Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____   _____

Ruogu Hu   Date

Enzymatic and Biochemical Characterization of the JARID1C Jumonji Demethylase

By

Ruogu (Roc) Hu

B.S., Sun Yat-Sen University, 2008

Graduate Division of Biological and Biomedical Sciences (GDBBS)

Biochemistry, Cell and Developmental Biology (BCDB) Program

_____
Xiaodong Cheng, Ph.D.
Advisor

_____
Haian Fu, Ph.D.
Committee Member

_____
William G. Kelly, Ph.D.
Committee Member

_____
Eric Ortlund, Ph.D.
Committee Member

_____
Paula M. Vertino, Ph.D.
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____
Date

Enzymatic and Biochemical Characterization of the JARID1C Jumonji Demethylase

By

Ruogu (Roc) Hu

B.S., Sun Yat-Sen University, 2008

Advisor: Xiaodong Cheng, Ph.D.

An Abstract of

A thesis submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Master of Science

Graduate Division of Biological and Biomedical Sciences (GDBBS)

Biochemistry, Cell and Developmental Biology (BCDB) Program

2011

Abstract

Enzymatic and Biochemical Characterization of the JARID1C Jumonji Demethylase

By

Ruogu (Roc) Hu

Human diseases result from epigenetic changes in addition to genetic alterations in the cell. A wide range of cancers and neurological disorders have been linked to aberrant patterns of covalent modifications in chromatin. For example, methylations of DNA and histones affect transcription and are controlled by methyltransferases and demethylases, among which Jumonji demethylases play important roles. Jumonji proteins belong to a superfamily of 2-oxoglutarate (2-OG)-Fe (II)-dependent dioxygenases and reverse histone lysine methylation. The evolutionarily conserved JARID1 (Jumonji AT-Rich Interactive Domain 1) family of Jumonji demethylases catalyzes demethylation of histone H3 lysine 4 (H3K4me2/3, a mark for active transcription). Their JmjC domains bear enzymatic activity, and such substrate specificity is potentially affected by the additional chromatin-reader domains (such as Plant Homeodomain, or PHD), representing a common mechanism in regulation of chromatin modifiers. In this study, I biochemically characterized the enzymatic activity of JARID1C on H3K4 methylation with different histone peptides, and explored the binding specificity of the two JARID1C PHD domains. An N-terminal JARID1C recombinant fragment was active on synthetic histone H3K4me2/3 peptides. The end product of demethylation is mono-methylated H3K4. This activity was affected by peptide length, yet independent of the methylation status of H3K9. Additionally, I measured the Michaelis-Menton kinetics of JARID1C demethylation in a formaldehyde release assay and showed that the R and S form of a 2-OG analogue, 2-hydroxyglutarate, could inhibit JARID1C activity with similar IC_{50} values (0.8 and 0.7 mM, respectively). Using bacterially expressed JARID1C PHD domains in pull-down and isotherm titration calorimetry assays, I showed that PHD1 domain preferentially associates with unmodified H3K4 peptide; but such binding is very weak compared to other known interactions between PHD domains and histone peptides. Similar results were obtained for PHD2. These results revealed the properties of JARID1C catalysis and interaction with histones. JAIRD1C is implicated in X-linked mental retardation and cancers. This work provides a biochemical perspective of the protein which complements current *in vivo* data in the literature. In the long run, it will help to elucidate mechanisms and roles of histone demethylation in human diseases, potentially leading to development of novel epigenetic targets for therapeutics.

Enzymatic and Biochemical Characterization of the JARID1C Jumonji Demethylase

By

Ruogu (Roc) Hu

B.S., Sun Yat-Sen University, 2008

Advisor: Xiaodong Cheng, Ph.D.

A thesis submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Master of Science

Graduate Division of Biological and Biomedical Sciences (GDBBS)

Biochemistry, Cell and Developmental Biology (BCDB) Program

2011

Acknowledgments

I would like to sincerely thank Dr. Xiaodong Cheng, Dr. Xing Zhang, and all members of the Cheng laboratory for their continued training and valuable guidance on this work. I am truly grateful of the kindness of and encouragement from my committee members, whose helpful advice lead me through the project. I would also like to thank all laboratories in the biochemistry department which provided facilities for my experiments, especially the Ortlund laboratory and the Cummings Laboratory. I appreciate the priceless help and friendship from my fellow students in the BCDB and other programs of the GDBBS, without which I could not work through these years. Finally, I thank my family members for being empowered by their unconditional love and support.

Table of Contents

List of Figures

List of Tables

Chapter I: Background and Introduction

1.1 Epigenetic changes in human diseases and methylation in the chromatin

Epigenetic changes refer to the reversible and heritable alterations in gene expression due to changes beyond DNA sequence, such as post-synthetic, covalent modifications of chromatin components. Various enzymes are responsible for such post-translational modifications (Berger, 2007) . As a key signal for gene activation and silencing, chromatin methylation is regulated by methyltransferases and demethylases. On DNA, DNA methyltransferases (DNMTs) catalyze cytosine methylation of CpG dinucleotides in mammals (Cheng and Blumenthal, 2008). Non-CpG methylation is also observed (Burden et al., 2005; Ramsahoye et al., 2000). On histones, the majority of known methylation marks under dynamic regulation are on the N-terminal tails of histone H3 and H4. Histone methyltransferases add methyl groups on histone lysine or arginine residues. Histone lysine demethylases remove methyl groups from methyl-lysines. Figure 1 lists several important histone lysine methyltransferases and histone lysine demethylases (Shi, 2007). The diversity of such enzymes underlies the specificity and reversibility of methylation, providing a framework to understand where complex gene regulation takes place. Furthermore, continued studies of these histone modifiers by biochemistry and structural biology reveal their potential as drug targets.

In eukaryotes, methylation marks established by related enzymes have different roles based on their context. For instance, tri-methylated H3 lysine 4 (H3K4me3) is a well-studied mark for transcriptional activation by recruitment of chromatin remodelers (Li et al., 2006; Schneider et al., 2004) and RNA polymerase II to initiate transcription (Santos-Rosa et al., 2003). Similarly, H3K9me2 is a repressive mark associated with transcriptional silencing (Nakayama et al., 2001; Perillo et al., 2008; Sims Iii et al., 2003), which is also involved in heterochromatin packaging and spreading (Bannister et al., 2001). On DNA, methylation of cytosines in mammalian CpG islands is

a hallmark of gene silencing (Eden et al., 1998). The coordination of these marks is thought to be crucial for proper control of gene expression, often achieved by interaction between methylation enzymes and readers (Cedar and Bergman, 2009). For example, a model for *de novo* methylation in cancer involves an H3K27me3 methyltransferase EZH2 and DNMT3A/B (Vire et al., 2006). In addition, the complexity of methylation-mediated gene regulation is further increased by the reversibility of methylation marks, which results from histone demethylases, particularly a large number of Jumonji proteins (Lan et al., 2008), as well as DNA demethylation pathways (Kangaspeska et al., 2008; Metivier et al., 2008). These DNA and histone modifications work collectively and constitute complex histone codes to regulate gene expression (Collins and Cheng, 2010; Gamble and Kraus, 2007).

1.2 Histone demethylation and modes of transcriptional control

1.2.1 Jumonji demethylases

Jumonji-proteins belong to a family of 2-oxoglutarate (2-OG)-Fe (II)-dependent dioxygenases with a variety of substrates. The first Jumonji gene, *jmj*, was discovered to be important for mouse development, as a *jmj* mutant generated an abnormal neural groove on the neural plate, forming a cross ("Jumonji" in Japanese) with the normal neural groove (Takeuchi et al., 1995). A large group of proteins bearing a conserved Jumonji (JmjC) domain have been identified, many of which are important histone lysine demethylases (Takeuchi et al., 2006).

Currently, three classes of enzymes are characterized as histone demethylase: PADI4, LSD1, as well as the JmjC-domain-containing proteins (Klose et al., 2006). The first identified H3K4 demethylase, Lysine-specific demethylase 1 (LSD1) targets only mono- and di-methyl lysines. Importantly, JmjC-domain containing demethylases can catalyze demethylation of tri-methylated lysines along with di/mono-methylated lysines, expanding the variety of methyl-lysine substrates in addition to other demethylases (Klose et al., 2006). Jumonji proteins contain the JmjC domain,

which folds into eight beta - sheets and contains residues required for Fe(II) and α-ketoglutarate binding. During demethylation reaction, a highly reactive oxoferryl species is produced and it hydroxylates hydroxylates the methyl group on the nitrogen atom of lysine side chain, and causes the spontaneous removal of the methyl group as formaldehyde (HCHO), yielding an unmodified lysine and converting 2-OG to succinate as a byproduct (Figure 2).

Jumonji proteins have more diverse targets other than histone lysines. The diversity of Jumonji protein substrates can be exemplified by the following proteins. Human Jmjd6 catalyzes lysyl-hydroxylation of U2AF65, a protein related to RNA splicing (Webby et al., 2009). Human ABH2/3, homologs of an *E.coli* DNA repair enzyme AlkB, remove methyl groups from 1-methyladenosine, 1-methylguanine, 3-methylcytosine, and 3-methylthymine in DNA (Yang et al., 2008). In addition, human Tenth-Eleventh-Translocation 1 (TET1) converts 5-methylcytosine into 5-hydroxyl-methylcytosine in DNA (Tahiliani et al., 2009).

1.2.2 Crosstalk between methylation marks mediated by Jumonji proteins

The notion that histone methylation on a particular lysine can be affected by the methylation status of another lysine on the same protein extends our understanding of how substrate specificity of Jumonji demethylases can be achieved in a chromatin context. Previous studies in the lab on PHF8 and KIAA1718 Jumonji demethylases show that these enzymes have promiscuous JmjC domains that are active on multiple H3 lysines and also contain PHD domains (Plant Homeodomain) that bind to H3K4me3. However, the difference in the length of a linker between PHD and JmjC domain defines which lysine they access and demethylate (Horton et al., 2010). When PHF8 is locked on a histone peptide via H3K4me3 - PHD binding, PHF8 has a linker that positions the JmjC domain on H3K9me2 of the same peptide allowing demethylation. However, when KIAA1718 binds to H3K4me3 via its PHD domain, it has a longer linker that keeps the JmjC domain away from H3K9me2, thus inhibiting its activity on this particular lysine. More convincingly, when

KIAA1718 was engineered with a PHF8 linker to replace the native linker, it gains activity on the

same H3K4me3-K9me2 peptide. This again fits a model where histone code
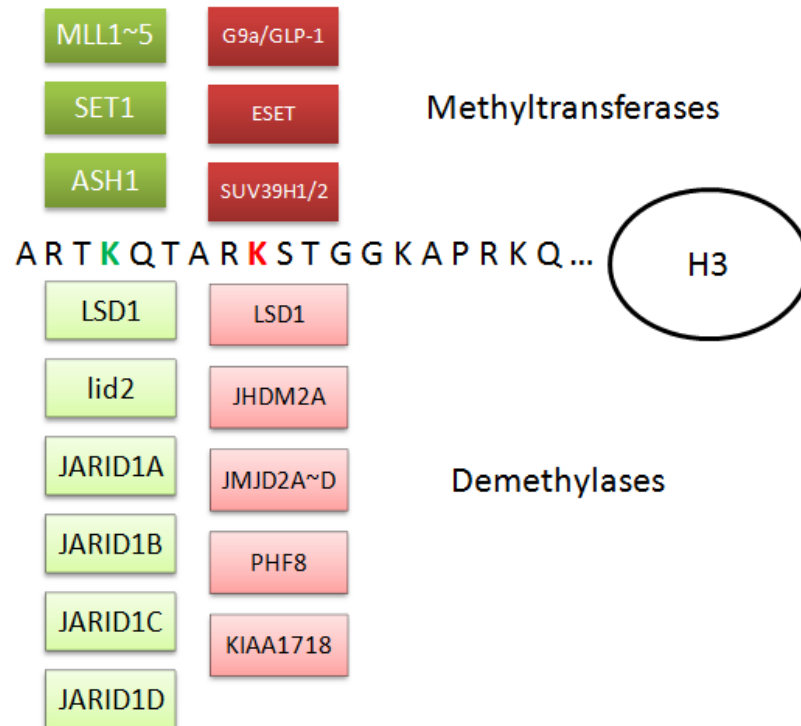
Figure 1: Examples of H3K4 and H3K9 methyltransferases and demethylases. The amino acid sequence in the middle is histone H3 N-terminal tail which undergoes extensive post translational modifications. This graph summarizes important histone lysine methyltransferases and demethylase that act on histone H3K4 and H3K9, which mark active and inactive transcription at promoters of genes in eukaryotes.
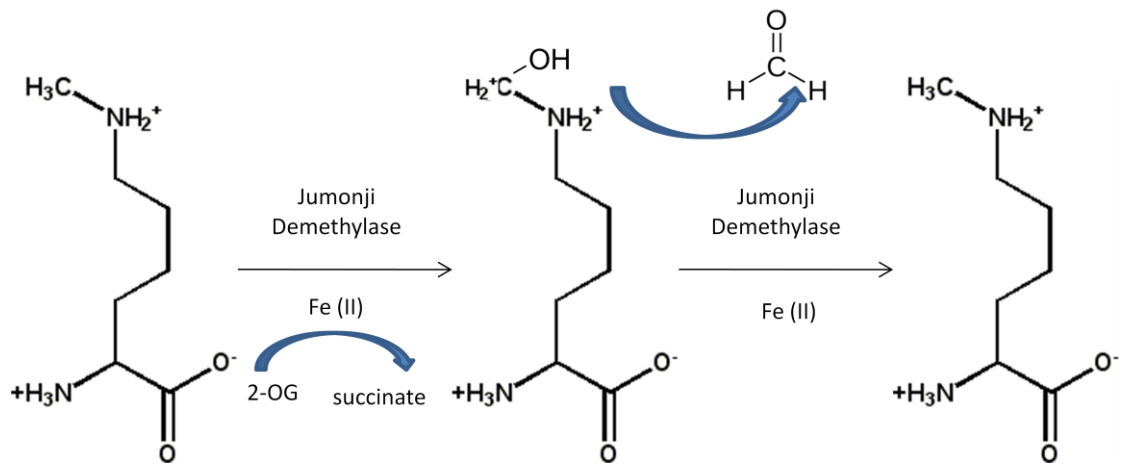
Figure 2: Mechanism of 2-OG dependent demethylation catalyzed by Jumonji proteins.

readers such as PHD domains can impact the *de facto* substrate specificity of a histone code eraser such as JmjC domain- carrying enzymes.

Remarkably, each JARID1 protein contains 2-3 PHD domains, which were reported to bind methylated histone lysines (Iwase et al., 2007; Li et al., 2008; Wang et al., 2009). At present, the interplay between JARID1 PHD and JmjC domains has been studied for Lid2, a fission yeast JARID1 ortholog shown to bear H3K4me3 demethylation activity (Li et al., 2008). In this study, the authors showed that Lid2 PHD2- and PHD3-deletion mutants of *S. pombe* fail to silence the *inner repeats* (*imr*) gene, which is normally repressed by a complex containing Lid2. Unlike the case of PHF8/KIAA1718, no current biochemical or structural data directly supports the physical link between Lid2's PHD and JmjC domains. Yet this example does provide grounds for the hypothesis that JARID1 protein can mediate crosstalk between H3K4 and H3K9 methylation, a model similar to that of PHF8/KIAA1718.

1.3 The JARID1 family histone demethylases

1.3.1 JARID1 proteins have multiple domains and target H3K4 methylation

JARID1 family demethylases are transcriptional repressors (Cloos et al., 2008). They reverse the active gene mark H3K4me3 at transcription start sites of target genes, and exist as members of multi-unit complex in mammalian chromatin. Their JmjC domains share sequence homology in the key amino acid residues bearing cofactor binding capability (Figure 3). While JmjC domains have been demonstrated to account for demethylation activity, the additional domains in JARID1s may mediate the interactions between these demethylases and interacting proteins (Li et al., 2008).

JARID1A (RBP2) binds to retinoblastoma protein, a tumor suppressor gene (Benevolenskaya et al., 2005; Gutierrez et al., 2005; Klose et al., 2007), and it is found to be over-expressed in sub-population of drug-tolerant cancer cells (Sharma et al., 2010). JARID1B (PLU1)

over-expression is found in breast cancer (Xiang et al., 2007; Yao et al., 2009). JARID1C (SMCX) mutations that affect its demethylase activity are found in mental retardation patients (Iwase et al., 2007; Tahiliani et al., 2007). JARID1D (SMCY) is an important male-specific antigen (Secombe and Eisenman, 2007).

Although Jumonji demethylases have different histone lysine substrates, they share highly conserved Jumonji domains (JmjC; Figure 3B), the eraser of histone methyl marks. Distinct substrate specificity determines the role of each demethylase *in vivo* and raises the question how specificity is regulated. A tempting answer to this question is the previously-mentioned PHF8/KIAA1718 model, considering the additional reader domains (such as PHD and ARID) in JARID1 proteins. The example of Lid2 provides some support for this model (see below).

1.3.2 The fission yeast JARID1 homolog Lid2 has demethylation activity on H3K4me2/3 and inversely correlates with H3K9 methylation

As a fission yeast JARID1 ortholog, Lid2 is also a bona fide H3K4me3 demethylase, as mutation in the JmjC domain abolishes this activity both *in vitro* and in yeast (Li et al., 2008). Lid2 PHD2 domain preferentially binds to H3K9me3 as reported by pull-down data, yet how this interaction would impact Lid2's JmjC domain function (if any) is not clear. Fission yeast studies did show mutually exclusive H3K4 and H3K9 methylation in heterochromatin regions such as centromeres (Li et al., 2008). A test for such relationship using purified Lid2/JARID1 proteins and doubly methylated histone H3 peptides could uncover mechanistic details this demethylation process. It is possible that the physical and functional link between H3K4 and H3K9 via Lid2 is more complex than (but not exclusive of) a simple function of the length between JmjC and PHD domains, because Lid2 resides in a complex with many other proteins in yeast. This information will be useful in extending to our understanding of human JARID1 proteins.
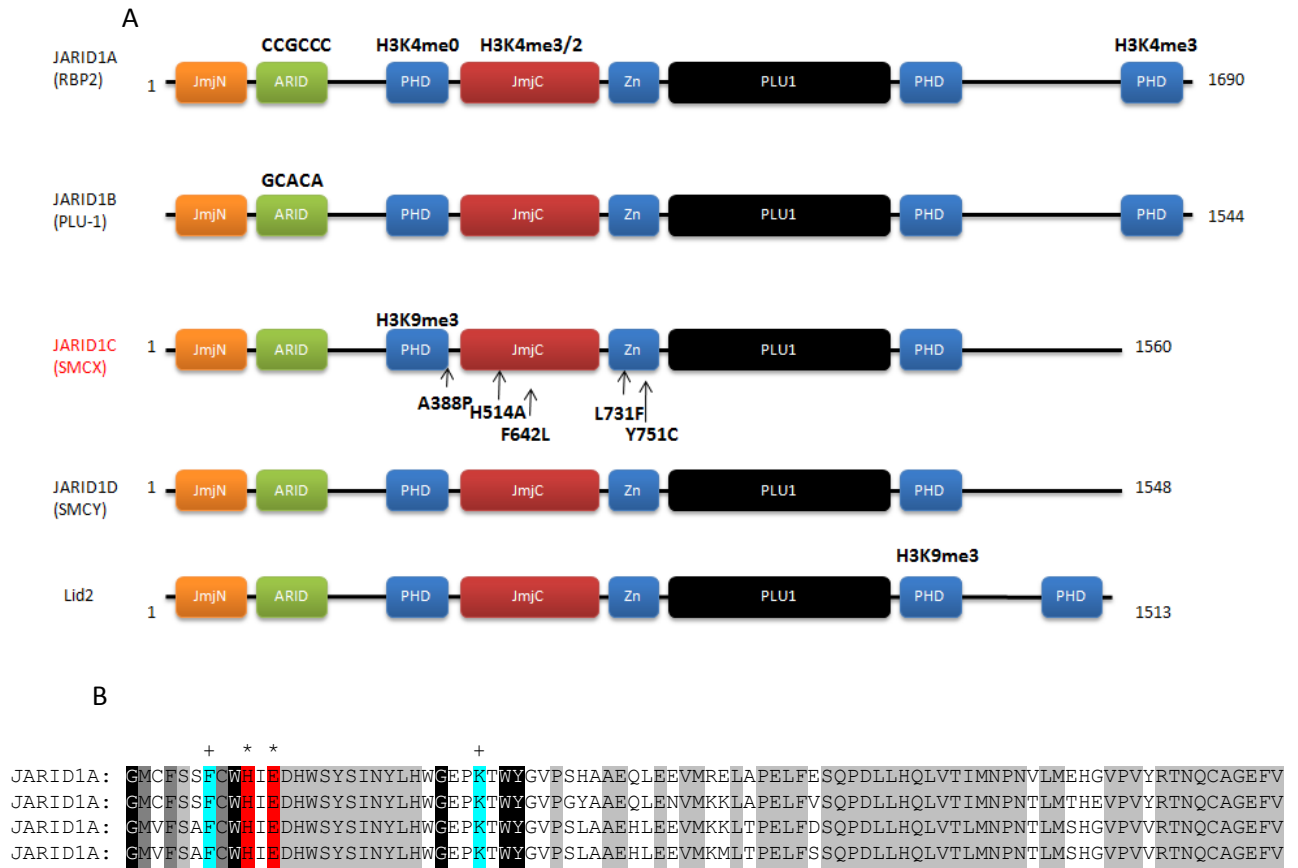
Figure 3: Summary of known JARID1 proteins. Panel A: Domain organization in the JARID1 family. DNA sequences and histone H3 methylation marks were listed above their respective binding domains. JARID1C mutations in XLMR were also listed. Panel B: Conserved cofactor binding residues in JARID1 JmjC active sites. Multiple sequence alignment of the JmjC domains of 4 human JARID1 proteins highlights the conservation of key residues which are Fe (II)-binding (*) and 2-OG-binding sites (+).

1.4 Significance

Diverse diseases are caused by enzymatic modifiers of the chromatin (Lan et al., 2008). Because such proteins control basic cellular processes such as gene expression, mis-regulation of these enzymes has been implicated in many diseases that can be as severe as developmental and neurological disorders (Qiu et al., 2010; Tahiliani et al., 2007). The capability of histone demethylases to fine tune gene expression via remodeling chromatin states make them critical control points in biological processes and draws significant research efforts in the field. An increasing number of histone lysine demethylase inhibitors have been reported. For example, the oncometabolite 2-hydroxyglutarate (2-HG), an analogue of 2-oxylglutarate (2-OG, or αKG), has been shown to inhibit a wide range of 2-OG dependent oxygenases including Jumonji demethylases (Xu et al., 2011). We therefore tested 2-HG's inhibition on JAIRD1C in our study. The long term goal of this project is to elucidate how gene regulation is achieved through the complex network of chromatin modifications. This biochemical study of histone demethylase will not only contribute to the understanding of epigenetic regulation and the underlying the pathology of human diseases, but also provide information for therapeutic development with epigenetic targets such as designing histone demethylase inhibitors.

Chapter II: Enzymatic properties of JARID1C

2.1 Introduction

2.1.1 Enzymatic properties of H3K4me1/2 demethylases

As mentioned in Chapter I, LSD1 belongs to the monoamide oxdase family and was the first histone demethylase identified, and can only act on H3K4 mono- and di-methylaed states, but not the tri-methylated state. The reaction is a flavin-dependent oxidative process (Shi et al., 2004) and the exact mechanism remains debated (Culhane and Cole, 2007). LSD1 needs at least the first 16 amino acids of histone H3 to display catalytic activity on H3K4 (Forneris et al., 2007), showing that peptide length in *in vitro* demethylation assays can be an important factor and is therefore considered in our study. Despite substrate specificity, LSD1 resembles Jumonji proteins in that its activity is also affected by additional domains. The N-terminal SWIRM domain contributes to protein stability; an insertion in the catalytic domain is critical for both for the demethylase activity and the interaction with CoREST, a corepressor complex containing LSD1 (Chen et al., 2006). Additionally, LSD1 has been proposed to also catalyze H3K9 demethylation *in vivo* in an androgen-dependent manner (Metzger et al., 2005). This result, although not corroborated by *in vitro* biochemical studies, shows that LSD1 acts like Jumonji proteins such as JMJD2A/PHF8 (Couture et al., 2007; Horton et al., 2010) to demethylate multiple marks based on chromatin context. Such context can be small motifs of protein sequences (in the case of JMJD2A) or distal modifications on the nucleosome read by the demethylases (in the case of PHF8). Thus, the chemistry and mode of regulation of LSD1 shares similarity to Jumonji proteins.

2.1.2 Enzymatic properties of H3K4me2/3 demethylases

Not all Jumonji proteins can act on tri-methylated histone lysines. For example, PHF8/KIAA1718 have activity only on H3K9me1/2 (Feng et al., 2010) but not H3K9me3, similar to LSD1's ability in terms of substrate methylation states. In this regard, JARID1 proteins are different because they can target tri-methyl lysine as their substrate, similar to the JMJD2 family of Jumonji proteins (Tan et al., 2008). Previous *in vivo* studies reported that the end product of JAIRD1C catalyzed reaction is H3K4me1 (Tahiliani et al., 2007), and the reaction utilizes the same Fe(II)-2-OG-dependent oxidative mechanism of all Jumonji proteins. However, no detailed biochemical studies were performed to characterize the enzymatic properties of JARID1C. In this work, we utilized different *in vitro* assays to measure the activity of JARID1C and measured its Michaelis-Menten kinetic parameters. Mass Spectrometry and a fluorescence-based formaldehyde release assay were used as read outs for demethylation reactions as previously described (Horton et al., 2010). In addition to kinetic characterization, we compared the activity of JARID1C on H3K4me3 with or without H3K9me3 on the same peptide molecule to test our model. In addition, two DNA sequences were reported to bind known ARID domains of two proteins, Dead-ringer in *D.drosophila* and JARID1B in human (Iwahara et al., 2002; Tu et al., 2008). Because ARID domains share high similarity, we also tested the binding of these DNA sequences to JARID1C and the effect of JARID1C's activity. Finally, as stated in Chapter I, 2-HG has been shown inhibitory toward known Jumonji demethylases, so we also tested the inhibitory effect of these 2-OG analogues on JARID1C catalysis and determined their IC_{50} values.

2.2 Results

2.2.1 JARID1C (residues 1-839) is active on H3K4me3 peptides in vitro

JARID1C is a large protein with 1560 amino acids. In order to purify an active recombinant construct, we tested several combinations of tags and protein length, and found that JARID1C (1-839) with a His-SUMO tag (construct number: pXC911) was the best behaved during purification

and therefore could be recovered in suitable amounts. JARID1C (1-839) has been shown to retain enzymatic activity *in vivo* and encompasses the N-terminal domains: JmjN-ARID-PHD1-JmjC-Zn, and part of PLU-1 (Figure 4A). We could purify this protein to over 90% homogeneity based on SDS-PAGE analysis (Figure 4B) and the protein is active on histone peptides in our assays. The purified protein was used for all subsequent biochemical analysis. Using mass spectrometry as our detection method (Figure 5; see materials and methods), the optimal reaction conditions for our purified JAIRD1C (1-839) has been determined to be MES buffer pH 6.5 and a NaCl concentration of 50mM (Figure 6A). Given this condition, JAIRD1C can demethylate H3K4me3 to H3K4me2 and H3K4me1, and such activity is dependent on peptide length (Figure 6B). Demethylation activity was not observed on (1-10) H3K4me2 (Figure 6C); yet activity increases as the H3 peptide length go from 1-12, 1-21, to 1-24, with 50% conversion times of H3K4me3 substrate being approximately 10, 5, and 4 minutes, respectively.

2.2.2 JARID1C's activity on H3K4 is independent of H3K9 methylation status on the same peptide

As we know, H3K9 methylation and H3K4 methylation at eukaryotic gene promoters has opposite effects in terms of transcriptional activity. Given that JARID1C's PHD1 domain has been reported to bind to H3K9me3, we tested the effect of H3K9 methylation status on H3K4 demethylation by JARID1C. We compared JARID1C's activity on two H3K4me3 peptides, one bearing H3K9me3 and the other bearing H3K9me0. We observed that modification at H3K9 does not affect the rate of demethylation on H3K4 for JARID1C (Figure 7), which supports that the way in which JARID1C JmjC domain accesses its substrate H3K4me3 is not altered and therefore catalysis is unaffected by modification at H3K9. This suggests that either the physical link between PHD1 and JmjC is flexible and therefore PHD1 binding to H3K9me3 does not affect JmjC-substrate interaction, or that PHD1 does not bind to H3K9me3 at all, assuming a relatively rigid

link between PHD1 and JmjC. We tested the binding of PHD1 to H3K9 and will discuss it in

Chapter III.
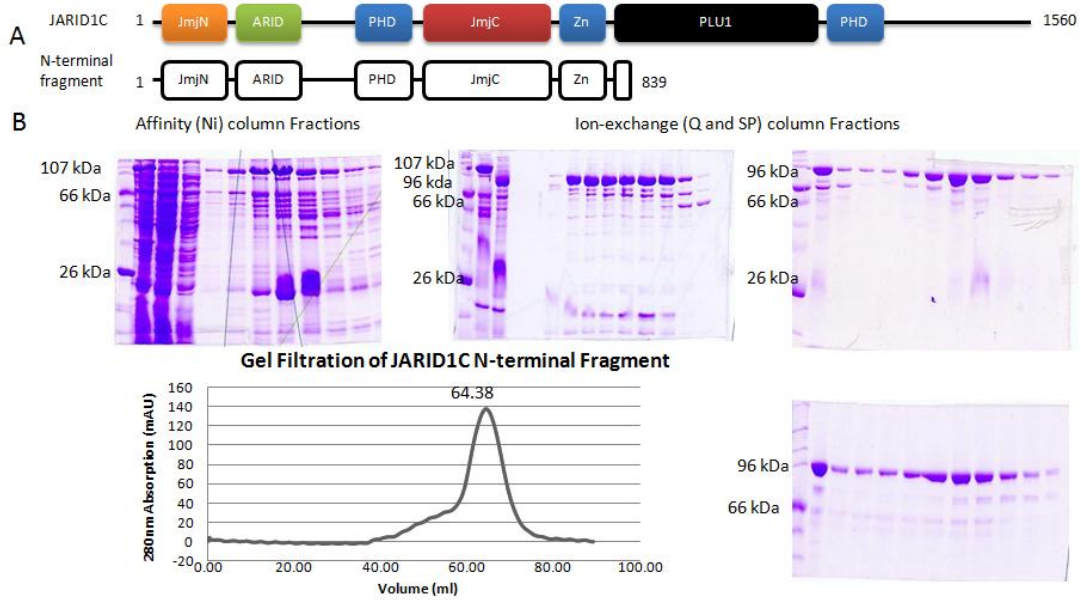
Figure 4: Purification of JARID1C N-terminal fragment (residues 1-839). The cartoons in panel A are schematic view of the JARID1C protein and our construct. Panel B shows representative gels corresponding to fractions collected during purification after going through respective columns. The first lane on the left of each gel is NEB Broad Range Protein Marker (2 - 212 KDa). The protein sizes are 107 kDa and 96 kDa, before and after cleavage of His-SUMO (see Q column fractions). The gel filtration graph shows the relationship between protein size and elution volume.
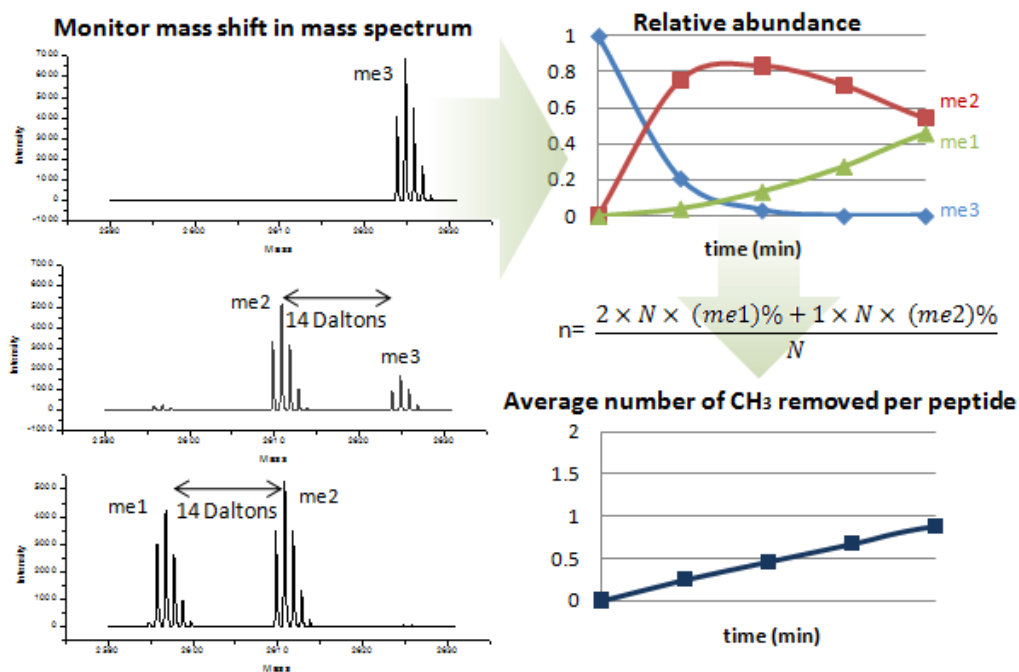
Figure 5: An example of analysis and presentation of JARID1C reaction data from mass spectrometry. An example mass spectrum at different time points was shown on the left. The area under each peak represents the amount of peptide species corresponds to the peaks. This information was used to calculate the relative abundance of each peptide species on top right. The average number of methyl group removed per peptide (n) was converted from relative abundance based on the equation shown. The number of peptides (N) is an arbitrary number and cancels itself off during calculation and therefore is independent of the result.
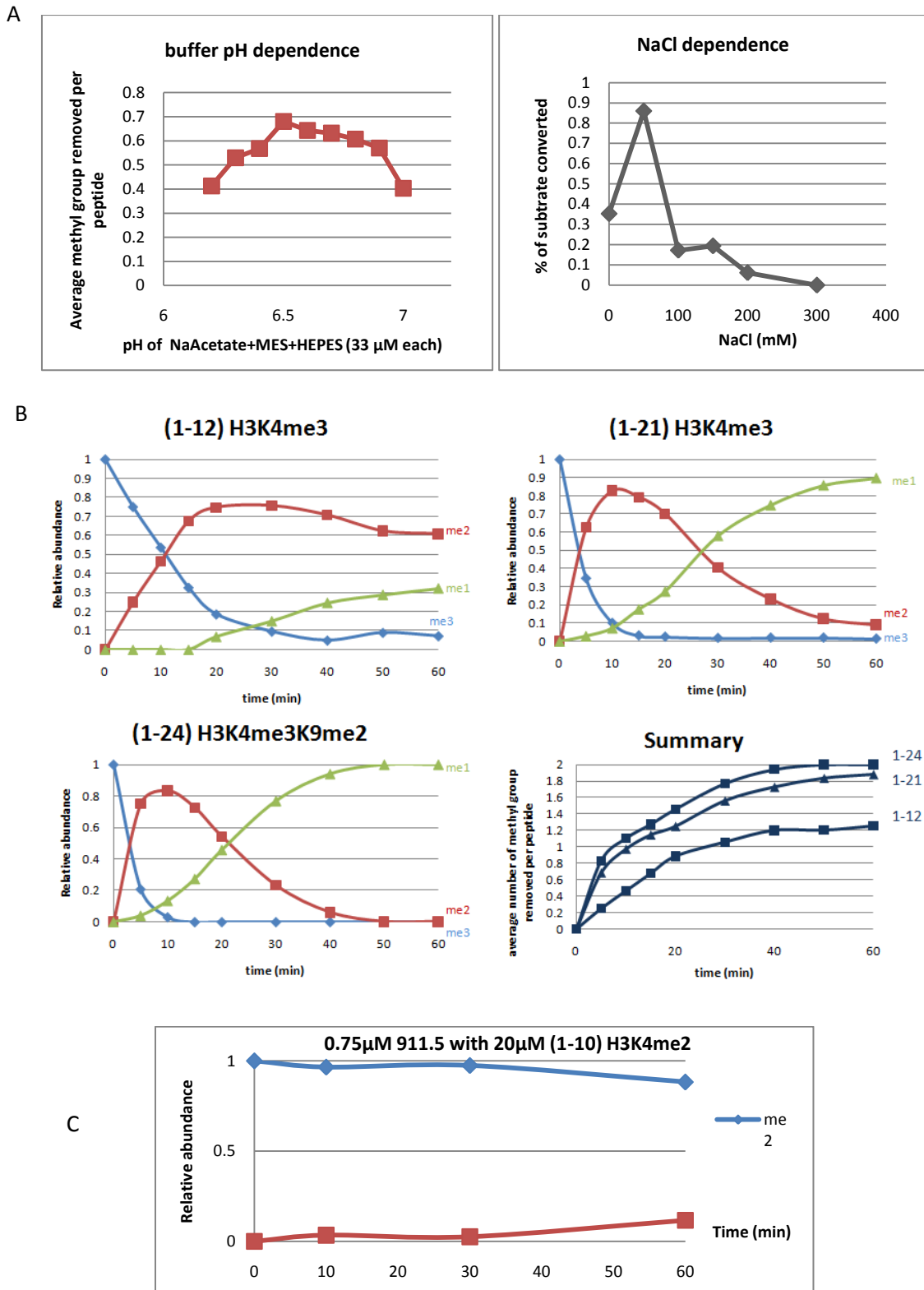
Figure 6: Optimization of JARID1C reaction condition (A) and the effect of peptide length on the rate and extent of reaction (B). For pH and NaCl dependence, data was collected from 5 min time point of a reaction containing 0.75 µM JARID1C (1-839) + 10 µM (1-24) H3K4me3K9me2. Data in B was collected in the same condition except 20 µM of peptide was used and samples for a full time course of 1 h were collected. C shows that the enzyme had no detectable activity on 1-10 H3K4me2.

2.2.3 Effect of two ARID-binding DNA sequences on JARID1C

As suggested in introduction, AIRD domain has been shown to interact with sequence-specific DNA (Iwahara et al., 2002; Tu et al., 2008). Given the ARID domain of JARID1C shares high level of homology to other ARID domains, we also tested binding of two know DNA sequences (see materials & methods on page 33) associated with other ARID domains. The DNA sequence that binds to Dead-ringer ARID domain is rich in AT and was named AT-rich DNA. The DNA sequence that binds to JARID1B ARID domain is rich in GC and was named GC-rich DNA. Figure 7B shows that neither AT-rich nor GC-rich DNA binds to JARID1C in the gel shift assay. Consistent with this observation Figure 7C shows that when AT-rich or GC-rich DNA is added to the JARID1C demethylation reaction, the rate was not affected compared to reaction without any DNA added. Therefore, in our assay conditions, we did not see the effect of these two specific DNA sequences on JARID1C's activity. The putative binding DNA sequences for JARID1C's ARID domain remain to be investigated.

2.2.4 JARID1C's Michaelis-Menten kinetic characteristics

We used a fluorescence-based formaldehyde release assay to determine the kinetic characteristics of JARID1C. Jumonji demethylases remove methyl groups on histone peptides in the form of formaldehyde, which can be converted by formaldehyde dehydrogenase (FDH) to formate. This process is coupled with the reduction of NAD+ to NADH, and the fluorescence generated by NADH can be monitored to estimate the rate of these coupled reactions. We used recombinantly purified FDH enzymes as well as APAD+ (a more stable analogue of NAD+(Roy and Bhagwat, 2007)) in these reactions. We first determined the K_m for 2-OG to be around 5 μM (Figure 8A), and then used a 2-OG concentration of 200 μM (a concentration at which reaction velocity begins to be saturated by sufficient amount of 2-OG) to determined K_m for peptide substrate (~ 3 μM; Figure 8B). The turnover number (k_{cat}) of JAIRD1C is approximately 2.5 ~ 3 per

minute in our experimental system, which is within 2-fold of the estimate based on mass spectrometry (around 2 ~ 5 per minute). Such number shows that our JARID1C (1-839) is faster than most known Jumonji histone demethylases which have published k_{cat} of around 0.1 per minute for PHF8/KIAA1718(Horton et al., 2010) and 0.01 per minute for JMJD2A and JMJD2D (Couture et al., 2007).

A Michaelis-Menten kinetics measurement was also performed for FDH itself, using various concentrations of formaldehyde standards, in a buffer condition identical to that for JARID1C demethylation (MES pH 6.5, 50 mM NaCl). The k_m of FDH for formaldehyde is around 60 μM and the k_{cat} of FDH under this reaction condition is about 17 per minute (Figure 8C). This turnover number indicates that the rate of FDH-catalyzed formaldehyde conversion is roughly 8 times faster than JARID1C-catalyzed formaldehyde formation, thereby eliminating the concern of inefficient enzyme coupling and validating our experimental approach and results.
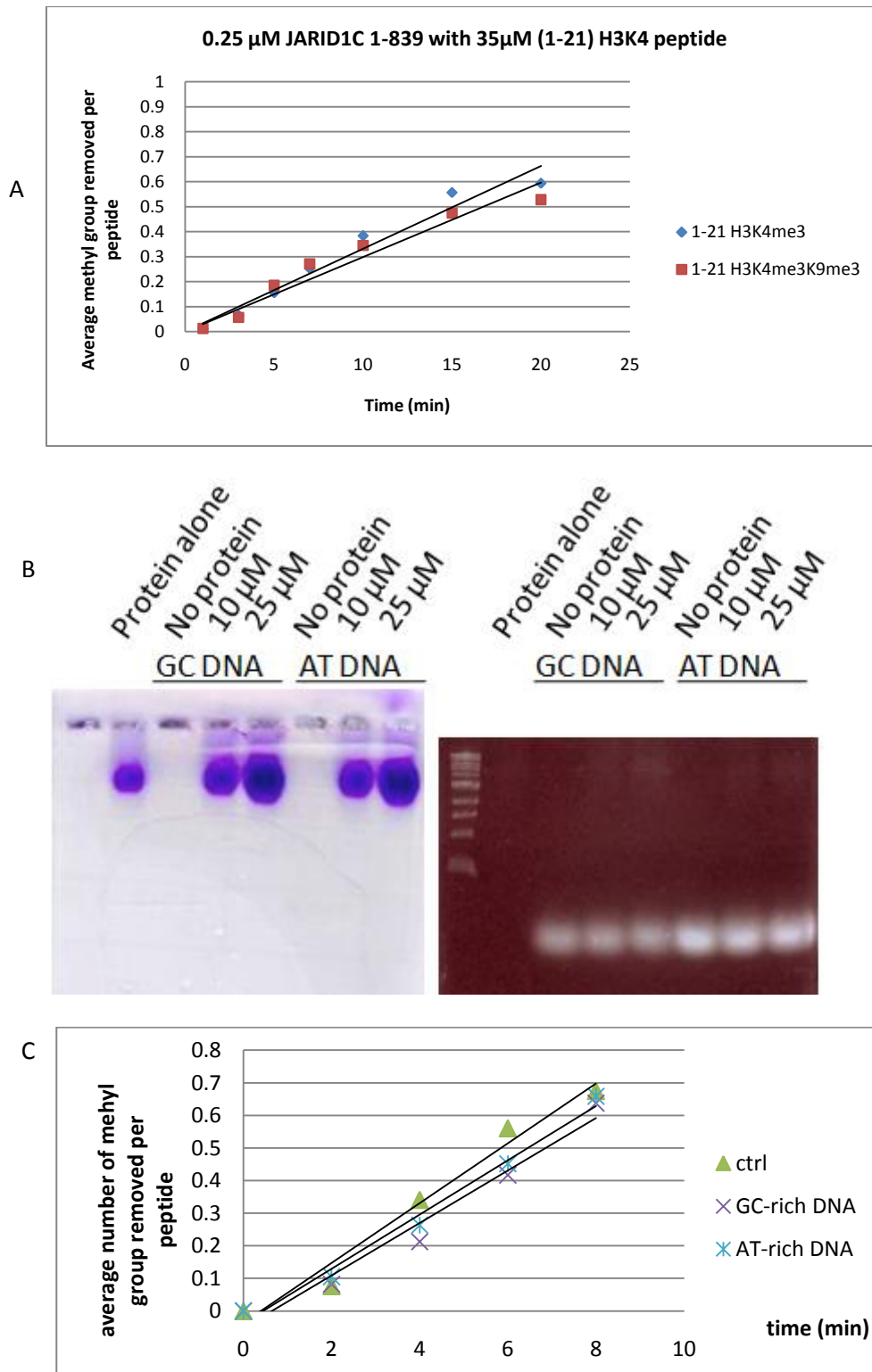
Figure 7: Effect of histone peptide and DNA on the activity JARID1C towards H3K4me3. Panel A shows the time course of reactions with 0.25 μM JARID1C (1-839) and 35 μM (1-21) H3K4me3K9me3 or 1-21 H3K4me3 peptides. No apparent change in the rate of reaction was observed. Panel B shows that no interaction of varying amounts of JARID1C and DNA was detected by gel-shift assay. Left and right are the same gel stained by Coomasie blue (to show proteins) and ethidium bromide (to show DNA), respectively. No DNA bands shifted to where the protein bands resided. Panel C shows that the reaction rate was not changed by addition of DNA sequences.
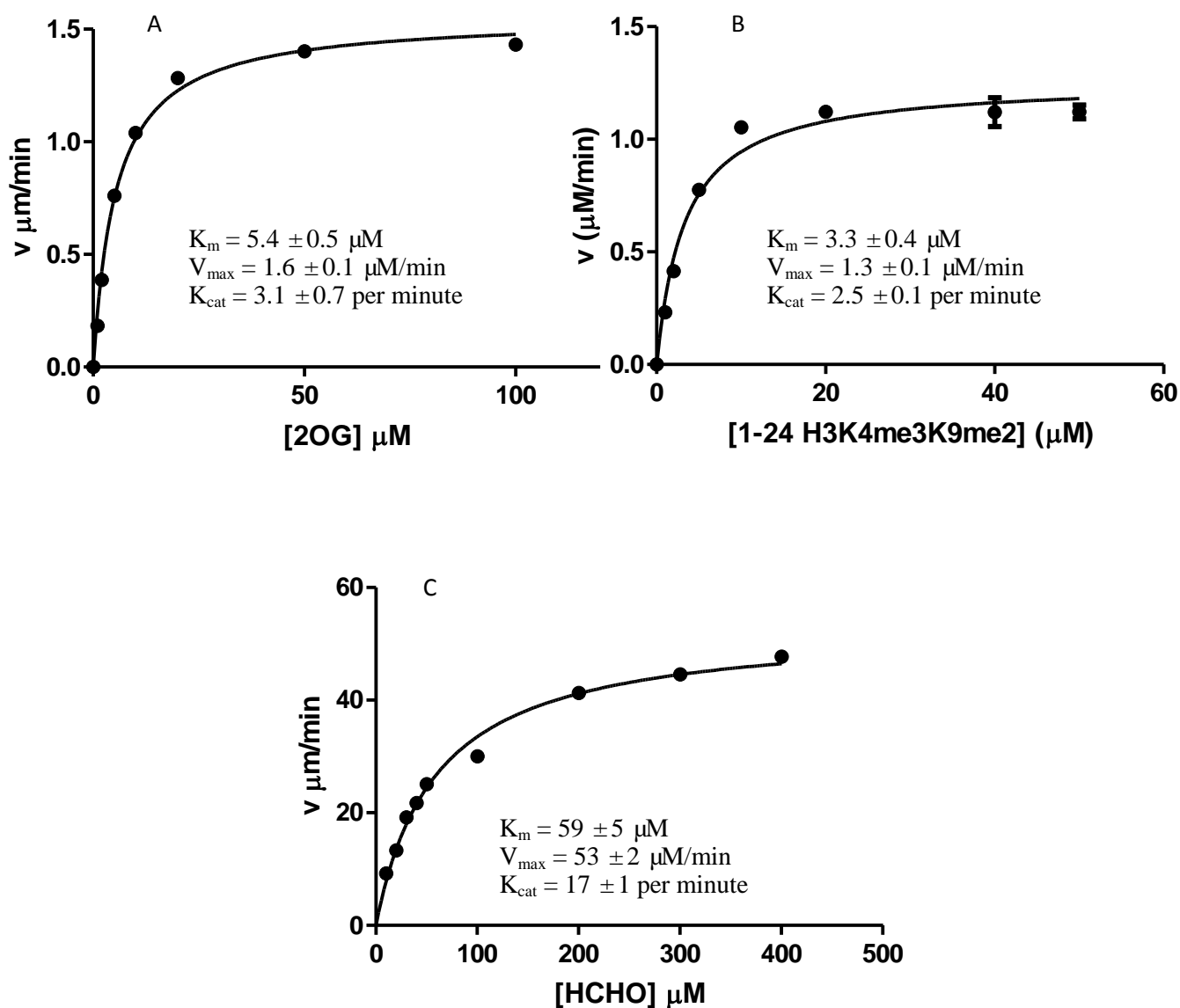
Figure 8: Michaelis-Menten kinetics of JARID1C and FDH. Kinetics of JARID1C was determined by varying concentrations of 2-OG (A) or peptide (B). Initial velocity was plotted against 2-OG or peptide concentration and fitted with the Michaelis-Menten equation. Different concentrations of formaldehyde standards were made fresh and used to get initial velocities (C). The turnover number of FDH is fast enough to ensure the clearance of formaldehyde from JARID1C catalysis is not delayed by FDH conversion. Experiments in panel A and C were in duplicates and experiment in B was in triplicates.

2.2.4 Inhibitory effects of three 2-OG analogues on JARID1C activity

Two analogues of 2-OG are the R and S forms of 2-HG. Previous studies show that R-2HG is the predominant form in the human body in some cases of cancer and is reflective of the oxidative stress in the cell (Ward et al., 2010). Both forms show inhibitory effects toward JMJD2A, 2C, and 2E, which potentially explain the disrupted gene expression in cancer. We tested the inhibitory effect of R and S 2-HG on JARID1C using our fluorescence-based assay. NOG (Figure 9A), an analogue of 2-OG and a known Jumonji demethylase inhibitor, was used for comparison. We observed that the IC_{50} value of NOG is roughly 50-fold lower than both R and S 2-HG (Figure 9B), indicating that neither form of 2HG was as good as NOG in inhibition of JARID1C demethylation. The IC_{50} values are not much different between the two 2-HG forms. However, the S form of 2HG is able to inhibit the v_{max} to close to zero at high concentrations, similar to what NOG does (100% inhibition), while the R form of 2HG can only inhibit v_{max} to about one third of original value (67% inhibition) at high concentrations. Collectively, these results reveal a potential inhibitory role of 2HG on JARID1C, but only at high inhibitor concentrations.
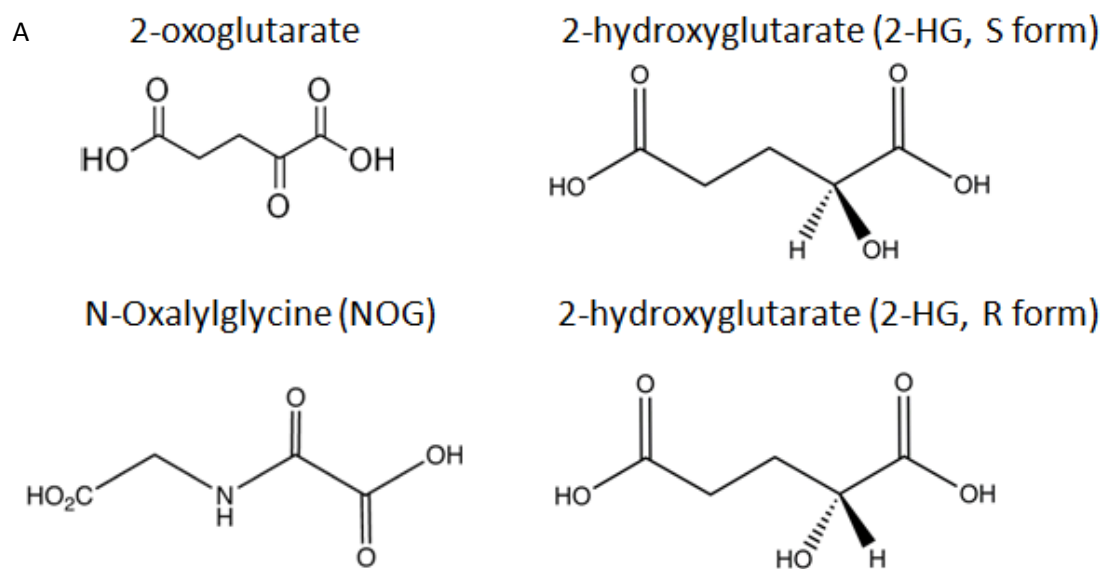
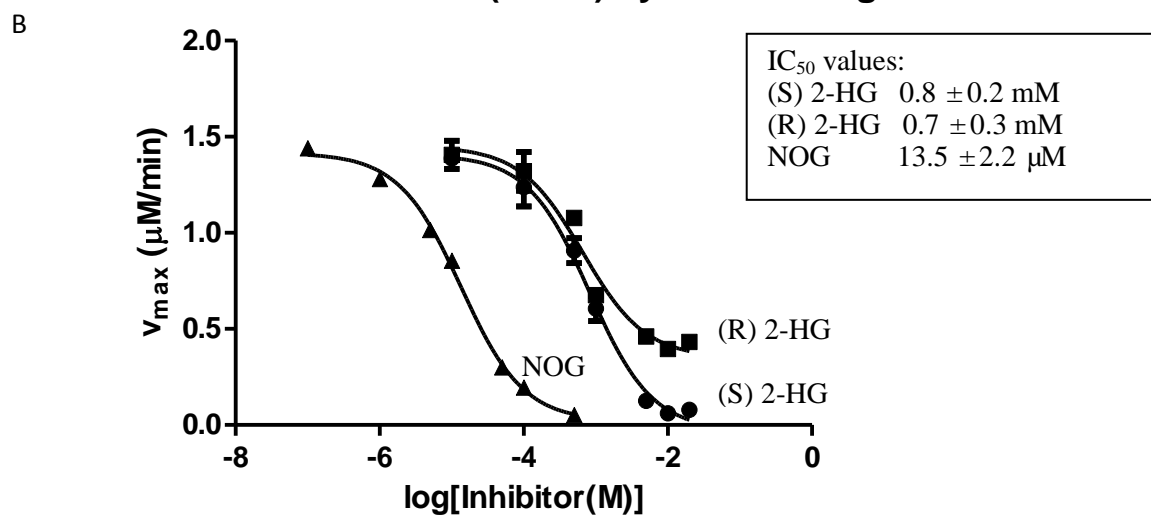Figure 9: Structures (A) and dose-response curve for inhibition of JARID1C activity (B) by 2-OG analogues. IC$_{50}$ values were used to compare the potencies of these inhibitors. Experiments for NOG and (S) 2-HG were in duplicates and experiment for (R) 2-HG was in triplicates.

2.3 Discussion

Our *in vitro* results revealed that JARID1C's basic biochemical properties are consistent with known JARID1C data in the literature. Our JARID1C (1-839) recombinant protein demethylates H3K4me3/2 to H3K4me1 as all other JARID1 family members. This proteins' activity on histone peptides positively correlates with peptide length. JARID1C is very active among Jumonji histone demethylases with a relatively high k_{cat}. However, unlike PHF8, whose activity on H3K9me2 is dependent on the methylation status of H3K4 (Feng et al., 2010), JARID1C does not differentiate the methylation status of H3K9 in terms of its activity on H3K4me3 (Figure 7), since in our assay the doubly methylated H3K4me3K9me3 peptide bears the same level reactivity as th3 singly methylated H3K4me3 peptide. The JARID1C (1-839) construct we have been using contains PHD1 but not PHD2; therefore, we could only test the effect of PHD1 on JmjC domain's activity using this construct. Although the PHD2 domain has not been reported to associate with histone peptides or any other protein, we did make JARID1C constructs that contain the PHD2 domain in the hope of exploring PHD2's effect on activity. We also tested the binding of PHD2 to histone peptides in Chapter III. Unfortunately, our JARID1C constructs containing PHD2 were mostly expressed in the insoluble portion of the bacterial cell lysate and were not purified in suitable amounts. While this is not uncommon for large proteins, it may be worthwhile to switch to another host for protein expression in the future, such as insect (sf9) cells. Unlike bacterial expression systems, such eukaryotic expression system can provide mechanisms for post-translational modification and different sets of molecular chaperones so that the recombinant protein could be expressed in a more native environment as in human cells. Meanwhile, we did not detect interaction between JARID1C and the two known ARID-binding DNA sequences, and no effect on activity was observed. A lot of factors can be taken into consideration for future experiments, such as length of DNA, different fragments of JARID1C protein, the type of detection method other than gel shift assay, etc. This preliminary attempt did show that the ARID domain of

JARID1C might not resemble the ARID domain of JARID1B in terms of interaction with DNA. Knowledge of this interaction is needed before we make any educated guess regarding the effect of such interaction on demethylation reaction.

The inhibition of (R/S) 2-HG on JARID1C activity suggests a potential connection between the mis-regulation of histone demethylases and cancer, at the level of enzyme inhibition instead of transcription or post-translational modifications. (R) 2-HG is accumulated in the cell in malignant glioblastoma and acute myeloid leukemia, sometimes with a concentration up to 10 mM (Chowdhury et al., 2011) . Other histone demethylases such as JMJD2A, JMJD2C, JARID1B, as well as a *C. elegans* KIAA1718 have been reported to be inhibited by both R and S form of 2-HG with different inhibition constant (k_i) values (Xu et al., 2011). Given that the IC_{50} of 2-HG in our reaction system is only about 50-fold of that of NOG, a potent known inhibitor, high concentration of 2HG *in vivo* is likely to cause a real inhibitory effect on JARID1C, which itself has been correlated with several diseases (Kleine-Kohlbrecher et al., 2010). This idea will be supported if data is available in the future to show that high levels of 2-HG accumulate in clinically relevant disease associated with JARID1C. Nevertheless, IC_{50} values of inhibitors for a certain protein can be compared only in the exact same assay conditions, but cannot be compared across different experimental systems. In this regard, a future experiment worthwhile is k_i measurement instead of the IC_{50} measurement done here for 2-HG on JARID1C. Although it would require a significant amount of enzyme, k_i values reflect intrinsic properties of proteins and are comparable across studies, thereby being a direct assessment of the 2-HG inhibitors' potencies in the cell.

2.4 Materials and Methods

2.4.1 Cloning of JARID1 constructs for subsequent analysis

N-terminal fragments of JARID1C were PCR amplified from commercial cDNA source (ATCC) using VENT or Phusion DNA polymerase (NEB) and with primers bearing restriction sites.

PCR fragments were digested with corresponding restriction enzymes and cloned into various expression vectors. We mainly used plasmids that provide hexahistidine (6×His), His-SUMO, or GST tags as well as protease cleavage sites for the removal of tags, such as the pET system and PEGX plasmids. We transformed these recombinant constructs into *E. coli* strain BL21DE3 for protein expression. This process was repeated to generate various versions of the JARID1 proteins (JARID1C 1-789, 1-808, 1-839, and 1-1249, each tagged with GST, His, and His-SUMO). Cloned constructs were confirmed by sequencing of plasmid DNA and induction of the desired protein visualized by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie staining. We picked the best expressed and purified construct (1-839, His-SUMO tagged) for all subsequent biochemical assays.

2.4.2 Purification and identification of active JARID1C constructs

JARID1C constructs were bacterially expressed and purified using fast protein liquid chromatography (FPLC). Typically, 4 to 12 liters of LB media was used to grow bacteria from starting culture, first at 37 ℃ to an OD_{600} value between 0.6 and 1, at which point 1 mM of MgCl2 and 1 mM of ZnCl2 were added and the temperature was set to 16 ℃. After 1~2 hours, when the liquid culture temperature dropped to 16 ℃, 0.4 mM of IPTG was added to induce the expression of recombinant protein overnight. Bacterial culture was centrifuged to collect the cells, which were washed and resuspended with nickel column buffer A (20 mM NaHPO4, 5% glycerol, 500 mM NaCl). Protease inhibitor (PMSF) was added to the cell resuspension and cells were broken by passing through a French Press twice on ice. The cell lysate was centrifuged for at least 1.5 hours and filtered by a 4 micron filter before loading on the first column. Affinity columns (nickel), ion-exchange columns (Q and SP), and gel filtration (sizing) columns were used to achieve the level of purity we want for subsequent analysis. The His-SUMO tag was removed in the purification process by Ulp-1 protease digestion. The final JARID1C (1-839) protein was eluted off the gel

filtration column in 20 mM MES 6.5, 5% glycerol, and 150 mM NaCl and concentrated using a

50kd MWCO Millipore concentrator to about 5 mg/ml, which was determined by OD_{280} reading.

Purified protein was frozen immediately in liquid nitrogen and stored in 10 μl aliquots at -80 ℃.

### 2.4.3 Optimization of JAIRD1C's enzymatic assays and mass spectrometry analysis

Purified JARID1 proteins were subjected to *in vitro* enzymes activity assays, where the

enzyme was incubated with histone peptide substrates in optimized buffer conditions as described

(Christensen et al., 2007; Klose et al., 2007; Li et al., 2008; Tahiliani et al., 2007), containing

factors essential for JmjC catalysis such as iron, 2-OG and ascorbic acid. We measured activity

assays with histone peptides by Matrix Assisted Laser Desorption Ionization – Time of Fly

(MALDI-TOF) mass spectrometry using a Bruker Ultra FlexII TOF/TOF instrument (Biochemistry

Department). Reactions were started by adding peptide to pre-incubated JARID1C with cofactors at

37 ℃. At different time points, 2 μl of sample was withdrawn and mixed with 5 μl of 0.1% Tri-

fluoric acid (TFA) to stop the reaction. These mixtures representing reaction progression at each

time points were spotted on a stainless steel mass spectrometry sample plate and mixed with equal

amounts of 60~70% ACAN dissolved in isopropanol. After the sample dried out, we mounted the

sample plate in the MALDI-TOF instrument and recorded the sample mass spectra and looked for

peptide mass changes in 14, 28, or 42 Daltons before and after reaction, corresponding to mono-,

di-, and tri-methyl groups. By plotting the relative abundance of each peptide species

(corresponding peak area of a an peptide species divided by total peaks area) versus reaction time,

we could determine the time needed for 50% substrate conversion to product, an indicator for

robustness of enzymatic activity. By integrating the relative abundance of each species into a single

line plot (average number of methyl group (Me) removed per peptide), we could present the

reaction course in a straightforward manner. By measuring activity in this way in different pH

buffer and NaCl concentration, we determined the optimal reaction condition for JARID1C (1-839)

to be 50 mM MES pH 6.5/6.6, 50 mM NaCl. Specifically, when we tested different buffer pH values, a mixture of NaAcetate, MES, and HEPES buffer (33.3 mM each, 100 mM total) was adjusted to different pH in the range from 6.0 to 7.0 (compatible with the pKa values of these compounds so that buffering power is retained). This allows us to eliminate the effects of different types buffering molecules and look at only the effect of changes in pH.

2.4.4 Preparation of (1-21) H3K4me3H3K9me3 doubly-methylated peptide

A doubly methylated H3K4me3K9me3 peptide was needed for comparison with H3K4me3 of the same length. Previous studies in the lab showed that (1-24) H3K4me3K9me3 doubly methylated peptide does not work well with mass spectrometry, and commercial sources of such peptide are of less than optimal purity. Since (1-21) H3K4me3 peptides as well as an H3K9 methyltransferase dim-5 was readily available in the lab, we use dim-5 and S-Adenosyl methionine (SAM) to methylate the K9 of (1-21) H3K4me3 peptide using published protocol (Zhang et al., 2002) and obtained the (1-21) H3K4me3K9me3 peptide. The methylation reaction was terminated by adding MES pH 6.5 buffer, which is compatible with subsequent demethylation reaction. The addition of tri-methyl group on K9 is shown using mass spectrometry. As a control, S-Adenosyl-L-homocysteine (SAH) was used instead of SAM to treat the (1-21) H3K4me3 peptide under the same reaction condition and no methylation on K9 was observed; and this peptide was used to compare with the doubly methylated peptide for JARID1C's demethylation activity. In neither case the peptides were purified because the methylation reaction buffer condition was compatible with demethylation reaction and addition of H3K9me3 was observed (not in control).

A

B
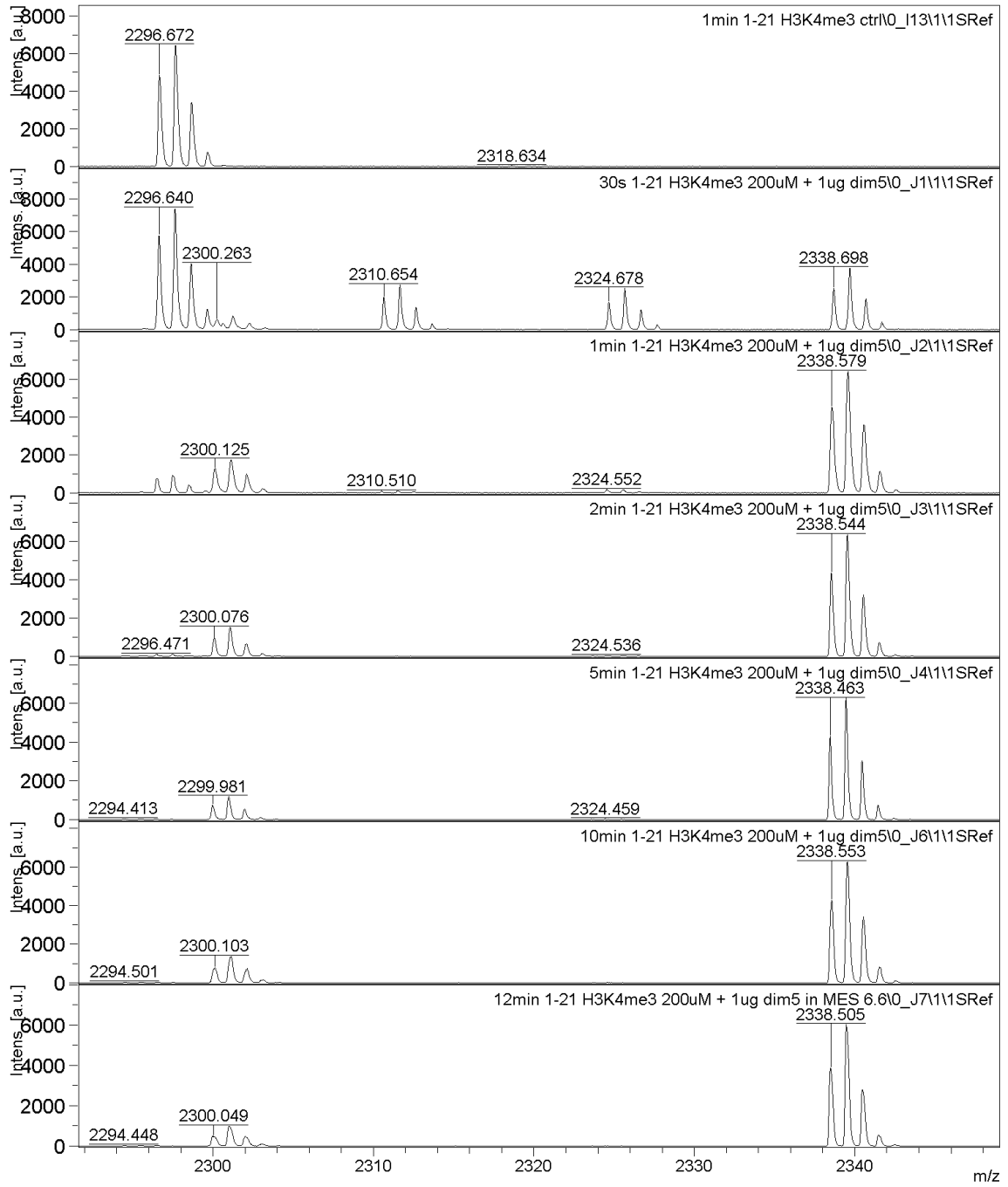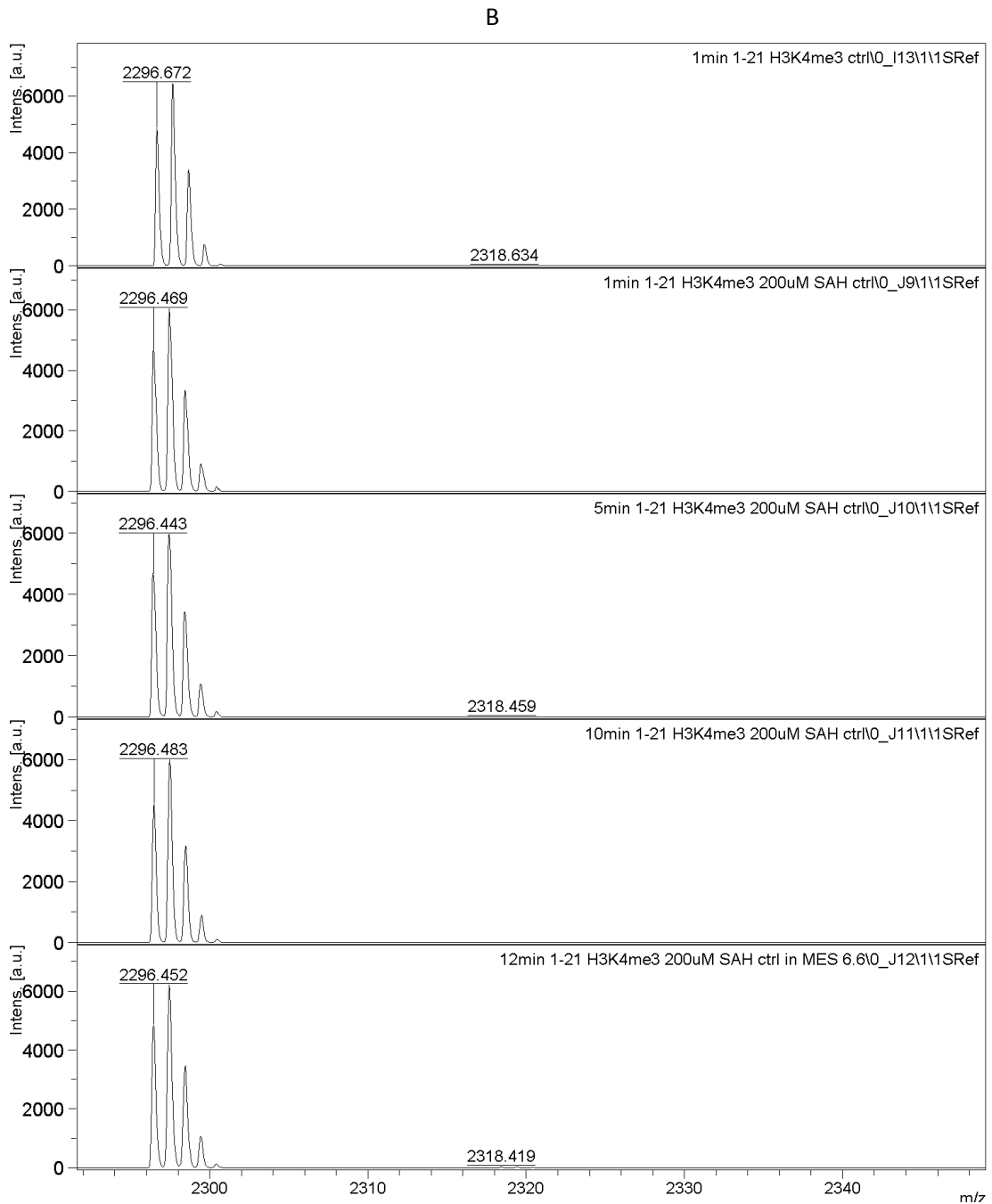
Figure 10: Mass Spectrum of 1-21 H3K4me3 peptide methylated at K9 by Dim-5 methyltransferase. SAM was a cofactor (methyl donor) utilized by Dim-5, and SAH is the by-product and inhibits methylation. 1-21 H3K4me3 peptide was converted to doubly methylated H3K4me3K9me3 in the presence of SAM as the mass increases by 52 Daltons, which is equivalent to three methyl groups (A). In the presence of SAH, however, it is not methylated as the mass did not change (B).

2.4.5 Purification of formaldehyde dehydrogenase (FDH)

A His-tagged FDH construct in *E. coli* strain BH249 cells was cultured in LB media at 37 ℃ to OD_{600} between 0.6 and 1.0 and the temperature was set at 16 ℃. After 1~2 hours when the culture liquid temperature dropped to 16 ℃, 0.4 mM of IPTG was added to induce the expression of recombinant protein overnight. Bacterial culture was centrifuged to collect the cells, which were washed and resuspended with 20 mM HEPES pH 7.0, 300 mM NaCl. Protease inhibitor (PMSF) was added to the cell resuspension and cells were broken by passing through a French Press twice on ice. The cell lysate was centrifuged for at least 1.5 hours and filtered by a 4 micron filter before loading on the first column. Affinity columns (Nickel), ion-exchange columns (Q), and gel filtration (S75) columns were used to purify FDH to over 99% purity based on SDS-PAGE (Figure 11). The final FDH protein was concentrated using a 30 kd MWCO Millipore concentrator to about 10mg/ml (determined by OD_{280} reading). Purified protein was frozen immediately in liquid nitrogen and stored in 30 μl aliquots at -80 ℃.

2.4.6 Calibration of fluorescence intensity and formaldehyde concentration and kinetics data measurement

Each round of FDH-JARID1C coupled reaction releases one molecule of methyl group which is converted to formate, while one molecule of APAD+ is converted to APADH. Therefore there is a 1:1:1 molar ratio among the amounts of APADH, formaldehyde released, and peptides converted. By saturating the fluorescence signal (RFU) under known concentrations of formaldehyde standards with FDH and APAD+, we could calculate a conversion factor between RFU and formaldehyde concentration (equivalent to converted peptide concentration), and therefore were able to derive velocity of demethylation.

For kinetic measurements, reactions were performed in 40 μl volume and in the same condition as those measured by mass spectrometry, except that FDH related components were

added (2mM Ascorbic Acid, 1mM 2-OG, 0.1 mM Fe(II), 20mM MES pH 6.5, 50 mM NaCl, 20 μM H3K4me3K9me2, 0.5 μM JARID1C (1-839), 0.6 mM APAD+, 10 μg FDH (3μM)). A 37˚C pre-incubation of JARID1C, FDH, with all other cofactors and buffers for 15 minutes was followed by addition of peptide and APAD+ to initiate the reaction. A 10 minute reaction at 37˚C was monitored on a BioTek Synergy™ 4 Hybrid Microplate Reader. Each set of experiments were done at varying concentrations of peptide substrate (from 0 to 50 μM) or 2-OG (from 0 to 100 μM), and fluorescence intensity for each corresponding peptide/2-OG conentration were plotted against time to give a series of curves (Figure 12). The curve for zero concentrations were set as background and subtracted from all other fluorescence curves at different substrate/2-OG concentrations. The slopes of initial linear range of each adjusted fluorescence curve was calculated and converted to initial velocities. Experiments were done in triplicates and initial velocities were then plotted against peptide/2-OG concentrations and fit with the Michaelis-Menten equation using GraphPad Prism 5.0.

2.4.7 IC_{50} measurements using fluorescence-based formaldehyde release assay

To test the inhibitory effect of (R/S) 2-HG and NOG on JARID1C demethylation, we added varying concentrations of these inhibitors and measured the decrease in v_{max} using the formaldehyde release assay. Experimental set up was the same as above except that inhibitors were added simultaneously with 2-OG and incubated with enzyme at 37 ℃ for 15 minutes. Experiments were done in duplicates or triplicates and the v_{max} at each concentration of inhibitor was plotted against the logarithm of inhibitor concentration. IC_{50} values were calculated in GraphPad Prism 5.0 using a dose (log(inhibitor))-response inhibition model.

Affinity (Ni) column fractions

66 kDa
42 kDa
26 kDa

Ion-exchange (Q) column fractions

66 kDa
42 kDa
26 kDa

Gel filtration (S200) factions

66 kDa
42 kDa
26 kDa
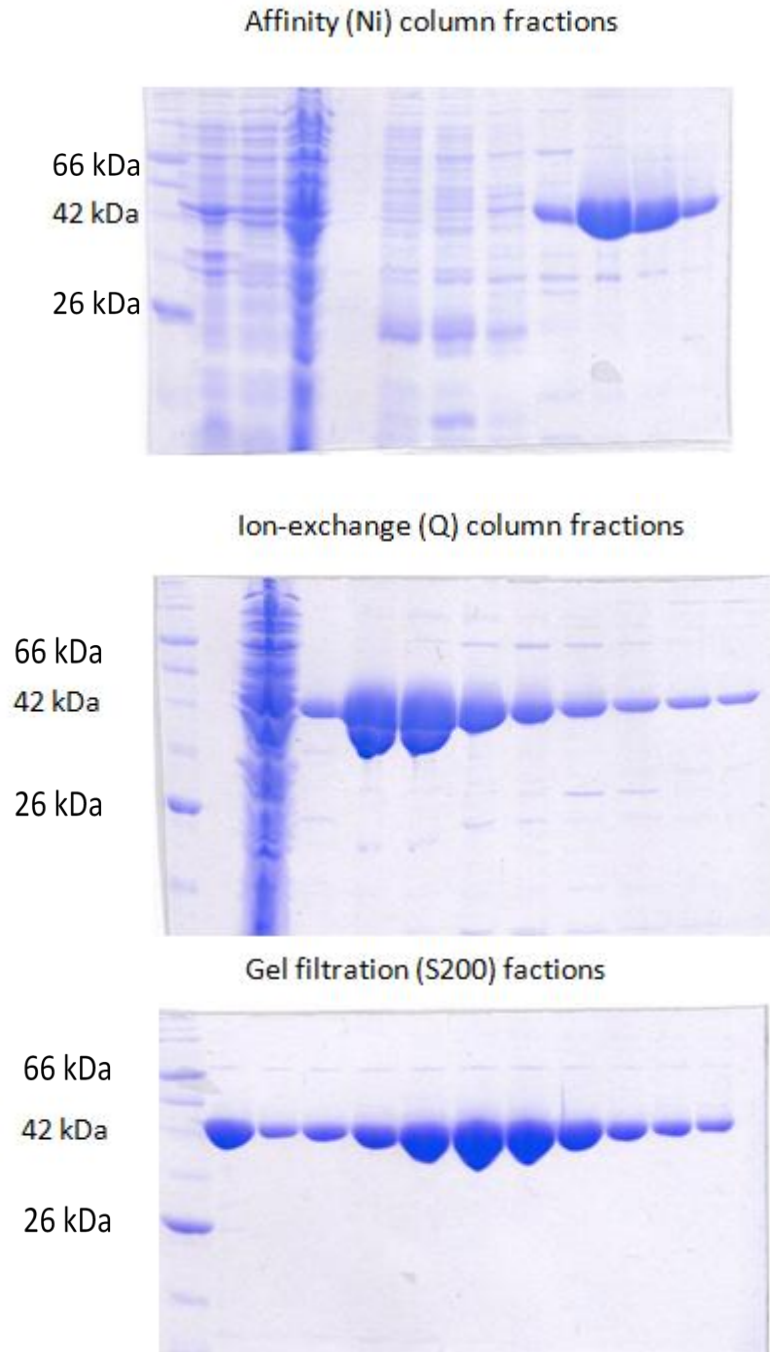
Figure 11: Purification of FDH. Shown are fractions collected from each step of column purification, run on SDS-PAGE. The first lane on the left of each gel is NEB Broad Range Protein Marker (2 - 212KDa). The size of FDH is about 42 kDa and the protein was purified to homogeneity. The buffers used and detail protocol are in materials & methods.
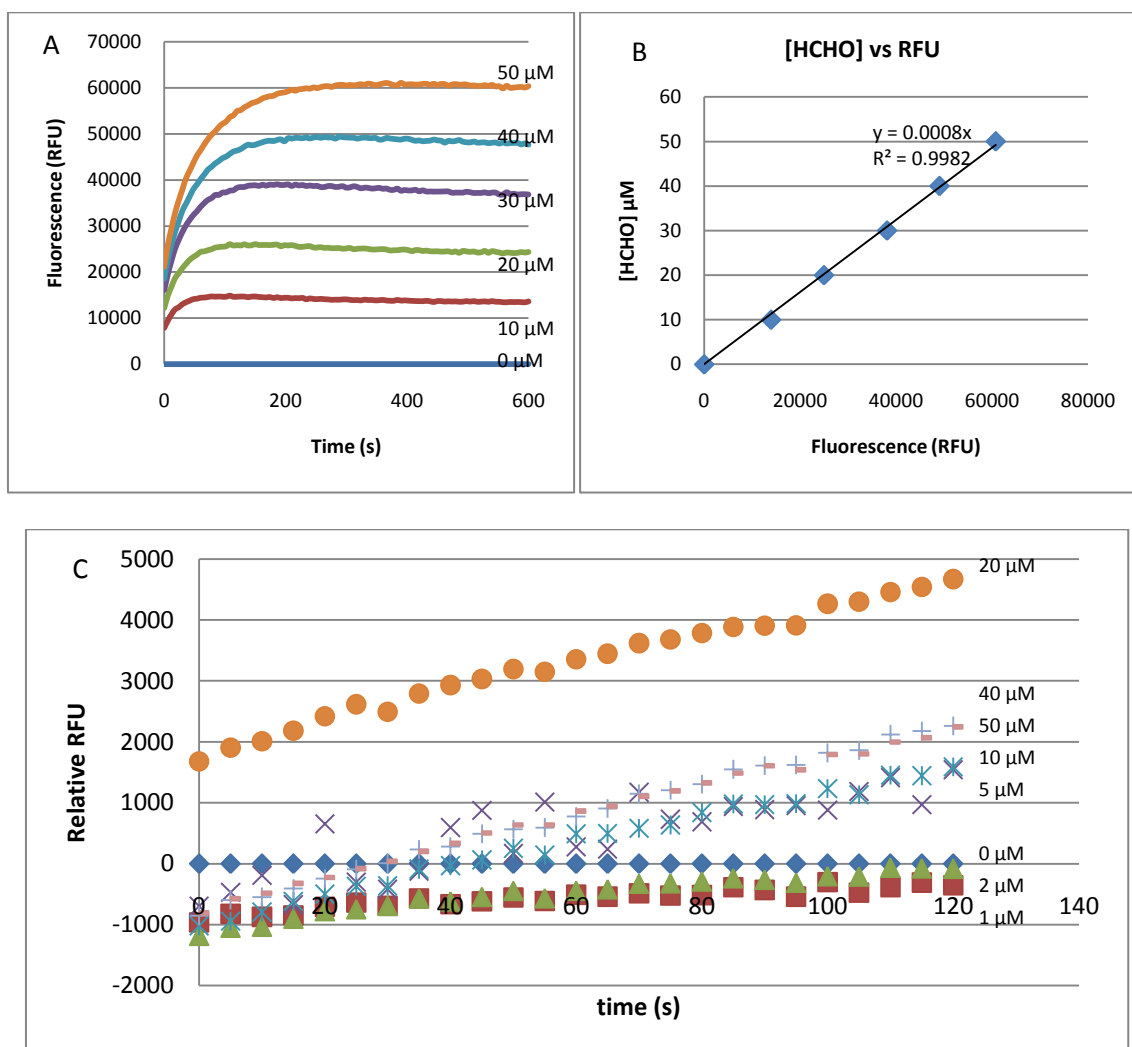
Figure 12: Calibration of fluorescence intensity and kinetics data analysis. A: time course of fluorescence intensity as FDH converts different concentrations of formaldehyde. B: formaldehyde concentration was plotted against the average fluorescence intensity in corresponding curves in A after saturation. The slope of this line is the conversion factor of RFU to formaldehyde concentration (equivalent to converted peptide concentration). C: A sample kinetics data showing different slopes corresponds to different peptide concentration. These slopes were converted to velocities using the above conversion factor and plotted against peptide concentration and fitted with the Michaelis-Menten equation.

2.4.8 Gel-shift assay for JARID1C (1-839) with ARID-binding oligonucleotides

Plus and minus strands of DNA sequences that were shown to bind ARID domains of Dead-ringer and JARIDB were obtained from Sigma and annealed into oligonucleotide duplexes in 20 mM Tris 7.5, 50 mM NaCl. (Dead-ringer ARID-binding or "AT-rich" DNA: 5'-CCTGTATTGATGTGG-3', 3'-GGACATAACTACACC-5'; JARID1A ARID-binding or GC-rich DNA: 5'-GCTCCCGCCCCACG-3', 3'-CGAGGGCGGGGTGC-5'). Different concentrations of JARID1C (1-839) protein (0 to 25 μM) was incubated with the 5 μM of DNA duplex in20 mM MES pH 6.5, 50mM NaCl, 10% glycerol. The reaction was kept on ice for 30 minutes and the samples were loaded on a 1.2% agarose gel. Electrophoresis was performed in 0.5X TBE buffer at 50 V for 10 ~ 40min.

2.4.9 Activity assays of JARID1A with ARID-binding DNA sequences

Activity assays were performed as mentioned in 2.4.3, except that the ARID-binding DNA sequences in 2.4.8 was added to the reactions with a 1:10 molar ratio of DNA to protein (0.25 μM enzyme, 2.5 μM DNA, 20 μM (1-24) H3K4me3K9me2). The average number of methyl group removed per peptide was calculated and plotted. Comparison was made between reaction without DNA, with Dead-ringer ARID-binding DNA, and with JARID1B ARID-binding DNA (Figure 7C).

Chapter III: Binding properties of JARID1C PHD domains

3.1 Introduction

3.1.1 Background on PHD domains

The PHD (Plant-homeodomain) domain is a type of Zn finger motif that in most cases bears the C4HC3 architecture and coordinates 2 or 3 Zinc ions (Baker et al., 2008). A large number of PHD domain containing proteins exist across species from yeast human (Bienz, 2006), many of which associate with the chromatin. Thus the PHD was initially thought to media the interaction (Baker et al., 2008). This is indeed the case as many PHD domains were found to bind histones. In fact, it has been suggested that the PHD histone-binders are like other chromatin binding domains (e.g., chromodomains, bromodomains, Tudor, etc) that have their own set of rules for interpretation of modifications in the chromatin (Taverna et al., 2007). Besides histone binders, a subset of PHD domains were found to act as an E3 ubiquitin ligase, though some were found in mitochondria and not associated with the chromatin (Lu et al., 2002; Uchida et al., 2004; Yonashiro et al., 2006).

3.1.2 Known PHD binders to the chromatin and JARID1 PHD domains

The majority of known PHD domains that bind to histones recognize the modification states of H3K4. There are two major groups: H3K4me0 binders, and H3K4me3 binders. These two groups of PHDs interact with the H3 tail using different mechanisms. The H3K4me3 binders use an aromatic cage to embrace the tri-methyl H3K4 and the H3K4me0 binders utilize a series of surface interactions to bind to the peptide(Taverna et al., 2007). Some H3K4me3 binders such as BPTF and ING2 recognize H3R2 at the same time (Li et al., 2006; Shi et al., 2006). Based on sequence alignment of some examples of PHD domains, JARID1C's PHD1 and PHD2 lack several key residues that would form the aromatic cage for H3K4me3, and are more similar to H3K4me0-binding PHDs in terms of amino acid identities at positions that interact with the peptide (Figure

13). In addition, JARID1C PHD2 has an extra stretch of amino acids preceding the last two conserved cysteine residues, different from the rest of the PHD domains. This information suggests that JARID1C's PHD1 and PHD2 are unlikely to bind H3K4me3.

Besides these H3K4 interacting PHDs, H3K9 was also reported to associate with several PHD domains, which are mainly in the JARID1 family. As stated in Chapter I, JARID1C PHD1 and Lid2 PHD2 were the only two PHD domains reported to bind H3K9me3 through pull-down assays. In comparison, PHD-H3K4 binding was supported by a wide range of methods (pull-down, Isothermal titration calorimetry (ITC), and crystal structures, etc) by many groups. Our biochemical results in Chapter II did not provide a direct support on JARID1C PHD1's binding to H3K9me3. Also, no published construct information was available regarding Lid2 PHD2. Therefore, we decided to test JARID1C PHD1 binding to H3K9me3 using both pull-down and ITC assays. These results would complement our enzymatic analysis and help explain JAIRD1C's catalytic behavior on H3K4me3 in response to H3K9 methylation.

H3K4me0 Binders

```
hBHC80/484-543       : GDIHEDFCS--VCRK-----SGQLLMCDT--CS-
hCHD4_PHD2/367-425   : TD-HQDYCE--VCQQG-----GEIILCDT--CP-
hJARID1A_PHD1/289-350: NFVDLYVCM--FCGRG--NNEDKLLLCDG--CD-
```

H3K4me3 Binders

```
hJARID1A_PHD3/1603-1660: SDDENAVCAAQNCQRP-CKDKVDWVQCDGG-CD-
hBPTF/2863-2925        : KKDTKLYC---ICKTP-YDESKFYIGCDR--CQ-
hING1/349-410          : DPNEPTYC---LCNQVSYGE---MIGCDNDECPI
PHF8/1-60              : MASVPVYC---LCRLP-YDVTRFMIECDM--C--
```

Others

```
spLID2_PHD2/1089-1150  : TIRKKKGCIFCFCRLPESG---VMIECEI--CH-
hJARID1C_PHD1/320-382  : QFIESYVCR--MCSRG--DEDDKLLLCDG--CD-
hJARID1C_PHD2/1181-1255: ASSTTSICC---VCGQVLAG--AGALQCDL--CQ-
```

```
RVYHLDCLDPPLKTIP---------------KGMWICPRCQDQMLKKEEAI
RAYHMVCLDPDMEKAP---------------EGKWSCPHCEKEGIQWEAKE
DSYHTFCLIPPLPDVP---------------KGDWRCPKCVAEECSKPRE-
```

```
EWFHQVCVGVSPEMAE---------------NEDYICINCAK---------
NWYHGRCVGILQSEAEL--------------IDEYVCPQCQSTEDAMTVL-
EWFHFSCVGLNHKP-----------------KGKWYCPKCRGENEKTMDKA
DWFHGSCVGVEEEKAAD--------------IDLYHCPNCEVLHGPS
```

```
EWYHAKCLKMSKKKLRQ--------------DEKFTCPICDYRVEIPR--
DNYHIFCLLPPLPEIP---------------KGVWRCPKCVMAECKRPPEA
DWFHGRCVSVPRLLSSPRPNPTSSPLLAWWEWDTKFLCPLCMRS-RRPRL-
```

Figure 13: Sequence alignment of known examples of PHD domains that bind to histone H3. The majority recognized H3K4me3 or H3K4me0. The "h" in front of gene names means human.
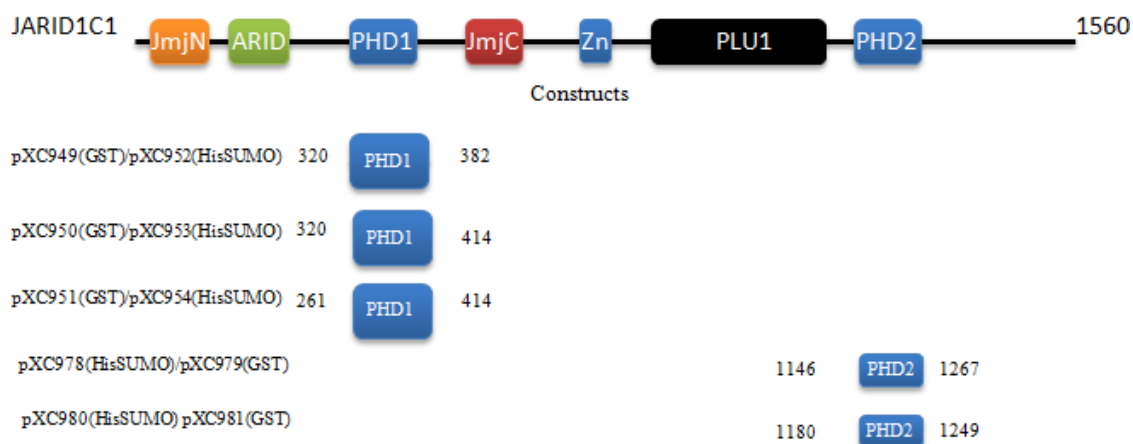
Figure 14: Summary of JARID1C PHD constructs. The pXC numbers on the left were plasmid identification numbers in the lab (pXC949, pXC952, etc). The numbers preceding and following each PHD domain cartoons are amino acid residue numbers of JARID1C.

Domain	Range	Forward Primer (BamHI-AseI)	Reverse Primer (EcoRI-NotI)
PHD1	Q320-A382	ATGGATCCATTAATCAGTTT ATTGAGTCATATGTCTGC	ATGAATTCGCGGCCGCTCAGGC TTCTGGGGGCCGCTTACACTC
PHD1	Q320-H414	ATGGATCCATTAATCAGTTT ATTGAGTCATATGTCTGC	ATGAATTCGCGGCCGCTCAATG CACGGGCATGTTGAAGTAGTC
PHD1	D261-H414	ATGGATCCATTAATGATAA GGAGGGGCCTGAGTGTCCC	ATGAATTCGCGGCCGCTCAATG CACGGGCATGTTGAAGTAGTC
PHD2	A1181-R1249	ATGGATCCATTAATGCCTCC TCTACAACCTCTATC	ATGAATTCGTCGACTCAGCGCA TACACAGTGGACACAG
PHD2	V1146-R1267	ATGGATCCATTAATGTGATC GTGGCCTTCAAGGAG	ATGAATTCGTCGACTCATCTCT GCAGGGCTACCAGCAG

Table 1: Summary of JARID1 PHD constructs. DNA sequences are from 5' end to 3' end.

3.2 Results

3.2.1 Purification of JARID1C PHD domains

Yang Shi's group (Iwase et al., 2007) showed that JARID1C PHD1 binds to H3K9me3 yet they did not detect JARID1C PHD2 binding to a histone peptide in the same experimental system. Based on PHD domain sequence alignment and JARID1C PHD constructs made by Yang Shi's group, we made the following PHD1 and PHD2 constructs (Figure 14). Among these constructs, His-SUMO-PHD1 (320-382) and His-SUMO-PHD2 (1181-1249) were most soluble and contained less aggregation or degradation during purification (Figure 14), which allowed us to purify large amounts of protein in high purity (Figure 15). This makes them suitable for Pull-down and particularly ITC experiments.

3.2.2 Pull-down and ITC results of JARID1C PHD1 and PHD2

The best expressed and purified PHD1 and PHD2 proteins were used in pull-down and ITC experiments. We only observed binding when the binding buffer contained NaCl concentrations as low as 50 mM. Under this condition, both PHD1 and PHD2 exhibited a similar pattern of binding to the various H3 methylation marks we tested (Figure 16). H3K4 tri-methylation appeared to disrupt binding, yet H3K9me2 and H3R2me2 (A or S, Figure 16) appeared to enhance the binding.
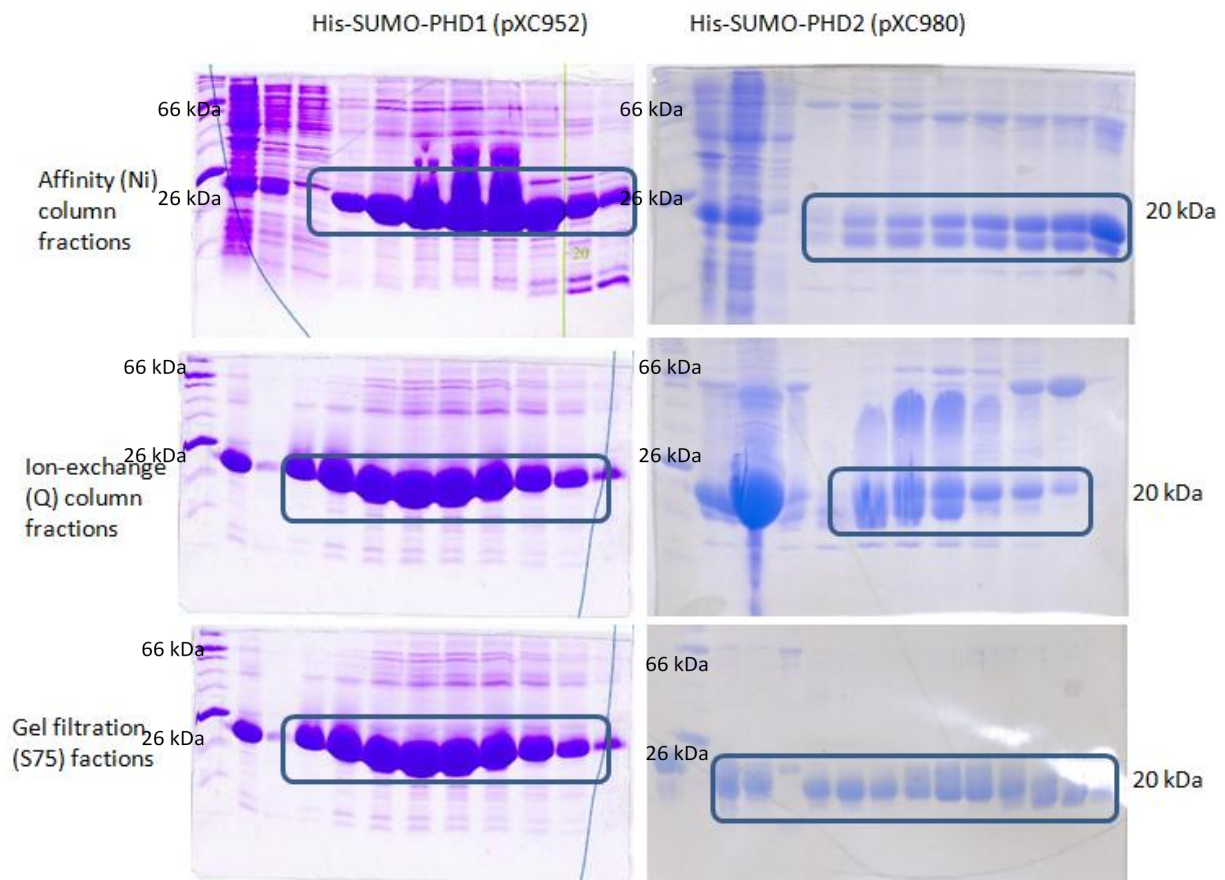
Figure 15: Purification of JARID1C PHD1 and PHD2. Shown are representative gels corresponds to fractions collected during purification after going through respective columns. The target proteins were marked with round corner boxes. The first lane on the left of each gel is NEB Broad Range Protein Marker (2 - 212KDa), showing the correct size of the target proteins (20 kDa). The buffers used and detail protocol are in materials & methods.
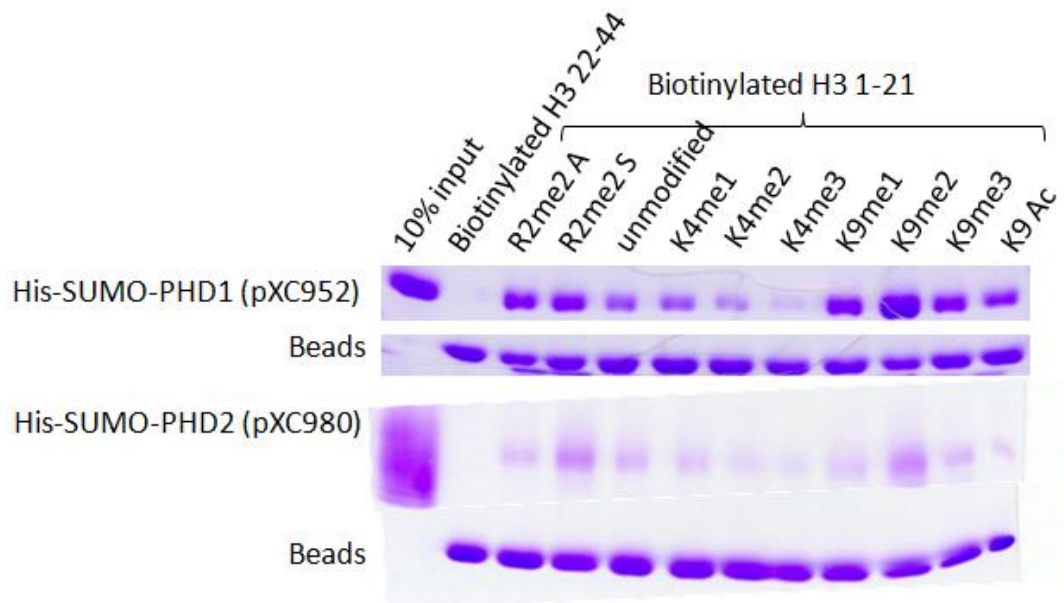
Figure 16: Pull down of JARID1C PHD1 and PHD2. Similar pattern of binding was observed for PHD1 and PHD2 where H3K4me3 appears to disrupt protein's association with peptide. The "A" and "S" following H3R2me2 mean "asymmetric" and "symmetric" di-methylation.

However, published PHD binding data were usually performed in salt concentration of above 50 mM such as 150 mM (Lan et al., 2007; Matthews et al., 2007; Wang et al., 2009). Based on our results, at 50 mM we could not rule out the possibility of non-specific interaction. ITC experiment was performed for PHD1 in order to verify our observation in pull-down assays (Figure 18). Not suprisingly, at the same NaCl concentration (50 mM), we observed that neither H3K4 methylation nor H3K9 methylation caused any meaningful changes in binding affinity (Figure 17), because in all cases the estimated dissociation constants (k_d) were very large (serveral hundred micromolar to millimolar range) compared to other known PHD-H3 interactions (from a few to less than a hundred micromolar) (Taverna et al., 2007). Therefore, we did not have sufficient evidence to claim any specific interaction between JARID1C PHD domains and histone H3. Given the behavior of PHD2 was similar to PHD1 based on pull-down assays, we did not proceed to test PHD2 binding using ITC.
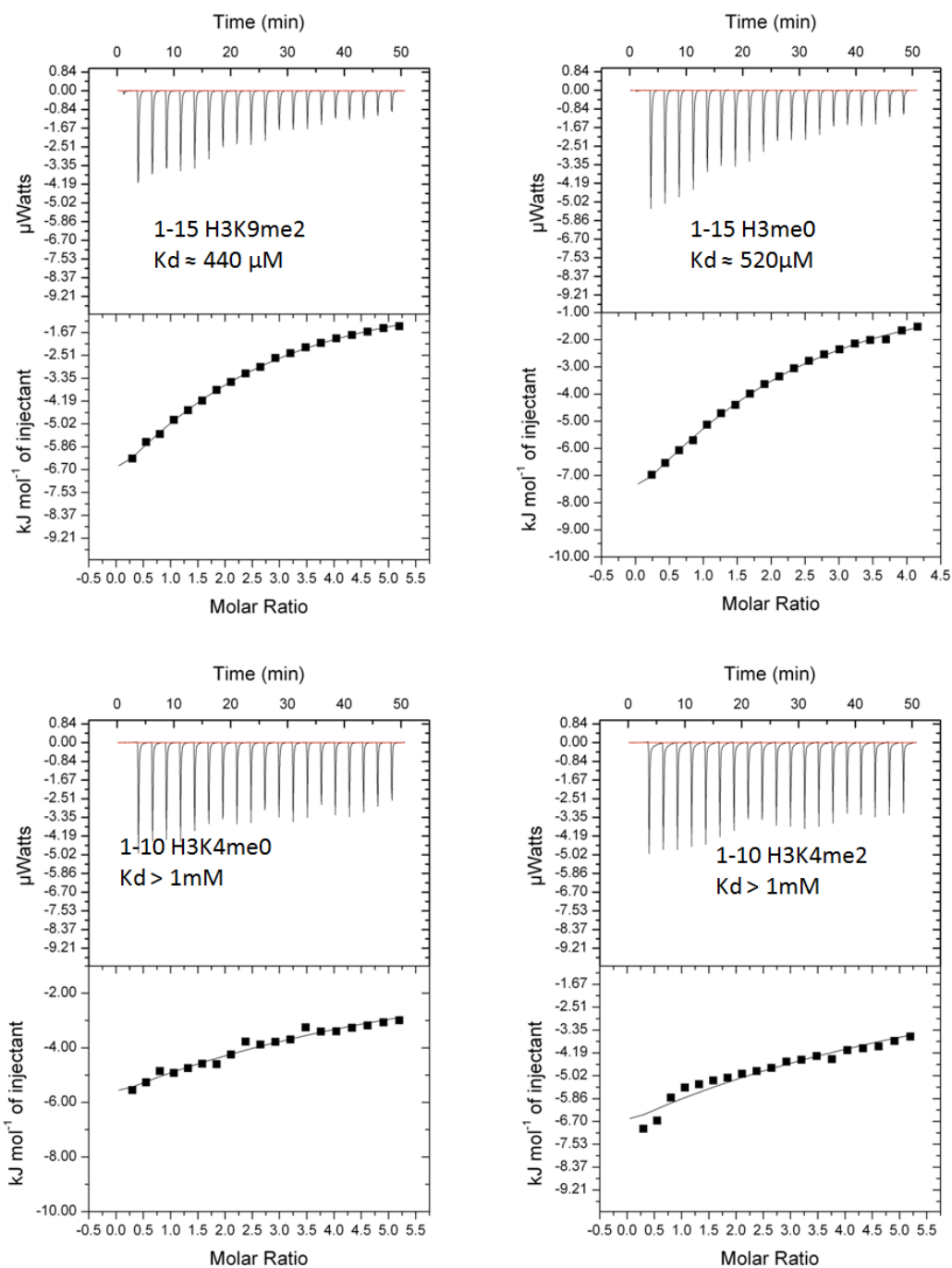
Figure 17: ITC of JARID1C PHD1 with various H3 peptides. Estimated dissociation constants for each peptide were shown. The titration was done using above 200 µM of protein and as high as 5 mM of peptide. The lack of top and bottom of the sigmoid curve indicates weak binding, as suggested by the large k_d values.

3.3 Discussion

Our pull-down and ITC results indicated that JARID1C PHD domains bind to histone peptides at unmodified H3K4 with very low affinities and behave close to non-specific interactions, which are not comparable to previously identified H3K4me0/3-binding PHD domains, whose k_d values fall in low micromolar ranges (Ruthenburg et al., 2007). While additional experimental approaches could be tried to verify this observation (such as fluorescence polarization or NMR titration), this phenomenon agrees with our data in Chapter II where the JmjC domain activity on H3K4me3 is independent of H3K9 methylation status, assuming a relatively constrained spatial relationship between PHD1 and JmjC. This would suggest a separate role for the PHD domains, possibly involving recruiting proteins other than histone H3; and the PHF8 model of cross talk between H3K4 and H3K9 is not sufficient to explain how JARID1C interacts with these two marks. As mentioned in conclusion and future directions, more detailed biochemical data such as a crystal structure of JARID1C will play a critical role in answering the above questions.

3.4 Materials and Methods

3.4.1 Cloning of JARID1C PHD constructs for subsequent analysis

Different length PHD1 and PHD2 domains of JARID1C were cloned using various primers (Table 1). The PHD domains were PCR amplified from commercial cDNA source (ATCC) using VENT or Phusion DNA polymerase (NEB) and with primers bearing restriction sites. PCR fragments were digested with corresponding restriction enzymes and cloned into expression vectors. We mainly used plasmids that provide His-SUMO and GST tags as well as protease cleavage sites for the removal of tags, such as the pET system and PEGX plasmids. We transformed these recombinant constructs into *E. coli* strain BL21DE3 for protein expression. This process was repeated to generate all versions of the PHD domains proteins (Figure 14). Cloned constructs were

confirmed by sequencing of plasmid DNA and induction of the desired protein visualized by

sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie staining.

3.4.2 Purification of JARID1C PHD domains

Out of the many constructs, pXC952 (His-SUMO PHD1) and pXC980 (His-SUMO PHD2)

were gave rise to most soluble protein and were purified in large amounts. Typically, 2 to 6 liters of

LB media was used to grow bacteria from starting culture, first at 37 $^\circ$C to an OD_{600} value between

0.6 and 1, at which point 1 mM of MgCl2 and 1 mM of ZnCl2 were added and the temperature was

set to 16 ℃. After 1~2 hours, when the liquid culture temperature dropped to 16 ℃, 0.4 mM of

IPTG was added to induce the expression of recombinant protein overnight. The bacterial culture

was centrifuged to collect the cells, which were washed and resuspended with nickel column buffer

A (20 mM NaHPO4, 5% glycerol, 500 mM NaCl). Protease inhibitor (PMSF) was added to the cell

resuspension and cells were broken by passing through a French Press twice on ice. The cell lysate

was centrifuged for at least 1.5 hours and filtered by a 4 micron filter before loading on the first

column. Affinity columns (nickel), ion-exchange columns (Q), and gel filtration (sizing) columns

were used to achieve the level of purity we want for subsequent analysis. The His-SUMO tag was

retained as previous studies in the lab show that PHD domains with His-SUMO tags behave well in

binding assays (Lan et al., 2007). The final PHD1 and PHD2 proteins were eluted off the gel

filtration column in 20 mM Tris 8.0, 5% glycerol (No glycerol for ITC experiments), and 150 mM

NaCl and concentrated using a 10kd MWCO Millipore concentrator to at least 10 mg/ml,

determined by OD_{280} reading. Purified protein was stored on ice and used for experiments as soon

as possible (within 1 week).

3.4.3 In vitro binding assays (Pull-down) of JARID1C PHD

Biotinylated histone peptides (0.5~1 μg) were incubated with 5~10 μg of purified recombinant His-SUMO tagged JARID1C PHD1 or PHD2 for 4~12 h at 4 °C in binding buffer (20 mM Tris pH 8.0, 50 mM NaCl). Binding was not observed at an empirical NaCl concentration of 150 mM but only detected by Coomassie blue staining when NaCl concentration dropped to 50 mM. After incubation of peptide and protein, Streptavidin beads (Upstate 16-126) were washed three times and added to the mixture. The mixture containing peptide, protein, and beads was gently rotated for 1 h, after which beads were centrifuged and supernatant was removed. The pull-down was subjected to Coomassie blue staining and boiled for 5 minutes. Samples were run on an SDS-PAGE for analysis.

3.4.4 Isothermal titration calorimetry (ITC)

His-SUMO-tagged fusion of the JARID1C PHD1 was eluted and concentrated in 20mM Tris pH 8.0 buffer, 50 mM NaCl, 0.5 mM TCEP from gel-filtration chromatography. Extensively lyophilized H3 1–10/15 peptides were dissolved in the same buffer. ITC measurements were carried out from 100–250 μM protein concentration, with 1.5–5 mM peptide concentration, on a GE MicroCal Auto-ITC$_{200}$ instrument at 25 °C. Reference titrations of protein into buffer were performed to control for heat of dilution and non-specific binding. Binding constants were calculated by fitting the data using the ITC data analysis module of Origin 7.0 (OriginLab Corporation). We compared the binding of JARID1C PHD1 to (1-15) H3K9me0, (1-15) H3K9me3, (1-10) H3K4me0, and (1-10) H3K4me2.

Chapter IV: Crystallization Trials and Activity Screens of a Jumonji Protein MINA53

4.1 Introduction

Besides work on the JARID1C protein, crystallization trials and activity screens were performed for another Jumonji-protein, human Myc-induced nuclear antigen 53kDa (MINA53). MINA53 has been proposed to have histone lysine demethylation activity, although current results in the literature have not reached a consensus (Lu et al., 2009; Okamoto et al., 2009). The protein was also named mineral dust-induced gene (mdig) (Zhang et al., 2005), because a high mRNA level was detected in alveolar macrophages from coal miners who have been exposed to mineral dusts compared to normal subjects. Localization data by immunofluorescence reveal that MINA53 exists in the nucleus and partly in the nucleolus (Tsuneoka et al., 2002). Later studies also reveal MINA53's c-Myc-independent expression in tumors (Komiya et al., 2009). RNAi treatment of MINA53 in several cancer cell lines correlates with suppression of cell proliferation (Tsuneoka et al., 2002).

MINA53 is an Interleukin 4 (IL4) repressor and associates with its promoter in complex with NFAT and Oct transcription factors (Okamoto et al., 2009). Okamoto et al showed that by modulating IL4 expression, MINA53 controls T helper type 2 bias (which is the propensity of naïve $CD4^+$ T cells to differentiate into IL4-secreting T helper type 2 cells), highlighting that MINA53 function involves changing the transcriptional activity of downstream genes. Interestingly, MINA53 binds to IL-4 promoter and has negative correlation with IL-4 expression, while PRMT1 enhances/positively correlates with IL-4 expression via an intermediate interactor NIP45 (Fathman et al., 2010). Therefore, as a demethylase, MINA53 potentially antagonize the catalytic substrates of PRMT1, such as methylated-NIP45 or H4R3me2 (Mowen et al., 2004).

In addition, microarray data suggests that genes potentially responsible for cell transformation and tumorigenicity in cancer, such as growth factors and receptors EFGR and HGF,

are regulated by MINA53(Komiya et al., 2009). Moreover, a recent study proposed that MINA53 up-regulation in lung cancer correlates with a decrease in H3K9me3 levels and promotes ribosomal RNA expression (Lu et al., 2009). Nevertheless, no evidence for MINA53's direct substrate has been provided, leaving the possibility that MINA53 may indirectly contribute to H3K9 demethylation. MINA53 bears a JmjC domain and no other known protein domains, which indicates that potential binding partners might be required for detecting MINA53's enzymatic activity *in vivo*. *In vitro* approaches are also needed to assess whether MINA53's demethylation activity in given assay conditions is physiologically possible. Collectively, it can be hypothesized that excessive MINA53 in cancer mis-regulates the expression of downstream genes by involvement in demethylation-dependent processes.

4.2 Results

4.2.1 Protein purification

The full length human MINA53 (residues 1 - 465) was cloned and purified using the same fast protein liquid chromatography (FPLC) protocols as described in Chapters II and II. The protein was soluble in bacteria cell lysate and could be purified in large amounts and high homogeneity (Figure 18). Purified protein was used for crystallization trials and activity screens.
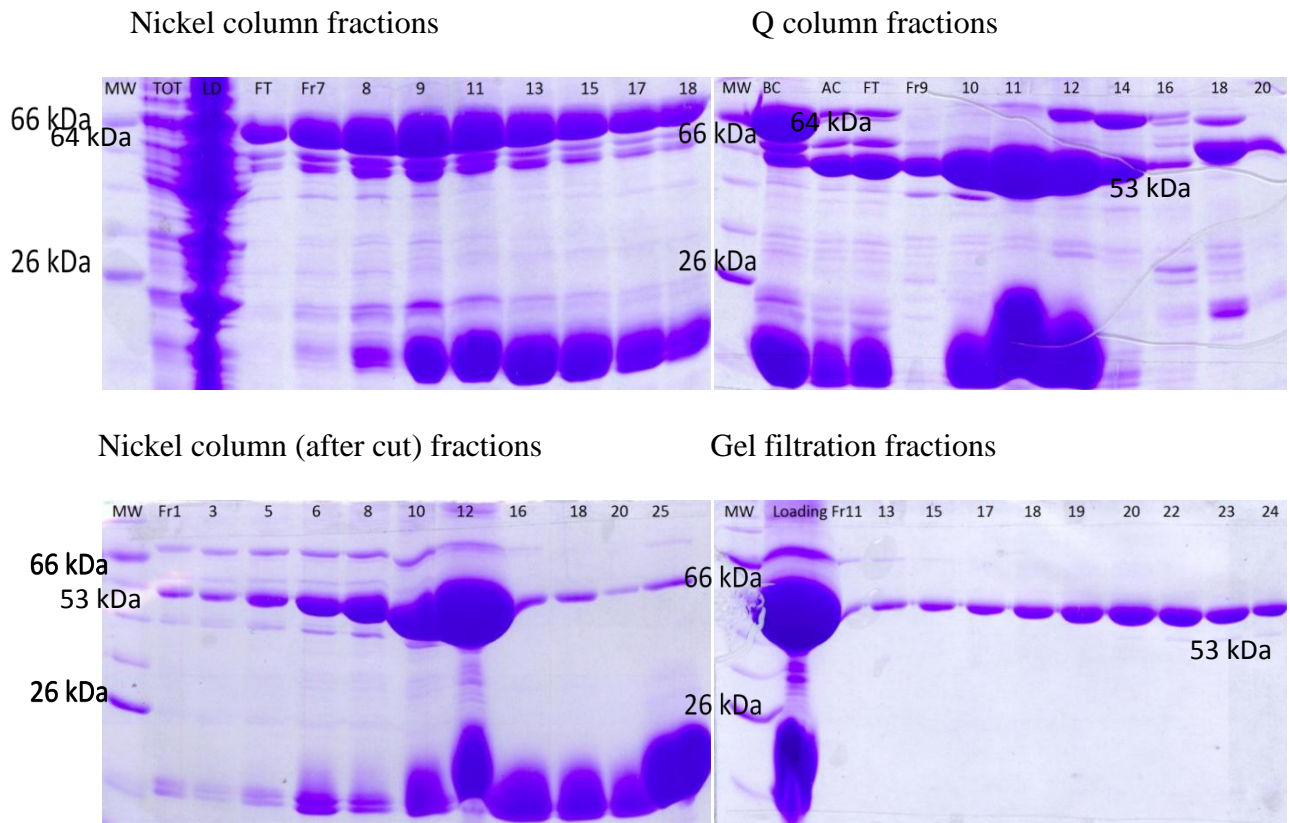
Figure 18: Purification of MINA53. Shown are representative gels of protein samples from each step of column purification. Fr and numbers: fraction number; LD: Loading sample; FT: Flow-through sample; MW: NEB Broad Range Protein Marker (2-212KDa). The second Nickel column was performed after Ulp-1 cleavage to remove the His-SUMO tag.

4.2.2 Protein crystallization

Purified MINA53 protein was crystallized under various conditions. Initial screens produced crystals on several conditions and the largest crystals in these conditions were picked up and diffracted on X-ray beam. Diffraction patterns confirmed that they are protein crystals but diffracted poorly (Table 2), and two optimization approaches were performed (Table 3). The first approach was to make solution blocks containing varied pH and salt conditions around the original screen solution (see materials & methods). These conditions were tested for crystallization. The second approach was to use protease to treat purified MINA53 during crystallization, in the hope that small fragments of the protein would be trimmed off and give rise to more homogeneous and stable fragments. In addition, different cryo-protectants for were tried. The results of both approaches and cryo-protectant screens are summarized in Table 3 and sample crystal pictures are shown in Figure 19. Despite changes in crystallization solution, crystal appearance, and cyro-protectant, no improvement of crystal quality in terms of diffraction was achieved.

4.2.3 Activity assay

Current information of MINA53, as suggested in section 4.1, did not provide clear-cut evidence regarding MIAN53's enzymatic activity and substrate specificity. Therefore, we started to screen for MINA53's activity using a number of different peptides (mostly histone peptides, as the majority of Jumonji protein substrates are histones). Among all the methylated-lysine (Table 4A) and methylated-arginine (Table 4B) peptides we tried, none of them showed mass changes caused by MINA53 and correspond to loss of a methyl group. Therefore demethylase activity was not detected for MINA53 on these histone peptides in our assay conditions. We will discuss the implications of this result in the following section.
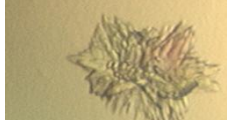
Table 2: Crystallization screen hits of MINA53
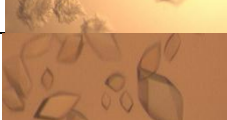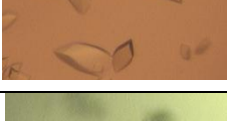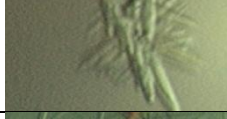
Buffer condition	Set up	Protein source	Screen	Put on beam	Data collection	Crystal
0.1M Bis-Tris pH 5.5, 2.0M AM	small sitting drop	pXC719 2nd prep, 10 mg/ml, with 2-OG	Index screen A3	yes		
0.1M Tris pH 8.5, 2.0M AM	small sitting drop	pXC719 2nd prep, 10 mg/ml, with 2-OG	Index screen A6	Yes		
2.0M AM	small sitting drop	pXC719 7th prep, 10 mg/ml, with or without 2-OG	C-main screen C8			
0.1M Na Acetate pH 4.6, 2.0M Na Formate	small sitting drop	pXC719 7th prep, 10 mg/ml, with or without 2-OG	C-main screen C10			
25% Ethylene Glycol	small sitting drop	pXC719 7th prep, 10 mg/ml, with 2-OG	C-main screen E3	yes	yes 7.5Anstrom, Space group $P3_2 2 1$	
10% EtOH, 1.5M NaCl	small sitting drop	pXC719 7th prep, 10 mg/ml, with or without 2-OG	C-main screen E8			
Wiz2 #19 2.0M Na/K Phosphate, 0.1M Phosphate-citrate 4.2	small sitting drop	pXC719 7th prep, 10 mg/ml, with or without 2-OG	C-main screen F2			
1.0M LiSO4, 0.1M Na Citrate pH 5.6, 0.5M AS	small sitting drop	pXC719 7th prep, 10 mg/ml, with or without 2-OG	C-main screen F3			
0.1M MES pH 6.5, 2.0M NaCl, 0.2M Na/K Phosphate	small sitting drop	pXC719 7th prep, 10 mg/ml, with or without 2-OG	C-main screen F9			
0.1M Na HEPES pH 7.5, 1.5M AM	small sitting drop	pXC719 7th prep, 10 mg/ml, with or without 2-OG	B-main screen A2	yes		

Table 3: Crystallization optimization of MINA53

Buffer condition	Set up	Source	Protease Digestion	Put on beam
0.1M Na Acetate pH 5.0~5.2, 1.8 ~2.1M Na Formate	small sitting drop	pXC719 7th prep, 10 or 22 mg/ml, with or without 2-OG	no	yes
0.1M BisTris pH5.6~6.4, 1.5~2.2M AM	small sitting drop	pXC719 7th prep, 10 or 22 mg/ml, with or without 2-OG	no	
0.1M BisTris / HEPES / Tris pH 5.5~8.5, 1.7~2.1M AS	large sitting drop	pXC719 8th prep, 10 mg/ml, with or without 2-OG	no	
15~30% Ethyleen Glycol	large and small sitting drop	pXC719 8th prep, 10 mg/ml, with or without 2-OG, both fresh and old	with or without 100:1 (mass ratio) trypsin / chymotrypsin digestion	
0.1M Na Acetate pH 4.8~5.8, 1.5~2.3M Na Formate	small sitting drop	pXC719 8th prep, 10 mg/ml, with or without 2-OG	with or without 100:1 (mass ratio) trypsin / chymotrypsin digestion	
0.1M Na Acetate pH 4.6~5.8, 0.1~0.6M Mg Formate	small sitting drop	pXC719 8th prep, 10 mg/ml, with or without 2-OG	with or without 100:1 (mass ratio) trypsin / chymotrypsin digestion	yes

Figure 19 Panel A

Na Formate

	1.8 M	1.9 M	2.0 M	2.1 M	

pH of 0.1M Na Acetate

10 mg/ml

22 mg/ml

4.2

4.6

5.0

5.4
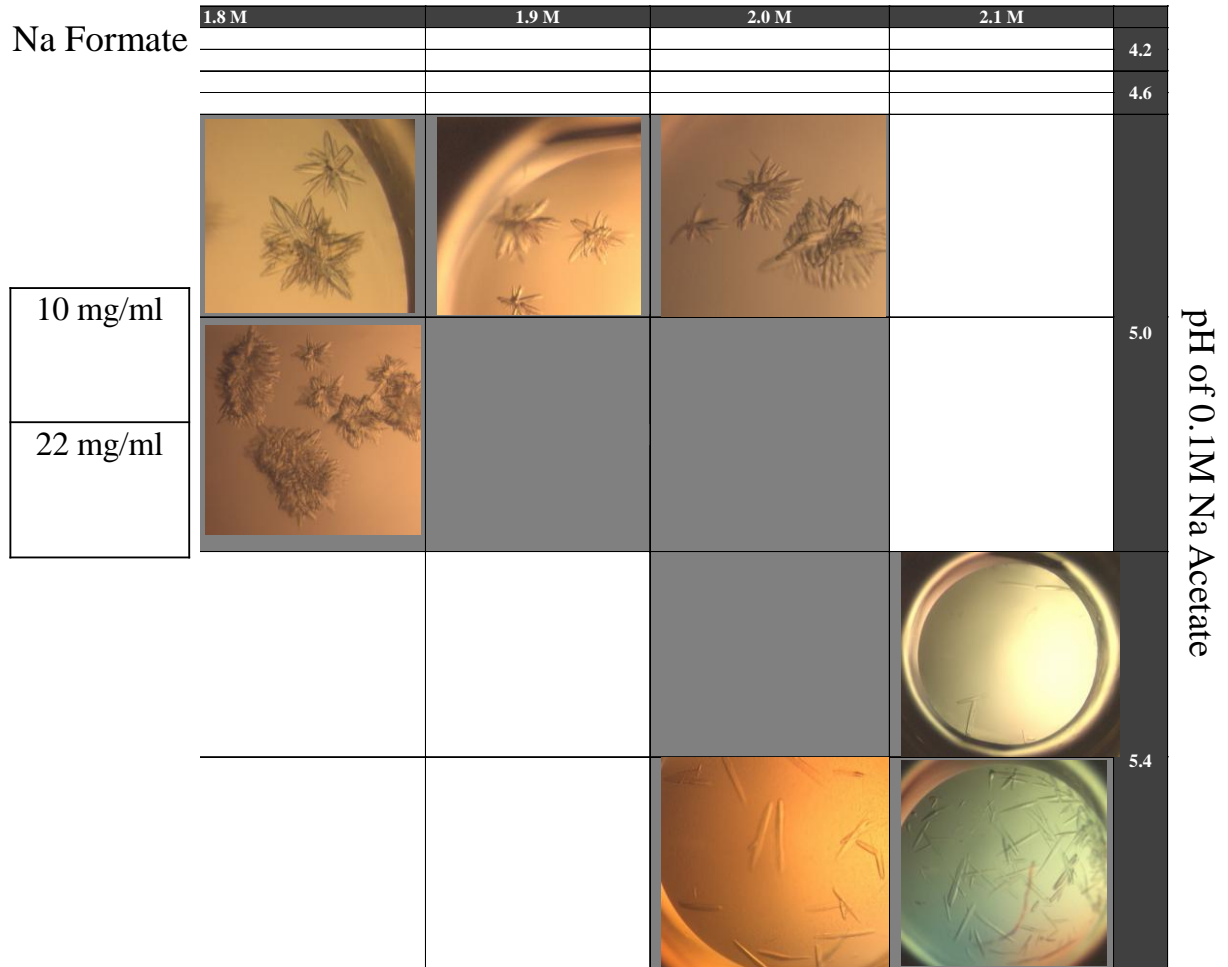
Figure 19 Panel B

Na Formate

	1.7 M	1.9 M	2.1 M	2.3 M	pH of 0.1M Na Acetate

Figure 19 Panel C
Mg Formate
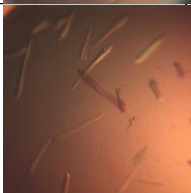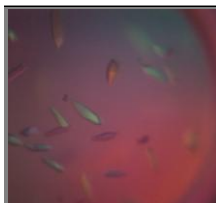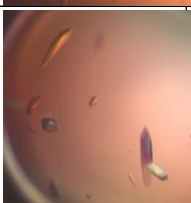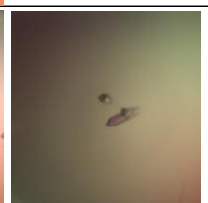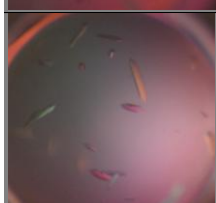
	0.1 M	0.2 M	0.3 M	0.4 M	
					4.6
					5.0
11 mg/ml no protease					5.4
11 mg/ml with 100:1 (mass ratio) trypsin					5.8

pH of 0.1M Na Acetate

Figure 19 Panel D
Mg Formate

Figure 19: Sample pictures of MINA53 crystals after optimization. The four tables in panel A, B, C, and D show different combinations of pH, salt, and protein for crystallization. Each cell in the main table represents a well on the crystallization plate. There are two protein drops in each well, as indicated by the box on the left of each table. Cells in white color correspond to wells without crystals. Cells in grey color correspond to well with crystals but pictures not taken. Cells with pictures show representative crystals.

Table 4: MINA53 activity tests using methylated peptides

Table 4A: List of methylated-lysine histone H3 peptides tested for MINA53's activity

MINA53 Substrates	1-21	1-15	1-24		10-27	1-35			21-44		
	K4 me3	K9 me2	K9 me2	K4me3 K9me2	K18 me2	K27 me1	K27 me2	K4 me3 K27 me2	K36 me1	K36 me2	K36 me3
Positive control		PHF8	PHF8	PHF8			KIAA 1718	KIAA 1718		KIAA 1718	

Table 4B: List of methylated-arginine peptides tested for MINA53's activity (The methylated arginine residues in bold and underline. At the time of writing this thesis, I did not know the degree of methylation for MeCP2 and R3* peptides or the type of methylation (symmetrical or asymmetrical) for H3/H4 peptides.)

Peptide names	Sequence
MeCP2	TG**R**GR GRPKG S
H4(1-16)R3me2	SG**R**GK GGKGL GKGGA K
H4(1-20)R3me2	SG**R**GK GGKGL GKGGA KRHRK
H3(1-18)R2me2	A**R**TKQ TARKS TGGKA PRK
H3(1-18)R2me1	A**R**TKQ TARKS TGGKA PRK
R3*	GGKGG FGGKG GFGG**R** GGFG

4.3 Discussion

A crystal structure of MINA53 (PDB ID: 2xdv) was deposited into the Protein Data Bank (T.krojer et al., 2010; to be published) at the time of this work. The structure was solved to 2.57 Angstrom and the construct used was shorter than the full-length construct that we used, lacking the 26 amino acid residues at the N-terminus. This information implies that the N-terminal of MINA53 is likely to be flexible and therefore interfere with crystallization. Such exam shows that a change in construct can dramatically alter crystal quality. In addition, the reported MINA53 protein was crystallized with other small compounds/metal ions such as 2-OG, 1, 2-ethanediol, manganese, and nickel. In our study, we had only added 2-OG during the crystallization process. Thus, it also shows that additive screens especially cofactors or analogues of the cofactors remain to be considered for crystallizing an enzyme.

The fact that our *in vitro* assays of MINA53 with histone peptides showed no demethylase activity brought up many possibilities. In one of the reported MINA53 assays (Lu et al., 2009), histones isolated from BEAS-2B cells and eukaryotically expressed MINA53 were used. Therefore, our histone peptide substrates and bacterially express protein may fail to reconstitute the necessary interaction between the real enzyme and substrate in the cell. Potentially, using insect (sf9) cells to express our recombinant MINA53 could help recover enzymes that are functional. Additionally, future experiments to systematically screen a wide range of pH buffers and types/concentrations of salts will be needed. As more functional data of MINA53 becomes available, it will help narrowing-down MIN53's targets for biochemical characterization.

4.4 Materials & Methods

4.4.1 Molecular Cloning

His-SUMO-tagged, GST-tagged, and non-tag versions of MINA53 full length constructs were made. The human MINA53 gene was PCR amplified from commercial cDNA source (ATCC) using VENT or Phusion DNA polymerase (NEB) and with primers bearing restriction sites. PCR fragments were digested with corresponding restriction enzymes and cloned into expression vectors. We transformed these recombinant constructs into *E. coli* strain BL21DE3 for protein expression. This process was repeated to generate all versions of the MINA53 proteins. Cloned constructs were confirmed by sequencing of plasmid DNA and induction of the desired protein visualized by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie staining.

4.4.2 Protein purification

Among our MINA53 constructs, pXC719 and pXC869 produced the most soluble and pure protein and were used for purification. Typically, 2 to 4liters of LB media was used to grow bacteria from starting culture, first at 37 °C to an OD_{600} value between 0.6 and 1, at which point the temperature was set to 16 °C. After 1~2 hours, when the liquid culture temperature dropped to 16 °C, 0.4 mM of IPTG was added to induce the expression of recombinant protein overnight. The bacterial culture was centrifuged to collect the cells, which were washed and resuspended with nickel column buffer A (20 mM NaHPO4, 5% glycerol, 300 mM NaCl). Protease inhibitor (PMSF) was added to the cell resuspension and cells were broken by passing through a French Press twice on ice. The cell lysate was centrifuged for at least 1.5 hours and filtered by a 4 micron filter before loading on the first column. Affinity columns (nickel), ion-exchange columns (Q), and gel filtration (sizing) columns were used to achieve the level of purity we want for subsequent analysis. The His-SUMO tag was removed by protease Ulp-1 digestion during the purification process, usually after the Nickel column. The final MINA53 protein was eluted off the gel filtration column in 20 mM Tris 8.5, 5% glycerol, and 150 mM NaCl and concentrated using a 10kd MWCO Millipore

concentrator to at least 10 mg/ml, determined by OD_{280} reading. Purified protein was stored on ice and used for crystallization as soon as possible (within 1 week). Aliquots of protein was frozen in -80 ℃ and taken out for activity assays.

4.4.3 Crystallization

Based on known structure of the JmjC domain, 5mM 2-OG was added to purified MINA53 before crystallization. Proteins were set up on 96-well trays (Hampton Research) either manually or by a Phoenix Liquid Handling System (Art Robbins Instruments). The solution sets used for initial screening include: B-Main, C-Main, Synergy, and Index (Hampton Research). Typically, 0.6 ~ 1 µl of protein was mixed with equal amounts of crystallization solution. The mixtures were deposited as either sitting or hanging drops into the wells on trays. The wells were sealed with plastic tapes and kept at 16°C. Trays were observed under a light microscope at different times after set up to for the appearance of crystals. For conditions in which crystals or needle crystals were observed after a period of time, more conditions were made for optimization of crystals. Typically, these conditions varied in pH and salt concentrations around a certain range of the original solution. Crystallization was tried on both small trays and large hanging drop trays. Results were recorded on scoring sheets.

4.4.4 Demethylation Assays

Demethylation assays were carried out according to protocols used in the lab for PHF8 (Horton et al., 2010). Methylated-lysine histone peptides ((1-10) H3K4, (1-15/24) H3K9, (1-21) H3K18, (1-35) H3K27, (21-44) H3K36, each with mono-, di-, and tri-methyl groups; table 4A) and methylated-arginine peptides (Table 4B) were tested. Twenty micromolar of peptide substrates were incubated with 10 µM MINA53 protein in a 10/20 µl

reaction solution containing 20 mM pH buffer, 10 mM 2-OG, 50 μM $Fe_2(NH_4)(SO_4)_2$, and 2 mM ascorbic acid at 30/37 ˚C overnight. BisTris pH 5.5/6.0/6.4 and Tris pH 7.0/7.5/8.5 were tested. Reactions without MINA53 were used as negative controls. Reactions with PHF8 and KIAA1718 were used as positive controls for H3K9 and H3K27 peptides. Two sets of samples were collected at 2hrs and overnight and subject to mass spectrometry analysis (as described in Chapter II) to identify peak shift (mass changes, if any) of the peptides. No activity was detected in this assay for all peptides that we tested.

Chapter V: Conclusions and Future Directions

5.1 Substrate specificity of JARID1 demethylases

One way to compare Jumonji proteins is to categorize them into two groups: those with JmjC and other known domains, and those with JmjC-only but without other recognizable domains (Klose et al., 2006; Tsukada et al., 2006). Current knowledge of Jumonji demethylases has reveal histone demethylation specificities toward H3K4me2/3 (Secombe and Eisenman, 2007), H3K9me1/2/3 (Couture et al., 2007; Horton et al., 2010) , H3K27me2/3 (van Haaften et al., 2009), and H3K36me2/3 (Couture et al., 2007; Ng et al., 2007). These Jumonji demethylases mostly carry auxiliary domains other than JmjC domain. The only JmjC-domain-only protein identified to bear enzymatic activity was NO66, which was reported to target H3K4 and H3K36 methylation dependent on the protein's interaction with other transcription factors (Eilbracht et al., 2004; Sinha et al., 2010). However, currently no additional data published by other groups was available to support this discovery. From the above information, we observe a correlation between Jumonji demethylases' enzymatic activity and their capability to interact with the chromatin, a potential paradigm by which JARID1 proteins follow.

In our study, we analyzed the enzymatic properties of JARID1C and performed the Michaelis-Menten kinetics study of a Jumonji H3K4 demethylase. We found that JARID1C is the fastest Jumonji demethylase among those with published k_{cat} values, as mentioned in Chaper II. JMJD2E is also a fast Jumonji demethylase (Katoh and Katoh, 2007; Kawamura et al., 2010), although a k_{cat} value has not been reported and therefore we did not compare it with JARID1C. One interesting question is why does a Jumonji H3K4 demethylase might possess a faster reaction rate than Jumonji H3K9/36 demethylases, by a magnitude of more than 10 (compared with PHF8/KIAA1718) or even 100 fold (compared with JMJD2A)? While errors in of *in vitro* experimental set up could be a factor to consider, there are other implications. The JARID1C

protein we used here contained intact N-terminal domains; but the PHF8 and JMJD2A proteins used in other studies contained only the JmjC and neighboring domains (PHD for PHF8, and JmjN for JMJD2A). This possibly means that the activity of JmjC domains is fully retained in more complete protein structures. It will be also interesting to see if the activity differences have any biological meaning. For instance, assuming even in intact proteins JARID1C is still faster than the other demethylases, does is imply that its faster removal rate of an active mark is needed for high efficiency of gene silencing? What is the fate of the end product – H3K4me1 – of JARID1 demethylation, a distinct mark for enhancers (Blair et al., 2011) or a substrate for LSD1 to be further demethylated? How is the fate determined? Similar scenarios also exist for H3K9 and other histone lysine demethylation. Although not yet reported, a systematic experimental set up will be needed for such investigations, considering the complex network of interactions among Jumonji demethylases' and the chromatin as well as the native environment in the cell. Meanwhile, our knowledge of the catalytic properties of JARID1C and other Jumonji proteins also directly applies to enzyme inhibitor studies aimed at drug discovery.

5.2 Mechanistic insights into the interplay among JARID1C's domains

As numerous known histone modifiers highlight the mode of gene regulation through reading, writing, and erasure of chromatin marks, a multi-domain protein such as JARID1C naturally prompts people to hypothesize about its interaction with chromatin through the reader domains. We did not find evidence in this study to support our initial model, where the enzymatic activity of JmjC domain can be affected by interactions via JARID1C's interaction domains such as PHDs. However, much work could be done to explore the possibilities. Interestingly, the ARID domain has been shown to bind sequence-specific DNA in various proteins, including JARID1 homologues such as JARID1A (Tu et al., 2008) and JARID1B (Scibetta et al., 2007). The sequence homology and structural similarity among ARID domains makes JARID1C's ARID domain a

candidate for future investigation in terms of binding specificity and effect on enzymatic activity (Scibetta et al., 2007; Tu et al., 2008; Yao et al., 2009; Yao et al., 2010). Although our experimental observations from this study and in the literature did not point to a same answer to the binding specificity of JARID1C PHD1 and did not detect interaction between JARID1C and known ARID-binding DNA sequences, functional studies in cell biology will suggest protein/nucleic acid targets for biochemical characterization. An example of JARID1 PHD domain mediating protein-protein interaction in the chromatin comes from Msc1, a fission yeast JARID1 homolog. The C-terminal PHD domain of Msc1 was found required for Msc1's association of the Swr1 complex, which is responsible for histone H2A.Z deposition into the chromatin in *S.pombe* (Yang et al., 2010). In this case, the PHD3 domain of Msc1 acts as an E3 ubiquitin ligase. Therefore, the search for JARID1C's PHD domain function is open to a variety of targets.

Particularly, structural studies will help explain the questions left in this study. Currently no Jumonji H3K4 demethylase structure was published (Upadhyay and Cheng, 2011), and given our ability to purify JARID1C (1-839) protein in sufficient amounts, crystallization and structural determination of this protein should be possible. Such structural information will directly allow the spatial orientation and overall architecture of the N-terminal domains of JARID1C to be visualized. Our PHD1 and PHD2 construct yields quality proteins after purification, and PHD1 and PHD2 domain crystal structures will also provide valuable hints on their binding partners when compared to known PHD domain structures. Co-crystal structure of protein and substrate/binding partner and even inhibitors would further reveal the interaction of JARID1 and Jumonji demethylases, and in turn provide information regarding the regulation of their enzymatic properties. Given the presence of several putative DNA- binding and chromatin-interacting domains, we predict that exciting models picturing JARID1 proteins' interaction with intact chromatin structure will come out as more data becomes available.

5.3 JARID1 demethylase as potential drug targets

Inhibitors of enzymes in the chromatin have been proved effective in disease treatment. For example, LSD1 (lysine-specific demethylase 1) inhibitors have been well studied (Hou and Yu, 2010) and HDAC (histone deacetylase) inhibitors have even been approved by FDA for cancer treatment (Bolden et al., 2006; Minucci and Pelicci, 2006). JARID1 H3K4 demethylases generally function as repressors (Secombe and Eisenman, 2007), and they have been linked to a wide range of cancer types (Blair et al., 2011). Thus, there exists a strong rationale for JARID1 proteins to be targets for cancer treatment.

Mechanistic and structural studies revealed that most of the reported Jumonji protein inhibitors are metabolic products or structural analogues of 2-OG and act as 2-OG competitors (Hamada et al., 2010; Hou and Yu, 2010). Currently no data has been available to assess the inhibition of JARID1 proteins by natural or synthetic inhibitors. By measuring IC_{50} values of (R/S) 2-HG on JARID1C demethylation, we provided the first 2-HG inhibition data on a Jumonji H3K4 demethylase and showed that 2-HG is about 50 fold less inhibitory than the known inhibitor NOG in our formaldehyde-release assay conditions. The reported IC_{50} of 2-HG for JMJD2A (24 μM for R form, 26 μM for S form) and JMJD2C (79 μM for R from, 97 μM for S form) (Chowdhury et al., 2011) showed that both forms of 2-HG are as inhibitory as NOG (17 μM for JMJD2A, 14 μM for JMJD2C) toward these two enzymes. One possible explanation is the difference in intrinsic structural and catalytic properties between H3K4 (JARID1C) and H3K9 (JMJD2A/C) demethylases, although we currently don't have data to support this idea. Alternatively, this difference possibly results from the amount of 2-OG used in the inhibition assays. The k_m values of JARID1C, JMJD2A and JMJD2C for 2-OG are 5, 4, and 6 μM, respectively, which are very close. In our assay for JARID1C, we used 200 μM of 2-OG so that the initial velocity is starting to be saturated. In the assay for JMJD2A/C, 10 μM of 2-OG was used. Although it was higher than the k_m for 2-OG, it had not yet saturated the initial velocity and was 20 fold lower than our 2-OG concentration. As a

result, a lower amount of inhibitors were needed to inhibit 50% activity of the demethylases and apparent IC_{50} values would be smaller. Nevertheless, as stated above, IC_{50} values are not directly comparable across assays. Currently available k_i data as mentioned in Chapter II suggests that 2-HG inhibits JMJD2A/C at micromolar range (similar to NOG; measured by formaldehyde release assay) (Chowdhury et al., 2011) but inhibits JARID1B and *C. elegans* KIAA1718 at millimolar range (measured by mass spectrometry) (Xu et al., 2011). In formaldehyde release assay, inhibition of initial velocity was measured (as in our study), while in mass spectrometry, inhibition on the amount of substrate converted was measured. As k_i values can be compared across assays, future experiments to systematically measure k_i values of these demethylases including JARID1C and other characterized Jumonji proteins are needed. In addition to 2-HG, comparison of the potencies of different types of inhibitors toward different group of histone demethylases will also add to our understanding of their catalytic properties.

Collectively, multiple levels of regulation (enzymology, post-translational modification, and protein-protein or protein-nucleic-acid interaction) highlight how Jumonji demethylases integrate into the chromatin and fine tune gene expression. The results of this project provide further details toward a continually-refined-model of understanding histone demethylation in human diseases, potentially leading to therapeutic development with epigenetic targets such as designing histone demethylase inhibitors.

Appendix

Summary of all JARID1C and other constructs generated

pXC #	Protein	Sequence	Tag	Expression	Solubility	Notes
719	MINA53	1-465	His-SUMO	++	++	
841	MINA53	1-465	GST	+	+	
869	MINA53	1-465	No tag	++	+++	Could be purified without Ni column (directly load on SP)
722	JMJD4	47-463	His-SUMO	-	-	
840	JMJD4	47-463	GST	+	-	
848	NO66	1-641	GST	-	-	
877	NO66	1-641	His	-	-	The start codon was disrupted by AseI-NdeI fusion site and therefore no protein will be expressed
903	JARID1A	1-759	His	+	-	
904	JARID1A	1-759	GST	+	-	
905	JARID1A	1-808	His	+	-	
906	JARID1A	1-808	GST	+	-	
907	JARID1C	1-789	His	+	-	
908	JARID1C	1-789	His-SUMO	+	-	
909	JARID1C	1-789	GST	+	-	
910	JARID1C	1-839	His	+	-	

911	JARID1C	1-839	His-SUMO	+	+	Soluble and enzymatically active protein
912	JARID1C	1-839	GST	-	-	
909	JARID1C	1-789	GST	+	+	
909	JARID1C	1-789	GST	+	+	
909	JARID1C	1-789	GST	+	+	
949	JARID1C	320-382	GST	+	-	
950	JARID1C	320-414	GST	+	-	
951	JARID1C	261-414	GST	+	+	Protein is soluble but with a lot of degradation bands
952	JARID1C	320-382	His-SUMO	++	+++	
953	JARID1C	320-414	His-SUMO	+	-	
954	JARID1C	261-414	His-SUMO	+	-	
955	JARID1C	320-700	GST	-	-	
956	JARID1C	320-700	His-SUMO	-	-	
978	JARID1C	1181-1249	GST	+	-	
979	JARID1C	1146-1267	GST	+	-	
980	JARID1C	1181-1249	His-SUMO	++	++	
981	JARID1C	1146-1267	His-SUMO	+	+	The protein will elute off a HisTrap HP column at over 1 M of imidazole
983	JARID1C	1-1249	His-SUMO	+	-	
1014	JARID1C	1-1249	His-SUMO	+	-	

References

Baker, L.A., Allis, C.D., and Wang, G.G. (2008). PHD fingers in human diseases: Disorders arising from misinterpreting epigenetic marks. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis *647*, 3-12.

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120-124.

Benevolenskaya, E.V., Murray, H.L., Branton, P., Young, R.A., and Kaelin, W.G., Jr. (2005). Binding of pRB to the PHD protein RBP2 promotes cellular differentiation. Mol Cell *18*, 623-635.

Berger, S.L. (2007). The complex language of chromatin regulation during transcription. Nature *447*, 407-412.

Bienz, M. (2006). The PHD finger, a nuclear protein-interaction domain. Trends in Biochemical Sciences *31*, 35-40.

Blair, L.P., Cao, J., Zou, M.R., Sayegh, J., and Yan, Q. (2011). Epigenetic Regulation by Lysine Demethylase 5 (KDM5) Enzymes in Cancer. Cancers *3*, 1383-1404.

Bolden, J.E., Peart, M.J., and Johnstone, R.W. (2006). Anticancer activities of histone deacetylase inhibitors. Nat Rev Drug Discov *5*, 769-784.

Burden, A.F., Manley, N.C., Clark, A.D., Gartler, S.M., Laird, C.D., and Hansen, R.S. (2005). Hemimethylation and non-CpG methylation levels in a promoter region of human LINE-1 (L1) repeated elements. J Biol Chem *280*, 14413-14419.

Cedar, H., and Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. Nat Rev Genet *10*, 295-304.

Chen, Y., Yang, Y., Wang, F., Wan, K., Yamane, K., Zhang, Y., and Lei, M. (2006). Crystal structure of human histone lysine-specific demethylase 1 (LSD1). Proc Natl Acad Sci U S A *103*, 13956-13961.

Cheng, X., and Blumenthal, R.M. (2008). Mammalian DNA methyltransferases: a structural perspective. Structure *16*, 341-350.

Chowdhury, R., Yeoh, K.K., Tian, Y.-M., Hillringhaus, L., Bagg, E.A., Rose, N.R., Leung, I.K.H., Li, X.S., Woon, E.C.Y., Yang, M.*, et al.* (2011). The oncometabolite 2-hydroxyglutarate inhibits histone lysine demethylases. EMBO Rep *12*, 463-469.

Christensen, J., Agger, K., Cloos, P.A., Pasini, D., Rose, S., Sennels, L., Rappsilber, J., Hansen, K.H., Salcini, A.E., and Helin, K. (2007). RBP2 belongs to a family of demethylases, specific for tri-and dimethylated lysine 4 on histone 3. Cell *128*, 1063-1076.

Cloos, P.A., Christensen, J., Agger, K., and Helin, K. (2008). Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease. Genes Dev *22*, 1115-1140.

Collins, R., and Cheng, X. (2010). A case study in cross-talk: the histone lysine methyltransferases G9a and GLP. Nucl Acids Res *38*, 3503-3511.

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

Culhane, J.C., and Cole, P.A. (2007). LSD1 and the chemistry of histone demethylation. Curr Opin Chem Biol *11*, 561-568.

Eden, S., Hashimshony, T., Keshet, I., Cedar, H., and Thorne, A.W. (1998). DNA methylation models histone acetylation. Nature *394*, 842.

Eilbracht, J., Reichenzeller, M., Hergt, M., Schnolzer, M., Heid, H., Stohr, M., Franke, W.W., and Schmidt-Zachmann, M.S. (2004). NO66, a highly conserved dual location protein in the nucleolus and in a special type of synchronously replicating chromatin. Mol Biol Cell *15*, 1816-1832.

Fathman, J.W., Gurish, M.F., Hemmers, S., Bonham, K., Friend, D.S., Grusby, M.J., Glimcher, L.H., and Mowen, K.A. (2010). NIP45 controls the magnitude of the type 2 T helper cell response. Proc Natl Acad Sci U S A *107*, 3663-3668.

Feng, W., Yonezawa, M., Ye, J., Jenuwein, T., and Grummt, I. (2010). PHF8 activates transcription of rRNA genes through H3K4me3 binding and H3K9me1/2 demethylation. Nat Struct Mol Biol *17*, 445-450.

Forneris, F., Binda, C., Adamo, A., Battaglioli, E., and Mattevi, A. (2007). Structural basis of LSD1-CoREST selectivity in histone H3 recognition. J Biol Chem *282*, 20070-20074.

Gamble, M.J., and Kraus, W.L. (2007). Visualizing the histone code on LSD1. Cell *128*, 433-434.

Gutierrez, G.M., Kong, E., and Hinds, P.W. (2005). Master or slave: The complex relationship of RBP2 and pRb. Cancer Cell *7*, 501-502.

Hamada, S., Suzuki, T., Mino, K., Koseki, K., Oehme, F., Flamme, I., Ozasa, H., Itoh, Y., Ogasawara, D., Komaarashi, H.*, et al.* (2010). Design, Synthesis, Enzyme-Inhibitory Activity, and Effect on Human Cancer Cells of a Novel Series of Jumonji Domain-Containing Protein 2 Histone Demethylase Inhibitors. Journal of Medicinal Chemistry *53*, 5629-5638.

Horton, J.R., Upadhyay, A.K., Qi, H.H., Zhang, X., Shi, Y., and Cheng, X. (2010). Enzymatic and structural insights for substrate specificity of a family of jumonji histone lysine demethylases. Nat Struct Mol Biol *17*, 38-43.

Hou, H., and Yu, H. (2010). Structural insights into histone lysine demethylation. Current Opinion in Structural Biology *20*, 739-748.

Iwahara, J., Iwahara, M., Daughdrill, G.W., Ford, J., and Clubb, R.T. (2002). The structure of the Dead ringer-DNA complex reveals how AT-rich interaction domains (ARIDs) recognize DNA. EMBO J *21*, 1197-1209.

Iwase, S., Lan, F., Bayliss, P., de la Torre-Ubieta, L., Huarte, M., Qi, H.H., Whetstine, J.R., Bonni, A., Roberts, T.M., and Shi, Y. (2007). The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. Cell *128*, 1077-1088.

Kangaspeska, S., Stride, B., Metivier, R., Polycarpou-Schwarz, M., Ibberson, D., Carmouche, R.P., Benes, V., Gannon, F., and Reid, G. (2008). Transient cyclical methylation of promoter DNA. Nature *452*, 112-115.

Katoh, Y., and Katoh, M. (2007). Comparative integromics on JMJD2A, JMJD2B and JMJD2C: preferential expression of JMJD2C in undifferentiated ES cells. Int J Mol Med *20*, 269-273.

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

Kleine-Kohlbrecher, D., Christensen, J., Vandamme, J., Abarrategui, I., Bak, M., Tommerup, N., Shi, X., Gozani, O., Rappsilber, J., Salcini, A.E.*, et al.* (2010). A functional link between the histone demethylase PHF8 and the transcription factor ZNF711 in X-linked mental retardation. Mol Cell *38*, 165-178.

Klose, R.J., Kallin, E.M., and Zhang, Y. (2006). JmjC-domain-containing proteins and histone demethylation. Nat Rev Genet *7*, 715-727.

Klose, R.J., Yan, Q., Tothova, Z., Yamane, K., Erdjument-Bromage, H., Tempst, P., Gilliland, D.G., Zhang, Y., and Kaelin, W.G., Jr. (2007). The retinoblastoma binding protein RBP2 is an H3K4 demethylase. Cell *128*, 889-900.

Komiya, K., Sueoka-Aragane, N., Sato, A., Hisatomi, T., Sakuragi, T., Mitsuoka, M., Sato, T., Hayashi, S., Izumi, H., Tsuneoka, M.*, et al.* (2009). Mina53, a novel c-Myc target gene, is frequently expressed in lung cancers and exerts oncogenic property in NIH/3T3 cells. J Cancer Res Clin Oncol.

Lan, F., Collins, R.E., De Cegli, R., Alpatov, R., Horton, J.R., Shi, X., Gozani, O., Cheng, X., and Shi, Y. (2007). Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. Nature *448*, 718-722.

Lan, F., Nottke, A.C., and Shi, Y. (2008). Mechanisms involved in the regulation of histone lysine demethylases. Curr Opin Cell Biol *20*, 316-325.

Li, F., Huarte, M., Zaratiegui, M., Vaughn, M.W., Shi, Y., Martienssen, R., and Cande, W.Z. (2008). Lid2 Is Required for Coordinating H3K4 and H3K9 Methylation of Heterochromatin and Euchromatin. Cell *135*, 272-283.

Li, H., Ilin, S., Wang, W., Duncan, E.M., Wysocka, J., Allis, C.D., and Patel, D.J. (2006). Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. Nature *442*, 91-95.

Lu, Y., Chang, Q., Zhang, Y., Beezhold, K., Rojanasakul, Y., Zhao, H., Castranova, V., Shi, X., and Chen, F. (2009). Lung cancer-associated JmjC domain protein mdig suppresses formation of tri-methyl lysine 9 of histone H3. Cell Cycle *8*, 2101-2109.

Lu, Z., Xu, S., Joazeiro, C., Cobb, M.H., and Hunter, T. (2002). The PHD domain of MEKK1 acts as an E3 ubiquitin ligase and mediates ubiquitination and degradation of ERK1/2. Molecular Cell *9*, 945-956.

Matthews, A.G., Kuo, A.J., Ramon-Maiques, S., Han, S., Champagne, K.S., Ivanov, D., Gallardo, M., Carney, D., Cheung, P., Ciccone, D.N*., et al.* (2007). RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. Nature *450*, 1106-1110.

Metivier, R., Gallais, R., Tiffoche, C., Le Peron, C., Jurkowska, R.Z., Carmouche, R.P., Ibberson, D., Barath, P., Demay, F., Reid, G*., et al.* (2008). Cyclical DNA methylation of a transcriptionally active promoter. Nature *452*, 45-50.

Metzger, E., Wissmann, M., Yin, N., Muller, J.M., Schneider, R., Peters, A.H.F.M., Gunther, T., Buettner, R., and Schule, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature *437*, 436-439.

Minucci, S., and Pelicci, P.G. (2006). Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer. Nat Rev Cancer *6*, 38-51.

Mowen, K.A., Schurter, B.T., Fathman, J.W., David, M., and Glimcher, L.H. (2004). Arginine methylation of NIP45 modulates cytokine gene expression in effector T lymphocytes. Mol Cell *15*, 559-571.

Nakayama, J.-i., Rice, J.C., Strahl, B.D., Allis, C.D., and Grewal, S.I.S. (2001). Role of Histone H3 Lysine 9 Methylation in Epigenetic Control of Heterochromatin Assembly. Science *292*, 110-113.

Ng, S.S., Kavanagh, K.L., McDonough, M.A., Butler, D., Pilka, E.S., Lienard, B.M.R., Bray, J.E., Savitsky, P., Gileadi, O., von Delft, F*., et al.* (2007). Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity. Nature *448*, 87-91.

Okamoto, M., Van Stry, M., Chung, L., Koyanagi, M., Sun, X., Suzuki, Y., Ohara, O., Kitamura, H., Hijikata, A., Kubo, M*., et al.* (2009). Mina, an Il4 repressor, controls T helper type 2 bias. Nat Immunol *10*, 872-879.

Perillo, B., Ombra, M.N., Bertoni, A., Cuozzo, C., Sacchetti, S., Sasso, A., Chiariotti, L., Malorni, A., Abbondanza, C., and Avvedimento, E.V. (2008). DNA oxidation as triggered by H3K9me2 demethylation drives estrogen-induced gene expression. Science *319*, 202-206.

Qiu, J., Shi, G., Jia, Y., Li, J., Wu, M., Dong, S., and Wong, J. (2010). The X-linked mental retardation gene PHF8 is a histone demethylase involved in neuronal differentiation. Cell Res *20*, 908-918.

Ramsahoye, B.H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A.P., and Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proc Natl Acad Sci U S A *97*, 5237-5242.

Roy, T.W., and Bhagwat, A.S. (2007). Kinetic studies of Escherichia coli AlkB using a new fluorescence-based assay for DNA demethylation. Nucleic Acids Research *35*, e147.

Ruthenburg, A.J., Li, H., Patel, D.J., and David Allis, C. (2007). Multivalent engagement of chromatin modifications by linked binding modules. Nat Rev Mol Cell Biol *8*, 983-994.

Santos-Rosa, H., Schneider, R., Bernstein, B.E., Karabetsou, N., Morillon, A., Weise, C., Schreiber, S.L., Mellor, J., and Kouzarides, T. (2003). Methylation of Histone H3 K4 Mediates Association of the Isw1p ATPase with Chromatin. Molecular Cell *12*, 1325-1332.

Schneider, R., Bannister, A.J., Myers, F.A., Thorne, A.W., Crane-Robinson, C., and Kouzarides, T. (2004). Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. Nat Cell Biol *6*, 73-77.

Scibetta, A.G., Santangelo, S., Coleman, J., Hall, D., Chaplin, T., Copier, J., Catchpole, S., Burchell, J., and Taylor-Papadimitriou, J. (2007). Functional Analysis of the Transcription Repressor PLU-1/JARID1B. Mol Cell Biol *27*, 7220-7235.

Secombe, J., and Eisenman, R.N. (2007). The function and regulation of the JARID1 family of histone H3 lysine 4 demethylases: the Myc connection. Cell Cycle *6*, 1324-1328.

Sharma, S.V., Lee, D.Y., Li, B., Quinlan, M.P., Takahashi, F., Maheswaran, S., McDermott, U., Azizian, N., Zou, L., Fischbach, M.A*., et al.* (2010). A Chromatin-Mediated Reversible Drug-Tolerant State in Cancer Cell Subpopulations. Cell *141*, 69-80.

Shi, X., Hong, T., Walter, K.L., Ewalt, M., Michishita, E., Hung, T., Carney, D., Pena, P., Lan, F., Kaadige, M.R., *et al.* (2006). ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. Nature *442*, 96-99.

Shi, Y. (2007). Histone lysine demethylases: emerging roles in development, physiology and disease. Nat Rev Genet *8*, 829-833.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., and Casero, R.A. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell *119*, 941-953.

Sims Iii, R.J., Nishioka, K., and Reinberg, D. (2003). Histone lysine methylation: a signature for chromatin function. Trends in Genetics *19*, 629-639.

Sinha, K.M., Yasuda, H., Coombes, M.M., Dent, S.Y.R., and de Crombrugghe, B. (2010). Regulation of the osteoblast-specific transcription factor Osterix by NO66, a Jumonji family histone demethylase. EMBO J *29*, 68-79.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., *et al.* (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324*, 930-935.

Tahiliani, M., Mei, P., Fang, R., Leonor, T., Rutenberg, M., Shimizu, F., Li, J., Rao, A., and Shi, Y. (2007). The histone H3K4 demethylase SMCX links REST target genes to X-linked mental retardation. Nature *447*, 601-605.

Takeuchi, T., Watanabe, Y., Takano-Shimizu, T., and Kondo, S. (2006). Roles of jumonji and jumonji family genes in chromatin regulation and development. Dev Dyn *235*, 2449-2459.

Takeuchi, T., Yamazaki, Y., Katoh-Fukui, Y., Tsuchiya, R., Kondo, S., Motoyama, J., and Higashinakagawa, T. (1995). Gene trap capture of a novel mouse gene, jumonji, required for neural tube formation. Genes Dev *9*, 1211-1222.

Tan, H., Wu, S., Wang, J., and Zhao, Z.K. (2008). The JMJD2 members of histone demethylase revisited. Mol Biol Rep *35*, 551-556.

Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat Struct Mol Biol *14*, 1025-1040.

Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. Nature *439*, 811-816.

Tsuneoka, M., Koda, Y., Soejima, M., Teye, K., and Kimura, H. (2002). A novel myc target gene, mina53, that is involved in cell proliferation. J Biol Chem *277*, 35450-35459.

Tu, S., Teng, Y.-C., Yuan, C., Wu, Y.-T., Chan, M.-Y., Cheng, A.-N., Lin, P.-H., Juan, L.-J., and Tsai, M.-D. (2008). The ARID domain of the H3K4 demethylase RBP2 binds to a DNA CCGCCC motif. Nat Struct Mol Biol *15*, 419-421.

Uchida, D., Hatakeyama, S., Matsushima, A., Han, H., Ishido, S., Hotta, H., Kudoh, J., Shimizu, N., Doucas, V., Nakayama, K.I., *et al.* (2004). AIRE functions as an E3 ubiquitin ligase. J Exp Med *199*, 167-172.

Upadhyay, A.K., and Cheng, X. (2011). Dynamics of histone lysine methylation: structures of methyl writers and erasers. Prog Drug Res *67*, 107-124.

van Haaften, G., Dalgliesh, G.L., Davies, H., Chen, L., Bignell, G., Greenman, C., Edkins, S., Hardy, C., O'Meara, S., Teague, J., *et al.* (2009). Somatic mutations of the histone H3K27 demethylase gene UTX in human cancer. Nat Genet *41*, 521-523.

Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J.M., *et al.* (2006). The Polycomb group protein EZH2 directly controls DNA methylation. Nature *439*, 871-874.

Wang, G.G., Song, J., Wang, Z., Dormann, H.L., Casadio, F., Li, H., Luo, J.-L., Patel, D.J., and Allis, C.D. (2009). Haematopoietic malignancies caused by dysregulation of a chromatin-binding PHD finger. Nature *459*, 847-851.

Ward, P.S., Patel, J., Wise, D.R., Abdel-Wahab, O., Bennett, B.D., Coller, H.A., Cross, J.R., Fantin, V.R., Hedvat, C.V., Perl, A.E., *et al.* (2010). The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate. Cancer Cell *17*, 225-234.

Webby, C.J., Wolf, A., Gromak, N., Dreger, M., Kramer, H., Kessler, B., Nielsen, M.L., Schmitz, C., Butler, D.S., Yates, J.R., 3rd, *et al.* (2009). Jmjd6 catalyses lysyl-hydroxylation of U2AF65, a protein associated with RNA splicing. Science *325*, 90-93.

Xiang, Y., Zhu, Z., Han, G., Ye, X., Xu, B., Peng, Z., Ma, Y., Yu, Y., Lin, H., Chen, A.P., *et al.* (2007). JARID1B is a histone H3 lysine 4 demethylase up-regulated in prostate cancer. Proc Natl Acad Sci U S A *104*, 19226-19231.

Xu, W., Yang, H., Liu, Y., Yang, Y., Wang, P., Kim, S.-H., Ito, S., Yang, C., Wang, P., Xiao, M.-T.*, et al.* (2011). Oncometabolite 2-Hydroxyglutarate Is a Competitive Inhibitor of [alpha]-Ketoglutarate-Dependent Dioxygenases. Cancer Cell *19*, 17-30.

Yang, C.G., Yi, C., Duguid, E.M., Sullivan, C.T., Jian, X., Rice, P.A., and He, C. (2008). Crystal structures of DNA/RNA repair enzymes AlkB and ABH2 bound to dsDNA. Nature *452*, 961-965.

Yang, Y., Hu, L., Wang, P., Hou, H., Lin, Y., Liu, Y., Li, Z., Gong, R., Feng, X., Zhou, L.*, et al.* (2010). Structural insights into a dual-specificity histone demethylase ceKDM7A from Caenorhabditis elegans. Cell Res *20*, 886-898.

Yao, W., Peng, Y., Chen, Q., and Lin, D. (2009). 1H, 13C, 15N backbone and side-chain resonance assignments of the Bright/ARID domain from the human histone demethylase JARID1B. Biomol NMR Assign *3*, 85-87.

Yao, W., Peng, Y., and Lin, D. (2010). The flexible loop L1 of the H3K4 demethylase JARID1B ARID domain has a crucial role in DNA-binding activity. Biochem Biophys Res Commun *396*, 323-328.

Yonashiro, R., Ishido, S., Kyo, S., Fukuda, T., Goto, E., Matsuki, Y., Ohmura-Hoshino, M., Sada, K., Hotta, H., Yamamura, H.*, et al.* (2006). A novel mitochondrial ubiquitin ligase plays a critical role in mitochondrial dynamics. The EMBO journal *25*, 3618-3626.

Zhang, X., Tamaru, H., Khan, S.I., Horton, J.R., Keefe, L.J., Selker, E.U., and Cheng, X. (2002). Structure of the Neurospora SET Domain Protein DIM-5, a Histone H3 Lysine Methyltransferase. Cell *111*, 117-127.

Zhang, Y., Lu, Y., Yuan, B.Z., Castranova, V., Shi, X., Stauffer, J.L., Demers, L.M., and Chen, F. (2005). The Human mineral dust-induced gene, mdig, is a cell growth regulating gene associated with lung cancer. Oncogene *24*, 4873-4882.