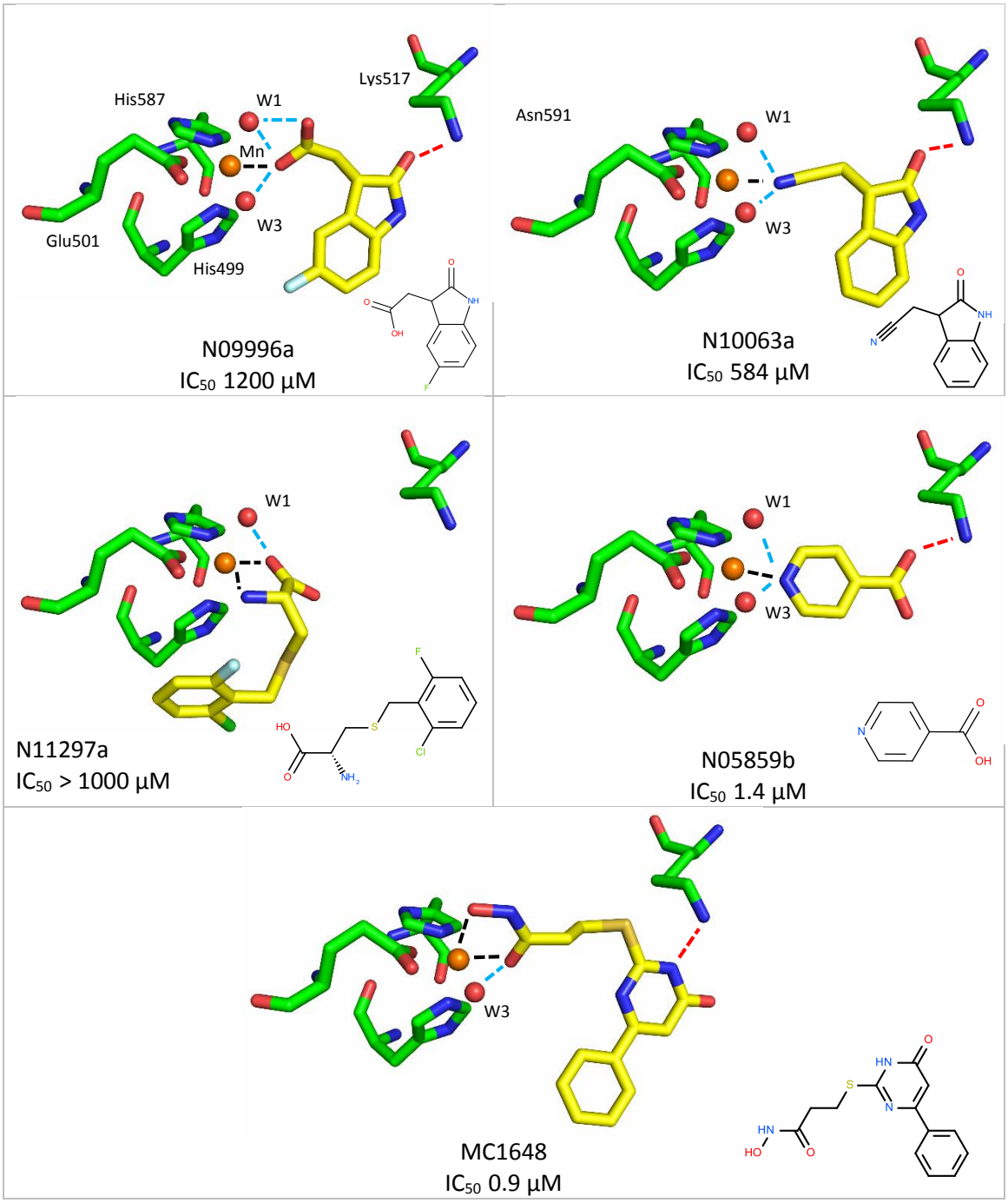



**Figure 51. Crystal structures of fragments bound to JARID1B with different metal binding groups.** The $IC_{50}$ values were determined in an AlphaScreen activity assay in duplicates in an 11-point 3-fold dilution series. MC1648 is shown for reference. The metal binding triad His587, Glu501 and His499 is shown together with the $Mn^{2+}$ coordinating water molecules, shown as red spheres and labelled W1 and W2 for comparison. Hydrogen bond interactions with the residue side-chains are depicted as red dashed lines, solvent mediated interactions as cyan dashed lines and metal coordination in black dashed lines.

Fragments presented in Figure 51 inhibit JARID1B with $IC_{50}$ values not exceeding 500 µM, with the exception of N05859, however the underlying electron density strongly supports the modelled conformation with full occupancies for N09996a, N19963a, N05859b and

occupancy of 0.83 for N11297a after the refinement. Compounds N09996a and N10063 are both based on the same lysine interacting oxindole scaffold with monodentate chelation of the active site metal; with a carboxylic acid for N09996a and a nitrile motif for N19963a which is especially interesting since no inhibitors with this feature have been published for JmjC-type oxygenases. The difference in the compound activity could be attributed to the metal binding motif with the nitrile having 2-fold lower $IC_{50}$ values for JARID1B of ~500 µM. N11297 is inactive at the tested concentrations and was found to bind to the metal site with a carboxylic acid oxygen and nitrogen of primary amine which displaces two $Mn^{2+}$ bound water molecules. Interestingly the 3-fluorochlorobenzene group of N11297 points to the outside of the binding pocket. Finally N05859b, a compound similar to 2,4-PDCA, binds to the metal via the pyridine ring and the carboxylic acid forms a hydrogen bond with the Lys517. Fragment N05859b has an $IC_{50}$ of 1.4 µM against JARID1B making it the most potent fragment of the tested set.

Since both carboxylic acid (N09996a) and nitrile based (N19963a) oxindole scaffolds place themselves in similar locations in the active site as described in Figure 52 we have decided to design follow-up compounds replacing the hydroxamic acid group with the nitrile and carboxylic acid motifs as shown in Table 38. The use of metal binding motifs represented by N11297 and N05859b was not attempted.

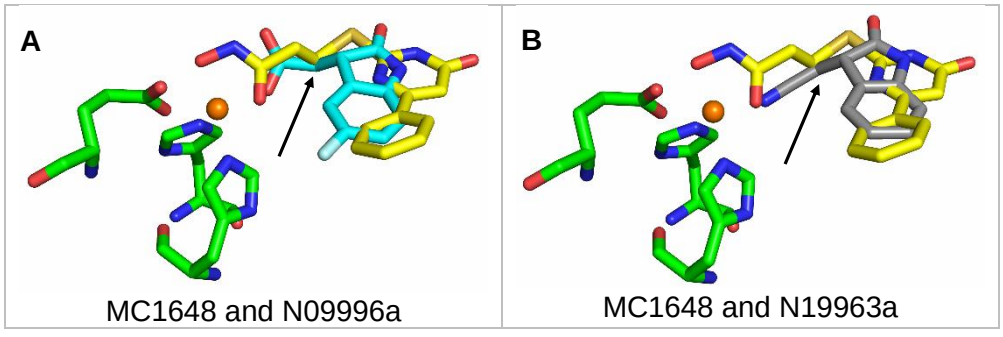



**Figure 52. Overlay of the crystal structure of JARID1B in complex with hydroxamic acid MC1648 (yellow) with A) JARID1B in complex with N09996a (cyan) and B) JARID1B in complex with N19963a.** The black arrow indicates the position of the carbon where fragment linking will be attempted replacing the hydroxamic acid metal chelation group with the nitrile- and carboxylic acid motifs.

We have decided to test a one- and two- carbon linker length between the 2-thiouracil and the metal binding group which allowed to identify possible metal binding poses as determined by docking and flexible compound alignment in ICM [218] (Figure 53). The conformations of both one- and two- carbon linker with the nitrile metal binding motif (MC3949 and MC3960) overlap well with the existing structure of hydroxamic acid MC1648 by docking (Figure 53). The carboxylic acid derivatives of MC1648 with the linker length of one (MC3948) and two (MC3963) also have shown conformations that would allow metal binding. The compounds used in the experiments described below were synthesised by our collaborators in the Prof. A. Mai laboratory at Sapienza University, Rome.

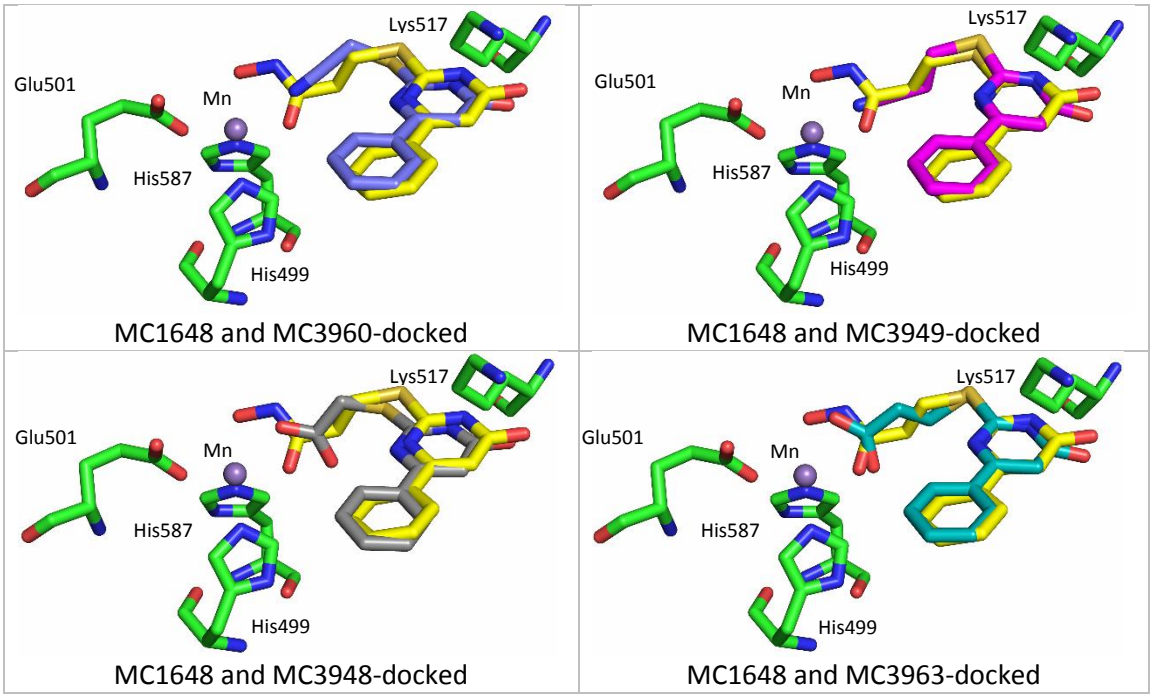



**Figure 53. Docking of carboxylic acid and nitrile analogues of MC1648 into the crystal structure of JARID1B in complex with MC1648 (shown as yellow sticks). The docking was performed in ICM with presence of solvent** [218]**.**

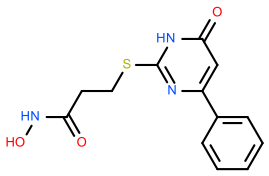
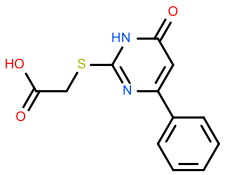
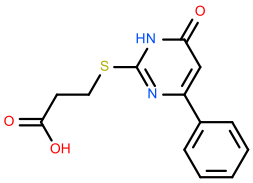
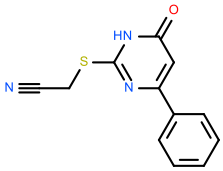
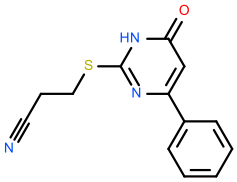
| | IC$_{50}$ [µM] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | RapidFire | | Alpha Screen | | | | | |
| | NO66 | MINA53 | FBXL11 | JMJD1B | JARID1BA | JMJD2A | JMJD2D | JMJD3 |
|  MC1648 | 166.0 | 6.0 | 30.8 | 34.3 | 0.9 | 8.7 | 10.5 | 25.3 |
|  MC3948 | | | >100 | >100 | 326.7 | | | >100 |
|  MC3963 | | | >100 | >100 | >100 | | | >100 |
|  MC3960 | | | >100 | >100 | 103.3 | | | >100 |
|  MC3949 | | | >100 | >100 | >100 | | | >100 |

**Table 38. Series of carboxylic acid and nitrile based analogues exploring alternative metal binding motifs. IC$_{50}$ values were determined by assays as indicated on top of the table.**

Both carboxylic acid and nitrile analogues are significantly less potent than the MC1648 prototype. Nitrile and carboxylic acid derivatives of MC1648 with a single carbon linker resulted in an IC$_{50}$ value for JARID1B of ~100 µM and ~326 µM, respectively, which is a 4-5- fold improvement from the IC$_{50}$ of 584 µM and 1200 µM for the nitrile- and carboxylate-

metal chelating fragments respectively (see Figure 51). The two carbon linker analogues did not inhibit any of the other JmjC-type oxygenases at tested concentration of 100 μM. In conclusion, these activity data suggest that the hydroxamic acid metal binding motif is the most suitable for this particular scaffold.

***Crystal structures of carboxylic acid and nitrile metal binding analogues of MC1648***

The compounds presented in Table 38 were soaked into JARID1B crystals which resulted in crystal structure complexes of MC3960 and MC3948, the one-carbon linker nitrile and carboxylic acid derivatives of MC1648 respectively (Figure 54). The 2-thiouracil motif of MC3960 and MC3948 forms the same hydrogen bond interactions as observed in the case of MC1648. Both compounds chelate the metal ion in a monodentate manner with the carboxylic acid oxygen in MC3948, and nitrile nitrogen in MC3960. The nitrile motif of MC3960 does not place itself perfectly radially to the active site $Mn^{2+}$ but at an angle of ~55˚ as shown in Figure 55. Interestingly, in the structure of MC3948 residue Arg98 points towards the metal thereby displacing water molecules, and interacts with the carboxylic acid oxygen of MC3948. In the nitrile derivative MC3960 the Arg98 residue does not move in relation to the MC1648 crystal structure thereby keeping water molecules (coloured in orange in Figure 54) at their position.

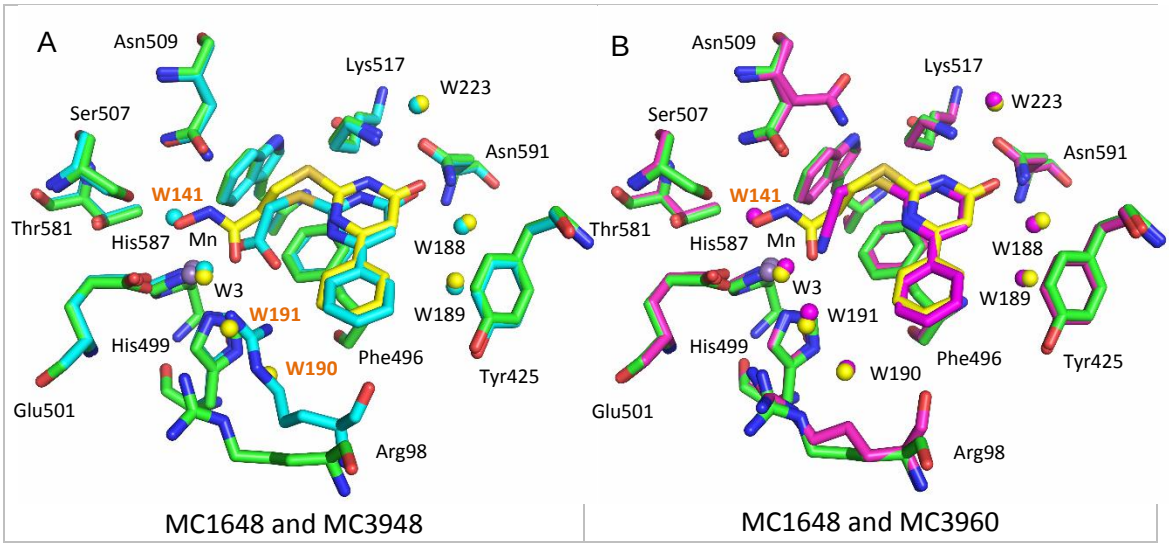



**Figure 54. Comparison of JARID1B crystal structures in complex with A) MC1648 (ligand and water molecules shown in yellow sticks and spheres respectively, residues in green sticks) and MC3948 (ligand, residues, and water molecules shown as cyan sticks and spheres) and B) MC1648 and MC3960 (ligand, water molecules and residues in magenta).** The difference in positions of water molecule between the corresponding structures is highlighted in orange.

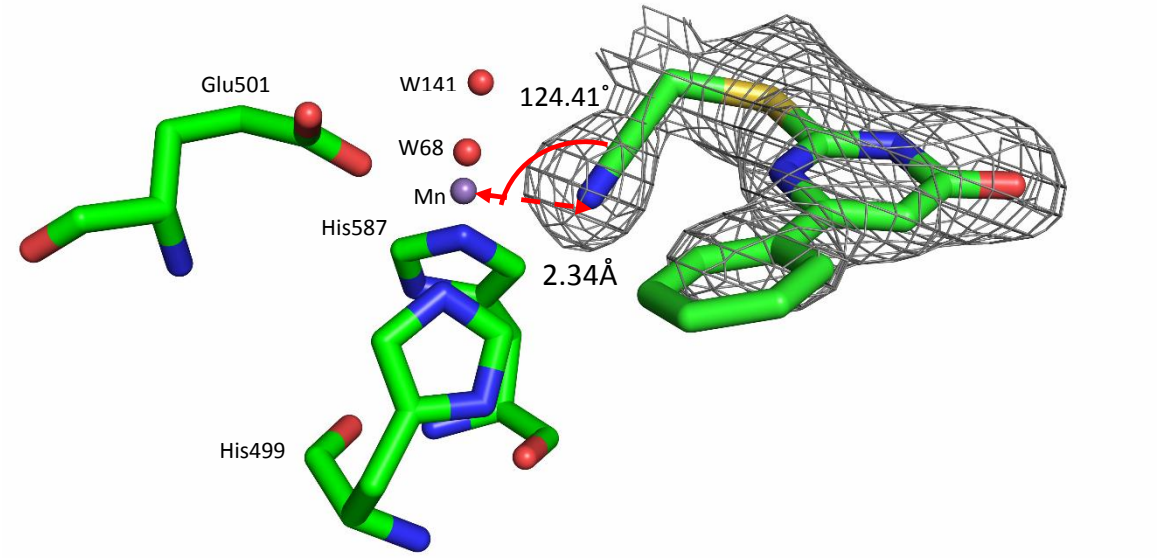



**Figure 55. The binding mode of MC3960 in the 2OG binding pocket of JARID1B showing the distance between the $Mn^{2+}$ and the nitrile nitrogen and the angle of the nitrile motif in relation to $Mn^{2+}$.** The Fo-Fc OMIT electron density map is contoured to 3σ.

### *Structure-based optimisation of MC1648 based on the GSK-J1 template*

Since modifications of the metal binding motif resulted in lower $IC_{50}$ values in comparison to the MC1648 precursor, we decided to focus on modifications to the phenyl ring of MC1648. The GSK-J1 chelates the metal via its pyridyl-pyrimidine core, mimics the 2OG binding with an aminopropanoic acid moiety and mimics the peptide with the benzazepine

motif [77]. Based on an overlay of the JARID1B crystal structures of MC1648 with the demethylase inhibitor GSK-J1 (unpublished) we have attempted to replicate an aromatic feature of GSK-J1 with the second series of MC1648 analogues by placing a phenoxy group on the meta and para positions of the MC1648 phenyl ring and keeping the hydroxamic moiety intact. The activity of compounds of this resulting series is shown in Table 39 and the GSK-J1 comparison in Figure 56. Interestingly GSK-J1 displaces a conserved solvent DMSO molecule present in the MC1648 structure (Figure 56) indicating that exploration of this location of the binding site is possible.

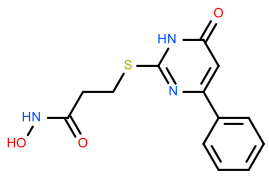
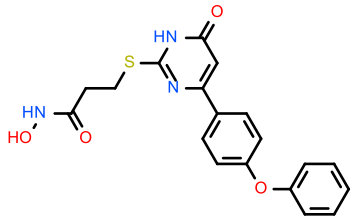
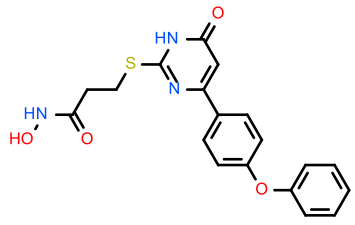
| | $IC_{50}$ [µM] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | RapidFire | | Alpha Screen | | | | | |
| | NO66 | MINA53 | FBXL11 | JMJD1B | JARID1BA | JMJD2A | JMJD2D | JMJD3 |
|  MC1648 | 166.0 | 6.0 | 30.8 | 34.3 | 0.9 | 8.7 | 10.5 | 25.3 |
|  MC3970 | | | >100 | | 20.6 | | | >100 |
|  MC3971 | | | 67.3 | | 9.5 | | | 71.1 |

**Table 39. Activity of the MC1648 analogues with a phenoxy group on the meta- and para- positions of the phenyl ring.**

Both analogues MC3970 and MC3971 appear less potent than MC1648 and result in $IC_{50}$ values against JARID1B of 20.6 µM and 9.5 µM, respectively. Thus, the para substitution

is more active of the two, with meta-substitution having no detectable inhibition at the highest tested concentration of 100 µM with FBXL11 and JMJD3.

We have observed reduced solubility of the compounds in crystallisation solution which was improved when compounds where first dissolved in ethylene glycol cryoprotectant and then in crystallisation solution mix which was added to the crystal as a soaking solution. We were unable to obtain crystal structures of MC3970 and MC3971 in complex with JARID1B, despite multiple soaking attempts using manual method described in section 2.3 no electron density was observed in the active site. However, in the docking simulations of MC3971 and MC3970 shown in Figure 56 C and D the phenoxy group overlaps with the position of DMSO in the MC1648 structure and the aromatic benzene ring of GSK-J1 indicates a possible binding mode.

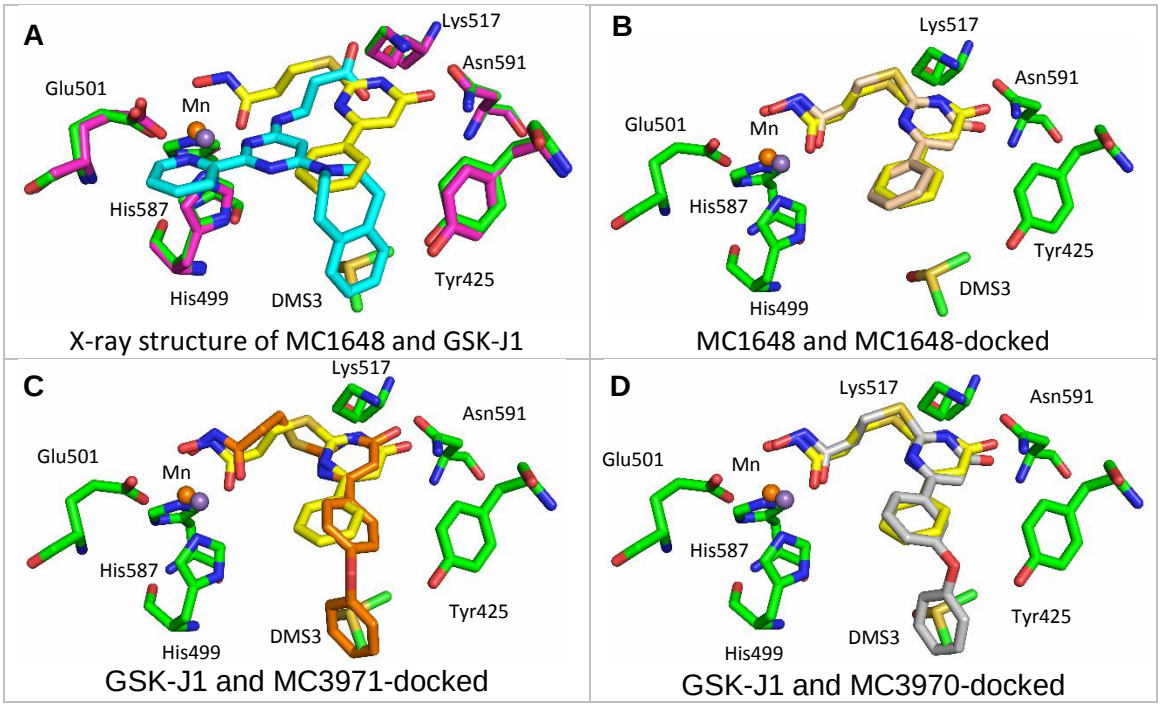



**Figure 56. A. Overlay of crystal structures of GSK-J1 (ligand cyan and residues magenta) and MC1648 (ligand yellow and residues green) bound to the 2OG binding pocket in JARID1BA. B. Comparison of crystal structure of MC1648 with the docked pose of MC1648 shows an RMSD of 0.50 Å increasing the confidence in the following docking results. Docking included receptor water molecules and a restrained hydroxamic acid metal binding motif. C. Comparison of GSK-J1 with the docking pose of MC3971. D. Comparison of GSK-J1 with the docking pose of MC3970.**

## Cellular demethylation assay

In order to assess if inhibitors in series 3 have JARID1B inhibition activity in cells we have used a miniaturised cellular immunofluorescence assay as reported in [161]. The assay overexpresses JARID1B transiently which results in near complete demethylation of H3K4me3 methyl mark (see Figure 58 DMSO control). Inhibition of JARID1B would therefore manifest itself in high levels of the H3K4me3 methyl mark, which is indeed the case when cells are treated with small molecules such as KDOAM25 (Figure 57) or transfected by an inactive JARID1B mutant enzyme instead. We have initially attempted to transfect HeLa cells with JARID1B, but the number of transfected cells have been low. We therefore resorted to U2OS cells which were shown to transfect well [Dr C. Yapp, personal communication].

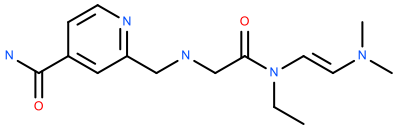


**Figure 57. Structure of KDOM25**

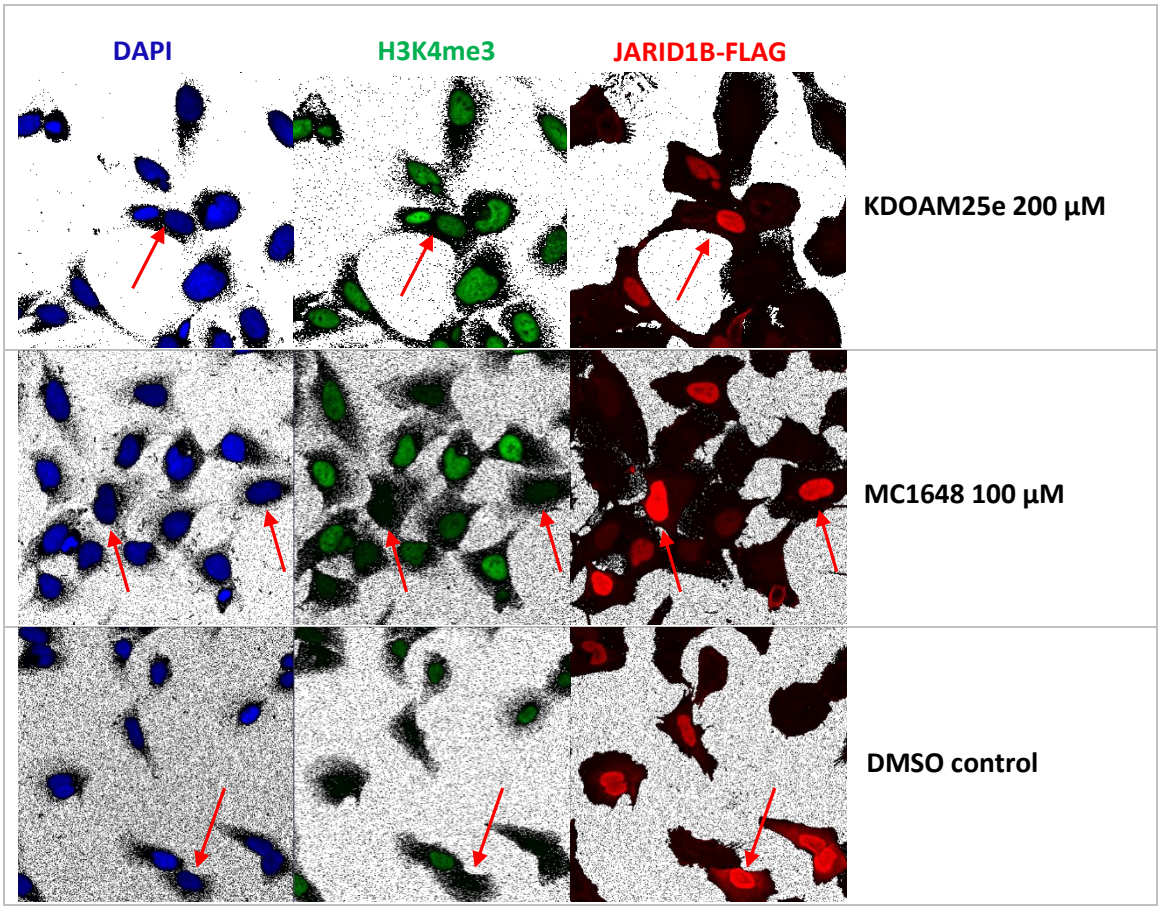


**Figure 58. MC1648 displays no increase in the H3K4me3 methyl mark, whereas KDOAM25 positive control shows levels of H3K4me3 comparable to untransfected cells.** The U2OS cells in immunofluorescence images were stained with DAPI nuclear stain (blue), a specific antibody against H3K4me3 (green) and Flag-tag antibody (red).

The $EC_{50}$ was determined for KDOAM25 to be 52 ± 30 µM and both MC1648 and MC3970 showed a marginal increase in the fluorescence intensity suggesting either fluorescence of the compounds or increase in H3K4me3 methylation (Figure 59). Compound MC3971 showed a dose-dependent increase of fluorescence signal at 100 and 50 and 25 µM concentrations indicating either concentration dependent increase of fluorescence signal due to compound fluorescence or accumulation of H3K4me3 methylation as a result of weak inhibition of JARID1B demethylation activity in cells.

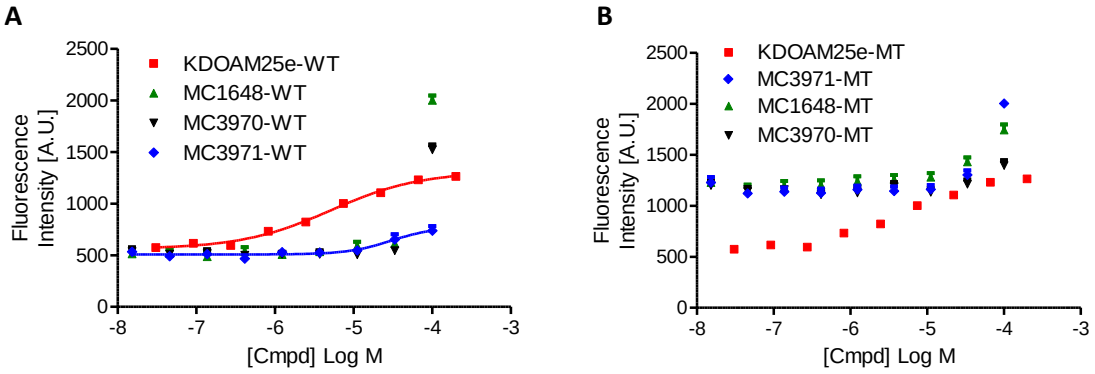



**Figure 59. A. $EC_{50}$ curves for positive control KDOAM25e, MC1648 and MC3970/71 in cells overexpressing JARID1B. KDOAM25 shows $EC_{50}$ of 52 µM.** Marginal increase in fluorescence intensity is observed for MC3971 (blue) and to lesser extend for MC1648 (green), but not MC3971 (black). Data point at concentration of 100 µM of MC1648 and MC3970 are outliers due to increased background fluorescence observed in those wells likely due to incomplete wash of the antibody in these two wells. **B. $EC_{50}$ curves for MC1648, MC3970 and MC3971 for inactive JARID1B mutant control.** All three compounds increase fluorescence intensity. Values are mean and SEM.

## Discussion

The chemical optimisation of the MC1648 structure aimed to improve the activity and selectivity of inhibition against JARID1B. Based on the available fragment complexes we have explored two alternative metal binding motifs using a carboxylic acid and a nitrile group which turned out to have significantly lower activity than the MC1648 starting point. It is likely that the hydroxamic moiety is in fact the optimal metal binding group for the 2-thiouracil scaffold. The remaining parts of the molecule amenable to chemical modification

were the 2-thiouracil motif and the phenyl ring. We did not attempt to modify the 2-thiouracil motif since it generated favourable hydrogen bond interactions with Lys517 and Asn591, instead we decided to focus on the modification of the phenyl ring of MC1648. Based on the crystal structure of GSK-J1 with JARID1B we have added a phenoxy group on the meta- and para- positions of the phenyl ring which based on our docking simulations should place itself in the position of the phenyl ring in the GSK-J1 benzazepine motif displacing the DMSO molecule found in the crystal structure of MC1648 (Figure 55). The resulting compounds MC3970/71 have shown lower solubility at 100 mM concentration than the MC1648 precursor which could be the reason why we were unable to generate a crystal structure with the established soaking method despite multiple attempts. Since crystal structures of MC3970/71 would test the hypothesised binding mode co-crystallisation or other soaking methods would be another path to explore in pursuit of the crystal structures of MC3970/71. Despite the fact that these analogues did not show increased inhibitory activity the phenyl ring of MC1648 remains a good starting point for further chemical optimisation. The cellular demethylation experiments need to be repeated with control for the fluorescence of tested compounds, which could cause increased levels of fluorescence intensity and is an alternative explanation other than low activity of the hydroxamic acids MC1648, MC3970/71 in osteosarcoma U2OS cells. The observed increased methylation of inactive JARID1B mutant control with increasing concentration of KDOAM25 can be explained as inhibition of endogenous JARID1B levels, or possibly as increase of methylation as a result of compound-induced cell death, or as mentioned previously by fluorescence caused by the compound. It is also possible that suspected low inhibition of demethylation is a result of poor cellular permeability of the compounds. The permeability of the compounds could be tested by Caco-2 permeability assay [219], which is frequently used in industry, or by attempting to detect compound using LC-MS of the cell extract. The latter experiment would have to be carefully designed in order to eliminate detection of compound which is still present in the cell media. We would

anticipate further studies in the cellular effect with HeLa cells and cell lines in which inhibition of JARID1B have shown to reduce proliferation such as multiple myeloma cells.

### 4.1.4. Effects on cellular proliferation

In order to assess the effect of identified compounds on cellular proliferation selected representatives of FBXL11 inhibitors series 1 (MC3126, MC3054, MC3096), MINA53 inhibitors series 2 (MC3095, MC3098, MC2204) and JARID1B series 3 (MC1648, MC3970, MC3971, MC1711) were profiled in 5 human multiple myeloma cell lines: MM1S, JIM3, U226, KMS11 and JJN3. The assays were performed by N. Wu and the results are described briefly for completion in this chapter.

**Introduction**

Compounds in series 3 inhibit JARID1B, however the selectivity profile indicates, albeit weaker, inhibition of MINA53 and JMJD2A. Compounds selectively inhibiting JARID1B such as the Epitherapeutics EPT-103182 have been shown to have antiproliferative effect in multiple myeloma *in vitro* and promising activity *in vivo* as described in 1.6. Compounds in series 2 inhibit MINA53 with 10-30- fold selectivity as described in 4.1.2. High levels of MINA53 are found in multiple myeloma as reported by Thakur *et al.* [220] but no inhibition studies have been reported to date. Compounds in series 1 inhibit FBXL11 and likely other enzymes in KDM2/7 family as discussed in paragraph 4.1.1 and there are currently no reports on inhibitory effect of KDM2/7 on the proliferation of human multiple myeloma.

**Cellular proliferation**

Two compound concentrations of 100 μM and 50 μM were tested and the proliferation was measured by PrestoBlue® cell viability staining after 5 and 7 days of incubation. Compounds from series 1, despite good *in vitro* activity (0.2 – 0.5 μM) for FBXL11 and a good selectivity profile, demonstrated no significant effects on proliferation of any of the cell lines tested. Compounds in series 2 have been shown to inhibit MINA53 *in vitro* with $IC_{50}$ value of 1 μM and at least 17-fold selectivity against other demethylases. Although inhibition of MINA53 could result in suppression of cellular proliferation [127] [221], a 1.5-2- fold increase in cellular proliferation was observed after 5 and 7 days in all tested human

multiple myeloma cell lines. More experiments would be necessary to discern the many questions posed by this result. Compounds in series 3 showed time-dependent inhibition of proliferation of MM1S, JIM3 and JJN3 cells in a range of 30-95% at 100 µM and 50 µM concentrations and 90% at 100 µM for U226 cell line with no effect at 50 µM concentration. Inhibition of proliferation of KMS11 cells did not exceed 30%. Dose response studies in JIM3 cell line show $GI_{50}$ of ~100 µM after 7 days of incubation and of ~75 µM after 10 days of incubation for MC3970, $GI_{50}$ of ~90 µM at day 5 and 7 but could not be determined at day 10 for MC3971. MC1648 and MC1711 did not show any inhibition in the tested concentration range. Initial data from JJN3 cell line show the $GI_{50}$ of 50 ± 7 µM for MC3970 and 40 ± 8 µM for MC3970 after 5 days of incubation (Figure 60).

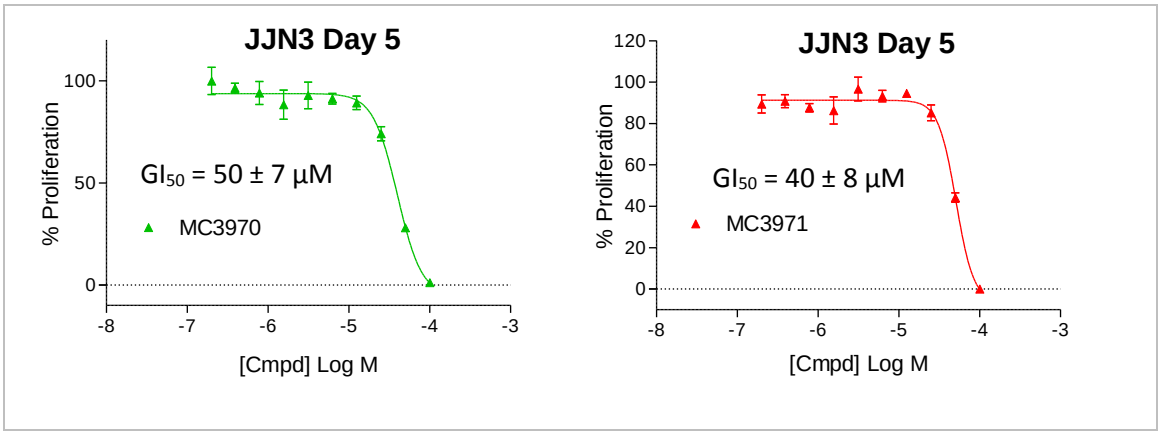



**Figure 60. $GI_{50}$ of the proliferation of the human multiple myeloma JJN3 cell line measured after 5 days of incubation with MC3970 and MC3971.** The dose response was set up in triplicates based on 2-fold dilution of 100 µM starting concentration. Cell proliferation was measured by presto blue cell viability reagent relative to the DMSO control. Data points are mean ± standard deviation of 4 replicates.

## Discussion

Compounds in Series 1 did not have an effect on the proliferation of the cells and Series 2 in fact appeared to induce proliferation. Since inhibition via siRNA knockdown of MINA53 has not been tested in multiple myeloma, but displayed antiproliferative effects in other cancers (see 1.3) the induced proliferation effect was unexpected. It is likely that the compounds in Series 2 developed in this study are not selective enough to attribute this effect to the inhibition of MINA53. It is also possible that a methyl-ester pro-drug form would be required in order to increase the cellular permeability and therefore increase the

active concentration of the compounds inside cells. In vitro activity of compounds in Series 3 have translated to inhibition of proliferation of multiple myeloma cell lines or cytotoxicity, in line with the data from the Epitherapeutics compound. The effect is cell line specific and appears in a time-dependent manner, which could be due to differences in permeability between the cell lines. Although the *in vitro* $IC_{50}$ value of MC3970/71 against JARID1B is lower than that of MC1648 the antiproliferative activity of MC3970/71 is better than for MC1648. It is possible that in addition to inhibition of JARID1B MC3970/71 inhibits other enzymes which could result in reduced proliferation, and/or the phenoxy group is metabolised in cells and the metabolites are potent inhibitors, and/or there are differences in cellular permeability between MC1648 and MC3970/71. Further experiments such as tests for cytotoxicity and cellular permeability would be necessary in order to answer many questions posed by these results.

# 5. Discussion and future directions

In order to characterise inhibitors of JmjC-type oxygenases we have first profiled the selected representatives of the family with libraries of fragments in activity assays and then obtained crystal structures of JARID1B with the identified fragment hits. We have then pursued development of assays and crystallography system for NO66 and MINA53 ribosomal histidinyl hydroxylases, for which no specific inhibitors have been reported. Finally, we have obtained crystal structure complexes of JARID1B with the identified inhibitors and performed initial chemical optimisation based on the available JARID1B fragment crystal structure complexes.

Despite the fact that significant effort has been committed to crystallisation of NO66 and MINA53, we were unable to obtain reproducible high resolution diffracting (< 2.5 Å) crystals. We largely relied on constructs optimised previously at the SGC exploring various downstream crystal optimisation methods. This optimisation has led to reproducible 3.0 Å diffraction of MINA53 crystals, where presence of 2OG was required for stabilisation, and sporadic 2.0 Å diffraction of NO66 crystals. Although low resolution (>3 Å) crystal structures of MINA53 with inhibitors are possible to obtain, the resulting electron density (if any) will likely result in an ambiguous ligand model of limited value. In search for reproducible crystallisation of NO66 future studies should explore shorter constructs of NO66 which have proven to generate higher resolutions structures. An interesting avenue of future work would therefore be further construct modification which has been recently explored at the SGC with the surface mutations of the NO66 (Ala167-Asn641) and MINA53 (Ala26-Val464) constructs designed in order to manipulate crystallisation contacts. This approach called ‘epitope screening’ has already proven successful for one of the 2OG oxygenases JMJD1B, where a construct with crystals diffracting to 2.3-3.0 Å was modified with double mutation to diffract to 1.7 Å [Dr M. Fairhead personal communication]. The promising results with diffraction of crystals of MINA53 surface mutants (~2.3 Å resolution) could make it amenable to fragment soaking, should the

diffraction prove reproducible. This early success suggests that surface mutations are a possible avenue of further exploration not only with NO66 and MINA53 but also with other 2OG oxygenases, hopefully generating reproducible, highly diffracting crystals amenable to compound soaking. It is possible that the fragment soaking efforts will identify novel allosteric binding sites, where initial data suggest a proof of principle for JmjC-type oxygenases by the crystal structure of JARID1B with N10072a. The modulation of the activity with compound N10072a has to be confirmed by a mode-of-action enzyme kinetics characterisation in order to assess whether N10072a is indeed an allosteric inhibitor, or whether the compound is simply not visible in the active site by crystallographic studies but modulates the activity through the active site binding. Further allosteric sites should also be also assessed by mode of inhibition studies and since activity of JmjC-type enzymes can depend on the oligomerisation state it is possible that compounds binding in the domains involved in oligomerisation can result in inhibition of the enzymatic activity.

Reproducible high resolution diffraction of crystals is a prerequisite for successful X-ray based fragment screening campaigns, as represented here by the JARID1B soaking system. JARID1B and JMJD2D [Prof F.V. Delft personal communication] soaking systems are one of the first attempts on fragment soaking with JmjC-type oxygenases which have generated an unprecedented number of small molecule complexes for JmjC-type oxygenases. The high throughput X-ray based fragment soaking protocol developed at the Diamond synchrotron substantially increases the potential of X-ray crystallography as a primary screen. The protocol first limits the amount of compound material required, by using the Echo acoustic dispenser (Labcyte) to transfer compounds directly onto the wells with crystals (for example just a ~20 nl compound transfer volume is required for a final 12.5 mM soaking concentration) and secondly uses robot assisted mounting allowing up to 150 crystals to be mounted in an hour (as in the case of JARID1B soaking), substantially increasing the throughput of this previously labour intensive exercise.

Another major achievement in recent years is the development of miniaturised activity assays for most of the JmjC-type enzymes allowing for compound selectivity to be easily assessed. The availability of assays has previously been a bottleneck as exemplified by GSK-J1 inhibition of JARID family of enzymes which was unknown at the time of the publication [222]. Furthermore, the availability of high throughput mass spectrometry system RapidFire has enabled the development of label free activity assays for demethylases and ribosomal hydroxylases. Availability of high quality mass spectrometry assays allows for additional comparison of *in-vitro* $IC_{50}$ between orthogonal activity assays such as luminescence based AlphaScreen or fluorescence based FDH-coupled enzymatic assays which are prone to false positives. The family-wide activity screen of JmjC-type demethylases provided us with a method to identify fragments with inhibitory activity and promising selectivity profiles.

Compounds in series 1-3 were initially identified in the activity assays of MINA53 and NO66. The family-wide profiling identified series 1 as active inhibitors of FBXL11, series 2 as inhibitors of MINA53 and series 3 as inhibitors of JARID1B, hence a dramatic shift from initial “assignment” as MINA53 and NO66 inhibitors. Early selectivity screening therefore has provided us with crucial information on the identification of the potential target of the molecule. The crystal structure of JARID1B with compound MC3095, which does not directly bind to the metal ion, despite structural differences between the active sites provides information on possible non-metal chelating binding mode with MINA53/NO66. The development of non-metal chelating MINA53 inhibitors based on MC3095 is an interesting avenue for future work. Hydroxamic acids in series 3 have proven to be especially amenable to crystallographic studies. The activity assays of synthesised carboxylic- and nitrile-metal binding analogues indicated improved activity from their fragment starting points, but not over the MC1648 starting point, and two other analogues of MC1648 designed based on existing JARID1B complex with GSK-J1 have been tested

in human multiple myeloma cell lines resulting in time-dependent inhibition of proliferation possibly due to cytotoxicity, which needs to be further evaluated.

We anticipate that the activity profiles determined in this work for two fragment libraries together with the crystal structures from this and future fragment soaking ‘campaigns’ for the JmjC-type oxygenases will form a fragment characterisation set for this class of enzymes. The structural and activity information of identified chemotypes could be used to generate structure-activity hotspots identifying site-specific differences and leading to more informed structure-based drug discovery of these exciting targets.

# Supplementary Information

## A. Constructs, expression and purification

All cloning was performed by Biotechnology group at the SGC.

### NO66 and MINA53 expressed in *E. coli*

Constructs for bacterial expression of NO66 and MINA53 were available at the SGC. NO66 (Ala167-Asn641 and Ser183-Asn641) and MINA53 (Met1-Val464 and Ala26-Val464) were subcloned into pNIC28-Bsa4 and pNIC-CTHF vectors, for expression with Tobacco Etch Virus (TEV) protease cleavable N- and C- terminal histidine tags, respectively. The plasmids were transformed using standard protocol into BL21(DE3)-R3-pRARE2 cell line, a phage-resistant derivative of BL21(DE3) carrying a pRARE2 plasmid to enhance expression of eukaryotic proteins that contain codons rarely used in *E. coli*.

*Expression*

Overnight cultures (100 ml) were prepared from glycerol stocks or freshly transformed cells using terrific broth (TB) supplemented with 100 µg/ml Kanamycin. 10 ml of overnight culture was then added per 1L TB supplemented with 100 µg/ml Kanamycin and 4 ml glycerol. The cultures were grown in incubators at 160 rpm, 37 °C until average OD600 of 2.5 and cooled for 30 min in incubator at 18 °C before they were induced with 0.1 mM IPTG at 18 °C overnight. The cells were harvested by centrifugation at 3000xg for 20 min and the supernatant discarded. The cell pellet was either frozen at -20˚C or processed immediately. Cell pellets from 1L expression, were re-suspended in 50 ml buffer containing 50 mM HEPES pH 8.0, 500 mM NaCl, 10 mM Imidazole, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP) and 5 % (v/v) glycerol, supplemented by 1 µl benzonase nuclease (Sigma Aldrich) and 50 µl of protein inhibitors cocktail III (Merck).

*Purification*

The cells were disrupted by sonication followed by centrifugation for 45 minutes at 23800xg and the supernatant collected for affinity chromatography. Proteins expressed in

*E. coli* were purified by nickel affinity chromatography at 4˚C, using a gravity column and a stepwise gradient of imidazole. All fractions from the immobilised affinity chromatography (IMAC) step were analysed on 4 - 12 % SDS-PAGE. The eluted fractions were pooled and concentrated using Amicon 30 kDa cut-off concentrators to 5 ml and filtered through a 0.22 µm PVDF filter before it was injected manually onto a size-exclusion chromatography column (S200) using ÄKTA xpress equilibrated in 10 mM HEPES pH 7.5, 500 mM NaCl (or 150 mM NaCl for MINA53), 5 % (v/v) glycerol and 0.5 mM TCEP. The fractions from S200 columns were analysed on 4 - 12 % SDS-PAGE and affinity tag removed by Tobacco Etch Virus (TEV) protease (30 µg per mg of protein) incubated overnight at 4˚C. The TEV protease and uncleaved protein was removed by Ni-sepharose and the mass of the cleaved proteins verified by electrospray mass ionization time-of-flight mass spectrometry (Agilent LC/MSD). The TEV cleaved proteins were concentrated in Amicon 30 kDa spin concentrators and the protein concentrations determined with NanoDrop (Thermo Scientific) using the absorption at 280 nm and the estimated extinction coefficient.

**Biotinylated NO66 and MINA53 expressed in *E. coli***

For *in vivo* biotinylation we have used protocols developed at the SGC [198]. Briefly, NO66 (Ala167-Asn641) and MIN53 (Met1-Val464) were subcloned into the pNIC-Bio2 vector with an N-terminal 10-Histidine-tag and a C-terminal biotin acceptor site. The plasmids were transformed into a BL21(DE3)-R3-BirA cell line, which also contains the pCDF-BirA plasmid for co-expression of BirA, a biotin-protein ligase that covalently adds a biotin moiety to the lysine residue in the C-terminal acceptor site. Biotinylated proteins were expressed and purified as described above, but 50 µM biotin was added at the stage of induction with IPTG and the affinity tag was not removed.

**NO66 and MINA53 expressed in *Sf9 cells***

For expression in insect cells, MINA53 (Met1-Val464, Ala5-Val464, Ala26-Val464, Met1-Val464, Ala5-Val464, Ala26-Val464) and NO66 (Ala167-Asn641, Met1-Asn641, Ala6-Asn641, Gln116-Asn641 and Ala167-Asn641) were amplified from an MGC cDNA clone (accession BC014928 – MINA53 and BC011350 – NO66), and subcloned into the pFB-LIC-Bse and pFB-CT10HF-LIC baculovirus transfer vectors (Bac-to-Bac) with tobacco etch virus protease-cleavable N- and C-terminal 10His-flag-tag, respectively, both with ampicillin resistance. Generation of recombinant baculoviruses, insect cell culture, and infections were performed by the Biotechnology group at the SGC according to the manufacturer instructions (Invitrogen). Briefly, the constructs were transformed and proteins expressed in *Sf9* cells and the cells collected 72 hours post infection. The cells were harvested by centrifugation at 400xg for 20 min and the supernatant discarded. The cell pellet was either frozen at -20˚C or processed immediately.

Proteins expressed in Sf9 cells were purified using the same procedure as described for *E. coli* expressed proteins, but the nickel affinity step was performed using a batch method. After washing with a stepwise gradient of imidazole, the resin was transferred to a gravity column and proteins eluted with 300 mM imidazole.

**MINA53 and NO66 surface mutants**

Expression, purification, and crystallisation of surface mutants was performed by Dr M. Fairhead. Briefly, all of the mutants and the wild type proteins were expressed in *E. coli* and purified with Ni-affinity chromatography using a stepwise gradient of imidazole. The imidazole was then removed by desalting on PD10 column (GE Healthcare) and the histidine tag removed by TEV protease followed by Ni-affinity column. Proteins were concentrated in Milipore centricon 10 kDa MW concentrator to average concentration of 12 mg/ml.

## B. Crystallisation and structure determination

### Crystallisation and structure determination of NO66 and MINA53

Crystallisation of MINA53 and NO66 was performed with protein expressed and purified from *E. coli* or Sf9 cells using the 3 drop sitting drop vapour diffusion method at either 20 or 4°C. Crystallisation drops were set up with protein at a concentration of 5-15 mg/ml (NO66) and 10-40 mg/ml (MINA53) and for co-crystallisation both proteins were pre-incubated with 2 mM compound and 2-8 mM $MnCl_2$. The protein compound mix was transferred onto a 96-well Swiss CI crystal plate (SWISSCI, Neuheim, Switzerland) with mosquito robot (TTP Labtech) and mixed with precipitant solution in a 2:1, 1:1 or 1:2 ratio resulting in a final volume of 150-300 nl. JCSG, LFS, MIDAS, HIN and HCS sparse matrix crystallisation screens and individual components of the optimisation screens were obtained from Hampton Research and Molecular Dimensions. Crystallisation plates were imaged with a Minstrel HT system (Rigaku) and viewed in TexRank [223]. Crystals of NO66 and MINA53 appeared after 2-14 days. Crystals were soaked in cryoprotectant containing either 25% (v/v) ethylene glycol or 25% (v/v) glycerol in well solution, and then flash cooled in liquid nitrogen at 100°K. Crystals were first screened for diffraction on in-house rotating anode X-ray generator (Bruker). Diffracting crystals were selected for screening and data collection at Diamond Light Source. The datasets were collected at Diamond Light Source beamlines I02, I03, I04, I04-1 and I24 on samples at 100°K.

The processing for NO66 datasets was done using the CCP4 program suite and the phasing performed by molecular replacement based the PDB ID 4DIQ in Phaser v2.5.2 [224]. MINA53 datasets in complex with 2OG was processed in Xia2 and was solved with PHENIX MRage pipeline [224], [225] based on all 3 available crystal structures PDB: 2XDV, 4BU2, 4BXF. Iterative model building with COOT [226] and refinement with cycles of PHENIX [225] and BUSTER v2.10.1 [227] resulted in final models. Data collection and refinement statistics can be found in D. Table 1 below.

**Crystallisation of surface mutants of NO66 and MINA53**

Vapour diffusion crystallisation plates were prepared with 12 mg/ml purified protein (as described above) in JCSG sparse matrix screen from Molecular Dimensions. Crystallisation was performed without supplementing of the protein with any compounds or divalent metals.

**Crystallisation and structure determination of JARID1B**

Purified JARID1B based on deletion construct (Phe26-Leu101-GlyGlyGlyGly-Ala374-Ile502) was available at the SGC from Dr C. Johansson. JARID1B was crystallised using the 3 drop sitting drop vapour diffusion method at 4°C in 0.1 M HEPES pH 7.5, 0.8 M potassium phosphate dibasic, 0.8 M sodium phosphate monobasic supplemented with 2 mM $MnCl_2$. Crystallisation drops (150 nl or 300 nl) were set up with a mosquito robot (TTP Labtech) using protein at a concentration of 8-9 mg/ml in 2:1, 1:1 and 1:2 protein to precipitant ratios, with and without addition of 20 nl of 1:100 (v/v) dilution of a seed stock. Preparation of seed stock was performed as described below. 30 – 100 crystals were reproducibly grown in Swiss CI plates (SWISSCI, Neuheim, Switzerland) in the uniform condition described above. Crystals were flash frozen in liquid nitrogen at 100 K. Data were collected at Diamond Light Source beamlines I04-1 or I02 and I04. Datasets were processed using Xia2 [228] followed by molecular replacement and initial refinement in DIMPLE pipeline [229] based on the PDB ID: 5A1F. Iterative model building with COOT [226] and refinement with PHENIX [225] and BUSTER v2.10.1 [227] resulted in final models. Data collection and refinement statistics can be found in the Table 40 below.

**Matrix Micro Seeding and seeding**

Seed stock was prepared from crystals obtained from optimisation screens (see main text) and homogenised with a seed bead (Hampton Research) using standard laboratory vortex at full speed in 100-150 μl corresponding reservoir solution. The seeds were vortexed several times for 1 minute with 1 min ice incubation intervals and frozen at -80˚C. The

matrix micro seeding was done following a previously described method [208]. Briefly, the protein mix was transferred onto a 96-well Swiss CI crystal plate with mosquito robot (TTP Labtech) and mixed with precipitant solution in a 1:2, 1:1 or 2:1 ratio resulting in a final drop volume of 150-300 nl followed by 20 nl transfer of concentrated or 1:10, 1:100 (v/v) dilution of seed stock to the crystallisation drops.

**Manual soaking**

In order to determine maximum tolerated compound concentration crystals were transferred in a series of drops with 2-fold dilution of compound with final concentration of 10-25 mM and 25% (v/v) ethylene glycol based on crystallisation well solution. Solution with maximum compound concentration for which no morphological changes to the crystal were observed, was then added directly to Swiss CI plate (SWISSCI, Neuheim, Switzerland) sub well containing crystals. Crystals were incubated with the soaking solution for 15 min to 120 min and mounted and flash frozen in liquid nitrogen at 100 K.

MC3970 and MC3971 where first dissolved in 100% ethylene glycol cryoprotectant in 2-fold dilution with maximum concentration of 10 mM. The crystallisation well solution was then added to final concentration of 25% ethylene glycol. The crystallisation solution mix was used as a soaking solution. The remaining procedure was performed as described above.

**Semi-automated soaking**

Crystallisation was performed following a protocol developed by Lab36, Diamond Light Source, Harwell, UK. Briefly, crystals grown in 300 nl well volumes were soaked with addition of 100 nl of 100% ethylene glycol, followed by addition of compounds to final concentration of 12.5 mM or 20 mM using Echo Acoustic Dispenser (Labcyte). Crystals were incubated for 2-3 h. Crystal harvesting was assisted by a robotic plate positioning system developed by N. Wright, which allowed for 150 crystals to be mounted in an hour.

## C. JARID1B Table 1

Statistics for the highest-resolution shell are shown in parentheses.

| | 2OG | Succinate | Fumarate | Acetate | Pyruvate | Malate |
|---|---|---|---|---|---|---|
| Wavelength (Å) | 0.9173 | 0.9173 | 0.9173 | 0.9795 | 0.9795 | 0.9795 |
| Resolution range (Å) | 71.17 -2.15 (2.227 - 2.15) | 71.07 -2.1 (2.175 - 2.1) | 64.28 -2.47 (2.558 - 2.47) | 57.07 -1.87 (1.937 - 1.87) | 57.13 -2.03 (2.103 -2.03) | 61.57 -1.86 (1.926 - 1.86) |
| Space group | *P*$6_5$22 | *P*$6_5$22 | *P*$6_5$22 | *P*$6_5$22 | *P*$6_5$22 | *P*$6_5$22 |
| Unit cell a(Å),b(Å),c(Å) α β γ | 142.34 142.34 152.26 90 90 120 | 142.14 142.14 152.1 90 90 120 | 141.68 141.68 153 90 90 120 | 142.18 142.18 152.15 90 90 120 | 142.29 142.29 152.65 90 90 120 | 142.2 142.2 152.47 90 90 120 |
| Total reflections | 962030 (95849) | 1057061 (102944) | 623357 (64550) | 1488447 (144030) | 1173428 (114395) | 1516065 (144566) |
| Unique reflections | 49913 (4898) | 53306 (5232) | 33072 (3235) | 75058 (7363) | 59212 (5808) | 76431 (7507) |
| Multiplicity | 19.3 (19.6) | 19.8 (19.7) | 18.8 (20.0) | 19.8 (19.6) | 19.8 (19.7) | 19.8 (19.3) |
| Completeness (%) | 100.00 (100.00) | 100.00 (100.00) | 99.99 (100.00) | 99.99 (100.00) | 99.99 (100.00) | 99.99 (99.96) |
| Mean I/sigma(I) | 29.51 (5.25) | 19.90 (1.50) | 10.31 (1.51) | 23.02 (1.35) | 18.29 (1.50) | 15.64 (1.22) |
| Wilson B-factor | 38.12 | 45.47 | 57.76 | 37.06 | 40.40 | 36.15 |
| R merge † (%) | 0.06963 (0.6652) | 0.1085 (2.232) | 0.2032 (2.196) | 0.08522 (2.097) | 0.1325 (2.432) | 0.1153 (2.334) |
| R-meas | 0.07154 | 0.1114 | 0.209 | 0.08747 | 0.136 | 0.1184 |
| CC1/2 | 1 (0.954) | 0.999 (0.571) | 0.994 (0.64) | 1 (0.563) | 0.999 (0.598) | 0.999 (0.533) |
| CC* | 1 (0.988) | 1 (0.853) | 0.999 (0.884) | 1 (0.849) | 1 (0.865) | 1 (0.834) |
| R-work‡ | 0.1830 (0.2454) | 0.1977 (0.3096) | 0.1909 (0.2835) | 0.1991 (0.3230) | 0.2253 (0.3196) | 0.2061 (0.3341) |
| R-free§ | 0.2259 (0.3105) | 0.2337 (0.3581) | 0.2263 (0.3297) | 0.2333 (0.3387) | 0.2559 (0.3701) | 0.2402 (0.3563) |
| Number of non-hydrogen atoms | 4263 | 3935 | 3855 | 4242 | 3904 | 4161 |
| macromolecules | 3708 | 3667 | 3727 | 3750 | 3762 | 3770 |
| ligands | 66 | 50 | 60 | 72 | 69 | 68 |
| water | 489 | 218 | 68 | 420 | 73 | 323 |
| Protein residues | 454 | 465 | 452 | 453 | 453 | 453 |
| RMS (bonds) †† | 0.008 | 0.009 | 0.009 | 0.008 | 0.008 | 0.008 |
| RMS (angles) †† | 1.09 | 1.10 | 1.20 | 1.06 | 1.18 | 1.10 |
| Ramachandran favoured (%) | 98 | 97 | 96 | 98 | 98 | 98 |
| Ramachandran outliers (%) | 0 | 0.22 | 0.44 | 0 | 0 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Clashscore | 5.14 | 4.12 | 5.93 | 3.74 | 4.12 | 5.98 |
| Average B-factor ($^2$) | 44.80 | 50.10 | 65.30 | 44.40 | 48.10 | 43.40 |
| macromolecules | 43.20 | 50.00 | 65.20 | 43.30 | 47.80 | 42.80 |
| ligands | 47.90 | 53.80 | 77.10 | 49.70 | 54.40 | 49.40 |
| solvent | 56.00 | 50.90 | 56.20 | 53.50 | 56.70 | 48.80 |

*Table continues below*

| | **D-2HG** | **L-2HG** | **N10072a** | **N05859b** | **N10042a** | **N09996a** |
|---|---|---|---|---|---|---|
| Wavelength (Å) | 0.9173 | 0.9173 | 0.9795 | 0.9795 | 0.9173 | 0.9173 |
| Resolution range (Å) | 47.88 -1.89 (1.958 - 1.89) | 47.69 - 2.05 (2.123 - 2.05) | 29.71 - 2.099 (2.174 - 2.099) | 31.03 - 1.84 (1.906 - 1.84) | 71 - 2.06 (2.134 - 2.06) | 70.91 - 1.87 (1.937 - 1.87) |
| Space group | *P*6$_5$22 | *P*6$_5$22 | *P*6$_5$22 | *P*6$_5$22 | *P*6$_5$22 | *P*6$_5$22 |
| Unit cell a(Å),b(Å),c(Å) α β γ | 142.08 142.08 152.47 90 90 120 | 141.8 141.8 151.41 90 90 120 | 141.37 141.37 153.105 90 90 120 | 141.58 141.58 151.6 90 90 120 | 142.01 142.01 152.2 90 90 120 | 141.81 141.81 151.83 90 90 120 |
| Total reflections | 1438569 (143766) | 1108750 (112488) | 106135 (10333) | 1551312 (154963) | 1116035 (112525) | 1475583 (147490) |
| Unique reflections | 72769 (7169) | 56673 (5560) | 53089 (5188) | 77744 (7641) | 56316 (5510) | 74509 (7316) |
| Multiplicity | 19.8 (20.1) | 19.6 (20.2) | 2.0 (2.0) | 20.0 (20.3) | 19.8 (20.4) | 19.8 (20.2) |
| Completeness (%) | 99.97 (99.80) | 100.00 (100.00) | 99.90 (99.03) | 99.96 (99.82) | 99.99 (99.98) | 99.99 (99.96) |
| Mean I/sigma(I) | 18.85 (1.36) | 15.78 (1.30) | 24.65 (4.16) | 14.77 (1.29) | 18.67 (1.46) | 19.62 (1.30) |
| Wilson B-factor | 35.77 | 41.76 | 36.84 | 32.17 | 45.37 | 33.93 |
| R merge † (%) | 0.1097 (2.566) | 0.1434 (2.398) | 0.02047 (0.153) | 0.1372 (2.354) | 0.1067 (2.196) | 0.1251 (2.608) |
| R-meas | 0.1126 | 0.1472 | 0.02894 | 0.1408 | 0.1096 | 0.1284 |
| CC1/2 | 1 (0.539) | 0.999 (0.624) | 0.999 (0.933) | 0.999 (0.566) | 0.999 (0.576) | 1 (0.56) |
| CC* | 1 (0.837) | 1 (0.877) | 1 (0.982) | 1 (0.85) | 1 (0.855) | 1 (0.847) |
| R-work‡ | 0.1949 (0.3144) | 0.1970 (0.3175) | 0.1913 (0.2445) | 0.1950 (0.3054) | 0.1956 (0.3088) | 0.1820 (0.3028) |
| R-free§ | 0.2261 (0.3392) | 0.2372 (0.3568) | 0.2256 (0.2745) | 0.2315 (0.3113) | 0.2310 (0.3149) | 0.2193 (0.3650) |
| Number of non-hydrogen atoms | 4266 | 4256 | 4136 | 4343 | 4089 | 4226 |
| macromolecules | 3768 | 3762 | 3761 | 3727 | 3727 | 3752 |
| ligands | 73 | 73 | 87 | 79 | 88 | 113 |
| water | 425 | 421 | 288 | 537 | 274 | 361 |
| Protein residues | 453 | 453 | 453 | 453 | 454 | 455 |
| RMS (bonds) †† | 0.008 | 0.008 | 0.008 | 0.007 | 0.008 | 0.031 |
| RMS (angles) †† | 1.10 | 1.15 | 1.11 | 1.09 | 1.11 | 2.38 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Ramachandran favoured (%) | 97 | 98 | 97 | 98 | 98 | 98 |
| Ramachandran outliers (%) | 0.44 | 0.22 | 0.66 | 0 | 0.44 | 0.22 |
| Clashscore | 4.25 | 7.58 | 5.73 | 3.63 | 4.96 | 8.21 |
| Average B-factor ($^2$) | 43.50 | 48.30 | 42.50 | 40.00 | 54.50 | 39.70 |
| macromolecules | 42.30 | 47.10 | 42.00 | 38.20 | 54.00 | 38.80 |
| ligands | 47.70 | 54.50 | 49.90 | 44.70 | 59.70 | 47.50 |
| solvent | 53.70 | 57.80 | 47.00 | 51.80 | 60.40 | 46.60 |

*Table continues below*

| | **N10063a** | **N11297a** | **E48115b** | **N09954a** |
|---|---|---|---|---|
| Wavelength (Å) | 0.9173 | 0.9795 | 0.9278 | 0.9278 |
| Resolution range (Å) | 95.58 - 1.75 (1.813 - 1.75) | 64.74 - 2.01 (2.082 - 2.01) | 96.35 - 2.39 (2.475 - 2.39) | 71.03 - 2.55 (2.641 - 2.55) |
| Space group | $P6_522$ | $P6_522$ | $P6_522$ | $P6_522$ |
| Unit cell a(Å),b(Å),c(Å) α β γ | 141.88 141.88 152.08 90 90 120 | 141.99 141.99 152.29 90 90 120 | 143.3 143.3 152.89 90 90 120 | 142.07 142.07 151.05 90 90 120 |
| Total reflections | 1806515 (179167) | 1193588 (121676) | 912305 (91435) | 588078 (59670) |
| Unique reflections | 90823 (8915) | 60272 (5963) | 37197 (3620) | 29907 (2922) |
| Multiplicity | 19.9 (20.1) | 19.8 (20.4) | 24.5 (25.3) | 19.7 (20.4) |
| Completeness (%) | 99.99 (99.96) | 99.49 (100.00) | 100.00 (100.00) | 100.00 (100.00) |
| Mean I/sigma(I) | 19.37 (1.35) | 13.71 (1.32) | 28.89 (5.17) | 19.03 (1.42) |
| Wilson B-factor | 30.45 | 40.15 | 44.15 | 59.41 |
| R merge † (%) | 0.1009 (2.463) | 0.1438 (2.216) | 0.0974 (0.904) | 0.1716 (2.608) |
| R-meas | 0.1036 | 0.1476 | 0.09947 | 0.1761 |
| CC1/2 | 1 (0.55) | 0.999 (0.65) | 0.999 (0.936) | 0.999 (0.631) |
| CC* | 1 (0.843) | 1 (0.888) | 1 (0.983) | 1 (0.88) |
| R-work‡ | 0.1960 (0.3184) | 0.1943 (0.2940) | 0.1850 (0.2543) | 0.1911 (0.3057) |
| R-free§ | 0.2300 (0.3287) | 0.2319 (0.3042) | 0.2148 (0.3306) | 0.2465 (0.3909) |
| Number of non-hydrogen atoms | 4162 | 4277 | 4019 | 3912 |
| macromolecules | 3727 | 3752 | 3772 | 3770 |
| ligands | 82 | 97 | 86 | 74 |
| water | 353 | 428 | 161 | 68 |
| Protein residues | 453 | 453 | 448 | 453 |
| RMS (bonds) †† | 0.033 | 0.008 | 0.009 | 0.008 |
| RMS (angles) †† | 2.54 | 1.10 | 1.18 | 1.23 |
| Ramachandran favoured (%) | 97 | 98 | 98 | 95 |
| Ramachandran outliers (%) | 0 | 0 | 0.66 | 0.66 |
| Clashscore | 7.93 | 4.65 | 5.56 | 9.16 |
| Average B-factor ($^2$) | 38.00 | 47.10 | 52.70 | 65.30 |

| | | | | |
|---|---|---|---|---|
| macromolecules | 37.20 | 45.80 | 52.40 | 65.40 |
| ligands | 42.20 | 54.50 | 61.40 | 70.20 |
| solvent | 45.40 | 56.80 | 53.80 | 57.30 |

*Table continues below*

| | **MC3095** | **MC1648** | **MC3962** | **MC3960** | **MC3948** |
|---|---|---|---|---|---|
| Wavelength (Å) | 0.9763 | 0.9795 | 0.9795 | 0.9795 | 0.9795 |
| Resolution range (Å) | 95.6 - 1.95 (2.02 - 1.95) | 51.85 - 1.87 (1.937 - 1.87) | 70.86 - 1.97 (2.04 - 1.97) | 70.89 - 2.02 (2.092 - 2.02) | 71.05 - 2.12 (2.196 - 2.12) |
| Space group | *P*6$_5$22 | *P*6$_5$22 | *P*6$_5$22 | *P*6$_5$22 | *P*6$_5$22 |
| Unit cell a(Å),b(Å),c(Å) α β γ | 141.97 141.97 152.03 90 90 120 | 142.11 142.11 151.67 90 90 120 | 141.71 141.71 150.91 90 90 120 | 141.79 141.79 150.89 90 90 120 | 142.09 142.09 151.24 90 90 120 |
| Total reflections | 1305566 (129987) | 1487026 (144870) | 1256367 (128540) | 1169939 (115034) | 1019326 (101088) |
| Unique reflections | 66099 (6498) | 74171 (7245) | 63439 (6220) | 58989 (5803) | 51493 (5037) |
| Multiplicity | 19.8 (20.0) | 20.0 (20.0) | 19.8 (20.7) | 19.8 (19.8) | 19.8 (20.1) |
| Completeness (%) | 100.00 (100.00) | 99.26 (98.63) | 100.00 (100.00) | 100.00 (99.98) | 100.00 (100.00) |
| Mean I/sigma(I) | 18.67 (1.72) | 15.61 (1.30) | 16.08 (1.73) | 15.37 (1.48) | 13.56 (1.33) |
| Wilson B-factor | 38.20 | 33.44 | 37.33 | 39.15 | 39.69 |
| R merge † (%) | 0.1119 (2.216) | 0.1312 (2.223) | 0.1333 (2.071) | 0.1466 (2.27) | 0.188 (2.461) |
| R-meas | 0.1148 | 0.1346 | 0.1368 | 0.1505 | 0.193 |
| CC1/2 | 0.999 (0.585) | 0.999 (0.572) | 0.999 (0.634) | 0.999 (0.568) | 0.999 (0.531) |
| CC* | 1 (0.859) | 1 (0.853) | 1 (0.881) | 1 (0.851) | 1 (0.833) |
| R-work‡ | 0.1869 (0.3030) | 0.1902 (0.3200) | 0.1797 (0.2889) | 0.1848 (0.3013) | 0.1829 (0.2837) |
| R-free§ | 0.2237 (0.3373) | 0.2234 (0.3720) | 0.2142 (0.3178) | 0.2186 (0.3314) | 0.2159 (0.2941) |
| Number of non-hydrogen atoms | 4344 | 4361 | 4136 | 4037 | 4139 |
| macromolecules | 3791 | 3737 | 3718 | 3690 | 3742 |
| ligands | 132 | 130 | 133 | 84 | 114 |
| water | 421 | 494 | 285 | 263 | 283 |
| Protein residues | 461 | 455 | 457 | 453 | 457 |
| RMS (bonds) †† | 0.008 | 0.007 | 0.008 | 0.008 | 0.009 |
| RMS (angles) †† | 1.17 | 1.09 | 1.14 | 1.10 | 1.18 |
| Ramachandran favoured (%) | 98 | 98 | 98 | 97 | 97 |
| Ramachandran outliers (%) | 0 | 0 | 0.22 | 0.22 | 0.65 |
| Clashscore | 4.46 | 4.39 | 4.69 | 3.82 | 3.22 |
| Average B-factor (Å$^2$) | 44.80 | 40.20 | 42.80 | 44.40 | 44.60 |
| macromolecules | 43.40 | 38.70 | 41.90 | 43.80 | 44.00 |

| | | | | | |
|---|---|---|---|---|---|
| ligands | 50.10 | 47.70 | 53.10 | 53.40 | 51.50 |
| solvent | 56.20 | 50.30 | 49.90 | 49.90 | 49.70 |

**Table 40 Data collection and refinement statistics for JARID1B.**

† $R_{merge} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the $i$th measurement of reflection $hkl$ and $I(hkl)$ is the mean value of $I_i(hkl)$ for all $i$ measurements.

‡ $R_{work} = \sum_{hkl} \left| |F_{obs}| - |F_{calc}| \right| / \sum_{hkl} |F_{obs}|$, where $F_{obs}$ is the observed structure factor and $F_{calc}$ is the calculated structure factor.

§ $R_{free}$ is the same as $R_{cryst}$ except calculated with a subset (5%) of data that were excluded from the refinement calculations.

††Engh Huber (1991).

## D. MINA53/NO66 Table 1

Statistics for the highest-resolution shell are shown in parentheses.

| | **MINA53 (2OG)** | **NO66** |
|---|---|---|
| Wavelength (Å) | 0.9763 | |
| Resolution range (Å) | 77.03 - 3.09 (3.2 - 3.09) | 59.2 - 1.95 (2.02 - 1.95) |
| Space group | $P12_11$ | $P2_12_12$ |
| Unit cell<br>a(Å),b(Å),c(Å)<br>α β γ | 88.53 154.06 208.06<br>90 96.41 90 | 81.32 151.52 105.09<br>90 90 90 |
| Total reflections | 354182 (31144) | 2427624 (150757) |
| Unique reflections | 100025 (9722) | 95122 (4037) |
| Multiplicity | 3.5 (3.2) | 33.0 (37.3) |
| Completeness (%) | 98.46 (96.13) | 99.91 (99.33) |
| Mean I/sigma(I) | 13.71 (1.91) | 20.12 (2.69) |
| Wilson B-factor | 85.61 | 31.50 |
| R merge † (%) | 0.06581 (0.6077) | 0.7836 (2.525) |
| R-meas | 0.07777 | 0.7963 |
| CC1/2 | 0.998 (0.791) | 0.824 (0.681) |
| CC* | 1 (0.94) | 0.951 (0.9) |
| R-work‡ | 0.1935 (0.2971) | 0.1901 (0.3147) |
| R-free§ | 0.2532 (0.3588) | 0.2350 (0.3512) |
| Number of non-hydrogen atoms | 20721 | 8298 |
| macromolecules | 20216 | 7428 |
| ligands | 67 | 36 |
| water | 438 | 834 |
| Protein residues | 2552 | 929 |
| RMS (bonds) †† | 0.012 | 0.008 |
| RMS (angles) †† | 1.69 | 1.11 |
| Ramachandran favoured (%) | 87 | 96 |
| Ramachandran outliers (%) | 3.5 | 0.22 |
| Clashscore | 21.38 | 4.83 |
| Average B-factor ($^2$) | 90.90 | 41.30 |
| macromolecules | 91.30 | 40.90 |
| ligands | 81.50 | 43.20 |
| solvent | 75.40 | 44.00 |

**Table 41 Data collection and refinement statistics for NO66 and MINA53.**

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the $i$th measurement of reflection $hkl$ and $I(hkl)$ is the mean value of $I_i(hkl)$ for all $i$ measurements.

‡ $R_{work} = \sum_{hkl} \left| |F_{obs}| - |F_{calc}| \right| / \sum_{hkl} |F_{obs}|$, where $F_{obs}$ is the observed structure factor and $F_{calc}$ is the calculated structure factor.

§ $R_{free}$ is the same as $R_{cryst}$ except calculated with a subset (5%) of data that were excluded from the refinement calculations.

††Engh Huber (1991).

## E. Differential Scanning Fluorimetry (DSF)

We have followed a protocol developed at the SGC which uses SYPRO orange fluorescent dye in a readout [199]. All DSF experiments were performed in DSF buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 1:1000 (v/v) SYPRO Orange Dye). NO66 and MINA53 were diluted to 1 μM in DSF buffer supplemented with 50 μM $NiCl_2$ for NO66 or 50 μM $MnCl_2$ for MINA53. The solution was then plated with multichannel pipette onto a 96-well PCR plate 19.5 μl per well. Compounds in 96-well plates were available at concentration of 0.4 mM, 1 mM or 2 mM and 0.5 μl of each compound in a plate well was mixed by pipetting up and down 10 times. DSF was conducted using a Stratagene Mx3005P qPCR instrument (Agilent) with the temperature range from 20˚C to 90˚C with the gradient of 3˚C/min. The nonlinear Boltzmann fitting used to extract the thermal shift values was initially performed in GraphPad Prism (GraphPad Software, San Diego, California, USA), and then in an automated analysis macro developed for this work in Microsoft Excel 2010, substantially speeding up the process.

2OG and $MnCl_2$ titration experiments with MINA53 were performed in DSF buffer (50 mM HEPES buffer pH 7.5, 50 mM NaCl, 1:500 (v/v) SYPRO Orange Dye). The 96-well plate was prepared with 25 μl of two-fold eight-point dilution of 2OG (along the columns) and $MnCl_2$ (along the rows) with starting concentration of 100 μM each diluted in 50 mM HEPES buffer pH 7.5, 50 mM NaCl buffer. The $9^{th}$ column of the 96-well plate contained only dilution series of $MnCl_2$ and no 2OG. MINA53 (Ala26-Val464) was diluted to 4 μM in DSF buffer and 25 μl of the solution was transferred directly to the pre-prepared plate. The assay was conducted and analysed as described above.

## F. AlphaScreen activity assays

AlphaScreen (Amplified Luminescence Proximity Homogeneous Assay) inhibition assays were carried out as previously reported [191], [215] and the details on the individual assay conditions are presented in Table 42. Purified proteins were available at the SGC. AlphaScreen General IgG detection kit was purchased from Perkin Elmer. Assays were carried out in 384-well Proxiplates plus from Perkin Elmer. HEPES buffer was purchased from Life Technologies, Ferrous ammonium sulphate (FAS) ((NH4)2[Fe(SO4)2].6H20 from Riedel-deHaen) was used as a source of $Fe^{2+}$ diluted in 20 mM HCl to 400 mM concentration and then to 1 mM in deionised water. Bovine serum albumin (BSA) was obtained from Sigma Aldrich (A7030). L-ascorbic acid (LAA) and 2OG were dissolved in deionised water and both obtained from Sigma-Aldrich. FAS, 2OG and LAA were prepared fresh each day.

Assays were performed at room temperature (RT) (25° C) and compound solvent is DMSO. Compounds were dispensed to plates with Echo 550 Acoustic Dispenser (Labcyte). 5 µl of 2X enzyme solution in assay buffer was added to the plates and incubated for 15 min at RT. 5 µl of 2X substrate prepared in assay buffer with LAA, 2OG and FAS with concentrations in Table 42, and added to the plate to start the reaction. Enzymatic reactions were incubated as described in Table 42 and the reaction was stopped by addition of stop solution (30 mM EDTA, 800 mM NaCl in assay buffer). 5 µl of AlphaScreen beads previously incubated with the appropriate antibody (Table 42) was added by multichannel pipette and plates were left to incubate for at least 2h. The luminescence was detected using a BMG Labtech Pherastar FS plate reader. The data was normalised to the positive control (100 µM of 2,4-PDCA) and analysed in GraphPad Prism. $IC_{50}$ determination was done with nonlinear regression curve fitting.

Counter-screen was performed as the main assay described above with two exceptions: 1) enzyme solution was replaced with no-enzyme assay buffer solution, and 2) the peptide

substrate H3K9(Me3) substrate was replaced with 15 nM of H3K9(Me2)-21mer; ARTKQTARK(2me)STGGKAPRKQLA-GGK Spacer-Biotin (Anaspec) product peptide.

| Enzyme | FBXL11 (M1-P517) N-terminal His6-tagged | JARID1B (L882-S1761) N-terminal His6-tagged | JMJD1B (L882-S1761) N-terminal His6-tagged | JMJD2A (M1-L359) | JMJD2D (M1-G373) N-terminal His6-tagged | JMJD3 (D1141-R1641) |
|---|---|---|---|---|---|---|
| Expression system | Insect Cells | Insect Cells | Insect Cells | *E. coli* | *E. coli* | *E. coli* |
| Assay Buffer | 50 mM HEPES pH 7.5, 0.01% (v/v) Tween-20, 0.1% (w/v) BSA. *80 nM Non-Biotinylated H3K36Me2 substrate was added to FBXL11 assay buffer in order to improve the antibody readout. | | | | | |
| 2OG (µM) | 10 | 5 | 5 | 10 | 10 | 10 |
| FAS (µM) | 10 | 10 | 10 | 1 | 1 | 10 |
| LAA (µM) | 100 | 100 | 100 | 100 | 100 | 100 |
| Peptide (nM) | 100 | 100 | 60 | 30 | 30 | 60 |
| Protein Conc. (nM) | 25 | 2 | 0.4 | 4 | 2 | 0.1 |
| Peptide Description | H3K36(Me2)-21mer; biotin-SAPATGGVK(Me2)KPHRYRPGTVAL (Anaspec) | H3K4(Me3)-21mer; ARTK(3me)QTARKSTGGKAPRKQLA-GGK Spacer-Biotin (Anaspec) | H3K9(Me2)-21mer; ARTKQTARK(2me)STGGKAPRKQLA-GGK Spacer-Biotin (Anaspec) | H3K9(Me3)-21mer; ARTKQTAR-K(Me3)-STGGKAPRKQLA-GGK-Biotin (Anaspec) | | H3K27(Me3)-21mer; Biotin-KAPRKQLATKAAR(Kme3)SAPATGG (Anaspec) |
| Antibody | Anti-Histone H3K36me1 antibody (Abcam AB9048) | Anti-H3K4me2 antibody (CST 9725S) | Anti H3K9me1 antibody (Abcam Ab8896) | Anti H3K9me2 antibody (Abcam Ab1220) | | Anti H3K27me2 antibody (Millipore 07-452) |
| Solvent Conc. (%) | 0.5 | 0.5 | 0.5 | 0.5 | 1 | 0.5 |
| Reaction time (min) | 30 | 20 | 5 | 20 | 8 | 10 |
| Bead Donor (mg/ml) | 0.02 | | | | | |
| Bead Acceptor (mg/ml) | 0.02 | | | | | |
| Antibody Conc. (µg/ml) | 0.0075 | 1 in 2500 (v/v) | 0.05 | 0.05 | 0.05 | 0.4 |

| | |
|---|---|
| Bead Incubation time (min) | 120 |

**Table 42 Conditions of the AlphaScreen assays.**

## G. RapidFire activity assays

**NO66/MINA53 hydroxylation assay**

Compounds for dose response or single concentration experiments were dispensed using an Echo 550 acoustic dispenser (Labcyte, Sunnyvale, CA) into 384-well HiBase plates (Greiner Bio-one) or 96-well plates (Greiner Bio-one) with the last two columns of the plate dedicated to 1% (v/v) DMSO control and 100 mM 2,4-PDCA control respectively. The plate was backfilled to a total of 50 nl DMSO. Peptides used in the assays were NO66 substrate -Rpl8 (Asn205 -Thr224, 20-mer; NPVEHPFGGGNHQHIGKPST) and MINA53 substrate Rpl27a (Gly31 -Pro49, 20-mer; GRGNAGGLHHHRINFDKYHP) synthesized by Peptide Synthetics (Fareham, Hampshire, UK). Other reagents were from commercial sources as described in section F. FAS was prepared as described in section F.

The hydroxylation assays for NO66 and MINA53 were performed using optimized buffers. NO66 enzyme was diluted to 375 nM in 50 mM MES pH 7.0, 150 mM NaCl. MINA53 enzyme was diluted to 187.5 nM in 50 mM HEPES pH7.5, 50 mM NaCl. The assay was performed in the following order: (1) addition of 40 µl enzyme solution to compound in the desired plate format, (2) 15 min incubation at room temperature (RT) with compounds, (3) start of the reaction by addition of 10 µl of peptide solution (250 µM FAS, 500 µM LAA, 25 µM 2OG and 25 µM peptide in respective assay buffer), (4) 30 or 60 min assay incubation with MINA53 or NO66 respectively, (5) reaction stopped by addition of 5 µl 10% (v/v) formic acid. The reaction plates were either processed immediately on RapidFire mass spectrometer or frozen at -20˚C.

Time-course assays for NO66 and MINA53 at final assay concentrations were carried out in 96 2ml deep-well plates (Greiner Bio-one) as described above by scaling the total assay volume to 1 ml. Enzyme reactions were initiated by addition of 250 µl of 5x final concentration of enzyme to 1.0 ml of 1.25x final concentration of substrate. Hydroxylation of peptide substrates was monitored in real-time by programing the RapidFire to sample

the enzyme reaction at two minute intervals. The enzyme reactions were sampled over a period of 35 min to generate enzyme progress curves.

$K_M$ and $V_{max}$ determination for 2OG and Fe was performed by preparing enzyme progress curves at 8 concentrations in four replicates each. The concentrations were prepared as 8-point 2-fold dilution series with starting concentration of 20 µM. Each time series was based on 6 time points (0, 1, 5, 10, 20 and 30 min for NO66 and 0, 1, 3, 5, 7 and 10min for MINA53). Assays were performed in 96-well plates (Greiner Bio-one) and the reaction was started by addition of 10 µl concentrated enzyme solution (3000 nM of NO66 or 1500 nM of MINA53) in respective assay buffer to 40 µl of peptide solution (500 µM LAA, and 12.5 µM peptide) in respective assay buffer with excess of 2OG (62.5 µM) or FAS (125 µM) and the reactions were stopped by addition of 10 µl of 10% (v/v) formic acid. First four reaction data points starting with 1 min time point were used for estimation of the reaction velocities for each enzyme.

The peptide $K_M$ and $V_{max}$ was determined from time series with 1, 2, 4, 5, 6, 8, 10, 12 and 15 µM concentrations of Rpl27a peptide for MINA53 and Rpl8 peptide for NO66. Assays were performed in 96 2ml deep-well plates (Greiner Bio-one) and the reaction started by addition of 250 µl of 5X concentrated enzyme (1500 nM for both NO66 and MINA53) in respective optimized assay buffer to 1000 µl of 1.25X concentrated substrate (62.5 µM FAS, 125 µM LAA and 12.5 µM 2OG) in respective assay buffer and stopped directly on the RapidFire at time points of 0, 15, 30, 45, 60, 90, 120, 180s for MINA53 and NO66. Reaction velocities were calculated from data points for MINA53 at 15, 30 and 45s time points or 15, 30, 45 and 60s for 6 µM peptide concentration, and 15, 30, 45 and 60s time points for NO66. Standard curves were based on the peptide concentrations from 0 time point.

### NO66 and MINA53 demethylation assay

Demethylation assays for NO66 and MINA53 were carried out by mixing 20 μl of peptide substrate mix (50mM MES pH 7.0, 100 μM 2OG, 100 μM FAS, 200 μM LAA and 20 μM peptide) with 20 μl of 4 μM MINA53 or NO66 to the final concentrations of 50 μM 2OG, 50 μM FAS, 100 μM LAA, 10 μM peptide and 2 μM enzyme (MINA53/NO66). The reaction was incubated for 2 h at room temperature (25˚C) and stopped with addition of 5 μl of 10% (v/v) formic acid and frozen in -20˚C. The JMJD2C, JMJD1A, JMJD3, FBXL11 and JARID1C controls were prepared as described above with incubation times of 1 h. Table 43 indicates the peptide substrates for each enzyme. The readout and data analysis was performed as described below.

| Enzyme | Peptide substrate |
|---|---|
| JMJD2C | H3K9(Me3)-15mer; ARTAQTARK(Me3)STGGIA (Peptide Protein Research) |
| JMJD1A | H3K9(Me2)-21mer; ARTKQTARK(Me2)STGGKAPRKQLA (Peptide Protein Research) |
| JMJD3 | H3K27(Me3)-17 mer; LATKAARK(Me3)SAPATGGVK (Peptide Protein Research) |
| FBXL11 | H3K36(Me2)-21mer; SAPATGGVK(Me2)KPHRYRPGTVAL (Peptide Protein Research) |
| JARID1C | H3K4(Me3)-21mer; ARTK(Me3)QTARKSTGGKAPRKQLA (Peptide Protein Research) |
| MINA53 | Rpl27a; GRGNAGGLHHHRINFDKYHP, (Peptide Protein Research) |
| NO66 | Rpl8; NPVEHPFGGGNHQHIGKPST, (Peptide Protein Research) |

**Table 43 Peptide substrates for the RapidFire assay.**

### FBXL11 demethylation assay

The FBXL11 (M1-P517; N-terminal His6-tagged) assay was performed by Dr A. Tumber. Compound plates and dilution series were prepared using Echo acoustic dispenser (Labcyte) as described in NO66 and MINA53 hydroxylation assay above. Briefly, 25 μl of FBXL11 diluted to 300 nM in 50 mM MES pH 7.0 buffer was added to the compounds in 384-well plate (Greiner) format and incubated for 15 min at 25˚C. The reaction was started

with addition of 25 µl of the substrate solution (200 µM LAA, 20 µM FAS, 20 µM 2OG and 20 µM H3K36me2 peptide substrate (Peptide Protein Research, see Table 43)) and incubated for 60 min at 25˚C. The reaction was stopped by addition of 5 µM of 10% formic acid and the reaction plate analysed as described below with the signal from charge state +4.

**RapidFire and data analysis**

Samples were aspirated under vacuum onto a C4 SPE cartridge on a RapidFire RF360 high throughput sampling robot (Figure 61). Aqueous wash for 5500 ms (0.1% (v/v) formic acid in mass spectrometry grade water) was followed by elution with organic buffer for 6000 ms (85% (v/v) acetonitrile in water, 0.1% (v/v) formic acid) onto an Agilent 6530 Q-TOF. The sample sipper was washed with deionised water and 100% acetonitrile, and the SPE cartridge was washed with aqueous wash buffer in order to remove any cross contamination between sample injections. We have seen maximum signal for charge state +4 in Rpl8 and +5 in Rpl27a. Ion chromatogram data for these two charge states were extracted for both hydroxylated and non-hydroxylated peptide and peak area was integrated using Agilent RapidFire Integrator v. 3.6.

For the $IC_{50}$ or single concentration inhibition studies the data was normalised to the 100 µM 2,4-PDCA positive control and the % inhibition calculated in Microsoft Excel according to the formula:

$$\%\ inhibition\ (x) = 1 - \frac{x\ -\ Average\ (positive\ control)}{Average\ (\ DMSO\ signal)\ -\ Average\ (positive\ control)}$$

where *x* represents an assay data point.

For the enzymatic constant determination the % conversion was calculated in Microsoft Excel. Data was analysed using GraphPad Prism and nonlinear curve fitting was

performed with Michaelis–Menten equation for enzyme kinetics studies and sigmoidal dose response curve with variable slope for the determination of the $IC_{50}$ values.

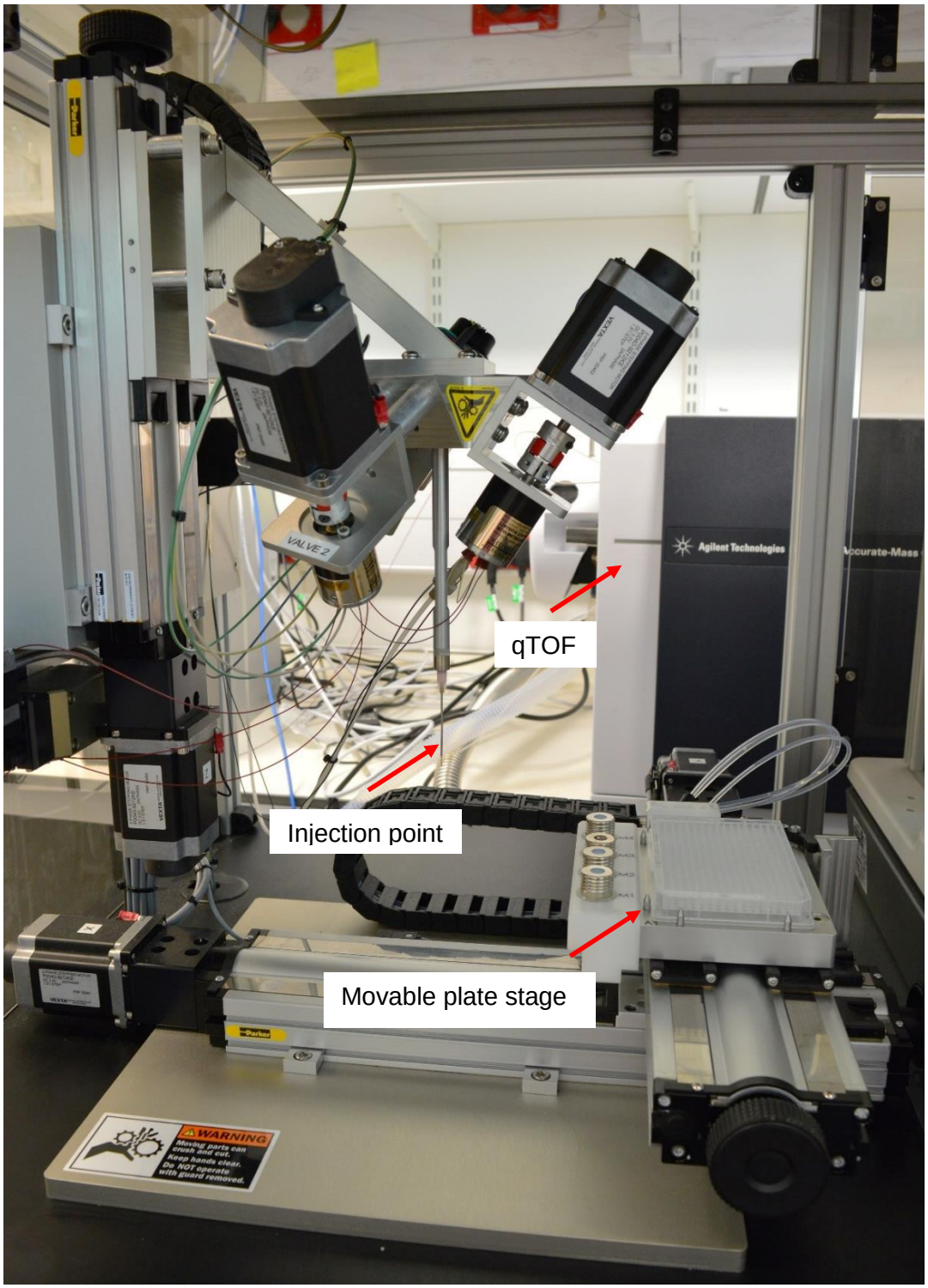



**Figure 61. Automated sample injection system of the RapidFire mass spectrometer.**

## H. Biolayer Interferometry

Biolayer interferometry experiments were performed on ForteBio Octet RED384 instrument using 16 sensors at 25 °C in 50 mM HEPES, 50 mM NaCl, 0.05% (v/v) TWEEN supplemented with 10 μM $MnCl_2$ salt as a substitute for Fe. Biotinylated MINA53 was attached to Super Streptavidin (SSA) Biosensors (FórteBio) by incubation for 160-s at 2 μM concentration. Compound sample was prepared in assay buffer by 2-fold dilution with starting concentration of 25 μM dispensed 50 μl per well. Measurements were performed using a 120-s baseline in assay buffer followed by 300-s association in compound sample and 300-s dissociation step in assay buffer on a black 384-well plate (Greiner). Reference signal was generated as described above with the sensors without the immobilised protein in order to remove the effect of unspecific binding to the sensor tips. Reference subtraction to correct for baseline drift and nonspecific binding was performed with the FórteBio data analysis software. $K_D$ calculations were performed using the Analysis software (ForteBio).

## I. Cellular assays

*Demethylation assay*

Assay was performed and analysed by Dr C. Yapp and this DPhil candidate (R. Nowak). U2OS/HeLa cells were grown, transfected, immunostained and imaged as described in [161]. Following antibodies were used: anti H3K4me3 (Diagenode, CI5410003), Monoclonal anti flag M2 Ab produced in mouse (Sigma Aldrich, F3165), Alexa Fluor 488 goat anti-rabbit IgG (Life Technologies, A11034), Alexa Fluor 568 goat- anti mouse IgG1(γ1) (Life Technologies, A21124).

*Cellular proliferation assay*

Assay was performed by N. Wu and analysed by this DPhil candidate. Briefly, human multiple myeloma cell lines MM1S, JIM3, U226, KMS11 and JJN3 were cultured in a 96-well plate at a density of 2000 cells per well and the viability was assessed using PrestoBlue® reagent (Thermo Fisher) using a FluroStar Optima (BMG) according to the manufacturer's instructions. Data was normalised to a vehicle control and plotted using GraphPad Prism.

## J. Small molecule docking

The docking was performed in ICM (Molsoft, San Diego, US) with the presence of solvent [218]. The crystal structure of JARID1B in complex with MC1648 (with MC1648 ligand removed), was used as a docking model for unconstrained docking with compounds MC1648, MC3949, MC3960, MC3963 and MC3948. The DMS3 molecule and MC1648 ligand were removed from MC1648 model which was used as a docking template for unconstrained docking of MC3971 and MC3970. Docked compounds were created in ICM [218]. Five docked conformations were generated each run and the top scoring conformation was chosen.

## K. Phylogenetic tree of 2OG oxygenases

Conserved catalytic JmjC domain of the human of $Fe^{2+}$- and 2OG- dependent oxygenases was selected as the basis of phylogenetic analysis. To allow for diversity, sequence of JmjC domain was expanded from N-terminus and C-terminus side by adding amino acids representing one secondary structural feature (alpha helix or beta sheet) based on secondary structure prediction (Psipred Server [230]). Multiple sequence alignment servers were used (Clustal Omega, ClustalW, ClustalW2, MUSCLE, MAFFT) and the results assessed based on assessment of the alignment of conserved metal binding residues. The final alignment was performed using MAFFT v6.850 server (http://www.ebi.ac.uk/Tools/msa/mafft/) [231] and corrected manually using ICM Pro (Molsoft LLC. San Diego, US). The bootstrap N-J tree was generated from the alignment file using ClustalX program v1.81 with 1000 bootstrap seeds [232]. The tree was drawn in PHYLODRAW v0.8 [233] and then processed in Adobe Illustrator.

# Bibliography


[1] M. A. Dawson and T. Kouzarides, “Cancer epigenetics: from mechanism to therapy.,” *Cell*, vol. 150, no. 1, pp. 12–27, Jul. 2012.

[2] P. Blancafort, J. Jin, and S. Frye, “Writing and rewriting the epigenetic code of cancer cells: from engineered proteins to small molecules.,” *Mol. Pharmacol.*, vol. 83, no. 3, pp. 563–76, Mar. 2013.

[3] K. J. Falkenberg and R. W. Johnstone, “Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders.,” *Nat. Rev. Drug Discov.*, vol. 13, no. 9, pp. 673–691, Aug. 2014.

[4] T. Lu, M. W. Jackson, B. Wang, M. Yang, M. R. Chance, M. Miyagi, A. V Gudkov, and G. R. Stark, “Regulation of NF-kappaB by NSD1/FBXL11-dependent reversible lysine methylation of p65.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 1, pp. 46–51, Jan. 2010.

[5] B. S. Mann, J. R. Johnson, M. H. Cohen, R. Justice, and R. Pazdur, “FDA approval summary: vorinostat for treatment of advanced primary cutaneous T-cell lymphoma.,” *Oncologist*, vol. 12, no. 10, pp. 1247–52, Oct. 2007.

[6] R. Frye, M. Myers, K. C. Axelrod, E. A. Ness, R. L. Piekarz, S. E. Bates, and S. Booher, “Romidepsin: A New Drug for the Treatment of Cutaneous T-Cell Lymphoma | CJON,” *Clinical journal of oncology nursing*, Apr-2012. [Online]. Available: https://cjon.ons.org/cjon/16/2/romidepsin-new-drug-treatment-cutaneous-t-cell-lymphoma. [Accessed: 01-Aug-2015].

[7] R. Sanchez, J. Meslamani, and M.-M. Zhou, “The bromodomain: from epigenome reader to druggable target.,” *Biochim. Biophys. Acta*, vol. 1839, no. 8, pp. 676–85, Aug. 2014.

[8] P. Filippakopoulos, J. Qi, S. Picaud, Y. Shen, W. B. Smith, O. Fedorov, E. M. Morse, T. Keates, T. T. Hickman, I. Felletar, M. Philpott, S. Munro, M. R. McKeown, Y. Wang, A. L. Christie, N. West, M. J. Cameron, B. Schwartz, T. D. Heightman, N. La Thangue, C. A. French, O. Wiest, A. L. Kung, S. Knapp, and J. E. Bradner, “Selective inhibition of BET bromodomains.,” *Nature*, vol. 468, no. 7327, pp. 1067–73, Dec. 2010.

[9] J. E. Delmore, G. C. Issa, M. E. Lemieux, P. B. Rahl, J. Shi, H. M. Jacobs, E. Kastritis, T. Gilpatrick, R. M. Paranal, J. Qi, M. Chesi, A. C. Schinzel, M. R. McKeown, T. P. Heffernan, C. R. Vakoc, P. L. Bergsagel, I. M. Ghobrial, P. G. Richardson, R. A. Young, W. C. Hahn, K. C. Anderson, A. L. Kung, J. E. Bradner, and C. S. Mitsiades, “BET bromodomain inhibition as a therapeutic strategy to target c-Myc.,” *Cell*, vol. 146, no. 6, pp. 904–17, Sep. 2011.

[10] B. Xu, K. D. Konze, J. Jin, and G. G. Wang, “Targeting EZH2 and PRC2 dependence as novel anticancer therapy.,” *Exp. Hematol.*, May 2015.

[11] S. K. Verma, X. Tian, L. V LaFrance, C. Duquenne, D. P. Suarez, K. A. Newlander, S. P. Romeril, J. L. Burgess, S. W. Grant, J. A. Brackley, A. P. Graves, D. A. Scherzer, A. Shu, C. Thompson, H. M. Ott, G. S. Van Aller, C. A. Machutta, E. Diaz, Y. Jiang, N. W. Johnson, S. D. Knight, R. G. Kruger, M. T.

McCabe, D. Dhanak, P. J. Tummino, C. L. Creasy, and W. H. Miller, “Identification of Potent, Selective, Cell-Active Inhibitors of the Histone Lysine Methyltransferase EZH2.,” *ACS Med. Chem. Lett.*, vol. 3, no. 12, pp. 1091–6, Dec. 2012.

[12] J. E. Campbell, K. W. Kuntz, S. K. Knutson, N. M. Warholic, H. Keilhack, T. J. Wigle, A. Raimondi, C. R. Klaus, N. Rioux, A. Yokoi, S. Kawano, Y. Minoshima, H.-W. Choi, M. Porter Scott, N. J. Waters, J. J. Smith, R. Chesworth, M. P. Moyer, and R. A. Copeland, “EPZ011989, A Potent, Orally-Available EZH2 Inhibitor with Robust in Vivo Activity.,” *ACS Med. Chem. Lett.*, vol. 6, no. 5, pp. 491–5, May 2015.

[13] C. D. Allis, S. L. Berger, J. Cote, S. Dent, T. Jenuwien, T. Kouzarides, L. Pillus, D. Reinberg, Y. Shi, R. Shiekhattar, A. Shilatifard, J. Workman, and Y. Zhang, “New nomenclature for chromatin-modifying enzymes.,” *Cell*, vol. 131, no. 4, pp. 633–6, Nov. 2007.

[14] C. Johansson, A. Tumber, K. Che, P. Cain, R. Nowak, C. Gileadi, and U. Oppermann, “The roles of Jumonji-type oxygenases in human disease.,” *Epigenomics*, vol. 6, no. 1, pp. 89–120, Feb. 2014.

[15] S. B. Rothbart and B. D. Strahl, “Interpreting the language of histone and DNA modifications.,” *Biochim. Biophys. Acta*, vol. 1839, no. 8, pp. 627–43, Aug. 2014.

[16] A. J. Bannister and T. Kouzarides, “Regulation of chromatin by histone modifications.,” *Cell Res.*, vol. 21, no. 3, pp. 381–95, Mar. 2011.

[17] J. Cadet and J. R. Wagner, “TET enzymatic oxidation of 5-methylcytosine, 5-hydroxymethylcytosine and 5-formylcytosine.,” *Mutat. Res. Genet. Toxicol. Environ. Mutagen.*, vol. 764–765, pp. 18–35, Apr. 2014.

[18] Z. Liutkevičiūtė, E. Kriukienė, J. Ličytė, M. Rudytė, G. Urbanavičiūtė, and S. Klimašauskas, “Direct decarboxylation of 5-carboxylcytosine by DNA C5-methyltransferases.,” *J. Am. Chem. Soc.*, vol. 136, no. 16, pp. 5884–7, May 2014.

[19] J. Bonnet, D. Devys, and L. Tora, “Histone H2B ubiquitination: signaling not scrapping.,” *Drug Discov. Today. Technol.*, vol. 12, pp. e19–27, Jul. 2014.

[20] Y. Shiio and R. N. Eisenman, “Histone sumoylation is associated with transcriptional repression.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 23, pp. 13225–30, Nov. 2003.

[21] Y.-M. Xu, J.-Y. Du, and A. T. Y. Lau, “Posttranslational modifications of human histone H3: an update.,” *Proteomics*, vol. 14, no. 17–18, pp. 2047–60, Sep. 2014.

[22] D. J. Slade, S. Horibata, S. A. Coonrod, and P. R. Thompson, “A novel role for protein arginine deiminase 4 in pluripotency: the emerging role of citrullinated histone H1 in cellular programming.,” *Bioessays*, vol. 36, no. 8, pp. 736–40, Aug. 2014.

[23] C. J. Nelson, H. Santos-Rosa, and T. Kouzarides, “Proline isomerization of histone H3 regulates lysine methylation and gene expression.,” *Cell*, vol. 126, no. 5, pp. 905–16, Oct. 2006.

[24] A. Böttger, M. S. Islam, R. Chowdhury, C. J. Schofield, and A. Wolf, “The oxygenase Jmjd6--a case study in conflicting assignments.,” *Biochem. J.*, vol. 468, no. 2, pp. 191–202, Jun. 2015.

[25] Y. Shi, F. Lan, C. Matson, P. Mulligan, J. R. Whetstine, P. A. Cole, R. A. Casero, and Y. Shi, “Histone demethylation mediated by the nuclear amine oxidase homolog LSD1.,” *Cell*, vol. 119, no. 7, pp. 941–53, Dec. 2004.

[26] E. Metzger, M. Wissmann, N. Yin, J. M. Müller, R. Schneider, A. H. F. M. Peters, T. Günther, R. Buettner, and R. Schüle, “LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription.,” *Nature*, vol. 437, no. 7057, pp. 436–9, Sep. 2005.

[27] S. M. Kooistra and K. Helin, “Molecular mechanisms and potential functions of histone demethylases.,” *Nat. Rev. Mol. Cell Biol.*, vol. 13, no. 5, pp. 297–311, May 2012.

[28] T. Maes, E. Carceller, J. Salas, A. Ortega, and C. Buesa, “Advances in the development of histone lysine demethylase inhibitors.,” *Curr. Opin. Pharmacol.*, vol. 23, pp. 52–60, Jun. 2015.

[29] J. Meiser, D. Weindl, and K. Hiller, “Complexity of dopamine metabolism.,” *Cell Commun. Signal.*, vol. 11, no. 1, p. 34, Jan. 2013.

[30] W. Ge, A. Wolf, T. Feng, C. Ho, R. Sekirnik, A. Zayer, N. Granatino, M. E. Cockman, C. Loenarz, N. D. Loik, A. P. Hardy, T. D. W. Claridge, R. B. Hamed, R. Chowdhury, L. Gong, C. V Robinson, D. C. Trudgian, M. Jiang, M. M. Mackeen, J. S. McCullagh, Y. Gordiyenko, A. Thalhammer, A. Yamamoto, M. Yang, P. Liu-Yi, Z. Zhang, M. Schmidt-Zachmann, B. M. Kessler, P. J. Ratcliffe, G. M. Preston, M. L. Coleman, and C. J. Schofield, “Oxygenase-catalyzed ribosome hydroxylation occurs in prokaryotes and humans.,” *Nat. Chem. Biol.*, vol. 8, no. 12, pp. 960–2, Dec. 2012.

[31] Y. Tsukada, J. Fang, H. Erdjument-Bromage, M. E. Warren, C. H. Borchers, P. Tempst, and Y. Zhang, “Histone demethylation by a family of JmjC domain-containing proteins.,” *Nature*, vol. 439, no. 7078, pp. 811–6, Feb. 2006.

[32] X. Wu, J. V. Johansen, and K. Helin, “Fbxl10/Kdm2b recruits polycomb repressive complex 1 to CpG islands and regulates H2A ubiquitylation.,” *Mol. Cell*, vol. 49, no. 6, pp. 1134–46, Mar. 2013.

[33] L. Liu, H. Kim, A. Casta, Y. Kobayashi, L. S. Shapiro, and A. M. Christiano, “Hairless is a histone H3K9 demethylase.,” *FASEB J.*, vol. 28, no. 4, pp. 1534–42, Apr. 2014.

[34] S. Watanabe, K. Watanabe, V. Akimov, J. Bartkova, B. Blagoev, J. Lukas, and J. Bartek, “JMJD1C demethylates MDC1 to regulate the RNF8 and BRCA1-mediated chromatin response to DNA breaks.,” *Nat. Struct. Mol. Biol.*, vol. 20, no. 12, pp. 1425–33, Dec. 2013.

[35] S. Wellmann, M. Bettkober, A. Zelmer, K. Seeger, M. Faigle, H. K. Eltzschig, and C. Bührer, “Hypoxia upregulates the histone demethylase JMJD1A via HIF-1.,” *Biochem. Biophys. Res. Commun.*, vol. 372, no. 4, pp. 892–7, Aug. 2008.

[36] D. Yamada, S. Kobayashi, H. Yamamoto, Y. Tomimaru, T. Noda, M. Uemura, H. Wada, S. Marubashi, H. Eguchi, M. Tanemura, Y. Doki, M. Mori, and H. Nagano, “Role of the hypoxia-related gene, JMJD1A, in hepatocellular carcinoma: clinical impact on recurrence after hepatic resection.,” *Ann. Surg. Oncol.*, vol. 19 Suppl 3, pp. S355–64, Jul. 2012.

[37] T. Osawa, R. Tsuchida, M. Muramatsu, T. Shimamura, F. Wang, J. Suehiro, Y. Kanki, Y. Wada, Y. Yuasa, H. Aburatani, S. Miyano, T. Minami, T. Kodama, and M. Shibuya, “Inhibition of histone demethylase JMJD1A improves anti-angiogenic therapy and reduces tumor-associated macrophages.,” *Cancer Res.*, vol. 73, no. 10, pp. 3019–28, May 2013.

[38] R. N. MacKinnon, G. Kannourakis, M. Wall, and L. J. Campbell, “A cryptic deletion in 5q31.2 provides further evidence for a minimally deleted region in myelodysplastic syndromes.,” *Cancer Genet.*, vol. 204, no. 4, pp. 187–94, Apr. 2011.

[39] M.-H. Tang, V. Varadan, S. Kamalakaran, M. Q. Zhang, N. Dimitrova, and J. Hicks, “Major chromosomal breakpoint intervals in breast cancer co-localize with differentially methylated regions.,” *Front. Oncol.*, vol. 2, p. 197, Jan. 2012.

[40] M. Brauchle, Z. Yao, R. Arora, S. Thigale, I. Clay, B. Inverardi, J. Fletcher, P. Taslimi, M. G. Acker, B. Gerrits, J. Voshol, A. Bauer, D. Schübeler, T. Bouwmeester, and H. Ruffner, “Protein complex interactor analysis and differential activity of KDM3 subfamily members towards H3K9 methylation.,” *PLoS One*, vol. 8, no. 4, p. e60549, Jan. 2013.

[41] S.-M. Kim, J.-Y. Kim, N.-W. Choe, I.-H. Cho, J.-R. Kim, D.-W. Kim, J.-E. Seol, S. E. Lee, H. Kook, K.-I. Nam, H. Kook, Y.-Y. Bhak, and S.-B. Seo, “Regulation of mouse steroidogenesis by WHISTLE and JMJD1C through histone methylation balance.,” *Nucleic Acids Res.*, vol. 38, no. 19, pp. 6389–403, Oct. 2010.

[42] P. Sroczynska, V. A. Cruickshank, J.-P. Bukowski, S. Miyagi, F. O. Bagger, J. Walfridsson, M. B. Schuster, B. Porse, and K. Helin, “shRNA screening identifies JMJD1C as being required for leukemia maintenance.,” *Blood*, vol. 123, no. 12, pp. 1870–82, Mar. 2014.

[43] W. Ahmad, A. A. Panteleyev, and A. M. Christiano, “The molecular basis of congenital atrichia in humans and mice: mutations in the hairless gene.,” *J. Investig. Dermatol. Symp. Proc.*, vol. 4, no. 3, pp. 240–3, Dec. 1999.

[44] P. Trojer, J. Zhang, M. Yonezawa, A. Schmidt, H. Zheng, T. Jenuwein, and D. Reinberg, “Dynamic Histone H1 Isotype 4 Methylation and Demethylation by Histone Lysine Methyltransferase G9a/KMT1C and the Jumonji Domain-containing JMJD2/KDM4 Proteins.,” *J. Biol. Chem.*, vol. 284, no. 13, pp. 8395–405, Mar. 2009.

[45] L. Hillringhaus, W. W. Yue, N. R. Rose, S. S. Ng, C. Gileadi, C. Loenarz, S. H. Bello, J. E. Bray, C. J. Schofield, and U. Oppermann, “Structural and evolutionary basis for the dual substrate selectivity of human KDM4 histone demethylase family.,” *J. Biol. Chem.*, vol. 286, no. 48, pp. 41616–25, Dec. 2011.

[46] Y. Katoh and M. Katoh, “Comparative integromics on JMJD2A, JMJD2B and JMJD2C: Preferential expression of JMJD2C in undifferentiated ES cells,” *Int. J. Mol. Med.*, vol. 20, no. 2, pp. 269–273, Aug. 2007.

[47] P. J. Pollard, C. Loenarz, D. R. Mole, M. A. McDonough, J. M. Gleadle, C. J. Schofield, and P. J. Ratcliffe, “Regulation of Jumonji-domain-containing histone demethylases by hypoxia-inducible factor (HIF)-1alpha.,” *Biochem. J.*, vol. 416, no. 3, pp. 387–94, Dec. 2008.

[48] T.-D. Kim, S. Oh, S. Shin, and R. Janknecht, “Regulation of tumor suppressor p53 and HCT116 cell physiology by histone demethylase JMJD2D/KDM4D.,” *PLoS One*, vol. 7, no. 4, p. e34618, Jan. 2012.

[49] Q.-J. Zhang, H.-Z. Chen, L. Wang, D.-P. Liu, J. A. Hill, and Z.-P. Liu, “The histone trimethyllysine demethylase JMJD2A promotes cardiac hypertrophy in response to hypertrophic stimuli in mice.,” *J. Clin. Invest.*, vol. 121, no. 6, pp. 2447–56, Jun. 2011.

[50] L. Ye, Z. Fan, B. Yu, J. Chang, K. Al Hezaimi, X. Zhou, N.-H. Park, and C.-Y. Wang, “Histone demethylases KDM4B and KDM6B promotes osteogenic differentiation of human MSCs.,” *Cell Stem Cell*, vol. 11, no. 1, pp. 50–61, Jul. 2012.

[51] J. C. Black and J. R. Whetstine, “Too little O2 Too much gain.,” *Cell Cycle*, Jul. 2015.

[52] D. Hanahan and R. A. Weinberg, “Hallmarks of cancer: the next generation.,” *Cell*, vol. 144, no. 5, pp. 646–74, Mar. 2011.

[53] J. C. Black, E. Atabakhsh, J. Kim, K. M. Biette, C. Van Rechem, B. Ladd, P. D. Burrowes, C. Donado, H. Mattoo, B. P. Kleinstiver, B. Song, G. Andriani, J. K. Joung, O. Iliopoulos, C. Montagna, S. Pillai, G. Getz, and J. R. Whetstine, “Hypoxia drives transient site-specific copy gain and drug-resistant gene expression.,” *Genes Dev.*, vol. 29, no. 10, pp. 1018–31, May 2015.

[54] R. M. Labbé, A. Holowatyj, and Z.-Q. Yang, “Histone lysine demethylase (KDM) subfamily 4: structures, functions and therapeutic potential.,” *Am. J. Transl. Res.*, vol. 6, no. 1, pp. 1–15, Jan. 2013.

[55] Y.-W. Chin and S.-Y. Han, “KDM4 histone demethylase inhibitors for anti-cancer agents: a patent review.,” *Expert Opin. Ther. Pat.*, vol. 25, no. 2, pp. 135–44, Mar. 2015.

[56] J. Christensen, K. Agger, P. A. C. Cloos, D. Pasini, S. Rose, L. Sennels, J. Rappsilber, K. H. Hansen, A. E. Salcini, and K. Helin, “RBP2 belongs to a family of demethylases, specific for tri-and dimethylated lysine 4 on histone 3.,” *Cell*, vol. 128, no. 6, pp. 1063–76, Mar. 2007.

[57] R. J. Klose, Q. Yan, Z. Tothova, K. Yamane, H. Erdjument-Bromage, P. Tempst, D. G. Gilliland, Y. Zhang, and W. G. Kaelin, “The retinoblastoma binding protein RBP2 is an H3K4 demethylase.,” *Cell*, vol. 128, no. 5, pp. 889–900, Mar. 2007.

[58] S. Iwase, F. Lan, P. Bayliss, L. de la Torre-Ubieta, M. Huarte, H. H. Qi, J. R. Whetstine, A. Bonni, T. M. Roberts, and Y. Shi, “The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases.,” *Cell*, vol. 128, no. 6, pp. 1077–88, Mar. 2007.

[59] D. J. Seward, G. Cubberley, S. Kim, M. Schonewald, L. Zhang, B. Tripet, and D. L. Bentley, “Demethylation of trimethylated histone H3 Lys4 in vivo by JARID1 JmjC proteins.,” *Nat. Struct. Mol. Biol.*, vol. 14, no. 3, pp. 240–2, Mar. 2007.

[60] E. V Benevolenskaya, H. L. Murray, P. Branton, R. A. Young, and W. G. Kaelin, “Binding of pRB to the PHD protein RBP2 promotes cellular differentiation.,” *Mol. Cell*, vol. 18, no. 6, pp. 623–35, Jun. 2005.

[61] L. DiTacchio, H. D. Le, C. Vollmers, M. Hatori, M. Witcher, J. Secombe, and S. Panda, “Histone lysine demethylase JARID1a activates CLOCK-BMAL1 and influences the circadian clock.,” *Science*, vol. 333, no. 6051, pp. 1881–5, Sep. 2011.

[62] J. D. E. de Rooij, I. H. I. M. Hollink, S. T. C. J. M. Arentsen-Peters, J. F. van Galen, H. Berna Beverloo, A. Baruchel, J. Trka, D. Reinhardt, E. Sonneveld, M. Zimmermann, T. A. Alonzo, R. Pieters, S. Meshinchi, M. M. van den Heuvel-Eibrink, and C. M. Zwaan, “NUP98/JARID1A is a novel recurrent abnormality in pediatric acute megakaryoblastic leukemia with a distinct HOX gene expression pattern.,” *Leukemia*, vol. 27, no. 12, pp. 2280–8, Dec. 2013.

[63] S. V Sharma, D. Y. Lee, B. Li, M. P. Quinlan, F. Takahashi, S. Maheswaran, U. McDermott, N. Azizian, L. Zou, M. A. Fischbach, K.-K. Wong, K. Brandstetter, B. Wittner, S. Ramaswamy, M. Classon, and J. Settleman, “A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations.,” *Cell*, vol. 141, no. 1, pp. 69–80, Apr. 2010.

[64] S. U. Schmitz, M. Albert, M. Malatesta, L. Morey, J. V Johansen, M. Bak, N. Tommerup, I. Abarrategui, and K. Helin, “Jarid1b targets genes regulating development and is involved in neural differentiation.,” *EMBO J.*, vol. 30, no. 22, pp. 4586–600, Nov. 2011.

[65] X. Niu, T. Zhang, L. Liao, L. Zhou, D. J. Lindner, M. Zhou, B. Rini, Q. Yan, and H. Yang, “The von Hippel-Lindau tumor suppressor protein regulates gene expression and tumor growth through histone demethylase JARID1C.,” *Oncogene*, vol. 31, no. 6, pp. 776–86, Mar. 2012.

[66] C. Akimoto, H. Kitagawa, T. Matsumoto, and S. Kato, “Spermatogenesis-specific association of SMCY and MSH5.,” *Genes Cells*, vol. 13, no. 6, pp. 623–33, Jun. 2008.

[67] G. Perinchery, M. Sasaki, A. Angan, V. Kumar, P. Carroll, and R. Dahiya, “Deletion of Y-chromosome specific genes in human prostate cancer.,” *J. Urol.*, vol. 163, no. 4, pp. 1339–42, Apr. 2000.

[68] L. J. Walport, R. J. Hopkinson, M. Vollmar, S. K. Madden, C. Gileadi, U. Oppermann, C. J. Schofield, and C. Johansson, “Human UTY(KDM6C) is a male-specific Nε-methyl lysyl demethylase.,” *J. Biol. Chem.*, vol. 289, no. 26, pp. 18302–13, Jun. 2014.

[69] K. B. Shpargel, J. Starmer, D. Yee, M. Pohlers, and T. Magnuson, “KDM6 demethylase independent loss of histone H3 lysine 27 trimethylation during early embryonic development.,” *PLoS Genet.*, vol. 10, no. 8, p. e1004507, Aug. 2014.

[70] R. Margueron and D. Reinberg, “The Polycomb complex PRC2 and its mark in life.,” *Nature*, vol. 469, no. 7330, pp. 343–9, Jan. 2011.

[71] S. Canovas, J. B. Cibelli, and P. J. Ross, “Jumonji domain-containing protein 3 regulates histone 3 lysine 27 methylation during bovine preimplantation development.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 109, no. 7, pp. 2400–5, Mar. 2012.

[72] W. Zhao, Q. Li, S. Ayers, Y. Gu, Z. Shi, Q. Zhu, Y. Chen, H. Y. Wang, and R.-F. Wang, “Jmjd3 inhibits reprogramming by upregulating expression of INK4a/Arf and targeting PHF20 for ubiquitination.,” *Cell*, vol. 152, no. 5, pp. 1037–50, Mar. 2013.

[73] T. Yasui, J. Hirose, S. Tsutsumi, K. Nakamura, H. Aburatani, and S. Tanaka, “Epigenetic regulation of osteoclast differentiation: possible involvement of Jmjd3 in the histone demethylation of Nfatc1.,” *J. Bone Miner. Res.*, vol. 26, no. 11, pp. 2665–71, Nov. 2011.

[74] T. Burgold, N. Voituron, M. Caganova, P. P. Tripathi, C. Menuet, B. K. Tusi, F. Spreafico, M. Bévengut, C. Gestreau, S. Buontempo, A. Simeone, L. Kruidenier, G. Natoli, S. Casola, G. Hilaire, and G. Testa, “The H3K27 demethylase JMJD3 is required for maintenance of the embryonic respiratory neuronal network, neonatal breathing, and survival.,” *Cell Rep.*, vol. 2, no. 5, pp. 1244–58, Nov. 2012.

[75] K. Jepsen, D. Solum, T. Zhou, R. J. McEvilly, H.-J. Kim, C. K. Glass, O. Hermanson, and M. G. Rosenfeld, “SMRT-mediated repression of an H3K27 demethylase in progression from neural stem cell to neuron.,” *Nature*, vol. 450, no. 7168, pp. 415–9, Nov. 2007.

[76] F. De Santa, M. G. Totaro, E. Prosperini, S. Notarbartolo, G. Testa, and G. Natoli, “The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing.,” *Cell*, vol. 130, no. 6, pp. 1083–94, Sep. 2007.

[77] L. Kruidenier, C. Chung, Z. Cheng, J. Liddle, K. Che, G. Joberty, M. Bantscheff, C. Bountra, A. Bridges, H. Diallo, D. Eberhard, S. Hutchinson, E. Jones, R. Katso, M. Leveridge, P. K. Mander, J. Mosley, C. Ramirez-Molina, P. Rowland, C. J. Schofield, R. J. Sheppard, J. E. Smith, C. Swales, R. Tanner, P. Thomas, A. Tumber, G. Drewes, U. Oppermann, D. J. Patel, K. Lee, and D. M. Wilson, “A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response.,” *Nature*, vol. 488, no. 7411, pp. 404–8, Aug. 2012.

[78] Q. Zhang, H. Long, J. Liao, M. Zhao, G. Liang, X. Wu, P. Zhang, S. Ding, S. Luo, and Q. Lu, “Inhibited expression of hematopoietic progenitor kinase 1 associated with loss of jumonji domain containing 3 promoter binding contributes to autoimmunity in systemic lupus erythematosus.,” *J. Autoimmun.*, vol. 37, no. 3, pp. 180–9, Nov. 2011.

[79] S. Manna, J. K. Kim, C. Baugé, M. Cam, Y. Zhao, J. Shetty, M. S. Vacchio, E. Castro, B. Tran, L. Tessarollo, and R. Bosselut, “Histone H3 Lysine 27

demethylases Jmjd3 and Utx are required for T-cell differentiation.," *Nat. Commun.*, vol. 6, p. 8152, Jan. 2015.

[80] G. L. Sen, D. E. Webster, D. I. Barragan, H. Y. Chang, and P. A. Khavari, "Control of differentiation in a self-renewing mammalian tissue by the histone demethylase JMJD3.," *Genes Dev.*, vol. 22, no. 14, pp. 1865–70, Jul. 2008.

[81] T. Shaw and P. Martin, "Epigenetic reprogramming during wound healing: loss of polycomb-mediated silencing may enable upregulation of repair genes.," *EMBO Rep.*, vol. 10, no. 8, pp. 881–6, Aug. 2009.

[82] A. Svotelis, S. Bianco, J. Madore, G. Huppé, A. Nordell-Markovits, A.-M. Mes-Masson, and N. Gévry, "H3K27 demethylation by JMJD3 at a poised enhancer of anti-apoptotic gene BCL2 determines ERα ligand dependency.," *EMBO J.*, vol. 30, no. 19, pp. 3947–61, Oct. 2011.

[83] M. E. McLaughlin-Drubin, C. P. Crum, and K. Münger, "Human papillomavirus E7 oncoprotein induces KDM6A and KDM6B histone demethylase expression and causes epigenetic reprogramming.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 108, no. 5, pp. 2130–5, Mar. 2011.

[84] K. Agger, P. A. C. Cloos, L. Rudkjaer, K. Williams, G. Andersen, J. Christensen, and K. Helin, "The H3K27me3 demethylase JMJD3 contributes to the activation of the INK4A-ARF locus in response to oncogene- and stress-induced senescence.," *Genes Dev.*, vol. 23, no. 10, pp. 1171–6, May 2009.

[85] H. Agherbi, A. Gaussmann-Wenger, C. Verthuy, L. Chasson, M. Serrano, and M. Djabali, "Polycomb mediated epigenetic silencing and replication timing at the INK4a/ARF locus during senescence.," *PLoS One*, vol. 4, no. 5, p. e5622, Jan. 2009.

[86] M. Barradas, E. Anderton, J. C. Acosta, S. Li, A. Banito, M. Rodriguez-Niedenführ, G. Maertens, M. Banck, M.-M. Zhou, M. J. Walsh, G. Peters, and J. Gil, "Histone demethylase JMJD3 contributes to epigenetic control of INK4a/ARF by oncogenic RAS.," *Genes Dev.*, vol. 23, no. 10, pp. 1177–82, May 2009.

[87] M. Priolo, L. Micale, B. Augello, C. Fusco, F. Zucchetti, P. Prontera, V. Paduano, E. Biamino, A. Selicorni, C. Mammì, C. Laganà, L. Zelante, and G. Merla, "Absence of deletion and duplication of MLL2 and KDM6A genes in a large cohort of patients with Kabuki syndrome.," *Mol. Genet. Metab.*, vol. 107, no. 3, pp. 627–9, Nov. 2012.

[88] G. van Haaften, G. L. Dalgliesh, H. Davies, L. Chen, G. Bignell, C. Greenman, S. Edkins, C. Hardy, S. O'Meara, J. Teague, A. Butler, J. Hinton, C. Latimer, J. Andrews, S. Barthorpe, D. Beare, G. Buck, P. J. Campbell, J. Cole, S. Forbes, M. Jia, D. Jones, C. Y. Kok, C. Leroy, M.-L. Lin, D. J. McBride, M. Maddison, S. Maquire, K. McLay, A. Menzies, T. Mironenko, L. Mulderrig, L. Mudie, E. Pleasance, R. Shepherd, R. Smith, L. Stebbings, P. Stephens, G. Tang, P. S. Tarpey, R. Turner, K. Turrell, J. Varian, S. West, S. Widaa, P. Wray, V. P. Collins, K. Ichimura, S. Law, J. Wong, S. T. Yuen, S. Y. Leung, G. Tonon, R. A. DePinho, Y.-T. Tai, K. C. Anderson, R. J. Kahnoski, A. Massie, S. K. Khoo, B. T. Teh, M. R. Stratton, and P. A. Futreal, "Somatic mutations of the histone H3K27 demethylase gene UTX in human cancer.," *Nat. Genet.*, vol. 41, no. 5, pp. 521–3, May 2009.

[89] C. Kandoth, M. D. McLellan, F. Vandin, K. Ye, B. Niu, C. Lu, M. Xie, Q. Zhang, J. F. McMichael, M. A. Wyczalkowski, M. D. M. Leiserson, C. A. Miller, J. S. Welch, M. J. Walter, M. C. Wendl, T. J. Ley, R. K. Wilson, B. J. Raphael, and L. Ding, "Mutational landscape and significance across 12 major cancer types.," *Nature*, vol. 502, no. 7471, pp. 333–9, Oct. 2013.

[90] G. L. Dalgliesh, K. Furge, C. Greenman, L. Chen, G. Bignell, A. Butler, H. Davies, S. Edkins, C. Hardy, C. Latimer, J. Teague, J. Andrews, S. Barthorpe, D. Beare, G. Buck, P. J. Campbell, S. Forbes, M. Jia, D. Jones, H. Knott, C. Y. Kok, K. W. Lau, C. Leroy, M.-L. Lin, D. J. McBride, M. Maddison, S. Maguire, K. McLay, A. Menzies, T. Mironenko, L. Mulderrig, L. Mudie, S. O'Meara, E. Pleasance, A. Rajasingham, R. Shepherd, R. Smith, L. Stebbings, P. Stephens, G. Tang, P. S. Tarpey, K. Turrell, K. J. Dykema, S. K. Khoo, D. Petillo, B. Wondergem, J. Anema, R. J. Kahnoski, B. T. Teh, M. R. Stratton, and P. A. Futreal, "Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes.," *Nature*, vol. 463, no. 7279, pp. 360–3, Jan. 2010.

[91] Y. Gui, G. Guo, Y. Huang, X. Hu, A. Tang, S. Gao, R. Wu, C. Chen, X. Li, L. Zhou, M. He, Z. Li, X. Sun, W. Jia, J. Chen, S. Yang, F. Zhou, X. Zhao, S. Wan, R. Ye, C. Liang, Z. Liu, P. Huang, C. Liu, H. Jiang, Y. Wang, H. Zheng, L. Sun, X. Liu, Z. Jiang, D. Feng, J. Chen, S. Wu, J. Zou, Z. Zhang, R. Yang, J. Zhao, C. Xu, W. Yin, Z. Guan, J. Ye, H. Zhang, J. Li, K. Kristiansen, M. L. Nickerson, D. Theodorescu, Y. Li, X. Zhang, S. Li, J. Wang, H. Yang, J. Wang, and Z. Cai, "Frequent mutations of chromatin remodeling genes in transitional cell carcinoma of the bladder.," *Nat. Genet.*, vol. 43, no. 9, pp. 875–8, Sep. 2011.

[92] J. Van der Meulen, V. Sanghvi, K. Mavrakis, K. Durinck, F. Fang, F. Matthijssens, P. Rondou, M. Rosen, T. Pieters, P. Vandenberghe, E. Delabesse, T. Lammens, B. De Moerloose, B. Menten, N. Van Roy, B. Verhasselt, B. Poppe, Y. Benoit, T. Taghon, A. M. Melnick, F. Speleman, H.-G. Wendel, and P. Van Vlierberghe, "The H3K27me3 demethylase UTX is a gender-specific tumor suppressor in T-cell acute lymphoblastic leukemia.," *Blood*, vol. 125, no. 1, pp. 13–21, Jan. 2015.

[93] P. Ntziachristos, A. Tsirigos, G. G. Welstead, T. Trimarchi, S. Bakogianni, L. Xu, E. Loizou, L. Holmfeldt, A. Strikoudis, B. King, J. Mullenders, J. Becksfort, J. Nedjic, E. Paietta, M. S. Tallman, J. M. Rowe, G. Tonon, T. Satoh, L. Kruidenier, R. Prinjha, S. Akira, P. Van Vlierberghe, A. A. Ferrando, R. Jaenisch, C. G. Mullighan, and I. Aifantis, "Contrasting roles of histone 3 lysine 27 demethylases in acute lymphoblastic leukaemia.," *Nature*, vol. 514, no. 7523, pp. 513–7, Oct. 2014.

[94] H.-J. Choi, J.-H. Park, M. Park, H.-Y. Won, H.-S. Joo, C. H. Lee, J.-Y. Lee, and G. Kong, "UTX inhibits EMT-induced breast CSC properties by epigenetic repression of EMT genes in cooperation with LSD1 and HDAC1.," *EMBO Rep.*, p. e201540244, Aug. 2015.

[95] J.-H. Kim, A. Sharma, S. S. Dhar, S.-H. Lee, B. Gu, C.-H. Chan, H.-K. Lin, and M. G. Lee, "UTX and MLL4 coordinately regulate transcriptional programs for cell proliferation and invasiveness in breast cancer cells.," *Cancer Res.*, vol. 74, no. 6, pp. 1705–17, Mar. 2014.

[96] K. Fortschegger and R. Shiekhattar, "Plant homeodomain fingers form a helping hand for transcription.," *Epigenetics*, vol. 6, no. 1, pp. 4–8, Jan. 2011.

[97] C. Huang, Y. Xiang, Y. Wang, X. Li, L. Xu, Z. Zhu, T. Zhang, Q. Zhu, K. Zhang, N. Jing, and C. D. Chen, “Dual-specificity histone demethylase KIAA1718 (KDM7A) regulates neural differentiation through FGF4.,” *Cell Res.*, vol. 20, no. 2, pp. 154–65, Mar. 2010.

[98] J. D. Stender, G. Pascual, W. Liu, M. U. Kaikkonen, K. Do, N. J. Spann, M. Boutros, N. Perrimon, M. G. Rosenfeld, and C. K. Glass, “Control of proinflammatory gene programs by regulated trimethylation and demethylation of histone H4K20.,” *Mol. Cell*, vol. 48, no. 1, pp. 28–38, Oct. 2012.

[99] M. F. Arteaga, J.-H. Mikesch, J. Qiu, J. Christensen, K. Helin, S. C. Kogan, S. Dong, and C. W. E. So, “The histone demethylase PHF8 governs retinoic acid response in acute promyelocytic leukemia.,” *Cancer Cell*, vol. 23, no. 3, pp. 376–89, Mar. 2013.

[100] S. Sinha, R. K. Singh, N. Alam, A. Roy, S. Roychoudhury, and C. K. Panda, “Alterations in candidate genes PHF2, FANCC, PTCH1 and XPA at chromosomal 9q22.3 region: pathological significance in early- and late-onset breast carcinoma.,” *Mol. Cancer*, vol. 7, p. 84, Jan. 2008.

[101] F. Abidi, M. Miano, J. Murray, and C. Schwartz, “A novel mutation in the PHF8 gene is associated with X-linked mental retardation with cleft lip/cleft palate,” *Clin. Genet.*, vol. 72, no. 1, pp. 19–22, Jun. 2007.

[102] A. Koivisto, S. Ala-Mello, S. Lemmelä, H. Komu, J. Rautio, and I. Järvelä, “Screening of mutations in the PHF8 gene and identification of a novel mutation in a Finnish family with XLMR and cleft lip/cleft palate,” *Clin. Genet.*, vol. 72, no. 2, pp. 145–149, Aug. 2007.

[103] C. J. Schofield and P. J. Ratcliffe, “Oxygen sensing by HIF hydroxylases,” *Nat. Rev. Mol. Cell Biol.*, vol. 5, no. 5, pp. 343–354, May 2004.

[104] J. Aragonés, P. Fraisl, M. Baes, and P. Carmeliet, “Oxygen sensors at the crossroad of metabolism.,” *Cell Metab.*, vol. 9, no. 1, pp. 11–22, Jan. 2009.

[105] N. Zhang, Z. Fu, S. Linke, J. Chicher, J. J. Gorman, D. Visk, G. G. Haddad, L. Poellinger, D. J. Peet, F. Powell, and R. S. Johnson, “The asparaginyl hydroxylase factor inhibiting HIF-1alpha is an essential regulator of metabolism.,” *Cell Metab.*, vol. 11, no. 5, pp. 364–78, May 2010.

[106] M. Yang, W. Ge, R. Chowdhury, T. D. W. Claridge, H. B. Kramer, B. Schmierer, M. A. McDonough, L. Gong, B. M. Kessler, P. J. Ratcliffe, M. L. Coleman, and C. J. Schofield, “Asparagine and aspartate hydroxylation of the cytoskeletal ankyrin family is catalyzed by factor-inhibiting hypoxia-inducible factor.,” *J. Biol. Chem.*, vol. 286, no. 9, pp. 7648–60, Mar. 2011.

[107] S. Karttunen, M. Duffield, N. R. Scrimgeour, L. Squires, W. L. Lim, M. L. Dallas, J. L. Scragg, J. Chicher, K. A. Dave, M. L. Whitelaw, C. Peers, J. J. Gorman, J. M. Gleadle, G. Y. Rychkov, and D. J. Peet, “Oxygen-dependent hydroxylation by FIH regulates the TRPV3 ion channel.,” *J. Cell Sci.*, vol. 128, no. 2, pp. 225–31, Jan. 2015.

[108] S. Andersen, T. Donnem, H. Stenvold, S. Al-Saad, K. Al-Shibli, L.-T. Busund, and R. M. Bremnes, “Overexpression of the HIF hydroxylases PHD1, PHD2, PHD3 and FIH are individually and collectively unfavorable prognosticators for NSCLC survival.,” *PLoS One*, vol. 6, no. 8, p. e23847, Jan. 2011.

[109] S. G. C. Kroeze, J. S. Vermaat, A. van Brussel, H. H. E. van Melick, E. E. Voest, T. G. N. Jonges, P. J. van Diest, J. Hinrichs, J. L. H. R. Bosch, and J. J. M. Jans, “Expression of nuclear FIH independently predicts overall survival of clear cell renal cell carcinoma patients.,” *Eur. J. Cancer*, vol. 46, no. 18, pp. 3375–82, Dec. 2010.

[110] M. N. Khan, T. Bhattacharyya, P. Andrikopoulos, M. A. Esteban, R. Barod, T. Connor, M. Ashcroft, P. H. Maxwell, and S. Kiriakidis, “Factor inhibiting HIF (FIH-1) promotes renal cancer cell survival by protecting cells from HIF-1α-mediated apoptosis.,” *Br. J. Cancer*, vol. 104, no. 7, pp. 1151–9, Mar. 2011.

[111] Y. Liu, H. Higashitsuji, H. Higashitsuji, K. Itoh, T. Sakurai, K. Koike, K. Hirota, M. Fukumoto, and J. Fujita, “Overexpression of gankyrin in mouse hepatocytes induces hemangioma by suppressing factor inhibiting hypoxia-inducible factor-1 (FIH-1) and activating hypoxia-inducible factor-1.,” *Biochem. Biophys. Res. Commun.*, vol. 432, no. 1, pp. 22–7, Mar. 2013.

[112] D. a Hsia, C. G. Tepper, M. R. Pochampalli, E. Y. C. Hsia, C. Izumiya, S. B. Huerta, M. E. Wright, H.-W. Chen, H.-J. Kung, and Y. Izumiya, “KDM8, a H3K36me2 histone demethylase that acts in the cyclin A1 coding region to regulate cancer cell proliferation.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 21, pp. 9671–6, May 2010.

[113] M.-Y. Youn, A. Yokoyama, S. Fujiyama-Nakamura, F. Ohtake, K. Minehata, H. Yasuda, T. Suzuki, S. Kato, and Y. Imai, “JMJD5, a Jumonji C (JmjC) domain-containing protein, negatively regulates osteoclastogenesis by facilitating NFATc1 protein degradation.,” *J. Biol. Chem.*, vol. 287, no. 16, pp. 12994–3004, Apr. 2012.

[114] M. a Jones, M. F. Covington, L. DiTacchio, C. Vollmers, S. Panda, and S. L. Harmer, “Jumonji domain protein JMJD5 functions in both the plant and human circadian systems.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 50, pp. 21623–8, Dec. 2010.

[115] X. Huang, S. Zhang, H. Qi, Z. Wang, H.-W. Chen, J. Shao, and J. Shen, “JMJD5 interacts with p53 and negatively regulates p53 function in control of cell cycle and proliferation.,” *Biochim. Biophys. Acta*, May 2015.

[116] A. Ishimura, K. Minehata, M. Terashima, G. Kondoh, T. Hara, and T. Suzuki, “Jmjd5, an H3K36me2 histone demethylase, modulates embryonic cell proliferation through the regulation of Cdkn1a expression.,” *Development*, vol. 139, no. 4, pp. 749–59, Feb. 2012.

[117] S. Oh and R. Janknecht, “Histone demethylase JMJD5 is essential for embryonic development.,” *Biochem. Biophys. Res. Commun.*, vol. 420, no. 1, pp. 61–5, Mar. 2012.

[118] C. J. Webby, A. Wolf, N. Gromak, M. Dreger, H. Kramer, B. Kessler, M. L. Nielsen, C. Schmitz, D. S. Butler, J. R. Yates, C. M. Delahunty, P. Hahn, A. Lengeling, M. Mann, N. J. Proudfoot, C. J. Schofield, and A. Böttger, “Jmjd6

catalyses lysyl-hydroxylation of U2AF65, a protein associated with RNA splicing.," *Science*, vol. 325, no. 5936, pp. 90–3, Jul. 2009.

[119] M. Mantri, T. Krojer, E. A. Bagg, C. A. Webby, D. S. Butler, G. Kochan, K. L. Kavanagh, U. Oppermann, M. A. McDonough, and C. J. Schofield, "Crystal Structure of the 2-Oxoglutarate- and Fe(II)-Dependent Lysyl Hydroxylase JMJD6.," *J. Mol. Biol.*, vol. 401, no. 2, pp. 222–211, May 2010.

[120] X. Hong, J. Zang, J. White, C. Wang, C.-H. Pan, R. Zhao, R. C. Murphy, S. Dai, P. Henson, J. W. Kappler, J. Hagman, and G. Zhang, "Interaction of JMJD6 with single-stranded RNA.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 33, pp. 14568–72, Aug. 2010.

[121] Y. F. Lee, L. D. Miller, X. Bin Chan, M. a Black, B. Pang, C. W. Ong, M. Salto-Tellez, E. T. Liu, and K. V Desai, "JMJD6 is a driver of cellular proliferation and motility and a marker of poor prognosis in breast cancer.," *Breast Cancer Res.*, vol. 14, no. 3, p. R85, Jan. 2012.

[122] J. He, L. Shen, M. Wan, O. Taranova, H. Wu, and Y. Zhang, "Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes," *Nat. Cell Biol.*, vol. 15, no. 4, pp. 373–384, Mar. 2013.

[123] X. Ding, H. Pan, J. Li, Q. Zhong, X. Chen, S. M. Dry, and C.-Y. Wang, "Epigenetic Activation of AP1 Promotes Squamous Cell Carcinoma Metastasis," *Sci. Signal.*, vol. 6, no. 273, pp. ra28–ra28, Apr. 2013.

[124] Z.-Q. Xi, J.-J. Sun, X.-F. Wang, M.-W. Li, X.-Z. Liu, L.-Y. Wang, X. Zhu, F. Xiao, J.-M. Li, Y. Gong, and L.-F. Guan, "HSPBAP1 is found extensively in the anterior temporal neocortex of patients with intractable epilepsy.," *Synapse*, vol. 61, no. 9, pp. 741–7, Sep. 2007.

[125] A. Noma, R. Ishitani, M. Kato, A. Nagao, O. Nureki, and T. Suzuki, "Expanding role of the jumonji C domain as an RNA hydroxylase.," *J. Biol. Chem.*, vol. 285, no. 45, pp. 34503–7, Dec. 2010.

[126] J. Eilbracht, S. Kneissel, A. Hofmann, and M. S. Schmidt-Zachmann, "Protein NO52--a constitutive nucleolar component sharing high sequence homologies to protein NO66.," *Eur. J. Cell Biol.*, vol. 84, no. 2–3, pp. 279–94, Mar. 2005.

[127] M. Tsuneoka, Y. Koda, M. Soejima, K. Teye, and H. Kimura, "A novel myc target gene, mina53, that is involved in cell proliferation.," *J. Biol. Chem.*, vol. 277, no. 38, pp. 35450–9, Sep. 2002.

[128] H. Ishizaki, H. Yano, M. Tsuneoka, S. Ogasawara, J. Akiba, N. Nishida, S. Kojiro, S. Fukahori, F. Moriya, K. Matsuoka, and M. Kojiro, "Overexpression of the myc target gene Mina53 in advanced renal cell carcinoma.," *Pathol. Int.*, vol. 57, no. 10, pp. 672–80, Oct. 2007.

[129] C. Suzuki, K. Takahashi, S. Hayama, N. Ishikawa, T. Kato, T. Ito, E. Tsuchiya, Y. Nakamura, and Y. Daigo, "Identification of Myc-associated protein with JmjC domain as a novel therapeutic target oncogene for lung cancer.," *Mol. Cancer Ther.*, vol. 6, no. 2, pp. 542–51, Feb. 2007.

[130] X.-P. Tan, Q. Zhang, W.-G. Dong, X.-W. Lei, and Z.-R. Yang, “Upregulated expression of Mina53 in cholangiocarcinoma and its clinical significance.,” *Oncol. Lett.*, vol. 3, no. 5, pp. 1037–1041, May 2012.

[131] K. Teye, M. Tsuneoka, N. Arima, Y. Koda, Y. Nakamura, Y. Ueta, K. Shirouzu, and H. Kimura, “Increased expression of a Myc target gene Mina53 in human colon cancer.,” *Am. J. Pathol.*, vol. 164, no. 1, pp. 205–16, Jan. 2004.

[132] Y. Zhang, Y. Lu, B.-Z. Yuan, V. Castranova, X. Shi, J. L. Stauffer, L. M. Demers, and F. Chen, “The Human mineral dust-induced gene, mdig, is a cell growth regulating gene associated with lung cancer.,” *Oncogene*, vol. 24, no. 31, pp. 4873–82, Jul. 2005.

[133] T. Mori, K. Okamoto, Y. Tanaka, K. Teye, T. Umata, K. Ohneda, K. Tokuyama, M. Okabe, and M. Tsuneoka, “Ablation of Mina53 in mice reduces allergic response in the airways,” *Cell Struct. Funct.*, Jun. 2013.

[134] N. Yosef, A. K. Shalek, J. T. Gaublomme, H. Jin, Y. Lee, A. Awasthi, C. Wu, K. Karwacz, S. Xiao, M. Jorgolli, D. Gennert, R. Satija, A. Shakya, D. Y. Lu, J. J. Trombetta, M. R. Pillai, P. J. Ratcliffe, M. L. Coleman, M. Bix, D. Tantin, H. Park, V. K. Kuchroo, and A. Regev, “Dynamic regulatory network controlling TH17 cell differentiation,” *Nature*, vol. 496, no. 7446, pp. 461–468, Mar. 2013.

[135] Y. Lu, Q. Chang, Y. Zhang, K. Beezhold, Y. Rojanasakul, H. Zhao, V. Castranova, X. Shi, and F. Chen, “Lung cancer-associated JmjC domain protein mdig suppresses formation of tri-methyl lysine 9 of histone H3.,” *Cell Cycle*, vol. 8, no. 13, pp. 2101–9, Jul. 2009.

[136] C. Wang, Q. Zhang, T. Hang, Y. Tao, X. Ma, M. Wu, X. Zhang, and J. Zang, “Structure of the JmjC domain-containing protein NO66 complexed with ribosomal protein Rpl8,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 71, no. 9, pp. 1955–1964, Aug. 2015.

[137] R. Chowdhury, R. Sekirnik, N. C. Brissett, T. Krojer, C.-H. Ho, S. S. Ng, I. J. Clifton, W. Ge, N. J. Kershaw, G. C. Fox, J. R. C. Muniz, M. Vollmar, C. Phillips, E. S. Pilka, K. L. Kavanagh, F. von Delft, U. Oppermann, M. A. McDonough, A. J. Doherty, and C. J. Schofield, “Ribosomal oxygenases are structurally conserved from prokaryotes to humans.,” *Nature*, vol. 509, no. 7505, pp. 422–6, Jun. 2014.

[138] Q. Chen, L. Zhang, B. de Crombrugghe, and R. Krahe, “Mesenchyme-specific overexpression of nucleolar protein 66 in mice inhibits skeletal growth and bone formation.,” *FASEB J.*, vol. 29, no. 6, pp. 2555–65, Jun. 2015.

[139] Q. Chen, K. Sinha, J. M. Deng, H. Yasuda, R. Krahe, R. R. Behringer, and B. de Crombrugghe, “Mesenchymal Deletion of Histone Demethylase NO66 in Mice Promotes Bone Formation.,” *J. Bone Miner. Res.*, vol. 30, no. 9, pp. 1608–17, Sep. 2015.

[140] I. J. Clifton, M. A. McDonough, D. Ehrismann, N. J. Kershaw, N. Granatino, and C. J. Schofield, “Structural studies on 2-oxoglutarate oxygenases and related double-stranded beta-helix fold proteins.,” *J. Inorg. Biochem.*, vol. 100, no. 4, pp. 644–69, Apr. 2006.

[141] M. A. McDonough, C. Loenarz, R. Chowdhury, I. J. Clifton, and C. J. Schofield, “Structural studies on human 2-oxoglutarate dependent oxygenases,” *Curr. Opin. Struct. Biol.*, vol. 20, no. 6, pp. 659–672, Dec. 2010.

[142] Y. Tao, M. Wu, X. Zhou, W. Yin, B. Hu, B. de Crombrugghe, K. M. Sinha, and J. Zang, “Structural Insights into Histone Demethylase NO66 in Interaction with Osteoblast-specific Transcription Factor Osterix and Gene Repression.,” *J. Biol. Chem.*, vol. 288, no. 23, pp. 16430–7, Jun. 2013.

[143] W. W. Yue, V. Hozjan, W. Ge, C. Loenarz, C. D. O. Cooper, C. J. Schofield, K. L. Kavanagh, U. Oppermann, and M. A. McDonough, “Crystal structure of the PHF8 Jumonji domain, an Nepsilon-methyl lysine demethylase.,” *FEBS Lett.*, vol. 584, no. 4, pp. 825–30, Feb. 2010.

[144] S. S. Ng, K. L. Kavanagh, M. A. McDonough, D. Butler, E. S. Pilka, B. M. R. Lienard, J. E. Bray, P. Savitsky, O. Gileadi, F. von Delft, N. R. Rose, J. Offer, J. C. Scheinost, T. Borowski, M. Sundstrom, C. J. Schofield, and U. Oppermann, “Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity.,” *Nature*, vol. 448, no. 7149, pp. 87–91, Jul. 2007.

[145] K. S. Hewitson, N. Granatino, R. W. D. Welford, M. A. McDonough, and C. J. Schofield, “Oxidation by 2-oxoglutarate oxygenases: non-haem iron systems in catalysis and signalling.,” *Philos. Trans. A. Math. Phys. Eng. Sci.*, vol. 363, no. 1829, pp. 807–28; discussion 1035–40, Apr. 2005.

[146] C. C. Thinnes, K. S. England, A. Kawamura, R. Chowdhury, C. J. Schofield, and R. J. Hopkinson, “Targeting histone lysine demethylases - Progress, challenges, and the future.,” *Biochim. Biophys. Acta*, May 2014.

[147] C. Loenarz and C. J. Schofield, “Physiological and biochemical aspects of hydroxylations and demethylations catalyzed by human 2-oxoglutarate oxygenases.,” *Trends Biochem. Sci.*, vol. 36, no. 1, pp. 7–18, Jan. 2011.

[148] N. R. Rose, M. A. McDonough, O. N. F. King, A. Kawamura, and C. J. Schofield, “Inhibition of 2-oxoglutarate dependent oxygenases.,” *Chem. Soc. Rev.*, vol. 40, no. 8, pp. 4364–97, Aug. 2011.

[149] E. C. Y. Woon, A. Tumber, A. Kawamura, L. Hillringhaus, W. Ge, N. R. Rose, J. H. Y. Ma, M. C. Chan, L. J. Walport, K. H. Che, S. S. Ng, B. D. Marsden, U. Oppermann, M. A. McDonough, and C. J. Schofield, “Linking of 2-oxoglutarate and substrate binding sites enables potent and highly selective inhibition of JmjC histone demethylases.,” *Angew. Chem. Int. Ed. Engl.*, vol. 51, no. 7, pp. 1631–4, Feb. 2012.

[150] R. Sekirnik, N. R. Rose, A. Thalhammer, P. T. Seden, J. Mecinović, and C. J. Schofield, “Inhibition of the histone lysine demethylase JMJD2A by ejection of structural Zn(II).,” *Chem. Commun. (Camb).*, no. 42, pp. 6376–8, Nov. 2009.

[151] O. N. F. King, X. S. Li, M. Sakurai, A. Kawamura, N. R. Rose, S. S. Ng, A. M. Quinn, G. Rai, B. T. Mott, P. Beswick, R. J. Klose, U. Oppermann, A. Jadhav, T. D. Heightman, D. J. Maloney, C. J. Schofield, and A. Simeonov, “Quantitative high-throughput screening identifies 8-hydroxyquinolines as cell-active histone demethylase inhibitors.,” *PLoS One*, vol. 5, no. 11, p. e15535, Jan. 2010.

[152] K. Salnikow, S. P. Donald, R. K. Bruick, A. Zhitkovich, J. M. Phang, and K. S. Kasprzak, “Depletion of intracellular ascorbate by the carcinogenic metals nickel and cobalt results in the induction of hypoxic stress.,” *J. Biol. Chem.*, vol. 279, no. 39, pp. 40337–44, Sep. 2004.

[153] A. Ploumakis and M. L. Coleman, “OH, the Places You’ll Go! Hydroxylation, Gene Expression, and Cancer.,” *Mol. Cell*, vol. 58, no. 5, pp. 729–741, Jun. 2015.

[154] R. A. Cairns and T. W. Mak, “Oncogenic isocitrate dehydrogenase mutations: mechanisms, models, and clinical opportunities.,” *Cancer Discov.*, vol. 3, no. 7, pp. 730–41, Jul. 2013.

[155] R. Chowdhury, K. K. Yeoh, Y.-M. Tian, L. Hillringhaus, E. A. Bagg, N. R. Rose, I. K. H. Leung, X. S. Li, E. C. Y. Woon, M. Yang, M. A. McDonough, O. N. King, I. J. Clifton, R. J. Klose, T. D. W. Claridge, P. J. Ratcliffe, C. J. Schofield, and A. Kawamura, “The oncometabolite 2-hydroxyglutarate inhibits histone lysine demethylases.,” *EMBO Rep.*, vol. 12, no. 5, pp. 463–9, May 2011.

[156] W. Xu, H. Yang, Y. Liu, Y. Yang, P. Wang, S.-H. Kim, S. Ito, C. Yang, P. Wang, M.-T. Xiao, L. Liu, W. Jiang, J. Liu, J. Zhang, B. Wang, S. Frye, Y. Zhang, Y. Xu, Q. Lei, K.-L. Guan, S. Zhao, and Y. Xiong, “Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases.,” *Cancer Cell*, vol. 19, no. 1, pp. 17–30, Jan. 2011.

[157] A. M. Intlekofer, R. G. Dematteo, S. Venneti, L. W. S. Finley, C. Lu, A. R. Judkins, A. S. Rustenburg, P. B. Grinaway, J. D. Chodera, J. R. Cross, and C. B. Thompson, “Hypoxia Induces Production of L-2-Hydroxyglutarate.,” *Cell Metab.*, vol. 22, no. 2, pp. 304–11, Aug. 2015.

[158] W. M. Oldham, C. B. Clish, Y. Yang, and J. Loscalzo, “Hypoxia-Mediated Increases in l-2-hydroxyglutarate Coordinate the Metabolic Response to Reductive Stress.,” *Cell Metab.*, vol. 22, no. 2, pp. 291–303, Aug. 2015.

[159] A. L. Harris, “A New Hydroxy Metabolite of 2-Oxoglutarate Regulates Metabolism in Hypoxia.,” *Cell Metab.*, vol. 22, no. 2, pp. 198–200, Aug. 2015.

[160] N. R. Rose, S. S. Ng, J. Mecinović, B. M. R. Liénard, S. H. Bello, Z. Sun, M. A. McDonough, U. Oppermann, and C. J. Schofield, “Inhibitor scaffolds for 2-oxoglutarate-dependent histone lysine demethylases.,” *J. Med. Chem.*, vol. 51, no. 22, pp. 7053–6, Nov. 2008.

[161] R. J. Hopkinson, A. Tumber, C. Yapp, R. Chowdhury, W. Aik, K. H. Che, X. S. Li, J. B. L. Kristensen, O. N. F. King, M. C. Chan, K. K. Yeoh, H. Choi, L. J. Walport, C. C. Thinnes, J. T. Bush, C. Lejeune, A. M. Rydzik, N. R. Rose, E. A. Bagg, M. A. McDonough, T. J. Krojer, W. W. Yue, S. S. Ng, L. Olsen, P. E. Brennan, U. Oppermann, S. Müller, R. J. Klose, P. J. Ratcliffe, C. J. Schofield, and A. Kawamura, “5-Carboxy-8-hydroxyquinoline is a broad spectrum 2-oxoglutarate oxygenase inhibitor which causes iron translocation,” *Chem. Sci.*, vol. 4, no. 8, p. 3110, Jul. 2013.

[162] X. Luo, Y. Liu, S. Kubicek, J. Myllyharju, A. Tumber, S. Ng, K. H. Che, J. Podoll, T. D. Heightman, U. Oppermann, S. L. Schreiber, and X. Wang, “A selective inhibitor and probe of the cellular functions of Jumonji C domain-containing

histone demethylases.," *J. Am. Chem. Soc.*, vol. 133, no. 24, pp. 9451–6, Jun. 2011.

[163] S. Hamada, T. Suzuki, K. Mino, K. Koseki, F. Oehme, I. Flamme, H. Ozasa, Y. Itoh, D. Ogasawara, H. Komaarashi, A. Kato, H. Tsumoto, H. Nakagawa, M. Hasegawa, R. Sasaki, T. Mizukami, and N. Miyata, "Design, synthesis, enzyme-inhibitory activity, and effect on human cancer cells of a novel series of jumonji domain-containing protein 2 histone demethylase inhibitors.," *J. Med. Chem.*, vol. 53, no. 15, pp. 5629–38, Aug. 2010.

[164] R. Schiller, G. Scozzafava, A. Tumber, J. R. Wickens, J. T. Bush, G. Rai, C. Lejeune, H. Choi, T.-L. Yeh, M. C. Chan, B. T. Mott, J. S. O. McCullagh, D. J. Maloney, C. J. Schofield, and A. Kawamura, "A cell-permeable ester derivative of the JmjC histone demethylase inhibitor IOX1.," *ChemMedChem*, vol. 9, no. 3, pp. 566–71, Mar. 2014.

[165] G. Rai, A. Kawamura, A. Tumber, Y. Liang, J. L. Vogel, J. H. Arbuckle, N. R. Rose, T. S. Dexheimer, T. L. Foley, O. N. King, A. Quinn, B. T. Mott, C. J. Schofield, U. Oppermann, A. Jadhav, A. Simeonov, T. M. Kristie, and D. J. Maloney, "Discovery of ML324, a JMJD2 demethylase inhibitor with demonstrated antiviral activity." National Center for Biotechnology Information (US), 16-Sep-2013.

[166] C. Jin, L. Yang, M. Xie, C. Lin, D. Merkurjev, J. C. Yang, B. Tanasa, S. Oh, J. Zhang, K. A. Ohgi, H. Zhou, W. Li, C. P. Evans, S. Ding, and M. G. Rosenfeld, "Chem-seq permits identification of genomic targets of drugs against androgen receptor regulation selected by functional phenotypic screens.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 25, pp. 9235–40, Jun. 2014.

[167] L. Anders, M. G. Guenther, J. Qi, Z. P. Fan, J. J. Marineau, P. B. Rahl, J. Lovén, A. A. Sigova, W. B. Smith, T. I. Lee, J. E. Bradner, and R. A. Young, "Genome-wide localization of small molecules.," *Nat. Biotechnol.*, vol. 32, no. 1, pp. 92–6, Jan. 2014.

[168] R. Hashizume, N. Andor, Y. Ihara, R. Lerner, H. Gan, X. Chen, D. Fang, X. Huang, M. W. Tom, V. Ngo, D. Solomon, S. Mueller, P. L. Paris, Z. Zhang, C. Petritsch, N. Gupta, T. A. Waldman, and C. D. James, "Pharmacologic inhibition of histone demethylation as a therapy for pediatric brainstem glioma.," *Nat. Med.*, vol. 20, no. 12, pp. 1394–1396, Nov. 2014.

[169] B. Heinemann, J. M. Nielsen, H. R. Hudlebusch, M. J. Lees, D. V. Larsen, T. Boesen, M. Labelle, L.-O. Gerlach, P. Birk, and K. Helin, "Inhibition of demethylases by GSK-J1/J4.," *Nature*, vol. 514, no. 7520, pp. E1–2, Oct. 2014.

[170] M. Labelle, T. Boesen, M. Mehrotra, Q. Khan, and F. ULLAH, "Inhibitors of histone demethylases." Google Patents, 2014.

[171] K. S. England, A. Tumber, T. Krojer, G. Scozzafava, S. S. Ng, M. Daniel, A. Szykowska, K. Che, F. von Delft, N. A. Burgess-Brown, A. Kawamura, C. J. Schofield, and P. E. Brennan, "Optimisation of a triazolopyridine based histone demethylase inhibitor yields a potent and selective KDM2A (FBXL11) inhibitor," *Med. Chem. Commun.*, vol. 5, no. 12, pp. 1879–1886, Nov. 2014.

[172] L. Wang, J. Chang, D. Varghese, M. Dellinger, S. Kumar, A. M. Best, J. Ruiz, R. Bruick, S. Peña-Llopis, J. Xu, D. J. Babinski, D. E. Frantz, R. A. Brekken, A. M. Quinn, A. Simeonov, J. Easmon, and E. D. Martinez, “A small molecule modulates Jumonji histone demethylase activity and selectively inhibits cancer growth.,” *Nat. Commun.*, vol. 4, p. 2035, Jan. 2013.

[173] J. Sayegh, J. Cao, M. R. Zou, A. Morales, L. P. Blair, M. Norcia, D. Hoyer, A. J. Tackett, J. S. Merkel, and Q. Yan, “Identification of small molecule inhibitors of Jumonji AT-rich interactive domain 1B (JARID1B) histone demethylase by a sensitive high throughput screen.,” *J. Biol. Chem.*, vol. 288, no. 13, pp. 9408–17, Mar. 2013.

[174] J. Baell and M. A. Walters, “Chemistry: Chemical con artists foil drug discovery.,” *Nature*, vol. 513, no. 7519, pp. 481–3, Sep. 2014.

[175] J. B. Baell and G. A. Holloway, “New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays.,” *J. Med. Chem.*, vol. 53, no. 7, pp. 2719–40, Apr. 2010.

[176] S. Kubicek, R. J. O’Sullivan, E. M. August, E. R. Hickey, Q. Zhang, M. L. Teodoro, S. Rea, K. Mechtler, J. A. Kowalski, C. A. Homon, T. A. Kelly, and T. Jenuwein, “Reversal of H3K9me2 by a small-molecule inhibitor for the G9a histone methyltransferase.,” *Mol. Cell*, vol. 25, no. 3, pp. 473–81, Feb. 2007.

[177] A. K. Upadhyay, D. Rotili, J. W. Han, R. Hu, Y. Chang, D. Labella, X. Zhang, Y.-S. Yoon, A. Mai, and X. Cheng, “An analog of BIX-01294 selectively inhibits a family of histone H3 lysine 9 Jumonji demethylases.,” *J. Mol. Biol.*, vol. 416, no. 3, pp. 319–27, Feb. 2012.

[178] H. M. Berman, “The Protein Data Bank,” *Nucleic Acids Res.*, vol. 28, no. 1, pp. 235–242, Jan. 2000.

[179] C. W. Murray and T. L. Blundell, “Structural biology in fragment-based drug design.,” *Curr. Opin. Struct. Biol.*, vol. 20, no. 4, pp. 497–507, Aug. 2010.

[180] R. J. Hall, P. N. Mortenson, and C. W. Murray, “Efficient exploration of chemical space by fragment-based screening.,” *Prog. Biophys. Mol. Biol.*, vol. 116, no. 2–3, pp. 82–91, Jan. 2014.

[181] D. Kozakov, D. R. Hall, S. Jehle, L. Luo, S. O. Ochiana, E. V Jones, M. Pollastri, K. N. Allen, A. Whitty, and S. Vajda, “Ligand deconstruction: Why some fragment binding positions are conserved and others are not.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 112, no. 20, pp. E2585–2594, Apr. 2015.

[182] J. Tsai, J. T. Lee, W. Wang, J. Zhang, H. Cho, S. Mamo, R. Bremer, S. Gillette, J. Kong, N. K. Haass, K. Sproesser, L. Li, K. S. M. Smalley, D. Fong, Y.-L. Zhu, A. Marimuthu, H. Nguyen, B. Lam, J. Liu, I. Cheung, J. Rice, Y. Suzuki, C. Luu, C. Settachatgul, R. Shellooe, J. Cantwell, S.-H. Kim, J. Schlessinger, K. Y. J. Zhang, B. L. West, B. Powell, G. Habets, C. Zhang, P. N. Ibrahim, P. Hirth, D. R. Artis, M. Herlyn, and G. Bollag, “Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 105, no. 8, pp. 3041–6, Mar. 2008.

[183] G. Bollag, P. Hirth, J. Tsai, J. Zhang, P. N. Ibrahim, H. Cho, W. Spevak, C. Zhang, Y. Zhang, G. Habets, E. A. Burton, B. Wong, G. Tsang, B. L. West, B. Powell, R. Shellooe, A. Marimuthu, H. Nguyen, K. Y. J. Zhang, D. R. Artis, J. Schlessinger, F. Su, B. Higgins, R. Iyer, K. D'Andrea, A. Koehler, M. Stumm, P. S. Lin, R. J. Lee, J. Grippo, I. Puzanov, K. B. Kim, A. Ribas, G. A. McArthur, J. A. Sosman, P. B. Chapman, K. T. Flaherty, X. Xu, K. L. Nathanson, and K. Nolop, "Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma.," *Nature*, vol. 467, no. 7315, pp. 596–9, Sep. 2010.

[184] S. B. Shuker, P. J. Hajduk, R. P. Meadows, and S. W. Fesik, "Discovering High-Affinity Ligands for Proteins: SAR by NMR," *Science (80-. ).*, vol. 274, no. 5292, pp. 1531–1534, Nov. 1996.

[185] Z. Chilingaryan, Z. Yin, and A. J. Oakley, "Fragment-based screening by protein crystallography: successes and pitfalls.," *Int. J. Mol. Sci.*, vol. 13, no. 10, pp. 12857–79, Jan. 2012.

[186] B. J. Davis and D. A. Erlanson, "Learning from our mistakes: the 'unknown knowns' in fragment screening.," *Bioorg. Med. Chem. Lett.*, vol. 23, no. 10, pp. 2844–52, May 2013.

[187] D. Patel, J. D. Bauman, and E. Arnold, "Advantages of crystallographic fragment screening: functional and mechanistic insights from a powerful platform for efficient drug discovery.," *Prog. Biophys. Mol. Biol.*, vol. 116, no. 2–3, pp. 92–100, Jan. 2014.

[188] A. D. Morley, A. Pugliese, K. Birchall, J. Bower, P. Brennan, N. Brown, T. Chapman, M. Drysdale, I. H. Gilbert, S. Hoelder, A. Jordan, S. V Ley, A. Merritt, D. Miller, M. E. Swarbrick, and P. G. Wyatt, "Fragment-based hit identification: thinking in 3D.," *Drug Discov. Today*, vol. 18, no. 23–24, pp. 1221–7, Dec. 2013.

[189] E. Zartler, C. Swain, and S. Pearce, "Fragment library design the evolution of fragment based lead discovery.," *Drug Discov. World*, no. Winter 2012/13.

[190] M. Congreve, R. Carr, C. Murray, and H. Jhoti, "A 'Rule of Three' for fragment-based lead discovery?," *Drug Discov. Today*, vol. 8, no. 19, pp. 876–877, Oct. 2003.

[191] A. Kawamura, A. Tumber, N. R. Rose, O. N. F. King, M. Daniel, U. Oppermann, T. D. Heightman, and C. Schofield, "Development of homogeneous luminescence assays for histone demethylase catalysis and binding.," *Anal. Biochem.*, vol. 404, no. 1, pp. 86–93, Sep. 2010.

[192] K. S. Hewitson, B. M. R. Liénard, M. A. McDonough, I. J. Clifton, D. Butler, A. S. Soares, N. J. Oldham, L. A. McNeill, and C. J. Schofield, "Structural and mechanistic studies on the inhibition of the hypoxia-inducible transcription factor hydroxylases by tricarboxylic acid cycle intermediates.," *J. Biol. Chem.*, vol. 282, no. 5, pp. 3293–301, Feb. 2007.

[193] J. Strelow, W. Dewe, P. W. Iversen, H. B. Brooks, J. A. Radding, J. McGee, and J. Weidner, "Mechanism of Action assays for Enzymes." Eli Lilly & Company and the National Center for Advancing Translational Sciences, 01-Oct-2012.

[194] Y. Zhang and J. Skolnick, “TM-align: a protein structure alignment algorithm based on the TM-score.,” *Nucleic Acids Res.*, vol. 33, no. 7, pp. 2302–9, Jan. 2005.

[195] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, and J. A. McCammon, “Electrostatics of nanosystems: application to microtubules and the ribosome.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 98, no. 18, pp. 10037–41, Aug. 2001.

[196] A. M. Anger, J.-P. Armache, O. Berninghausen, M. Habeck, M. Subklewe, D. N. Wilson, and R. Beckmann, “Structures of the human and Drosophila 80S ribosome.,” *Nature*, vol. 497, no. 7447, pp. 80–5, May 2013.

[197] M. Yang, A. P. Hardy, R. Chowdhury, N. D. Loik, J. S. Scotti, J. S. O. McCullagh, T. D. W. Claridge, M. A. McDonough, W. Ge, and C. J. Schofield, “Substrate selectivity analyses of factor inhibiting hypoxia-inducible factor.,” *Angew. Chem. Int. Ed. Engl.*, vol. 52, no. 6, pp. 1700–4, Feb. 2013.

[198] T. Keates, C. D. O. Cooper, P. Savitsky, C. K. Allerston, C. Phillips, M. Hammarström, N. Daga, G. Berridge, P. Mahajan, N. A. Burgess-Brown, S. Müller, S. Gräslund, and O. Gileadi, “Expressing the human proteome for affinity proteomics: optimising expression of soluble protein domains and in vivo biotinylation.,” *N. Biotechnol.*, vol. 29, no. 5, pp. 515–25, Jun. 2012.

[199] F. H. Niesen, H. Berglund, and M. Vedadi, “The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability.,” *Nat. Protoc.*, vol. 2, no. 9, pp. 2212–21, Jan. 2007.

[200] O. Gileadi, S. Knapp, W. H. Lee, B. D. Marsden, S. Müller, F. H. Niesen, K. L. Kavanagh, L. J. Ball, F. von Delft, D. A. Doyle, U. C. T. Oppermann, and M. Sundström, “The scientific impact of the Structural Genomics Consortium: a protein family and ligand-centered approach to medically-relevant human proteins.,” *J. Struct. Funct. Genomics*, vol. 8, no. 2–3, pp. 107–19, Sep. 2007.

[201] A. Mai, S. Massa, D. Rotili, S. Simeoni, R. Ragno, G. Botta, A. Nebbioso, M. Miceli, L. Altucci, and G. Brosch, “Synthesis and biological properties of novel, uracil-containing histone deacetylase inhibitors.,” *J. Med. Chem.*, vol. 49, no. 20, pp. 6046–56, Oct. 2006.

[202] Y. Abdiche, D. Malashock, A. Pinkerton, and J. Pons, “Determining kinetics and affinities of protein interactions using a parallel real-time label-free biosensor, the Octet.,” *Anal. Biochem.*, vol. 377, no. 2, pp. 209–17, Jun. 2008.

[203] M. Sakurai, N. R. Rose, L. Schultz, A. M. Quinn, A. Jadhav, S. S. Ng, U. Oppermann, C. J. Schofield, and A. Simeonov, “A miniaturized screen for inhibitors of Jumonji histone demethylases.,” *Mol. Biosyst.*, vol. 6, no. 2, pp. 357–64, Feb. 2010.

[204] S. E. Hutchinson, M. V Leveridge, M. L. Heathcote, P. Francis, L. Williams, M. Gee, J. Munoz-Muriedas, B. Leavens, A. Shillings, E. Jones, P. Homes, S. Baddeley, C.-W. Chung, A. Bridges, and A. Argyrou, “Enabling lead discovery for histone lysine demethylases by high-throughput RapidFire mass spectrometry.,” *J. Biomol. Screen.*, vol. 17, no. 1, pp. 39–48, Jan. 2012.

[205] M. Plant, T. Dineen, A. Cheng, A. M. Long, H. Chen, and K. a Morgenstern, “Screening for lysine-specific demethylase-1 inhibitors using a label-free high-throughput mass spectrometry assay.,” *Anal. Biochem.*, vol. 419, no. 2, pp. 217–27, Dec. 2011.

[206] H. B. Brooks, S. Geeganage, S. D. Kahl, C. Montrose, S. Sittampalam, M. C. Smith, and J. R. Weidner, “Basics of Enzymatic Assays for HTS.” Eli Lilly & Company and the National Center for Advancing Translational Sciences, 01-Oct-2012.

[207] A. McPherson and B. Cudney, “Searching for silver bullets: An alternative strategy for crystallizing macromolecules,” *J. Struct. Biol.*, vol. 156, no. 3, pp. 387–406, Dec. 2006.

[208] A. D’Arcy, F. Villard, and M. Marsh, “An automated microseed matrix-screening method for protein crystallization.,” *Acta Crystallogr. D. Biol. Crystallogr.*, vol. 63, no. Pt 4, pp. 550–4, Apr. 2007.

[209] G. C. Ireton and B. L. Stoddard, “Microseed matrix screening to improve crystals of yeast cytosine deaminase,” *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 60, no. 3, pp. 601–605, Feb. 2004.

[210] X. Zhou, Y. Tao, M. Wu, D. Zhang, and J. Zang, “Purification, crystallization and preliminary crystallographic analysis of histone lysine demethylase NO66 from Homo sapiens.,” *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.*, vol. 68, no. Pt 7, pp. 764–6, Jul. 2012.

[211] V. Naumov, W. N. Price, S. K. Handelman, and J. F. Hunt, “Engineering surface epitopes to improve protein crystallization,” US20130123467 A1, 16-May-2013.

[212] W. N. Price, Y. Chen, S. K. Handelman, H. Neely, P. Manor, R. Karlin, R. Nair, J. Liu, M. Baran, J. Everett, S. N. Tong, F. Forouhar, S. S. Swaminathan, T. Acton, R. Xiao, J. R. Luft, A. Lauricella, G. T. DeTitta, B. Rost, G. T. Montelione, and J. F. Hunt, “Understanding the physical properties that control protein crystallization by analysis of large-scale experimental data.,” *Nat. Biotechnol.*, vol. 27, no. 1, pp. 51–7, Jan. 2009.

[213] Z. S. Derewenda, “Rational protein crystallization by mutational surface engineering.,” *Structure*, vol. 12, no. 4, pp. 529–35, Apr. 2004.

[214] T. Suzuki, H. Ozasa, Y. Itoh, P. Zhan, H. Sawada, K. Mino, L. Walport, R. Ohkubo, A. Kawamura, M. Yonezawa, Y. Tsukada, A. Tumber, H. Nakagawa, M. Hasegawa, R. Sasaki, T. Mizukami, C. J. Schofield, and N. Miyata, “Identification of the KDM2/7 Histone Lysine Demethylase Subfamily Inhibitor and its Antiproliferative Activity,” *J. Med. Chem.*, p. 130821120910007, Aug. 2013.

[215] N. R. Rose, E. C. Y. Woon, A. Tumber, L. J. Walport, R. Chowdhury, X. S. Li, O. N. F. King, C. Lejeune, S. S. Ng, T. Krojer, M. C. Chan, A. M. Rydzik, R. J. Hopkinson, K. H. Che, M. Daniel, C. Strain-Damerell, C. Gileadi, G. Kochan, I. K. H. Leung, J. Dunford, K. K. Yeoh, P. J. Ratcliffe, N. Burgess-Brown, F. von Delft, S. Muller, B. Marsden, P. E. Brennan, M. A. McDonough, U. Oppermann, R. J. Klose, C. J. Schofield, and A. Kawamura, “Plant growth regulator daminozide is a selective inhibitor of human KDM2/7 histone demethylases.,” *J. Med. Chem.*, vol. 55, no. 14, pp. 6639–43, Jul. 2012.

[216] R. Copeland, *Enzymes*, Second Edi. John Wiley & Sons, Inc., 2000.

[217] M. Dokmanovic, C. Clarke, and P. A. Marks, “Histone deacetylase inhibitors: overview and perspectives.,” *Mol. Cancer Res.*, vol. 5, no. 10, pp. 981–9, Oct. 2007.

[218] R. Abagyan, M. Totrov, and D. Kuznetsov, “ICM. A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation,” *J. Comput. Chem.*, vol. 15, no. 5, pp. 488–506, May 1994.

[219] R. B. van Breemen and Y. Li, “Caco-2 cell permeability assays to measure drug absorption.,” *Expert Opin. Drug Metab. Toxicol.*, vol. 1, no. 2, pp. 175–85, Aug. 2005.

[220] C. Thakur and F. Chen, “Current understanding of mdig/MINA in human cancers,” *Genes Cancer*.

[221] M. Yu, J. Sun, C. Thakur, B. Chen, Y. Lu, H. Zhao, and F. Chen, “Paradoxical roles of mineral dust induced gene on cell proliferation and migration/invasion.,” *PLoS One*, vol. 9, no. 2, p. e87998, Jan. 2014.

[222] L. Kruidenier, C. Chung, Z. Cheng, J. Liddle, K. Che, G. Joberty, M. Bantscheff, C. Bountra, A. Bridges, H. Diallo, D. Eberhard, S. Hutchinson, E. Jones, R. Katso, M. Leveridge, P. K. Mander, J. Mosley, C. Ramirez-Molina, P. Rowland, C. J. Schofield, R. J. Sheppard, J. E. Smith, C. Swales, R. Tanner, P. Thomas, A. Tumber, G. Drewes, U. Oppermann, D. J. Patel, K. Lee, and D. M. Wilson, “Kruidenier et al. reply.,” *Nature*, vol. 514, no. 7520, p. E2, Oct. 2014.

[223] J. T. Ng, C. Dekker, M. Kroemer, M. Osborne, and F. von Delft, “Using textons to rank crystallization droplets by the likely presence of crystals.,” *Acta Crystallogr. D. Biol. Crystallogr.*, vol. 70, no. Pt 10, pp. 2702–18, Oct. 2014.

[224] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, and R. J. Read, “Phaser crystallographic software.,” *J. Appl. Crystallogr.*, vol. 40, no. Pt 4, pp. 658–674, Aug. 2007.

[225] P. D. Adams, P. V Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, and P. H. Zwart, “PHENIX: a comprehensive Python-based system for macromolecular structure solution.,” *Acta Crystallogr. D. Biol. Crystallogr.*, vol. 66, no. Pt 2, pp. 213–21, Feb. 2010.

[226] P. Emsley and K. Cowtan, “Coot: model-building tools for molecular graphics,” *Acta Crystallogr. Sect. D*, vol. 60, no. 12 Part 1, pp. 2126–2132, Dec. 2004.

[227] G. Bricogne, E. Blanc, M. Brandl, C. Flensburg, P. Keller, W. Paciorek, P. Roversi, A. Sharff, O. S. Smart, C. Vonrhein, and T. O. Womack, “BUSTER.” Cambridge, United Kingdom: Global Phasing Ltd., 2011.

[228] G. Winter, C. M. C. Lobley, and S. M. Prince, “Decision making in xia2.,” *Acta Crystallogr. D. Biol. Crystallogr.*, vol. 69, no. Pt 7, pp. 1260–73, Jul. 2013.

[229] S. J. Fisher, K. E. Levik, M. A. Williams, A. W. Ashton, and K. E. McAuley, “SynchWeb: a modern interface for ISPyB.,” *J. Appl. Crystallogr.*, vol. 48, no. Pt 3, pp. 927–932, Jun. 2015.

[230] D. W. A. Buchan, S. M. Ward, A. E. Lobley, T. C. O. Nugent, K. Bryson, and D. T. Jones, “Protein annotation and modelling servers at University College London.,” *Nucleic Acids Res.*, vol. 38, no. Web Server issue, pp. W563–8, Jul. 2010.

[231] K. Katoh and H. Toh, “Recent developments in the MAFFT multiple sequence alignment program.,” *Brief. Bioinform.*, vol. 9, no. 4, pp. 286–98, Jul. 2008.

[232] J. D. Thompson, T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins, “The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools.,” *Nucleic Acids Res.*, vol. 25, no. 24, pp. 4876–82, Dec. 1997.

[233] J. H. Choi, H. Y. Jung, H. S. Kim, and H. G. Cho, “PhyloDraw: a phylogenetic tree drawing system.,” *Bioinformatics*, vol. 16, no. 11, pp. 1056–8, Nov. 2000.